

Matters of Interpersonal Trust

A thesis submitted to The University of Manchester for
the degree of Doctor of Philosophy
In the Faculty of Humanities

2018

Andrew J. Kirton

School of Social Sciences

Table of Contents

Abstract	6
Declaration	7
Copyright Statement	7
Acknowledgements	8
Dedication	10
Table of Figures	11
1 Matters of interpersonal trust	12
1.1 Matters of trust & the trust stance	12
1.2 Domains of trust & competence/willingness.....	13
1.3 Two-place	14
1.4 Basic/one-place & voluntariness.....	17
1.5 Practical & testimonial.....	21
1.6 The anticipatory element & the interpersonal element	23
1.7 The trust-betrayal connection & vulnerability.....	27
1.8 Challenging interpersonal import as necessary for matters of trust	35
1.8.1 ‘Trust that’	42
2 The Reliance Plus view of trust and distrust	45
2.1 Reliance plus betrayability.....	45
2.2 REL+ and narrowing down betrayability.....	48
2.2.1 Baier: the ‘Goodwill’ account	48
2.2.2 Holton: the ‘Participant Stance’ account.....	51
2.2.3 Hawley’s ‘Commitment’ account.....	55
2.3 REL+ as an account of the trust stance	58
3 The argument from uncertainty: against REL+	64
3.1 Setting out the argument	64
3.2 Responses	69
3.2.1 Dissolving uncertainty into trust/distrust.....	70
3.2.2 Dissolving uncertainty into reliance/non-reliance.....	73
3.2.3 The third category response	75
3.3 Summing up.....	76
4 The argument from distrust & the active/stative distinction	77
4.1 The Commitment account.....	77
4.2 An initial problem with COMMIT: non-direct commitments	78
4.3 Non-reliance is unnecessary for distrust	80
4.4 The argument from distrust.....	83
4.5 REL+ conflates an active and a stative sense of ‘trust’	84
4.5.1 It is possible to actively trust and distrust at the same time.....	84
4.5.2 Activities, states and the progressive aspect.....	85
4.5.3 Why suppose that ‘trust’ can be both reliance and assurance?.....	87
4.5.4 Setting out the new picture of the trust stance: active and stative layers.....	90
4.6 Active/stative in the literature	91
4.7 Responses to the argument & stative distrust	94
4.7.1 COMMIT _{DIR} is not satisfied in the example cases	94
4.7.2 NON-REL is not satisfied because risk is unnecessary for reliance	95

4.7.3	NON-REL is too underspecified / is an activity	102
4.7.4	Distrust is not wariness	103
4.7.5	Qualifying wariness	105
4.8	Summing up	106
5	Reliance	107
5.1	Negating reliance and negating predictive attitudes.....	107
5.2	The acceptance view and anticipation	110
5.3	Smith's 'plan-theoretic' account	112
5.3.1	Smith's account of reliance.....	113
5.3.2	Internal & External reliance	115
5.3.3	Promiscuity: past, background, negative states of affairs.....	117
5.3.4	Luck cases: a dilemma for Smith's account.....	121
5.4	Reliance as role placement in an activity	124
5.4.1	States of affairs on the role placement view.....	130
5.4.2	Accounting for intuitions & problems in other accounts	132
5.4.3	Progressive and habitual reliance	136
6	Dependence	139
6.1	Rely / depend	139
6.2	Successful activities / proper functioning.....	140
6.3	Needs & concerns: instrumental & non-instrumental (fundamental)	142
6.4	Need & concern routed in reliance & dependence	144
7	The trust stance: the rolling schema	148
7.1	Active/stative & the voluntariness problem.....	149
7.2	Active/stative vs. two-place/three-place	152
7.3	Perception colouring	157
7.4	Situational vulnerabilities & 'entrusting': the structure of matters of trust..	160
7.5	Prior assurance trust as a reason for active trust	166
7.6	The stative rolling schema: retrospective, counterfactual, hypothetical.....	170
7.6.1	Retrospective trust stances	172
7.6.2	Putting pressure on the Commitment account	174
7.6.3	Counterfactual & hypothetical trust stances.....	178
7.7	The rolling schema model.....	179
8	The interpersonal import of matters of trust: attachment.....	188
8.1	Recap: betrayability.....	188
8.2	'Right kind of motive' views	189
8.3	The con-artist & (non-)instrumental concern over motives.....	190
8.4	Interpersonal attachment	192
8.5	To the group	196
8.6	Betrayability varies with attachment concern, not with stative trust.....	202
8.7	Summing up	204
	Conclusion.....	206
	Bibliography	207

Word count: 79,362

Abstract

This thesis defends an account of what it is to trust other people, and what gives matters of trust a characteristic interpersonal or normative importance to us. Trust is an attitude of the ‘trust stance’; a more general attitude we take toward others in matters of trust, that includes distrust. Matters of trust are situations we trust/distrust others in. I put forward an account of the trust stance, that explains why matters of trust have interpersonal importance to us.

Chapter 1 introduces the key questions to be addressed by the account. I outline how trust can be tied to specific actions, but can also be a general attitude we have about a person, or people. I set out how trust is standardly conceived as an anticipatory/predictive attitude, that also involves interpersonal import. That import is glimpsed in the possibility of betrayal by those we trust, and I point toward existing accounts of betrayal. I present arguments against accounts of trust that take it to be purely predictive, i.e. those of the rational choice/game-theoretic tradition.

Chapter 2 introduces the dominant philosophical view of trust, which holds that to trust is to rely on another, such that we can be betrayed by her. I call this the ‘Reliance plus’ (REL+) view. I offer a critical overview of three prominent REL+ accounts, from Baier (1986), Holton (1994), and Hawley (2014). I illustrate how an account of distrust that Hawley endorses, of ‘betrayable’ non-reliance on another, results from REL+.

Chapter 3 presents an argument against REL+. I argue it cannot allow for the possibility of uncertainty about another, where uncertainty is a trust stance attitude between trust and distrust. Uncertainty is possible, so REL+ must be false.

Chapter 4 presents another argument against REL+. The argument is that distrust cannot be a product of non-reliance, so REL+ must be false. I argue that REL+ fails because it ignores a distinction between two senses of ‘trust’: an activity of reliance, and a mental state of assurance. Distrust is *only* an attitude of wariness, opposed to *assurance*, rather than reliance. I defend the claim that reliance requires practical dependence on what is relied on. I build upon in this claim in the next chapter.

Chapter 5 defends an account of reliance as an activity, in support of the active/stative trust distinction from chapter 4. I evaluate Smith’s (2010) account of reliance, which endorses practical dependence. I argue that Smith’s account faces a dilemma, showing the account is either incomplete, or that it renders reliance impossible. I defend an ‘role placement in activity’ account of reliance, that avoids the dilemma.

Chapter 6 defends a distinction between reliance and dependence in general. Where reliance involves practical dependence, I argue that dependence is a matter of fundamentally needing something as a matter of functioning and wellbeing. My account of the concept comes into play in chapter 8.

Chapter 7 sets out a more detailed account of the ‘stative’ trust stance attitudes. I use the active/stative distinction to address a question over whether we can trust voluntarily, and the relation between specific and general trust. I set out the concept of a situational vulnerability, that the trust stance attitudes are ‘about’, and which can result from reliance on another. I defend an account of the trust stance as a ‘rolling schema’: an anticipatory framework that involves interpreting another’s motives toward us, in respect of situations of vulnerability.

Chapter 8 argues that the interpersonal import of trust is a product of our felt need for secure attachments to individuals, and to belong to a group. I explain the relationship between social dependence on others and ‘betrayability’.

Andrew J. Kirton, PhD Philosophy, “Matters of Interpersonal Trust”. 2018
The University of Manchester

Declaration

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright Statement

- i. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the “Copyright”) and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.
- ii. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.
- iii. The ownership of certain Copyright, patents, designs, trade marks and other intellectual property (the “Intellectual Property”) and any reproductions of copyright works in the thesis, for example graphs and tables (“Reproductions”), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.
- iv. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property University IP Policy (see <http://documents.manchester.ac.uk/display.aspx?DocID=24420>, in any relevant Thesis restriction declarations deposited in the University Library, The University Library’s regulations (see <http://www.library.manchester.ac.uk/about/regulations/>) and in The University’s policy on Presentation of Theses.

Acknowledgements

Thanks to my friends who non-consensually lent their names for my example cases.

My supervisors: Thomas Smith, Joel Smith, Paula Satne. Tom: endless thanks for the years of support, critical insight, creativity, thought provoking discussions and direction. It'd be impossible to identify specific points in this thesis where Tom's help contributed. Tom's 'Promises and Obligations' module on my masters course was responsible for getting me obsessed with the subject of trust in the first place, so he is really to blame for this. Joel's guidance whipped me into shape (or a mould approximating one) at various points throughout the PhD, and gave me more motivation to take myself seriously as a person worth listening to on the subject. Paula for her support and guidance on my work in the final year. Thanks so much to all three of you.

My mum Angela and dad Michael for looking after me during the final two weeks. And, for the prior 30ish years. Their values, insights into and curiosity about human nature are woven into this. Thanks for the relatively secure attachment, you two. Also my brother Liam and his partner Catherine Smith (who came to the rescue after a laptop charger failure in my last 48 hours!), and my sister Naomi and her partner James Pickup, for helping support that sense of security and encouraging me. I should be able to spend less time staring into the middle distance on family days out now guys.

Justina Berskyte for her love and support, especially in the highly stressful final months. Also for giving up so much of her time to proofread drafts in the last month. Her terrifyingly sharp brain helped refine this in many places and I can't thank her enough.

Jack Casey for many discussions in the past three or so years, directly and indirectly about trust, about the madness of humans/the world at the minute. Also for particularly helpful input in setting out the role of 'the middle ground'/uncertainty section. Jack reassured me of its argumentative force. For also proofreading during the final weeks. 10/10 Would PhD with again.

Maeve MacPherson for her support, encouragement and laughs throughout the PhD and since the masters course in 2012/13. Thank you for the many times you helped me (/pushed me) in matters of anxiety and mental health.

Leo Tarasov and Rob Knowles for much support, encouragement and advice on how to navigate this mad undertaking over the years. Your cheering me on (figuratively, via electronic messages) in the final couple of weeks was a great boost and helped me cope with the terror.

Pilar Lopez-Cantero for her encouragement (getting me applying for way more things than I otherwise would have), laughs, insightful thoughts about love and relationships, even if I spent most of those discussions saying 'it's all about attachment!' (thanks for humouring me).

The wider Manchester Philosophy PhD crew 2014-2017 for encouragement, insightful comments, and great laughs about lots of things, particularly Abigail Connor, Chris Ovenden, Fred Horton, Sylvia Barnett, Joey Montgomery, Jamie Henson, Andy

Routledge, Aaron Wilson, Nathan Duckett, Wen Chin Lung (honorary crew member).

Ana Alecsandru for indulging mutual trust obsessions in many great discussions, and kindly letting me join in her world of trust in International Relations. Getting to talk about this stuff to people in other disciplines (non-philosophers!) really helped my ideas.

Graham Stevens and Phil Meadows for their comments and encouragement on early drafts of chapters. The further Manchester Philosophy staff who supported/encouraged my work and offered suggestions at various points: Helen Beebee, Julian Dodd, Catherine Abell, Ann Whittle, John O'Neill, David Liggins.

Katherine Hawley for her many insights into the subject, and her encouragement and helpful comments on earlier drafts of work. Her work along with Annette Baier's has been absolutely pivotal and inspiring, and I wouldn't have arrived at the ideas I put forward here without it.

Monique Wonderly whose work on attachment I found at a late stage, and which helped slot in a puzzle piece I had been trying to mash into place. I had been struggling to find any philosophical treatments of attachment, and hers appeared at just the right time. Many thanks to her for being so helpful when I got in touch, and kindly sending her PhD thesis and then-forthcoming papers on to me.

Non-philosophy friends who I have bounced ideas around with, whose support has kept me sane, and who stopped me from being a total hermit:

Ali McDowall for great laughs and lunchtime chats (a lot of which were about Bob's Burgers) during the years working on this, and for many intuitions and insights about the underlying emotions that drive people, that I have leeches from. For also being very supportive and giving helpful practical advice in matters of dealing with anxiety and obsessive perfectionism.

Ruth Wilkinson for years of support/encouragement/laughs/telling me to stop worrying about things so many times I have lost count.

Lawrie Roberts, Katie Flanagan, Camilla Rosario Smith (our discussions about personality disorders fed into this in many ways), Amy McDowall and Suzy Eyden, who are each fantastic for putting up with me/supporting me in many ways in the last four years.

Shout out to Malvern(/Manchester) hard boys: Matt Mahabadi, Nick Jones, Chris Williams. Never forget Dysons 1999-2006 (I won't due to the lingering PTSD/anxiety). Matt and Chris: our Whatsapp group discussions were a great testing ground for a lot of these ideas.

Dalton Ellis Hall friends: Thank you Amanda Källstig for unceasing encouragement and concern, particularly in the final months. Max Gibbs for his support and understanding when I was super busy. I had a lot of useful discussions on this subject (and laughs) with Amanda, Moozhan Shakeri and Melissa Markauskas. Melissa's perspective on matters of social structures and oppression involved great levels of insight, and helped sharpen up my ideas in this. Thank you Melissa for your support and encouragement.

Dedication

For Annette Baier, who died in November 2012, just before I first read her work. I regret never getting the chance to meet her. Her work has a rare combination of razor sharp insight and air-punching vitality. Go and read her stuff, if you haven't already.

Also for my 5-year old nephew Elliot, and my friend Matt's new baby (Miles). I suspect that if we're going to figure out how humanity can do better, then we need a better grasp on what drives people to work together in doing good things. We also need to know what drives us to fracture, and to do terrible and self-defeating or self-destructive things. It seems to me that trust and distrust play a big part in any answer to that. I hope that Elliot and Miles grow up in a wiser world in that respect.

Table of Figures

Figure 1: The Prisoner's Dilemma.	38
Figure 2: The trust stance and reliance stance under REL+	59
Figure 3: Trust/distrust and neither-nor under REL+	62
Figure 4: Denying uncertainty is distinct.....	69
Figure 5: The third category response to uncertainty.....	75
Figure 6: Active and stative trust under the trust stance.....	90
Figure 7: Active/stative & Two-place/Three-place trust	155

1 Matters of interpersonal trust

In this chapter, I introduce the central question this thesis seeks to answer: what it is to trust or distrust another person on some matter. I also give an overview of the central philosophical issues that arise in addressing this question, and the different positions in the trust literature taken on these issues. As part of this, I also offer an overview of the positions I endorse, and the key concepts I appeal to, in the overall account of interpersonal trust and distrust that I defend in this thesis.

1.1 Matters of trust & the trust stance

What is it to trust or distrust another person? The presence of other people impacts upon us in a certain way. As a result, we must trust or distrust others sometimes. In a situation where the question of whether to trust or distrust another is at issue, we are confronted with a *matter of trust*. The matter of trust is the situation or issue itself. In these situations, we respond to or regard others with an attitude of trust, distrust, or some shade in between these. These attitudes belong to what I call the *trust stance*. This includes trust and distrust as the opposed ways we respond to others in matters of trust. Between trust and distrust though we also have a ‘middle ground’ attitude: a feeling of being unsure, unsettled, or uncertain about another.

These trust stance attitudes share a characteristic concern about how another will respond in a matter of trust. However, they each involve a different way of anticipating how another will respond; with trust involving anticipating the matter as settled, distrust involving anticipating the matter as unsettled and decidedly fraught, and uncertainty as anticipating it could go one way or the other. This gives us two key concepts that will shape the enquiry. A *matter of trust* as the situation or issue, and *the trust stance* as the general stance we take in matters of trust; that involves attitudes of trust, distrust and uncertainty.

There is a characteristic concern we have about other people in matters of trust. That concern is not a purely theoretical concern, in being a matter of what we ought to believe about others, or what we can predict they will do. It is also not purely practical; a matter of what we ought to do. Because we have the capacity for interpersonal trust and distrust, matters of trust have a kind of import. The trust stance attitudes are ways of adopting a

position on the question: ‘can I live alongside this person?’. This question is made pressing when we recognise that, as Annette Baier puts it, those we live alongside have a power over the things we care about, such that these things, “can be left or put within the striking power of others” and may be taken advantage of (1986 : p.235). The things we care about will broadly include the projects and goals we want to pursue, which stem from the people and objects and other things we are attached to. What we care about will also include our own sense of efficacy or agency in looking after the things we care about. It may be that others we live alongside will simply not treat the things we care about with enough due concern, or competence. They may also not let us exercise our independent agency in taking care of some of those things ourselves. Those things we care about, in falling under the control of others, may thereby come to harm, through their recklessness or carelessness, even if this carelessness isn’t borne out of any specific malice toward us. By asking ourselves whether other individuals we encounter will respect our desire to determine our own ends, and help us pursue our projects, and to preserve and nurture what we are attached to, we thereby ask ourselves whether we trust those individuals.

Matters of trust thus have a kind of interpersonal import. This thought is the starting basis for many views of trust and distrust in the philosophical literature. Typically, views of trust proceed by identifying some sort of basic attitude for the anticipatory element of trust, and giving an account of the interpersonal import. Before discussing ways the anticipatory and interpersonal elements are accounted for (sections 1.6 and 1.7), though, I will first cover some of structural features of matters of trust, and the trust stance attitudes.

1.2 Domains of trust & competence/willingness

Matters of trust tend to be about the possibility of another’s performing some action in some particular scenario. As a result it is common, in the trust literature, to endorse a view of trust as essentially a three-place predicate (Baier 1986; Castelfranchi and Falcone 2010 : p.29; Faulkner 2007a; Hawley 2014 : p.1; Hieronymi 2008; Holton 1994 : p.67; Horsburgh 1960 : p.343; Jones 1996, 2004; Levi 1998; Simpson 2011 : p.327). Typically, this will mean analysing the trust attitude as one such that X trusts Y, in respect of some action φ . Matters of trust, on this model, are where the performance of some action is at issue, in a particular situation, or *domain of interaction*.

This makes sense: it is commonplace that we trust people in specific respects and not others. Just because we trust someone, this doesn't mean we trust them infinitely, across all domains of interaction that we could possibly have. Often, that we do trust someone in some respect, is tied to the fact that another's competences lie in certain areas, and we trust her in those areas. We trust people in respect of their areas of competence, e.g. a midwife during childbirth, a music journalist in recommending records, a friend in being able to look after our cats while we're away. Given that matters of trust are domain specific like this – we apportion our trust to people in certain matters – the three-place predicate, tied to an action, seems highly plausible, as a formalisation of the attitude of trust.

It isn't generally enough to make some domain a matter of trust, just based on a recognition of competence in that domain. Trust of another in some domain requires attributing some sort of willingness on her part to perform. While the midwife might be an expert on helping others through the birthing process, and the music journalist might be capable at identifying records particularly worth listening to, expertise is nothing without the willingness to follow through on the ability to help. The jaded music journalist may, despite having gained her position through having good taste and expert analysis, have become unwilling to bother recommending good records, instead promoting whatever releases big labels pay her to promote. Similarly, the friend who could look after the cat, might be lazy. A lack of motivation to act will thus defeat the possibility of making that a matter of trust. Because of these thoughts concerning competence and willingness, authors therefore suppose that matters of trust revolve around another's competence and willingness to act (e.g. Baier 1986 : p.259; Hawley 2014 : p.1; Jones 1996 : p.4; Rowe and Calnan 2006 : p.377).

1.3 Two-place

While it is standard to treat trust as a three-place predicate, and matters of trust as tied to specific actions (e.g. feeding the cat, delivering the baby, recommending a good record), we do nonetheless talk of trusting people more broadly than in one specific respect. We can, it seems, simply *trust* others. We can construe trust *simpliciter*, as a two-place predicate such that X trusts Y (Becker 1996 : pp.44–45; Faulkner 2015 : p.424). Baier calls this form

of trust 'plain trust' (and 'plain distrust'), since it isn't relative to some particular matter (1986 : p.259). You might for example say that you simply trust your friends or colleagues, where this doesn't appear to amount to saying that you trust your friends or colleagues to do specific things.

We can also express a strong kind of two-place trust. Horsburgh, for example, says, "The [general] sense is much wider. It may be illustrated by the trust in B that A expresses when he says, 'I'd trust him with my life'. It implies absolute confidence in B's good faith." (1960 : p.343). Arguably, of those we say we generally trust, where we leave the domain unspecified, we only trust a few in the strong sense Horsburgh alludes to. The kind of deep two-place trust that Horsburgh raises implies a level of intimacy with the other.

There are various positions in the trust literature on how we ought to understand the relationship between two-place and three-place trust. Those who view the three-place form as in some sense primary, sometimes suggest that even though we might say of another that we simply trust her, this can be elliptical for saying that we trust another with a range of more specific things (Smolkin 2008 : p.446). The point noted above, that just because we trust someone we don't thereby trust them infinitely, may lend credence to this position. Faulkner emphasises that even when we might be said to trust another generally, there are always limits to what we trust her to do (Faulkner 2015 : p.425).

Of those who we trust generally, such as our friends and colleagues, we clearly do not trust them in *all* matters of import to us, because often the respects in which we do trust them are limited to the contexts and activities we share with them. As such the 'elliptical' interpretation of the two-place formulation, may account for such general trust. It may, for the same reason, account for deep trust of the 'I'd trust him with my life' sort. Even those we trust deeply, such as (in an ideal world) our close friends, family, and romantic partners, we don't trust on all matters of importance to us, e.g., in recommending good records, or to fix our boiler when it breaks. We apportion trust, even in these two-place cases, based on where another's competence and willingness lies, and everyone's competence and willingness has limits. Our general/deep two-place trust needn't become contaminated with distrust because of the fact that there are limits. It is simply that certain matters ought to be closed off, as avenues on which you *can* or should adopt a trust stance

on those matters. Otherwise, you will just be led to disappointment, possibly through no fault of the one you trusted.

We typically expand our trust to the limits of what we recognise another is competent in, and willing to help us out in. Even then, of the trust within those limits, it is still sensible to evaluate what others say or do. This isn't because we should necessarily be constantly wary of the motives of those around us, but because everyone can make mistakes on such matters. When your friend insists that it was at her birthday *two* years ago, not *three* years ago, where you had that embarrassing incident involving the cup of tea and her mother, you can fail to trust her on this, because you are convinced that it was in fact three years ago when that incident happened. That you disagree needn't imply you think your friend is a liar, and out to mislead you. She may just be convinced you are wrong.

There are further ways of drawing the distinction between two-place and three-place trust, though, and these would cast doubt on the claim that two-place is elliptical for three-place trust. Holton for example argues that two-place trust is really describing a phenomenon that he calls a 'trusting relationship', and this appears to involve more than just having a set of domains of interaction, in which we three-place trust another:

If you and I trust each other in various ways over time, and our trust is not betrayed, we will be likely to build a trusting relationship. That is not to say that there will be some particular thing that I will trust you to do. Rather it is to say that I will in general be more ready to trust you: partly because I am confident that you will not betray that trust, and partly because, having trusted you before, further trust becomes appropriate. A trusting relationship makes a greater range of trust available to me. (1994 : p.71)

The crucial thought in this passage, that casts doubt on the 'elliptical' proposal, is that when we two-place trust another, we are readier to trust her in specific three-place respects. In two-place trust of the sort Holton describes, we feel as though we have stepped beyond a set of specific domains of interaction, and now are 'comfortable' with the other person. The trust we extend to the person, instead of being confined to one domain, now permeates or bleeds into other possible domains. This sense of cross-boundary trust is captured in Horsburgh's 'I'd trust him with my life', noted above. It implies that, given some situation we may not even have thought of, even one where we are at the gravest sort of risk, we would have absolute assurance in his competence and willingness to help us. Two-place isn't elliptical for a range of three-place trusts, if it

describes a *disposition* to trust in specific three-place respects, as many of these we won't even have considered yet.

1.4 Basic/one-place & voluntariness

Also opposed to the 'elliptical' position is Hertzberg's (1988). His view takes inspiration from Wittgenstein's remarks in *On Certainty*, concerning the central role of trust in learning about the world. Wittgenstein focuses particularly on the kind of dependence that a child has on adults to induct her into language use, and forming beliefs about the world. Hertzberg builds on Wittgenstein's remarks to argue that trust is a more basic attitude involving a readiness to believe others, not on a particular matter, but across matters. Again, this is much like Holton's notion of a trusting relationship: being readier to trust a particular person in specific matters, as glimpsed in the kind of faith expressed in Horsburgh's 'I'd trust him with my life'. But, Hertzberg is pointing out a different form of trust to that Holton and Horsburgh point out. He is unpacking a more fundamental kind of trust; not one that we have to particular people, but that we have toward people in general.

This type of fundamental trust is central to our lives, because dependence on other people is an unavoidable feature of our lives (I return to this point shortly). For *dependents*, e.g. children, this dependence is immediate and ever-present. Children have little choice but to trust by default. It's through this that children develop basic abilities to help them navigate the world. As children, we learn what particular words refer to, what is safe to eat, who is safe to approach, and how we conduct the most basic tasks involved in feeding, cleaning, and looking after ourselves.

But, we also have little choice but to trust the adults we find as guides. Small children are often not in a position to pick and choose who they trust on what matters: they have to be willing to trust adults across matters. But part of growing and maturing - becoming an independent, empowered person - is learning *who* we should trust in *what* domains. We learn where it is safe to make roots, thereby anchoring us in a group of individuals we can feel safe around.

By paying attention to the basic sort of dependence on others found in childlike trust, we also glimpse the fact of our ongoing dependence on the community around us as adults. We can't simply opt out of belonging to *some* community for our language, our knowledge, our wellbeing, or help in pursuing many of our goals. Life without this basic trust is not simply the Hobbesian picture of a state of war of every person against every other person; it is one with the most rudimentary form of language, and a drastically limited worldview.

The basic form of trust we have toward our community, is discussed by Herman (2001) and Jones (2004). Their concern with basic trust is in how it becomes disrupted by traumatic events. Jones points out that there must be a basic form of trust we have toward people in general, underpinning the trust we have to particular individuals and in specific matters, because this is the very thing that is shaken in the indiscriminate violence of terrorist attacks.

Herman's concern is to offer a psychotherapeutic method for overcoming trauma. As part of this method, we glimpse the importance of basic trust toward others. Herman argues that what is key to overcoming trauma is restoring our basic trust in others. Part of this process requires the suffering and status of the victim, as someone wronged, to be validated by her community. Without this, the trauma victim cannot readily trust others, because her sense of threat is 'stalled'; stuck on playing back the traumatic event, and trying to integrate it or understand it, in light of her need to belong to that community. Herman says that, in abuse and trauma, "the profound disruption in basic trust, the common feelings of shame, guilt and inferiority ... all foster withdrawal from close relationships. But the terror of the traumatic event intensifies the need for protective attachments." (Herman 2001 : p.56). A restored sense of basic trust follows, once the community validates both that a victim did experience a traumatic event, and that her experience is a matter of concern for the community.

So, it seems that a willingness to two-place trust particular people, is underpinned by a more basic or fundamental type of trust toward people in general, or a community, which can be profoundly shaken as a result of trauma. This basic trust underpins our willingness to two-place individuals in a community, and thus to three-place trust others on specific matters. While basic trust is brought into focus when it is disrupted by trauma, we also glimpse it when we simply reflect on the mundane aspects of community living.

Community living is made possible by this basic, unconscious trust we have toward others. Thomas illustrates how it underpins day-to-day interactions in communities:

We need to be able to have basic trust (as I shall say) in complete strangers. Much of what we do in life it would be well-nigh impossible to do if we could not trust many who are complete strangers to us. We would scarcely order a meal in a restaurant or ask a stranger for the time or walk down the street if we thought for a moment that every stranger in life was out to harm us - if we lived in a world in which basic trust among all complete strangers was not possible. (Thomas 1989 : p.34)

These thoughts echo Hertzberg's position. Without being able to trust others at the most ordinary level required of being able to live with others - what certain words mean, what the world is like, what other's motives generally are - we can't engage with others, or live alongside them in a community. This is why basic trust being disrupted by traumatic events is so damaging: it fosters withdrawal from a community that helps us to live comfortably, and moreover, *understand* the world around us.

Along these lines, Govier (1992a : p.55) points out how, when that basic trust is replaced with basic distrust, our sense of reality can become entirely warped. An inability to treat any reported fact as settled, or anyone's motives as generally honourable, leads to constant second-guessing. We start to constantly scan offers of help from those around us for hidden motives.

It's plausible to conclude, from these brief reflections on the relationship between the fundamental need for basic trust, that when we find ourselves with a basic distrust of the majority, we cling to those individuals who we do two-place trust. It follows that isolated worldviews could grow around small besieged-feeling groups, containing individuals who only trust one another, while distrusting others beyond them. Furthermore, as Herman points out, the terror of widespread distrust intensifies the need for protective attachments. So, even if these islands of two-place trust foster unhealthy or self-destructive worldviews, the distrust of those outside of their small community, will keep them holding firm to the worldview shared by those they trust.

So, just as we have seen that two-place trust is not simply elliptical for a set of three-place trusts, it appears that there is a basic trust that is not elliptical for a set of two-place trusts. We could also call this basic trust 'one-place'. Faulkner (2015), for example, refers to it in

this way. He suggests that one-place trust picks out the disposition to two-place trust others, rather than being elliptical for saying whether we already do two-place trust a range of specific individuals. My ‘besieged-feeling community’ example would explain why: the individuals in this community are not one-place trusting, even if there is a set of individuals they two-place trust.

At this point, a question is raised concerning how much we can *voluntarily* trust others. It’s commonplace, for example, that people offer imperative phrases such as “Trust me.” But the question arises as to what the function of such imperatives are, and how we ever choose to trust another on a specific matter, if we do not already two-place trust her. We’ve seen above that two-place trust appears to inform our willingness to trust another across different matters. Furthermore, we will be unable to trust another in a two-place way, unless we already have some basic trust. Lagerspetz (2015) raises a similar thought, in arguing that Baier’s (1986) three-place account of trust misses the fundamentality of basic trust; “not all trustful relations seem to be goods-related in this way,” and the starting point of an account of trust should be ‘simple’/basic trust, “considering that we are unlikely to trust people *with* anything unless we trust *them*” (Lagerspetz 2015 : p.51).

I will be arguing for a view of the trust stance where three-place trust is not contingent on prior two-place trust of another, nor on prior one-place trust of others. In chapter 4 I will argue for a distinction between two senses of ‘trust’ that sheds light on why. One sense of trust is an activity that is within our control to undertake, and another is a mental state of assurance about another. My suggestion is that while we *can* voluntarily undertake the activity of trusting another, we *cannot* voluntarily adopt the mental state of assurance that another will come through for us. In other words, we *can* decide to put ourselves in a vulnerable position toward another, but we *cannot* decide to feel fine about that. The process of rebuilding two-place trust and even basic trust, involves rebuilding the state of basic assurance about others. We do so via specific individuals. This, furthermore, involves undertaking three-place active trust of another, even when we lack the mental state or feeling of assurance that she will fulfil it.

In sections 7.1 and 7.2 I discuss the voluntariness question and its relation to the widely recognised two-place/three-place distinction. I will argue that the voluntariness question

is better resolved by the active/stative trust distinction I put forward, rather than the two-place/three-place one that is commonplace in the literature.

For now, I will point out that there simply is the possibility of voluntarily trusting another. Lagerspetz is right to point out that three-place accounts like Baier's miss the fundamentality of basic trust. However, while it is right that we might be less likely to trust another *with* anything (three-place), unless we already do trust them (two-place), it is not required that we do. Baier (1986 : p.259) notes, for example, that we can have no choice but to trust others sometimes on specific matters, when all other options have exhausted themselves. I may desperately need someone to drive me to A&E, and the only person on hand to help is my next-door neighbour who I distrust. Still, I am desperate. Should my neighbour fulfil my trust in this matter, I may come to change my view of her. Holton (1994) also points out instances where we can voluntarily trust, despite lacking assurance that our trust will be fulfilled. While Lagerspetz is right to note that it is sometimes unlikely we will trust people *with* anything unless we trust *them*, it is not impossible or unreasonable to trust people in a three-place way, despite lacking the more general trust picked out by the two-place form, or even the basic/one-place form. Trusting another in a three-place way, beyond our two-place or even one-place form, is a way of rebuilding trust.

1.5 Practical & testimonial

Here I draw attention to a distinction between types of trust that has already come into play. This is in terms of a distinction between broad categories of domains of interaction. Just as we trust others to perform specific actions based on a supposition of competence and willingness, we also believe what others say on this basis. When you fail to trust your friend that the embarrassing incident (involving you, the cup of tea, and her mother) was two years and not three years ago, or when the schoolchildren no longer trust what the teacher tells them, what you and the schoolchildren are failing to do is trust the assertions or testimony of another on some matter. Again, this failure to trust needn't be because the other is suspected of lying, such that you distrust her. It could be that you suspect she is mistaken about a fact of the matter. Nonetheless, you fail to trust her.

The distinction at issue is this: just as we can specifically/three-place trust another to perform certain actions, we can also trust her *assertions*, or her testimony. Testimonial trust (or interpersonal epistemic trust) is the subject of a vast literature, e.g. (Adler 1994; Audi 2006; Coady 1994; Faulkner 2007b, 2011, Hinchman 2005a, 2005b; Lackey, J. & Sosa, E. (eds.) 2006; McCraw 2015; Zagzebski 2012). Testimonial trust is often discussed as a distinct phenomenon from trusting another in some specific action. In trusting another to perform a specific action, X trusts Y to φ . Instead, accounts of testimonial trust are concerned with cases where X trusts Y that p , whereas p is what Y asserts.

No doubt there are key differences between the two phenomena. Practical trust involves trusting Y to bring some desired state of affairs about, as a result of her action. Testimonial trust on the other hand involves trusting Y that some fact about the world holds (that Y may or may not have had a hand in making true). But while the distinction between testimonial and practical trust is worth noting, they are not the completely isolated phenomena that the division in the literature between discussions of each, would have us believe. For one, when we trust Y that p , and when we trust Y to φ , it's true that in both cases we treat that matter (whether p , whether Y will φ) as *settled*. We would anticipate that p is the case, and that Y will φ . The same type of positively anticipatory attitude is at issue in either case. This is also true when we fail to trust Y that p is the case, or that she will φ : we regard the matter as unsettled. When we distrust Y, and she tells us p or that she will φ , we may scan for ways that ulterior motives could be behind her proposal. I also take it that in both cases of trusting testimony and trusting another to act, these are matters of import to us, partly because of the practical consequences that can befall us. Our own welfare can be at risk in either case; either from misinformation about how the world is, or from Y's possible failure to φ , when Y's φ 'ing matters to us.

A further point to note is that in any given scenario, practical and testimonial trust are intertwined and enable one another. They simply do not occur in isolation of one another. For one, testimonial trust also involves trusting another in respect of a certain action: telling the truth. X trusts Y to provide accurate information, or to be honest and competent. But similarly, 'practical' trust, when directed toward future actions, often involves trusting the testimony of another that she *intends* to perform a specific action. If we are trusting another that she *has* φ 'd, this also involves trusting her testimony, about the past action she has completed. I take it that, in any given scenario, the two forms of

trust are so closely intertwined that it would be misguided to discuss them as isolated phenomena. An account of matters of trust and the trust stance ought to pay attention to both.

In section 7.1 I discuss an example of testimonial trust, in the context of the puzzle of when trust can be voluntary, that I raised at the end of the previous section. It appears that testimonial trust appears, in one way, to be unable to be voluntarily given, unlike practical trust. But, I will use the active/stative distinction that I argue for in chapter 4 (mentioned at the end of the previous section), to explain why this appears to be so.

1.6 The anticipatory element & the interpersonal element

In the first section, I referred to how the trust stance attitudes had an anticipatory element. I have repeatedly made tacit appeal to that anticipatory quality of trust, when I have said that trust involved anticipating some matter as ‘settled’. We can now flesh out our understanding of the way in which the trust stance involves an anticipatory quality.

In very simple terms, trust involves anticipating good things from another. And correspondingly, distrust involves anticipating bad things from another. When we are somewhere in between trust and distrust of another we may not be able to anticipate one way or another, despite trying to. Trust and distrust involve opposed ways of anticipating another’s actions within a matter of trust. What is common to the trust stance attitudes then is an anticipatory element.

In the trust literature, this anticipatory element is pinned to various ‘predictive’ type attitudes. Some authors identify the anticipatory element along the lines of what philosophers call a ‘cognitive’ or belief-like attitude. Here the anticipatory element of trust (/the trust stance) is construed a type of subjective probabilistic expectation of another’s performance (Dasgupta 1988; Gambetta 1988), as a type of belief (Castelfranchi and Falcone 2010; Hardin 1998; Hieronymi 2008; Hill and O’Hara 2006; Keren 2014; Levi 1998), and as a type of confidence (Pettit 1995). In contrast there are accounts that identify the anticipatory element as a product of ‘non-cognitive’, or affective attitudes. Here we have the anticipatory element of trust (/the trust stance) construed as a type of optimism (Jones 1996), an emotional attitude comparable with faith (Lahno 2001), a form

of hopefulness (McGeer 2008), a feeling of security about another (Becker 1996) and as a Strawson-style (1974) reactive attitude we take toward others actions (Helm 2014). This is just a selection of accounts to help give an impression of the variety of ways that the anticipatory element of the trust stance is understood.

The anticipatory element is not exhaustive of the trust stance (though this is disputed: see section 1.8 below). As mentioned above, matters of trust and the trust stance attitudes we take concerning them, possess a characteristic interpersonal importance. That interpersonal importance is defined in various ways (discussed in more detail in the following section), but the central thought is that we take matters of trust seriously in a way beyond merely predicting someone will/won't perform some action. As such it isn't enough to predict or anticipate that someone will act, to have a trust stance attitude about her. Simply having a predictive attitude wouldn't be enough to make the matter of her possible acting a matter of trust. This is obvious for one broad category of actions: those that we are indifferent to. We don't treat it as a matter of trust that our neighbour gets in her car and drives to work in the morning, because it doesn't matter to us whether she does.

Even if we stipulate that what is at issue is whether the other's acting would be instrumentally beneficial for us, this still doesn't suffice for making it a matter of trust that she does so act. And, likewise, it isn't enough to predict that someone *won't* perform the action that would be instrumentally beneficial for us, or to predict that she will act in some way that would be detrimental to us, for us to thereby distrust her.

This is the thinking behind a dominant philosophical view of trust that I will be discussing more in depth in subsequent chapters, particularly 2, 3 and 4. This is the 'reliance plus' (REL+) view. This view stems from Baier (1986). Baier suggests that trusting someone is like relying on her in some respect, but it isn't *just* relying on her. It is relying on her with the susceptibility to being betrayed or let down by her (1986 : p.235). We could for instance, rely on our neighbours' regularly-timed leaving for work in the morning, as a guide for own morning schedule. This wouldn't entail we trust our neighbour to leave for work at that time. The lesson from such cases is that trust is reliance on another's acting, plus an interpersonal factor that allows us to be *betrayed* by the one we're relying on, in case of her non-performance. In other words, it isn't enough to risk a goal of yours on

the possibility that someone acts, for you to thereby trust her to act. Instead, there is an interpersonal importance to matters of trust. Baier's key observation is that this importance can be gleaned from the difference in reactions to disappointed trust (feelings of wronging, betrayal, being let down, interpersonal disappointment) and frustrated reliance (simple frustration and disappointment).

REL+ is a pervasive view in the trust literature. Yet considering the accounts referred to above, we can see that *reliance* is just one way of construing the anticipatory element of the trust stance. But because it is so widespread, REL+ will be the target of my enquiry in ensuing chapters. Using it as a foil, I will argue that *reliance* cannot fully account for the actual nature of the anticipatory element of the trust stance. Interrogating the view will prove to be philosophically productive, providing a rich seam to mine in revealing what the nature of the trust stance attitudes are, and so what is really at issue in matters of trust.

REL+ takes seriously the thought that there is an interpersonal importance to matters of trust. Many of those accounts of the anticipatory element of the trust stance, mentioned above, also take this thought seriously. REL+ provides one way of drawing attention to that interpersonal importance; by drawing a contrast between 'merely relying' on another, and trustingly-relying on her to act. Yet this is just one way of drawing the distinction, just as *reliance* is one way of accounting for that anticipatory element. There is a difference between a generic anticipation of another's acting, and trusting/distrusting her. I offer some simple contrast cases of trust and mere anticipation about the possibility of another performing an instrumentally beneficial act, to illustrate that the 'mere'/'interpersonal' contrast is a generic one.

Food Stealing Flatmate. John anticipates that his flatmate Elis will keep putting food in the shared fridge that can be taken without Elis noticing.

Ailing Flatmate. Maria is ill and contagious, but her flatmate Izzy has told her not to worry and that she will do her food shopping for her until she gets better. Maria is unwell, but also anticipates that Izzy is true to her word, so gives Izzy her debit card and pin number for the time being.

Food Stealing Flatmate is a case of ‘mere anticipation’ of another’s instrumentally beneficial act, while *Ailing Flatmate* is a case of trusting another to act. Maria and John could be relying on their respective others (Izzy and Elis), or have one of the other types of anticipatory attitudes referenced above. The key intuition is that in *Food Stealing Flatmate*, John’s anticipatory attitude (toward Elis’ continuing to put food in the fridge) doesn’t have the requisite kind of interpersonal salience to make the matter of Elis continuing to put food in the shared fridge a matter of trust. In *Ailing Flatmate* though, Maria’s attitude toward Izzy does have the requisite kind of interpersonal salience to amount to trust, and Izzy buying the food is a matter of trust between them. So, matters of trust are not simply matters of anticipating that someone will perform an instrumentally beneficial act.

The following cases bring out that it also isn’t enough to anticipate an instrumentally detrimental act in order to distrust another. Again, this is because these are not yet matters of trust.

Wary Flatmate: Elis suspects John’s reassurances about the missing food are lies and that John has been stealing from him.

Cash Stealing Flatmate: Izzy has been exploiting the opportunity presented by Maria’s illness by stealing cash from Maria, and is now stressed and pessimistic, because she anticipates that Maria will figure this out soon and cancel her card.

Wary Flatmate is a case of distrusting another, while *Cash Stealing Flatmate* is a case of merely anticipating another’s instrumentally detrimental act. In *Wary Flatmate*, Elis’ suspicion concerning John clearly does have the requisite kind of interpersonal salience to amount to distrust of John. There is something in addition to Elis’ anticipating that John has taken the food despite his assertions to the contrary, that makes his attitude amount to distrust. On the other hand, Izzy’s anticipating that Maria will cancel her card doesn’t amount to distrust of Maria. The matter of her cancelling her card is not a matter of trust, because it lacks the interpersonal salience that Elis’ attitude possesses.

The important point from each of the above cases is that matters of trust have some sort of interpersonal import. That interpersonal import cannot be purely a matter of whether

the other person performs some instrumentally beneficial or detrimental act, because those are at issue in the Food & Cash stealing flatmate cases above.

1.7 The trust-betrayal connection & vulnerability

The thought that trust is connected to harms such as betrayal is, as mentioned above, the insight that Baier offers to differentiate cases of mere reliance from those of trust. The trust-betrayal connection is widely appealed to in the wider trust literature, and not just in the literature that construes trust as a form of reliance. What it provides is a test of sorts for whether something is a matter of trust, or simply of mere anticipation of another's beneficial/detrimental acting. This also applies with distrust: Hawley points out that the betrayal connection also captures how *distrust* is an attitude with the same characteristic interpersonal import, because we can be betrayed by those we distrust (2014 : p.13). For example, it seems right to say that Elis, despite his anticipating that John has been lying to him, could still be betrayed by John, should Elis discover that John has been lying. Even when we distrust another, we are still responding to a matter of trust.

Because trust and distrust share the ability to be betrayed, it might be thought that, with the trust-betrayal connection, we have the beginning of an account of the interpersonal import of matters of trust and the trust stance attitudes. Matters of trust are those involving a susceptibility to being betrayed. While it is certainly a starting point, the trust-betrayal connection doesn't yet get us an explanation of why matters of trust involve some sort of interpersonal import. Pointing out that matters of trust involve a susceptibility to betrayal gives us a good angle on what sort of import is at issue, but it doesn't offer an account of it. Rather, susceptibility to betrayal gives us a useful placeholder for a condition of what it is for an attitude to belong to the trust stance. To arrive at an account of the interpersonal import though would require looking at the underlying reasons why matters of trust create a susceptibility to betrayal. It would also arguably require looking more closely at what sort of harm betrayal is. With a better understanding of betrayal – what sort of structural features are common to betrayals, what the effects of them are, and what the characteristic thoughts and feelings that result from them – we might be able to shed light on the anticipatory nature, and interpersonal importance, of the trust stance.

Remarkably, despite the widespread adoption of the trust-betrayal connection, it is rare to find a sustained treatment of betrayal in the philosophical literature on trust. I will briefly mention two rare exceptions – Jackson (2000) and Shklar (1984) – who both note that literature and history, rather than philosophy, are more fruitful in attempting to understand the nature of betrayal (drama, whether fictional or non-fictional, high and low brow, tends to revolve around the dynamics of trust and betrayal). Both also offer their own useful characterisations of the nature of betrayal. Shklar for instance, draws on dictionary definitions of betrayal, in order to paint a vivid picture of the emotional turmoil that accompanies betrayals:

Betrayal [...] is to place another person “in the power of an enemy, by treachery or disloyalty,” and also “to prove false to, to disappoint the hopes or expectations of.” This, I think, should include breaking an appointment that means much to the other person, neglecting those who depend on our care, and talking maliciously about our friends. [...] There is, as the dictionary does show us, an irreducible experience in betrayal: desertion. That brings into play the greatest of childhood anxieties, the fear of abandonment. In quitting a bonded group, an equally primeval fear is stirred: of the failure to distinguish kin and stranger, the latter almost always called “enemy” as well. To reject a blood relationship for a new and alien association, or for none at all, is to deny the most elementary of social ties. (Shklar 1984 : p.139)

Where Shklar connects betrayal with neglect and abandonment, Jackson on the other hand suggests that we ought to distinguish between betrayal proper and abandonment. Noting that literature provides more detailed illustrations of betrayal than philosophy, he draws on an example case from *Sense and Sensibility*:

The spirit of instrumentality and deception imbues the entire relationship; it colors all the betrayer’s actions and words. Whatever compliments or tenderness the violator extends to the truster is suspect [...] Further, because of the engagement of betrayers, they operate with an awareness, a foreknowledge of what the inevitable violation will do to the truster. The more that a violation is marked by these features the more we see it as betrayal.

Although Willoughby initially entered into the relationship with this attitude, it quickly vanished as he fell in love with Marianne. Willoughby does not see his trust relationship with Marianne in instrumental terms; his actions are not the calculated manoeuvring of a traitor setting up his victim. What Willoughby does instead is to unilaterally withdraw from the relationship, indifferent to (or at least insufficiently concerned with) the effect on Marianne. He abandons his care of the relationship of trust. While this is clearly a violation of her trust, it is not betrayal. He is not engaged in the relationship to manipulate it; rather, he disengages from it entirely. (Jackson 2000 : pp.85–86)

While Shklar's point seems correct, that in betrayal there is an element of felt abandonment, Jackson's view is also convincing, in drawing a contrast between two different sorts of act; abandonments and betrayals. It seems right to say that there is a difference in kind, based on the different actions and attitudes of the violator of trust, in the run up to that violation. In a case of abandonment, the abandoner will simply withdraw from what was presumably a relationship that was otherwise not marked by deception and manipulation. The abandoner may simply stop caring about furthering the relationship at a certain point, so up and leave. A betrayer however, as Jackson says, seems to do more; to imbue the relationship with prior deception, to lull the trustor into a position where she can be exploited for some gain.

However, Jackson arguably places too much emphasis, in characterising the nature of a betrayal, on there being prior intent of deception on the part of the violator. There needn't be prior intent of deception, to call the violation a betrayal, at least in everyday vernacular. A cheating spouse may have had no prior intention of deceiving her significant other before betraying his trust. If the cheater instead gives into selfish temptation, this seems sufficient to call what the cheater does a betrayal of trust, and there needn't have been any "calculated manoeuvring of a traitor setting up his victim". Of course, the result of the betrayal may be that, for the betrayed partner, all the previous loving acts of the cheater are now tainted and coloured with a spirit of deception. But the cheater needn't have actually been deceiving, for us to call her action a betrayal.

Furthermore, abandonments and betrayals are not mutually exclusive. An abandonment could also be called a betrayal of trust in at least some contexts, on the presumption that prior trust involves a recognised expectation of not being abandoned or left unfulfilled. It is plausible to imagine those with strong fears of abandonment viewing perceived abandonments as betrayals of trust. An act can be both an abandonment, in being a case where someone ups and leaves another without warning, and also a betrayal, if there was a presumed mutual recognition between trustor and trustee, that upping-and-leaving was a matter of concern for the trustor. For example, think of the parent who 'goes out for cigarettes' and never comes back, leaving the child with a lingering sense of both abandonment and betrayal.¹

¹ Thanks to Justina Berskyte for this example.

