Morality justifies motivated reasoning in the folk ethics of belief☆

Corey Cusimano [a,*], Tania Lombrozo [b]

[a] *University Center for Human Values and Program in Cognitive Science, Princeton University, Princeton NJ 08540, USA*
[b] *Department of Psychology, Peretsman Scully Hall, Princeton University, Princeton, NJ 08540, USA*

A R T I C L E   I N F O

A B S T R A C T

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 considerations affect how people evaluate others' beliefs. First, the moral value of a belief affects the evidential threshold required to believe, such that morally beneficial beliefs demand less evidence than morally risky beliefs. Second, people sometimes treat the moral value of a belief as an independent justification for belief, and on that basis, 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.

1. 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 (Austen, 2001, p. 164). But which sister is believing as she should? Is it Elizabeth, who takes herself to follow 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, 1877; Locke, 1690; Russell, 2013), such that "it is wrong always, everywhere, and for anyone, to believe anything upon insufficient evidence" (Clifford, 1877, 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 relative, 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.[1] In particular, we ask whether an individual's moral obligations to others influence how observers evaluate what that individual should believe. And if so, how do these moral considerations interact with judgments of the individual's 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 (Ståhl, Zaal, & Skitka, 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 (Pennycook, Cheyne, Koehler, & Fugelsang, 2020). 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 (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.

1.1. 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, 1877; Foley, 1987; Milkman, Chugh, & Bazerman, 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., 2020; Stanovich & West, 1998).

A large body of research suggests that lay people 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 principle, they and others are likely to be biased (e.g., that people similar to 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, & Garcia-Marques, 2012, for a review). Indeed, one limit to people's tendency to form motivated beliefs is an apparent desire to justify their beliefs with evidence (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, 2016; Kunda, 1990).

People's desire for objective and impartial evidence-based reasoning appears to extend to how they evaluate others. For instance, 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). People also tend to prefer that others share their beliefs (Abelson, 1986; Golman, Loewenstein, Moene, & Zarri, 2016; but see discussion in Gervais, 2014, and Heiphetz, Spelke, & Young, 2015 for possible exceptions) 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, & Ross, 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 others in all its forms.

1.2. The case for moral encroachment on belief evaluation

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, Kristel, Beth, Green, & Lerner, 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), and may be central to judgments of personal identity as well (Strohminger, Knobe, & Newman, 2017). 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 admirable alternative. In other words, if people treat morality as a legitimate basis for belief, 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, 1877). However, inaccurate beliefs can also 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, Fahrbach, & 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 them, 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

---

[1] By "belief" we have in mind a mental state constituted by an attitude of assent to, or endorsement of, some proposition. People possess a rich conception of belief. For instance, people commonly use beliefs to explain behavior (Malle, 2004). They also believe that beliefs are caused by information about objective reality (e.g., Ross & Ward, 1996), are voluntarily controllable (Cusimano & Goodwin, 2019), are less controllable in the face of strong evidence (Cusimano & Goodwin, 2020), and affect other mental states such as motivation (Tenney et al., 2015), to name a few (c.f. D'Andrade, 1987). Features of this lay concept may vary across cultures or depart from expert theories of belief. For instance, while cognitive scientists generally agree that beliefs are uncontrollable (e.g., Epley & Gilovich, 2016), lay people in the United States tend to think that they are controllable (Cusimano & Goodwin, 2019). We assume here that our participants' intuitive concepts of belief share the properties identified in the work cited above.

so can help people 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 overly-optimistic after deciding on a course of action (when optimism would be motivating and useful), but not when deliberating, 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). While this study provides a promising demonstration that people will prescribe non-evidential belief, it is technically 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 the outcome that the belief is about. Because Tenney et al.'s studies did not stipulate or measure the evidence available to each agent across cases, participants may have prescribed beliefs that they took to be best supported by the evidence. These studies are therefore silent on the question of whether people endorse non-evidential norms for belief. However, 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 one is supposed to think well; c.f. Cohen & Rozin, 2001; Keller, 2004; McManus, Kleiman-Weiner, & Young, 2020; 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; but see Gardiner, 2018). 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 (Nesson, 1985; Wells, 1992) or when relevant statistical evidence is caused by race-based inequality or discrimination (Tetlock et al., 2000). This work raises the possibility that people think others ought to ignore evidence when reasoning about what to believe as well.

However, while people may endorse dismissing evidence in certain decision-making contexts, it is not clear that 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 though, by the participant's own lights, those beliefs are consistent with the evidence. 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 that person's 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 expedient or morally offensive those beliefs are. However, because this work has neither controlled for, nor directly measured, the evidential status of the beliefs in question, it is ultimately inconclusive with respect to whether people prescribe motivated beliefs. We address this limitation in the studies reported below. Beyond this methodological issue, however, prior work has not addressed the question of precisely how evidential and non-evidential norms interact when people evaluate beliefs. Investigating this is the second major contribution of this paper, to which we now turn.

1.3. 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 reasoning. These two routes by which morality can influence belief are derived by distinguishing between two kinds of evaluations that people typically make of their own and others' beliefs during belief formation (Fig. 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 considerations could 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.

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 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, they 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

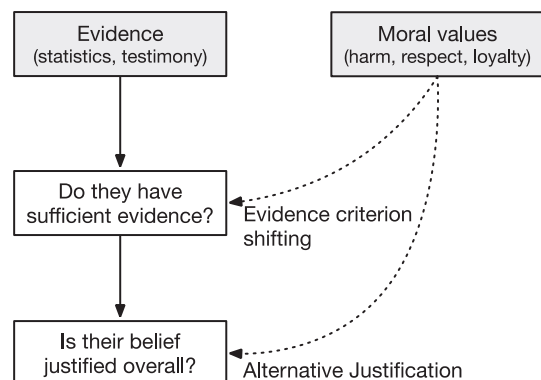**Fig. 1.** Moral values could affect what beliefs people prescribe to others in two ways. According to the "evidence criterion shifting" hypothesis, morality generates an evidential double-standard for morally beneficial vs risky beliefs. According to the "alternative justification" hypothesis, moral benefit results in prescribed motivated belief by justifying belief on non-evidential grounds.

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; Pinillos, 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 "evidence criterion shifting hypothesis."

According to the evidence criterion shifting hypothesis, people will hold others to higher standards of evidence for adopting morally risky beliefs but lower standards of evidence for adopting morally beneficial 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 considerations could affect people's evaluations of others' belief, also depicted in Fig. 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 beneficial 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 clearly inconsistent with, or challenged by, the evidence (Stroud, 2006).

The evidence criterion shifting 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 considerations are weighed against the evidentiary status of a belief such that adopting the morally beneficial belief may demand that someone sacrifice their duty to adopt a belief that is supported by the evidence. By contrast, according to the evidence criterion shifting 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.

1.4. The current studies

To summarize: Past work has documented that people often engage in motivated reasoning despite subscribing to what appear to be evidential norms for belief. At the same time, people weigh moral considerations heavily when they evaluate others' behavior and character, and moreover, in some cases, people prescribe overly-optimistic beliefs

to others. Yet, there is uncertainty regarding whether people think motivated reasoning is ever a permissible way to form a belief, and about how evidential and non-evidential considerations interact when people evaluate beliefs. 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 considerations impact people's belief prescriptions – through "evidence criterion shifting" or an "alternative justification."

To this end, we provided participants with vignettes in which the objective evidence favors one belief (e.g., that a friend is guilty) but some salient moral consideration favors the opposing belief (e.g., that the friend is innocent). If people demand that others form evidence-based beliefs, then they should prescribe beliefs that exactly correspond to an objective assessment of the evidence. Moreover, they should not judge two people who form beliefs based on the same evidence any differently, even if those individuals vary in whether they have strong moral reasons to adopt a particular belief unsupported by the evidence. However, if people evaluate others' beliefs based in part on moral considerations, then they will prescribe beliefs that deviate from an objective evaluation of the believer's evidence towards the morally beneficial belief. Our vignettes have one other notable feature; namely, that they describe situations, and make salient moral considerations, that are both common in everyday life and relevant to prior work documenting motivated reasoning. Thus, if we observe that participants prescribe motivated reasoning to the characters in these stories, we have some reason to think that they similarly export these prescriptions to the real people they think about in their real lives.

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 individual, the two individuals shared all of their evidence with each other before each forming the same belief (either on the basis of their shared evidence or despite it). Table 1 summarizes how we operationalized prescribed motivated reasoning, evidence criterion shifting, and alternative justification within each of these designs. 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 indicate the believer lacks evidence for.

1.5. Transparency

All sample sizes, exclusion criteria, and statistical analyses were preregistered. All analyses are reported as either planned or exploratory. All sample sizes represent the number of participants who completed the study and passed exclusion criteria. Participants recruited for one study were excluded from participating in any of the other studies. 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://osf.io/ygpr8/. All studies reported in this paper were approved by the Office of Research Ethics at Princeton University.

Table 1

Operationalizations for prescribed motivated reasoning, evidence criterion shifting, and alternative justification in Studies 1–3. The optimistic belief refers to the belief favored by moral considerations (e.g., believing that a friend is innocent).

