

Running ahead: MORALITY JUSTIFIES MOTIVATED REASONING

Morality justifies motivated reasoning in the folk ethics of belief

Corey Cusimano and Tania Lombrozo

Department of Psychology, Princeton University

Author Note

Corey Cusimano, University Center for Human Values and Program in Cognitive Science, Princeton University; Tania Lombrozo, Department of Psychology, Princeton University.

Materials, data, pre-registrations, and analysis scripts can be accessed at <https://tinyurl.com/osf-mpmr>.

The authors wish to thank Sara Abramowitz, Jon Baron, Renee Bolinger, and Andrew Chignell for helpful feedback on this project.

Correspondence concerning this article should be addressed to Corey Cusimano, Department of Psychology, Peretsman Scully Hall, Princeton University, Princeton, NJ 08540. E-mail: cusimano@princeton.edu.

Abstract

When faced with a dilemma between believing what is supported by an impartial assessment of the evidence (e.g., that one's friend is guilty of a crime) and believing what would better fulfill a moral obligation (e.g., that the friend is innocent), people often believe in line with the latter. But is this how people think beliefs ought to be formed? We addressed this question across three studies and found that, across a diverse set of everyday situations, people treat moral considerations as legitimate grounds for believing propositions that are unsupported by objective, evidence-based reasoning. We further document two ways in which moral evaluations affect how people prescribe beliefs to others. First, the moral value of a belief affects the evidential threshold required to believe, such that morally good beliefs demand less evidence than morally bad beliefs. Second, people sometimes treat the moral value of a belief as an independent justification for belief, and so sometimes prescribe evidentially poor beliefs to others. Together these results show that, in the folk ethics of belief, morality can justify and demand motivated reasoning.

Keywords: Belief; Moral judgment; Lay epistemics; Optimism; Motivated reasoning

Introduction

In an early scene from Jane Austen's *Pride and Prejudice*, sisters Jane and Elizabeth Bennett review everything they know about their new neighbors, Mr. and Miss Bingley. Though their evidence is the same, they arrive at different conclusions: Jane finds Miss Bingley pleasing and thinks her likely to be a charming neighbor; Elizabeth is far from convinced. In this exchange and others, Elizabeth accuses Jane of being *too* good: “*You wish to think all the world respectable,*” she insists (2001, p. 164). But which sister is believing as she should? Is it Elizabeth, who follows the evidence? Or is it Jane, who gives others the benefit of the doubt?

This example illustrates the competition between *evidential* and *non-evidential* considerations in forming and evaluating beliefs. On some philosophical views, beliefs are justified by evidence and evidence alone (Clifford, 1886; Locke, 1690; Russell, 1948/2013), such that “it is wrong always, everywhere, and for anyone, to believe anything upon insufficient evidence” (Clifford, 1886, para 17). So if Jane and Elizabeth have the same evidence, they should arrive at the same belief, and that belief should be based on their evidence alone. But others argue that what one believes ought to be affected by what is morally good or prudent to believe. For example, believing in God (James, 1937; Pascal, 1852), giving a friend the benefit of the doubt (Stroud, 2006), or forming beliefs about minorities that ignore salient race- or sex-based statistics (Basu, 2019; Bolinger, 2018), could be justified despite seeming to violate normal evidentiary standards (see Bolinger, 2020, and Chignell, 2018, for reviews). On these views, Jane might be justified in her optimistic belief because moral or prudential considerations are on her side. Furthermore, she might be justified in arriving at a different belief from Elizabeth's if – for example – Jane has a greater moral obligation towards Miss Bingley as a closer friend or potential sister-in-law, even though Jane and Elizabeth's evidence is the same.

In the current paper we take up the question of whether and how non-evidential considerations influence people's everyday evaluations of others' beliefs. In particular, do an individual's moral obligations to others influence the beliefs we think that individual should hold? And if so, how do these moral considerations interact with how that person ought to evaluate his or her evidence? The normative counterparts to these questions have been hotly debated within contemporary epistemology (Bolinger, 2020; Chignell, 2018), but relatively little is known about how people in fact tend to balance evidential and non-evidential considerations when judging what they or others ought to believe.

Addressing this lacuna in the psychology of belief evaluation is important for a number of reasons. First, the norms that people endorse for belief predict their evaluations of others: For instance, people who moralize forming beliefs based on logic and evidence negatively evaluate others they suspect of holding illogical or non-evidential beliefs (e.g., Ståhl, Skitka, & van Proojen, 2016). Second, norms for belief appear to affect how people form and update beliefs of their own. For instance, people who believe that they and others ought to reason in actively-open minded ways tend to have more scientific and fewer superstitious beliefs (e.g., Pennycook et al., 2019). Thus, if it is the case that people endorse non-evidential norms for belief, especially norms that justify or demand motivated reasoning, this could help explain why people hold motivated beliefs in the first place (e.g., Cohen, 1981; Koehler, 1996), and sometimes tolerate them in others. We return to these points in the General Discussion. Below, we review prior work on how people evaluate beliefs and then set the stage for our inquiry.

The case for evidence-based norms on belief

The dominant view in epistemology and the philosophy of science is that beliefs are only justified if they are in proportion to, and based on, evidence (see Chignell, 2018, and Douglas, 2016, for reviews). In theory, evidence-based reasoning is the best way to acquire knowledge, and knowledge in turn forms the basis for optimal behavior (see discussion in Baron, 2008; Clifford, 1886; Foley, 1987; Milkman et al., 2009). And indeed, in practice, people who unbiasedly search for evidence, and change their mind when they acquire it, tend to perform better on problem solving tasks and have more accurate beliefs than those who do not (e.g., Baron, 2019; Haran, Ritov, & Mellers, 2013; Pennycook et al., 2019; Stanovich & West, 1998).

A large body of research suggests that ordinary people, who may be unaware of the evidence and arguments above, also treat impartial, evidence-based reasoning as the only legitimate basis for forming beliefs. For instance, people typically judge that their beliefs reflect an objective and impartial assessment of the evidence (Ross & Ward, 1996). Even when people acknowledge that, in theory, they and others are likely to be biased (e.g., that people like them tend to be overly optimistic or let their identities affect their judgment), they often deny that they were influenced by these biases when asked about any particular judgment that they have made (Ehrlinger, Gilovich, & Ross, 2005; Hansen, Gerbasi, Todorov, Kruse, & Pronin, 2014; Pronin, Gilovich, & Ross, 2004; Pronin, Lin, & Ross, 2002; West, Meserve, & Stanovich, 2012). They also claim to prefer to hold beliefs that are logical and based on evidence, with a recent survey reporting that the majority of participants (78%) agreed with statements such as, “It is important to me personally that I can justify my beliefs using rational arguments and evidence” (Ståhl et al., 2016). Consistent with this, people tend to update their beliefs when they discover errors or biases in their reasoning (see Wegener, Silva, Petty and Garcia-Marques, 2012, for a review). Indeed, one limit to people’s tendency to form motivated beliefs is an apparent desire to keep

their beliefs evidentially-justified (Kunda, 1990). Thus, even though people often engage in motivated reasoning, they seem not to recognize that they do so, and think it would be bad if they did reason that way (Epley & Gilovich, 2006; Kunda, 1990).

People's desire for objective and impartial evidence-based reasoning appears to extend to how they evaluate others. Indeed, norms of impartiality are common in many decision-making contexts. For instance, even though people want their friends to be partial towards them (Shaw, DeScioli, Barakzai, & Kurzban, 2017), they dislike when others treat their friends favorably (and try themselves to hide partiality from others; Shaw, Choshen-Hillel, & Caruso, 2018). In the domain of belief evaluation, prior work demonstrates that people often react negatively towards others they think are biased (Kennedy & Pronin, 2008), illogical (Ståhl et al., 2016), or wrong (Molnar & Loewenstein, 2020). Moreover, people tend to favor others with beliefs like their own (e.g., Abelson, 1986; Golman, Loewenstein, Moene, & Zarri, 2016) and exhibit an "objectivity bias" such that they assume that anyone who has the same information that they do will form identical beliefs (Robinson, Keltner, Ward, Lee, 1995; Rogers, Moore, & Norton, 2017; Ross & Ward, 1996). These findings suggest that people are committed to strict, evidential norms for belief, such that beliefs should be based on evidence and evidence alone, with no role for loyalties or other commitments that might arise from friendship or other social obligations that potentially differ across people. In other words, this line of reasoning – which motivates a large and varied set of work in psychology – predicts that people ought to reject motivated reasoning in all its forms.

The case for moral encroachment

Despite the findings reviewed above, there are reasons to think that moral considerations enjoy a special status during belief evaluation. A growing body of work suggests that moral values, such as concerns about harm, justice, respect, or loyalty, play a cardinal role in people's lives, holding a protected status that outweighs and resists comparison to nonmoral considerations (Baron & Spranca, 1997; Tetlock et al., 2000). Consistent with this idea, moral virtues, such as being loyal, just, or kind, are seen as more important to judging someone's character than are epistemic virtues, such as being logical or intelligent (Goodwin, Piazza, & Rozin, 2014). Thus, if holding an evidence-based belief would cause harm or signal disloyalty relative to a non-evidential belief, people may recommend that others adopt the non-evidential, morally preferable alternative. In other words, they may judge Jane Bennett in the right, and Elizabeth in the wrong.

As noted above, one reason to adopt accurate and evidence-based beliefs is that they often generate optimal behavior (Baron, 2008; Clifford, 1886). However, inaccurate beliefs can sometimes be useful. Being overly optimistic can make someone happier (Carver & Scheier, 2014) or motivate them in academics (e.g., Multon, Brown, & Lent, 1991), athletics (e.g., Moritz, Feltz, Fahrback, & Mack, 2000), or personal health (e.g., Holden, 1991; see Bandura & Locke, 2003, for a review). Even lying to oneself can be helpful, as people are better at deceiving others when they have self-deceived (Schwardmann & van der Weele, 2019). If observers believe that unrealistic optimism or self-deception could have positive downstream consequences for the believer or for those around her, then they may associate that belief with moral value and prescribe it to others on that basis.

Consistent with this reasoning, Tenney, Logg, and Moore (2015) found that many people believe that optimism improves motivation, and that because of this, it can help them fulfill their

goals. Tenney et al. (2015) asked participants what someone should believe either when that person was deliberating about what decision to make or after she had made her decision and now needed motivation to follow through. Participants reported that she should be more optimistic after deciding (vs. during deliberation), consistent with the notion that people prescribe optimism in part based on whether they think it will benefit the believer (see also Armor, Massey, & Sackett, 2008). However, this finding is consistent with a portrait of purely evidence-based belief evaluation. Participants may have inferred that someone who decided on a course of action had more evidence that licensed optimism relative to someone who was still deliberating about what to do. It is also possible that people interpret self-fulfilling beliefs as *evidentially* self-fulfilling, given that adopting a motivating belief provides additional reason to predict a successful outcome. Because Tenney et al.'s studies did not stipulate or measure the evidence available to each agent across cases, participants could have been prescribing precisely those beliefs that they took to be supported by the evidence in each case. These studies are therefore silent on the question of whether people endorse non-evidential norms for belief, but they nicely demonstrate that observers readily reason about the downstream consequences of belief, and thus set the stage to test whether people will prioritize these consequences over the evidence.

Beyond the downstream consequences of belief, beliefs could also have intrinsic value that warrants adopting or avoiding them. For instance, an evidence-based belief may be morally bad because it constitutes a form of disloyalty towards one's friends or family (of whom you are supposed to think well; c.f. Cohen & Rozin, 2001; Keller, 2004; Stroud, 2006). Alternatively, as some philosophers have argued, making race- or sex-based inferences may fail to treat minorities with due respect, even when they are based on evidence (e.g., Basu, 2019; Gendler, 2011; Moss, 2018; but see Gardiner, 2018). Indeed, some work has shown that people will disregard certain

kinds of evidence when making decisions, such as when it would lead them to convict someone of a crime based on statistical evidence (e.g., Nesson, 1985; Wells, 1992) or when relevant statistical evidence is caused by race-based inequality or discrimination (Tetlock et al., 2000).

While people may in fact dismiss evidence to avoid committing objectionable behaviors, it is less clear if they think evidence ought to be dismissed when forming beliefs. Cao, Kleiman-Weiner, and Banaji (2019) report that people will condemn certain beliefs as unjustified even while, on the basis of evidence, they hold the same beliefs. For instance, many people will infer that a surgeon is more likely to be male than female, which is consistent with Bayesian reasoning given unequal base rates favoring men. However, when told that another person believes the same thing, people will condemn their belief. People may do this because the belief violates a moral norm to which they are more sensitive when evaluating others than themselves, an interpretation consistent with our hypothesis. However, as the authors note, people may have condemned others because they assumed that those others formed their beliefs in an evidentially poor way (e.g., baseless sexist assumptions) rather than via (their own) rational reasoning. In this case, people's condemnation would be consistent with the thesis that people evaluate beliefs on the basis of how well people's beliefs reflect their evidence.

In sum, prior work provides preliminary evidence that people evaluate others' beliefs in part based on how beneficial or morally offensive those beliefs are. However, because this work has not controlled for, nor directly measured, the evidential status of the beliefs in question, it is ultimately inconclusive with respect to whether people prescribe beliefs for non-evidential reasons. We address this limitation in the studies reported below. Beyond this methodological limitation, however, prior work has not addressed the question of precisely how evidential and

non-evidential norms interact when people prescribe beliefs. Investigating this is the second major contribution of this paper, to which we now turn.

Two ways morality could license motivated belief

We will call the discrepancy between what someone ought to believe based on an objective assessment of the evidence and what someone ought to believe when taking into account other (in this case, moral) considerations “prescribed motivated reasoning.” However, observing that people prescribe motivated reasoning raises the question of exactly how people think moral considerations interact with evidentiary ones. Epistemologists and moral philosophers have argued that there are two potential (non-exclusive) ways that concerns about the moral value of a belief could justify motivated belief. These two routes to moral influence are derived by distinguishing between two kinds of evaluations that people typically make of their own and others’ beliefs during belief formation (Figure 1). First, people assess whether they or the other person has sufficient evidence for a belief. This may then be one of several inputs into a second, all-things-considered normative evaluation of the belief, which determines whether the belief is, on the whole, justified, permissible, or otherwise “good.” In principle, moral concern can impact people’s evaluations of others’ beliefs at either of these points (Bolinger, 2020; Kim & McGrath, 2019, *inter alia*; Pace, 2011). That is, morality could influence a belief’s quality by changing what counts as “sufficient evidence” for belief (e.g., as proposed by Bolinger, 2018; James, 1937; Pace, 2011) or by weighing moral considerations directly against evidentiary ones (e.g., Pascal, 1852; Stroud, 2006). We motivate each of these ideas below.

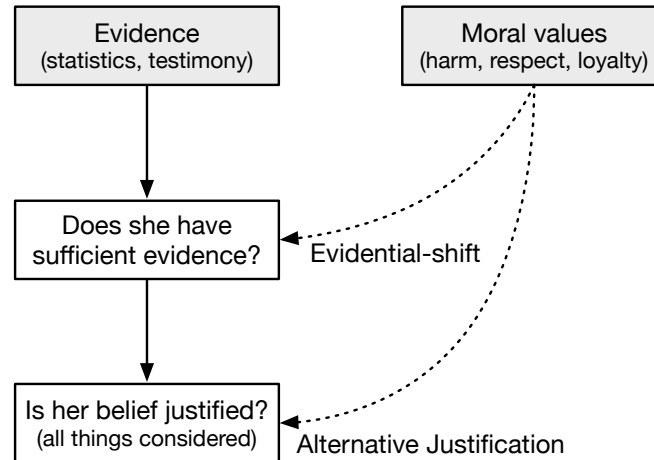


Figure 1. Moral values may affect what beliefs people prescribe to others in two ways.

According to the “evidential shift” hypothesis, morality generates an evidential double standard for morally good vs bad beliefs. Moral benefit as an “alternative justification” results in prescribed motivated belief by justifying belief independent of evidential considerations.

To motivate the idea that moral considerations could influence what constitutes sufficient evidence, we turn to prior work on belief formation that has investigated how people set a criterion for how much, and what quality, of evidence is considered sufficient before stopping inquiry. This work finds that what people set as the proper standard of evidence for a belief varies across people and contexts (Kruglanski, 2004; Kruglanski & Webster, 1996; Mayseless & Kruglanski, 1987). For instance, when people must make a decision quickly, people will weigh sparse evidence more heavily and feel justified adopting a belief on the basis of less total information (Kruglanski & Freund, 1983). However, when it is important to form an accurate belief (as opposed to withholding belief), the same evidence will result in lower confidence and a longer time before acceptance (Mayseless & Kruglanski, 1987). This is one way that people can form beliefs on varying amounts of evidence, yet still think that their beliefs are all supported by “sufficient” evidence. In principle, altering one’s decision criterion across contexts reflects a rational way to allocate attentional resources when faced with many decisions (McAllister, Mitchell, & Beach, 1979) and, consistent with this rational analysis, people evaluate others’

beliefs in light of these risk-dependent evidential standards (e.g., McAllister, et al, 1979; Pinollos, 2012). In practice, however, this often results in pernicious double-standards such that people set lower thresholds for evidence for desired beliefs than they do for undesired beliefs (Ditto & Lopez, 1992). This in turn facilitates motivated reasoning even while people believe that they are forming beliefs in an evidentially-admirable way (Ditto & Lopez, 1992; Gilovich, 1991). We hypothesize that it is precisely this kind of double-standard threshold-shifting that people will sometimes knowingly, morally prescribe, and we refer to this as the “evidential shift hypothesis.”