It's for these two reasons that Shklar's definition of 'simple betrayal' is more appealing than Jackson's, which requires too much presumed scheming on the part of the betrayer. Shklar's is as follows:

The one irreducible experience surely is having an expectation disappointed. For a simple act of betrayal, one person should have both intentionally convinced another person of his future loyalty and then deliberately rejected him. The latter then, not surprisingly, feels betrayed. (Shklar 1984 : pp.140–141)²

This analysis would explain why we're tempted to call what a cheating partner does an act of betrayal, even if her motives in doing so are just selfish and careless. It just requires that the cheater have previously intentionally convinced her partner of future fidelity, and then deliberately carried out the opposing action anyway, due to whatever motives. It also explains why an abandonment could also be called a betrayal, if the issue of abandonment was raised and recognised as of great import to the trustor.

I thus suggest we endorse Shklar's definition of 'simple betrayal', as it seems to capture the essence of disappointed trust. At this stage, though, I primarily draw attention to these accounts of betrayal as a way of helping flesh out our picture of the interpersonal import of matters of trust, given that the trust-betrayal connection is standardly appealed to without much elaboration. It's clear that once we start to unpack the nature of betrayal itself, we reveal that violations of trust are complex and worthy of more interrogation.

The focus on betrayal in the trust literature also shouldn't make us lose sight of the various ways that matters of trust can be disappointed in an interpersonally significant way. Shklar for instance also alludes to "breaking an appointment that means much to the other person" as a distinct phenomenon, such that we can presume matters of minor import wouldn't necessarily warrant being called betrayals. We can conclude that one dimension on which interpersonal import varies is just how important the matter of trust is, to the trustor.

Another dimension of import is *who* the trustee is, to the trustor. Shklar's and Jackson's descriptions of betrayal place the phenomenon within the stance of close relationships,

² I take it that the sense of 'expectation' that is disappointed is *normative* rather than predictive. Simply predicting that another will φ doesn't suffice for it being a matter of interpersonal import, and thus betrayal.

alliances and so on, those that make salient notions like loyalty and fidelity. This makes sense, intuitively. Shklar's remark that in a simple betrayal, the betrayer must have convinced the betrayed of her "future loyalty" is not hyperbolic. The matter itself – exactly what is trusted – will also tend to vary with the nature of the relationship.

Matters of trust, though, can still occur between non-intimates, and such trust can still be violated, breached, or disappointed. Baier perhaps invokes a helpful notion here. At the same time as offering the trust/reliance distinction on the basis that trust can be betrayed, she says our trust might be simply "let down" (1986 : p.235). Those occupying professional roles can disappoint our trust by e.g. providing incorrect or misleading information, but it might be more appropriate to say we are let down by the individuals in question, in these cases. A "let down" might also be a more appropriate description of some sorts of disappointed trust between intimates. For instance, if a friend fails to show up to your regular lunch appointment without explaining why, you might say you were let down rather than betrayed.

The question arises as to whether these harms in disappointed trust are all technically betrayals of trust, or different categories of wrongdoing altogether. For now, I do not offer a systematic ruling on what is required for a disappointment to count as a betrayal, or a let-down, or an abandonment. I simply note the distinctions between these violations of trust, to point out that taking the trust-betrayal connection seriously, doesn't require endorsing the view that all forms of disappointment *are* betrayals. I suggest that we use the trust-betrayal connection to characterise how there is a family of interpersonally-focused harms associated with disappointed trust, of which betrayal is the flag-bearer. In chapter 8 I will further discuss the possible differences between these harms.

The trust-betrayal connection characterises the interpersonal import of the trust stance in terms of what kinds of harms it leaves us open to. But, we can also characterise the interpersonal import in terms of *how* the trust stance leaves us open to such harms. In this respect, appealing to the notion of *vulnerability* is a common way of construing the interpersonal import of matters of trust. This is also a notion that seems to stem from how Baier describes the interpersonal quality of trust, when she describes trust as an attitude of *accepting vulnerability* toward another (1986 : p.235). Vulnerability has since become a common shorthand for characterising that interpersonal import of matters of

trust for a number of authors, e.g. (Govier 1992b; Hill and O'Hara 2006; Jones 1996, 2004; Lahno 2001; McLeod 2015; Pettit 1995). Lahno for instance, rather than focusing on the possibility of betrayal, focuses on vulnerability to harm that results from engaging with others in matters of trust. Focusing on the notion of *vulnerability* rather than the trust-betrayal connection, then, may be useful as a way of encapsulating the various interpersonally-focused harms that we open ourselves up to in matters of trust.

Again, as I noted when introducing the notion of betrayal, we should recognise that vulnerability is only a characterisation of how matters of trust involve interpersonal import. I suggest that unlike the trust-betrayal connection, though, vulnerability as a concept loses explanatory momentum, once we try to lever it as a way of differentiating between cases of anticipating beneficial/detrimental acts, and matters of trust. That's because there is a sense in which, by simply planning on the basis that others may perform instrumentally beneficial or detrimental acts, we are vulnerable. John the food stealing flatmate is vulnerable to not having food if Elis doesn't buy any. If X is said to be vulnerable to Y, this can plausibly just mean that X is at risk from lessened wellbeing from Y's action. This could apply to John's anticipating/planning on Elis' continuing to buy food that he can steal. Hawley for instance suggests that, "the notions of accepted vulnerability plus foregoing the attempt to reduce such vulnerability capture roughly the notion of reliance" (2014 : p.8). As such, even *mere* non-betrayable reliance could involve vulnerability. However, it also seems right to say that the kind of vulnerability we take to be at issue in matters of trust is distinct from the way that John is 'vulnerable' to not having food, if Elis doesn't continue to buy any. John might be vulnerable to lessened wellbeing if Elis doesn't buy food, but he isn't vulnerable to interpersonal harm by Elis, in the form of a betrayal of trust.

These brief reflections suggest that, if we are taking *vulnerability* as the concept that fundamentally explains the interpersonal import of matters of trust, we will need to do more to specify what sort of vulnerability is at issue. Vulnerability as a concept is broad enough to apply even to those who exploit and use against others. By exploiting another, we can be vulnerable to being attacked or found out, and losing everything. If we're committed to vulnerability being the concept at issue in bestowing matters of trust with an interpersonal import, then, we need to specify what sort of vulnerability it is. In effect,

we still need an account of the interpersonal import of matters of trust, beyond noting that we are vulnerable in matters of trust.

That said, I will later appeal to the concept of vulnerability. I will also be invoking it to help explain the trust-betrayal connection. I will argue that it is right to point out that in matters of trust we are vulnerable, because this is an important structural feature of matters of trust. I will use ‘vulnerability’ to describing a feature of our *situation*. In a matter of trust, we must be ‘open’. There must be a weakness or gap in our armour, such that we can be gotten at, or ‘struck’, by another. This will be because things we care about can be affected by another’s action.

A weakness resulting from our situation does, I take it, roughly capture the concept of vulnerability. It equates to an ‘opening in our defences’, through which we or things we care about can be taken, or damaged. We might not even notice such weaknesses or openings. A military will look for vulnerabilities in the opposing force (that hopefully have gone unnoticed by that opposing force) in order to gain the advantage on the battlefield.³ We can imagine a unit tasked with carrying out a targeted killing using phrases to the effect of “that’s where he’s vulnerable”, meaning an opening or weakness.

It isn’t enough to have such a *situational* vulnerability to another, for her actions toward us in respect of that vulnerability, to thereby become a matter of trust. Thus, it isn’t enough to have a situational vulnerability toward another, for us to be able to adopt trust stance attitudes toward her in respect of that. The trust stance requires interpersonal import. But, I will also argue that the interpersonal import of matters of trust results from there being a further kind of weakness we are open to in matters of trust. As such, the concept of a situational vulnerability can also shed light on what the interpersonal import of a matter of trust is. It results from an additional feature of the situation of those who have a trust stance attitude.

I will offer a brief example to illustrate what the additional feature is. This amounts to a quick sketch of my broader account of the trust stance, as I point out the additional kind of vulnerability that is at issue in matters of trust. Suppose that two people - Xavier and Yolanda - don’t have robust passwords on their bank accounts. As such, both could lose

³ This sort of usage of ‘vulnerable’ is found in (Simpson 2011 : p.331).

all of their money to hackers. This is enough to say that Xavier and Yolanda are vulnerable, in some sense. Suppose also that both are unaware of their vulnerability to having their money stolen. Xavier and Yolanda aren't aware that a lack of a robust password is a security weakness. As such, neither are paying attention to how another person could, if she wanted to, take all their money.

Xavier, though, has a high level of basic trust toward people. It wouldn't occur to him that someone would be capable or motivated to take his money from his account. Yolanda on the other hand, doesn't trust anyone at all. Moreover, she long ago stopped caring about others' opinions of her. Her money is from ill-gotten means. It seems reasonable to suppose that Xavier and Yolanda share the same basic situational vulnerability: their lax password protection. But, it also seems reasonable to say that Xavier is vulnerable to others in an additional way. Because of his trust, he can be hurt, in a way that Yolanda long ago stopped being able to. Where Xavier can be struck by being robbed, in a way that cuts to his core, Yolanda can only be struck by having something that she wants, be taken from her, or prevented access to. As such, Xavier's situational weakness is a *window* on a deeper vulnerability he possesses. A further sort of concern, or thing that matters to him, is left exposed, by his lax password. That someone takes his money means something additional to him. Behind Yolanda's situational weakness, though, simply lies her money. Her money is meaningful for her as means to more things that she wants. So, by taking her money, you can disrupt what she wants to do, and leave her angry and frustrated. But should you do this, it won't mean the same thing to her as it would to Xavier. Xavier may feel angry, but he will also feel betrayed and disillusioned by the action, including at the fact that no one warned him about the need for a secure password.

In chapter 8 I will put forward an account of what is behind Xavier's window of vulnerability and so what the interpersonal import of matters of trust stems from. I suggest that Xavier has a concern for his need to form attachments to others, and this is exposed by his lax password, and what is struck at when someone robs him. Attachment concern is a form of emotional dependence on other people: a non-instrumental need to form social bonds with individuals, and to thereby feel a sense of belonging to a group. This represents an additional vulnerability for Xavier, because it makes him liable to be betrayed and hurt, should the community he is attached to, strike at him. Yolanda on the

other hand treats nothing as a matter of trust. She trusts no one, and has no attachment to the group. When she is struck by having her money taken, she admonishes herself, and possibly directs her fury out at the people she already feels no need for, beyond potential instrumental need.

Windows of vulnerability exist because of our being physically unable to prevent actions that could hurt us. There are endless ways in which people can strike at us or things we care about. We needn't be aware of these gaps in our defences, or pay attention to them. When we have high levels of basic trust, the possibility that another might exploit certain gaps, simply isn't salient.

Relying on another to perform an action is one way of creating a window of vulnerability to that attachment concern. In section 4.7.2 and chapter 5 I argue for an account of reliance that goes toward explaining how relying on another to perform some action φ , leaves the success of something we are trying to achieve, in the hands of another's acting. From that reliance, some end we are trying to bring about, can be obstructed by another failing to φ ; we create an opening for another to hinder our progress. Depending on whether we have a deeper attachment concern like Xavier, we can also be struck in the way characteristic of a matter of trust. It can mean something to us that another fails to φ when we rely on her to φ , beyond a frustration that something that we wanted to happen has now fallen through. What would happen is that she – the one we were relying on – hasn't come through for us, in a time when we needed her to.

1.8 Challenging interpersonal import as necessary for matters of trust

The discussion in the preceding section, concerning the connection between matters of trust, betrayal and vulnerability, is intended to flesh out the claim that matters of trust are partly defined by carrying some sort of interpersonal importance. The position that interpersonal import is a necessary feature of a matter of trust, though, runs contrary to a range of accounts of trust in the literature. To use three examples:

Trust is a fundamentally cognitive notion. To trust or distrust others is to have some presumption of knowledge about them. [...] A common claim about trust is that it is inherently normative. Unless it is merely an abstruse definitional move, this claim is not well grounded and even appears to be false. (Hardin 1998 : p.11)

I am using the word ‘trust’ in the sense of correct expectations about the *actions* of other people that have a bearing on one’s own choice of action when that action must be chosen before one can *monitor* the actions of those others. (Dasgupta 1988 : p.51)

[...] trust – a particular expectation we have with regard to the likely behaviour of others – is of fundamental importance. [...] trust (or, symmetrically, distrust) is a particular level of the subjective probability with which an agent assesses that another agent or group of agents will perform a particular action, both *before* he can monitor such action (or independently of his capacity to ever be able to monitor it) *and* in a context in which it affects his *own* action [...] When we say we trust someone or that someone is trustworthy, we implicitly mean that the probability that he will perform an action that is beneficial or at least not detrimental to us is high enough for us to consider engaging in some form of cooperation with him. (Gambetta 1988 : p.217)

Hardin’s claim above, that trust is not inherently normative, amounts to the claim that the interpersonal quality of trust and distrust, is not a necessary ingredient in the trust stance attitudes.⁴ This is a position, also reflected in Dasgupta’s and Gambetta’s remarks, typically found in the trust literature that is, broadly speaking, informed by a rational choice theory, or game-theoretic perspective.

It is perhaps telling that the typical example case of trust that these perspectives consider, is not the sort of case, found in the *flatmate* examples from section 1.6, of people who live alongside one another in close proximity. Instead, the game-theoretic views typically involve interactions between unknown individuals, groups, or institutions. In these interactions, we can be tempted to suppose that interpersonal import is missing. But, because it seems sometimes plausible to use ‘trust’ to characterise interactions between unknowns of this sort, it follows that the overtly interpersonal or normative features of the cases between intimates, are not necessary features of trust. Those features we could suppose are a side-effect of the underlying machinery of trust, because there are cases of trust – those between unknowns – where those features are absent.

Against this thought, I take it that we can argue it is much less clear whether we should describe interactions between unknowns in the typical Prisoner’s Dilemma kinds of case, as trust-based. As such they serve as poor paradigm cases of trust. Compare this with the flatmate contrast cases. In *Food Stealing Flatmate* (John predicting Elis continues to buy

⁴ My phrase ‘interpersonally important’ is intended to encapsulate this normativity; it opens the possibility of harms and wrongs associated with disappointed trust i.e. betrayal.

food he can take) and *Cash Stealing Flatmate* (Izzy predicting Maria will cancel the card she is stealing from), these are cases of ‘mere prediction’ or ‘mere reliance’ concerning whether someone is likely to perform an instrumentally beneficial/detrimental act. In those cases it is just not appropriate to say that John trusts Elis to continue providing food, or that Izzy distrusts Maria regarding cancelling her card. On the other hand, *Ailing Flatmate* (Maria trusting Izzy with her debit card) and *Wary Flatmate* (Elis suspecting John is lying) are clear cases of trust and distrust. In those, the interpersonal quality is undoubtedly present. If the phenomenon we are attempting to understand is interpersonal trust and distrust, and it appears that our willingness to apply the concept of trust or distrust depends on there being interpersonal import/normative machinery, then we must suppose that matters of interpersonal trust inherently involve some sort of interpersonal import. It follows that even in the ‘distanced’ cases that the rational choice/game-theoretic perspective is concerned with, we have reason to suppose that *if* we are inclined to think they are examples of trust, then those examples *must* somewhere contain an interpersonal quality.

It may be that a candidate case is not in fact describing a matter of trust, but instead a matter of mere prediction or risk. But then we can’t presume that reflecting on such a case of mere prediction or risk reveals anything about the nature of interpersonal trust or distrust. It is instead describing a separate phenomenon. We have more warrant to deduce that matters of trust involve an interpersonal quality, because in the flatmate cases where a matter of trust is present and where it is absent, the interpersonal quality coincides with the trust or distrust attribution. Yet the game-theoretic ‘pure prediction’ view of trust, which is reflected in Hardin’s, Dasgupta’s and Gambetta’s positions above, gives both necessary and sufficient conditions for John’s attitude toward Elis. But we are inclined to think that case is not a case of interpersonal trust at all. As such, we have reason to believe the game-theoretic view misses something essential for making a matter one of trust.

It is also telling that the game-theoretic view of trust as pure prediction typically appeals to, or makes overtures toward, interpersonally loaded concepts like cooperation and betrayal. In a Prisoner’s Dilemma game, for example, there are two agents presented with two courses of action. The game is set up as follows:

Figure 1: The Prisoner's Dilemma.⁵

		PLAYER 2	
		COOPERATES	DEFECTS
PLAYER 1	COOPERATES	$R=3, R=3$ Reward for mutual cooperation	$S=0, T=5$ Sucker's payoff, and temptation to defect
	DEFECTS	$T=5, S=0$ Temptation to defect, and sucker's payoff	$P=1, P=1$ Punishment for mutual defection

The Prisoner's Dilemma is intended to depict the basic structure of matters of human cooperation. Players are presented with two courses of action, for how to go about procuring points. These points stand in for some outcome that each player desires, where the higher the number of points, the more desirable the outcome. How many points each player gets, depends on what course of action the other player pursues. If both players choose 'cooperate' then both receive some reward. If one defects and the other cooperates, the defector takes everything (the other receiving the 'sucker's payoff', i.e. nothing). If both defect, then both receive the 'punishment' payoff. The act of pursuing the lesser payoff is typically labelled as 'cooperate', whereas pursuing the larger payoff (the selfish option) is labelled 'defect' (synonymous with 'betray') (Axelrod 2006 : pp.7–8). As such, the most instrumentally desirable outcome is had, when we 'defect', and the other 'cooperates'.

My point against the game-theoretic view concerns what warrants the application of concepts such as 'cooperate' and 'defect', to the different courses of action in the Prisoner's Dilemma. By applying these concepts, found in the realm of trust, distrust, and betrayal, this indicates that the game aims to shed light on matters of trust. This is even

⁵ This table based on (Axelrod 2006 : p.8).

though the game is supposed to be describing a scenario where the machinery of pure prediction/risk is at play.

But, we can question whether the Prisoner's Dilemma *does* shed light on matters of trust. We can ask, for instance, what warrants using the label 'cooperates' for an act of choosing the lesser payoff? Imagine that Player 1, say, chooses the lesser payoff, out of motives that would align with John's or Izzy's. After all, a free-loader who knows to exploit others to just enough of a degree that the target doesn't move elsewhere, is more successful than one that exploits the other, to the point that they leave. Player 1 might know this, and it forms the basis of his playing style. It could be inconsequential to Player 1 that Player 2 may receive some payoff too, when 1 chooses the lesser payoff. Player 1 may just know that if the larger payoff is pursued, this closes off the possibility of future payoffs with this target. Yet, we don't characterise what a freeloader or parasite is doing as 'cooperating' with their target, in the same way we wouldn't describe John's choosing to only take enough of Elis' food to not give the game away, as cooperating with Elis. Even if the action of a freeloader sometimes bears a *structural* resemblance to a genuine instance of cooperation, we aren't warranted in using an interpersonally/normatively loaded concept like *cooperation*, to describe what is happening when one player chooses the lesser payoff. That is, without knowing more about the agent's motives in that scenario.

As a result, theorists sympathetic to the game-theoretic model of interpersonal interaction cannot appeal to the model itself as revealing that matters of trust aren't inherently normative (or: that they can be purely calculative). By presuming that the concepts of cooperation and defection can be applied to acts of calculated manoeuvring, without further argument, we have already re-described matters of trust in a way that ignores our intuitions about cases like the *Flatmate* contrast cases. This makes it the case that, whatever scenario the game-theoretic model is describing, this scenario bears only a structural resemblance to matters of trust, as they appear in real life. That's because, by supposing that we are permitted to call an act of choosing the lesser payoff (for whatever reason) an act of cooperation, we have already sanctioned an interpretation of an action with interpersonal or normative import like 'cooperating', as consisting in a calculated manoeuvre in the pursuit of self-interest. Deducing then that the interpersonal import in matters of trust contains nothing inherently normative, is short-sighted. It fails to recognise that our starting point was to say that actions that appear to have, on the

surface, interpersonal or normative salience, are in fact actions that have no interpersonal or normative salience.

It is also the case that authors in the game-theoretic tradition typically make overtures to the interpersonal import of matters of trust, in their broader discussions of trust. Dasgupta for example appeals to the interpersonally vivid concept of betrayal (1988 : p.53) and emphasises the role of external pressures or “enforcement agencies”, which create incentives among individuals to fulfil trust. One such enforcement agency, Dasgupta notes, could be “society at large. Social ostracism, and the sense of shame that society can invoke, are examples of such punishment” (1988 : p.50).

The role of enforcement agencies is a common concern in the rational choice/game-theoretic, post-Hobbesian tradition. It stands to reason that it should be, since we have already seen that trustors, on this tradition, fit the psychological profile of John or Izzy. There is arguably little possibility, on this view, of such a thing as the two-place trusting relationship, or of one-place basic trust toward a community (as discussed above in sections 1.3 and 1.4). For such agents, there is no consequence to defection beyond (possibly) no longer having the other player as a source of procurement of resources. It therefore makes sense (it is rational) for these agents, given an abundance of other players, to continually ‘cash out’, and defect. But then, the ‘tragedy of the commons’ results: all will continually receive the ‘punishment’ outcome. It would therefore make sense for all players in this society to create an external agency to bootstrap their own and other’s incentives into favouring cooperation. In effect, these individuals cannot trust or rely on themselves to do the thing that makes sense for them to do: to co-operate. So, in a Hobbesian move, they pool together to create an institution or set of practices, that will incentivise themselves, and everyone else, toward this outcome.

Appealing to the concept of social ostracism or shame as performing the function of an external enforcement agency, though, involves an appeal to a new set of motivations in the players. We aren’t creating a social institution or practice (a justice system, police force, etc.), the presence of which would threaten consequences that tip the balance in favour of a choice – co-operate – that would *otherwise* not be chosen. It rather amounts to a rewiring of the motivational make-up of the individual players; it is an overture to the players having a non-instrumental *need* to belong to a social group, and resulting social

emotions associated with feeling one may be ostracised. The original players didn't have this, though. Hence, the problem of needing an enforcement agency. The original players only had an instrumental need to have another agent to play against, in order to receive any points. This set-up led to the very problem we are now positing the enforcement agency of society itself/social ostracism/shame, to resolve. Players like John or Izzy were not moved by social ostracism or a sense of shame; they were moved by taking what they can get from others. If we supposed that social incentives were already salient for John and Izzy, we would not have faced the very problem – everyone defecting – that led to us *positing* such motivations, as a solution. In effect, we are not talking about the same players that we were to begin with. Players that are moved by shame, don't have need of an external enforcement agency.

Hardin's account also gestures toward something interpersonal. His account is that trust is not just a prediction about whether another will perform some instrumentally beneficial action. He suggests that when I trust you to φ , I trust you to φ from "a reason in some way grounded in me [...] I do not trust you to [φ] if I have no reason to suppose you do so somehow on my behalf. I trust you with respect to some action if your reason for doing it is to take me into account in some relevant way. Typically, your reason will be that it is in your interest to maintain our relationship. Hence, my trust in you is typically encapsulated in your interest in fulfilling my trust." (Hardin 1998 : p.12). As such, his account of trust is that it involves a supposition of shared or 'encapsulated' interest: it is in your interest to fulfil my trust, because it is in your interest to maintain our relationship.

While this inches Hardin's account into the domain of interpersonal import, it isn't definitively placed within it. Depending on which of his phrases we focus on, Hardin's conception of trust could amount to an attitude of cynical knowingness; a positive prediction that another will act, with full awareness that she does so in order to continue getting what she wants from us ("it is in your interest to maintain our relationship"). This arguably falls short of capturing the interpersonal import of an attitude like that Maria has toward Izzy. On the other hand, Hardin's conception could characterise the attitude Maria has toward Izzy ("you do so somehow on my behalf"; "to take me into account in some relevant way"), but this is more due to the vagueness of the descriptions allowing our intuitions, that trust *does* involve interpersonal import, to fill in the blanks of what is meant.

It is also telling that Hardin makes an overture to the notion of *trustworthiness* – but not trust – being inherently normative (Hardin 1998 : p.11, 1996). This is puzzling, though. By making a concession to trustworthiness as inherently normative, while still holding that trust isn't, this creates an implausible disconnect between the attitude of trust, and what the attitude is 'about'. The question is raised: at what point does another's trustworthiness matter in that normatively/interpersonally significant way, if our trust of her doesn't? Presumably, if trustworthiness is normatively significant, we would blame others and ourselves for breaching trust.⁶ If we were to suppose that *trust* wasn't normatively significant, but trustworthiness was, then perhaps the trustor would have no claim against the trustee who breaches her trust, but bystanders *would*. They would sanction the trustee, because it is important to be trustworthy. This is clearly absurd: if broken trust matters to anyone, it is the trustor. It stands to reason that, when we suffer a betrayed trust, this is a matter of normative significance to us. It therefore seems that, once we hold trustworthiness to be 'inherently normative', it would follow that *matters of trust*, and the trust stance attitudes, are also inherently normative.

In chapter 8 I will further discuss Hardin's account, as part of defending my attachment-based account of the interpersonal import in matters of trust (my account suggests that Dasgupta's overture to social ostracism and shame, is well-motivated).

1.8.1 'Trust that'

Here I raise a point about different usages of the term 'trust', that may shed light on confusions over whether a particular case is one of interpersonal trust, or mere prediction. The concept 'trust' has a broader application beyond the domain of interpersonal trust. This because we can also have trust stance attitudes toward our own judgements. As such, we could reach a judgement about the likelihood of another performing some action, and

⁶ On this front, we can compare trustworthiness with 'creditworthiness', as assessed by credit ratings agencies. Creditworthiness is not interpersonally normatively significant, as entering into a loan contract isn't either. The 'worthiness' in question isn't *deservingness* of credit, because it is just the conclusion of a risk assessment. This brings out that trust and trustworthiness involve something more than this. Thanks to my supervisor Thomas Smith for raising this comparison.

describe ourselves as trusting this. However, this is not *interpersonal* trust, where the target of our trust stance attitudes is an individual or group of persons.

Along these lines, a distinction between “trust that” and “trust in” is drawn by (Hill and O’Hara 2006 : pp.1725–1727). Lagerspetz also notes the difference, in drawing a distinction between interpersonal ‘normative’ trust and non-normative trust:

While some uses of the word ‘trust’ are normative in the way just described, there are others that surely are not. Some of these instances of ‘trusting’ are simply predictive beliefs with no clear moral implications, as when I trust that the weather will stay warm or that the Euro will not fall in relation to Pound Sterling. The crucial question is how plausible it would be to construe nonfulfilment as betrayal. However, I am not suggesting that these other cases are not ‘really’ instances of trust after all. That I leave to the reader to consider. Differentiating between cases is philosophically important, but legislating about language use is not. (2015 : p.20).

There are interesting questions to ask about the overlap between the concept of interpersonal trust, and trust in a judgement concerning an event/state of affairs. My purpose in pointing out the difference is just to note a possible source of the confusion in the rational choice/game-theoretic literature over whether interpersonal trust carries normative significance (interpersonal import).

I will use ‘trust that’ to characterise the non-normative trust directed toward states of affairs. When we trust that p , this could be directed toward a judgement that another person Y will φ /would φ /has φ ’d. But, when X trusts *that* Y will φ , this isn’t yet a case where X trusts Y to φ . It reflects X ’s trust of a judgement. The matter of trust then isn’t yet obviously interpersonal, because the concern is with her own judgements and plans. It’s possible that in the case of the Food Stealing Flatmate, John trusts *that* Elis will put food in the fridge. This doesn’t entail that John thereby trusts Elis *to* put food in the fridge. The usage of ‘trust’ in ‘trust that’ picks out a slightly different attitude to interpersonal trust.

No doubt ‘trust that’ can also be present in a case of interpersonal trust, insofar as whether or not X trusts Y to φ will be a factor in whether X trusts that Y will φ . The key point is that even if it is appropriate to describe a case as one of ‘trust that’, it doesn’t follow that it must also be a case of interpersonal trust.

But, I don't claim that an instance of 'trust that' in everyday life is *always* only used to report an attitude of this sort of propositional trust. It seems right to say that, for example, one might say to another "I trust that you'll come through for me on this", where this would report an instance of interpersonal trust. I only mention the 'trust you'/'trust that' contrast as a way of drawing attention to a possible source of confusion, over whether interpersonal trust is not essentially different from a prediction about whether certain states of affairs will obtain. As such we can sometimes use this distinction as a guide to determining why our judgement may be clouded about a particular case. Furthermore, much of what Hardin et al. say about interpersonal trust is not true of interpersonal trust, but is true of 'trust that'.

2 The Reliance Plus view of trust and distrust

In this chapter I will set out the dominant view of trust in the philosophical literature. This is the view I call the ‘Reliance Plus’ (or ‘REL+’) view. It claims that to trust another is to rely on her to act, where that reliance is invested with interpersonal or normative significance.

In section 2.1 I will set out the generic form of a REL+ view. REL+ views account for what I referred to in chapter 1 as the anticipatory and interpersonal elements of the trust stance, by pinning the anticipatory element to reliance on another, and the interpersonal element to the ability to be betrayed, in case of disappointed reliance. In section 2.2 I offer a critical overview of three prominent REL+ accounts: Baier’s (1986) ‘Goodwill’ account, Holton’s (1994) ‘Participant Stance’ account, and Hawley’s (2014) ‘Commitment’ account. In the last section of this chapter – section 2.3 – I illustrate how committing to the generic REL+ view also commits us to a broader account of the trust stance, i.e. it also commits us to a view of distrust. This is a view of trust and distrust that Hawley (2014) and Helm (2014) commit to.

In chapter 3 I set out an argument against this REL+ account of the trust stance. This argument illustrates that REL+ is false. In chapter 4 I set out a further argument against the REL+ view of the trust stance, that brings out the underlying reason why REL+ is false.

2.1 Reliance plus betrayability

In this section I set out the generic form of a REL+ view. REL+ is the view that trust is an attitude consisting of “reliance plus some extra factor” (Hawley 2014 : p.5). The extra factor is, roughly, an ability to be betrayed by the one you rely on. I will frequently use the term ‘betrayability’ to stand in for this extra factor.

REL+ represents a species of views of trust, that are pervasive in the literature (Baier 1986 : pp.234–235; Hawley 2014 : pp.2–3; Helm 2014 : pp.199–205; Hieronymi 2008 : p.215; Holton 1994 : p.66; Jones 1996 : p.14; McLeod 2015; Walker 2006 : pp.80–83). The essential characteristic of a REL+ view is to take trust to be a species of the more generic attitude *reliance*. REL+ says that trust is a species of reliance, albeit an

interpersonally/normatively loaded variety. I refer to the view as ‘Reliance plus’, then, because it tells us that trust is reliance with an extra interpersonal quality, that sets trust-reliance apart from other types of reliance. REL+ accounts for that interpersonal quality by drawing on the intuition that trust is connected to the harm or wronging of betrayal. Trust is interpersonally salient because when we trust another, we are susceptible to being betrayed by her. As such the thought is that trust is reliance, plus conditions that enable the possibility of betrayal.

As we saw in section 1.7, the connection between trust and betrayal is immediately apparent, when we reflect on everyday cases of trust. Trusting your partner to remain faithful, for instance, entails the possibility of being betrayed should she stray. Trusting your friend to keep an embarrassing secret entails the possibility of his betraying your trust in telling your secret to others. The ability to suffer betrayal as a kind of action or event, or a feeling that tends to result from such actions, is a defining characteristic of trust.

REL+ can be traced to Baier (1986 : p.234). Baier points out a connection between trust and reliance (or: the ‘trust-reliance’ connection). The connection, when unpacked amounts to this claim: trust entails reliance, but reliance doesn’t entail trust. In other words, reliance necessary but is not sufficient for trust. There is, as a result, a contrast between trust and ‘mere reliance’. Baier uses a comparison case to motivate this claim. There is an infamous piece of philosophical folklore that Kant was so predictable in his routines that his neighbours could rely on his walking for keeping time (1986 : p.235). The fact that Kant’s neighbours could rely on his walking doesn’t entail that they trusted Kant to walk, though. The trust-betrayal connection explains what is missing: the fact that Kant’s neighbours aren’t betrayed by Kant just in case he doesn’t go for his walk.

If we imagine now that the neighbours *trust* Kant to walk every day, it seems that they still rely on his walking. But, their reliance now seems normatively/interpersonally loaded, in a way that their *mere* reliance on him isn’t. It seems right to say that Kant can betray his neighbours when they trust him to walk, but not when they merely rely on him walking. It therefore seems that, by reflecting on this case, we can draw out a general point about the difference between merely relying on another to φ and trusting another to φ . The difference between mere reliance and trust, is whether one can be betrayed by the one

you rely on, in case she doesn't φ . When merely relying on her, one isn't betrayable by her. But when trusting her, one is betrayable by her.

Not every instance of disappointed trust need amount to a betrayal. As mentioned in section 1.7, we might also simply be 'let down' on some occasions. What is important is that there is a difference between mere reliance and trust, in that trust is interpersonally or normatively significant in some way. The trust-betrayal connection provides a useful way of characterising what that significance is. Trust, rather than mere reliance, is an attitude that opens up the possibility of interpersonally-focused wrongings or disappointments. The stereotypical or most vivid such disappointment we can suffer in trust is betrayal. As such focusing on the notion of betrayal is useful to help bring into focus the somewhat vague (though correct) remark that trust has interpersonal import. By saying that trust has 'betrayability', then, this doesn't mean that any instance of trust, must be a case where X relies on Y φ 'ing, such that Y's not φ 'ing counts as a betrayal. It could be that Y lets X down, or abandons X. It's plausible that the sort of harm that results from Y's failure to φ , will depend on what the matter of trust is, how significant that matter is for X, and in what way Y fails to φ . Then, the significance of the matter for X may well vary with who Y is to X. What betrayability denotes is that X's attitude to Y carries the interpersonal import of trust, such that X is susceptible to the harms characteristic of disappointed trust, such as betrayal.

As stated, variants of REL+ are often explicitly and tacitly endorsed in the literature on trust. This should come as no surprise since, as we've seen, the view follows from two intuitive claims about trust: that trust is connected to betrayal, and that reliance is necessary but not sufficient for trust. These are intuitions concerning what I referred to in chapter 1 as the anticipatory element and the interpersonal element of trust. On the one hand, the trust-betrayal connection accounts for the interpersonal element; it explains the interpersonal import of matters of trust. On the other hand, the trust-reliance connection accounts for the anticipatory element, because reliance on another seems to involve anticipating her fulfilling that reliance.

REL+ is also useful in that it provides a *method* for analysing trust. Candidate accounts of trust can be assessed by how well they provide convincing conditions for reliance and betrayability. And, the reliance condition is very easy to satisfy: it can just be stipulated

that it holds in an example case. Critically evaluating a candidate REL+ account thus tends to involve raising doubts about how well the proposed account of betrayability holds up to counter-examples. So, how we ought to understand trust comes to rest on putting forward the correct account of betrayability. That the view is so widespread is no doubt partly due to the appeal of the procedural method for analysing trust that it grants.

2.2 REL+ and narrowing down betrayability

Following the REL+ paradigm has allowed trust theorists to rule out certain candidate accounts of betrayability. As such, REL+ is not only well supported by intuition, but has arguably been philosophically productive. Following the REL+ paradigm has allowed theorists to make important claims about what the interpersonal import of trust consists in. We can see the progress made on this score, in a thread of argument running through Baier (1986), Holton (1994) and Hawley (2014). Each puts forward a REL+ account offering a different formulation of the betrayability component. In the following sections I will offer a critical overview of this thread of arguments. The conclusion from this, I will argue, is that Hawley's account of betrayability manages to crystallise the underlying structure of an issue of betrayal, and thus a matter of trust.

2.2.1 Baier: the 'Goodwill' account

Baier (1986) contains several different formulations of an account of trust. The first of these is largely responsible for the REL+ paradigm in the literature. It's that account I'll discuss in this section. Baier's first formulation of a view of trust is as reliance on another's goodwill toward one (1986 : p.234). Let's call this the 'Goodwill' account.

When we use the Goodwill account to explain the case of Kant's neighbours, the account is convincing. Kant's neighbours merely relied on Kant's walking in going about their day, and this seems to lack any of the normative or interpersonal significance that trusting Kant would involve. It is possible, of course, that they merely rely on his walking as they would a clock. If we instead suppose that they rely on his goodwill toward them as the motive of his walking, then I take it that we are more inclined to agree that the neighbours trust Kant to walk. That we are so inclined suggests that, as an attitude with some sort of representational content, trusting someone involves representing the one you trust as

having some level of (perhaps minimal) goodwill toward you. So, what is missing from mere reliance, but present in trust as a type of reliance, is presumed goodwill toward you. Trust, therefore, seems to be reliance, plus (presumed) goodwill.

The Goodwill account, though, is insufficient for capturing trust. This is because it doesn't capture betrayability. Holton shows this by a counterexample: a con-artist (or 'trickster') relies on your goodwill toward him, yet doesn't trust you (1994 : p.65). It is possible to rely on another's goodwill toward you, as the motive for her action, while precisely not trusting her. A con-artist, after all, feigns trusting, just to solicit the sort of action from others that will be beneficial to his own ends. The type of reliance that a con-artist has on his mark is, surely, exploitative. This is opposed to the kind of interpersonally significant reliance found in trust, which, as intuition suggests, involves the possibility of being betrayed. So, even in stipulating that in trust we rely on another's goodwill, we don't guarantee that one can be betrayed in so relying. We can rely on another's goodwill while not trusting her. The con-artist objection is a decisive one in showing that the 'reliance on another's goodwill' account is insufficient for trust.⁷

Holton also argues that presuming another has goodwill toward us is not a necessary component of trust. He gives three examples to show this (1994 : p.65). One involves an estranged couple, who can trust one another to look after their children, without having goodwill toward one another. Holton then anticipates a response to this case; even though the couple have no goodwill toward one another, they have goodwill to the child. As such, the response is that, even if the couple have no goodwill toward one another, goodwill is still *relevant* to trust. It's just that, in some cases, we presume the person we trust has goodwill toward the thing we trust her with, rather than us.

Holton raises a further case, intended to show that goodwill is *irrelevant* to trust. He suggests that when we are forced to surrender to an enemy while under gunfire, by waving a white flag, we can trust the enemy to not shoot us. But, Holton claims, we would also suppose that the enemy has no goodwill toward us, as part of our trust attitude. This

⁷ I will argue in chapter 8 that the underlying reason why the Goodwill account fails is because, when we rely on another's goodwill toward us, whether she has goodwill is of *instrumental* concern to us; it is playing a practical role in achieving our goal. A matter of trust though must involve a *non-instrumental* concern for another's goodwill (or something like it) toward us.

example is intended to show that trust needn't involve any presumption of goodwill toward us. He suggests that, "Our enemies' restraint might be grudging, driven simply by a belief that it is wrong to shoot someone who has surrendered. To impute goodwill here would be to deprive the notion of all content. Nevertheless, talk of trust is not out of place" (1994 : p.65).

While I take it that the con-artist case is definitive in showing that Baier's Goodwill account is insufficient for trust, Holton's arguments against the necessity of a concern for goodwill in trust are much less successful. In respect of the estranged couple case, I take it that, if there is any level of interpersonal trust between them, then there is some presumed goodwill. Each would presumably feel guilt toward the other, should the child come to harm under their care. If not, though, it may be a case that each 'trusts that' the other will look after the child (as per section 1.8.1), rather than trusts the other *to* look after the child.

In respect of the 'white flag' example, this does not conclusively show that a concern for another's goodwill is irrelevant to trust. It might be that Holton's dispute is with the connotations of 'goodwill' as maybe implying some personal warmth toward you when you surrender. We can suppose the enemy doesn't have personal warmth, because he does not know you, and has *just* engaged in combat with you. We could, on that front, simply use a different term that doesn't have the connotations of personal warmth, but that is still within the realm of other-regarding attitudes, such as *respect*. I take it that respect equates to something like goodwill toward another, based on her status as another human.

We can therefore respond to Holton's dispute. Insofar the enemy is *moved* in a moral sense by your surrender, then he must have some goodwill toward you as another human being. After all, a soldier that doesn't care about upholding proper rules of engagement has no goodwill toward enemy combatants. In that sense, if the soldier is simply adhering to the rules of engagement on the basis that a punishment could befall him should he not, then he is not apparently concerned about you, and so has no goodwill toward you. Overall, Holton's claim that imputing goodwill to the enemy's motives 'deprives the notion of content' can be easily disputed. The only plausible reading of the white flag case where you trust the enemy, involves appealing to something like his respect for you as another

combatant. Moreover, respect is certainly - like goodwill - within the realm of other-regarding concepts.

In the next section I discuss Holton's own account, where I argue that his account involves a tacit appeal to goodwill as a relevant concern.

2.2.2 Holton: the 'Participant Stance' account

Holton's positive account of trust is built to deal with the con-artist case. Where Baier pinpoints an attitude of the trustee that we rely on when we trust (the trustee's goodwill toward us), Holton suggests that we do not rely on a specific attitude of a trustee. He supposes that while we rely on a trustee to do something, what actually makes our reliance trusting, and so what is missing from the con-artist's reliance on his mark, is an additional disposition toward having *reactive attitudes* of a trustee, just in case she disappoints our trust (1994 : pp.67–68).

The notion of a reactive attitude comes from Strawson, who conceives of the reactive attitudes as those ways we respond (react) positively or negatively to the good or ill will manifest in another's actions, in a way that manifests our own engagement in an interpersonal relationship with her (1974 : pp.5–6). Such attitudes include blame, resentment and gratitude. Crucially though, betrayal is a reactive attitude that we have toward manifest ill will. Since our goal is to secure betrayability, our preferred account of trust should therefore require that a trustor has the sort of psychology that makes any sort of reactive attitude appropriate. This is what Holton suggests we build into our account of trust. As such, his account of trust is that it is reliance from the kind of stance we adopt toward others when we are disposed toward reactive attitudes from their actions. Holton adapts Strawson's construal of this stance as an "attitude (or range of attitudes) of involvement or participation in a human relationship" (1974 : p.9), and gives it the more general label of a *participant stance*. On Holton's account, then, trust is reliance on another φ 'ing, plus a participant stance toward the one we rely on (1994 : p.66). We can call this the 'Participant Stance' account for short.

The Participant Stance account appears to avoid the con-artist counterexample, because the con-artist fails to take a participant stance toward you. This is because he is not

disposed to reactive attitudes to you, in the matter of you e.g. giving him your bank details. While he may be disposed to frustration at your possible failure to give your bank details, this isn't a reactive attitude in the sense that Strawson conceives of reactive attitudes. Reactive attitudes are those we have in our engagement or participation in an interpersonal relationship with another, not just attitudes we have due to things going well or badly for us, whatever our projects or ends are. A con-artist would experience those, but doesn't experience the reactive attitudes, as they come from engagement in interpersonal relationships.

Here we can raise a point against Holton's previous argument, that goodwill was unnecessary for trust, discussed in the previous section. The Participant Stance account seems to involve an appeal to goodwill. That is because his account invokes Strawsonian reactive attitudes. As we have just seen, these are ways we respond to another, based on perceived good or ill will in her actions. They are attitudes reflecting involvement in human relationships. As such, the Participant Stance account provides a 'Trojan horse' for concern about another's goodwill to re-enter the picture. This is not an issue for the account itself, because we saw that Holton's arguments, against the necessity of considerations of goodwill, were unpersuasive. Holton does note, though, that he is open to the accusation of inconsistency, and he addresses this in a footnote (1994 : p.66 footnote 9). There, he points out that Strawson later stipulates that the reactive attitudes can result from our concern about the *regard* in which others hold us. How regard is distinct from goodwill, though, is left unqualified. Presumably, a concern about regard is similar to a concern about respect (which I discussed above). So, a concern about regard is within the realm of concern about goodwill and ill-will. We are concerned about where we stand with another person. Thus, invoking regard doesn't obviously remove goodwill, or something related to it, from the picture.

I also suggest that the Participant Stance account, as a way of securing betrayability, can be accused of being relatively unexplanatory. It is right to point out that a con-artist doesn't have the requisite attitude toward his mark, that allows him to be betrayed. Holton identifies the missing attitude as a participant stance toward the mark. However, the account doesn't tell us why the con-artist doesn't adopt a participant stance, or why he couldn't. I claimed above that the con-artist doesn't adopt a participant stance, because he doesn't have an interpersonal relationship with the mark. But we don't yet have a

principled reason for supposing that a con-artist *couldn't* have an interpersonal relationship with his mark. As such, pointing out that the con-artist doesn't have the required attitude to be betrayable by his mark, doesn't shed light on why the con-artist doesn't trust his mark.

Certainly, the account identifies a missing necessary attitude, that would be required for the con-artist to be betrayable. But we already recognised that the con-artist lacked such an attitude, because he wasn't betrayable. This was how it served as a useful counter-example to the Goodwill account. The participant stance account avoids the example by brute stipulation that, if one is not betrayable, then one doesn't have a participant stance. It thus asks us to infer from the lack of betrayability in the con-artist, that the con-artist doesn't satisfy the conditions of the account. But it remains to be seen whether a con-artist could also have a participant stance. In fact, it seems plausible to imagine a con-artist being disposed to react *at some level* with interpersonal reactive attitudes, even if he may not be warranted in doing so.

A related point against the participant stance condition, is that it is too *general* to account for the interpersonal import of matters of trust in particular. Hawley (2014) argues that while the Participant Stance account captures a necessary condition for trust, it doesn't provide sufficient conditions for trust. It is indeed necessary that, in trust, we adopt a stance toward another from which reactive attitudes are appropriate. But, it is still true that we can rely on another from a participant stance, but not be disposed to *betrayal* specifically. If that is right then the Participant Stance account doesn't capture trust, because it doesn't get us betrayability.

Hawley gives an example to illustrate that reliance from a participant stance isn't sufficient for trust (Hawley 2014 : pp.7–8). I will re-frame it slightly. Suppose that Xavier comes to rely on his partner Yolanda cooking for him. He also adopts a participant stance toward her in his reliance, because he expresses gratitude to her when she cooks. However, should Yolanda not cook for Xavier one time, it isn't yet appropriate that we, or he, call this a betrayal of trust. That's because it isn't yet enough for Xavier to be relying on Yolanda cooking from his participant stance, for it to thereby be a matter of trust that she does cook.

I take it that this example is persuasive. Xavier relying on Yolanda to cook while adopting a participant stance toward her, isn't enough for Xavier to be betrayed, should she not cook. The participant stance, after all, isn't an attitude that is exclusively adopted toward another in a matter of trust. It is simply the general attitude we adopt toward another such that we are disposed to interpersonal reactive attitudes, which are responses to manifest good or ill will in her actions. Considering this point, I suggest we can explain why it is intuitive that Xavier can express *gratitude* to Yolanda for cooking, while it not follow he thereby must feel *betrayal* should she not cook.

Yolanda's continued cooking for Xavier can manifest good will toward him, and so he can express gratitude toward her for it. He would do so, because her cooking can amount to a repeated series of affectionate gestures. But it doesn't follow that not cooking can amount to a *failure* to cook or be a gesture of *ill will*, and so a failure to express respect to Xavier. Just because another has undertaken a schedule of affectionate actions, we aren't warranted in supposing that a failure to maintain that schedule of affectionate actions, thereby expresses ill will (perhaps Yolanda has had a busy day at work). But betrayal is a reactive attitude to perceived ill will. So, it doesn't follow yet that Yolanda's failure to carry out the affectionate act is going to be perceived by Xavier as an expression of ill will, and thus a betrayal.

However, there are two features of the case that, I think, stop the example being definitive in showing the Participant Stance account is inadequate. For one, I think our intuitions can be swayed toward thinking that Xavier *could* receive Yolanda's not cooking as a gesture of ill will, or a lack of good will. So, he could be treat it as a matter of trust, due to her not cooking. Thus, the example doesn't work as a case of participant stance without betrayability. This is because it arguably is a feature of human nature that a schedule of affectionate actions being maintained, can generate a normative/interpersonal expectation of those actions being continued. As a result, *not* performing an affectionate action one time, when we have come to anticipate it, can be received as manifesting a lack of concern/affection/good will toward us, or even as manifesting ill will.⁸ This inches us toward perceiving something like a betrayal, or perhaps a feeling of being let down, should

⁸ Hawley recognizes this - her account, discussed in the next section, explains this. It does so by allowing that we can incur *commitments*, which lie at the heart of matters of trust, via repeated actions.

that schedule be broken. Consider how even though as adults we know we shouldn't *expect* it of our friends and family that they celebrate our birthday, we might still feel hurt, abandoned, betrayed, or let down, should they fail to remember to celebrate our birthday (think of their efforts to make you feel disappointed and forgotten, prior to the reveal of your surprise party).

The 'repeated kind gestures' point indicates that the example may not be a case of participant stance without betrayability, as Hawley supposes it is. A further feature of the case can sway us toward the same conclusion. This stems from Xavier's *reliance* on Yolanda's cooking. We can argue that given Xavier's reliance on Yolanda's cooking, her cooking doesn't just represent a kind gesture for Xavier. If Xavier has come to rely on her cooking, this indicates he now has a practical need for it: he hasn't made alternative arrangements for dinner.⁹ Yolanda's cooking is now an action that serves to ensure something else he needs is brought about. If Yolanda *recognises* his reliance on her cooking, and makes no effort to assuage it, he may be right in treating this as a matter of trust.

This last point, though, is exactly what Hawley's positive account can explain. That's because the case begins to look like one where Yolanda has made a *commitment* to Xavier, to continue cooking.

2.2.3 Hawley's 'Commitment' account

The two features of the Yolanda's cooking case, that I suggested could make it possibly unpersuasive as an argument against the participant stance being a sufficient condition for betrayability, in fact help serve to motivate Hawley's positive account. The betrayability condition that Hawley puts forward can explain why these two features draw our intuitions toward thinking that Yolanda's repeated cooking could become a matter of trust to Xavier. That's because, by undertaking repeated gestures, and carrying on with apparent recognition of another's reliance on those actions, we can incur commitments to others. As such, the gap that is identified by Hawley in the Participant Stance account, between the participant stance and betrayability, is arguably filled by the concept she appeals to in her account of betrayability.

