



Honesty and Bad Faith

Nathan Robert Howard, Philosophy, University of Toronto, nathan.howard@utoronto.ca

N. G. Laskowski, Philosophy, University of Maryland, College Park, ngl.philosophy@gmail.com

An appealing account of dishonesty subsumes it under the paradigm of lying. However, the account faces clear trouble from a wide range of cases, including cases of bullshit and brazen dishonesty. Such cases show not only that lying is inessential to dishonesty but also that honesty requires more than the absence of dishonesty. We propose an alternative account grounded in norms associated with games. On our account, dishonesty is better understood in terms of cheating or of breaking the rules of a game; being honest requires upholding those rules.

Honesty and Bad Faith

NATHAN ROBERT HOWARD
Philosophy, University of Toronto

and

N. G. LASKOWSKI
Philosophy, University of Maryland, College Park

What is it to be honest with someone? A natural answer goes something like this: you're honest with them when you're *truthful* with them – that is, when you're deliberately telling them how it is. For similar reasons, it is equally natural to suppose you're dishonest with someone just when and because you're *lying* to them – that is, when you're deliberately being untruthful. The naturalness of these claims suggests the tempting analysis that dishonesty is, essentially, lying and honesty is, essentially, truthfulness.

Philosophical focus on paradigmatic cases of a phenomenon like this is often helpful. Discussions about *beneficence*, for example, often point to cases of charitable giving.¹ Likewise, discussions about *misogyny* focus on what Incels and Men's Rights Activists are stirring up in the darker corners of the internet.² Discussions about *slurs* can focus on the sort of words that members of the KKK use.³ Because lying is paradigmatic of dishonesty, this method makes it natural to suppose that all dishonesty is, fundamentally, a kind of lying – and similarly for honesty and truthfulness.

However, excessive focus on paradigmatic examples of a phenomenon can also lead us to mistake their incidental features for essential characteristics. It may seem that leopards essentially have spots until we realize that black panthers are leopards. We argue that the natural answer, which has seemingly tempted all recent *analyses of honesty*, commits this mistake.⁴ While lying is paradigmatic of dishonesty, it is inessential to dishonesty – likewise for truthfulness and honesty.

¹ Consider the evolution of collective thinking that took philosophers from Singer's (1972; 2016) famous shallow pond example to the development of effective altruism, a movement of folks committed to promoting the most good they can. See Berkey (2021) for discussion.

² See Manne 2017.

³ Start with Hom (2008) and subsequent discussion. See Foster (2023) for what a more expansive discussion can do for our theorizing on this topic.

⁴ We'll often use the phrase "analysis of honesty" as shorthand for inquiry into the nature of *both* honesty and dishonesty.

Preliminary challenges to the natural answer rooted in this mistake are easy to spot. For example, it forces us to say that if someone plays a card game honestly, then they play it truthfully. It also forces us to say that unless you lie to someone, you're not, strictly speaking, being dishonest with them. Implications like these move us to reconsider the natural answer.

We think that honesty and dishonesty instead are better understood through *trust*, rather than *truth*. Honesty should instead be understood in terms of the trust between players required to play a game. According to the view we defend, acting honestly is, very roughly, pursuing an activity according to rules you're entrusted to follow. To act dishonestly is to violate that trust by culpably breaking those rules.

We begin with some preliminary challenges to the natural account from cases of bullshit and of brazen theft and cheating. Accommodating these cases in a theory of dishonesty leaves us with the following question: what about these activities – speech (in the case of lies and bullshit), property ownership (in the case of brazen theft) or game playing (in the case of brazen cheating) – allows them to be pursued dishonestly? We argue that pursuing each of these activities involves trusting others to follow the rules associated with the activity. As a result, these activities can be pursued honestly when this trust is vindicated by respecting the relevant rules.

This model allows us to preserve what's natural in the natural answer by arguing that truthfulness is a norm of assertoric speech (or entailed by its norms) without also adopting the natural answer's procrustean limitations, which hamper its explanation of forms of honesty or dishonesty that involve neither truthfulness nor mendacity. The natural answer may nevertheless seem so obviously true as to suggest that its core can be preserved simply by refining it to address these limitations. Indeed, some philosophers have attempted to preserve the natural answer in exactly this way. We criticize these attempts along the way.

I. STAGE SETTING

Contemporary philosophy largely neglects the analysis of honesty.⁵ This is perhaps because many take for granted what we'll call the *familiar-but-flawed analysis* of honesty, the kind of analysis we get by refining the natural answer: “the virtue of

⁵ Consider Miller's (2017, pp. 605–606) recent characterization of honesty's place in scholarship: “No one in philosophy has paid much attention to the virtue of honesty in recent years. We find several recent books about compassion, empathy, and love. Very lively discussions have been had in the journals about modesty and about courage. There are multimillion dollar projects funding new work on humility, on understanding, and on faith.... Yet as far as I can tell, only one article in a philosophy journal has appeared in several decades which discusses this virtue at any length.”

honesty is the virtue of being disposed, centrally and reliably, to not intentionally distort the facts as the agent sees them.”⁶ In a similar spirit, it is claimed, “I will argue that the wrongs of dishonesty and indiscretion are intertwined; where honesty is the virtue of telling the truth, discretion is the virtue of withholding it.”⁷ Likewise, it has been said that “The notion of truth is clearly central to our understanding of honesty.⁸ But we typically think of honesty and truth as merely indirectly related. We think of honesty as directly implying *truthfulness* but not truth itself.” And those who act honestly act from “respect for the right not to be deceived.”⁹ The familiar-but-flawed analysis is often tacitly assumed elsewhere.

Its core is that honesty and dishonesty are forms of truthfulness and lying, respectively. Since (part of) what it is to *lie* is to intentionally deceive,¹⁰ the familiar-but-flawed analysis suggests:

The Lying Model of Dishonesty What it is for S to ϕ dishonestly toward another is for S to ϕ with the intention to deceive them.

The Lying Model captures a natural way of explaining what makes certain actions dishonest, such as Bernie Madoff’s decades-long Ponzi scheme, Lance Armstrong’s doping in professional cycling, or Richard Nixon’s involvement in the Watergate scandal. If your friends and family are anything like ours, conversation about dishonesty will centre on intentional deception.

Inverting the Lying Model suggests that what it is for S to ϕ *honestly* toward another is for S to ϕ without intentional deception – or, as we’ll put it, is for S to ϕ *truthfully* toward them:

The Truthfulness Model of Honesty What it is for S to ϕ honestly toward another is for S to ϕ truthfully.

⁶ Miller 2023, p. 31. As we’ll see, Miller nuances this view over several chapters.

