Trust and Trustworthiness

J. ADAM CARTER

COGITO Epistemology Research Centre, University of Glasgow

1. Introduction

What is the relationship between trust and trustworthiness? The question is a fraught one, not least because philosophers of *trust* have tended to focus on three-place trust (S trusts X with ϕ), whereas philosophers of *trustworthiness* have been primarily concerned with one-place trustworthiness – viz., (S is trustworthy).[1]

This mismatch in focus presents challenges for those who want their accounts of trust and trustworthiness to be mutually illuminating. And it also prompts deeper methodological questions, such as whether we ought to be trying to understand trust in terms of trustworthiness (as some philosophers have[2]) or trustworthiness in terms of trust (as others have)[3]?

Though there is little consensus here, a widespread underlying assumption is that the central phenomenon of interest on the trustee's side is *dispositional* (viz., trustworthiness) whilst the central phenomenon of interest on the trustor's side is *non-dispositional* (viz., trust). An important byproduct of this assumption is that the evaluative norms of principal interest on the trustor's side regulate trusting attitudes and performances whereas those on the trustee's side regulate dispositions to respond to trust.

The aim here will be to highlight some unnoticed problems with this asymmetrical picture – and in particular, how it elides certain key evaluative norms on both the trustor's and trustee's

---

[1]As Hardin (1996) notes, a further complication is that 'Many discussions of trust run trust and trustworthiness together, with claims about trust that might well apply to trustworthiness but that seem off the mark for trust' (1996, 28).

[2]See, e.g., O'Neill (2018, 293), Ashraf et al. (2006, 204), McLeod (2020, sec. 1),

[3]For example, according to Wright (2010), trustworthiness requires that the trustee 'acknowledges the value of the trust that is invested in them ... and to use[sic.] this to help rationally decide how to act' (2010, sec. 3.b.). Other accounts of trustworthiness in terms of trust are found in Williams (2000) and Potter (2002, 205).

side the satisfaction of which are critical to successful cooperative exchanges – and to show that a symmetrical, 'achievement-first' approach to theorising about trust and trustworthienss (and their relation to each other) has important advantages by comparison. The view I develop is guided by a structural analogy with practical reasoning. Just as practical reasoning is working as it should only when there is realisation (knowledge and action) of states (belief and intention) with reverse directions of fits (mind-to-world and world-to-mind), likewise, cooperation between trustor and trustee is functioning as it should only when there is an analogous kind of realisation on both sides of the cooperative exchange – viz., when the trustor 'matches' her achievement in trusting (an achievement in *fitting reliance to reciprocity*) with the trustee's achievement in responding to trust (an achievement in *fitting reciprocity to reliance*). An upshot of viewing cooperation between trustor and trustee as exhibiting achievement-theoretic structure is that we will be better positioned to subsume trustworthiness (and its cognates on the trustee's side), like trust, under a wider suite of evaluative norms that regulate attempts, dispositions, and achievements symmetrically on both sides of a cooperative exchange, with 'matching achievements' as the gold standard.

Here is the plan. §2 clarifies and criticises the kind of asymmetric picture that is embraced in the philosophy of trust and trustworthiness, which privileges performances (and norms regulating them) on the trustor's side and dispositions (and norms regulating them) on the trustee's side. §3 develops an analogy between practical reasoning and cooperation in order to motivate an alternative picture, on which trusting and trustworthiness are better understood as having achievement-theoretic structure with reverse directions of fit. §4 builds on this picture in order to defend symmetrical evaluative norms – norms of *success*, *competence*, and *aptness* – on both sides, and §5 considers and responds to some potential objections.

2. Trust and Trustworthiness: doing versus being?

In the philosophy of trust, trust and trustworthiness are characteristically theorised about in a way that largely takes for granted that what is of central philosophical interest on the *trustor*'s side of a cooperative exchange is a kind of doing – *trusting* – where as on the trustee's side, what's of central interest is the trustee's *being* a certain way (i.e., being trustworthy) on account of *possessing* some dispositional property or properties.[4] That philosophers of trust are

---

[4]Possessing a dispositional property is not itself a matter of being in a mental state or behaving some way. Rather, dispositional properties 'provide the possibility of some further specific state or behaviour' (see, e.g., Mumford 2016) in certain conditions. Accordingly, the various 'accounts of trustworthiness' on the market are not aiming to give

interested blue – on the trustor's side of a cooperative exchange – in occurrences of trust (or doings) shouldn't be surprising given that, as Faulkner (2015) puts it, 'Most philosophical discussion of trust focuses on the three-place trust predicate: X trusting Y to ϕ'[5], instances of which will always involve, on the trustor's part, more than merely possessing some propensity to trust the trustee[6] but their (in fact) doing so in a given case.[7]

One the trustee's side, however, the focus philosophically has been squarely on disposition possession: what disposition (or dispositions) line up with *being* trustworthy? Answers here vary. For example, according to a family of views defended by Annette Baier (1986), Karen Jones (2012), and Zac Cogley (2012), trustworthiness is to be identified with a disposition to fulfil commitments, in conditions under which one has those commitments, and *in virtue of* goodwill towards the trustor. For Diego Gambetta (1988), the trustworthy person needn't be disposed to fulfill the commitments they have out of good will; they simply must be disposed to fulfil their commitments, whatever they are, 'willingly'. More minimalistically, Christoph Kelp and Mona Simion (2021) identify trustworthiness with the disposition to fulfill one's obligations *simpliciter*, and not necessarily through any distinctive motivation or accompanying attitude. More weakly, for Katherine Hawley (2019), the relevant disposition 'trust-

---

an account of any mental state, attitude, or behaviour; they are aiming to characterise accurately the dispositional property that they take trustworthy people but not others to possess.

[5]Apart from this observation registering what has attracted philosophical focus, there is the question of whether three-place trust is more explanatorily fundamental than one-place trust; Faulkner (2015) denies that this is the case. For an overview of this debate contesting whether three-place trust is more fundamental, see, along with Faulkner (2015), Carter and Simion (2020, §1.b).

[6]It might seem as though a focus on dispositions (rather than doings) would be pertinent equally in the case of *trust*, at least for proponents of doxastic accounts of trust who elect to embrace, independently, a dispositional account of the nature of belief (e.g., Schwitzgebel 2013); on such accounts, believing a proposition (such as, in our case of interest, the proposition that the trustee will take care of things as entrusted) just amounts possessing one or more particular behavioural dispositions pertaining to the target proposition. Does this observation call into doubt the asymmetry highlighted? There are several reasons to think not. First, and importantly, the scope of the asymmetry claim is that, as a historical point, focus (on the trustor's side) has been almost entirely on doings, and on the trustee's side on, being a certain way. The possibility space of a substantive view on the trustor's side which would invite attention to trusting as a belief-cum-disposition is compatible with the larger observation. But, perhaps more importantly, it's not entirely clear that a proponent of doxastic theories of trust which embrace a dispositional account of belief will de facto be focusing on the mere possessing (rather than exercising) of dispositions in so far as the object of interest is three-placed trust. It is worth noting further that at least liberal dispositionalists about belief allow that forming representations (including with propositional content) can be essential to belief; the liberal dispositionalist's thesis is just that what would make such token representaitons essential to belief is their grounding of behaviour patterns (Schwitzgebel 2019, §1.3). Thanks to an anonymous referee for prompting further clarification on the option space consisting of doxastic accounts of trust combined with dispositional theories of belief.

