Enough skill to kill: Intentionality judgments and the moral valence of action

Steve Guglielmo *, Bertram F. Malle

Brown University, Department of Cognitive, Linguistic, and Psychological Sciences, 89 Waterman St., Providence, RI 02912, United States

ARTICLE INFO

ABSTRACT

Extant models of moral judgment assume that an action's intentionality precedes assignments of blame. Knobe (2003b) challenged this fundamental order and proposed instead that the badness or blameworthiness of an action directs (and thus unduly biases) people's intentionality judgments. His and other researchers' studies suggested that blameworthy actions are considered intentional even when the agent lacks skill (e.g., killing somebody with a lucky shot) whereas equivalent neutral actions are not (e.g., luckily hitting a bull's-eye). The present five studies offer an alternative account of these provocative findings. We suggest that people see the morally significant action examined in previous studies (killing) as accomplished by a basic action (pressing the trigger) for which an unskilled agent still has sufficient skill. Studies 1 through 3 show that when this basic action is performed unskillfully or is absent, people are far less likely to view the killing as intentional, demonstrating that intentionality judgments, even about immoral actions, are guided by skill information. Studies 4 and 5 further show that a neutral action such as hitting the bull's-eye is more difficult than killing and that difficult actions are less often judged intentional. When difficulty is held constant, people's intentionality judgments are fully responsive to skill information regardless of moral valence. The present studies thus speak against the hypothesis of a moral evaluation bias in intentionality judgments and instead document people's sensitivity to subtle features of human action.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

Some observations about moral judgment are uncontroversial. Foreseeability has been widely observed as a requirement for folk responsibility judgments—people normally do not consider someone responsible for outcomes that the person was unable to anticipate (Shaver, 1985). Equally significant, Hamilton (1978) recognized the role of obligation, as people are blamed only for negative outcomes that they were obligated to prevent. Finally, Weiner (1995) observed that such obligation is meaningful only if the outcome was controllable by the person—that is, if he or she could have intentionally prevented it. Thus, blame for negative events arises when the person *should have* and *could have* prevented it (Malle, Moses, & Baldwin, 2001).

These findings highlight people's rational, rule-following assignments of responsibility and blame (cf. Nichols & Mallon, 2006). Admittedly, things are not always so clear-cut. Alicke (2000) and others have shown that extraneous variables can alter moral judgment beyond rational rules (Schnall, Haidt, Clore, & Jordan, 2008). But there is little doubt that such rules exist and operate in many instances.

One of the most important rules governing the assignment of blame is that intentional moral transgressions—when a person intentionally performs an immoral action—amplify blame (Cushman, 2008; Heider, 1958; Ohtsubo, 2007; Shaver, 1985). Unintentional harmful

* Corresponding author. Address: Brown University, Department of Cognitive, Linguistic, and Psychological Sciences, 89 Waterman St., Providence, RI 02912, USA. Tel.: +1 401 863 2727.

E-mail address: Steve_Guglielmo@Brown.edu (S. Guglielmo).
URL: http://research.clps.brown.edu/mbq/guglielmo/ (S. Guglielmo).

behavior may elicit blame if the agent could have and should have prevented the harm (Weiner, 1995). But actually having a desire and intention to bring about harm, and exerting effort to realize this intention, is the worst offense in any social community. Excuses sometimes avert blame for unintentional harm; only justifications can possibly avert blame for intentional harm (Tedeschi & Reiss, 1981), and except for rare cases in which a compelling reason for the harmful act is available (e.g., the dentist hurting her patient), full blame applies.

Judgments of intentionality, thus, serve as a central input to judgments of blame. Schematically, the social perceiver takes two judgment steps, in order (Guglielmo, Monroe, & Malle, 2009):

1. Determine: Is the negative event intentional?
2. If Yes → examine the actual intention or goal; then assign proper blame.
   If No → examine obligation and ability to prevent (foreseeability and controllability); then assign proper blame.

Knobe (2003a, 2003b) challenged this fundamental order and proposed instead that the badness or blameworthiness of an action can influence people's intentionality judgments. In particular, Knobe argued, the same behavior that is seen as unintentional when performed without moral implication (e.g., shooting and hitting a bull's-eye) may in fact be seen as intentional when performed with moral implication (e.g., shooting and killing another person). If true, such a pattern would cast serious doubt on rational models of both blame assignment and intentionality judgments. People would not, as traditionally believed, assess intentionality to designate blame but would instead assess blame to designate intentionality.