⁹ I defend the claim that reliance involves practical need, or dependence, in section 4.7.2, and chapter 5.

Hawley's positive account is concerned with pinpointing the specific type of "worldly situation to which (dis)trust is an appropriate response" (Hawley 2014 : p.9) and therefore what makes the possibility of betrayal *appropriate*. Betrayability then is secured by the trustor's belief that this betrayal-enabling situation obtains. The betrayal-enabling condition Hawley identifies as underpinning matters of trust, is that the trustor believes the trustee has made an interpersonally, normatively binding *commitment* to acting, via explicitly promising or implicitly encouraging the trustor's reliance on her to act (2014 : pp.10–12). Commitments, she explains, are broader than explicit promises:

Aren't there obvious, immediate counterexamples to the claim that trust and distrust presuppose commitment of this kind? We often trust people to do things which we know they have not explicitly promised to do. I trust my friends not to steal the cutlery when they come to dinner, and, at least in some areas, I trust strangers to let me walk unhindered. To make the account plausible I must use a very broad notion of commitment: commitments can be implicit or explicit, weighty or trivial, conferred by roles and external circumstances, default or acquired, welcome or unwelcome. In particular I will take it that mutual expectation and convention give rise to commitment unless we take steps to disown these. (2014 : p.11)

In addition to this, we can trust those who make commitments to third parties. By making a commitment to Y, we thereby make ourselves open to being trusted by Z. Hawley defends this using an example of trusting a friend who makes a promise to your daughter to pick her up from a party (2014 : p.11). You can trust the friend in this respect, despite the commitment being to your daughter and not to you.

So, on this account, when we trust another, we not only rely on her to act, but we also have a belief that she has made a commitment to acting, whether that is a commitment to *us* or to someone or other. We can call this the 'Commitment' account.

The appeal of the Commitment account is plain. By isolating the *kind* of situation where susceptibility to betrayal is appropriate, it builds on the groundwork set by the Participant Stance account, to tell us something about trust specifically. The Participant Stance account secured the *possibility* of betrayal. But it only did so, because it secured the *possibility* of that participant stance being adopted toward a matter of trust. The Yolanda's cooking case was designed to show that in relying from a participant stance we don't necessarily rely in the manner needed for trust, because we can still fail to be betrayable

from the participant stance. That's because we can still rely on another without taking her to have made a commitment.

In the Yolanda's cooking case, the two grounds on which we might have thought it was unpersuasive as an example of a participant stance without trust (repeated gestures of affection, and possibility of Xavier thinking she has recognised his reliance on her) are features that would lead us to attribute a belief to Xavier, that Yolanda has made a commitment to cooking. Philosophical accounts of promising (a species of commitment), tend to emphasise the importance of *invited* reliance/trust in generating promissory obligations, such that the reliance is recognised (Friedrich and Southwood 2011; Pink 2009; Scanlon 1990; Thomson 1990). In this respect, we can imagine that it wouldn't be completely unreasonable for Xavier to suppose that Yolanda has implicitly encouraged his reliance on her, by at least not assuaging his continued reliance on her cooking, following recognition of the fact of his reliance. What the commitment account does is crystallise a certain underlying structure to interactions found in matters of trust. This structure, roughly, is that another has led us to believe, or is acting in such a way that indicates, that she recognises the respect in which we could lose out from her not doing what she presents herself as doing. There may be some fact of the matter as to whether she *has* so acted. But, this is less important to the fact that she is *trusted*, than the trustor's perception that she has so acted.

I think we see how intuitive and explanatorily powerful the appeal to the notion of commitment is, when we reflect on how the notion of commitment closely maps onto Shklar's definition of the simple act of betrayal, mentioned in section 1.7. This was of "having an expectation disappointed," where "one person should have both intentionally convinced another person of his future loyalty and then deliberately rejected him" (1984 : pp.140–141). If Shklar's definition is right (and I take it to be highly intuitive), the notion of a commitment or a promise then can't fail to secure betrayability. That's since, at the core of the notion of betrayal, there is the archetype or structure of a promise being made, and then broken.

In summary, then; widespread adoption of REL+ in the philosophical literature has arguably enabled progress in understanding the nature of trust, because it has allowed us to reveal what the interpersonally normative aspect of trust consists in. By separating out

the interpersonally normative aspect of trust – betrayability - from its predictive aspect, we can systematically scrutinise what betrayability requires. Considering Hawley’s Commitment account, it appears that Baier’s Goodwill and Holton’s Participant Stance accounts were circling a submerged aspect of trust, that the Commitment account helps bring to the surface. We’ve thereby shown what was exerting a pull on our intuitions about betrayability in different cases, by showing that our intuitions on this seem to track with our intuitions about whether a commitment has been made. REL+, then, has arguably been philosophically productive. Given that, we appear to be well justified in adopting it as a general account of trust.

2.3 REL+ as an account of the trust stance

Once we commit to the idea that trust is reliance plus betrayability, a certain view of distrust falls out of this. This view of distrust falls out of the REL+ framework once we assume that distrust is in some sense an attitude that is *similar* or *symmetrical* to trust. Distrust is like trust in the same way that disbelief is similar to belief. Disbelieving something is similar to but not the same as believing it, and an account of belief ought to shed light on how.

As set out in section 1.1: distrust, being similar to trust, exists in the trust stance. The trust stance contains the attitudes that are ‘trust-y’ in nature. This obviously includes trust and distrust, as attitudes we can have toward another in matters of trust. Something being a matter of trust, after all, is also necessarily a matter of possible distrust. Trust and distrust are in some sense defined by their standing in opposition to one another. We know what it is like to trust someone, by how it is opposed to what it is like to distrust them. Because of their similarity, we ought to be able to explain distrust within the same framework we use to explain trust.

As stated above, a way of explaining distrust under the REL+ framework in fact falls out of the basic theoretical commitment of REL+. The basic theoretical commitment, of course, is that trust is built out of reliance. If we are assuming that trust is like distrust, then we should presume that distrust is grounded in a similar attitude to trust; one on same plane or in the same stance as reliance. We should assume, therefore, that the trust stance is built out of the ‘reliance stance’. We can see how this looks on Figure 2 below:

Figure 2: The trust stance and reliance stance under REL+

	<i>Reliance</i>	<i>Non-reliance</i>	
<i>Betrayability</i>	Trust	Distrust	(TRUST STANCE)
<i>Non-betrayability</i>	Mere Reliance	Mere Non-reliance	(RELIANCE STANCE)

This illustrates that the basic theoretical commitment underpinning REL+ does afford a way of accounting for distrust. In fact, it affords only *one* way of plausibly accounting for distrust. The only conceptual space that makes sense for it is betrayable *non-reliance*. This is the account of distrust that Hawley in fact endorses, in her Commitment account (2014 : p.10). I will explain what the motivation for endorsing this view of distrust is.

Distrust can't be *mere* non-reliance. We don't rely on people in all sorts of respects, yet this doesn't mean that we distrust them in those respects. We simply don't rely on many of the actions others undertake in their day. This doesn't entail that we distrust others in respect of those actions. That's because, for many of those actions, we will be simply unaware that they are being carried out. Of those we are aware of, we will be indifferent to most of them. As with the trust/reliance distinction, there is some interpersonally loaded aspect to distrust that elevates it beyond mere non-reliance on another's action.

Yet while distrust is interpersonally loaded, it isn't interpersonally loaded *reliance*. That just gets us trust again. So, given that distrust is interpersonally loaded, but isn't interpersonally loaded *reliance*, the only remaining conceptual space that makes sense to place distrust in is betrayable non-reliance.

Could we give a different account of distrust, while endorsing the basic theoretical commitment of REL+: that trust is reliance plus betrayability? Arguably not. For one, that trust and distrust are in some sense symmetrical – the mirror images of one another – compels us to endorse a reliance stance-based account of distrust. Recall the analogy between trust/distrust and belief/disbelief. Trust and distrust seem to be related in the same way that belief and disbelief are. This would imply that an account of the positive attitude (trust, belief) would use the same groundwork for an account of the opposing

negative attitude (distrust, disbelief). It would be surprising for instance if disbelief turned out to be a fundamentally different type of attitude from belief. In the same way, it would be surprising if distrust turned out to be a fundamentally different type of attitude from trust. If we were to, on the one hand, commit to a view of trust as a type of reliance, yet commit to a view of distrust as grounded in e.g. *hope*, we would have to ask why it isn't reasonable to ground *trust* in hope also; such that trust was a variety of hope. If there were a reason for supposing that distrust was based in some other kind of attitude besides reliance we would have to wonder whether that same reason also suggested that trust wasn't a variety of reliance either. So, taking distrust to be grounded in some different type of attitude to trust, would just call REL+ into question. Thus, given the way trust and distrust are related, we have *prima facie* reason to suppose that, by endorsing REL+ as an account of trust, we must endorse a reliance stance account of the trust stance more broadly.

This means that we are restricted to this account of distrust on REL+, because the reliance stance has only two options: reliance and non-reliance. As noted, distrust is interpersonally salient, in a similar way to trust. But since we've given over interpersonally salient (betrayable) *reliance* to trust, we have only one option for another betrayable attitude: non-reliance.

We're therefore compelled to adopt the 'non-reliance plus betrayability' account of distrust, once we endorse REL+. This is because trust and distrust are similar, and reliance/non-reliance is a binary. But that we are compelled to adopt the non-reliance plus betrayability account needn't be a problem. Such an account of distrust is an attractive one, for two main reasons.

First, 'non-reliance plus betrayability' accords with our intuitions about the interpersonal/normative aspect of distrust. It's already been noted that when we distrust another, we regard her in a way that casts an interpersonal/normative/ethical light on our dealings with her. As with trust though, we could reasonably pin this to *betrayability*, because we can be betrayed by those we distrust. Consider a case where, for example, a jealous husband hacks into his wife's email, phone, social media accounts etc. and finds that she has been cheating on him after all. In such a case, he is betrayed, even though he already distrusts her. Similarly, distrusted politicians can still betray or let down

constituents by failing to do what those constituents expect of them (Hieronymi 2008 : p.229). Hawley gives an example in support of the thought that distrust can be betrayed; Jesus distrusted Judas, yet was still able to be betrayed by him (2014 : p.13). So that REL+ offers an account of distrust as betrayable non-reliance appears to accord with our intuitions: we can be betrayed by those we distrust.

Second, this account of distrust accords with our intuitions about the predictive aspect of distrust. We've seen that while trusting someone casts them in a positive light, distrusting someone casts them in a negative light. Trust involves a certain kind of optimism about another, and distrust a kind of pessimism. This also applies to our understanding of trustworthiness. If you are trustworthy, this is (generally) a good thing, whereas being untrustworthy is (generally) a bad thing. That trust and distrust cast others in opposing lights cannot be due to them having a differing sort of interpersonal component, though, since they appear to have that as a common ground: they both leave us susceptible to betrayal. Given that the other component of trust is reliance and distrust non-reliance, we should infer that the optimism of trust is due to its involving reliance, and the pessimism of distrust due to its non-reliance.

Tying the optimism of trust and the pessimism of distrust to the reliance/non-reliance distinction seems to make a lot of sense. For one, reliability is preferable to unreliability. Recall the Kant example: even if Kant's neighbours merely rely on Kant, it is still preferable to them that he walks. If Kant's walking was an unreliable measure of time they wouldn't wish to rely on his walk for their timekeeping. It also makes sense that trustworthy individuals are reliable in some interpersonally loaded way, and untrustworthy individuals are unreliable in some interpersonally loaded way.

The account of distrust that falls out of REL+ then is an attractive one. Trust involves optimism, and distrust involves pessimism, by the fact that trust involves reliance, whereas distrust involves non-reliance. There is good reason for thinking this, because reliability is connected to predictability. Kant was predictable in his routines and this made him reliable. This *trust stance* version of REL+ tells us that trust and distrust are distinct attitudes – the mirror image of one another – in terms of having opposing predictions of another's performance of some action. While the REL+ schema leaves us one way of explaining distrust, that way seems to be natural and intuitive.

Now we can see that, in addition to committing to a view of distrust, committing to REL+ commits us to a view of what *neither* trust-nor-distrust of another amounts to. As per Figure 2 above, the trust/distrust and ‘neither-nor’ divide runs along the ‘trust stance’ and ‘reliance stance’ divide. In other words, all forms of ‘neither-nor’ reside in the reliance stance. See Figure 3:

Figure 3: Trust/distrust and neither-nor under REL+

	<i>Reliance</i>	<i>Non-reliance</i>	
<i>Betrayability</i>	Trust	Distrust	(TRUST/DISTRUST)
<i>Non-betrayability</i>	Mere Reliance	Mere Non-reliance	(NEITHER-NOR)

This figure is intended to illustrate how the distinction between trust/distrust and neither trust-nor-distrust, on REL+, runs along the betrayable/non-betrayable attitude distinction. This way of accounting for neither-nor seems to make sense. Once we switch off betrayability, we are left with mere reliance and mere non-reliance, which clearly amount to neither trust-nor-distrust. That Kant’s neighbours merely relied on Kant meant they didn’t trust him. But nor did they distrust him. Also, that we simply don’t rely on others doesn’t mean we distrust them. Mere reliance and mere non-reliance on others reside outside of the trust stance. Thus, it seems right that neither trust-nor-distrust can be pinned to the non-betrayable ‘reliance stance’, consisting of reliance and its absence.

I’ve illustrated that if we assume REL+, then a certain account of distrust and neither-nor results. Such accounts are also endorsed in the literature. Hawley for example sets out her account of distrust as non-reliance plus betrayability, along similar lines of reasoning as that set out above, concerning the symmetry between reliance/trust and non-reliance/distrust. She supports extending the REL+ schema to explaining distrust by noting that distrust can involve the expectation of betrayal (2014 : p.13), and ties distrust to an expectation of unfulfilled commitment (2014 : p.1). In other words, her account of distrust is as a predictive expectation of betrayal.

A similar view of trust and distrust is endorsed by Helm (2014). While Hawley explicitly sets out her account of distrust, Helm intends that his account of trust extends *mutatis*

mutandis to explaining distrust (2014 : p.214 footnote 6). Hawley's and Helm's accounts differ in the details. Nonetheless, their accounts share common ground in the respects at issue here. Helm's account presumes that trust is normatively significant reliance and that trust and distrust involve predictive expectations akin to optimism and pessimism about another's performance. It also assumes that trust and distrust are similar in respect of containing the same commitment to a normative standard, and where that normative standard is what makes us betrayable (Helm 2014 : pp.199–205). The view of trust, distrust and neither-nor set out in Figure 2 and Figure 3 above is thus embodied by at least two prominent REL+ accounts of trust.

In the next chapter, I present an argument against the view of the trust stance that we have just arrived at. Given that this is the view of the trust stance that falls out of REL+, this argument works against REL+.

3 The argument from uncertainty: against REL+

In this chapter I present an argument against REL+. REL+ cannot be correct as an account of trust, because the view of the trust stance that results from it is unable to accommodate an attitude that exists in the trust stance: uncertainty of another. As such, a necessary consequence of REL+ is false. Therefore, REL+ is false.

3.1 Setting out the argument

The argument from uncertainty is a *reductio ad absurdum* against REL+. It is used to motivate the claim that the trust stance cannot be grounded in the reliance stance, for the reason that the reliance stance doesn't afford the number of attitudes needed to ground those of the trust stance.

The argument is this. If REL+ is true, it is impossible to be betrayable by another *on a matter of trust*, while being neither trusting nor distrusting of her. This was clear on Figure 2 and Figure 3 in the previous chapter. So, if REL+ is true, it is impossible for there to be a further attitude in between trust and distrust. But having a sort of middle ground attitude between trust and distrust is possible. We can be in between trust and distrust of another when we are *uncertain* of her. What sort of attitude do I mean by being uncertain of someone? Consider the following case:

New Colleague

A new person, Sophie, starts at your workplace. You make a point to introduce yourself and have a chat. Sophie seems very nice and you find yourself sharing a few jokes, getting to know a bit more about her, and you tell her a bit about yourself. You also bring her out for drinks after work with the rest of your team, where you get her up to speed with the office gossip. You have a good first couple of weeks getting to know her more, and she quickly becomes a trusted member of the team.

In the next week you catch yourself feeling oddly hesitant about Sophie. In the lunchroom you see her behave strangely before she takes her food from the fridge. Then while having a chat with her in the lunchroom she makes some

jarringly disparaging remarks about colleagues who you know very well and hold in high esteem. You can't quite understand why she would make these comments. You're keeping an open mind about her, even though you can't ignore that you now feel a bit less sure than you did. If pressed, you wouldn't say that you trust her, though you would have done previously. But you wouldn't say you distrust her either; that would be too strong. You're more inclined to say that you're unsure or uncertain of Sophie. Depending on what instances of her behaviour you recall and what reasons for that behaviour you posit, your attitude about her could veer from some trust to perhaps some distrust. Nonetheless, your attitude remains somewhere in between trust and distrust of her. You subtly raise your concern with some of your colleagues and find that most of them also share your uncertainty about her (though one or two are simply indifferent, not caring or wanting to get involved in 'office politics').

A couple more weeks go by and Sophie has now consistently turned up to work late. One day last week she was missing from the office for several hours during work time. You sneak a glance at some timesheets she has on her desk, that it seems she has yet to submit to your manager (who rarely closely monitors anyone's arrival and departure times). You notice that the times don't correspond at all to the actual times she has been in. You also can't help but notice that your other colleagues are much less talkative with each other. You suspect that her disparaging remarks have helped cultivate a bad, tense atmosphere. Your lunch has also gone missing from the shared fridge on a couple of occasions. You decide to raise that with her and she denies all knowledge, offering a longwinded and insincere seeming explanation, which you don't buy at all. You now recognise that you are no longer keeping an open mind, because you're not uncertain or unsure of her. Instead, you simply distrust her. You're sure that she can't be trusted.

When you are uncertain of Sophie, I suggest that you are neither trusting-nor-distrusting of her, yet you are 'betrayable' – i.e., open to being interpersonally wronged or harmed in the manner characteristic of the trust stance – by her. For example, you can be betrayed by her in the period you are uncertain, should you see her steal your food from the fridge. Your attitude toward her there would involve the same characteristic interpersonal or normative salience as your attitude when you initially trust her, and when you later distrust her. It isn't that your attitude is outside of the trust stance, in the mere reliance/non-

reliance stance. As such, we have an attitude that constitutes a form of neither-nor, *yet* is betrayable. This is not possible on REL+, as we saw on Figure 2 and Figure 3. But this suggests that REL+ is false.

While this is an imagined case, the kind of scenario it presents isn't hard to envisage, or anything out of the ordinary. What's being described highlights some everyday mechanics of our interactions with others. It's commonplace to feel something in between trust and distrust about another. We can refer to this sort of middle ground attitude in different ways. Uncertainty of another is how I refer to it. We could also refer to it as a feeling of being unsettled. We can say we have some doubts or suspicions, are hesitant, unsure, ambivalent about another. Different idioms and phrases are associated with such an attitude: keeping an open mind, reserving judgement, or seeing how things pan out. In each case, what we report is that we neither trust nor distrust the other, *yet* our attitude and interactions with her are still at some level concerning a matter of trust. In other words, our attitude toward the other is within the trust stance, even if our attitude amounts to neither trust nor distrust of her.

The possibility of being neither-nor about the other is acknowledged in various places in the trust literature. It is sometimes claimed that trust and distrust are contraries but not contradictories (Govier 1992b : p.18, 1998 : p.121; Jones 1996 : p.15). What this means is that the absence of trust needn't imply the presence of distrust, and vice versa. If trust and distrust were contradictories it would be impossible to have an attitude of neither-nor towards the other, because the absence of one attitude would necessarily entail the presence of the other. By negating distrust, you would affirm trust, and vice versa. But, we can both fail to trust and fail to distrust the other, because we can be neither-nor of her concerning some matter.

However, a crucial distinction is often missed between two types of neither-nor. REL+, as we saw in the previous chapter, accounts for one category of neither-nor; that where our attitude is outside of the trust stance. REL+ is built to account for this sort of neither-nor. Recall Baier's example of Kant's neighbours' 'mere reliance' on him. In that case, Kant's neighbours' attitude toward Kant was neither-nor. They didn't trust him because they merely relied on him. But this didn't mean that they distrusted him either. Their attitude toward Kant wasn't a betrayable one; it was outside of the trust stance altogether.

The possibility of being neither-nor toward another is part of the foundation of the view itself.

Hawley in her REL+ account endorses this way of drawing the distinction between trust, distrust and neither-nor, claiming that just as trust is not mere reliance, distrust is not mere non-reliance. She says:

Distrust is not mere absence of trust. Moreover, distrust is not even mere absence of reliance. Like trust, distrust has a normative dimension. The distinction between trust and mere reliance shows in our different reactions to misplaced trust (betrayal) and misplaced reliance (disappointment). [...] So distrust is richer than mere nonreliance, just as trust is richer than reliance. Just as we should distinguish trustworthiness from mere reliability, we should distinguish untrustworthiness from mere unreliability: colleagues who do not buy me champagne are unreliable in this respect, but not thereby untrustworthy. Just as there is a middle ground between trust and distrust, there is a middle ground between trustworthiness and untrustworthiness—in the clearest case, inanimate objects merit neither trust nor distrust, so they are neither trustworthy nor untrustworthy. (2014 : p.3)

The neither-nor middle ground that Hawley is referring to then is the same one that Baier's 'Kant's neighbours' example points out. It is the neither-nor where our trust/distrust concerns – the concern that makes us betrayable - are absent. This is of course going to be the case with most matters, because the majority of others' actions/possible actions won't be matters of trust. We will be unaware or indifferent to most of the things that others do/intend to do.

Uncertainty is a different kind of neither-nor from the sort that REL+ is built upon, though. And, it is a problem for REL+, because it is an attitude distinct from trust and distrust, that nonetheless belongs in the trust stance. Being uncertain of another amounts to neither trusting-nor-distrusting another, yet with the possibility that we are betrayed by her. This is so in the New Colleague case.

Furthermore, remarks from different authors suggest that the existence of this betrayable type of neither-nor is not in question. Govier for example notes the sort of attitude I mean:

Trust and distrust are contraries, not contradictories. To say that it is the case that one trusts another person is not always to say that one distrusts him. We may

neither trust nor distrust another, either because we lack relevant knowledge or because the evidence and feelings we have are mixed. (1998 : p.121)

What Govier says above is that we can neither trust nor distrust someone because we have mixed feelings; somewhere between trust and distrust. This is a distinct kind of neither-nor from that which we've noted above, where others' actions just doesn't matter to us in the way characteristic of the trust stance. In having mixed feelings, it can matter to us what another does, in the trust stance way. But our attitude just doesn't amount to trust or distrust. It sits somewhere in the middle.

Lagerspetz also acknowledges this middle ground attitude, which he refers to as suspicion. This occurs in the context of a critique of Baier's 'accepted vulnerability' account of trust. He says that, in being geared toward finding a theory of what are the "appropriate objects and motives for trusting" (Lagerspetz 2015 : p.49), Baier's account belies a tradition in trust theory, of treating the lack of trust as the default position toward others. He suggests that, on this tradition, distrust *or suspicion* is treated as the starting point of human relationships:

The idea of looking (or shopping for worthy objects of trust implies a presumption in favour of distrust: distrust is the default stance until an assessment of the reasonableness of trusting has been produced. One might perhaps want to describe this initial attitude as neutral, but 'suspicious' might be the more apt description insofar as the agent is supposed to perceive the situation as implying that it is *just as possible as not* that the other has hostile designs. (2015 : pp.49–50)

There is contrast drawn here between attitudes of distrust and suspicion, but also (crucially) with a 'neutral' attitude. Suspicion is taken to be the attitude of perceiving another as *possibly* having 'hostile designs' or motives, but *also possibly not*. Neutral, presumably, would be the more apt description for an attitude outside of the trust stance altogether, held when another's motives do not concern us as a matter of trust.

So far this is enough to help grasp what attitude I mean by uncertainty of another. But why exactly would uncertainty of another be problematic for REL+? Recall Figure 3:

	<i>Reliance</i>	<i>Non-reliance</i>	
<i>Betrayability</i>	Trust	Distrust	(TRUST/DISTRUST)
<i>Non-betrayability</i>	Mere Reliance	Mere Non-reliance	(NEITHER-NOR)

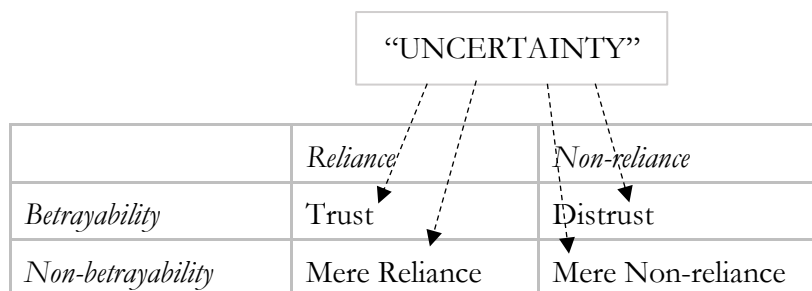
Here we see that it is impossible for neither-nor to transcend the boundary between the betrayable and non-betrayable attitudes, set out by REL+. In other words, it is impossible for there to be an attitude which is betrayable, yet isn't trust or distrust. Uncertainty of another *is* such an attitude. So REL+ must be wrong-footed. If it were true, an actual attitude wouldn't exist. But this attitude does exist. So, REL+ is false.

3.2 Responses

The problem arises because the New Colleague case presents us with a further attitude to place in the trust stance, which is an attitude of betrayable neither-nor. But there is no further space for an additional betrayable attitude, because distrust has taken up the only other space *left* in the trust stance, once we've accounted for trust. This is the space for a betrayable form of *non-reliance*. As such one possible line of response could have been to deny that distrust should be placed in this category, since this would free up a space for uncertainty. But, this response was dealt with in section 2.3. Denying that distrust should be placed in the category for a betrayable form of non-reliance would just throw REL+ into doubt.

It seems that the most obvious way of responding to the argument is to deny that uncertainty is a distinct attitude; to deny that there is a betrayable form of neither-nor. We could e.g. explain uncertainty away as one or both of the 'mere' attitudes, or as one or both of the trust stance attitudes. In any case, denying uncertainty amounts to keeping the binary categories in the betrayable and non-betrayable rows, i.e. to keeping the setup seen on Figure 2 and Figure 3. This response then explains away or dissolves uncertainty into one or more of the existing categories of attitude, represented in Figure 4 below:

Figure 4: Denying uncertainty is distinct



First, I consider the prospects of explaining away uncertainty as a combination of the trust stance attitudes. Second, I will do the same but for the reliance stance attitudes.

3.2.1 Dissolving uncertainty into trust/distrust

The underlying thought in this type of response is that, once we drill down further beyond the surface appearances of cases like New Colleague, what we find is either trust, distrust, or some sort of mixture of trust and distrust, rather than a distinct sort of attitude. There are a few variants of this response, which I consider in turn.

The ‘averaging out’ response

One way we could motivate a response to the argument from uncertainty is by pointing out that trust is in some sense ‘domain specific’, as noted in section 1.2. This is the idea that we trust people in some respects but not others. Govier for example says that “Both trust and distrust are often relativized to role or context: we may, for example, trust someone as our dentist but not as an accurate commentator on political affairs” (1992b : p.18). So just because we trust someone in one matter, or possible domain of interaction, does not entail that we trust her in other domains.

Perhaps then the New Colleague case is one where in some respects you trust Sophie, and in others you distrust her. The *averaging out* of all these respects is overall neither trust-nor-distrust. But the average results from you trusting or distrusting her in specific respects.

The quick counter to this is to point out that your attitude of uncertainty about Sophie also concerns specific things she does (or may have done). For example, you are initially uncertain about her with your lunch. Eventually though you just distrust her, after you confront her about it. So, if we pursue the ‘averaging out’ response, the problem of uncertainty arises again once we point out that uncertainty occurs at the level of specific domains. You can be uncertain of Sophie in respect of whether she’d taken your lunch.

The ‘simultaneity’ response

A similar line of response, which again suggests that what we have is a mixture of trust and distrust, is to say that uncertainty is just *simultaneous* trust and distrust. This is a distinct response from the ‘averaging out’ one. The averaging out response left open the charge that we can be uncertain in specific domains. Simultaneity plugs this. We can explain the appearance of uncertainty in specific domains, as the result of you simultaneously trusting and distrusting *in those domains*. So, your uncertainty of Sophie in respect of taking your lunch is the result of you both trusting and distrusting her in that respect. There may be some reason for favouring this response when we recall that uncertainty could also be referred to as having mixed feelings, or being ambivalent.

There are two problems with this response. First, it just strikes as implausible that we can simultaneously trust and distrust, in some specific respect. It is difficult to interpret this as a coherent psychology; akin to both believing and disbelieving p at the same time. Second, it seems that if we do presume that we can trust and distrust in some respect simultaneously, then *so much for REL+*. Simultaneity on REL+ has to be explained by our being in *logically* contradictory states; simultaneous reliance and non-reliance *in some specific respect*. In other words, both relying *and not relying* on another for something. This violates the law of non-contradiction, which dictates that a statement and its negation cannot both be true at the same time. Recall that non-reliance on another φ ’ing amounts to it *not* being the case that one is relying on another φ ’ing, as opposed to relying on her not- φ ’ing. As such, ‘X is relying on Y φ ’ing’ and ‘it is not the case that X is relying on Y φ ’ing’ (‘X is not relying on Y φ ’ing’) are contradictory statements. But if X is to both trust Y to φ and distrust Y to φ at the same time, on REL+, then these statements must both be true.

In some domains/areas of discourse, denying the law of non-contradiction may be an attractive course of action, due to problems that arise when pinning down the truth of statements about those domains. However, doing so in this domain - whether a subject is or isn’t relying on another - would be an *ad hoc* move, until we have a reason to think that the law of non-contradiction restricts us when describing reliance states. As yet we have no independent justification for supposing that it can be true we are both relying and not relying on another φ ’ing. Instead, questions of reliance appear clear cut: if it is false that you are relying on Sophie to meet her workplace obligations, then it is true that you are *not* relying on Sophie to meet her workplace obligations.

The 'oscillation' response

Instead of suggesting that there is simultaneous trust and distrust, we can say that uncertainty is the result of an oscillation between trust and distrust. Rather than simultaneous trust and distrust of Sophie in respect of your lunch, there is cycling between trust and distrust of her in that respect. Again, the example itself provides some motivation for this response. It says that when you are in that uncertain state concerning her, that your attitude can veer from trust to distrust depending on what behaviours you recall, and what sort of reasons you posit for that behaviour.

The oscillation response though is suspect. In the specific respect of Sophie taking your lunch, it doesn't seem that your uncertainty is due to your flipping from trust to distrust of her in that respect. You simply aren't sure of her in respect of whether she would take it or not. There isn't a clear period where you trust and then distrust her on this.

Of course to this one may press the point about your attitude veering depending on how you construe it. But the veering only shows that uncertainty is *unstable*, easily tipping into one or the other state either side of it; not that it doesn't exist as a distinct state. In fact, the veering provides more motivation to suppose that there *is* a distinct state between trust and distrust.

If the oscillation picture is right, what we really have in uncertainty (in a specific respect) is transitions between discrete states occurring. But it seems odd to say that, *whenever* we transition from trust to distrust or vice versa (an everyday occurrence), *oscillating or not*, that there is a moment where our attitude clearly changes from trust to distrust. So in other words, where at $t-0.00[\dots]1$ we trust and at time $t+0.00[\dots]1$ where we distrust (or vice versa). It may be that there are rapid instances of transition from one to the other, when we quickly have it shown to us that someone is (dis)trustworthy by something they've done. Insisting that we always rapidly do this in *every* case, without further argument, is ad hoc.

The 'it's distrust' response

Presumably the response that uncertainty is just trust is implausible. However, it isn't immediately implausible that uncertainty is just distrust. Rather than locating a distinct attitude in uncertainty, what we have found is just another case of distrust.

One problem with the 'it's distrust' response is how to make sense of why it isn't *obvious* that uncertainty is distrust, without this collapsing into a variation of the mixture responses above. In New Colleague there is a period where you clearly trust and eventually where you clearly distrust Sophie. But in between this there is a period where it seems you neither trust-nor-distrust her. It isn't clear how to explain what the qualitative difference is between that neither-nor period and the eventual distrust of her, without appealing to it being intermingled with trust at some level.

A further problem with the 'it's distrust' response is that this would mean your non-reliance on Sophie in fact occurs much earlier than it would seem to. Remember that if we are committed to REL+, and arguing that that period of uncertainty is distrust (without also endorsing some mixture or intermingling response) then this cashes out in terms of you beginning to *not* rely on Sophie in some respects. In other words, it would require that you withdraw from Sophie in various respects much sooner than you in fact seem to do in the example. When you are uncertain of her in respect of your lunch you don't stop putting your lunch in the shared fridge. Your uncertainty isn't tied to any non-reliance on her in that sense. But on REL+, distrust is tied to non-reliance. This gives us reason to suppose that your uncertainty isn't just distrust.

3.2.2 Dissolving uncertainty into reliance/non-reliance

While we seem to have exhausted the prospects of explaining uncertainty in the trust stance attitudes on REL+, there is the possibility that we explain it away as belonging to the reliance stance. If we can do this then the pressure that uncertainty exerted on REL+ is deflated. Uncertainty turns out to not be betrayable after all.

To do this though we need to make a case for uncertainty being outside of the trust stance. This seems difficult. That said, one way to motivate the response that uncertainty is outside of the trust stance might be to reflect on how the attitude is sometimes referred to: keeping an open mind, being unsure, reserving judgement, having doubts and so on. These seem to put what we might call a 'cognitive spin' on your attitude. In other words,

it could be argued that, from the sounds of things, when you are uncertain you are not so much engaging with another in adopting a trust stance attitude to them, but instead are stepping back from that, perhaps momentarily. You are attempting to figure out whether someone is trustworthy or not, rather than engaging with them in the direct interpersonal manner of trust and distrust.

But even if we are in some sense stepping back, this stepping back needn't take us outside of the trust stance. In other words, stepping back and trying to figure out someone doesn't render us non-betrayable by them. In the New Colleague case for example you are still betrayable by Sophie even when you start to have doubts about her. She can still betray you by e.g. taking your food.

This response can be made more pressing if we compare the kind of stepping back you do with Sophie from the kind of stepping back that Kant's neighbours might do with Kant, if he began to noticeably deviate from his regular walking time. In that case the neighbours would also presumably step back, assessing whether Kant's walking was still reliable. But this is a different kind of stepping back from that which you do in the New Colleague case. It is qualitatively different, because for Kant's neighbours any stepping back would be assessing whether it was instrumentally prudent for them to continue relying on his walking. They are trying to figure out whether he will continue to walk or not. In the New Colleague case you are 'stepping back' not in the sense that you are trying to figure out whether Sophie will or won't do something. You are stepping back and trying to figure *her* out. You have suspicions about her, and are trying to determine what her *motives* concerning you are, rather than figuring out whether she will or won't do some specific action that would be instrumentally beneficial or detrimental for you.

So, uncertainty is a trust stance attitude because, as the New Colleague case indicates, it involves being concerned about how others will act, where this concern is not purely a matter of whether relying or not on someone would be instrumentally beneficial. The trust stance involves a non-instrumental concern about how others will act (see chapter 8 for further defence of this point), and this is what uncertainty involves.

3.2.3 The third category response

An additional response might be to deny that reliance/non-reliance is a binary option. Instead, we could suppose that there is a ‘neither reliance nor non-reliance’ category of attitude (indicated by ??? in figure below). This would mean we had an additional category of non-betrayable attitude. Uncertainty would no longer be a problem with this extra category, because now we have the available conceptual space in which to ground a corresponding additional betrayable attitude.

Figure 5: The third category response to uncertainty

	<i>Reliance</i>	???	<i>Non-reliance</i>	
<i>Betrayability</i>	Trust	Uncertainty	Distrust	(TRUST STANCE)
<i>Non-betrayability</i>	Mere Reliance	???	Mere Non-reliance	(RELIANCE STANCE)

The issue with this response is that it violates the law of excluded middle, which states that for any statement p , either p or $not-p$ must be true. Recall the simultaneity response; if we assert that ‘X is relying on Y φ ’ing’, but also assert that ‘it is not the case that X is relying on Y φ ’ing’ (/‘X is not relying on Y φ ’ing’), we violate the law of non-contradiction. The third category response instead amounts to asserting that neither is it true that ‘X is relying on Y φ ’ing’, nor is it true that ‘X is not relying on Y φ ’ing’. Neither p (‘X is relying on Y φ ’ing’) nor $not-p$ (‘X is not relying on Y φ ’ing’) is true.

A similar point applies here to the one made in the simultaneity response, regarding the law of non-contradiction. In some domains/areas of discourse, denying the law of excluded middle may be an attractive course of action, due to problems that arise when pinning down the truth of statements made about those domains. However, doing so in this domain – whether a subject is or isn’t relying on another - would be an *ad hoc* move, until we have a reason to think that the law of excluded middle is inadequate for describing reliance states. As yet, we have no independent justification for supposing that it can be true we are neither relying nor not relying on another φ ’ing. Again: questions of reliance appear clear cut: if it is false that you are relying on Sophie to meet her workplace obligations, then it is true that you are *not* relying on Sophie to meet her workplace obligations.

3.3 Summing up

The argument from uncertainty shows that REL+ leads to a *reductio ad absurdum*. It cannot account for the everyday possibility of being uncertain of another, where we neither trust-nor-distrust another, yet where that neither trust-nor-distrust is still interpersonally loaded in the same way that trust and distrust are. Committing to REL+ though requires that neither trust-nor-distrust never be interpersonally loaded in the same way as trust and distrust. As such, we should revise our paradigm theory of trust, taking lessons from why REL+ failed. We should seek a different kind of base attitude in which to ground the ‘trust stance’ of attitudes (trust, uncertainty and distrust); one that doesn’t lead to the same *reductio*.

4 The argument from distrust & the active/stative distinction

Like the argument from uncertainty, the argument from distrust shows that REL+ is false. It goes further though, in illustrating that there is a distinction between two senses of ‘trust’: active *reliance* and stative *assurance*. This shows how reliance can ‘come apart’ from the trust stance, and provides us with a base for characterising what the base anticipatory attitudes in the trust stance are.

As we saw in the previous chapter, committing to REL+ means committing to a view of distrust as a type of non-reliance. In this chapter I show that distrust is not a product of non-reliance. Instead, X distrusts Y in respect of φ ’ing, only if X is wary of Y’s motives. This wariness attitude is not a product of non-reliance, because it can also be present when X relies on Y to φ . This allows us to infer that there is a form of attitudinal trust, distinct from reliance, that opposes this wariness. I call that attitude assurance trust.

The resulting picture of the trust stance is that we have an activity of trusting someone, where we actively rely on another. But this is underpinned by anticipatory attitudes. These are where we find assurance trust, uncertainty, and distrust as a form of wariness. The argument from uncertainty will show that there is no corresponding ‘active’ distrust, only an attitude of wariness.

4.1 The Commitment account

For the argument from distrust, I will be using Hawley’s Commitment account of trust and distrust as a foil. This is because it provides a clear pairing of conditions for both trust and distrust, which give a solid basis for constructing example cases of both. However, I am using the examples to show how REL+ is flawed, not that the Commitment account in particular is flawed. These example cases will ultimately allow me to illustrate the redundant role of *non-reliance* in attributing distrust to a subject.

The Commitment account, recall, is the following:

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. (Hawley 2014 : p.10)

This gives us the following ‘plus’ condition, common to both trust and distrust:

COMMIT: X believes that Y has a commitment, to someone, to φ

This is conjoined with a non-reliance condition, to make distrust:

NON-REL: X doesn’t rely on Y to meet her commitment

Commitments here are understood as akin to promises to perform certain actions. We can take it then that ‘X relies/doesn’t rely on Y to meet her commitment’, when unpacked, becomes:

NON-REL: X doesn’t rely on Y to φ

Where φ is the action that Y has made a commitment to perform.

To simplify the argument, I will use cases where the commitments in question are explicit promises, and not ones conferred by “roles and external circumstances” or convention (Hawley 2014 : p.11).

4.2 An initial problem with COMMIT: non-direct commitments

First, I deal with an initial problem with the Commitment account. This arises due to a problem with the formulation of COMMIT. The default formulation makes room for the possibility of trust and distrust toward those who make commitments to ‘third parties’. This allows that the commitment at issue not be *to* the subject X, but to someone or other. But allowing that it is to someone or other, makes the account vulnerable to a type of counterexample. An example of this type is given below:

News Report:

COMMIT: I see it reported on the news that Trump has a commitment to meet with Merkel.

NON-REL: I don't rely on Trump to meet Merkel.

It doesn't follow that I distrust Trump from this example, even though the conditions for distrust are satisfied. This means that the initial formulation of the commitment account is off.

A quick way of avoiding the above sort of counterexample is to restrict COMMIT to dealing with *direct* commitments. This is perhaps unsatisfactory in terms of accounting for how, as several authors agree (Helm 2014; Holton 1994; Jones 2004), we can trust or distrust others in 'third party' or non-direct trust cases. Such cases are those like the one Hawley gives, where your daughter's friend promises to pick her up from a party, and you can take a trust stance toward the friend (Hawley 2014 : p.11).

We can assuage this concern, I think. It may be that the 'third party' cases, that intuitively strike us as matters of trust, are really an odd sort of case where the commitment at issue is at *some* level to you. Consider for instance that both Holton's and Hawley's non-direct trust cases involve commitments of care, by someone else, to a dependent of yours. Holton's case is one where you trust your estranged partner to look after the child you had together (Holton 1994 : p.65), even if the partner has made no overture or commitment to you in this respect. You can still trust the partner to look after the child, according to Holton.

This and the daughter-friend case share the feature of being about dependents. This is telling. They are cases where someone takes on the burden of care for the third-party – the dependent – whose welfare is normally your responsibility. This suggests that these are cases where the *role* another adopts plays a part in her garnering commitments. The friend and partner take on the role of carer for someone of whom you are normally playing or sharing that role, and so with it incur a commitment to *you*. It seems right to suggest that by taking on certain roles we can incur certain commitments (and as noted above, Hawley suggests this). Taking on a temporary role of carer would arguably incur commitments to the normal carer. Even if we agree with our intuitions that these third-party cases are genuinely cases of trust *of* the committers, we have reason to suppose that the commitment implicate you at some level. The *News Report* case above is the *reductio* of

the thought that there is no need for the commitment to be to you at some level in order to make it a matter of trust.¹⁰

As such, we can make the commitment condition the following:

COMMIT_{DIR}: X believes that Y has a commitment *to her* to φ

Where a commitment being *to X* can be in the indirect way of the estranged spouse and daughter's friend cases.

4.3 Non-reliance is unnecessary for distrust

In this section we move on to cases that put further pressure on the Commitment account of distrust. Ultimately, it will be shown that NON-REL (non-reliance) is unnecessary for distrust.

Let's return to the planned meeting between Trump and Merkel. Suppose that Merkel is known to have problems with acid reflux, and that before the meeting Trump promises to bring her some antacid tablets.

Antacid Tablets:

COMMIT_{DIR}: Merkel believes Trump has a commitment to her, to bring her some antacid tablets for the meeting.

NON-REL: Merkel doesn't rely on Trump to bring her antacid tablets; she brings some to the meeting.

It doesn't follow, from these two conditions being satisfied, that Merkel distrusts Trump. The reason is that we do not know *why* Merkel has brought her own antacid tablets. To

¹⁰ Alternatively, perhaps it is that Holton's and Hawley's cases are not ones where we trust *the committer*, but trust *that* the committer will such-and-such (see section 1.8.1). Either way, we can resolve the problem case, because either solution indicates that the COMMIT condition should be finessed to say that the commitment must be to the subject of the trust stance, at some level.

see this more clearly, imagine the following variation of NON-REL, compatible with how it is specified above:

Needed Change:

NON-REL: Merkel doesn't rely on Trump to bring her antacid tablets; she brings some to the meeting. She realised she needed change for the coffee machine in the meeting venue, and it was convenient to buy tablets in order to break into a bank note.

The *Needed Change* variant of *Antacid Tablets* doesn't suffice for distrust of Trump. The reason why Merkel doesn't rely on Trump to bring the antacids is due to reasons unrelated to her trust stance of Trump; it is for reasons of practical convenience. To reinforce this point, imagine this further variation of the NON-REL condition:

Jacket Pocket:

NON-REL: Merkel doesn't rely on Trump to bring her antacid tablets; She brings some to the meeting, because en route to the meeting, she discovers some in her jacket pocket.

Now, it is just a matter of circumstance that leads to Merkel not relying on Trump to bring the antacids. It clearly doesn't follow that she distrusts Trump, due to her non-reliance. If it did, then we should suppose that, upon her discovering the antacids, her attitude becomes distrust. This is implausible. To drive this point home: it should follow that if, later, when near the meeting venue, Merkel discovers she has lost the antacids she found in her jacket pocket, she again relies on Trump. If the Commitment account is right, she therefore switches from distrusting to trusting Trump. Whether her attitude is one of trust or distrust thus hinges upon whether she has a hold of her antacids. But this is obviously implausible. We need to know more about Merkel's attitudes toward Trump to determine whether she distrusts him.

This indicates that what is missing in the Commitment account's formulation for distrust is a stipulation about the reason for non-reliance. In both the *Needed Change* and *Jacket*

Pocket cases, Merkel's non-reliance is due to factors that don't touch upon whether she regards Trump as untrustworthy. To bring this out, consider a new variant on a case:

Suspects Scheming.

NON-REL: Merkel doesn't rely on Trump to bring her antacid tablets; she brings some to the meeting. She suspects Trump has no intention of bringing any, and wants to embarrass her on the world stage, by leaving her at the mercy of her indigestion in the meeting.

Now, when we ascribe to Merkel a wariness of Trump's motives, the case is clearly one of distrust. By stipulating the reason for her bringing some tablets as relating to her feeling that she should *avoid* relying on Trump to bring some, we get a case of clear distrust. In addition, her reason for avoiding relying on Trump is not just because he is forgetful. Instead, it is because she thinks that Trump made the promise with the intent of deceiving her, to make her look weak (she knows that he likes to make other leaders look weak in comparison to him). Her wariness attitude here is quintessential distrust.

This move, however, spells the end for the view that distrust is built out of non-reliance. Her non-reliance on Trump isn't doing any of the work in attributing distrust to her. We have three variations of *Antacid Tablets* where Trump has made a commitment. All three of these cases involve non-reliance. Only one of them – *Suspects Scheming* – is a case that involves distrust. This is because we invoke a reason for the non-reliance: an attitude of wariness of the promisor, or suspecting ill motives behind the promisor's promise.

We might conclude from this that COMMIT conjoined with NON-REL is insufficient for distrust, rather than non-reliance being irrelevant. Rather, we just need to add a further condition – the *reason* for non-reliance, as wariness. This would preserve the REL+ view of distrust, and therefore the REL+ view of trust.

However, non-reliance is in fact *unnecessary* for distrust. Merkel's non-reliance on Trump in *Suspects Scheming* is a consequence of her distrust, not a component part of it. Consider how the other two cases – *Needed Change* and *Jacket Pocket* – are compatible with the possibility of Merkel having that wary, distrustful attitude. If she also had that distrustful

attitude in those cases, we would say she distrusted Trump. Her non-reliance would be due to other factors, rather than her wariness, but this wouldn't affect whether we called her distrustful. The non-reliance as such is irrelevant to whether there is *distrust* or not. It is irrelevant whether Merkel doesn't rely on Trump for a reason unrelated to her having a wary attitude. All that *is* relevant is that she has a wary attitude.

We can drive this home even further by pointing out how this distrustful attitude can even be present in cases where she *relies* on Trump to bring the antacids. Consider another case:

No Other Option:

REL: Merkel suspects Trump has no intention of bringing any antacids and just wants to embarrass her on the world stage by leaving her at the mercy of her indigestion in the meeting. Still, Merkel has no other options, having no antacids on her person, and no time to pick some up en route. As such, she relies on Trump to fulfil his promise.

We would still say that Merkel distrusts Trump in this case, even though she *relies* on him. This example is significant for several reasons. Ultimately, it illustrates why REL+ is not correct.

4.4 The argument from distrust

No Other Option illustrates that distrust is not constituted by non-reliance. Merkel distrusts Trump in this example. Yet, she relies on Trump nonetheless. This tells us that non-reliance is not necessary for distrust, because distrust can be present even with reliance. Distrust therefore just consists in an attitude of wariness about another's motives.

From this, the 'argument from distrust' can be made:

- P1. If trust is reliance plus, then distrust is non-reliance plus.
- P2. It is not the case that distrust is non-reliance plus.

C. It is not the case that trust is reliance plus.

Premise 1 follows from the core claim of REL+, that trust is reliance plus. This premise was shown to be true in section 2.3. There, it was argued that, if we commit to the view that trust is reliance plus, then we are committed to the view that distrust is non-reliance plus. Premise 2 has just been demonstrated by the preceding example cases. The conclusion, that ‘trust is reliance plus’ is false, follows by *modus ponens*.