[7]For instance, as Baier (1986) puts it, in theorising about trust we are centrally concerned with 'one person trusting another with some valued thing' (1986, 236); likewise, Hawley (2014) maintains that trust is 'primarily a three-place relation, involving two people and a task' (2014, 2). This is not to say that trusting has only been of interest as a three-placed predicate; on the contrary both two- and one-place trust have received discussion, and it remains a live question whether one relation is more fundamental than the others (see Faulkner 2015 for discussion of one- and two-placed trust as fundamental). Rather, the point is just that we can easily how interest in trust qua the activity of trusting, and not merely qua being a certain way, would be expected given that attention to trust as three-placed is found as widely as it is in the throughout discussions of trust by, among others, e.g., Baier (1986), Holton (1994), Jones (1996), Faulkner (2007), Hieronymi (2008) and Hawley (2014); for an overview, see Carter and Simion (2021, §1.b.).

worthiness' refers to is best framed negatively – viz., as a disposition to *avoid unfulfilled* commitments. By contrast with all of these views, Nancy Potter (2002), insists that the relevant disposition lining up with trustworthiness should be understood as a full-fledged moral virtue – one that consists in being disposed to respond to trust in appropriate ways.[8]

Notice that, despite their differences, these accounts all sign on, at least tacitly, to the idea that trustworthiness (on the trustee's side) and trust (on the trustor's side) are subject centrally to norms governing a (mere) 'being' to a 'doing'.[9] And thus, to the extent that these accounts are conducive to theorising about trustworthiness and trust in terms of each other, it will be as a being illuminates a doing, (or vice versa). A byproduct of this approach in focus is that other norms (on each side) are given much less consideration; that is to say, if the evaluative norms of principal interest on the trustor's side are norms that regulate trusting attitudes and performances, whereas those on the trustee's side regulate the trustee's dispositions to respond to trust, then what is inevitably going to be theoretically suppressed are whatever norms might govern, respectively, dispositions (on the trustor's side) and active performance through which dispositions are manifested (on the trustee's side).

While there is no doubt that being trustworthy corresponds with possessing *some* disposition or dispositions, so likewise does *being a competent trustor*, e.g., being one who trusts in ways that don't too often lead to betrayed trust.[10] And by the same token, just as trusting is itself not a disposition but an activity or performance, so likewise is the trustee's *manifestation of trustworthiness* when taking care of things as entrusted[11], viz., when actually reciprocating the trust placed in her (as opposed to merely 'being the sort of person' who would take care of things as entrusted).

It is worth asking: Is there any good reason that would justify the status quo here – viz., what has thus far been an asymmetrical focus on the dispositional property of trustworthiness and not on the trustee's performance of manifesting trustworthiness through reciprocity?

Perhaps there would be if the disposition of the trustee (rather than any performance on the

---

[8] According to Potter (2002), trustworthy persons "[...]give signs and assurances of their trustworthiness" and "They take their epistemic responsibilities seriously' (2002, 174–75; Cf., for criticism, Jones 2012, 75–76) and Kelp and Simion (2021). Note, however, that Jones (2012) is more sympathetic to the idea of trustworthiness as a virtue in the special case of what she calls 'rich' rather than 'basic' trustworthiness. See, e.g., (2012, 79).

[9] Or, alternatively, with reference to Vendler/Kenny classes – as an occurence (trust) to a state (trustworthiness). See, e.g., Verkuyl (1989).

[10] For a defence of this way of thinking about competent trust, see Carter (2020).

[11] The locution 'as entrusted' is meant to encompass views on which the trustee counts as taking care of things as entrusted only if doing so in a particular way, including, e.g., out of goodwill (Baier 1986; Jones 1996) or in conjunction with a belief that one is so committed (e.g., Hawley 2014). The present discussion – which is theoretically neutral on this point – is compatible with opting for either such kind of gloss.

trustee's part) features essentially in a plausible specification of what one aims at in trusting, and thus, in explaining when trust is successful. Or, alternatively, if the evaluative norms that are so often elided (e.g., manifestations on the trustee's side, or competence or dispositions on the trustor's side) are somehow less important to cooperative exchanges going well. Neither line of thought is promising.

The former is implicit in what Carolyn McLeod (2020) takes to be a platitude about trust, which is that 'Trust is an attitude that we have towards people whom we hope will *be trustworthy*, where trustworthiness is a property not an attitude'[12]. Variations of this idea are seen in Elizabeth Fricker's (2018) claim that '[...] One is not really trusting unless one adopts an attitude of optimism to the proposition that the trustee *is trustworthy*' (2018, 6). Likewise, as Russell Hardin (2002) puts it, 'trusting someone in some context is simply to be explained as merely the expectation that the person will most likely *be trustworthy*' (2002, 31). And perhaps most directly, proponents of doxastic accounts of trust (Hieronymi 2008; McMyler 2011) straightforwardly identify trust with a belief that the trustee *is trustworthy*.[13]

Of course, we seek out a trustworthy person when initially deciding *whether* to trust or forbear from trusting; in this respect, Onora O'Neill is right that 'where we aim [...] to place and refuse trust intelligently *we must link trust to trustworthiness*' (2018, 293). But when we actually trust someone, the relevance of the trustee's simply 'being a certain way' independent of their actually performing in a way that manifests how we perhaps hope or believe they are (i.e., trustworthy) – is not clear at all.

When I trust you to pay back the loan, I rely on you to pay it back, making myself vulnerable to your betrayal.[14] Suppose you *do* then pay it back. Is my trust successful? Not necessarily, says the proponent of the idea that trustworthiness is of special interest in understanding trust. In trusting, I aim *not just* that you take care of things any way, but take care of things *as entrusted*, which (on this line of thought) involves your 'being trustworthy'.

This is partly right. But it gets an important thing wrong. Just as my trust isn't thereby successful if you merely take care of things *any* old way (e.g., by attempting to betray me, but in doing so accidentally pay back the loan – then only my *reliance* would be successful), like-

---

[12]See McLeod (2020, sec. 1, my italics).

[13]For a related though less standard kind of doxastic account, see, e.g., Keren (2014, 2019).

[14]For various expressions of the idea that trust essentially involves subjecting oneself to risk of betrayal, see, along with Hardin (1992), e.g., Baier (1986, 244), McLeod (2020, sec. 1), Nickel and Vaesen (2012, 861–62), Carter (2020, 2301, 2318–19), Carter and Simion (2020, sec. 1.a), Becker (1996, 45, 49). Dasgupta (1988, 67–68), Dormandy (2020, 241–42), Kirton (forthcoming), O'Neill (2017, 70–72), Potter (2020, 244), and Hinchman (2017). Cf., Pettit (1995, 208).

wise, my trust misses the mark if you simply *are* trustworthy but (perhaps due to bad luck) *don't* pay back the loan. But crucially – even more – there is a sense in which my trust *still* misses the mark if you (i) pay back the loan; (ii) are trustworthy; but (iii) your paying back the loan doesn't manifest your trustworthiness (e.g., perhaps despite being trustworthy you pay back the loan on this occasion under threat or through some kind of manipulation by a third party).[15]

The foregoing suggests that (on the trustee's side) the largely exclusive theoretical focus on *trustworthiness* qua disposition (as opposed to, e.g., focusing on performances of trustworthiness) isn't going to be justified simply in light of the importance of this disposition to understanding the aims of trusting. After all, the trustor (through trusting) aims not just at the trustee merely *being* a certain way – or even at the trustee doing a certain thing *while at the same time being a certain way* – but at the trustee *achieving* a certain thing, viz., succeeding in taking care of things *through* their trustworthiness.