According to the evidential-shift hypothesis, people will hold others to higher standards of evidence before adopting morally bad beliefs but lower standards of evidence when adopting morally good beliefs (see Bolinger, 2020, for a review of recent normative theories defending this double-standard reasoning). For instance, if people believe that friends are obligated to “give each other the benefit of the doubt,” then they may judge that, when it is somewhat ambiguous whether a person did something bad or not, a loyal friend has to collect a lot of evidence before forming a negative belief about them, but only a little evidence before forming a positive belief about them, relative to a neutral observer. Critically, this would not be because the believer has greater prior confidence that her friend is innocent. A neutral observer, who lacked a moral obligation to be loyal but who had all the same information, would be permitted (and likely expected) to have a symmetrical criterion for forming positive or negative beliefs about the person in question.

The second (but not incompatible) way in which moral concerns could affect people’s evaluations of others’ belief, also depicted in Figure 1, is by providing an “alternative justification” for belief. On this account, when deciding what others ought to believe, people

weigh moral considerations against other qualities of the belief, including whether there is sufficient evidence for it. Thus, holding constant observers' judgments that someone has sufficient or insufficient evidence for belief, they may decide whether someone ought to form a belief on the basis of whether it would be morally good to do so. Keeping with our introductory example, Jane could subscribe to the idea that a moral obligation to see the best in others justifies believing them charming and kind, even when that belief is inconsistent with, or challenged by, the evidence (Stroud, 2006).

The evidential shift hypothesis and the alternative justification hypothesis generate unique predictions about how evidential and non-evidential concerns interact with each other. According to the alternative justification hypothesis, moral concerns are weighed against the evidentiary status of a belief such that adopting the morally good belief may demand that someone sacrifice their duty to adopt a belief that is supported by the evidence. By contrast, according to the evidential shift hypothesis, it is possible for morality to affect belief prescription while simultaneously fulfilling an obligation to only believe what one has "sufficient evidence" to believe (James, 1937; Pace, 2011). However, they can also operate in parallel. It is therefore possible that people believe that moral considerations affect both the evidentiary status of beliefs as well as their overall quality, only one of these, or neither of them (and therefore that morality plays no legitimate role in evaluating what someone ought to believe). The studies below were capable of adjudicating between each of these possibilities.

The current studies

To summarize the work reviewed above, we know that people often engage in motivated reasoning despite subscribing to what appear to be evidential norms for belief. At the same time,

we know that moral considerations are central to our evaluations of others, and that in some cases, people prescribe optimistic beliefs to others. However, we know little about whether people think motivated reasoning is ever a permissible way to form a belief, and less about how evidential and non-evidential considerations interact in the evaluation of belief. We hypothesize that people evaluate others' beliefs in part based on their moral quality – including how harmful or constructive the belief is, and how respectful or loyal the belief is – thereby predicting that people will sometimes report that another person ought to engage in motivated reasoning. Our studies are further designed to identify how moral concerns impact people's belief prescriptions – through an “evidential shift” or an “alternative justification.”

A major challenge in addressing our research questions is varying the perceived moral demands on someone's belief while holding evidence constant. Across our studies we do this in two distinct ways. In Study 1, we compared participants' prospective judgments about what someone ought to believe when evidence and moral demands conflict with their judgments of what an objective, “advanced AI” (that is not subject to moral demands) would believe based on that person's total evidence. In Studies 2 and 3, participants made retrospective evaluations about beliefs formed by two individuals who differ in how socially close they are to someone (and therefore how subject they are to moral demands to think well of that person). To equalize the evidence held by each, the two believers shared all of their evidence with each other before both forming the same belief (either on the basis of their shared evidence or despite it). Across all studies, we consistently document that people (i) prescribe beliefs that depart from what a more objective observer would believe, (ii) evaluate the evidential quality of a belief differently based on the belief's moral qualities, and (iii), in some circumstances, prescribe beliefs that they know the believer lacks evidence for.

Transparency

All sample sizes, exclusion criteria, and statistical analyses were preregistered. All analyses are reported as either planned (i.e., preregistered) or exploratory. All sample sizes represent the number of subjects who completed the study and passed exclusion criteria. All studies were conducted using jspsych (de Leeuw, 2015). The JavaScript code used to run these studies, data, analyses (annotated R scripts), and pre-registrations for each study are available on OSF: <https://tinyurl.com/osf-mpmr>. All studies reported in this paper were approved by the Office of Research Ethics at Princeton University.

Study 1

Study 1 had two primary aims. The first aim was to test the hypothesis that, when a particular belief is morally preferable, people sometimes prescribe motivated reasoning – that is, they maintain that a person ought to hold a belief that differs from an “objective” belief based on the same evidence. The second aim was to determine whether the influence of moral considerations occurs by the provision of an alternative justification for belief or through an evidential shift.

Participants read one of six vignettes designed to elicit moral concerns that could plausibly override evidential ones (see discussion below and overview in Table 1). After reading their assigned vignette, participants reported what a perfectly objective, impartial, and rational observer would believe (in this study, an “advanced artificial intelligence” with all the evidence held by the main character of the vignette). Participants then reported what the main character in the vignette, who is putatively subject to moral concerns, ought to believe. We reasoned that

discrepancies between these two believers would reflect participants legitimizing morality's influence on belief, whereas correspondence would reflect a commitment to impartial evidence-based reasoning. To help us assess the alternative justification and evidential shift hypotheses, participants also indicated the range of beliefs consistent with the evidence for each believer, as well as their endorsement of the relevant moral norm for each vignette,

Methods

Participants. We recruited 839 adults (441 reported female, 395 reported male, mean age 38) from Amazon Mechanical Turk (MTurk). An additional 144 participants were excluded for failing at least one of three comprehension questions (explained below). For all studies, participation was restricted to users with a US-based IP address and a 95% approval rating based on at least 500 prior tasks.

Stimuli. Prior work in psychology, moral philosophy, and epistemology has produced a catalog of mundane situations in which moral considerations plausibly override what one ought to believe on the basis of impartially evaluating the evidence (see e.g., Basu, 2019; Bolinger, 2018, 2020; Cao et al, 2019; Stroud, 2006; Pace, 2011). We generated six vignettes which comprise a diverse and representative set of cases from this literature (see Table 1). Two vignettes investigate cases in which someone possesses base-rate information about someone based on their sex or race (Race, Sex). The remaining vignettes look at instances in which someone's relationship with another person generates social obligations to remain optimistic in the face of pessimistic circumstances. In "Friend," someone learns that his childhood friend is under investigation for drug possession. In "Marriage," a newlywed husband learns that the base rate of divorce for his demographic is 70%. And in "Cancer," someone learns that her husband has been diagnosed with a rare cancer that few recover from. We reasoned that each of these

vignettes would generate conflicts between what the evidence warranted and what social obligations demanded. And finally, in “Bully,” a teacher is presented with evidence that a student is (and will continue to be) poorly behaved, however, she knows that the belief she adopts will have implications for his future success as a student. We reasoned that, if we observe evidence for these phenomena in many of these cases, that this would reflect a broad tendency for people to integrate moral considerations into their evaluation of belief rather than reflecting idiosyncratic features of any individual vignette.

Table 1

Overview of six vignettes used in Studies 1-3.

Vignette	Belief	Evidence	Moral Consideration	Study
Bully	New student will behave poorly tomorrow.	Behaved poorly on first day of class. Older siblings were poorly-behaved.	Obligation to treat every student as having high potential.	1, 2
Cancer	Husband will survive his cancer. ¹	Studies show that only 15% of those diagnosed live past 1 year.	Optimism will improve well-being for husband and family.	1, 2
Marriage	Main character and wife will eventually divorce.	Study showing that 70% of similar marriages end in divorce.	Vowed life-long commitment. Optimism facilitates a better relationship.	1, 2, 3
Friend	Friend is not guilty of possessing cocaine. ¹	Drugs found in dorm room, rumors of friend associating with drug dealers.	Friend requests benefit of the doubt, loyalty to friend demands trust.	1, 2
Race	Approaching Black man is dangerous.	In this neighborhood, 80% of young Black men are in a dangerous gang.	Respecting others demands you do not judge based on group statistics.	1
Sex	Approaching woman is surgeon rather than nurse. ¹	In this particular dental practice, 90% of surgeons are male.	Respecting others demands you do not judge based on group statistics.	1

Note. Half of the vignettes used beliefs that stated optimistic outcomes (marked by 1 above). In the analyses below, participants responses to these items were reversed so that greater endorsement was pessimistic, matching the other vignettes.

Procedure. Participants were randomly assigned to read one of the six vignettes described in Table 1. All vignettes featured situations in which the main character acquires strong but inconclusive evidence for a proposition that they have a moral reason to reject. For

instance, in “Friend,” the main character, Adam, learns that his childhood friend John is under investigation after cocaine was found in his dorm room:

Adam and John grew up together on the same block and attended the same elementary school, middle school, and high school. Throughout this time, they were very good friends. They helped each other in school and supported each other in hard times. After high school, they started college at different schools at opposite ends of the country. Even though they no longer saw each other, they kept in touch by talking on the phone a couple times a month. They both continued to think of each other as close friends. A few months after starting college, Adam learns from a mutual friend that John is in trouble at his college. Apparently, the campus police found a small bag of cocaine in John's dorm room and are now investigating him for known possession of a controlled substance.

John insists to Adam that he is innocent and that, because they are friends, he ought to trust him:

When Adam asks John about the rumor, John admits that he is being investigated but then says, “I know it looks bad but please believe me that it isn't mine. You're one of my closest friends and I need someone on my side”. Adam and John smoked marijuana together in high school, but Adam has never seen John do any hard drugs before.

However, Adam later gains evidence that John is guilty:

Over the next few days, Adam learns more about John's situation from some mutual friends who attend John's school. Adam learns that John does not share his dorm room with anyone else at his school. Additionally, people have seen John hanging around known drug dealers off campus. And lastly, John's class attendance has been poor since the semester began. Although everyone is speculating about whether John is guilty or innocent, the case will not be settled for a couple of months.

At the end of each vignette, participants were told that the main character (e.g., Adam) is wondering what to believe and is weighing everything they have recently learned, which suggests a pessimistic belief, as well as what they have moral reason to believe, which suggests an optimistic belief.

Attributions of bias. Immediately after being told that the character is wondering what to believe, participants reported what they thought the main character was most likely to believe. For instance, in the Friend vignette, participants read, “Taking into account Adam's experiences with John growing up, as well as all the other things that Adam has learned about John's situation right now (including all the rumors he has heard), what do you think Adam is most likely to believe about John's innocence?” Participants were further instructed to indicate what they think is most likely regardless of what they think the character ought to believe. Participants responded to this question using a slider representing the subjective probability toward a proposition (e.g., “John is innocent”) anchored at 0% and 100% with 5% intervals. After responding to this question, participants were told, “You just indicated what you think [the main character] is most likely going to believe in this situation. All the remaining questions will ask you to make different judgments about [the main character]’s situation and his beliefs.” We took these steps to ensure that participants did not confuse subsequent prescriptive judgments as questions that were asking them to predict how biased the main character was likely to be.

Most accurate estimate. Participants then reported the most accurate estimate the main character could make in light of the evidence they have. To estimate what would be objectively most accurate, participants were told to imagine that the main character’s mind was uploaded to an advanced AI that is able to “detect, catalog, and synthesize” all of the character’s information and experiences. For instance, in the Friend vignette, participants read:

Imagine that Adam gets his brain scanned by an *advanced artificial intelligence* (AI). This AI is able to detect, catalog, and synthesize all of Adam's experiences with John growing up, as well as all the other things that Adam has learned about John's situation right now (including all the rumors he has heard). This AI is able to simulate what a *perfectly detached perceiver* would estimate based solely on the information downloaded from Adam's brain (including Adam's beliefs, memories, and emotions). What do you think that this advanced AI would estimate about the probability that John is innocent (i.e., that the cocaine found in his dorm was not his) based solely on the information it downloaded from Adam's brain?

Participants then reported their answer on the same 0-100% scale that they used on the previous page.

Evidence-based bounds on belief. Next, participants reported the most optimistic and most pessimistic estimates that could be considered “consistent with” and “based on” the evidence available. For instance, if the participant had just reported that the best estimate is 30%, then they would have read:

You have just indicated that, based on the evidence that the advanced AI/Adam has, the best estimate that John is innocent is 30%. There is some uncertainty in this estimate based on the evidence available. In light of this, the advanced AI can also compute the most pessimistic (i.e., the lowest probability that John is innocent) and the most optimistic (i.e., the highest probability that John is innocent) estimates based on the evidence available.

Participants were instructed to report what these estimates would be “if the advanced AI made sure its estimates are based on, and consistent with, all the information that it has from

[character's] brain.” A 0%-100% range input appeared below containing three slider handles (see Figure 2). One handle was immovable and set to the estimate they had indicated to be most accurate. Up to 15%-points to the left and right were two handles that represented evidence-bound pessimism and optimism.¹ Participants could move these handles in five-point increments from the anchors (0% or 100%) to the “most accurate” value handle.

After participants submitted their judgments for the AI, instructions appeared below directing the participant to now report what the evidence-based bounds would be for the main character (e.g., Adam). We used the deviation between participants' estimates for the AI and the character as a measure of the “evidential threshold shift” hypothesis.

¹ If the participant provided a “most accurate” estimate that was 15% or less away from one of the scale anchors (0% or 100%), then the evidence bound slider was set equal to the scale anchor. It could still be moved between the anchor and the most accurate estimate. If the participant set the most accurate estimate to 0% or 100% then the relevant evidence-bound anchor was set equal to the most accurate estimate and could not be moved.

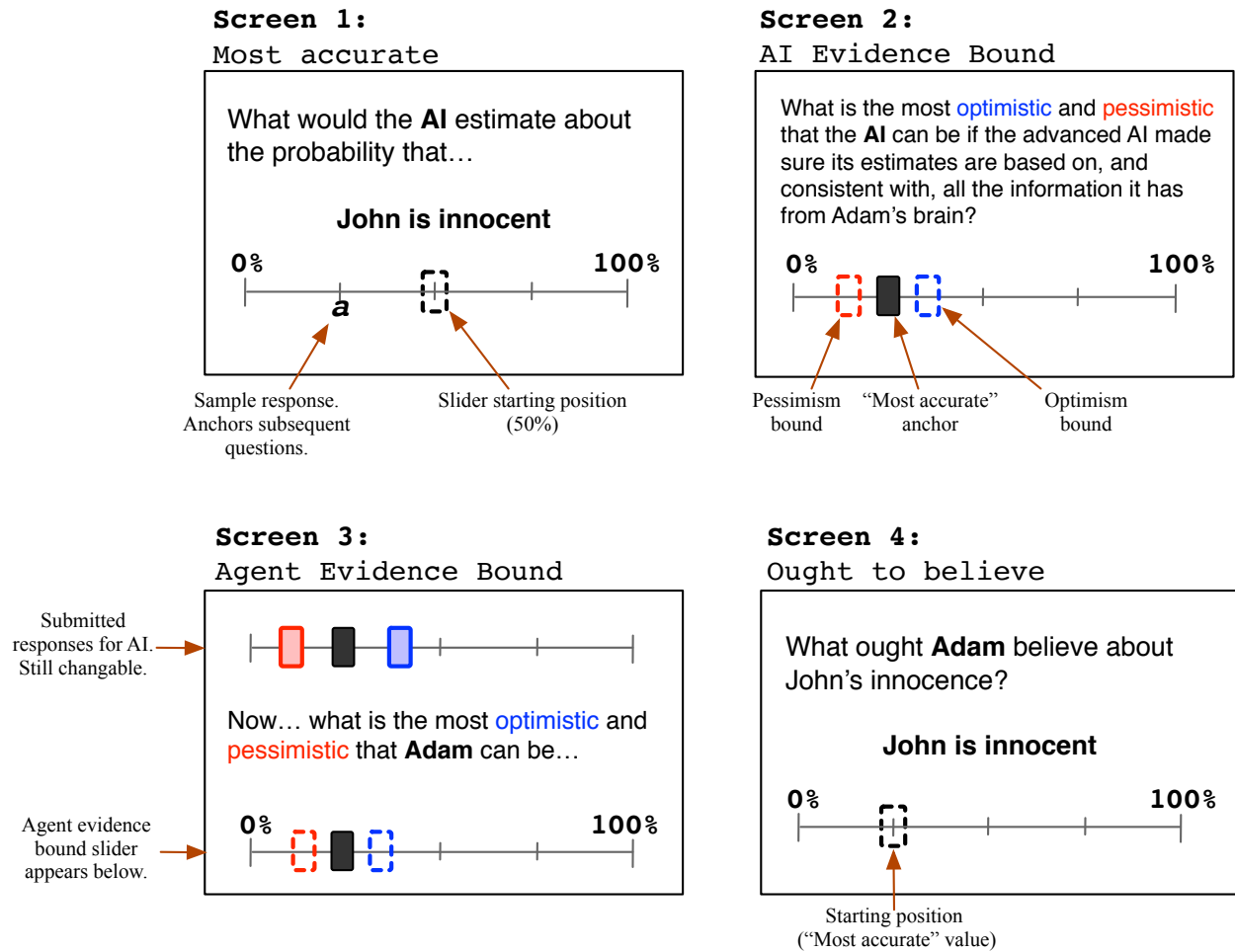


Figure 2. Method used to measure normative judgments in Study 1. Once a participant selected a most accurate estimate (a), this estimate was used to anchor all subsequent responses.