The fault of the traditional account of blame, following Knobe's argument, lies in its assumption about how intentionality judgments are made. Both philosophical theories (e.g., Mele, 1992; Mele & Sverdlik, 1996; Searle, 1983) and psychological theories of intentionality (Malle & Knobe, 1997) subscribe to the valence-neutral model sketched in Fig. 1. According to this model, for positive, negative, and neutral behaviors alike, people process five information components that all have to be present for a behavior to be considered intentional. If even one component is missing, the behavior does not count as intentional.

Knobe claims that this five-component model is correct only for neutral actions but does not hold for negatively valenced actions (we will return to the question of morally positive actions later in this article). Knobe's evidence for his claim falls into two sets. The first set challenges the

necessity of the *intention* component for judgments of intentionality. According to the standard intentionality model, a behavior is judged as intentional only if the agent actually intended to perform that behavior (Adams, 1986; Malle & Knobe, 1997). In contrast, Knobe (2003a) provided data suggesting that people consider a known but *unintended* side effect intentional if that effect is negatively valenced (e.g., harming the environment, risking the lives of soldiers, decreasing sales). This effect has been replicated numerous times (Cushman & Mele, 2008; McCann, 2005; Nadelhoffer, 2006a; Nichols & Ulatowski, 2007), and we refer to this set of findings as the "side-effect challenge."

The second set of findings questions the necessity of the *skill* component for judgments of intentionality. According to the standard intentionality model, people judge a behavior as intentional only if the agent has reliable ability or skill to produce that behavior (Malle & Knobe, 1997; Mele & Moser, 1994; Thompson, Armstrong, & Thomas, 1998). Knobe (2003b) showed that an agent's unskilled neutral action (e.g., a lucky shot to win a contest) is not viewed as intentional but an equivalent unskilled *immoral* action (e.g., a lucky shot to kill someone) is very much seen as intentional. Once more, other researchers have replicated this effect (Nadelhoffer, 2004, 2005; Sousa & Holbrook, 2010), and we refer to this set of findings as the "skill challenge."

In a separate paper, we have analyzed the validity of the side-effect challenge (Guglielmo & Malle, in press), particularly its conditions of occurrence and the components of intentionality it reveals. We concluded that, once comparisons between morally valenced and nonvalenced cases are made truly parallel, and once people can express their judgments in their own terms, the side-effect finding disappears. People are in fact keenly sensitive to the basic components of intentionality that the standard model postulates—belief, desire, and intention—and do not seem to be biased by the moral valence of a side-effect. Moreover, when freely characterizing known but unintended side effects, people do not actually label them intentional, and so they do not make judgments of intentionality without a prior judgment of intention.

The skill challenge, however, still looms. To examine this challenge we will try to identify the conditions under which such findings occur and clarify their implications for theories of intentionality, theories of blame, and for the prospect of valid *mens rea* judgments in the law. For if the negative valence of a defendant's action biases jurors toward "seeing" intent in the action, we would seriously question their fair capacity to assess such intent (Nadelhoffer, 2006b).
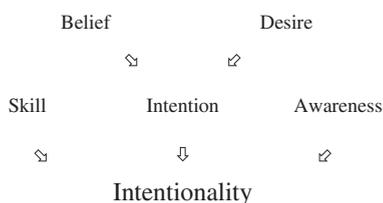
1.1. Knobe's original study

Knobe's (2003b) original skill challenge derived from a study in which four components of intentionality (see Fig. 1) were held constant but skill was varied. Together with a manipulation of valence, Knobe employed a 2 (skill: high/low) × 2 (valence: neutral/negative) design. The specific vignettes were as follows:

Belief   Desire

Skill   Intention   Awareness

Intentionality

Fig. 1. A model of the folk concept of intentionality.

1.1.1. Neutral condition

Jake desperately wants to win the rifle contest. He knows that he will only win the contest if he hits the bull's-eye. He raises the rifle, gets the bull's-eye in the sights, and presses the trigger.

[Skill:] Jake is an expert marksman. His hands are steady. The gun is aimed perfectly...

[Low skill:] But Jake isn't very good at using his rifle. His hand slips on the barrel of the gun, and the shot goes wild...Nonetheless,

the bullet lands directly on the bull's-eye. Jake wins the contest.

Did Jake intentionally hit the bull's-eye?