⁷ White 2022, p. 6.

⁸ Dougherty 2024, p. 1.

⁹ Um 2024, p. 292.

¹⁰ See Section 1 of the *Stanford Encyclopedia of Philosophy* entry on lying: <https://plato.stanford.edu/entries/lying-definition/>. We’ll later draw a (pedestrian) connection between lies and norms. An ample body of literature on norms of assertion offers varied perspectives on the rules that govern assertoric speech. At the end of the day, it doesn’t matter which of these accounts is correct. Our account relies only on the existence of such norms and their connection to lying.

While perhaps initially appealing, we'll show how these two models are too narrow. The Lying Model imagines too close a connection between dishonesty and mendacity. Likewise, the Truthfulness Model ties honesty too closely to truthfulness. These ties handicap their accounts of honest or dishonest activities where concern for the truth is irrelevant.

II. TRUTH'S INCIDENTAL RELATIONSHIP TO HONESTY

Some basic, perhaps even obvious, cases show how anchoring an analysis of honesty in truth makes it unacceptably narrow. While these cases challenge the familiar-but-flawed analysis's *extensional* inadequacy – *i.e.*, its failure to account for the full range of honest and dishonest activity – this challenge isn't our focus for, as we will see, the analysis can be tweaked endlessly. Extensional counterexamples are often received by enthusiasts as cause for tinkering with the model rather than for rejecting it. Consequently, we present these cases chiefly to motivate our explanatorily superior alternative. As we'll see, contrasted against this alternative analysis, tweaks to the familiar-but-flawed analysis look unacceptably *ad hoc*. Anchoring honesty and dishonesty in claims about truth is unacceptably narrow because only activities that involve *representation*, broadly understood, can exhibit truth. So it would seem that only things that represent, *representational* acts and so on, can be honest or dishonest.

But this claim is false. When you cheat at a game, you exploit others' good-faith acceptance of the rules of the game. So cheating is dishonest. But cheating doesn't (essentially) involve representation. So it seemingly cannot be connected to truth in the way that the Lying Model requires. Likewise, *theft* is dishonest. But theft doesn't involve representing anything as true or false, accurate or inaccurate. So the idea that dishonesty essentially involves lying is a non-starter because it cannot account for cheating and theft's dishonesty.

Or so you'd be entitled to think. But this elemental difficulty hasn't deterred some. They have tried to understand fundamentally non-representational acts such as cheating and stealing as forms of other wrongful, intentional misrepresentation. They argue, roughly, that stealing is dishonest because it involves the thief making it seem as if they are the rightful owner of the stolen property when they are not the rightful owner.¹¹ Likewise, cheating is dishonest because it involves “distorting the facts of what game is even being played and what its rules are.”¹²

¹¹ Miller 2023, p. 45.

¹² Miller 2017, p. 264.

There are at least two places to resist this response. The first is to question whether cheating and stealing are, necessarily, representational acts in the manner that the proposal requires. We doubt this but this doubt engages in opaque and slippery questions about the nature of representation probably better avoided here. So, we'll merely register some skepticism and oppose the move on different grounds.

A second challenge to subsuming cheating and stealing under the Lying Model comes from *brazen dishonesty*.¹³ Our opponents claim that an act of cheating or stealing is dishonest only when the act involves misrepresenting the rules governing games or property. But this simply isn't true of every case of cheating and stealing. Consider the following:

Shameless Golf Cheat Imagine that Tiger Woods is playing golf with someone notorious for cheating at golf, namely, Donald Trump.¹⁴ On one of the greens, Trump picks up his ball, looks Tiger square in the eye, walks up to the cup, and drops it in. If confronted, Trump would say, "Yeah, I cheated. What are you going to do about it, Paper Tiger Woods?" Tiger may throw up his hands in disgust and despair, but there's nothing he can do.¹⁵

Trump is clearly being dishonest in this case; after all, he's cheating. Yet he's not misrepresenting his cheating to Tiger or anyone else. Since the cheating is brazen, it involves no misrepresentation whatsoever. Consequently, dishonest actions need not involve misrepresenting the content of norms.

The case is particularly illuminating because Trump's aims depend on *not* misrepresenting "what game is even being played and what its rules are" but on representing those rules *as they are* in order to violate them. For example, "suppose" that Trump doesn't cheat to win at golf; he is indifferent to winning. Rather, he cheats to show his social power and impunity. Impunity is shelter from rightful punishment. Consequently, showing Trump's impunity requires that he break the rules, so that he deserves punishment but does not receive it. Misrepresenting the rules of golf stifles this aim, since by clouding the rules, Trump clouds his impunity, and so his show of power.

¹³ Miller (2023, p. 48) discusses similar cases.

¹⁴ For a small sampling of Trump's long history of cheating at golf, see Reilly (2020) and <https://www.palmbeachpost.com/story/sports/2023/02/03/trump-and-golf-fancy-resorts-a-list-partners-cheating-at-highest-level/69857594007/>.

¹⁵ Miller (2021) discusses cheating at golf; though not *brazen* cheating, as we've defined it. Some, such as Roberts and West (2020), might contest our claims.

Similar cases of brazen rule-breaking involve theft, where a Trump-like bandit wants to display his impunity by openly stealing another's property. In the cases that we imagine, the theft does not misrepresent the rules governing property. Indeed, the fact that the bandit wrongfully possesses *another's* property is precisely the point of the bandit's theft, much as the brazen wrongfulness of Trump's cheating is its point. Imagine stealing another's name tag: there's no representation that *another's* name tag is one's rightful property.

Indeed, Trump's recent (as of January 2026) claim of power and control, amounting to a claim of ownership, of Venezuelan oil assets provides a real world analogue. Trump has openly asserted control over assets that are not his and has done so without pretending that the ordinary rules governing property and legal process confer any such entitlement. By executive order, his administration blocked judicial attachments, liens, and creditor claims against Venezuelan oil funds, declaring ordinary legal processes null and void and placing the assets under executive control for policy purposes. No one imagines that these actions aim to misrepresent ownership. Their purpose is rather to display the power to take and hold them regardless. Like stealing another person's name tag, the wrongfulness is not concealed or denied; it is made conspicuous.

Brazen cheating and theft involve no misrepresentation. But they are dishonest: cheating (theft) is dishonest, so when Trump brazenly cheats and steals, he's brazenly dishonest. So misrepresentation is not essential to dishonesty.

Some might prefer to bite the bullet and deny that Trump's cheating is dishonest precisely because it's brazen rather than give up the link between dishonesty and misrepresentation. But it seems plain that cheating a competitor is dishonest. After all, by competing against you, they can expect you to play by the rules – that is, they are entitled to trust you to play by the rules. By cheating, you subvert this expectation. This subversion is dishonest.

