

Chapter No.

TRUST IN WITTGENSTEIN

Olli Lagerspetz and Lars Hertzberg

0. Introduction

There are important reasons why the work of Ludwig Wittgenstein ought to be central to any philosophical study of trust. This is partly because of what he explicitly says about this and related issues (such as knowledge and certainty), and partly because of the applicability of his more general insights.

Wittgenstein's various contributions to philosophy are, generally speaking, related to certain basic insights about the role of language. He thinks we should not presuppose that the meaning of a given expression is established by identifying its referents, the phenomena that it stands for. Instead, we should look at the uses of the expression in various contexts of human life; in particular, its *different* uses. For our present purposes, it is profitable to look at Wittgenstein's discussions of psychological expressions. He encourages the reader not to engage in introspection in order to capture the essence of given psychological states.

Applying this to the present case: What we need to realize here is that, by invoking the language of trust and betrayal, we do not simply identify facts out there. Rather we invoke a certain perspective. We are invited to see someone's behavior in a certain light. Thus the question to ask is not, "what are the proper criteria for applying the word 'trust' to psychological or behavioral phenomena?" but instead, "what is the role, in human interaction, of speaking about trust?"

The aim of the present paper is to present relevant strands of Wittgenstein's work and to relate them to the contemporary discussion on trust. Three main Wittgensteinian themes will be identified: (1) a general methodological point about how to investigate psychological expressions; (2) a critique of the idea that our grasp on reality can ultimately be spelled out by means of factual assertions; and (3) Wittgenstein's own emphasis, especially in *On Certainty*, on the logical and epistemic role, in learning and judging, of our dependence on other people. However, in our discussion we will be going beyond what Wittgenstein said. We hope to show that his methods and insights may be fruitfully used to shed light also on themes that he did not address directly.

1. On Psychological Expressions

A. Is Trust a Psychological State?

Grammatically, trust is a state that an individual is in, a state somehow relating to another individual (as in "X has trust in Y"). However, so far this description does not specify what *kind* of a state trust is. By "state", one may mean a person's current state of mind or body, but also dispositional states, even such 'states' as that of being rich or poor. The point of speaking of a state in all these cases is that some lasting state of affairs pertaining to an

individual is used as a point of reference when something about the individual's life is explained or judged. Thus the fact that I rely on someone's promise may, in some context, be explained and judged in terms of the fact that I trust her.

But is trust a *psychological* state? For instance, when, if ever, can a feeling of trust be identified?¹

According to Annette Baier, trust "has a special 'feel', most easily acknowledged when it is missed, say, when one moves from a friendly 'safe' neighbourhood to a tense insecure one."² However, an obvious problem with this description is that it does not really specify a feeling of trust. Baier is, instead, describing a case of distrust. This highlights a general difficulty about pinpointing specific feelings or thoughts that might constitute trust as a psychological state. Often trust is, on the contrary, characterized by the *absence of* certain feelings and thoughts such as suspicion and fear.

Also, when I trust someone it is not, for the most part, something I think about; rather the opposite, as we shall see. Nor do I mostly do so because I have, at some point in time, formed the judgment that the person in question is trustworthy. If I am short-changed at the grocery store, but only notice it afterwards, it will be correct to say I had trusted the shop attendant to hand me the correct change. However, perhaps I had never addressed the question in my mind. And I need not have harbored specific feelings about him either.

In his *Philosophical Investigations*, Wittgenstein takes up an analogous question. He discusses the "feeling of familiarity" that may strike us, for instance, when we enter the same old room after a long absence. He points out that, apart from such situations, familiar objects in our surroundings do not typically give us a feeling of familiarity. It is easier to get at a feeling of strangeness at the sight of unfamiliar objects, even though we do not have that feeling whenever we see something unfamiliar.³ Thus even if feelings of familiarity do exist, their presence or absence is not typically the criterion of what is familiar to us.

This is part of a more general theme addressed in *Zettel*. In that work, Wittgenstein distinguishes between cases of believing, feeling, thinking, and so on, that involve "genuine duration" and others that do not.⁴ A psychological state has genuine duration if it makes sense to ask whether the state endures from one moment to the next. We can imagine asking "do you still feel the pain?" and a few minutes later, "what about now?" Also, if a state has genuine duration, we can imagine it being interrupted. Being in pain is a clear case of genuine duration. On the other hand, suppose I ask you "are you still a Republican?" and again, after five minutes of silence, "what about now?" This could hardly be understood as anything but a joke unless, of course, you have received some quite extraordinary revelation in the meantime. Nor could we imagine someone being interrupted in being a Republican. The state of being a Republican, then, does not have genuine duration.⁵

Anything that can be measured with a stopwatch would be a clear case of genuine duration. This may not always be possible, though, contrary to what Wittgenstein suggests, since the beginning or end of the state may be diffuse, and thus it may not be clockable. If the answer to the question

“Do you still feel it?” is “I’m not sure” that does not necessarily rule out its being a case of genuine duration.

For instance, sleep, physical pain, intense thinking, and intense expectation have genuine duration. Knowledge, ability, understanding, and intention, on the other hand, do not.⁶ They are not interrupted when a person is asleep or in a faint. One can be interrupted in thinking or planning, but not in intending.⁷

An important connecting theme in the *Investigations* is Wittgenstein’s emphasis on the fact that in many cases, the intelligibility of attributing a feeling or attitude to someone is dependent on that person’s overall situation. Thus we can meaningfully attribute love, or hope,⁸ or grief⁹ to a person only because that person’s situation as a whole makes these descriptions applicable. We cannot correctly and intelligibly attribute feelings of grief to a person unless she has, for instance, recently lost a loved one, regardless of what else may be true about her states of mind at a given point in time. The presence or absence of a mental state at a given point in time does not settle the issue on its own. For instance, someone at a cinema may feel like crying at the death of the film’s (fictitious) main character, but this will not qualify as genuine grief unless some very peculiar explanation is produced.

Let us here interject a point of methodology. The foregoing remark is not an attempt to exclude unexpected cases by setting up ordinary language as a norm for language use, but a reminder of the fact that intelligible uses of the word “grief” presuppose *some* plausible anchoring in the person’s situation.

