The Journal of Ethics
https://doi.org/10.1007/s10892-026-09576-6

Signalling and Moral Ambiguity

Claire Benn[1]

Received: 12 August 2025 / Accepted: 19 August 2026
© The Author(s) 2026

Abstract
Moral agents often act under conditions of informational asymmetry: while they may know that their actions are permissible, observers may lack the information needed to reach the same conclusion. This paper explores the ethical significance of such asymmetries. Drawing inspiration from Signalling Theory, I introduce the concept of *normative signalling*: the idea that agents have moral reasons to choose actions that not only are permissible but also appear unambiguously so to others. I argue that morally ambiguous actions can generate anxiety and undermine trust, allow or demand harmful and unnecessary interventions, and undermine moral education. By analysing a range of cases, I show that reducing moral ambiguity through normative signalling can be a matter of moral responsibility. Further, I show that the strength and scope of this responsibility—like other moral responsibilities for example to be beneficent or kind—is dependent on context, sensitive to costs and so lies on a spectrum from the supererogatory to the morally obligatory. Highlighting the informational value of ethical action, I demonstrate that, in a world where our actions are not—and never can be—completely transparent to others, ethics is fundamentally incomplete if it fails to take into account how our actions play the role of *evidence* of our moral commitments rather than the mere fulfilment of them. It is therefore of moral importance that not only are our actions permissible but that they are *seen* to be permissible.

Keywords Ethics · Permissibility · Trust · Ambiguity · Uncertainty · Signalling · Information

✉ Claire Benn
cmab3@cam.ac.uk

[1] University of Cambridge, Cambridge, UK

Published online: 01 October 2026

1 Introduction

It has been well-established that what we ought to do when we possess all the relevant facts is often different to what we ought to do when we don't.[1] In this paper, I argue that what we ought to do must also be sensitive to the facts that *others* possess, specifically the mistaken beliefs that others have about our moral behaviour. Such mistaken beliefs easily arise because doing what is morally permissible is not the same as being *seen* to do what is morally permissible. The moral ambiguity of our actions does not rely on or imply any disagreement between us and an observer about the moral norms in play, only whether or not we have in fact abided by them. It is simply that our *observable* behaviour is often consistent with a wide range of actions and can therefore give an observer reason to believe that we have done something that is impermissible, even when we haven't. Given that some of our actions are morally ambiguous in this way, a normative question arises: is an agent morally *responsible* for the mistaken beliefs of others about the permissibility of the agent's actions? It is this question that this paper seeks to address. My conclusion is that agents do indeed have a moral responsibility to choose actions that not only are permissible but also appear unambiguously so to others. Like other moral responsibilities, such as being beneficent or kind, this responsibility is unlikely to be an absolute requirement, of equal strength in every context. Rather, I highlight that reducing moral ambiguity is a serious moral consideration that must be taken into account, while noting that (like almost all our moral responsibilities) the scope and strength of this responsibility will depend on other ethical considerations.

This paper is laid out as follows. In Sect. 2, I offer some cases in which the issue of moral ambiguity intuitively constrains moral behaviour. In Sect. 3, I analyse these cases as instances of informational asymmetries, taking inspiration from Signalling Theory, which has primarily been developed and deployed in non-moral domains. I propose a key but often overlooked moral responsibility I call 'normative signalling': that an agent, in addition to doing what is permissible, ought to do what is *unambiguously* permissible.[2] Specifically, this norm encourages us to acknowledge the informational value of our actions and to find ways to reduce informational asymmetries by signalling our commitment to moral constraints through the selection of actions that reduce moral ambiguity. In Sect. 4, I ground this responsibility in the broader principle that, morally speaking, we ought not to mislead others and demonstrate three moral benefits that this normative signalling realises: it can prevent anxiety and promote trust, avoid harms caused by unnecessary interventions, and support moral education. In Sect. 5, I identify considerations (like unreasonableness and observer responsibility) that limit the agent's responsibility to normatively signal and also consider cases where it might be all-things-considered better *not* to reduce moral

[1] For more on objective and subjective oughts, see (Brandt 1959; Ewing 1947; Parfit 1987; Jackson and Pargetter 1986; Jackson 1986; Gibbard 2005). It also applies to moral and non-moral facts. For more on the difference between moral and descriptive uncertainty, see (Jackson and Smith 2006; Zimmerman 2009).

[2] Whether or not this constraint should be folded into our understanding of what is permissible or, instead, considered supererogatory (i.e. *beyond* duty) is explored in Sect. 6. Until then, and for simplicity, I assume that my proposed constraint stands separate from the permissibility of an act.

ambiguity. In Sect. 6, I explore the force of this responsibility, in particular whether it should be folded into our accounts of permissibility itself. Drawing these elements together, this paper demonstrates that we act in a world where our actions are not—and never can be—completely transparent to others; and that our ethical theory is fundamentally incomplete if it fails to take into account how our actions play the role of *evidence* of our moral commitments rather than the mere fulfilment of them.

2 Morally Ambiguous Actions

The clearest statement of how the moral ambiguity of our actions should be taken into account is found in Judaism. On Shabbat (the Jewish holy day of rest that runs from Friday sunset to Saturday sunset), there are 39 'melachot' ('father' or primary categories of work) prohibited by the Torah. These include fieldwork such as sowing, ploughing and reaping; making materials such as shearing, weaving and sewing; and other activities such as writing, building and kindling a fire.[3]

In addition to these Torah prohibitions, there are rabbinic prohibitions.[4] These render certain acts impermissible on Shabbat even if they are not proscribed by the Torah. They do so for one of four reasons (Chait 1992). The first reason is that the act in question resembles a Torah prohibition: the Torah forbids the pressing of grapes and olives on Shabbat; there is therefore a rabbinic prohibition on the squeezing of oranges. The second reason is that the act in question is likely to lead to a violation of a Torah prohibition: the Torah prohibits the picking of fruit on Shabbat; there is therefore a rabbinic prohibition on the smelling of fruit attached to a tree because drawing the fruit close enough to smell is likely to pull it from the tree. The third reason is that the act would violate the spirit of Shabbat: while washing dishes on Shabbat is not prohibited by the Torah, there is a rabbinic prohibition on washing dishes that are not needed for Shabbat as this would not be part of the goal of rest.

The fourth reason is the one of particular interest: *to avoid the suspicion that you have performed or will perform an act prohibited by the Torah*. The washing of laundry is prohibited by the Torah on Shabbat but hanging it out is not. However, hanging it out is prohibited by the rabbinic tradition because it may lead people to conclude that the clothes had been washed on Shabbat. Let's call this example *Shabbat*.

Here, the rabbinic tradition recognises and responds to the fact that our behaviour is often ambiguous—even when we do something otherwise permissible, it might nevertheless look like we have done something impermissible—and that this should affect what we ought to do. While this is the clearest statement of this concern in an ethical tradition, such a concern applies beyond the context of Jewish religious observance. Consider these more secular examples.

[3] For an informative overview, see https://www.chabad.org/library/article_cdo/aid/102032/jewish/The-39-Melachot.htm and https://www.ou.org/holidays/shabbat/the_thirty_nine_categories_of_sabbath_work_prohibited_by_law/.

[4] There are also binding requirements derived from custom (minhag). As these are not needed to illustrate the problem, I leave them aside.