Crucially, however, subverting this expectation doesn't involve *falsifying* it, as the familiar-but-flawed analysis would have it. Expecting you to play by the rules is something close to believing that you should play by the rules of the game. And it's certainly true that you should play by the rules of golf when you play golf.

Cheating subverts this expectation, whether it's brazen or not. Now it's reasonable to wonder whether Trump's golf partners *really* expect him to play by the rules. After all, he's a famous cheat. Without this expectation, we lose this provisional explanation of cheating's dishonesty. But we'll deal with this complication, and many others, as we develop this suggestion for, as we'll see, we take this kind of subversion to be the root of dishonesty. It is easily mistaken for a kind of lying, which is precisely the mistake underlying the familiar-but-flawed analysis.

III. TOWARD AN ALTERNATIVE VIEW: THE PROPERTIES COMMON TO HONESTY AND DISHONESTY

None of the challenges we've raised is plainly decisive. Our aim in presenting them isn't direct refutation but rather to illustrate the range of honest and dishonest activities. Any adequate theory of honesty must identify what unifies these activities, making them candidates for honesty or dishonesty.¹⁶ The familiar-but-flawed view holds that this common feature is representation. Our cases cast doubt on this idea. But the most effective rebuttal is a better alternative. In what follows, we develop such an alternative, which accommodates the full range of cases while preserving what is attractive in the natural view.

This alternative model begins by reconsidering the common property shared by honest and dishonest activities. Consider the following case. Suppose that Jim agrees to bring John coffee every morning at 6am for the next month if John gives Jim his tickets to the ballet. Deep down inside, Jim knows that he won't be able to get up that early: he's a night owl not an early bird. On some level, he's fully aware of how his weakness for ballet often leads him to make bad decisions. Only last month he took out an ill-advised second mortgage to fund a week at the Bolshoi. He's done unspeakable things for just one more fix of Swan Lake. Consequently, Jim knows better than to agree to John's arrangement because Jim knows that he'll disappoint. Nevertheless, Jim's love of ballet clouds his judgment with unearned optimism, leading him to agree.

Jim is making a "lying promise," to use Kant's term. He's agreeing to an arrangement that he, at least on some level, knows that he will not fulfil. But characterizing this agreement as a "lying promise" isn't literally correct. Broadly speaking, Jim makes a *promise* only in the sense that he commits himself to something and it's *lying* only in the sense that he knows better than to make the promise, since he knows that he will not follow through, though self-deception somewhat clouds this fact. But lying requires intentional misrepresentation. Because Jim is unintentionally self-deceived, he does not intentionally mislead John.

Likewise, Jim isn't promising anything to John, strictly speaking. Rather, Jim and John enter into an *agreement* in which one person gets coffee and the other gets tickets.¹⁷ Rather than criticize Jim for making a "lying promise" it's closer to the truth to say that Jim enters into an agreement with John in *bad faith*.

¹⁶ Cf. Miller's (2023, p. 24) discussion of the "unification challenge."

¹⁷ See Gilbert (1993) for an argument that promises and agreements are distinct. See de Kenessey (2023) for an account of what links them.

Broadly liberal traditions in moral and political philosophy have long recognized the importance of agreements to what we can properly expect of each other. On the political side, the concept of an *unspoken agreement* underlying legitimate state authority grounds the social contract tradition. On the moral side, it helps explain how moral rules are compatible with – perhaps even a precondition of – human freedom. But we needn't accept either of these lofty ideas to suspect that institutions such as assertion and property and games depend on something like an unspoken agreement.¹⁸

One straightforward argument for this suspicion comes from the practice of promising.¹⁹ When we agree to do something, we create a *pro tanto* duty for ourselves to do it. But this duty relies on the convention of using certain sounds or gestures to express and thereby create the agreement. So, there is a conceptually or metaphysically prior proto-agreement – what we've been calling an unspoken agreement – that certain sounds and gestures constitute promises and so create certain duties.

Why do unspoken agreements have the normative power that they do? This important question is orthogonal to our discussion. Perhaps unspoken agreements can be powerful because they're tied to what we would agree to, hypothetically speaking, were we fully rational. Perhaps they're grounded in what we would agree to if we're pursuing our deepest values. Perhaps they're underwritten by prudential necessity. We needn't decide between these views here.

Rather, we wish only to underscore how unspoken agreements underpin a wide range of actions, such as promising, for we think that only *then* can honesty and dishonesty be properly understood. Consider a familiar Kantian thought experiment, which asks us to envision a world where everyone constantly lies to others. This world seems improbable, if not impossible. This is because everyone would be aware of everyone else's constant deception, preventing us from relying on another's word. But successfully lying to someone requires them to rely on your word. Consequently, lies exist only in societies where people generally tell the truth and expect each other to do so.

Property and games similarly depend on unspoken agreements. In a world of thieves without conventions governing the rightful acquisition and transfer of possessions, property amounts to no more than mere physical possession. Likewise, if neither of the two would-be chess players respects the rules of chess, they are not playing chess but merely playing with chess pieces.

¹⁸ Cf. Goodin 2010.

¹⁹ See Raz (1986, ch. 7) and Narveson (1994), among many others, for this line of argument.

Because neither cheating nor stealing intrinsically involve representation, the common property linking activities with the potential for honesty seems not to be the property of being representational, *contrary to* the natural answer and its advocates. Rather, the common property linking these activities seems to be that they are subject to unspoken agreements about the norms internal to those activities.

It's tempting to rush from this link to a flawed analysis. Perhaps because honest actions are ones governed by norms, one acts dishonestly by violating those norms. When someone lies, for example, they violate an unspoken agreement to assert truths. When they steal, they claim rights of possession that they deny to others. When they play a game, they insist that others uphold the very rules they're breaking. Dishonest action, on a first pass, seems to fit the following pattern:

The Norm Model of Dishonesty What it is for S to ϕ dishonestly toward someone is for S's interaction with them to violate a norm of ϕ -ing.

But this can't be right. Dishonesty is (*pro tanto*) morally objectionable. But mere rule-breaking is not – indeed, rule-breaking can be praiseworthy. This is because rules are metaphysically “cheap.” We can create them out of thin air by conventions, fiat, and contingent social arrangements. These “cheap” rules typically lack what some call the “authoritative normativity” characteristic of moral norms.²⁰ Correspondingly, violating those rules is not morally objectionable. So breaking “cheap” rules can't suffice for dishonesty.