4.5 REL+ conflates an active and a stative sense of ‘trust’

Here, I’ll set out *how* REL+ is false. I argue that it conflates two senses of ‘trust’: an activity of reliance, and a state of assurance. We’re warranted in supposing that trust also is a state of assurance, because this is the type of trust that is opposed to the state of wariness, found in distrust.

4.5.1 It is possible to actively trust and distrust at the same time

We’ve demonstrated that distrust is wariness about another’s motives, not non-reliance. This wariness attitude is necessary and sufficient to attribute distrust to a subject. That wariness can come apart from reliance and non-reliance. It can even be present with reliance on another, as seen in *No Other Option*. But *No Other Option*, because it involves reliance, is supposedly a case where Merkel *trusts* Trump. This raises the question: how is that example tenable, such that it is possible to trust and distrust another at the same time? This is generally taken to be impossible, if we recall the claim that trust and distrust are contraries but not contradictories (Govier 1992b : p.18, 1998 : p.121; Jones 1996 : p.15).

The contraries-not-contradictories claim is in line with the picture of the trust stance on REL+: it is possible to neither trust-nor-distrust another, but not possible to trust and distrust another at the same time. It might be thought that, given the contraries-not-contradictories claim is generally accepted, then by concluding with a case that is apparently one of simultaneous trust and distrust, I have provided a *reductio* of my own position.

This is not so, because the contraries-not-contradictories claim is false. *No Other Option* would only be a *reductio* of my argument, if the kind of trust and distrust at issue in *No Other Option* involved simultaneous reliance and non-reliance, thus violating the law of non-contradiction. But distrust is not non-reliance, because it is an attitude of wariness. And, we haven't shown, via the argument from distrust, that trust can't be a form of reliance. Instead, what we can conclude is that there are two senses of 'trust': *active* reliance, and *stative* assurance.

4.5.2 Activities, states and the progressive aspect

The active/stative distinction is one that cleaves along different metaphysical/ontological categories: processes and states, with activities belonging to the broader categories of processes (Mourelatos 1978 : p.422).¹¹ These are different categories of temporal entities, things that occupy time in different ways. Typically, that something is a process/activity or a state is signified in a language by verb *aspect*. Aspect is a grammatical system that encodes how something (designated by a verb) undergoes or occupies a period of time. Aspect as such encodes the way an entity occupies the *fabric* of time.¹² Activities can admit of the progressive or imperfective aspect, where typically, in English, a verb is given the suffix '-ing', e.g. 'I am running'. This is unlike states, which do not admit of the progressive or imperfective. They are complete entities, seen from outside as unchanging, e.g. 'I believe that Paris is in France'.

Crowther, in discussing Vendler's (1957) way of distinguishing between activity / accomplishment / achievement / state categories (discrete categories of verbs which can and can't admit of the progressive and perfective), draws the contrast like so:

¹¹ For further discussion of the distinctions between activities, processes, events, and states, and the relationship with verb aspect see (Crowther 2011; Parsons 1989; Steward 2012, 2013; Vendler 1957). Vendler's system for distinguishing between types of verb according to whether they are activities, accomplishments, achievements, or states, has been superseded by recognizing the role of aspect. The aspectual system instead helps us designate whether some temporal entity belongs to which of the more fundamental ontological categories of process / event / state. The relevant distinction for my purposes here is just between states and processes, with assurance being a state, like belief and knowledge, and reliance being an activity, like running.

¹² Aspect is distinct from grammatical tense, which encodes how a subject stands in relation to something (designated by a verb) in time, i.e. the *location* in time.

I might intelligibly answer the question “What are you doing?” with “I am running” or “I am walking,” but not “I am knowing that p” or “I am believing that p.” The latter verbs, [Vendler] says, are statives; they single out a standing condition or a state, a way that someone or something is or can be. Like particular material objects, states such as knowledge and belief do not have temporal duration in the sense of having temporal parts or successive temporal phases over which they unfold. Running and walking, by contrast, are things that exist by developing or unfolding over a period of time. (Crowther 2011 : p.5)

Reliance, in being an activity, falls under the temporal category of a process. Reliance is something we do, in proceeding to act (Hawley 2014 : p.4; Holton 1994 : p.74). It requires activity to be going on, unfolding over a period of time, in order for it to be true that we are relying.¹³ Like activities such as running, it admits of the progressive aspect, i.e. “I’m relying on you”/“I was relying on you”. Assurance on the other hand falls under the category of a state: it does not admit of the progressive. We can say “I am assured of you,” but not “I am assuring of you,” or “I am being assured of you”.¹⁴

When we actively trust another by relying on her, our trust can admit of the progressive. We can say “I am trusting you on this/to φ ”. The progressive corresponds to how we are reporting that we are *doing* something. On the other hand, when we want to report that we trust in a *stative* sense, we don’t use the progressive; we say “I trust you on this/to φ ”.

Crucially, we can note that distrust cannot admit of the progressive. “I am distrusting you” is ungrammatical, whereas “I distrust you” (or “I don’t trust you,” given relevant contextual factors to pull off the implicature that this means distrust as opposed to the absence of trust) is fine, just because distrust is only stative and not active, unlike trust.

Invoking this distinction illustrates how it is in a sense possible to both trust and distrust another at the same time. This is because we can trust another in an *active* sense, by relying on her, and yet distrust her while relying on her. That’s because distrust is *only* a stative attitude, not an *absence of reliance activity* (as REL+ required it to be). This is the basis on which REL+ is false: it conflates two different senses of ‘trust’: an act of reliance, and a state of assurance.

¹³ See chapter 5 for further defence of this.

¹⁴ We could on the other hand say, “You are assuring me,” when we are describing how someone else is influencing our level of stative trust.

4.5.3 Why suppose that ‘trust’ can be both reliance and assurance?

We might still be wondering what reason we have for supposing that there are two senses of trust, one referring to a form of reliance and another referring to a form of assurance, as opposed to concluding that trust is just assurance. The reason is that the active/stative distinction best accounts for the shape of the phenomena we’re discussing.

For one, we haven’t, in the argument from distrust, also given an argument that shows that trust is *only* a form of assurance. The argument from distrust was able to go through because the Commitment account of distrust was vulnerable to counterexamples in the account of *distrust* it offers. It was vulnerable to those counterexamples because there is no sense of ‘distrust’ *except* an attitude of wariness. However, the Commitment account of *trust* is not vulnerable to equivalent counterexamples. Consider the following variant of *Antacid Tablets* – where Merkel relies on Trump instead:

Antacid Tablets (reliance version):

- COMMIT_{DIR}: Merkel believes Trump has a commitment to her, to bring her some antacid tablets for the meeting.
- REL: Merkel relies on Trump to bring her some antacid tablets.

Immediately, it is intuitive that Merkel trusts Trump to bring her some tablets. This conclusion falls out of the case without having to say anything more about Merkel’s attitudes concerning Trump’s motives. This is because there is a sense of ‘trust’ that refers to a form of reliance. This is the type of trust that REL+ theorists have fixed upon, and the Commitment account provides a highly plausible explanation of when that type of reliance typically occurs; i.e. when someone has made a commitment to us.

As such, the active reliance/stative assurance trust distinction accommodates the intuitions that motivated REL+ in the first place. Yet, it also accounts for the discovery that distrust is just a type of wariness. Recall the intuitions motivating the inference that distrust was non-reliance, which were:

1. Trust is like reliance on another, but not mere reliance on another.
2. Trust and distrust involve an interpersonal import: they are both ‘betrayable’ attitudes.
3. Distrust and trust are opposed attitudes to one another; they involve opposed *anticipatory* elements.

Intuition 1 – the trust-reliance connection - is the main motivator behind REL+. The active/stative distinction respects this intuition, because it allows that there is a form of reliance we refer to as ‘trust’. In combination with intuition 2 - the trust-betrayal connection - and the observation that we can be betrayed while distrusting another, this led us to suppose that trust was reliance, plus betrayal conditions. That conclusion, in combination with intuition 3, led us to conclude that distrust was non-reliance (plus betrayal conditions). All we have to do is point out that the mistake made by REL+ theorists was assuming from intuition 1 that ‘trust’ *only* refers a form of reliance. This assumption was the baton taken up by authors following Baier in the REL+ tradition. These authors miss that, while we use ‘trust’ to refer to a kind of reliance on another, we can use ‘trust’ can refer to a kind of anticipatory mental state. What we’ve done with the argument from distrust is isolate that this anticipatory mental state is separate phenomenon from reliance trust. That is because, given that distrust is wariness, we are permitted to infer, as per intuition 3, that there is an attitude opposed to wariness of another. An obvious candidate for this is assurance of another. This assurance is the stative sense of ‘trust’, that is opposed to distrust.

Once we reflect on the trust stance in everyday reality, this inference is highly plausible. An everyday sense of ‘trust’ does refer to an attitude of *assurance* about another. Consider this variation of *Antacid Tablets (reliance version)*:

Assured Merkel:

REL: Merkel is assured that Trump has no intention of misleading her in his promise to bring her antacids. As such, she relies on Trump to bring her some antacid tablets.

What we would be tempted to say about this case, compared to *No Other Option*, is that Merkel both trusts Trump in the sense that she is assured of him, and also trusts him in her reliance on him. She trusts him in terms of her attitudes, in what she is inclined to think about his motives and his intentions. She also trusts him in terms of her actions, by proceeding to rely on him to bring her antacids. So, we have the active sense of trust, also underwritten by the stative sense of trust.

However, we shouldn't conclude from this case that stative assurance trust can only be present with active trust. It can be present without it. Just as we saw that wariness could be present with active reliance trust, so too can assurance trust be present with non-reliance. See for instance this case where Merkel's assurance is combined with her non-reliance in *Needed Change*:

Assured but Needed Change

NON-REL: Merkel is assured that Trump has no intention of misleading her in his promise to bring her antacids, however doesn't rely on him to bring her antacid tablets. She realised she needed change for the coffee machine in the meeting venue, and it was convenient to buy tablets on the way, in order to break into a bank note.

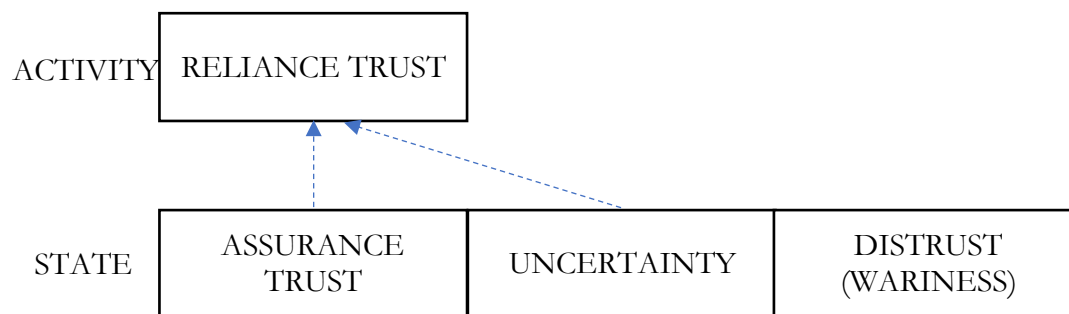
This illustrates how active trust comes apart from stative trust. It's not necessary that we actively trust when we assuredly trust too.

Finally, returning briefly to the contraries-not-contradictories claim: it's possible to both actively trust and distrust another at the same time (*No Other Option*). However, it is not possible to be both stative assurance trusting, and distrusting of another (though we could be somewhere in between these, in being uncertain). Simultaneous assurance and wariness is incoherent (try to imagine a case where Merkel is both assured of Trump yet wary of Trump). As such, the active/stative trust distinction accounts for the intuition underpinning the contraries-not-contradictories claim, while also explaining how it isn't completely correct.

4.5.4 Setting out the new picture of the trust stance: active and stative layers

The active reliance/stative assurance trust distinction also answers the challenge posed by the argument from uncertainty. That argument concluded by telling us that the trust stance is not grounded in the reliance stance. The argument from distrust reveals why: distrust isn't non-reliance, but is instead an attitude of wariness. This provides a place for uncertainty: uncertainty sits comfortably between trust as assurance and distrust as wariness, as the state of being neither assured nor wary of another; just unsure. As a result, we can set out the picture of the trust stance that is revealed by the arguments from uncertainty and distrust: the active/stative trust distinction:

Figure 6: Active and stative trust under the trust stance



The figure above is intended to convey the following:

1. The trust stance thus consists of two 'layers': something we *do* and a mental state of how we *feel* about another.
2. Active reliance trust – where X relies on Y to φ – sits in the activity layer. There is only one category in this layer, with no equivalent 'active distrust' or 'active uncertainty'.
3. The stative layer is where we find the anticipatory trust stance attitudes. These are feelings we have of others, and dispose us toward certain actions.
4. The arrows running from the state layer to the active layer convey a disposition one has, to actively trust another when under each category of attitude. Short arrows convey a strong disposition and longer arrows a weak/tenuous one.

Assurance trust involves a strong disposition to rely. When we are assured of another, we are ready and willing to actively trust her. So, there is a short arrow from assurance to reliance. The longer arrow in uncertainty conveys that we are more weakly disposed to rely when uncertain. The absence of an arrow from distrust to reliance conveys that we need some external reason to rely, whether this is through having no other option but to rely (à la the *No Other Option* example), or through having additional practical incentives to rely, e.g. of wishing to build a more trusting relationship.

This figure also illustrates the way in which, as the previous section (4.5.3) noted, the commitment account has an asymmetric susceptibility to counterexamples, in its formulations of trust and distrust. When we are posed an example case of trust that satisfies the COMMIT and REL conditions, the case will immediately be one of trust. That is because there is an active sense of ‘trust’, consisting in reliance on another. In a sense, when we look for trust in such an example case, we hit the activity layer. We’re satisfied that such a case is one of trust (of a sort). We don’t need to dig deeper through to the stative layer. For distrust on the other hand, we won’t be satisfied, just from COMMIT and NON-REL, that the case is one of distrust. That’s because there is no such thing as distrust on the ‘activity layer’. We need to dig further, into the stative layer. Once we find the subject’s wariness, our intuitions are satisfied that the case is one of distrust. And, as we’ve seen in *No Other Option*, that wariness can even be present along with the activity of trust.

4.6 Active/stative in the literature

The active/stative trust distinction has not gone completely unnoticed in the trust literature. In this section I raise some instances where the distinction has been recognised.

The most prominent and well developed account that recognises the active/stative trust is found in Hieronymi (2008). Hieronymi argues that there is both an act of *merely entrusting* oneself to another and what she refers to as a ‘purists’ notion of trust, as a trusting belief. She arrives at this conclusion through evaluating Holton’s REL+ view, and showing that it encounters difficulty in explaining a type of problem case where we cannot voluntarily decide to trust. This forces us to distinguish between trust as a state and trust as an

action.¹⁵ Her own account, distinguishing between an act of entrusting and a trusting belief, thus mirrors two of the core elements of the account I put forward. For instance, she says that “the purist understands trust as a certain kind of confidence and understands trusting actions as actions performed in or from that kind of confidence.” (Hieronymi 2008 : p.228) What I refer to as ‘assurance’ is what I presume picks out the same attitude Hieronymi identifies as ‘confidence’, and trusting actions can be understood as picking out the same phenomena as active trusting.

However, our accounts differ seemingly on the basis that Hieronymi doesn’t wish to consider the active form as a fully-fledged type of trust in its own right. Rather, she says that it is plausible to adopt the ‘purists notion’ of trust - the trusting belief - as the primary form, such that “it is plausible to [...] say that one trusts another to the degree that the one has and acts from a trusting belief, even in cases of entrusting”, [sic] (2008 : p.232) and that “the degree of one’s trust tracks the strength of one’s trusting belief,” (2008 : p.230). If this is right, then trust is in a sense *really* the stative, rather than the active form. The active form is a derivative form.

I think we should deny this view. My view is that the term ‘trust’ can just be ambiguous between an activity and a state. We shouldn’t put either of these on a pedestal, such that if we are trying to determine whether a case is *really* one of trust, we appeal to whether there is one of these forms in particular. Instead, we should recognise that both forms can be referred to depending on how we use the term ‘trust’ in different contexts and with different sentence forms. Recognising both therefore helps us better make sense of the broad phenomena of the trust stance and matters of trust, because it makes more sense of everyday language surrounding it.¹⁶

Active/stative trust is more explicitly recognized by Castelfranchi & Falcone (2010), who suggest that trust can be a “psychological attitude of X towards Y relative to some possible desirable behavior or feature” but also a “decision and the act of relying on, counting on, depending on Y” (2010 : p.18). On their view, the *decision* to trust is also

¹⁵ I discuss this type of problem case and why the active/stative distinction resolves it in section 7.1

¹⁶ In section 8.6 I show that Hieronymi’s argument in favour of construing the stative form as primary, is an unsound one.

brought under the category the activity. While this may be right, the account of the active form of trust I put forward needn't require this; the decision can be the decision to undertake the active form.

Pettit (1995) suggests a three-fold distinction between types of trust, one in which we can glimpse the active/stative distinction:

The most general usage of [trust] in this connection would equate trust with confidence that other people will treat you reasonably well: confidence that they will not waylay or cheat you, for example. We speak in this sense of trusting our fellow citizens or trusting the institutions under which we live. A somewhat less general usage would link it with confidence that other people are reliable under certain tests: they will treat you well, in the event of your placing yourself in their hands. We speak in this sense of trusting the police or trusting the courts. A third usage, more specific still, would associate trust, not with a detached confidence that people are reliable under such tests, but with putting that confidence to the test: with actually placing yourself in the hands of others. (1995 : p.204)

The first two attitudes appear to be two-place stative trust and three-place stative trust respectively. The first is general confidence, and the second is confidence concerning a specific matter.¹⁷ The third attitude Pettit describes as *active* reliance, because he describes each of these three attitudes as forms of reliance, even the attitudes of confidence. This is a mistake that stems from the fact that REL+ tends to be taken as an axiom in trust theory; one that tempts authors to fold the positive anticipatory aspect of the trust stance into reliance, for which it is ill-suited. Nonetheless, I take it that Pettit recognizes the importance of the active/stative distinction, albeit misconstruing it slightly.

Faulkner (2011 : p.23) gives a nice summary of the active/stative distinction. He distinguishes between an action of “putting oneself in a position of depending on something happening or someone doing something” and an attitude “towards this dependence”. Horsburgh (1960 : pp.343–344) is an early case of the distinction being recognized. However, he seems to suggest, similarly to Hieronymi, that trust is really reserved for an attitude of confidence, with reliance as something we do as a result of trust. Marušić (2017) has recently put forward an account that distinguishes between trust as an attitude or judgement, and reliance as an action. Again, like Hieronymi and

¹⁷ Holton similarly blurs boundaries between the active/stative and two-place/three-place distinction: see section 7.2.

Horsburgh, he takes it that trust is really the attitude of judgement, which he takes to be a belief to the effect that another will φ (2017 : pp.4–5).

4.7 Responses to the argument & stative distrust

We can deny the argument from distrust by denying that the Trump/Merkel counterexamples are legitimate problem cases for the Commitment account of distrust, and thus for REL+. Those examples purported to be problems, because they satisfied both COMMIT_{DIR} and NON-REL, yet still were not cases of distrust. So, if it could be argued that one or both of those conditions weren't satisfied, then my charge that the Commitment account is flawed as an account of distrust, fails to get off the ground. The first three responses I consider are of this sort. The last response takes issue with my construal of distrust as wariness.

4.7.1 COMMIT_{DIR} is not satisfied in the example cases

We could argue that the commitment condition is undermined in the *Antacid Tablets* case and/or one or more of its variants. We could suggest this based on the stipulation in those cases that Merkel doesn't rely on Trump to bring the antacid tablets, because she brings some herself. From this, we could argue, the COMMIT_{DIR} condition no longer holds; the thought being that if Merkel does bring her own tablets, then she wouldn't believe that Trump still has a commitment to bringing her some.

We could motivate this response by appealing to the thought that Merkel bringing her own tablets would relieve Trump, in her mind, of the commitment to bring her any. As such, once she had brought her own tablets, she wouldn't believe Trump still had a commitment to doing so. Compare with the thought that, in order for a promise to obligate a promisor to φ , it has to be in some sense *accepted* by a promisee in order to obligate (Friedrich and Southwood 2011 : p.280; Owens 2006 : p.73; Thomson 1990 : p.298). Another way of putting this is to say it must be 'taken up' by the promisee. Without this uptake, the promise is expressed, but isn't binding. We might also be tempted to construe uptake as the promisee actually proceeding to rely on the promisor to fulfil the promise/to meet her commitment. If this is plausible, then the Antacid cases fail to present a problem for the Commitment account of distrust, because they don't

present cases of distrust. They present cases of mere non-reliance; the ‘plus’ condition doesn’t hold.

There are two issues with this line of response, however. First, it is not obviously true that uptake of a promise by a promisee requires the promisee to rely on the promisor. Thomson suggests that uptake having occurred, and so the promisor being bound to act, is a matter of whether the promisee may *at some point* rely on the promisor to meet her commitment. Presumably, once the promise has been made such that the promisee would be warranted in relying on it (even if she never does), the promisor is “morally at risk” (Thomson 1990 : p.298). In other words, it isn’t necessary that Merkel relies on Trump, in order for Trump to have the commitment to bring her tablets. So, we can suppose that Merkel’s belief that Trump has the commitment, would not be extinguished by her *not* relying on him to bring the tablets.

A deeper problem with that line of response is that undermines the Commitment account of distrust. If we argue that the COMMIT_{DIR} condition no longer holds in the *Antacid Tablets* case, on the basis that NON-REL holds, then the problem that arises is that, by the lights of the Commitment account, distrust is impossible. This is because, if our reason for suggesting that COMMIT_{DIR} doesn’t hold is Merkel’s non-reliance on Trump, then *reliance* must be a necessary condition of the COMMIT_{DIR} condition holding. The cost of avoiding the counterexamples then would be undermining the account of distrust altogether.

4.7.2 NON-REL is not satisfied because risk is unnecessary for reliance

We could argue that NON-REL is not satisfied, on the basis that Merkel buying her own tablets does not suffice for *not* relying on Trump to bring her some. It only would do so on the assumption that reliance entails risk, i.e. in having the success of your goal depend on another’s acting. I will refer to this sort of risk as *practical dependence*.

Hawley suggests that reliance “needn’t imply risk or vulnerability”, because she adopts her account of reliance from Holton (Hawley 2014 : p.4). Holton suggests that to rely on something happening is to plan on the supposition that it will happen (1994 : p.65). This understanding of reliance is also adopted by Alonso (2014, 2009, 2016). On this view,

reliance is an attitude akin to that which some philosophers call acceptance (Bratman 1992; Cohen 1989), where to accept a proposition is “to have or to adopt a policy of deeming, positing, or postulating that p , i.e. of including that proposition or rule among one’s premises for deciding what to do or think in a particular context, whether or not one feels it to be true that p ,” (Cohen 1989 : p.368). This is synonymous with Holton’s ‘planning on the supposition that’ p . As such, we can call the view of reliance adopted by Hawley, Holton and Alonso the *acceptance* view of reliance.

Acceptance that p is an attitude of taking-it-for-granted that p . It involves treating certain states of affairs as fixed, within a plan. Taking it for granted that someone will act certainly seems to be a feature of relying on someone to act. When Merkel relies on Trump to bring her antacids, it seems right to say that she takes it for granted that he will bring her antacids. If acceptance suffices for reliance, this is all that Merkel’s reliance on Trump consists in.

But there is a crucial way in which acceptance can’t be sufficient for reliance. This is because we can accept things in contexts where it would be inappropriate to say we rely on those things. That’s because we accept certain things we are trying to militate against, in order to reduce the possibility they hinder our goals. For instance, Bratman sets out an example of acceptance, that involves militating against going over budget in a construction project (1992 : p.6). By proceeding on the basis that costs for individual aspects of the project will be at the top of their estimates, we ensure we don’t go over our total budget. But, we wouldn’t say we thereby *rely* on spending at the top of our estimates. So, acceptance is not sufficient for reliance.

More obvious counterexamples can be raised: I might accept that I will be mugged if I walk home too late at night, thereby ensuring I take steps to reduce the possibility of being mugged (e.g. going home earlier, or via a different route). I don’t rely on being mugged, though. Similarly, the planning department of a major city may accept that a hurricane is likely to hit at some point in future, and so take steps to build in defences for their city. But it isn’t true that the city council is relying on a hurricane hitting the city. That’s because reliance on something requires more than just acceptance. Acceptance may be necessary, but it is arguably not sufficient for reliance on something. What else is required is that that the premise we accept must be thereby what our aims come to depend

upon.¹⁸ This explains why Merkel does not rely on Trump, when her having tablets does not depend on Trump bringing them (even if she prefers to have tablets). It also explains why it is not the case that we aim to spend over budget, that I aim to get mugged, that the city aims to be hit by a hurricane.

Reliance on something is a possible consequence of acceptance. We must sometimes accept certain things which we *need* to be the case, in the face of uncertainty over whether they are, in order to bring about what we are aiming at. This is a feature of many of the standard cases of reliance, particular those in the vein of ‘relying on a rope’ to hold our weight, when ascending/descending a cliff (a typical example, found in e.g. (Alonso 2014 : pp.1–2; Black 2004 : p.272; Holton 1994 : p.68)). We accept that the rope will hold our weight, in the sense of proceeding on the premise that it will, even if we are uncertain that it will. But it is not the case that acceptance of something is sufficient for reliance on it. Acceptance of things in order to militate against them demonstrates this.

We have further good reasons to hold firm to our intuition that reliance does require practical dependence, contra the acceptance view of reliance. I will go through these in turn.

There is a quick response to the objection that NON-REL is not satisfied in the example cases. We can say that, if *Antacid Tablets* is not a case of non-reliance, then it must be a case of reliance. But if it is a case of reliance rather than non-reliance, it is not clear what then *does* count as non-reliance. Our categories of reliance and non-reliance thus seem to demand elusive standards. And again, if we try to deny that NON-REL is satisfied in what appears to be a typical case of non-reliance, this raises the threat of distrust being made impossible on the Commitment account. If a highly plausible instance of non-reliance is not an instance of non-reliance, but the opposite, we’re left wondering what could meet the standards of non-reliance, and thus, on the Commitment account, distrust.

Now, I discuss a case that Hawley offers, which is intended to illustrate how reliance needn’t imply practical dependence:

¹⁸ For authors endorsing this position, see also (Marušić 2017; Smith 2010).

...relying on someone to do something needn't mean putting your fate in their hands. You can rely upon me to bring enough food for everyone at the picnic whilst nevertheless bringing plenty of food yourself, because you don't want to seem ungenerous: you're acting on the supposition that I will bring lots of food, and indeed this partly explains the large quantities you bring along. Thus reliance in this sense needn't imply risk or vulnerability. (Hawley 2014 : p.4)

The main problem with the picnic example is that it is quite quickly described, so it isn't clear how it shows that reliance needn't entail practical dependence. I think the example can be responded to, though. Call the one who relies while bringing enough food 'Xavier', and the one being relied upon to bring enough food 'Yolanda'. We should ask in what sense Xavier *does* rely on Yolanda to bring lots of food, if Xavier brings enough food for everyone?

Imagine how the picnic case plays out, if Yolanda fails to bring enough food for everyone, despite Xavier doing so anyway. Suppose Xavier admonishes Yolanda for not bringing enough food, by saying 'I was relying on you to bring food for everyone'. If Xavier brought enough food for everyone though, Yolanda could rightly respond 'but you brought food for everyone anyway'. If it's plausible that Yolanda's response serves as a way of countering or undermining Xavier's admonishment, this suggests our concept of reliance involves practical dependence. Of course, if Yolanda had promised Xavier that she would bring food regardless, then simply pointing out that no harm was done won't be enough; Xavier could rightly admonish Yolanda for failing to uphold a promise. But then, something Xavier wanted – Yolanda upholding her promise – was dependent on Yolanda doing something: upholding her promise.

It could be that Xavier wants everyone going to the picnic, to themselves bring sufficient food for everyone at the picnic. So Xavier brings food for (Xavier, Yolanda, Zadie), and so does Yolanda, and so does Zadie. If that is the case, then the example does still involve practical dependence. That's because what Xavier wants is dependent on Yolanda and Zadie following through on certain actions.

The claim that reliance doesn't involve practical dependence is not well supported by reflecting on the literature. The typical examples in the literature, of X relying on Y to φ , without exception (besides the purported picnic case) involve cases where X's goal depends on Y's φ 'ing. Holton's drama class example involves X having the goal to fall,

but not fall to the floor, thus hurting herself. So, X relies on others to catch her, as X cannot satisfy both goals at once, unless others intervene. Baier's 'Kant's neighbours' case is one where the neighbours' timekeeping depends on Kant's walking; they risk having their plans go awry in case Kant goes for his walk at a different time.

Many authors explicitly appeal to practical dependence as a feature of reliance. Baier for instance compares her trust as reliance, with other forms of practical dependence on other's actions:

Trust which is reliance on another's good will, perhaps minimal good will, contrasts with the forms of reliance on others' reactions and attitudes which are shown by the comedian, the advertiser, the blackmailer, the kidnapper-extortioner, and the terrorist, who all *depend* on particular attitudes and reactions of others *for the success of their actions*. (Baier 1986 : pp.234–235) [emphases added]

Black proposes an account of reliance that requires a condition to be met such that, for A to rely on B φ 'ing, A must have a goal that is realised if and only if B φ 's (Black 2004 : p.271).¹⁹ Smith's account of reliance is that we rely on states of affairs, when those states of affairs are necessary conditions for the success of our plans (Smith 2010 : p.146). Railton's characterisation of reliance has it as a primitive mental process, that occurs prior to representational judgement states like belief, which is involved in goal directed activity. Among the examples he gives is of an infant, learning to walk. He suggests that, "It is perfectly natural to say that she now tacitly relies upon her sense of balance to keep her upright while she focuses on the task of propelling herself toward her goal" (Railton 2014 : p.124). This is an intuitive case of reliance; it involves an object of sorts (sense of balance) playing a necessary role in helping the infant to achieve a goal (walking). If the sense of balance falters, then the walking does too. The infant relies on her sense of balance, in virtue of being practically dependent on it.

¹⁹ Black's 'goal dependence' condition is too demanding, due to the 'if and only if' requirement. It isn't necessary that Y's φ 'ing *wholly* determines whether X achieves some goal, in order for X to rely on Y to φ . For instance, in the example that Black uses to illustrate his account, MacDonald is holding a rope for Jones, who is at the bottom of a cliff, while the tide is coming in. The action Jones is relying on MacDonald for - 'holding the rope' - is not sufficient for Jones to achieve his goal of scaling the cliff and so avoiding drowning. This is because Jones also needs to climb the rope to avoid drowning. Y's φ 'ing can be a necessary condition for the success of X's goal; it needn't be a sufficient one.

Hawley herself is among those who explain example cases of reliance by appealing to practical dependence. This includes the case of relying on your daughter's friend to pick her up from the party, where Hawley invokes practical dependence to cement the fact that you are reliant on the friend. She says:

Suppose you rely upon the friend to keep this promise: you drink several glasses of wine, making it impossible for you to safely drive and fetch your daughter yourself. (Hawley 2014 : p.11)

Hawley also appeals to practical dependence when describing how we come to rely on another to provide us lunch:

I rely on you to provide my lunch: I anticipate that you will do so, and I don't make alternative arrangements. (Hawley 2014 : p.2).

If practical dependence is not necessary to reliance, then pointing out that you drink several glasses of wine is irrelevant to whether you are reliant on the friend. It would also be irrelevant that we don't make alternative arrangements for lunch, to the fact that we rely on someone else to provide it (compare with the picnic example). But it is hard to know what the alternative explanation would be for your reliance on the friend, if not as a matter of something you want (having your daughter returned safely) be now dependent on the friend's following through on her acting. Your dependence is cemented by your actions rendering you unable to intervene yourself.

These are not the only instances where Hawley invokes practical dependence underpinning reliance. In an example of not relying on your colleagues to buy you champagne, she equates deciding not to rely on someone to φ , with taking care of the matter yourself, thus defeating your practical dependence on another:

Likewise, deciding not to rely upon someone to whom you take a participant stance—deciding to buy your own champagne—need not mean distrusting that person. (2014 : pp.7–8)

When discussing Jones' (2004 : p.6) account of trust, as "accepted vulnerability to another person's power over something one cares about", Hawley also equates *vulnerability* with reliance:

I will take it that the notions of accepted vulnerability plus foregoing the attempt to reduce such vulnerability capture roughly the notion of reliance. (2014 : p.8)

If reliance doesn't entail practical dependence, though, then why would vulnerability (equated with risk in Hawley's explanation of the picnic case) and foregoing reducing vulnerability capture reliance? It only could if risk was a necessary feature of reliance.

There is a more fundamental problem with Hawley's suggestion that reliance needn't involve risk or practical dependence. This is because the intuition motivating REL+ in the first place seems to require that it does. As such, if we deny that reliance involves risk, then REL+ fails to get off the ground.

Recall the intuition that trust is not mere reliance on others. This is on the basis that reliance allows us to be merely frustrated, whereas trust allows us to be betrayed. In either case, reliance on others is connected to frustration of goals and our responses to that, whether that is in terms of mere poorly chosen reliance, or interpersonal harm from those who let down our reliance. This presupposes risk of some form. For that connection between reliance and frustration to exist, reliance must entail practical dependence.

If it wasn't the case that reliance was connected to practical dependence, we wouldn't be warranted in supposing that the comparison between mere disappointed reliance and betrayed trust, reveals anything about the nature of trust. Trust involves risk of betrayal, where betrayal is interpersonally important disappointment. But if *reliance* needn't involve risk of a disappointed goal, then we would have no reason to suppose trust was grounded in reliance. The practical dependence involved in reliance thus seems required to motivate REL+.

Practical dependence also seems to account for how it can be an *imposition* to trust others to do certain things, and so how we incur responsibilities in being trusted. We should wish to avoid relying on those who don't invite our reliance, because we don't want to blame them in case our plans fail. The imposition of trust is a point that Hawley places much emphasis on in the Commitment account (2014 : p.7).

If we choose to give up practical dependence as a necessary feature of reliance, then reliance seems to be just a prediction or presupposition that something will happen. But again, consider how reliance can be disappointed or frustrated when what we rely on happening doesn't take place. That reliance can be disappointed presupposes that something we wanted has now fallen through, or is at risk of falling through, not just that something we predicted has not happened. A prediction can be wrong, and we can be *relieved* sometimes by predictions being proven wrong. On the other hand, my reliance can be disappointed, but it makes no sense to say my reliance can be *relieved*. It can be fulfilled, of course, but fulfilment relates to satisfaction. If it were true that reliance was simply a presupposition that something will happen, then we could say that I rely on suffering recurring bouts of migraines in the storm season, or that I rely on being mugged if I walk home alone late at night. But I don't rely on these things. I wouldn't be disappointed or frustrated in any respect if my migraines didn't happen one year, or if I got home safe. I would be relieved.

Even if Hawley is right that *reliance* doesn't entail risk, we would have good reason to come up with another concept that *did* characterise how we can be dependent on certain things for the success of our goals, or pursuing our wants. Due to our natures as practical, goal-directed agents, it's simply *useful* that we have a concept to describe how, when we aim at and pursue certain states of affairs, we presume certain conditions hold steady, and *thereby* put ourselves at practical risk from those conditions changing. We might as well suppose that reliance is that concept.

4.7.3 NON-REL is too underspecified / is an activity

Another way of responding to the example cases is to suggest that we ought to have given a narrower definition of non-reliance than the one NON-REL puts forward. The one NON-REL gives is the *absence* of reliance. A more robust analysis of non-reliance might be put forward, along the lines of an interpretation of non-reliance that Hawley gives at one point, as consisting in *avoiding* reliance:

Let's understand nonreliance as a refusal to accept vulnerability, and/or a continuing attempt to reduce such vulnerability. (Hawley 2014 : p.8)

This response is also akin to denying a feature of the active/stative trust account I've given above. It says that there *is* such a thing as distrust on the 'activity' layer of the trust stance. That's because non-reliance is a kind of activity, in being a kind of avoidance activity.

In respect of the thought that non-reliance and thus distrust can be an activity: what motivated construing *reliance trust* as an activity in its own right was the observation that we can rely, even when we are uncertain or wary of another. Reliance is not tethered to an attitude of assurance of another. It is a different category of mental phenomena; part of action, rather than an inner dispositional state, or feeling. Avoidance actions we interpret *as* avoidance on the basis that they are accompanied with, and motivated by, wariness.²⁰ As such, the 'avoidance analysis' of non-reliance collapses into imputing wariness to distrusters, and thus appealing to stative distrust.

4.7.4 Distrust is not wariness

I have argued that distrust is only obviously present in the 'Trump/Merkel example cases when Merkel has an attitude of wariness about Trump. As such, distrust involves in some sense anticipating negative motives from Trump. Hawley though argues against a construal of distrust as involving negative expectations about another's motives. She argues this when considering whether Jones' (1996) account of trust could in principle be extended to cover distrust.

For reference, Jones' (1996) account of trust is the following:

Trust is an attitude of optimism that the goodwill and competence of another will extend to cover the domain of our interaction with her, together with the expectation that the one trusted will be directly and favorably moved by the thought that we are counting on her. The attitude of optimism is to be cashed out not primarily in terms of beliefs about the other's trustworthiness, but rather - in accordance with certain contemporary accounts of the emotions - in terms of a distinctive, and affectively loaded, way of seeing the one trusted. This way of seeing the other, with its constitutive patterns of attention and tendencies of interpretation, explains the willingness of trusters to let those trusted get dangerously near the things they care about. (1996 : p.4)

²⁰ Excepting avoidance for reasons related to not wishing to *impose* on another, or give them responsibility over something.

Jones' (1996) account of trust contains the elements of the account I have set out above in section 4.5.4, with a distinction between stative assurance trust, and active reliance trust. My 'assurance trust' arguably picks out the attitude that Jones refers to as an attitude of optimism/an affectively loaded way of seeing the one trusted. Active reliance trust isn't explicitly isolated by Jones, but is alluded to, when she refers the attitude of optimism as being about how the trustee will regard our 'counting on' her. Furthermore, optimism involves a positive disposition/willingness to count on another (/let another get near to something cared about); corresponding to the thought that assurance makes us willing to rely. Where my account differs from Jones' is in how I stipulate 'counting on' is itself a type of trust, which can be present absent the attitude of optimism: we can trust in spite of our distrust. Despite this, the broad outline of Jones' 'Affective Optimism' account is one I am sympathetic to.

Hawley suggests however that Jones' Affective Optimism account falls down when we attempt to extend it in accounting for distrust. This is because it would require supposing that, in distrusting another, we hold negative expectations of another's motives, by imputing to her ill will against us. She argues first that anticipating another has ill will toward us, or will attempt to frustrate our interests, is not necessary for distrust of her. This is because someone who lies and cheats can be distrusted, yet needn't bear ill will toward others, and doesn't care about another's interests (2014 : p.6).

It is right that someone who lacks concern for others/has no sympathy for others, would not necessarily bear ill will toward others, in a particular circumstance.²¹ If exploiting another by lying or cheating happens to be the best means to obtaining what one wants, this doesn't necessarily mean that one 'has it in for' the other.

The question is whether this point puts into question the claim that distrust is wariness of another's motives. It doesn't, because wariness of another needn't involve supposing that another genuinely bears us ill will, in the sense of wishing us harm or 'having it in'

²¹ Though callousness often goes hand in hand with being vengeful/grudge-bearing. Callous people still have wants, and when others get in the way of pursuing those wants, or attempt to frustrate them, ill will can result. Selfishness unchecked by sympathy for others can lead to long term ill will.

for us. Wariness just requires that we are on guard, anticipating that another will attempt to take advantage of us, or cause us harm in some way. Distrust is the anticipation of such actions. Being wary of another is compatible with feeling that she has callous disregard for us. In fact, wariness toward people like this is an especially appropriate attitude to have. That's because lacking concern/sympathy for others will open the possibility of acting in exploitative ways against others, that would otherwise be closed off by the possession of a conscience. In the eyes of those who lack concern for others, people present as opportunities to exploit for instrumental gain.

4.7.5 Qualifying wariness

It is also worth mentioning that Hawley argues that anticipating negative motives from another is not sufficient for distrust. Considering this point will help clarify what I mean by wariness. She uses the following example:

Suppose that a deeply honourable person campaigns to have me imprisoned for my crimes. I cannot rely on this person to help me, moreover I know that she bears me ill-will and is actively trying to frustrate my goals. But my attitude to her needn't amount to distrust, for she is straightforward and honest in her campaigning. (Hawley 2014 : p.6)

We can note that the example involves adopting the mind-set of a criminal, who recognizes that she is on the wrong side of the law, yet recognizes virtues in others like honourability, which go hand in hand with trustworthiness. If the criminal regards the person who campaigns to have her imprisoned as doing so in a straightforward way, not resorting to underhand tactics, then it is right to say that that subject wouldn't distrust the campaigner. I would argue that neither, then, is the criminal *wary* of the campaigner. The criminal would no doubt be monitoring the campaigner's progress and the extent to which the campaigner, as Jones puts it, is 'dangerously near' the things the criminal cares about (e.g. her own ill-gotten gains). But the criminal's monitoring/keeping track of the campaigner is not sufficient for wariness. Wariness involves something closer to paranoia. It involves the attitude that the other is doing things *beyond* your awareness, not making her specific intentions apparent. While someone can make clear her dislike of you, if she also makes her intentions and courses of action against you open knowledge, and plays by the known rules, then you won't be *wary* of her. What she is doing won't be a matter of trust for you: she is already telling you that she is acting against you, and is doing so in

your full view. As such you can have no sense that she is manoeuvring against you *beyond your awareness*.

It is worth noting that this point about wariness, is related to why the Commitment account was so intuitive. COMMIT being satisfied transforms an interaction we have with another, into one where we are open to being misled, toyed with, let down, but also helped. By anchoring the anticipatory attitudes of the trust stance around the locus of commitment, we isolate the type of situation where another can mislead us, and manoeuvre against us beyond our awareness, exploiting our false sense of security. Playing by the rules and campaigning openly doesn't constitute hidden manoeuvring. Wariness of another involves a feeling that she can't be pinned down, that she is not as she may present herself outwardly, that her actions may be hidden and come at your expense. When you are wary of another, you are anticipating, but have yet to have confirmation of, actions at your expense. Commitments, in being related to promises, involve assurances to the effect that no such actions are forthcoming. Yet when we distrust another – are wary of her – we treat those assurances with suspicion, as deployed as a smokescreen. I'll elaborate on this picture of distrust in chapter 7 (section 7.7).

4.8 Summing up

The argument from distrust illustrates that REL+ is not true. It also illustrates *how* REL+ is not true; it takes reliance to be the only form of trust. But this is just trust in an active sense. There is also trust in a stative assurance sense. Distrust is only a stative attitude, standing in opposition to assurance, and we see this in our repeat failure to pin distrust to non-reliance.

5 Reliance

Why has the active/stative distinction not been recognized by those authors sympathetic to REL+? The most plausible reason is that, following Baier, many authors take the claim that to trust is to rely, but not merely to rely, as an axiom of trust analysis. This fails to recognise that there is a form of trust that isn't based in reliance, but in assurance. Reliance is something we do, and we can be more inclined to rely on those we assuredly trust.

In this chapter, I defend an account of reliance, as an activity. This makes reliance a distinct phenomenon from stative predictive attitudes. As such, this forms part of a defence of the 'active/stative' distinction, in my account of the trust stance.

In the first section, I raise a point about different ways we can negate certain attitudes, that illustrates why it is wrong to suppose that non-reliance could ground distrust. In the second section, I argue that acceptance cannot serve as a predictive attitude. This means that reliance, construed as a form of acceptance, doesn't ground the anticipatory attitude of trust. In the third section, I discuss Smith's (2010) account of reliance, and argue that it faces several problems. In the last section I use these problems, in combination with those noted about the acceptance view, to defend a new account of reliance, that I call the *role placement* view.

5.1 Negating reliance and negating predictive attitudes

In this section I raise a point of comparison between reliance and predictive attitudes, that sheds light on why reliance is not such an attitude, and so why it was ill-suited to ground the anticipatory element of the trust stance. This meant REL+ succumbed to the argument from uncertainty and argument from distrust.

The difference between reliance and predictive attitudes can be glimpsed in how we are able to negate the presence of positive predictive attitudes (so; belief, trust, assurance, confidence), in a way we cannot negate reliance. With the predictive attitudes, we can negate the presence of the positive using the prefix 'non-', giving us non-trust, non-belief, non-confidence and non-assurance. This is the same with reliance; we can negate the presence of reliance using 'non-', as we saw in section 2.3. The use of the prefix 'non-' simply serves to indicate the absence of the attitude.

However, reporting the absence of those positive predictive attitudes (so; non-belief, non-trust, non-assurance and non-confidence) is not yet to report the *presence* of a negative predictive attitude. We can report the presence of a negative predictive attitude using the prefixes 'dis' and 'un-'.

We can negate trust and belief with both 'dis' and 'non-'. Distrust and disbelief pick out a negative predictive attitude, whereas non-trust and non-belief pick out the absence of the positive predictive attitude. We can do the same with assurance and confidence, but with 'un-' instead of 'dis'. Here we have un-assurance and non-assurance, un-confidence and non-confidence. In the case of each of these predictive attitudes, the prefixes 'un-' or 'dis' serve to indicate that a different sort of negation is at issue from 'non-'. It's not just that the presence of the positive attitude is being negated, but that the subject has a negative prediction. In the case of disbelieving that *p*, distrusting another, being unassured or being unconfident, the subject anticipates the negative, rather than not anticipating the positive.

Reliance cannot be negated in these two different ways. We can negate the presence of reliance with 'non-', but we can't pick out any corresponding negative attitude by adding an 'un-' or 'dis' prefix. There is no attitude corresponding to unreliance or disreliance, and these terms fail to refer to any attitude that exists. If there were such a thing as disreliance, then we would understand this as something distinct from just the absence of reliance. Ayer notes that different prefixes used for negating are useful to distinguish between different kinds of 'lacking':

Prefixes like 'un-' and 'in-' count in English, in many instances, as negative signs; but to say, for example, of someone that he is ungenerous is not to negate what is expressed by saying of him that he is generous. If it is false that he is ungenerous, it does not follow that he is generous: he may be neither the one nor the other but something between the two. [...] What these and other such examples show is that, while the presence in an English sentence of what is counted as a negative sign does often have the effect of negating the statement which is expressed by the remainder of the sentence, this is not always so. To assume that negative signs served only to reverse the truth-values of the statements which were expressed by the sentences into which they were introduced would be to overlook some niceties of actual usage. (1952 : p.801).

The predictive attitudes are, like the predicate ‘generosity’ in Ayer’s example, amenable to different sorts of negation. Reliance, however, isn’t. This indicates that reliance isn’t a predictive attitude, because in the case of the predictive attitudes, we are able to take a positive (optimistic) or negative (pessimistic) stance toward some possibility. In other words, we can take a stance that some event will or will not occur, or that something is or isn’t true. And, the presence of the negative attitude is distinct from just the absence of the positive attitude. So, when we negate the presence of predictive attitudes, we need to be able to further distinguish between the absence of the positive, and the presence of the negative.

The ‘dis’ and ‘un-’ prefixes help us pick out the presence of the negative, in cases where simply reporting that we ‘don’t believe *p*’ or ‘don’t trust *her*’ would be ambiguous. In belief reports for example, when we report that “Angela doesn’t believe that Borneo is a large island of the Malay Archipelago”, this could mean that Angela has no such belief about Borneo (perhaps she is not even aware of Borneo or Malay, or what an archipelago is). Alternatively, it could mean that Angela *disbelieves* that Borneo is a large island of the Malay Archipelago.

Similarly, when we report that “Angela doesn’t trust Boris to feed her cat”, in the sense that she isn’t assured Boris will or would feed her cat, we face the same possibility of this meaning different things. It could mean that Angela lacks a trust stance attitude about this, or that she distrusts Boris in this respect (e.g. she might say of Boris “I don’t trust him when it comes to looking after my cat”, containing the implication that she has some wariness about him).

We can’t negate Angela’s reliance on Boris in the same way we can negate her trust of him. If we report that Angela is not relying on Boris to feed her cat, then we just report that Angela is *not* doing something: relying on Boris to feed her cat. Unlike the equivalent ‘non-’ negations with the predictive attitudes, this doesn’t leave open the possibility that Angela also *disrelies* or *unrelies* on Boris; that she in some sense ‘negatively relies’ on him. That’s because there is no such thing as a negative reliance attitude. It may be that “Angela is not relying on Boris to feed her cat” implies that she is no longer assured Boris is capable or responsible around her cat, when previously she was assured of this. But notice

that this is to seek an explanation for the fact of her *not* relying, not asking whether she disrelies.

This sheds light on why REL+ fails. The fact that reliance doesn't admit of 'dis-' or 'un-' type negation should indicate that it is unlike the predictive attitudes. It should come as no surprise that *dis*trust couldn't be grounded in *non*-reliance. An attitude like disreliance – a corresponding negative version of reliance, rather than the absence of it – would have been needed to ground distrust. However, there is no such attitude, and this is because reliance is not a predictive attitude.

