It's The Thoughts That Count:

Mindreading Moral Principles Underlies Judgments of Moral Praise

Clayton R. Critcher

University of California, Berkeley

Erik G. Helzer

Wake Forest University

David Tannenbaum

University of California, Los Angeles

David A. Pizarro

Cornell University

Authors' Note

Clayton R. Critcher, Haas School of Business, University of California, Berkeley; Erik G. Helzer, Department of Psychology, Wake Forest University; David Tannenbaum, Anderson School of Management, University of California, Los Angeles; David A. Pizarro, Department of Psychology, Cornell University.

Correspondence concerning this article should be sent to Clayton R. Critcher at ClaytonCritcher@haas.berkeley.edu.

SUBMITTED ON: November 22, 2012

Abstract

In order to make sense of an agent's morally-relevant actions, motives, and character, people engage in *mindreading*—attempts at inferring an agent's mental states. The present paper proposes a *mindreading moral principles* (MMP) account, which outlines a new route by which people mindread mental contents in order to evaluate agents' character. People attempt to mindread agents' moral principles—the mental antecedents that precede a morally-motivated action. Perceivers then praise agents to the extent they are assumed to appreciate a principle that would morally motivate their action. Studies 1a-1c presented participants with moral dilemmas that offered a moral agent a choice between two competing courses of action. How much praise the agents received for making either choice was mediated by the extent to which they were assumed to appreciate the moral principle that would justify that choice. Furthermore, the MMP account predicts that certain non-decision-related cues (i.e., features of the agent or the agent's context) should change how the agent is evaluated for different actions because they change what moral principles the agent likely appreciates. Studies 2-4 varied features of the decision-maker (e.g., capacity for rational thought) or decision context (e.g., time duress) that enhanced or diminished the perceived likelihood that an agent would appreciate different moral principles. This changed how much praise the agent received for each action. Implications of the MMP account for previous and future research in moral psychology in particular, and social judgment more generally, are discussed.

KEYWORDS: moral judgment, person perception, mindreading, theory of mind

It's the Thoughts That Count:

Mindreading Moral Principles Underlies Judgments of Moral Praise

How do people decide whether a particular actor is deserving of moral praise? Traditional theories of moral responsibility hold that an individual is responsible for an action when, among other things, the individual desires, intends, and causes an outcome. Praise judgments go beyond judgments of whether the agent should be credited (in the causal sense) for his behavior to include whether he should receive *moral* credit (in the characterological sense) as well. Because of this, judgments of praise should be sensitive to the motivations that underlie behavior (Fedotova, Fincher, Goodwin, & Rozin, 2011; Gray, Young, & Waytz, 2012). For a behavior to warrant moral praise, the agent must have not only intended and brought about the action and its consequence, he must have also performed the act for reasons that are themselves praiseworthy.

For instance, suppose that during the holiday season Steve donates to a charity that is sponsored by his workplace. How much praise does Steve deserve for his behavior? The answer is probably "it depends." While his coworkers may know that Steve intended, desired, and caused this behavior (donating to charity), this information may not be enough to convince them that Steve deserves praise. To be sure, information about whether an action or outcome was intended or not is essential (Baird & Astongton, 2004; Cushman, 2008; Karniol, 1978; Knobe, 2004; Miller et al., 2010; Piaget, 1932; Young et al., 2007; Young & Saxe, 2008; Yuill, 1984; Yuill & Perner, 1988), but people typically want to know more; they want to understand the motives or reasons that went into this decision (Critcher, Inbar, & Pizarro, in press; Reeder, 2009; Reeder et al., 2004; see also Monroe & Reeder, 2011). Was he truly motivated by a charitable sentiment (and thus deserving of praise), or was he trying to impress the administration, ease a guilty conscience, or merely appear prosocial? Note that all four of these

possibilities are equivalent in whether Steve did or did not intend the behavior, but each has different implications for the question of moral praise.

Mindreading Informs Moral Praise

This paper is concerned with how perceivers bridge the inferential gap between intentional actions and judgments of praise. Of course, identifying people's underlying reasons and motives for their intentional actions is a tricky task. Unlike behavior, the contents of another's mind are not directly observable (e.g., Pronin, 2008). Although sometimes moral agents make their reasons for acting explicitly clear ("I'm donating to help fight muscular dystrophy because one of the donors is going to win a new car!"), such explicit statements are likely rare. As such people must frequently engage in *mindreading* (Reeder, 2009) to make sense of why a person behaved as he or she did. As Reeder put it, "Intentional acts open a window to theory of mind... [in which] the perceiver is looking for a coherent narrative that explains the known facts" (p. 3-4). People must engage in mindreading to "fill in the blanks" to infer the underlying reasons and motives for any given action.

Given the difficulty of the problem, how exactly do people get inside the head of others to develop a coherent explanation for why they behave as they do? A look to the literature suggests several "questions" people, in effect, ask before deciding whether people acted for praiseworthy reasons. After reviewing these previously-demonstrated approaches to mindreading, we put forth a new question that guides the mindreading process. We suggest that people attempt to mindread moral principles—those mental precursors that reflect the moral motivation behind an action—to determine whether moral praise is warranted. As we will develop, this outlines a qualitatively new way in which mindreading unfolds and also opens the door to a new class of features that affect which agents, in which circumstances, receive moral

praise. To place our proposal in context, we first consider how people have been shown to mindread reasons for intentional action, as demonstrated in past work:

"Did the agent make explicit his or her selfish motives?" Mindreading is simplest when an agent makes it quite clear that his moral action was a product of non-praiseworthy motives. Consider the *switch version* of the trolley problem, in which a moral agent must decide whether to save five individuals by diverting a runaway trolley onto another track, thereby killing one person. Although most people judge the utilitarian solution to this dilemma favorably (Hauser, 2006), their judgments of permissibility are greatly reduced if the decision maker first says he "doesn't give a damn about saving those five men, but this is [his] chance to kill that bastard!" (Mikhail, 2002, 2011). Similarly, work on moral side-effects indicates that people do not readily dole out praise for good consequences (such as helping the environment) if the actor makes explicitly clear that he was interested only in profits and not in helping the environment (Sripada, 2012; see Knobe, 2003). In these examples, mindreading is facilitated by the agents' explicit statements that their actions were driven by amoral (and sometimes immoral) reasons.

"Even with no explicit statement, is there reason to suspect selfish motives?" In other cases, a moral actor does not make explicit his or her reasons for acting, but social perceivers use information in the situation to identify ulterior motives underlying the behavior. People seem to be particularly attuned to features of a situation that explain why seemingly good behaviors should actually be reinterpreted as self-serving (Fein, 1996). Merely recognizing that ulterior motives exist, even without confirmation that they actually drove behavior, seems to be enough to diminish praise. In one study (Critcher & Dunning, 2011), participants considered philanthropy that was heralded by *Slate Magazine* as among the most generous acts of the year. As participants spent more time mindreading why these actions may have occurred, they

generated more and more reasons why these seemingly-selfless acts may have been performed selfishly. As a result, their impressions of these philanthropists deteriorated.

"Did the person have selfish reasons to act in a different way?" Instead of focusing directly on the incentives to act as the person did, perceivers can instead ask whether there were non-praiseworthy reasons to act in a different way. That is, if one still does good when there are incentives to do bad, then one must have very strong motives to do good (and thus deserve more praise). For example, Reeder and Spores (1983) found that a man who placed money in a collection jar received more praise when his romantic interest had encouraged him to steal the charity money than when she had encouraged him to make the donation. In the former case, the augmentation principle implies that he had a particularly strong interest in helping charity (see also Kelley, 1971; Kruglanski, Schwartz, Maides, & Hamel, 1978). This involves an indirect form of reasoning in which one makes inferences about principled intent not by directly inferring its presence, but by inferring its presence given what behavior was forgone. The perceiver infers strong moral motives given that the temptations to be resisted.

Mindreading Moral Principles: A New Form of Mindreading

What is common to the above examples is that praise judgments will be affected depending on whether selfish motives are assumed to point toward or away from the chosen course of action. Whether through explicit statements of selfishness, suspicions aroused by what a selfless-seeming behavior actually reflects, or attunement to immoral forces that are succumbed to or ignored, perceivers get a sense for how actions were (or were not) a product of candidate selfish motives. Our account suggests a different, and complementary, approach to mindreading. Instead of considering only whether and how the *wrong* reasons might have or would have guided behavior, we propose people also try to more directly answer whether the

right reasons were behind an action. Furthermore, we detail how this occurs.

We suggest that in determining praise, people ask what *moral principles* the agent likely appreciated prior to acting. In so doing, the perceiver attempts to directly confront whether the agent's action was done for a morally-praiseworthy reason. We define moral principles as those cognitive or emotional precursors to moral behaviors that offer a morally good justification for a specific behavior. For example, a flash of concern for the welfare of vulnerable children (moral principle) may precede Steve's decision to donate money to a children's charity (moral behavior). Alternatively, a conviction that it would be preferable to hurt one to save many more (moral principle) may precede a decision to fire one worker if it means that the jobs of ten others can be saved (moral behavior). Our argument is that perceivers look for evidence that an agent had principles like these in mind when deciding to act, and then base their praise judgments accordingly.

"Did the person appreciate the relevant moral principle when acting?" Our argument is based on a simple point that has non-obvious implications for how moral judgments unfold: For a certain moral principle to be seen as the reason for a person's action, the person should be assumed to appreciate that principle when deciding to act (Malle, Knobe, O'Laughlin, Pearce, & Nelson, 2000). Stated differently, if a person seems unlikely to appreciate a moral principle, it is unlikely to be the basis for subsequent action. If people attempt to mindread moral principles in this way, it would be particularly noteworthy because it would suggest the importance of certain *extradecisional* features—cues that are not specific to what moral decision a person made (e.g., whether or not Steve donates to charity), but that reside in the decision context (e.g., whether Steve makes his decision under time duress, what Steve's visual perspective was when deciding). That is, features that themselves may have little or nothing to do with a person's moral character,

but that are assumed to make different moral principles more or less focal in a person's mind, may change how much praise the agent gets for each action. For example, the moral principles people may be assumed to appreciate under rushed conditions may differ from those they are assumed to appreciate under less time constraints; we argue this should change how much praise agents receive for actions under these two circumstances. In short, praise may be conditional on believing the person appreciated the relevant moral principle when acting.

Overview of the Present Studies

We conducted six studies, using three different moral dilemmas, to test whether people mindread moral principles in determining praise. In each dilemma, a moral agent is confronted with a choice between pursuing two actions. Each scenario pits a utilitarian decision (maximizing total welfare that also causes harm in the process) against a deontological decision (refusing to violate a moral rule). We employed dilemmas of this sort because both courses of action have clear matching principles—deontology-backed aversions to direct harm and utilitarianism-derived justifications for promoting the greater good (Nichols & Mallon, 2006)—that facilitate a test of whether people rely on mindread moral principles in forming moral evaluations.[1] Our model, applied to this context, is depicted in Figure 1.