This observation serves neatly to counter the second idea envisaged in favour of the asymmetrical picture under consideration; it's simply not evident that the evaluative norms that the asymmetric picture de facto elides (e.g., manifestations of dispositions on the trustee's side, or competence or dispositions on the trustor's side) are somehow less important to successful trust exchanges going well than are 'doings' on the trustor's side and 'being' on the trustee's side which have received the brunt of the philosophical focus so far. At minimum, getting these de facto elided norms into better focus will open up additional ways to evaluate what goes wrong (and right) in a given cooperative exchange.

This insight offers us a new vantage point for revisiting the relationship between trust and trustworthiness – and how to theorise about both in connection to each other and to appreciate some important performance-theoretic *symmetries* between the trustee/trustee sides of a cooperative exchange.

---

[15]Coercion isn't essential to making this kind of point; for example, the trustworthy person might be such that her success (in taking care of things as entrusted) doesn't manifest her trustworthiness not because she lacked the opportunity to do so (as would be the case if she were coerced) but rather due to the abnormal presence of luck in accounting for the success. The underlying idea here – one that has been defended variously by Sosa (2007), Greco (2010, Ch. 5), Pritchard (2012), and Zagzebski (1996) – is that a success doesn't manifest one's reliable disposition (construed as an ability, virtue, or competence) if that success is unusually due to luck. How to unpack 'unusually' (alternatively: abnormally) is a contested point, one that features centrally in discussions in virtue epistemology of achievement, luck, and credit. See, e.g., Turri et al. (2019, sec. 5, §7).

3. Structural analogies with practical reasoning

Whereas mere *reliance* is successful just in case the person relied on takes care of things *any way*, the success conditions for trust are more demanding. Where we've got to so far is that in *trusting*, as opposed to merely relying, the trustor aims through trusting that the trustee *manifests her trustworthiness* in successfully taking care of what the trustor relies on her to do. And this involves, on the part of the trustee, a kind of *success through trustworthiness* – viz., an *achievement* in trustworthiness, rather than, say, responding to the trustor's trust in such a way as to avoid betraying that trust just by luck.

Let's now take this working idea – that the trustor aims in trusting at the trustee's *achievement* in trustworthiness – even further. Just consider that when the trustor *herself* attains this aim (i.e., the aim that trustee's taking care of things manifests her trustworthiness) – this might on some occasions of cooperation be down to dumb luck; the trustor might foolishly trust the one trustworthy person in the village of tricksters, but this lone trustworthy person might then manifest her trustworthiness full well in taking care of things.[16]

Trust is *successful* here. *And* the trustee exhibits an achievement in fulfilling the trust placed in her through trustworthiness. But, in this situation, there is no symmetrical achievement (success that manifests a trusting competence) on the *trustor's* side, even though trust is successful. And in this respect, there is an important sense in which the cooperation itself still falls short; the cooperation does not match 'achievement to achievement', but matches merely success (by the trustor) to achievement (by the trustee).

Of course, the symmetry can be regained if we simply shore up the performance on the trustor's side.[17] Suppose it is *not* simply through good fortune but *through the trustor's competence* (to trust *successfully* reliably enough) that the trustor trusts successfully. In such a case, the relevant trust is not just successful but 'apt' in that the successful trust manifests

---

[16]In this example, we are to suppose that the trustor is not manifesting any trusting competence here, but rather, simply and naively trusting and just happens to be lucky. The structure of the case is importantly different from a case where the trustee *does* manifest competence, in a normal environment, but could have easily trusted someone who was mistaken in that environment. In the latter case, the structure is analogous to that of a 'fake barn case' – and that is a case which, at least within performance-theoretic epistemology, there is no barrier to attributing the success to the ability and thus to attributing achievement. For discussion, see, e.g., Sosa (2007, Ch. 2), Littlejohn (2014), Carter (2016), Pritchard (2009), Jarvis (2013), and Kallestrup and Pritchard (2014).

[17]Note that the idea of trust as a kind of performance is compatible with the thought that – as Nguyen (forthcoming) holds in recent work – trusting essentially involves forbearing from the taking up of certain attitudes, such as questioning. This is because forbearing can itself feature in aimed performances. Consider, for example, the opera singer's performance of a piece, which (if it is to succeed in its aim) must include silence at the right parts, a form of forbearing from singing. See, e.g., Sosa (2021, Ch. 3), who discusses this point in connection with our evaluation of suspension of judgment as a performance type.

the trustor's competence to bring that success about reliably. This *apt* (and not merely successful) trust, an achievement of trusting, on the trustor's side would then match the trustee's achievement in trustworthienss. And now *cooperation* is functioning well in that the cooperation between the two *falls short on neither side* of the cooperative exchange.

An analogy is useful here between (i) the symmetrical picture just described of cooperation working well; and (ii) Tim Williamson's (2017) view of practical reasoning working well. According to Williamson, a practical reasoning system is working well when and only when one acts on what one knows.[18] One is in a position to act on what one knows only if one 'realises' two kinds of states, with *reverse directions of fit* (mind-to-world and world-to-mind). Accordingly, on Williamson's picture, practical reasoning is not functioning as it should if there is a defect on either on the mind-to-world side (i.e, mere belief rather than knowledge) or on the world-to-mind side (i.e., mere intention rather than action).[19]

The working analogy so far is this: just as practical reasoning is functioning well only when we have symmetrical realisation (knowledge and action) of states (belief and intention) with reverse directions of fits, likewise, cooperation between trustor and trustee is functioning well only when we have an analogous kind of *symmetrical realisation on both sides of the cooperative exchange* – viz., when the trustor 'matches' her achievement in trusting with the trustee's achievement in responding to trust.

This working analogy can be extended further, by considering how the trustor's and trustee's matching achievements, when cooperation is working well, are themselves (like knowledge and action) *realisations of attempts with reverse direction of fit*.[20] To a first approximation:

---

[18]This idea, originating in Williamson (2002), is given a sustained defence in his (2017) with some further updates in (2021). Whereas Williamson encourages us to view the idea that practical reasoning's working well is a matter of acting on knowledge in the service of a wider criticism of the centrality of belief-desire psychology as explanatorily central, the core normative idea that, in practical reasoning, one should act only on what one knows has received defences by (along with Williamson) Hawthorne and Stanley (2008), Stanley (2005), Fantl and McGrath (2002). For an overview, see Benton (2014, sec. 2.a).

[19]I say 'intention' here rather than 'desire' as standing in for botched action to reflect Williamson's updated (2017) structural analogies. In *Knowledge and its Limits* (2002), practical reason's working well was also understood in terms of acting on knowledge. This picture was meant to replace belief-desire psychology as the centre of intelligent life. However, the original (2002) version of the analogy maintained that belief stood to knowledge (on the mind-to-world side) as desire stood to action (on the world-to-mind side). The updated picture assimilates desire to belief – i.e., belief about what is good – e.g., (see, e.g., Lewis 1988; Price 1989) – and replaces 'desire' with 'intention' in the analogy. Thus, the updated picture holds that belief stands to knowledge as intention to action. For critical discussion, see Miracchi and Carter (2022).