5.2 The acceptance view and anticipation

Authors tend to treat reliance as an anticipatory/predictive attitude, or they run it together with anticipation/prediction of reliance being fulfilled. Pettit for instance suggests that trust can be a form of confidence, and that all forms of trust are forms of reliance (1995 : p.204). Hawley takes trust and distrust to be reliance and non-reliance respectively, and suggests that, “trust involves anticipation of action” (2014 : p.7) while “distrust involves an expectation of unfulfilled commitment” (2014 : p.1). This section will reinforce a point made in the previous section and in the prior chapter, that even if we endorse a common understanding of reliance on another to φ , as akin to taking it for granted that she will φ , this doesn't entail anticipating that she will φ . We can rely on her to φ while anticipating that she won't. Reliance is not an anticipatory attitude.

As I discussed in 4.7.2, the view of reliance that Holton puts forward is as an attitude akin to the acceptance of a proposition. This is in the sense of acceptance as per e.g. (Bratman 1992; Cohen 1989; Engel 1998, 1999). To accept p is to take it for granted that p is true, in some context. For example, when we rely on the rope to hold our weight, we take it for granted that it will. Thereby, we can rely on it to ascend/descend some obstacle.

I argued in 4.7.2 that acceptance is insufficient for reliance. This is on the basis that acceptance doesn't entail risk, or practical dependence on the thing relied on. What differentiates reliance from acceptance here is that reliance entails practical dependence on something. Despite this, I take it that acceptance is similar to reliance, but is not the same thing as reliance. It is similar because there is a sense in which relying on someone

to φ involves treating it as settled that she will φ , which is also a feature of acceptance. But, importantly, treating the matter of her φ 'ing as settled doesn't entail actually anticipating that she will φ . We saw this was true in the *No Other Option* case of the previous chapter. Merkel in some sense treated the matter of Trump bringing the antacids as settled, by proceeding to act on the basis that he would bring the antacids. However, she actually felt a sense of impending disappointment about this.

Still, we might be sympathetic to the acceptance view of reliance in particular, and argue that surely to *accept* that p is to anticipate that p is true. But, even if we are sympathetic to the acceptance view of reliance, acceptance that p doesn't entail anticipating p , or any positive prediction that p is true. In fact, acceptance is much like reliance in being a type of activity.

For one, it is not necessary that we have a positive predictive expectation that p is true in order for us to accept p . Examples adapted from Kaplan (1981) can help illustrate this point. Kaplan argues that there is no minimal threshold of confidence that p is true, that is necessary for us to accept p . For instance, we can, deep down, think that some project we are pursuing is highly likely to include many problems, perhaps even to fail. Nonetheless, we can press on regardless, by proceeding as though the flaws are not present. As an example, I may be anticipating that there are a fair number of points in this thesis on which I am mistaken or have defended poorly. Nonetheless, I do not proceed as though these points are mistaken or defended poorly; I treat them as settled, and treat each one as though they were true and well defended. Thus, it isn't necessary that I have a positive prediction that p is true, in order to accept that p is true.

Examples with raised stakes make this point even clearer. When I am trapped at the bottom of a cliff with the tide coming in, I must accept that the flimsy looking rope offered to me by my friend at the top of the cliff will hold my weight and proceed to climb it, even though in normal circumstances I would not bank on it doing so. Practical incentives will determine whether we should accept some proposition, because acceptance itself is a kind of doing, which we undertake for some end. It is not an attitude of positive anticipation.

That reliance is a form of activity is glimpsed in how those endorsing the acceptance view of reliance tend to make remarks that indicate it also involves activity. For example, Hawley describes reliance as “acting on the supposition that” (2014 : p.4). Holton remarks, “so I can decide to act on the supposition that you will catch me. That is to decide to rely on you.” (1994 : p.69). Other authors also offer accounts of reliance that take it to be an activity; I drew attention to these in sections 4.6 and 4.7.2.

5.3 Smith’s ‘plan-theoretic’ account

Nonetheless, two prominent accounts of reliance – Smith’s (2010) and Alonso’s (2014) – treat it as an attitude that operates within plans. On these accounts, reliance is an attitude we adopt toward propositions, within plans that we form prior to action. These are plans construed in the Bratman (1987) sense of branching structures of intentions, that we formulate when figuring out how to proceed in working towards some goal. This seems to make reliance something that is only open to relatively cognitively sophisticated agents, who can adopt representational attitudes with propositional content about states of the world, and to undertake complex planning.

However, I argue that this sort of ‘plan-theoretic’ view of reliance is mistaken, because reliance is found within and requires activity, rather than just planning to act. We only rely on things once proceed to execute plans/undertake goal-driven activity. Furthermore, goal-driven activity is something that relatively cognitively simple agents can undertake. So, they can rely. Reliance needn’t require sophisticated propositional planning, it is found in the flow of activity.

In this section I will discuss Smith’s (2010) account of reliance, in order to motivate two closely related points. One is that reliance is a product of our goal-directed activity, not a product of a prior plan. Another is that reliance is often *upon* things in the world, and not toward representational contents. I’ll argue that Smith’s account encounters several problems, which show that our concept of reliance is a product of how we pursue goals, not what our plans for pursuing them are.

I discuss Smith’s account, rather than Alonso’s. This is because Alonso’s is an acceptance view, and these have already been shown to be lacking, because they do not explain the

practical dependence involved in reliance (defended in section 4.7.2). Smith's, however, does take seriously the thought that reliance involves practical dependence. As such, it is preferable to an acceptance view.

5.3.1 Smith's account of reliance

Here I summarise Smith's account of reliance. Smith puts forward the view that we rely on states of affairs obtaining, within plans we have. The account is built from two conditions, which I have labelled [Success Depends] and [Intentions Excluded], to help sum up the intuitions behind each:

[Success Depends]: *E* obtaining is necessary for the successful realisation of *A*'s plan
[Intentions Excluded]: *A* is not intending to realise *E*
(2010 : pp.144–146)

The account then sets out what it is for *A* to rely upon the obtaining of a state of affairs *E*. [Success Depends] is intended to capture how reliance involves practical dependence. We can use one of Smith's examples to illustrate how it does this. In order to rely on the library being open late, intuition suggests that you must have a plan, the success of which depends upon the library being open late. For instance, it may be that you have just realised at this late hour that you need to return a book by today's closing to avoid a fine. As such the goal you have – avoiding the fine – is hinging on the library being open late. [Success Depends] is poised to capture this intuition.

However, [Success Depends] by itself generates too many cases of reliance. When we have some plan to achieve a goal, there are typically many actions that we intend to perform in bringing that goal about. These actions being performed, or sub-intentions being executed, will *also* qualify as states of affairs that we recognise as needing to obtain for our goal to be realised. They are the states of affairs whereby *we* perform those actions. However it is counter-intuitive that we rely on those parts of our plans that we intend to bring about ourselves, even if these are also technically necessary for the success of our plan.

Consider this point in respect of the library example. If I plan to avoid the fine for the late library book by going to the library late and returning the book, the states of affairs that are generated by [Success Depends] will also include states of affairs whereby I perform certain actions. For instance, I will have to somehow get to the library, using some mode of transport. Once there, I will have to physically deposit the book. However, it would be odd to say that, in intending to get to the library, I'm relying on the state of affairs obtaining whereby *I* use some mode of transport. It is also odd to say that I'm relying on the state of affairs obtaining whereby *I* physically deposit the book. Suggesting that I do, seems to generate a certain alienation from my own actions, because now I rely on it being the case that I do certain things. But typically, unless we have some reason to think we are disconnected from our actions (e.g. through a lapse of memory or loss of executive control over our actions), we don't rely on ourselves doing certain things. So, in isolation, [Success Depends] creates too many instances of reliance. What we rely upon seems to be a matter of those things that are not up to us; that are beyond our own actions, and that we have ceded control over.

Ceding control to what we rely on is a feature of many of the example cases of reliance in the literature, and those we have considered in discussing cases of trust. For instance, in relying on another person to perform some action, we cede control to that person to perform that action. Kant's neighbours ceded control over their timekeeping to Kant's routine habits. As such, we have reason to think that [Success Depends], by itself, will generate too many cases of reliance. There is a way in which reliance must involve control being ceded to the thing relied on, such that our goal being brought about depends on things *beyond* ourselves and/or our actions actually playing a role in bringing that goal about.

Thus, [Intentions Excluded] is intended to narrow down the states of affairs we rely upon, to only those that are matters on which we have ceded control over. [Intentions Excluded] allows us to rule out those states of affairs, that [Success Depends] otherwise encompasses, that we are intending to bring about in our plan. For instance, above we identified that alienation from my own actions appeared to result were I to say that I rely on myself e.g. depositing the book in order to avoid the fine. Depositing the book is an action I *intend* to perform, not one that I rely on happening. If we relied upon those states of affairs whereby we perform certain actions, we imply that we are not in control of

those actions. By ruling out these states of affairs, our account of reliance returns only those cases where we have ceded control over whether a state of affairs obtains. This aligns more closely with our intuitions concerning reliance. Overall then, Smith's account of reliance appears to be well motivated, and explains our intuitions.

5.3.2 Internal & External reliance

Smith draws a crucial distinction between two types of reliance (2010 : pp.144–146). The distinction lies in whether the subject is aware of whether a state of affairs satisfies the two conditions at the heart of the account: [Success Depends] and [Intentions Excluded].

One of these types of reliance is when a subject *A* *regards* [Success Depends] and [Intentions Excluded] as true, such that *A* has a propositional attitude representing a state of affairs *E*. This is when *A* *internally* relies upon *E*.²² The other type of reliance is instanced when [Success Depends] and [Intentions Excluded] just *are* true of a state of affairs. This is when *A* *externally* relies upon *E*: externally, because it is beyond *A*'s awareness of what she is relying upon.

What we have then is two types of reliance. One of these is an internal form that consists in our attitude toward a state of affairs in our plan, that we *regard* as meeting [Success Depends] and [Intentions Excluded]. Another consists in *our* standing in relation to a state of affairs; one that *actually* meets [Success Depends] and [Intentions Excluded].

Smith uses the internal/external reliance distinction to account for the possibility that we can be relying on states of affairs without realising that we are. We can use the library example again to illustrate the intuition motivating this.²³ Suppose that, in my plan to get

²² 'A internally relies on *E*' is elliptical for 'A has a reliance attitude toward a proposition *p* that represents *E* such that it satisfies [Success Depends] and [Intentions Excluded]'.
²³ Smith's own case is one where a woman named Mary, who knows nothing about aeronautical engineering, relies on the 'stationary valve plate' of a plane she plans to take from Los Angeles to New York. I avoid using this example because, based on a distinction between the concepts of relying on something and *depending* on it, that I argue for in chapter 6, this could qualify as a case of Mary depending on the plane/valve plate, rather than relying on it. That's because the valve plate working becomes of non-instrumental concern for Mary: she fundamentally needs it to work, rather than needs it just for a goal she is pursuing, as (presumably) the plane will crash

to the library, I internally rely on my car's functioning properly, because I regard the state of affairs whereby my car functions properly as satisfying [Success Depends] and [Intentions Excluded]. But it is also the case that I don't know very much about what it is for a car to function properly. In reality, for my car to function properly, the fuel injector needs to supply the engine cylinders with fuel, which is then subject to combustion. Let's call this state of affairs <Fuel-inject>. It is thus *true* of <Fuel-inject> that it is a necessary state of affairs for my plan to be realised, and also one I do not intend to realise myself. Thus, <Fuel-inject> *in fact* satisfies both [Success Depends] and [Intentions Excluded], and I am *externally* relying on the fuel injector working. This is regardless of whether I know what a fuel injector is, or what role it plays in the car's engine functioning properly. Once I have a plan to take my car to the library, I am externally relying on those conditions that satisfy [Success Depends] and [Intentions Excluded], whether I regard those conditions as such.

We can suppose that internal reliance should 'track' external reliance, if we are properly attuned to what conditions satisfy [Success Depends] and [Intentions Excluded]. Suppose that I read about the inner workings of car engines and come to learn about fuel injection. I will then not only be externally relying on <Fuel-inject>, but also internally.

The claim that we can externally rely on certain states of affairs can seem puzzling. It might be that we sometimes refer to ourselves as having been relying on certain things we didn't realise we were relying on. It can be plausible to say that I didn't realise I was relying on <Fuel-inject>, before I learned about fuel injectors. We might also refer to ourselves as having been unknowingly relying, in cases of mistaken identity. These are cases where we don't realise that we are relying on X as opposed to Y. Suppose, for example, that Matt takes it he has been relying on his employee Chris for the weekly financial reports, so has been internally relying on Chris. Instead of Chris, though, another employee – Nick – has been completing the weekly financial reports for Matt, because Chris is too lazy. I think it is plausible that there is a sense in which Matt relies on Nick completing the financial reports, rather than Chris doing so, and the internal/external reliance distinction may be helpful here.

without it working. The library example I give illustrates the same point Smith wishes to make, without this complication.

A problem with external reliance however is that it arguably makes reliance too ‘promiscuous’. In other words, we discover that there are many more things we are relying on than seems intuitive. Some of these discovered reliances may be plausible, as in the above fuel injector and mistaken identity cases. But others are much less plausible. Internal reliance ought to track external reliance, because by reflecting on our plans, we thereby come to *regard* ourselves as externally relying on things. Regarding ourselves as externally relying on things, is to internally rely on them. This has the consequence that we also internally rely on many more things than we should suppose, if we (accurately) recognise we are externally relying on them.

In the next section I set out different types of case that result from this ‘promiscuity’ problem. The objection ultimately seems to result in it being true that we are relying on an infinite number of states of affairs in our plans. This makes Smith’s account unattractive, since it is inappropriate or implausible to say that we rely on so many states of affairs in a mundane plan to e.g. do some food shopping. However, I will then argue in the next section (5.3.4) that the promiscuity objection leads to a dilemma for Smith’s account. The dilemma shows that Smith’s account is either incomplete, or that it makes reliance impossible.

5.3.3 Promiscuity: past, background, negative states of affairs

The promiscuity problem facing Smith’s account is that an innumerable amount of states of affairs in fact satisfy [Success Depends] and [Intentions Excluded], beyond what seems plausible. One type of state of affairs that does is certain *past* states of affairs.

It isn’t stipulated in Smith’s account that *E* must be a state of affairs that is yet to be realised. As such, we could suppose that past states of affairs that satisfy both [Success Depends] and [Intentions Excluded] are also what we externally (and so potentially internally) rely on. These can also be states of affairs that are both necessary for our goals, and ones that, being already settled/fixed (having happened in the past), we *cannot* intend to realise ourselves. We could argue therefore that we rely on states of affairs obtaining such as <our having been born> and <that we performed all the necessary past actions to enable this goal to be pursued>, within our plans, since these are also states of affairs that necessarily need to obtain, for particular goals to be realised. It may therefore turn

out that, in *all* our plans, we are each relying on the state of affairs <that our parents met> (hereafter: <Parents-met>), among an infinite number of other past states of affairs that had to obtain in the past, in order for our plan to be realised in future. This problem arises because, for us to figure in the goals we intend to bring about, then these long-past states of affairs, that enabled us to exist, had to obtain. They had to obtain, for us to be parts of the states of affairs we are trying to realise, in our plans.

This result, that we rely on infinite past states affairs, is counter-intuitive. It would be an odd result of our account of reliance that, given some mundane intention to e.g. do some food shopping, we would be reporting accurately were we to say “I’m relying on having been born”, as part of that food shop.

We may wish to appeal to the internal/external distinction to explain why we don’t rely on these past states of affairs. While it is true that we externally rely on them, we perhaps don’t have the cognitive capacity to pay attention to many of the states of affairs we do rely on externally. So, we don’t rely on these internally. But, the issue presented by these past states of affairs (and the further ‘promiscuous’ types discussed below) is not resolved by appealing to cognitive capacity to hold all the states of affairs we rely on, in awareness. Once we have noted that these past states of affairs satisfy [Success Depends] and [Intentions Excluded], then we should come to rely on them internally, no matter whether we forget that we do.

For example, Matt can be relying on Chris to fill in the financial reports, without paying attention to this at every instant. It can be true of Matt that he is relying on Chris, even when he puts this to the back of his mind, and focuses on other aspects of running his business. And, once Matt is made aware of the fact that he relies on <Parents-met>, in running his business, this will permeate through to all his plans. This is because internal reliance tracks external reliance. The internal/external distinction will therefore not help us rule these past states of affairs out as reliances. Once we recognise such past states of affairs as both necessary and unable-to-intend, we come to rely on them internally as well as externally. That we don’t think of ourselves as relying on such past states of affairs in our plans, though, requires some explanation, which Smith’s account doesn’t give.

It might be thought that we can resolve this promiscuity issue by just stipulating that the states of affairs we can rely on are ones that are yet to be realized. As such, we could make it so that we read [Success Depends] in a certain way. [Success Depends] as it stands is “*E* obtaining is necessary for the successful realisation of *A*’s plan”. Now, we could stipulate that for ‘*E* obtaining’, we restrict our account to states of affairs that haven’t already obtained. We therefore limit candidate states of affairs we rely on to those that may obtain in future.

The difficulty with this suggestion is that it will rule out too many states of affairs that we can otherwise rely on. We *can* rely on certain states of affairs having obtained in the past, when those are states of affairs that remain to be seen as having been realised. It can therefore be true that we are relying on states of affairs having obtained in the past, albeit those that we are both unsure as to whether they occurred, and that our plan depends upon (i.e. they satisfy [Success Depends] and [Intentions Excluded]). For example:

Hidden key

John is on the train, returning to his flat after visiting his parents in another part of the country, but he has left his keys back at his parents’ home. He forewarns his flatmate Elis that he has left his keys at his parents’ home, and sends Elis a message asking him to hide their spare key just outside the door. John needs Elis to do this before Elis goes off to work, as by the time John arrives back, Elis won’t be able to let him in until much later. Elis says he will leave the key in the agreed spot. He then tells John that he’s about to head off to work so won’t be contactable. Once John arrives back at the flat, Elis is at work, and he has to rely on Elis having left their spare in the agreed hiding place.

The Hidden Key case contains a state of affairs that we can call <Key-left>, i.e. Elis having left the key before he went off to work. This is a state of affairs that had to obtain in the past, which is necessary for the successful bringing about of John’s plan, and which John cannot intend to realise himself. It is therefore like the state of affairs ‘having been born’ (<Been-born>) in being a past state of affairs that satisfies [Success Depends] and [Intentions Excluded]. Yet, <Key-left> is unlike <Been-born>, because our intuitions strongly suggest that John relies on <Key-left>, but not <Been-born>, in his plan. So,

our suggestion that we should restrict the states of affairs at issue in [Success Depends] to states of affairs that may obtain in the future, as a way of avoiding the consequence that we rely on states of affairs like <Been-born> in all our plans, would be unhelpful: it would rule out too much. That's because it would rule out states of affairs like <Key-left>, that we can plausibly rely on in our current plans, despite their having obtained in the past.

We might be tempted to suggest, of the Hidden Key case, that <Key-left> differs from states of affairs like <Been-born>, in that John is currently relying on the lasting effect of <Key-left>. That's because John's plan to get inside the home not only implicates <Key-left> as a state of affairs he relies on, but it also implicates the persisting effects of <Key-left>. This is because John needs it to be the case that <Key-left> hasn't since been 'overwritten' by further events, such as a magpie finding the key and taking it back to its nest.

The thought here is that we can treat events like <Key-left> and <Been-born> differently, such that there is some way we can refine Smith's conditions for reliance. This is to avoid the result that <Been-born> is a state of affairs we are relying on, in all of our plans. However, pointing out that we rely on the persisting effects of <Key-left> will not help us rule out <Been-born> as a state of affairs we rely on. This is because the continuing lasting effects of <Been-born> also satisfy [Success Depends] and [Intentions Excluded]. All our plans depend upon, not just <Been-born>, but the lasting effect of <Been-born>; being alive and out in the world, able to pursue our plans. So, this suggestion will not help us narrow down, within the framework of Smith's account, what makes <Key-left>, but not <Been-born>, a ripe candidate state of affairs for reliance. Just ruling out states of affairs that obtained in the past as viable candidates, will do away with ones we can plausibly rely on, like <Key-left>.

The point just made, about lasting effects of past states of affairs being relied upon, helps to indicate a further category of states of affairs, that pose a promiscuity problem for Smith's account. I'll refer to these as 'background' states of affairs. Consider that there will be an infinite amount of fine-grained physical conditions that have to be in place, for us to be able to pursue any of our plans. This will include very fundamental states of affairs that have to persist, such as <that current physical laws of nature continue to

hold>. If physical laws were to change such that e.g. atoms could no longer form bonds and molecules, then our plans would not be able to be realised. The plans we pursue in everyday life are also dependent upon the continuing availability of food and water (necessary for the lasting effects of <Been-born> to continue being sustained), and these depend upon many background biochemical processes. These processes represent states of affairs that must be realised for our plans to be successful, and that we do not intend to realise ourselves (because in most cases we cannot). As a result, we externally rely on these states of affairs in all our plans. Then, once we recognise this fact, we internally rely on them too, in all our plans. Again, as with past states of affairs such as <Been-born>, it seems implausible that we are relying on such background states of affairs, in every one of our plans.

The kind of ‘background’ states of affairs discussed in the previous section are further multiplied, once we consider the possibility of ‘negative’ states of affairs that have to continue to obtain. For example, for any plan we have, we rely on it being the case that a disabling illness doesn’t take hold of us, at each instant of time within that plan. Within our plans, there are an infinite number of ‘not being the case that’ states of affairs that will meet [Success Depends] and [Intentions Excluded]. Yet, intuitively, these are not things that we rely on, in our plans. Smith’s account lacks an explanation of why such a narrow selection of states of affairs, that otherwise meet the conditions for reliance, are the privileged selection that strike a chord with our intuitions.

5.3.4 Luck cases: a dilemma for Smith’s account

An additional category of promiscuous reliances affects Smith’s account. This is the category of ‘lucky’ states of affairs. What do I mean by lucky states of affairs? Consider the following variant of the Hidden Key case:

Elis forgets

John takes it that Elis leaving the key for him is necessary for being able to get back into the flat. Elis does put the spare key in his back pocket on his way out of the flat, with the intention of placing it in the hiding place before he goes to work. After locking the door, however, Elis forgets to leave the key in the hiding

place. He takes it with him. But then, through a random sequence of events, the key finds its way back to the exact hiding spot just in time for John to find it. The random sequence of events is this: Elis gets halfway to work, and the key falls out of his back pocket. At that time, a magpie that nests near to Elis & John's flat happens to be looking for food, right where the key falls out of Elis' pocket; a location that is 4 miles from the flat. The magpie picks up the key to return to the nest. However, it falls out of the magpie's nest, and lands on the pavement on the street. A passing child kicks the key, and it lands in the hiding spot, ready for John to find it. John does find it, and enters his flat.

Each step in the convoluted causal history following Elis' putting the key in his back pocket, appears to count as a state of affairs that satisfies [Success Depends] and [Intentions Excluded]. Each step is, after all, *in fact* necessary for John to be able to enter his flat, and John could not intend to bring each step about himself. However, it is implausible that such 'lucky' states of affairs could be ones that John is externally relying on, within his plan to get home. We don't tend to consider random lucky events as possibilities, when determining what we are relying upon, in our plans. But then Smith's account of reliance must be wrong-footed, because it leads to implausible reliances.

Perhaps we could simply bite the bullet, and suggest that these states of affairs – Elis dropping the key, the magpie bringing it back, it falling out of the nest, the child kicking it – are all externally relied upon. They meet the requirements posed by [Success Depends] and [Intentions Excluded], after all. If we wish to bite the bullet, though, a more fundamental problem with Smith's account arises. This is the problem that, if we wish to argue that lucky states of affairs can be necessary for goal realisation, we are led to conclude that there are arguably *no* states of affairs we can rely on, because we can't take any to be *necessary* for goal realisation.

Consider why we would initially identify a particular state of affairs like Elis leaving the key for John <Key-left> as necessary for John to enter his flat. We do so because we tacitly rule out random lucky alternatives, as salient possible contingencies. In other words, we suppose that <Key-left> is necessary, because we rule out the possibility that Elis will drop the key, that the magpie will bring it back, that it will fall out of the nest, and that there will be a child to kick it back into place. Yet, if we bite the bullet on cases

like *Elis forgets*, we admit that we can be wrong to rule these random lucky alternatives out. It's possible, after all, that they could be the states of affairs that bring about our plan. But then, there is a crucial concession to random possibilities. Once we consider any one particular sequence of lucky events as a salient possibility, we open the door to considering an endless number of others. This undermines our ability to identify one such sequence of states of affairs as *necessary* for goal realisation. Thus, there no states of affairs we can rightly regard as meeting [Success Depends]. As such, we don't externally rely on *any* states of affairs. Because internal reliance ought to track external reliance, then if we are accurately reflecting on our plans, we should recognise that we rely on nothing. If we do regard ourselves as relying on anything, this is an error.

So, the luck cases present a dilemma for Smith's account. If we deny that they are reliances, then Smith's account must be implausible and incomplete. If we try to accept that they are reliances, we make a concession to random lucky possibilities. But if we do that, we are no longer warranted in calling any particular state of affairs necessary for our goal to be brought about. The possibility of any state of affairs meeting [Success Depends] is therefore undermined, and thus the possibility of external reliance, and accurate internal reliance.

This dilemma arises because Smith's account holds that reliance is ultimately a relation between an agent and a state of affairs. The state of affairs is one that is necessary for the realisation of the agent's goal (and which satisfies [Intentions Excluded]). Notice, however, that the states of affairs we identify as necessary for the agent's goal, needn't be those we identify as necessary for the success of the agent's own *actions*, which are directed toward that goal. There is a difference between the states of affairs that are necessary for the agent's end goal (that her actions are directed toward) to be successfully realised, and those that are necessary for the agent's *goal-directed activity* being fruitful. In the *Elis forgets* case, John's activity (i.e. asking Elis to leave the key) doesn't go toward achieving his goal. Rather, luck does, so the random states of affairs that occur through luck are necessary for John's goal being successfully realised. John's activities toward achieving his goal are ineffective in bringing about his goal. Yet this doesn't matter because the random lucky events of *Elis forgets* transpire.

This points to the underlying problem with Smith's account. What John relies on, on Smith's account, isn't a question of what John *does* toward his goal. At first glance, it appears to be that what John relies on, on Smith's account, is a matter of what John does. But, that is because we approach particular example cases with assumptions that certain possibilities, like those in *Elis forgets*, are not salient. We go into the *Elis forgets* case with a working schema for what events are likely to occur; what events will follow what. Then, when we present the possibility of lucky events realising a goal, regardless of what the agent pursuing that goal does, we show our inability to legislate over which states of affairs will turn out to be the ones that were necessary, for the agent's goal being realised. The floodgates are thus opened to no particular state of affairs being necessary, because we now have no principled reason to say that e.g. the child kicking the key into the hiding place was necessary, vs. e.g. a cat playing with the key and flinging it into the hiding place, or even the door to the flat unlocking itself before John arrives back, triggered by a strong gust of wind and a design flaw.

The problem lies in making John's own goal-directed activity irrelevant to what he is relying on. I take it that we are more inclined to say that John *was* relying on Elis to leave the key, as opposed to that John was relying on the lucky states of affairs. I also presume we are not mistaken about this. I also take it that John is relying on Elis to leave the key, regardless of whether luck intervenes. This can only be so if what John relies on is a matter of what John's activity toward his goal consists in. His reliance on Elis' leaving the key can go unfulfilled, yet his goal can be successfully brought about anyway. The problem of lucky states of affairs shows that Smith's account cleaves a gap between reliance and the agent's goal-directed activity.

5.4 Reliance as role placement in an activity

In this section I present my positive account of reliance, which is poised to avoid the dilemma affecting Smith's account, while also respecting the intuition that reliance involves practical dependence. It does this by holding that reliance is not an attitude we adopt toward states of affairs, within a plan. Rather, reliance is a product of our goal-directed activity.

To illustrate the account, I will use a variation of the ‘library book’ example from above, where I aimed to return my book before the library closed, to avoid a fine:

Library break-in

It is now late at night, past the time of the library closing, but I know that to avoid the fine I must place the book I have in the returns box, before morning. I bring my night-vision goggles and lock-pick, so I can break in to the library. Suppose that I regard the library having an open point of entry as necessary for me to avoid the fine. I need to deposit the book in the library’s returns box, after all. But, I also intend to ensure that the library has an open point of entry, by picking a lock.

Reliance is the result of goal-directed activity. Specifically, it involves particular sorts of mental processes and corresponding physical processes. Because it springs from the process of pursuing goals, whereby we exploit things in the environment around us for different practical purposes. As a result, it can take a different sort of *object* from a state of affairs. It can take actual objects (e.g. the lock-pick), people (e.g. an accomplice), properties of objects and people (e.g. e.g. the sturdiness of the lock-pick, your accomplice’s tenacity), ‘abstract’ objects (e.g. your sense of touch in detecting when a pin has been ‘popped’ in the lock) and conditions of the environment (e.g. the cover of darkness). The things we rely on are what we are in sensory contact with, at some level.

To help encapsulate the view: we rely based on what objects, people, features of the environment, we are in the process of harnessing, for pursuit of our ends, in goal-directed activities. We rely on things insofar as we place those things in roles, that serve to further the pursuit of our goals. What we rely on is a matter of what we are placing in a role toward achieving our goal. We rely on *things*, to perform *roles*, toward *goals*. Reliance therefore has a structure: agent-object-role-goal.

‘Harnessing’ and ‘placing’ are not meant literally. By harnessing I don’t mean attaching a physical harness to an object, and by placing I don’t mean physically putting an object in a particular spatial location (though this is sometimes done in order to place in the figurative sense at issue). What I mean is that, by acting in a way such that our own activity

toward a goal is informed by the causal inputs of certain things, such that those things are primed to contribute to bringing about our goals, then we rely on those things. In other words, we make *use* of objects of reliance. The sense of ‘using’ at issue encompasses using from a distance, indirectly, not (necessarily) through directly physical manipulation.

To help explain what I mean: recall how Kant’s neighbours relied on Kant for their timekeeping. When they rely on Kant, they harness Kant’s walking. Kant’s walking is a feature of their environment that is already going on, available to be made use of. To do this - to harness it - all Kant’s neighbours need to do is time monitor for/check for Kant’s walking as a trigger for when to undertake their own actions. So, by acting in some sense ‘around’ Kant’s walking, Kant’s walk informs their own routines. It takes on the role of regulating their own activities, in serving their goal of having their own activities be properly co-ordinated. By acting in a certain way, they place a regular feature of their environment— Kant’s walking – in a role, within their goal-directed activity.

Suppose that we rely on our alarm clock to wake us up in the morning. To rely on it, we place it in the role of waking us up, refraining from implementing other measures to wake us up, when proceeding to go to bed. Our goal is to wake up with enough time to get ready in the morning, so we are not late for work.

So, when we rely on something, we create a role – a gap in the progression toward our goal – that an object of reliance, through its properties, is placed to fulfil. In a sense then, harnessing/placing means making use of the objects of reliance, by acting around them.

We can use the library break-in case as an additional source of examples. Suppose that I am now outside of the library, equipped with my lock-picking kit. In this case, I am in the process of looking for a lock I can pick open. This is a complex process, involving physical and mental activities. I scan my environment for salient candidate locks and make inferences about likely locations of less secure locks. What I am relying on, are all those things I am harnessing, that I have placed in a role, in respect of how I am going about pursuing my current goal (depositing the book). Things I am harnessing include my sensory faculties, my night-vision goggles, and my lock-pick. These are primed and ready to be activated in the roles given to them, when required. I also harness the shadows outside of the library, to move about unseen. I harness these things by undertaking

coordinated, complex movements, that are ‘shaped’ (informed by, directed by) their input. I am *sensitive* to the placement of the things I am harnessing, in respect of there being a sensory link between my activity and these things. Placing an object in a role, means acting in a coordinated way, so as to be able to make use of the effects that object is able to have on the environment. This makes it so the object is ready to causally contribute to your goal being realised.

This view of reliance, as an activity of role placement, allows us to articulate what sort of phenomenon reliance is, in a way that respects our everyday reliance talk. We can see how we can talk about e.g. relying on our bank cards to purchase goods, that we are relying on our friend to meet us for lunch, that we are relying on our satnav to get us home. Our bank cards are, by how we exploit them, placed in a role of allowing us to access goods and services in many domains. Our broader goal-directed activity is to acquire things we need. Our friend has a role in ensuring we eat lunch together (something we want to do). And, our satnav has the role of giving us instructions on which direction to take when driving, with the goal being to arrive at our destination. Even though, in each of these cases, we must undertake coordinated goal-directed activity so that these things are harnessed in some way, we don’t have to overly ‘mentalise’, what we do in relying on those things. Reliance isn’t purely a product of what is in the head, because it is more fundamentally a product of how we harness things in goal-directed activity.

The role placement view doesn’t require us to talk in terms of reliance being an attitude held toward a proposition, representing a state of affairs. We don’t have to therefore suppose that, when I am relying on the night-vision goggles I use during my break-in, I am, throughout the break-in, holding a representational attitude toward *it being the case that* the night vision goggles will allow me to see. To rely on the night-vision goggles, I don’t need to hold conceptual representations of the night-vision goggles and their relationship with my seeing. If I were to be lost in the flow of activity such that I forget entirely that I am wearing night-vision goggles, this wouldn’t defeat the fact that I am relying on them at that point. Instead, the fact that I am relying on the night-vision goggles simply falls out of the *way* I am pursuing my goal; that I am harnessing the night-vision goggles, proceeding in such a way that they are influencing my activity, which is built around and shaped by their presence.

As Railton (2014) suggests, reliance can be a ‘low-level’, pre-conceptual mental process, not at the level of conscious awareness, that can be manifested in relatively cognitively unsophisticated agents. The role placement view corresponds with this. All that is required for an agent to be able to rely, is that the agent is in sensory contact with things in its environment, and able to undertake goal-directed activity, geared toward harnessing those things. Railton gives examples of primitive reliance on sensory faculties; an infant relying on her sense of balance to walk, and a hamster relying on his sense of touch when navigating a passageway in the dark (2014 : p.124). In these cases, it seems natural to say that the hamster and infant are relying on these sensory faculties, without holding conceptual representations of those faculties. What is true however is that the infant, by acting in a way that is informed and directed by the sensory feedback from her sense of balance, accords her sense of balance a role in pursuing her activity of walking. The same is true of the hamster.

I suggest that the features of these primitive cases, that tempt us to attribute the concept ‘reliance’ to them, are features replicated in more cognitively sophisticated forms of relying. As such there is a common structure underpinning reliance, as a form of activity, regardless of how much that activity is informed by high-level conceptual representations. For one, the role placement of the hamster’s sense of touch mimics the role placement of the lock-pick, as a tool I harness in probing the library’s entrances, and the role placement of the shadows that I move within, to avoid being spotted.

An aspect of the sensory link between our activity and the things we rely on, is being disposed to ‘look to’, ‘return to’, or ‘appeal to’ what we are harnessing, just in case our activity stalls or begins to veer off target. This disposition reflects having placed things in practical roles. By ‘looking to’ or ‘returning to’ I mean turning our attention toward the thing relied on, when our goal-directed activity stalls or veers off target. It can also involve ongoing mental processes of *monitoring* or checking that the thing relied on remains placed in its role. When it isn’t clear why some activity veers off target, we can be disposed to check whether this is due to something failing to perform a role. Should that be the case, we may attempt to re-engage the relied-on object in performing the role it was placed in.

Suppose, for instance, that I am looking for a lock to pick, making use of my night-vision goggles. Should the night-vision goggles glitch, so that I can’t see my way around in the

dark, my activity will be disrupted. I will then direct my attention toward the resulting lack of input from an object (the night vision goggles) that was playing a role (enhancing vision) in enabling me to pursue my goal. I may attempt to get the night-vision goggles to work again. That is, if I wish to continue pursuing my goal in the same way I was before, by continuing to give that pair of night-vision goggles a role. I may try to press on without relying on the night-vision goggles, allowing my eyes to adjust to the darkness. My non-reliance on the night-vision goggles will be a matter of my activity no longer affording the night-vision goggles a role, because I am pursuing my goal without using them.

It's possible that we attempt to re-engage the relied-on object by trying to continue in the same way as before. We might do this when our car starts to falter; we try to carry on driving toward our destination. Similarly, when Railton's hamster is relying on its sense of touch to navigate a maze in the dark, the hamster is paying attention to the sensory feedback from its feet and whiskers. This feedback informs the hamster's ability to make progress. Should it lose sensation from its whiskers, the hamster will be much less capable, probably at a loss. It will likely try to proceed in the same way, but it will be doing so through a habitual routine of actions, that are normally informed by anticipated input from its whiskers. Because of the loss of this input, its actions will be thrown off target and misjudged. After a time of attempting to proceed in a way that is geared around the input from whiskers, and being continually thrown off, the hamster will learn to attempt to disengage the input from its whiskers, and may try to harness other means of navigating. To no longer be relying on its whiskers, the hamster will have to adjust how it proceeds in activity, so that a practical role for input from its whiskers is no longer a required part of that activity.²⁴

²⁴ I've said previously that reliance isn't an anticipatory attitude. Yet I have mentioned that reliance involves anticipating inputs from objects. But there is an important distinction at issue, in terms of what is anticipated. To rely on X to φ , we have to anticipate the *possibility* of X φ 'ing. Minimal levels of anticipation make reliance possible. To rely on X to φ , we have to anticipate the possibility that X can perform the role of φ 'ing. But the level of anticipation required can be minimal: we don't need to anticipate that X *will* perform that role, just that it *could*. As our pessimism about X's being able to φ increases, we will begin to increasingly monitor it, and will likely remove it from the role. Once we anticipate that X positively cannot perform the role, we are unable to rely on it.

This disposition to ‘return to’ or appeal to the object of reliance is also present when we are relying on people to do things. Suppose that Justina is relying on her colleague Cam to warn any of Justina’s clients that appear at her office today, that she is unable to hold any meetings today, as she has too heavy a workload. When Justina’s clients keep appearing at her office, she will look to Cam and attempt to re-engage her in that role of warning clients of her unavailability (by, e.g., asking her whether she has been doing this, and if not, to ask her please remember to do it). That’s because Justina has placed Cam in the role of warning clients, so that she can meet her workload. When this activity begins to be disrupted by clients, she looks to Cam.

The object-role-goal structure of reliance also explains how we can avoid relying solely on particular objects. We can place objects in such a way that the role being fulfilled doesn’t rest on a particular object performing that role. That’s because we can implement fail-safes, that are placed to fulfil a role, should one object fail to perform the role. For example, suppose that when you are writing a long document, you need to be able to save your work. As such, you rely on something to store your work. Now, suppose you have a system that creates two backup copies of your writing, on two different storage devices, on a rolling basis. When you save a copy of your current document to your own hard-drive, it also simultaneously uploads to a cloud storage drive. As a result, there isn’t one particular drive that is placed to fulfil the role of saving your work. We therefore wouldn’t say that you’re relying solely on the physical drive to save your work, or the cloud drive. You do rely on the overall system to save your work, but not a particular drive within that system. That’s because there are two devices placed in that role.

So it can be with relying on people. Suppose Justina, in addition to telling Cam that she cannot see any clients today, also tells Ruth. As such, Justina has placed both Cam and Ruth in the role of warning her clients of her unavailability. She doesn’t solely rely on either Cam or Ruth to perform that role, but she relies on at least one of them to perform that role.

5.4.1 States of affairs on the role placement view

The role placement view also allows that we rely on states of affairs à la

the propositional reliance of acceptance and Smith's internal/external reliance. I suggest that what we can do, in the flow of activity, is rely on 'potentials'; possibilities of enabling conditions for our goals, being made actual/realised.²⁵

We can rely on *there being* certain conditions of the environment (e.g. there being a lock to pick), or some particular object *being* some way (e.g. the last lock I find being amenable to picking). This can be because, when we anticipate the *possibility* of something being so, that would enable us to progress toward our goal, what we can do is place that potential in a practical role. We do that by acting in such a way that, *if* that condition is satisfied, then our goal will be furthered. This still means that reliance is a product of our activity, rather than as a proposition or state of affairs within a plan. The activity is just shaped around the potential for that condition of the environment being so.

For instance, when I am in the process of looking for a locked window or door, I am seeking out a potential opening; one that I can make into an actual opening, using my lock-pick. What I am doing is looking for a feature of the local environment – a locked window or door that can be opened – to play a role in bringing about my goal. As such, I am relying on *there being* a locked window or door that can be opened, because this potential is what I have placed in a practical role, based on how I am pursuing my goal of getting in. This sort of reliance is picked out when we say that we rely on 'there being' such-and-such, or '<an object> being' such-and-such. Reliance on potentials also follows the same shape as the role placement/harnessing model. These potentials are what we anticipate as, at some minimal level, *possibly* holding. We needn't be assured to any particular extent that these potentials/enabling conditions are in place, in order to rely on them being in place. We could also anticipate them *not* holding, yet still rely on them being in place. That's because, to rely on them, we just need to be acting in a way that means we place an enabling condition in a practical role, that will need to be activated at some point, if our goal is to be carried off.

We have an enabling condition in the library break-in case: I am relying on *there being* a lock I can pick. That's because, based on how I am proceeding toward my goal, in looking for locks to pick, I am looking for a lock to fulfil the practical role of a lock I can pick.

²⁵ As in the *Hidden Key* case from before, these can be conditions that remain to be seen as holding, even if they are the product of past events.

That role is there, just as a result how I am currently seeking to gain entry. This doesn't mean I am relying on a particular lock being able to be opened. That's because, in the process of seeking one, I am relying on the potential for there being one. Once I have tried all the locks but one, I am relying on that final lock being the one I can open.

The possibility of reliance on potentials requires that I refine my stipulation that we are in sensory contact with those things we rely on, because it can be that we aren't *yet* in sensory contact with what we are relying on. It can be that we haven't yet had the thing we are relying on, actually make a causal contribution. Yet we can infer the future possibility of that potential being brought about. What reliance on potentials amounts to is proceeding in activity in such a way that a gap is created – a role to be fulfilled – that we judge the potential will fulfil, once it is actualised. This involves minimal level of anticipation of the potential: it involves anticipation of the possibility of that potential, being actual. When we are judging things accurately, our goal-directed activity will be pulled off, because that potential, when realised, will be able to causally contribute to our goal. That's due to our having created a role for it. We can act in a way that creates the possibility of future sensory contact with that potential we are relying on, without yet having sensory contact with that potential.

5.4.2 Accounting for intuitions & problems in other accounts

The role placement view accounts for the intuition underpinning the acceptance view, that relying involves 'taking it for granted' that some object will do something, or that some state of affairs will be the case. On the role placement view, though, 'taking it for granted' or 'accepting' falls out of a mode of goal-directed activity, rather than requiring a subject to have a mental attitude of assuming a state of affairs obtains. When we rely on a rope to ascend a cliff, for instance, we proceed in a mode of acting, such that the rope plays a practical role in our pursuit of our goal to ascend the cliff. Our mode of acting involves using the rope, to ascend the cliff. Taking it for granted that the rope will perform this role, in this sense, doesn't have to be particularly cognitively elaborate. It isn't a matter of our representing a state of affairs, in a propositional attitude (i.e. 'relying upon it being the case that the rope holds our weight throughout the ascent'), that we assume holds, before proceeding to climb. Taking it for granted, instead, falls out of how we act. We can climb the rope in a panic. Once we reach the top, step back and reflect

on what role the rope was playing, by how we'd acted, we can describe what we were doing in terms of having taken it for granted that the rope will allow us to ascend the cliff. Just climbing the rope can occur prior to this propositional, reflective, plan-theoretic 'taking it for granted' attitude. And, climbing the rope, no matter how hesitant we are, is how we rely on it.

The role placement view also respects the intuition that reliance involves practical dependence (defended in 4.7.2). Reliance involves placing objects in roles, so that they are primed to contribute to furthering some goal. When we rely on things, we rely on them actually performing those roles. Our goal-directed activity being fruitful, then, depends on those things performing those roles. This point also accounts for how 'taking it for granted' isn't *enough* for reliance. For example, the city planner takes it for granted that the town will at some point be hit by a hurricane. Taking this for granted doesn't entail reliance on being hit by a hurricane. That's because reliance requires harnessing things, possibilities, in the pursuit of our ends, so that they contribute to furthering our goals. The city planner doesn't harness the hurricane for some end; she takes steps to mitigate the impact of the possible hurricane on the city.

The role placement view may also help explain why external reliance, in Smith's account, could be attractive, when considering certain types of case. External reliance was invoked to make sense of cases where we e.g. might say we are relying on the car's fuel injector, as part of relying on the car to function properly. By relying on the car to work, so performing the role of getting us from A to B, the fuel-injector is a component part of this working. The fuel injector itself has a role in the functioning of the engine, and thus the car. The role placement view sheds light on why we may have had this intuition. By placing the car in that practical role, the fuel injector is in a sense also part of that role, by being a part of the object playing that role.

However, the role placement view also explains the intuition that there is something not quite right in saying that a person who has no knowledge of the car's inner workings is relying on the fuel injector. This is because the person who has no knowledge of the car's inner workings is not placing the *fuel injector* per se in a practical role, even if she places the car in a practical role. She doesn't place the fuel injector in a role because her goal-directed activity is not sensitive to, informed by, determined by, the fuel injector. Should

the car break down, she will be at a loss as to what to do next. She will not return to the fuel injector or attempt to engage with it, because she won't know what it is. Role placement requires a sensory link to those things we place in roles, and an aspect of that is the disposition to look to whether those things have retained their placement, should our goal-directed activity hit a complication.

External reliance was also supposed to explain how we can be relying on particular objects in mistaken identity cases. One way the role placement view can shed light on such cases is as follows. It may be that, for example, Matt takes it that the employee he is relying on for the financial reports, is called Chris. But in fact, this employee is called Nick. So, even though Matt may attach the name 'Chris' to the employee he looks to for the financial reports, the actual name of this employee is not as relevant as to Matt's reliance, as to which person he has actually placed in the role of delivering the financial reports (where role placement in this case also takes on a literal sense, within the workplace context). That's because, essential to reliance, is the practical activity structure of harnessed object-role-goal. What object has been placed in the role is more important than what term the agent uses to refer to that object. We're warranted in saying that Matt relies on Nick, rather than Chris, because Matt *looks to* Nick for those reports; Chris is a different employee altogether.²⁶ Matt is just mistaken about what the object he relies on is called. This mistake may of course lead, in the long run, to difficulties in keeping the actual employee he relies on – Nick – effectively placed in that role. That's because the name Matt uses to refer to that employee will potentially lead to confusion from others, about who needs to do the financial reports. Chris may start filling in the financial reports, hearing that Matt is counting on him, and Nick may presume he's been removed from that task, but not been told directly, so step back.

The role placement view also avoids the complications of Smith's view, that reliance seems to be extremely promiscuous. In terms of past states of affairs like 'having been born' (<Been-born>), the role placement view can account for why this is not a condition we are relying upon for the pursuit of a goal. That's because we don't, through our activity, place the potential for <Been-born> being true in a practical role in bringing

²⁶ If Matt says "I rely on Chris for my financial reports", my intuition is that he has said something false, because he actually relies on Nick. In Matt's statement, 'Chris' still serves to pick out Chris, rather than Nick.

about our goals. Rather, the possibility of our having goals in the first place is premised on our having been born. <Been-born> (and its continuing effects) are conditions that enable the possibility of the sort of goal-directed activity that reliance springs from. Furthermore, the role placement view is also able to explain why we can rely on certain past states of affairs, like those in cases like *Hidden key*. In that case, John relies on Elis to have left the key to their flat in the agreed hiding place. These are cases where we create a role in our current goal-directed activity, for a past state of affairs being realised. Elis leaving the key has a role to play, based on how John proceeds toward his aim of getting back into the flat.

The role placement view also accounts for our intuitions that suggesting background states of affairs, like ‘photosynthesis continuing to occur’, and ‘the Earth continuing to revolve around the Sun’, are not obviously relied upon. These potentials are not typically given a practical role – an opening to fulfil – in our activity. Instead, they underpin the possibility of our activity as a whole, as a background condition of all of it. Unless we harness photosynthesis or the Earth’s continual travel around the Sun for some particular goal we are pursuing, then there is something not quite about saying we rely on these things.

‘Negative’ states of affairs, such as ‘not being struck by a debilitating illness’, can be potentials we are relying on. That is so, if we are at some level disposed to check or monitor for the these negative potentials being overridden, such that the condition we wish to avoid comes about, in some activity we are pursuing. Someone with a fear of being unwell in public may, in leaving the house to go food shopping, be monitoring and checking for indications that she is about to be unwell, throughout her shop. In doing this, she is checking that what she is relying on – not being ill, or being well – remains fulfilling the role it has. Similarly, in the *Library break-in* case, I am relying on the lock-pick not breaking, if I’m disposed to monitor that this condition remains fulfilled.

The dilemma presented by ‘lucky’ states of affairs for Smith’s account doesn’t threaten the role placement view. That dilemma resulted because Smith’s account allowed that reliance was not a matter of what activity the agent pursues toward her plan. The role placement view avoids this by ensuring that reliance is grounded in our goal-directed activity. It is a matter of what things we harness, by giving them practical roles in pursuit

of our goals. As such, whether lucky circumstances befall us in carrying out our plans is irrelevant to what we were relying on, even if those lucky circumstances turn out to be necessary for our goal being realised.

5.4.3 Progressive and habitual reliance

The role placement view, in holding that reliance is borne from activity, also helps explain an important feature of our reliance talk. What it explains is how we can refer to reliance in both the *progressive* and *habitual* aspect. Recall the discussion of verb aspect from section 4.5.2. Reliance is an activity. This makes it unlike assurance, which is a state. Because it is an activity, we report our reliance using the progressive aspect. This is also how we report activities that are unfolding. Just as we say that we are running, that we are writing a book, we can say we are relying on our satnav to guide us to the destination. Here, the progressive verb form indicates that we are undertaking a kind of activity.