Ought to believe. On the next page, participants reported what the agent ought to believe. The default slider value was set to the “most accurate” estimate that the participant had previously provided. We operationalized prescribed motivated reasoning as deviation from this anchor in the direction of the morally preferable anchor. Furthermore, we considered ought judgments that fell outside of the range consistent with the character’s evidence (which participants had provided on the prior screen), as evidence for the alternative justification hypothesis.

Moral quality of the belief. On the next screen, participants reported their agreement with a series of statements about the moral value of the main character's beliefs in the vignette. Two statements measured commitment to the moral norm we hypothesized would be most operative in the vignette (e.g., "All else being equal, it is morally good to give your friend the benefit of the doubt"). Participants reported their agreement on a 7-point scale (anchored at "strongly disagree" and "strongly agree"). Intermingled with these questions were two distractor items about how pragmatically beneficial an optimistic belief would be in these situations (e.g., "It would make Adam feel good to believe that John is innocent"). As per our pre-registration, we treated these as distractor items and did not analyze them. See Appendix A for a list of all of the items across the six vignettes.

Finally, participants reported their sex and age and were debriefed.

Results

We first transformed participants' responses so that for all vignettes, higher probability estimates reflected more pessimistic (and therefore morally undesirable) beliefs, and lower estimates reflected more optimistic (and morally desirable) beliefs. We also averaged together the two moral concern items (r s ranged from .36 to .72 across vignettes, p s < .001²). Means and standard errors across vignette are shown in Table 2.

Prescribed motivated reasoning. We submitted participants' ought and most accurate judgments to a 2 x 6 mixed-design ANOVA, regressing estimates on judgment type (ought vs most accurate), vignette, and the interaction of judgment type and vignette. As expected,

² The moral concern items correlated weakly with each other in the Bully, $r(147) = 0.43$, $p < .001$, 95% CI [0.28, 0.55], Marriage, $r(140) = 0.35$, $p < .001$, 95% CI [0.2, 0.49], and Sex vignettes, $r(130) = 0.36$, $p < .001$, 95% CI [0.2, 0.50]. Our results do not change when using the composite score or either of the individual items. Detailed analyses of these items is available in the Supplemental Materials.

participants' ought estimates ($M = 59, SD = 28$) were significantly more optimistic than their judgments of what is most accurate ($M = 69, SD = 23$), $F(1, 833) = 145.47, p < .001$, with 53% of participants giving ought estimates higher than most accurate estimates (Figure 2A). The magnitude of this difference varied by vignette, as revealed by a significant judgment x vignette interaction, $F(5, 833) = 11.08, p < .001$. Follow-up analyses showed that, on average, participants prescribed motivated reasoning to others in five out of the six vignettes -- that is, in these vignettes, participants reported that the main character ought to hold a belief that differed from what would be most accurate for them to believe ($ps < .001$). The exception was the Sex vignette, for which the intended moral norm was against using gender-based statistical information to infer someone's status (e.g., assuming that a female healthcare worker is not a doctor).

We next computed participants' "prescribed motivated reasoning" by calculating the difference between participants' ought and most accurate estimates. When we regressed these difference on vignette, moral concern, and their interaction, we observed that, across vignettes, greater prescribed motivated reasoning was strongly associated with greater moral concern, $F(1, 827) = 32.34, p < .001$.

Table 2

Means (and standard deviations) for all main dependent measures in Study 1.

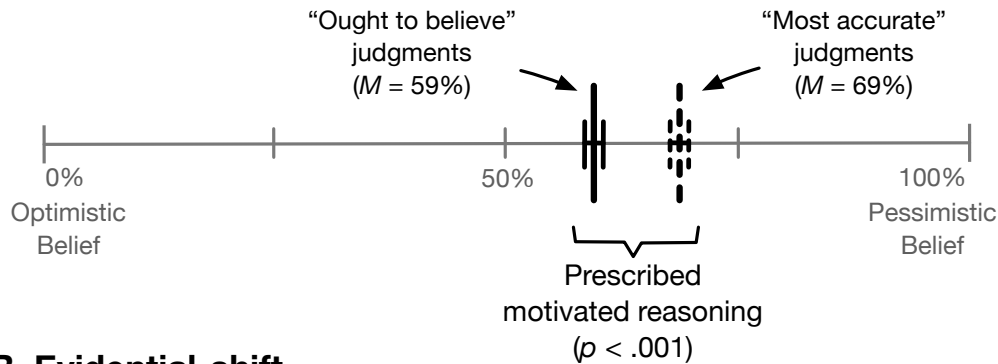
Vignette	Moral Concern	Most Accurate	Should Believe	Evidence-based Pessimism Boundary		Evidence-based Optimism Boundary	
				<i>AI</i>	<i>Character</i>	<i>AI</i>	<i>Character</i>
Bully	4.91 (1.05)	83 (16)	75 (20) ^a	91 (13)	92 (13)	62 (20)	61 (19)
Cancer	5.05 (0.97)	76 (20)	64 (31) ^a	88 (15)	86 (17)	61 (25)	58 (25) ^b
Friend	4.35 (1.00)	61 (26)	56 (24) ^a	77 (23)	75 (22)	42 (26)	33 (24) ^b
Marriage	5.26 (0.84)	51 (20)	32 (27) ^a	70 (19)	70 (18)	32 (20)	27 (21) ^b
Race	5.01 (1.08)	70 (19)	61 (22) ^a	83 (17)	83 (17)	47 (23)	46 (23)
Sex	4.90 (1.07)	70 (24)	69 (25)	81 (21)	82 (21)	55 (26)	53 (25)
Overall	4.91 (1.04)	69 (23)	59 (28) ^a	82 (20)	81 (20)	50 (26)	46 (26) ^b

Notes. Moral concern ratings were made on a 1-7 rating scale. Higher numbers indicate greater moral concern. All other judgments were made on a 0-100 scale, where higher numbers indicate higher subjective probability toward a pessimistic belief.

(a) Should and Most accurate judgments significantly different at $p < .05$ threshold.

(b) AI and character evidence-boundary significantly different at $p < .05$ threshold.

A. Prescribed motivated reasoning



B. Evidential-shift

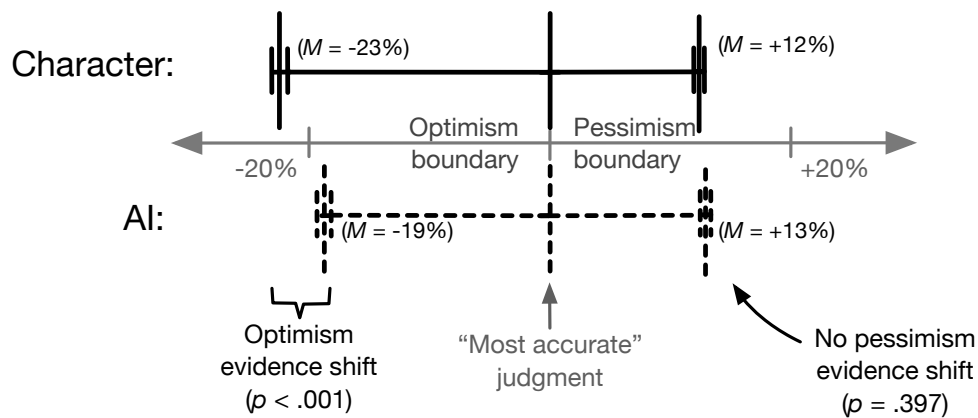


Figure 3. Main results from Study 1 averaged across all vignettes. (A) Mean estimates (error bars: ± 1 standard error) of what would be most accurate to believe in light of an objective assessment of the evidence (dotted) and what the main character ought to believe (solid). (B) Mean estimates (error bars: ± 1 standard error) of most optimistic and pessimistic that the character (solid) and AI (dotted) could be.

Evidential shift hypothesis. Participants also reported that the main characters in the vignettes were licensed to evaluate the evidence differently from the perfectly impartial, but equally informed, observer. We submitted participants' evidential optimism bounds for the AI and the main character to a 2x6 mixed-design ANOVA, regressing optimism bound on target (AI vs character), vignette, and the interaction of target and vignette. As expected, participants believed the character was licensed to be more optimistic ($M = 46$, $SD = 25$) than the objective

AI ($M = 50$, $SD = 26$), $F(1, 833) = 44.29$, $p < .001$, with 42% of participants giving higher optimism bound estimates for the character than for the AI. The strength of this character-AI difference varied across vignette, as revealed by a significant target x vignette interaction, $F(5, 833) = 5.44$, $p < .001$. Follow-up analyses showed that, on average, participants indicated a lower optimism bound for the character than for the AI in Cancer, Friend, and Marriage ($ps \leq .037$), but not in Bully, Race, and Sex ($ps \geq .071$).

It is possible that participants believed characters were licensed to be less confident in general rather than licensed to evaluate the evidence in a directional, motivated way. If this were the case, the range of beliefs licensed by the evidence would be greater for both optimistic and pessimistic beliefs. However, when we performed the same analysis on participants' evidence bounds for pessimistic belief, we observed no corresponding difference between how pessimistic the character ($M = 81$, $SD = 20$) and AI ($M = 81$, $SD = 20$) could be, $F(1, 833) = 0.72$, $p = 0.397$ (Figure 2b).

Finally, we observed mixed support for an association between an evidential shift and moral concern. Differences in evidence-bound optimism estimates were associated with prescribed inaccuracy, $F(1, 827) = 13.66$, $p < .001$; however, there was no association between these differences and our measure of moral concern, $F(1, 827) = 0.004$, $p = .948$.

Alternative justification. Lastly, our experimental design allowed us to test whether participants prescribe beliefs to others in a way consistent with the alternative justification hypothesis. Recall that the alternative justification hypothesis predicts that people will prescribe beliefs to others even when those beliefs unequivocally violate the evidence. Because participants explicitly reported what range of beliefs can be considered “consistent with” and “based on” the evidence, we examined how many participants prescribed beliefs that fall outside

this bound. This analysis revealed that, of the 53% of participants who prescribed more optimistic beliefs than what would be most accurate, 32% of these individuals (about 17% of the full sample) prescribed optimistic beliefs that fell outside the range licensed by the evidence. This proportion varied across vignette (see Table 3). In some vignettes, such as Friend, most participants who prescribed motivated optimism still prescribed optimism that fell inside evidentiary-bounds. However, in some vignettes, such as Cancer and Marriage, nearly half of participants who prescribed motivated reasoning did so by prescribing beliefs that were – by their own lights – unsupported by the evidence (44% and 50%, respectively).

Table 3

Percentage of participants showing evidence for the prescribed motivated reasoning, evidential shift, and alternative justification hypotheses across vignette in Study 1.

<i>Vignette</i>	<i>N</i>	Evidential Shift		Prescribed Motivated Reasoning Beyond Evidence	
		<i>Optimism</i>	<i>Pessimism</i>	<i>Overall</i>	
Bully	149	37%	16%	52%	13% (26%)
Cancer	132	32%	17%	54%	23% (44%)
Friend	139	58%	25%	56%	9% (15%)
Marriage	142	51%	30%	66%	33% (50%)
Race	145	38%	23%	56%	14% (26%)
Sex	132	37%	27%	34%	8% (22%)
Overall	839	42%	23%	53%	17% (32%)

Note. Pessimism evidential shift is kept here for comparison; on average there was no difference between AI and character evidence-licensed pessimism.

Discussion

Study 1 found support for prescribed motivated reasoning, the evidential-threshold shifting hypothesis, and the alternative justification hypothesis. In support of prescribed motivated reasoning, participants routinely indicated that another person ought to hold an inaccurate belief, with the degree of inaccuracy positively related to the moral benefit of holding

the inaccurate belief. We observed this across five of our six vignettes, suggesting that morality affects the beliefs people will prescribe to others across a wide variety of situations (as long as there is a strong enough moral norm). By contrast, these results are inconsistent with predictions that people think others ought to always reason objectively and solely on the basis of their evidence.

The “Friend” vignette in particular provides an important demonstration of prescribed motivated reasoning. In this vignette, the main character forms a belief about something that occurred in the past – whether the friend performed bad behaviors or not. Because of this, there is no plausible way this belief could be self-fulfilling. This finding therefore demonstrates that an (overly) optimistic belief does not have to be potentially self-fulfilling for people to prescribe it to others (c.f. Tenney et al., 2015). Rather, the moral benefit of trusting a friend’s testimony over the accumulated evidence engenders an obligation to give him the benefit of the doubt.

Study 1 also uncovered evidence for the evidential-shift hypothesis: Participants reported that the main character, who had a moral reason to be optimistic, was permitted to evaluate the evidence they had differently than an impartial observer, but only in a way that licensed the morally desirable optimistic belief. This effect was driven by three out of the six vignettes, Cancer, Marriage, and Friend (Table 2). In these vignettes, the character has a social obligation to discount certain kinds of evidence (such as statistical evidence) or more heavily weight certain kinds of evidence (such as testimony). In two vignettes, Race and Bully, we observed evidence for prescribed motivated reasoning but failed to observe evidence for evidential shift. We did not observe any evidence for evidence-shifting in the Sex vignette; however, this may have reflected the fact that people do not seem to think moral concerns ought to play any role when making inferences about others’ status based on sex information, as we failed to find evidence for

prescribed motivated reasoning in this vignette as well. Although the participants who exhibited evidential shifting were also more likely to prescribe motivated reasoning, we did not observe an association between participants' judgments of moral concern about the beliefs and their evaluations of the evidence. This result was unexpected, and may have reflected the observation that, unlike all-things-considered belief prescriptions, there is a limit to how plausibly the evidence in our vignettes could be shifted. This is consistent with the observation that participants' average evidence-shifting was much smaller (4%) than their average prescribed motivated reasoning (10%). Nevertheless, our results show clear evidence for motivated evidence shifting in the predicted direction.

Finally, Study 1 demonstrated support for the alternative justification hypothesis. About one-third of participants who prescribed motivated reasoning prescribed a belief that fell beyond the range of beliefs they considered consistent with the evidence. This was most pronounced in two vignettes, Cancer and Marriage, in which almost half of participants who prescribed motivated reasoning did so by prescribing a belief they knew to be inconsistent with the evidence. However, it's worth noting that our measure of alternative justification (a prescribed belief outside the evidence bounds) is very conservative, and therefore likely underestimates the extent to which alternative justification played a role. Many participants may have prescribed motivated reasoning because they thought the moral quality of the belief directly justified doing so, yet still prescribed a belief that was licensed by the evidence. We address this in Studies 2 and 3, by testing whether evaluations of the belief's moral quality predict variation in evaluations of a belief after accounting for changes in evidential quality.

Study 2

In Study 2, participants read vignettes similar to those used in Study 1. However, instead of prescribing beliefs to someone who is currently reasoning about what to believe, they evaluated the belief that a character had already formed – either the evidence-based belief, or the optimistic belief. The character was either socially close to the target of the belief (e.g., a friend or spouse), and so had a moral reason to adopt the optimistic belief, or socially distant (e.g., a friend of a friend), so had a weaker moral obligation, but access to the same evidence. Participants were then asked to evaluate the extent to which formed beliefs were justified, were permissible, were supported by sufficient evidence, constituted knowledge, and were moral. These methodological changes allowed us to conceptually replicate and extend the findings from Study 1. First, evaluating the beliefs of two people who share information but differ in their social obligations (vs. a person and an AI) more clearly isolates the role of social and moral demands, and better reflects real-life belief evaluation. Second, by asking participants to make judgments of people who already hold beliefs (vs. judgments of which beliefs people should form), we can directly test whether observers permit believers to base their belief formation in part on moral concerns. This alleviates a possible concern in Study 1, namely that observers want others to form morally beneficial beliefs, but to be unaware that they are doing so. Third, because the socially close and socially distant believers must each justify their formed beliefs on *some* grounds, we can test whether moral justifications for belief are viewed as more justifying than non-moral justifications (e.g., a strong preference to hold a belief), thereby further isolating the relevance of morality for belief evaluation. Finally, by investigating people's evaluations of already-formed beliefs, we can measure the prescribed motivated reasoning, evidential shift, and alternative justification hypotheses using multiple, triangulating measures that differ from those employed in Study 1.

In line with the prescribed motivated reasoning hypothesis, we predicted that people's judgments of a belief's "justifiability" and "permissibility" – which correspond closely to all-things-considered evaluations of a belief – would be a function of both the belief (whether it was evidence-based or optimistic) and the believer (socially close or socially distant). Specifically, we predicted that for the socially close believer, relative to the socially distant believer, the optimistic belief would be rated more favorably, and the pessimistic, evidence-based belief more harshly, even though both believers had access to the same evidence. In line with the evidential-shift hypothesis, we predicted that people would be more willing to say that morally preferable beliefs, compared to neutral or immoral beliefs, were supported by "sufficient evidence," and that when the belief turned out to be true, that the character had "knowledge" of the outcome (c.f. Starmans & Friedman, 2012; Kim & McGrath, 2019, *inter alia*). We further predicted that these effects would be greater to the extent the belief was judged morally preferable. Finally, in line with the alternative justification hypothesis, we predicted that the perceived moral quality of the belief would predict judgments of the belief's "justifiability" and "permissibility," above and beyond its evidential quality. These findings would constitute a conceptual replication of Study 1, and therefore offer converging evidence for our claims.

Methods

Participants. We recruited 1,021 adults (638 reported female, 524 reported male, mean age 40) from MTurk. An additional 145 participants were excluded for failing at least one of three comprehension checks (explained below).