[20]The language of 'direction of fit' is originally usually credited to Anscombe (1957), as a way of characterising a distinction between theoretical and practical intentional mental states. Theoretical mental states aim at representing things as they are (e.g., beliefs) and practical mental states aim at getting things done (i.e., desires). Realisation (i.e., success), for a cognitive (or theoretical) intentional mental state involves fitting mind-to-world; realisation for a practical mental state (e.g., desire, intention, etc.) involves fitting world-to-mind. A central 'lesson' direction-of-fit theorists (e.g., Smith 1994; Velleman 2000; cf. Frost 2014) have taken from Anscombe's initial discussion is that intentional mental states are characterisable along the mind-to-world or world-to-mind faultline. However, the kinds

whereas the trustor aims not just to rely, but to *fit her reliance to the trustee's reciprocity*, the trustee (as such) aims to *fit her reciprocity to the trustor's reliance*.[21]

When the trustor attempts, but fails, to fit her reliance to reciprocity, what is residual is a kind of botched trust. When the trustee attempts, but fails, to fit her reciprocity to reliance, what is residual is a kind of botched reciprocity. (Compare with Williamsons's suggestion that belief is a kind of 'botched knowledge' and mere intention 'botched' action).[22]

On this wider picture, then, in any two-way cooperative system, trust stands to apt trust as reciprocity to apt reciprocity (reciprocity that succeeds through trustworthiness) in a way that is broadly analogous to how (in a practical reasoning system, for Williamson) belief stands to knowledge (viz., *apt* belief[23]) as intention to action (i.e., *apt* intention[24]). And, further, just as practical reasoning's working well requires a match between not merely belief and intention but between knowledge and action; cooperation working well requires a match between not *mere* but *apt* trust and reciprocity.

The tables below represent these key analogies:

of things to which direction of fit talk is applicable are not limited to intentional mental states. For example, according to Searle (1979), statements and predictions have a *word-to-world* direction of fit, whereas commands and promises have a *world-to-word* direction of fit. It's worth noting further that the very thought that things other than mental states can have directions of fit is actually perfectly compatible with Anscombe's initial idea, which is that what *makes* intentional states like beliefs and desires have the directions of fit is that they have normative realisation conditions; beliefs *aim* (constitutively, not intentionally) at a certain kind of realisation, the same for desires. A similar normative reading is due to Platts (1980; for discussion see Frost 2014, 449–50). What this suggests, then, is that – at least in so far as we follow progenitors of DOF theory such as Anscombe and Searle, there is no barrier to viewing attempts more generally (including, e.g., trust and its reciprocation) with constitutive aims as admitting of directions of fit in so far as they have specifiable normative realisation conditions.

[21] I say 'reverse' direction of fit for ease of presentation, given that 'reliance-to-reciprocity' and 'reciprocity-to-reliance' are ostensibly reverse directions of fit. That said, it might have been more precise to describe the kind of direction of fit here as lining up even better with what Searle (1979) calls 'double direction of fit'. The reason here is that – in the unique case of cooperation – the realisation of one entails the realisation of the other.

[22] In more recent work, Williamson (2021) has distinguished between what he calls *local failure* and *global failure* in a cognitive system. As he puts it, 'when a belief fails to constitute knowledge, it is a local failure, a defect in that particular belief. But when a cognitive system is too prone to produce such beliefs short of knowledge, it is a global failure, a defect in the system as a whole' (2021, 7). One philosophical question, of interest in recent work on epistemological dilemmas (e.g., Hughes 2019), concerns how to evaluate mixed cases that feature combinations of global success + local failure and local success + global failure; of particular interest is whether such cases feature mere normative conflict (e.g., Simion, forthcoming) or epistemological 'dilemmas' for a thinker. While it goes beyond the present aims to take this issue up, it is worth registering that an analogous global/local failure distinction arises in the philosophy of trust – where we can envision cases of local failure + global success and vice versa.

[23] The idea that knowledge is type-identical with apt belief has advantages in epistemology; see Sosa (2007, 2010), Greco (2010, Chs. 5-6) and Zagzebski (1996) for some notable defences of this position. Although I find this view plausible, the identification of knowledge with apt belief – while if fits snugly with the proposal developed here – isn't essential to it. For some criticism of the identification of knowledge with apt belief, see, e.g., Lackey (2007), Pritchard (2007), Kelp (2013), Kornblith (2004), and Kallestrup and Pritchard (2014).

[24] For defences of the view that action is fruitfully understood as apt intention, see Sosa (2015, Ch. 1) and Miracchi and Carter (2022).

Table 1: Practical reasoning: realisations and attempts

Practical reasoning	Fitting mind-to-world	Fitting world-to-mind
Functioning well	knowledge (realisation)	action (realisation)
	belief (attempt)	intention (attempt)

Table 2: Cooperation: realisations and attempts

Cooperation	Fitting reliance-to-reciprocity	Fitting reciprocity-to-reliance
Functioning well	trustor's apt trust (realisation)	trustee's apt reciprocity (realisation)
	trust (attempt)	reciprocity (attempt)

One important feature of Williamson's 'knowledge-action' centric picture of practical reasoning is that it is meant to contrast with a competing picture (see, e.g., Humberstone 1992) that takes attempts – belief and desire – *rather than their realisations* as the core explanatory mental attitudes at the centre of intelligent life.[25] Attempts at knowledge and attempts at action retain a place in this picture, but it is their realisations, rather than the attempts themselves, that are of comparative theoretical importance.

The picture of cooperation suggested here likewise gives primacy to realisations over their attempts. That is, the present picture rejects the trustee's performance (trust), a mere attempt at realisation by fitting reliance to reciprocity, and the trustee's disposition (trustworthiness) to fit reciprocity to reliance are the most theoretically important notions in a wider picture of cooperation. Rather, we should think of the importance of the trustor and trustee's matching achievements of trust and trustworthienss in cooperation as broadly analogous to the importance of action and knowledge (as opposed to mere belief and desire) in practical reasoning.

---

[25]On the kind of view embraced by Humberstone (1992), it is also possible to accept the structural analogy on which belief stands to knowledge as desire to action. However, such a structural analogy would (on the belief-desire centred picture) begin with belief and desire as 'direction of fit mirror images', from which we would then 'solve upward' in the analogy to get the result that belief stands to knowledge as desire to action. Resisting this picture is the central argument in Williamson (2017), who suggests we begin with knowledge and action as direction of fit mirror images and then solve 'downward', filling in the relevant attempts. For a criticism of the role of 'mirrors' in both Williamson and Humberstone's approaches, see Miracchi and Carter (2022).

4. Symmetric evaluative normativity: trustor and trustee

In the good case where cooperation is working as it should, the trustor matches her achievement in trusting with the trustee's achievement in responding to trust. In both of these achievements (of apt trust and apt reciprocity) *competence* is manifested in success.

