

Some Moral Benefits of Ignorance

(Forthcoming in *Philosophical Psychology*)

This is the penultimate draft. Please cite official version.

The Socratic maxim that the recognition of our ignorance is the beginning of wisdom has profound significance for our understanding of society. The first requisite for this is that we become aware of men's necessary ignorance of much that helps him to achieve his aims.

—Friedrich Hayek, *The Constitution of Liberty* (1960)

ABSTRACT

When moral philosophers study ignorance, their efforts are almost exclusively confined to its exculpatory and blameworthy aspects. Unfortunately, though, this trend overlooks that certain kinds of propositional ignorance, namely of the personal costs and benefits of altruistic actions, can indirectly incentivize those actions. Humans require cooperation from others to survive, and that can be facilitated by a good reputation. One avenue to a good reputation is helping others, sticking to moral principles, and so forth, without calculating the personal costs of doing so, e.g., saving someone from a burning building without calculating how personally costly or beneficial it would be. These actions are indirect moral benefits (partly) resulting from that kind of propositional ignorance.

1 | Introduction

Moral philosophers generally frame ignorance in exculpatory or culpability terms. Ignorance can excuse: Michael isn't morally blameworthy for poisoning Ann if he didn't know, and couldn't have known, the bottle labeled 'sugar' was actually poison; his ignorance excuses an otherwise

blameworthy act. Nearly everyone accepts this. More controversial, however, is whether one can be culpably ignorant (Smith 1983). Michael looks culpable for poisoning Ann if he knew that the poison and sugar bottles were beside each other but didn't check the bottle before adding it to her coffee. Contemporary moral philosophers rarely study ignorance outside these issues.

There are, though, some exceptions—though their aims are distinct from that of the paper—in the literature: Rescher (2009: 12) discusses the exculpatory role of ignorance under conditions where everyone lacks knowledge, e.g., under global skepticism. DeNicola (2017: 68) explores the ethics of ignorance with respect to the right and duty to be informed or to remain ignorant, e.g., reading someone's diary against her wishes. And Julia Driver (1989) argues that ignorance about one's abilities and accomplishments are a necessary condition for humility.

Despite exceptions, moral philosophers haven't noticed moral benefits to ignorance of a certain variety: morally laudable actions incentivized by reputation boosts, which in turn, are facilitated by propositional ignorance of the personal costs and benefits of those actions, e.g., someone helping a stranger in need without calculating the personal benefits and costs. Actions of this sort, predicated on ignorance of the personal benefits and costs, provide some evidence that one helps someone in need, not for the personal benefits, but for the sake of the individual helped. And as a result, this can confer a reputational boost to the rescuer. The thesis of this paper can be roughly stated as:

Actions like keeping promises, rescuing a child from a fire, and honesty are morally good. Those actions can be incentivized by, among other things, boosts to one reputation to the extent that those actions are performed under (varying degrees of) propositional ignorance of the personal costs and benefits of performing those actions.

To quickly clarify, throughout the paper we use the word ‘cooperate’ to refer to many kinds of behaviors that include aiding a stranger in need, promise keeping, sticking to moral principles, and whatnot. What we mean by ‘cooperation’ is broadly pro-social behaviors—rescuing a child from a burning building, sticking to a moral principle when inconvenient, and so on. To clarify our thesis, we should next explore different kinds of ignorance.

(1.1) *What is Ignorance?*

Although the literature on the nature of ignorance is too much to be surveyed here, we can note that the traditional conception holds ignorance is a lack of knowledge: ignorance of p is to lack knowledge that p (LeMorvan 2011; Zimmerman 2008). There is also the new view definition of ignorance as the lack of true belief (Peels 2010; Goldman and Olsson 2009). Both definitions are *propositional* in nature, that is, lacking knowledge of, or belief in, a true proposition.

However, many philosophers agree these definitions of ignorance are too limited (Arfini 2020). Someone can be ignorant of how to perform a task if they lack know-how (Nottelmann 2016), they can be ignorant of subject matter (Brogaard 2016), and ignorance can even be an actively upheld falsehood, e.g., choosing to remain ignorant out of bias, and bad epistemic practices rewarded by institutions and society (Kassar 2018). And finally, some philosophers argue that epistemic agents can have complete ignorance by wholly lacking doxastic attitudes toward a proposition, along with investigative ignorance, where someone remains ignorant of something to better investigate it (Haas and Vogt 2015; Peels and Pritchard 2020).

For our purposes, however, we are exclusively concerned with propositional ignorance of the personal costs and benefits of certain moral actions—either lacking knowledge or true belief as to the personal costs and benefits, as a matter of degree, of undertaking an action, e.g., jumping into a river, to save a drowning child. And this is because we think that propositional ignorance is a fruitful way to model the kind of ignorance operant in such cases.

(1.2) *An Initial Objection*

Unfortunately, it may initially look as though our thesis is doomed from the start because most people unavoidably engage in some amount of calculating the personal costs and benefits of moral actions, even if only tacitly. However, if that is so—it looks plausible—this is a problem for the claim that ignorance of those costs and benefits enables morally good actions. And so it would look as though the kind of ignorance we need for our thesis is unavailable.