Moral influences on belief evaluation	Study 1	Studies 2–3
Prescribed motivated reasoning	Indicating that a person ought to adopt a belief that is more optimistic than the belief that it would be most accurate for that person to believe based on their evidence.	An overly-optimistic belief is judged more justified and permissible for a morally-motivated believer than for a believer motivated by a mere preference, even though their evidence is the same. Likewise, an evidence-based belief is judged to be less justified and permissible for a believer who has a moral reason to be overly-optimistic than for a believer who has a mere preference to be overly-optimistic.
Evidence criterion shifting	Indicating that some optimistic beliefs are consistent with the evidence for a person subject to moral constraints, but not for an objective AI who has the same evidence.	An overly-optimistic belief is more likely to be judged to meet the criterion for "sufficient evidence" for a morally-motivated believer than for a believer motivated by a mere preference, even though their evidence is the same. Likewise, an evidence-based belief is *less* likely to be judged to meet this criterion for a believer who has a moral reason to be overly-optimistic than for a believer who has a mere preference to be overly-optimistic.
Alternative justification	Indicating that a person ought to believe something that falls outside of the range of beliefs consistent with that person's evidence.	Whether the moral quality of belief predicts how justified and permissible the belief is judged to be after accounting for its evidential quality.

2. 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 impartial 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 evidence criterion shifting.

Participants read one of six vignettes designed to pit moral considerations against evidential ones (see discussion below and overview in Table 2). 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 more likely to be subject to moral obligations) ought to believe. We reasoned that discrepancies between these two believers, in favor of morally-beneficial optimism, would reflect participants legitimizing morality's influence on belief, whereas correspondence would reflect a commitment to impartial evidence-based reasoning. Indeed, by pitting moral reasons against evidentiary ones, we can infer that, if people prescribe morally-beneficial beliefs, it is not because they did so on the basis of the evidence. To help us assess the alternative justification and evidence criterion shifting 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.

2.1. Methods

2.1.1. Participants

We recruited 839 adults (441 reported female, 395 reported male, 3 unreported, mean age 38 years) from Amazon Mechanical Turk (MTurk). An additional 144 participants were excluded for failing at least one of three comprehension questions (described 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.

2.1.2. 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; Pace, 2011; Stroud, 2006). We generated six vignettes which comprise a diverse and representative set of cases from this literature (see Table 2). 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 from which few recover. 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 gains evidence that a student is (and will continue to be) a poorly behaved bully; 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 prescribed motivated reasoning in many of these cases, 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.

2.1.3. Procedure

Participants were randomly assigned to read one of the six vignettes described in Table 2. 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

C. Cusimano and T. Lombrozo

Table 2
Overview of six vignettes used in Studies 1–3.

Vignette	Belief	Evidence	Moral consideration	Study
Bully	New student will behave poorly tomorrow.[1]	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.	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.[1]	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.	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.[1]	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.	In this particular 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 pessimistic outcomes (marked by 1 above). In the analyses below, participants responses to these items were reversed so that greater endorsement was optimistic, matching the other vignettes.

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.[2]

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.

2.1.3.1. Attributions of bias. Immediately after being told that the main character is wondering what to believe, participants reported what they

---

[2] We presented the story across several pages to make it easier to read. Comprehension questions were shown between important breaks in the story. For instance, in the Friend scenario presented here, participants answered two questions following the first major paragraph quoted above (Q1: How long have Adam and John been friends? A1: Since elementary school, Q2: What happened after high school (when they started college)? A2: They remained good friends). After reading that Adam and John spoke to each other, participants answered another question (Q3: What did the campus police apparently discover in John's dorm room? A3: Cocaine). We followed this general structure for all scenarios. See Supplemental Materials and materials on OSF for full text of the stories and comprehension questions.

thought that this person 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 towards 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 [Adam] is most likely going to believe in this situation. All the remaining questions will ask you to make different judgments about [Adam]'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.

2.1.3.2. 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. Throughout the rest of the study, we yoked together the main character and AI when referring to this estimate. For instance, on subsequent screens, participants would read statements like the following, "You have just indicated that, based on the evidence that the advanced AI/Adam has, the best estimate that John is innocent is X%" (where X would correspond to the value they had reported as being most accurate). We did so in order to reinforce the idea that this estimate applied to the main character too, and thus deviation from this estimate would entail biased or inaccurate reasoning.

2.1.3.3. 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
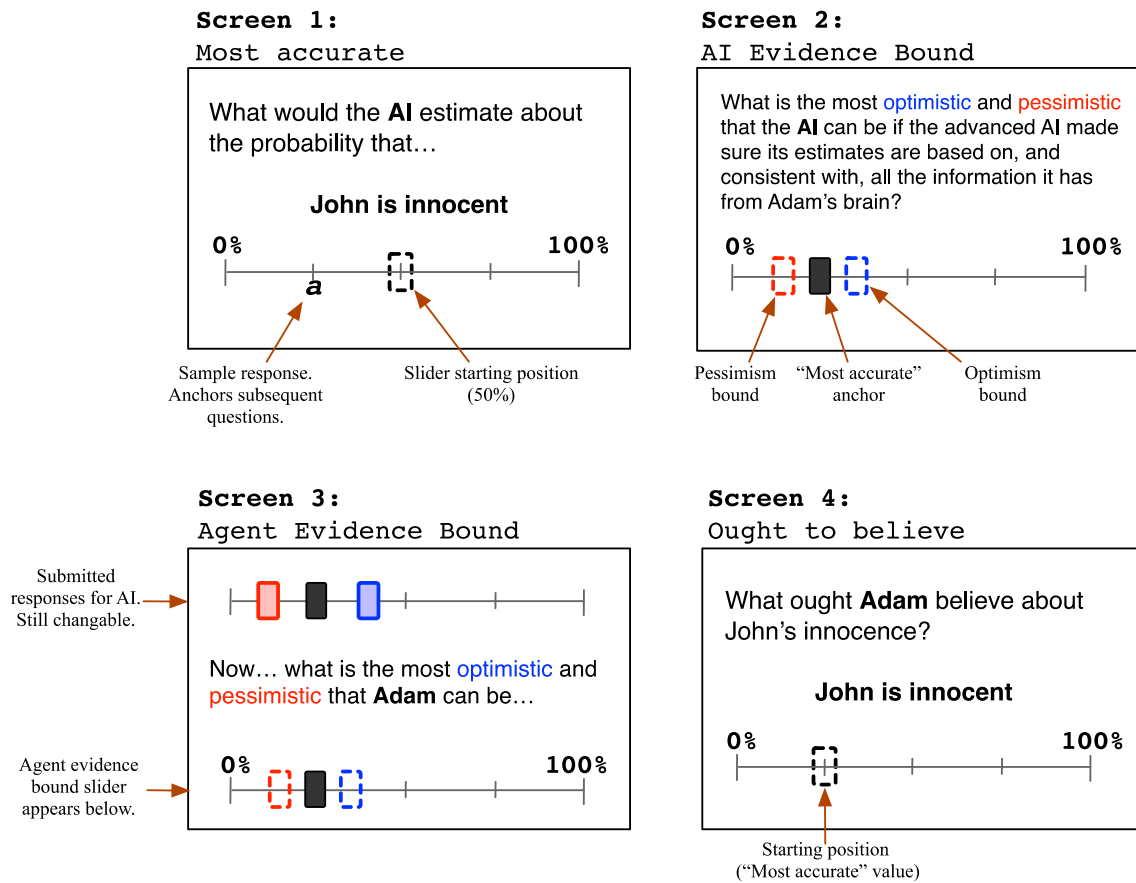
Screen 1:
Most accurate

What would the **AI** estimate about the probability that…

John is innocent

0%   100%

a

Sample response. Anchors subsequent questions.

Slider starting position (50%)

Screen 2:
AI Evidence Bound

What is the most optimistic and pessimistic that the **AI** can be if the advanced AI made sure its estimates are based on, and consistent with, all the information it has from Adam's brain?

0%   100%

Pessimism bound

"Most accurate" anchor

Optimism bound

Screen 3:
Agent Evidence Bound

Submitted responses for AI. Still changable.

Now… what is the most optimistic and pessimistic that **Adam** can be…

Agent evidence bound slider appears below.

0%   100%

Screen 4:
Ought to believe

What ought **Adam** believe about John's innocence?

John is innocent

0%   100%

Starting position ("Most accurate" value)

Fig. 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.

Fig. 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.[3] 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 evidence criterion shifting hypothesis.

2.1.3.4. Ought to believe. On the next page, participants reported what the main character 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.

2.1.3.5. 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.

2.2. Results

We first transformed participants' responses so that, for all vignettes, higher probability estimates reflected more optimistic/morally desirable beliefs and lower estimates reflected more pessimistic/morally undesirable beliefs. We also averaged together the two moral

---

[3] 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.

Table 3
Means (and standard deviations) for all main dependent measures in Study 1

Vignette	Moral quality	Most accurate	Should believe	Evidence-based pessimism boundary		Evidence-based optimism boundary	
				AI	Character	AI	Character
Bully	4.91 (1.05)	17 (16)	25 (20)[a]	9 (13)	8 (13)	38 (20)	39 (19)
Cancer	5.05 (0.97)	24 (20)	36 (31)[a]	12 (15)	14 (17)	39 (25)	42 (25)[b]
Friend	4.35 (1.00)	39 (26)	44 (24)[a]	23 (23)	25 (22)	58 (26)	67 (24)[b]
Marriage	5.26 (0.84)	49 (20)	68 (27)[a]	30 (19)	30 (18)	68 (20)	73 (21)[b]
Race	5.01 (1.08)	30 (19)	39 (22)[a]	17 (17)	17 (17)	53 (23)	54 (23)
Sex	4.90 (1.07)	30 (24)	31 (25)	19 (21)	18 (21)	45 (26)	47 (25)
Overall	4.91 (1.04)	31 (23)	41 (28)[a]	18 (20)	19 (20)	50 (26)	54 (26)[b]

Notes. Moral quality 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 towards 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.

consideration items (rs ranged from 0.36 to 0.72 across vignettes, ps < .001).[4] Means and standard errors across vignette are shown in Table 3.