Design. Study 2 used a 2 (Social Distance) x 2 (Presence of moral reason) x 4 (Vignette) mixed design. Participants read one of four vignettes from Study 1 that readily enabled our social distance manipulation: Bully, Cancer, Friend, or Marriage.

Within each vignette, participants read about the character in Study 1 who, by virtue of being socially close to the target, has a legitimate moral reason to hold an optimistic belief about them (moral reason condition), as well as another character who, in light of being socially distant from that person, lacked a justifiable moral reason to hold an optimistic belief (non-moral reason condition). For instance, the close friend in the Friend vignette weighed the evidence against his obligation to trust his friend whereas the socially distant observer weighed the evidence against her “strong preference” to be optimistic (see exact text below). We stipulated that the socially distant observer had a strong preference to adopt the optimistic belief because, without such an explanation, participants are likely to fill in an explanation and may do so in a way that introduces confounds. For instance, when the socially distant observer adopts a putatively inaccurate belief, they may infer it was on the basis of prior, hidden evidence, or on the basis of a shared moral commitment. Stipulating that a non-evidential belief was caused by a preference eliminates both possibilities. However, as noted above, it also allows us to test whether a moral reason to adopt a motivated belief is seen as more justifiable than a bare motivation to do so. See Table 4 below for full text of this manipulation in the Friend vignette.

At the end of the vignette, participants learned either that these two people adopt an evidence-based belief (despite their moral or non-moral reason to adopt an optimistic belief) or they adopt an overly-optimistic belief (on the basis of their moral or non-moral reason); this comprised our Belief manipulation. For instance, in the Friend vignette, when the two characters adopted the evidence-based belief, they believed that John is guilty, and did so after “setting aside” their moral obligation or strong preference. When they adopted their preferred belief, they did so either because of their moral obligation or because of their strong preference. See Table 4 for the text in the Friend vignette.

Table 4

Text used to manipulate the presence of a preferred belief or dis-preferred belief across conditions in the Friend vignette in Study 2.

Social Distance Condition	Belief Condition	
	Optimistic Belief	Evidence-based belief
Socially Close (Moral reason to maintain optimism)	Adam says to Jasmine that, because he is close friends with John, he knows that he has an obligation to trust John and to give him the benefit of the doubt. After all, that is what a true friend would do, and it is what John asked him to do.	Adam says to Jasmine that, because he is close friends with John, he knows that he has an obligation to trust John and to give him the benefit of the doubt. After all, that is what a true friend would do, and it is what John asked him to do.
	Adam thinks about everything that he and Jasmine discussed, and thinks about being loyal to John. In the end, based on his loyalty to John, Adam decides to believe that John is innocent.	Adam thinks about everything that he and Jasmine discussed, and thinks about being loyal to John. In the end, Adam sets aside his loyalty to John and decides to believe that John is guilty.
Socially distant (Non-moral reason to maintain optimism)	Jasmine is no longer friends with John and so has no obligations towards him. But, it makes her upset to think that any person would get involved in drugs, so she has a strong preference to believe that John is innocent.	Jasmine is no longer friends with John and so has no obligations towards him. But, it makes her upset to think that any person would get involved in drugs, so she has a strong preference to believe that John is innocent.
	Jasmine thinks about everything she and Adam discussed, and thinks about her preference. In the end, based on her preference, Jasmine decides to believe that John is innocent.	Jasmine thinks about everything she and Adam discussed, and thinks about her preference. In the end, Jasmine sets aside her preference and decides to believe that John is guilty.

Procedure. Vignettes were minimally modified from Study 1 so that the main character discusses what they should believe with someone who, over the course of the conversation, ends up with the same relevant information.³ For instance, in the Friend vignette, participants would have read the following after reading the vignette excerpted above in Study 1:

³ An ideal manipulation would involve two characters who have the same evidence about a target individual, but where one has a strong moral obligation to believe a certain way, and the other has no obligation whatsoever. Instead, our design involves a manipulation of social distance, where the distant character still has some relationship to the target, as we feared that participants

Adam decides to talk about all of this with his old friend Jasmine. Jasmine was friends with Adam and John all her childhood. However, when the three of them went to college she became overwhelmed with work and fell out of touch with them. Because of this, she no longer considers herself to be very good friends with John or Adam. And indeed, when Adam has brought Jasmine up to John, he no longer considers her a close friend either. Adam shares everything he recently learned about John with Jasmine. Because Jasmine already knew Adam and John well from their shared childhood, she already knew about John's behavior in elementary school, middle school, and high school.

Participants were further instructed to “Assume that Adam has shared everything that he knows about John and John's situation with Jasmine so that they now have the exact same relevant information.” To reinforce the idea that the two believers have the same information, participants reported what an objective observer would estimate based on their shared information. For instance, in the Friend vignette, participants read, “Assume that Adam has shared everything that he knows about John and John's situation with Jasmine so that they now have the exact same relevant information. Based on this information, what would a perfectly objective perceiver estimate about the probability that John is innocent?” As in Study 1, participants provided an estimate on a 0-100% rating scale with 5%-point increments.

After learning what each character in the vignette decides to believe, participants responded to eight questions about that character and their belief. Two questions measured

would judge it implausible that a character who was completely uninvolved could have the same information about the target individual as a close friend or a spouse would. In light of this, the characters who occupy the socially-distant role likely still had some social obligations toward the person they eventually form a belief about, and indeed, our results below bear this out. For this reason, it is more accurate to say that we manipulated the relative social distance and therefore the relative strength of social obligations on belief. We thank Liz Harman and Melissa Lane for encouraging us to clarify this point.

overall belief quality, including whether the belief was justified for the target to hold (e.g., “How justified is Adam to believe that John is innocent?”) and whether it was permissible for the target to hold (e.g., “How permissible is it for Adam to believe that John is innocent?”). Two questions measured the moral quality of the belief, including whether the belief was loyal (or helpful, e.g., “How loyal to John is Adam's belief that John is innocent?”) and whether the person’s reasoning was morally good (e.g., “How morally good is Adam's reasoning about whether John is innocent or not?”). Two questions measured the evidentiary quality of the belief, including whether the person had sufficient evidence for their belief (e.g., “Is Adam's evidence sufficient for his belief that John is innocent?”), and whether they should have obtained more evidence before coming to their belief (e.g., “How much more evidence should Adam have acquired before coming to believe that John is innocent?”). Finally, two questions measured the moral quality of the character, including whether they were trustworthy (e.g., “How trustworthy do you find Adam to be?”) and whether they were a good person (e.g., “How good of a person is Adam?”).

Participants responded to these questions using 7-point rating scales with higher values indicating more positive evaluations. The order that participants learned about each character’s belief (and evaluated them for that belief) was counterbalanced across participants. The eight questions above were presented in a random order for each judgment target.

Participants then read that the belief that the two characters adopted turned out to be true. For instance, in the Friend vignette, if the characters adopted the belief that “John is innocent,” then it turned out that, in fact, he was innocent. Likewise, if the characters adopted the belief that “John is guilty,” then it turned out that he was guilty. For each character, participants reported their agreement (1: strongly disagree, 7: strongly agree) with statements claiming that the characters “knew” the outcome (e.g., “Adam knew that John was guilty.”).

Lastly, participants reported their sex and age and were debriefed.

Results

Data preparation and analysis plan. We created composite measures of the moral quality of belief, the moral quality of the believer, the evidential quality of belief, and overall belief quality. The two-item pairs for each question passed our preregistered criterion ($r > .50$) for analyzing such a composite across both evidence-based belief and optimistic belief conditions. There were two exceptions: The two item measures correlated with each poorly in the Bully, $r(512) = 0.49, p < .001, 95\% \text{ CI } [0.42, 0.55]$, and Friend, $r(520) = 0.31, p < .001, 95\% \text{ CI } [0.23, 0.38]$ vignettes. Because our results are largely the same when analyzing these judgments independent of one another (see Supplemental Materials), we present results for the composite measure here for all measures.⁴

We subjected each of our five DVs – moral character of believer, moral quality of belief, evidential quality of belief, attributions of knowledge, and overall belief quality – to a series of 2 (Belief: Optimistic vs Evidence-based) x 2 (Reason: Moral vs Non-moral) x 4 (Vignette) fully-crossed ANOVAs. A main effect of belief would reflect that, generally, one belief (e.g., the evidence-based belief) was better for the characters to hold than another. A main effect of reason could reflect a few things, including participants' inferences that the target actually had better grounds for their belief (e.g., because they know them better) or worse grounds for their belief (e.g., due to a duty to be more diligent in their belief formation). However, the key prediction made by the prescribed motivated reasoning hypothesis is a belief-by-reason interaction, as that

⁴ All analyses yield identical results with one exception. In the Friend vignette, we observed a significant Reason x Belief interaction when analyzing participants' judgments of how "loyal" the resulting belief was, $F(1, 259) = 141.36, p < .001, \eta^2_G = 0.14$, but not of how "morally good" the believer's reasoning was, $F(1, 259) = 2.50, p = .115$. See the Supplemental Materials for more detail.

would demonstrate that the quality of the believer and belief depends on whether the believer had a moral reason to adopt a particular belief and whether they honored that reason or not.

Results are shown in Figure 4. For ease of communicating the findings, we refer to the two conditions in the reason manipulation as the “close observer” for the character who was socially close (e.g., friend, wife, husband, or teacher), and therefore had a social obligation to form an optimistic belief, and the “distant observer” for the character who was relatively socially more distant (e.g., friend of friend, friend of teacher, etc.) and had merely a strong preference to adopt an optimistic belief.

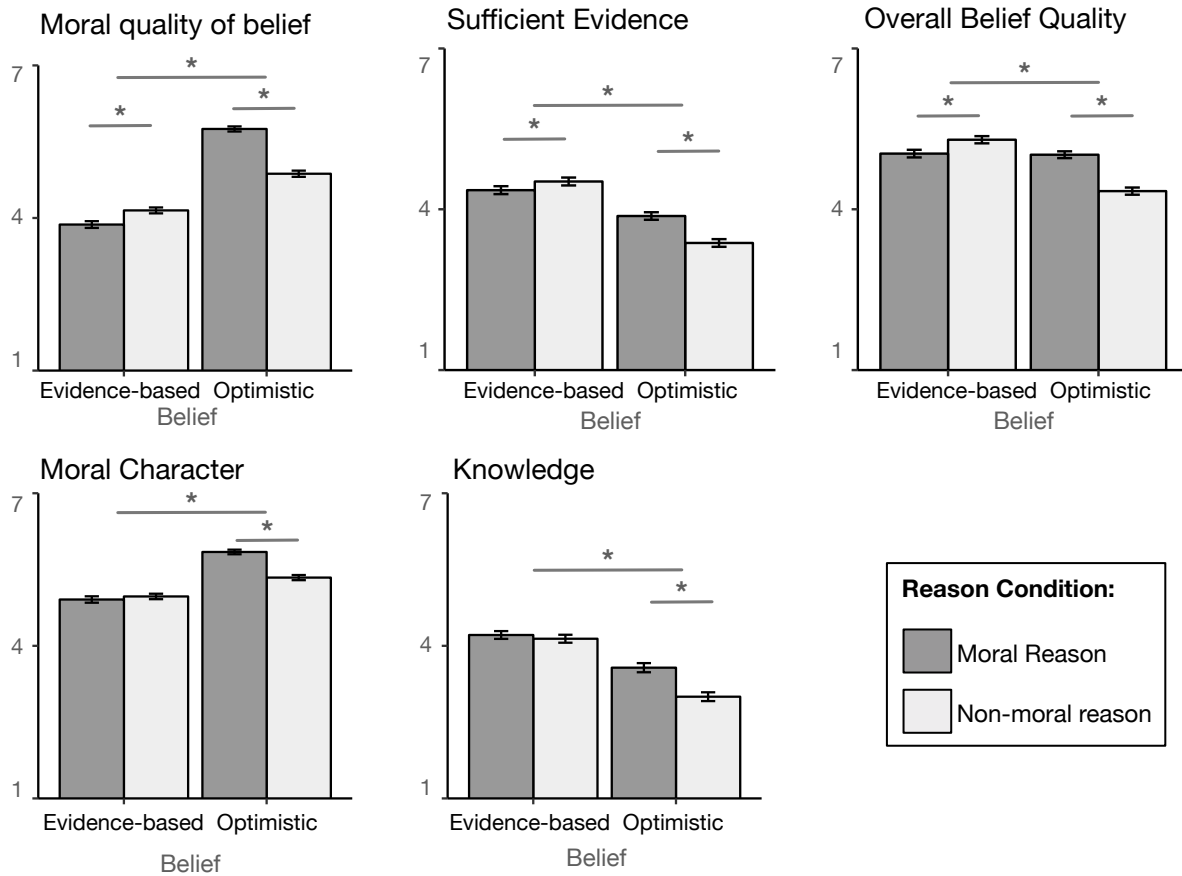


Figure 4. Means and standard error (error bars: ± 1 standard error) across condition for ratings of moral quality of belief, moral quality of believer's character, evidential quality, knowledge, and overall belief quality in Study 2. Higher values indicate more positive evaluations. * $p < .005$

Moral evaluations of belief and believer. On average, participants rated the optimistic belief ($M = 5.31$, $SD = 1.33$) to be morally better than the evidence-based belief ($M = 4.01$, $SD = 1.42$), $F(1, 1013) = 331.23$, $p < .001$, $\eta_G^2 = 0.21$, and, on average, participants rated the quality of the socially-close character's belief ($M = 4.80$, $SD = 1.63$) no differently than the distal character's belief ($M = 4.51$, $SD = 1.39$), $F(1, 1013) = 70.71$, $p < .001$, $\eta_G^2 = 0.01$. However, these two main effects were qualified by a significant Belief x Reason interaction, $F(3, 1013) = 23.1$, p

$< .001$, $\eta_G^2 = 0.05$. In the evidence-based condition, participants rated the socially close character's belief to be morally worse ($M = 3.87$, $SD = 1.51$) than the distal character's belief ($M = 4.15$, $SD = 1.32$), $F(1, 512) = 31.66$, $p < .001$, $\eta_G^2 = 0.01$. By contrast, in the optimistic belief condition, participants rated the socially-close character's optimism to be morally superior ($M = 5.75$, $SD = 1.13$) to the socially-distant character's optimism ($M = 4.87$, $SD = 1.37$), $F(1, 501) = 313.14$, $p < .001$, $\eta_G^2 = 0.12$.

Evaluations of the believers' moral character followed a similar pattern. On average, participants rated the two individuals as morally better when they were optimistic ($M = 5.59$, $SD = 1.11$) compared to when they formed beliefs based on the evidence ($M = 4.94$, $SD = 1.33$), $F(1, 1013) = 94.44$, $p < .001$, $\eta_G^2 = 0.07$, while rating the socially-close individual slightly morally better ($M = 5.37$, $SD = 1.33$) than the distal individual ($M = 5.15$, $SD = 1.20$) overall, $F(1, 1013) = 48.32$, $p < .001$, $\eta_G^2 = 0.01$. However, these main effects were qualified by a Belief x Reason interaction, $F(1, 1013) = 80.34$, $p < .001$, $\eta_G^2 = 0.01$. Participants did not judge the socially-close person to have worse character ($M = 4.91$, $SD = 1.43$) than the distal character ($M = 4.97$, $SD = 1.22$) when they both adopted the evidence-based belief, $F(1, 512) = 1.64$, $p = 0.202$, $\eta_G^2 = 0$. By contrast, when the two people both adopted the optimistic belief, participants judged the socially-close individual to have better character ($M = 5.85$, $SD = 1.02$) than the socially-distant individual ($M = 5.34$, $SD = 1.14$), $F(1, 501) = 145.18$, $p < .001$, $\eta_G^2 = 0.05$.

Evaluation of evidence and knowledge. To test for the evidential shift hypothesis, we next analyzed participants' evaluations of the believers' evidence and knowledge. As predicted, participants' judgments of the evidential-quality of the two people's beliefs followed a pattern similar to their moral judgments. On average, participants rated the socially-close and socially-distant persons' evidence-based beliefs ($M = 4.44$, $SD = 1.68$) to be evidentially better than their

optimistic beliefs ($M = 3.62$, $SD = 1.62$), $F(1, 1013) = 85.31$, $p < .001$, $\eta_G^2 = 0.07$. On average, there was a minor, but significant difference in the evidentiary quality of the socially-close character's ($M = 4.12$, $SD = 1.65$) and socially-distant character's ($M = 3.95$, $SD = 1.74$) belief, $F(1, 1013) = 18.94$, $p < .001$, $\eta_G^2 = 0$. As predicted, we observed a significant Belief x Reason interaction, $F(1, 1013) = 75.56$, $p < .001$, $\eta_G^2 = 0.01$. When participants judged the evidence-based belief, the socially close character was seen as having less sufficient evidence for their belief ($M = 4.36$, $SD = 1.68$) relative to the socially distant character ($M = 4.52$, $SD = 1.67$), $F(1, 512) = 9.6$, $p = .002$, $\eta_G^2 < 0.01$. By contrast, when they both adopted the optimistic belief, the distant character ($M = 3.37$, $SD = 1.61$) was now seen to have less sufficient evidence for their belief relative to the close character ($M = 3.87$, $SD = 1.58$), $F(1, 501) = 79.8$, $p < .001$, $\eta_G^2 = 0.03$.