1.1.2. Negative condition

Jake desperately wants to have more money. He knows that he will inherit a lot of money when his aunt dies. One day, he sees his aunt walking by the window. He raises his rifle, gets her in the sights, and presses the trigger.

[Skill:] Jake is an expert marksman. His hands are steady. The gun is aimed perfectly...

[Low skill:] But Jake isn't very good at using his rifle. His hand slips on the barrel of the gun, and the shot goes wild...Nonetheless,

the bullet hits her directly in the heart. She dies instantly.

Did Jake intentionally kill his aunt?

As expected, when skill was high, both valence conditions elicited intentionality judgments in most people (79% for neutral, 95% for negative). When skill was low, however, the neutral condition showed the predicted drop in intentionality judgments to 28% whereas the negative condition showed persistently high intentionality judgments by most people (76%).

Scholars have interpreted the skill challenge findings to reflect two related phenomena—one specific, the second more general:

1.1.3. Skill neglect

For negative actions, skill is not a necessary condition of intentionality. Nadelhoffer (2005, p. 350) argues that "skill and control are not necessary conditions of the folk concept of intentional action." More specifically, Sousa and Holbrook (2010) assert that "people ignore the skill component...when judging the intentionality of immoral actions" (p. 368). According to the skill neglect interpretation, variations in skill should have little to no impact on intentionality judgments about negative actions.

1.1.4. Moral bias

Negative actions are judged more intentional than neutral actions, and this is because of their immorality or blameworthiness. Moral bias is a more general phenomenon than skill neglect and has been offered as an explanation for a range of findings in the literature. For example, Knobe (in press) suggests that "people's judgments about whether a given action truly is morally good or bad can actually affect their intuitions about what that action caused and what mental states the agent had" (p. xx) With respect to intentionality judgments, Nadelhoffer (2006b, p. 208) argues there is a "biasing effect that moral considerations have on folk ascriptions of intentional action," which may be because "observers engage in a biased information search to support a desired blame attribution" (Alicke, 2000, p. 567). One way in which moral bias may manifest is via skill neglect. But other processes may contribute to moral bias as well—for example, when considering negative actions, people may adopt more lenient thresholds for other intentionality components (Pettit & Knobe, 2009) or may ignore certain components altogether (Alicke, 2008).

In this article, we challenge both the skill neglect and moral bias claims. We begin with the more specific claim that people neglect the skill condition of intentionality when judging immoral actions. We argue that, in the extant studies, even the low-skill agent who acted immorally (e.g., shot his aunt) skillfully performed a basic action that may count as killing intentionally. Study 1 shows that intentionality judgments drop when the skill of this basic action becomes doubtful. Studies 2 and 3 directly compare high- and low-skill actions and show a strong effect of skill on intentionality judgments, for both positive and negative actions.

We next examine the more general claim that an immoral action is more likely to be judged intentional than a neutral one specifically because of the former's blameworthiness. Study 4 shows that Knobe's (2003b) original neutral action (hitting the bull's-eye) was actually more specific and more difficult than the negative action (killing). Once we equate specificity and difficulty of the actions (e.g., comparing hitting the aunt's heart with hitting the bull's-eye), the intentionality asymmetry disappears. Finally, Study 5 combines the previous insights and shows that if no skillful basic action is mentioned and people consider the specific action of hitting the aunt's heart, hardly anyone judges the action to be intentional, even though the agent is assigned a great deal of blame.

In these studies, we focus primarily on the negative/low-skill condition, as this condition lies at the core of both the skill neglect and the moral bias interpretation of the skill challenge. Moreover, this condition leads to distinct predictions between the skill challenge and the standard intentionality model. Specifically, both the skill neglect and moral bias interpretations predict that for negative actions, intentionality judgments will remain high and insensitive to variations in skill. In contrast, the standard model predicts that skill, but not valence, reliably influences intentionality.

2. Study 1

In Knobe's (2003b) original study, people judged a wobbly, wayward shot as intentional when it fulfilled the plan to kill another person. Why would people interpret this behavior as intentional even if it lacked skill? According to the skill neglect interpretation, once people encounter a highly negative action, they ignore information about skill. Consequently, variation in skill has little to no impact on intentionality judgments for such actions.