2.2.1. Prescribed motivated reasoning

We submitted participants' ought and most accurate judgments to a 2 × 6 mixed-design ANOVA, with judgment type (ought vs most accurate) as a within-subjects factor and vignette as a between-subjects factor. As expected, participants' ought estimates ($M = 41$, $SD = 28$) were significantly more optimistic than their judgments of what is most accurate ($M = 31$, $SD = 23$), $F(1, 833) = 145.47$, $p < .001$, with 53% of participants giving ought estimates higher than most accurate estimates (Fig. 3A). 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 significantly from what participants reported 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).[5]

We next computed the degree to which each participant prescribed motivated reasoning by calculating the difference between participants' ought and most accurate estimates. When we regressed these difference scores on vignette, moral consideration, and their interaction, we observed that, across vignettes, greater prescribed motivated reasoning

---

[4] The moral consideration 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.

[5] Though not the focus of our investigation, readers may wonder about the relationship between what participants predicted the characters would believe, what would be accurate to believe, and what the characters ought to believe. In the Bully and Marriage vignettes, participants predicted that the characters would not be motivated believers (ps > 0.41) but that they should be (ps < 0.001), and that they should be more motivated than expected (ps < 0.001). In the Friend and Race vignettes, participants estimated that the character would be about as motivated as they prescribed the person to be (ps > 0.11). In the Cancer scenario, participants predicted that the character would be motivated, $t(131) = 4.04$, $p < .001$, but that she should be even more motivated than they predicted that she would be $t(131) = -2.39$, $p = .018$. And finally, in the Sex scenario, participants did not prescribe, nor predict, any motivated reasoning. In no vignette did we observe (on average) participants attributing more motivated reasoning than they prescribed. These results suggest a partial replication of previous findings, reported by Armor et al. (2008), that in many cases in which people prescribe unrealistic beliefs to others, they prescribe beliefs beyond what they expect others to hold.

was strongly associated with greater moral consideration, $F(1, 827) = 32.34$, $p < .001$.

2.2.2. Evidence criterion shifting hypothesis

Consistent with the evidence criterion shifting hypothesis, participants reported that the main characters were licensed to evaluate the evidence differently from the perfectly impartial, but equally informed, observer (the AI). We submitted participants' evidential optimism bounds for the AI and the main character to a 2 × 6 mixed-design ANOVA, with target (AI vs character) as a within-subjects factor and vignette as a between-subjects factor. As expected, participants believed the character was licensed to be more optimistic ($M = 54$, $SD = 25$) 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 ≤ .037), but not in Bully, Race, and Sex (ps ≥ .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 ($M = 19$, $SD = 20$) and AI ($M = 19$, $SD = 20$) could be, $F(1, 833) = 0.72$, $p = .397$ (Fig. 3B).

Finally, we observed mixed support for an association between evidence criterion shifting and moral consideration. 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 the strength of moral considerations in that vignette, $F(1, 827) = 0.004$, $p = .948$.

2.2.3. 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 may 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 4). In some vignettes, such as Friend, most participants who

A. Prescribed motivated reasoning
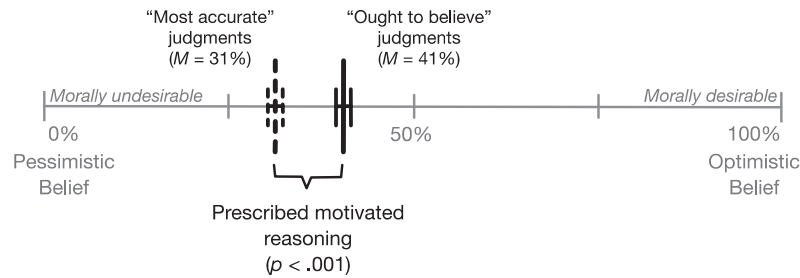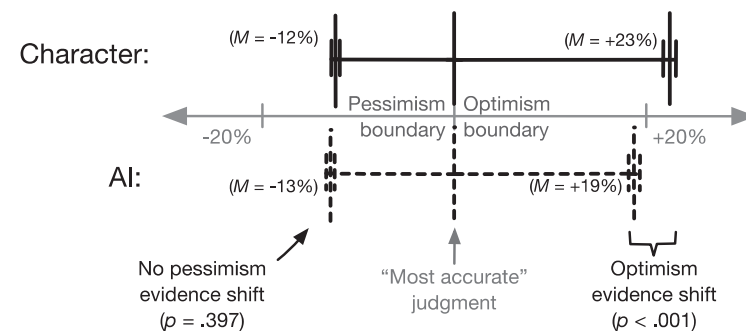

B. Evidential-shift

Fig. 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.

Table 4

Percentage of participants showing evidence for the prescribed motivated reasoning, evidence criterion shifting, and alternative justification hypotheses across vignettes in Study 1.

Vignette	N	Evidence criterion shifting		Prescribed motivated reasoning	
		Optimism	Pessimism	Overall	Beyond evidence
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 evidence criterion shifting is kept here for comparison; on average there was no difference between AI and character evidence-licensed pessimism.

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).

2.3. Discussion

Study 1 found support for prescribed motivated reasoning, the evidence criterion 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 salient and 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. That is, there is no way that forming a belief about whether the friend is guilty could affect the friend's past behavior. This finding therefore demonstrates that an (overly) optimistic belief does not have to be potentially self-fulling for people to prescribe it to others. Rather, the moral benefit of trusting a friend's testimony over the accumulated evidence engenders an obligation to give the friend the benefit of the doubt.

Study 1 also offers some support for the evidence criterion shifting 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 belief. This difference was driven by three out of the six vignettes, Cancer, Marriage, and Friend (Table 3). 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 evidence criterion shifting. 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 considerations 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 evidence criterion shifting were also more likely to prescribe motivated reasoning, we did not observe an association

between participants' judgments of the strength of moral consideration 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 evidential standards in our vignettes could be shifted. This is consistent with the observation that participants' average evidence criterion shifting was much smaller (4%) than their average prescribed motivated reasoning (10%). And finally, asking participants to think about the range of beliefs consistent with the evidence may have felt strange for them to do. We address this concern in Study 2 by using a more natural way of measuring the perceived evidential quality of a belief. Nevertheless, observing evidence criterion shifting in the majority of scenarios in which we observed prescribed motivated reasoning offers tentative support for the evidence criterion shifting hypothesis.

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 reported would be inconsistent with the main character's evidence. However, 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 perceived evidential quality.

3. 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, participants evaluated the beliefs that two characters had already formed. As in Study 1, one character was socially close to the target that the belief was about (e.g., a close friend or spouse). Based on findings from Study 1, we hypothesized that, because certain beliefs are treated as morally beneficial in close relationships, participants would attribute a moral reason for that character to adopt that belief. However, participants also read about a socially distant character in the same situation (e.g., a more distant friend) who plausibly shared the same relevant information as the socially close character. Because this new character was socially distant, we hypothesized that they would be seen as having weaker (or possibly no) moral obligations to adopt specific beliefs. These two characters – the socially close character with a moral reason to be optimistic, and a socially distant character who lacks one – then adopt either a desirable overly-optimistic belief, or they adopt an undesirable, but evidence-based belief. Participants evaluated the extent to which each person's beliefs were justified, permissible, supported by sufficient evidence, constituted knowledge, and were moral. We predicted that participants' evaluations of others' beliefs would take into account both whether the believer had a strong moral reason to be optimistic (manipulated here by social distance), and whether the believer acquiesced to those demands (and so formed a desirable belief) or did not (and so formed an evidence-based belief).

This design naturally gives rise to the worry that social distance and moral obligation are confounded with one another. This is a serious worry, and we detail below how we address it. However, taking on this risk confers an important benefit; namely, that it better reflects how moral considerations during belief evaluation likely operate in everyday life. Putative moral reasons to adopt certain beliefs about others seem to correlate with how close one is to those about whom one is reasoning. And indeed, people tend to adopt motivated beliefs about others whom

they are close to rather than about others whom they are distant to. Thus, if we can successfully address the worry that social distance may bring with it differences in information, and so isolate the unique contribution of moral considerations on people's evaluations, then we gain a richer understanding of how people evaluate others' beliefs in contexts with which they are familiar, and in which motivated reasoning actually does seem to occur.

These methodological changes also allow us to conceptually replicate and extend the findings from Study 1 in other ways. For instance, 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 considerations. This alleviates a possible concern in Study 1, namely that observers want others to form biased 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 (in this case, a strong preference to hold a belief), thereby further isolating the relevance of morality for belief evaluation. And finally, by investigating people's evaluations of already-formed beliefs, we can measure the prescribed motivated reasoning, evidence criterion shifting, and alternative justification hypotheses using multiple, triangulating measures that differ from those employed in Study 1 (see Table 1 and paragraph below).