There was a robust association between the perceived moral quality of the belief and judgments that the believer's evidence was sufficient. We regressed judgments of evidential quality on participants' moral evaluations of the belief separately for each belief condition (evidence-based and optimistic) across the four vignettes, yielding eight tests total. These regressions included a Reason x Moral quality interaction, and by-subject random intercepts of evidentiary quality. Across all eight tests, morality significantly predicted participants' judgments concerning whether the believer had sufficient evidence for their belief ($bs > 0.25$, $ts > 4.64$, $ps < .001$). The worse the moral quality of the belief (e.g., the less helpful or loyal the belief was judged to be), the more strongly participants indicated that the believer lacked sufficient evidence for the belief.

Participants' attributions of knowledge largely recapitulated their evaluations of the evidentiary quality of belief. On average, participants attributed greater knowledge to believers in the evidence-based belief condition ($M = 4.18$, $SD = 1.76$) compared to the optimistic belief

condition ($M = 3.29$, $SD = 2.00$), $F(1, 1013) = 71.81$, $p < .001$, $\eta_G^2 = 0.06$. On average, participants attributed more knowledge to the socially-close character ($M = 3.90$, $SD = 1.91$) than to the socially-distant character ($M = 3.58$, $SD = 1.95$) belief, $F(1, 1013) = 82.8$, $p < .001$, $\eta_G^2 = 0.01$. However, similar to the results above, we observed a significant Belief x Reason interaction, $F(1, 1013) = 50.07$, $p < .001$, $\eta_G^2 = 0.01$. When participants were told that the believers' evidence-based belief was true, there was no difference in knowledge attributed to the close ($M = 4.21$, $SD = 1.74$) and distant characters ($M = 4.14$, $SD = 1.78$), $F(1, 512) = 2.67$, $p = .103$, $\eta_G^2 < 0.01$. However, when participants were told that the optimistic belief turned out to be true, participants attributed more knowledge to the close character ($M = 3.57$, $SD = 2.02$) than they did to the distal character ($M = 3.00$, $SD = 1.95$), $F(1, 501) = 111.64$, $p < .001$, $\eta_G^2 = 0.02$.

Overall belief quality. To test for prescribed motivated reasoning, we next investigated participants' judgments of overall belief quality, which was a composite measure of their judgments that the belief was justified and that it was permissible for the believer to hold that belief. On average, participants judged the optimistic belief to be less justified and permissible ($M = 4.68$, $SD = 1.51$) than the evidence-based belief ($M = 5.17$, $SD = 1.58$), $F(1, 1013) = 35.05$, $p < .001$, $\eta_G^2 = 0.03$, while the socially close character's beliefs were rated slightly better on average ($M = 5.03$, $SD = 1.53$) than the distal character's beliefs ($M = 4.82$, $SD = 1.59$), $F(1, 1013) = 29.67$, $p < .001$, $\eta_G^2 = 0.01$. As above, these main effects were qualified by a significant Belief x Reason interaction, $F(1, 1013) = 155.22$, $p < .001$, $\eta_G^2 = 0.03$. When the two characters adopted the evidence-based belief, the close character was seen as less justified and given less permission ($M = 5.04$, $SD = 1.62$) for doing so than the distal character ($M = 5.30$, $SD = 1.52$), $F(1, 512) = 23.00$, $p < .001$, $\eta_G^2 = 0.01$. By contrast, when they both adopted the optimistic belief, the close character was now seen as more justified for doing so and viewed as more permitted to

do so ($M = 5.02$, $SD = 1.43$), relative to socially-distant character ($M = 4.34$, $SD = 1.51$), $F(1, 501) = 167.03$, $p < .001$, $\eta_G^2 = 0.05$.

These all-things-considered judgments of belief quality may have reflected differences in participants' evaluations of the evidential quality of the husband and friend's beliefs, or they could have reflected those differences in addition to independent moral concerns participants have about the beliefs. This latter possibility is predicted by the alternative justification hypothesis. To test whether the moral quality of the belief predicted overall justifiability and permissibility, we regressed participants' overall judgments on their ratings of the belief's evidentiary quality and on the belief's moral quality. We did this separately for the evidence-based belief condition and the optimistic belief condition for each of the four vignettes. These regressions included a Reason x evidentiary quality and Reason x belief moral quality interactions as well as random by-subject intercepts of overall belief quality. Unsurprisingly, evidentiary quality predicted overall belief quality in both the evidence-based belief conditions ($bs > 0.35$, $ts > 6.80$, $ps < .001$) and the optimistic belief conditions ($bs > 0.33$, $ts > 5.12$, $ps < .001$) across all four vignettes. However, even when accounting for differences in evidentiary quality, the moral quality of the belief independently predicted overall judgments in the evidence-based belief conditions ($bs > 0.20$, $ts > 3.42$, $ps < .001$) and in the optimistic belief conditions ($bs > 0.33$, $ts > 2.65$, $ps < .008$) across all four vignettes.

Variation across vignette. All of the results we report above were qualified by by-Vignette interactions (see by-vignette results in Appendix B). Specifically, while we observed clear evidence for prescribed motivated reasoning and for evidential-shifting when averaging across all four vignettes, these findings were driven by results in three out of the four vignettes we tested. Similar to our findings in Study 1, we observed clear evidence for prescribed

motivated reasoning and for evidential shifting in all vignettes except the Bully vignette. In Bully, two results were consistent with our predictions. First, we observed the predicted interactions on the perceived moral quality of the belief and believer ($ps < .05$), suggesting that we effectively manipulated the perceived moral obligation of the two believers in this vignette. Second, as noted above, we observed statistically significant associations between the moral quality of the belief and other properties of the belief. The moral quality of the belief, such as how helpful and constructive the teacher's belief was, was strongly associated with how sufficient the believer's evidence was in both the evidence-based belief ($b = 0.51, se = 0.06, t = 7.84, p < .001$) and optimistic belief ($b = 0.39, se = 0.06, t = 6.85, p < .001$) conditions. Additionally, the moral quality of the belief predicted how justified and permissible the belief was judged to be, even after accounting for differences in evidential quality in both the evidence-based belief ($b = 0.20, se = 0.06, t = 3.42, p < .001$) and optimistic belief ($b = 0.45, se = 0.06, t = 7.31, p < .001$) conditions, in line with the alternative justification hypothesis. However, we did not observe the predicted Reason x Belief interactions on overall belief evaluation, sufficient evidence, or knowledge ($ps > .05$).

Discussion

Study 2 provided additional evidence that people treat morality as a legitimate basis on which to evaluate belief. Mirroring participants' "ought" judgments in Study 1, Study 2 showed that a moral reason to be optimistic, which we manipulated by varying someone's social distance to the person they were forming a belief about, increased the perceived permissibility and justifiability of morally desirable optimistic beliefs and lowered the permissibility and justifiability of evidence-based, but morally undesirable beliefs. As a part of this, we found that a distant observer, who lacked a moral reason to form an optimistic belief, was not licensed to do

so on the basis of a strong preference to form the belief. Thus, Study 2 confers some additional evidence that morality provides either a unique – or at least a stronger – justification for motivated reasoning compared to the reasons that underlie many mundane cases of actual motivated reasoning.

Study 2 also provided additional evidence for the evidential-shift and alternative justification hypotheses. Mirroring the finding from Study 1 that the morally-justified characters were licensed to evaluate the evidence more optimistically, Study 2 found that optimistic beliefs were seen as more evidentially supported for the characters with moral reasons to be optimistic than for characters who had mere preferences to be optimistic. Likewise, having a moral reason not to adopt a belief, on average, increased the evidence required relative to having a strong preference not to adopt that belief. And mirroring the finding from Study 1 that some participants prescribed beliefs that they knew to be unsupported by the evidence, Study 2 found that the moral quality of the belief – for instance, how helpful or loyal it was – predicted evaluations of the overall quality of the belief even when accounting for differences in the evidential quality of the belief.

However, there is a skeptical interpretation of our findings in Study 2 that our design fails to rule out. Specifically, it is possible that some participants did not accept the premise in the vignettes that the morally-motivated and non-morally-motivated characters had the same relevant information, and instead believed that the morally-motivated character still had privileged optimism-licensing evidence. For instance, perhaps in the divorce vignette, participants thought that the husband did not tell his friend about how strongly he loves his new wife, and they consider the strength of love a piece of evidence about the likelihood of divorce that favors optimism. This private information could make optimistic beliefs more justified than pessimistic

beliefs relative to someone with nearly the same information but lacking that specific evidence. Because we only asked participants to estimate the most accurate estimate based on both characters' information, we are unable to detect whether this happened, and therefore to definitively rule it out as an alternative explanation.

We address this worry in Study 3. Specifically, we took additional steps in the text of the vignette to equalize the information held by the morally-motivated and non-morally-motivated believers and then also asked participants to independently report the state of the evidence they thought that each person had. We could then test whether participants on average provide different estimates to the morally-motivated and non-morally-motivated believers, and more crucially, restrict our analyses to only those who provide comparable evidence judgments for the two individuals. Although this method has the advantage of addressing the worry stated above, this method also has several drawbacks. First, asking participants to separately estimate judgments from the perspective of the morally-motivated and non-morally motivated character is likely confusing in light of prior textual evidence that they should both have the same relevant information. Resolving this confusion may lead participants to infer differences between these individuals that they might otherwise not have inferred. Second, it is possible that some people will provide estimates of what is most accurate taking into account their judgments that the information each person has ought to be weighed differently (which may have been affected by the moral status of each individual; see Study 1). Both of these kinds of participants would be excluded from our analysis. For this reason, Study 3 represents an especially conservative test of our predictions.

Study 3

The primary goal of Study 3 was to test whether people evaluate beliefs as predicted by the *alternative justification* and *evidential shift* hypotheses even after removing all participants who report differences in the information held by the morally-motivated and non-morally-motivated characters. To this end, we conducted a lightly-modified replication of the “marriage” vignette from Study 2. This vignette showed the strongest effect of moral concern in Study 2, and we reasoned that it would yield the most power to reliably detect any differences in belief evaluation after removing a potentially large proportion of our recruited sample.

A secondary goal of Study 3 was to examine which properties of belief people might associate with moral justifications to adopt non-evidential beliefs or to raise the evidential threshold for belief. As discussed in the Introduction, prior literature suggests that people apply moral norms to beliefs because they expect such beliefs to bring about morally appropriate behavior, or potentially because they treat certain attitudes as intrinsically good or bad. These motivations are not mutually exclusive but may both contribute to the sense that someone has a moral obligation to adopt a particular belief. Lastly, we investigated a possible alternative explanation for our results; namely, that people who appear to have little control over their belief are permitted to believe differently than others. For instance, people could think that a newlywed, relative to a neutral observer, is unconsciously or uncontrollably biased in his reasoning. In this case, observers may judge that he is not responsible for what he believes and therefore not appropriately subject to standards of belief that demand evidence-based reasoning.

Methods

Participants. We recruited 233 adults (126 reported female, 106 reported male, mean age 39) from MTurk. An additional 65 participants were excluded for failing at least one of three comprehension checks.

Design. This study used a 2 (Reason: moral reason vs non-moral reason) x 2 (Belief: evidence-based vs optimistic) within-between design.

Procedure. The procedure for Study 3 followed that of Study 2 with few changes, described below. The full text of the study is available in Appendix C (as well as the Supplemental Materials).

Participants read the marriage vignette from Study 2, in which a 19-year-old newlywed named Brian learns that 70% of newlyweds like him end up getting a divorce within five years of marrying. Unsure what to think, he confides in his longtime friend Patrick who knows him and his relationship well. For instance, Brian tells Patrick, “how he feels about Maya, what he likes and dislikes, what his gut tells him, and everything he hopes will come to pass.” This was done to ensure that even “emotion-” or “intuition-” based evidence held by Brian was shared with his friend. In addition to this, participants read that Brian and Patrick both agreed that Brian and Maya are similar to other recently married couples, but also that Brian and Maya “may grow and change in unexpected ways over the next few years.”

Participants then reported what would be the most accurate estimate that Brian and Maya will divorce based on the information they each had. Specifically, participants reported what would be “most accurate” based on the information contained in Brian’s mind and separately in Patrick’s mind. As in Study 2, participants reported their answer using a 0-100% scale with intervals every 5%. This procedure mirrored the procedure in Study 2 except that participants provided most accurate ratings for Brian and Patrick individually, yielding two potentially distinct ratings.

After evaluating Brian and Patrick’s evidence, but before learning what they actually decide to believe, participants made “in principle” judgments about Brian and Patrick’s beliefs.

Specifically, participants answered four questions for each character, probing obligation (e.g., “To what extent does Brian have a moral obligation to believe the best of his marriage?”), behavioral impact (e.g., “If Brian is in fact pessimistic about his marriage, how difficult would it be him to act in all respects as if he is genuinely optimistic about his marriage?”), intrinsic wrongness (e.g., “If Brian could successfully act in all respects as if he is genuinely optimistic about his marriage, would it still be morally wrong for him to in fact be secretly pessimistic?”), and voluntary control (e.g., “To what extent does Brian have voluntary control over what he believes about his marriage?”). As noted above, these questions were included as exploratory measures of what considerations impact people’s judgments that beliefs are subject to moral norms. Participants responded to these questions using 7-point rating scales. All eight questions were shown on the same page in four pairs of two, such that the version of the question for Brian and Patrick were always adjacent. The four pairs were shown in a random order for each participant, however, questions (iii) and (iv) were always shown together in order to highlight the contrast between them. Finally, between participants we counterbalanced whether, within each question pair, participants first responded for Brian or for Patrick.

Participants then learned what Brian and Patrick believed about Brian’s marriage. Participants were randomly assigned to learn that Patrick and Brian adopt the evidence-based belief that there is a 70% chance of divorce in the next five years (evidence-based belief condition) or the optimistic belief that there is a 0% chance of divorce in the next five years (optimistic belief condition; between-participants manipulation). Whether they adopted the desirable or undesirable belief, the two characters did so either because of, or despite a moral reason (Brian) or a non-moral preference. For instance, Brian had an obligation to maintain an optimistic attitude about his relationship despite the new information he has just gained. By

contrast, Patrick strongly dislikes the notion of divorce in general and so has a strong preference to believe that Brian will not get divorced. Identical to the procedure used in Study 2, participants evaluated each of their beliefs' (i) moral quality, (ii) evidentiary quality, (iii) and overall quality, as well as each person's (iv) moral character, and finally, (v) whether each person "knew" the outcome.

Results

Most accurate estimate judgments. Participants on average thought that Brian (the husband) had slightly different evidence than Patrick (the friend). In the evidence-based belief condition, participants indicated that his evidence suggested a 53% chance ($SD = 21\%$) of divorce, compared to Patrick's 61% ($SD = 16\%$), $t(111) = -4.80, p < .001, 95\% \text{ CI } [-11.6, -4.83]$. In the optimistic belief condition people attributed Brian 53% ($SD = 19\%$) compared to Patrick's 58% ($SD = 18\%$), $t(120) = -2.97, p = .004, 95\% \text{ CI } [-8.26, -1.66]$. Following our pre-registered procedure, we removed all participants who gave estimates for Brian and Patrick that differed by more than 10% (2 points on our 21-point, 0-100% scale). This left 142 participants of our original 233 (61%). Within this sample, participants in the evidence-based belief condition provided "most accurate" estimates for Brian ($M = 59\%, SD = 18\%$) and Patrick ($M = 60\%, SD = 18\%$) that were not significantly different, $t(64) = -1.90, p = .062, 95\% \text{ CI } [-2.53, 0.07]$. Similarly, participants in the optimistic-belief condition gave "most accurate" estimates for Brian ($M = 57\%, SD = 18\%$) and Patrick ($M = 58\%, SD = 17\%$) that were not significantly different, $t(76) = -1.49, p = .141, 95\% \text{ CI } [-2.13, 0.31]$. Thus, based on the evidence, divorce is more likely than not, and therefore a pessimistic belief is more warranted than an optimistic one. We conducted all of the analyses below on this subset of our original sample of participants.

Analysis plan. As in Study 2, we created composite measures of the moral quality of belief, the moral quality of the believer, the evidential quality of belief, and overall belief quality. The two-item pairs for each question passed our preregistered criterion ($r > .50$) across both evidence-based belief and optimistic belief conditions, with two exceptions for which the independent judgments yielded similar results.⁵ We therefore subjected each of these composites, as well as participants' knowledge attributions, to a series of 2 (Belief: Optimistic vs Evidence-based) x 2 (Reason: Moral vs Non-moral reason) fully-crossed ANOVAs. For readability, we will describe differences in ratings across the Reason manipulation as between Brian, the husband (who has a moral obligation to be optimistic), and Patrick, the friend (who lacks a moral obligation to be optimistic).

Analyses replicating Study 2. Because our main analyses replicated Study 2, we report a summary here; a full description of our results is available in the Supplemental Materials. As expected, we observed predicted reason x belief interactions for moral evaluations of the belief and believer, evaluations of the believer's evidence, and overall evaluations of how justified and permissible the belief is ($ps \leq .005$; see Figure 5). When Brian (the husband) and Patrick (the friend) adopted the evidence-based belief, participants judged Brian's belief to be morally worse, less satisfied by the evidence, less justified or permissible, and Brian as morally worse than

⁵ In the optimistic-belief condition, judgments of the belief's permissibility and justification correlated weakly ($r = .29, p < .001$), as did judgments of how loyal and morally good the agent's reason was ($r = .43, p < .001$). Our results are largely the same when analyzing the two items for each judgment individually (see Supplemental Materials for details). However, the two items differed in their pattern of correlation with other items. For instance, the justificatory status of the belief was significantly predicted by morality after accounting for differences in evidential quality ($b = 0.24, SE = 0.09, t = 2.54, p = .011$), but permissibility was only marginally so ($b = 0.17, SE = 0.09, t = 1.86, p = .062$). Likewise, how "morally good" the belief was significantly predicted the overall belief quality after accounting for evidential quality ($b = 0.21, SE = 0.06, t = 3.45, p = .001$), but how "loyal" the belief was did not ($b = 0.06, SE = 0.06, t = 1.08, p = .279$). We present results for the composite measure here for readability.