Fortunately for us, this objection misses the mark: propositional ignorance of the personal costs and benefits of action—the lack of knowledge or lack of true belief—comes in degrees. We can have a better or worse idea of the costs and benefits of engaging in altruistic behavior. And to the extent that we engage in altruistic actions, without thinking too much about it, we can effectively signal to others that, *ceteris paribus*, we are more ignorant of the costs and benefits of doing good than we would otherwise, had we taken more time to calculate them. We recognize, of course, there are cases where one cannot help but have robust knowledge of the costs and benefits of an action absent calculations. For our thesis to be valuable, we only need there to be cases where calculating would better uncover these costs and benefits, but where we would be more ignorant of them sans calculation. There are obviously cases of that sort such the above objection isn't as compelling as it may seem initially.

(1.3) *Ignorance and Ignoring*

Astute readers may wonder whether there are moral benefits to ignorance of the personal costs and benefits of acting morally, or whether the moral benefits are from *ignoring* those costs and benefits. For example, someone with knowledge that rescuing a stranger trapped in a burning building would be personally more costly than beneficial, but who rushed into the burning building to save them anyway, is more morally laudable than someone who does the same, but without first calculating those costs and benefits. The former accepted that the personal costs outweighed the benefits, but acted anyway deliberately, while the latter at most only accepted them unwittingly. Someone who knows a moral action would be more costly than beneficial to them, but ignores this and acts morally anyway, has good moral character; we don't know if the latter moral actor does.

However, there are some reasons to explore ignorance of those costs and benefits, rather than exploring the value of ignoring those costs and benefits. First, if one is ignorant of the personal costs and benefits of a moral action, they aren't likely to be tempted to refrain from helping based on how costly helping would be; if they were ignorant that their morally good actions would be personally costly, with minor benefits, there would be cases where they wouldn't be as tempted to refrain from performing those actions. And to the extent that such actions are morally good, this is a moral benefit of ignorance even if it doesn't necessarily reflect well on the moral agent responsible. The fact is that some people would avoid morally good actions if the personal costs outweighed the benefits; not everyone is motivated by self-interest, obviously, but many are. This highlights that while some people perform morally good actions for selfish

reasons, unrelated to the salient moral reasons, their actions can nonetheless have morally good by-products, indirectly facilitated by ignorance.

(1.4) *Paper Layout*

First, we review the evidence that people highly value their reputations, and then explain costly signaling theory. Second, we explain how cooperating without first calculating personal costs and benefits can boost one's reputation. Third, we survey some moral benefits of acting under ignorance of those costs and benefits, like aiding a stranger in distress and keeping a promise. Finally, we address a couple objections that ignorance is incidental to those moral benefits.

2 | Reputations Matter

A valuable survival skill—our evolutionary superpower—afforded humans is cooperation. Most species, like beavers and wolves, cooperate on limited tasks with biological relatives. Humans, in contrast, will cooperate with nearly anyone on nearly any task given they benefit somehow and believe their cooperation partners trustworthy enough. One of the downsides of depending on cooperation for survival is potential vulnerability to cheaters: individuals who take without contributing their fair share. As Sperber and Baumard point out:

[Humans] depend for their survival and welfare on frequent and varied cooperation with others. In the short run, it would often be advantageous to cheat, that is, to take the benefits of cooperation without paying the costs. Cheating however may seriously compromise one's reputation and one's chances of being able to benefit from future cooperation. In the long run, cooperators who can be relied upon to act in a mutually beneficial manner are likely to do better in what may be called the 'cooperation market' (2012: 495).

We need not strictly formulate the nature of cheating to appreciate that rational cooperators will guard against it—to be cheated makes one a sucker, and no one wants to be a sucker. One way to address cheating is tracking reputations. We know individuals often prefer ‘jail time, amputation of limbs, and death to’ forms of reputational damage like acquiring a reputation as a Nazi or as a child molester (Vonasch et. al., 2018: 604). Children as young as five will choose to preserve a good reputation by avoiding ‘cheating at the cost of losing a highly desirable prize’ (Fu et. al. 2016: 277). Reputations facilitate cooperation (Henrich 2015) and mate choice (Miller 2007). And people avoid sharing ‘fake news’ to protect their reputations (Altay, Hacquin, and Mercier 2020). This is strong evidence that people value a good reputation.

How are reputations managed? Signaling theory, discovered in evolutionary biology (Zahavi 1975) and economics (Spence 1973), gives a partial account, by explaining some behavior as a signal of otherwise opaque qualities. Peacocks are a classic example: their colorful plumage is costly to grow and undercuts predator avoidance. Engagement rings are another: real ones are expensive, and fakes don’t fool jewelers; they are a costly signal of a sincere intent to marry. A *stable* signaling system has the following features:

1. Some individuals have a quality hard to perceive directly that could be reliably signaled.
2. Some observers would stand to gain from accurate information about that quality.
3. Without costly signals, signalers could deceive observers about the quality, benefiting at the observers’ expense.
4. The cost of the signal must benefit the signaler (Bird and Smith 2005: 224).

