Navigating Vagueness:
Rule-Following and The Scope of Trust

Eli Benjamin Israel

Abstract: In this paper, I address a fundamental challenge in the philosophy of trust: how to account for trustee discretion in scenarios that fall outside explicitly defined expectations. I argue that this challenge reveals vagueness as an inherent feature of trusting relationships, often leading to disagreements between trustors and trustees. To resolve this, I propose a novel account of trust grounded in rule-following, shifting the object of trust from particular actions to adherence to rules constitutive of relationships. By focusing on relationships and their constitutive rules, I provide here a framework for understanding and minimizing disagreements through communication and engaged practice. Furthermore, I argue this framework sheds light on the dynamics of trust-related reactive attitudes, in particular, distinguishing between betrayal and mere disappointment.

Keywords: Trust, Trustworthiness, Rule-following, Vagueness, Interpersonal Relationships, Norms.

1. Introduction

Let's begin by considering the following case:

> *The Friend-Caregiver*: You trust your close friend Mel to care for your three-year-old
>
> son while you're away on an urgent business trip. Your son requires precise
>
> medication administration, so you provide Mel with detailed instructions, including
>
> the exact dosage she should give him each day. However, shortly after you leave,
>
> there's an unexpected shortage of the medication. Mel finds out she has only half of

the amount your son will need for the duration of your trip. Neither you nor the doctor can be reached, so Mel must decide what to do.

It may be hard for one to grasp the scope of your trust in such a scenario. Although you trusted Mel to administer your son's medication in a specific way, trust seems to allow some malleability that can't be reduced to the particular action or task at hand. Mel has some discretion as to how to fulfill her responsibilities toward you, and some courses of action may constitute a breach of trust while others would reasonably fulfill it.

This case illustrates the kind of discretion we often find in trusting relationships, concerning how trustees fulfill their responsibilities toward trustors by exercising their agency and choosing alternative appropriate ways to act without breaking trust. As Annette Baier writes:

> When we are trusted, we are relied upon to realize what it is for whose care we have some discretionary responsibility, and normal people can pick up the cues that indicate the limits of what is entrusted… This discretionary power will of course be limited by the limits of what is entrusted and usually by some other constraints. (Baier 1986, 236–37)

Baier is right that trusting relationships often encompass some discretion on the trustee's part.[1,2] Nevertheless, she is deliberately vague about the extent of these discretionary powers (again, because it is hard to grasp the scope of trust). She argues that there are some general boundaries about what is entrusted, which we can assume are mostly set by or with the trustor (although some may be externally imposed, such as social expectations, general moral obligations, etc.), but she takes it to be given to the trustee's discretion to judge how to fulfill

---

[1] Under certain circumstances, the friend might be able to exercise discretion beyond the mere knowledge of one's affairs, being trusted to act on the relevant worry, even without consulting me.
[2] Throughout the paper I use "trusting relationships" to denote different instances of trust, but this is not to assume a relationship characterized by trust, like the ones we have with family members, friends and romantic partners.

their responsibilities within those boundaries. This vagueness, which is inherent to trusting relationships, usually goes unnoticed and without significant bumps. However, there can be borderline cases where it becomes more difficult for trustors and trustees to know what counts as fulfillment or a break of trust. This can occur either because the agents have difficulties in picking up the relevant cues that limit the scope of trust (as in cases of neurodivergent individuals, or when interacting with people from a different culture, etc.) or because the scenario in which trustees find themselves performing changes so drastically, in ways that couldn't be imagined. Such cases illustrate problems at intrapersonal and interpersonal levels.

On an intrapersonal level, vagueness affects trustors and trustees differently. When trustors are unsure whether their trust has been fulfilled, their ability to appropriately respond and discern when it's suitable to adopt reactive attitudes characteristic of trust—such as feeling gratitude when trust is fulfilled, or betrayal and personal disappointment when it is breached—is challenged. Trustees, on the other hand, may find themselves uncertain about what is required from them as trustworthy agents, not knowing how to act to appropriately respond to the trust placed in them.[3] Should they search for alternative ways to address the trustor's needs or limit themselves to the expectations that were originally set? If the former, to what extent should they do so? How much leeway do they have? Ultimately, where is the line between what they're committed to within a trusting relationship and what they may choose to do, but only as a supererogatory act?

On an interpersonal level, we find that trustors and trustees often disagree about the appropriate expectations they may have of each other (Jones 2017, 104; Hawley 2019, 77;

---

[3] This aspect of trust is thoroughly discussed within trust-responsive theories of trust, such as Faulkner (2011; 2017); Jones (2012; 2017; 2019); and McGeer and Pettit (2017). I am not, however, committed (neither do I have to be, for the sake of this paper) to responsiveness as a fundamental aspect of trust to say that trustees may strive for ways of being responsive to the trust of others and struggle when they face challenges in doing so.

Stewart 2024, 63). This interpersonal manifestation of the problem is particularly common, as people are much quicker in identifying problems in others' conduct rather than attribute it to their own epistemic uncertainties. Such disagreements are not only unfortunate, but they may deeply strain our relationships (Jones 2017, 104).[4]

These problems of vagueness in trusting relationships reveal significant limitations in our current understanding of trust, as the existing accounts fail to properly account for the scope of trust, as well as to explain why it often seems vague to us, especially in discretionary situations.[5] In this paper, I address this lacuna by proposing a novel account of trust. I argue that the scope of trust is not determined by particular actions or even a domain of action, but by the rules constitutive of the trusting relationship. With this relational account of trust, we may not only understand vagueness but also find ways to address and navigate it, as well as minimize disagreements with our fellow trustors and trustees about the scope of trust.

My argument goes as follows. I start, in Section 2, by briefly clarifying why trust, rather than mere reliance, is the appropriate framework for addressing discretionary situations. In section 3, I further explain the problems of vagueness regarding the scope of trust and the disagreements that may arise from it. In section 4, I discuss why standard accounts of trust (action-based and domain-based) fall short of addressing these problems. In section 5, I introduce my novel account of trust, arguing that to trust someone is to rely on them to follow the rules constitutive of our relationships with them. In section 6, I demonstrate how my account accommodates discretion, explains the vagueness inherent to trust, and provides us with tools to address it and minimize disagreements through communication and practice. Finally, in section 7, I offer some conclusive remarks on how this framework illuminates the

---

[4] The problem I aim to address doesn't lie in the trustee's discretionary powers *per se*. Baier rightly points out that trustee discretion and vulnerability to their agential decisions is characteristic of trusting relationships and essential for achieving the levels of cooperation we strive for in such relations (1986, 239). However, what is risky is the vagueness or apparent lack of a practical principle to guide the trustee when discretion is necessary. Namely, a way for the trustee to know what actions are included within the scope of trust.

[5] More on this in section 4.1.

naturof trust-related reactive attitudes, particularly helping us distinguish between cases that warrant feelings of betrayal and those that merit mere disappointment.

2. A Brief Clarification

Before we begin, it might be helpful to say a few words about why trust — rather than mere reliance — is the correct framework to understand this vagueness about the expectations we have from each other, particularly in the context of practical discretion as situations unfold in unexpected ways. This step is important because reliance is at least part of what is happening in the Friend-Caregiver case. You are able to go on your business trip because Mel's taking care of your son, and so you act on the basis of it.[6] This way, reliance is a more basic attitude by which you consider Mel's actions within your own practical reason.

Reliance consists of a cognitive attitude that guides one's reasoning, such that "relying on p involves a disposition to, among other things, deliberate on the basis of p, plan on the basis of p, act on the basis of p, and draw conclusions from p, provided the relevant conditions are satisfied." (2014, 166). Such guidance is sensitive to the pragmatic considerations relevant to what one is trying to achieve (Alonso 2014, 169; Marušić 2017, 3; Holton 1994). Importantly, reliance is a reflexive attitude one owns for the sake of reason guidance. Similarly to one's attitude toward a self-formed belief (that is, one that you didn't form by means of trusted testimony) that proves itself to be false, when reliance proves wrong, failing to achieve one's desired ends, one can only address themselves for making a bad judgment and failing to assess the situation properly.

---

[6] In this brief and preliminary section of the paper, I'm focusing mostly on one aspect of the distinction between trust and reliance: the dutiful nature of trust and the accountability that follows it. For a comprehensive overview of the literature on the trust-reliance distinction, see, Goldberg (2020). For recent opposing views, see Marušić (2017) and Thompson (2017).