Jaywalker: A pedestrian steps out onto the road in front of you. You are morally prohibited from hitting them with your car when you can safely stop. You stop a mere inch from them. Insofar as you haven't touched them and you know that you were never in danger of hitting them, you seem to have met your obligation. However, your behaviour is troubling. When you stop so close to them, the pedestrian will be increasingly scared that you won't stop in time.[5]

Promise: Your friend calls you in a panic: their debut concert is tonight (a huge professional opportunity where they will showcase their talent in front of important people in the music industry) and the guitarist who would normally accompany them has just cancelled. They ask you if you could step in: you are an amazing musician and already know the set list. Their set starts at 7pm. You promise you will be there by then. It would have been just as easy and convenient for you to turn up at 6:30 or 6:45, but instead you choose to turn up at 6:59pm. This seems permissible: you promised to be there and you are. However, your friend is understandably angry about your turning up so close to 7pm in part because it made them increasingly worried that you would break your promise and not show up at all.

Crucible: You have been cast as John Proctor in your local am-dram production of *The Crucible*. You decide to practice in the local park. While this seems permissible (the park is public and has no prohibition against aspiring thespians practicing there), it nevertheless seems problematic to practice the scene where John grabs Elizabeth by the throat without doing it in costume or otherwise making it clear that you are rehearsing a play, lest someone walking through the park feels they need to intervene believing that you are attacking someone.

Rescue: You are on duty at mountain rescue and a call comes in over the radio that some hikers are stranded. Just before you head off to rescue them, you think about pressing send on the tweet you had just composed: in it you joke (in response to comments about your hermetic lifestyle, which have a grain of truth) that the lonely job of mountain rescue isn't solitary enough so you are calling it a night and heading off to your even more isolated cabin in the countryside (which has no phone or radio reception). However, you know your off-duty colleague would have heard on the radio about the hikers and is also a keen Twitter follower. While it is in general fine to post a joke on Twitter, given that you actually do your job and save the hikers, it seems like the better thing to do would be to not send the tweet, in case your colleague takes it upon themselves to put themselves in danger by trying to rescue the hikers, assuming that you had really left your post.

[5] It is also perhaps marginally less safe insofar as a small error can lead to something impermissible when you stop close to the pedestrian as opposed to stopping sooner. But let's assume that both stopping in good time and stopping mere inches away are deterministically safe. Even in a case where both you and the pedestrian are certain that you could stop in time *should you so choose*, it is still unnerving from the point of view of the pedestrian who doesn't know whether or not you are so choosing.

Phone: You are supposed to pick up your kids after school. Your partner has an important presentation at work at that time. It seems permissible to turn your phone off (it is your phone after all), as long as you actually do pick up your kids. However, turning your phone off seems problematic as you know your partner, upon being unable to get through to you if they call, will be put in the position of not knowing whether to abandon their presentation in order to go to the school on the assumption that you have forgotten (you often turn your phone off when you go to the gym after work, which you would normally do today).

Mask: During the SARS-CoV-2 pandemic, mask wearing was shown to be important in combatting the spread of the virus and as such was mandated in many places. You decided to make your own mask from appropriate fabric at home. You remember you have some leftover fabric printed with a picture of your own face (from a silly pre-COVID craft project), which perfectly matches your skin tone and replicates your mouth. If you make a mask out of this fabric it will look to almost everyone that you aren't wearing a mask. You can make a mask out of this fabric that meets all the requirements for a mask that helps stop transmission. However, it is likely to make people think you aren't wearing one at all. As such, it would be better to wear a mask where it is more obvious you are wearing one.

Thus, we have in hand a host of intuitive examples that demonstrate how our actions can be morally ambiguous from the point of view of an observer. They also go beyond this: they suggest that this moral ambiguity is *morally* important. To unpack this further, I turn now to how to understand the informational value that actions can have.

3 Signalling Moral Compliance

The moral ambiguity of the above examples reveals the existence of an *informational asymmetry* between the agent and the observer in each case: the agent knows that they have done (and would only do) what is permissible, but the observer doesn't know that. Thus, the problem is one of communication: in the presence of this informational asymmetry, how can a conscientious agent successfully convey the permissibility of their actions when this is not directly observable?

Luckily, we do not have to start from scratch in answering this question. Signalling Theory was developed explicitly to deal with 'reducing informational asymmetry between two parties' (Connelly et al. 2011: 40). Informational asymmetries emerge given that, as in our examples, the mere possession of some feature or property by an individual does not entail that others will be certain they possess it. In such cases, a communication problem arises: one party has access to some information that is not directly observable or accessible but wants to convey that information somehow to another. Signalling Theory is useful for describing the behaviour that emerges in such situations that result in the reduction of informational asymmetries (Connelly et al. 2011; Spence 2002).

This communication problem—of how to successfully convey the presence or absence of certain features that are not directly observable—has application in a broad range of domains.[6] In biology, Signalling Theory can explain, for example, the brightly coloured appearance of poisonous frogs: poisonous frogs need to communicate their toxicity to deter predators as, if they didn't, the lesson learned by eating them would only benefit future generations of frogs (if it benefits any at all). The bright outlandish colours clearly signal the frog's lack of palatability (Maan and Cummings 2012).

In economics, signalling explains how certain behaviour is beneficial because of its *informational value*: it enables one party to convey important information (or convey the truthfulness of what they say), which can be beneficial for a variety of reasons (Veblen 1899; Spence 1973). Take education. Going to university had long been theorised as a rational investment for both individuals and government under the assumption that it increases economic growth through improving worker productivity. Contrary to this, Michael Spence (one of the founding theorists of Signalling Theory in economics) emphasised that education is a good *signal* of the unobservable property of a job candidate's quality: less capable candidates are more likely to have failed to obtain a university qualification than more capable candidates (Spence 1973; see also Weiss 1995).

Signalling can thus explain actions that are costly, including seeming extravagance. In anthropology, signalling has been used to explain the rationality of throwing a large feast with no expected reciprocation: while a great expense to the host, it is an honest signal of the host's competence and expertise as a hunter or provider (the feast would be infeasibly costly for those who do not have such abilities) (Hayden 2009). Similarly, while the bright colours of a poisonous frog make it more likely to be seen by predators (which is a cost to prey), it makes them less likely to be eaten by predators. The very visibility of the frog in fact communicates to predators that they should not eat it: a highly visible, non-poisonous frog would not, after all, have lasted long.[7]

Across applications of Signalling Theory, there are several core components: a signaller[8] and a receiver; a property that is present in or known to the signaller but cannot be directly observed by the receiver; an informational asymmetry between the two parties; and a signal which is an observable means by which that information is transmitted between them thereby reducing the informational asymmetry.

[6] Signalling Theory in fact reveals two cases where the problem of communication arises: where there is uncertainty and where there is cheating. The latter has been the focus of much of the philosophical engagement, which has examined cases of deception in biological organisms, where (according to some) the sender sends (systematic) misinformation that benefits the sender, often at the expense of the receiver; see for example (Skyrms 2010; Artiga and Paternotte 2017; Fallis and Lewis 2021; Krstić 2023).

[7] Excepting those mimics who benefit derivatively from this reasoning.

[8] I use the term 'signaller' as the British English variant following the American English 'signaler' sometimes used in economics and biology (Spence 1973, 2002; Smith and Bird 2003; Smith and Harper 1995; Grafen 1990). The term 'sender' is used by some economists (Bagwell 1995) and those in the philosophy of biology, in particular those who emerge from the Lewisian signalling game tradition (Fallis and Lewis 2021; Artiga and Paternotte 2017).

For example, in the biological case: the frog is the signaller, the predator the receiver, the poisonousness of the frog the property that cannot be directly observed, and the garish colours the signal that communicates the property reducing the informational asymmetry. We can see the same components in the economics example of education: the candidate is the signaller, the prospective employer the receiver, the quality of the candidate is the unobservable property, and a university education the signal that communicates that property.