Individuals employ high-fidelity signals—costly signals are in that set—to convey that they are trustworthy and giving. One method is to cooperate without calculating. We explore that idea, and the evidence for it, in the next section.

3 | Uncalculating Cooperation and Reputation Boosts

Humans engage in signaling. Here we focus on a specific kind of signaling with concrete moral benefits: cooperating with and helping others without first knowing that one would personally benefit from it. Others who observe this uncalculating altruistic behavior register it as a costly signal that the actor can be relied upon to help others for the sake of helping, instead of only helping when it is personally beneficial. If one helps strangers in need without first establishing that one would personally benefit, e.g., receiving a large reward or publicity for one's business, one strongly signals they can be relied on to help others, conferring a reputation boost.

The reputation boost we gain from helping without calculating solves a puzzle of uncalculating altruistic behavior. Why would someone rational help someone likely never able to repay them without even calculating how costly it would be? The same puzzle applies to firmly sticking to one's moral principles or keeping a tiresome promise when it would be easier to simply ignore the moral principle or to break the promise. Brackett the issue of whether we have an imperfect duty of beneficence: if we didn't gain anything but helped others without knowing the personal costs and benefits, we risk taking on a costly task. This looks practically irrational, and yet this kind of behavior is commonplace.

We know part of the explanation is reputation boosts. When people are observed, they tend to act in a less calculating way, suggesting that cooperating without calculating the benefits and costs are motivated (partially) by the desire to gain a reputation boost:

[We] experimentally test the hypothesis that people avoid calculating the costs of cooperation because of reputational concerns. Across two experiments, we demonstrate that *when people's decision-making processes are observable to others, they behave in a less calculating way*. This observation suggests that they use *uncalculating cooperation to gain reputational benefits, and not merely as an efficient way to avoid the (nonsocial) costs of calculating*. Thus, we provide the first experimental evidence, to our knowledge, for the key prediction of the reputation account ... [Observers] perceive uncalculating cooperation as a reliable signal and trust uncalculating cooperators with more money (Jordan et. al., 2016: 8659—my emphasis; see also: Hoffman et. al. (2015); Bereczkei et. al. (2010)).

Uncalculating cooperation reliably signals trustworthiness (Jordan et. al., 2016: 8659). A good explanation for this is that people who help without first calculating whether they will benefit can be trusted; they helped without knowing whether they would benefit from it, as opposed to someone who only helps when personally beneficial, e.g., if Bob keeps his promises even when overall costly, he is more trustworthy than, *ceteris paribus*, Sally who only keeps her promises for an even bigger favor in return.

People who cooperate without first calculating the personal costs and benefits *prima facie* lack knowledge and true belief of those costs and benefits; it looks like they lack enough time, if they help without thinking twice, to form even true beliefs about the personal costs and benefits of cooperating. This accords with what we know of extreme altruists who claim that when 'they

decided to help, the cognitive processes they describe are overwhelming *intuitive, automatic and fast* (Rand et. al., 2014: 4—emphasis mine). This evidence fails to decisively prove individuals who help without calculating lack either knowledge or true beliefs of the costs and benefits; but it looks like many folks would lack the time needed to calculate those costs and benefits.

There is a possible objection here: one may worry that if they rely on heuristics to decide when to cooperate, they may have already calculated the benefits and costs of helping would be, such that they don't need to calculate anything the split-second before coming to the aid of someone. The problem with this worry though is that people often cooperate without calculating in the moment in too many situations for prior calculations to be a viable approach; it looks like there are too many varied situations where people decide to cooperate without calculating for them to have the benefits and costs worked out prior. With the background in place, we examine some examples of morally beneficial ignorance.

We should quickly clarify here that when we're talking about costs and benefits, there are two important caveats in play. First, costs and benefits should be understood as costs and benefits *to the person who helps or sticks to their principles* without prior calculation. There are, clearly, costs and benefits for other parties too, but that isn't our focus. Second, when we talk of the costs and benefits of cooperating without first calculating, we mean the benefits minus reputational gains accrued from uncalculated altruistic actions.

How do others know someone is engaged in calculated behavior? Sometimes we can infer what others are thinking; sometimes not. People often rely on reaction times as an approximation of whether someone is engaged in calculating behavior. Choosing to cooperate quickly implies,

but fails to prove, clearly, an individual lacked the time required to determine whether helping would be more beneficial than costly. This isn't foolproof but is often good enough (Evans and Rand 2019: 70; Capraro 2017). Although there will be instances where, instead of calculating whether helping would be more personally beneficial than costly, they are determining if they can help, and are willing to help if possible, a quick decision removes some doubt that one is only helping because it is more beneficial than costly.

Finally, we need to pin down how individuals decide what to do when making a quick decision. No doubt this is a hard question to answer, but we have empirical work on it: people internalize strategies *usually* advantageous in regular social interactions. When individuals are in odd one-off situations, like seeing a stranger in distress, people often revert to their social intuitions that work in usual social situations. Moral psychologists call this *the social heuristics hypothesis*: people revert, in unusual circumstances, to social heuristics that facilitate successful navigation of their usual social interactions (Rand et. al. 2014: 2; Isler et. al., 2018).

