

## Trust in Education

The philosophy of trust is a relatively small subfield. Nonetheless, it contains within it many important insights. Our contention in this paper is that careful study of this subfield can bring with it many insights that can and should be used to reconsider a variety of arguments that have been brought forward in the literature on the philosophy of education and other areas of educational research. Our focus is largely destructive in section 2 of the paper. We object to two recent papers from the philosophy of education literature arguing that they fail to recognise important insights from the philosophical literature on trust. In section 3 we say something more positive, noting the routes forward.

### 1. What do philosophers say about trust?

Much work in the philosophy of trust has proceeded down three (interconnected) lines, two of which will detain us here. These are: teasing apart different, closely related concepts associated with trust; exploring the moral dimensions of these closely related concepts; providing analyses of these different concepts. We shall consider the first and third of these.

In the first vein, we find philosophers carefully distinguishing between a range of concepts. Relevant to us here will be the following. First, we have three-placed trust, such that ‘x trusts y to perform some action F’. An example here is a case whereby a teacher trusts a particular pupil to tell the truth. Second, we have the concept of trustworthiness, such that ‘x is trustworthy’. A teacher or pupil may, for instance, be trustworthy. Note, here, that we should be careful to distinguish being trustworthy—a trait of a particular individual—from the perception *that* they are trustworthy, had by another. Schematically, we should be careful to distinguish ‘x is trustworthy’ from ‘y believes x to be trustworthy’. From some points of view, the distinction is obvious: a con-artist is likely not trustworthy, but will trade upon the fact that they are perceived to be in order to execute the con. As obvious as this point may seem when made in this way, as we shall see later on it has important downstream consequences.

Further distinctions are important, too. For instance, we should be careful to distinguish the concept of *merely* relying on some agent to perform a task from the concept of trusting them to perform that task ((Baier 1986, 234; Hieronymi 2008, 215; Holton 1994, 2–3; Jones 1996, 14; Jones 2004, 4; McLeod 2006; O’Neill 2002, 15; Potter 2002, 3–4; Pettit 1995, 205)). The latter, but not the former, is likely to bring with it a moral dimension. For instance, and to borrow an example from Hawley (2014: 2):

Suppose you regularly bring too much lunch to work, and leave the excess for others to eat. Suppose you do this because you’re bad at judging quantities, not because you’re keen to feed your colleagues. I rely on you to provide my lunch: I anticipate that you will do so, and I don’t make alternative arrangements. But this reliance should not amount to trust: you would owe me no apology if you ate all the food yourself, and I ought not to feel betrayed by this, even if I felt disappointed (and hungry).

Thus, we have trust (a rich, morally loaded notion), and we have mere reliance (which lacks the moral dimension—though, to be sure, we may sometimes *speak* of this in terms of trust, as in ‘we

trust the shelf to hold the vase'). We also have what we might call a generalised 'attitude of trust', such that we might say that 'x trusts y'. For instance, perhaps the teacher can be said simply to 'trust a pupil'. Last, here, we should be careful to distinguish a lack of three-placed trust (/trustworthiness/general trust) from the presence of distrust. For, whilst it is certainly true that a teacher new to a class may not yet trust the children within it to (for instance) tidy away unbidden at the end of the day, it would be odd to say that the teacher thereby *distrusts* them to do so. Or, to give another, more frivolous example (again from Hawley 2014: 3):

My colleagues have never bought me champagne, so in particular I do not rely upon them to buy me champagne next Friday. But it would be wrong, even offensive, to say that I distrust my colleagues in this respect—after all, they have not offered to buy me champagne next Friday, and there is no social convention that they do so. If they did buy me champagne unexpectedly, I ought to be grateful, but I need not feel remorse about my earlier decision not to rely on them. Indeed, it would be bad manners for me to suggest in retrospect that I should have trusted my colleagues to buy me champagne, or to apologise for my earlier non-trust.

Much as trust must be earned, so too does distrust.

In addition to the teasing apart of these different concepts, philosophers working on trust have sought to try to analyse these concepts once they are teased apart. To give a feel for the kinds of analysis provided, consider the first, offered by Jones (2013: 70-1), of the concept of trustworthiness. Consider two agents, B and A, with some domain of interaction, D.