Similarly, the ethical examples proposed above, despite their disparate details (which I will return to), share many of these important features. There is you, the moral agent (the signaller). There is also an observer (the receiver): your neighbour, the pedestrian, your concert-giving friend, a person walking through the park, your off-duty colleague, your partner, and your fellow endurers of the pandemic.[9] In each case, the observer does not know an important piece of information, specifically, whether or not you have done (or will do) something impermissible (here the agent's moral conscientiousness is the property that can't be directly observed): washing your clothes on Shabbat, hitting the pedestrian on the road with your car, breaking your promise, hurting someone, and failing to rescue the hikers, to pick up your kids or wear a mask. And finally, there is in each case a less ambiguous course of action (the signal): not hanging out your laundry on Shabbat, stopping further away from the jaywalker, making clear you are acting in a play, not posting the joke, keeping your phone on and wearing a mask that looks like a mask.

Thus, in this paper, taking inspiration from Signalling Theory, I propose that normative theory incorporate the idea of, what I will call, *normative* signalling: that informational asymmetries exist in the moral sphere because even when we only do what is permissible, our actions can seem impermissible to an observer; and that we can reduce this asymmetry by signalling our commitment to and compliance with normative constraints by choosing less ambiguous courses of action.[10] As such, normative signalling emphasises that even among *equally permissible* actions, we have reason to prefer some over others because of their greater *informational* value.

Drawing on Signalling Theory is valuable. It allows us to conceptualise our original, morally ambiguous examples as informational asymmetries and identifies the informational value of non-ambiguous actions as signals of our moral conscientiousness. However, other key aspects of Signalling Theory perhaps do not apply in the moral domain as straightforwardly. One is cost.

The cost of a signal is often thought to be an important factor in determining an 'honest' signal (that is, in ensuring that the receiver has good reason to trust the signals

[9] As this paper is laying out the foundations of incorporating the communicative elements of ethical behaviour, only one observer is considered, leaving aside multi-observer situations.

[10] The term 'signalling' has been used in normative contexts for example by Renée Jorgensen Bolinger (2020). However, her focus is on 'signals of aggression' specifically and on the difference such signals make to whether an observer's reasonable belief of threat can justify mistaken self-defence. My work differs by focusing on the *agent's* responsibility to choice less ambiguous courses of action in order to signal their compliance. There has also been work on signalling *commitment*, though these have tended to articulate signals that communicate commitment to a particular religious affiliation (Bulbulia and Schjoedt 2010; Bulbulia et al. 2013) or to intimate relationships (Yamaguchi, Smith, and Ohtsubo 2015) and so on, rather than obedience to moral constraints per se.

of an agent) (Connelly et al. 2011). It has been used for example to explain in biology how evolutionarily stable strategies are arrived at (Zahavi 1975; Grafen 1990; Johnstone 1997). As one example: Zahavi in his seminal work asks why females might prefer a male with an evolutionary disadvantage, for example the more impressive the plumage of a peacock, the less good it will be at evading predators. His answer is that the tail of a peacock is a signal of its underlying health and fitness. This is a highly costly signal (in terms of energy and reducing its ability to evade predators) but it is an honest signal because only a male with the underlying qualities of health and fitness could produce an impressive tail in the first place. This costliness is correlated with its trustworthiness. Similarly, in economics, when it comes to imperfectly observable actions, it is sometimes assumed that increases in cost will lead to better information (Miyagawa, Miyahara, and Sekiguchi 2008).

What is the equivalent 'cost' in the case of the moral examples I have given? In some of the examples, normative signalling would necessarily involve some cost in terms of time, effort, resources and so on: for example, making it clear that you were acting in a play through using costumes or signs and so on could be monetarily expensive, while turning up somewhat earlier than exactly when you promised in order to reassure your friend that you will keep your promise is comparatively costly in terms of your time. However, in some cases, it seems like no additional costs might exist. Perhaps it doesn't inconvenience you at all to turn up earlier in the case of *Promise* or to brake sooner in *Jaywalker*. Nevertheless, this is in keeping with arguments made about the relevance of costs by those like Maynard Smith who proved that 'honest signals must be costly if there is a conflict of interest between signaller and receiver, but that cost-free signals can be honest if there is no such conflict' (Smith 1991: 1035). And in many of the cases I raise, interestingly, there is no conflict of interest between signaller and receiver, that is between the agent and observer. This is because of the specifics of the benefits of normative signalling (which I discuss in the next section).[11]

These benefits underpin the explicitly normative nature of my proposal: that agents in many circumstances *ought* to avoid courses of action even if they are permissible, because they are morally ambiguous. In light of this, I should emphasise that, while normative signalling draws on the structure provided by Signalling Theory in the ways described above, it also moves beyond it. Established Signalling Theory, in biology and economics, provides an explanatory perspective on why some strategies exist and why they are stable and successful. It is a mistake to suppose that Signalling Theory *requires* or even *recommends* that the poisonous frog has garish colours. Signalling Theory does, nevertheless, explain the *value* of these strategies and how it can *benefit* the signaller and/or receiver, and thus *justifies* the extra cost of employ-

[11] There is much more to be said about the relationship between cost, normative signalling and ethical action in general: for example, even when the interests of agent and observer are aligned, are more costly signals more honest? Are the costs of normative signals efficacy costs or strategic costs (to use the terminology of (M.J. Smith and Harper 1995)? There is, however, not enough space in this paper to do justice to exploring in detail the extensive debates in Signalling Theory about the exact role and nature of relevant costs. I focus here on establishing that informational asymmetries exist in normative interactions and that normative signals are a key means of reducing them, so key that we have moral reason to normatively signal.

ing them: even if education is not a rational investment in the sense that it increases worker productivity, it is nevertheless a rational investment because it is a good signal of the underlying capabilities of a job candidate. Nevertheless, I acknowledge that I am moving beyond its original deployment by proposing an explicitly normative recommendation: that we morally *ought* to signal compliance to existing normative constraints by choosing courses of action that reduce ambiguity and informational asymmetries between the agent and the observer. To see why this is justified in the moral domain, we must explore the moral value of normative signalling. I begin by situating it within a more basic moral responsibility and then explore the specific moral benefits of normative signalling in more detail.

4 The Value of Signalling

4.1 The Responsibility not to Mislead

The moral value of normative signalling is grounded in the broader ethical responsibility we have towards the epistemic state of others: specifically, the moral responsibility not to *mislead* others in their beliefs. This broader principle covers several well-established areas of moral norms, including those governing testimony, lying, deception and manipulation.

The normative signalling is, I propose, similar to these norms insofar as they all take seriously that we can cause others to come to believe things, which in turn generates moral responsibilities. My discussion of morally ambiguous actions emphasises that our actions (similarly to our statements, in the case of testimony) can generate an entitlement-to-believe in others, in virtue of constituting evidence from which an observer can make certain justified inferences.[12] And so, similarly to testimonial obligations, we can have responsibilities with respect to our actions so that we take into account how we might mislead others. This responsibility not to mislead can also explain why the force of this responsibility might vary, as it does with other general moral responsibilities, with the costs and stakes of ambiguity (something I explore in more depth in Sect. 6). And yet, moral ambiguity and normative signalling cannot be subsumed by these existing areas of moral responsibility. I briefly explore why.

A key example of an existing duty not to mislead the beliefs of others is the literature on lying. According to this literature, lying has several key components, a central one of which is that for something to be a lie, the agent must, surely, *say something* (Isenberg 1964; Primoratz 1984; Mahon 2016). This is most obvious in the case of uttering something with clear linguistic meaning like "this giant carved horse is just a peace offering." However, it can include sign language, Morse code or semaphore as these also convey linguistic meaning, because these too involve 'conveying meaning through conventional signs' (Siegler 1966: 128). These non-verbal utterances can be lies precisely because they are, nevertheless, *asserting* things. Similarly, on many accounts, testimony presupposes assertion. The cases of moral ambiguity I

[12] For more on testimony, in particular on reduction vs anti-reductionist views, see (Fricker 1994; Coady 1992).

have raised, on the other hand, do not necessarily involve asserting anything. No one has said anything to anyone in the case of *Jaywalker* for example (or made use of anything to convey linguistic meaning). The cases of moral ambiguity of interest here involve our *actions*, rather than what we say.