To illustrate the point, imagine cheating at a game like solitaire by moving cards around in a way that's prohibited by the game's rules. Cheating like this isn't dishonest even though it violates the rules that govern solitaire. So rule-breaking doesn't suffice for dishonesty. It's merely necessary. The Norms Model is missing an ingredient.

Contrast cheating at solitaire to cheating at competitive chess. Cheating in the latter case is dishonest but not in the former, yet both involve cheating at a game by breaking “cheap” only weakly “normative” rules. So what's the difference? Answering this question is the key to finding the missing ingredient.

We propose that the difference is the following. When you intentionally cheat at solitaire, you've given yourself permission to do so, in some sense. You're both the judge and jury. If you've permitted yourself to cheat, you're not entitled to expect yourself

²⁰ This kind of normativity is also sometimes called “genuine” or “robust.” However, unlike some, we are skeptical of the philosophical significance of the distinction supposedly marked by these terms (Howard and Laskowski forthcoming).

not to cheat. But that's not true when you cheat at competitive chess. Your opponent normally expects you to follow the rules in a way that you're not expecting yourself to follow the rules when cheating at solitaire. That's because when you participate in an activity governed by an unspoken agreement, you participate in the unspoken agreement. By choosing to play chess, you participate in the agreement governing chess, which requires you to play by the rules. Similarly for, e.g., assertion and games.²¹

Both chess and solitaire can give rise to cheating, but only in the chess case do we see an added element of *counting on* or *trusting* the other person to follow those rules. In the case of solitaire, you exempt yourself from the unspoken agreement that governs solitaire. Conversely, chess opponents don't typically exempt you from this agreement. Consequently, when you persist in competing at chess while cheating, you're willfully violating the rules of chess while simultaneously relying on the agreement between chess players to uphold those rules. In short, you're playing chess in *bad faith*. And bad faith, we think, is the essence of dishonesty.

IV. THE TRUST MODEL OF DISHONESTY

A. Bad Faith

Here is the big idea, in broad strokes: only activities governed by certain norms can be assessed for honesty or dishonesty. One performs an activity honestly by upholding the norms that govern it; culpable violations of those norms are dishonest. Such violations are dishonest because they breach others' *entitlement to trust* that you will uphold the relevant norms, given that you are engaging in an activity structured by them. Violating that trust constitutes a form of *bad faith*, while upholding the norms vindicates it and thereby manifests *good faith*. In short, honesty and dishonesty are best understood in terms of good- and bad-faith participation in these unspoken agreements.

²¹ Indeed, since unspoken agreement about norms makes activities like assertion and chess possible, it's not a stretch to suggest that the relevant sense of dependence between these rules and activities is *constitution* – part of what makes a token action an instance of assertion, for example, is that the action is inherently subject to the truth norm (or any other norm that entails it, such as the knowledge norm – see Williamson 2000, pp. 238–269). If so, it's clear that one is subject to the truth norm when one asserts a claim: an utterance counts as an assertion *only if* it is subject to the truth norm. However, the reader need not accept this further claim provided they agree that, e.g., by playing chess, one participates in the unspoken agreement governing its norms, and so one is (defeasibly) subject to the norms of chess by playing it.

There's a lot to unpack here, so let's focus on the notion of *bad faith*. Sartre's discussion of bad faith is probably the most prominent in philosophy. But Sartre uses the term in a technical and narrow sense to characterize a kind of self-deception about one's existential freedom. This doesn't match the common usage according to which, for example, a widget-maker signs an agreement in bad faith if they know that they cannot uphold it. The widget-maker can act in bad faith in this sense even if they fully grasp their "existential freedom." So, Sartre's ideas will not serve us here.

Rather, what it is for someone to act in bad faith, in the colloquial sense, is for them to freely participate in an agreement while intentionally refusing to follow it. Because bad faith presupposes agreement, it offers a promising lens through which to examine the nature of dishonesty for, as we've just suggested, honest activities rest on unspoken agreements. When someone uses an agreement to achieve their goals without abiding by it, they exploit both the agreement and its other participants. This exploitation is what makes participation in the agreement *morally* objectionable in the particular manner of dishonesty – something that, as we've seen, cannot be secured by mere rule-breaking. Thus, when someone engages in these activities while rejecting the rules that make them possible, they act in bad faith and so act dishonestly.

Some examples will help make these abstract claims concrete. By intentionally lying, one breaks a norm that governs assertion. By breaking the norm, you fail to live up to the unspoken agreement underlying the norms that govern assertion. Moreover, because you've intentionally chosen to assert, you've entered into this unspoken agreement. But you've entered into the agreement with the intention to break it, by intending to lie. So your lying counts as bad faith because you've entered into an agreement while planning to break it.

Likewise, by intentionally cheating, you break a norm that governs the game. By breaking the norm, you fail to live up to the unspoken agreement underlying the game's governing norms. Moreover, because you've intentionally chosen to play, you've entered into this unspoken agreement. But you've entered into the agreement with the intention to break it, by intending to cheat. So your cheating counts as bad faith because you've entered into an agreement while planning to break it.

Similarly, by intentionally stealing, you break a norm that governs property. By breaking the norm, you fail to live up to the unspoken agreement underlying property's governing norms. Moreover, because you've intentionally chosen to participate in the system of property – by claiming the stolen goods as your own – you've entered into this unspoken agreement. But you've entered into the agreement with the intention to break it, by intending to steal. So your stealing counts as bad faith because you've entered into an agreement while planning to break it.

So understood, bad faith can, perhaps, be described as a kind of self-misrepresentation. For example, by lying you represent yourself as accepting the rules that govern assertion. But you don't end up following those rules. Because bad faith can be described as creating a kind of false expectation, and so a kind of *misrepresentation*, our proposal echoes the "natural answer." But this similarity is superficial. Properly distinguishing the proposed view from the natural answer therefore requires distinguishing two types of expectations.²²

In a *descriptive* sense, we sometimes expect things to behave in a certain way, akin to a future-directed belief. Evidence for these descriptive expectations is often inductive or stochastic. Their content is typically descriptive. For example, when someone expects the sun to rise tomorrow, they don't think that the sun *ought* to rise tomorrow in any non-epistemic sense.

By contrast, we can also expect things – typically individuals but also institutions – to behave in a certain way in a *normative sense*. The content of a normative expectation is typically normative. For example, when someone expects their child to say please and thank-you in this sense, they broadly think that they ought to say please and thank-you.

We can distinguish these two kinds of expectation by their connections to different clusters of non-cognitive attitudes. For example, failures of descriptive expectations characteristically license surprise or bafflement. Failures of normative expectation, in contrast, typically license disappointment or outrage. Of course you can be both baffled and outraged at the same time, so these two forms of expectation are compatible.