Cooperation doesn't always go so well. It falls short – at least to some extent – if we have anything short of achievement on either the trustor or trustee's side. In some cases, cooperation doesn't fall short by much, as when the trustor matches successful and competent but inapt (i.e., Gettiered) success to the trustee's achievement.[26] The trustor could do far worse. Successful but incompetent trust falls short of Gettiered trust on the trustor's side, as does competent but unsuccessful trust.[27] On the bottom rung on the trustor's side, we have trust that is neither competent nor successful, e.g., the betrayal of the gullible.

Likewise, on the trustee's side, falling just short of achievement is a kind of *Gettiered reciprocity*; suppose the trustee manifests her trustworthiness in assiduously entering the correct bank details online to pay back the loan she was entrusted to pay back, but succeeds only because a fortuitous electronic glitch (good luck) *corrects for an initial glitch* (bad luck) that would have diverted the funds to the wrong account.[28]

The trustee could do far worse. For one thing, she could have *not* manifested trustworthiness in responding to the trust placed in her, but succeeded just by luck. In such a case, suppose she *intends* to wire the money to the wrong account but only accidentally wires it to the right one.

Whereas the first loan case is a case of Gettiered reciprocity, the second is successful but incompetent reciprocity. Two remaining categories, lower down the rung on the trustee's side are: unsuccessful and competent reciprocity (i.e., exactly like the Gettiered reciprocity case *without* the second stroke of good luck), and – at the very bottom rung – incompetent and unsuccessful reciprocity (e.g., the trustee intends to wire money to the wrong account, and – failing in reciprocity – succeeds in betrayal.)

---

[26] Performances that are successful and competent but inapt have a 'Gettier' structure, where the success is disconnected from the good method used. For discussion, see Sosa (2007, Ch. 2; 2010, 467, 474–75) and Greco (2009, 19–21; 2010, 73–76, 94–99). Cf., Pritchard (2012, 251, 264–68).

[27] The performance-theoretic analogy with virtue epistemology holds that successful but incompetent trust and competent but unsuccessful trust fall short of apt trust in a way that is analogous to how unjustified true beliefs and justified false beliefs both fall short of knowledge. See, for discussion, Sosa (2007, 2010, 2015).

[28] For discussion of this kind of 'double luck' structure in relation to Gettier cases, see, e.g., Zagzebski (1994); see also Pritchard (2007) on what he calls 'intervening' veritic luck.

The above picture shows not only the many ways that cooperation can be defective (by less or greater degree) by matching anywhere from *just* less to *much* less than achievement on either the trustor's or trustee's side. But it also reveals an important *normative symmetry* on both sides.

By 'normative symmetry' what I mean is that the relevant *attempts* on each side (fitting reliance to reciprocity on the trustor's side, fitting reciprocity to reliance on the trustee's side) are such that we can evaluate each for the very same three things: (i) *success*; (ii) *competence*; and (iii) *aptness*.[29] And, moreover, it is specifically by *failing* to satisfy combinations of these norms that performances on the trustor and trustee's side fall short of achievement to whatever extent that they do.

The symmetrical picture of evaluative norms on each side is accordingly as follows:

	On the truster's side	**On the trustee's side**
Direction of fit attempt	Reliance-to-reciprocity (trust) by means of reliance	Reciprocity-to-reliance reciprocity (by means of responding to trust)
success norm	S's trusting X with ϕ is better if successfully reciprocated than if not; S's trusting X with ϕ is successfully reciprocated iff takes care of ϕ as entrusted.	X's reciprocating S's trust with ϕ is better if X successfully reciprocates S's trust with ϕ than if not; X successfully reciprocates S's trust with ϕ iff X takes care of ϕ as entrusted.
competence norm	S's trusting X with ϕ is better if S trusts X with ϕ competently than if S does not.	X's reciprocating S's' trusting X with ϕ is better if X reciprocates S's trust with ϕ competently than if X does not.

---

[29]This is not to say that we can evaluate performances only along these three dimensions. For instance, a more recent innovation in the theory of performance normativity recognises the normative standard – beyond mere aptness – of *full aptness* (e.g., Sosa 2015; 2021; Carter 2021; see also Carter and Sosa 2021). To a first approximation, a performance is fully apt iff it is not merely apt, but also guided to aptness by an apt (second-order) risk assessment that it would (likely enough) be apt. For discussion of this difference (as applicable to trust) see Carter (2020). For ease of presentation, I have articulated the relevant achievement matching (on the side of trustor and trustee) in terms of aptness on each side; however, see §5 (Objections and Replies) for some additional discussion.

	On the truster's side	**On the trustee's side**
aptness norm	S's trusting X with ϕ is better if S trusts X with ϕ aptly than if S does not.	X's reciprocating S's trusting X with ϕ is better if X reciprocates S's trust with ϕ aptly than if X does not.

This symmetrical picture offers us a number of advantages. For one thing, our guiding idea that cooperation between trustor and trustee is working as it should when both sides match achievement to achievement can now be restated as an *aptness norm on cooperation*, one that is formulated *in terms of* trustor and trustee satisfying respective evaluative norms of aptness: a cooperative trust exchange E between trustor and trustee is better than it would be otherwise if E is apt;[30] E is apt iff trustor and trustee satisfy their respective aptness norms.[31]

Secondly – and this bring us back to where we started – it should now be even more evident why focusing principally on a disposition (trustworthiness) on the trustee's side but not on the trustor's side (and vice versa for performance) is going to be somewhat arbitrary. From a wider view that takes in and evaluates the trust exchange in full, neither has any special status, even though both are essential to cooperation going well. They are, in a bit more detail, essential to cooperation going well in a way that is roughly analogous to how our beliefs and intentions (or: dispositions to form intentions) are important to practical reasoning going well. Both deserve attention, but should be appreciated as attempts *at* realisations, where the realisations of those attempts are what's needed in good practical reasoning as well as (*mutatis mutandis*) in good cooperation.

Thirdly, by transitioning to a symmetrical picture of the evaluative normativity of trust – with achievement matching achievement as the gold standard – we are better positioned to see the importance of questions that have been so far obscured. Perhaps most conspicuously here

---

[30] It is worth noting that an aptness norm on cooperation offers, additionally, a standard for evaluating wider trust networks, in terms of the cooperative trust exchanges that comprise them; in this way, an evaluation of a network (consisting in individuals) can be assessed against the guiding value of apt cooperation – and not merely by values that pick out individual-level metrics.

[31] The idea that cooperation itself admits of an aptness norm suggests that cooperation is a kind of multi-agent performance itself. A natural way of thinking of this is as an irreducibly collective property of cooperators engaged jointly in a trust exchange. While I am sympathetic to this kind of gloss, I want to stress that we needn't be committed to it. The crux of the idea – viz., that cooperation is apt iff its individual cooperators perform aptly – is also compatible with a 'summativst' gloss, on which the cooperation has the relevant property (i.e., aptness) iff all its individual members have that property. For relevant recent discussion of these points, see Lackey (2021) and Broncano-Berrocal and Carter (2021). For a discussion of aptness as applicable to groups, see Kallestrup (2016).

are questions about the competence norm of trust. After all, we have a grip on *apt trust* only by understanding competent trust, and this involves a clear view of those dispositions of the trustor that lead them to trust *successfully* reliably enough. Other questions invited by the symmetrical picture involve the evaluative normativity of cooperation generally. Even if 'aptness on both sides' of the trustor/trustee divide implies that the cooperative exchange itself is apt, it remains an open question how to evaluate certain cooperation permutations that involve at least one norm violation on one side. For example: is cooperation working better if the trustor matches success without competence to the trustee's achievement or competence without success to the trustee's achievement?