Similar to the cases I propose, deception can involve a broader set of non-assertive actions: for example, pretending to have a limp in order that others think you are incapacitated (Siegler 1966: 129). This might also include wearing a wedding ring when you are not married or wearing a police uniform when you are not a police officer (Mahon 2016). We can deceive others through our conduct, our appearance and other non-assertive acts (Chisholm and Feehan 1977). This is closer to the cases I have presented, where the agents in *Shabbat*, *Jaywalker* and others risk generating a false belief in the observer through their conduct, despite the fact that their conduct is entirely permissible.

However, it should be noted that the specific cases I am discussing cannot be subsumed under already existing accounts of deception. This is because intentionality is a key part of most definitions of deception and the cases I am concerned with do not require an intention to mislead. In morally ambiguous acts, what matters is that it is *foreseeable* that the observer will form a false belief about the permissibility of the agent's conduct, regardless of whether the agent intends this effect. To return to *Jaywalker*, the driver need not intend the pedestrian to think that they will be hit by the car; in fact, the driver (and indeed all the agents in the cases discussed) might actively intend and hope that the observer only believes them to be conscientious (i.e. abiding by the moral norms in play), as they in fact are. Nevertheless, my account highlights their moral ambiguity and the mistaken beliefs of others that are predictable as a result.

Accounts of manipulation often identify the wrong of manipulation as influencing another's beliefs, desires, emotions or actions in a way that bypasses their rational agency (Noggle 1996) or that 'exploit[s] cognitive, emotional, or other decision-making vulnerabilities' (Susser, Roessler, and Nissenbaum 2019: 27). Again, the cases I am discussing do not make any assumptions about bypassing rational agency or exploiting cognitive vulnerabilities and therefore are not covered by existing work on manipulation. In fact, as we will see in Sect. 5.1, it is a condition of the responsibility having some bite in a specific instance that the mistaken beliefs of the observer are rationally reasonable.

My argument also differs from existing philosophical discussion on norms against misleading because these are concerned with actions or speech that mislead others about the world more generally. My concern is far narrower and more specific. Morally ambiguous actions are those that misrepresent *the moral status* of the agent's actions: it appears to the observer that the agent has done, is doing or will do some-

thing morally *im*permissible despite the fact that the agent knows that they have done something permissible.[13]

Just as with lying or deceiving, there are always limits to our responsibility not to mislead others and this responsibility can therefore vary in its force depending on the costs, stakes and context. I explore those limits in Sect. 5 and the force of our responsibility to signal in Sect. 6. But first, I turn now to exploring three specific moral benefits of not misleading people about our moral conduct: decreasing anxiety and increasing trust; reducing harm; and supporting moral education.

4.2 Anxiety and Trust

One benefit is that normative signalling avoids the negative psychological harms that arise from morally ambiguous actions. Consider those cases where the observer is also the person who would be harmed or whose rights would be infringed if the agent were in fact doing something impermissible. This is the case in *Jaywalker*: as you drive closer and closer to the pedestrian before stopping, they will get increasingly anxious that you are about to hit them with your car. Here, even though the agent only does what is permissible (in *Jaywalker*, actually stopping before you hit the pedestrian), their morally ambiguous behaviour gives rise to justified anxiety on the part of the observer. In such cases, signalling compliance can reduce the anxiety of the observer. The pedestrian in *Jaywalker* will be much less anxious if you stop several feet from them, rather than just an inch. Much ink has been spilt on the question of whether putting someone at *risk* of harm even if they are not in fact harmed can constitute a wrong (Hansson 2013; Oberdiek 2017; Thomson 1986). Here, cases of moral ambiguity highlight that increasing someone's *perceived* risk can wrong them even if we do not actually put them *at risk*.

Even when the observer isn't anxious about their own safety, the observer might be anxious that the agent will do something impermissible. By signalling her understanding of and commitment to the constraints in play, an agent demonstrates her moral trustworthiness in general. Additional communication aids in the building of appropriate trust. This will facilitate smoother interpersonal interactions. Take our examples of *Rescue* and *Phone*: here choosing the less ambiguous course of action makes the observer trust that you are doing what you need to do (and will do in the future). In *Rescue*, by choosing not to send the joke social media post about heading to your cabin, your colleague is able to trust that you will in fact be doing your duty in rescuing the hikers. By leaving your *Phone* on and being potentially contactable, your partner can trust that you have *not* gone to the gym and will be picking up the kids, so they can concentrate on their presentation instead.

[13] In virtue of this, there are some interesting connections between my proposal for the value of normative signalling and the literature on integrity. In particular, many theorists who write about integrity believe that we ought to concern ourselves with how we present our actions before our epistemic community because it can constitute important information about ethics itself (Calhoun 1995). Similarly, my proposal is that we must take into account the epistemic position of others in our community or in specific relationships with us when acting morally. I thank a reviewer for this journal for prompting me to note the connection. I leave it for future work to unpack this further.

Signalling compliance also helps to build trust on a societal level. Morally ambiguous actions could undermine society-wide norms by making it appear as if people are not, in general, abiding by them. We can see this from *Mask*: by choosing to wear a mask that is obviously a mask, rather than run the risk of being mistaken for someone who isn't wearing one, the agent signals and reinforces the mask-wearing norm *in general*, without which others might be less inclined to wear masks themselves.

As *Promise* illustrates, normative signalling is also important in order to maintain important social practices like promises. It undermines the very institution of promising to make the promisee unsure whether or not you are going to keep the promise you made.[14]

4.3 Harm Reduction

The second reason that signalling compliance is valuable is that, in the face of moral ambiguity, observers might feel entitled, or even required, to intervene. Importantly, they might be subjectively right given the information available to them. Risks of harm can be enough to justify intervention and *believing* someone is at risk of harm can subjectively justify or require intervening.

When the observer is the person at risk of harm, moral ambiguity might license them to defend themselves given the perceived risk. Harming the agent in the case of self-defence can be permissible in many cases, even when the observer is mistaken about the risks (Bolinger 2020).[15] When the observer is a third party, they might be permitted or required to intervene on behalf of another when they believe that person to be at risk. We can see this in the case of *Crucible*, where a passer-by might justifiably intervene when they see you appearing to choke your co-actor. But it also occurs in the case of *Rescue*, where your colleague intervenes by going out to rescue the hikers themselves because they believe you have failed to do your job. In both cases, the observer believes that they need to intervene in order to protect others from harm that would occur if you were doing something impermissible.

Thus, moral ambiguity is likely to lead to unnecessary interventions which in turn involve bringing about unnecessary harm or risk of harms. Your colleague puts themselves at risk when they set out to save the hikers in your stead. This risk is required *had you actually failed to go*; but, given that you did go, the risks to your colleague are unnecessary. Similarly, in *Crucible* and in a self-defence case, the observer undertakes risks upon themselves unnecessarily in attempting to intervene and stop the harm from occurring.

Not only are these harms unnecessary, these mistaken interventions arise because the agent places the observer in a situation where the observer is then *subjectively* permitted, justified and perhaps required to do something that is in fact *objectively* wrong: it is objectively wrong for the passer-by in the park to tackle you to the ground risking physical harm to you and themselves, given that you are only play-acting in *Crucible*; however, given their reasonable belief that you are attacking someone, it

[14] A particularly prominent concern on Expectation Theories of promises, e.g. (Scanlon 1998).