B is trustworthy with respect to A in domain of interaction D, if and only if she is competent with respect to that domain, and she would take the fact that A is counting on her, were A to do so in this domain, to be a compelling reason for acting as counted on.

To put some flesh on the bones here: a teacher might be trustworthy with respect to telling the truth to their pupils in the classroom; she would be such if and only if she is competent with respect to telling the truth to their pupils in the classroom, and would take the fact that the pupils are counting on her to tell the truth, were they to do so, to be a compelling reason to tell the truth in the classroom.

Two features of this analysis are worth bringing out. The first feature is that this analysis rightly surfaces the fact that, within the classroom setting, what matters is whether our putative teacher is trustworthy *in that setting*. A teacher, Hilary, may be entirely deceitful in their love-life and so utterly untrustworthy within the domain of matters of the heart; nonetheless, they may be perfectly trustworthy within the domain of classroom interactions. (Whether Hilary would be *perceived* to be trustworthy if their private life were made public is a different question.)

The second feature worth noting is that described as 'trustworthy' here are *agents*. Thus, if we were to take this analysis to be complete and total, we ought not to describe testimony, qualifications, or processes as trustworthy. At least, we might *describe* them as such, but we should

understand this as shorthand for speaking about their authors or the agents that create them. Thus, and to give but one illustrative example, what *Margret says* is not (properly speaking) trustworthy; rather, Margret is.

Both features of trustworthiness that are implicit in Jones' account, are ones that we think generalise to at least some degree. For instance, just like trustworthiness, we think that three-placed trust is obviously thought of as being rather narrow: if  $x$  trusts  $y$  to  $F$ , that tells us nothing about whether  $x$  trusts them to  $G$ . To illustrate: a teacher may trust a pupil to tell the truth, but not to look after the class rabbit. And, in addition, we do not think that the very general trust that we adverted to a moment ago, such that  $x$  simply trusts  $y$ , is present in very many cases at all. For, quite simply, to say that ' $x$  trusts  $y$ ' is *perfectly* general; it places no limit on the number of domains in which  $x$  trusts  $y$  to perform a given action. That being so, it would seem to follow that if  $x$  simply trusts  $y$ , then  $x$  trusts  $y$  to  $F$  in every conceivable domain. And this, we think, is wholly unlikely.

For, as is clear in most analyses of trusting  $y$  to  $F$ , this requires a belief that  $y$  has some commitment or other with respect to  $F$ -ing. Thus, and just to illustrate, Hawley (2014: 10) argues powerfully in favour of a view of trusting  $y$  to  $F$  along the following lines:

To trust someone to do something is to believe that she has a commitment to doing it, and to rely upon her to meet that commitment. To distrust someone to do something is to believe that she has a commitment to doing it, and yet not rely upon her to meet that commitment.

Assuming something like Hawley's account is correct (see Tallant (2017) for criticism and minor amendment), it then follows that it is highly unlikely that we will simply *trust* another, in full generality. Simply, and returning to our classroom case, the pupil will lack the full range of commitments required to be trusted in full generality.

**In taking this position, that trust is most typically to be understood as this three-placed trust ( $x$  trusting  $y$  to perform some action,  $F$ ), Hawley is following in a well established tradition. For instance, Baier (1986: 237), understands trust in terms of an agent  $A$  trusting another,  $B$ , with the care of something valued,  $C$ , and offers an account such that '[t]rust, on the analysis I have proposed, is letting other ...take care of something the truster cares about, where such "caring for" involves some exercise of discretionary powers' (1986: 40).**

Similarly, Faulkner (2007) also distinguishes trust from mere reliance (labelling the former as 'affective trust', the latter as 'predictive trust), giving an analysis of trust (in the affective sense), as:

**' $A$  trusts  $S$  to  $\varphi$  ... if and only if (1)  $A$  knowingly depends on  $S$   $\varphi$ -ing; (2)  $A$  expects  $S$ 's knowing that  $A$  depends on  $S$   $\varphi$ -ing to motivate  $S$  to  $\varphi$ ' (2007: 882)**