The goal of Studies 1a-1c is to test whether MMP in part explains judgments of moral praise. In this way, Studies 1a-1c take the form of typical moral judgment studies, offering information about an agent who is confronting a moral dilemma, with no additional information as to what moral principles the particular agent is likely experiencing. In order to test our MMP account, we assessed whether participants' praise judgments for different courses of action were mediated by their beliefs about the agent's appreciation of relevant moral principles at the time of acting. If so, this would provide evidence consistent with the idea that people essentially ask,

"Did the agent appreciate the relevant moral principle when acting?", in determining how much moral praise to grant. Such results would be consistent with this expanded perspective of mindreading, but would not yet offer the definitive causal test that such mindreading is definitely influencing moral judgments.

Studies 2-4 offer a causal test of our model. The studies also demonstrate how our MMP approach makes novel predictions about how certain extradecisional features (extraneous to the decision itself) change praise judgments for different courses of action. In particular, these studies manipulated uncontrollable features of the agent or the agent's situation that were assumed to affect the moral principles an agent would likely appreciate: whether the agent was said to have a brain injury that affected his ability to feel or to reason (Study 2), whether the decision was made under time constraints (Study 3), and who was focal in the agent's visual field (Study 4). If moral judgments are influenced by mindreading moral principles, perceivers should offer praise to the extent they believe an agent is experiencing the moral principle that serves as the basis for their action. The MMP account predicts that these cues will be used to infer whether the agent appreciated the moral principle that would justify their action. As such, we predict that these extradecisional factors will moderate the relationship between the agent's behavior and judgments of praise. Without the present theoretical perspective, it would be difficult to predict or explain these patterns.

Studies 1a, 1b, and 1c

According to our mindreading moral principles account, people praise agents for actions to the extent that they could have plausibly followed from a matching moral principle. In light of the considerable body of previous research that has focused on actions being right or wrong (even when no cues about motives or moral principles are provided), readers may assume that

mindreading is only engaged when mental state information is readily available or when, as discussed above, additional information provides a clue as to what selfish motives may have been in play. Consider the finding that people think it is more praiseworthy to save a sick child in need of a costly liver transplant versus letting the child die, even if saving the child means foregoing resources that could have been used to save even more lives in the future (Critcher et al., in press; Tetlock et al., 2000). The MMP account stipulates that even in this case, when no information about either moral principles or selfish motivations is explicitly provided, people will still attempt to infer what moral principles the agent was likely appreciating as a guide to allocating praise.

In Studies 1a-1c, we use three moral dilemmas, two of which have been used in previous research. Before learning how the moral agent resolved each dilemma, participants indicated to what extent they thought the agent would spontaneously experience each of two moral principles (one utilitarian, one deontological). Each principle forms the basis for a different course of action (one utilitarian, one deontological). Then, participants were randomly assigned to learn that the moral agent actually chose the utilitarian or deontological course of action. Finally, participants indicated their moral evaluations of the agent.

In each study, we predicted that the amount of praise an agent would receive for one action or the other (the between-subjects manipulation) would be mediated by the degree to which the agent was assumed to appreciate the matching moral principle. That is, we predicted that the extent to which a moral agent would receive praise for a deontological (or utilitarian) action would be mediated by the extent to which they were assumed to have appreciated the accompanying deontological (or utilitarian) principle.[2] Because our account suggests that only the matching moral principle is informative for moral evaluation, we predicted that the extent to

which a moral agent was assumed to appreciate the competing moral principle (the principle that would justify the foregone choice) would not mediate the effect of action on moral evaluation.

Method

Participants and design. Participants in all three studies were recruited from Amazon.com's Mechanical Turk. Ninety-four participants were recruited in Study 1a; 95, in Study 1b; 107, in Study 1c. In each study participants were randomly assigned to one of two decision conditions: *utilitarian* or *deontological*.

Procedure and materials. In each study, participants considered a different moral dilemma whereby a moral agent was confronted with two options, one justified by utilitarian calculus and one supported by deontological reasoning (i.e., maximizing lives saved vs. adhering to moral rules against harming). After learning the details of the dilemma, participants were asked to indicate the degree to which each principle—the relevant utilitarian and deontological principles—would be salient to the agent in the situation. At that point, participants learned that the agent made either the deontological or the utilitarian decision. Finally, participants completed a moral evaluation scale.

***Moral dilemmas*.** In Study 1a, participants read a moral dilemma about Jewish townspeople hiding in a secret basement while Nazi soldiers searched the town above (e.g, Greene, Nystrom, Engell, Darley, & Cohen, 2004). The townspeople were maintaining careful quiet, for the Nazis would kill anyone they discovered. Suddenly, a small baby in the arms of a townsperson, Jack, began to bawl. Left unabated, the crying would attract the attention of the Nazis, which would result in the certain death of all of the townspeople. Only by smothering the child would the crying stop and the townspeople be saved, though this would kill the child (utilitarian decision); otherwise, Jack could let the child continue to cry, though the Nazis would

then find and kill the hidden townspeople (deontological decision).

In Study 1b, participants read a modified version of Tetlock et al.'s (2000) "sick Johnny" moral dilemma. A hospital director, Robert, must decide whether to spend $3 million of the hospital's limited resources to save the life of a sick five-year-old named Johnny. Spending the money to save Johnny would prohibit the hospital from updating hospital infrastructure, updates that could be used to save many future lives. Thus, Robert must choose whether to let Johnny die in order to save more lives in the future (utilitarian decision) or spend the money and thereby save the life of Johnny (deontological decision: what Tetlock et al. [2000] called a "taboo tradeoff").

Study 1c introduced a new dilemma not used in previous research. Participants read about a high-level military commander working to root out Al Qaeda terrorist cells in Afghanistan. Intelligence from a mole had led the commander, Michael, to a rural inn on the Ukraine-Poland border. There, a meeting of top Al Qaeda leaders planning a 9/11-style attack was to take place. Several of these men were among the FBI's "Most Wanted Terrorists." The night of the meeting, Michael looked down at the inn from the surrounding mountains and could clearly see the expected Al Qaeda leaders enter the inn. He also saw their translator, an innocent man kidnapped by the terrorists and forced to work for them against his will. Michael had to decide whether to recommend an airstrike, which would kill all of those present, the terrorists and the innocent (utilitarian decision). Justifying why a pure utilitarian would likely support the strike, we added that "if a strike is not ordered now, it is doubtful that one will occur in time to stop the 9/11-style attack." Alternatively, Michael could decide not to order the strike (deontological decision).

These dilemmas and choices, which are also used in Studies 2-4, are summarized in Table 1 for quick reference.

Moral principles. Participants then completed measures that asked them to indicate to what extent the agent likely "appreciated, experienced, or possessed" each of two relevant moral principles. The measure specified that "different moral principles may be at the forefront of [the agent's] mind" and participants should indicate "to what extent you believe he is experiencing each sentiment as he is confronted with this situation." Participants responded on nine-point scales anchored at 1 (*not at all)* and 9 (*is experiencing strongly*) for both the utilitarian and deontological principle. For each study, the wording of the principle was modified to be scenario-appropriate. In Study 1a, participants indicated to what extent Jack experienced two principles: "It is morally wrong to actively kill a child" (deontological), and "By killing this child, I could save everyone" (utilitarian). In Study 1b, participants indicated to what extent Robert experienced each principle: "It is morally wrong or troubling to let a child die" (deontological), and "By letting the child die, the hospital could actually save money which would allow it to ultimately save many more lives." (utilitarian). In Study 1c, participants rated to what extent Michael was likely experiencing each principle: "It is morally wrong to kill innocent civilians regardless of the circumstances" (deontological), and "It is morally right to stop the terrorists from killing thousands of people, even if it means killing an innocent person in order to stop the worse tragedy" (utilitarian).[3] We note that we followed precedent in assuming that deontological principles would be experienced as morally-laden rules proscribing certain actions instead of as a conscious appreciation of philosophical (e.g., Kantian) logic (Broeders et al., 2011; Greene, 2007; Greene, Morelli, Lowenberg, Nystrom, & Cohen, 2008).[4]

***Moral evaluation*.** Participants then offered moral evaluations of the agent. Participants responded to five items on 8-point scales. They indicated to what extent the agent: is a bad versus good person, has a bad versus good conscience, is or is not "in the wrong," has

blameworthy versus praiseworthy character, and is in general a moral versus immoral person. After reverse-scoring negatively-worded items, we averaged the judgments into a *praise composite* such that higher numbers reflected greater praise for the agent (Study 1a: $\alpha = .86$; Study 1b: $\alpha = .93$; Study 1c: $\alpha = .78$).

Results and Discussion

Although the three moral dilemmas may (and did) differ in whether the deontological or the utilitarian decision leads to more positive moral evaluations, the MMP account predicts that these differences in moral evaluation should stem from the same mediator. More specifically, one action should receive more praise than another to the extent that the agent is assumed to appreciate one moral principle more than the other. In both Studies 1a and 1b, participants offered more praise to the actor who performed the deontological action. In Study 1c, participants praised the utilitarian actor more (see Table 2, Model 1). More specifically, in Study 1a, Jack was praised more when he made the deontology-backed decision to spare the child's life ($M = 5.87$, $SD = 1.55$) than the utilitarian decision to save more lives by killing the child ($M = 4.63$, $SD = 1.35$), $t(92) = 4.14$, $p < .001$. In Study 1b, Robert was offered more praise when he made the deontology-backed decision to save the child ($M = 6.47$, $SD = 1.26$) instead of the utilitarian decision to let the child die ($M = 3.97$, $SD = 1.72$), $t(93) = 8.11$, $p < .001$. In Study 1c, participants judged Michael as more moral when he made the utilitarian decision to order the strike on the inn, thereby killing an innocent civilian in the process ($M = 5.64$, $SD = 1.32$), as opposed to when he made the deontology-backed decision not to order the strike ($M = 5.13$, $SD = 1.24$), $t(105) = 2.05$, $p = .04$.

If agents are praised to the extent that they are assumed to experience the matching moral principles, then we should see a similar pattern of results to that observed above on the measures

of principle appreciation. And indeed, participants thought that the agents in Studies 1a and 1b would more strongly experience the deontological principle than the utilitarian principle, and participants in Study 1c thought that the agent would more strongly appreciate the utilitarian than the deontological principle. More specifically, participants thought that Jack in Study 1a would be experiencing the deontological principle ("It is morally wrong to actively kill a child") more strongly (M = 8.29, SD = 1.13) than the utilitarian principle (M = 7.08, SD = 2.32), paired $t(94) = 4.94$, $p < .001$. Participants thought that Robert in Study 1b would more strongly experience the deontological principle ("It is morally wrong or troubling to let a child die": M = 8.04, SD = 1.29) than the utilitarian principle (M = 6.06, SD = 2.50), paired $t(96) = 7.24$, $p < .001$. Study 1c participants thought that Michael would more strongly experience the utilitarian principle ("It is morally right to stop the terrorists from killing thousands of people…": M = 7.47, SD =1.84) than the deontology-backed principle (M = 6.16, SD = 2.36), paired $t(107) = 4.56$, $p < .001$.