[15] Bolinger discusses cases of a person who, with no intent to harm, walks close behind a woman in a car park, causing her to have the reasonable, though mistaken, belief that he might attack her.

is subjectively permitted (perhaps required) for them to intervene. When we cause others to believe certain things that are mistaken, we change what they subjectively ought to do, making them increasingly distant from what is objectively right. Insofar as any objectively wrongful act is problematic when it is avoidable, the agent ought not to put observers in situations where those observers justifiably ought to do what is objectively wrong. This is an additional way in which we can morally wrong others, and one that is only explainable through considering the informational asymmetries that can exist between ourselves and observers and thus our responsibility for the mistaken beliefs of others, as explored in this paper.

Thus, normative signalling can reduce unnecessary harms and risks from taking place. Moreover, it can help avoid further objective wrongdoing by reducing the gap between what is objectively right and subjectively justified for others to do.

4.4 Education

The third benefit of normative signalling concerns the role our behaviour has in shaping the moral beliefs and moral behaviour of others.[16] Our actions have important educative implications. Recall *Shabbat*: the rabbinic tradition prohibits actions that could lead someone to think that you have performed or will perform an act that is prohibited by the Torah. When someone sees you hanging out your laundry on Shabbat, it is reasonable for them to assume that you have washed your laundry on Shabbat. If they know that this is prohibited, this may lead them to conclude that you have done something impermissible. But another reason to be concerned in this case is that, *if they are not certain about what is prohibited*, they may conclude that washing laundry on Shabbat is permitted.[17] After all, if they (mistakenly) infer that you washed your laundry on Shabbat, and they believe (correctly) that you keep Shabbat, they would understandably conclude that washing laundry on Shabbat must not be prohibited. This is why, as Aron Moss states, in the Jewish tradition, the educative aspect of moral practice means 'your morality is my business' (Moss 2011).

Thus, in general, when someone who does not know what is permissible and impermissible observes the agent performing morally ambiguous behaviour, they might conclude that a morally impermissible act is in fact permissible. This is perhaps particularly important in times of moral change when what is morally permissible or impermissible is unknown or in flux. In the case of *Mask*, the importance of unambiguously wearing a mask during a pandemic is increased in the face of confusion or rejection of the evidence of the effectiveness of mask-wearing.

In the *Shabbat* case, those involved are part of the same community. However, the educational value of normative signalling obtains even when we do not have any special relationship to the observer. In the other examples presented, we can see the variety of roles the observer might have: the person at risk (as in *Jaywalker*); someone

[16] For just one example of the importance of role models in moral learning, see (Moberg 2000).

[17] This is reflected in Halakhah in the distinction between 'Cheshad' and 'Mar'it Ayin'. Cheshad concerns causing someone to think you have done wrong (and so to view you negatively). Mar'it Ayin, on the other hand, states that one may not perform certain actions that appear to be forbidden even if one is in fact doing them in a permissible way. It is often reserved for those actions that could cause someone to think that something which is forbidden is in fact permitted.

who has a right that you do something (*Promise*); someone with a duty to intervene if you are doing something impermissible (*Crucible*); someone who has role-based responsibilities (*Phone*); a member of our social (*Mask*) or religious (*Shabbat*) community. In each case, our responsibilities might be stronger or weaker depending on what we owe to that observer: we plausibly have stronger responsibilities when it comes to our partners and promisees than to strangers. Yet even the members of our moral community—perhaps the broadest class of observers—have an interest in our signalling our compliance, because everyone has an interest in not being misled as to the moral facts.

Thus, in this section, I have shown that normative signalling—avoiding courses of action that are morally ambiguous, even if they are permissible—has significant moral value in many cases. Notably, this moral value is not dependent on motive. We can see this by contrasting my account of normative signalling with another place in which the term 'signalling' appears in discussions of ethics: in the concept of 'virtue signalling'. The term 'virtue signalling' picks out acts (such as changing a profile picture or publicly signing a petition) that are conspicuous, though often impotent, demonstrations of holding certain values (Levy 2020; Tosi and Warmke 2016). Such acts are often thought to be problematic or deserving of derision. We might therefore be concerned that normative signalling also promotes the performance of actions that focus on the mere *appearance* of being good. Being motivated to be *seen* in a moral light does seem to be morally problematic. However, 'virtue signalling' concerns actions that fail to be a good indication of any actual underlying moral conscientiousness, which is partly why they are derided. Changing a profile picture or signing a petition doesn't necessarily mean anything about the agent's underlying commitments. Thus, while it can *display* one's virtues or values, virtue signalling fails to be an honest indication of one's actual values. Normative signalling, on the other hand, focuses on courses of actions that do in fact reduce moral ambiguity in virtue of being less compatible with interpretations of impermissibility. Furthermore, even if someone is motivated by the desire to appear morally conscientious, the benefits of normative signalling as outlined in this section still accrue so long as the actions *actually* reduce moral ambiguity: it is as a direct result of choosing a less morally ambiguous course of action that anxiety is avoided, trust built, harm reduced, and moral learning achieved. This is not mere talk, but actions with concrete and specific benefits through the resolution and reduction of informational asymmetries.

5 The Limits of Signalling

We have seen the value of normative signalling. But, as with many other norms and responsibilities, there are limits.

5.1 Reasonableness of Mistaken Beliefs

If we ruled out any course of action where the observer's observations were merely *compatible* with an impermissible course of action, this norm would be far too strong. Just because it is *possible* that the agent did something impermissible does not entail

that it is reasonable for the observer to assume that the agent has done something impermissible.

The reasonableness of the observer's mistaken belief could be factored in in different ways. One would be to consider the agent's known (non-moral) preferences. If it would be irrational for the agent to do what is impermissible, then we ought not to worry about the agent's behaviour even if it is ambiguous in the sense of it being compatible with an impermissible course of action. Suppose for example in *Rescue* your colleague knew that you in fact hated the cabin in the woods (an inheritance you were planning on selling anyway) and would prefer to do anything, including the job you love, rather than stay in that spider-infested dump. Thus, while it is *possible* that you could actually have gone to the cabin as you joke, your colleague knows you would not do so. In such a case, we could say that an agent need not avoid morally ambiguous courses of action if the impermissible action it is compatible with is dominated by a permissible alternative.

The observer can also be expected to incorporate other sources of information (for example, whether the agent has done that impermissible action before or has only ever done what is permissible) to estimate subjective probabilities, setting aside morally ambiguous actions where it is more likely that the agent has done what is permissible than impermissible.[18]

Thus, it is important to understand the agent's behaviour as *evidence*. Responsibility for the negative consequences of moral ambiguity might lie with an observer if they irrationally or unreasonably (given the relevant likelihoods and risks of the situations) insist on assuming that the agent has done something impermissible because it is consistent with the agent's action, despite the fact that other evidence indicates that it is incredibly unlikely. This is particularly pertinent when mistaken beliefs about the actual conduct of the agent arise from problematic prior beliefs: for example, assuming that the black person reaching into his pocket is getting out a gun rather than his wallet while you are at an ATM (Bolinger 2020).

Furthermore, the reasonableness of the observer's mistaken beliefs is key when considering the agent's responsibilities for the moral education of others. The rabbinic prohibitions on morally ambiguous actions are in part dependent on the *likelihood* (and reasonableness) that someone will learn something mistaken about what is prohibited.[19]

5.2 Responsibilities of the Observer to Disambiguate

The focus of this paper has been on the opportunities and responsibilities of the agent to choose less ambiguous courses of actions. However, the *observer* can also have responsibilities in the case of moral ambiguity. For example, consider a scenario in

[18] We can see that this could also increase our responsibility to avoid ambiguity. If you are chronically late to things, you might have a greater responsibility to turn up not at the last minute in the case of *Promise* because it is very reasonable for the promisee to believe that you will be late (again).