Indeed, this general approach of understanding trust as having the form of one agent trusting another to perform some action is extremely widespread. As Faulkner (2015: ) puts it:

**‘Almost without exception, philosophical discussion of trust focuses on the three-place trust predicate: X trusts Y to  $\varphi$  (see, for instance, Baier 1986; Faulkner 2007; Hawley 2014; Hieronymi 2008; Holton 1994; Jones 1996).’**

**It is similarly typical to distinguish rich, morally loaded trust, from mere reliance.**

We recognise, of course, that we may *speak* in ways that appear in tension with our argument here. That is, we might say to a pupil that we ‘trust them’. And in doing so we may signal that we don’t *merely* trust them to perform one specific action. That’s fine. We take it that people speak ‘fast and loose’ on a regular basis. In this particular case, we assume that what is being adverted to is the fact that the teacher trusts the pupil to do what is required of them *in most classroom cases*. But that, of course, is not a different concept; saying that we ‘trust someone’ in this sense, is just a shorthand way of indicating that, in most cases we can think of offhand, y is trusted to F.

Equally, we think it quite right, as implied by Jones’ analysis, that instances of *x being trustworthy* or *x being trusted to F*, are, properly speaking, to be understood as making commitments about *agents*. To be sure, there might be some reason to think that we occasionally include inanimate objects and technologies when we make these claims about trust and trustworthiness (see Nickel et al (2010); Tallant (2019)). Nonetheless, when we say that we ‘trust y to F’, or similar, it is the agents performing the actions to whom we should typically be taken to refer. And it is not the case that testimonies, or processes, or what have you, can, properly speaking, be treated as trusted to perform an action, or regarded as trustworthy. To see why to think this is correct, we simply point to the fact that the natural reactive attitudes to cases of a breach of trust include a feeling of moral betrayal. Quite simply: testimony cannot morally betray us—nor can processes, or similar. Only agents are properly viewed as the subject of moral approbation and so agents (and only agents) are the entities that should be described as trustworthy, or trusted to perform certain actions. That being so, testimony, processes, and the like, are not the sorts of thing that can be trusted, trustworthy, or distrusted.

**Of course, it’s very natural to suppose that other kinds of entity—most obviously, institutions (groups, collectives)—can also be trusted. We might naturally speak of trusting a bank to look after our money; trusting a company to ship goods on our behalf; trusting a school with the education of our children, and so on. How can these be accommodated within a focus on trust that treats trust as relating (moral) agents?**

Here, we think that one of two routes could be pursued. The first is to understand such institutions as kinds of collective agents. This might require of us that we treat these institutions as genuine entities and attribute to them, perhaps in addition to their members, genuine agency. Thus, we might say, a particular school is to be treated *as* an entity and, over and above the moral agency of its staff and pupils, the school *itself* has moral agency—and may, accordingly, be trusted or distrusted to perform particular actions. This would be a challenging road to the extent that it might require us to re-conceptualise our relationships with such collectives.

Instead, we prefer to follow aspects of Hawley's (2017) approach to trust in institutions and collectives generally. There, Hawley argues that we can make do without the notion of trust in groups and collectives. Key, we think, are Hawley's arguments in section 8 of her paper. As she notes, in cases where a group behaves in such a fashion that we might initially think of as constituting a breach of trust, our reactive moral attitudes tend to target specific individuals. Thus, if a company misbehaves, we seek to pin the blame on an individual, as when a CEO is forced to stand down in the event of a scandal. Similarly, if a school fails to educate its children, it will be particular figures within the school upon whom we look to lay the blame (perhaps the Head Teacher will be the subject of our ire; perhaps it will be the governing body, etc.). Ultimately, it is *their* fault that the children have not been educated properly. What both of these descriptions of trust in institutions have in common, however, is that they treat apparent instances of trust in institutions as instances of trust in specific individuals.<sup>1</sup> We'll follow Hawley's suggestion and adopt this approach to trust in institutions.

This concludes our quick sketch of the current state of play in the literature in the philosophy of trust. We now want to engage in a careful analysis of some recent work that has been carried out in the philosophy of education. We suggest that this understanding of trust and related concepts does much to undermine that recent work in the philosophy of education. We noted above, we then close with some suggested routes forward.