Thus we do not suppose that limits of intelligibility can be laid down in the abstract. In suggesting that it might not be intelligible to say certain things in a certain type of situation, we are appealing to the reader’s ability to imagine herself confused if someone were to utter certain words in those kinds of circumstances. This does not entail that her confusion would have to be irremediable. For instance, she might come to understand the words as a joke, a metaphor, a sign of poor command of English, or the like. Still, the fact that our initial response would be one of confusion may help draw our attention to certain features of the use of the word.

Turning to trust in the light of these considerations, first of all it seems clear that the state of trusting does not have genuine duration. In other words, trust is not for the most part manifested as a particular state that occupies one’s mind. The presence of trust must rather be established by looking for an overall pattern in a person’s thinking and acting: a pattern in the weave of life, to apply what Wittgenstein said about grief.¹⁰

Someone’s trust in another may show itself in her being relaxed in his company, as well as in things she does *not* do, such as not taking certain precautions. It may show in her thinking about the future, perhaps plans that rely on information from the person she trusts; or perhaps indeed in the complete absence of a plan where one might be expected. But it does not need to involve specific feelings towards him or explicit thoughts about him. Again we do not deny that, in the right circumstances, it may be natural to describe someone’s state of mind as involving feelings of trust. But the presence or absence of certain psychological states does not settle the issue

on its own. On the contrary, the feelings can plausibly be described as ones of trust only if the appropriate context is in place.

B. Is Trust a Disposition?

At this juncture, someone might suggest a way out of the problem. The reader might think that trust is, or involves, a behavioral disposition instead of (merely) a psychological state. Views along these lines are mostly adopted in current literature. Trust is defined in terms of expectations of a cooperative disposition in others; expectations that may or may not be accompanied by specific feelings. On that view, my trust in a person involves my explicit *or implicit* expectation that she will not take advantage of me but will, rather, behave in ways that are beneficial or at least not harmful to me.¹¹ The presence of such positive expectations (or the absence of negative ones) is, then, seen in the fact that I am prepared to enter into some kind of cooperative relation with the person in question.

Some degree of cooperation is involved in almost all human pursuits. Thus, if the above definition were taken literally, almost all normal human activities would involve an element of trust.

Quite a number of writers have been driven to that extreme position. Baier claims, “[w]e inhabit a climate of trust as we inhabit an atmosphere and notice it as we notice air, only when it becomes scarce or polluted.”¹² For instance, she claims, in a library we trust our fellow library users to be looking for books and not victims between the library stacks.¹³ Still according to Baier, the wife that goes to bed next to her husband trusts that a brain disease will not unexpectedly turn him into a mad aggressor.¹⁴ (The example involves another problem too: if the husband did become violent through a sudden onset of brain disease, she probably would not consider that a betrayal of her trust.) Niklas Luhmann claims, “[o]ne who goes unarmed among his fellow men puts trust in them.”¹⁵ Without such background trust, Lorraine Code sums up, society would “simply fall apart.”¹⁶

In most of these cases, “trust” just appears to be equated with the absence of outright distrust. But if this is done categorically and regardless of the circumstances, trust is turned into a trivial element of almost all interaction, indeed of many cases where people do nothing more than keep out of each other’s way. I have not formed any specific judgment (good or bad) about most of the individual men and women who constitute humankind. Should we now say I trust them all since I go unarmed among them? But the fact remains that people make a distinction between different cases of human co-ordination. More remains to be said about what is expressed when such a relation is described as one of trust.

Suppose I go to the library and do not think of my fellow library users at all, and suppose that this library is not known to be badly crime infested. And suppose just nothing out of the ordinary happens. In what way will it be helpful now to say I trusted the other patrons, as opposed to saying I was hardly aware of them at all? Perhaps the suggestion is that my ordinary life in a society will involve generalized trust in most people around me. But again, we would need an illuminating description of the difference between saying this, and just saying that I typically do not *distrust* other people unless I have some reason to.

Here someone might think that the definition of trust as a cooperative disposition should be supplemented by additional criteria. For instance, Diego Gambetta suggests that my favorable expectation from another qualifies as trust if the perceived probability of cooperation is >0.50 .¹⁷ One difficulty with such specifications is that they risk ruling out some cases that might plausibly be described as ones of trust, while still including implausible ones. The quoted definition would rule out the (surely numerous) cases where we are either completely unaware of our trust or have no precise idea of the probabilities. On the other hand, it is applicable to a manipulative attitude that might more properly be described as suspicious.

However, there is a more fundamental objection. It appears to us that, in looking for additional criteria for a proper definition of 'trust', one inevitably runs into problems similar to the ones indicated in Section 1.A. The definition of trust as a psychological state and the definition of trust in terms of expectations involve the same problematic assumption: namely, the idea that the presence of the relevant state, disposition, or behavior could be established neutrally, regardless of context.

At this juncture, our suggestion is then not, say, "Trust is the expectation of good will *plus* something else." What we need is not a more precise classification of various types of human attitudes. We rather need to see *why* and *when*, in human interaction, relationships are described as ones of trust. Thus the focus is moved from the classification of different cases to an inquiry into the way the concept itself enters our lives.

For instance, it seems to us that the question to be asked is not, "Do we, or do we not, trust our fellow members of society, our fellow library users, and so on?" but instead, "What (if anything) are we saying if we answer this question one way or the other?" In other words, what can be achieved by such talk; in what kinds of contexts may it be illuminating to talk in these ways about our relations to other people?

To suggest a possible case, a generalized claim about trust in society might have a point in making a contrast between different societies, say, one in which people always lock their front doors, and one in which they do not. *If such talk is intelligible, it will be so because a specific contrast is invoked.*

Trust is a pattern in the weave of life. To call something a pattern is to see it under the aspect of meaningfulness and purpose, not as a haphazard combination of elements. We propose the question: what are we doing when we recognize such a pattern and point it out to others?