Let's consider an objection: this project looks like it wrongly conflates rationality and morality. Perhaps the most rational choice isn't always the most moral one and vice versa. However, we aren't conflating rationality and morality. In illustrating the rationality of rescuing a stranger in a burning building, though perhaps personally costly, by appealing to the reputational boosts afforded by acting on that ignorance, we highlight some of the moral benefits to that ignorance. We focus on the rationality of altruistic actions to demonstrate that propositional ignorance of the costs and benefits of those actions can have morally good outcomes.

In this section, we explore several benefits of ignorance: altruism, adhering to moral principles, keeping promises, and honesty. There are no doubt other benefits, but these four should suffice to illustrate how the ignorance we explored, earlier in the paper, can indirectly incentivize good moral actions either as a substitution for, or in concert with, other motivations.

(4.1) *Altruism*

Altruism is the act of coming to the aid of a stranger, often when they will likely not be able to return the favor; perhaps one will never see them again. Even though this behavior is rationally puzzling, since rescuing a stranger in need can be costly, even granting the real possibility that one will be repaid or have the favor returned, not to mention strangers one will likely never see again, the empirical evidence indicates that (in some cases), a major part of the explanation for why people help strangers, without calculating the personal costs and benefits, even when they could, is a costly signal that can boost one's reputation.

While this looks like (at least) a partial solution to the rational puzzle, we should highlight the moral benefits that accrue from cooperating without calculating the personal benefits and costs of doing, even if we aren't morally required to help a stranger in distress, as it may hinge on how dangerous the helping is. Saving the life of an innocent person is undeniably morally good. An underlying premise in Singer's argument for charitable giving is that 'suffering and death from lack of food, shelter, and medical care are [morally] bad' (1971: 231). This premise struck Singer as too obvious to require arguments. This premise is the backbone of perhaps the most famous thought experiment in applied ethics:

Imagine you're walking in a park, past a shallow ornamental pond, and you notice that a small child has fallen in and seems to be in danger of drowning. You look around: where are the parents or the babysitter? But there's nobody, except you and the child. *What should you do? Of course, you think, I must rush in and save the child.* Then you remember you're wearing your favorite, quite expensive pair of shoes, and they'll get ruined if you rush into the pond. Is that a reason for not saving the child? I'm sure you'll say: no it isn't; *you just can't compare the life of a child to the cost of a pair of shoes, no matter how expensive*¹ (my emphasis).

When I teach Singer's drowning child thought experiment in class, I find nearly every student, with few exceptions, strongly believe they should rescue the child, and that comparing a pair of fancy shoes to the life of a child is morally wrong. I suspect part of their response is explained by the fact that the topic was broached in a social context, like a philosophy class, as students are being observed and evaluated by professors and fellow students, and there is social pressure to agree it would be morally awful to let the child drown. This isn't of course to claim students would *only* agree saving the child is morally required because of social pressure. There are other reasons, too, like apprehending a moral truth that they ought to save the child, but instead that it is no doubt part of the explanation (Simler and Hanson 2018).

Although we think about the costs of saving a drowning child in settings like class and seminar rooms, presumably most of us would act, without thinking twice, in the presence of a drowning child (Jordan et. al., 2016; Hoffman et. al. 2015; Rand et. al., 2012). And we find support for this in *The Life You Can Save*, where Singer makes a point that nicely aligns with the empirical stuff we've canvassed:

¹ Peter Singer gives this version of the thought experiment on the website: TheLifeYouCanSave.com. We do though find *slightly* different versions of the thought experiment in Singer (1972) and (2009). Though, for our purposes, these versions are virtually identical.

When we start talking about global poverty, I ask my students what they think you should do in this situation. Predictably, they respond that you should save the child. “What about your shoes? And being late for work?” I ask them. *They brush that aside. How could anyone consider a pair of shoes, or missing an hour or two at work, a good reason for not saving a child’s life?* (2009: 3-4).

Even if students would only save the child for reputational reasons—an extreme assumption for the sake of argument—the result of this reputational motive would be the saving of a child. This is a morally good result even if the rescuer isn’t motivated by the salient moral reasons to save the child; it is consistent with the moral psychology literature canvassed earlier that moral reasons (partially) explain why people would save the child. This is a clear moral benefit of our ignorance of the costs and benefits of cooperating, since ignorance is a catalyst to conferring the reputational benefits that, in some cases anyway, motivate the moral good actions.

(4.2) *Moral Principles*

Sticking to one’s moral principles, without knowing how personally costly it can be, is another rational puzzle: in terms of personal costs, sometimes, it would be easier just to abandon one’s moral principles and offer a plausible rationalization to make one look good and allow them to avoid the cost of sticking to those principles. We know some people do that, but the question is why everyone doesn’t (Uhlman et. al., 2009; Ditto et. al., 2009).