To connect moral evaluations with principle appreciation, we then tested whether appreciation of the matching principle mediated the effect of decision on moral evaluation. We created two new variables: *appreciation of matching principle* and *appreciation of competing principle*. When the agent made the utilitarian [deontological] decision, the matching principle was the utilitarian [deontological] principle. The other principle was labeled as the competing principle. Thus, when the agent made the utilitarian [deontological] decision, the competing principle was the deontological [utilitarian] principle. These two variables will be relevant for testing our preferred and an artifactual mediation model, respectively:

Mindreading moral principles mediation model. To determine whether the agent was praised to the extent that he was appreciating the matching principle, we regressed the praise composite on Jack' decision and the matching principle appreciation variable. In all three

studies, assumed appreciation of the matching moral principle mediated the effect of decision on moral evaluation (see Table 2, Model 2). Using Preacher and Hayes's (2008) bootstrapping technique with 10,000 resamples, we found the 95% confidence interval of this indirect effect did not include 0 for Study 1a, [-.3256, -.0580], Study 1b, [-.4770, -.0991],[5] and Study 1c, [.0078, .1430]. Thus, although we provided no direct information about the agent's principles or motives, participants nonetheless made inferences about the principles that were salient to the agents, and allocated praise to the extent that they found evidence that the agent's decision emerged from an appreciated principle. This shows that mindreading reasons may not only affect moral judgments by leaning on explicitly stated or inferred selfish motives. Instead, perceivers seem to ask themselves whether the matching moral principle was likely in place.

An alternate mediation model. For each study, we also tested an alternative (and less interesting) explanation for our results. By this artifactual account, people do not mindread moral principles in deciding whether to offer praise; instead, people judge one or the other behavior as the right course of action, and then assume that the agent would experience moral principles that encouraged this good behavior but not the principles that would encourage the "incorrect" course of action. In other words, this account suggests participants treat the principle appreciation measures as a question of whether a person would be inclined or not inclined to take that particular course of action. If people simply make inferences about an agent's moral principles to match what they see as the more and less preferable course of action, then participants' moral judgment should not merely be mediated by the matching moral principle, but also (negatively) by the *competing* moral principle. In other words, whereas our MMP account predicts that the matching moral principle should be the sole mediator (because people are merely interested in whether an action could have been motivated by the matching moral principle), this alternative

account predicts that both the matching *and* the competing moral principles should be mediators.

To better appreciate the intuition behind this artifactual interpretation, consider Study 1a participants who learned that Jack smothered the child to death. Jack was not much praised for this decision. By our account, Jack receives little praise for this utilitarian action because participants do not think the utilitarian principle (the matching moral principle) would be focal in Jack's mind in this context. By the alternative account, Jack is also not praised because participants think that any reasonable person in this context would have accessible a very strong deontological principle to avoid smothering the child (the competing moral principle). The two accounts then are most easily distinguished by looking at this competing moral principle, because the accounts differ in whether they suggest that it too should mediate the effects. Thus, we reran our mediation models to include both the matching and the competing moral principles to see if the matching moral principle would continue to mediate the effect, but that the competing moral principle would fail to. As shown in Table 2, Model 3, for no study was there a significant negative effect of competing moral principle on moral evaluation.[6] In fact, the effect in Study 1a even trended in the opposite direction. Furthermore, Peacher and Hayes's (2008) bootstrapping procedure found that the 95% confidence interval on the indirect effect through competing moral principle included 0 for Study 1a, [-.0475, .2932], Study 1b, [-.3429, .0157], and Study 1c, [-.0340, .2134].

Although Studies 1a-1c provide evidence consistent with the possibility that people assign praise by mindreading moral principles, and we attempted to rule out an artifactual interpretation of our results, the correlational approach does not establish the necessity of our MMP model. The remaining studies build on these findings by using experimental manipulations that permit causal tests of our hypotheses. Furthermore, the subsequent studies show how the

MMP approach points to heretofore unidentified influences on moral judgment. In particular, if the moral principles a moral agent is assumed to appreciate help to define what actions are praiseworthy versus not, it follows that extradecisional influences on principle appreciation should be able to change which actions do or do not receive praise. In the remaining three studies, we identify three cues that participants believed would change the principles that a person would appreciate. By changing the experienced principles, they should change how much praise different actions would merit.

Study 2: Emotional or Rational Deficits

In Study 2, we varied an extradecisional feature of a moral agent—the nature of a brain deficit he was said to have. For some participants, the agent was said to have a rational deficit, meaning the agent was only able to rely on emotional impulses to guide his sense of right and wrong. For other participants, they learned the agent had an emotional deficit, meaning the agent could only rely on calculations to differentiate right from wrong. We varied the nature of these deficits because they permitted a clean test of our theoretical account (not because of an interest in directly generalizing our findings to moral judgments of those with neural disorders). In particular, we expected that participants would believe that the moral agent would experience different moral principles depending on whether he had an emotional or rational deficit.

Participants in Study 2 considered the Nazi-baby dilemma used in Study 1a, in which a Jewish townsperson must decide whether to actively kill a baby whose crying will alert Nazi soldiers to the hidden location of Jewish townspeople. In high-conflict personal moral dilemmas of this variety, people tend to quickly appreciate or experience a negative affect-backed deontological principle (e.g., “Killing a child is wrong!...”; Koenigs et al., 2007; Nichols & Mallon, 2006; Valdesolo & DeSteno, 2006), that with more time and reasoned thought is

supplemented or replaced by a utilitarian principle (e.g., "…but by killing now, I could save the lives of so many of these people!"; Greene, 2009; Greene et al., 2001; Greene, Morelli, Lowenberg, Nystrom, & Cohen, 2008; Kahane, Wiech, Shackel, Farias, Savulescu, & Tracey, in press). Our perspective remains agnostic as to whether deontological and utilitarian principles actually or always map onto these properties (see Baron, 2011; Kahane et al., in press); it only matters that in many moral dilemmas (including the ones we used), people *intuit* these properties. If our participants have the intuition that the deontological principles would be more easily appreciated by the emotion-intact agent, and that the utilitarian principle would be more easily appreciated by the reason-intact agent, then our MMP account predicts that the two agents should be differentially praised for deciding to kill (utilitarian) or not kill (deontological). By the MMP account, differences in how much praise each agent receives for each decision will be mediated by assumed appreciation of the matching moral principle.

Method

Participants and design. Four hundred fifty-eight Cornell University undergraduates were randomly assigned to one of four conditions in a 2 (intact faculty: emotion or reason) X 2 (decision: utilitarian or deontological) full-factorial design. Participants received course credit for their participation.

Procedure. As in Study 1a, participants read the moral dilemma about Jewish townspeople hiding from Nazi soldiers in a basement. But this time, we included a manipulation that was designed to impact the moral principles that the moral agent, Jack, could reasonably appreciate. The manipulation did not change any details about the moral decision itself.

Those in the *reason intact* condition were told that Jack was "missing the part of his brain that allows him to have strong emotional impulses that signal what is morally right or wrong.

Instead, all he can do is use rational calculation to calculate what is the right thing to do." In this way, it was noted Jack was "like a computer." Those in the *emotion intact* condition were told that Jack's deficit kept him from "engaging in rational calculations to arrive at his decision. Instead, all he can do is use his strong emotional impulses that signal what is morally right or wrong." In both conditions it was stressed that Jack was simply "born this way."[7]

Participants answered two questions about whether Jack would appreciate the two relevant moral principles, that "killing [the child] is wrong" and that "by killing the child he could save more people." Both responses were made on 8-point scales anchored at 1 (*not at all*) and 8 (*completely*).

Participants then learned that Jack let the baby continue to cry (deontological decision) or that Jack smothered the baby, thereby stopping the baby from crying, but killing him in the process (utilitarian decision). Finally, participants made judgments on 8-point scales about whether Jack deserved moral praise (α = .82), indicating whether Jack: was a good person, should be praised, had a good moral conscience, had blameworthy moral character (reverse-scored), was an immoral person (reverse-scored), and was "in the wrong" (reverse-scored).

Results

Participants' intuitions about what moral principles Jack would experience depended on which faculty was left intact. A 2 (intact faculty: emotion-intact or reason-intact) X 2 (principle: utilitarian or deontological) mixed-model ANOVA, with only the second factor measured within-subjects, showed that principle appreciation depended on the type of neurological deficit, $F(1, 456) = 238.18$, $p < .001$. (The means for the principle appreciation and praise measures are in Table 3.) Participants assumed that reason-intact Jack appreciated the utilitarian principle more so than did emotion-intact Jack, $t(456) = 8.61$, $p < .001$. Emotion-intact Jack was instead

assumed to appreciate the deontological principle more so than reason-intact Jack, $t(456) = 14.90$, $p < .001$.

As expected, praise judgments followed a similar pattern. The intact faculty X decision interaction was also significant, $F(1, 454) = 8.52$, $p = .004$. Jack was praised more for killing the child when he possessed reason compared to when he possessed emotion, $t(454) = 1.97$, $p = .05$. In contrast, Jack was praised more for not killing the child when he possessed emotion than when he possessed reason, $t(454) = 2.16$, $p = .03$.

To test the hypothesized mediation model (see Figure 2), we again created a *matching principle* appreciation variable that reflected the extent to which Jack was assumed to appreciate the principle that matched his ultimate behavior. When Jack made the utilitarian [deontological] decision, this variable reflected the extent to which Jack could appreciate the utilitarian [deontological] principle. We submitted the praise composite to a two-way 2 (intact faculty) X 2 (decision) ANCOVA, with appreciation of the matching principle as a covariate. Showing the effect of our proposed mediator, Jack was praised more to the extent he was assumed to appreciate the matching moral principle, $F(1, 453) = 14.30$, $p < .001$. Providing evidence consistent with full mediation, the intact faculty X decision interaction dropped to non-significance, $F < 1$. We tested the indirect effect of our manipulations (specifically, their interaction) on moral evaluation through the assumed appreciation of the matching principle. By the same Preacher and Hayes's (2008) bootstrapping procedure used in Studies 1a-1c, the 95% confidence interval for this effect did not include 0, [.2718, 1.1144]. It is worth noting that in this full model, the main effect of decision observed in Study 1a—that in general people tend to think that it is better to avoid actively killing the child—still emerged, $F(1, 453) = 7.42$, $p = .01$. This indicates that although the effect of the deficit manipulation on praise was fully mediated

through mindreading of moral principles, mindreading of moral principles was not the *only* factor that influenced how much Jack was praised for one action or the other. We will return to the significance of this distinction in the General Discussion.