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 evidence criterion shifting 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 as morally beneficial. 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.

3.1. Methods

3.1.1. Participants

We recruited 1,021 adults (573 reported female, 443 reported male, 2 reported non-binary, 3 unreported, mean age 40 years) from MTurk. An additional 145 participants were excluded for failing at least one of three comprehension checks (explained below). We expected after exclusions to have roughly 250 participants per scenario, with 125 for each of the two between-participant conditions within each scenario. 125 participants per cell yields 90% power to detect small differences ($d = 0.30$) within each cell.

3.1.2. Design

Study 2 used a 2 (Belief: evidence-based vs optimistic) × 2 (Reason to be optimistic: moral vs preference) × 4 (Vignette) mixed design. Participants read one of four vignettes from Study 1 that readily enabled our reason 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 putative moral

reason to hold an optimistic belief about that person (moral reason condition), as well as another character who, in light of being socially distant from that person, lacks a strong, justifiable moral reason to hold an optimistic belief.[6] Because the socially distant character lacks a strong moral reason to be optimistic, we stipulated that their reason to be optimistic was based on their preference (preference reason condition). This comprised our Reason manipulation. For instance, the close friend in the Friend vignette weighed the evidence he had against his obligation to trust his friend, whereas the socially distant observer (a friend of a friend) weighed the evidence against her preference not to think that people ever get involved in drugs (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 non-evidential 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 the result of the believer prioritizing their 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 5 below for the 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 reason or preference to adopt an optimistic belief) or that they adopt an overly-optimistic belief (on the basis of their moral reason or preference). 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 5 for the text in the Friend vignette.

3.1.3. Procedure

Vignettes were minimally modified from Study 1 so that the socially close character discusses what they should believe with the new, socially distant character until the socially distant character ends up with all the same relevant information. For instance, in the Friend vignette, participants read the following after reading the vignette excerpted above in Study 1:

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

---

[6] 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 moral obligation whatsoever. Instead, our design in Studies 2 and 3 involves a manipulation of social distance where the distant character is not perfectly removed from the situation. Instead, they still have some close tie to the target, and therefore plausibly are seen as having some social obligations to hold certain beliefs. We made this choice because it would be implausible that a character who was completely distant from the situation could have the same information as a close friend or a spouse would. We use the language of possessing a moral reason to be optimistic or not to reflect differences in the presence of unique, and uniquely strong moral reasons to hold particular beliefs. We thank Liz Harman and Melissa Lane for encouraging us to clarify this point.

Table 5
Text used to manipulate the presence of a preferred belief or dis-preferred belief across conditions in the Friend vignette in Study 2.

Reason condition	Belief condition	
	Optimistic belief	Evidence-based belief
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 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 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, Adam sets aside his loyalty to John and decides to believe that John is guilty.
Preference 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 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 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, Jasmine sets aside her preference and decides to believe that John is guilty.

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 further 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 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.

3.2. Results

3.2.1. Data preparation and analysis plan

We created composite measures of the moral quality of the belief, the moral quality of the believer, the evidential quality of the belief, and overall belief quality. The two-item pairs for each question passed our preregistered criterion ($r > 0.50$) across both evidence-based belief and optimistic belief conditions. There were two exceptions: The two-item measures correlated with each other poorly in the Bully, $r(512) = 0.49, p < .001$, 95% CI [0.42, 0.55], and Friend, $r(520) = 0.31, p < .001$, 95% 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.[7]

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) × 2 (Reason: Moral vs Preference) × 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 have more information) 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 × Reason interaction, as that 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 Fig. 4.

3.2.2. 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 moral reason belief ($M = 4.80, SD = 1.63$) no differently than the preference reason belief ($M = 4.51, SD = 1.39$), $F(1, 1013) = 70.71, p < .001, \eta_G^2 = 0.01$. However, these results were qualified by a significant Belief × Reason interaction, $F(3,1013) = 23.1, p < .001, \eta_G^2 = 0.05$. In the evidence-based condition, participants rated the moral reason belief to be morally worse ($M = 3.87, SD = 1.51$)

---

[7] All analyses yield identical results with one exception. In the Friend vignette, we observed a significant Reason × Belief interaction when analyzing participants' judgments of how "loyal" the resulting belief was, $F(1, 259) = 141.36, p < .001$, 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.

than the preference reason 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 moral reason belief as morally superior ($M = 5.75, SD = 1.13$) to the preference reason belief ($M = 4.87, SD = 1.37$), $F(1, 501) = 313.14, p < .001, \eta_G^2 = 0.12$.

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 moral reason individual slightly morally better ($M = 5.37, SD = 1.33$) than the preference reason 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 × Reason interaction, $F(1, 1013) = 80.34, p < .001, \eta_G^2 = 0.01$. Participants did not judge the moral reason individual to have worse character ($M = 4.91, SD = 1.43$) than the preference reason individual ($M = 4.97, SD = 1.22$) when they both adopted the evidence-based belief, $F(1, 512) = 1.64, p = .202$. By contrast, when the two characters both adopted the optimistic belief, participants judged the moral reason individual to have better character ($M = 5.85, SD = 1.02$) than the preference reason individual ($M = 5.34, SD = 1.14$), $F(1, 501) = 145.18, p < .001, \eta_G^2 = 0.05$.

3.2.3. Evaluation of evidence and knowledge

To test for the evidence criterion shifting hypothesis, we next analyzed participants' evaluations of evidential quality and knowledge. As predicted, participants' judgments of the evidential-quality of the two beliefs followed a pattern similar to their moral judgments. On average, participants rated the evidence-based beliefs ($M = 4.44, SD = 1.68$) to be evidentially better than the 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 moral reason ($M = 4.12, SD = 1.65$) and preference reason ($M = 3.95, SD = 1.74$) beliefs, $F(1, 1013) = 18.94, p < .001$. As predicted, we observed a significant Belief × Reason interaction, $F(1, 1013) = 75.56, p < .001, \eta_G^2 = 0.01$. When evaluating the evidence-based belief, participants rated the moral reason belief as evidentially worse ($M = 4.36, SD = 1.68$) than the preference reason belief ($M = 4.52, SD = 1.67$), $F(1, 512) = 9.60, p = .002, \eta_G^2 < 0.01$. By contrast, when evaluating the optimistic belief, they now rated the preference reason belief ($M = 3.37, SD = 1.61$) as evidentially worse than the moral reason belief ($M = 3.87, SD = 1.58$), $F(1, 501) = 79.8, p < .001, \eta_G^2 = 0.03$.

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 × Moral quality interaction, and by-participant 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 (i.e., 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 in the moral reason condition ($M = 3.90, SD = 1.91$) compared to the preference reason condition ($M = 3.58, SD = 1.95$), $F(1, 1013) = 82.8, p < .001, \eta_G^2 = 0.01$. However, similar to the results above, we observed a significant Belief × Reason interaction, $F(1, 1013) = 50.07, p < .001, \eta_G^2 = 0.01$. When participants were told that the evidence-based belief was true,
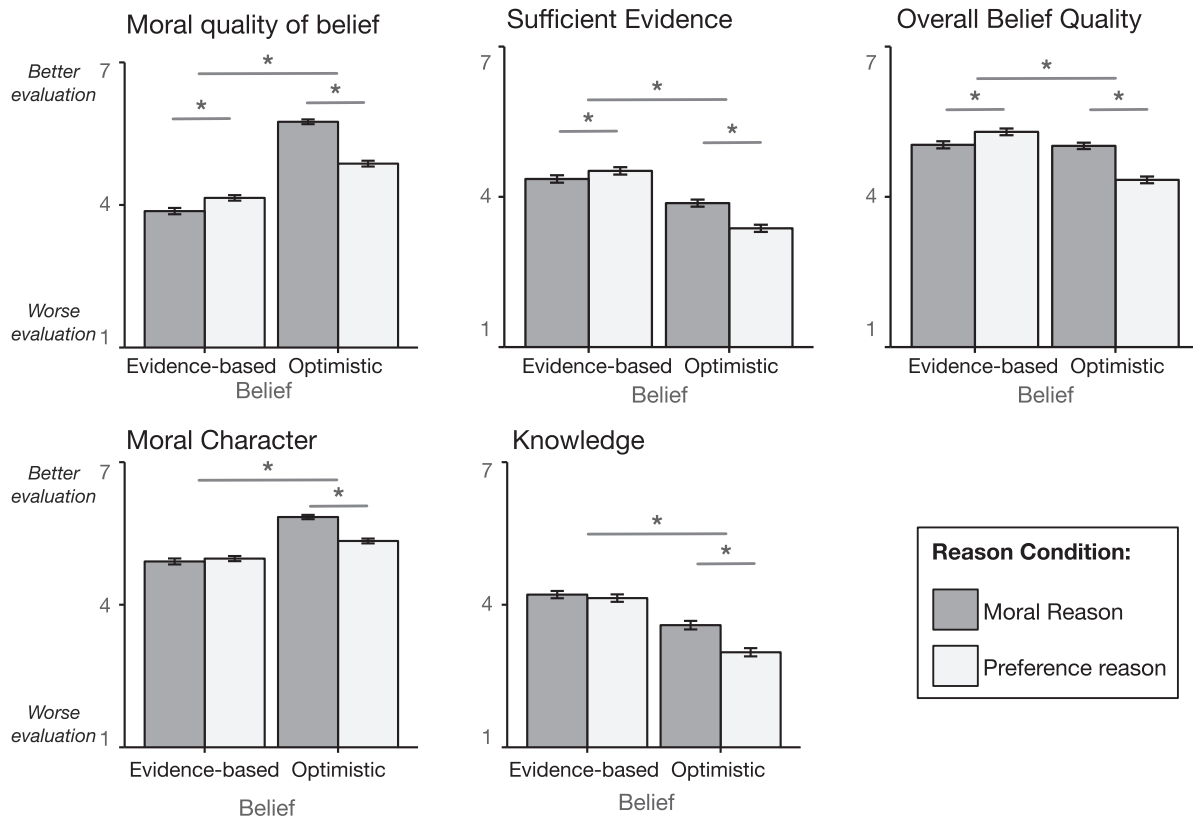
Fig. 4. Means and standard error (error bars: ±1 standard error) across condition for ratings of moral quality of belief, moral quality of believer's character, evidential quality, knowledge, and overall belief quality in Study 2. Higher values indicate more positive evaluations. * $p < .005$.