Of course, we don't always report that we are *relying*. We sometimes say we *rely*. This may indicate that reliance can also be a state. However, because reliance is a form of activity, this verb form doesn't serve to indicate a mental state. Rather, it indicates a habit: it is taking the *habitual* aspect. Just as we can say we run, or that we write books, we can say we rely on our satnav to guide us. By using the stative form, we report a feature of our broader activities over time; a kind of habit or regularity. When we say 'I rely on my satnav', we use the habitual aspect to report that the satnav regularly takes on a role, as part of our broader activity over time.

That said, 'I rely on my satnav' can also be read using the simple present tense. This would report that I am doing something in the moment, but conveying a sense of instantaneous time, rather than a sense of time unfolding (as though I am reporting on my activity more detachedly; 'outside' of time). To distinguish the habitual from the simple present, we can bring out the habitual aspect using 'regularity' indicating adverbs such as 'often', 'sometimes'. For example, 'I sometimes rely on my satnav to guide me' would convey the habitual. That said, when we use the simple present, in 'I rely on my satnav', this still reports an activity. In addition, our ability to report reliance using both progressive and habitual aspects should reflect that reliance is a form of activity.

Smith's account construes reliance as an attitude found in a plan, as do acceptance views. These views of reliance are thus supposed to construe reliance as a mental state, because plans are construed as mental states. It might be argued that these views still explain how the progressive and habitual is used in our reliance talk, without invoking the claim that reliance is an activity, rather than a mental state. That's because we can also report our intentions and plans using the progressive aspect. We can say, for example, that we are intending to go to the doctors tomorrow, or planning on visiting our friend soon. Because such expressions are commonplace, the argument can be made that, because we can report intentions and plans using the progressive, and these are mental states, then the progressive being used to report reliance needn't indicate that an activity is going on. Thus, it seems that use of the progressive is compatible with reliance being a state, rather than an activity.

However, we can argue that, when we are reporting truthfully that we are intending to read that book, or planning on visiting that friend, this does serve to indicate that we are engaged in form of activity. By using the progressive aspect to describe our intention, we report we are now in process of pursuing our goal, in regulating our activities in line with future achievement of that goal. This needn't mean we are *actively* pursuing the goal, by undertaking actions directly toward that goal. We may be passively pursuing it, perhaps still figuring out the steps to take. For instance, I needn't be in the process of driving to the doctors, or booking the train tickets to see my friend. We can be passively pursuing a goal, but still be pursuing it. We might be waiting for certain events to unfold. Waiting typically involves mental activities of monitoring for passing of time, looking for particular 'windows of opportunity' to open, in which particular actions can then have an effect in bringing us closer to our goal. When I am intending to go to the doctors tomorrow, active pursual of the goal will involve taking certain steps, like setting an alarm now to wake me up in time for the opening of the surgery, so I can book an appointment for the same day. I will also keep track of my other activities so that they are compatible with me waking up in time, e.g. not drinking too much alcohol tonight, which would make me more likely to sleep through or ignore the alarm. This is, at least, if I am reporting accurately that I am intending to go to the doctors tomorrow.

In effect, then, that we can report our intentions and plans using the progressive doesn't serve to undermine my claim that reliance is a form of activity. That's because we have reason to suppose that, when we use the progressive to report an intention or plan, we are reporting that we are in the process of pursuing some goal, in the sense of taking steps toward it, and refraining from activity that would be incompatible with it.

6 Dependence

In this brief chapter, I draw a distinction between the concepts of reliance on things and dependence on things, when applied to agents as subjects. This distinction will become more important in chapter 8, as I outline what the interpersonal import of the trust stance consists in. The interpersonal import of the trust stance, I argue, stems from interpersonal dependence, not just interpersonal reliance.

My proposal in this chapter summarised: The kind of dependence at issue in reliance is due to an agent's choice to pursue a goal in some way. Relied-on things are of instrumental need, so of instrumental concern to the agent. A more basic sense of dependence applies to things the agent needs non-instrumentally. So, having those things is of non-instrumental concern for the agent, rather than of concern because they play a practical role in the activity she is undertaking.

I thus suggest that there is a distinction between the sense of dependence that is at issue in reliance, and being dependent on things more generally. While reliance involves practical dependence on something, that is by your own activity. It is dependence by virtue of what you are trying to do, and how you go about doing it. Being *dependent* upon something, however, describes a more pressing relation. It isn't by virtue of some particular activity that you depend on that object performing some role. Instead of your activity being dependent on that object, it's the case that *you* are dependent on it. Dependence on an object involves a much deeper concern for procuring that object and having it accessible. That object's presence is not just of instrumental concern, but of a deeper constitutive concern – a non-instrumental concern – based on the kind of thing you are. Again, this distinction will be used in the account of interpersonal import of the trust stance, presented in chapter 8.

6.1 Rely / depend

Many authors use the concepts of reliance and dependence interchangeably. In the trust literature, Baier (1986) for example, suggests that trust, “seems to be reliance on [another's] good will toward one,” (1986 : p.234), and later says that, “when I trust another, I depend on her good will toward me.” (1986 : p.235). Baier's phrasing suggests

that reliance on another's goodwill is not a distinct phenomenon from dependence on another's goodwill. Mäkelä and Townley (2013 : p.viii) discuss trust in terms of both reliance and dependence, using these terms interchangeably. Tummolini et al. (2012 : p.590) describe reliance on another in terms of being dependent on another's action. McCraw is explicit in taking the concepts to be equivalent, when he says "I see no real substantial difference here between relying on someone and depending on someone. Thus, I shall take reliance and dependence to be equivalent." (McCraw 2015 : p.4).

This seems permissible. When we reflect on some cases of reliance, it can seem plausible to also describe these in terms of dependence. Consider the primitive reliance cases that Railton (2014) discusses, where a hamster relies on its sense of touch to navigate a maze, and where an infant relies on its sense of balance to walk (Railton 2014 : p.124). We could possibly say that the hamster depends on its sense of touch, and the infant depends on the sense of balance, without drastically altering what is meant. Similarly, when we say that nowadays most rely on their smartphones, we could also say that most depend on their smartphones.

In this chapter I defend a conceptual distinction between reliance and dependence. However, I don't intend to offer a set of conditions that stipulate whether it is correct or not to use 'rely' or 'depend' in a particular context. My proposal is instead to outline a distinction between a core sense of reliance and a core sense of dependence, when applied to agents. It can be that in everyday language we often use each term interchangeably to pick out these concepts. What is important is that there is a difference between core usages of the two concepts. I also think everyday language often does reflect this difference, with differing connotations associated with use of either term. Reliance is a concept that operates in the domain of our goal-directed activities. Such activities are self-determined, and we rely on things when we harness them in pursuit of doing things. Dependence, however, is a concept that operates in the domain of our core functioning, regardless of what activity or goal we choose to pursue. What we depend upon tends to *determine* what goals we pursue, while reliance is the result of how we *choose* to pursue those goals.

6.2 Successful activities / proper functioning

I've argued that reliance has a certain structure: agent-object-role-goal. When we rely on some object, that object plays a practical role for us, in some goal-directed activity. When we depend on some object, however, the 'role-goal' part of the structure is 'internalised'; brought within the constitution or functioning of the agent itself. As such, the 'role' part is no longer a matter of contributing to a goal-directed activity; we don't talk about what we depend on in relation to a goal. That's because the thing we depend on is not needed in relation to a particular form of activity we have decided to pursue. Instead, it is needed for our functioning as the kind of thing that we are.

Dependence, therefore, is closer to home than reliance. I take it that this is reflected in the weightier relationship evoked in saying that X depends on Y, compared to saying that X relies on Y. Consider the connotations in saying X depends on alcohol, versus the connotations in saying X relies on alcohol. Alcohol dependence is a condition that affects our functioning. Dependence on alcohol implies that an agent structures her activities, and ultimately her life, around the procurement of alcohol. Alcohol dependence gives rise to felt needs, due to internal biological and psychological reward structures being affected. That gives rise to an impulse for alcohol drinking, driving the felt need for alcohol. This impulse will then determine what sorts of goals we pursue, and how we pursue them. We will structure our activities around the presence of alcohol, and avoid activities where alcohol won't be present. Activities that don't involve alcohol may not even be salient to us, due to how central the continued presence of alcohol is to our daily functioning.

On the other hand, we needn't imply this kind of relationship with alcohol in saying we are relying on alcohol. For instance, we can be relying on alcohol in the progressive sense, so for a current goal. This would mean that we are placing alcohol in a practical role in our current activity. Perhaps we are at a social gathering and need something to help allay our anxiety. In that case, we could place alcohol in that role, as a way of getting us through that gathering. Now, alcohol is helping us to achieve our goal of getting through the gathering without continual feelings of anxiety. Thus, we are relying on it.

That we are relying on alcohol to get us through the gathering, needn't mean that we are dependent upon alcohol more generally. That's because this is one activity we are choosing to pursue, and choosing to harness alcohol within it. It's still up to us whether

we choose to pursue that activity, and how we choose to pursue it. We can choose to use alcohol to help us get through it, or to ride out the feelings of anxiety, or just to not go to the gathering. Relying on something falls out of how we conduct some activity, where pursuing that activity is up to us.

Something relied upon can turn into something we depend upon, though. This can be a gradual process. Recall that reliance can also admit of the habitual aspect. Now, we take a broader overview of what we do, and how we do it. When we say that we rely on alcohol in the habitual sense, this could mean that we attend many gatherings, and tend to use alcohol as a way of getting through these gatherings. Depending on how many gatherings we attend, this role could become deeply embedded in our life. Now, it would be more difficult for us to simply choose to no longer attend gatherings. If we have no other means of getting through these gatherings than by using alcohol, our relationship with alcohol begins to become less take-it-or-leave-it.²⁷

However, in such a case it still isn't clear yet whether it is the alcohol or gatherings that we are dependent upon. It may be that the need for social connection is more pressing, and our use of alcohol is a by-product. On the other hand, it could be that our continued use of alcohol has begun to restructure our inner functioning, such that we are now in possession of a felt *need* for alcohol. As a result, we can start to seek out activities related to alcohol. Now, it's the gatherings we can take or leave; what's crucial is that we have some excuse for alcohol. As such, the felt need for alcohol drives what activities we pursue. This is where we begin to depend on alcohol, rather than rely on it.

6.3 Needs & concerns: instrumental & non-instrumental (fundamental)

These thoughts indicate that we can also shed light on the reliance/dependence distinction in using a distinction between two senses of 'need'.²⁸ A basic sense of X

²⁷ I take it that, in the case of alcohol, the continuous form 'relying' can also sometimes report the habitual. When asked 'how are you doing these days', and we say that we're 'relying on alcohol', we mean to suggest something worrying; implying a relationship with alcohol that could be approaching dependence.

²⁸ I'm only concerned with the concept of *felt* need. We can also need things in an external sense, beyond our awareness. Someone whose sense of thirst is absent still has a need of water. In normal functioning, our felt needs will be reliable indicators of our actual biological needs. Also, what we *actually* need can sometimes be not what we *feel* as

needing Y is that, without Y, X would suffer some sort of harm, or loss (Wiggins and Dermen 1987 : p.63; Wonderly 2016 : p.228). X needing Y is distinct from simply X's *desiring* Y. I can desire to cook a meal for my friends, because I would enjoy doing so. Desiring to cook a meal for my friends, however, doesn't entail I *need* to cook for them. I may simply be excited by the idea of cooking, suffering no harm or loss if I don't. Invoking the concept of need, however, casts the issue of cooking a meal for my friends, as a pressing concern for me.

Given that the concept of need operates in the logical space surrounding reliance and dependence, the concept of 'concern' is also brought into play here. Concern is a concept describing a matter of felt importance; an import that is associated with needing something. Concern is associated with what matters to us. When I need to cook a meal for my friends, this isn't a mere desire, because the matter of cooking a meal takes on a kind of pressing concern for me. That I cook a meal has a characteristic emotional weight or import to me, beyond the feeling that it would be good or pleasant to do so. So, I need to cook a meal, and the matter of cooking a meal takes on a characteristic concern. It matters to me that I cook a meal.

Needs can be both instrumental and non-instrumental (Thomson 1987 : pp.6–7; Wiggins 1991 : pp.9–10). Because concern is intertwined with need, this means that I can also have instrumental and non-instrumental *concerns* for things. When a need is instrumental, it is necessary for some broader goal or activity to be successfully carried out. For instance, I could need to cook a meal for my friends, as part of completing a court-ordered alcohol dependence recovery program, one that requires me to express gratitude to people in my life. The matter of cooking a meal for my friends becomes of instrumental concern to me. Were it that I didn't need to pursue the goal of expressing gratitude, based on court orders (and cooking a meal were not the identified means to pursuing this goal), it would not be a matter of concern to me. I probably wouldn't do it.

When it is of instrumental concern for me that I cook a meal for my friends, this does not have the weight of importance were it be of *non-instrumental* concern for me that I do it. When it is of mere instrumental concern that I cook, it gains some import by proxy,

though we need. It can be that we *actually* need to stop drinking, rather than to have another drink.

from a more pressing matter; the goal being achieved. Having a *non*-instrumental need for something, on the other hand, affects us more keenly. I will also call non-instrumental needs ‘fundamental’ needs.

Fundamental needs are those we have, regardless of what activities we choose to pursue. This means we can’t simply stop needing them, by abandoning pursuing some activity. Fundamental needs will still be pressing, regardless of what activities we choose to pursue. They will in fact tend to determine what activities we choose to pursue, because of how pressingly they are felt. For instance, suppose I instead have a fundamental need to cook a meal for my friends. I feel that, if I don’t, I will be extremely upset and anxious, where this isn’t due to some goal being put at risk by failing to cook. Failing to cook the meal would in itself cause me emotional disturbance, not because I feel a *goal* slipping out of my reach, but because it simply matters to me that I do. What is slipping out of reach is a scenario more amorphous than a specific goal; it is something like ‘my being happy and functioning’. The matter of me cooking has its own importance or resonance, not due to some further end it serves having importance. I simply *need* to do it, to remain content and settled, or for the feeling of relief it will bring.

6.4 Need & concern routed in reliance & dependence

Reliance and instrumental need are not equivalent concepts. Nor are dependence and fundamental need. However, they are closely related. In this section I will elaborate upon how they are related. I will also elaborate on how non-instrumental/fundamental needs and instrumental needs are related. I suggest that instrumental needs are needs-by-proxy. They gain the status of need, by serving to satisfy some fundamental need.

Through relying on things, we come to instrumentally need them. The thing we rely on - that we need to perform some role - thus becomes a matter of instrumental concern for us. We need the thing we rely on, in whatever respect we rely on it. However, not all instrumental needs are things we thereby rely on, because we can instrumentally need to perform actions ourselves. As we’ve seen when discussing the [Intentions Excluded] condition of Smith’s account of reliance (section 5.3.2), reliance involves needing something outside of ourselves to perform some role: whatever we cede control to, to take care of an aspect of our activity. The ‘cooking for friends’ example though is one of

instrumentally needing to carry out an action for myself. I don't thereby rely on myself to perform this action, unless I am alienated from my own actions for some reason. That said, I will be relying on the cooking of the meal as a means to achieving the end I have in mind, insofar as the cooking takes on practical role in bringing about my end. There will also be, in cooking for my friends, things I am relying on while pursuing that end. I will rely on the cooker to cook, insofar as this has been given a role in my activity, directed toward my goal. As such, the cooker working becomes a matter of instrumental need, and so instrumental concern.

Just as reliance and instrumental need are not equivalent, nor are dependence and fundamental need. Though I can have a fundamental need to cook a meal for my friends, I do not, strictly speaking, depend on me cooking a meal for my friends. Just as reliance involves instrumental need for things outside of ourselves, dependence involves fundamental need for things outside of ourselves. Gaining and maintaining access to the things we depend upon, becomes a matter of non-instrumental concern for us.

When I fundamentally need to cook a meal for my friends this is because I have some sort of dependence, that cooking the meal would serve to satisfy. In this case, let's stipulate that I fundamentally need others' approval. I have a fundamental need for it, because my ability to function normally requires that I have approval. I am upset and unsettled without it. As such, I am dependent on other's approval. I am not relying on it, because it doesn't play a practical role in something I am trying to achieve. It is just something I need for my continued functioning, or wellbeing. Cooking a meal is something I fixate on as a fundamental need, and so is of fundamental concern, because it feeds this dependence.

How are fundamental and instrumental needs related to one another? I suggest that part of the answer to this lies in how fundamental need for something overlays meaning, import, and emotional salience upon situations we are in. When I fundamentally need to cook a meal for my friends, this determines what I am concerned about in my current situation, and what I am concerned about happening in the near future. The non-instrumental concern stemming from my fundamental need to cook thereby becomes *routed* through different identified possible means to satisfy that need. This routing, when it passes through an identified means, generates an instrumental need.

Instrumental needs are needs-by-proxy. When, for example, I have a fundamental need to cook for my friends, cooking becomes a matter of concern. I instrumentally need the cooker to work, because my non-instrumental concern to cook the meal is routed through this, as an identified means. When we have a fundamental need for something, instrumental needs are generated, based on our identifying possible ways of fulfilling that fundamental need. That gives an overlay of import on the situation. Then, certain salient possibilities, such as the cooker failing to switch on, my friends being late, will possess import, by virtue of having the potential to obstruct me in cooking for my friends. Our concern is in procuring that fundamental need, and matters that threaten that also become matters of concern.

I have said that my fundamental need to cook isn't strictly speaking what I am dependent upon. That's because my fundamental need to cook in fact stems from a deeper fundamental need for other's approval. That drives me to identify opportunities to gain approval. The goal of cooking a meal, that I fixate on, is simply a product of that deeper fundamental need, identifying some means to being satisfied.

Fundamental needs can serve some purpose, despite being non-instrumental. But, the concern that results from a fundamental need is much more pressing than that for an instrumental need. That's because fundamental needs serve, not some practical purpose, but some more basic core purpose. Fundamental needs are those that flow from the realm of basic biological functioning, and thus evolutionary advantage. As such, the concern generated from a fundamental need points us in the direction of pursuing an outcome that will typically be beneficial to us in a more basic homeostatic sense. For example, the characteristic *impulse* to drink when thirsty is the product of our inner wiring, and the functional role of that impulse is set down in basic biology. The impulse being reliably sensitive to dehydration allows us to maintain an inner equilibrium, maintaining biological structures and processes that constitute us.

Similarly, the logic of our emotional wiring serves a more basic functional role, that has evolved to keep us safe. For example, fear and wariness keep us safe from threats. In the example where I have a fundamental need to cook for my friends, this is because the act of cooking for my friends has immanent emotional salience. It has this immanent

emotional salience because of what purpose cooking for my friends would serve, in restoring an equilibrium in my emotional balance. This may reflect an unconscious association between cooking a meal and preventing something terrible from happening to me, such as being abandoned by my friends. That reflects the operation of underlying emotion schema which has been roughly hewn through evolutionary predispositions, then fine-tuned by experience. That schema has a role in keeping me safe; avoiding threats, and keeping me in proximity of sources of safety and nourishment.

So, even fundamental needs are, in some sense, instrumental. Fulfilling them serves the purpose of satisfying them and restoring a sense of contentment. But this is due to serving some larger evolutionary or biological function. Many fundamental needs: food, water, oxygen, keep us alive. Others contribute to our emotional and physical wellbeing. That they do so is due to those roughly hewn evolutionary predispositions we are furnished with, and which experience sets to work on shaping and rerouting.

However, fundamental needs and instrumental needs are distinct in terms of the concern associated with them. The concern that comes from a fundamental need is distinct from the kind of concern felt for an instrumental need. What we can't do with a fundamental need is directly lose it, by choosing a different means to achieving some goal, or abandoning pursuing some goal altogether. We have limited choice over needing it, in having it matter to us.

Here again then is the essential difference between being reliant on something and being dependent upon it. When we are reliant upon something, it is instrumentally needed as part of a goal we are pursuing. We may not need to pursue that goal, or need to pursue it in the way we are doing so. On the other hand, when we are dependent upon something, it is fundamentally needed because of the kind of thing we are, not because of the kind of activities we pursue.

This account of dependence on something, as involving a non-instrumental concern for it, will be put to use in chapter 8, in defending my account of the interpersonal import of matters of trust.

7 The trust stance: the rolling schema

In this chapter I refocus attention again on the trust stance attitudes, and flesh out the account of the trust stance outlined so far. So far I have defended a distinction between the active and stative senses of trust, and suggested that the ‘stative’ layer consists of attitudes of assurance (stative trust), uncertainty, and wariness (distrust). In this chapter I elaborate on the stative layer, and ultimately motivate an account of the stative layer, that I will sketch out in the last section of the chapter. I call this account the *rolling schema* model.

To build toward the rolling schema model, I first discuss the difference between the active/stative distinction defended in chapter 4, and the closely related distinction of two-place/three-place trust (as raised in sections 1.2 and 1.3). I argue in sections 7.1 and 7.2 that the active/stative distinction is often confused for the two-place/three-place distinction, and set out how they are different.

This directs our attention toward an unusual feature of the stative trust stance. This is the way in which prior stative trust stance attitudes ‘colour’ our perceptions of interactions with others. I defend this perception colouring quality in section 7.3. The thought behind this is that attitudes of the trust stance give us an overlay of interpretative/anticipatory information, about what another is doing, in a way that bleeds from one domain of interaction to another. Following this, I take a closer look at what the stative attitudes, in their perception colouring, appear to be *about*: namely, how another is likely to act, in a situation of vulnerability (as raised in section 1.7).

In section 2.2.3 and also in section 4.7.5, when discussing the nature of distrust as an attitude of wariness, I explained why the Commitment account was highly intuitive. This related to the point, raised in section 1.7, that being ‘vulnerable’ (in an everyday sense of the term) to another’s actions, is not enough to have a matter of trust with her. It isn’t enough to just be aware that another can ‘get’ to you, via a situational vulnerability, for you to treat her doing so as a matter of trust. That’s because the anticipatory attitude you adopt toward her, in that respect, can lack the interpersonal import of the trust stance. You can anticipate that she will attempt to get to you, even be fearful or worried that she will. But, it won’t matter in the way characteristic of the trust stance. This import was

what the Commitment account appeared to identify. It identified what transforms a situation from one of mere prediction, to interpersonal import.

However, in this chapter, in section 7.6.2, I will argue that there must be something more basic than concept of commitment underpinning matters of trust. I'll raise cases where we appear to be able to have that interpersonal import, that put pressure on the concept of commitment. These are cases where it would stretch the concept out of recognition to suppose that a commitment was at issue. Rather, there is something more basic than commitment: a concern about another being able to be pinned down. In the next chapter I will defend an account of what that concern consists in.

7.1 Active/stative & the voluntariness problem

In this section, I set out how the active/stative trust distinction resolves a question, concerning to what extent we can voluntarily trust another. Holton (1994) suggests that the two/three-place distinction can explain why trust appears to be voluntary sometimes but not at other times. I'll argue that in fact the active/stative distinction explains why trust appears to be voluntary and non-voluntary at different times: active trust is voluntary, but stative trust isn't. In the next section I argue that the active/stative distinction cuts across the two-place/three-place distinction.

My proposal to resolve the question over whether we can voluntarily trust – i.e. whether we can take someone up on her request, when she says “trust me on this” – is as follows. The circumstances in which we *can* choose to trust are those where we can simply decide to rely on others, e.g. rely on others to catch us in the drama class ‘trust circle’ (Holton 1994). We can proceed to trustingly rely, because we can, all other things being equal, control how we act, and reliance is a mode of acting (see section 4.7.2 and chapter 5). However, there are circumstances in which we *can't* choose to trust, are those where we can't choose to be assured, e.g. that we will be caught by those in the trust circle. So in one sense, we can decide to trust. But in another sense, we can't. That's because of the ambiguity inherent in ‘trust’: it can mean an act – relying, accepting vulnerability, putting ourselves in a situation of vulnerability - and a state of assurance about the situation that results from that act – the actual state of *being* vulnerable.

REL+ does not have the active/stative distinction. But, it answers the question of how it is that we can decide to trust others in some circumstances, by pointing out that we can decide to rely. The problem that REL+ encounters is that there are circumstances where we *can't* decide to trust others, or where reliance seems to be insufficient for the type of trust at issue.

Such a problem case is found in Baker (1987), where the trust of a friend who has been accused of a crime is at issue. Let's call the accused Millie, and her friend Andrea. Millie asks Andrea whether she trusts her that she is innocent. It would be fair to say that active reliance trust here isn't enough for the type of trust that Millie asks of Andrea. A friend accused of a crime wouldn't want her friend to just *rely* on her being innocent; where that innocence is in some sense remaining to be seen. It remains to be seen, of course, when we rely on those in the trust circle to catch us, whether they will in fact catch us. What a friend wants to know is that we believe her; that there is no doubt in our mind about her having not committed the crime, and of her good character. The kind of trust Millie wants from Andrea is trust in the assurance sense. Anything less than assurance from Andrea would suggest that Andrea was uncertain or wary of Millie, or that Andrea suspects Millie should perhaps doubt her own memory. Without Andrea's assurance, then, Millie's own assurance of her innocence might be shaken. Millie may begin to doubt herself, and this would hinder her own ability to cope with the situation. Active reliance trust in this case is not sufficient for Millie's needs. The kind of trust she wants from Andrea is that which Andrea cannot give voluntarily. That's where stative assurance trust can be appealed to.

Sometimes of course, active reliance trust is all that can be asked for. Compare Millie and Andrea's situation with that between a couple dealing with infidelity. Suppose Xavier has cheated on his partner Yolanda. Xavier asks for a second chance from Yolanda, and Yolanda grants this. Active reliance trust is what she can grant. It cannot be assurance trust; his betrayal has seen to that. But with Yolanda's active reliance trust, it is up to Xavier to continually demonstrate that he is someone she can now be assured of. Xavier can't demand that assurance; it has to be earned through repeated demonstration of being worthy of assurance.

Compare this with what the accused friend Millie is asking of Andrea. The level of assurance Millie and Xavier are entitled to demand from their respective other are not equivalent. Millie, insofar as Andrea is her friend, is warranted in presuming that there is a background of assurance of one another, and so seeks confirmation of Andrea's assurance trust of her. Xavier on the other hand, cannot ask Yolanda to be genuinely assured of him. He can ask for a second chance to prove himself worthy of Yolanda's assurance, by asking her to once again put herself in the position where she is vulnerable to him, by actively trusting him.

As such, we can see the usefulness of the active/stative trust distinction in making sense of the voluntariness of trust. In some scenarios, we can only ask for active trust from another – something that *can* be voluntarily given. Assurance trust though is earned, and not able to be given voluntarily.²⁹ We can to some extent 'control' our stative attitudes, since it is possible to choose to not act on the urge to check or verify that comes from being unsure or wary of another. However, we can't choose to not have the feeling that something is amiss.

Because REL+ can only take 'trust' to mean active reliance trust, it has a difficulty making sense of what Millie can ask of Andrea. Holton's own solution to Baker's problem case, in trying to remain steadfast to REL+, involves invoking a causal link between relying on the friend to be knowledgeable and sincere, and thereby coming to believe what she says. He suggests that, by relying on the friend to be knowledgeable and sincere, "the belief [that she is innocent] follows from the reliance in the sense that I would be failing to act on the supposition that she speaks knowledgeably and sincerely if I did not believe what she said" (Holton 1994 : p.74).

²⁹ Demanding or requiring assurance trust from another can, when taken to an extreme, amount to a form of abuse, since it amounts to asking another to ignore the compass that helps her navigate potentially dangerous or harmful situations. 'Gaslighting' takes this to the extreme, as form of emotional abuse, where the doubts one has are dismissed by another as being the result of one's own poor judgement or faulty senses. If this is taken seriously by the unassured party, it eventually leads to a loss of assurance trust in her own senses, thus being unable to feel assured about what is real, and a paralysis concerning how to proceed and what to rely on sets in. No doubt in some cases our stative trust stance dispositions - levels of assurance or wariness about others in different situations - are unreliable guides to actual threats, prone to disproportionate responses. But often that can also be due to experience priming us to expect certain kinds of behavior from others (section 7.7 offers an account of why).

However, this solution involves some equivocation between two different senses of ‘belief’. The sense of belief (indicated by ‘believe’) in the second half of the sentence, is not the sense of belief that involves assurance that someone is telling the truth, and so being convinced that what she says is true. It is the sense of ‘believe’ found in a more colloquial everyday sense of choosing to believe someone, which actually means to accept what another says, by proceeding to act on its basis. This is not the sense of belief as an attitude of being convinced that p is true. Andrea *would* fail to *act* on the supposition that Millie is speaking knowledgeably and sincerely if she did not choose to believe Millie in the colloquial sense. Andrea *wouldn’t* fail to *act* on the supposition that Millie is speaking knowledgeably and sincerely if Andrea failed to be convinced. Simply choosing to be convinced that p is something we can’t do, and I take it that the active/stative trust distinction helps shed light on why we needn’t suppose we can be voluntarily convinced.³⁰ Instead, I suggest that what we can choose to do is to suspend checking whether what another has said is true; to refrain from acting on our sense of being unsure. Andrea may do this, while keeping her actual doubts hidden from Millie, for her sake.

As such I take it that the active/stative trust distinction resolves the puzzle over how sometimes trust seems voluntarily able to be given but sometimes not, that REL+ cannot satisfactorily resolve. We can decide to actively reliance trust, but not to statively assurance trust.

7.2 Active/stative vs. two-place/three-place

Holton suggests that the reason for any confusion over whether we can decide to trust is because we fail to recognise a distinction between two-place and three-place trust. To trust another in a two-place sense is, Holton tells us, to have what he calls a trusting relationship with another. Three-place trust on the other hand is to trust someone in a

³⁰ For discussion concerning whether we can decide to believe, see e.g. (Bennett 1990) and (Hieronymi 2006). (Ginet 2001) gives an account of what can constitute deciding to believe p where this consists in you *counting on* it being the case that p . To do this is to stake things – ends that you have - on p being the case. In other words, to decide to believe p is to continue to act as though p is true, where doing so puts one’s goals or plans at risk. Here we have a definition of deciding to believe that seems to just be deciding to rely on p ’s truth. This corresponds to deciding to actively trust another.

specific situation, with some action. We can't choose to have a trusting *relationship*, but we can choose to trust someone in a specific situation, i.e. the drama class trust circle. Holton says:

I have treated trust as a three-place relation holding between two people and an action. Contrast this with a two-place relation like friendship or hatred: if I am really your friend, I am your friend simpliciter. I am not your friend for bus driving, or for clothes buying, or whatever. Nevertheless, it seems that there is often a two-place relation that stands behind cases of trust. I have a background relationship of trust with some people; and it is in virtue of this that I would trust them to do a whole range of things for which they have not yet proven their trustworthiness. Can I choose to have such a relationship? Surely not. Compare the case of friendship. I choose my friends; but I cannot choose to be a close friend of yours straight-off. I can choose to start along the path that will lead to that relationship; and I can control, to some extent, how quickly we move along it. However, close friendships are things which develop over time. Similarly with trusting relationships. If you and I trust each other in various ways over time, and our trust is not betrayed, we will be likely to build a trusting relationship. That is not to say that there will be some particular thing that I will trust you to do. Rather it is to say that I will in general be more ready to trust you: partly because I am confident that you will not betray that trust, and partly because, having trusted you before, further trust becomes appropriate. A trusting relationship makes a greater range of trust available to me. As with a close friendship, so with a trusting relationship: I cannot simply decide to have one. But that does not mean that I cannot sometimes decide to trust. (1994 : pp.70–71)

Holton's suggestion, then, is that we cannot will ourselves to have a trusting relationship, while we can decide to trust someone in a specific respect. He supposes that this is what leads to confusion over whether trust can be willed or not.

It isn't clear if this explanation is supposed to shed light on the accused friend case, though. Presumably, Holton doesn't intend to use it to shed light on why there was confusion over whether Andrea could voluntarily trust Millie's innocence, since Holton argues Andrea can in fact voluntarily trust Millie that that she is innocent, in that case. That being said, we could still ask whether the problem in the accused friend case is resolved by pointing out that Andrea cannot voluntarily choose to have a two-place trusting relationship with Millie, but can choose to trust her in a specific respect. Clearly, this isn't what the problem is. Andrea and Millie are *already* friends, and we can presume already have a trusting relationship. The problem is that, in the circumstances, Andrea's *stative* assurance trust of Millie in a particular respect - her innocence - is shaken by the accusations against Millie. She can't voluntarily trust Millie in this particular respect,

because she can't voluntarily be assured. But this assurance is the type of trust that Millie wants from Andrea. The two-place/three-place distinction does not do any work in resolving why Andrea cannot voluntarily trust Millie.

Here, I will set out how the active/stative distinction is not the same as two-place/three-place. In other words, active/stative and two-place/three-place come apart. Pointing this out helps us to further unpack the nature of the trust stance.

We can assurance trust another in both a two-place and a three-place sense. Millie can two-place trust Andrea in a stative assurance sense, where this describes how she generally feels about Andrea. Millie is assured of Andrea, in a general sense. Millie can also two-place trust Andrea in an active sense: Millie trusts Andrea, in the sense that there are many domains of interaction in which Millie habitually relies on Andrea. Her activities are, in some sense, enmeshed with Andrea's.

We can also see three-place trust being cleaved along the lines of the active/stative distinction. Let's hone in on one of the particular domains of interaction between Millie and Andrea: Millie relies on Andrea to bring her toiletries while she is incarcerated. She thus three-place trusts Andrea, in an active sense. However, it also happens that Millie is assured that Andrea will bring her toiletries. She doesn't feel any sense of uncertainty over this. So, she also three-place trusts Andrea to bring her toiletries in a stative assurance sense. Figure 7 (below) illustrates how, I suggest, the active/stative and two-place/three-place distinction cut across one another.

Holton's suggestion is that the puzzle over whether trust is voluntary can be resolved if we notice that we can't decide to trust in a two-place sense, but can in a three-place sense. However, we can deny this claim. We can decide to two-place trust another in an active sense. In fact, if we can decide to trust in a three-place sense, we can also decide to trust over a broad range of things. It is just rare that we do. Normally, we are not inclined to do so; our willingness to is tempered by not knowing a person well, and so not yet feeling assured about the other person in a two-place sense.³¹ But, if circumstances require it, we can decide to actively trust in a two-place sense. The example below will illustrate this.

³¹ Section 7.7 defends an account of why this can be so.

Figure 7: Active/stative & Two-place/Three-place trust

		No. of places	
		Two-place “X trusts Y”	Three-place “X trusts Y to φ ”
Active / Stative	Active	X is relying on Y, in many domains; her activity is enmeshed with Y’s, or is following Y	X relies on Y for some particular purpose, or to perform a particular action φ
	Stative	X is assured of Y in a general sense, in no particular domain	X is assured that Y will or would φ should she rely on her to/while she relies on her to

Wilderness stranger

Belinda has fallen while out hiking in a wilderness, loses her memory, and finds herself alone. A stranger calling himself Kevin, also out hiking in the wilderness, finds her in a state of some distress. Kevin tells her he will help her get back to the nearest town where they can go to a hospital and police station, so that an appeal can be put out. Belinda’s head is foggy, she is frail and requires a lot of assistance in hiking back, a trek that will take several days. This requires her to repeatedly rely on Kevin for help over a broad domain of interactions; manoeuvring around obstacles, help in setting up her tent and camping gear. Belinda’s state of confusion means she feels continually unassured and guarded, but she recognizes that she has to trust Kevin in the sense of actively rely on him for a large range of things. She is forced, in some sense, to adopt a state of childlike dependence on him: she follows him, and needs him, for her survival.

This example is one where we can decide to actively trust in a two-place sense (Belinda could, after all, choose to go off on her own). Belinda trusts Kevin, by following him, and allowing him to lead her back to a place of safety. Belinda actively two-place trusts him: she allows herself to be led by Kevin across many domains. Yet, she doesn’t two-place *statively/assuredly* trust him. If anything, she is uncertain of Kevin.

I take it this is also roughly what Yolanda's active trust of Xavier will consist in, after Xavier cheats on her, should she give him a second chance. Yolanda would proceed to let him back into her life, in many of the domains of interaction that she and he previously were engaged. She actively two-place trusts him, but there will now be underlying feelings of uncertainty in each domain. It won't be one single three-place domain of interaction in which she chooses to rely on him once again, but it will clearly be difficult for her to be open in these domains of interaction, due to the lack of two-place stative trust.

We might wish to deny that the *Wilderness stranger* case represents active two-place trust, because it can be given a three-place reading, e.g. Belinda chooses to actively trust Kevin to *look after her*. But, we can counter this response. If we argue that it isn't an example of active two-place trust, on the basis that it can be given a three-place reading, this threatens to make any sort of two-place trust elliptical for three-place. That's because, presumably, all candidate cases of two-place trust ("Millie trusts Andrea") can be given a very generic three-place reading, by way of giving a very generic characterization of the sort of concern we have, in trusting others ("Millie trusts Andrea to *be loyal*").

The kind of situation encountered in *Wilderness stranger* is rare in adult life. We often don't need to rely on people in such a broad range of domains, all at once. We rarely *need* the level of assistance that Belinda does, because as able-bodied adults privileged with high levels of physical and cognitive capability, we often don't need assistance looking after ourselves. We also aren't inclined to trust people who we don't already feel two-place assurance of (à la Horsburgh's 'I'd trust him with my life' from section 1.3), in situations like Belinda's, that would leave us so able to be taken advantage of. However, when we find ourselves in a situations where we approach something resembling a child-like level of dependence on others, we are forced to actively trust in a number of domains.

In normal circumstances, broad two-place assurance of another lags behind our assurance in particular domains of interaction. Trusting relationships involve both active trust over a broad range of domains, and assurance in those domains. But, such relationships are built based on active trust in specific domains being fulfilled, gradually, tentatively, over time.³² This is possible because fulfilled trust in one domain gets the ball rolling on more

³² In section 8.4 I raise the example of individuals with William's Syndrome, who are trusting-by-default of new people, due to a strong drive for social approach. Such

general assurance of another. Three-place stative trust of another – assurance of her in one domain – bleeds into further domains.

7.3 Perception colouring

This cross-domain bleeding results from a key characteristic of the trust stance, that determines what three-place attitudes we have. Trust or distrust of another on some specific matter is typically the result of our two-place trust stance toward her. It thus isn't the case that the two-place stance we have is a product of what three-place attitudes we have toward her: the two-place tends to determine the three-place. The claim I defend here is that a two-place trust stance *colours* our perception of another and her actions, so that we three-place trust or distrust her on specific matters as a result.

By 'colours', I mean that the trust stance *interprets* what another is likely to do, or is currently doing. We understand what another is doing in some specific situation, or is likely to do in some possible situation, in light of a general trust stance attitude we already hold toward her. For example, when we say of someone, "I'd trust him with my life", we report that, presented with almost *any* given situation, we will three-place trust him with specific actions. Because it is in respect of any given situation, I don't need to have already thought of every possible situation and arrived at a three-place attitude about him. Prior general trust of him means I already interpret his actions in novel situations in a positive light. And, I would be assured of him, under any new domain of interaction. So, I am reporting that, should I find myself in a situation of extreme risk or vulnerability, I would trust him in the specific respects that the situation demanded I trust him. My trust of him colours my perception of what motives he has in current interactions, and in future ones. I will take what he says as truthful. I will interpret his advice, guidance, and offers to perform certain actions, as genuine attempts to help me toward my desired end, rather than attempts to undermine me.

This interpretive or perception colouring quality of the trust stance is what gives the trust stance attitudes a characteristic stubbornness. Assurances from those you already distrust are not going to be instantly effective, because the assurances can be interpreted as further

individuals (typically children) readily, for example, jump into cars with strangers, without any fear of danger (Spiegel 2010).

attempts to mislead you. When we already two-place trust another though, such assurances will typically be unnecessary, barring contexts where her competence is in question. Similarly, when we already two-place trust another, any evidence of possible untrustworthiness is less readily integrated than if we already two-place distrust her. When we already two-place trust another, the possibility that she has hidden motives, such that she intends to undermine us in some particular situation, simply isn't salient, and can be unthinkable. That thought either doesn't enter our minds, or when it does, we dismiss it. We presume that there are no unseen machinations or violations she has concealed from us.³³

This perception colouring feature also determines what counts as evidence of another's trustworthiness. When we already distrust another, we interpret her particular actions and verbal assurances as masking ulterior motives and actions. What she does is interpreted in light of our prior two-place distrust toward her, so that a promise to help you can be taken as insincere, or as another indicator of her untrustworthiness.

A recent real-world example of the perception colouring quality of the trust stance, in interpreting actions, can be seen in the backfiring of Facebook's attempt to clamp down on the spreading of 'fake news' over its network (Cellan-Jones 2017).³⁴ Facebook's original measure for clamping down on fake news articles was to attach a 'red flag' icon to links that were to non-verified, dubious, or misleading news sources. While this was intended to warn users that the link they were about to follow was untrustworthy in comparison to mainstream news sources, the red flag measure backfired. This is because users who were already inclined to look at such misleading news articles only did so, because they already two-place distrusted the mainstream news media and social media platforms. As a result, they interpreted the red flag marker when attached to particular articles, as a further attempt to quash a dissenting truthful view. In light of the users two-place distrust of the mainstream news media and social media platforms, the actions by the social media platform, that ought to have been taken as indicators of the platform's trustworthiness, were taken as further evidence of its untrustworthiness. While these

³³ In light of this, we can see why Millie would be upset if Andrea didn't completely believe her that she was innocent. If Andrea truly did two-place assurance trust her, then any accusations or countervailing evidence against her innocence would surely be interpreted as mistaken.

³⁴ Articles containing fabricated or misleading claims but presented as genuine.

actions were intended to help guide users to trustworthy news sources, these actions were interpreted as a further attempt to hide the truth. Their existing two-place trust stance attitudes coloured their perceptions of Facebook's actions.

The perception colouring quality of the trust stance attitudes, influences how we interpret past, present and possible future actions of others, in terms of their motives. The two-place trust stance is not just present-looking and forward-looking, but also backward looking. Past actions, previously seen as signs of trustworthiness, are coloured in light of untrustworthiness uncovered in the present. When two-place distrust another, we are inclined to think she has already undertaken actions against our interests, or is in the process of doing so, or would do/intends to do so in the future. As such, two-place trust stance attitudes don't simply influence our interpretation of what another is *doing*, but also what she has done. An article on video evidence of American police killing black citizens, offers a useful metaphor for how our prior two-place trust stance attitudes colour or perceptions of others past, present, and future actions:

To the officers who back an expansion of body-camera use, the growing pains are worth the still-unknown cost. To Joe Gamaldi, the Houston union vice president, the greatest benefit of cameras lies in their potential to exonerate officers in an era when more people are distrusting of the official police narrative. "I believe that video will clear them more than not," he said. His view is rooted in the philosophy that cops don't lie as often as all the publicized videos suggest, that the public sees most of the bad but little of the good, that there isn't much more iceberg hidden beneath the dark water. "*How big you think the iceberg is depends on how much you trust them,*" said professor Stoughton. "If you don't trust the police at all, you think that that's a huge iceberg and we only see a little bit of it. If you tend to strongly trust the police, even if we're not seeing all of it, we're seeing most of it." The truth emerges only as the water continues to recede. We don't know what we haven't seen. (Samaha 2017) [emphasis added]

The 'iceberg' metaphor deployed here is powerful one, in helping illustrate the perception colouring quality of the trust stance attitudes. Depending on how much we trust the American police, we will anticipate the number of offences committed by them differently. When we trust another, we think well of her, and don't anticipate that she is concealing matters of concern from us. She is perceived as being open to us, and we'll be assured that any actions that we *aren't* able to glimpse, are not going to be detrimental to our wellbeing (or the wellbeing of those we care about). They will be innocent actions. We also, as a result, interpret what she is currently doing in some situation, in a positive light. On the other hand, when we distrust another, we anticipate concealment of matters

of concern from us (there is iceberg hidden below the surface). We also interpret what she is currently doing as scheming, against our interests.

Lagerspetz (2015 : pp.82–83) and Jones (2013) both recognise the significance of the perception colouring quality of the trust stance attitudes in this respect. Lagerspetz takes this to indicate that trust lies at the heart of human reasoning, rather than being a specialised variety of it:

Trust guides our perception of the facts: it ‘reveals itself in patterns in the agent’s interpretations of reasons’, for example in the fact that we tend to interpret information as evidence in favour of a person we trust. This is an important point, because it suggests an idea of trust as something enmeshed in our cognitive pursuits, not something additional and separable from them. (2015 : p.83)

I suggest this point can shed light on the trust stance in the following way. Without the possibility of general two-place (and one-place) trust stance attitudes, there would typically be no way of making sense of *why* we have some specific three-place trust stance attitude toward another. That’s because the specific three-place trust attitude we adopt toward another in some new domain of interaction with her, is typically determined by the two-place stative trust stance attitude we already have toward her. For example, if we already two-place distrust the police, we interpret what the police are currently doing in issuing a press release, claiming to be conducting an internal investigation of charges of brutality, as presenting a front that allows them to conceal and hide their manoeuvring. On the other hand, if we two-place assurance trust the police, we take this at face value. There is no more iceberg hidden beneath the surface of their actions.

Below, in sections 7.6 and 7.7, I will defend what I call a ‘rolling schema’ account of the stative trust stance. I defend the rolling schema account, via a look at specific cases of *prior, retrospective, counterfactual and hypothetical* three-place stative trust stances. I use this to help explain the relationship between one-place, two-place and three-place trust, and how the trust stance involves a perception colouring quality.

7.4 Situational vulnerabilities & ‘entrusting’: the structure of matters of trust

Before defending the rolling schema model, I will further unpack a concept I make use of in setting it out, that explains what the trust stance attitudes are fundamentally attitudes *about*. This is the notion of a ‘situational’ or ‘situation of vulnerability (which I first raised in section 1.7). I suggest that the notion of entrusting, defended by Baier, can help us here. With the notion of entrusting, we glimpse the basic skeletal structure of a matter of trust. In a matter of trust there must be some way in which another can get at us, via a situational vulnerability we have. This emotional logic of the trust stance involves awareness of this, and a concern about whether another is going to harm us in this respect.

Considering the perception colouring discussion, I will argue that two-place and three-place assurance trust can occur prior to active reliance (section 7.5 below). Assurance trust can also be about the possibility of active reliance (section 7.6 below). This, in combination with the preceding discussions of the active/stative and two-place/three-place distinction, indicates what the stative trust stance is: it is a reading of a situation, where we take a stance on how another will act toward us, given a certain situation where we are vulnerable to her. This vulnerability we can incur through relying on her to perform an action, and so being practically dependent on her action (per the ‘role placement’ view of reliance defended in section 5.4).

But we can also incur a situational vulnerability by simply having something we care about ‘dangerously near’ (Jones 1996 : p.4) or within the ‘striking power’ of others (Baier 1986 : p.235). And, we can have a stative trust stance toward another, in respect of such vulnerability. Consider how we incur such a vulnerability, when another learns an embarrassing secret about us, without our having told her. We aren’t thereby actively trusting her with that secret, but we can still have a stative trust stance toward her in respect of that secret.

As such, the stative trust stance is about a situation of vulnerability, whether or not this vulnerability results from active reliance. But, we also needn’t currently *be* vulnerable. We can have a stative trust stance toward another concerning the possibility of being vulnerable to her, *prior* to that situation of vulnerability. Similarly, we can have a stative trust stance toward another concerning the *possibility* of a situation of vulnerability, whether or not this takes place. What we can do is project forward, or imagine a certain situation of vulnerability, and have a trust stance attitude about it. I defend this further in

section 7.6, in setting out the possibility of trust stance attitudes concerning retrospective, counterfactual, and hypothetical situations of vulnerability.

To help unpack the structure of a situational/situation of vulnerability, I will discuss an analysis of three-place trust that is distinct from the three-place action or ‘reliance’ model typically used so far, where X trusts Y to φ . The alternative model is Baier’s (1986) ‘entrusting’ model. Considering this will help shed further light on how we ought to understand the structure of matters of trust. I suggest that the entrusting model helps us gain a higher vantage point on what is at issue in matters of trust and why our trust stance attitudes tend to be tied to specific actions. Specifically: we trust others in respect of performing specific actions, because those actions have the potential to impact things we care about, given some situation where we are vulnerable to how another acts. This also explains why, for example, three-place trust is often understood in terms of *reliance* on another; we rely on others in respect of specific actions. That’s because actively trusting another by relying on her for something, involves putting ourselves in a situation of vulnerability.

Baier (1986 : p.236), rather than tying trust to a specific action, prefers to analyse trust in terms of a relation to a specific object or *thing*, that has been entrusted to another. Thus, the model of trust being proposed is that, when we say ‘X trusts Y’, this means ‘X trusts Y with some valued thing C’. The valued thing C is something that X cares about, such that C’s fate is a matter of concern for X, and which she has placed in the ‘care of’ Y. As a result, Y has the capacity to either take care of C, or abuse the power made available to her by having access to C. Baier describes this as placing the cared-about thing within the ‘striking power’ of another, and she intends that the analysis extends to all matters of trust. We can call this the ‘entrusting’ model.