Next, we tested the alternative mediation model that moral evaluation would also be influenced by assumed appreciation of the competing moral principle. Like in Studies 1a-1c, there was no significant effect of assumed competing principle appreciation on moral evaluation, $F < 1$. Furthermore, Preacher and Hayes's (2008) bootstrapping technique showed that there was no significant indirect effect on moral evaluation through assumed appreciation of the competing moral principle: The 95% confidence interval of the indirect effect included 0, [-.4034, .2995].

Discussion

Consistent with the MMP account, the moral agent was praised if he was assumed to appreciate the moral principle that would justify his action (i.e., the matching moral principle). That is, participants offered praise to the extent it was plausible that an action followed from a moral principle. The MMP perspective does not predict that social perceivers mindread by relying upon how much a moral agent is assumed to appreciate just any moral principle. And indeed, moral praise stemmed only from assumed appreciation of the matching moral principle; evaluations of the moral agent were unaffected by how much he was assumed to appreciate the moral principle that would justify the forgone behavior.

In the absence of direct information about what moral principles an agent would experience, Study 1a participants assumed the agent would experience the deontological principle against killing the child, and praised the agent more for that action. But when we introduced in Study 2 an extra factor (i.e., emotion or reason brain deficits) that was assumed to change what moral principles the agent would appreciate, moral evaluations for each decision

shifted accordingly.

In considering the usefulness of the MMP perspective, it is worthwhile to consider the present findings in light of recent developmental psychology research. Danovitch and Keil (2008) found that even young children reported that an emotionally-deficient computer would be a worse moral advisor than a rationally-deficient computer. This suggests that people may prize emotional sentiments over rational calculation as a source of moral knowledge. Consistent with this possibility, participants in Study 1a thought that the action driven by the emotional sentiment was superior. But participants in the present study showed no tendency to see the emotion-intact person as more moral than the reason-intact person. Instead, what shifted with the intact faculty manipulation was the moral evaluation each action prompted. Participants seemed to care little that moral agents appreciate one type of principle or the other, but instead that moral agents act on whatever principles they were able to appreciate.

One limitation of these results is that our measures of principle appreciation may have walked participants through a reasoning process—i.e., the hypothesized meditational pathway—that they would not have spontaneously employed. To address this concern, Study 3 assessed beliefs about principle appreciation in a pretest (to confirm that the manipulations did change the extent to which the agent was assumed to appreciate each principle) but assessed only moral evaluations in the main sample. In this way, participants provided moral evaluations without having had their attention first drawn to moral principles.

Study 3: Time to Deliberate

In Study 3, we returned to Tetlock et al.'s (2000) dilemma about a hospital director who must decide whether to spend a large sum of money to save a sick child (i.e, the same dilemma used in Study 1b). Study 3 built on the previous study in two ways. First, Study 3 manipulated a

feature of the decision making context that is likely to factor into many real world moral decisions. In particular, we manipulated whether a moral agent had to decide quickly or could engage in more extensive thought. As noted earlier, many deontological decisions are driven by quickly-appreciated, affect-backed principles. In contrast, utilitarian logic may be more easily appreciated only after longer deliberation and reflection (Greene et al., 2004, 2008). For Study 3, we assumed (and empirically confirmed) that participants would think that a rushed agent would more easily experience the deontological prohibition against letting the child die than the utilitarian principle that by letting the child die more lives could be saved in the future. In contrast, with more time to think, people should expect that an agent would appreciate both the deontological principle and utilitarian principle. This suggests that when the decision maker is rushed, perceivers should offer relatively more praise for the deontological decision (when the deontological principle is better appreciated than the utilitarian principle), but that this gap should be diminished when he has extensive time to deliberate (when both principles are appreciated). That is, given that only under rushed conditions is principle appreciation predicted to differ, it is under these (rushed) conditions that moral praise should be most sensitive to action.

Second, we measured assumed principle appreciation and moral evaluation in different samples. Although this approach does not allow us to directly test for mediation, it addresses the concern that our principle appreciation measures themselves may have suggested a type of mindreading in which participants would not have otherwise engaged.

Method

Participants and design. Two hundred fifty-eight Cornell University students were randomly assigned to one of four conditions in a 2 (speed: rushed or lengthy) X 2 (decision:

utilitarian or deontological) full-factorial, between-subjects design. Participants received course credit for their participation.

Procedure. Participants read a modified version of Tetlock et al.'s (2000) "sick Johnny" moral dilemma. In our version, two hospital directors must each decide whether to spend $3 million of the hospital's limited resources to save the life of a sick five-year-old named Johnny. Spending the money to save Johnny would prohibit the hospital from updating hospital infrastructure, updates that could be used to save many future lives. Thus, the hospital's co-directors — Robert and Alan — must independently choose whether to let Johnny die in order to save more lives in the future (utilitarian) or save the life of Johnny (deontological: "taboo tradeoff").

By chance, Alan was at the hospital when this situation arose, whereas Robert was initially unreachable. By the time hospital officials could reach and explain the situation to Robert, he had no time to engage in careful deliberation and had to make a quick decision based on his immediate gut instinct. In contrast, Alan had many hours to engage in careful, thorough reflection.

In a pretest, we confirmed our assumption that people believe it takes more deliberation to appreciate the utilitarian principle. We presented 126 people with the moral dilemma and asked them to what extent Robert (rushed deliberation) and Alan (lengthy deliberation) would appreciate the deontological ("find it troubling to kill a person") and utilitarian ("realize that by letting the person die, the hospital would actually save money which would allow it to save many more lives") principle. Confirming our assumption, a 2 (speed) X 2 (principle) ANOVA, measured entirely within-subjects, returned a significant interaction, $F(1, 125) = 36.49$, $p < .001$. People assumed that Robert, who had no time to deliberate, would better appreciate the

deontological principle than the utilitarian principle, paired $t(125) = 7.11$, $p < .001$. In contrast, people assumed that Alan, who had time to engage in more lengthy deliberation, would appreciate both the deontological and the utilitarian principles equally, $t < 1$.

In our main sample, participants learned of the existence and constraints of both directors, but learned of the decision of either Robert or Alan. Participants made five judgments about the target, each on 7-point, Likert-type scales. Participants indicated whether he should be praised (versus blamed), had a good (versus bad) moral conscience, was a good (versus bad) person, was the type of person who would be a good (versus bad) friend, and was a moral (versus immoral) person ($\alpha = .86$).

Results and Discussion

We submitted the praise composite to a 2 (speed) X 2 (decision) ANOVA. As expected, the interaction was significant, $F(1, 252) = 4.80$, $p = .03$ (see Table 3). Robert, who had to make a decision immediately, was praised much more for saving Johnny's life than for letting Johnny die, $d = .63$: $t(252) = 5.07$, $p < .0001$. In contrast, Alan, who had considerable time to think about his decision — received only somewhat more praise for the same decision, $d = .25$: $t(252) = 1.97$, $p = .05$. Thus, although the significant interaction is consistent with the mindreading moral principles approach, the fact that there was still a slight preference for the deontological action (saving the sick child's life) shows that mindreading moral principles is not the *only* influence on moral judgments. As Tetlock et al. (2000) would predict, people may find trading off a child's life for money to be morally distasteful in itself. This is true on top of the effect accounted for by mindreading moral principles. The present approach makes the unique prediction that this tendency to offer more praise for saving the child's life will be even stronger when the moral agent is forced to make his decision quickly—that is, when the agent is assumed to experience

the deontological but not the utilitarian principle.

Note that this pattern of results is inconsistent with an alternative prediction that when a decision is made quickly under situational duress, the act may be seen as less intentional and thus less useful in assessing blame or praise (see Monroe & Reeder, 2011). To the contrary, it was under rushed conditions that the agent's decision was seen to offer a more diagnostic, differentiated moral signal. Our pretest indicated that only under time duress should there be a difference in how much the deontological versus utilitarian principle would be appreciated; thus, it is under such time pressure that moral evaluations should most differ depending on the agent's action. Furthermore, given we measured beliefs about principle appreciation and praise with different samples, our results suggest that people spontaneously rely on principle appreciation in crediting targets. That is, the predicted interaction on moral evaluation emerged even though the measures never drew participants' attention to moral principles.

Relatedly, it is worth noting that decision speed was a useful cue even though the agent himself did not have control over whether he had a long or short time to deliberate. The present study can be contrasted against recent research that has examined what is signaled when moral agents arrive at moral decisions quickly or slowly of their own accord (Critcher et al., in press). In the present research, the length of time participants had to deliberate was instead governed by the situation. In this way, deliberation time in the present study is not an endogenous variable that is revealing of the agent's competing motives (Critcher et al., in press), but is an exogenous cue that serves as an indirect signal of what moral principles the agent could have been experiencing. These mindread principles altered how much praise each action prompts.

Study 4: Visual Salience

Study 4 moved beyond person (Study 2) or situational (Study 3) factors that *limited* the

agent's perceived ability to appreciate a moral principle. Study 4 examined a heretofore unstudied factor that might be seen *to enhance* the appreciation of one of two competing moral principles: the agent's visual perspective. Study 4 used the terrorist-inn dilemma introduced in Study 1c, in which an agent must decide whether to bomb an inn containing both high-level al-Qaeda terrorists and innocent civilians. We varied the agent's visual perspective, such that as he was making his decision, either a terrorist or an innocent bystander loomed large in the agent's visual field. We hypothesized that participants would assume that when the bystander was visually salient, the deontological principle against actively taking human life would loom large; when the terrorist was visually salient, the utilitarian principle justifying killing some to save more would loom large instead. We again predicted that agents would be praised to the extent that they were assumed to experience the matching moral principle.

Participants and design. Two hundred nineteen students at the University of California, Berkeley, were randomly assigned to one of four conditions in a 2 (salient target: terrorist or innocent bystander) X 2 (decision: utilitarian or deontological) full-factorial, between-subjects design.

Procedure. We modified the terrorist-inn scenario used in Study 1c to permit our manipulation of visual salience. Participants read about two high-level military commanders working to root out Al Qaeda terrorist cells in Afghanistan. The same information about terrorists and an innocent bystander in a rural inn was provided. The night of the meeting, Michael and Matt looked down at the inn from separate vantage points in the surrounding mountains. From Michael's lookout, the only person he could see through a window was the nervous-looking innocent translator. From Matt's lookout, the only person he could see through a window was a terrorist "who is #3 on the FBI's 'Most Wanted Terrorist' list". We reminded

participants that despite their different vantage points "both Michael and Matt know who all is in the room." Michael and Matt each had to decide independently whether to recommend an airstrike, which would kill all of those present.

At this point, participants rated the extent that both Michael and Matt would appreciate and experience the relevant deontological ("One should not kill innocent people regardless of the circumstances") and utilitarian ("One must stop people from killing thousands of people, even if one must kill an innocent person to do this") principles.