there was no difference in knowledge attributed in the moral reason condition ($M = 4.21, SD = 1.74$) and the preference reason condition ($M = 4.14, SD = 1.78$), $F(1, 512) = 2.67, p = .103, \eta_G^2 < 0.01$. However, when participants were told that the optimistic belief turned out to be true, participants attributed more knowledge in the moral reason condition ($M = 3.57, SD = 2.02$) compared to the preference reason condition ($M = 3.00, SD = 1.95$), $F(1, 501) = 111.64, p < .001, \eta_G^2 = 0.02$.

3.2.4. 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 moral reason beliefs were rated slightly better on average ($M = 5.03, SD = 1.53$) than the preference reason 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 moral reason character was seen as less justified (and the believer less permitted) ($M = 5.04, SD = 1.62$) compared to the preference reason 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 overly optimistic belief, the moral reason character was now seen as more justified and permitted ($M = 5.02, SD = 1.43$) relative to preference reason 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 evidentiary quality of the moral reason and preference reason beliefs, or they could have reflected those differences in addition to evaluations of the moral

quality of the belief. 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 quality 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 Reason × evidentiary quality and Reason × belief moral quality interactions as well as random by-participant intercepts of overall belief quality. Unsurprisingly, evidentiary quality strongly 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.

3.2.5. Variation across vignette

All of the results we report above were qualified by by-vignette interactions (see by-vignette results in Appendix B Table B1). Specifically, while we observed clear evidence for prescribed motivated reasoning and for evidence criterion 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 evidence criterion 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 (i.e., how helpful and constructive participants rated the teacher's belief to be), 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 evidentiary 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. Although these correlational results are consistent with prescribed motivated reasoning, we did not observe the predicted Reason × Belief interactions on overall belief evaluation, sufficient evidence, or knowledge ($ps > .05$).

3.3. 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. We also found that a distant observer, who lacked a moral reason to form an optimistic belief, was not licensed to form an overly-optimistic belief 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 evidence criterion shifting 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. We observed further evidence of evidence criterion shifting from participants' knowledge attributions. In the overly-optimistic condition, participants reported that the moral reason character had more knowledge of the eventual outcome, but this advantage disappeared when the two characters formed the evidence-based belief. And finally, 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.

Participants also made inferences about the moral quality of the believer that incorporated information about both the moral and evidentiary quality of that person's belief. This finding provides some evidence against the claim that people prefer not to interact with others that they perceive as biased. However, it is possible that these differences in character evaluation did not solely reflect the beliefs that the main characters had formed, but also reflected other facts about how those individuals had reasoned. For instance, in this study (as well as Study 3, below), the characters in the moral reason condition actively dismiss one of their moral obligations. It could be that this dismissal of a moral consideration signals poor character in a way disconnected from whether they formed a morally beneficial belief. Our studies do not tease apart character inferences based on perceptions of the agent's broader values from inferences based on how they reason about which beliefs to form. For this reason, this study provides weak support for the idea that adopting an evidence-based belief, on its own, is sometimes sufficient to

tar perceptions of someone's character.

There is a strong 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, reinforced throughout the study, that the socially close and socially distant characters had the same relevant information. Instead, according to the skeptical interpretation, perhaps some participants thought that the socially close 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 provide the most accurate estimate based on both characters' information, we are unable to detect whether this happened, and therefore unable 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 socially close and socially distant characters. We 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 preference-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 directly 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 preference-motivated character is likely confusing in light of repeated 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 participants will provide estimates of the character's current evidential standing by 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.

4. Study 3

The primary goal of Study 3 was to test whether people evaluate beliefs as predicted by the *alternative justification* and *evidence criterion shifting* hypotheses even after removing all participants who report differences in the information held by the socially close (moral reason) character and the socially distant (preference reason) character. To this end, we conducted a lightly-modified replication of the "marriage" vignette from Study 2. Participants rated this vignette as having the strongest moral reasons for optimism in Studies 1 and 2, and we reasoned that it would yield the most power to reliably detect any differences in belief evaluation after removing part of our recruited sample.

A secondary goal of Study 3 was to examine which properties of belief people might associate with moral quality. As discussed in the Introduction, prior literature suggests that people may evaluate beliefs as morally beneficial because they expect such beliefs to bring about morally good behavior, or potentially because they treat certain beliefs 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.

4.1. Methods

4.1.1. Participants

We recruited 233 adults (126 reported female, 106 reported male, 1 unreported, mean age 39 years) from MTurk. An additional 65 participants were excluded for failing at least one of three comprehension checks.

4.1.2. Design

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

4.1.3. 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 Table C1 (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 [his wife], 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 that each of the two characters 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 prospective judgments about what Brian and Patrick should believe. Specifically, participants answered four questions for each character, probing (i) obligation (e.g., "To what extent does Brian have a moral obligation to believe the best of his marriage?"), (ii) behavioral impact (e.g., "If Brian is in fact pessimistic about his marriage, how difficult would it be for him to act in all respects as if he is genuinely optimistic about his marriage?"), (iii) 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 (iv) 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 properties of a belief people associate with moral quality. 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 versions of the question for Brian and Patrick were always adjacent. The four pairs were shown in a random order for each participant, however, questions (ii) and (iii) were always shown together (and in order) to highlight the contrast between them. Finally, across participants we counterbalanced whether, within each question pair, participants first responded for Brian or for Patrick.

Participants then learned what Brian (moral reason condition) and

Patrick (preference reason condition) 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 in spite of, a moral reason (Brian, moral reason condition) or a preference (Patrick, preference reason condition). For instance, Brian feels as if he has 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 belief's (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.

4.2. Results

4.2.1. Most accurate estimate judgments

Prior to discarding any participants from our analyses, we found that, on average, participants thought that the husband, Brian (moral reason condition) had slightly different evidence than the friend, Patrick (preference reason condition). 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% CI [$-11.6, -4.83$]. In the optimistic belief condition, people attributed to Brian an evidence-based estimate of 53% ($SD = 19\%$) compared to Patrick's 58% ($SD = 18\%$), $t(120) = -2.97$, $p = .004$, 95% 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% 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% 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.

4.2.2. 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 the belief's overall quality. The two-item pairs for each question passed our preregistered criterion ($r > 0.50$) across both evidence-based

belief and optimistic belief conditions, with two exceptions for which the independent judgments yielded similar results.[8] 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 Preference) fully-crossed ANOVAs. For readability, we will describe differences in ratings across the Reason manipulation as between Brian, the husband (who has a moral reason to be optimistic), and Patrick, the friend (who has a preference to be optimistic).

4.2.3. 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 the 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 Fig. 5). When Brian (the husband) and Patrick (the friend) adopted the evidence-based belief, participants judged Brian's belief to be morally worse, less supported by the evidence, less justified or permissible, and Brian as morally worse than 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 ($ps < .001$).

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, $F(1, 140) = 0.50$, $p = .482$. However, on average participants did attribute more knowledge to the husband ($M = 3.58$, $SD = 1.93$) than to the friend ($M = 3.31$, $SD = 1.96$), $F(1, 140) = 8.42$, $p = .004$. We also failed to observe a significant Belief x Reason interaction, $F(1, 140) = 0.68$, $p = .41$.

4.2.4. 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, who is not a part of the relationship but has a preference to be optimistic. 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 there is a 0% chance of divorce ($M = 4.03$, $SD = 1.81$), $t(139.47) = -1.08$, $p = .282$, 95% 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% CI [-0.66, 0.41]. By comparison, even though participants attributed the same evidence to Brian, the husband, they evaluated his beliefs in a way

that revealed strong consideration of his personal obligations. 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% 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% CI [-1.74, -0.65]. These differences are striking because participants reported that, objectively, Brian's evidence favored a pessimistic belief (i.e., a 70% chance of divorce) rather than an optimistic one. Therefore, participants judged Brian to be more justified, both evidentially and overall, to adopt a belief that was objectively evidentially unsupported compared to a belief that was objectively evidentially supported. Beyond the comparison to Patrick, participants' evaluations of Brian provide independent evidence for the potent impact morality has on belief evaluation.