Recall, however, the description of the agent's low skill in Knobe's vignette: "[Jake] raises his rifle, gets her in the sights, and presses the trigger. But Jake is not very good at using his rifle. His hand slips on the barrel of the gun, and the shot goes wild…" Although the agent is described as "not very good," his lack of skill does not manifest until *after* he has pressed the trigger. So there is still a basic action—pressing the trigger—that is performed skillfully. This act requires very little skill and is clearly intentional. According to Mele (2003), performing any intentional action (such as pulling a trigger) that leads to a morally charged outcome (e.g., someone's death) may be sufficient for people to deem the outcome intentional. Slipping on the barrel (and thus displaying lack of skill) may thus be irrelevant to people if the crucial intentional action of pressing the trigger has already been skillfully performed. Further, if pressing the trigger counts as killing, then killing is skillful (and intentional) as well.

It follows that if the initial act of pressing the trigger were portrayed as lacking skill, killing should lack skill as well, and intentionality judgments should drop as a result. Study 1 therefore compared the standard condition in the shooting vignette (which had *high* evidence of skill, as the agent first pulls the trigger and slips after) with two alternate conditions that varied the skillful performance of the basic act of pulling the trigger. The first presented *moderate* evidence of skill (the agent slips *before* the act of pulling the trigger); the second presented *low* evidence of skill (there was no mention at all of the agent pulling the trigger). These new conditions should make participants doubt the skillful performance of the basic action of pulling the trigger and, by hypothesis, lower the rate of intentionality judgments for the critical action of killing the aunt. According to the skill neglect claim, by contrast, skill is irrelevant to intentionality for highly immoral actions, so intentionality judgments should be equivalent across the three conditions.

2.1. Method

In Knobe's original vignette, the agent's lack of skill was demonstrated at both the dispositional-level ("Jake isn't very good at using his rifle") and the performance-level ("His hand slips on the barrel of the gun and a shot goes wild"). For two reasons, our new conditions in Study 1 (and, typically, those in the remaining studies we present) focus on the latter type of skill. First, performance-based skill can be manipulated more precisely in an experimental context. Second, and most important, we suspect that this type of skill—regarding how a particular action is performed—is most relevant for intentionality judgments (which are, after all, made with respect to particular actions). Dispositional-based skill is relevant only insofar as it informs performance-based skill (i.e., unskilled agents tend to perform actions unskillfully).

Participants were 225 undergraduate students who completed a larger computer-presented questionnaire and received partial course credit in return. Each was randomly assigned to one of three conditions. In the "high skill" condition (identical to Knobe's original vignette), the agent "raises his rifle, gets her in the sights, and pulls the trigger.

But Jake isn't very good at using his rifle. His hand slips on the barrel of the gun and the shot goes wild." In the "moderate skill" condition, Jake's "hand slips on the barrel of the gun even before he pulls the trigger." To provide a tougher test, this new *moderate skill* condition omitted the sentence that mentions Jake's dispositional lack of skill ("Jake isn't very good at using his rifle"). Thus, only the slipping before pulling the trigger could cast doubt on the agent's skill and reduce intentionality judgments. In the "low skill" condition, Jake "raised his rifle and got her in the sights, but his hand slipped on the barrel of the gun and a shot went off." In all three conditions, the bullet hits the aunt directly in the heart and she dies instantly. All participants then answered with Yes or No the standard intentionality question, "Did Jake kill his aunt intentionally?"[1]

2.2. Results

When Jake slipped after he pulled the trigger (high skill), 93% said that the act of killing was intentional; when he slipped before he pulled the trigger (moderate skill), 71% said the killing was intentional. When there was no mention of the trigger being pulled (low skill), just 42% said the killing was intentional. Intentionality ratings, therefore, differed substantially across conditions, χ^2 (2, $N = 225$) = 51.5, $p < .001$. A loglinear analysis showed that the *high skill* condition differed from the *moderate skill* condition, which differed from the *low skill* condition, zs = 3.7 and 2.9 (both ps < .01), ds = .61 and .56, respectively.

2.3. Discussion

In contrast to the skill neglect claim, even when the outcome is held constant (and is highly negative), people's intentionality judgments are sensitive to variations in skill. In particular, people are notably less likely to call the action intentional when the agent's lack of skill becomes evident earlier in the causal chain of events and has a causal effect on the action. These findings suggest that one reason for the high intentionality ratings in Knobe's negative/low skill condition (and the *high skill* condition in this study) was that the purportedly "unskilled" agent successfully performed a basic action that led to the killing—namely, pressing the trigger. The agent had enough skill to perform this basic action intentionally, and to the extent that pulling the trigger counts as killing, the agent had enough skill to intentionally kill. In the present study, however, when the act of pressing the trigger was less controlled (*moderate skill* condition) or was absent (*low skill* condition), the killing was no longer constituted by a skillful basic action and its perceived intentionality dropped substantially. It is worth noting that, to minimize demand characteristics, we did not tell participants outright that Jake failed to pull the trigger in the *low skill* condition. Thus, some people may have believed that Jake did pull the trigger (similar to the *moderate skill* condition), which may account for

---

[1] The exact formulation of all vignette materials used in the current studies can be viewed online at http://research.clps.brown.edu/SocCogSci/skill/.

the fact that 42% still judged the act of killing intentional. We return to this issue in Section 7.