2. The Use of 'Trust'

A. Trust and Betrayal

The difficulty of giving general criteria for the state, belief, disposition, or attitude that might be involved in trusting is importantly connected with one of Wittgenstein's central methodological insights. His discussion of the meaning of psychological expressions is closely related to his general critique of what is known as the Augustinian picture of language, presented in the opening pages of the *Investigations*: the picture where language at bottom consists of names that have referents. Wittgenstein wants us to get away from the idea that the most basic use of language consists in simply

reporting how things are. We are encouraged to regard language as a set of tools, used in a variety of ways to cope with different situations.¹⁸

Clearly a similar point holds for both mental and behavioral phenomena. The roles of what in some respect may count as instances of the *same* attitude or behavior may be quite different depending on the situation. This is an instance of Wittgenstein's insight that, if we want to get clear about the meaning of a word, it is of little help simply to focus on some entity that the word stands for. In many cases there is no such common entity, and even when there is, there remains the question under what aspect the entity is being referred to when the word is used. Instead, we should look at the role of the word in human interaction.

Then how is the word "trust" used? In what kinds of situation do we intelligibly ascribe trust to others or to ourselves?

Probably the most natural situation where we speak of trusting is one where someone's trust has been betrayed, or perhaps kept in the face of severe temptation. Thus the ideas of trust and *betrayal* belong together.

This is certainly not to say that trust, at bottom, includes a deposit of suspicion. Yet the point is that, in order for an attribution of trust to be intelligible, one must be able to invoke some imaginable disappointment. The disappointment may be due to active betrayal, but also to incompetence or thoughtlessness. (In many cases, incompetence and thoughtlessness will also involve betrayal. This is true when a person has accepted a task though he should have realized he was not qualified to perform it.)

Thus by saying that I trust my friend I vouch for him against a possible suspicion. This is not simply a neutral description of my thinking or behavior. Instead, I am *justifying* my thinking and behavior (as well as *his*) against a suggestion of disappointment. The issue, then, is not just that of finding a fitting description of the situation. By choosing to describe the situation in a certain way I may in fact change it.

Suppose I invite a good friend for dinner. Do I also trust that he is not going to pocket the family silver when I am not looking? The reader might reply, "Of course!" But if you were to ask this question in a real life situation, I would probably not answer "Of course!" but, "What do you mean?" In other words, what makes you ask this silly question? Is there something I should know? To say I trust my friend not to steal from me is to imply that he *might* do it.

On the surface, this point is reminiscent of Baier's influential definition of trust. She describes trust as "accepted vulnerability":

When I trust another, I depend on her good will toward me. [...] Where one depends on another's good will, one is necessarily vulnerable to the limits of that good will. One leaves others an opportunity to harm one when one trusts, and also shows confidence that they will not take it. [...] Trust then, on this first approximation, is *accepted vulnerability to another's possible but not expected ill will* (or lack of good will) toward one.¹⁹

Trust, then, according to Baier is, "awareness of risk along with confidence that it is a good risk."²⁰

What is right in this description is the emphasis on the conceptual relation between trust and the risk of betrayal. On the other hand, it involves a troubling ambiguity. It is not clear from whose perspective Baier is viewing the situation.

If I trust my friend, from my own point of view I have not made myself vulnerable by inviting him. I am not leaving my friend an opportunity to harm me, any more than I would if my friend was a locksmith and I failed to have safety locks installed to keep him out. To speak of opportunity here implies a certain idea of what he wants to do. And indeed, if in the end he does steal from me, from my reaction it will be obvious that I do not regard his behavior as something I had *accepted*.

On the other hand, the police officer who investigates the case will note that I turned my back and gave my (ex-) friend the chance to line his pockets with valuables. This brings up the fact that in very many cases it is such a divergence between perspectives (between those of different individuals, or those of one individual at different points in time) that provides a room for talk about trust.

We shall return to this point shortly, but first we need to address an obvious objection.

Granted that I do not *think* of my friend's behavior in terms of vulnerability, might one not nevertheless say that I *objectively* make myself vulnerable to harm? After all, I am not taking measures to prevent him.

In answer to this, one needs to point out that risks, probabilities, and possibilities are not entities that exist out there on their own. They are, instead, something we invoke in the context of our practical reasoning. We hope to get things done (or undone), we fear other things, and we realize that certain obstacles and prohibitions stand in the way of our action. Possibilities are not invoked haphazardly but in the light of human aspirations and expectations. To speak of some risk as "objective" is to claim, for instance, that I ought to consider a scenario to which I have not paid sufficient attention.

In one sense it will be 'possible' for my friend to steal from me, just as it will be 'possible' for him, in the abstract, to do a number of other completely unexpected things. Saying "Joe stole from his own friend" will not involve a logical problem (the way that saying "Joe stole from himself" might), nor is there anything physically to stop him. But clearly, no one could challenge a person's trustworthiness merely by bringing up considerations of what is logically conceivable or physically feasible. Any such discussion could be relevant in the first place only if there were reasons for suspicion.

The fact that a scenario is imaginable in *some* sense, then, does not necessarily make it possible in the sense relevant to the present example. The meaningfulness of speaking of vulnerability depends, not on what might be construed as imaginable in some abstract sense, but on what is the point that might be made by invoking a scenario in a context.²¹ Thus in a sense we do not distrust someone *because* we consider certain things possible for him. On the contrary, the fact that we consider certain things possible for him is *an expression of* our distrust.

By saying that it is possible for my friend to steal from me we would normally be voicing our suspicion that he might want to do it. To say

that it is possible for him *not* to do it sounds even odder. This last point may sound surprising since, according to a widely held philosophical view, whatever is the case must by definition also be possible. But possibilities are not objective entities that can meaningfully be said to exist regardless of our reasons for invoking them.

To describe someone's attitude as trusting implies a risk of betrayal. But unless my friend actually tries to steal from me, there will be no obvious reason why I should agree that there ever was a risk. (On the other hand, a spectator may, with or without good reason, harbor suspicions against my friend that are alien to *me*.) For these reasons, it would typically not be meaningful for me to say, out of the blue, that I trust my friend not to pocket valuables from the house. I could only say it meaningfully as a reply to what I can recognize as an intelligible expression of suspicion.