Trust, on the other hand, shares the basic structure of reliance and is also directed at practical ends, but adds to it a normative component attributed to the trustee. Therefore, while we can rely on both objects and agents, trust is exclusively agential. What it is exactly in the trustee that we regard in trust — their good will (Baier 1986; Jones 1996), commitment (Hawley 2019; 2014), stance as morally responsible agents (Holton 1994), or relational rule-following capacities (as I argue in this paper) — is a matter of dispute. However, it is fairly uncontroversial that trust is primarily directed at the trustee and their normative duty to act in certain ways.[7] When we merely rely on X to φ, we act on the basis of φ, but when we trust X to φ, we also hold that X owes us φ and is accountable for failing to φ.

This distinction becomes important in the cases I'm interested in, as there are not only expectations that can be misplaced or practically wrong judgments. Your relationship to Mel's conduct with your son is not one of mere reliance, as it does not end with you traveling based on that someone is taking care of your son. Rather, it primarily involves Mel's duty to take care of your son after she took upon herself such commitment. You regard her as owing you this care and expect it to be carried out properly, while she, in turn, considers herself committed to you in this way. Furthermore, you both acknowledge her accountability for deliberate failures in caring for your son, rendering it more than a wrong judgment on your part. Thus, the challenges of navigating discretionary scenarios such as this, as well as the potential for disagreement, stem from this vagueness about the scope of this normative component, and thus, trust is the appropriate framework for addressing these issues.

---

[7] My use of normativity here is rather thin, as I want to avoid any necessary moral dimensions to trust, in order to accommodate trusting relationships that are morally neutral, or even morally corrupt. An example of the latter comes from Hawley, mentioning instances of "honor among thieves" (Hawley 2012, 4).

3. Vagueness and Disagreement in Trust

In this section, I further clarify the problems of vagueness regarding the scope of trust and the disagreements that may arise from it. As I highlighted in the introduction, cases such as the Friend-Caregiver, where the trustee must use their discretion, show us that it's unclear what is the scope of trust and what it means to act in fulfillment of trust. If Mel was trusted to administer the child's medication as prescribed, this attitude of trust assumes, at least, that Mel *can* administer the child's medication as prescribed. However, given the medication shortage, it's not up to Mel whether to successfully fulfill this task, and we supposedly could consider trust moot in this scenario, as it often happens when it's not up to the trustee anymore to fulfill the task.

For example, imagine you trust me to buy us tickets for a concert, but when I got to the box office, I found out that they were sold out. In such a scenario, it is not up to me anymore to fulfill the task, and so the fact that you trust me to buy the tickets becomes moot — my failure to buy the tickets can't be considered a break of trust once we both realize the entrusted task is not up to me anymore.[8] However, we would judge the situation much differently if, going back to the Friend-Caregiver case, Mel decides that the medication shortage exempts her from administering any medication to your son. You may reasonably consider this a break of trust, and feel disappointed or perhaps even betrayed. This is because, as Baier rightly says, the trustee has "discretionary responsibility" (1986, 236). That is, they are normatively *required* (trust-wise, as opposed to other ways in which such obligation could

---

[8] Of course, we can fill in the details to make this completely wrong if, for example, you know there is another box office just down the street. However, considering this is not the case and there is no way of buying the tickets, or no reasonable such way — say the only way is through a scalper at exorbitant rates — my failure to obtain tickets couldn't be considered a breach of trust, even though I failed to do something you trusted me to do. Furthermore, I think it is reasonable to say that depending on the nature of our relationship or your level of excitement about the concert, I may or may not need to check upon other options to fulfill what is trust-wise required from me.

come about, such as on moral grounds) to use discretion to find a way to address the child's medical needs.[9]

Accounting for the required alternative action in the Friend-Caregiver case as the way of fulfilling trust is difficult if it's not what Mel was entrusted with. In other words, if discretionary scenarios were exceptionary, they could not be considered ways of fulfilling trust. Trust, as established in the previous section, is generally understood as a form of reliance, and so its fulfillment requires conformity between what the trustor initially relied on and the eventual outcome. For instance, when rock climbing, if I rely on a rope to support my weight, my reliance is only successful if the rope indeed carries my weight.[10] Similarly, if I trust B to φ, my trust is only successful if B (the trustee) actually φ-ies.[11] Thus, the trustor can't be said to trust someone if they do not guide their reasoning based on the possible discretionary scenarios as well. For example, if you trust Mel merely to administer your son's medication as prescribed, administrating it in an alternative way given the shortage could not be something you rely on, and thus, could not be the way of fulfilling trust. It perhaps could be the thing to do, but not an act that fulfills trust.[12]

Yet, I argue, that's exactly what it is. Administering the medication in some way is entailed by the trusting relationship you hold with Mel. Therefore, we can't say that the object of trust in this case (what Mel was trusted to do) is the specific administration of the medication as prescribed, but it must be broader so that an alternative course of action may fulfill the

---

[9] This is not to say that there are no other moral grounds for Mel's duty to help the child, and these may come hand in hand with the normative responsibility that comes from trust. What I highlight in this is that Mel's failure to use discretion in this case wouldn't only be immoral but would also be a breach of trust, of the kind that could justify you in feeling betrayed by Mel.

[10] This example is from Holton (1994) — also reproduced in Marušić (2017).

[11] Most accounts of trust follow this schema, viewing trust as a three-place relation, contrasted with two-place accounts, which view trust's basic structure as "A trusts B." The distinction between three- and two-place trust was first introduced by Horsburgh (1960). However, it gained prominence later in accounts such as those of Baier (1986); Holton (1994); Jones (1996); Hardin (2002); Hieronymi (2008); and Hawley (2014). For recent opposed views, see Domenicucci and Holton (2017) and Faulkner (2015).For the purposes of this paper, I bracket this debate, adopting the prevalent three-place view without further defense.

[12] The distinction could neither be waived by requiring the fulfillment of the task to the best of one's abilities. In the tickets' case, if there's only one ticket available, buying only one doesn't seem to make things any better.

relation. Although you asked Mel to administer the medication as prescribed, she is required to meet some more general expectation, say, that she will take care of your son's medical needs according to the circumstances while you're away."[13] This is what I mean by the scope of trust — whatever objects would count for the fulfillment of trust in a given trusting relationship — and that, as has already been noted, is quite blurry and vague.[14]

There is a fundamental difference between the concert tickets and the Friend-Caregiver cases. Yet, merely looking at the relation of trust as it is spelled out in these cases—trusting another to buy concert tickets or administer a child's medication per specific instructions, respectively — does not provide us with an account for the broader expectations that may arise when circumstances change. It remains vague what it is about these scenarios that shape their particular scope of trust.

Now, some vagueness is inherent to trust—this is quite uncontroversial.[15] Recently, Elizabeth Stewart made this point salient in her discussion about disagreements over domains of trust. Stewart points out that we are simply unable to anticipate every scenario or expectation in trusting relationships. Trustors can't foresee all potential scenarios that may unfold to their trustees, and so they can't make explicit all their relevant expectations. Likewise, trustees can't anticipate how things may go, and so they can't account for all possible ways in which they might act within a trusting relationship (2024, 74–75). This, Stewart argues, puts trustees in an "awkward position," where they must make practical decisions with unclear boundaries, and leads to different forms of disagreements and "mismatches" in trusting relationships. We may understand this awkward position as the vagueness we find about the scope of trust. Without clear expectations for the changing circumstances that a trusting relationship may

---

[13] Stewart (2024, 74)refers to this as the domain of trust.
[14] I purposely avoid reducing objects of trust to mere actions, as we may think of cases where our expectations from the trustee aim at their beliefs, emotions, character traits, etc.
[15] This has been widely emphasized in the literature, from Baier (1986), who laid down the basic intuitions we are still working with on the nature of trust, to recent works where this has been explicitly flagged, such as in Holton (2023); Stout (2022); Stewart (2024).

face, both trustors and trustees may have different views on whether discretion should be exercised, to what extent, and how such discretion should be applied.