[19] For example, Rabbi Meir prohibits his student from washing a tiled floor (despite it being permitted by the Torah) lest people mistake it for a dirt floor (which is prohibited by the Torah) ("Tractate Shabbat", fol. 151b). However, washing a dead body is allowed, as the Talmud assumes people will not mistake it for a floor.

which the observer could determine whether or not the agent had done something impermissible by finding another point of observation. For example, suppose that your colleague in *Rescue* could go and dig out the work-issued device that lets them track your car, which would tell them you had headed out to the mountains to rescue the hikers rather than to your cabin. In such a case, their responsibility to easily and straightforwardly disambiguate your actions is likely to reduce your responsibility to do so.

The above reminds us that ambiguity is often relative to cost. Thus, two courses of action might be indistinguishable relative to a certain cost and yet be distinguishable relative to a greater cost (let us assume that as the cost increases, observations become more precise). As in many other areas of ethics, the cost to the individual could be in terms of time, money, resources, effort, reputation or many other things that make life go well or are considered valuable.[20] Which level of cost is the relevant point of assessment can depend on many things. For example, when the stakes are high, we may want to avoid any course of action that could be mistaken for an impermissible one (that is, that are indistinguishable at the lowest cost). In other scenarios, a weaker constraint might be appropriate, whereby we only rule out those courses of action that cannot be distinguished *cheaply*, according to a context-dependent understanding of 'cheap'. When our actions could be disambiguated (or at least, disambiguated cheaply) by the observer, the responsibility to do so might rest on the observer rather than the agent. There is no straightforward answer to how costs and benefits here might be weighed. This is an area that is notoriously difficult in ethics more broadly. Instead, I wish to draw attention to the fact that whether or not an action *is* morally ambiguous will depend, on occasion, on how much the agent or the observer is willing to do to disambiguate them.

Thus, while it is in general *better* to avoid moral ambiguity, it might sometimes be the observer's responsibility to do so. The issue of communication and signalling comes therefore with attendant concerns about justice: moral ambiguity can lead to moral harms and moral wrongs but the costs of reducing that moral ambiguity might rightly fall on the observer rather than the agent.[21]

5.3 Positive Effects of Ambiguity

It should also be noted that there are cases where moral ambiguity is all-things-considered *positive* and therefore signalling, in such cases, is not something that the agent ought to avoid. Consider the following examples.

First are cases where you have done something morally *im*permissible. If the observer does not know that the action is impermissible, they may well be likely to

[20] The costs involved here often make a difference to the normative categorisation of an act, as acts of extreme self-sacrifice for miniscule benefits to others might well be prohibited and acts of extreme benefit to others for little to no cost to the agent might be required. This is something that is well discussed in the literature on supererogation in particular (see for example Benn 2017).

[21] This is something Bolinger identifies in her discussion of self-defence. She emphasises that reasonable mistake depends on 'whether it is reasonable to ask the [defender] to bear the risk of a perceived attack, or to make the apparent aggressor bear the risk that the [defender] was mistaken' (Bolinger 2020; quoting Ripstein 1996: 691).

emulate it. By making your action more ambiguous, they might be more likely to emulate a better course of action than the one you actually performed. This mirrors the education case discussed earlier. By reducing ambiguity when you do what is permissible and increasing it when you do something impermissible, you might lead other people to be more likely to copy permissible behaviour.

Second are cases where you have done something morally *permissible* and even morally worthy. Even in such cases, it might sometimes be beneficial to disguise those actions. For example, suppose you gave money to charity: there are occasions where moral ambiguity (such as failing to publicise that you have done so) could be positive, if you hoped to avoid, say, pressure on a friend who you know cannot afford to do the same. Similarly, suppose you give money to a beneficiary, but again disguise this fact: this can have the positive overall impact of not making that person feel that they are the object of charity, allowing them to save face and benefit without having to be uncomfortably conscious of their own need and precarity.

Thus, while the positive effects outlined above do generally attend the reduction of moral ambiguity, the connection is context-dependent and contingent. As outlined in the example *Mask*, during the COVID-SARS-2 pandemic, most countries required citizens to wear masks in order to curb the spread. However, some studies claimed that certain kinds of masks had negative impact (certain kinds of mesh fabrics were hypothesised to transform large droplets into small droplets increasing their spread (Fischer et al. 2020)). How does this impact what an agent ought to do in the situation where an observer will only know that someone is or is not wearing a mask, and not what kind of mask it is? On the one hand, it might still be better to wear any mask, even a mesh one that has direct negative health impacts, rather than no mask, in order to encourage others to wear masks, strengthening the mask-wearing norm overall. On the other hand, it might be better to wear no mask than a mesh mask (thereby unambiguously violating the norm) because it will then allow others to take alternative precautions (like keeping their distance) that they would fail to do under the mistaken assumption that you were wearing an effective mask.

6 The Force of Signalling

We have seen how normative signalling is morally valuable, though the responsibility to signal might be limited in some cases. So, what is the strength of this responsibility? What kind of 'ought' does it imply? How should it be incorporated into normative theory? Here are two answers.

The first view is that normative signalling is a second-order moral consideration. If the choice is between actually hitting the pedestrian and stopping so close as to make them *believe* you will hit them, you should do the latter. It is not a matter of flipping a coin. After all, signalling addresses the *perception* of an action's permissibility. Other things being equal, first-order constraints take priority over second-order constraints and, intuitively, considerations that traditionally determine permissibility take priority over those concerning appearance of permissibility. Also, intuitively, reactive attitudes, such as blame, differ in the strength apt for first- and second-order violations: we ought to blame more strongly the person who was actually strangling someone in

the park than the person practicing in *Crucible* who made you *believe* that they were strangling someone.

The idea that equally permissible actions can have other normative differences is not limited to considerations of appearance. Supererogationists have long argued that not all permissible actions are equal: some are also supererogatory (beyond the call of duty). Similarly, we could take the value of signalling to indicate that while two actions might both be permissible, that does not mean that they are both of equal value: less ambiguous actions might still be morally better than more ambiguous ones.[22]

The second view folds signalling into the notion of permissibility itself. On this view, some actions typically seen as permissible are in fact impermissible when they are morally ambiguous. Normative signalling therefore functions like other moral constraints such as *do not harm*: it constrains what is permissible, but when one constraint comes into conflict with another, it can be justified to violate it. After all, while all things being equal we ought not to harm someone, it is nevertheless permissible to harm them for all sorts of reasons, such as in surgery to benefit them more greatly overall, or in order to stop them from harming someone else. This is because it is not always possible to achieve or maximise every moral value.

Take safety. It might be that, in order to signal their compliance, agents would need to choose an action that is objectively *less* safe. Suppose, for example, that the safest course of action a driver could take is to speed up and closely circumnavigate a pedestrian who steps out onto the road. This leaves us with a dilemma: should the driver behave in order to maximise the objective safety of the pedestrian or instead respond to the pedestrian's subjective perception of their safety? In such cases, and others where two values conflict, we have to decide which to prioritise and to what extent. Indeed, it might turn out where the responsibility to normatively signal conflicts with other responsibilities (for example, to preserve the objective well-being of others) it might be good—and even on occasion required—*not* to signal. Nevertheless, in the absence of such conflicts, on this view, it can be impermissible to follow courses of action that are unnecessarily morally ambiguous.

This might strike some as a very strong claim as, arguably, violations for failing to normatively signal should be met with milder reactive attitudes than failures to do

[22] Another possible parallel case comes from the literature on indeterminacy: we perhaps ought to choose those actions that are determinately permissible over those that are indeterminately permissible. In this case, the norms governing acting in cases where some actions are indeterminate are secondary, and derivative, upon the norms that determine whether or not an act is permissible in the first place.

what is otherwise permissible.[23] Nevertheless, this does not entail that such violations are permissible. It can be impermissible to expose someone to risk of harm and yet such exposures might warrant milder reactions than actual imposition of harm. Just like *risks* of harm, the moral significance of how our actions will be understood by others is derivative of other ethical concerns. In this sense, the first and second view agree: it is only because *harming* is morally wrong that *appearing to harm* is also problematic. The first view sees this as a separate concern, while the second incorporates this concern back into our judgement of what is permissible.