I would find it silly to say, of any of my friends individually, that I trust them not to steal from me. But on some occasions, perhaps, there still is something to be said for the general claim that, as a rule, we trust our friends not to do such things. It may, for instance, be a way to admonish someone who has tried to steal from a friend. The meaningfulness of such descriptions will depend on how credible they are in establishing the new perspective.

Thus by invoking the language of trust and betrayal, we do not simply identify facts, possibilities, or risks that exist out there. Instead we take up a certain perspective. In the sequel, it will be argued that this perspective is an *ethical* one.

B. Trust and Risk Taking

In much the same way as the connected word "betrayal", our talk of trust invokes an ethical perspective on human action. To describe a relation as one of trust is already to claim that the breach of that relation would constitute betrayal.

However, we are struck by the almost programmatic avoidance of that very perspective in current literature. There the assumption is, instead, that philosophers should define "trust" in an ethically neutral way and perhaps *subsequently* address the question whether it might be wrong to betray someone's trust.

In that literature, trust is usually described as a form of risk taking or risk management. Thus, according to Gambetta,

[w]hen we say we trust someone or that someone is trustworthy, we implicitly mean that the probability that he will perform an action that is beneficial or at least not detrimental to us is high enough for us to consider engaging in some form of cooperation with him.²²

Baier's notion of accepted vulnerability is similar to this view. Trudy Govier's account is also closely related:

Trusting another, we are willing to go ahead without a guarantee. We feel that we can rely or depend on the other, even though there

is always some possibility that he or she will act in unexpected ways, or even betray us.²³

Given the risks connected with “go[ing] ahead without a guarantee” the reader may ask why we should trust others at all. A conclusion might be: if we are to trust, the rationally required amount of good grounds should, in each case, be proportionate to the value of what is at risk. Govier indeed reaches this conclusion, only adding that the calculus should “in some cases” be tempered by “ethical or prudential considerations”, which are somehow superimposed.²⁴ Trust can be slight, moderate, or complete.²⁵ For instance, “[t]o accept a man's help carrying packages across a busy street, a woman needs to trust him, but slight trust will be enough—unless the packages contain exceedingly valuable items.”²⁶

Thus it is a curious fact that several writers, in trying to give an account of trust, seemingly end up describing what is more correctly characterized as a kind of cheerful suspiciousness. On one natural way of reading them at least, important cases of real, genuine trust simply fall out of the picture. (D. Z. Phillips has noted a similar tendency among philosophers of religion trying to give an account of religious trust.²⁷) As was argued in the previous section, trust is typically characterized by the fact that we *do not consider* the possibility that we might be let down. In contrast, the definitions quoted above would be applicable to a sort of cynical calculus characteristic of coercive or manipulative relations. These are relations where trust is manifestly missing.

This criticism applies to much of the theoretical work on trust carried out within Game Theory but also, in less formal terms, to other contributions to what may safely be described as the current mainstream.²⁸ According to Gambetta, to quote an extreme example, a slave-owner's *trust* in his slaves may simply amount to trust in the fact that the slaves are not going to rob him of their work force by committing mass suicide.²⁹ This description strikes us as nothing short of (an inadvertent) *reductio ad absurdum* of the position that many authors are advancing.

C. Two Approaches to Human Interaction

What attracts philosophers to such paradoxical views? One possible explanation is that there is a clash here between two ways of adjudicating what is at stake. There are two ways to address the question what should count as *justified* trust.

First, beliefs may be evaluated from an instrumental point of view. On this view they guide a person in her attempts to cope with reality. Someone's holding a specific belief is a good thing for her if it helps her achieve her goals and stay unharmed. The beliefs she holds about other human beings can also be considered from this perspective. In this perspective, other people constitute one subset of the various entities that she has to deal with in her efforts to get by in the world; entities that are, to be sure, particularly intricate and that may have an uncommonly powerful influence on her chances of success.

When someone's belief in the favorable intentions of another is being considered from this perspective, this belief may or may not be a good thing. Whether it is depends on the outcome in each particular case.

What a person expects from others is here simply seen as a matter of the projections she makes on the basis of their past behavior. For this reason, any blame for possible disappointments lies squarely with the believer herself. If her erroneous prediction about someone's behavior was one that could have been avoided, it shows deficient judgment on her part.

This means that, from this perspective, no room exists for a notion of trust as an ethical relation, irreducibly involving two individuals.

When regarded under the aspect of trust, however, disappointments may be judged differently. This is connected with the internal relation between trust and betrayal. If A's relation to B is rightfully to be called a case of trust (*that is, if A had a right to trust B*) this means that certain ways of behaving on the part of B will constitute a betrayal. If A is disappointed in B, we *may* judge that B had betrayed her trust, thus implying that B, *not* A, is the one to blame.

The question whether what B did was indeed a betrayal of trust can only be answered by considering the relation between A and B. For instance, we may need to ask whether B was A's parent, teacher or friend, whether through promises or through his past behavior he had permitted A to count on him, or, for instance, whether what A was counting on was something that anyone could count on from another in the circumstances. Thus, people would normally take it to be self-evident that a stranger offering directions will not willfully mislead, or that a passer-by helping you climb out of a well will not suddenly let go. (It should be clear that for something to constitute a betrayal does not depend on there having been any explicit undertaking.)

My faith that someone will not betray me goes beyond expecting him to do or not to do certain *specific* things; it involves resting assured that he will be mindful of my wellbeing. Trust, in this sense, is an open-ended relation between two individuals.

On the other hand, the present analysis will not apply to cases where a person's invocation of trust is considered groundless. If a kidnapper says she "trusts" the abducted child's parents not to contact the police after its release, the kidnapper will not be in a position to accuse them of betrayal if they do. She has no right to expect them to keep their promise.

In the end, what would be called for in adjudicating the issue are neither epistemological nor strategic considerations, but a moral judgment. If B betrays A's trust, B will not be absolved by the argument that A should have seen it coming, any more than the fact that I have been a scoundrel in the past gives me a license to go on being a scoundrel. From this perspective, the question whether a person's trust was misplaced is primarily a judgment *about the person trusted* and not about the person having the trust.