4.2.5. Exploratory hypotheses

We next examined participants' judgments of the two characters' beliefs prior to learning what they both believe. Even though participants had reported that the husband and friend held the same evidence, these participants 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 was 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 had over his own ($M = 3.96$, $SD = 1.84$), $t(141) = 2.97$, $p = .003$, 95% CI [0.17, 0.84].

We next calculated husband-minus-friend difference scores for these four judgments. Differences in perceived obligation to be optimistic correlated with differences in perceived intrinsic wrongness of pessimism, $r(140) = 0.22$, $p = .01$, 95% CI [0.06, 0.38], as well as differences in the anticipated deleterious effect of such pessimism on behavior, $r(140) = 0.28$, $p < .001$, 95% CI [0.13, 0.43]. Differences in the anticipated effects of pessimism on behavior also correlated with differences in the perceived intrinsic wrongness of pessimsm, $r(140) = 0.40$, $p < .001$, 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 judgments predicted participants' later evaluations of the husband's and friend's beliefs. We calculated husband-minus-friend or friend-minus-husband difference scores such that greater positive differences indicated participants' judgments that 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 scores and participants' prior judgments about the characters' moral obligations 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% CI [0.03, 0.35], as well as whether the husband had sufficient evidence for his belief, $r(140) = 0.19$, $p = .02$, 95% 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% CI [-0.12, 0.2], or with differences in the perceived moral character of the two people, $r(140) = 0.07$, $p = .44$, 95% CI [-0.10, 0.23].

---

[8] In the optimistic-belief condition, judgments of the belief's permissibility and justification correlated weakly ($r = 0.29$, $p < .001$), as did judgments of how loyal and morally good the agent's reason was ($r = 0.43$, $p < .001$). 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.
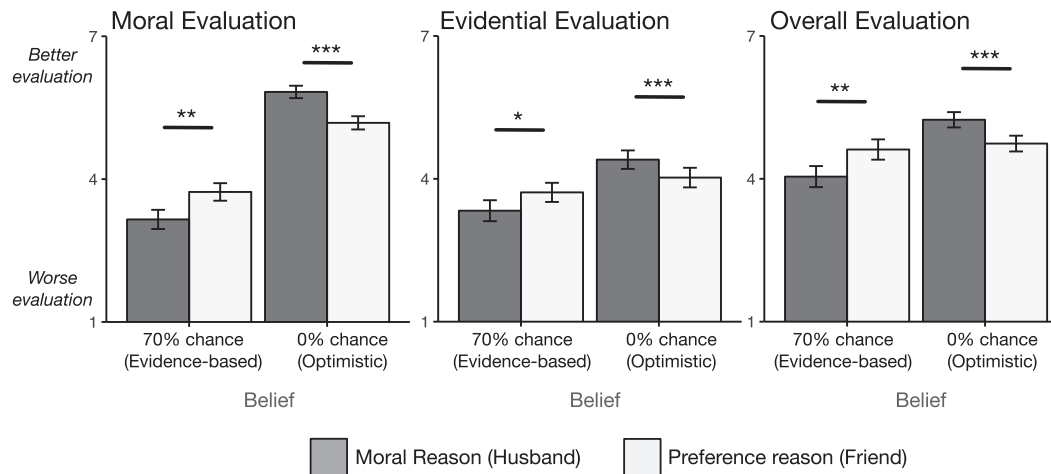
Fig. 5. Means and standard error (error bars: ±1 standard error) across condition for ratings of moral quality, evidential quality, and overall belief quality in Study 3. * $p < .05$, ** $p < .01$, *** $p < .001$.

4.3. Discussion

Study 3 replicated the prescribed motivated reasoning, evidence criterion shifting, and alternative justification findings 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 objectively 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 documented what features of a belief participants associate with a moral obligation to hold that belief. Before learning which beliefs the two characters adopted, participants were asked to make several in-principle judgments about each character and their potential beliefs. As expected, participants reported that, despite having the same evidence, the husband was more obligated to remain optimistic than the friend. This difference in perceived obligation predicted how participants evaluated the two characters later on. Additionally, 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, reported that he was obligated to be optimistic.

Finally, Study 3 ruled out another possible skeptical interpretation of our findings; namely, that participants are judging the socially close believer differently because they attribute to that individual less control over what they believe. The reasoning behind this objection is that someone may be more permitted to hold a non-evidential belief when their particular situation causes in them 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, slightly more control to the socially close believer than they did to the socially distant believer.

5. 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 (Golman et al., 2016), predicts that people will demand that others set aside moral concerns to 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 legitimate influences on belief. This would entail that people sometimes prescribe motivated reasoning to others. We then articulated two ways that 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 "evidence criterion shifting 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 prescribe motivated reasoning to others, and further whether they do so in line with the evidence criterion shifting 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 beneficial 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 beneficial 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 an overly-optimistic belief compared to someone else with the same information but who lacked a moral justification (and so adopted the overly-optimistic belief on the basis of a strong preference). Moreover, when both people adopted an evidence-based belief, the believer who disregarded a moral benefit to do so was judged to be less justified than someone who merely gave up a preference to do so. And finally, in Study 3, participants reported that, even though a spouse and a friend held the same evidence about the objective chances of the spouse's divorce, the spouse had a stronger obligation to remain optimistic about the marriage than the friend did. Taken together, 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.

Consistent with the evidence criterion shifting hypothesis,

participants also evaluated others' beliefs by applying evidential double-standards to them. 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 similarly rated more consistent with the evidence for the main character compared to the impartial observer. Studies 2 and 3 provided converging evidence using different measures of perceived evidential quality. In these studies, participants judged that overly-optimistic beliefs were more likely to pass the threshold of "sufficient evidence" when the believer had a good moral reason to adopt those beliefs, compared to a believer who adopted the same beliefs based on a mere preference. Likewise, participants judged that beliefs that disregarded a good moral reason were less likely to have sufficient evidence compared to beliefs that disregarded a preference. Importantly, these differences in evidentiary quality arose even though the two beliefs were backed by the same objective information. Finally, Study 2 (though not Study 3), documented evidence criterion shifting using an indirect measure of evidence quality, namely, attributions of knowledge. In sum, these findings document that one reason why an observer may prescribe a biased belief is because moral considerations change how much evidence they deem necessary to hold the belief in an evidentially satisfactory way.

Finally, these 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. Study 1 documented 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. Studies 2 and 3 revealed a more subtle way in which moral considerations directly affect belief evaluation. Across all the vignettes, the moral quality of the belief – such as how helpful or loyal the belief was – predicted participants' evaluations of how justified and permissible the belief was to hold even after accounting for the evidential quality of the belief. Thus, people will sometimes prescribe a belief to someone knowing that the belief is unsupported by that person's evidence because the belief confers a moral benefit.

Though the focus of the current investigation was to determine whether people prescribed motivated reasoning in certain common situations, it is important to note that these studies documented substantial evidence that people think beliefs ought to be constrained by the evidence. In Study 1, participants prescribed beliefs that were close to what they thought was best supported by the evidence (Table 3). Indeed, on average, participants prescribed beliefs that were pessimistic (i.e., closer to the evidence) rather than optimistic (i.e., morally preferable) (Fig. 3). Additionally, Studies 2 and 3 documented a strong association between the perceived evidentiary quality of the belief and judgments of the belief's permissibility and justifiability. Specifically, the less sufficient the believer's evidence, the less justifiable and permissible it was for them to hold the belief. Thus, while these 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.

5.1. Alternative explanations

Two alternative explanations for our findings stem from the observation that we manipulated moral obligation by changing the social distance between two believers. 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 that one is forming a

belief about either (1) makes one's belief more likely to be self-fulfilling, or (2), creates a reason to be more diligent and therefore more withholding of belief in general. We address each of these two concerns below.

5.1.1. Self-fulfilling beliefs

Adopting a belief can make certain outcomes more likely. For instance, adopting an optimistic belief could cause one to feel happier, try harder at some task, or bring about a beneficial outcome. We hypothesized that, sometimes, people treat these effects as constituting moral reasons to adopt a specific belief. For instance, if adopting an optimistic belief about a spouse's prognosis could improve their prognosis, then this benefit may constitute a morally good reason to adopt the optimistic belief. However, when the outcome in question is also what the belief is about, as it is in this example, then the belief is potentially a "self-fulfilling" belief. Self-fulfilling beliefs could confound moral reasons to adopt a belief with evidential reasons to adopt the belief. This can happen if participants attribute to the believer of a self-fulfilling belief the additional belief that their belief is self-fulfilling, which would then entail that this person has more evidence (in the form of the belief that they hold) in favor of the outcome that they have formed a belief about. For instance, participants may infer that, if the newlywed husband in Study 3 makes it more likely that he will not get divorced by adopting the belief that he has a 0% chance of divorce, then the observation that he has formed this belief may constitute additional evidence that he has a 0% chance of divorce. His friend who adopts the same belief would not have access to this additional evidence because the friend's belief does not affect the husband's outcome. If participants reason about beliefs in this way, then it is possible that the cases in which people seem to be endorsing non-evidential grounds for belief are really cases in which participants are inferring the presence of new evidence stemming from the self-fulfilling belief.