I do not adjudicate here as to which view ought to be adopted. The purpose of this paper is to put the importance of how our actions will be interpreted and the value of normative signalling squarely on the moral map. But to speculate, I believe that normative signalling (just like other moral norms such as beneficence) will sometimes be so important—and that failures to signal so harmful—that it is required, while other times it will be what we ought to do or even supererogatory. Which it is will depend on several factors: the positive consequences of signalling; how reasonable the mistaken belief of the observer is; the seriousness of the harm if you had in fact done what the observer believes you to have done; the smallness of the cost to the agent to avoid that ambiguity; the negative consequences of that mistaken belief; what we might owe to certain kinds of observers; and so on. When signalling is easy, the downstream harms great and the concerns reasonable, it may well amount to a moral requirement to signal.[24] And more broadly: the cheaper it is, the greater the potential harms, the more reasonable the concerns, then the stronger the agent's responsibilities are. Again, this is similar to other concerns such as putting others at risk of harm: when the risk to others is so high or the costs of avoiding that risk so small, avoiding the imposition of that risk is morally required. However, it can also be supererogatory to avoid putting another at risk when the risks are small and the costs are high.

I therefore take it to be plausible that the strength of the injunction to normatively signal lies on a spectrum that runs from supererogatory to morally obligatory, just

[23] It might also be objected that making it required to take into account the mistaken beliefs of others is a problematically high *rational* bar for agents. Indeed, other requirements—like articulacy on virtue ethics accounts—have been criticised for being too rationally rigorous. Nevertheless, I believe that any rational bar in the case of normative signalling is relatively low and that the value and limits of normative signalling are sufficient to justify any rationalism needed on the part of the agent. The agent, in the case of normative signalling, need not be aware of *why* actions are permissible or impermissible, or what makes them good; and they need not be able to articulate this. Instead, all they need is to be aware that certain courses of action simply *are* permissible and be able to predict that their behaviour might nevertheless *appear* impermissible to a reasonable observer. This does demand that they take into account points of view that are not their own, including the mistaken beliefs of others. However, while this might be a rational cost to the agent (it is epistemic work after all), as I argued in §4, the moral value of doing so (within the limits set out in §5) justifies this. I acknowledge that I assume from the start that it is not too much to ask, rationally speaking, that an agent evaluate courses of action as permissible or impermissible (and do only what is permissible). All I add here is the extra step of taking into account whether they are also *appearing* to do what is permissible. They would only need to determine which permissible courses of action would have the same points of observation to a reasonable observer as impermissible courses of action. I thank the reviewer for this journal who prompted me to address this concern.

[24] Note that in my bracketing off the problem of non-conscientious, deceptive agents, I have also set aside the issue of finding a signal that is hard to fake. As noted earlier, in some circumstances, more costly signals can be more honest signals, which might affect whether that cost is required for an agent to undertake in order to avoid moral ambiguity.

like many other morally significant concerns. Nevertheless, except in the relatively few cases where it is all-things-considered morally better *not* to signal, it is morally better to signal.

7 Conclusion

Spence, who popularised Signalling Theory and won a Nobel Prize in economics for his efforts, was asked by 'a somewhat incredulous questioner (actually, a journalist) whether it was true that you could be awarded the Nobel Prize in Economics for simply noticing that there are markets in which certain participants do not know certain things that others in the market do know' (Spence 2002: 435). Spence responded that Signalling Theory is significant because it provides a systematic 'attempt to capture *informational* aspects of market structure' (Spence 2002: 435). The value of Signalling Theory is thus in part because it is able to expose and explain a common thread that runs through a wide range of economic and social interactions: the existence and resolution of informational asymmetries (Connelly et al. 2011: 42). This paper is valuable for a similar reason: it provides a systematic exploration of some of the ethical dimensions of informational asymmetries, specifically how moral ambiguity is morally consequential and thus, in turn, how avoiding moral ambiguity can be a matter of moral responsibility. As such, it uses ideas from Signalling Theory (as has been done in other disciplines, see (Smith and Bird 2005)) to provide a framework that brings together a wide range of otherwise disparate examples, and demonstrates that the ethical issues in all of them arise from the epistemic mediation of other people's interpretation of our actions: specifically, the informational asymmetries that emerge because, as with the fitness and health of peacocks, or the poisonousness of frogs, or the underlying capabilities of a job candidate, there is something that is not possible to observe directly in moral interactions, namely, the conscientiousness of the agent.

Thus, this paper, by thinking about the informational value of moral actions, adds to traditional models of moral decision-making in philosophy. Even when the long-held assumption of perfect information has been challenged (as it has been in recent years), the focus has been to determine how the *agent's* ignorance, false beliefs or uncertainty ought to inform what that agent subjectively ought to do. This paper asks how the ignorance, false beliefs and uncertainty of *others* ought to inform our moral decision-making. Given that our actions are always transparent to ourselves, it is easy to overlook this question and act as if our actions are as transparent to others. Yet, just as we have grappled with the impact of our own epistemic limits, we must also address the epistemic limits of others with respect to our own conduct. We must determine what costs the agent ought to bear to protect others from the costs and harms of moral ambiguity. This demands that the agent takes into account the reasonableness, likelihood and consequences of the mistaken beliefs of others that arise from and concern the agent's own conduct. The informational value of ethical behaviour is an additional moral consideration that constrains what we ought to do. It is not enough to avoid what is impermissible; we must also care about the *perception* of impermissibility. This idea might initially have rankled: after all, why should we care about appearances or what others mistakenly think about us, especially if

we know that we have only done what is permissible? However, this has been a key consideration in other domains, such as religious practices (as in the *Shabbat* case) or classically, within the legal system. This paper has shown that we should care about it in ethics too: that the appearance of doing wrong, even when we haven't, is of serious moral concern.

Declarations

Conflict of interest I have no funding or conflicts of interest to declare. I did not receive support from any organisation in the preparation of this manuscript and have no relevant financial or non-financial interests to disclose.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Artiga, Marc, and Cédric Paternotte. 2017. Deception: A functional account. *Philosophical Studies* 175 (3): 579–600. https://doi.org/10.1007/s11098-017-0883-8.

Bagwell, Kyle. 1995. Commitment and observability in games. *Games and Economic Behavior* 8 (2): 271–280. https://doi.org/10.1016/S0899-8256(05)80001-6.

Benn, Claire. 2017. Supererogation, optionality and cost. *Philosophical Studies* 175 (10): 2399–2417. https://doi.org/10.1007/s11098-017-0965-7.

Bolinger, Renée Jorgensen. 2020. The moral grounds of reasonably mistaken self- defense. *Philosophy and Phenomenological Research* 103(1): 140–156. https://doi.org/10.1111/phpr.12705.

Brandt, Richard. 1959. *Ethical Theory*. New Jersey: Prentice-Hall.

Bulbulia, Joseph, Quentin Atkinson, Russell Gray, and Simon Greenhill. 2013. Why do religious cultures evolve slowly? The cultural evolution of cooperative calling and the historical study of religions. In *Mind, Morality and Magic: Cognitive Science Approaches in Biblical Studies*, ed. István Czachesz and Risto Uro, 197–212. Durham: Acumen.

Bulbulia, Joseph, and Uffe Schjoedt. 2010. Religious culture and cooperative prediction under risk: Perspectives from social neuroscience. *Religion and Reason: Religion, Economy, and Cooperation* 49:35–60. https://doi.org/10.1515/9783110246339.35.