Patrick ($ps < .05$). By contrast, when Brian and Patrick adopted the optimistic belief, they judged Brian's belief to be morally better, to have more-sufficient evidence, and to be more justified and permissible, and they judged Brian to have better moral character ($ps < .05$). Finally, when regressing overall belief quality on evidential evaluations and moral evaluations, both independently predicted overall belief quality across target and condition ($ps < .001$).

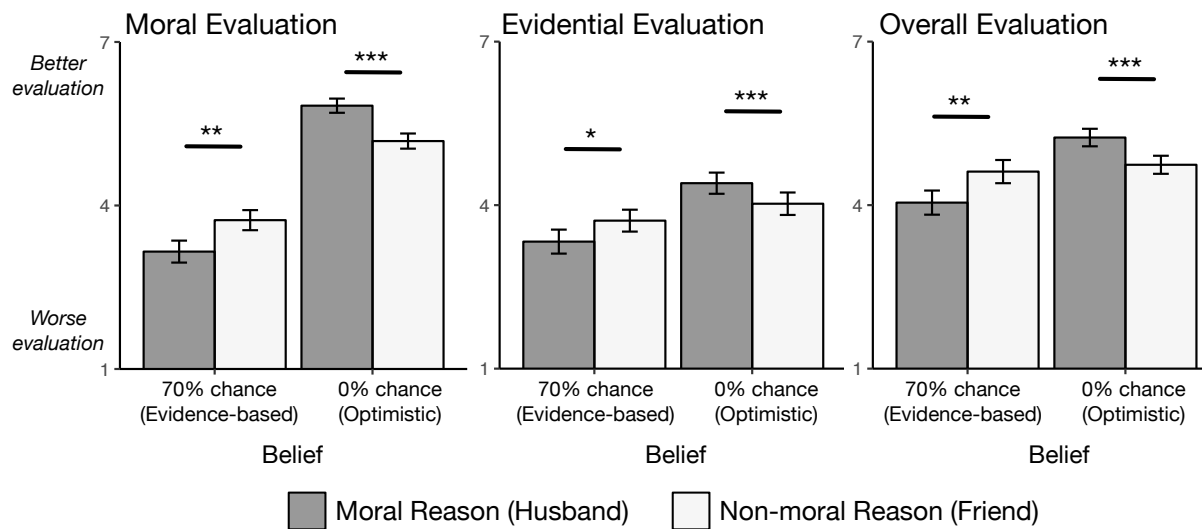


Figure 5. Means and standard error (error bars: ± 1 standard error) across condition for ratings of moral quality, evidential quality, and overall belief quality in Study 3. * $p < .05$, ** $p < .01$, *** $p < .001$

Unlike in Study 2, we did not observe the predicted pattern of results for participants' attributions of knowledge. On average, attributions of knowledge did not differ across the optimistic belief ($M = 3.55$, $SD = 2.03$) and evidence-based belief ($M = 3.32$, $SD = 1.85$) conditions, $F(1, 140) = 0.50$, $p = .482$, though on average participants did attribute more knowledge to the husband ($M = 3.58$, $SD = 1.93$) than to the friend ($M = 3.31$, $SD = 1.96$), $F(1, 140) = 8.42$, $p = .004$, $eff < .01$. We also failed to observe a significant belief x reason interaction, $F(1, 140) = .68$, $p = .41$.

Across-belief comparisons. Comparisons of participants' evaluations across the belief conditions offers a revealing look at how the moral norms operate in this vignette. Consider Brian's friend, Patrick. Based on Patrick's evidence, participants indicated that the most accurate estimate was about a 60% chance of divorce. On average, however, participants reported comparable ratings for the extent to which his evidence was sufficient to believe that there is a 70% chance of divorce ($M = 3.72, SD = 1.62$) as a 0% chance of divorce ($M = 4.03, SD = 1.81$), $t(139.47) = -1.08, p = .282, 95\% \text{ CI } [-0.88, 0.26]$. And likewise, he was judged equally justified in believing that there is a 70% ($M = 4.62, SD = 1.72$) or a 0% ($M = 4.74, SD = 1.45$) chance of divorce, $t(126.13) = -0.46, p = .644, 95\% \text{ CI } [-0.66, 0.41]$. By comparison, even though participants attributed the same evidence to the husband, Brian, evaluations of his beliefs were strongly dictated by the norms governing his relationship. For instance, he was judged to have less sufficient evidence for the belief that he has a 70% chance of divorce ($M = 3.33, SD = 1.77$) than the belief that he has a 0% chance of divorce ($M = 4.40, SD = 1.71$), $t(134.05) = -3.65, p < .001, 95\% \text{ CI } [-1.65, -0.49]$. And likewise, he was seen as more justified in believing he had a 0% chance of divorce ($M = 5.24, SD = 1.41$) than a 70% chance ($M = 4.05, SD = 1.78$), $t(120.99) = -4.36, p < .001, 95\% \text{ CI } [-1.74, -0.65]$. Despite the fact that Brian's evidence favored a pessimistic belief (i.e., a 70% chance of divorce), he was judged as having more sufficient evidence, as well as being more permitted and justified, to adopt the optimistic belief he has objectively less evidence for. Thus, beyond the comparison to Patrick, participants' evaluations of Brian provide independent evidence for the impact of morality on belief.

Exploratory hypotheses. We also examined participants' evaluations of the situational features of the husband's and friend's beliefs. Even though participants reported that the evidence held by the husband and friend was the same, they still reported that the husband had a

stronger obligation to be optimistic ($M = 4.32$, $SD = 1.78$) than the friend did ($M = 2.15$, $SD = 1.76$), $t(141) = 12.31$, $p < .001$, 95% CI [1.82, 2.52]. Similarly, participants reported that the husband's pessimism would have a more deleterious effect on his behavior ($M = 4.39$, $SD = 1.43$) than the friend's pessimism would on his ($M = 3.50$, $SD = 1.73$), $t(141) = 5.88$, $p < .001$, 95% CI [0.59, 1.19], and that the husband's pessimism would more intrinsically wrong ($M = 3.15$, $SD = 1.91$) than the friend's pessimism ($M = 2.25$, $SD = 1.80$), $t(141) = 6.47$, $p < .001$, 95% CI [0.63, 1.18]. Finally, participants judged that the husband had more voluntary control over his belief ($M = 4.47$, $SD = 1.68$) than the friend did ($M = 3.96$, $SD = 1.84$), $t(141) = 2.97$, $p = .003$, 95% CI [0.17, 0.84].

We next calculated husband-minus-friend difference scores for these four belief properties. Perceived obligation to be optimistic correlated with the intrinsic wrongness of pessimism, $r(140) = 0.22$, $p = .01$, 95% CI [0.06, 0.38], as well as the deleterious effect of pessimism on behavior, $r(140) = 0.28$, $p < .001$, 95% CI [0.13, 0.43]. Participants who tended to think that beliefs have a deleterious effect on action also tended to think that, even if those beliefs did not have that effect, the beliefs are intrinsically wrong, $r(140) = 0.40$, $p < .001$, 95% CI [0.25, 0.53]. When we regressed differences in obligation on differences in perceived control, connection to action, and intrinsic wrongness, only differences in the perceived bad effects of pessimism on action uniquely predicted differences in moral obligation, $b = 0.29$, $SE = 0.10$, $t = 2.76$, $p = .006$. Attributions of control did not correlate with obligation, $r(140) = 0.07$, $p = .42$, 95% CI [-0.1, 0.23], the intrinsic value of belief, $r(140) = -0.01$, $p = .88$, 95% CI [-0.18, 0.15], or the effect of belief on behavior, $r(140) = -0.11$, $p = .18$, 95% CI [-0.27, 0.05].

These in-principle judgments predicted participants' evaluations of the husband and friend's formed belief. We calculated husband-minus-friend or friend-minus-husband difference

scores such that greater positive differences indicated participants believing a pessimistic attitude was worse for the husband to hold or that an optimistic attitude was better for the husband to hold. We then estimated the correlation between these ratings and participants' differences in the perceived moral obligation to be optimistic. As expected, differences in obligation correlated with differences in how justified and permissible the husband's belief was relative to the friend's, $r(140) = 0.19, p = .02, 95\% \text{ CI } [0.03, 0.35]$, as well as whether the husband had sufficient evidence for his belief, $r(140) = 0.19, p = .02, 95\% \text{ CI } [0.02, 0.34]$. However, perceived obligation did not significantly correlate with the moral value of the belief, $r(140) = 0.04, p = .63, 95\% \text{ CI } [-0.12, 0.2]$, or with differences in the perceived moral character of the two people, $r(140) = 0.07, p = .44, 95\% \text{ CI } [-0.1, 0.23]$.

Discussion

Study 3 replicated the prescribed motivated reasoning, evidence shifting, and alternative justification effects from Study 2 even after ensuring that participants attributed the same evidence to the socially-close and socially-distant characters. Moreover, we observed that the newlywed husband was seen to have more sufficient evidence for, and more justification to hold, the optimistic belief about his marriage (vs. the pessimistic belief) despite the fact that his evidence numerically favored the pessimistic belief. And, as before, moral evaluations of the belief – in this case judgments that the belief exhibits loyalty – were strongly associated with differences in participants' evaluations of the evidentiary quality of the belief as well as their evaluations of the belief's justifiability.

Study 3 also extended the findings from Study 2 to better understand the role of moral considerations in belief evaluation. Before learning which belief the two individuals adopted, participants were asked to make several "in principle" judgments about each character's belief.

As expected, participants believed that, despite having the same evidence, the husband was more obligated to remain optimistic than the friend. This difference in obligation significantly predicted participants' evaluations later on. Secondly, participants anticipated that having an optimistic belief would help the husband (more-so than the socially distant friend) behave the way he ought to behave, and partly on that basis seemed to say that he was obligated to be optimistic. And lastly, Study 3 ruled out another possible interpretation of our finding; namely, that participants are judging the socially-close believer differently because they attribute to him less control over his belief. The reasoning behind this objection is that someone may be more permitted to hold a non-evidential belief when his particular situation engenders an unintentional and uncontrollable bias. This objection was not borne out by our data. Consistent with prior work showing that people attribute to others a great deal of control over belief (Cusimano & Goodwin, 2019, 2020), participants attributed high control to both believers and, more importantly, more control to the socially-close believer than they did to the socially-distant believer.

General Discussion

Dilemmas about what to believe based on the evidence and based on morality are commonplace. However, we know little about how people evaluate beliefs in these contexts. One line of reasoning, consistent with past work documenting both an objectivity bias (Ross & Ward, 1996), and an aversion to discrepant and inaccurate beliefs in others (e.g., Golman et al., 2016), predicts that people will demand that others set aside moral concerns and form beliefs impartially and solely on the basis of the evidence. However, we hypothesized that people will sometimes reject a normative commitment to evidence-based reasoning and treat moral considerations as a legitimate influence on belief. We then articulated two ways people could integrate moral and

evidential value into their evaluations of belief. First, they could treat moral considerations as shifting the evidential decision criterion for a belief, which we called the “evidential-shift hypothesis.” Or, they could treat a belief’s moral quality as an alternative justification for belief that they weigh against its evidential quality, the “alternative justification” hypothesis. Our studies were capable of detecting whether people treat moral concerns as a legitimate basis for belief, and further whether people prescribe motivated reasoning to others in line with the evidential shift hypothesis or the alternative justification hypothesis.

Across all studies, participants routinely indicated that what a believer ought to believe, or was justified in believing, should be affected by what would be morally good to believe. In Study 1, participants on average reported that what someone ought to believe should be more optimistic (in favor of what is morally good to believe) than what is objectively most accurate for that person to believe based on their evidence. The extent to which participants prescribed these optimistic beliefs was strongly associated with the amount of moral benefit they thought an optimistic belief would confer, as measured by abstract statements such as “All else being equal, it is morally good to give your friend the benefit of the doubt.” In Studies 2 and 3, participants reported that someone who would gain a moral benefit by being optimistic was more justified in adopting that belief against the evidence compared to someone else with the same information but who lacked a moral reason for believing against the evidence (and only adopted the optimistic belief because they had a strong preference to do so). Moreover, when both people adopted the morally undesirable belief, the believer who gave up a moral benefit for doing so was judged as less justified than someone who did not give up a moral benefit. These results provide strong evidence against the idea that people always demand that others form beliefs based on an impartial and objective evaluation of the evidence.

Additionally, participants reported that moral concerns affected the standards of evidence that apply to belief, such that morally-desirable beliefs require less evidence than morally-undesirable beliefs. In Study 1, participants reported that, relative to an impartial observer with the same information, someone with a moral reason to be optimistic had a wider range of beliefs that could be considered “consistent with” and “based on” the evidence. Critically however, the broader range of beliefs that were consistent with the same evidence were only beliefs that were more morally desirable; morally undesirable beliefs were not more consistent with the evidence. In Studies 2 and 3, participants agreed more strongly that someone who had a moral reason to adopt a desirable belief had sufficient evidence to do so compared to someone who lacked a moral reason, even though they formed the same belief on the basis of the same evidence. Likewise, on average, participants judged that when someone adopted the morally undesirable belief, they were more often judged as having insufficient evidence for doing so relative to someone who lacked a moral reason (again, even though they formed the same belief on the basis of the same evidence). Finally, in Study 2 (though not in Study 3), these judgments replicated using an indirect measure of evidentiary quality; namely, attributions of knowledge. In sum, these findings document that one reason people may prescribe a motivated belief to someone is because morality changes how much evidence they consider to be required to hold the belief in an evidentially-sound way.

Finally, our studies were capable of detecting whether or not people thought that moral considerations could justify holding a belief beyond what is supported by the evidence – that is, whether moral reasons constitute an “alternative justification” for belief that can possibly outweigh evidential considerations for belief. In Study 1, we observed morality playing this role in two out of the six vignettes that we examined. When prescribing beliefs to a newlywed who is

trying to judge his chance of divorce, about half of participants who prescribed a motivated belief reported that the newlywed should believe something that was, by their own lights, inconsistent with his evidence. We observed similar findings in the Cancer vignette. Studies 2 and 3 revealed a more subtle way in which moral concerns independently and directly affect belief. Across all the vignettes, the moral quality of the belief – such as how helpful or loyal the belief was deemed to be – predicted how justified participants rated the belief even after accounting for variation in how evidentially justified the belief was judged. Thus, it appears that people will sometimes prescribe a belief to someone knowing that the belief is not supported by the person's evidence, and that they will do so because the belief confers a moral benefit.

Though the focus of the current investigation was to determine whether we can detect morally-prescribed motivated reasoning, it is important to note that we also found substantial evidence that people think beliefs ought to be constrained by the evidence. In Study 1, average ought-to-believe estimates were closely related to what people thought would be most accurate for the believer to believe (Table 2). Indeed, on average, ought-to-believe estimates were still pessimistic (i.e., closer to the evidence) rather than optimistic (i.e., morally preferable) (Figure 3). In Studies 2 and 3, we found a strong association between the perceived evidentiary quality of the belief and judgments of the belief's permissibility and justifiability. Specifically, the more insufficient the believer's evidence, the less justifiable and permissible it was for them to hold the belief. Indeed, even in these relatively constrained experimental contexts, the evidentiary quality of beliefs explained more variance in the belief's justification and permissibility than did the independent contribution of morality. Thus, while our findings show that people will integrate moral considerations into their belief evaluations, they also show that people think others should balance these considerations with the evidence that they have.

Alternative explanations

Our experimental procedures and results allow us to rule out a variety of alternative explanations for our results. Specifically, two alternative explanations stem from the observation that we manipulate moral obligation by changing the social distance between two individuals. Rather than social distance affecting the moral norms that apply to one's belief, as we hypothesize, it could instead be the case that being close to the person you are forming beliefs about either (1) makes your beliefs more likely to be self-fulfilling, or (2), creates a reason to be more diligent in one's reasoning, and therefore more withholding of beliefs in general. We address each of these two alternatives below.

Self-fulfilling beliefs. Self-fulfilling beliefs are beliefs that, by virtue of being adopted, are more likely to come true. If people attribute to believers the knowledge that certain beliefs are self-fulfilling, then they may infer that someone who has adopted a self-fulfilling belief will have more evidence in favor of that belief compared to someone who has adopted a similar, but non-self-fulfilling belief. For instance, if the newlywed husband in Study 3 makes the optimistic outcome more likely by adopting an optimistic belief, then he will have slightly better evidence for his optimistic belief compared to his friend (whose belief is not self-fulfilling). If participants evaluate beliefs in this way, then it is possible that some of the cases in which people are putatively endorsing non-evidential grounds to prescribe or permit belief are really cases in which participants are inferring the presence of additional evidence stemming from the adoption of the self-fulfilling belief.

Several findings from the studies above speak against this proposal. First, in the marriage vignette specifically, this alternative explanation would predict that – because the husband's beliefs about his marriage are self-fulfilling – people will attribute to the husband more evidence

for these beliefs (no matter what they are) than they will to the friend. However, as we observed in Studies 2 and 3, when the husband adopts the pessimistic belief (that there is a 70% chance of divorce), participants judge his belief as worse than the friend who adopts the same belief. Yet, if he gained more evidence for this outcome by adopting this belief then participants should have judged him more positively. We also documented prescribed motivated reasoning in cases where the belief could not be self-fulfilling: All of our findings in Studies 1 and 2 replicated in the Friend vignette, for which the relevant belief (i.e., whether the friend is guilty or innocent) is about something that occurred in the past. On this basis, the possible self-fulfilling nature of belief is insufficient to explain the findings we report.