Fourthly, given that *distrust* no less than trust can be successful, competent and apt, the normative symmetry we find on the trustor and trustee's side invites us to consider – analogically – what stands to distrust as trust to reciprocity[32], and to consider how to best characterise parallel symmetrical norms that would regulate – symmetrically with successful, competent and apt distrust (on the side of the trustor) – also forbearance on the side of the trustee.

5. Objections and replies

(a) Objection

It's not entirely clear that *cooperation* functions as it should only if we get a kind of 'aptness' on both the trustor and trustee's side (viz., matching achievements of fitting reliance-to-reciprocity and reciprocity-to-reliance, respectively). Might this not be too strong? For example, suppose a trustworthy trustee does what she is trusted to do (*vis-à-vis* the trustor) *because* she is trustworthy; isn't that enough for the cooperative exchange to plausibly get 'full marks'? Why should we think there must also be, as the view here suggests, something extra that the trustor does on top of this, which is to trust the trustworthy trustee *because* she is a competent trustor?

Reply. On the supposition, in the above case described, that the trustor's trust was successful but *not* through trusting competence, then regardless of how well the trustee performs (even if ideally), we are still left with a kind of Gettier-structure on the trustor's side; that is, there will be a kind of credit-reducing luck applicable to the trustor's trusting in such a case, even

---

[32] Analogically, this will involve some form of forbearance from reciprocity. One point of note here: forbearing from reciprocity, in a sense that would be analogous with distrust, would involve refraining from accepting the obligation incurred with being entrusted with something, as opposed to, accepting the obligation and then betraying the trust. For discussion of forbearance in trusting (and distrusting), see Carter (2022).

if such credit-reducing luck is not going to be equally applicable to trustee's manifestation of trustworthiness in responding to the trustor's trust. A satisfying assessment of the kind of trust exchange just described should be able to make to make sense of *both* (i) what goes right in the above case (which is very much, particularly given the performance of the trustee), while (ii) also having some explanation available for what comes up short on the trustor's side. Performance-theoretic assessment can deliver both results straightforwardly; recall that, on the view defended, the cooperative exchange itself involves two performances (one of fitting reliance to reciprocity, the other reciprocity to reliance); the conjunction of the two performances (exhaustive of the cooperative exchange) gets *very near* full marks precisely because nearly all evaluative norms on both sides are met; success, competence and aptness norms are satisfied on the trustee's side; success and competence but not aptness norms are satisfied on the trustor's side.

(b) Objection

Here is an objection that presses in the other direction. Even if we concede that Gettier structures (viz., featuring success and competence, but not success *because* competence) on either the trustor's or trustee's side of a cooperative exchange would implicate a kind of credit-reducing luck that is incompatible with aptness[33], there still remains a (different) kind of credit-reducing luck – *environmental luck*[34] – that is widely thought to be compatible with aptness (and even if Gettier-style luck is not); this is just the sort of luck at play in fake barn cases in epistemology, whereby one's success is unsafe but only because one is in an inhospitable environment that includes near-by error possibilities.[35] The very idea that environmental luck is compatible with aptness (even if Gettier-style luck is not) would seem to call into doubt whether 'achievement matching', understood as matching aptness to aptness on the trustor and trustee's side, respectively, really is the 'gold standard' when it comes to a cooperative exchange. Why not, after all, think that cooperation is functioning as it should only if performances on both side succeed in a way that safeguards against *both* Gettier-style as well

---

[33]For a dissenting view on this point, see Littlejohn (2014).

[34]See, e.g., Pritchard (2005) for the distinction between environmental epistemic luck and intervening epistemic luck.

[35]Note that the idea that environmental luck is compatible with aptness, while a point that has been made with reference to fake barn cases in epistemology, is not specifically an epistemological point; it is meant to apply equally to performances more generally. For example, to use an example from Sosa (2015), suppose a trainee pilot very easily could have woken up in a flight simulator, but instead luckily wakes up strapped in a real cockpit; in such a case, the pilot (assuming they are well-trained) can shoot targets competently and aptly, where the success manifests competence, and this is so even though very easily they could have woken instead in a simulator in which case they would have been shots that do not succeed in hitting (real) targets. For further discussion of this case, see also Sosa (2010, 467-8).

as environmental luck?

Reply: The are two key points to make in response to the above. The first will have us revisit Williamson's own structural analogies: consider again the idea that a practical reasoning system is working properly only when one acts on what one knows; this requires no defect either on the mind-to-world side (i.e, mere belief rather than knowledge) or on the world-to-mind side (i.e., mere intention rather than action). Even so, this needn't require maximal, or even extremely comprehensive knowledge with respect to object of the target intention (this is so even of one embraces a strong infallibilist thesis, eg., that knowledge = probability 1).[36] What this observation indicates is that a charitable reading of Williamson's own analogy is not that practical reason is working as it should only if there is no defect *whatsoever* (and of which falling short of maximal knowledge, or perhaps failing to know *that* one knows, etc., would constitute such a defect on the mind-to-world side) but, rather, that each side meet a salient standard that will be secured, minimally, by the realisation of the attempt one makes in fitting mind-to-word and world-to-mind in practical reasoning. And by parity of reasoning from this point to our own case of interest (featuring trustor and trustee), we can see how an analogous kind of position is going to be comparatively more plausible (on the achievement-matching view of cooperation working well) than would be a view that demands (on the trustor's and trustee's sides) *more* than aptness, in order to eliminate luck or chance that goes beyond mere Gettier-style luck that is incompatible with aptness. That said – and here is the second line of reply to the objection – *even if one were attracted to a more demanding picture* (on which mere aptness, on each side, isn't enough for cooperation to go as it should), the good news is that the tools of performance normativity offer the resources to model exactly what such a more demanding (and environmental-luck *excluding*) achievement-matching picture would look like. To a first approximation, the key idea would be to distinguish *mere aptness* from *full aptness* – where a fully apt performance, a higher form of achievement, is not merely apt, but also such that not easily would it have been inapt.[37] Then, the idea available within the framework (for one tempted to a stronger picture than what I've opted for here), would be to hold that cooperation between trustor and trustee attains an even higher (performance-theoretic)

---

[36] See, e.g., Williamson (2000).

[37] See Sosa (2015, Ch. 3). This is a simple statement of the idea. A more theoretically involved articulation of full aptness, found in Sosa (2015, Ch. 3; see also Carter 2020) would require distinguishing between mere apt trust and fully apt trust with reference to trusting *metacompetences* – competences to trust in ways that would not easily lead to inapt trust. Fully apt trust can then be stated (within a performance-theoretic framework), in terms of metacompetence as follows: trust is fully apt just in case it is guided to aptness by a meta-apt risk assessment that not too easily would the trust have been inapt. These details however take us beyond what. is needed to articulate the core idea that full aptness, even if not mere aptness, precludes environmental luck by de facto requiring safety.

quality when the trustor and trustee match achievements of *full* aptness on each side, in such a way as to gain aptness *safely*.