Several findings from the studies above speak against this skeptical proposal. First, if the socially close character's beliefs are treated as self-fulfilling, and therefore as evidentially self-supporting, then this feature of their beliefs ought to apply to pessimistic beliefs just as it does to optimistic ones. However, as we observed in Studies 2 and 3, when the husband adopts the pessimistic belief, participants judge his belief as worse than the friend who adopts the same pessimistic belief, directly contradicting this prediction. Put another way, a self-fulfilling account predicts that close others will always be judged as better evidentially situated than distant others. Thus, the statistical interaction between believer and belief, documented in Studies 2 and 3, rules out this interpretation. And second, in Studies 1 and 2, prescribed motivated reasoning, evidence criterion shifting, and alternative justification were all supported in the Friend scenario. In this scenario, the relevant belief concerned something that occurred in the past, namely, whether the cocaine that had been discovered belonged to the friend or not. Because the belief concerns something in the past, neither an optimistic nor pessimistic belief could affect its likelihood of being true. Thus, while it is possible that, in some vignettes, participants could treat self-fulfilling beliefs as evidentially self-supporting, this potential confound cannot fully explain our results.

5.1.2. Norms of due diligence

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 your knowledge 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 – even if you otherwise have similar reasons to think that you usually leave them down. 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; Pinillos, 2012), and people in these situations actually do engage in more thorough reasoning when the risks are high (Fiske & Neuberg,

1990; Kunda, 1990; Mayseless & Kruglanski, 1987; Newell & Bröder, 2008; 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 does, and therefore require more evidence before becoming too confident. Likewise, in Studies 2 and 3, perhaps participants think that people should reason more diligently about those to whom they are close compared to those to whom they are distant, and this norm explains why believers were evaluated negatively for adopting pessimistic beliefs. In sum, perhaps changes in social distance affect how diligent one must be when reasoning, rather than affecting whether one ought to reason in a motivated or biased way.

However, norms of diligence also fail to fully explain results from these studies. In Study 1, a due diligence explanation would predict that a wider range of beliefs 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 evidence criterion shifting only for more morally-desirable beliefs, not for more morally-undesirable beliefs, inconsistent with predictions based on due diligence. Similarly, in Studies 2 and 3, a due diligence explanation would predict that, based on the same amount of information, the socially close character 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 having a distal one, does not impose a demand to be careful in your beliefs, but instead imposes a demand to be partial.

5.2. Implications for motivated reasoning

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 (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 certain decisions because of a normative commitment to not rely on certain kinds of evidence (Nesson, 1985; Wells, 1992), or because of a normative commitment to prioritize deontological concerns over utility-maximizing concerns (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; Cao et al., 2019; Metz, Weisberg, & Weisberg, 2018; 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 to others.

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, McGonigal, Richards, Butler, & Gross, 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 motivated reasoning. 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 that one has.

Beyond this general observation about why motivated reasoning may come about or persist, our results also hint at a possible mechanism for how moral norms for belief 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 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 in Karma, 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, Tabatabaeian, Fleming, Sulzmann, & Holbrook, 2019, for some preliminary findings consistent with this proposal; see James, 1937, and Pace, 2011, for fuller discussion of how to evaluate evidence for morally beneficial beliefs). Although this is speculative, it naturally follows from the findings presented here and presents a valuable direction for future research.

5.3. Moderating prescribed motivated reasoning

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 is that the relevant moral norms in those vignettes did not outweigh demands to be accurate. Indeed, participants reported on average the least moral concern for the Sex vignette, raising the possibility that a putative demand to favor a particular belief in that scenario was not strong enough to override the norm to be objective.[9] Likewise, in the Bully vignette, there could have been strong reasons to be diligent and accurate that directly competed with reasons to be partial, but which we had not foreseen when constructing our materials. 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. This explanation is speculative, but it is consistent with prior work documenting that people temper their recommendations for over-optimism when the risks outweigh the potential benefits. For instance, Tenney et al. (2015) found that people were less likely to prescribe optimism to others when those others were in the process of making a decision compared to when a decision had already been made. This was presumably because making a decision on wrong information is unnecessarily risky in a way that over-optimism after a decision is not. In general, these considerations suggest that, just as people are sensitive to the benefits of accuracy and bias when setting their own reasoning goals (c.f. Kruglanski, 2004), it is likely that they incorporate the comparative advantages of accuracy and bias when prescribing beliefs to others.

Though we tested a wide range of scenarios in the current studies, the range of morally beneficial beliefs was still relatively limited. Specifically, many of the scenarios we tested invoked moral obligations that stem from one's close personal relationships. However, it is possible that people will sometimes endorse moral demands that extend to distant others and that outweigh the normal demands to be partial towards

---

[9] Our suggestion that people do not associate the act of making an inference about another person's social status on the basis of sex with immorality is consistent with prior findings that people often do exactly that (e.g., Cao et al., 2019).

one's friends and family. For instance, if someone's friend has been accused of sexual assault, it is possible that observers will no longer prescribe giving that friend the benefit of the doubt. Instead, one's moral obligations to the potential victims may demand either being perfectly objective or perhaps even weighing the alleged victim's testimony more heavily than the friend's. As this example highlights, the moral reasons that sometimes justify motivated beliefs in our studies may be outweighed by reasons that confer different kinds of moral benefits (beyond the possible benefit of accurate reasoning discussed above).

Importantly, which moral norms are salient to observers, and indeed whether observers moralize mental states at all, differs across individuals, religious communities, and cultures (Graham et al., 2013). For instance, it may be that Christians are more likely to demand of others that they form respectful beliefs about parents (irrespective of the evidence) compared to Jews, because Christians (relative to Jews) are more likely to judge disrespectful attitudes as morally wrong and under the believer's control (Cohen & Rozin, 2001). Likewise, conservatives may be more likely to demand partial beliefs about friends or authority figures in light of their tendency to attach greater value to these moral norms (Graham, Haidt, & Nosek, 2009). Future work would benefit from rigorously documenting what beliefs people moralize, and in what situations people believe motivated reasoning will be beneficial.

A final moderating factor that we did not explore in our studies concerns the extent to which epistemic rationality may be valued differently across individuals. Some prior work has suggested that people vary in their intuitive commitment to objective, logical, and evidence-based reasoning (Pennycook et al., 2020; Ståhl et al., 2016). If these individual differences reflect the degree to which individuals intrinsically value epistemic rationality, then on average these individuals should be less sensitive to changes in the moral benefits of motivated reasoning. However, prior work measuring commitment to rationality has not investigated why certain individuals tend to value epistemic rationality more than others. This omission is important because there are potentially many reasons why someone may categorically reject bias – morally motivated or otherwise (Chignell, 2018). In our view, this remains an underexplored, but valuable, domain of research.

To summarize, it is likely the case that whether people prescribe motivated reasoning to others reflects a complex integration of (i) situational demands to be accurate, (ii) situational demands to adopt a morally beneficial belief (where more than one moral norm may come into play, and where such norms are likely to vary across culture), and (iii) individual differences in the extent to which people value accuracy and objectivity over other qualities of belief. Our results suggest that a large proportion of people feel the tug of moral benefits of belief in at least some common social scenarios, but much work remains to be done.

5.4. Prescribing motivated reasoning for moral or non-moral reasons

The studies above provide strong support for the claim that, in the lay ethics of belief, morality can justify motivated reasoning, therein raising the question of whether moral value is the only kind of non-evidential consideration that people explicitly endorse in belief formation. Specifically, it raises the question of whether people think others should adopt beliefs that are merely useful (but not morally beneficial). We found that moral considerations were treated as a better justification for motivated reasoning compared to mere preferences (Studies 2–3), but these studies do not definitively rule out the possibility that a large personal benefit could also justify motivated reasoning in the eyes of observers. 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.

6. 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 participants prescribed motivated reasoning to others, reported that morally beneficial beliefs require less evidence to be justified, and that, in some circumstances, a morally beneficial 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. These results also suggest another reason that motivated beliefs emerge and persist: People think that they ought to.

Declaration of Competing Interest

None.

Appendix A

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

The first two items measure endorsement of moral considerations. 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
		Moral Reason	Preference reason	Moral Reason	Preference reason	
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]	*

Table B1 (*continued*)

Vignette	Dependent Measure	Evidence-based belief		Optimistic belief		Predicted interaction
		Moral Reason	Preference reason	Moral Reason	Preference reason	
Friend	Moral Quality of Belief	4.41 (1.53)	4.34 (1.35)	6.15 (0.90)	5.29 (1.22)[b]	*
	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]	*
Marriage	Moral Quality of Belief	3.99 (1.29)	4.27 (1.23)[a]	5.62 (0.87)	4.64 (1.17)[b]	*
	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 preference 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 preference 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.

Table C1
Belief adoption text in each condition in Study 3.

Social distance condition	Belief condition	
	Optimistic belief	Evidence-based belief
Moral reason	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 *no chance* 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 *a 70% chance* that he and Maya will get divorced.
Preference reason	Although Patrick knows Brian very well, he has no obligations towards him or Maya. However, thinking about *anyone* 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 towards him or Maya. However, thinking about *anyone* getting divorced makes him sad, so he has a strong preference to believe that no one will ever get divorced.
	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 *no chance* 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 *a 70% chance* that Brian and Maya will get divorced.

Appendix D. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cognition.2020.104513.

References

Abelson, R. P. (1986). Beliefs Are Like Possessions. *Journal for the Theory of Social Behaviour, 16*, 223–250. https://doi.org/10.1111/j.1468-5914.1986.tb00078.x.