Baier rightly notes that "normal people can pick up the cues that indicate the limits of what is entrusted" (1986, 236). This is probably true for most standard relationships and standard ways in which a situation could unfold. Not only because of the ability of the agents involved to pick up the relevant cues, but also because of how social conventions develop historically to characterize what we can expect of each other.[16] It would be almost absurd if Mel did not understand that she must do something to continue assisting your son with his medication. In the same way, in a society where monogamy is the prevalent norm, it would be absurd if your new partner would take it to be vague whether your relationship is monogamous or not just because you never talked about it. You may very reasonably trust them to not engage in other romantic/sexual relationships based on implicit norms you share as members of a certain society.

Nevertheless, genuine vagueness, and so genuine disagreement, is still to be found. Consider for example the following example. Suppose you consider me a good friend of yours and you invite me to your birthday party. If we are indeed good friends, then one could say I am not truly free to turn down an invitation to your birthday party without a very good reason.[17] If we are indeed good friends, you can reasonably trust me to come, and feel let down if I don't. However, much like trust, many of our relationships are vague, and while you may consider me a good friend I might take you more as an acquaintance, feeling no obligation to come to your party. This mismatch about what is our relationship causes us to perceive the scope of

---

[16] I'm not entirely comfortable with Baier here. When she relies on what "normal people" are usually capable of doing, not only she waives the issues that may arise from vagueness on the grounds of common sense, but also excludes appropriate treatment of the issues that may arise to neurodivergent people, who may struggle with interpreting social cues or implicit expectations. A more inclusive account, and what I believe my account ultimately does, would allow us to address the vagueness in each particular trusting relationships.
[17] I borrow this example from Hawley (2019, 77).

trust relevant to our relationship differently.[18] Furthermore, there may also be genuine vagueness, and so genuine disagreement, about how to correctly assess the situation and not only act but act in the right way. Suppose Mel understands her discretionary responsibility and that she must find an alternative way to administer the medication. She thinks of how she should do that and figures she has only two options — giving your son half the prescribed dosage each day to stretch the medication until you return, or giving him the prescribed dosage until it runs out, leaving him unmedicated for a few days.

Trust-wise, those options are not equally acceptable, although both attend to what was entrusted. You're not allowed to use your discretion in whichever way you see fit, or flip a coin to decide between the options. When we trust someone, we trust them to do well with what we have entrusted to them. Likewise, holding them trustworthy implies seeing them as capable of making the right decision within the scope of trust. We often describe trusting relationships as taking the other person to act as we would under the same circumstances, and we may hold reactive attitudes that are characteristic of breaches of trust when they fail to act as we judge right.

In conclusion, we often find vagueness about or within the scope of trust, and that leads to disagreements that may strain trusting relationships. The challenge, so, is to find a way to understand trust that can properly account for its scope, and at the same time account to its scope properly, but help us understand why it is so difficult to determine it. This will hopefully help us address vagueness and minimize disagreement.[19]

---

[18] I take that the status of our relationship, and so the commitments I hold to you, are not solely dependent on my judgment and I might not see myself obliged to come when I actually am. In other words, more details could be filled in, so I want to make clear — this is a case of an ambiguous relationship, and we can find good reasons for your assessment as well as mine.

[19] Thus, my view is that disagreements/mismatches are mere symptoms of vagueness, and so, are a secondary issue.

4. Why Current Accounts of Trust Will Not Do?

The shortcomings of the prevailing accounts in addressing the problems of vagueness mostly concern how they conceptualize the objects of trust (i.e., what the trustor trusts their trustee with). As highlighted in 3.1., our difficulty in understanding discretionary situations in trust involves the apparent mismatch between what one is relying on to happen when they trust another and the discretionary action that fulfills their trust. The action relied on represents the object of trust. Such objects are accounted for under most accounts as a fundamental component of trust.[20]

4.1. Challenges to the Action-Based Model

Following Baier (1986), most accounts understand trust in the form of "A trusts B to φ," where φ stands for the performance of a particular action or task.[21] I call this the action-based model.[22] As Jason D'Cruz notes, this model is typically assumed rather than argued for, likely due to Baier's influence (2018, 241). Thus, we should not be too surprised if it does not stand the test of time as we proceed with exploring new challenges about trust.

One way to argue for this model is by following the parallel between trust and reliance: since reliance pertains to specific actions, trust, as a form of reliance, is assumed to operate similarly. However, this approach may only account for more general forms of trust in an

---

[20] I purposely focus here only on three-place accounts, that is, accounts that hold the basic structure of trust to be schematically represented as "A trusts B to C," as we have established that for successful reliance within the attitude of trust we must have something of practical relevance (i.e., an action or state of affairs that is relied upon) at the other end. This is also the most prevalent position in the literature. Contrasting views, however, have gained traction in recent years, partly due to concerns similar to those I advance in this paper. For example, see Faulkner's (2015) defense of a one-place view, regarding trust as a non-directed attitude or capacity of the trusting person; and two-place views, as defended by Domenicucci and Holton (2017), that emphasize the relationship between trustor and trustee, allowing room for discretionary power and vulnerability by underdetermination of trust's objects.

[21] Baier is not exactly precise about that, and at some points she alternates between performance of actions and care for items, which she refrains from providing a more clear definition of. For the sake of simplicity, I take her concern for entrusted actions to primarily define her view, as it also represents how she is generally understood in this literature.

[22] See, for example, Hardin (2002); Baier (1986); and Hawley (2014), who offer prominent action-based account and have made this the prevalent position in the literature.

aggregative manner, viewing them as compilations of numerous action-specific instances. This perspective would either treat discretionary situations as exceptional or argue for the implicit nature of many trusting relations to grasp unfolding expectations and discretionary scenarios (Hawley 2019, 77).[23] For example, we might say that you trust Mel to administer medicine as prescribed, but this act of trust (to perform a specific action) comes with several implicit attendant acts. If you trust Mel to administer injections, you implicitly trust her to do so safely, avoid injecting dangerous air bubbles, call an ambulance if your son has an adverse reaction, and also to act sensibly if the medicine is unavailable.[24]

While this suggestion seemingly expands trust to encompass discretionary situations, it requires conceiving of trust as an exhaustive list of commitments the trustee must undertake. This approach is problematic for two reasons. First, it fails to capture the phenomenology of trust—we don't typically conceive of particular instances of trust as merely representative of numerous contingent commitments. Second, if there is discretion indeed, an action-based account must accommodate not only hypothetical scenarios but also different, sometimes even contrary, actions that the trustee might adopt while still fulfilling the trust. Therefore, such a model has difficulty in maintaining coherence about what the trustor relies upon while accommodating the numerous and varies implicit expectations.

An alternative way of conceiving action-based trust, though not explicitly defended in the literature to my knowledge, articulates action at a more general level. For instance, one might say that you trust Mel to nurse your son or to look after his health care. According to this view, saying that you trust Mel to administer the medicine as prescribed constitutes an overspecification of your trust. This version of the view is appealing in that it acknowledges the problems inherent to the action-based accounts and leans closer toward the domain-based

---

[23] For potential issues of accounting for general trust in this way, see Israel (in review).
[24] I thank an anonymous reviewer for suggesting this alternative version of the action-based account in the Friend-Caregiver case.

model we will discuss shortly. Nevertheless, I contend that it is incapable of providing a thorough account, as trust requires us to be able to ultimately rely on specific actions, as well as hold trustees accountable for specific actions that may fall under the general action but fail to fulfill our trust. For example, suppose I trusted Mel to look after my son's health care, but Mel, having read lately about alternative medicine, decides instead to treat my son with new homeopathic drops. In this case, she has clearly violated my trust even if she indeed looked after his health care needs (assuming the drops worked).[25]

4.2. Challenges to the Domain-Based Model

Another model regarding the nature of object of trust defines trust schematically as "A trusts B in D," where D is a particular domain of interaction (D'Cruz 2018; Jones 1996, 4; 2019, 958). I call this the domain-based model. While its proponents argue for its stronger emphasis on the agent (the trustee) rather than on actions (Jones 2019, 958), the domain of trust is still reducible to particular actions (and attitudes) that a trustee shows competence and cares for (McGeer 2008, 250; Stewart 2024, 65). This account is certainly more appealing to us, as it acknowledges the vague scope of trusting relationships and addresses trustee discretion as contained within the domain. This is supported by the fact that we can trust and distrust the same person within different domains of action or within certain practical limitations (D'Cruz 2018, 246).