The natural answer focuses on descriptive expectations. In particular, it focuses on how lying creates false descriptive expectations – this is what makes many lies deceptive. But lying doesn't automatically falsify normative expectations: people remain entitled to expect you not to lie, in the sense that they remain entitled to believe that you ought not lie, even when you do. Therefore, instead of *falsifying* the normative expectation that speakers should be sincere or truthful, lying *violates* that expectation (*de re*, not necessarily *de dicto* but perhaps *non-specific de re*).²³ What characterizes dishonesty isn't *false* expectation, contrary to the natural answer, but *violated* expectation.

To illustrate, consider that when Jim enters into an agreement with John in bad faith, Jim violates John's normative expectation for Jim to deliver John's coffee – that's what

²² For more on this distinction between kinds of expectation, see Jones (2004) and Hawley (2014).

²³ See Howard (2021) for an account of the moral significance of the non-specific *de re* in moral intentions.

makes Jim's participation in the agreement dishonest. But Jim's bad faith doesn't – or needn't – falsify any of John's descriptive expectations. For example, John might know full well that Jim is a night owl, so that he's unlikely to rise to the challenge of getting the coffee on time. He doesn't *project* or *forecast* that Jim will follow through, he merely hopes that he will. In that case, Jim doesn't give John a false descriptive expectation but still acts dishonestly by agreeing in bad faith.

These two senses of expectation correspond to two senses of misrepresentation. Jim doesn't misrepresent himself to John in the sense of *falsifying* John's descriptive expectations about how he, Jim, will behave. After all, John might know that Jim is unlikely to hold up his end of the bargain. But Jim's breach of their initial agreement misrepresents his commitment to the agreement in the sense of *violating* John's normative expectations. This kind of misrepresentation is not a species of lying or mendacity. Jim's dishonesty lies in culpably breaking their agreement, not in making one of John's beliefs false. Consequently, although bad faith might involve a kind of misrepresentation, it's not the kind of misrepresentation central to the “natural answer.”

Our proposal also resembles the “natural answer” in a second respect. The latter not only holds that lying is the *paradigmatic* example of dishonesty but also that the fact that lying is dishonest is *fundamental*, for it analyzes other instances of dishonesty as forms of lying, broadly understood. The dishonesty of lying is a kind of basic posit underlying the natural answer.

While our proposal is consistent with the idea that lying is the paradigmatic example of dishonesty, it offers an analysis both of why lying is dishonest and of why lying is a paradigmatic form of dishonesty. Truth-telling is widely recognized as *an* if not *the* internal norm of assertion. Liars exploit this norm while participating in the agreement to uphold it by participating in declarative speech. This observation allows us to subsume the dishonesty of lies to a more basic and general conception of dishonesty: lying culpably violates others' normative expectations borne of reliance on assertion's internal standards, *i.e.*, their expectation for you to tell the truth or their entitlement to trust you to tell it.²⁴ Consequently, lying's dishonesty is not a brute fact, but a special case of the more general analysis of dishonesty.

Second, speech is a paradigm of social engagement. If an act is dishonest because it violates an important kind of social trust, then a type of speech that breaks that

²⁴ We favor the account of trust of Hawley (2014, p. 10) according to which “To trust someone to do something is to believe that she has a commitment to doing it, and to rely upon her to meet that commitment.”

trust would naturally emerge as paradigmatic of dishonesty, given that speech is paradigmatic of social interaction more generally. Consequently, our proposal can not only explain why the natural answer seems so natural; it can also explain its motivation by ratifying the centrality of lying to the genus of dishonest action, more broadly.

Here's our official statement of this view:

The Trust Model of Dishonesty What it is for S to ϕ dishonestly toward R is for S's ϕ -ing to *violate* the trust to which R is entitled under the unspoken agreement that governs ϕ -ing, by *culpably breaking* its norms.

The Trust Model improves on the Norms Model. The institution of property relies on an agreement that one's possessions should be one's property, with allowances for borrowing, loans, renting, and so on. By possessing something that isn't one's property, one breaks this agreement. The Kantian thought highlights how thieves simultaneously claim the rights of property, and so participate in the institution of property, yet try to exempt themselves from its duties. By wrongfully possessing another's property, the thief enters into a kind of bad faith agreement with others concerning the norms that regulate property, violating a special kind of trust to which others are entitled by the institution of property.

Likewise for cheating in a game. Games are goal-oriented activities in which participants voluntarily agree to abide by rules that artificially restrict the most efficient methods of achieving those goals.²⁵ For instance, the quickest way to get the ball in the hole in golf would be to pick it up and drop it in. However, golfers are prohibited from doing so by the rules. The existence of these rules depends on players' expectations that others will abide by them.

The Trust Model also accounts for the challenging cases presented at the outset of the paper. When someone asserts, you're entitled to trust them to be sincere, and bullshit is insincere speech.²⁶ When you become aware of your intention to cheat at solitaire, you are in effect giving yourself permission to cheat. Consequently, you are not entitled to trust yourself *not* to cheat. So, you don't act dishonestly. This contrasts with *Shameless Golf Cheat* where Tiger is aware of Trump's intention to cheat, but Tiger does *not* permit Trump to cheat. Consequently, Tiger is entitled to trust that Trump won't cheat, so Trump acts with brazen dishonesty by cheating.

A second attraction of this approach is its *positive* analysis of honesty – it offers an appealing suggestion about what it takes to be honest, beyond not being dishonest:

²⁵ For an influential account of games along these lines, see Suits (1978).

²⁶ See Frankfurt (2005) for an influential theory of bullshit.

The Trust Model of Honesty What it is for S to ϕ honestly toward R is for S's ϕ -ing to vindicate the trust to which R is entitled under the unspoken agreement that governs ϕ -ing, by following its norms.

Famously, golfer Bobby Jones lost an important golf tournament when he logged a penalty that no one else witnessed. This was an honest action. On the Trust Model, it's an honest action because Jones vindicates the trust of his fellow golfers by following the rules of the game when he logs his penalty stroke. This seems like exactly the right thing to say about Jones's action.

It's also true that Bobby Jones' representation of his score was truthful. So, the familiar-but-flawed account also accords with this paradigmatic case of being honest (if it didn't, it wouldn't even be a contender). But it faces trouble from games like poker because it takes truthfulness as part of the brute, primitive analysis of honesty. Poker players often bluff, misrepresenting the strength of their cards. Do you play poker dishonestly by bluffing? Clearly not. But the familiar-but-flawed analysis predicts otherwise, since it holds that misrepresentation, "lying" suffices for dishonesty.