## 2. Arguments in the philosophy of education

In this section of the paper we explore two recent arguments that seek to explore the intersection of the philosophy of education and the philosophy of trust. We suggest that both fail, though they fail for different reasons. Our suggestion is that the first paper fails to properly tease apart and appreciate certain concepts associated with trust, leading to the conclusion that the arguments developed are unsound. The second paper, we suggest, though attempting to engage with the literature, fails to properly grasp what is at the core of certain key positions.

### 2.1 Webster

Webster (2018) marshals a number of arguments in order to show that: 'educators need to be trustworthy and that this involves more than the demonstrated evidence of appropriate qualifications and compliance to professional standards and codes of conduct' (2018: 161). In order to show that this follows, Webster reasons as follows.

We must realise that those who are to be taught must, to at least some degree, trust their teachers. As he puts it: 'being trustworthy, is considered here to be important for educators if society and especially students are to become our willing 'moral hostages' to undergo educative experiences for their own benefit' (2018: 156).

Broadly, we find what Webster says on this score plausible, though we note here that Webster seems to be unclear on what exact kind of trust is in play here. As Webster says: 'trust is invested

---

<sup>1</sup> Of course, it's consistent with this claim that we nonetheless *rely* on institutions, groups and other collectives. And, as above, we might speak in terms of trusting those things. Nonetheless, this is to be understood in terms of reliance, rather than the rich, morally loaded notion of trust.

when those who are being trusted have some freedom of discretion or choice. Trust has the hope or faith that trustees will use their freedom well to pursue the interests of the trusting' (2018: 156). As we noted above, fully general trust is almost never realised, so it cannot be *that* to which Webster is referring. That being so, we assume that they have in mind three-placed trust. That gives us that result that:

'trusting someone to F is invested [occurs] when those who are being trusted to F have some freedom or discretion or choice. Trusting someone to F has the hope or faith that the person trusted will use their freedom well to pursue the interests of the person trusting them to F.'

That seems to sit a little uneasily. It might be utterly against Smith's personal interests to have Jones correctly report them to the authorities for tax evasion, though it is still entirely plausible that Smith trusts Jones 'to do the right thing', and, in doing so, to report them to the authorities. The simple point being that trusting someone to F need not be in the least connected to their pursuing our interests.

In any case, this is only a small and in-passing concern. Our more pressing concern is to do with what Webster goes on to say about trustworthiness, and why this cannot be the product merely of the demonstrated evidence of appropriate qualifications and compliance to professional standards and codes of conduct' (2018: 161). As Webster puts it:

Trustworthiness needs to be 'won' or 'earned' and requires an ethical relation with the trustee beyond what any confidence in qualifications or competencies can provide (2018: 156)

But this, it seems to us, is a simple confusion. Whether or not x is trustworthy is a fact that is independent of any winning or earning that may occur. Rather, *being worthy of trust* is a state that an individual can occupy. To pivot back to Jones' account:

B is trustworthy with respect to A in domain of interaction D, if and only if she is competent with respect to that domain, and she would take the fact that A is counting on her, were A to do so in this domain, to be a compelling reason for acting as counted on.

To say, as Webster does, that one can 'win' or 'earn' this kind of property simply reflects a conceptual confusion. Either x is competent and would behave in a given way if called upon, or they would not. There is no 'earning', here. But, that trustworthiness requires 'winning' or 'earning', and that it requires an 'ethical relationship' with the trustee that goes beyond mere confidence in qualifications, is, so far as we can see, all that Webster provides us with by way of argument that what is required in the educational context is not the mere generation of confidence, but trustworthiness.

Thus, although it may turn out that Webster is correct (perhaps confidence and the different forms of trust are very different from one another (we suspect so, see Smith (2005), Earl (2016)); perhaps it really is the case that trust, especially trustworthiness, is important within the

educational context), no argument that turns on such a misconception will prove useful in the long run.

## 2.2 White

A second example in this space concerns a paper by White (2018). White seeks to show that ‘student partnership can easily undermine relations of authority and trust between students and academic teachers’ (2018: 163). She here has in mind the idea that ‘when it comes to making and evaluating changes to curricula, student voice and student engagement should be built in to academic decision-making’ (2018: 165).