This model has clear advantages in accounting for the varying scope of trust and the inclusion of implicit expectation. As for the first, unlike actions, trust domains can vary significantly in their breadth of expectations. Stewart argues that overly narrow domains can make trust impossible to satisfy, while overly wide domains leave trustors vulnerable to exploitation (2024, 82). For example, a narrow trust domain in a professional relationship might focus

---

[25] I thank an anonymous reviewer for pointing out this possible conjecture of the view.

solely on specific job-related tasks, potentially overlooking important aspects of collaboration. On the other hand, an overly broad trust domain in a personal relationship might leave one vulnerable to manipulation or disappointment.

An additional important aspect that approaches the situations that interest us is that domains of trust may also include expectations that are not explicitly recognized by either the trustor or the trustee until a specific situation arises, but are within the types of action and arrays of skills encompassed by the relevant domain (D'Cruz 2018, 245; Stewart 2024, 74). This gray area embedded in the idea of domains of trust is appealing as it allows for discretion, as well as for the vagueness we attempt to identify here, and as situations unfold unexpectedly, implicit expectations can lead to disagreements about whether the trustee's action falls within the domain.

Nevertheless, by considering domains of trust to be sets of actions expected by trustors, the domain-based model faces two main problems. First, it provides trustees with no way of following up with the exhaustive list of actions that may be comprised within the trust domain, coming back to the immediate impasse we found in action-based accounts.[26] In this way, although domain-based accounts acknowledge vagueness, taking trust to be posited across a domain, they don't offer both trustors and trustees pragmatic means of addressing it and considering whether a particular course of action fits within the domain.

Second, when defining the scope of trust solely by the trustor's expectations, trustees have no say in what they may be entrusted to do, particularly in discretionary situations. What falls within the scope of their commitment is not accounted for as their own decision. The scope of trust stands, but derives solely from the trustor and what they consider to be within the domain of trust. This undermines, at least partially, the trustee's agency in the trusting

---

[26] See, section 3.1.

relationship, as they can only express their approval or disapproval of being trusted with the relevant domain of interaction, but have no way of shaping it. However, even in such case, it has been argued that it is to the trustor to recognize that the trustee is not interested in being trusted with a particular domain, and choose to respect and defer to their preference (D'Cruz 2019, 935).

5. Trust as Relationship Rule-Following

In this section, I offer a novel framework of trust. I introduce here a relational framework of trust based on the notion of rule-following, according to which the scope of trust is delineated by the rules constitutive of trusting relationships, setting the grounds for tackling the problems of vagueness in the following section.

5.1 Rules in Trusting Relationships

I propose an account of trust that aims to overcome the limitations of action-based and domain-based models. This account regards the objects of trust as the rules constitutive of a trusting relationship. Specifically, to trust someone is to rely on their rule-following capacities, to the extent that those rules define the relationship between trustor and trustee. To act in fulfillment of trust, in turn, is the quality of aptly recognizing the rules constitutive of the relationship and having those guiding the trustee's actions.

Understanding trust as an attitude that occurs within a framework of interpersonal and social relationships may seem like the intuitive take, but it has not been sufficiently considered in the literature. Trust is often thought of as a mental attitude grounded on whether others will follow through on their commitments to us, but that doesn't presuppose any relationship with

our counterparts.[27] This seems pretty straightforwardly false, not only because we may better deal with problems like the ones posed by vagueness when adding this foundational aspect to trust as I will argue here, but also because this way of viewing trust just doesn't resonate with our most basic intuitions about this attitude: Trust doesn't happen in a vacuum. It requires an interpersonal or social context where agents connect and where trust may arise.[28]

I argue we may consider such contexts as relationships. Accordingly, when we trust someone, we necessarily hold some relationship with them. Now, I don't take relationships here in the strictest sense, denoting only the kind of thick connections we have with family, friends, and loved ones. Our interdependence extends beyond these intimate spheres to include thin connections as well, like the ones we hold with colleagues, fellow citizens, or members of the same community—each carrying its own set of special duties and responsibilities. (Scheffler 2002, 97).[29]

---

[27] The language of commitment primarily comes from Hawley's (2014; 2019) account of trust, which is one I particularly like for its generality, but it is not essential to my argument here — In the same way, you may think of this as a goodwill (Baier, 1986) or a stance that drives the trustee's actions toward us (Holton, 1994).

[28] An exception to the tendency in the literature of overlooking this relation context can be found in Thomas Simpson's genealogical account of trust (2012). Simpson acknowledges the social context necessary for explaining trust, and defines the latter as reliance on one's "freely cooperative behaviour" (2012, 558). Nevertheless, his account does not provide us with a unified mechanism to understand the variety of trusting relations we engage in throughout our social lives. Although Simpson's view is comfortable with "plural forms of trust" (2012, 561), I believe my account is capable of having the cake and eating it too. Namely, by locating the objects of trust in the rules constitutive of our relationships, we have a unified mechanism of trust based on the idea of rule-following, while also accommodating plurality in content through the much more intuitive idea that we must recognize plural forms of relationships. Similarly, Domenicucci and Holton (2017) acknowledge the centrality of relationships to trust. They contend that the extension of trust is primarily determined by the nature and quality of the relationship between the trusting parties. According to this view, trust is a "cluster concept" that involves a host of attitudes and expectations that are shaped by the relationship itself. This seems like a more accurate way to characterize trust and contextualize the interpersonal and social contexts where this attitude arises (2017, 154–56). However, Domenicucci and Holton infer from the relational determination of trust that it is fundamentally a two-place relation, represented as "A trusts B."# This inference, I argue, is unwarranted. A two-place formulation not only fails to account for trust as reliance but also inadequately represents the mutual nature of trusting relationships and the specific determinants of attitudes and expectations that arise from these relationships.

[29] I borrow here some useful terminology from Avishai Margalit (2004; 2017), who highlights the distinction between thick relationships, the kind we have with family, friends, and lovers; and thin relationships, which we have with people we don't know. I believe it's safe to say we use the language of trust for both, and it seems appropriate to do so. What makes a relationship thick or thin is how meaningful it is, but this distinction also reflects on the nature of the rules constitutive of those relationships, as thicker relationships usually require more from their participants (think, for example, how the friendship with Mel makes her more committed to your son's health than if she were merely a hired caregiver).

Relationships, to the extent that they can accommodate trust, have a normative component, and some norms or rules of conduct apply — ways in which we are committed to acting with each other as part of the relationship.[30] This aligns with Raz, who defines relationships as "constituted by norms which determine what conduct is appropriate between people in the relationships, between, for example, parents and children, as well as towards other people's children, etc" (2009, 316). These norms or rules (we may use those interchangeably in this context) aren't merely part of relationships but are constitutive of them, as Raz argues that knowledge of the norms is knowledge of the nature of the relationship, which guides our behavior within it (2009, 317).[31]

It's worth clarifying what I mean here by rules constitutive of the relationship, examining how the same idea works, considering also its application to other domains, like games. Reiland argues that constitutive rules function by regulating antecedently existing actions while making possible new forms of action within the constituted practice (2020, 142–43). In relationships, these rules regulate basic interpersonal interactions and enable actions distinct to the relationship. Just as moving chess pieces following the rules of chess is what makes the game of chess, following the rules constitutive of the relationship is what properly makes the relationship, defining its nature and the possibilities within it.[32] This reinforces what we found in Raz's account: relationships aren't reducible to behavioral patterns but are fundamentally normative in character. This parallel extends to how relationship rules, like

---

[30] Here, again, the normative component should be understood very minimally, merely in terms of correctness of one's actions relative to the rule. It may include moral normativity when the relevant relationship is such that it is founded on moral principles, but it may also include morally neutral or even repugnant rules. Think, for example, of the normativity found in the relationship between chess players where they are required to follow the rules of chess (Boghossian 2012, 29), or the normativity found in the relationship between a mafia don and his consiglieri, where they can trust each other with horrible things that are constitutive of the relationship between mafia members, what we often call "honor among thieves"(Hawley 2012, 4).

[31] Elsewhere, Raz further argues that relationships are individualized by the norms/duties of the relationship (1989, 19), reinforcing their constitutive role.