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. https://doi.org/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. https://doi.org/10.1007/bf01044377.

Bandura, A., & Locke, E. A. (2003). Negative self-efficacy and goal effects revisited. *Journal of Applied Psychology, 88*(1), 87–99. https://doi.org/10.1037/0021-9010.88.1.87.

Baron, J. (2008). *Thinking and deciding* (ed.). New York, NY: 4Cambridge University Press.

Baron, J. (2019). Actively open-minded thinking in politics. *Cognition, 188*, 8–18. https://doi.org/10.1016/j.cognition.2018.10.004.

Baron, J., & Spranca, M. (1997). Protected values. *Organizational Behavior and Human Decision Processes, 70*, 1–16. https://doi.org/10.1006/obhd.1997.2690.

Basu, R. (2019). Radical moral encroachment: The moral stakes of racist beliefs. *Philosophical Issues, 29*(1), 9–23. https://doi.org/10.1111/phis.12137.

Baumeister, R. F. (1989). The optimal margin of illusion. *Journal of Social and Clinical Psychology, 8*(2), 176–189. https://doi.org/10.1521/jscp.1989.8.2.176.

Bolinger, R. J. (2018). The rational impermissibility of accepting (some) racial generalizations. *Synthese,* 1–17. https://doi.org/10.1007/s11229-018-1809-5.

Bolinger, R. J. (2020). The varieties of moral encroachment. *Philosophical Perspectives.* https://doi.org/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. https://doi.org/10.1177/0956797618805750.

Carver, C. S., & Scheier, M. F. (2014). Dispositional optimism. *Trends in Cognitive Science, 18*(6), 293–299. https://doi.org/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.* Prometheus: Amherst, MA.

Cohen, A. B., & Rozin, P. (2001). Religion and the morality of mentality. *Journal of Personality and Social Psychology, 81*(4), 697–710. https://doi.org/10.1037/0022-3514.81.4.697.

Cohen, L. J. (1981). Can human irrationality be experimentally demonstrated. *Behavioral and Brain Sciences, 4*(3), 317–331. https://doi.org/10.1017/s0140525x00009092.

Chignell, A. (2018). The Ethics of Belief. The Stanford Encyclopedia of Philosophy). Metaphysics Research Lab, Stanford University. Retrieved from https://plato.stanford.edu/archives/spr2017/entries/ethics-belief/.

Cusimano, C., & Goodwin, G. P. (2019). Lay beliefs about the controllability of everyday mental states. *Journal of Experimental Psychology: General, 148*(10), 1701–1732. https://doi.org/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.* https://doi.org/10.1037/pspa0000198.

D'Andrade, R. (1987). A folk model of the mind. In D. Holland, & N. Quinn (Eds.), *Cultural models in language and thought* (pp. 112–148). Cambridge, England: Cambridge University Press.

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. https://doi.org/10.1037/0022-3514.63.4.568.

Douglas, H. (2016). Values in science. In P. Humphreys (Ed.), *The Oxford handbook of philosophy of science* (pp. 609–630). New York: Oxford University Press.

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. https://doi.org/10.1177/0146167204271570.

Epley, N., & Gilovich, T. (2016). The mechanics of motivated reasoning. *Journal of Economic Perspectives, 30*(3), 133–140. https://doi.org/10.1257/jep.30.3.133.

Epley, N., & Gilovich, T. (2016). The mechanics of motivated reasoning. *Journal of Economic Perspectives, 30*(3), 133–140. https://doi.org/10.1257/jep.30.3.133.

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. https://doi.org/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. https://doi.org/10.1007/978-3-319-95993-1_11.

Gervais, W. M. (2014). Everything is permitted? People intuitively judge immorality as representative of atheists. *PLoS One, 9*(4), e92302.

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. https://doi.org/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. https://doi.org/10.1037/a0034726.

Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, S. P., & Ditto, P. H. (2013). Moral foundations theory. In *Advances in experimental social psychology* (pp. 55–130). Elsevier. https://doi.org/10.1016/b978-0-12-407236-7.00002-4.

Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology, 96*(5), 1029–1046. https://doi.org/10.1037/a0015141.

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. https://doi.org/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.

Heiphetz, L., Spelke, E. S., & Young, L. L. (2015). In the name of god: How children and adults judge agents who act for religious versus secular reasons. *Cognition, 144*, 134–149.

Holden, G. (1991). The relationship of self-efficacy appraisals to subsequent health related outcomes: A meta-analysis. *Social Work in Health Care, 16*(1), 53–93. https://doi.org/10.1300/j010v16n01_05.

James, W. (1937). *The will to believe, and other essays in popular philosophy.* Green and Co: Longmans.

Keller, S. (2004). Friendship and belief. *Philosophical Papers, 33*(3), 329–351. https://doi.org/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. https://doi.org/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. https://doi.org/10.1017/s0140525x00041157.

Kruglanski, A. W. (2004). *The psychology of closed mindedness.* Psychology Press.

Kruglanski, A. W., & Freund, T. (1983). The freezing and unfreezing of lay-inferences: Effects on impressional primacy, ethnic stereotyping, and numerical anchoring. *Journal of Experimental Social Psychology, 19*, 448–468. https://doi.org/10.1016/0022-1031(83)90022-7.

Kruglanski, A. W., & Webster, D. M. (1996). Motivated closing of the mind: "seizing" and "freezing". *Psychological Review, 103*, 263–283. https://doi.org/10.1037/0033-295X.103.2.263.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin, 108*, 480–498. https://doi.org/10.1037/0033-2909.108.3.480.

de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a web browser. *Behavior Research Methods, 47*(1), 1–12. https://doi.org/10.3758/s13428-014-0458-y.

Lobato, E. J. C., Tabatabaeian, S., Fleming, M., Sulzmann, S., & Holbrook, C. (2019). Religiosity predicts evidentiary standards. *Social Psychological and Personality Science.* https://doi.org/10.1177/1948550619869613, 194855061986961.

Locke, J. (1690). *An essay concerning human understanding.* Oxford: Clarendon.

Malle, B. F. (2004). *How the mind explains behavior.* Cambridge, MA: MIT Press.

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. https://doi.org/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. https://doi.org/10.1016/0030-5073(79)90027-8.

McManus, R. M., Kleiman-Weiner, M., & Young, L. (2020). What we owe to family: The impact of special obligations on moral judgment. *Psychological Science, 31*(3), 227–242. https://doi.org/10.1177/0956797619900321.

McPhetres, J., & Zuckerman, M. (2017). Religious people endorse different standards of evidence when evaluating religious versus scientific claims. *Social Social and Personality Science, 8*(7), 836–842. https://doi.org/10.1177/1948550617691098.

Metz, S. E., Weisberg, D. S., & Weisberg, M. (2018). Non-scientific criteria for belief sustain counter-scientific beliefs. *Cognitive Science, 42*, 1477–1503. https://doi.org/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. https://doi.org/10.1111/j.1745-6924.2009.01142.x.

Molnar, A., & Loewenstein, G. F. (2020). The Othello effect: People are more disturbed by others' wrong beliefs than by different beliefs. *SSRN Electronic Journal.* https://doi.org/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. https://doi.org/10.1080/02701367.2000.10608908.

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. https://doi.org/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. https://doi.org/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. https://doi.org/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. https://doi.org/10.2307/41330856.

Pascal, B. (1852). *Pensées*. Magdeleine: Dezobry et E.

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. (2020). On the belief that beliefs should change according to evidence: Implications for conspiratorial, moral, paranormal, political, religious, and science beliefs. *Judgment and Decision making, 15*, 476–498.

Pinillos, Á. (2012). Knowledge, experiments and practical interests. In J. Brown, & M. Gerken (Eds.), *Knowledge ascriptions* (pp. 192–219). Oxford: Oxford University Press. https://doi.org/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. https://doi.org/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. https://doi.org/10.1177/0146167202286008.

Robinson, R. J., Keltner, D., Ward, A., & Ross, L. (1995). Actual versus assumed differences in construal:" naive realism" in intergroup perception and conflict. *Journal of Personality and Social Psychology, 68*(3), 404. https://doi.org/10.1037/0022-3514.68.3.404.

Rogers, T., Moore, D. A., & Norton, M. I. (2017). The belief in a favorable future. *Psychological Science, 28*(9), 1290–1301. https://doi.org/10.1177/0956797617706706.

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. https://doi.org/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. https://doi.org/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. https://doi.org/10.1037/0022-3514.91.1.143.

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. https://doi.org/10.1037/0096-3445.127.2.161.

Starmans, C., & Friedman, O. (2012). The folk conception of knowledge. *Cognition, 124*(3), 272–283.

Strohminger, N., Knobe, J., & Newman, G. (2017). The true self: A psychological concept distinct from the self. *Perspectives on Psychological Science, 12*(4), 551–560. https://doi.org/10.1177/1745691616689495.

Stroud, S. (2006). Epistemic partiality in friendship. *Ethics, 116*, 498–524. https://doi.org/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. https://doi.org/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. https://doi.org/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. https://doi.org/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. https://doi.org/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. https://doi.org/10.2307/1339610.

Wegener, D. T., Silva, P., Petty, R. E., & Garcia-Marques, T. (2012). The metacognition of bias regulation. In P. Brïnol, & K. De Marree (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. https://doi.org/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. https://doi.org/10.1037/a0028857.