Other norms of belief. Norms of due diligence could explain, in principle, why two people with the same evidence should hold different beliefs. For example, if there is nothing in your car, then you may be justified to assume that you left the windows down based on knowing that you usually do. But if you left a child in the car, then you have a reason to double check before deciding whether you did or did not. Prior work shows that people believe one's diligence in belief formation should vary according to the risk imposed by a false belief (McAllister et al., 1979; Pinollos, 2012), and people in these situations actually do engage in more thorough reasoning when the risks are high (Newell & Bröder, 2008; Fiske & Neuberg, 1990; Kunda, 1990; Mayseless & Kruglanski, 1987; Payne, Bettman, & Johnson, 1993). Thus, perhaps the main characters in Study 1 have a wider range of beliefs consistent with the evidence than the AI because they need to be more diligent in their reasoning than the AI, and therefore require more evidence before becoming too confident in any specific belief. In Studies 2 and 3, perhaps participants believed that people should reason more diligently about their loved ones than about distant others, and this is why believers were evaluated negatively for adopting pessimistic

beliefs. That is, perhaps changes in social distance affect how diligent one must be while reasoning, rather than affecting whether one ought to reason in a motivated or biased way.

However, norms of diligence also fail to explain the pattern of findings we observed. In Study 1, a diligence account would predict that a wider range of beliefs in general would be consistent with the evidence, such that a wider range of morally-undesirable beliefs would also be permitted for the characters but not the AI. However, we observed this only for more morally-desirable beliefs, not for morally-undesirable beliefs. Similarly, in Studies 2 and 3, a diligence account would predict that, based on the same amount of information, the socially-close actor would be less justified to adopt *any* belief, whether optimistic or pessimistic. Yet, the statistical interaction we observe rules this out: Whereas socially-close observers were judged poorly for adopting the morally-undesirable (but evidentially-better) belief, these differences were either attenuated or reversed for the morally-desirable, optimistic belief. Thus, our data suggest that being a person rather than an AI, or having a close relationship as opposed to a distal one, does not impose a demand to be careful and accurate in your beliefs, but rather a demand to be partial.

Implications

Psychologists have long speculated that commonplace deviations from rational judgments and decisions could reflect commitments to different normative standards for decision making (rather than merely cognitive limitations or unintentional errors; e.g., Cohen, 1981; Koehler, 1996; Tribe, 1971). This speculation has been largely confirmed in the domain of decision making, where work has documented that people will refuse to make decisions they know to be accurate because of a normative commitment to not rely on certain kinds of evidence (Nesson, 1985; Wells, 1992), or refuse to make utility-maximizing decisions based on deontological moral considerations (Baron & Spranca, 1997; Tetlock et al., 2000). And yet, there

has been comparatively little investigation in the domain of belief formation. While some work has suggested that people evaluate beliefs in ways that favor non-objective, or non-evidential criteria (e.g., Armor et al, 2008; Gao et al., 2019; Tenney et al., 2015), this work has failed to demonstrate that people prescribe beliefs that violate what objective, evidence-based reasoning would warrant. To our knowledge, our results are the first to demonstrate that people will knowingly endorse non-evidential norms for belief, and specifically, prescribe motivated reasoning.

Our results therefore warrant a fresh look at old explanations for irrationality. Most relevant are overconfidence or optimism biases documented in the domain of close relationships (e.g., Baker & Emery, 1993; Srivastava et al., 2006) and health (e.g., Thompson, Sobolew-Shubin, Galbraith, Schwankovsky, & Cruzen, 1993). Past work has suggested that the ultimate explanation for motivated reasoning could derive from the downstream benefits for the believers (Baumeister, 1989; Murray & Holmes, 1997; Taylor & Brown, 1988; Srivastava et al., 2006; but see Neff & Geers, 2013, and Tenney et al, 2015). Our findings suggest more proximate explanations for these biases: That lay people see these beliefs as morally beneficial and treat these moral benefits as legitimate grounds for belief. Thus, overconfidence or over-optimism may persist in communities because people hold others to lower standards of evidence for adopting morally-beneficial optimistic beliefs than they do for pessimistic beliefs, or otherwise treat these benefits as legitimate reasons to ignore the evidence one has (c.f., Metz, Weisberg, & Weisberg, 2018).

Our results also suggest a novel explanation for how moral norms for belief could facilitate motivated reasoning. Specifically, people could acknowledge that one of their beliefs is supported by less total evidence compared to their other beliefs, but judge that the belief

nevertheless satisfies the demand for sufficient evidence just as well because the standards for evidence are lower in light of the belief's moral quality. As a result, they may not judge it necessary to pursue further evidence, or to revise their belief in light of modest counter-evidence. As an example, people could recognize that a belief in God, or a belief that all things happen for a reason, is supported by little objective evidence, but at the same time believe that the little evidence they have nevertheless constitutes sufficient evidence in light of the moral benefit that the belief confers (see McPhetres & Zuckerman, 2017, and Lobato et al., 2019, for some preliminary findings consistent with this proposal; see James, 1937, and Pace, 2011, for fuller discussion of how to evaluate evidence for moral beliefs). Although this is speculative, it naturally follows from the findings presented here and presents a valuable direction for future research.

Limitations and future directions

Though we have demonstrated that people prescribe motivated reasoning to others under some conditions, we have not offered a comprehensive treatment of the conditions under which this occurs. Indeed, we did not observe prescribed motivated reasoning in the Bully vignette in Study 2, or the Sex vignette in Study 1, despite their similarity to the other vignettes. One straightforward explanation for this is because the relevant moral norms in those vignettes did not outweigh norms to be accurate. Indeed, participants reported the least moral concern for the Sex vignette.⁶ And, in the Bully vignette, while we observed that participants thought that the teacher could do good for the student by adopting an optimistic belief, there may have been

⁶ The idea that people do not associate moral anxiety with making inferences about others' social status on the basis of sex is consistent with prior findings that people often do exactly that (e.g., Cao et al., 2019).

competing moral considerations at play. For instance, participants may have believed that the teacher had a moral responsibility to be clear-eyed about the bully in order to protect the other students. Lending some support to this explanation, we observed that the evidence in favor of the pessimistic belief (that the child was in fact a bully and would misbehave again) was stronger compared to all the other vignettes. Because the evidence was less ambiguous, other competing reasons to be accurate may supplant demands to weigh the evidence preferentially or adopt knowingly delusional beliefs. This explanation is speculative and so it is possible that there are factors beyond the strength of competing moral considerations that moderate the effect of prescribed motivated reasoning. In light of the clear normative value that people give moral considerations, a valuable goal for future research will be to delineate when these values demand motivated reasoning and when they do not.

Our findings show strong support for our claim that morality can justify motivated reasoning, therein raising the question of whether moral value is the only kind of non-evidential consideration that people treat as relevant to belief. Specifically, it raises the question of whether people think others should adopt beliefs that are merely useful (but not morally good). We found that moral considerations were considered a better justification for motivated reasoning than was satisfying a strong preference (Studies 2-3), but these studies do not definitely rule out the possibility that a large benefit could also justify belief. Some philosophers have famously argued in favor of this possibility, as when Pascal (1852) concluded that, despite a paucity of evidence, he ought to believe that God exists or else risk incalculable suffering after death. Whether people judge that these kinds of benefits can justify motivated belief warrants further investigation.

A final limitation stems from our focus on average trends within a relatively homogenous sample. Even within our sample, we found variation in whether participants prescribed motivated

reasoning, and in whether they exhibited the hallmarks of an evidential shift or alternative justification. In light of this, as well as other work demonstrating that people vary in the norms they hold for belief (Ståhl et al., 2016), an important direction for future research should be documenting the sources and consequences of this variation.

Conclusion

People often engage in motivated reasoning in situations where there are good moral reasons to adopt a belief that is not supported by their evidence. Do people think that they and others are believing poorly in these situations, or do they believe that moral reasons are legitimate grounds for belief? The present findings show strong support for the latter. Across three studies, many people prescribed motivated reasoning to others, reported that morally good beliefs require less evidence to be justified, and that, in some circumstances, a morally good belief can be justified even in the absence of sufficient evidence. These results overturn a long-standing assumption that people believe others ought to be impartial, and base their beliefs on the evidence, under all circumstances. Lastly, they suggest another reason that motivated beliefs emerge and persist: People think that they ought to.

References

- Armor, D. A., Massey, C., & Sackett, A. M. (2008) Prescribed optimism: Is it right to be wrong about the future? *Psychological Science*, 19, 329-331. doi: 10.1111/j.1467-9280.2008.02089.x
- Austen, J. (2001). *Pride and prejudice*. Broadview Press.
- Baker, L. A., & Emery, R. E. (1993). When every relationship is above average: Perceptions and expectations of divorce at the time of marriage. *Law and Human Behavior*, 17(4), 439-450. doi:10.1007/bf01044377
- Bandura, A., & Locke, E. A. (2003). Negative self-efficacy and goal effects revisited. *Journal of Applied Psychology*, 88(1), 87-99. doi:10.1037/0021-9010.88.1.87
- Baron, J. (2019). Actively open-minded thinking in politics. *Cognition*, 188, 8-18. doi:10.1016/j.cognition.2018.10.004
- Baron, J., & Spranca, M. (1997). Protected values. *Organizational Behavior and Human Decision Processes*, 70, 1-16. doi:10.1006/obhd.1997.2690
- Basu, R. (2019). Radical moral encroachment: The moral stakes of racist beliefs. *Philosophical Issues*, 29(1), 9-23. doi:10.1111/phis.12137
- Baumeister, R. F. (1989). The optimal margin of illusion. *Journal of Social and Clinical Psychology*, 8(2), 176-189. doi:10.1521/jscp.1989.8.2.176
- Bolinger, R. J. (2018). The rational impermissibility of accepting (some) racial generalizations. *Synthese*, 1–17. doi: 10.1007/s11229-018-1809-5
- Bolinger, R. J. (2020) The varieties of moral encroachment. *Philosophical Perspectives*. doi:10.1111/phpe.12124

- Cao, J., Kleiman-Weiner, M., & Banaji, M.R. (2019). People make the same Bayesian judgment they criticize in others. *Psychological Science*, 30, 20-31.
doi:10.1177/0956797618805750
- Carver, C. S., & Scheier, M. F. (2014). Dispositional optimism. *Trends in Cognitive Science*, 18(6), 293-299. doi:10.1016/j.tics.2014.02.003
- Clifford, W. K. (1877). The ethics of belief. In T. Madigan (Ed.), *The ethics of belief and other essays*. Amherst, MA: Prometheus.
- Cohen, L. J. (1981). Can human irrationality be experimentally demonstrated. *Behavioral and Brain Sciences*, 4(3), 317-331. doi:10.1017/s0140525x00009092
- Cohen, A. B., & Rozin, P. (2001). Religion and the morality of mentality. *Journal of Personality and Social Psychology*, 81(4), 697-710. doi:10.1037/0022-3514.81.4.697
- Cusimano, C., & Goodwin, G.P. (2019). Lay beliefs about the controllability of everyday mental states. *Journal of Experimental Psychology: General*, 148(10), 1701-1732.
doi:10.1037/xge0000547
- Cusimano, C., & Goodwin, G.P. (2020). People believe that others have more voluntary control over beliefs than they themselves do. *Journal of Personality and Social Psychology*.
doi:10.1037/pspa0000198
- de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a web browser. *Behavior Research Methods*, 47(1), 1-12. doi:10.3758/s13428-014-0458-y
- Ditto, P. H., & Lopez, D. F. (1992). Motivated skepticism: Use of differential decision criteria for preferred and nonpreferred conclusions. *Journal of Personality and Social Psychology*, 63, 568-584. doi: 10.1037/0022-3514.63.4.568

- Douglas, H. (2016). Values in Science. In Humphreys, P (Ed.), *The Oxford Handbook of Philosophy of Science* (pp. 609–630), New York: Oxford University Press.
- Epley, N., & Gilovich, T. (2016). The mechanics of motivated reasoning. *Journal of Economic Perspectives*, 30(3), 133-140. doi:10.1257/jep.30.3.133
- Ehrlinger, J., Gilovich, T., & Ross, L. (2005). Peering into the bias blind spot: People's assessments of bias in themselves and others. *Personality and Social Psychology Bulletin*, 31(5), 680-692. doi:10.1177/0146167204271570
- Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. In *Advances in Experimental Social Psychology* (pp. 1-74). Elsevier. doi:10.1016/s0065-2601(08)60317-2
- Foley, R. (1987). *The theory of epistemic rationality*. Cambridge, MA: Harvard University Press.
- Gardiner, G. (2018). Evidentialism and moral encroachment. In *Believing in Accordance with the Evidence* (pp. 169-195). Cham: Springer International Publishing. doi:10.1007/978-3-319-95993-1_11
- Gendler, T. S. (2011). On the epistemic costs of implicit bias. *Philosophical Studies*, 156(1), 33-63. doi:10.1007/s11098-011-9801-7
- Gilovich, T. (1991). *How we know what isn't so: The fallibility of human reason in everyday life*. New York, NY: Free Press.
- Golman, R., Loewenstein G., Moene, K.O., & Zarri, L. (2016). The preference for belief consonance. *Journal of Economic Perspectives* 30, 165-187. doi:10.1257/jep.30.3.165

- Goodwin, G. P., Piazza, J., & Rozin, P. (2014). Moral character predominates in person perception and evaluation. *Journal of Personality and Social Psychology*, 106, 148–168. doi: 10.1037/a0034726
- Hansen, K., Gerbasi, M., Todorov, A., Kruse, E., & Pronin, E. (2014). People claim objectivity after knowingly using biased strategies. *Personality and Social Psychology Bulletin*, 40(6), 691-699. doi:10.1177/0146167214523476
- Haran, U., Ritov, I., & Mellers, B. A. (2013). The role of actively open-minded thinking in information acquisition, accuracy, and calibration. *Judgment and Decision Making*, 8(3), 188-201.
- Holden, G. (1991). The relationship of self-efficacy appraisals to subsequent health related outcomes: A meta-analysis. *Social Work and Health Care*, 16(1), 53-93. doi:10.1300/j010v16n01_05
- James, W. (1937). *The will to believe, and other essays in popular philosophy*. Longmans, Green and Co.
- Keller, S. (2004). Friendship and belief. *Philosophical Papers*, 33(3), 329-351. doi:10.1080/05568640409485146
- Kennedy, K. A., & Pronin, E. (2008). When disagreement gets ugly: Perceptions of bias and the escalation of conflict. *Personality and Social Psychology Bulletin*, 34, 833-848. doi: 10.1177/0146167208315158
- Kim, B., & McGrath, M. (Eds.). (2019). *Pragmatic encroachment in epistemology*. Routledge.
- Koehler, J. J. (1996). The base rate fallacy reconsidered: Descriptive, normative, and methodological challenges. *Behavioral and Brain Sciences*, 19(1), 1-17. doi:10.1017/s0140525x00041157

- Kruglanski, A. W. (2004). *The psychology of closed mindedness*. Psychology Press.
- Kruglanski, A. W., & Webster, D. M. (1996). Motivated closing of the mind: “seizing” and “freezing.”. *Psychological Review*, 103, 263-283. doi: 10.1037/0033-295X.103.2.263
- Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, 108, 480–498. doi: 10.1037/0033-2909.108.3.480
- Lee, J., & Holyoak, K. J. (2020). “But he’s my brother”: The impact of family obligation on moral judgments and decisions. *Memory & Cognition*, 48(1), 158-170. doi:10.3758/s13421-019-00969-7
- Locke, J. (1690). *An essay concerning human understanding*. Oxford: Clarendon.
- Lobato, E. J. C., Tabatabaeian, S., Fleming, M., Sulzmann, S., & Holbrook, C. (2019). Religiosity predicts evidentiary standards. *Social Psychological and Personality Science*, 194855061986961. doi:10.1177/1948550619869613
- Mayseless, O., & Kruglanski, A. W. (1987). What makes you so sure? Effects of epistemic motivations on judgmental confidence. *Organizational Behavior and Human Decision Processes*, 39, 162-183. doi:10.1016/0749-5978(87)90036-7
- McAllister, D. W., Mitchell, T. R., & Beach, L. R. (1979). Contingency-model for the selection of decision strategies: Empirical-test of the effects of significance, accountability, and reversibility. *Organizational Behavior and Human Performance*, 24(2), 228–244. doi:10.1016/0030-5073(79)90027-8
- McPhetres, J., & Zuckerman, M. (2017). Religious people endorse different standards of evidence when evaluating religious versus scientific claims. *Social Psychological and Personality Science*, 8(7), 836-842. doi:10.1177/1948550617691098