Here someone might object: this may be true as far as it goes, but it does not change the fact that, if A had been less trusting, things would not have turned out as badly as they did. However, for A to be less trusting would simply mean that A did not trust B. But the fact is that she did. For her, there *was* no room for plausible suspicion: that is what it means to trust someone. True, it is probably unwise to rely on people whom we know to be untrustworthy. But the fact that someone's trust is betrayed will not as such show that she acted unwisely, given the extent of her knowledge *at the time*.

There is simply no account of rationality that guarantees immunity against misfortune and ill will. The question here is, instead, how we should judge B's performance *given* the fact that A trusted him.

Out of these two approaches, mainstream theorists invariably adopt the instrumental perspective. This may be surprising, as it looks unpromising as a way to approach moral relations. However, it is in accordance with a rationalist self-understanding where philosophy is perceived as a disinterested and in some sense scientific pursuit. This perspective gives the impression of offering a neutral vantage point from which to assess human behavior.

In contrast, accounts given from the perspective of trust may give rise to the suspicion that normative views are being proposed under the guise of conceptual analysis. It looks as if they committed us to the view that trusting is *a priori* always a good thing. Thus, both Baier and Govier have reacted to what they perceive as a Wittgensteinian normative argument in favor of less reflection and more trust.³⁰ In contrast with what she sees as idealizing descriptions, Baier claims that “[o]nly if we had *reason to believe* that the most familiar types of trust relationships were morally sound would breaking trust be any more *prima facie* wrong than breaking silence.”³¹

However, this is to misconstrue the analysis just proposed. As is clear from our example (where A is betrayed by B), we are not proposing that trust can never be misplaced. Nor are we claiming that each and every relation that someone, for whatever reason, chooses to call trust is beyond moral criticism. People simply do distinguish between trust on the one hand and, for instance, naïveté or manipulation on the other; and by *calling* a person's expectations from another “trust” the speaker commits herself to the view that to let that person down would constitute a betrayal. Thus, normativity is not brought in by philosophical theory; it is already there when someone chooses to describe a relationship as one of trust.

Baier has suggested that Olli Lagerspetz's account of trust³² is expressive of a cultural bias, due to his lack of exposure to life in large American cities³³, and Govier has made an analogous suggestion invoking Lars Hertzberg's North European value background³⁴. But surely it must be possible to judge a philosophical attempt to clarify what it means to speak about trust independently of the writer's cultural environment. Whether to recommend that people on the whole be more trustful or more suspicious is, we suppose, a matter of world-view. This is bound to reflect one's experiences in life. (Although it is not obvious what conclusions one *must* be drawing from experience. Suppose someone refuses to give up on people despite having been deceived time and again. Some of us might consider her stupid or self-deceived, while others would find her attitude an admirable sign of strength. Both responses, we would argue, may be intelligible.) But such recommendations should be kept separate from the attempt to get clear about the sense of the remarks that people may make about trust. Calling something a betrayal of trust is a negative remark, whether it is made in Åbo, Watts, or in the streets of Baghdad.

3. Certainty and Evidence

A. Introduction

In her essay, “Trust as an Affective Attitude,” Karen Jones describes trust as an attitude of optimism about the goodwill and competence of the other.³⁵ Her description agrees with other mainstream accounts in representing trust as an essentially risky undertaking that stands in need of justification. In the present context, however, what interests us is her emphasis on the relation between trust and the notion of evidence. She writes:

Trust restricts the interpretations we will consider as possibly applying to the words and actions of another. When we can—and sometimes even if doing so requires ingenuity—we will give such words and actions a favorable interpretation as consistent with the goodwill of the other. Trusting thus functions analogously to blinkered vision: it shields from view a whole range of interpretations about the motives of another and restricts the inferences we will make about the likely actions of another. Trusting thus opens one up to harm, for it gives rise to selective interpretation, which means that one may be fooled, that the truth might lie, as it were, outside one’s gaze.³⁶

As the one trusted “is viewed through the affective lens of trust”, the person who trusts him (the victim of this distortion) will be prepared to risk dependence on him “often on the basis of the smallest evidence.”³⁷ Given this characterization of trust as a distortion of available data, the author maintains that we need to find some justification for the fact that we trust others at all.

However, it is not obvious that such a need exists. Jones’s description of trust as a case of blinkered vision will primarily be applicable to cases of ill-advised trusting (bad judgment, naïveté, self-deception, recklessness, or the like). But these are not cases that anyone should wish to justify. On the other hand, the description is not going to cut ice with a person who does not already think there is a problem about her trust. She will simply not think of herself as overly optimistic about her trusted friend. If her friend is trustworthy, trust will be a matter of realism, not of optimism. Distrust, on the other hand, will come out as pessimism; or perhaps, as optimism regarding one’s superior skills in detecting falsity. Suppose the cited passage is travestied as follows:

Distrust restricts the interpretations we will consider as possibly applying to the words and actions of another. When we can—and sometimes even if doing so requires ingenuity—we will give such words and actions *an unfavorable* interpretation as consistent with the *ill will* of the other. *Distrust* thus functions analogously to blinkered vision: it shields from view a whole range of interpretations about the motives of another and restricts the inferences we will make about the likely actions of another. *Distrust* thus opens one up to harm, for it gives rise to selective interpretation, which means that one may be fooled, that the truth might lie, as it were, outside one’s gaze.

This description strikes us as just as applicable as the previous one. Generally speaking, it is hardly reasonable to suggest that anyone should try to settle, in the abstract and regardless of the situation at hand, whether trust or distrust is the more rational attitude.

However, Jones's description does highlight the intimate relation between the notions of reasoning and evidence on the one hand, and the notions of trust and distrust on the other. Only her account, as it were, turns the relation upside down.

Jones offers us an account where the starting point is a recognizable standard of good reasoning based on sound evidence. Trust enters the picture as a disturbing element. (Jones concedes that distrust too sometimes disturbs reasoning in an analogous way.³⁸) Thus in this picture, the fully rational person attends to *the* evidence (which is already there) and, uninfluenced by others, takes it for what it is worth, neither more nor less.

In contrast, we would maintain that the ability to trust is *constitutive of there being any such thing* as valid reasoning and sound evidence in the first place.