Borrowing from Hawley, White (2018: 170) proceeds as follows.

Hawley argues that the appropriateness of relations of trust or distrust hinges on an explicit or implicit commitment from the trusted towards the trustee.

To trust someone to do something is to believe that she has a commitment to doing it, and to rely on her to meet that commitment. To distrust someone to do something is to believe that she has a commitment to doing it, and yet not rely upon her to meet that commitment. (Hawley, 2014, p. 10)

The problem of the university teacher’s authority turns on the teacher’s commitment and purposes. It is likely that the university teacher understands her commitment in epistemic terms or in terms of attitude towards developing a sense of intellectual curiosity. The student, however, is encouraged by higher education policies and practices to understand the university teacher’s commitment on a different horizon of skill development in preparation for paid employment. The bureaucratic manager is keen to understand the academic’s commitment in terms of targets, various league table positions and research evaluation scores. The different horizons are not brought together: the concept of student partnership (insofar as it attempts to involve students in academic decisions) is insufficiently strong to fuse together instrumental student and instrumental academic interests.

She goes on to note that that this difference in commitments

encourages the student to adopt a distrustful stance towards the teacher and to regard the academic’s interest as distinct from the student’s interest. In other words, a non-moral, thin, functional conception of higher education precludes trust, and therefore precludes the recognition of authority, which in turn renders impossible an acquisition of authority by students through education. (2018: 170)

But whilst we commend White's engagement with the literature on the philosophy of trust, we don't see how this point follows. At the crux of the issue for us is the question: what kind of trust is precluded if different parties have, as she suggests, different commitments and interests?

To start to put flesh on the bones of our concern here, let us agree with White that each of teacher, pupil and bureaucrat have different commitments. This, in and of itself, is unsurprising. We all—each of us—have different commitments.

Next, we borrow an example from White (2018: 169).

If I trust you to teach me Locke's theory of just property acquisition, I might assume you have the right sort of motive for teaching me Locke. This might be your intellectual interest in Lockean political philosophy.

Now, let us focus in on the case: suppose that the student trusts us to teach them Locke. White is right, they might assume that we have particular motivations. But we think that we can set that to one side. If we are trusted to teach the student Locke's theories, then it follows from Hawley's view, that we have a commitment to teaching them Locke's theories. Now our motivations for teaching it here are, we think, irrelevant. If the student believes that we have a commitment to teaching them Locke's theories and relies on us to do so, then they trust us to teach them those theories. If we *are* in fact committed to teaching them Locke's theories, and the student can and does rely on us to teach them those theories, then the student was right to place their trust in us. It does not matter, then, whether we are motivated by avarice, intellectual curiosity, or any other guiding light; if the student (correctly) believes that we have a commitment to F-ing and relies on us to F, and we F, then we were rightly trusted to F.<sup>2</sup>

What goes wrong with White's argument? So far as we can tell, White wants to argue that if our commitments are different in this space, then we do not have a situation where distrust will emerge. But White's argument here fails to notice that each of student, lecturer and bureaucrat in this case, though they may have a different impression of the total set of the lecturer's commitments, do not disagree that the lecturer has one very particular commitment: to teaching the student Locke's theories. It isn't enough for White to show that the commitments in this case are different; White would need to show that each of student, lecturer and bureaucrat *do not believe* that the lecturer has a commitment to teaching the student Locke's theories. And since White does not provide any empirical evidence about the belief states of those involved (indeed, beyond the content of policy and reports White does not provide any empirical evidence at all) there is no evidence to support their view. It remains unclear, then, why distrust is supposed to follow from differing understandings of a lecturer's commitments (and, given the dearth of empirical data, it also remains unclear whether the different parties do *as a matter of fact* have different understandings of a lecturer's commitments).

---

<sup>2</sup> At least, this is correct if Hawley's account of trust is correct—something that White assumes. We're happy to make the same assumption here.



From some points of view, this ought to be unsurprising. Very generally, we all the time trust other agents to perform actions despite their having a very different total set of commitments from our own. The authors have very different total commitments from right-wing members of the political spectrum, but we still trust them to not physically assault us in the street. We do not distrust them on *this* score. So much is entirely natural. For distrust to arise *in that domain*, and with respect to that action, we would have to be shown much more than merely that we have a different set of interests and a different total set of commitments.