Next, participants were told that they would consider only one of the commanders, either Michael (innocent bystander salient) or Matt (terrorist salient). Participants learned that the agent either decided to order the attack (utilitarian decision) or not (deontological decision). Participants rated either Michael or Matt on the same five items used in all studies. All but one item required reverse-scoring (praiseworthy vs. blameworthy character) so that higher numbers would reflect greater praise ($\alpha = .79$).

Results and Discussion

Participants believed that visual salience would impact the relative appreciation of the competing principles: In submitting the four principle appreciation measures to a 2 (salient target: terrorist or bystander) X 2 (principle: deontological or utilitarian) repeated-measures ANOVA, a significant interaction emerged, $F(1, 218) = 224.09$, $p < .001$. Michael, who was looking at the innocent bystander, was assumed to appreciate the deontological more than the utilitarian principle, paired $t(218) = 16.14$, $p < .001$. Matt, who was looking at a terrorist, was instead assumed to appreciate the utilitarian more than the deontological principle, paired $t(218) = 3.02$, $p = .003$. (See Table 3 for all means.)

We then tested whether moral praise depended on our two between-subjects

manipulations, decision and visual salience. As predicted, a 2 (decision) X 2 (visual salience) interaction emerged, $F(1, 215) = 4.20$, $p = .04$. When the terrorist was visually salient, Matt was praised more for ordering the strike than failing to order it, $t(215) = 2.12$, $p = .04$. In contrast, when the innocent bystander was visually salient, there was a non-significant tendency in the opposite direction: Michael was praised directionally, but not significantly, more for failing to order the strike, $t < 1$.

Using a similar analytic strategy to Studies 1a, 1b, 1c, and 2, we assessed the mediation model that offers the more direct test of our MMS account (see Figure 3). The more the target was thought to appreciate the matching moral principle, the more he was praised, $F(1, 214) = 6.57$, $p = .01$. With the proposed mediator (matching moral principle) entered as a covariate, the decision X visual salience interaction no longer predicted praise, $F < 1$. Thus, the influence of the visual salience manipulation on judgments of one action versus the other was fully accounted for by the mindread matching principle. We used Preacher and Hayes' (2008) bootstrapping procedure to test the indirect effect through matching principle appreciation. The 95% confidence interval of this effect did not include 0, [.0204, 1.1239].

It is worth nothing that although unlike Studies 2 and 3, the results of Study 4 did not yield a significant main effect of decision, $F < 1$. That is, once the impact of mindreading moral principles was accounted for (i.e. the decision X extradecisional factor interaction), there was no significant tendency to see one action as better than another. That said, there still remained a non-significant tendency to see the utilitarian decision as more praiseworthy than the deontological decision (Study 1c). This can be seen in looking at the asymmetric strength of the simple effects. Nonetheless, the significant interaction does show that mindreading moral principles was the only significant predictor of moral judgment. There was no significant

tendency to view one type of action as more permissible or forbidden. Participants instead relied on who was salient in the agent's visual field to infer what moral principles the agent experienced, and this inference guided judgments of praise. This finding was foreshadowed in Study 1, in which this terrorist-inn dilemma (Study 1c) was the only scenario for which moral principles *fully* mediated the effect of decision on moral evaluation.

Although no support for the alternate mediation model through the competing moral principle was found in four previous attempts, we again assessed whether appreciation of the competing moral principle would also affect moral evaluation. When we changed the covariate in the above ANCOVA from assumed appreciation of the matching moral principle to the competing moral principle, the covariate no longer predicted praise, $F < 1$. Preacher and Hayes's (2008) bootstrapping procedure provided no evidence that there was an indirect effect on the praise composite through assumed appreciation of the competing principle, [-.2283, .5734].

Whereas the actual influence of the features manipulated in Studies 2 and 3 (emotion vs. reason and decision speed) has been the subject of previous research, Study 4 introduced a novel feature, visual perspective. Recent research, though, has examined the role of mental imagery in moral judgment. Amit and Greene (2012) found that one reason people find it more acceptable to kill one person in order to save five people (a utilitarian action) when that involves flipping a switch (switch dilemma) as opposed to pushing the single victim to his death (footbridge dilemma) is that people are more likely to create a vivid mental image of the victim in the footbridge versus the switch dilemma. If one treats presence in one's visual field as analogous to more vivid mental imagery, then Amit and Greene's (2012) study could be cited as support for the reasonableness of our participants' intuitions. What is important for the present purposes is that social perceivers assume that visual perspective affects moral principle appreciation, and

perceivers use this information in assigning moral praise. Future research could more directly assess how visual perspective affects moral decision-making.

General Discussion

How do perceivers assess whether an agent is worthy of moral praise? In addition to previously-identified mindreading strategies that have focused on the possible influence of selfish motives, perceivers also try to mindread what moral principles the agent appreciates. Praise is offered when the agent's actions are consistent with an appreciated principle; praise is withheld when the agent is unlikely to have appreciated the matching moral principle. This suggests perceivers may be constantly vigilant for positive signs that an action could plausibly have followed from a moral principle.

Studies 1a-1c showed that mindread moral principles inform judgments of praise. The studies employed moral dilemmas similar to those used in much previous research on moral reasoning. They described a moral choice confronting an agent and focused squarely on details of the action instead of the moral cognitions or beliefs of the agent. Consistent with the MMP account, the extent to which an agent was praised for each course of action was mediated by the extent to which the agent was assumed to appreciate the relevant moral principle (but not the competing moral principle).

Studies 2-4 offered experimental tests of our model by varying features that were assumed to shift an agent's appreciation of relevant moral principles. Study 2 varied whether an agent was said to have a deficit in emotion or reason; Study 3 varied whether an agent was rushed in his decision; Study 4 manipulated who was salient in an agent's visual field. These manipulations affected what principles the agent was assumed to appreciate and thus how much praise the agent received. Studies 2 and 4 showed that mindreading of principles mediated the

effects on praise, whereas Study 3 suggested that moral judgments draw on mindreading of moral principles spontaneously, even when participants' attention was not focused on principle appreciation by the measures. These studies support a causal direction in our mediation model (i.e., that assumed principle appreciation influences the praiseworthiness of actions) and shows how the MMP account can be useful in identifying and explaining extradecisional influences on moral evaluation.

Our studies highlight a new way by which people lean on contextual information not to help them determine whether actions are caused by the person or the situation—the historical focus of attribution work (e.g., Kelley, 1967)—but instead to help them *identify* the underlying meaning of a behavior. Trope (1986) noted that many behaviors are inherently ambiguous (e.g., an emotional facial expression), and people rely on information about the situation (e.g., the fact that the emoter just won a bet) to disambiguate what the behavior was. The mindreading moral principles account similarly emphasizes that people may look to contextual factors to help resolve ambiguity about a behavior's underlying meaning. The present work details one general way in which this disambiguation may unfold: The context provides cues about what moral principles are likely to be salient, which affects the perceived moral connotation of the behavior.

In thinking about why perceivers mindread moral principles to understand whether actions are principled and thus praiseworthy, it is helpful to consider the functions that moral evaluations ultimately serve. Part of the reason perceivers form moral evaluations of others is on the assumption that this will be a helpful guide in predicting their future moral or immoral behavior. That is, at one level, moral judgment concerns character judgment (Gray et al., 2012; Pizarro & Tannenbaum, 2011), which reflects one's more stable proclivity to do good or bad. This idea connects to the present research if one assumes that a principle-backed moral behavior

is more likely to signal future moral acts than is a behavior performed without an appreciation of the matching moral principle. In this way, even if the source of a moral principle is outside of the agents themselves (e.g., the visual salience of the terrorist in Study 4), the mere fact that the actor then decided to act on the salient principle may be informative as to whether he will act morally again in the future. That is, even though different situations may make different principles salient, and thus different actions praiseworthy, attention to these (shifting) cues is in the service of pinpointing who has stable moral character.

In light of recent findings that moral judgments can be pushed around by influences as trivial and incidental as hypnotically-induced disgust (Wheatley & Haidt, 2005), humorous film clips (Valdesolo & DeSteno, 2006), a bitter beverage (Eskine, Kacinik, & Prinz, 2011), and odious "fart spray" (Inbar, Pizarro, & Bloom, 2012; Schnall, Haidt, Clore, & Jordan, 2008), our depiction of moral reasoners as engaging in a sophisticated mindreading process may seem inconsistent. Is it odd that moral perceivers both try to infer what moral principles an agent appreciated before acting but also make judgments that are a function of incidental sensory experiences? A similar apparent contradiction was considered by Simonson (2008), who asked how it is that people's preferences show signs of being constructed at the moment they are asked to report those preferences, even as people's underlying preferences show signs of stability. Simonson's resolution applies to both his question about preferences and ours about moral judgment: The error is in thinking that the psychological process must be characterized by one or the other, for in actuality both can apply. Moral evaluation may be shaped by fairly sophisticated processes like mindreading moral principles even as such judgments are also influenced by incidental, biasing influences.

On this point, we should note that we did not predict (nor did our findings suggest) that

mindreading of moral principles is the *only* influence on moral judgment. In Studies 1a and 1b, mindreading of moral principles partially mediated effects on praise, and in Studies 2 and 3 the deontological action was still praised more than the utilitarian action even after controlling for mindreading of moral principles. That our manipulations' (i.e., neural deficits and visual salience) interactive influence on moral evaluation was fully mediated by mindread principles indicates that MMP *fully* accounts for these variables' influence on praise. But when main effects of decision lingered, this means that (unsurprisingly) other features affect moral evaluations as well. For example, although in Study 3 the relative praiseworthiness of funding sick Johnny's surgery compared to letting the child die was weaker when the hospital director had more time to consider his decision (and thus more time to come to appreciate the utilitarian principle justifying letting the child die), participants still thought it was relatively worse to trade off a child's life for money. Notwithstanding, in some circumstances mindreading moral principles can fully account for moral evaluations (Studies 1c and Study 4). One reading of this variability is that mindreading moral principles may always underlie moral evaluations, but sometimes other influences may matter as well.

One implication of the present findings is that in order to more fully understand moral judgment, interest in the question "What features of an action make it permissible or impermissible?" needs to be supplemented with "What features of a decision-making context will impact the salience of moral principles?" In our studies, participants' intuitions about moral principles conformed to certain patterns that need not (and likely do not) apply in all such situations. For example, although participants in Study 2 intuited a relationship between deontology and emotion, in other contexts it may be the utilitarian principles that are emotion-rich (Baron, 2011). Whereas participants in Study 3 intuited that deontological thinking is

quickly-appreciated, in other moral dilemmas it is actually the utilitarian principle that is appreciated more quickly and intuitively (Kahane et al., in press). Of course, there need not be a one-to-one correspondence between the actual properties of moral principles and the properties that social perceivers assume. Put differently, the validity of our findings does not hinge on people having accurate predictions about what moral principles spring to mind. But if future research documents the nature of beliefs about the factors affecting what moral principles a person appreciates in the moment, it will become easier to apply the present model to predict praise judgments.