Rule consequentialism roughly holds that ‘the right action is not the action that results in the best overall consequences, rather it is the action performed in accordance with the set of rules which maximize the good’ (Driver 2011: 87). This is, clearly, an oversimplification—there is a heated debate among consequentialists and their critics (Law 1999; Arneson 2005), over how to

plausibly characterize this view—but is sufficient for us. Suppose that rule consequentialism is the correct moral theory that, along with the appropriate rules, captures moral obligations and prohibitions. It would be morally good for people to abide by those morally prescribed rules. However, even if this picture of things aligns with the moral facts, it fails to fully explain why some people stick to the moral rule even if they could personally benefit by violating them.

An appeal to the moral facts doesn't provide a *psychological* explanation for why it is that people would stick to a rule, especially when violating it in a specific case may confer personal benefits on them. We should pause here to note that one can be neutral about whether a full explanation would require an appeal to moral facts; however, it is beyond the scope of this paper to explore that live possibility (Cuneo and Shafer-Landau 2014). One viable psychological explanation is that by sticking to the moral principles without calculating the personal benefits and costs first, it signals trustworthiness to others. It can be costly to stick to moral principles, and beneficial to violate them. To the extent it is morally good to stick to a set of consequentialist rules, there are moral benefits from the salient ignorance.

The same applies to respecting rights. As Nozick roughly argues, we shouldn't think of rights as an end state to be achieved by weighing costs and benefits, but we should instead treat them as 'side-constraints' on what is morally permissible; rights as side-constraints operate outside the purview of the costs and benefits (Nozick 1974: 28-29). Admittedly, this view of rights isn't the only one in the literature (Wenar 2005; Van Duffel 2021), it is sufficient for our purposes. It should be clear that rights have (at least) some features Nozick ascribes to them, even if those features only hold defeasibly, e.g., they constraint the permissibility of actions in the absence of strong enough countervailing moral factors.

Why wouldn't someone rational only respect others' rights when the benefits to them outweigh the costs? Some agents will violate the rights of others when respecting them is too costly, and violating them beneficial, e.g., Tony could steal from his mom, with dementia, without anyone knowing. We would be naïve to expect that rational agents would always respect the rights of others, even when too personally costly. And yet there are people who look rational, but who nonetheless respect the rights of others as their default decision—they don't bother calculating the costs and benefits of that respect. A viable, if only partial, explanation of this likely involves reputational incentives. We respect the rights of others, thereby signaling one is committed to the rights of others and trustworthy, even if personally costly (Everett et. al., 2016).

(4.3) Keeping Promises

Though philosophers disagree on what explains why promises morally bind (Heuer 2012), they accept a consensual promise to someone is defeasibly morally binding, except when fulfilling it would be impossible without doing something morally worse, or where it would be impossible simpliciter to fulfil it. Keeping a promise can be personally costly to someone who freely made it—what appeared, at the time, to be an easily kept a promise could become costly. Suppose that Jack promises to be godfather to his best friend's son, Marvin. Unfortunately, a few years later, Marvin's parents are killed in a plane crash without anyone else to care for Marvin. We would morally expect that Jack, because of his promise, would have a duty to care for Marvin, despite the child's medical issues requiring substantial time and attention. In this case, Jack keeping his promise to care for Marvin would be personally costly. To the extent that Jack lacks stronger countervailing moral reasons, Jack has a moral obligation to care for Marvin.

When presented with this thought experiment, most readers will no doubt have a strong moral intuition that *Jack ought to care for Marvin*, since Jack freely made a promise to his best friend, even if Jack wouldn't have otherwise had a moral obligation to care for Marvin in the event his parents died without a substitute caretaker. However, if Jack could ignore his promise to his best friend, and not care for Marvin, without it personally costing him too much, shouldn't he for reasons of practical rationality? Of course, it would be morally monstrous for Jack to ignore his promise to his best friend. One morally ought to keep promises freely made in the absence of overriding moral factors, e.g., Jack had his own children, before Marvin's parents died, such that Jack couldn't afford to care for Marvin too.

Why doesn't everyone break their promises when it is too costly to keep them? Although there are many explanations here, one viable (even if only partial) explanation is that their reputation could take major damage. People do not like it when they see someone only keep their promises when personally beneficial, but not when it would be too personally costly. How can someone trust an individual who is *that* calculating? It looks like they aren't trustworthy.

We should pause to explore an important distinction implicit here: we may trust someone more to the extent that they have worked out the costs and benefits of making a promise, such that they only make a promise they are reasonably sure they can keep. This is a kind of commitment trustworthiness. In contrast to someone who, after making a promise, only keeps their promise when it is personally beneficial, or not too personally costly, to keep. This is kind of fulfillment trustworthiness. The first kind of trustworthiness clearly has value: we wouldn't trust someone who made promises chaotically before thinking about whether they even had the ability to keep them, e.g., perhaps if they made that promise, they wouldn't be able to keep, for reasons outside

their control. Put that kind of trustworthiness aside. We are instead focused on the latter kind of trustworthiness for a couple reasons. The first is that the aim of the paper is to explore some of the moral upsides to cost-benefit ignorance. The second is that there will be some individuals who may be tempted, based on what they know of the personal costliness of keeping a promise, despite having the ability to fulfill a promise, to bail on the promise because they simply do not want to incur the costs of keeping their promise.