### 3. Avenues for future research

We have, to this point, focused primarily upon a negative research program—arguing that two recent attempts to bring together insights from the literature on the philosophy of education and the philosophy of trust fall short in their attempts. Nonetheless, the papers appear as a part of a special issue on trust in education, and we agree wholeheartedly with the editorial piece that introduces the collection, in which Haynes (2018) notes that the connection between trust and schooling is an area of research well worth exploring.

In this section of the paper, we want to try to draw some general lessons; lessons based both on the material from section 2 of the paper, but also from other areas that have sought to explore the trust in the world. In particular, we want to draw out two lessons as pressing. The first is the import of a theoretical model; the second is the importance of asking the right empirical question. In the abstract both points are someone anodyne, but given the history of research on trust both are worth making.

As adverted to in 2.1, having in hand an appropriate theoretical model of the different concepts associated with trust, as well as a detailed understanding of those concepts, is crucial if we are to make headway in areas of applied research—in this case, research in education. There are interesting parallels here with recent work by Donati and Tallant (forthcoming) in the lessons that the business literature can learn from philosophical work on trust.

As Donati and Tallant note:

Though there is no firmly agreed and explicit definition within the business literature, a fairly widely adopted approach is that advocated by Mayer et al. (1995: 712):

‘Trust is the willingness to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to the trustor, irrespective of the ability to monitor or control that other party’.

As they go on to point out, ‘trust’ as it is referred to here, cannot be any of trustworthiness, general trust, because it speaks of trusting a partner to perform specific actions, or three-placed trust. It cannot be three-placed trust for the simple reason that the analysis would be prone to counterexample. To borrow from Donati and Tallant:

suppose that company C regularly, mistakenly, and contradicting their own statements of what is ‘in the box’, ship their mobile phones with cases—that are described in all promotional materials as to be purchased separately. The company does this because its production line is a mess, not because it is keen to provide the cases to customers. A buyer, aware of this, might rely upon the company to provide them with a phone *and* case, without making an additional purchase of a case. In other words, they anticipate the case being shipped and make no alternative arrangements to get hold of one. This reliance should not amount to trust: the company would owe the buyer no apology if the phone turned up without a case, and the buyer ought not to feel betrayed by this, even if they felt disappointed.

Donati and Tallant go on to discuss a range of other theoretical models from the business literature, pointing out in each case that the model does not suffice.

Now in 2.1 we explained why we thought that Webster’s view of trustworthiness is problematic. But there is a more general concern lurking here: simply, Webster does not provide a theoretical model of trust such that we can test his many claims about it. There are a number of claims that would require further testing if we had such a model to hand. For instance:

‘the trust which makes professional autonomy and academic freedom possible can also be understood as potentially making an important contribution for promoting democratic practices’ (Webster, 2018: 153)

Again, what is this notion of trust? Is it two-placed trust? We suspect not, since it does not seem likely that what makes professional autonomy and academic freedom possible is *perfectly general* trust. It cannot be trustworthiness, since trust and trustworthiness are so very different (the former but not the latter is relational, for instance). It could be that Webster thus means to appeal to the notion of ‘trusting someone *to* do something’ But, in that case, much more detail is required. Not least of which: by whom and to do what? And, given a model of what three-placed trust consists in, does the resulting claim sound plausible? That is, does it sound plausible to say that it is x’s trusting y to F which makes professional autonomy and academic freedom possible [and] can also be understood as potentially making an important contribution for promoting democratic practices? We confess that we struggle to see what ‘F’ will be here that will do the job. Nonetheless, we would welcome an attempt to spell this out.

Thus, just as Tallant and Donati press the business literature for a clear and precise model of trust that carefully takes into account the philosophical literature on trust, so we too think that work in the philosophy of education ought similarly to model trust carefully and draw thoroughly and deeply on the philosophical literature on trust.