Future research may also explore whether the MMP approach can explain why certain features of actions turn an otherwise permissible action into an impermissible one (Baron & Spranca, 1997; Mikhail, 2007). For example, people typically find it permissible to kill one person in order to save five if flipping a switch produces this outcome (switch dilemma), but not when pushing a fat man to his death does the same (footbridge dilemma). In explaining this divergence, researchers have identified how the kill and no-kill actions take different forms in each scenario (e.g., Greene et al., 2009; Waldmann & Dieterick, 2007). But instead of explaining people's tendency to condemn murder in the footbridge dilemma (but not the switch dilemma) by referencing descriptive rules that govern judgments of actions (e.g., "Directly pushing someone to their death is morally outrageous!"), it may be helpful to consider how these same contextual variations may shift what moral principles are presumably appreciated. For example, intentionally applying personal force to a victim likely requires that the agent hold the victim in his visual field. If this visual salience is assumed to make the rule condemning harm salient (as in Study 4), this could explain why perceivers believe agents should take the deontological no-harm action. Thus, an understanding of how MMP affects moral evaluation may not be independent of

an understanding of how subtle differences between actions affect moral evaluation, for the permissibility of certain actions may be traced to what moral principles the agent confronting those actions would be assumed to appreciate.

One may ask whether the mindreading moral principles account can be modified to offer a mindreading *im*moral principles account. That is, if people receive praise to the extent that they are assumed to appreciate a moral principle that would provide a moral justification for an action, would it also be the case that people are only blamed more to the extent that they are assumed to have an immoral thought that is the motivator of their action? We suspect mindread immoral principles may be treated differently. Such principles, because they are counternormative, may instead be directly informative in assessing an agent's moral character. That is, if a person is assumed to have the immoral motivating principle, "It is moral to kill children for my own gain", then merely holding this principle (even before it drives an action) may be grounds for ascribing immoral character.[8] In contrast, positive moral principles (e.g., "It is immoral to harm a child") are normative, which may be why they were not a basis for praise in themselves. If they had been, then assumed appreciation of the competing moral principle, not merely assumed appreciation of the matching moral principles, would have positively predicted moral praise as well. This result never emerged. Instead, moral-principle-backed actions, and not assumed principles alone, appear necessary to receive praise (see Kruger & Gilovich, 2004). As future research develops the MMP account, it will be important to directly test the role of mindread immoral principles as well.

Even more generally, does the mindreading moral principles account lend itself to other types of person perception? In the moral domain, mindreading is central because perceivers are interested not merely in observed behavior, but in the moral motivations that preceded it.

However, this interest in the mental steps that precede behavior is likely to guide non-morally-relevant judgment processes as well. For example, a math teacher interested in judging her student's calculus ability would not merely want to know whether he answered a multiple-choice math question correctly, but whether the student arrived at that answer by way of the correct mathematical computation. If the student answers correctly in just one second, it may be assumed that there was no time to actually work through the complex integral the problem required. As a result, accolades for the student's calculus ability may be withheld. Thus, the specific mindread bases for evaluation may vary by domain, but the general spirit of our model may prove valuable in non-moral contexts as well.

Conclusion

In American jurisprudence on discrimination, there was a longstanding debate on whether the law condemns actions (e.g., legislation) that have discriminatory effects, or actions that are based on discriminatory motives. In Arlington Heights v. Metropolitan Housing Corp (1977), Justice Powell wrote: 1) that motives, rather than actions, should serve as the grounds for censure, but 2) that actions frequently provide good insight into one's reasons for acting. The present work reinforces the first conclusion, that people are judged not for actions themselves but based on the mental precedents of those actions. But the present studies also suggest that we should be even more tentative in embracing Powell's second conclusion. Instead of merely looking to actions, people are sophisticated mindreaders of moral principles. They are sensitive to subtle cues that frequently provide insight into whether an action could have stemmed from a proper moral principle.

References

Amit, E., & Greene, J. D. (2012). You see, the ends don't justify the means: Visual imagery and moral judgment. *Psychological Science, 23*, 861-868.

Arlington Heights v. Metropolitan Housing Corp. 429 U.S. 252 (1977).

Baird, J. A., & Astington, J. W. (2004). The role of mental state understanding in the development of moral cognition and moral action. *New Directions for Child and Adolescent Development, 103*, 37-49.

Baron, J. (2011). Utilitarian emotions: Suggestions from introspection. *Emotion Review, 3*, 286-287.

Baron, J., & Spranca, M. (1997). Protected values. *Organizational Behavior and Human Decision Processes, 70*, 1-16.

Bartels, D. M. (2008), Principled moral sentiment and the flexibility of moral judgment and decision making," *Cognition, 108,* 381-417.

Critcher, C. R., & Dunning, D. (2011). No good deed goes unquestioned: Cynical reconstruals maintain belief in the power of self-interest. *Journal of Experimental Social Psychology, 47*, 1207-1213.

Critcher, C. R., Inbar, Y., & Pizarro, D. A. (in press). How quick decisions illuminate moral character. *Social Psychological and Personality Science*.

Cushman, F. (2008). Crime and punishment: Distinguishing the roles of causal and intentional analyses in moral judgment. *Cognition, 108*, 353-380.

Cushman, F.A. and Greene, J.D. (2012) Finding faults: How moral dilemmas illuminate cognitive structure. *Social Neuroscience, 7*, 269-279.

Danovitch, J. H., & Keil, F. C. (2008). Young Humeans: The role of emotions in children's

evaluation of moral reasoning abilities. *Developmental Science, 11*, 33-39.

Eskine, K. J., Kacinik, N. A., & Prinz, J. J. (2011). A bad taste in the mouth: Gustatory disgust influences moral judgment. *Psychological Science, 22*, 295-299.

Fedotova, N. O., Fincher, K. M., Goodwin, G. P., & Rozin, P. (2011). How much do thoughts count? Preference for emotion versus principle in judgments of antisocial and prosocial behavior. *Emotion Review, 3*, 316-317.

Fein, S. (1996). Effects of suspicion on attributional thinking and the correspondence bias. *Journal of Personality and Social Psychology, 70*, 1164–1184.

Gray, K., Young, L., & Waytz, A. (2012). Mind perception is the essence of morality. *Psychological Inquiry*, *23*, 101-124.

Greene, J. D. (2009). Dual-process morality and the personal/impersonal distinction: A reply to McGuire, Langdon, Coltheart, and Mackenzie. *Journal of Experimental Social Psychology, 45*, 581-584.

Greene, J.D., Cushman, F.A., Stewart. L.E., Lowenberg, K., Nystrom, L.E., and Cohen, J.D. (2009) Pushing moral buttons: The interaction between personal force and intention in moral judgment. *Cognition*, *111*, 364-371.

Greene, J. D., Morelli, S. A., Lowenberg, K., Nystrom, L. E., & Cohen, J. D. (2008). Cognitive load selectively interferes with utilitarian moral judgment. *Cognition, 107*, 1144-1154.

Greene, J. D., Nystrom, L. E., Engell, A. D., Darley, J. M., & Cohen, J. D. (2004). The neural bases of cognitive conflict and control in moral judgment. *Neuron, 44*, 389-400.

Greene, J.D., Sommerville, R.B., Nystrom, L.E., Darley, J.M., & Cohen, J.D. (2001). An fMRI investigation of emotional engagement in moral Judgment. *Science*, *293*, 2105-2108.

Guglielmo, S., Monroe, A. E., & Malle, B. F. (2009). At the head of morality lies folk psychology. *Inquiry, 52*, 449-466.

Hauser, M. (2006). *Moral minds: How nature designed our universal sense of right and wrong*. New York: Ecco.

Inbar, Y., Pizarro, D. A., & Bloom, P. (2012). Disgusting smells cause decreased liking of gay men. *Emotion, 12*, 23-27.

Kahane, G., Wiech, K., Shackel, N., Farias, M., Savulescu, J., & Tracey, I. (in press). The neural basis of intuitive and counterintuitive moral judgment. *Social Cognitive and Affective Neuroscience.*

Karniol, R. (1978). Children's use of intention cues in evaluating behavior. *Psychological Bulletin, 85*, 76-85.

Kelley, H. H. (1967). Attribution theory in social psychology. In D. Levine (Ed.), Nebraska symposium of motivation (Vol. 15). Lincoln: University of Nebraska Press.

Kelley, H. H. (1971). *Attribution in social interaction.* New York: General Learning Press.

Knobe, J. (2003). Intentional action and side effects in ordinary language. Analysis, *63*, 190–193.

Knobe, J. (2004). Intention, intentional action and moral considerations. *Analysis, 64*, 181-187.

Koenigs, M., Young, L., Adolphs, R., Tranel, D., Cushman, F., Hauser, M., et al. (2007). Damage to the prefrontal cortex increases utilitarian moral judgments. *Nature*, *446*, 908–911.

Kruger, J., & Gilovich, T. (2004). Actions, intentions, and trait assessment: The road to self-enhancement is paved with good intentions. *Personality and Social Psychology Bulletin, 30*, 328-339.

Kruglanski , A. Schwartz , J. M. Maides , S. A. Hamel , I. A. (1978). Covariation, discounting, and augmentation: Toward a clarification of attributional principles. *Journal of Personality, 46*, 176-189.

Malle, B. F., Knobe, J., O'Laughlin, M. J., Pearce, G. E., & Nelson, S. E. (2000). Conceptual structure and social functions of behavior explanations: Beyond person-situation attributions. *Journal of Personality and Social Psychology, 79*, 309-326.

Mikhail, J. (2002). Aspects of the Theory of Moral Cognition: Investigating Intuitive Knowledge of the Prohibition of Intentional Battery and the Principle of Double Effect. Georgetown University Law Center Public Law & Legal Theory Working Paper No. 762385. Available at http://ssrn.com/abstract1/4762385

Mikhail, J. (2007). Universal moral grammar: Theory, evidence and the future. *Trends in Cognitive Science, 11*, 143-152.

Mikhail, J. (2011). *Elements of moral cognition: Rawls' linguistic analogy and the cognitive science of moral and legal judgment*. Cambridge, UK: Cambridge University Press.

Miller, M. B., Sinnott-Armstrong, W., Young, L., King, D., Paggi, A., Fabri, M., Polonara, G., & Gazzaniga, M. S. (2010). Abnormal moral reasoning in complete and partial colostomy patients. *Neuropsychologia, 48*, 2215-2220.

Monroe, A. E., & Reeder, G. D. (2011). Motive-matching: Perceptions of intentionality for coerced action. *Journal of Experimental Social Psychology, 47*, 1255-1261.