We conducted Study 2 to replicate Study 1's finding that intentionality judgments vary with the presence or absence of a skillful basic action (pulling the trigger) and to extend this finding to a slightly different vignette content. In one condition, mimicking Knobe's (2003b) original skill scenario, the agent was an expert marksman and performed the basic action of pulling the trigger. In the other condition, the agent had no experience with guns (as in Knobe's (2003b) low-skill scenario) but, more important, he was not described as skillfully performing the basic action of pulling the trigger (as in the *low skill* condition of Study 1: "his hand slipped on the barrel of the gun and the shot went wild"). Thus, the conditions are parallel to Knobe's two conditions in terms of dispositional skill but differ in the critical presence or absence of the basic intentional action. Skill neglect implies that, despite differences in dispositional and performance skill, people will judge the action equally intentional across conditions (since skill is not a necessary condition for intentionality) whereas the standard model predicts that substantially fewer people will regard the scenario as intentional when no basic skillful action was performed.

3. Study 2

3.1. Method

Participants were 34 adults who were spending time in public places in Providence, RI. They were approached by the experimenter and asked if they would be willing to complete a brief questionnaire. In the scenario, the protagonist, Frank T., had an ongoing dispute with his neighbor over a stretch of land and one day decided to shoot his neighbor in the body. In the *high skill* condition, "Frank T. was an expert marksman, so he pulled the trigger and directly hit the neighbor's body, causing significant injury." In the *low skill* condition, "Frank T. had no experience with guns, his hand slipped on the barrel of the gun, and the shot went wild. Nonetheless, the bullet bounced off a large boulder several feet away and hit the neighbor's body, causing significant injury."

Participants answered an intentionality question ("Did Frank T. intentionally shoot his neighbor in the body?") and a blame question ("How much blame does Frank T. deserve for his action?" on a 0–6 scale).[2]

3.2. Results

People were far more likely to judge that Frank intentionally shot his neighbor in the body when Frank had skill (100%) than when he had low skill (47%), χ^2 (1, $N = 34$) = 13.25, $p < .001$, $d = 1.17$. Blame ratings were higher in the *high skill* condition ($M = 5.53$) than in the *low skill* condition ($M = 4.50$), $t(33) = 2.19$, $p < .05$, $d = .74$. However,

within the *low skill* condition, blame ratings were high regardless of whether people judged the shot as intentional, $r = -.01$.

3.3. Discussion

Study 2, like Study 1, showed that assessments of intentionality are guided by information about skill, even for actions that are highly immoral (see Fig. 2). Whether killing (Study 1) or shooting someone's body (Study 2), the less skillfully the action was performed, the fewer people deemed the action intentional. In particular, when the skillful performance of the basic action of pulling the trigger was doubtful (as in the *low skill* conditions of Studies 1 and 2), fewer than 50% of participants viewed the action as intentional. In contrast, when this basic action was performed with skill (as in the *high skill* conditions of Studies 1 and 2), nearly everyone judged the action intentional. A follow-up study[3] further showed that people's graded judgments of an agent's skill predicted their assessments of intentionality ($r = .35$, $p = .06$). Thus, our findings contradict the skill neglect claim (that, for immoral actions, skill is not a necessary condition of intentionality) and instead suggest that people are keenly aware of variations in skill—as long as such variations are not undermined by skillful basic actions.

Rejection of the skill neglect claim would be even more convincing, however, if no basic action intervened in the scenario in the first place. In Study 3 we designed a vignette with a more straightforward action performance. Moreover, we sharpened the agent's intention to avoid any ambiguity about effort or commitment (see Malle & Knobe, 2001), and we included a positive valence condition.