B. "The Substratum of All my Enquiring and Asserting"

This theme is something of a *Leitmotif* in Wittgenstein's *On Certainty*. That work starts off with a discussion of G. E. Moore's attempted refutations of skepticism.³⁹ However, Wittgenstein does not really focus on skepticism but, instead, on the remarkable *fact* that we would find it impossible to doubt certain statements about the world and our own lives in it. I cannot seriously doubt that the world has existed for a long time before my birth or that all human beings have intestines. I cannot doubt that I have two hands or that I have never been to China.

These statements do not qualify as logical truths in a conventional sense, yet they are not straightforwardly empirical either. As Moore, too, points out, in cases of this kind I could not point to specific evidence that has settled the question for me.⁴⁰ I could not cite a letter by Napoleon to prove that the world is more than two hundred years old. I can take the letter to be authentic only if I already accept that the world existed at the time. "The belief that the world is more than two hundred years old" is then not a historical belief in any normal sense. If there are to be historical inquiries at all it is, instead, taken for granted that we do not doubt such things.⁴¹ As Wittgenstein notes, "I have a world-picture. Is it true or false? Above all it is the substratum of all my enquiring and asserting."⁴² He adds, "I say world-picture and not hypothesis, because it is the matter-of-course foundation for [the scientist's] research and as such goes unmentioned."⁴³

An analogous point is possibly developed in R. G. Collingwood's *Essay on Metaphysics*, where he puts forward a conception of metaphysics as the study of the "absolute presuppositions" of thought.⁴⁴

According to Wittgenstein, my convictions form "a system, a structure."⁴⁵ The system "is not so much the point of departure, as the element in which arguments have their life."⁴⁶ I do not explicitly learn such convictions, but I can later recognize that they are implicit in the things I do learn, "like the axis around which a body rotates." Like the axis of a globe floating in space, it does not support the body but instead "the movement around it determines its immobility."⁴⁷

The upshot is that it is misleading to think that we might find evidence for something completely obvious. By “evidence”, we mean facts that rightly incline us to settle *an open question* one way or another. This means, for instance, that a historian today could not seriously treat newspapers from 1944 as evidence for the fact that the Second World War took place. It would be easy enough to locate written material that implies there was a great war at the time. But to call it *evidence* would imply that serious disagreement exists about the matter. For anyone with an ordinary Western education such historical facts will count as at least as obvious as is the idea that old newspapers could be employed as historical sources in the first place.

This is not to say that no one could ever doubt that the Second World War happened. What we cannot doubt today perhaps some future generation will.⁴⁸ (For related reasons, Collingwood describes the study of the absolute presuppositions of thinking as an essentially historical investigation, since those presuppositions will be different during different historical periods.) Nor is it to say that there is nothing the historian could say to the professed skeptic. But insofar as such a thing as historical inquiry will exist, some facts of this general kind will be commonly treated as unquestionable.

Our ability to distinguish between serious and spurious disagreement is something we acquire by being brought up in a culture. We generally take as true what we find in textbooks. If in doubt, we consult other written sources or ask an expert. Thus “I learned an enormous amount and accepted it on human authority, and *then* I found *some* things confirmed or disconfirmed by my own experience.”⁴⁹

The child begins its learning process by believing the adults around it.⁵⁰ Without such obvious trust, the child could never develop the critical faculties that make genuine doubt possible. Reasoned doubt will always presuppose trust in other directions. But trust in the testimony and sound judgment of others is not only characteristic of children. In adults, it is not a residual “childlike thoughtlessness, innocence, and powerlessness.”⁵¹ On the contrary, as we grow older we learn to make *more* use of the information and expertise available around us. Indeed our ability to distinguish between serious research and nonsense is based on our ability to trust some sources. Our intellectual life as a whole is characterized by what one may call an epistemological division of labor.

In sum, our faith in the judgment and testimonies of others “is not, so to speak, hasty but excusable: it is *part of judging*.”⁵² While trust may involve intellectual distortion in some cases, the general fact that we tend to trust others is, on the contrary, constitutive of sound judgment. It is not a weakness to be excused or justified.

On the whole, *On Certainty*, like the *Investigations*, shows Wittgenstein as a ‘social’ thinker. Both works describe meaningful thought as an activity that arises and makes sense in the context of human interaction and interdependence.

C. The Pathology of Distrust

We see that there is an important asymmetry between the place of trust and distrust in human growth. No critical faculties will develop unless the child

starts by *not doubting*. (Analogously, there is an asymmetry between pretense and genuine communication. Wittgenstein points out that a child must learn a great deal before it can pretend, presumably including the use of the corresponding genuine expression.⁵³)

Does this imply that normal children trust their parents from the very start? Or only that they do not distrust them? Both answers might be right depending on the context of the question. As argued earlier, the point of using the language of trust is not primarily to arrive at an accurate classification of behavior or mental states but, instead, to invoke an ethical perspective. To say that a child trusts its parents is, for instance, to admonish parents who ignore its needs or talk ironically to it. Our present point is that normal human lives will, from the very start, involve relations of mutual dependence that are not questioned.

In a small child, the opposite of trust is typically not called doubt, distrust or suspicion but, perhaps, fear. An infant who did not start life with a basically trustful attitude towards those around him (that is, an infant whose attitude is an overwhelmingly *fearful* one) would be considered abnormal. It is doubtful whether one could coherently describe such an infant as suspicious at all.

Something analogous is true of adults. If each time we go to a restaurant our friend brought along a chemistry set to make sure the food is safe to eat, or if she always looked under the table to see if it is bugged, then, unless she could give us specific reasons, we would not just describe her as unusually suspicious. We would soon start worrying about her sanity.⁵⁴ In contrast, we will *not* consider her pathologically careless or gullible just because she *never* checks for poison or bugs except in extraordinary circumstances.

We normally go about our daily business without giving much thought to the strangers around us. Yet as soon as the need for interaction arises, most of the time we will incline in the direction of trust unless we have grounds for acting otherwise. If we would not ask a stranger in the streets of New York or Moscow to keep an eye on our camera, it is because we are conscious of the ubiquity of graspingness in big city cultures. Nevertheless we might ask that same stranger to help us call for an ambulance.