Nichols, S., & Mallon, R. (2006). Moral dilemmas and moral rules. *Cognition, 100*, 530-542.

Piaget, J. (1932). *The moral judgment of the child.* London: Kegan Paul, Trench, Trubner and Co.

Pizarro, D. A., & Tannenbaum, D. (2011). Bringing character back: How the motivation to

evaluate character influences judgments of moral blame. In P. Shaver & M. Mikulincer (Eds.), *The social psychology of morality: Exploring the causes of good and evil* (pp. 91–108). Washington, DC: American Psychological Association.

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods, 40,* 879–891.

Pronin, E. (2008). How we see ourselves and how we see others. *Science*, *320*, 1177–1180.

Reeder, G. D. (2009) Mindreading and dispositional inference: MIM revised and extended. *Psychological Inquiry, 20*, 73-83.

Reeder, G.D. & Spores, J.M. (1983). The attribution of morality. *Journal of Personality and Social Psychology, 44*, 736-745.

Reeder, G. D., Vonk, R., Ronk, M. J., Ham, J., & Lawrence, M. (2004). Dispositional attribution: Multiple inferences about motive-related traits. *Journal of Personality and Social Psychology, 86*, 530-544.

Schnall, S., Haidt, J., Clore, G. L., & Jordan, A. H. (2008). Disgust as embodied moral judgment. *Personality and Social Psychology Bulletin, 34*, 1096-1109.

Schopenhauer, A. (2009). *The two fundamental problems of ethics* (C. Janaway, Trans.). Cambridge, UK: Cambridge University Press. (Original work published 1841)

Simonson, I. (2008).Will I like a “medium” pillow? Another look at constructed and inherent preferences. *Journal of Consumer Psychology, 18*, 157−171.

Sripada, C. S. (2012). Mental state attributions and the side-effect effect. *Journal of Experimental Social Psychology, 48*, 232-238.

Tetlock, P. E., Kristel, O. V., Elson, S. B., 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.

Trope, Y. (1986). Identification and inferential processes in dispositional attribution. *Journal of Personality and Social Psychology, 93*, 239-257.

Valdesolo, P. & DeSteno, D.A. (2006). Manipulations of Emotional Context Shape Moral Decision Making. *Psychological Science, 17 (6),* 476-477.

Waldmann, M. R., & Dieterich, J. H. (2007). Throwing a bomb on a person versus throwing a person on a bomb: intervention myopia in moral intuitions. *Psychological Science, 18*, 247-253.

Wheatley, T., & Haidt, J. (2005). Hypnotically induced disgust makes moral judgments more severe. *Psychological Science, 16,* 780-784.

Young, L., Cushman, F., Hauser, M., Saxe, R., (2007). The neural basis of the interaction between theory of mind and moral judgment. *Proceedings of the National Academy of Sciences*, *104*, 8235–8240.

Young, L., & Saxe, R. (2008). The neural basis of belief encoding and integration in moral judgment. *NeuroImage, 40*, 1912-1920.

Yuill, N. (1984). Young children's coordination of motive and outcome in judgements of satisfaction and morality. *British Journal of Developmental Psychology, 2*, 73-81.

Yuill, N., Perner, J., (1988). Intentionality and knowledge in children's judgments of actors responsibility and recipients emotional reaction. *Developmental Psychology, 24*, 358–365.

FOOTNOTES

1. We test our hypotheses using these dilemmas not only because they pit actions with clear matching moral principles, but also for three additional reasons. First, there has been extensive research in moral psychology on sacrificial moral dilemmas of this type, largely in an effort to develop a descriptive account of moral judgment (Bartels, 2008; Cushman & Greene, 2012; Mikhail, 2007). Relying on similar methodologies permits comparisons between our investigations. Second, and relatedly, this previous research has typically focused on what features of *actions* change moral judgments. This offers a particularly conservative context in which to test our mindreading moral principles (MMP) account, given our interest in how inferred, but unobservable, principles may mediate moral judgments. Third, studying contexts in which moral agents have a choice between two courses of action permits us to pit our MMP account against a related account that suggests moral evaluations may be affected by mindread moral principles that would justify forgone choices. We detail this alternative model when we first test it.

2. It should be stressed that we rely on the *utilitarian* and *deontological* distinction simply to keep straight which principle matches with which behavior, not because our model treats utilitarian and deontological actions differently.

3. In Studies 1a and 1c, the deontological prohibition against killing is rooted in Kant's categorical imperative; in Study 1b, in Tetlock's notion of taboo tradeoffs.

4. It has even been suggested that this more psychologically-realistic route to deontological behavior is actually more praiseworthy than a dispassionate deduction from Kantian principles (Schopenhauer, 1841/2009).

5. It is worth noting that Bartels (2008) found, in an almost-identical dilemma, that people

indicated that they would smother the child in this context (utilitarian behavior). We find that participants praise the agent more for not smothering the child (deontological behavior). This highlights that studies that examine how people would resolve dilemmas are not a substitute for studies of moral evaluation. Instead, as the mindreading moral principles perspective predicts, praise is offered for the deontological behavior because they assume the deontological principle will be more focal than the utilitarian principle in the agent's mind.

6. The failure of the alternative model results whether the two mediators are tested simultaneously as candidate mediators or each separately as lone mediators.

7. We used two questions to check whether participants in fact believed that appreciation of the deontological and utilitarian principles stemmed from emotionality and reason, respectively. Participants indicated on 8-point scales whether a decent person whose morals told him he should [not] kill the baby would be influenced more by his emotional impulses (1) or dispassionate "mathematical" calculation (8). Participants indicated that a decent person's decision to kill the child would be driven more by mathematical calculation than by emotional impulses ($M = 6.05$, $SD = 1.69$), $t(457) = 19.58$, $p < .001$, but that a decision to let the child cry would be driven more by emotional impulse than mathematical calculation ($M = 2.54$, $SD = 1.72$), $t(457) = 24.41$, $p < .001$. These two tests against the midpoint (4.50) confirm our assumption that in this dilemma the utilitarian principle is assumed to be appreciated through reason, and the deontological principle, through emotion.

8. Recall that we use the word "principle" as a catchall term to refer to the moral cognitions and emotions that precede action. Thus, in considering immoral principles, this would not merely refer to misguided attempts at moral justifications, but the cognitions and emotions that are the immoral motivation behind blameworthy actions.

Table 1

Summary of Moral Dilemmas Used

Dilemma	**Summary**	**Utilitarian Action**	**Deontological Action**	**Studies**
Crying Baby	A Jewish townsperson, along with other Jewish townspeople, hide in a secret basement as Nazi soldiers search the town above. A baby in the basement begins to cry. If the baby continues to cry, it will attract the attention of the Nazi soldiers. The Nazis will kill any Jews—children or adults—that they discover.	The Jewish townsperson smothers the child to death.	The Jewish townsperson does not smother the child to death.	1a, 2
Sick Johnny	A hospital director must decide whether to save the life of a sick five-year-old Johnny by funding an expensive organ transplant. If the surgery is denied, Johnny will die, but the hospital will retain funds to improve hospital quality, thereby saving more lives in the future.	The hospital director denies the surgery.	The hospital director funds the surgery.	1b, 3
Terrorist-Inn	An American military commander must decide whether to launch an airstrike on a rural inn. Top Al Qaeda operatives are meeting inside. A strike on the inn would kill everyone (including an innocent bystander), but would stop the terrorists from launching a 9/11-style terrorist attack.	The military commander orders a strike on the inn.	The military commander does not authorize the strike.	1c, 4

Table 2

Regression Models Predicting Moral Praise (Studies 1a-1c)

	Study 1a			Study 1b			Study 1c		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
(Deontological) Decision	.40***	.29**	.34**	.64***	.51***	.44***	-.20*	-.16	-.10
Appreciate matching principle		.32**	.30**		.36***	.37***		.20*	.20*
Appreciate competing principle			.16			-.13			-.14

Note: Model 1 is the direct effect of the IV (Decision) on the DV (Praise). Models 2 and 3 add two possible mediators: assumed appreciating of the matching principle (Model 2) and the competing principle (Model 3).

Table 3

Assumed Moral Principle Appreciation and Moral Praise For Utilitarian and Deontological Principles and Behaviors

	Moral principle appreciation		Moral evaluation for behavior	
	Utilitarian	Deontological	Utilitarian	Deontological
Study 2: Skill intact				
Emotion	4.03 (2.05)$_b$	5.88 (1.87)$_a$	4.35 (1.28)$_a$	4.92 (1.04)$_c$
Reason	5.74 (2.21)$_a$	3.11 (2.09)$_c$	4.64 (0.91)$_{bc}$	4.59 (1.31)$_{ab}$
Study 3: Speed				
Rushed	5.31 (1.69)$_c$	6.79 (1.30)$_a$	5.14 (1.11)$_a$	6.07 (1.12)$_c$
Lengthy	6.26 (1.43)$_b$	6.32 (1.56)$_b$	5.22 (1.11)$_a$	5.61 (1.00)$_b$
Study 4: Visual Salience				
Innocent Bystander	4.34 (2.07)$_d$	7.53 (1.76)$_a$	5.16 (1.19)$_{ab}$	5.35 (1.10)$_{ab}$
Terrorist	6.76 (1.69)$_b$	6.17 (2.00)$_c$	5.49 (0.98)$_a$	5.03 (1.31)$_b$

Note. Within each distinct group of four means—those from a single study that describe the same dependent measure—means with a different subscript are significantly different, $p < .05$.