[32] Under such Razian conception of relationships, those are individualized by norms rather than by individuals. This also means that if the norms are fundamentally changed, we arrive at a whole new relationship with its characteristic norms, even when the participants are the same. For an opposing view, strictly using the term to merely thick connections and individualizing those by the individuals and their particular history, see Kolodny (2003; 2010).

rules of games, determine the roles and duties of those who participate. When we consider relationships carrying special duties and responsibilities, we see these deriving directly from the rules that constitute the relationship.

5.2. Rule-Following in Trust

Trust, I claim, is when we rely on others to follow the rules constitutive of our trusting relationship. Having trustors and trustees as fellow participants who choose to be in relationship with one another, by means of these rules, this framework adds to trust its necessary normative component, distinguishing it from mere reliance.[33] Particular normative duties, as well as accountability, apply to those who participate in a trusting relationship.

Under this account, we must regard others as capable and willing to follow the relationship's constitutive rules.[34] We hold many different relationships in our lives—with our lovers, professors, fellow soccer players, siblings, bankers, friends, and more—but not all of those are trusting relationships, or can even justifiably accommodate trust. For trust to come about, one must take the trustee to be genuinely committed to the relationship, with whatever rules it entails.[35] As Reiland argues regarding rule-constituted actions generally, performing such actions requires more than mere conformity to rules—it requires accepting or endorsing them

---

[33] A way to illustrate this distinction is by considering relationship rules we're not part of. Imagine inviting a friend to an event, expecting them to bring their romantic partner. You're relying on their commitment, but you're not in a position of trust regarding their loyalty. If they bring a secret lover instead, your reliance is failed, but you haven't been betrayed – their partner has. You relied on an expectation, but the trust broken wasn't yours.

[34] Although I endorse a doxastic view of trust (the prevalent view; see, Keren 2020) and think our regard of the trustee is a form of belief in this case, I'm purposely ambiguous about this here, as I think that non-doxastic views (e.g. (Lahno 2020; Jones 1996; Holton 1994) can equally benefit from my framework in addressing problems of vagueness.

[35] In this context, a possible objection to my view that trust presupposes a relationship is the case of relying on police officers in a foreign country. While police officers are bound by rules that establish how they should treat civilians, this case is better understood as one of mere reliance rather than trust. What would be missing in such a case is grounds for trustworthiness - Although we may have grounds to say that there is a relationship individualized by norms between a foreign citizen and a police officer, I simply don't know whether police officers in that country understand such relationship and actually take their duties seriously and so, can be trusted, or are, for instance, deeply corrupt. In such a scenario, we may blame foreign police officers for failing their duties, but reactive attitudes like betrayal, which are associated with trust, would be implausible. I thank an anonymous reviewer for bringing up this case.

(2020, 149). This explains why trusting relationships require more than mere interaction or coordination: they require that we understand trustees as genuinely accepting the rules that constitute the relationship.

Now, the idea of rule-following is very rich philosophically, particularly within Wittgenstenian circles, and it describes a practice embedded in shared patterns of behavior, customs, and institutions that give meaning to our actions. Although traditionally tied to the private language argument, it can be more broadly applied if we understand it referring to practices of social character where we are normatively constrained and guided by general practical rules in the performance of particular actions.

One way to understand rules is as general and abstract directives aimed at types of actions or state of affairs (Boghossian 2015, 6). By generality here I mean that it doesn't prescribe particular actions but merely guidance and constraints upon possible paths of action. Some generality of this kind is an important feature to the extent we are interested in the context of trust. Namely, it accommodates the distinction we are looking for, distancing from mere aggregation of particular actions available within the trusting relationship, and providing practical guidance while also allowing trustee discretion. However, as Reiland argues, generality of rules is not enough, as it lacks the normative component that is in effect through our interpersonal and social practices (2024, 779). Rules are contingently in force - they exist and apply because they are practically endorsed by participants. This is particularly relevant for understanding trusting relationships, where the rules that constitute them must be actively enacted or accepted by the parties as part of their decision to engage in a particular relationship.

Being irreducible to specific actions, a relational rule-based approach to trust avoids the problematicity found in action-based or domain-based models that attempt to specify the

actions one may take for acting in fulfillment of trust. It targets the idea that our conduct is normatively constrained by relevant rules, in the sense that there is some determination of what conforms with them. That is, relationship rules, when in force upon its participants, create a scope of permissible conduct that fulfills trust without determining particular applications. This explains the vagueness inherent to rule-following generally, and to the fulfillment of trust particularly. To follow a rule is to "act in accordance with it or [when] one's action is somehow otherwise a product of sensitivity to its demands" (Reiland 2024, 2).[36] As such, there are many ways in which following the rules may unfold, depending on one's grasp of the relevant rules and their ability to aptly act in accordance with them in different contexts.

When we apply this notion of trust to the Friend-Caregiver case, we see clearly how trust accommodates discretion within constraints. It would not be accurate to say you trust Mel simply to administer the medication as prescribed - that would miss how the rules constitutive of your relationship delineate the scope of permissible conduct. Perhaps it is a specification of what you expect from Mel under normal circumstances, or what you interpret would be a plausible application of those rules. Rather, you trust Mel to act within the scope established by the rules constitutive of your relationship, encompassing rules grounded in your friendship and those added by particular commitments, such as the parent-caregiver rules. Within this scope, what counts as appropriate action can vary with circumstances. When medication is sufficient, administering it as prescribed fulfills trust. However, in case of shortage, the way of following the rules constitutive of relationship changes: Mel might be required to administer the available dosage until it runs out, monitor your son's condition on the

---

[36] The emphasis on "tries" in this quote is mine. I claim that requiring mere attempt, rather than success, is important to create a distinction (an overlooked one, I may add) between two different ways in which a trustee may fail to act as expected in a trusting relationship. In this way, the sense of "rule-following" I use here is quite minimal, requiring only guidance by rules but not actual conformity of the actions (Boghossian 2012; Pettit 1990). It is possible to act in a trustworthy way while failing to conform (Reiland 2024, 2), and different reactions may follow in cases where one fails to conform or fails to be guided by the appropriate rules. I say more about this in section 5.

subsequent days, and potentially seek additional medication if it becomes available. As your friend, she might be expected to go beyond what a professional caregiver would do, such as going out of her way to get the medication needed for the remaining days. On the other hand, as your friend and not a professional caregiver, Mel might be expected to take your son to a hospital under complications where a professional caregiver would be trusted to take certain medical actions themselves.

Understanding what are the rules constitutive of a trusting relationship and how they should be applied is key to the fulfillment of trust. As rules of the relationship itself, rather than rules unilaterally established by either side, they should be assessed irrespective of what either trustors or trustees merely *think* is constitutive of the relationship. The correctness conditions for Mel's actions are external to the agents involved, having to do with the factual nature of your trusting relationship. Thus, Mel's best intentions to fulfill trust are not enough, nor is your own perspective on how she should act in a given situation determinative of the scope of trust. What matters are the rules that actually constitute your relationship. Mel must correctly be guided by these rules, and not merely think she is doing so.

Fulfilling trust by rule-following is not subject to individual interpretation but by how the trusting relationship was established, be it socially or interpersonally, by the relevant trustors and trustees, as they may agree on rules that constitute their particular interaction locally. Although these normative expectations derive from our social environment and often have their starting point at pre-defined social scripts, they allow some significant flexibility. Raz argues that beyond the "stable frameworks" of what it is to be a friend, or a parent or a son, the participants of relationships negotiate and personalize the meaning of their relationships while maintaining their social intelligibility. A father might express care differently than cultural norms suggest, yet still be performing the role of father; friends might rarely meet in person but establish that their relationship is characterized by sending each other memes on

social media - all that while still engaging in what is a friendship. Such flexibility, according to Raz, is a feature of relationships. They are both socially structured and "individually explored and established," creating a form that is "vital to social life" (2009, 317).

Given the voluntary nature of trusting relationships, grounded on the trustee's decision to guide their conduct by particular rules and the often active role they have in establishing those rules, we overcome here another disadvantage found in the domain-based view: the trustee isn't passive to the terms of the trusting relationship and has an agential role in entering and establishing its conditions. The relevant rules are such that when one commits to when they engage in a relationship, and failing to adhere to them can be held against the trustee. Therefore, being the basis of the commitments people have with each other, relationships are the (only) fertile ground for attitudes of trust.