In addition to this highly theoretical work, and branching out from the philosophy education a little, we think that empirical educational research would similarly benefit from careful interaction with the philosophical literature on trust, for, as has been acknowledged in the trust-measurement literature (e.g. Bauer and Freitag (2015), the philosophical literature is of direct relevance in empirical studies. For example, Guerro et al (2019: 116) report that

[d]evelopmental research on testimonial learning shows that around four years of age, children are less gullible than assumed, as they can be selective in their trust in information coming from different sources

However, the study that they engaged in tested only whether or not children could reproduce information given to them by their teacher within very particular contexts.

This reveals a number of issues. First, and to repeat a theme from earlier in the paper, properly speaking information is not such that it can be trusted (only agents are trusted). Second, it is not clear whether what we have here is *mere reliance*, or trust, and what the consequences of that difference would come to. For instance, would children better reproduce information (or be able to reproduce the information over a greater period of time, if the trust their teacher to tell them accurate information, rather than merely relying upon them to do so? We do not have answers to these questions, but think the question to be one of genuine pedagogical value. If the ability to retain and reproduce information is driven by genuine trust, as opposed to mere reliance, then that may require quite specific teaching strategies.<sup>3</sup> After all, as we noted towards the start of the paper, trust (unlike mere reliance) appears to have a moral dimension. It may be that generating trust thus requires different behaviours than does the generation of reliance.

**And, we think, this last point also enables us to say something important about the difference between the pre-conditions for trust, and trust itself (in its various forms). There may be some concern that, in the above, we've engaged in hair-splitting. Perhaps, goes the thought, we can agree that traits like competence, willingness to be disposed towards reliability and honesty, and so on, can in some sense be distinguished from trust. Nonetheless, it's pretty clear that the exemplification of these traits will *lead* to trust in most cases. And, if that's right, then how much do we really need to worry about the distinctions we've drawn?**

**Whilst we agree that there's very likely, typically, a causal chain that takes us *from* the exemplification of traits such as competence, a willingness to be disposed towards reliability, and honesty, *to* trust; nonetheless, it's important to hold the two apart. Consider the following. Suppose that we believed (on empirical grounds) that where children trust their teachers to tell them the truth, they perform better in classroom assessment. We thus consider a strategy of developing pedagogical approaches to generate trust. Suppose, too, that we then seek to test our hypothesis by testing levels of trust by looking for each of the traits just identified (competence, a willingness to be disposed towards reliability, and honesty) and comparing those against attainment.**

---

<sup>3</sup> The case is hardly unique. Other studies that we think claim to draw conclusions about trust, but could equally well be read as claims about reliability, include Corriveau and Harris (2009), Corriveau et al (2009), Shafto et al (2009), Pesh et al (2018) and Elashi and Mills (2014). We will not seek to substantiate this claim here, though recommend careful analysis to the reader.

Now suppose that the children in the test case do *not* in fact trust the teacher to tell them the truth because of culturally prevalent biases that they have picked up from parents and friends. Let us suppose that the teacher is an avowed atheist and that the children have been exposed to the kinds of biases against atheists that lead to atheists strongly associating atheists with distrust. (see Gervais, et al. 2011).

The teacher could be highly competent, willing to be disposed towards reliability, and honest; nonetheless, because of the presence of a confound—the biases of the children against atheists—we might not see the expected improvement in classroom assessment because the children did not in fact trust their atheist teacher. This would lead to our incorrectly rejecting our hypothesis that trust is correlated with classroom attainment, and with it our strategic approach to classroom pedagogy. But this would, of course, be quite the wrong conclusion to draw. The correct conclusion to draw is that trust was not formed and so our hypothesis is neither confirmed nor disconfirmed by the evidence.

#### 4. Conclusions

To recapitulate an earlier theme: we welcome the integration of research in the philosophy of education (and education research more widely) with work in the philosophy of trust. We think such work extremely important. Nonetheless, we think that in order to make valuable progress, more care needs to be taken with the philosophical literature on trust. Of course, we do not mean this to be the end of things. We think that there is much work to be done, both in the domains described and also along new avenues. For instance, untouched in the literature (so far as we can tell) is the question of whether it is even *possible* for teaching to occur in the absence of trust, and, for instance, how this might flow from competing analyses of the nature of what teaching is. There is much to be done.