(c) Objection

Setting aside whether the 'top end' of the symmetrical picture of normative evaluation (of trustor and trustee) is too demanding (or not demanding enough), a more basic question remains, concerning the theoretical value of the symmetrical picture in comparison with the asymmetrical picture. Even if we grant that the symmetrical picture opens up space for asking new (performance-evaluative) questions about the trustor and trustee that might otherwise be (de facto) suppressed by focusing centrally on the trustor's actions and the trustee's dispositions, it is not entirely clear why it should be thought to be of special *importance* to ask these otherwise elided questions.

Reply: Two points are relevant here. First, the proposal defended does not privilege the kinds of normative evaluations the asymmetrical picture elides; such norms are not given 'special status'. The idea is, rather, that from a wider view that takes in and evaluates the trust exchange in full, it is evident that the satisfaction of success, competence, and aptness norms on both sides (fitting reliance-to-reciprocity, and reciprocity-to-reliance) are critical to cooperation going well, and in such a way that broadly mirrors how – by way of analogy – success, competence and aptness on both sides (fitting mind-to-world and world-to-mind) are likewise important to practical reasoning going well. So, the right characterisation of the picture here not that something other than the norms of central focus to the asymmetrical picture are actually the more important ones, but rather, that we stand to gain from evaluating cooperative exchanges between trustor and trustee in a more ecumenical way (developed by the picture here) that isn't going to be artificially restricted to a limited focal range on what the trustor and trustee are doing (and what dispositions they have), and how what they do manifests those dispositions. The second point worth making here is that – even setting this aside – there are important theoretical payoffs that line up specifically with gaining a better grip on the otherwise elided norms and when they are met. To use but one example here – on the trustor's side – consider the competence norm. Evaluating trustors for competence implicates concern for trusting in ways that don't too often lead to betrayal. What properties of a trustor facilitate trusting competence? Can they be taught and cultivated? Answers to such questions are of special importance not only in better understanding skilled trusting (as opposed to dispositions in responding to trust) but also, more practically, in navigating environments with

high levels of misinformation, and wherein such skills are increasingly needed to meet basic objectives.

6. Concluding Remarks

The aim here has been to motivate and defend a new way of theorising about trust and trustworthiness – and their relationship to each other – by locating both within a broader picture that captures largely overlooked symmetries on both the trustor's and trustee's side of a cooperative exchange. The view I've defended here takes good cooperation as a theoretical starting point; on the view proposed, cooperation between trustor and trustee is working well when achievements in trust and responding to trust are matched on both sides of the trust exchange. In a bit more detail, the trustor 'matches' her achievement in trusting (an achievement in fitting reliance to reciprocity) with the trustee's achievement in responding to trust (an achievement in fitting reciprocity to reliance). From this starting point, we can then appreciate *symmetrical* ways that the trustor and trustee can (respectively) fall short, by violating what I've shown are symmetrical evaluative norms – of success, competence and aptness – that regulate the attempts made by both trustor and trustee. The overall picture was shown to have important advantages over the received way of theorising about how trust stands to trustworthiness, and it clears the way – by identifying key questions that have been obscured – to making further progress.[38]

References

Anscombe, G. 1957. *Intention*. Harvard University Press.

Ashraf, Nava, Iris Bohnet, and Nikita Piankov. 2006. "Decomposing Trust and Trustworthiness." *Experimental Economics* 9 (3): 193–208.

Baier, Annette. 1986. "Trust and Antitrust." *Ethics* 96 (2): 231–60.

Becker, Lawrence C. 1996. "Trust as Noncognitive Security about Motives." *Ethics* 107 (1): 43–61.

Benton, Matthew A. 2014. "Knowledge Norms." *Internet Encyclopedia of Philosophy*.

Broncano-Berrocal, Fernando, and J. Adam Carter. 2021. *The Philosophy of Group Polarization: Epistemology, Metaphysics, Psychology*. Routledge.

Carter, J Adam. 2021. "De Minimis Normativism: A New Theory of Full Aptness." *The Philosophical Quarterly* 71 (1): 16–36. https://doi.org/10.1093/pq/pqaa017.

Carter, J. Adam. 2016. "Robust Virtue Epistemology as Anti-Luck Epistemology: A New Solution." *Pacific Philosophical Quarterly* 97 (1): 140–55.

———. 2020. "On Behalf of a Bi-Level Account of Trust." *Philosophical Studies* 177 (8): 2299–2322. https://doi.org/10.1007/s11098-019-01311-2.

———. 2022. "Trust as Performance." *Philosophical Issues: A Supplement to Noûs.*

Carter, J. Adam, and Mona Simion. 2020. "The Ethics and Epistemology of Trust." *Internet Encyclopedia of Philosophy.*

Carter, J. Adam, and Ernest Sosa. 2021. "Metaepistemology." In *Stanford Encyclopedia of Philosophy.*

Cogley, Zac. 2012. "Trust and the Trickster Problem."

Dasgupta, Partha. 1988. "Trust as a Commodity." In *Trust: Making and Breaking Cooperative Relations*, edited by Diego Gambetta, 49–72. Blackwell.

Dormandy, Katherine. 2020. "Exploitative Epistemic Trust." In *Trust in Epistemology*, edited by Katherine Dormandy, 241–64.

Fantl, Jeremy, and Matthew McGrath. 2002. "Evidence, Pragmatics, and Justification." *The Philosophical Review* 111 (1): 67–94.

Faulkner, Paul. 2007. "A Genealogy of Trust." *Episteme: A Journal of Social Epistemology* 4 (3): 305–21.

———. 2015. "The Attitude of Trust Is Basic." *Analysis* 75 (3): 424–29. https://doi.org/10.1093/analys/anv037.

Fricker, Elizabeth. 2018. "Trust and Testimonial Justification." *Manuscript.*

Frost, Kim. 2014. "On the Very Idea of Direction of Fit." *The Philosophical Review* 123 (4): 429–84. https://doi.org/10.1215/00318108-2749720.

Gambetta, Diego. 1988. *Trust: Making and Breaking Cooperative Relations.* Blackwell.

Greco, John. 2009. "Knowledge and Success from Ability." *Philosophical Studies* 142 (1): 17–26.

———. 2010. *Achieving Knowledge: A Virtue-Theoretic Account of Epistemic Normativity.* Cambridge University Press.

Hardin, Russell. 1992. "The Street-Level Epistemology of Trust." *Analyse & Kritik* 14 (2): 152–76.

———. 1996. "Trustworthiness." *Ethics* 107 (1): 26–42.

———. 2002. *Trust and Trustworthiness*. Russell Sage Foundation.

Hawley, Katherine. 2014. "Trust, Distrust and Commitment." *Noûs* 48 (1): 1–20.

———. 2019. *How to Be Trustworthy*. Oxford University Press, USA.

Hawthorne, John, and Jason Stanley. 2008. "Knowledge and Action." *The Journal of Philosophy* 105 (10): 571–90.

Hieronymi, Pamela. 2008. "The Reasons of Trust." *Australasian Journal of Philosophy* 86 (2): 213–36.

Hinchman, Edward. 2017. "On the Risks of Resting Assured: An Assurance Theory of Trust." In *The Philosophy of Trust*, edited by Paul Faulkner and Thomas W. Simpson. Oxford: Oxford University Press.

Hughes, Nick. 2019. "Dilemmic Epistemology." *Synthese* 196 (10): 4059–90.