There is a moral benefit to ignorance of the personal costs and benefits of keeping a promise; keeping a promise is incentivized by, among other things, the reputational boost one may gain indirectly from that ignorance. To the extent we morally ought to keep the promises we make, unless there are strong counter moral reasons, this kind of ignorance is morally good.

(4.4) *Honesty*

Despite the apparent moral goodness of honesty, we know sometimes truth telling is personally costly as it can be more beneficial to lie, especially when there is little chance of getting caught, than to be honest. There is a question of why some people are honest when they could lie, and personally benefit but avoid the costs of honesty. A viable but only partial explanation is that honesty, especially when one could easily to get away with lying, and avoid the cost of honesty, can improve one's reputation, especially if one is honest without calculating the personal costs and benefits of it, e.g., Sally knows Tony could have lied but didn't, even if Tony is unaware of that. There is no denying there could be other explanatory factors—e.g., moral character, belief that one is subject to a divine command—but simply that reputational boosts, in many cases, do some work too. The ignorance of the benefits and costs of honesty has an indirect moral benefit to the extent that honesty is morally good.

Time pressure—where someone must answer quickly—is an example of how ignorance of the personal costs and benefits can promote honest because ‘most of our daily interactions are with friends, family, and coworkers, and thus they are repeated, being honest, *although costly in the short run, may be optimal in the long run ...*’ (Capraro 2017: 57; original emphasis; Critcher, Inbar, Pizarro 2012; Capraro et. al., 2019). Not only do people tend to be honest as a default strategy, but the pressure to answer quickly, presumably, impedes the chance to concoct a convincing lie. It can often be hard, in a split second, to tell whether the costs of lying outweigh the benefits. Literature is filled with realistic stories of folks who lied, believing they wouldn’t be discovered, only to find the lie was more costly than beneficial, e.g., it ended up backfiring when the truth was eventually and inevitably revealed. This is another example of how ignorance can facilitate something morally good.

(4.5) *When No One is Watching*

Here we may wonder why people perform costly, morally good acts, when no one else around. How could there be reputational pressure? In many cases people act altruistically, even when costly, though no one else is around to observe it. If these actions are motivated by reputational gains, we must wonder how the reputational account can (even partially) explain it.

There are couple reasons, backed by empirical research, to expect people to cooperate without calculating, even when no one is watching. First, people form social heuristics to aid navigating frequent interactions, with members of their social circle, that intuitively and automatically kick-in, even in one-off cases where someone needs assistance in a costly situation (Rand et. al.

2014; Isler et. al., 2018). People help others, when no one else is watching, because of intuitive automatic social heuristics from everyday interactions operating in one-off situations.

Second, consider why people dislike those who calculate before cooperating: it gives one reason to worry that someone will cooperate only because it will benefit them, not because they value the person they are helping. Someone may object there is a distinction between calculating the personal costs and benefits of an action, and only helping when it would be beneficial; someone may calculate but still help, despite how personally costly it is. Though this distinction matters, it doesn't address that others may worry that someone engaged in calculating prior to helping is only interested in helping when personally beneficial. To lessen the worry that someone is calculating for selfish reasons, not simply as a matter of course, one could cooperate *without* calculating, as a costly signal that one is trustworthy and altruistic. Therefore, it may be better to even avoid meta-calculations as to whether to calculate calculating the costs and benefits of helping for reputational reasons (Jordan et. al., 2016: 8661).

5 | Objections

There are a couple objections related to the broad worry that ignorance is merely incidental to producing morally good actions; we should care whether someone acts on their beliefs about the personal costs and benefits of altruistic actions, not whether they actually have the salient knowledge. Not only that, but we shouldn't really care if people calculate the costs and benefits of actions—there may be practical reasons for doing that—but whether they only act when it would be more personally beneficial than costly. On either version of this worry, then, the link between moral benefits and ignorance looks thin.

(5.1) *The Incidental Objection—The Belief Version*

Suppose someone already knows the costs and benefits of acting altruistically, say, in saving the child drowning in a shallow pond. Perhaps she is informed of the costs and benefits of altruism by an omniscient being but ignores that knowledge and acts altruistically regardless of how it comes out. We would value her actions to the same degree, if not more, than someone who acts altruistically under ignorant of the personal benefits and costs. What matters is not whether someone is ignorant of the costs and benefits of acting altruistically, but whether they act on their cost and benefit beliefs. In that case, ignorance looks incidental to the moral benefits.

There are a couple issues with this objection. First, people who have knowledge of the personal costs and benefits of altruistic action, by an omniscient being, can plausibly be tempted to act on that knowledge of those costs and benefits just in virtue of having the knowledge. Perhaps Harry would break his promise to his friend, to help him move out of state, had he known the process would be rife with hassles, e.g., the moving truck rental is cancelled last minute. This speaks poorly of Harry's moral character, admittedly, but this ignorance has moral benefits by incentivizing Harry to keep his promise. This cynical take doesn't apply to everyone, as some folks will keep costly and onerous promises anyway, but it applies to some individuals. It would be morally better, *ceteris paribus*, were they to remain appropriately ignorant.