#### References:

- Baier, A. 1986. 'Trust and Antitrust', *Ethics*, 96, 231–60.
- Bauer P.C, and Freitag M. 2015. 'Measuring Trust'. In Uslaner (ed.) *The Oxford Handbook of Social and Political Trust* Oxford: OUP, 15-36
- Corriveau K.H, Harris P.L, Meins E, Fernyhough C, Arnott B, Elliot L, Liddle B, Hearn A, Vittorini L, de Rosnay M. 2009. 'Young children's trust in their mother's claims: longitudinal links with attachment security in infancy', *Child Development*, 80:750-761
- Corriveau K.H, and Harris P.L. 2009. 'Choosing your informant: weighing familiarity and recent accuracy', *Developmental Science*, 12, 426-437
- Donati, D. and Tallant, J. forthcoming. 'Trust: from the philosophical to the commercial', *Philosophy of Management*, <https://doi.org/10.1007/s40926-019-00107-y>

- Elashi F.B and Mills, C.M 2014. 'Do children trust based on group membership or prior accuracy? The role of novel group membership in children's trust decisions', *Journal of Experimental Child Psychology*, 128, 88-104
- Faulkner, P. 2007. 'On telling and trusting', *Mind*, 116, 875–902.
- Faulkner, P. 2015. 'The Attitude of Trust is Basic', *Analysis*, 75, 424-9
- Gervais, W.M., Sharrif, A.F., and Norenzayan, A. 2011. 'Do you believe in Atheists? Distrust is central to anti-atheist prejudice', *Journal of Personality and Social Psychology*, 101, 1189-1206
- Guerrero S, Sebastián-Enesco C, Pérez N, Enesco I. 2019. 'Myths in science: Children trust but do not retain their teacher's information', *Journal of Applied Developmental Psychology*, 62, 116-121
- Hardin, R. 2002. *Trust and Trustworthiness*. Russell Sage Foundation
- Hawley K. 2014. 'Trust, distrust and commitment', *Nous*, 48, 1–20.
- Hawley K. 2017. 'Trustworthy Groups and Organisations'. In *New Philosophical Perspectives on Trust*, ed. P. Faulkner and T. Simpson. Oxford: OUP
- Haynes B. 2018. '**Trust and schooling**', *Educational Philosophy and Theory*, 50, 119-122
- Hieronymi P. 2008. 'The Reasons of Trust', *Australasian Journal of Philosophy*, 86, 213–236
- Holton R. 1994. 'Deciding to Trust, Coming to Believe', *Australasian Journal of Philosophy*, 72, 63–76.
- Jones K. 1996. 'Trust as an Affective Attitude', *Ethics*, 107, 4–25
- Jones K. 2013. 'Trustworthiness', *Ethics*, 123, 61–85
- McLeod, C. 2006. 'Trust', *Stanford Encyclopedia of Philosophy*
- Mayer R. C, Davis J.H, and Schoorman D.F. 1995. An integrative model of organizational trust. *The Academy of Management Review*, 20, 709–734
- Nickel P.J, Franssen M, & Kroes, P. 2010. 'Can we make sense of the notion of trustworthy technology?', *Knowledge, Technology and Policy*, 23, 429–444
- O'Neill O. 2002. *Autonomy and Trust in Bioethics* Cambridge: Cambridge University Press.
- Pesch A, Suarez S, and Koenig M.A. 2018. 'Trusting others: shared reality in testimonial learning' *Current Opinion in Psychology*, 23, 38-41

- Potter N.N. 2002. *How Can I Be Trusted?*, Lanham: Rowan and Littlefield
- Pettit P. 1995. 'The Cunning of Trust', *Philosophy and Public Affairs*, 24, 202–25
- Tallant J. 2017. 'Commitment in Cases of Trust and Distrust', *Thought*, 6, 261-7
- Tallant, J. 2019. 'You can trust the ladder, but you shouldn't', *Theoria*, 85, 102-118
- Webster, S. R. 2018. 'Being trustworthy: going beyond evidence to desiring', *Educational Philosophy and Theory*, 50, 152-162
- White M. 2018. 'Student partnership, trust and authority in universities', *Educational Philosophy and Theory*, 50, 163-173