6. Addressing Vagueness and Minimizing Disagreement

Before we proceed to examine whether my account is ultimately up to the challenges posed in this paper, let's recapitulate what exactly we are looking for. An account of trust that addresses the challenges described in section 2 must accommodate trustee discretion, matching between what the trustor entrusts to the trustee and the actions that may fulfill their trust; (provide a framework for understanding what falls within the scope of trust; and help us navigating through vagueness and minimizing disagreement.

As I outlined my view in the previous section, we can see how my account addresses the first aspect quite immediately. By directing trust to the rules constitutive of the trusting relationship, we arrive at a notion of trust that can be fulfilled in various ways, while constrained by the rules they are committed to as part of as long as the trustee is guided by

those rules while being sensitive to the dynamic nature of things. Therefore, we may accommodate trustee discretion within the scope of trust: we shift our focus from the outcome to whether the trustee is being guided by the rules they are committed to as part of this relationship. Now it remains to show in this section how my account successfully explains vagueness and helps minimizing disagreement.

My account of trust is useful in targeting the sources of vagueness, so that we can address it more efficiently. When we trust someone, we trust them according to the rules constitutive of our relationship. Likewise, when we act in trustworthy ways, the same rules practically guide us in our actions. Therefore, vagueness does not lie in the attitude of trust, but it concerns the nature of the relationship and the normative expectations it presupposes, regardless of how each of the parties involved perceives it.

How trustors and trustees act within a trusting relationship derives from how they perceive their relationship. This, I argue, is the source of the inherent vagueness in trust: the gap between how one perceives the relationship they are in and what it actually is.[37] Consider, for example, the following scenario: Think of this romantic couple, Jamie and Kit, who have been together for a couple of weeks but haven't got to do "the talk." That is, they have never discussed their exact status concerning sexual exclusivity, emotional involvement, plans, boundaries, etc. Although they may already have a trusting relationship, chances are that they will find themselves in situations where vagueness about the scope under which they can trust each other will become salient, and so they might be uncertain about how far their discretion may go with one another. This is due to the vague nature of their relationship and possible gaps concerning how each one of them understands it.[38] Similarly, in the Friend-Caregiver

---

[37] Boghossian, in his treatment of rule-following, highlights the possibility of vagueness in rules, such as in any content, and how vagueness about rules becomes more salient in borderline cases (Boghossian 2015, 8). This resonates with what we've been saying about instances of vagueness in trust, which, all other things being equal, are also more prone to appear in extreme circumstances.

[38] This is not to say that we must verbally and explicitly address the nature of our relationship. Some relationships are reasonably understood without needing to say much. Moreover, by mere practice, that is living

case, if you and Mel haven't been friends for long enough or if that's one of her first times taking care of a child with special medical needs, there is a greater chance that she will find vagueness about how she should act in a discretionary scenario such as the one we portrayed, simply by the fact that the nature of her friendship with you or the nature of your parent-caregiver relationship are still unclear to her.

The way to address this vagueness and minimize disagreements is to get a better grasp of the relationship. The most immediate approach is having "the talk." That is, communicating with our counterparts about how we understand our relationship and interpret its rules so that we can arrive at some shared understanding of the basic principles we should follow.[39, 40] This method, however, is not conclusive, as explicit communication will always be subject to gaps in understanding and fail to achieve enough generality so that the relevant rules can be applied in all situations that may unfold.[41]

A more conclusive way of addressing vagueness about relationships and their rules is through a notion of relationship rule-following as an enduring practice, achieved simply by participating in various relationships throughout our lives. This solution is inspired by Wittgenstein (1953, 86–88, §198-202), who, acknowledging the limitations found in grasping

---

by the relationship, one may come to know what rules are constitutive of it. I say more about the importance of practice, and its actual advantage over communication, in the following page.

[39] Emphasizing the power of communication in an initial shared assessment of the relationship is important to address the problems I found with Baier's view in footnote 14. There, I pointed out that when Baier waives the issues that may arise from vagueness on the grounds of common sense, she excludes appropriate treatment of these issues for people who struggle with interpreting social cues and implicit expectations, such as neurodivergent individuals.

[40] Stewart also argues for the importance of communication in avoiding and resolving disagreement in trust (2024, 70, 84). On the other hand, the limitations of communication in facilitating trust have been already pointed out in this literature. For example, Faulkner (2011, 5–6) argues that trustors and trustees may have different interests in communication, and so conveying how one understands rules and trying to achieve consensus about the nature of the trusting relationship would not necessarily clarify the nature of the relationship, but could instead be grounds for uncooperative efforts by one of the parties involved.

[41] In addition to these limitations I find to communicative efforts for addressing vagueness, contemporary accounts of rule-following follow Wittgenstein in rejecting the effectiveness of communication based on what's known as the regress argument (Reiland 2024, 3; Brandom 1994, 20–21; Mcdowell 1992, 265–66), which I mention here briefly. This argument basically states that adding interpretation by means of explicit instructions, will not suffice to mitigate vagueness, as those instructions can also be interpreted in multiple ways. I do, however, think there is value in communication, if not for the complete mitigation of vagueness, at least as a first step that will enable proper practice and participation.

rules by communication, proposes a more intuitive approach to grasping and applying rules correctly . By intuitive, I mean that it emphasizes the practice over the intellectual and interpretative process of trying to grasp rules propositionally and converting them into action.[42]

The idea is that we should focus not only on spelling out the nature and terms of our relationships, but on active, continued engagement in trusting relationships governed by rules. Through repeated practice, trustors and trustees develop an understanding of how to apply rules in various contexts. This practical engagement allows us to develop skills and intuitions that guide us toward better application of rules in unexpected situations.[43] Having this practice inherent to social and interpersonal relationships we hold with others, we find in them a mechanism of recurrent feedback, correction, and reinforcement that makes us better "relationship rule-followers," and so, more trustworthy agents. Furthermore, one's participative role makes it so that by practice they shape the relationship themselves, providing others with feedback, correction, and reinforcement of their conduct within the trusting relationship.

The idea that we minimize disagreement through training and participation may come across as disappointing. It would be really nice to have a more straightforward way of understanding the relationships we are in and laying out its scope clearly. However, we must also acknowledge that one is not born trustworthy but becomes trustworthy. When we engage in trusting relationships — beginning with the relationship with our parents in childhood, through our first romantic and sexual relationships, to our mature relations in long-standing relationships as well as our thin relationships with co-citizens and co-workers — we can

---

[42] For further discussion, see Child (2011, 142–46).
[43] For an application of the idea that skills and intuitions facilitate trustworthy behavior, particularly in sexual relationships, see Israel (2024).

develop a better grasp of what those relationships are, what rules they entail, and how they should be applied.

The significance of training and participation in building trustworthiness resonates with a valid concern posed by Katherine Hawley. Hawley argues that since trustworthiness doesn't require us to positively take any commitments, given the inevitable uncertainty about which commitments we will be able to follow through, the safest option concerning trustworthiness would be to avoid any commitments (2019, 96). This, however, has significant costs, leading to "a lonely, unrewarding life, one which benefits neither the recipe-follower nor the people around her" (2019, 120). While I share this concern with Hawley, it follows from my account that avoiding commitments (and so avoiding relationships) is not simply undesirable, but it jeopardizes one's chance of becoming a trustworthy agent.

7. Final Remarks: Looking Ahead to Reactive Attitudes

The approach to disagreements in trusting relationships that I laid out in this paper acknowledges the power of communication in addressing cases of vagueness and somewhat minimizing disagreements. However, there are no silver bullets. The path to conclusively addressing these issues lies in the ongoing practice of relationships. We must view the fulfillment of trust as a continuous endeavor, linked to our efforts to actively participate in the various social and interpersonal relationships that shape our lives.

It is precisely because there are no silver bullets and some degree of vagueness in trust is inevitable that violations or perceived violations of trust (i.e., when a trustor mistakenly believes the trustee has failed in rule-following) are almost bound to occur, especially between individuals who are still developing their relationship skills and rule-following capacities. This is something we not only acknowledge but accept — trusting relationships always involve risk and vulnerability (Becker 1996; Baier 1986). The risk is not only that our

27

trustees will deliberately choose to act contrary to their commitments, but also that there may

be a mismatch because they failed to understand the rules of the relationship, or because we

failed to grasp the rules we are entitled to expect them to follow.[44] In other words, things

could go wrong in so many ways even when everyone is acting in good faith. Thus, the risk

and vulnerability we find and accept in trust take a particular shape when viewed in the

context of our relationships and their rules. In instances of unmet expectations, both trustors

and trustees may bear responsibility, and this may often happen inadvertently.