Proponents of the analysis are aware of this obvious challenge. But rather than abandon the analysis, they take it to motivate a seemingly minor revision. For example, the following analysis has been proposed:

The Restricted Lying Model of Dishonesty (RLM) What it is for S to ϕ dishonestly toward R, when in "factual environments which agents take to involve expectations about the way things actually are" is for S to ϕ with an intention to misrepresent or deceive.²⁷

But what licenses gerrymandering the model in this way besides the need to avoid a counterexample? Why is it that an analysis of honesty can plausibly remain silent about so-called "factual environments"? What gives rise to such environments? Preserving the model in this way is akin to preserving utilitarianism by restricting the scope of utilitarianism to exclude inconvenient implications involving trolleys or the harvesting of organs.

By contrast, the Trust Model easily explains why it's ok to misrepresent one's position in poker but not in golf. That's because accurately representing your play is a norm of golf. But it is not a norm of poker. As a result, others are entitled to trust you to be truthful when playing golf but not poker.

²⁷ Miller 2023, p. 71.

When we participate in social practices like assertion, promising, game-playing, or property exchange, we implicitly invoke norms that make those practices possible. These norms do double duty. Not only do they help characterize the relevant practices. They also structure what we expect of each other, licensing a reciprocal trust between participants in the practice. For example, speakers trust one another to assert only what they believe; players trust one another to follow the rules; co-owners trust each other to respect property boundaries. Culpably violating these norms while still presenting oneself as a participant in the practice exploits the trust generated by those norms. This kind of bad faith is what makes such behavior dishonest – not because it distorts facts, but because it exploits the normative expectations to which others are entitled within the practice.

B Nuances

We've said that cheating is dishonest. But imagine that we unfreeze baseball legend Ted Williams and, once his confusion and fear of the modern world abates, we trot him onto the field and ask him to do his thing. The rules of baseball have changed since Ted's time. So, he's likely to unintentionally break them, violating, for example, the new pitch clock requiring pitchers to pitch the ball within a certain amount of time after stepping on the mound. But he's not playing baseball in bad faith. So, if breaking a game's rules is cheating and cheating is dishonest, then you can be dishonest without bad faith.

However, we don't really think Ted is being dishonest here. We doubt that breaking a game's rules suffices for cheating, we think only *culpable* rule-breaking does. Culpability here requires awareness of the rules or negligent ignorance. So, Ted isn't being dishonest.

A similar point applies in cases of unintentional misrepresentation. Suppose that Rosa is giving a museum tour and, relying on reputable training materials that nevertheless contain undiscovered falsehoods, misattributes a particular painting to Vermeer. Unbeknownst to her, the painting was recently reattributed to a lesser-known Dutch artist. The members of the tour are misled by her testimony. Nevertheless, Rosa is not being dishonest. She believes what she says and makes the claim in good faith, trusting her training and aiming to fulfill her role as a guide. The mere fact that she is mistaken does not make her dishonest. After all, dishonesty requires that she culpably betray the trust that museumgoers place in her – either by knowing that her claim is unsupported or by being negligent in failing to update her knowledge. As with Ted Williams, the absence of culpable rule- or norm-violation means that, despite the factual error, no dishonesty occurs. In both cases, the norms

characteristic of the activity are violated, but without culpability, and hence without dishonesty.

Our account implies that were Rosa's deception culpable, she *would* be dishonest. For example, rather than using apparently reliable (but ultimately misleading) training materials to guide her tour, she bases her tour on claims about the Dutch Golden Age found in her favorite fictional novel about Dutch Masters, which departs from actual art history.²⁸ In *that* case, it would be dishonest for Rosa to present herself as giving a tour of the museum. Participants on a museum tour are entitled to trust their guide to present the facts, at least to the best of the guide's ability. Rosa culpably violates that trust by using a novel rather than reputable training materials. So, when Rosa misattributes the painting's artist, the misattribution here is dishonest. At a minimum, whether Rosa's misattribution is dishonest is *correlated* with whether she violates the relevant party's trust to be a good tour guide. To this, we merely add the claim that the dishonesty is the product of the violation.

The cases of dishonesty we've used to illustrate our account feature relative agreement between the parties about the relevant rules, *e.g.*, Jim and John agree about the content of their arrangement and playing a game like golf or chess requires knowing the rules.²⁹ But matters are not always so clear. For example, two parties may take an activity to involve different rules. Some groups require calling each ball before sinking it when playing pool whereas some only require calling the 8-ball. Does a player of the latter group act dishonestly when they don't call each ball when they play the other group? It seems not, provided their ignorance about which rules apply is non-culpable. Consequently, non-culpable ignorance about a practice's norms, born of reasonable interpersonal variation in the interpretation and application of those norms, often exculpates rule-breakers. Conversely, if a rule-breaker seeks to exploit this confusion – for example, by strategically calling their shot only after the ball's direction is clear – then they are plainly not exculpatorily ignorant of the norms. After all, they can exploit the norms only because they *know* that they're in flux. So, they may be acting dishonestly. Something similar holds for practices governed by vague or indeterminate norms: one breaks the rules honestly only if one's confusion is non-culpable.

Brazen rule-breaking, we've argued, is dishonest when it violates a certain kind of trust. Nevertheless, we don't claim that *all* brazen rule-breaking is dishonest. For example, suppose that you, a non-golfer, join your friends for a round simply for the

²⁸ With thanks to a historically-informed referee for a minor correction involving the periodization of Dutch art.

²⁹ Many of these challenges come to us from a referee for *Philosophy and Phenomenological Research*.

camaraderie. You move balls, don't count your swings, pocket your ball when you get tired, and generally break the rules at your convenience. You're not dishonest though you're brazenly cheating. What's the difference between this case and the case of Trump?

Given the setup, it's natural to suppose your friends have given you leave to cheat, just as you do for yourself when you cheat at solitaire. Because they've let you cheat, they're not entitled to trust you not to. Likewise, when Tiger doesn't give Trump leave to cheat, Trump's cheating is dishonest. If Tiger did, it wouldn't be.

The case raises the important question of how and when we permit others to break the rules. For example, group projects in school are governed by norms. It's natural to divide work roughly evenly, for members to have a say in what part they accomplish, and for members to do their part, though circumstances often force us to negotiate. When a lazy group member says, "Don't rely on me; I'm irresponsible and disorganized!" and proceeds to contribute poorly to the project, are they being dishonest? It seems not. But the norms of collaboration seem to entitle participants to trust that others will do their part, and the lazy member's laziness seems to violate that trust, presenting a challenge for the Trust Model.