On Certainty and *Philosophical Investigations* do not contain sustained discussion of the ethical character of our mutual dependence. This is perhaps part of Wittgenstein's general reluctance to address ethical questions in his properly philosophical work.

However, some such discussion is included in the notebook remarks that appear in the posthumous collection *Culture and Value*. Wittgenstein comments on the problems of distrust. He points to the kinship between suspiciousness and insanity:

Madness doesn't *have* to be regarded as an illness. Why not as a sudden—more or *less* sudden—change of character?

Everybody is (or most are) mistrustful, & perhaps more so towards their relations, than towards others. Is there any reason for mistrust? Yes & no. Reasons can be given for it, but they are not compelling. Why shouldn't someone suddenly become *much* more

mistrustful of people? Why not *much* more withdrawn? or devoid of love? Don't people get like this even in the ordinary course of events?⁵⁵

In an earlier notebook, Wittgenstein writes about our inclination to shut ourselves off from others. He describes it as a moral shortcoming:

Someone who ... opens his heart to God in remorseful confession opens it for others too. He thereby loses his dignity as someone special [alternative translation: outstanding] & so becomes like a child. That means without office, dignity & aloofness from others. You can open yourself to others only out of a particular kind of love. Which acknowledges as it were that we are all wicked children.

It might be said: hate between human beings comes from our cutting ourselves off from each other. Because we don't want anyone else to see inside us, since it's not a pretty sight in there.

Of course you must continue to feel ashamed of what's within you, but not ashamed of yourself before your fellow human beings.⁵⁶

Shutting oneself off from others, and, accordingly, hate between human beings, is closely connected with lack of trust. On Wittgenstein's view, then, the significance of trust did not simply lie in its role in human thought and judgment, but many forms of human conflict had their source in, or were aggravated by, a lack of trust.

We have earlier referred to the grammatical connection between trust and betrayal. This connection highlights the responsibility of anyone rightfully considered an object of trust. In his notebooks, Wittgenstein draws attention to the other side of the relation. He suggests that there is an obligation to trust, or, rather, not to be distrustful. However, he is not offering this as the result of a grammatical investigation. It is expressive of Wittgenstein's understanding of his own life, and of what is important in human affairs.

4. Summary

We have discussed three themes in Wittgenstein's work that are particularly relevant for understanding the concept of trust.

(1) Wittgenstein's discussion of language and of psychological concepts in particular suggests that psychological expressions (such as "trust") are not primarily employed to refer to independently identifiable psychological states. This concept should instead be seen as part of an interaction where moral relations come into play. In particular, it is connected with the attribution of blame. Thus we need not expect the presence of a specific state of mind or a specific behavior pattern every time the word "trust" is applicable. Conversely, the fact that the word is truthfully used to describe a given case of human interaction does not imply that all similar cases may be so described regardless of the moral relations involved. (The main relations to consider are that between the truster and the trustee as well as that of the observer to them both.) This implies, in particular, a

criticism of the idea (by Luhmann, Baier, and others) that a generalized form of trust is necessary in order for social life to be possible.

(2) If philosophers assume that our grasp on reality is ultimately to be spelled out in terms of factual assertions, they will come to see trust as a matter of holding certain beliefs about the person trusted; beliefs that, in turn, will be taken to justify our confidence in what the person tells us. This widely held view has largely come to shape current accounts of trust. No room exists in them for the idea of genuine trust as a moral relation irreducibly involving two individuals. However, according to Wittgenstein, the starting point for this line of thought needs to be drawn into question. The sense of an assertion, he claimed, is dependent on the way it enters into a context of life. Hence the idea that our relation to reality is ultimately constituted by factual assertions is confused. This realization opens up for a non-reductive understanding of trust. In many cases the trust we have for another individual will be basic to the beliefs we come to form, not the other way round.

(3) In *On Certainty*, Wittgenstein is exploring the fact that we take some ways of thinking as self-evident without asking for further evidence. Skepticism is only possible against the background of massive general agreement. A certain dependence on knowledge claims and judgments by others is not only a practical necessity but indeed part of what it means to make reasoned judgments. Our general situation may be described as logical and epistemic division of labor. It is not illuminating to describe all instances of such division of labor as trust regardless of the contrasts one wants to make; nevertheless, ascriptions of trust as well as distrust make sense against this background of mutual dependence.

NOTES

1 Those who speak of trust as a feeling include Anthony Giddens, *Modernity and Self-Identity* (Cambridge: Polity Press, 1991), p. 36; Ann-Mari Sellaerberg, "On Modern Confidence," *Acta Sociologica* 25 (1982), pp. 39–48, see pp. 40, 45, 46; Trudy Govier, "An Epistemology of Trust," *International Journal of Moral and Social Studies* 8 (1993), pp. 155–174, see p. 156.

2 Annette Baier, *Moral Prejudices* (Cambridge, Mass.: Harvard University Press, 1994), p. 132.

3 Ludwig Wittgenstein, *Philosophical Investigations* (Oxford: Blackwell, 1953), I: § 596.

4 Ludwig Wittgenstein, *Zettel* (Berkeley: University of California Press, 1967), §§ 45, 78, 81. Also see §§ 46–47, 50, 76–77, 82–83, 85.

5 For a discussion of genuine duration, see Malcolm's contribution to D. M. Armstrong and Norman Malcolm, *Consciousness and Causality* (Oxford: Basil Blackwell, 1984), pp. 1–101; esp. pp. 79 ff.

6 Wittgenstein, *Zettel*, §§ 82, 45, 50.

7 *Ibid.*, § 50.

8 Wittgenstein, *Philosophical Investigations* I: § 583.

9 *Ibid.*, II: i, p. 174.

10 *Ibid.*, II: i, p. 174.

11 Annette Baier, "Trust and Antitrust," *Ethics* 96 (1986), pp. 231–260, see p. 235. Reprinted in Annette Baier, *Moral Prejudices* (Cambridge, Mass.: Harvard University Press, 1994), pp. 95–129; Diego Gambetta, "Can We Trust Trust?" in Diego Gambetta (ed.), *Trust. Making and Breaking Cooperative Relations* (Oxford: Blackwell, 1988), pp. 213–245, see p. 217; Trudy Govier, "Trust, Distrust, and Feminist Theory," *Hypatia* 7 (1992), pp. 16–33, see p. 17; Govier, "An Epistemology of Trust," p. 157; Karen Jones, "Trust as an Affective Attitude," in Clifford Williams (ed.), *Personal Virtues: Introductory Essays* (Basingstoke: Palgrave Macmillan, 2005), pp. 253–257, see p. 253.