The entrusting model clearly fits cases of trust where we physically hand over the things we care about to another. For instance, in giving your friend your cat for the weekend, you entrust your cat to her care. Your friend can then look after your cat, or harm it, or steal it from you, and so on. This seems right. Can the entrusting model extend to more standard three-place ‘action’ or reliance trust cases? I suggest that it can. But, we must do some work in identifying what the entrusted thing is in such cases. We also must adopt a charitable reading of what it is for Y to ‘take care of’ X’s valued thing C.

For example, we can ask what it is we are concerned about – what is at risk of the striking power of another – in the music journalist example from section 1.2. Plausibly, we can say that we are concerned about our limited free time and money not being wasted on bad music, insofar as avoiding that is why we’re turning to the music journalist for advice. We entrust our aim to find new music, to the music journalist. In doing so, we can be misled by her. Our time and resources can thus be wasted on following up her recommendations. In the midwife case, also from section 1.2, we are concerned about the bodily safety of the mother and baby, and we entrust that to the care of the midwife. She has the power to cause drastic harm to both, should she wish.³⁵

Should we prefer the three-place action/reliance model or the entrusting model? I suggest that we should take them to be complementary, rather than in competition. However, the entrusting model is attractive, because it makes explicit the background structure of matters of trust, that the action/reliance model leaves unstated. For example, the entrusting model draws our attention toward the vulnerabilities we have to another in trust; by virtue of what thing is placed within the striking power of another. Such vulnerabilities can be a result of relying on another to perform a role for us (per the role placement model of reliance). When we actively trust the music journalist to recommend good music, applying the entrusting model to the case highlights how our reliance on the journalist gives rise to specific situational vulnerabilities. These include our time and money being wasted, and the possibility of hearing bad music (something we’d prefer not to happen). These can result from our reliance, as we will be following the directions of the music journalist, who is taking on the role of pointing us in the direction of good music. In the midwife case, the entrusting model draws our attention toward the power that the midwife has over the fate of the mother and baby. The mother and baby are ‘in the hands of’ the midwife.

Of course, situational vulnerability is also a feature of the action model. That’s because we have our attention drawn toward the action that someone may fail to do when we trust her, and thus what our goal is at risk from. ‘Recommending good music’ or

³⁵ Underneath this, I think, is also a concern that the other respects our sense of efficacy. We are concerned about not being misled by others, and we prefer to feel effective in pursuing our goals. This is a typical feature of testimonial trust.

‘delivering the baby’, as actions, are in *some* sense within another’s ‘striking power’, because if these actions are left unperformed, something we want to happen will be unrealised. But one of the benefits of the entrusting model is that it illuminates why these actions are important to us, when this is a matter of trust. That’s because these are actions that serve to ‘take care’ of things that we are concerned about. We are concerned about our time not being wasted on listening to terrible records, and we are concerned about the bodily safety of the mother and baby. The entrusting model thus points out how, by relying, we often incur *further* vulnerabilities, that are not captured in ‘an action being left unperformed’. It is not just the possibility of a goal we wanted being left unrealised, that we are vulnerable to. The music journalist has power over whether we *spend money*, and the midwife can now *harm* the mother and baby. In other words, the entrusting model indicates how trustees have the power of *action*, not just *omission*.

Baier hints at this point by noting the importance of ‘discretionary powers’ being granted to others in matters of trust. She draws attention to the fact that, when we trust another to perform some action φ , this is rarely the *only* action that we trust another to perform (1986 : pp.236–237). For instance, when you trust someone to post a letter for you, you also trust her to bring it back, rather than simply leave the letter, should the mailbox be inaccessible for some reason. Similarly, when you trust your friend to look after your cat while you are away, there are a vast range of actions that you don’t explicitly state as being trusted to perform, but that you nonetheless would trust your friend to perform. Should the cat become gravely ill, you would trust your friend to not simply continue feeding it, but to call a vet. By also giving your friend access to your home as part of the cat feeding, you trust your friend to not abuse this power and overstep the boundaries of what she is trusted to do. The entrusting model draws our attention to how ‘X trusts Y to φ ’ is often elliptical for some range of actions that X trusts Y to do/refrain from doing, because there is presumed recognition, by the one we trust, of the different ways in which we make ourselves vulnerable to her actions, by relying on her, or by entrusting an object to her.

A deeper point concerning the power of the entrusting model, is that it helps to exhume the underlying emotional logic structure, in matters of trust. The entrusted object C is, I suggest, akin to what Helm calls, in his account of the structure of emotions, the *focus* of an emotion; a particular object that is important to us (Helm 2009, 2007). The concept

of a focus helps articulate how we respond to others based on their actions toward that object. An example that Helm uses illustrates this: when you see your car is at risk of damage from the children playing a ball game in the street, the focus of your emotion is the car itself, while the targets of your worry and anger are the children themselves (Helm 2014 : p.189). The entrusted object on Baier's account is what equates to focus of our emotion.

As I have already stated, having a situational vulnerability to another is not enough for that situation to be transformed into a matter of trust. Along these lines, we can note that an entrusted object and the focus of an emotion aren't equivalent concepts. We wouldn't say, for example, that Helm entrusts his car to the children. This is even though his car is the focus of his emotion of worry. His car is simply within harms reach of the children's actions. Entrusting something to another, we can suppose, involves some overture to another: it involves active trusting. When we entrust something to another, we rely on her to take care of it. We take it there is recognition from her of what we are placing within her striking power. This is the point I also made in section 2.2.3 when setting out how the Commitment account was highly intuitive as an account of betrayability. There is an underlying structure to matters of trust, that also involves taking it that another recognises our vulnerability to her. Just finding oneself in a situation where a possession or thing we care about is within harm's reach of another doesn't suffice for having entrusted it to them, and so doesn't suffice for that being a matter of trust, yet.

We can, for instance, be careless, and not pay attention to what we are leaving within harms reach of another. We might thoughtlessly leave our bike unlocked outside of the shop. This doesn't mean we entrust our bike to the care of those outside of the shop. Compare thoughtlessly leaving our bike unlocked with asking the person outside of the shop if he would be able to mind our bike for 5 minutes, because we have forgotten to bring our lock. There, we have a case of entrusting it to him. In both carelessly leaving our bike, and entrusting it to another, we would of course be aggrieved should our bike be stolen. But when we thoughtlessly leave it unlocked, we wouldn't take it as a violation of trust that has been extended to another.

This situational vulnerability model also extends to explaining how we are vulnerable to others in cases of 'testimonial' trust; where what is at issue is whether we trust what

another tells us (see section 1.5). This was true of the ‘music journalist’ case in the previous section. Our time and resources can be at stake in an instance of ‘testimonial’ trust; when what is at issue is whether we trust what another tells us. That’s because we can be ‘pointed in the wrong direction’; our actions will be directed by what she tells us. This can put us in new situations of vulnerability. Depending on what exactly we’re being told, it could also be that we’re being guided toward opening ourselves up to a much greater harm. The stranger on the street who advises us that the train station is down the dingy-looking alleyway may be directing us to a place where we can be robbed.

The conclusion we can draw from this discussion, I suggest, is that the entrusting model sheds light on the structure of a situational vulnerability, that are the skeletal structures underpinning matters of trust. Such a situation can result from active trust: where we actively change our situation so that we are vulnerable to another, in relying on her to act. But they can also be passive; features of the situation we find ourselves in i.e. when someone learns an embarrassing secret of ours. In section 7.6 I discuss further cases of passive situational vulnerabilities, that we take trust stance attitudes toward others in.

7.5 Prior assurance trust as a reason for active trust

That two-place assurance trust involves perception colouring, means that we can already trust another in respect of an possible action, prior to our reliance on her to do it. In this section, I show how this constitutes a further point against REL+, which cannot allow that trust can occur prior to reliance. Establishing the possibility of prior trust, though, helps us flesh out the picture of the stative trust stance being drawn in this chapter. The possibility of prior trust will be used in unpacking the ‘rolling schema’ model of the trust stance, that I sketch out in the final section of this chapter.

Trust can occur prior to reliance on another. That prior trust can *explain* why we rely on another. The active/stative trust distinction allows for this possibility. We can rely on someone because we already trust her, in the sense we can are already assured that she will fulfil our reliance on her in some respect. That assurance can explain why we choose to rely on her. REL+ cannot account for this. if trust just is reliance, then to say that we rely on another because we trust her, is circular.

To illustrate the possibility of prior trust, consider Holton's discussion of his 'drama class' case. There are three key claims he makes about this case:

- (i) We fall *because* we trust:
"You let yourself fall because you trust [the others] to catch you" (Holton 1994 : p.63).
- (ii) To rely is to act (fall):
"I can decide to act on the supposition that you will catch me. That is to decide to rely on you." (1994 : p.69)
- (iii) To fall is to trust:
"If you are like me, there is a moment at which you weigh up whether or not to let yourself fall. How does it feel at that moment? It feels as though you are deciding whether or not to trust. I think that we should take this feeling at face value: there are circumstances in which we can decide to trust." (1994 : p.63)

The act of falling is to rely on others to catch you. Then, to rely is to trust, given Holton's view of trust. It follows that when we say, in the first claim, that we fall *because* we trust, this amounts to saying that we *trust* because we trust. But, if trust simply means 'reliance', it is circular or trivial to say that we trust another to catch us *because* we trust another to catch us. If we disambiguate between two senses of trust though, the statement that we fall because we trust becomes reasonable to make. That's because prior assurance trust can be a reason for your deciding to rely on another to catch you. It's not necessary that trust means the reliance itself. In other words, your trust was *prior* to your reliance, and provides an explanation for why you decided to rely. Of course, this possibility contravenes the view that trust just is reliance.

Consider another example:

Cat Friends

Ali and Amy choose to rely on their friend Andy to look after their cats for the weekend, because they trust him to do this. That is, they trust him to do so in the assurance sense. They phone him beforehand to ask if he could take care of the cats this weekend. Andy asks why they would choose to rely on him rather than a professional cat-sitter, and they respond that it's because they trust him over a cat-sitter.

If we are committed to the view that trust just is reliance, then Amy and Ali's explanation as to why they approach Andy rather than a cat-sitter ought to make no sense as an explanation. That's because it would report that they choose to rely on him because they rely on him. But they aren't relying on Andy to look after the cats *before* they proceed to rely on him. The explanation is that the trust they report is prior to their reliance on him. It is their reason for relying on him. This is trust in the sense that they are assured of him, and confident in his capability as a cat-sitter.

Could somebody respond that Ali and Amy are already relying on Andy to look after the cats, prior to him (possibly) agreeing to look after the cats? This wouldn't be right. Ali and Amy are not yet relying on Andy to look after the cats. At this stage they're seeing if they're able to rely on Andy (they're not proceeding on the basis that he will; they're seeing if they *can* so proceed). They aren't yet relying on him, because he hasn't agreed to the role of cat-sitter yet. They are already assured of him as a candidate for this potential role, but he doesn't have it yet. Furthermore, the response that they are trustingly relying on him already is particularly problematic for the Commitment account, because Andy hasn't yet made a commitment to Ali and Amy to look after the cats this weekend.

Could somebody respond that, in the claim that we rely *because* we trust, 'because' is doing the work of identifying a constituent part of that reliance, rather than a reason for it? On this suggestion, trust is *part* of the reliance: it is part of the falling, or the handing over of the cats. In other words, it could be identifying the reason why we give the action the *description* of 'letting yourself fall'/'handing over the cats', rather than identifying a practical reason it is performed. 'Because' is sometimes used in this way when setting out why an object is one of a type, e.g. 'this is a triangle because it has interior angles adding up to

180°, or why a state of affairs/event is of a certain type, e.g. “he was poisoned because his blood contained cyanide”.

However, this isn't the usage of 'because' in the sentence “You let yourself fall because you trust them to catch you”. If we are using 'because' to identify a constituent part of the action, then 'you trust them to catch you', simply doesn't work as an explanation of why the event of letting yourself fall warrants the label 'letting yourself fall'. It may be right in the drama class case that your trusting another to catch is in some sense a part of your letting yourself fall, but it can't work as an explanation of *why* we are describing what you do as 'letting yourself fall', rather than something else. Trusting someone else to catch you presumably isn't a constituent part of an act of letting yourself fall, because you can let yourself fall in any number of circumstances beyond those of drama class trust exercises.

In any case, in the examples of 'constituent part' usages of 'because', there is a clear implication that what is in question is what we say about the object or event at issue, or what we label it. “He was poisoned because his blood contained cyanide” is an answer to the question of why we should say he was poisoned rather than suffered some other cause of death. It is not an explanation of why the poisoner did the poisoning in identifying the motive behind it (e.g. “He was poisoned because he needed silencing”). In “You let yourself fall because you trust them to catch you”, we are not asking whether we should call the event one of letting yourself fall, rather than something else. Instead, we are providing an explanation of why you let yourself fall, in terms of the reason why *you* chose to fall, rather than do something else.

An additional passage from Holton cements that we are referring to the *reasons* why you rely on another to catch, rather than the constituent *parts* of that reliance:

When I let myself fall, and do not put out my hand to save myself, I am relying on you to catch me. It could be that that is all I am doing: I might be relying quite without trust. I might hate the drama class, and regret the day I enrolled for it. [...] I might rely on you to catch me because this is the only way to complete the course and get the credit I need. [...] More likely, however, if I let myself fall I will trust you to catch me. This could be because I think that an atmosphere of trust is important for a drama class; and that one way to work towards such an atmosphere is to trust you. I could value the relationships that my trust will bring. But even if I do not, even if my aim is simply to get through the course, I could

still have reason to trust you. You might have explicitly invited my trust; or perhaps I think that you have done so implicitly by participating in the game. And, given that I already have reason to rely on you, I can think that your invitation to trust gives me a reason to take an attitude of trust towards you. [...] Spelling my reasons out like this makes them seem calculated. We should not think, however, that in general this is so. My reasons might be hidden or unthinking. *I might have always trusted*; perhaps I can imagine no alternative but to go on. (1994 : p.69) [emphasis added]

Again, if trust is reliance, how are we to interpret the remark emphasized, that one of your reasons for falling and relying on another to catch you, could be that you have always trusted the other to catch you? This makes little sense. You can't have trusted prior to the reliance, if trust is reliance itself. With the active reliance and stative assurance trust distinction, Holton's remark is unproblematic. We can just have always been assured.

7.6 The stative rolling schema: retrospective, counterfactual, hypothetical

The previous section contained cases of trust prior to reliance. In this section I raise cases of 'retrospective', 'counterfactual', and 'hypothetical' trust stance attitudes. In addition to being problematic for REL+, these will also motivate an account of the stative trust stance I call the 'rolling schema' model. Our stative trust stance attitudes are products of a rolling schema; a mental model we have of how others are likely to act toward us, in respect of situational vulnerabilities. It is a *schema* insofar as it is a mental framework that interprets and anticipates another's actions. It is *rolling* insofar as how we perceive a situation is coloured by our schema, and the schema is continually updated through experience of others in different situations.

A retrospective trust stance is one where only 'after the fact' we recognise that we have been the subject of a situational vulnerability, that another had an opportunity to take advantage of. Our trust stance attitude is directed toward this situation. A counterfactual trust stance is directed toward a situational vulnerability that *could* have arisen, had we not done something. A hypothetical trust stance is one we take toward an imagined scenario where we have some sort of situational vulnerability.

I present an example case that will be used to illustrate the possibility of retrospective, counterfactual and hypothetical trust stance attitudes.

Offering coat

Maeve has a live-in role supervising undergraduate students in a hall of residence. Part of the role requires her to be ‘on-call’ in the hall of residence overnight, ready to respond to emergencies. One of the problems she must deal with is students taking drugs on site. The university has a zero-tolerance approach to drugs, and if she finds any students doing drugs, she must immediately call the security team, who will attend the scene and deal with the situation. The students are made aware of the rules when they arrive at the beginning of term and are introduced to the live-in welfare team.

Maeve is called out around midnight by a student being disturbed by a group of other students outside of her window. He claims the group are making a lot of noise, and taking drugs. Maeve goes to investigate, finds the small group of students taking drugs, and confronts them. They recognise Maeve as a member of the welfare team. They know that they are now in trouble, with potential consequences extending to fines and possibly being expelled from their course. Maeve calls the university security team, who say they are on their way.

Maeve has in previous weeks had to fine and discipline one of the students in this group – Leo – who was causing noise at a late hour. Maeve and Leo are therefore familiar with one another to some extent. With this new penalty, Leo faces more serious consequences for his course. Leo begins to argue with Maeve, claiming that Maeve is out to get him and his friends specifically, that she could have given them a warning rather than immediately called security. Leo says that he and all of the other students distrust Maeve, and asks her what her problem is. Leo is clearly agitated and on edge – he is shivering uncontrollably – partly from the worry and adrenaline from being caught, partly from the substances he has been taking, but also because it is cold outside.

From previous dealings with Leo, Maeve suspects that he and his friends are behind many ‘pranks’ other students have suffered, including food being stolen and bedroom doors being loudly kicked late at night (before fleeing). Maeve thinks that when confronted Leo tends to lie and charm people in authority in

order to get away with his bad behaviour, and exploits being a part of a rowdy group of friends, to intimidate other students against reporting them (“snitching”). On this occasion, seeing Leo shivering and clearly worried about the consequences of being caught taking drugs, Maeve feels sorry for him, so hands Leo her coat to keep warm, until the security team get here, which will be around 5-10 minutes.

I will use this example and variants of it to help illustrate several key argumentative points about the trust stance:

- (i) The stance can be directed toward retrospective, counterfactual, hypothetical situations of vulnerability.
- (ii) The stative trust stance attitudes needn’t be about situations of reliance, because situations of vulnerability needn’t be matters of reliance.
- (iii) The stative trust stance attitudes we hold are products of a rolling schema; we anticipate what others are likely to do in respect of situational vulnerabilities, based on the model we have of their character, abilities, from our experiences of them and other people like them.

7.6.1 Retrospective trust stances

Retrospective variant:

After the situation is dealt with by security, Leo gives Maeve her coat back. It is only afterward that Maeve realizes her phone, keys and wallet were in the coat pockets, while Leo was wearing it. Had Maeve remembered that these were still in the coat before giving it to Leo, she absolutely wouldn’t have given Leo the coat without removing them first. She distrusts Leo not to the point she thinks Leo would attempt to directly attack her, or cause her trouble in way that is in plain view (Leo isn’t foolish, or out of control). But Maeve thinks that, had Leo thought about it (he was possibly too preoccupied by worry and cold), he would have taken the opportunity to toy with Maeve, given how obviously personally attacked he felt by her. Leo could simply have lied about taking the keys, wallet and phone.

I suggest that this is a case of Maeve distrusting Leo, in respect of him having access to her keys, wallet and phone (hereafter; ‘her belongings’). As such, this is a specific or ‘three-place’ distrust. That attitude of distrust, I take it, stems from a deeper two-place attitude that Maeve has about Leo, informed by her experiences of him so far. What makes the example particularly interesting, and problematic for a theory like REL+, is that it is about a situational vulnerability that has *already* happened, that neither Leo nor Maeve herself recognized as holding at the time. It is retrospective distrust.

Distrust isn’t the only trust stance attitude that can be ‘retrospective’ like this. The other stative attitudes - assurance and uncertainty – can also be held to others regarding retrospective situations of vulnerability. Imagine a different case, where Maeve gives her coat to her trusted friend Pilar while she is cold, and while it also unknowingly contains her belongings. When Pilar gives Maeve her coat back, and Maeve recognizes it had her belongings in the pockets, Maeve would not feel the same sense of alarm (and part relief) as she felt after getting the coat back from Leo. Instead, she would have a sense of assurance trust: a sense of safety or assurance that Pilar had them on her person. It would probably be unthinkable for Maeve that her belongings would be at risk, if Pilar had them.³⁶

Note that this is a case of *retrospective* distrust rather than ongoing/lingering distrust. That’s because it wasn’t true that Maeve held an attitude of distrust of Leo in respect of having possession of her belongings *at the time* he did. She might never have entertained the scenario before it occurred to her that it had taken place. So, this is not a case of Maeve *having* distrusted Leo (or *having* trusted Pilar). This is a case of Maeve now having an attitude of distrust of Leo (and an attitude of assurance of Pilar), albeit in respect of a situation that has now already passed.

The Leo and Pilar examples are problematic for REL+. We have a case of a trust stance attitude about a situation that has already taken place. That situation is where reliance or non-reliance (/avoided reliance, as per the revision considered in section 4.7.3) *could have* occurred. Maeve, had she been aware of the belongings, prior to handing over the coat,

³⁶ We could imagine a further case to illustrate *uncertainty* from chapter 3: Maeve gives the coat to her new colleague Sophie, who she isn’t yet sure about.

could have decided to rely on Leo, or Pilar, to look after those belongings. But neither Maeve (nor Leo, nor Pilar) were even aware of the presence of those belongings in the coat. So, it cannot be that Maeve's distrust attitude is a product of her reliance/non-reliance (/avoided reliance) on Leo/Pilar to look after her belongings. The situation in which she *could have* relied or not has now long since passed.³⁷ This is a trust stance attitude about a situation where the 'reliance stance' does not come into play. It is, rather, about a more general situational vulnerability: where another can harm us, or manipulate us, in some way.

7.6.2 Putting pressure on the Commitment account

In this section I highlight how retrospective cases like this put pressure on the Commitment account of the trust stance. This establishes a point I will elaborate on in the next chapter: that underneath the notion of commitment is a more basic concern we have about others in matters of trust, that seems to be hardwired into the trust stance attitudes. We are concerned that another takes seriously our situational vulnerability, when there is mutual recognition of that vulnerability.

As we've seen in sections 2.2.3 and chapter 4, the Commitment account holds that a three-place trust/distrust attitude toward another requires a belief that she has made a commitment to φ 'ing. The *Offering coat* example however casts doubt on such a belief as a necessary component of a trust stance attitude.

There are two features of the case that suggest trust and distrust needn't presuppose commitment. One of these concerns Leo's attitude toward Maeve. Leo, we can safely say, distrusts Maeve. For a start, he feels persecuted by Maeve; that she 'has it in for him' and his friends. He also suspects that she has a personal problem with him, rather than this incident being a case of carrying out the responsibilities of her job role. As such, he takes it that she is manoeuvring against him, beyond what he can expect of her from her job

³⁷ Though it is true that Maeve is now not relying on Leo to look after her belongings. But this further demonstrates how non-reliance, as we saw in chapter 4, does no work in generating the attitude of distrust. Non-reliance on another in some respect is just the absence of reliance on her in that respect. It is true that we are at each moment not relying on an infinite number of things and people. As such reporting that we don't rely on another in some respect tells us very little about our mental state.

role. He might wonder if she has been waiting for an opportunity to catch him taking drugs, for example.

If the Commitment account is right, his distrust of her in this case should be in respect of a commitment that she has made to him. But arguably Maeve hasn't made any commitment to Leo about the respects in which he distrusts her. If anything, she has made the opposite commitments, because at the start of term, it is made clear to all students that Maeve's job is to enforce a zero-tolerance drug policy. Leo doesn't distrust Maeve in respect of *that* commitment, however. So, the interpersonal salience of Leo's distrust of Maeve cannot be a product of his belief she has made commitments to him to e.g. not report him to security. It is a product of some more basic concern about how Maeve regards him, that underpins, comes prior to, is more basic than any commitments she makes. It is, I suggest, a product of an attitude of wariness about her, that she is undertaking hidden manoeuvring that would exploit any situational vulnerabilities he has.

A further feature of the case, that casts doubt on the necessity of commitments featuring in matters of trust, concerns Maeve's distrust of Leo. Maeve distrusts Leo in respect of him having access to her belongings. But Leo hasn't made any commitment to her to *not* take them. Maeve's distrust of Leo – like *his* distrust of *her* – is more basic, and comes prior to any commitments he makes. It is, again, an attitude of suspecting that Leo will undertake hidden manoeuvring to exploit situational vulnerabilities of his.

We could offer a response to these points on the basis that, as we saw in section 2.2.3, Hawley allows that commitment can take a very broad reading, seeing as commitments can be conferred by “roles and external circumstances” and “mutual expectation and convention” (Hawley 2014 : p.11). Hawley stipulates this when considering the sort of case I am discussing here: where we are simply vulnerable to another's actions, and treat her acting as a matter of trust, without the other person appearing to have made a commitment to not exploiting that vulnerability. The examples she uses in this respect are of trusting your friends not to steal your cutlery, when having them round for dinner, and trusting strangers to let you walk unhindered. In these cases, we could try to stipulate that there is something like a ‘background commitment’ at issue. As such, we can respond that perhaps Leo does have a commitment to not take Maeve's belongings, if we allow

commitments to be incurred by these very broad conditions. By not *yet* having tried to take Maeve's belongings, Leo incurs a commitment to not doing so.

I think we would be entitled to say that Leo hasn't made any commitment to Maeve to not take the belongings in her coat. I accept that we can suppose the "mutual expectation and convention" point allows that we can incur a commitment to another, through a process resembling one I discussed in setting out the Commitment account in section 2.2.3. This was where we had a process of presumed mutual recognition of what X was relying on Y to do.³⁸ This would account for why you can trust your friends to not take your cutlery, when they come to your house. The process of mutual recognition (perhaps low level and unconscious) that occurs when they see your cutlery, and you see them seeing this, and they see you seeing them (etc.), generates the presumed commitment that they will not take it. What we have here though is a case that the mutual expectation and convention point might not be able to account for. That's because Leo can't be a party to any process of mutual expectation, in respect of the action at hand. Not only is the situation now in the past, but Leo wasn't aware that the belongings were in the coat, and nor at the time, was Maeve. So, that process of mutual recognition can't take place. Nonetheless, Maeve has an actual, present distrust attitude about Leo in respect of taking her belongings; she currently is supposing that he would take them, had he known about them, and would keep this hidden from her. She finds herself distrusting *now*, not in the past, when the situation held.

What we can argue, I think, is that the perception colouring quality, discussed in section 7.3, of Maeve's prior distrust of Leo, means that her distrust attitude fills out new scenarios with anticipatory attitudes about what Leo will likely do. So, even though Leo hasn't made a commitment to not take the belongings in her coat, Maeve can anticipate that he would, and this can still be a trust stance that she adopts.

The 'broad commitment' response also seems unable to account for Leo's distrust attitude toward Maeve. It's implausible that his attitude is grounded in Maeve having

³⁸ This sort of process is at the heart of Scanlon's (1990) account of promising, and Black's (2004) discussion of whether obligations can be incurred through reliance. Both involve a set of conditions that stipulate an obligation to another to φ can only go through when X and Y mutually recognise that Y wants X to φ , and that X is encouraging (in some sense) Y to believe that X will φ .

made any of the commitments to him, that would explain his distrust of her. She has made many of the opposite commitments, in fact. Namely, calling security.

Now, we can consider the response that the commitments that are at issue, are extremely general: such as ‘not undertaking hidden manoeuvring’. But I would suggest that this commitment just makes explicit the basic form of the distrust/wariness attitude, rather than being a background commitment that is at issue in any matter of trust. If this is the commitment that Leo takes Maeve to have, and the commitment that Maeve takes Leo to have, then it is, in principle, a commitment like any other. But in fact, ‘not undertaking hidden manoeuvring’ is the generic structure of *any* commitment that could turn a new domain of interaction, into a matter of trust. ‘Not undertaking hidden manoeuvring’ is the generic thing we are concerned about in matters of trust.³⁹ Having to impute this commitment in every case where it seems difficult to find a commitment, would suggest that there is an underlying concern in matters of trust. Specific commitments are crystallised forms of such a concern, applied to specific domains of interaction. Making commitments, I suggest, is a method for eliciting this concern, and so turning a new domain of interaction into a matter of trust.

We can also put pressure on the commitment account in this scenario, by changing the case so that Maeve hands her coat to one of Leo’s friends – Jack – who she hasn’t encountered before. It’s implausible to suppose that Jack, who she has just encountered for the first time, has a commitment to Maeve to not take her belongings, just by virtue of circumstance. I take it that Maeve would, by virtue of Jack’s association with Leo, distrust Jack. But it is probably implausible that any generic commitments Leo has to Maeve (i.e. ‘not undertaking hidden manoeuvring’) would be transferred to Jack, just by association.

We might finally seek to respond that Leo and his friend Jack do have a commitment to Maeve, by virtue of the law, or a code of conduct they have signed. But there are several issues with this proposal. We can imagine a structurally similar circumstance between

³⁹ Making this just another commitment, as the focus of our trust stance, also doesn’t explain *why* this is the underlying concern at issue in matters of trust. In the next chapter, I will set out an account of why it matters to us, in the manner of trust stance attitudes, that others aren’t undertaking hidden manoeuvring against us.

individuals, where there is no formal law or code of conduct, in a 'state of nature'. Or, where the law or codes of conduct, are widely considered to be barely enforced. The requisite attitudes of assurance or wariness over a situational vulnerability would still be felt.

Furthermore, the response that Leo and Jack have a commitment to Maeve by virtue of the rule of law/code of conduct, rests on a dubious premise. This is the implied premise that, prior to the existence of the law or codes of conduct, it can't have been possible to trust/distrust another, in respect of a situation of vulnerability (where any explicit commitments had yet to be made). This is highly implausible. It is true that we can find ourselves in new types of situations of vulnerability, that the law or our institutions have yet to rule on. The law and our codes of conduct are playing catch up with different situations we where we can trust and distrust others.

In addition, I'd suggest that Maeve isn't thinking at the time about the commitments Leo or Jack have, by the rule of law. Her concern about what Leo or Jack might do, doesn't stem from a concern that they might break the law or a code of conduct. In 'live' situations, where things we care about are at risk from another's actions, the law is often far from our minds; we are absorbed in monitoring the situation. It is only later we appeal to the law, to redress a wrong done to us.

7.6.3 Counterfactual & hypothetical trust stances

We can also present a counterfactual variant of the case above. This shows that we can have a trust stance attitude in respect of a situational vulnerability that *could have* arisen. Again, this is without any reliance, or commitments.

Counterfactual variant:

After receiving the coat back from Leo, Maeve is struck by the fear that her belongings may have been in the pockets when she gave him the coat, so checks them. They aren't there. Thankfully, she realises that her keys and phone are in her front jeans pocket, and she left her wallet in her room.

Here is a case of Maeve distrusting Leo over a situational vulnerability that didn't even transpire, but could have done, had circumstances been different. Again, we can pose a counterpart case where Maeve has the same realisation, after receiving her coat back from her friend Pilar. Even though there may be a moment where she is unsure whether her belongings were in the coat, upon locating them and reflecting on the possibility that they *could have* been in Pilar's possession, she will have a sense of assurance about that scenario; she will have counterfactual trust of Pilar.

We can also give a case of a trust stance in respect of a hypothetical situation of vulnerability. This shows that we can have trust stance attitudes about imagined situations:

Hypothetical variant:

Sometime after the drug incident, Maeve is imagining the kinds of thing Leo might do in retaliation. She sometimes shares the kitchen facilities with students. She wonders about the possibility of Leo or his friend Jack entering the kitchen at a time when Maeve has left her cooking unattended. She imagines that they would tamper with her cooking, if they knew she wouldn't be there to see it.

Here again we have a case where Maeve distrusts Leo, in respect of a purely imagined circumstance. As such no reliance/non-reliance has yet to occur. It is simply a distrust attitude Maeve has about Leo, in respect of what he would do, given some situation. And, as with the retrospective and counterfactual cases, we can imagine a corresponding hypothetical assurance attitude that Maeve would have about Pilar, in the same situation. Maeve would anticipate that Pilar would leave Maeve's cooking alone, should Pilar find it unattended (perhaps even anticipating that Pilar would help, by taking responsibility for it, while Maeve was away).

7.7 The rolling schema model

What is the take-home lesson, from considering these different forms of stative trust stance attitude? I suggest that what they illustrate is that the trust stance attitudes are the products of an anticipatory framework we possess, which we apply unconsciously onto

different scenarios of situational vulnerability (whether retrospective, current, or anticipated). The broad framework is what I call a *rolling schema*. It is a mental model, continually applied and updated based on experience (hence ‘rolling’), that helps us to navigate actual, potential, and anticipated situations of vulnerability, in respect of who we are possibly vulnerable to. The rolling schema model takes seriously the lessons of the previous sections, on the relationship between two-place and three-place trust stances, and the perception colouring quality of the trust stance. It also attempts to account for these in a systematic way.

The cases of prior/retrospective/counterfactual/hypothetical stative trust stance attitudes, considered in the previous sections, are examples of a subject’s (Maeve’s) rolling schema being applied to different possible situations of vulnerability with different people. We take anticipatory (stative) trust stance attitudes toward people, in respect of those situations, based on our rolling schemas for those people. The trust stance attitudes that result from our rolling schema, allow us to ‘read’ situations and interpret other’s behaviours within them. They influence how willingly we allow ourselves to get ourselves into, or remain within, situations of vulnerability.

Our schemas are self-focused. They are about what people are prone to do, in respect of things that we are concerned about, in different situations. They encode anticipatory information about people’s motives and their abilities, and thus likely actions. Information contained within this schema, informs our one-place, two-place, three-place trust stance attitudes (see sections 1.2 - 1.4). At the most general or abstracted level, we will have a stative trust stance attitude toward *people in general*, based on our schema for how people generally are toward one another, in respect of a generic situation of vulnerability. This sort of attitude would constitute a one-place stative trust stance: whether we are generally assurance trusting, or distrusting (wary of others).

We can flesh out the schema so that it becomes more fine-grained and specific. One of the ways we can do so is by drilling down into the *type* of person we are considering. We develop schemas for types, such as man, woman, white person, child, student, politician, and so on. For instance, Maeve’s schema for ‘student’ will be different for her schema for ‘university security staff’. We will also develop schemas for types of person in terms of their relationship to us, and to others. I take it that we can have schemas for stranger,

friend, romantic partner, extended family member, and so on, based on our experiences of individuals under those types.

We also develop schemas for specific people, such as your colleague Sophie, the man operating the checkout, President Trump. For example, Maeve's schema for Leo is different from her schema for Pilar. She has a different two-place trust stance attitude toward each. It's clear that Maeve, for example, anticipates Leo and Pilar acting differently toward her, in respect of ways she could be vulnerable to them. That's because she has different schemas for Leo and Pilar in these respects, informed by her experiences of them. These could be experiences of them acting directly toward her, or toward other people. Through this, she forms assessments of their character, and motives. For instance, Maeve has experienced first-hand Leo acting selfishly and recklessly, and as having no fondness for her or people in her position. So, her schema for Leo paints a negative picture of him. Applying the schema to situations of vulnerability involving her (hypothetical, retrospective, etc.), this determines Maeve's trust stance toward Leo in those situations; she anticipates certain kinds of actions from him, based on his character, motives and abilities (she distrusts Leo, as a result, with her belongings). On the other hand, Pilar has been a good friend, so her schema for Pilar paints a positive picture of her. There are going to be very few situations where Maeve would not presume the best of Pilar, unless Pilar begins to act unfavourably toward Maeve.

Maeve also forms assessments of others' competence and willingness (per section 1.2) in certain respects. These assessments then inform the trust stance attitudes she adopts toward them, in situations where she could already be vulnerable, or could make herself vulnerable to them. Competence and willingness assessments inform what situations of vulnerability she puts herself into, in respect of others. For example, Maeve might understand people as generally not being competent in some specific matter. Therefore, she might then suppose *Pilar* is not competent in some respect. As such, there are limits to what she can trust Pilar in, even when she otherwise trusts Pilar deeply, due to her schema for her being so positive in terms of Pilar's character.

Our general assessments of where people's competence and motivations lie – our general schemas - become particularly relevant in institutional contexts, when trying to arrive at policies and procedures for good practice. For instance, in undertaking a complicated

structural engineering project, it will be sensible to have checks performed by multiple engineers on a set of plans, before those plans are approved. It would be irresponsible to solely trust the judgement of one engineer in such a context. This isn't because we have a schema for a specific engineer as having suspect motives, and so worthy of distrust. Rather, it is based on our schema for people being liable to error. It's because our schema has it that, given the scenario at issue, the possibility of an error of judgement slipping through – the costs of which could be catastrophic - is greatly reduced with more eyes covering the same problem.

While schemas can encode competence, they can also encode willingness. This might inform how we structure incentives in institutional contexts.⁴⁰ We know that it is laborious for engineers to seek many checks on work they otherwise want approved. However, it is catastrophic if this doesn't happen. So, we take steps to ensure that there are incentives in place to ensure that engineers in general, not specific ones, will seek the checks needed.

These points concerning competence and willingness are related to why I have earlier stated (first in sections 4.7.4 and 4.7.5) that distrust requires wariness of another in terms of her *motives*. When our schema encodes a lack of competence on some matter, this isn't enough yet to *distrust*. Maeve can *fail to trust* Pilar on the matter of e.g. fixing her boiler, and can feel anxiety about the possibility of relying on her in this domain. A failure to trust another doesn't imply distrust of her, even if it is associated with a feeling of anxiety about what would befall us, should we rely, or be in a position requiring her help. What Maeve would likely do, in recognising Pilar's helpful but misplaced desire to help, is close off this as avenue on which Pilar is to be given responsibility in future, lest Maeve herself be lead to disappointment. This needn't give rise to the stronger sort of anxiety, suspicion, needing to pin down whether Pilar is toying with her, or trying to exploit her, that is accompanied by the sense that Pilar is selfishly motivated. If Pilar fails to take seriously

⁴⁰ I take it that our schemas about people can be influenced by what others tell us. This can cause a spiral of selfishness, because hearing that others are selfish puts us on guard. If we read, from someone we take seriously as an authority, an account of people or specific sorts of people as at root selfish, our schema for people in general may be updated in kind, and we may come to treat others accordingly. It can be an effort to guard against people's views getting behind your eyes and reframing how you interpret others.

what the consequences of Maeve's reliance on her could be, in some domains where she herself is incompetent or liable to be waylaid by other desires, *then* Maeve may come to distrust Pilar.

This leads us to the thought that, when another takes seriously that there are limits to which she is, or can be trusted, this will cast her in a positive light. Recognising there are limits to what another can three-place trust you in, is in fact a way of encouraging two-place trust. You demonstrate a seriousness about how much risk another can be subject to, through your actions, and other's schemas about your motives and character will be updated. To give another example: you might be assured that you are free of sexually transmitted infections, though you haven't had this verified by a clinician. But for your new partner, it may be that she won't simply take your word for it, because she is aware that some infections are symptomless. Also, given that you are not yet entirely familiar with one another, the possibility that you will be more inclined to try to simply get what you want from the situation, is still a possibility that she can't yet discount. She is aware that there are incentives in place that would lead some people, perhaps not you, to be hasty. This needn't mean she generally distrusts you as a person, of course. But, allaying her concerns by having yourself checked by a clinician, would in fact be a way of demonstrating your trustworthiness, not just in this matter, but more generally. By respecting the limits to which others can or should trust us, we demonstrate that in fact we can be trusted on those matters. Our word will come to carry with it a level of assurance that we not trying to mislead, and also that we have taken measures to reduce the possibility that we have misjudged our own competence and willingness. Thus, we assure that there is a slim possibility of negative consequences resulting from trusting us on this matter.

Maeve's specific schemas for Leo and Pilar are updated on the fly by experiences of Leo and Pilar, hence they are *rolling*. Schemas are updated by building associations between people and their actions; making inferences of motivations and character. I suggest that just as her specific schemas for Leo and Pilar are updated by experiences of them, the schemas for the *types* of person that Leo and Pilar fall under are also influenced and updated, based on her experiences with Leo and Pilar. Leo and Pilar are representative of certain types of person, and so what they do is representative of what their types of person do.

The rolling schema model explains how it is that observing the way Y treats a third party Z can affect the stative trust stance attitudes *we* have toward Y. Our feelings about Y will change because our schema updates to show what sort of person Y is to other people. By observing that Y is that way to that type of person, then our schema for her in terms of her character and motivations, and her selflessness/selfishness, is updated. I take it that our schema for other people is also updated by how we *ourselves* act toward others. By seeing that we are an example of a person and that we act in a certain way, we recognise that people *can* act this way in some situation.

The schemas for *types* of people become applied to new individuals we encounter, as we judge individuals to fall under those types. We may, as a result, take a two-place trust stance toward a certain individual, just after encountering her, purely based on our schemas of certain types of person being unconsciously applied to her. For instance, we may have a schema that encodes people from a certain town as being generally very giving, so we might generally trust these individuals. I take it that this will also be true of the schemas we have for ‘relationships’; we apply these to individuals when we discover their relationships with other individuals we have a schema for. This is how Maeve arrives at a trust stance toward Jack, when encountering him for the first time as a friend of Leo’s. She applies the schema for *friend* and Leo to Jack, and comes to distrust him to some extent, as a result.

We also have schemas for certain types of situation, that encode where we are vulnerable in those situations. For example, we will have a schema for how we are vulnerable while waiting at a cashpoint. When our schema for ‘stranger’ is applied to this situation, we arrive at a stative trust stance attitude toward others in a situation of waiting at a cashpoint. This becomes finer grained, depending on how we judge a stranger, present in that situation, as falling under a certain type of person.

The rolling schema model explains the relationship between the one-place, two-place, three-place trust stances, and their perception colouring qualities. We develop schemas for types of person, and individual people. These get applied to specific situations, where we have a situational vulnerability, that another can take advantage of. We interpret what the other is doing in that situation, and what she may or may not be hiding, based on the two-place schema for her that we apply unconsciously. That gets applied based on our

experiences of her. It is also based on our experiences of the *types* of person that we have schemas for, that we judge her as falling under. So, given a situation where we are at a cashpoint with our trusted friend, we will three-place stative/assurance trust her in many unconscious ways, such as trusting her in respect of not attempting to sneak a glance at our pin number. However, when we are at a cashpoint next to someone we don't know, our trust stance toward that stranger will be informed by our schemas for the types of person they fall under.

In the next chapter I will defend an account of the interpersonal import of matters of trust – their betrayability – that informs how our assessment of another's character is at issue in the trust stance. I argue that at the heart of matters of trust, we have a non-instrumental concern about the extent to which others have a non-instrumental concern for us, i.e. the extent to which they care about us. I appeal to the notion of *interpersonal attachment* to capture this. The attachment account sheds light on what exactly our rolling schema is forming an assessment about, in assessing another's motives. If her motives, in a non-instrumental way, take us into consideration, then we form a positive appraisal of her. We come to trust her.

How do we come to form such a positive assessment of another's motives, and so come to trust her? The obvious answer is by specific three-place trusts being repeatedly fulfilled in the past. This is a standard answer in the literature on trust. However, I think this answer ignores the 'bigger picture', i.e. what specific trusts being fulfilled is emblematic of. A curious or unexpected way (because it doesn't appear to be 'rational'), in which we sometimes come to trust others has, I think, been unduly ignored. This is through displays of non-instrumental concern, or affection, or – putting it in more basic terms – people being nice to us, and looking out for us. This encourages the formation of social bonds or attachments to individuals, that underpin two-place trust relationships. The rolling schema model is well suited to explain this. Our schema for a person can be revised so we see them as non-instrumentally concerned about us, just through acts of care and affection.

Consider for example the following case. This comes from an account of a Ugandan woman named Eunice, who was taken from her village and held captive by the Lord's Resistance Army (LRA: a militia responsible for carrying out many abductions and

killings). She was forced to marry a male soldier named Bosco who had also been abducted by the LRA, and forced to fight for them. Their initial encounter causes Eunice to distrust Bosco. However, Bosco loves Eunice, and wants to make their marriage work, so he sets about trying to win her trust:

Eunice tried to escape once but was caught. She and her friend, a young woman like herself, had been washing clothes in a stream when they decided to run. Rebels stationed high in the trees working as lookouts spotted them. Bosco had defended her, prevented her from being killed.

On his forays away from camp, Bosco wanted to be back with Eunice, but he sensed her fear around him and set about trying to win her trust. When LRA leaders asked Eunice to participate on looting or abduction trips, Bosco would lie to them and say that he had ordered her to perform other tasks. She still occasionally had to go on those raids, and so he shielded her in battle. During food shortages in the camp, he scrounged up water and bread and brought it to their tent to surprise her. (Okeowo 2017 : p.25)

What we have in this example are unsolicited expressions of care and affection being used to win trust.⁴¹ I suggest that the rolling schema model of the trust stance, and attachment account of interpersonal import, explains this. Eunice's schema for Bosco was initially one that construed him as selfish, brutal, uncaring. However, as he undertook unsolicited acts of affection and protection, he acted in a way that indicated a non-instrumental concern for her situational vulnerabilities. Eunice wasn't relying on him in any of these respects, but she was in a situation of vulnerability, and Bosco continually displayed a non-instrumental concern about her in those situations. Eunice came to trust Bosco as a result, as her understanding of him - her schema - became radically revised. She later grew increasingly attached to Bosco and fell in love with him. Their bond deepened to the point that they worked together in escaping the LRA; an act that required an extreme level of trust and coordination in their joint activity.

To summarise how the rolling schema model underpins the trust stance: how we feel – what stative trust stance we have toward a people generally, type of person, specific

⁴¹ This suggests that expressions of affection are also a backdoor to persuasion, and thus to making someone more inclined to be vulnerable to you. This seems right. Manipulative people for instance, understand the power of promises, as a means for soothing others. Making a promise displays concern; it implies that you take seriously her need for assurance on some matter, that her assurance matters to you, and thus it has the power to elicit her assurance. Thus, insincere promises can be deployed as a way of further manoeuvring.

individuals, either in general situations or specific ones – will be informed by associations we build through experience of these types of person/specific people (perhaps direct, or via indirect testimony from others). We build certain schemas for certain individuals, or types of person, and apply those to schemas we have for types of situation. We come to anticipate actions in respect of our situational vulnerabilities, from different types of people.

8 The interpersonal import of matters of trust: attachment

In this final chapter I defend an account of the interpersonal import of matters of trust. I call this the Attachment account. My suggestion is that at the heart of any matter of trust, a non-instrumental concern for secure attachment to others must be at stake, at some level. The phenomenon of social attachment is a kind of dependence; we need to live alongside others for our emotional and physical wellbeing. In other words, we have, by our nature, a non-instrumental concern for others being non-instrumentally concerned about *us*. As such, the account harnesses the concepts I set out in chapter 6: the difference between an instrumental concern, and a non-instrumental concern, and the notion of dependence upon something.

This explains the betrayability of the trust stance attitudes, because in betrayal we see the possibility of being deserted by someone we have some level of attachment to. This means that the phenomenon of interpersonal trust, and the trust stance attitudes, revolve around our attachment security, at some level.

8.1 Recap: betrayability

The interpersonal import of matters of trust is glimpsed in the possibility of betrayal. Here I will briefly recap the discussion of betrayal and betrayability given so far. In section 1.7 I discussed the different ways that trust can be disappointed or harmed, in an interpersonally significant way. In section 2.2 I discussed how the REL+ literature has progressed based on improved accounts of betrayability, and I endorsed Hawley's Commitment account in this respect. I suggested that the Commitment account corresponded with an account of betrayal I endorsed in section 1.7: Shklar's (1984) definition of the 'simple act' of betrayal:

For a simple act of betrayal, one person should have both intentionally convinced another person of his future loyalty and then deliberately rejected him. The latter then, not surprisingly, feels betrayed. (Shklar 1984 : pp.140–141)

Betrayal, I suggested in section 1.7, was the flag-bearer for a family of related harms we can suffer in matters of trust. It is through the possibility of betrayal that we glimpse how matters of trust carry an importance to us, beyond the possibility of loss or gain resulting

from another's actions. Betrayal is standardly appealed to in the literature to characterise the interpersonal import of matters of trust. I suggest that if we are to seek an account of the interpersonal import, we must seek an account of betrayability.

We also need one that solves problems the Commitment account seemed to face in 7.6.2. There it seemed the concept of commitment was unable to explain cases where we have a trust stance attitude toward another in respect of some passive situational vulnerability. In these cases, the commitment we need to attribute amounts to the generic form of what the assurance or wariness attitude is about, i.e. that the other is not concealing actions against our interests. We don't yet, though, have any account of why another doing this against us, should matter to us, in the way that gives rise to betrayability. In effect, we need an account that fills in the gaps in the Commitment account and Shklar's definition, that explains why 'convincing another person of his future loyalty and then deliberately rejecting him', *matters* to us in the way that it then causes us to feel betrayed, as opposed to simply frustrated. For this, we need to look more closely at what matters to people, that explains why we can feel the vivid interpersonally oriented feelings of betrayal, and associated harms.

8.2 'Right kind of motive' views

To motivate the attachment account I first consider accounts of interpersonal import that attribute it to a concern about another's *motives*. I argue that these accounts fail, because they allow that concern to be an instrumental concern. This establishes a key argumentative point in favour of the attachment account: matters of trust must involve a non-instrumental concern about another's motives. This was seen in reflecting on the con-artist case, and the contrast 'Flatmate' cases considered in section 1.6.

The interpersonal import of a matter of trust is typically explained by appealing to the possibility of betrayal in a matter of trust. Without the possibility of betrayal, the matter is one of mere prediction or anticipation of another performing an instrumentally beneficial/detrimental act. The intuition that trust is distinct in terms of betrayability was thus the motivation behind Baier's 'Goodwill' account from section 2.2.1. The Goodwill account told us that the interpersonal import of matters of trust – betrayability – stemmed from the fact that when we trust, we rely on another's goodwill toward us. The underlying

thought behind this account was that matters of trust involve a concern about another's motives, in some way. This seemed at first glance to explain what was different about Kant's neighbours' reliance on Kant when they trusted him to walk, as opposed to merely relying on him walking.