Nevertheless, though the norms of collaboration normally license trusting others to do their part, this entitlement is defeasible. If the members continue collaborating with the lazy member, particularly if they do not persistently hold the lazy member to account, then the norms governing collaboration come under dispute. By allowing the lazy member to continue collaborating without criticizing them, the lazy member could reasonably expect to have been given leave to be lazy by the group. This ignorance exculpates their rule-breaking like how ignorance of which set of pool rules governs a game can exculpate rule-breaking, as discussed immediately above.

Conversely, if the group continues to criticize the free-riding, lazy member – insisting that they either contribute or quit – then the lazy member acts in bad faith and, in doing so, dishonestly exploits their team. This scenario arises only when the group is powerless to change its membership, as when a teacher forbids regrouping. In such cases the group is vulnerable to the free-rider's exploitation, and that exploitation is dishonest. After all, the free-rider claims credit they have not earned. Because everything is out in the open – the free-rider's intentional exploitation and the group's explicit criticism – there is no deception here. But, as we have seen, not all dishonesty involves deception. In this respect the free-rider resembles a brazen thief: someone who openly takes what they do not deserve. Theft is dishonest even when it is brazen, and likewise, wrongfully appropriating the fruits of others' labour as one's own, even when done openly, is dishonest.

Our explanation of why a dishonest action is (*pro tanto*) morally objectionable is that it violates a trust to which others are entitled by certain rules associated with that action. These rules can be *benign*, such as the rules that govern golf or chess. They can even be *beneficial*, such as the rules that govern a practice of considerable social good, such as assertion. But the rules that govern a practice can also be *unjust*. In those cases, facts about honesty may not seem closely tied to trust based on unjust rules. For example, while physicians are told to do no harm, they also have duties of disclosure to their patients. These two duties conflict when sharing a diagnosis could be harmful, *e.g.*, a terminal diagnosis. Prior to widespread recognition of the value of patient autonomy in the 1960s onward, physicians were known to withhold terminal diagnoses. Consequently, patients were not entitled to trust their physicians to disclose these diagnoses, given the rules that governed disclosure at that time did not require it in all cases. Withholding a terminal diagnosis did not therefore violate that trust. Nevertheless, it seems dishonest to withhold such a diagnosis.

However, whether the failure to disclose is in fact dishonest depends in part on how the terminal patient regards the rules governing medical practice and what they expect of their physician as a result. For example, if the patient accepts those rules, then it seems natural to deny that a failure to disclose is dishonest. It may not respect the patient's autonomy (or it may, in fact, if the patient endorses the rules), but it is no more dishonest than a failure to disclose one's hand in poker. On the other hand, if the patient rejects rules that permit physicians to withhold distressing diagnoses, then they do not participate in the local practice of medical disclosure, at least as governed by the rules of 1950s medical practice. Instead, they *are* entitled to trust their physician to disclose their diagnosis, grounded in a broader expectation for their physician to respect their autonomy. These two cases represent two extremes on a continuum. Surely most cases will fall in the middle, and judgments of (dis)honesty will be correspondingly vague. But in this respect our account closely tracks the facts on the ground.

We've advocated for the Trust Model by stressing its explanatory depth. But this explanatory depth also yields explanatory breadth. Some apparent kinds of dishonesty do not involve bad faith. Consider the following case:

Wall Street Suppose that a business is cooking the books.³⁰ The auditors are coming, and the boss knows it. So, the boss asks the accountant to lie about the cooked books, which the accountant promises to do. The accountant lies to the auditors when they show up, thereby fulfilling their promise to the boss.

³⁰ We owe this example to Taya Cohen.

Cases like this, which involve multiple relations to distinct agents, are particularly challenging for the *familiar-but-flawed analysis*. That's because the analysis holds that "the virtue of honesty is the virtue of being disposed, centrally and reliably, to not intentionally distort the facts as the agent sees them." In particular, the analysis is broadly neo-Aristotelian by presupposing that virtues are, first and foremost, *dispositions*.

Wall Street is particularly challenging for dispositional analyses because the analysis requires us to settle whether the accountant has the relevant disposition before we can judge whether their lie is honest. This requirement makes the approach too coarse-grained for either the lie expresses the agent's virtue, or it doesn't. If it does, then neo-Aristotelians face the challenge of explaining why the lie contains an element of vice, given its deception. If it doesn't, then neo-Aristotelians face the converse challenge of explaining why the lie contains an element of virtue. This dilemma pushes us to recognize a certain awkwardness in centering character when accounting for virtue.³¹

The source of this awkwardness is the simple fact that different sets of norms can overlap on a single action. The accountant's testimony to the auditors participates equally in both the practices of speech and promise-keeping. The lie keeps a promise, so it is honest in this respect. But it violates the norms of assertion, so it is dishonest in this respect. But these respects are clearly compatible. So the case leads us to believe that honesty and dishonesty are relational properties between actions and sets of norms.

Dispositions – at least the dispositions thought to underlie Neo-Aristotelian virtues – are not, by contrast, thought of as relational. They are naturally understood as monadic. This is why the Neo-Aristotelian approach has such trouble with the case.

By contrast, actions are especially well-suited to capturing honesty's relational nature for actions often have multiple patients. For example, a single transaction can benefit one party but harm another. The Trust Model incorporates this feature of actions. Just a single action can benefit one person but harm another, a single action can *vindicate* one person's trust while *violating* another's when their trust is grounded in different sets of norms. In that circumstance, our analysis predicts it will be *honest* relative to the first and *dishonest* relative to the second, ratifying common sense about *Wall Street*.³²

Another instance of this difficulty arises when someone "games the system." This happens when someone exploits a system of laws while respecting the law as written while contravening its spirit, especially when the laws allow for loopholes or

³¹ We supplement our skepticism with Hurka's (2013) forceful arguments for this position.

³² This feature also explains why there can still be honor among thieves.

are subject to interpretation. For example, suppose that a free-rider creates a chain of shell companies to exploit a tax loophole that reduces her liability to almost nothing, fully complying with the written tax code while deliberately undermining the code's purpose of fair contribution by the wealthy. The notion of exploitation here cannot be a *legal* one, for the behaviour is legal, by hypothesis. Rather, what's objectionable about the free-rider's behaviour is that they're expecting others to contribute taxes, so they can enjoy the public goods supported by those taxes, while contributing nothing themselves.³³

A second example helps illustrate the phenomenon.³⁴ Cricketers will sometimes walk from the crease even when they have not been given out by the umpire when they know that the ball touched the bat before it was caught. These cricketers intentionally break the rules of the game. But they are not acting dishonestly – after all, there is an advantage in waiting to see whether the umpire noticed that the ball touched the bat. Indeed, surely their rule-breaking is a means of being admirably honest! Our account may seem unable to capture either fact.