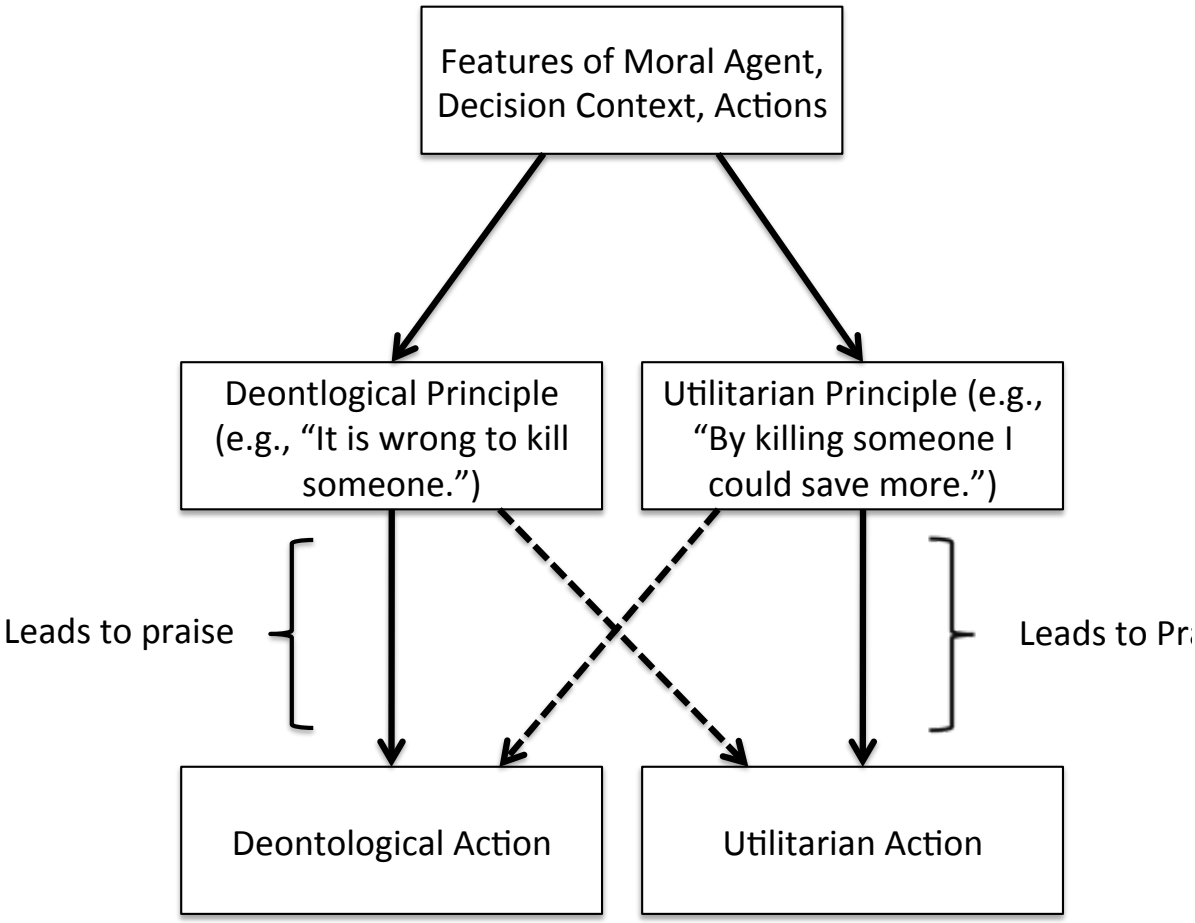

Figure 1. Mindreading moral principles account of praise, applied to dilemmas pitting deontological against utilitarian courses of action. Features of the moral agent, the decision context, and the possible actions themselves influence what principles it is assumed a moral agent appreciates. The moral agent can choose between competing actions. Moral agents are praised to the extent that they appreciate or experience the matching moral principle (those tied to each action by the solid arrow). Whether or not they appreciate the competing moral principle (those tied to each action by the dotted arrow) has no influence.

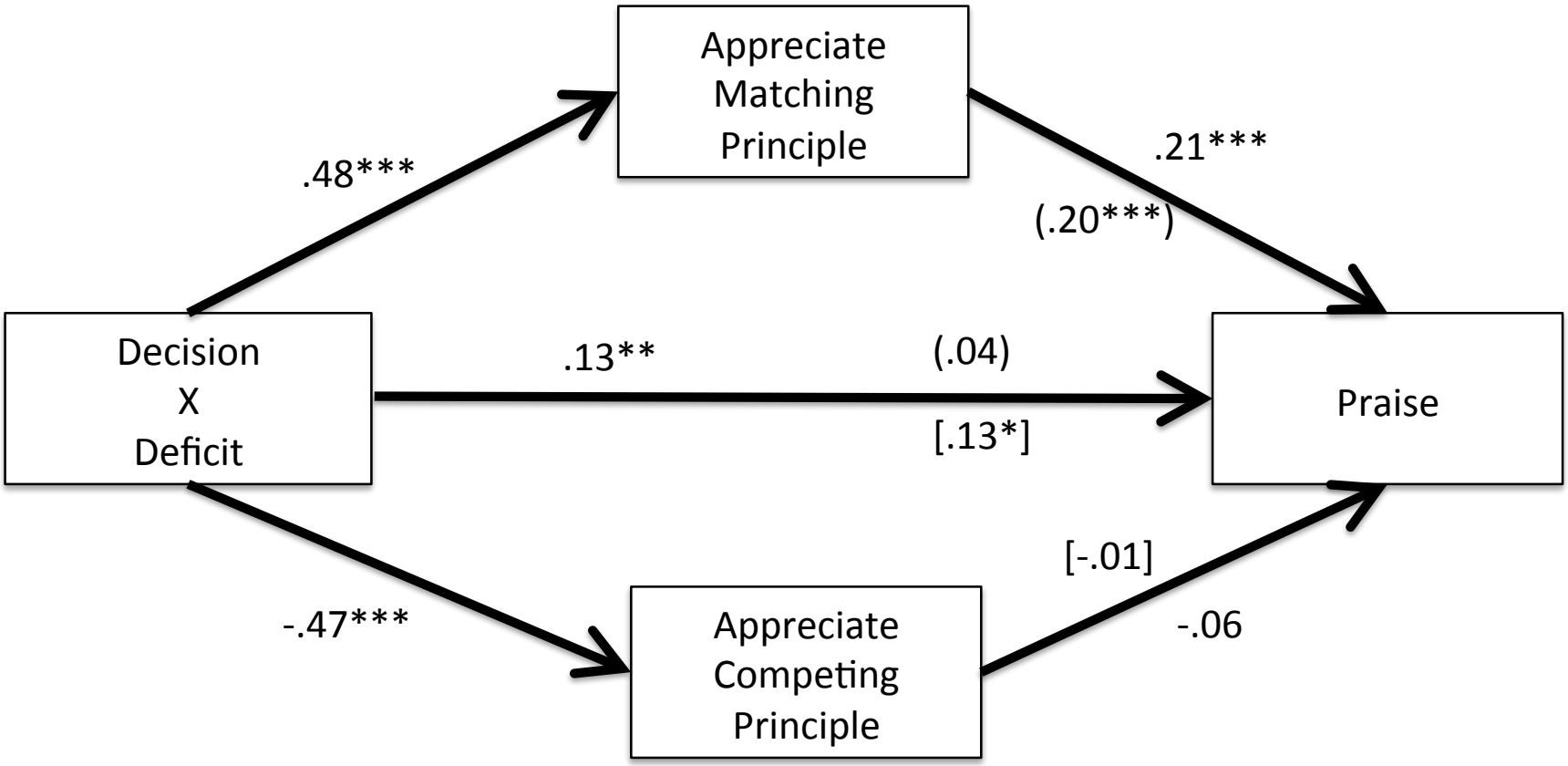

Figure 2. Assumed appreciation of the matching principle fully mediates the interactive influence of the manipulations (decision and visual salience) on praise. There is no similar indirect effect through assumed appreciating of the competing principle. All numbers are standardized betas. Standardized betas in parentheses or brackets come from the same model. (Study 2).

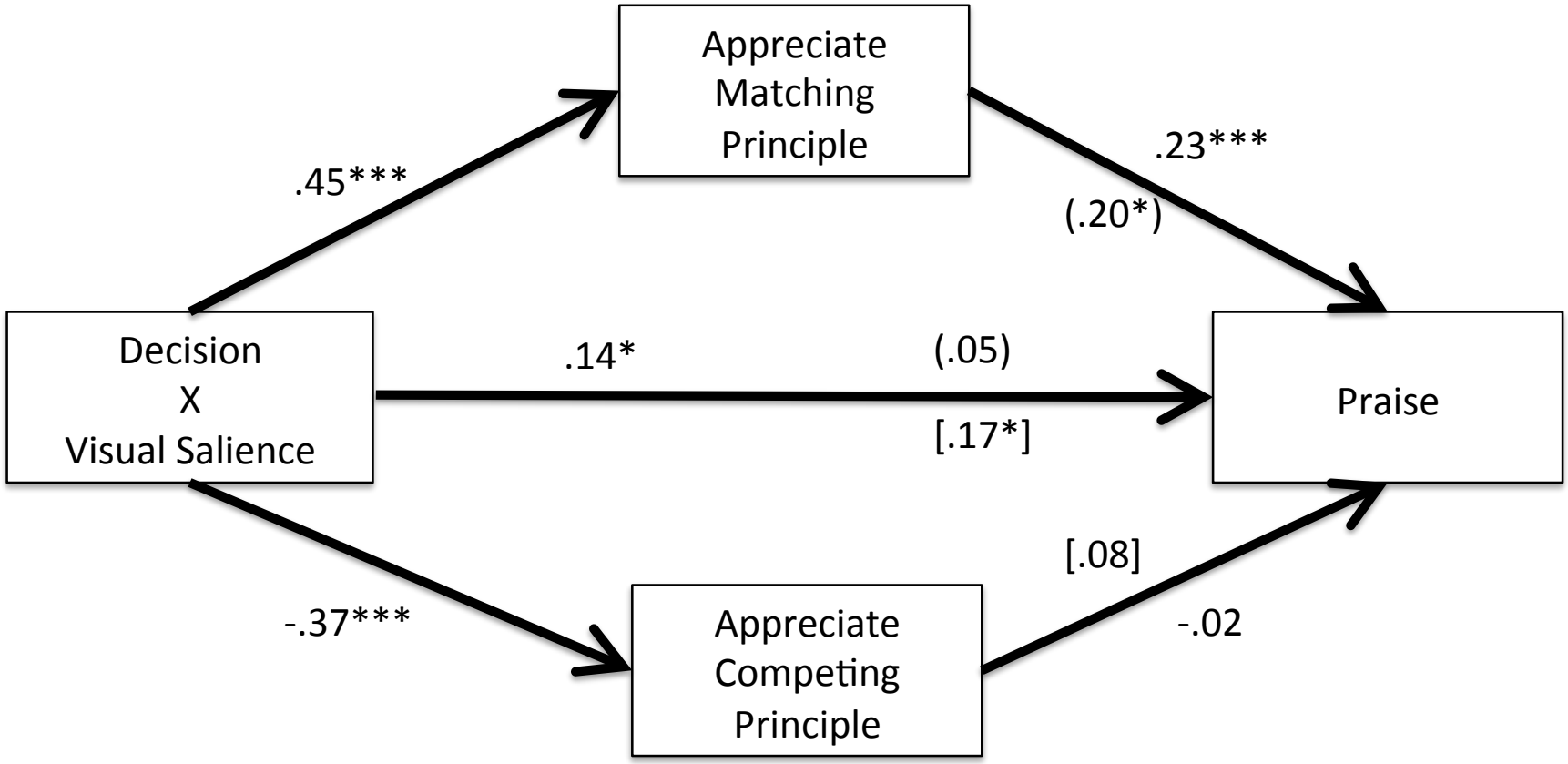

Figure 3. Assumed appreciation of the matching principle fully mediates the interactive influence of the manipulations (decision and visual salience) on praise. There is no similar indirect effect through assumed appreciating of the competing principle. All numbers are standardized betas. Standardized betas in parentheses or brackets come from the same model. (Study 4).