Calhoun, Cheshire. 1995. Standing for something. *The Journal of Philosophy* 92 (5): 235–260.

Chait, Rabbi Baruch. 1992. *The 39 avoth melcha of shabbath*. B&B Septimus Educational Publications.

Chisholm, Roderick M., and Thomas D. Feehan. 1977. The intent to deceive. *The Journal of Philosophy* 74 (3): 143–159.

Coady, Cecil Anthony John. 1992. *Testimony: A Philosophical Study*. Clarendon Press.

Connelly, Brian L., S. Trevis Certo, R. Duane Ireland, and Christopher R. Reutzel. 2011. Signaling theory: A review and assessment. *Journal of Management* 37 (1): 39–67. https://doi.org/10.1177/0149206310388419.

Ewing, Alfred C. 1947. *The Definition of Good*. Routledge.

Fallis, D., and P. J. Lewis. 2021. Animal deception and the content of signals. *Studies in History and Philosophy of Science* 87:114–124. https://doi.org/10.1016/j.shpsa.2021.03.004.

Fischer, Emma P., Martin C. Fischer, David Grass, Isaac Henrion, Warren S. Warren, and Eric Westman. 2020. Low-cost measurement of face mask efficacy for filtering expelled droplets during speech. *Science Advances* 6 (36): eabd3083-eabd3083. https://doi.org/10.1126/sciadv.abd3083.
Fricker, Elizabeth. 1994. Against gullibility. In *Knowing from Words*, ed. B. K. Matilal and A. Chakrabarti, 125–161. Kluwer Academic Publishers.
Gibbard, Allan. 2005. Truth and correct belief. *Philosophical Issues* 15:338–350.
Grafen, Alan. 1990. Biological signals as handicaps. *Journal of Theoretical Biology* 144 (4): 517–546. https://doi.org/10.1016/S0022-5193(05)80088-8.
Hansson, S. O. 2013. *The Ethics of Risk: Ethical Analysis in an Uncertain World*. New York: Palgrave Macmillan.
Hayden, B. 2009. Funerals as feasts: Why are they so important? *Cambridge Archaeological Journal* 19 (1): 29–52. https://doi.org/10.1017/S095977430900002X.
Isenberg, Arnold. 1964. Deontology and the ethics of lying. *Philosophy and Phenomenological Research* 24(4): 463–480. https://doi.org/10.2307/2104756.
Jackson, Frank. 1986. A probabilistic approach to moral responsibility. In *Logic, Methodology and Philosophy of Science VII*, ed. Ruth Barcan Marcus, Georg J. W. Dorn and Paul Weingartner, 351–365. Amsterdam: North-Holland.
Jackson, Frank, and Robert Pargetter. 1986. Oughts, options, and actualism. *Philosophical Review* 95:233–255.
Jackson, Frank, and Michael Smith. 2006. Absolutist moral theories and uncertainty. *Journal of Philosophy* 103 (6): 267–283.
Johnstone, Rufus A. 1997. The Evolution of Animal Signals. In *155–178*, ed. John R. Krebs and Nicholas B. Davies. Oxford: Blackwell.
Krstić, Vladimir. 2023. We should move on from signalling-based analyses of biological deception. *Erkenntnis* 90 (2): 545–565. https://doi.org/10.1007/s10670-023-00719-x.
Levy, Neil. 2020. Virtue signalling is virtuous. *Synthese*. https://doi.org/10.1007/s11229-020-02653-9.
Maan, Martine E., and Molly E. Cummings. 2012. Poison frog colors are honest signals of toxicity, particularly for bird predators. *The American Naturalist* 179 (1): E1–E14.
Mahon, James Edwin. 2016. The Definition of Lying and Deception. In *The Stanford Encyclopedia of Philosophy*, edited by Edward Zalta.
Miyagawa, Eiichi, Yasuyuki Miyahara, and Tadashi Sekiguchi. 2008. The folk theorem for repeated games with observation costs. *Journal of Economic Theory* 139 (1): 192–221. https://doi.org/10.1016/j.jet.2007.04.001.
Moberg, Dennis J. 2000. Role models and moral exemplars: How do employees acquire virtues by observing others? *Business Ethics Quarterly* 10 (3): 675–696.
Moss, Aron. 2011. "Does What I Appear to Be Doing Matter?". https://www.chabad.org/library/article_cdo/aid/1454116/jewish/Does-What-I-Appear-to-Be-ng-Matter.htm.
Noggle, Robert. 1996. Manipulative actions: A conceptual and moral analysis. *American Philosophical Quarterly* 33 (1): 43–55.
Oberdiek, J. 2017. *Imposing Risk: A Normative Framework*. Oxford: Oxford University Press.
Parfit, Derek. 1987. *Reasons and Persons*. Oxford: Oxford University Press.
Primoratz, Igor. 1984. Lying and the "methods of ethics." *International Studies in Philosophy* 16:35–57.
Ripstein, Arthur. 1996. Self defense and equal protection. *University of Pittsburgh Law Review* 57:685–724.
Scanlon, T. M. 1998. *What We Owe to Each Other*. Cambridge, Mass.: Harvard University Press.
Siegler, F. A. 1966. Lying. *American Philosophical Quarterly* 3:128–136.
Skyrms, Brian. 2010. *Signals: Evolution, learning, and information*. Oxford: Oxford University Press.
Smith, Eric A., and Rebecca Bliege Bird. 2003. Costly signaling and prosocial behavior. *Strong reciprocity: roots of cooperation and exchange*: 1–24.
Smith, Eric A., and Rebecca Bliege Bird. 2005. Signalling theory, strategic interaction, and symbolic capital. *Current Anthropology* 46 (2): 221–248.
Smith, John Maynard. 1991. Honest signalling: the Phillip Sydney game. Animal Behavior 42: 1034-1035.
Smith, John Maynard, and D. G. C. Harper. 1995. Animal signals: Models and terminology. *Journal of Theoretical Biology* 177 (3): 305–311. https://doi.org/10.1006/jtbi.1995.0248.
Spence, Michael. 1973. Job market signaling. *Quarterly Journal of Economics* 87 (3): 355–374.
Spence, Michael. 2002. Signaling in retrospect and the informational structure of markets. *The American Economic Review* 92 (3): 434–459.
Susser, Daniel, Beate Roessler, and Helen Nissenbaum. 2019. Online manipulation: Hidden influences in a digital world. *Georgetown Law Technology Review* 4 (1): 1–52.

Thomson, Judith Jarvis. 1986. *Rights, Restitution and Risk: Essays in Moral Philosophy*. Cambridge, Mass.: Harvard University Press.

Tosi, Justin and Brandon Warmke. 2016. Moral grandstanding. *Philosophy & Public Affairs* 44 (3): 197–217.

"Tractate Shabbat." *The William Davidson digital edition of the Koren Noé Talmud*. https://www.sefaria.org/texts/Talmud.

Veblen, Thorstein. 1899. *The Theory of the Leisure Class*. New York: Macmillan.

Weiss, Andrew. 1995. Human capital vs. signalling explanations of wages. *Journal of Economic Perspectives* 9 (4): 133–154. https://doi.org/10.1257/jep.9.4.133.

Yamaguchi, Mana, Adam Smith, and Yohsuke Ohtsubo. 2015. Commitment signals in friendship and romantic relationships. *Evolution and Human Behavior* 36 (6): 467–474. https://doi.org/10.1016/j.evolhumbehav.2015.05.002.

Zahavi, Amotz. 1975. Mate selection: A selection for handicap. *Journal of Theoretical Biology* 53:205–214.

Zimmerman, Michael J. 2009. *Living with Uncertainty: The Moral Significance of Ignorance*. Cambridge University Press.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.