Consider the first fact: does our account entail that the batter is dishonest because they intentionally violate the rules of cricket by walking from the crease? Whether our account implies this depends on whether walking from the crease in the case violates other cricketers' expectation for batters to stay in the crease in that circumstance. Since cricketers are seldom sanctioned, if ever, for leaving the crease in this circumstance and since walking from the crease works against the batter's self-interest, we find it hard to see how breaking this rule constitutes a *violation* of the other cricketers' trust. After all violations are, by definition, culpable and culpable wrongdoing is blameworthy. But the batter is not blameworthy. So we conclude that it's not a violation. Since it's not a violation, it's not dishonest.

Walking from the crease in this circumstance inverts gaming the system. Gaming the system involves following the rules in a way that is objectionable by another set of standards. By contrast, walking from the crease involves breaking the rules in a way that is admirable by a different standard – namely, the standard of fair play. Winning because the umpire missed a call is still winning – c.f. Maradona's "Hand of God" – but it is dissatisfying. And insofar as players count on one another to play the game well, not merely to comply with its rules, they can expect one another to act as though they are out even if the umpire has missed it. Walking from the crease vindicates that trust.

³³ Thanks to the editor at *Political Philosophy*.

³⁴ Thanks to a referee for *Political Philosophy*.

V. CONCLUSION

In shifting our focus from deception to trust, and from falsity to bad faith, we've offered a model of honesty that explains not only why lying is paradigmatic of dishonesty, but also why dishonesty need not involve lying at all. On the Trust Model, dishonesty is best understood as a culpable betrayal of the norms that make social practices possible, and honesty as the good faith vindication of those norms. This framework does justice to the full range of honest and dishonest actions – whether interpersonal or intrapersonal, spoken or unspoken, subtle or brazen – while avoiding the distortions imposed by a narrow focus on truthfulness. In doing so, we reject the natural answer and replace it with a unified view that accounts for why honesty matters.

ACKNOWLEDGMENTS

Thanks to audiences at The Honesty Project's Initial and Final Research Conferences, Pepperdine University, UC Davis, and University of Maryland, College Park. Thanks especially to Frank Hong, Robert Long, Nate Sharadin, Elliot Thornley, and Mitch Barrington for helpful discussion. Special thanks to Wes Sisco for written feedback. This project began when the authors received funding from the Templeton Foundation, whom we thank for support. The authors contributed equally to this article and are listed in the byline alphabetically by surname.

COMPETING INTERESTS

The authors declare that they have no competing interests.

REFERENCES

- Berkey, B. 2021. The philosophical core of Effective Altruism. *Journal of Social Philosophy*, 52 (1): 93–115. <https://doi.org/10.1111/josp.12347>
- de Kenessey, B. 2023. Promises, offers, requests, agreements. *Ergo*, 9: 34. <https://doi.org/10.3998/ergo.2623>
- Dougherty, M. 2024. Honesty and the truth: against subjectivism about honesty. *Journal of Value Inquiry*, Online First. <https://doi.org/10.1007/s10790-024-09990-9>
- Foster, J. 2023. Busting the ghost of neutral counterparts. *Ergo*, 10 (42): 1187–1242. <https://doi.org/10.3998/ergo.5179>
- Frankfurt, H. 2005. *On Bullshit*. Princeton, NJ: Princeton University Press.
- Gilbert, M. 1993. Is an agreement an exchange of promises? *Journal of Philosophy*, 90 (12): 627–49. <https://doi.org/10.2307/2940815>

- Goodin, R. 2010. Perverting the course of politics. *British Journal of Political Science*, 40 (4): 725–39. <https://doi.org/10.1017/S0007123410000207>
- Hawley, K. 2014. Trust, distrust and commitment. *Noûs* 48(1): 1–20. <https://doi.org/10.1111/nous.12000>
- Hom, C. 2008. The semantics of racial epithets. *Journal of Philosophy*, 105 (8): 416–440. <https://doi.org/10.5840/jphil2008105834>
- Howard, N. R. 2021. Moral fetishism and a third desire for what’s right. *Journal of Ethics and Social Philosophy* 2 (3): 370–381. <https://doi.org/10.26556/jesp.v20i3.1249>
- Howard, N. R. & Laskowski, N. G. Forthcoming. Robust vs formal normativity II, or: no gods, no masters, no authoritative normativity. In *The Oxford Handbook of Metaethics*, ed. David Copp & Connie Rosati. Oxford University Press.
- Hurka, T. 2013. Aristotle on virtue: wrong, wrong, and wrong. Pp. 9–26 in *Aristotelian Ethics in Contemporary Perspective*, ed. J. Peters. London: Routledge.
- Jones, K. 2004. Trust and Terror. Pp. 3–18 in *Moral Psychology*, ed. Peggy DesAutels and Margaret Urban Walker. Lanham MA: Rowman and Littlefield.
- Manne, K. 2017. *Down Girl: The Logic of Misogyny*. New York: Oxford University Press.
- Miller, C. 2017. Honesty. Pp. 237–273 in *Moral Psychology, Volume V: Virtue and Character*, ed. Walter Sinnott-Armstrong and Christian Miller. Cambridge, MA: MIT Press.
- Narveson, J. 1994. The agreement to keep our agreements. *Philosophical Papers*, 23 (2): 75–87. <https://doi.org/10.1080/05568649409506414>
- Raz, J. 1986. *The Morality of Freedom*. Oxford: Clarendon Press.
- Reilly, Rick. 2020. *Commander In Cheat: How Golf Explains Trump*. New York: Grand Central Publishing.
- Roberts, R. and R. West. 2020. The virtue of honesty: a conceptual exploration. Pp. 97–126 in *Integrity, Honesty, and Truth Seeking*, ed. Christian B. Miller and Ryan West. Oxford University Press.
- Singer, P. 1972. Famine, affluence and morality. *Philosophy & Public Affairs*, 1 (3): 229–43.
- Singer, P. 2016. *Famine, Affluence, and Morality*. New York: Oxford University Press.
- Suits, B. 1978. *The Grasshopper: Games, Life and Utopia*. Peterborough, Ont.: Broadview Press.
- Um, Sungwoo. 2024. Honesty: respect for the right not to be deceived. *Journal of Moral Education*, 53 (2): 292–306. <https://doi.org/10.1080/03057240.2023.2199347>
- White, P. Q. 2022. Honesty and discretion. *Philosophy and Public Affairs*, 50 (1): 6–49. <https://doi.org/10.1111/papa.12203>