- McManus, R. M., Kleiman-Weiner, M., & Young, L. (in press). What we owe to family: The impact of special obligations on moral judgment. *Psychological Science*.
- Metz, S. E., Weisberg, D. S., & Weisberg, M. (2018). Non-scientific criteria for belief sustain counter-scientific beliefs. *Cognitive Science*, 42, 1477-1503. doi:10.1111/cogs.12584
- Milkman, K. L., Chugh, D., & Bazerman, M. H. (2009). How can decision making be improved? *Perspectives on Psychological Science*, 4, 379-383. doi:10.1111/j.1745-6924.2009.01142.x
- Molnar, A., & Loewenstein, G. F. The Othello effect: People are more disturbed by others' wrong beliefs than by different beliefs. *SSRN Electronic Journal*. doi:10.2139/ssrn.3524651
- Moritz, S. E., Feltz, D. L., Fahrbach, K. R., & Mack, D. E. (2000). The relation of self-efficacy measures to sport performance: A meta-analytic review. *Research Quarterly for Exercise and Sport*, 71(3), 280-294. doi:10.1080/02701367.2000.10608908
- Moss, S. (2018). Moral encroachment. *Proceedings of the Aristotelian Society*, 118(2), 177-205. doi:10.1093/arisoc/aoy007
- Multon, K. D., Brown, S. D., & Lent, R. W. (1991). Relation of self-efficacy beliefs to academic outcomes: A meta-analytic investigation. *Journal of Counseling Psychology*, 38(1), 30-38. doi:10.1037/0022-0167.38.1.30
- Murray, S. L., & Holmes, J. G. (1997). A leap of faith? Positive illusions in romantic relationships. *Personality and Social Psychology Bulletin*, 23(6), 586-604. doi:10.1177/0146167297236003

- Neff, L. A., & Geers, A. L. (2013). Optimistic expectations in early marriage: A resource or vulnerability for adaptive relationship functioning. *Journal of Personality and Social Psychology*, 105(1), 38-60. doi:10.1037/a0032600
- Nesson, C. (1985). Evidence or the event on judicial proof and the acceptability of verdicts. *Harvard Law Review*, 98(7), 1357-1392.
- Newell, B., & Bröder, A. (2008). Cognitive processes, models and metaphors in decision research. *Judgment and Decision Making*, 3(3), 195.
- Pace, M. (2011). The epistemic value of moral considerations: Justification, moral encroachment, and James “Will To Believe”. *Noûs*, 45, 239-268. doi:10.2307/41330856
- Pascal, B. (1852). *Pensées*. Dezobry et E. Magdeleine.
- Payne, J. W. , Bettman, J. R. , & Johnson, E. (1993). *The adaptive decisionmaker*. Cambridge, England: Cambridge University Press.
- Pennycook, G., Cheyne, J. A., Koehler, D. J., & Fugelsang, J. A. (2019) On the belief that beliefs should change according to evidence: Implications for conspiratorial, moral, paranormal, political, religious, and science beliefs. *Preprint*.
- Pinillos, Á. (2012) Knowledge, experiments and practical interests. In *Knowledge Ascriptions*, Jessica Brown & Mikkel Gerken (eds), Oxford: Oxford University Press, pp. 192–219. doi:10.1093/acprof:oso/9780199693702.003.0009
- Pronin, E., Gilovich, T., & Ross, L. (2004). Objectivity in the eye of the beholder: Divergent perceptions of bias in self versus others. *Psychological review*, 111(3), 781-799. doi:10.1037/0033-295X.111.3.781

- Pronin, E., Lin, D. Y., & Ross, L. (2002). The bias blind spot: Perceptions of bias in self versus others. *Personality and Social Psychology Bulletin*, 28(3), 369-381.
doi:10.1177/0146167202286008
- Ross, L., & Ward, A. (1996). Naive realism in everyday life: Implications for social conflict and misunderstanding. In E. S. Reed, E. Turiel, & T. Brown (Eds.), *The Jean Piaget symposium series. Values and knowledge*. US: Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Russell, B. (2013). *An Inquiry into Meaning and Truth*. Routledge. doi:10.4324/9780203714331
(Original work published 1948)
- Schwardmann, P., & Van Der Weele, J. (2019). Deception and self-deception. *Nature Human Behavior*, 3(10), 1055-1061. doi:10.1038/s41562-019-0666-7
- Srivastava, S., McGonigal, K. M., Richards, J. M., Butler, E. A., & Gross, J. J. (2006). Optimism in close relationships: How seeing things in a positive light makes them so. *Journal of Personality and Social Psychology*, 91(1), 143-153. doi:10.1037/0022-3514.91.1.143
- Shaw, A., Descioli, P., Barakzai, A., & Kurzban, R. (2017). Whoever is not with me is against me: The costs of neutrality among friends. *Journal of Experimental Social Psychology*, 71, 96-104. doi:10.1016/j.jesp.2017.03.002
- Shaw, A., Choshen-Hillel, S., & Caruso, E. M. (2018). Being biased against friends to appear unbiased. *Journal of Experimental Social Psychology*, 78, 104-115.
doi:10.1016/j.jesp.2018.05.009
- Ståhl, T., Zaal, M. P., & Skitka, L. J. (2016). Moralized rationality: Relying on logic and evidence in the formation and evaluation of belief can be seen as a moral issue. *PLOS ONE*, 11, e0166332.

- Stanovich, K. E., & West, R. F. (1998). Individual differences in rational thought. *Journal of Experimental Psychology: General*, 127(2), 161-188. doi:10.1037/0096-3445.127.2.161
- Starmans, C., & Friedman, O. (2012). The folk conception of knowledge. *Cognition*, 124(3), 272-283.
- Stroud, S. (2006). Epistemic partiality in friendship. *Ethics*, 116, 498-524. doi:10.1086/500337
- Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin*, 103(2), 193-210. doi:10.1037/0033-2909.103.2.193
- Tenney, E.R., Logg, J.M., & Moore, D.A. (2015) (Too) optimistic about optimism: The belief that optimism improves performance. *Journal of Personality and Social Psychology*, 108, 377-399. doi:10.1037/pspa0000018
- Tetlock, P. E., Kristel, O. V., Beth, S., Green, M. C., & Lerner, J. S. (2000). The psychology of the unthinkable: Taboo trade-offs, forbidden base rates, and heretical counterfactuals. *Journal of Personality and Social Psychology*, 78, 853-870. doi:10.1037/0022-3514.78.5.853
- Thompson, S. C., Sobolew-Shubin, A., Galbraith, M. E., Schwankovsky, L., & Cruzen, D. (1993). Maintaining perceptions of control: Finding perceived control in low-control circumstances. *Journal of Personality and Social Psychology*, 64(2), 293-304. doi:10.1037/0022-3514.64.2.293
- Tribe, L. H. (1971). Trial by mathematics: Precision and ritual in the legal process. *Harvard Law Review*, 84(6), 1329-1393. doi:10.2307/1339610

- Wegener, D. T., Silva, P., Petty, R. E., & Garcia-Marques, T. (2012). The metacognition of bias regulation. In P. Brinol & K. DeMarree (Eds.), *Social Metacognition*. New York: Psychology Press.
- Wells, G. L. (1992). Naked statistical evidence of liability: Is subjective probability enough. *Journal of Personality and Social Psychology*, 62(5), 739-752. doi:10.1037/0022-3514.62.5.739
- West, R. F., Meserve, R. J., & Stanovich, K. E. (2012). Cognitive sophistication does not attenuate the bias blind spot. *Journal of Personality and Social Psychology*, 103(3), 506-519. doi:10.1037/a0028857

Appendix A

Items measuring belief's moral and other qualities in Study 1.

The first two items measure moral concern. The second two items were distractor items measuring how personally beneficial an optimistic belief would be.

Bully

Being a good teacher requires treating your students as individuals, and not making assumptions about them based on their background.

All else being equal, it is morally good for teachers not to assume that a child will be a poor student.

It would make Moira feel good to believe that the Chad will behave well tomorrow.

Moira will personally benefit if she believes that Chad will behave well tomorrow.

Cancer

Being helpful and supportive requires staying optimistic even in the face of bad news.

All else being equal, it is morally good to stay optimistic when helping friends and family through hard times.

It would make Amy feel good to believe that Jordan will recover.

Amy will personally benefit if she believes that Jordan will recover.

Friend

Being a loyal friend to someone requires giving that person the benefit of the doubt, even when doing so requires questioning what other people have said about them.

All else being equal, it is morally good to give your friend the benefit of the doubt.

It would make Adam feel good to believe that John is innocent.

Adam will personally benefit if he gives John the benefit of the doubt.

Marriage

Being a good spouse requires treating your relationship as unique, and not making assumptions about it or your partner based on others.

All else being equal, it is morally good for spouses to assume that their marriage will not end in divorce.

It would make Brian feel good to believe that he and Maya will never divorce.

Brian will personally benefit if he believes that he and Maya will never divorce.

Race

Treating others with respect means treating them as individuals, and not making assumptions about them based on their race.

All else being equal, it is morally good not to assume that a black man is dangerous.

It would make Dara feel good to believe that Michael is not dangerous.

Dara will personally benefit if she believes that Michael is not dangerous.

Sex

Treating others with respect requires treating them as individuals, and not making assumptions about them based on their sex/gender.

All else being equal, it is morally good not to assume that a woman has a low-status job instead of a high-status job.

It would make Brian feel good to believe that the approaching woman is Dr. Adem.

Brian will personally benefit if he believes that the approaching woman is Dr. Adem.

Items measuring the belief's moral quality in Study 2.

Bully

How morally good is *Moira/Benny's* reasoning about whether Chad will be well-behaved?

How helpful/constructive is *Moira/Benny's* belief that Chad will be well-behaved tomorrow?

Cancer

How morally good is *Amy/the oncologist's* reasoning about whether Jordan will recover?

How loyal to Jordan is *Amy/the oncologist's* belief that Jordan will recover?

Friend

How morally good is *Jasmine/Adam's* reasoning about whether John is guilty?

How loyal is *Jasmine/Adam's* belief that John is guilty?

Marriage

How morally good is *Brian/Patrick's* reasoning about whether *he/Brian* and Maya will get divorced?

How loyal to Maya is *Brian/Patrick's* belief that there is no chance that *he/Brian* and Maya will get divorced?

Items measuring the belief's moral quality in Study 3.

How morally good is *Brian/Patrick's* reasoning about whether *he/Brian* and Maya will get divorced?

How loyal to Maya is *Brian/Patrick's* belief that there is no chance that *he/Brian* and Maya will get divorced?

Appendix B

Table B1

Means (and standard deviation) for all dependent measures across the four vignettes in Study 2.

Vignette	Dependent Measure	Evidence-based belief		Optimistic belief		Predicted interaction
		<i>Moral Reason</i>	<i>Non-moral reason</i>	<i>Moral Reason</i>	<i>Non-moral reason</i>	
Bully	Overall Belief Evaluation	5.76 (1.11)	5.53 (1.44)	4.41 (1.47)	3.94 (1.50)	
	Moral Character	5.31 (1.15)	5.18 (1.16)	5.86 (0.93)	5.29 (1.17)	*
	Sufficient Evidence	5.15 (1.32)	4.97 (1.55)	3.30 (1.44)	3.08 (1.47)	
	Knowledge	4.93 (1.27)	4.64 (1.42)	2.56 (1.84)	2.53 (1.89) ^b	
	Moral Quality of Belief	4.06 (1.35)	4.09 (1.24)	5.10 (1.26)	4.37 (1.45) ^b	*
Cancer	Overall Belief Evaluation	5.50 (1.25)	5.49 (1.34)	5.45 (1.24)	4.57 (1.32) ^b	*
	Moral Character	5.22 (1.41)	5.06 (1.28)	6.07 (0.94)	5.42 (1.15) ^b	*
	Sufficient Evidence	5.03 (1.31)	5.12 (1.47)	4.34 (1.31)	3.66 (1.49) ^b	*
	Knowledge	4.75 (1.32)	4.92 (1.24)	3.39 (2.09)	2.77 (1.91) ^b	*
	Moral Quality of Belief	4.41 (1.53)	4.34 (1.35)	6.15 (0.90)	5.29 (1.22) ^b	*
Friend	Overall Belief Evaluation	5.27 (1.40)	5.36 (1.41)	4.76 (1.29)	4.23 (1.49) ^b	*
	Moral Character	4.85 (1.34)	4.83 (1.20)	5.55 (1.07)	5.14 (0.92) ^b	*
	Sufficient Evidence	4.28 (1.40)	4.29 (1.49)	3.24 (1.37)	3.00 (1.43) ^b	*
	Knowledge	3.80 (1.88)	3.82 (1.86)	3.53 (1.98)	2.72 (1.79) ^b	*
	Moral Quality of Belief	3.99 (1.29)	4.27 (1.23) ^a	5.62 (0.87)	4.64 (1.17) ^b	*
Marriage	Overall Belief Evaluation	3.65 (1.73)	4.82 (1.78) ^a	5.53 (1.38)	4.65 (1.62) ^b	*
	Moral Character	4.27 (1.55)	4.82 (1.22) ^a	5.94 (1.06)	5.53 (1.27) ^b	*
	Sufficient Evidence	2.98 (1.72)	3.70 (1.76) ^a	4.72 (1.65)	3.81 (1.89) ^b	*
	Knowledge	3.38 (1.91)	3.18 (1.95) ^a	4.84 (1.45)	4.01 (1.87) ^b	*
	Moral Quality of Belief	3.03 (1.51)	3.89 (1.40) ^a	6.21 (1.06)	5.25 (1.40) ^b	*

Notes. Participants made judgments on 1-7 rating scales; higher numbers mean more positive judgments.

(a) Participants provided more negative evaluation of the moral reason character than the non-moral reason character when they adopted the evidence-based belief ($p < .05$).

(b) Participants provided a more positive evaluation of the moral reason character than the non-moral reason character when they adopted the optimistic, desirable belief ($p < .05$).

Appendix C

Study 3 Vignette Text:

Brian and Maya are high school sweethearts. They met on the first day of class their first year of high school. They were both interested in musical theater and started volunteering and performing in school plays together. A year later, they revealed to each other that they had feelings for one another and started to date. They dated all throughout high school and then started college together. In their second year of college, though they were each only 19-years-old, Brian proposed to Maya and she said yes. They soon after went to the courthouse with a close friend of theirs and got married. During the ceremony, Brian promised to love, trust, and be faithful to Maya for the rest of his life.

About a month after the marriage, Brian stumbles across a recent scientific study on marriage. The study presents statistics about which marriages tend to last the longest and which ones quickly end in divorce. When Brian finds the data for his and Maya's age and relationship length, he reads that 70% of couples similar to him and Maya divorce within 5 years. Brian had never seen this statistic before. In fact, he had assumed that the divorce rate was much lower for couples his age. Brian is shaken. He thinks that his relationship with Maya is very good. But, based on the study he read, he also knows that just about everyone his age who got married thought the same thing about their own relationships.

Brian decides to talk about this with his friend, Patrick, who has known the two of them for as long as they have been together. Patrick asks Brian to list out all the things about his relationship with Maya that are good, and all the things that are bad. Brian does so, and goes on to tell Patrick about how he feels about Maya, what he likes and dislikes, what his gut tells him, and everything he hopes will come to pass. Finally, Brian tells Patrick about the study he just read, which Patrick also had not known about. In return, Patrick tells Brian everything he has observed about their relationship - the positive and the negative - from an outside perspective.

Brian and Patrick talk for a long time. Eventually, Patrick has all the same relevant information that Brian has about Brian and Maya's relationship. Brian cannot think of anything he personally knows that Patrick does not also now know, and vice versa. It seems like Brian and Maya's relationship is very similar to other couples who have recently married: they are deeply in love, agree on many things, but still have occasional fights over trivial things. Unlike many other recently married couples, however, they are very young and they have never dated anyone else. Brian and Patrick agree that Brian and Maya may grow and change in unexpected ways over the next few years.

Assume that Brian has shared everything he can think of with Patrick, and that Patrick has shared everything that he can think of too, so that they now have the exact same relevant information.

Study 3 Belief x Reason Manipulation Text:

Table C1

Belief adoption text in each condition in Study 3.

Social Distance Condition	Belief Condition	
	Optimistic Belief	Evidence-based belief
Socially Close (Moral reason to maintain optimism)	Brian says to Patrick that he has an obligation to believe that he and Maya will never divorce. After all, he made a vow to love and support her forever. If he doubts that will happen then it will be like he did not make that promise. Moreover, he knows that thinking positively will help him be a better husband.	Brian says to Patrick that he has an obligation to believe that he and Maya will never divorce. After all, he made a vow to love and support her forever. If he doubts that will happen then it will be like he did not make that promise. Moreover, he knows that thinking positively will help him be a better husband.
	Brian thinks about everything that he and Patrick discussed and thinks about his duty as Maya's husband. In the end, based on his duty to Maya, Brian decides to believe that within the next five years there is <i>no chance</i> that he and Maya will get divorced.	Brian thinks about everything that he and Patrick discussed and thinks about his duty as Maya's husband. In the end, Brian sets aside his obligation to Maya and decides to believe that within the next five years there is <i>a 70% chance</i> that he and Maya will get divorced.
	Although Patrick knows Brian very well, he has no obligations toward him or Maya. However, thinking about <i>anyone</i> getting divorced makes him sad, so he has a strong preference to believe that no one will ever get divorced.	Although Patrick knows Brian very well, he has no obligations toward him or Maya. However, thinking about <i>anyone</i> getting divorced makes him sad, so he has a strong preference to believe that no one will ever get divorced.
Socially distant (Non-moral reason to maintain optimism)	Patrick thinks about everything that he and Brian discussed and thinks also about his preferences. In the end, based on his preference, Patrick decides to believe that within the next five years there is <i>no chance</i> that Brian and Maya will get divorced	Patrick thinks about everything that he and Brian discussed and thinks also about his preferences. In the end, Patrick sets aside his preference and decides to believe that within the next five years there is <i>a 70% chance</i> that Brian and Maya will get divorced.