The Goodwill account isn't the only account of trust that attributes betrayability to a concern about another's motive. Jones' 'Affective Optimism' account (section 4.7.4) involves a concern that the other is "favourably moved by the thought that we are counting on her," (1996 : p.4). In other words, when we trust another, we take it that she is positively moved by the fact of our counting on (/relying on) her. We attribute to her a positive motive, such that a concern about that motive is part of the trust attitude. Similarly, if we recall Hardin's 'Encapsulated/shared interest' view (section 1.8), his account tells us that trust involves a presumption that "it is in [the trustee's] interest to maintain our relationship," (1998 : p.12). Even though Hardin doesn't suppose this attribution of a motive lends any normative significance to trust, we can treat it as a type of 'motive' account, like the Goodwill and Jones' Affective Optimism accounts.⁴²

8.3 The con-artist & (non-)instrumental concern over motives

The issue with 'right kind of motive' accounts is that they suffer from an objection that Holton raises against Baier's 'Goodwill' account. This was the con-artist case, from section 2.2.1. What the con-artist case showed was that reliance on another's goodwill toward one, as the motive for her action, was insufficient for trusting her. That's because a con-artist could possess this attitude, and a con-artist precisely doesn't trust you to act.

I suggest that the deeper reason why the con-artist objection proves fatal for 'right kind of motive' views, is that a con-artist has an *instrumental* concern about the motives of his mark. Even if the con-artist has a situational vulnerability toward his mark, in that he is

⁴² I am here cherry-picking from the Affective Optimism account. Based on how Jones construes optimism as an "affectively loaded way of seeing the one trusted," (1996 : p.4), the Affective Optimism arguably can escape the con-artist objection; a con-artist doesn't have an affectively loaded way of seeing his mark. I also presume that the 'affectively loaded way of seeing the one trusted' and concern that the other is 'favourably moved by the thought that we are counting on her', that Jones refers to, are circling the account I set out below.

liable to lose out on time, resources, and may be caught out, he doesn't have the right kind of concern about the motive of his mark, for it to be a matter of trust. This is why Jones' and Hardin's accounts also suffer from the same counterexample. It is of instrumental importance to the con-artist that his mark has favourable attitudes toward him, because those favourable attitudes are playing a practical role (as per the role placement view of reliance in 5.4) in motivating the mark to do useful things for the con-artist (e.g., handing over the bank details).

Sections 6.3 and 6.4 set out the difference between instrumental and non-instrumental concerns. When we have a non-instrumental concern for something, this is because it matters for our wellbeing, and we are liable to suffer emotional distress without it. The con-artist, though, has an instrumental concern that your motive toward him is favourable. Your motive toward him, as manifest in whether you do hand over your money, doesn't matter in relating to a basic need that others are favourable toward him. Matters of trust, though, invoke such a need. We are liable to suffer emotional distress – feelings of betrayal – where that distress stems from the fact that the other doesn't appear to care, or be concerned, about us. It is important to us, in a *non-instrumental* way, that we matter to the other.⁴³

Recall the discussions, from section 1.7, of Jackson's (2000) and Shklar's (1984) accounts of betrayal. Both accounts supported the thought that what another does, in a matter of trust, matters to us, from a concern about *desertion*. The possibility of being deceived matters to us, not simply because we have an instrumental need for others to be truthful and to perform beneficial actions. Rather, we have a non-instrumental concern that others are true to us; that they do not abandon us, or convince us of their loyalty, in order to reap some instrumental benefit. We want to matter non-instrumentally, to the other person.

Remarks from Shklar paint the non-instrumental need in vivid colours: “whenever our friends desert us that unquenchable uneasiness wells up in us, and we, however momentarily, are infants again,” (1984 : p.140), and, “the greatest of childhood anxieties, the fear of abandonment,” is stirred up in betrayals; “In quitting a bonded group, an

⁴³ A con-artist may be distressed if the mark doesn't hand over the bank details, but not because this indicates the mark doesn't care. It will more likely be because it is of import, for his own needs, that he gets access to some money.

equally primeval fear is stirred: of the failure to distinguish kin and stranger, the latter almost always called “enemy” as well. To reject a blood relationship for a new and alien association, or for none at all, is to deny the most elementary of social ties.” (1984 : p.139). An account of the interpersonal import of the trust stance must explain why others’ actions elicit this sort of concern in situations of vulnerability.

8.4 Interpersonal attachment

To explain that non-instrumental concern, I appeal to a concept that comes from developmental psychology: attachment. The concept of attachment is outlined by Bowlby (1969, 1973, 1980) and Ainsworth (1969) to characterise the emotional bond that children and caregivers form, such that the continued presence of the caregiver is a matter of emotional wellbeing for the child.

Bowlby’s studies of attachment centred on an scenario called the ‘strange situation’. The strange situation involves observation of a child playing in a room, in close proximity to her caregiver. The caregiver then leaves the room, and a stranger enters. The caregiver then re-enters the room. From observing different ways children respond to the caregiver upon re-entering, Bowlby and Ainsworth theorise that the child’s way of responding manifests different methods of securing the continued proximity of the caregiver and/or of coping with her felt need for her caregiver. In most cases (in healthy, or secure styles) this involves approaching the caregiver for reassurance, to secure continued proximity in future.

Underpinning each style of attachment is the same core phenomenon of emotional dependence on the caregiver. When a child is attached to a caregiver, she is emotionally dependent on her caregiver’s ongoing presence and proximity, at the level of a basic need. She needs the caregiver to be present, as a matter of felt security. Ainsworth describes how this need manifests in distress when the caregiver leaves, and the ensuing behaviours to secure the continued proximity of the caregiver:

To attract his mother to him, he can signal through crying, smiling, babbling or calling, through gestures such as raising his arms, and through a variety of other behaviors. As soon as he is mobile he can regulate proximity by approaching or

following his mother, and he can achieve contact by climbing up on her and can maintain contact by clinging. (1969 : p.1003)

Approach behavior may be activated either by distance [...] or by alarm. Alarm is an intense activator and is likely to evoke very rapid locomotion, in contrast to the leisurely approach likely under less intense circumstances. Furthermore, as the activating conditions become more intense, different forms of behaviors may be elicited. The alarmed child may scream in addition to approaching. [...] Whereas low-intensity behavior may be terminated merely by sight or sound of the mother - or by some kind of reassurance that she is nearby or likely soon to return - high-intensity behavior has close physical contact with the mother as its set-goal. (1969 : p.1004)

The concept of attachment from developmental psychology can shed light on the concern underpinning matters of trust, from childhood to adulthood. A persisting need to form attachments to other people and a wider group, gives rise to the interpersonal import seen in betrayability. There are clear echoes of the reactions of the child to abandonment by her caregiver, within the interpersonally focused reactions experienced in disappointed trust, as adults. Betrayal and abandonment, as Shklar and Jackson's illustrations point out, have the threat of desertion as their focus. I take it that this is because matters of trust, at their most basic level, hinge on our underlying need for continued proximity to others.⁴⁴ This stems from an attachment concern that the majority of us have, as a feature of our psychology and functioning.

My account of the interpersonal element of matters of trust and the trust stance then is the 'Attachment' account. The proposal is that for another's action to be matter of trust for us, it is necessary that our attachment concern is (indirectly) able to be impacted/shaken by another's action toward us, in respect of a situational vulnerability. If our attachment concern is not able to be impacted/shaken in how another acts toward us in respect of a situational vulnerability, then her acting is not a matter of trust for us. The trust stance attitudes have as their focus (à la Helm's notion, discussed in section 7.4) those things (e.g. our cat, the money in our bank account) and goals (e.g. the book being returned to the library) that we are concerned about, that are within the 'striking power' of another's actions, in active or passive situational vulnerabilities. But the trust stance attitudes have as their *deep focus* our non-instrumental need to have others non-

⁴⁴ Recall Dasgupta's appeal to the enforcement agency of 'society itself', and the threat of social ostracism resulting from shame, from section 1.8. These are only truly threats when we have a felt need to belong.

instrumentally concerned about us; to belong to a group, and to maintain proximity to those individuals within the group we have close relationships with. That need stems from our underlying non-instrumental concern for secure attachments.

When another fulfils our active trust, or helps us in situations of passive vulnerability, that can amount to an expression of non-instrumental concern for *us*, via things we are concerned about.⁴⁵ This was seen in the example case of Eunice and Bosco from section 7.7. When we are helped despite not relying on another, we can, if there are no ulterior motives detected, come to statively trust her. When our trust is let down, or betrayed, this gives rise to the characteristic feelings of desertion, and leads to distrust. Our need to belong, for others to stay in proximity, such that we non-instrumentally matter to them, is made pressing. As a result of the betrayal, we are shown that we do not matter to another. We are abandoned, used instrumentally, taken advantage of. As such, situational vulnerabilities are *windows* of vulnerability, to a deeper attachment concern we possess.

The concept of attachment I am using has its root in the concept outlined by Bowlby and Ainsworth to characterise the emotional bond between child and caregiver. However, I am applying the concept of attachment more broadly to describe the ongoing felt need for the presence of other people, and for assurance of their non-instrumental concern about us, that persists throughout life.⁴⁶ In this respect, this more general concept of attachment takes its core meaning from the phenomenon outlined in Bowlby and Ainsworth's Attachment theory. But, it is adapted to characterise a broader phenomenon of social dependence, that has its roots in the bond between child and caregiver. So, attachment is derived from the concept that describes the emotional bond between a child and caregiver, but applies to matters of trust more generally. Matters of

⁴⁵ I presume that being helped or having active trust fulfilled won't amount to a display of non-instrumental concern for us, if it's the result of self-interested motives. If we detect this, then we won't come to statively trust the one who helped (via a rolling schema as outlined in 7.7). Consider a case where we fall down in the street, and are helped up by a passerby. When we see that the passerby is a celebrity, who proceeds to admonish his trailing camera crew for missing the incident, then we won't take this act as kindly helping us, as much as helping boost his profile.

⁴⁶ The exception may be those with psychopathic traits, who have a largely instrumental concern about contact with other people (Kelly 2015b, 2015a, 2015c). I suggest this would be connected to why so few psychopaths can treat interactions with others as matters of trust, and why interactions with psychopaths are marked by manipulation and exploitation. Their motivations are more aligned with those of the agents described in the Prisoner's Dilemma (section 1.8).

interpersonal trust in the adult world, in the various of domains of interaction we have, are, in a sense, Bowlby's 'strange situation', writ large.

Wonderly (2016) puts forward a philosophical account of attachment, that nicely captures the essence of the concept. Wonderly outlines "security-based attachment" as a rich "mode of mattering". She suggests "that the relevant form of attachment involves a felt need for its object and particular relationship between the object and the attached agent's sense of security." Furthermore, this is a distinct attitude "from the more philosophically familiar notion of caring." (2016 : p.223). Wonderly summarises the core concept as follows:

In this form of attachment, the agent experiences a particular object as a felt need, such that her senses of well-being and general competence suffer without it. Unlike caring, this attitude is largely self-focused and marked by an integral connection between its object and the agent's felt security. (2016 : p.224)

Attachment is, like caring, a psychological phenomenon implicated centrally in things *mattering* to us. But, attachment involves a mode of mattering that is stronger than caring about something. Attachment is related to our proper functioning. Again, this places it within the realm of dependence, as I outlined in chapter 6. As such, when we are attached to someone, this is not a matter of just having an attitude of desire for the person (as per the traditional philosophical concept of desire), or an emotional attitude of caring about her. As Wonderly emphasises in the passage above, attachment involves a *self-focused* concern for the continued presence of the object.⁴⁷ It is not that we *necessarily* care about the other person, rather that *we* need her. When our felt need for another is so strong (perhaps due to a rolling schema that makes the possibility of abandonment especially salient), we can in fact fail to display care about *her*; being so focused on our own need of her, that she is neglected as a result. As mentioned in chapter 6, in being a non-instrumental need for another, this will give rise to goal-directed activity to secure her presence, when this is felt to be under threat. Precursors to this are glimpsed in early childhood, in the caregiver proximity-maintaining behaviour that Ainsworth outlines above. Wonderly summarises this succinctly: "in virtue of (what I feel as though is) my need [for the attachment object], I am tugged this way and that" (2016 : p.228).

⁴⁷ This corresponds to the thought I put forward in section 7.7 that the rolling schema involves a self-focused model of what others are like to *you* and those you are attached to.

The concept of attachment helps characterise how, post-childhood, the close personal relationships we have with others are tied to our sense of wellbeing and felt security. Those of us with an attachment concern have a disposition to form social bonds, and need close relationships, e.g. friends, family, romantic partners.

The strong relationship between trust and the concern for attachment, finds support in cases of individuals with Williams's Syndrome (WS), who are 'pathologically trusting'. Such individuals find it difficult to distrust others, and are driven to approach strangers with the level of affection reserved for close friends (Dobbs 2007; Doyle et al. 2004; Järvinen-Pasley et al. 2008; Moseley 2014). Empirical evidence suggests that this results from a lessened ability to detect threat of deception or ill will in facial expressions (Ng et al. 2015; Riby et al. 2014), such that WS individuals don't have the necessary competing evidence of untrustworthiness to temper a strong drive for social approach. Such individuals appear to have heightened levels of oxytocin – the hormone implicated in affiliative attachment (mother/child, romantic partners, friends etc.) – when around individuals, as a result (Dai et al. 2012). Ng et al. (2015), though, note that WS individuals are able to discriminate between trustworthy and untrustworthy individuals when reflecting on abstract examples of behaviour.

To summarise: Trust is assurance about a matter of trust. Distrust is wariness of another, also in respect of a matter of trust. Matters of trust are important, because they are situations where our attachment security – our need to matter to others – is at stake. Now, it may be intuitive that an attachment concern underpins matters of trust with those we already trust, but what about with strangers/non-intimates? It also seems much less intuitive that an attachment security is what we are concerned about in respect of those we *distrust*. In the next section I set out how attachment underpins matters of trust with non-intimates. In the final section, I set out how it underpins matters of trust with those we distrust.

8.5 To the group

I take it that there is a clear correspondence between two-place trusting relationships and the attachments involved in interpersonal relationships. However, I have suggested that

attachment concern gives rise to the interpersonal import of matters of trust more generally, because a matter of trust involves a window of vulnerability on a core attachment concern. Matters of trust, though, can arise with those we do not have a personal relationship with. As such, we require an explanation of how it is that we have our *attachment concern* at stake, such that we can trust or distrust those individuals we do not have a close relationship with.

My proposal here is that we are (often *via* our close attachments), attached more generally to a social group. A sense of secure belonging to a group is core to our security, and wellbeing. Matters of trust with non-intimates can strike at our attachment security, insofar as they reflect our secure attachment to the group. We have a social dependence on the group at large. This corresponds to the notion of basic or one-place trust, discussed in section 1.4.

We can motivate this by appealing to literature on the negative effect that social isolation has on our wellbeing. Brownlee (2013) puts forward a defence in favour of a right to social contact with others, such that deprivation is a violation of human rights. She argues that, “Having minimally adequate opportunities for decent or supportive interpersonal contact and social inclusion is both a necessary condition for a minimally decent human life and a constitutive part of such a life” (2013 : p.200). In defence of this she cites cases that illustrate the effects of social deprivation on astronauts, long-distance solo sailors, and prisoners left in solitary confinement. In each case, the effect of isolation constitutes not just emotional but also physical harm, such is the strength of the need for social contact. The behaviour of individuals subjected to isolation tends to become driven around the procurement of interactions with others, with long-distance sailors coming to “depend on radio and video communications for social contact” (2013 : pp.205–206).

Cacioppo & Patrick (2008) argue that loneliness and social isolation have a profoundly deleterious effect on physical and emotional wellbeing. They suggest that the “special balm of acceptance that [social] bonds provide, and the uniquely disturbing pain of rejection when they are denied, is what makes humans so highly attuned to social evaluation. We care deeply what others think of us, and this is why, of the ten most common phobias that cause people to seek treatment, three have to do with social anxiety: fear of speaking in public, fear of crowds, fear of meeting new people.” Like

Brownlee, they note the punishment of banishment/ostracism constitutes “the most severe punishment, short of torture or death,” and that this is why in modern prisons, “the penalty of last resort is solitary confinement.” (2008 : pp.10–11). The lasting effects of isolation take their toll in physical health, even in cardiovascular functioning (2008 : p.31), because of the stressful effect of ongoing feelings of abandonment resulting from isolation. The effect of such a stressor also affects our ability to focus on simple tasks, with loneliness causing an attentional deficit when faced with tasks requiring higher levels of concentration. In a study involving students who were required to distinguish between sounds, played into their ‘non-dominant’ ear (from which auditory input is harder to differentiate), the students who reported higher levels of loneliness found it harder to perform the task (2008 : pp.36–37). This is no doubt because social isolation is experienced as an ongoing stressor; it weighs on our motivations, as a pressing concern, to the point that we don’t have the cognitive capacity to devote to other tasks, when we feel the threat of social isolation.

The effect of lacking contact with *other people*, regardless of whether those are people we have a close relationship with, amounts to the effect of being deprived of something we are dependent on. This is regardless of whether the other person is someone we are closely attached to/who we have a personal relationship with. In a domain of interaction with an unknown, she has the capacity to make clear our lack of worth, so having the potential for manifesting our lack of worth to the group more generally. It is in that interaction with her we glimpse the possibility of being left behind by others. If an unknown is in a position to demonstrate concern for our situational vulnerabilities, then that situation can become a matter of trust; we have a window of vulnerability to our need to be anchored within a wider group. It is telling that the effect of being short-changed, let down, or betrayed, in an interaction with an unknown, will prompt the sort of assurance/soothing seeking behaviour descended from that we see in the child who seeks reassurance from her caregiver. We might appeal to those we have close attachments to, such as our friends, to validate that what the other person did to us was ‘not on’. What we are doing is seeking the comfort of secure bonds, to reassure us of our worth to the group. If our *friends* don’t validate us in that situation, we may be betrayed twice-over; first by the action of the unknown, then by our friends. We will likely feel rejection more strongly, as a result of the latter. The close relationships we have anchor us in a wider group.

Alternatively, we sometimes require the validation of the group more broadly. As discussed in section 1.4, Herman (2001) notes that the shattering of basic trust experienced in severe trauma, requires validation from the community as a whole, to be overcome. When the group fails to recognise and express non-instrumental concern for the victim's standing in the group, the victim can feel lingering betrayal and ongoing distress. What the group does to us, and how the group responds to what members do to us, matters to us non-instrumentally. It's plausible to suppose that the legal institutions we have to resolve disputes and rule on accusations of harm are, at some level, functions of our felt need for validation and assurance from the group.⁴⁸ When another exploits our position in a situation of vulnerability, or fails to express the non-instrumental concern we need in such a position, we seek the affirmation from the group that we do indeed matter (plus, that the other person is made to understand this). But when the group as a whole responds with indifference, or against us, we feel betrayed, because our lack of worth to the group then is communicated clearly.

The feeling that results from the group failing to express concern over our situational vulnerabilities, can come apart from the express commitments that have been made to us by that group. Nonetheless, we can feel betrayed - not as a result of broken commitment, but as a result of our lack of worth to those whom we have an attachment to (as seen in the actions of a member of that group). This was seen in section 7.6.2, when Leo's distrust of Maeve didn't appear to match up with the express commitments that had been made to him. His feeling of having been wronged, and resulting distrust, stemmed from a more basic concern about how others viewed him. An example from Herman illustrates a similar thought:

The imagery of [traumatic] events often crystallises around a moment of betrayal, and it is this breach of trust which gives the intrusive images their intense emotional power. For example, in Abram Kardiner's psychotherapy of the navy veteran who had been rescued at sea after his ship was sunk, the veteran became most upset when revealing how he had felt let down by his own side: "The patient became rather excited and began to swear profusely; his anger aroused clearly by incidents connected with his rescue. They had been in the water for a period of about twelve hours when a torpedo-boat destroyer picked

⁴⁸ This is opposed to the view of the law, set out in section 1.8, as an enforcement agency we designed to ensure that we more effectively *delay gratification*, as the 'defect' option is most preferential.

them up. Of course the officers in the lifeboats were taken off first. The eight or nine men clinging to the raft the patient was on had to wait in the water for six or seven hours longer until help came.”

The officers had been rescued first, even though they were already relatively safe in the lifeboats, while the enlisted men hanging onto the raft were passed over, and some of them drowned as they awaited rescue. Though Kardiner accepted this procedure as part of the normal military order, the patient was horrified at the realization that he was expendable to his own people. The rescuer’s disregard for the man’s life was more traumatic to him than were the enemy attack, the physical pain of submersion in the cold water, the terror of death, and the loss of the other men who shared his ordeal. (Herman 2001 : p.55)

In this case, the procedure to rescue officers first was, we can presume, an expressed commitment (or a consequence of one). However, the veteran in question experienced the moment of being passed up by the torpedo-boat rescue team as a betrayal, regardless of what had been explicitly communicated. His extreme situational vulnerability – being left at the threat of death in the water – was, by the navy’s actions, made clear to not be a matter of pressing concern. The feeling of betrayal in this case stems from his deep attachment to a wider group being completely shaken.

So, while we may not be tempted to say that we *trust* unknown individuals we encounter as part of living among the wider group, our interactions with these individuals can be *matters* of trust. That is insofar as it matters to us whether others manifest non-instrumental concern to us, due to our need for secure attachment to that group. So, we can have trust stance attitudes toward them. It can matter to us that we have a basic level of concern manifested toward us by the group we belong to.

This corresponds to a point made by Gudrun (2016), who notes the importance of low level gestures of goodwill and concern, in neighbourhoods comprised of disparate individuals, from diverse backgrounds. Gudrun points out from an ethnographic study that, while large diverse neighbourhoods don’t foster many personal trust relationships, neighbours place an importance on feeling “safe” (2016 : p.27). A large part of this feeling of safety hinges on conventions among residents of expressing minor gestures of social contact and recognition, such as nodding or greeting when walking on the same paths, fostering feelings of connection among different and relatively unknown individuals (2016 : p.31). While these neighbours don’t place importance on close two-place trust relationships with their neighbours, this isn’t as important to them as the feeling that they matter at a basic level to their fellow residents. This enables high levels of basic trust to

be hand, and a sense of safety as part of a group, even when surrounded by relative unknowns.

Views of trust in the literature (like those considered in section 1.8) that de-emphasise the interpersonal import of matters of trust, from focusing on interactions between strangers, are short-sighted. Just because matters of trust between strangers appear not to require much of a strong personal attachment, this does not entail that attachment isn't our underlying *concern* in our interactions with strangers. Our concern is with how others, who purport to be part of our group, regard us. We are concerned with whether *we matter* to the group or not. Even if these interactions occur in the detached exchanges of the marketplace, the possibility of such exchanges being construed as matters of trust, *presupposes* that we have an underlying concern about group attachment. The marketplace, after all, is a space of shared activity, governed by rules of fair exchange, buttressed by legal institutions to enforce them. Those rules are a reflection of a concern to maintain group cohesion, to ensure that we are validated in cases of being victim to another's harm. We need rules that 'contain' the fallout of upset resulting from selfish behaviour (which a well-formed rolling schema would suggest is an inevitability; see 7.7), to minimise retaliatory behaviour, and so we can remain assured about engaging in that shared activity. Onora O'Neill remarks that Baier's view, which emphasises interpersonal relationships, is "nostalgic," and not for "trust in a complex social world" (2015 : p.49). She fails to recognise that the beating heart of matters of trust, in a complex social world, is our need for a sense of belonging to, mattering to, a group, and the individuals within it.

The above discussion may explain how matters of trust can be had with those whom we already identify as belonging to a group we identify with. But what of those who we do not identify as possibly belonging to a group we identify with, as in a 'state of nature' case? My suggestion is that such an interaction can still implicate a *potential* for an attachment, either to the individual, or to the group this unknown belongs to. This, combined with the well-supported empirical claim that most of us have a disposition to seek out secure attachments, means that we can treat an interaction with a complete unknown as a matter of trust. On the other hand, if we are fearful, or wary, to the point will don't open ourselves to the opportunity to receive non-instrumental concern from others, then we may be unable to get the ball cannot get rolling on that potential

attachment. But, recall Eunice & Bosco from section 7.7: it may be that this unknown has the ability to help us out anyway, so riding roughshod over the defences we put up.

8.6 Betrayability varies with attachment concern, not with stative trust

In this final section, I address the thought that while the attachment account can plausibly explain the interpersonal import of matters of trust with those we already trust, and those we do not know but have an attachment to a group of which she is also a part, the account cannot explain how *distrust* has interpersonal importance, i.e. how we can be betrayed even by those we distrust.

The answer to this is hinted in the last paragraph of the previous section: it can still be important to us that we matter to those who we distrust, even if our distrust attitude is pulling us away from them. There is a feeling of tension in distrust that I suggest partly results from the underlying concern for belonging, with the protective anticipation that we ought to pull away, to protect ourselves from another. The attachment account helps explain this. Those we distrust, I suggest, have our attachment concern ‘in their teeth’, perhaps via their commitments.⁴⁹ Consider the accompanying feeling of warmth and relief, when someone we previously distrusted, turns out to have our concerns at heart. This motivates the sort of bonding process we see, noted by Hawley, when we express sorrow for having previously distrusted another, who instead fulfilled our trust (2014 : p.3). If there was no attachment concern implicating the one we distrusted, prior to our now being helped by her, then the process would not involve the *relief* that we *can* trust her. It would be closer to beginning a new friendship, from a starting position of how she regarded us not previously being a matter of non-instrumental concern to us.

Attachment concern underpins distrust, and the attachment account can help explain the claim that has been noted previously, that it is possible to be betrayed by those we distrust. We can, because we can have strong attachments to those we distrust. Stative assurance is not a pre-requisite of betrayability. To illustrate this I argue against a point that Hieronymi (2008) makes in defence of the opposite claim.

⁴⁹ See footnote 41, p. 186

Hieronymi draws a distinction between ‘trusting belief’ (i.e. stative trust) as a ‘primary’ or fundamental sense of trust, and ‘mere entrusting’ (an activity of reliance, i.e. active trust). She argues that trusting belief is the primary sense of ‘trust’. She uses the trust-betrayal connection as a premise (the degree to which we trust is the degree to which we can be betrayed). Then, she argues that degree to which we can be betrayed (statively trust), is the degree to which we trustingly believe. From this, she concludes, the degree to which we trust another is the degree to which we trustingly believe her in some matter (2008 : pp.228–230). The argument form is as follows:

P1. Degree of trust is the degree of betrayability

P2. Degree of betrayability is the degree of trusting belief

C. Degree of trust is the degree of trusting belief

We arrive at C through the transitive. But, the argument is unsound, because P2 is false. Instead, what determines the degree to which we are betrayable is the degree to which we have a concern stemming from an *attachment* to the other.

To support P2 Hieronymi compares two examples; (i) we trust a friend with a secret and “fully believe” that she is trustworthy, and (ii) where we have doubts about our friend. If the friend were to tell our secret to others, she argues that in the first case where we had a fully trusting belief, we are more betrayed than in the second case (2008 : p.230). Hieronymi also suggests that we often mitigate the possibility of betrayal by mitigating our trusting belief (2008 : p.231), i.e. we tell ourselves we never trusted anyway. This is why, she argues, the case of merely entrusting the secret to your friend has lesser betrayability than the first case.

While it might be right that being uncertain or distrustful of another can incline us to *begin* the process of *detaching* from her, which shields us against the painful feelings of betrayal, this is distinct from not having *assurance* about her. An element of uncertainty or distrust about a particular matter of trust doesn’t diminish the extent to which we can be betrayed on that matter. Consider Yolanda who strongly suspects her partner Xavier of cheating, and is nonetheless severely betrayed when her suspicions are confirmed. Perhaps after many repeated violations of trust, Yolanda will have begun to detach to the point that the

betrayals no longer sting. But this is not mitigating the extent of the stative trust attitude, it is mitigating the extent of her attachment.⁵⁰

There is an even bigger problem that results from the supposition that degree of betrayability is degree of stative trust, rather than the degree of attachment concern. It would allow that the sorts of harms that can be suffered by cheated on partners, are unproblematic, in cases where cheaters are distrusted anyway. At the extreme end, this would have the consequence that in cases of systematic abuse (e.g. of children by caregivers, women by abusive husbands), involving repeated betrayals, such abuses could become unproblematic. That is, if the victim can be said to distrust the abuser anyway. The abuser would be warranted in saying that it wasn't a betrayal that he committed, because he wasn't trusted. Betrayability is not a by-product of one's assurance, however. Rather, it is a product of your need for secure attachment. Betrayal is the turmoil that results from having someone you depend on, as a fundamental emotional source of security, make clear your lack of worth. It is a product of what that assurance was of, being shattered. The process of detachment following abuse and harm like that suffered in severe betrayal is wrenching, precisely because of the centrality of the need for attachment, to our wellbeing.

8.7 Summing up

The attachment account is well suited to explaining what the characteristic interpersonal import of matters of trust is, and what the trust stance attitudes are at a deeper level *about*. The account explains that matters of trust, and the trust stance attitudes, carry a characteristic interpersonal import, stemming from an underlying need for secure attachments, to individuals and to a group. This attachment concern being implicated in situations of vulnerability, gives rise to the possibility of interpersonal harms such as betrayal. Insofar as a situational vulnerability is a matter of trust, then we must have a window of vulnerability, such that another's actions can affect us on the level of our attachment security. That is whether the attachment is to her, or to the group of which she is a member. In effect, matters of trust have a characteristic interpersonal importance,

⁵⁰ Hieronymi, in endorsing P2, is inconsistent with her endorsement of the claim that distrust is able to be betrayed (2008 : p.229).

that presents opposing horizons. Beyond one lies increased contentment and a feeling of belonging. Beyond the other lies the pain of rejection.

Conclusion

I have defended an account of interpersonal trust and distrust, as attitudes of the ‘trust stance’; attitudes we take toward others in matters of trust. Matters of trust are situations of vulnerability, where others have the power to help or harm us. These matters can be products of active reliance on others, or a result of our situation. Matters of trust have a characteristic interpersonal importance to us, otherwise they are not matters of trust at all. That importance is glimpsed in the possibility of being betrayed by those who act against our interests in such matters. I have defended an account of that interpersonal importance, as being the product of our need for secure social attachments, to individuals or a group. This is a form of social dependence on others, and relates to our core wellbeing.

To support my account, I argued that the dominant ‘Reliance plus’ view of trust and distrust in current literature is flawed, for the primary reason that it fails to distinguish between an activity of trust (reliance on another) and a mental state of trust (assurance about her). Following this, I gave an account of reliance as involving practical dependence on things we harness in goal-directed activity. I then distinguished reliance from a more general notion of being *dependent* on things that relate to our core functioning and wellbeing. This account of dependence was used to support my attachment account of the interpersonal import of matters of trust.

I have also defended an account of the trust stance attitudes as the products of a rolling schema; an anticipatory framework we have for discerning whether individuals, types of people, or people in general, are motivated non-instrumentally and capable of helping us in respect of situations of vulnerability. We develop a sense of assurance of others (and so trust them), or wariness of their motives (and so distrust them). The rolling schema, in conjunction with the attachment account, suggests that the trust stance attitudes reflect what individuals and group we feel we matter to. This will be as manifest in their actions toward us, in respect of situations of vulnerability.

Bibliography

- Adler, J.E. (1994) Testimony, Trust, Knowing. *The Journal of Philosophy*. [Online] 91 (5), 264. Available from: doi:10.2307/2940754.
- Ainsworth, M.D.S. (1969) Object Relations, Dependency, and Attachment: A Theoretical Review of the Infant-Mother Relationship. *Child Development*. [Online] 40 (4), 969–1025. Available from: doi:10.2307/1127008.
- Alonso, F.M. (2016) Reasons for Reliance. *Ethics*. [Online] 126 (2), 311–338. Available from: doi:10.1086/683536.
- Alonso, F.M. (2009) Shared Intention, Reliance, and Interpersonal Obligations. *Ethics*. [Online] 119 (3), 444–475. Available from: doi:10.1086/599984.
- Alonso, F.M. (2014) What is reliance? *Canadian Journal of Philosophy*. [Online] 44 (2), 163–183. Available from: doi:10.1080/00455091.2014.919722.
- Audi, R. (2006) Testimony, Credulity and Veracity. In: Sosa, E. & Lackey, J. (eds.) *The Epistemology of Testimony*. [Online] Oxford University Press. Available from: doi:10.1093/acprof:oso/9780199276011.001.0001 [Accessed: 2 September 2017].
- Axelrod, R.M. (2006) *The evolution of cooperation*. rev. ed. New York, NY, Basic Books.
- Ayer, A.J. (1952) Negation. *Journal of Philosophy*. 49 (26), 797–815.
- Baier, A. (1986) Trust and Antitrust. *Ethics*. 96 (2), 231–260.
- Baker, J. (1987) Trust and Rationality. *Pacific Philosophical Quarterly*. [Online] 68 (1), 1–13. Available from: doi:10.1111/j.1468-0114.1987.tb00280.x.
- Becker, L.C. (1996) Trust as noncognitive security about motives. *Ethics*. 43–61.
- Bennett, O. (1990) Why is belief involuntary? *Analysis*. [Online] 50 (2), 87–107. Available from: doi:10.1093/analys/50.2.87.
- Black, O. (2004) Reliance and Obligation. *Ratio Juris*. 17 (3), 269–284.
- Bowlby, J. (1969) *Attachment and Loss, Volume I: Attachment*. [Online] London: The Hogarth Press and the Institute of Psycho-Analysis. Available from: <http://www.citeulike.org/group/884/article/633517> [Accessed: 17 May 2016].
- Bowlby, J. (1973) *Attachment and Loss: Volume II: Separation, Anxiety and Anger*. London: The Hogarth Press and the Institute of Psycho-Analysis.
- Bowlby, J. (1980) *Attachment and Loss, Volume III: Loss, Sadness and Depression*. London: The Hogarth Press and the Institute of Psycho-Analysis.
- Bratman, M. (1987) *Intention, Plans, and Practical Reason*. Center for the Study of Language and Information.

- Bratman, M.E. (1992) Practical Reasoning and Acceptance in a Context. *Mind*. 101 (401), 1–16.
- Brownlee, K. (2013) A Human Right Against Social Deprivation. *The Philosophical Quarterly*. [Online] 63 (251), 199–222. Available from: doi:10.1111/1467-9213.12018.
- Cacioppo, J.T. & Patrick, W. (2008) *Loneliness: human nature and the need for social connection*. 1st. ed. New York, Norton.
- Castelfranchi, C. & Falcone, R. (2010) *Trust theory: a socio-cognitive and computational model*. Wiley series in agent technology Chichester, West Sussex, U.K, J. Wiley.
- Cellan-Jones, R. (21st December 2017) Facebook ditches fake news warning flag. *BBC News*. [Online] Available from: <http://www.bbc.co.uk/news/technology-42438750> [Accessed: 26 December 2017].
- Coady, C.A.J. (1994) *Testimony*. [Online] Oxford University Press. Available from: doi:10.1093/0198235518.001.0001 [Accessed: 2 September 2017].
- Cohen, L.J. (1989) Belief and acceptance. *Mind*. 367–389.
- Crowther, T. (2011) The matter of events. *The Review of Metaphysics*. 3–39.
- Dai, L. et al. (2012) Oxytocin and Vasopressin Are Dysregulated in Williams Syndrome, a Genetic Disorder Affecting Social Behavior. Slattery, D.A. (ed.) *PLoS ONE*. [Online] 7 (6), e38513. Available from: doi:10.1371/journal.pone.0038513.
- Dasgupta, P. (1988) Trust as a Commodity. In: Gambetta, D. (ed.) *Trust: Making and Breaking Cooperative Relations*. Blackwell, pp.49–72.
- Dobbs, D. (8th July 2007) The Gregarious Brain. *The New York Times Magazine*. [Online] Available from: <http://www.nytimes.com/2007/07/08/magazine/08sociability-t.html> [Accessed: 2 November 2016].
- Doyle, T.F. et al. (2004) “Everybody in the world is my friend” hypersociability in young children with Williams syndrome. *American Journal of Medical Genetics Part A*. [Online] 124A (3), 263–273. Available from: doi:10.1002/ajmg.a.20416.
- Engel, P. (1998) Believing, holding true, and accepting. *Philosophical Explorations*. [Online] 1 (2), 140–151. Available from: doi:10.1080/10001998058538695.
- Engel, P. (1999) Dispositional belief, assent, and acceptance. *Dialectica*. 53 (3–4), 211–226.
- Faulkner, P. (2007a) A Genealogy of Trust. *Episteme*. [Online] 4 (03), 305–321. Available from: doi:10.3366/E174236000700010X.
- Faulkner, P. (2011) *Knowledge on trust*. New York, Oxford University Press.
- Faulkner, P. (2007b) On Telling and Trusting. *Mind*. [Online] 116 (464), 875–902. Available from: doi:10.1093/mind/fzm875.

- Faulkner, P. (2015) The attitude of trust is basic. *Analysis*. [Online] 75 (3), 424–429. Available from: doi:10.1093/analys/anv037.
- Friedrich, D. & Southwood, N. (2011) Promises and Trust. In: Sheinman, H. (ed.) *Promises and Agreement: Philosophical Essays*. Oxford University Press.
- Gambetta, D. (1988) Can We Trust Trust? In: Gambetta, D. (ed.) *Trust: Making and Breaking Cooperative Relations*. Blackwell, pp.213–237.
- Ginet, C. (2001) Deciding to Believe. In: Steup, M. (ed.) *Knowledge, Truth, and Duty*. [Online] Oxford University Press. Available from: <http://www.oxfordscholarship.com/view/10.1093/0195128923.001.0001/acprof-9780195128925> [Accessed: 27 May 2015].
- Govier, T. (1998) *Dilemmas of trust*. Montreal ; Ithaca, McGill-Queen's University Press.
- Govier, T. (1992a) Distrust as a Practical Problem. *Journal of Social Philosophy*. [Online] 23 (1), 52–63. Available from: doi:10.1111/j.1467-9833.1992.tb00484.x.
- Govier, T. (1992b) Trust, Distrust, and Feminist Theory. *Hypatia*. 7 (1), 16–33.
- Gudrun, J.T. (2016) Neither Trust Nor Distrust. *Nordic Journal of Migration Research*. [Online] 6 (1), 25–32. Available from: doi:10.1515/njmr-2016-0009.
- Hardin, R. (1998) 'Trust in Government' In: Braithwaite, V.A. & Levi, M. (eds.) *Trust and governance*. The Russell Sage Foundation series on trust New York, Russell Sage Foundation, pp.9–27.
- Hardin, R. (1996) Trustworthiness. *Ethics*. 107 (1), 26–42.
- Hawley, K. (2014) Trust, Distrust and Commitment. *Nous*. [Online] 48 (1), 1–20. Available from: doi:10.1111/nous.12000.
- Helm, B.W. (2007) *Emotional reason: deliberation, motivation, and the nature of value*. Cambridge studies in philosophy. Cambridge, Cambridge University Press.
- Helm, B.W. (2009) Emotions as evaluative feelings. *Emotion Review*. 1 (3), 248–255.
- Helm, B.W. (2014) Trust as a Reactive Attitude. In: Shoemaker, D. & Tognazzini, N. (eds.) *Oxford Studies in Agency and Responsibility, Volume 2*. [Online] Oxford University Press, pp.187–215. Available from: <http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780198722120.001.0001/acprof-9780198722120-chapter-10> [Accessed: 3 November 2016].
- Herman, J.L. (2001) *Trauma and Recovery*. Pandora.
- Hertzberg, L. (1988) On the attitude of trust. *Inquiry*. [Online] 31 (3), 307–322. Available from: doi:10.1080/00201748808602157.
- Hieronymi, P. (2006) Controlling attitudes. *Pacific Philosophical Quarterly*. 87 (1), 45–74.
- Hieronymi, P. (2008) The reasons of trust. *Australasian Journal of Philosophy*. [Online] 86 (2), 213–236. Available from: doi:10.1080/00048400801886496.

- Hill, C.A. & O'Hara, E.A. (2006) Cognitive Theory of Trust, *A. Wash. UL Rev.* 84, 1717.
- Hinchman, E.S. (2005a) Advising as inviting to trust. *Canadian journal of philosophy.* 35 (3), 355–385.
- Hinchman, E.S. (2005b) Telling as Inviting to Trust. *Philosophy and Phenomenological Research.* 70 (3), 562–587.
- Holton, R. (1994) Deciding to trust, coming to believe. *Australasian Journal of Philosophy.* [Online] 72 (1), 63–76. Available from: doi:10.1080/00048409412345881.
- Horsburgh, H.J.N. (1960) The Ethics of Trust. *The Philosophical Quarterly (1950-).* [Online] 10 (41), 343–354. Available from: doi:10.2307/2216409.
- Jackson, R.L. (2000) The Sense and Sensibility of Betrayal: Discovering the Meaning of Treachery Through Jane Austen. *Humanitas.* 13 (2), 72–89.
- Järvinen-Pasley, A. et al. (2008) Defining the Social Phenotype in Williams Syndrome: A Model for Linking Gene, the Brain, and Behavior. *Development and psychopathology.* [Online] 20 (1), 1–35. Available from: doi:10.1017/S0954579408000011.
- Jones, K. (2004) Trust and Terror. In: DesAutels, P. & Walker, M.U. (eds.) *Moral Psychology: Feminist Ethics and Social Theory.* Rowman & Littlefield, pp.3–18.
- Jones, K. (1996) Trust as an affective attitude. *Ethics.* 4–25.
- Jones, K. (2013) 'Trusting Interpretations' In: Mäkelä, P. & Townley, C. (eds.) *Trust: analytic and applied perspectives.* Value inquiry book series Nordic value studies Amsterdam, Rodopi, pp.15–29.
- Kaplan, M. (1981) Rational Acceptance. *Philosophical Studies.* 40 (2), 129–145.
- Kelly, J. (2015a) Psychopathy and Attachment Disorders? *Psychogendered.* 10 June 2015 [Online] Available from: <http://psychogendered.com/2015/06/psychopathy-and-attachment-disorders/> [Accessed: 7 November 2017].
- Kelly, J. (2015b) The Ivory Box - More on Attachment Disorders in the Psychopath. *Psychogendered.* 11 June 2015 [Online] Available from: <http://psychogendered.com/2015/06/the-ivory-box-attachment-disorders-in-the-psychopath/> [Accessed: 7 November 2017].
- Kelly, J.B. (2015c) *A Tale of Two Masks.* S.l., Blurb.
- Keren, A. (2014) Trust and belief: a preemptive reasons account. *Synthese.* [Online] 191 (12), 2593–2615. Available from: doi:10.1007/s11229-014-0416-3.
- Lackey, J. & Sosa, E. (eds.) (2006) *The Epistemology of Testimony.* [Online] Oxford University Press. Available from: doi:10.1093/acprof:oso/9780199276011.001.0001 [Accessed: 2 September 2017].
- Lagerspetz, O. (2015) *Trust, ethics, and human reason.* London, Bloomsbury Academic, an imprint of Bloomsbury Publishing Plc.

- Lahno, B. (2001) On the emotional character of trust. *Ethical theory and moral practice*. 4 (2), 171–189.
- Levi, M. (1998) A state of trust. *Trust and governance*. 1, 77–101.
- Mäkelä, P. & Townley, C. (eds.) (2013) *Trust: analytic and applied perspectives*. Value inquiry book series Nordic value studies Amsterdam, Rodopi.
- Marušić, B. (2017) Trust, Reliance and the Participant Stance. *Philosopher's Imprint*. [Online] 17 (17). Available from: <http://hdl.handle.net/2027/spo.3521354.0017.017> [Accessed: 21 November 2017].
- McCraw, B.W. (2015) The Nature of Epistemic Trust. *Social Epistemology*. [Online] 1–18. Available from: doi:10.1080/02691728.2014.971907.
- McGeer, V. (2008) Trust, hope and empowerment. *Australasian Journal of Philosophy*. [Online] 86 (2), 237–254. Available from: doi:10.1080/00048400801886413.
- McLeod, C. (2015) Trust. In: Zalta, E.N. (ed.) *The Stanford Encyclopedia of Philosophy*. Fall 2015. [Online] Available from: <http://plato.stanford.edu/archives/fall2015/entries/rust/> [Accessed: 19 July 2016].
- Moseley, T. (2014) *What Happens When You Trust Too Much*. *The Atlantic*. 12 May 2014 [Online] Available from: <http://www.theatlantic.com/health/archive/2014/05/going-to-work-with-williams-syndrome/361374/> [Accessed: 30 March 2016].
- Mourelatos, A.P. (1978) Events, processes, and states. *Linguistics and philosophy*. 2 (3), 415–434.
- Ng, R. et al. (2015) Reasoning About Trust Among Individuals With Williams Syndrome. *American Journal on Intellectual and Developmental Disabilities*. [Online] 120 (6), 527–541. Available from: doi:10.1352/1944-7558-120.6.527.
- Okeowo, A. (2017) *A Moonless, Starless Sky: Ordinary Women and Men Fighting Extremism in Africa*. [Online] Available from: https://nls.ldls.org.uk/welcome.html?ark:/81055/vdc_100050390016.0x000001 [Accessed: 2 January 2018].
- Owens, D. (2006) A Simple Theory of Promising. *Philosophical Review*. 115 (1), 51–77.
- Parsons, T. (1989) The progressive in English: Events, states and processes. *Linguistics and Philosophy*. [Online] 12 (2), 213–241. Available from: doi:10.1007/BF00627660.
- Pettit, P. (1995) The Cunning of Trust. *Philosophy and Public Affairs*. 24 (3), 202–225.
- Pink, T. (2009) Promising and Obligation. *Philosophical Perspectives*. 23 (1), 389–420.
- Railton, P. (2014) Reliance, Trust, and Belief. *Inquiry*. [Online] 57 (1), 122–150. Available from: doi:10.1080/0020174X.2014.858419.
- Riby, D.M. et al. (2014) Stranger danger awareness in Williams syndrome. *Journal of Intellectual Disability Research*. [Online] 58 (6), 572–582. Available from: doi:10.1111/jir.12055.

- Rowe, R. & Calnan, M. (2006) Trust relations in health care: developing a theoretical framework for the “new” NHS. *Journal of Health Organization and Management*. [Online] 20 (5), 376–396. Available from: doi:10.1108/14777260610701777.
- Samaha, A. (2017) *How Video Finally Proved That Cops Lie*. *BuzzFeed News*. News17 January 2017 [Online] Available from: <https://www.buzzfeed.com/albertsamaha/blue-lies-matter> [Accessed: 22 January 2017].
- Scanlon, T. (1990) Promises and Practices. *Philosophy and Public Affairs*. 19 (3), 199–226.
- Shklar, J.N. (1984) The Ambiguities of Betrayal. In: *Ordinary Vices*. Harvard University Press.
- Simpson, T.W. (2011) Robots, Trust and War. *Philosophy & Technology*. [Online] 24 (3), 325–337. Available from: doi:10.1007/s13347-011-0030-y.
- Smith, M.N. (2010) Reliance. *Noûs*. 44 (1), 135–157.
- Smolkin, D. (2008) Puzzles about trust. *The Southern Journal of Philosophy*. 46 (3), 431–449.
- Spiegel, A. (2010) *A Life Without Fear*. *Morning Edition*. NPR, 26 April 2010 [Online] Available from: <http://www.npr.org/templates/story/story.php?storyId=126224885> [Accessed: 30 March 2016].
- Steward, H. (2012) Actions as processes. *Philosophical Perspectives*. 26 (1), 373–388.
- Steward, H. (2013) Processes, Continuants, and Individuals. *Mind*. [Online] 122 (487), 781–812. Available from: doi:10.1093/mind/fzt080.
- Strawson, P.F. (1974) *Freedom and Resentment and Other Essays*. Methuen: London.
- Thomas, L. (1989) Trust and Survival: Securing a Vision of the Good Society. *Journal of Social Philosophy*. [Online] 20 (1–2), 34–41. Available from: doi:10.1111/j.1467-9833.1989.tb00005.x.
- Thomson, G. (1987) *Needs* International library of philosophy London ; New York, Routledge & Kegan Paul.
- Thomson, J.J. (1990) Giving One’s Word. In: *The Realm of Rights*. Harvard University Press.
- Tummolini, L. et al. (2012) A convention or (tacit) agreement betwixt us: on reliance and its normative consequences. *Synthese*. [Online] 190 (4), 585–618. Available from: doi:10.1007/s11229-012-0194-8.
- Vendler, Z. (1957) Verbs and Times. *Philosophical Review*. 66 (2), 143–160.
- Walker, M.U. (2006) *Moral repair: reconstructing moral relations after wrongdoing*. Cambridge, UK ; New York, Cambridge University Press.
- Wiggins, D. (1991) *Needs, values, truth: essays in the philosophy of value*. Aristotelian Society series 2nd ed. Oxford, Blackwell.

Wiggins, D. & Dermen, S. (1987) Needs, need, needing. *Journal of Medical Ethics*. [Online] 13 (2), 62–68. Available from: doi:10.1136/jme.13.2.62.

Wonderly, M.L. (2016) On being attached. *Philosophical Studies*. [Online] 173 (1), 223–242. Available from: doi:10.1007/s11098-015-0487-0.

Zagzebski, L.T. (2012) *Epistemic Authority: A Theory of Trust, Authority, and Autonomy in Belief*. [Online] Oxford University Press. Available from: <http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780199936472.001.0001/acprof-9780199936472> [Accessed: 1 September 2015].