Humberstone, I. Lloyd. 1992. "Direction of Fit." *Mind* 101 (401): 59–83.

Jarvis, Benjamin. 2013. "Knowledge, Cognitive Achievement, and Environmental Luck." *Pacific Philosophical Quarterly* 94 (4): 529–51.

Jones, Karen. 1996. "Trust as an Affective Attitude." *Ethics* 107 (1): 4–25.

———. 2012. "Trustworthiness." *Ethics* 123 (1): 61–85.

Kallestrup, Jesper. 2016. "Group Virtue Epistemology." *Synthese*, 1–19.

Kallestrup, Jesper, and Duncan Pritchard. 2014. "Virtue Epistemology and Epistemic Twin Earth." *European Journal of Philosophy* 22 (3): 335–57.

Kelp, Christoph. 2013. "Knowledge: The Safe-Apt View." *Australasian Journal of Philosophy* 91 (2): 265–78.

Kelp, Christoph, and Mona Simion. 2021. "What Is Trustworthiness?" *Manuscript*.

Keren, Arnon. 2014. "Trust and Belief: A Preemptive Reasons Account." *Synthese* 191 (12): 2593–2615.

———. 2019. "Trust, Preemption, and Knowledge." *Trust in Epistemology*.

Kirton, Andrew. forthcoming. "Matters of Trust as Matters of Attachment Security." *International Journal of Philosophical Studies*, forthcoming, 1–20.

Kornblith, Hilary. 2004. "Sosa on Human and Animal Knowledge." *Ernest Sosa and His Critics*, 126–34.

Lackey, Jennifer. 2007. "Why We Don't Deserve Credit for Everything We Know." *Synthese* 158 (3): 345–61.

———. 2021. *The Epistemology of Groups*. Oxford University Press, USA.

Lewis, David. 1988. "Desire as Belief." *Mind* 97 (387): 323–32.

Littlejohn, Clayton. 2014. "Fake Barns and False Dilemmas."

McLeod, Carolyn. 2020. "Trust." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Fall 2020. Metaphysics Research Lab, Stanford University.

McMyler, Benjamin. 2011. *Testimony, Trust, and Authority*. OUP USA.

Miracchi, Lisa, and J. Adam Carter. 2022. "Refitting the Mirrors: On Structural Analogies in Epistemology and Action Theory." *Synthese*. https://doi.org/10.1007/s11229-022-03462-y.

Mumford, Stephen. 2016. "Dispositions." In *Routledge Encyclopedia of Philosophy*, 1st ed. London: Routledge. https://doi.org/10.4324/9780415249126-N116-1.

Nguyen, C. Thi. Forthcoming. "Trust as an Unquestioning Attitude." *Oxford Studies in Epistemology*, Forthcoming.

Nickel, Philip J., and Krist Vaesen. 2012. "Risk and Trust." In *Handbook of Risk Theory*, edited by Sabine Roeser, Rafaela Hillerbrand, Martin Peterson, and Per Sandin. Springer.

O'Neil, Collin. 2017. "Betraying Trust." In *The Philosophy of Trust*, edited by Paul Faulkner and Thomas W. Simpson, 70–89. Oxford, UK: Oxford University Press.

O'Neill, Onora. 2018. "Linking Trust to Trustworthiness." *International Journal of Philosophical Studies* 26 (2): 293–300. https://doi.org/10.1080/09672559.2018.1454637.

Pettit, Philip. 1995. "The Cunning of Trust." *Philosophy and Public Affairs* 24 (3): 202–25.

Platts, Mark. 1980. "Ways of Meaning." *Mind* 89 (355): 454–56.

Potter, Nancy Nyquist. 2002. *How Can I Be Trusted?: A Virtue Theory of Trustworthiness*. Rowman & Littlefield.

———. 2020. "Interpersonal Trust." In *The Routledge Handbook of Trust and Philosophy*, edited by Judith Simon, 243–55. Routledge.

Price, Huw. 1989. "Defending Desire-as-Belief." *Mind* 98 (389): 119–27. https://www.jstor.org/stable/2255064.

Pritchard, Duncan. 2005. *Epistemic Luck*. Clarendon Press.

———. 2007. "Anti-Luck Epistemology." *Synthese* 158 (3): 277–97.

———. 2009. "Knowledge, Understanding and Epistemic Value." *Royal Institute of Philosophy Supplement* 64: 19–43. https://doi.org/10.1017/s1358246109000046.

———. 2012. "Anti-Luck Virtue Epistemology." *The Journal of Philosophy* 109 (3): 247–79.

Schwitzgebel, Eric. 2013. "A Dispositional Approach to Attitudes: Thinking Outside of the Belief Box." In *New Essays on Belief: Constitution, Content and Structure*, edited by Nikolaj Nottelmann, 75–99. London: Palgrave Macmillan UK. https://doi.org/10.1057/97

81137026521_5.

———. 2019. "Belief." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Fall 2019. Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/fall2019/entries/belief/.

Searle, J. R. 1979. *Expression and Meaning: Studies in the Theory of Speech Acts Cambridge University Press*. Cambridge.

Simion, Mona. Forthcoming. "Scepticism about Epistemic Dilemmas." In *Epistemic Dilemmas: New Arguments, New Angles*, edited by Kevin McCain, Scott Stapleford, and Matthias Steup. Routledge.

Smith, Michael. 1994. *The Moral Problem*. Blackwell.

Sosa, Ernest. 2007. *A Virtue Epistemology: Apt Belief and Reflective Knowledge, Volume I*. Oxford University Press.

———. 2010. "How Competence Matters in Epistemology." *Philosophical Perspectives* 24 (1): 465–75.

———. 2015. *Judgment & Agency*. Oxford University Press UK.

Stanley, Jason. 2005. *Knowledge and Practical Interests*. Oxford University Press.

Turri, John, Mark Alfano, and John Greco. 2019. "Virtue Epistemology." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Fall 2019. Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/fall2019/entries/epistemology-virtue/.

Velleman, David. 2000. *The Possibility of Practical Reason*. Oxford University Press.

Verkuyl, Henk J. 1989. "Aspectual Classes and Aspectual Composition." *Linguistics and Philosophy* 12 (1): 39–94.

Williams, Bernard. 2000. "Formal Structures and Social Reality." *Trust: Making and Breaking Cooperative Relations* 1: 3–13.

Williamson, Timothy. 2002. *Knowledge and Its Limits*. Oxford University Press on Demand.

———. 2017. "Acting on Knowledge." In *Knowledge First: Approaches in Epistemology and Mind*, edited by J. Adam Carter, Emma C. Gordon, and Benjamin W. Jarvis, 163–81. Oxford: Oxford University Press.

———. 2021. "Epistemological Ambivalence." In *Epistemic Dilemmas*, edited by Nick Hughes. Oxford: Oxford University Press.

Wright, Stephen. 2010. "Trust and Trustworthiness." *Philosophia* 38 (3): 615–27. https://doi.org/10.1007/s11406-009-9218-0.

Zagzebski, Linda T. 1994. "The Inescapability of Gettier Problems." *The Philosophical Quarterly (1950-)* 44 (174): 65–73.

———. 1996. *Virtues of the Mind: An Inquiry into the Nature of Virtue and the Ethical Foundations of Knowledge.* Cambridge University Press.