In describing the distinctive ways in which we are prone to react to unmet expectations,

Richard Holton writes, "When you trust someone to do something… you have a readiness to

feel betrayal should it be disappointed, and gratitude should it be upheld" (1994, 67).[45] Ther

is some true to this claim, but while betrayal and gratitude are indeed attitudes characteristic

of trusting relationships, most accounts don't tell us much about when they are appropriate,

and neither do they account for cases where unmet expectations amount only to

disappointment.[46]

I believe that the account of trust I lay out here can shed light on when these attitudes are

appropriate, particularly in differentiating between situations that warrant feeling betrayed

and those that merely justify disappointment. Furthermore, it reveals that reactive attitudes

such as betrayal are not merely responses to breaks of local trust, but to violations of the

---

[44] Risk and vulnerability are inherent to trusting relationships, in ways that attempts to mitigate them can undermine trust. For a related discussion, see Chae (2022) who discusses how such attempts through contingency plans damage the very trust they aim to protect.

[45] This view draws on Strawson's (1962) remarks on what he terms "the participant point of view." The notion that these attitudes are characteristic of trust became widely endorsed in the literature - see, for example, Hieronymi (2008, 216), Hawley (2014, 7), and Marušić (2017, 5).

[46] Two exceptions come from Collin O'Neil (2017), who argues that betrayal consists in the violation of a second-order obligation to vindicate the honor bestowed by an expectation of trustworthiness (2017, 88); and and Avishai Margalit (2017), who contends betrayal is a characteristic attitude of thick relationships. I find both accounts unsatisfactory. Regarding O'Neill, his accounts fails to grasp cases one seems to have proven themselves unworthy of trust, but it is the trustor who held the trustee to expectations that aren't within the relationship. Regarding Margalit, his view doesn't seem to distinguish cases where we are merely justified to feel disappointed, but not betrayed, in our thick relationships.Furthermore, their accounts do not provide us with tools for understanding justified disappointment.

relationship as a whole. This explains why certain breaches of trust feel so deeply personal and general, extending beyond the specific action or expectation that was violated. Consider, for example, a case of infidelity in a monogamous relationship. When someone cheats on their partner, it is not merely as a break of local trust regarding sexual exclusivity, but a violation of the relationship rules, and so it affects the relationship as a whole, warranting accusations such as "you betrayed *me*," or "you betrayed *our marriage vows*."

In this way, we may assess why some breaks of trust are more severe and personal than others. When a trustee violates a rule that is central to the relationship, it calls into question the entire basis of the trusting relationship. On the other hand, when a trustee makes an error in judgment while still attempting to adhere to the relationship's rules, it may lead to disappointment but not necessarily betrayal. Consider again the case of Mel, as she discovers the shortage in your son's medication. If you are a reasonable person, your reaction would likely be fundamentally different if she deliberately decided to stop medicating your son altogether due to the shortage, compared to if she had chosen the wrong alternative, giving him half the dosage. The difference stands even if the outcome to your son's health is the same (assume that an incomplete dosage is equivalent to no medication at all).

The reason for this is quite simple: You justifiably find it obvious that she could not take the shortage as a reason to completely stop your son's medication, and so you find her decision to be a betrayal of your relationship, which in turn, justifies you feeling betrayed. In the second case, you may recognize that she tried to fulfill her responsibility, but did it poorly, and so it seems like you are not justified in feeling betrayed.

We intuitively understand betrayal as a deliberate violation — an active decision not to follow the rules of the trusting relationship. Given that trustworthiness entails being guided by those rules, betrayal happens when trustees fail to live up to the trust placed in them, not

just by failing to do what is expected, but by failing to prove themselves worthy of being seen as trustworthy (O'Neil 2017, 88). In acts of betrayal, one shows disregard for the relationship rules, and so they are not worthy of our trust. This requires us to judge the unmet expectation to not be a result of vagueness. Were that the case, one would probably change their mind about the appropriateness of feeling betrayed.[47]

Feelings of betrayal are out of order when unmet expectations stem from vagueness, and this is where disappointment comes into play. While vagueness is inherent to trust, and it is hard to have it conclusively treated, we can't let trustees off the hook entirely when they fail to act guided by the rules due to vagueness. By entering into the trusting relationship and assuming a constitutive role in it, they hold normative responsibility.[48] This, in and of itself, allows for some normative resentment, even though less severe than what we find in instances where feelings of betrayal are warranted. Nevertheless, the unintentional character of their failure, and the possibility for correction and improvement by continuation of the trusting relationships, makes personal disappointment an appropriate reaction to such cases, and one that builds them up as better trustees.[49]

In conclusion, there is no clear-cut way of drawing the line where each attitude is more appropriate merely on the grounds of whether there is vagueness or not. We can think of variables that may make it more difficult to do so. For example, there are probably epistemic duties one has to grasp the relevant rules constitutive of a relationship, and those may turn an unintentional failure into a betrayal. These are questions I hope to address in future work. Nevertheless, the very fact that we can now make such distinctions and begin to explore the nuanced nature of trust-related reactive attitudes is a promising and distinguished feature of

---

[47] This understanding of betrayal closely aligns with conditions for blame.
[48] Otherwise, we would not be dealing with trust but with a case of mere agential reliance, that is, reliance on the actions of another person, but with no normative claim toward them.
[49] As I said in 4.1., the practice inherent to social and interpersonal relationships creates a mechanism of recurrent feedback, correction, and reinforcement that helps to shape better "relationship rule-followers."
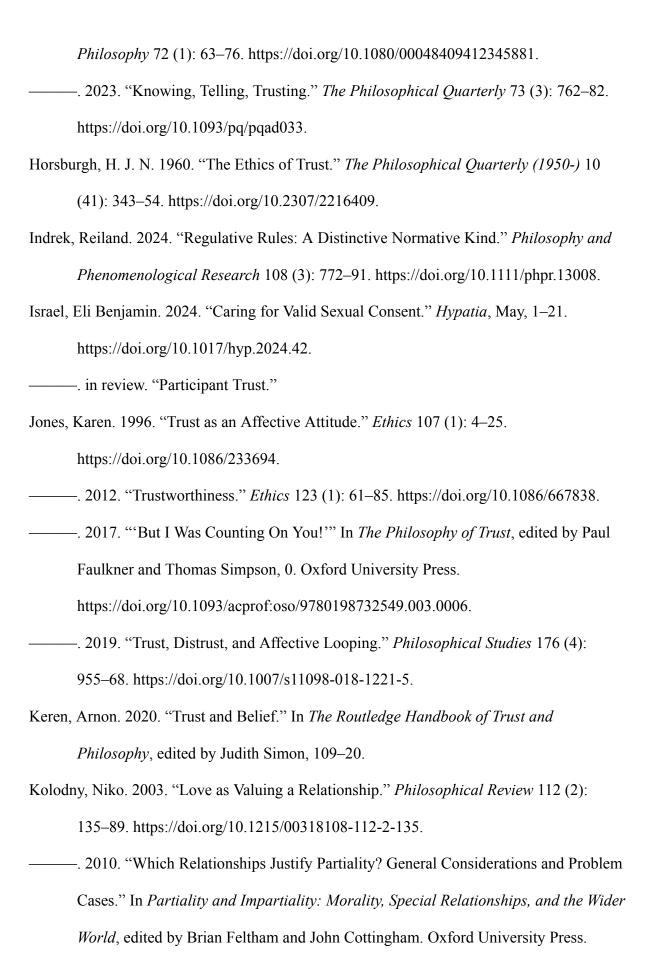
my account. That is, I believe it follows from it that taking trust as a "readiness" to feel

betrayed isn't enough, and we must say more about when it is fair and appropriate to adopt

such an attitude, and when we ought to insist on practical trust.

References

Alonso, Facundo M. 2014. "What Is Reliance?" *Canadian Journal of Philosophy* 44 (2):

163–83.

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

https://doi.org/10.1086/292745.

Becker, Lawrence C. 1996. "Trust as Noncognitive Security About Motives." *Ethics* 107 (1):

43–61. https://doi.org/10.1086/233696.