Second, although we may hope that people would, even if informed about the personal costs and benefits of moral action by an omniscient being, disregard that knowledge, and act accordingly regardless, there will be people who require an extra incentive, over and above what the moral facts are, to act morally. Of course, there will be individuals who act on the moral reasons, even if would be more personally costly; however, some folks won't be motivated to perform morally

good actions unless they have adequate incentives, and some incentives will be the reputational boost they receive from acting under appropriate ignorance.

(5.2) *The Incidental Objection—The Selfishness Version*

There is a related objection where one could object that the central issue isn't knowledge of the costs and benefits of their altruistic actions *per se*, but the fact that someone would choose to act altruistically only when the personal benefits outweighed the costs; the issue here isn't one that related to knowledge, but instead that they are selfish, only performing morally good actions when personally beneficial. Here too, ignorance looks incidental to morally good actions.

There are couple of issues here. First, there are selfish people who won't act altruistically as it would be too personally costly for them. That is obvious. However, if we could find a method to motivate such folks, despite their flawed moral character, to act altruistically, there would no doubt be moral value to doing that. It may not improve the moral character of those inclined to act selfishly, but there would be moral benefits, regardless, to motivating someone, even if only out of reputational concern, to act altruistically.

Second, dovetailing with the first reply, people have a tendency toward moral mediocrity: solid empirical evidence strongly implies many people attempt to be about as good or as bad as their peers and those in their social circle (Schwitzgebel 2019). However, if most people are morally mediocre, we should expect that many people won't discard their knowledge of the costs and benefits of altruism, especially when personally costlier, and act altruistically unless their peers and those around them would do likewise. The additional incentive of a reputation boost from ignorance of the personal costs and benefits of moral action may be what is needed.

6 | Conclusion

Our thesis is that there are moral benefits to ignorance. When we act altruistically but choose to remain ignorant of the personal costs and benefits, we robustly signal to others we can be relied on to act altruistically, even if personally costly. There will, of course, be cases where we already know the costs and benefits of acting—since knowledge and true belief about costs and benefits comes in degrees—but sometimes calculating before acting would inform someone of the personal costs and benefits of action. This ignorance signals to others we are trustworthy, and improve our reputations, indirectly incentivizing the ignorance of the costs and benefits of altruistic actions like helping a stranger in need, adhering to strict moral principles, keeping promises, and honesty. There are thus clear (indirect) moral benefits to ignorance.

REFERENCES

- Altay, Sacha, Anne Sophie Hacquin, and Hugo Mercier (2020). Why Do So Few People Share Fake News? It Hurts Their Reputation. *New Media and Society* [doi.org/10.1177/1461444820969893].
- Arfini, Selene (2020). Ignorance. Vlad Petre Glăveanu (ed.), *The Palgrave Encyclopedia of the Possible*. Palgrave Macmillan, DOI: 10.1007/978-3-319-98390-5_123-1.
- Arneson, Richard (2005). Sophisticated Rule Consequentialism: Some Simple Objections. *Philosophical Issues* 15 (1): 235–251.
- Bereczki, Tamas, Bela Birkas, and Zsuzsanna Kerekes (2010). Altruism Toward Strangers In Need: Costly Signaling in an Industrial Society. *Evolution and Human Behavior* (31, 2): 95-103.
- Bird, Rebecca and Eric Smith (2005). Signaling Theory, Strategic Interaction, and Symbolic Capital. *Current Anthropology* (46): 221-248.
- Brogaard, Berit (2016). Ignorance and Incompetence: Linguistic Considerations. In Rik Peels and Martijn (eds.), *The Epistemic Dimensions of Ignorance*. Cambridge University Press, pp. 57-80.
- Capraro, Valerio, Jonathan Schulz, and David G. Rand (2019). Time Pressure and Honesty in a Deception Game. *Journal of Behavioral and Experimental Economics* (79): 93-99.
- Capraro, Valerio (2017). Does the Truth Come Naturally? Time Pressure Increases Honesty in One-Shot Deception Games. *Economic Letters* (158): 54-57.
- Critcher, Clayton R., Yoel Inbar, David A. Pizarro (2012). How Quick Decisions Illuminate Moral Character. *Social Psychological and Personality Science* 4 (3): 308-315.
- Cuneo, Terence and Russ Shafer-Landau (2014). The Moral Fixed Points: New Directions for Moral Nonnaturalism. *Philosophical Studies* 171 (3):399-443.
- Ditto, Peter H., David A. Pizarro, David Tannenbaum (2009). Motivated Moral Reasoning. In Brian H. Ross (ed.) *Psychology of Learning and Motivation*. Academic Press, pp. 307-338.
- Driver, Julia (2011). *Consequentialism*. Routledge.
- (1989). The Virtues of Ignorance. *Journal of Philosophy* 86 (7): 373-384.
- Evans, Anthony M. and David G. Rand (2019). Cooperation and Decision Time. *Current Opinion in Psychology* (26): 67-71.
- Everett, Jim A. C., David A. Pizarro, and M. J. Crockett (2016). Inference of Trustworthiness from Intuitive Moral Judgments. *Journal of Experimental Psychology: General*. 145 (6): 772-787.
- Fu, Genyue, Gail D. Heyman, Miao Qian, Tengfei Guo, and Kang Lee (2016). Young Children With a Positive Reputation to Maintain Are Less Likely to Cheat. *Developmental Science* 19 (2): 275-283.
- Goldman, Alvin I. and Erik J. Olsson (2009). Reliabilism and the Value of Knowledge. In Adrian Haddock, Alan Millar & Duncan Pritchard (eds.), *Epistemic Value*. Oxford University Press, pp. 19-41.