12 Baier, "Trust and Antitrust," p. 234.

13 *Ibid.*, p. 234.

14 Baier, *Moral Prejudices*, p. 159.

15 Niklas Luhmann, *Trust and Power* (Chichester: Wiley, 1979), p. 25.

16 Lorraine Code, "Second Persons," in Marsha Hanen & Kai Nielsen (eds.), *Science, Morality, and Feminist Theory* (Calgary: University of Calgary Press, 1987), p. 377. Also see Govier, "Trust, Distrust, and Feminist Theory."

17 Gambetta, "Can We Trust Trust?" p. 218.

18 Wittgenstein, *Philosophical Investigations*, I: §§ 11–12, 14, 17, 23.

19 Baier, "Trust and Antitrust," p. 235. Emphasis added.

20 *Ibid.*, p. 236.

21 For a discussion of the notion of what can be imagined, see Lars Hertzberg, "Imagination and the Sense of Identity," in his *The Limits of Experience* (Helsinki: Acta Philosophica Fennica), 1994, pp. 96–112.

22 Gambetta, "Can We Trust Trust?" p. 217.

23 Trudy Govier, *Social Trust and Human Communities* (Montreal & Kingston: McGill-Queens University Press, 1997), p. 4.

24 Govier, "An Epistemology of Trust," pp. 167–168.

25 *Ibid.*, p. 157.

26 *Ibid.*, p. 167.

27 D. Z. Phillips, "On Trusting Intellectuals on Trust," *Philosophical Investigations* 25 (2002), pp. 33–53.

28 A position similar to Gambetta's is explicitly stated in the following contributions included in Gambetta (ed.), *Trust. Making and Breaking Cooperative Relations*. Partha Dasgupta, "Trust as a Commodity," pp. 49–72, see p. 51; David Good, "Individuals, Interpersonal Relations, and Trust," pp. 31–48, see p. 33; Keith Hart, "Kinship, Contract, and Trust," pp. 176–193, see pp. 186–187; Anthony Pagden, "Trust in Eighteenth-century Naples," pp. 126–141, see p. 129; Bernard Williams, "Formal Structures and Social Reality," pp. 3–13, see p. 8. Also see Govier, "An Epistemology of Trust"; Peter Johnson, *Frames of Deceit. A Study of the Loss and Recovery of Public and Private Trust* (Cambridge University Press, 1993), p. 15.

29 Gambetta, "Can We Trust Trust?" p. 219.

30 Annette Baier, "Reply to Olli Lagerspetz," in Lilli Alanen, Sara Heinämaa & Thomas Wallgren (eds.), *Commonality and Particularity in Ethics* (Basingstoke: Macmillan, 1997), pp. 118–122, see p. 121; Trudy Govier, "Trust and Testimony: Nine Arguments on Testimonial Knowledge," *International Journal of Moral and Social Studies* 8 (1993), pp. 21–39, see pp. 24–25, 32–33.

31 Baier, "Trust and Antitrust," p. 253. Emphasis added.

32 Olli Lagerspetz, "The Notion of Trust in Philosophical Psychology," in Lilli Alanen, Sara Heinämaa & Thomas Wallgren (eds.), *Commonality and Particularity in Ethics* (Basingstoke: Macmillan, 1997), pp. 95–117. Also see Olli Lagerspetz, *Trust: The Tacit Demand* (Dordrecht: Kluwer, 1998).

33 Baier, "Reply to Olli Lagerspetz," p. 121.

- 34 Govier, "An Epistemology of Trust," p. 172. Govier is referring to Lars Hertzberg, "On the Attitude of Trust," *Inquiry* 31 (1988), pp. 307–322. Reprinted in Hertzberg, *The Limits of Experience*, pp. 113–130.
- 35 Jones, "Trust as an Affective Attitude," p. 253 and *passim*.
- 36 *Ibid.*, p. 262.
- 37 *Ibid.*, p. 263.
- 38 *Ibid.*, p. 267.
- 39 G. E. Moore, "A Defence of Common Sense," in J. H. Muirhead (ed.), *Contemporary British Philosophy, 2nd Series* (London: George Allen & Unwin, 1925), pp. 192–233; "Proof of an External World," *Proceedings of the British Academy* 25 (1939), pp. 273–300. Both essays are reprinted in G. E. Moore, *Selected Writings*, ed. Thomas Baldwin (London: Routledge, 1993), pp. 106–133, 147–170.
- 40 Moore, "Defence...", *Selected Writings*, p. 118.
- 41 Ludwig Wittgenstein, *On Certainty* (New York: Harper & Row, 1972), § 188. Also see §§ 163, 182–192.
- 42 *Ibid.*, § 162.
- 43 *Ibid.*, § 167.
- 44 R. G. Collingwood, *An Essay on Metaphysics* (Oxford: Oxford University Press, 1979).
- 45 Wittgenstein, *On Certainty*, § 103.
- 46 *Ibid.*, § 105.
- 47 *Ibid.*, § 152.
- 48 Also see *ibid.*, §§ 96–99.
- 49 *Ibid.*, § 161. Emphases added.
- 50 *Ibid.*, §§ 160, 310.
- 51 Cf Baier, "Reply to Olli Lagerspetz," p. 121.
- 52 Wittgenstein, *On Certainty*, § 150. Emphasis added.
- 53 Wittgenstein, *Philosophical Investigations*, II: xi, pp. 228–229.
- 54 See Raimond Gaita, *Good and Evil: An Absolute Conception* (Basingstoke: Macmillan, 1991), p. 314.
- 55 Ludwig Wittgenstein, *Culture and Value*, revised ed. (Oxford: Basil Blackwell, 1998), p. 62e.
- 56 *Ibid.*, pp. 52e–53e .