Boghossian, Paul. 2012. "Blind Rule-Following." In *Mind, Meaning, and Knowledge:*

Themes from the Philosophy of Crispin Wright, edited by Annalisa Coliva, 27–48.

Oxford University Press.

———. 2015. "Rules, Norms and Principles: A Conceptual Framework." In *Problems of*

Normativity, Rules and Rule-Following, edited by Michał Araszkiewicz, Paweł Banaś,

Tomasz Gizbert-Studnicki, and Krzysztof Płeszka, 3–11. Cham: Springer

International Publishing. https://doi.org/10.1007/978-3-319-09375-8_1.

Brandom, Robert. 1994. *Making It Explicit: Reasoning, Representing, and Discursive*

Commitment. Harvard University Press.

Chae, Lee-Ann. 2022. "Trust and Contingency Plans." *Canadian Journal of Philosophy* 52

(7): 689–99. https://doi.org/10.1017/can.2023.8.

Child, William. 2011. *Wittgenstein*. Routledge Philosophers. Abingdon, Oxon ; New York:

Routledge.

D'Cruz, Jason. 2018. "Trust within Limits." *International Journal of Philosophical Studies*,

March. https://www.tandfonline.com/doi/full/10.1080/09672559.2018.1450080.

D'Cruz, Jason. 2019. "Humble Trust." *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 176 (4): 933–53.

Domenicucci, Jacopo, and Richard Holton. 2017. "Trust as a Two-Place Relation." In *The Philosophy of Trust*, edited by Paul Faulkner and Thomas Simpson, 0. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198732549.003.0009.

Faulkner, Paul. 2011. *Knowledge on Trust*. 1st edition. Oxford, GB: Oxford University Press.

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

———. 2017. "The Problem of Trust." In *The Philosophy of Trust*, edited by Paul Faulkner and Thomas Simpson, 0. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198732549.003.0007.

Goldberg, Sanford C. 2020. "Trust and Reliance." In *The Routledge Handbook of Trust and Philosophy*. Routledge.

Hardin, Russell. 2002. *Trust and Trustworthiness*. Russell Sage Foundation. https://www.jstor.org/stable/10.7758/9781610442718.

Hawley, Katherine. 2012. "What Are Trust and Distrust?" In *Trust: A Very Short Introduction*, by Katherine Hawley, 1st ed., 3–12. Oxford University PressOxford. https://doi.org/10.1093/actrade/9780199697342.003.0002.

———. 2014. "Trust, Distrust and Commitment." *Noûs* 48 (1): 1–20. https://doi.org/10.1111/nous.12000.

———. 2019. *How To Be Trustworthy*. Oxford ; New York, NY: Oxford University Press.

Hieronymi, Pamela. 2008. "The Reasons of Trust." *Australasian Journal of Philosophy*, June. https://doi.org/10.1080/00048400801886496.

Holton, Richard. 1994. "Deciding to Trust, Coming to Believe." *Australasian Journal of*

Philosophy 72 (1): 63–76. https://doi.org/10.1080/00048409412345881.

———. 2023. "Knowing, Telling, Trusting." *The Philosophical Quarterly* 73 (3): 762–82. https://doi.org/10.1093/pq/pqad033.

Horsburgh, H. J. N. 1960. "The Ethics of Trust." *The Philosophical Quarterly (1950-)* 10 (41): 343–54. https://doi.org/10.2307/2216409.

Indrek, Reiland. 2024. "Regulative Rules: A Distinctive Normative Kind." *Philosophy and Phenomenological Research* 108 (3): 772–91. https://doi.org/10.1111/phpr.13008.

Israel, Eli Benjamin. 2024. "Caring for Valid Sexual Consent." *Hypatia*, May, 1–21. https://doi.org/10.1017/hyp.2024.42.

———. in review. "Participant Trust."

Jones, Karen. 1996. "Trust as an Affective Attitude." *Ethics* 107 (1): 4–25. https://doi.org/10.1086/233694.

———. 2012. "Trustworthiness." *Ethics* 123 (1): 61–85. https://doi.org/10.1086/667838.

———. 2017. "'But I Was Counting On You!'" In *The Philosophy of Trust*, edited by Paul Faulkner and Thomas Simpson, 0. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198732549.003.0006.

———. 2019. "Trust, Distrust, and Affective Looping." *Philosophical Studies* 176 (4): 955–68. https://doi.org/10.1007/s11098-018-1221-5.

Keren, Arnon. 2020. "Trust and Belief." In *The Routledge Handbook of Trust and Philosophy*, edited by Judith Simon, 109–20.

Kolodny, Niko. 2003. "Love as Valuing a Relationship." *Philosophical Review* 112 (2): 135–89. https://doi.org/10.1215/00318108-112-2-135.

———. 2010. "Which Relationships Justify Partiality? General Considerations and Problem Cases." In *Partiality and Impartiality: Morality, Special Relationships, and the Wider World*, edited by Brian Feltham and John Cottingham. Oxford University Press.

Lahno, Bernd. 2020. "Trust and Emotion." In *The Routledge Handbook of Trust and Philosophy*. Routledge.

Margalit, Avishai. 2004. *The Ethics of Memory:* Cambridge, MA: Harvard University Press.

———. 2017. *On Betrayal*. Cambridge, Massachusetts: Harvard University Press.

Marušić, Berislav. 2017. "Trust, Reliance and the Participant Stance." *Philosopher's Imprint* 17 (17). http://hdl.handle.net/2027/spo.3521354.0017.017.

Mcdowell, John. 1992. "Meaning and Intentionality in Wittgenstein's Later Philosophy." *Midwest Studies In Philosophy*, Reprinted in Mind, Value, and Reality, 17 (1): 40-52 (253-278). https://doi.org/10.1111/j.1475-4975.1992.tb00141.x.

McGeer, Victoria. 2008. "Trust, Hope and Empowerment." *Australasian Journal of Philosophy* 86 (2): 237–54. https://doi.org/10.1080/00048400801886413.

McGeer, Victoria, and Philip Pettit. 2017. "The Empowering Theory of Trust." In *The Philosophy of Trust*, edited by Paul Faulkner and Thomas Simpson, 14–34. Oxford University Press. https://philarchive.org/rec/MCGTET-5.

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

Pettit, Philip. 1990. "The Reality of Rule-Following." *Mind* 99 (393): 1–21. https://doi.org/10.1093/mind/xcix.393.1.

Raz, Joseph. 1989. "Liberating Duties." *Law and Philosophy* 8 (1): 3–21.

———. 2009. "Interpretation: Pluralism and Innovation." In *Between Authority and Interpretation*, by Joseph Raz, 1st ed., 299–322. Oxford University PressOxford. https://doi.org/10.1093/acprof:oso/9780199562688.003.0012.

Reiland, Indrek. 2020. "Constitutive Rules: Games, Language, and Assertion." *Philosophy and Phenomenological Research* 100 (1): 136–59. https://doi.org/10.1111/phpr.12525.

———. 2024. "Rule-Following I: The Basic Issues (1st Edition)." *Philosophy Compass* 19

(1): e12900. https://doi.org/10.1111/phc3.12900.

Scheffler, Samuel. 2002. "Relationships and Responsibilities." In *Boundaries and Allegiances: Problems of Justice and Responsibility in Liberal Thought*, edited by Samuel Scheffler, 0. Oxford University Press. https://doi.org/10.1093/0199257671.003.0007.

Simpson, Thomas W. 2012. "What Is Trust?" *Pacific Philosophical Quarterly* 93 (4): 550–69. https://doi.org/10.1111/j.1468-0114.2012.01438.x.

Stewart, Elizabeth. 2024. "Negotiating Domains of Trust." *Philosophical Psychology* 37 (1): 62–86. https://doi.org/10.1080/09515089.2022.2144190.

Stout, Rowland. 2022. "Betrayal, Trust and Loyalty." *International Journal of Philosophical Studies* 30 (3): 339–56. https://doi.org/10.1080/09672559.2022.2121892.

Strawson, Peter. 1962. "Freedom and Resentment." *Proceedings of the British Academy* 48:187–211.

Thompson, Christopher. 2017. "Trust without Reliance." *Ethical Theory and Moral Practice* 20 (3): 643–55.