- Haas, Jens and Katja Maria Vogt (2015). Ignorance and Investigation. Matthias Gross and Linsey McGoey (eds.), *Routledge International Handbook of Ignorance Studies*. Routledge, pp. 17-25.
- Henrich, Joseph (2015). *The Secret of Our Success: How Culture Is Driving Human Evolution, Domesticating Our Species, and Making Us Smarter*. Princeton University Press.
- Heuer, Ulrike (2012). Promising—Part 1 & 2. *Philosophy Compass* 7 (12): 832-841.
- Hoffman, Moshe, Erez Yoeli, and Martin A. Nowak (2015). Cooperate without Looking: Why We Care What People Think and Not Just What They Do. *Proceedings of the National Academy of Sciences* 112 (6): 1727-1732.
- Ishler, Ozan, John Maule, Chris Starmer (2018). Is Intuition Really Cooperative? Improved Tests Support the Social Heuristic Hypothesis. *PLoS ONE* 13 (1): e0190560.
- Jordan, Jillian J., Moshe Hoffman, Martin A. Nowak, and David G. Rand (2016b). Uncalculating Cooperation Is Used to Signal Trustworthiness. *Proceedings of the National Academy of Sciences* 113 (31): 8658-8663.
- Kassar, Nadja El (2018). What Ignorance Really Is. Examining the Foundations of Epistemology of Ignorance. *Social Epistemology* 32 (5): 300-310.
- Law, Iain (1999). Rule-Consequentialism's Dilemma. *Ethical Theory and Moral Practice* 2 (3): 263-276.
- LeMorvan, Pierre (2011). On Ignorance: A Reply to Peels. *Philosophia* 39 (2): 335-344.
- Miller, Geoffrey F. (2007). Sexual Selection for Moral Virtues. *The Quarterly Review of Biology* 82 (2): 97-125.
- Nottelmann, Nikolaj (2016). The Varieties of Ignorance. In Rik Peels and Martijn (eds.), *The Epistemic Dimensions of Ignorance*. Cambridge University Press, pp. 33-56.
- Nozick, Robert (1974). *Anarchy, State, and Utopia*. New York: Basic Books
- Peels, Rik and Duncan Pritchard (2020). Educating for Ignorance. *Synthese* 198 (8): 7949-7963.
- Peels, Rik (2010). What is Ignorance? *Philosophia* 38 (1): 57-67.
- Rand, David G., and Ziv G. Epstein (2014). Risking Your Life Without a Second Thought: Intuitive Decision-Making and Extreme Altruism. *PloS one* (9, 10): e109687.
- Rand, David, Joshua D. Greene, and Martin A. Nowak (2012). Spontaneous Giving and Calculated Greed. *Nature* (489): 427-430.
- Schwitzgebel, Eric (2019). Aiming for Moral Mediocrity. *Res Philosophica* 96 (3): 347-368.
- Simler, Kevin and Robin Hanson (2018). *The Elephant in the Brain: Hidden Motives in Everyday Life*. Oxford University Press.
- Singer, Peter (2009). *The Life You Can Save: Acting Now to Stop World Poverty*. Random House.
- (1972). Famine, Affluence, and Morality. *Philosophy and Public Affairs* 1 (3): 229-243.
- Smith, Holly (1983). Culpable Ignorance. *Philosophical Review* 92 (4): 543-571.
- Spence, Michael (1973). Job Market Signaling. *Quarterly Journal of Economics* (87): 355-374.

- Sperber, Dan and Nicholas Baumard (2012). Moral Reputation: An Evolutionary and Cognitive Perspective. *Mind and Language* 27 (5): 495-518.
- Uhlman, Eric Luis, David A. Pizzaro, David Tannebaum, and Peter H. Ditto (2009). The Motivated Use of Moral Principles. *Judgment and Decision Making* 4 (6): 476—491.
- Van Duffel, Siegfried (2012). The Nature of Rights Debate Rests on a Mistake. *Pacific Philosophical Quarterly* 93 (1): 104-123.
- Vonasch, Andrew J., Tania Reynolds, Bo M. Winegard, and Roy F. Baumeister (2018). Death Before Dishonor: Incurring Costs to Protect Moral Reputation. *Social Psychology and Personality Science* 9 (5): 604-618.
- Wenar, Leif (2005). The Nature of Rights. *Philosophy and Public Affairs* 33 (3): 223-252.
- Zahavi, Amotz (1975). Mate Selection—A Selection for a Handicap. *Journal of Theoretical Biology* (53): 205-14.
- Zimmerman, Michael J. (2008). *Living with Uncertainty: The Moral Significance of Ignorance*. Cambridge University Press.