4. Study 3

Knobe (2003b) focused on the claim that morally blameworthy acts are considered intentional even when performed without skill. However, in one study, he examined a morally praiseworthy act—a soldier's aiming at a military device, hitting it, and thereby saving a number of lives. In this study, 92% of people found that even the "unskilled" soldier hit the device intentionally, just as the unskilled shooter killed his aunt intentionally. Thus, Knobe actually endorsed a general skill neglect claim, whereby people judge unskilled actions intentional if they are morally *significant*, whether negative or positive. We therefore extended our examination of skill neglect to include both a negative and a positive case, but within exactly the same context (in contrast to Knobe's comparison of an assault,

---

[2] Participants also answered a difficulty question ("How challenging is a shot at someone's body?" on a 0–6 scale), which showed that the actions were equivalent in difficulty and any differences must be due to the displayed skillful action performance.

[3] In this small study ($N = 31$) we tested whether people's explicit ratings about the agent's level of skill were predictive of their intentionality judgments. All participants read a vignette nearly identical to the *low skill* vignette from Study 2 and, in addition to answering questions about intentionality and blame, they rated Frank's skill ("How much skill does Frank T., have in using a rifle?"). Although skill judgments were quite low overall, they were strongly related to intentionality judgments, $r(30) = .35$, $p = .06$. The more skill people thought Frank had, the more likely they were to say that he intentionally hit the neighbor's body (even when controlling for people's blame ratings, this effect remained: $r(29) = .28$, $p < .08$).
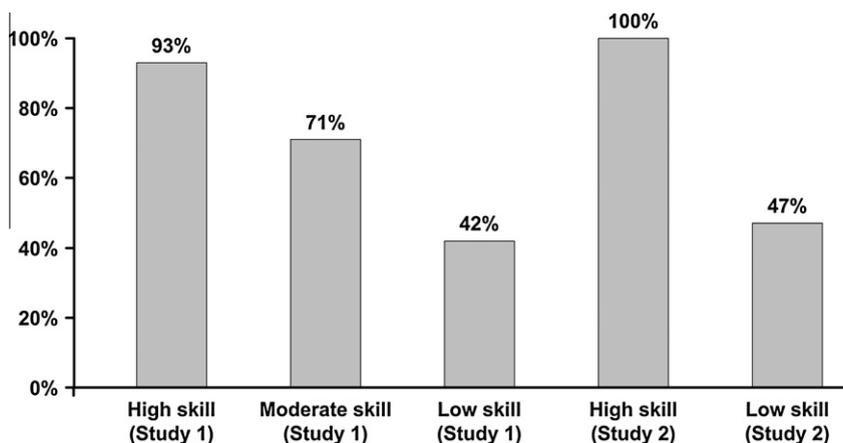
Fig. 2. Intentionality judgments as a function of skillful performance of a basic action (pulling trigger) in the course of performing an immoral action.

a contest, and a military action). That way, we ensured that the positive and negative actions were of identical difficulty and specificity.

The design of Study 3 thus crossed skill vs. low skill with positive vs. negative moral valence and assessed intentionality judgments (this time on a rating scale). The skill neglect claim predicts high intentionality ratings for both the low-skill negative and the low-skill positive actions—just as in Knobe's original findings. Conversely, the standard intentionality model predicts higher intentionality ratings for actions with skill than for actions with low skill, irrespective of moral valence.

4.1. Method

Participants were 120 undergraduate students who completed a larger computer-presented questionnaire and received partial course credit in return. We created a story context in which the morally significant action was more representative of everyday interaction, in which meanness or benevolence are more frequent than taking or saving lives. The background story common to all conditions was that George and his sister Lena reunite at their parents' house for Thanksgiving. Whereas George just got into medical school, Lena is unhappy in her marriage and recently lost her job. Over the course of the day, George and Lena get into a number of heated arguments. Later in the afternoon they play a game of darts. They split the first two games, and the third game is close until the end. "Who will win comes down to George's last shot. If he hits 11 or more, he wins; if he hits less than 11, Lena wins." An image of a dart board was included in the vignette.[4]

We manipulated moral valence by varying George's goal and its outcome. In the positive condition, "George

thinks of the difficult time Lena is having, and he really wants to let her win." After his shot, "the dart lands in the 6-point segment. Lena wins the game and is very happy." In the negative condition, "George doesn't care that Lena is having a difficult time; he really wants to beat her." After his shot, "the dart lands in the 14-point segment. George triumphs in his victory while Lena is sad."

The actor's skill was manipulated within the action performance. In the low-skill condition, "George has little control over his shots. He tries to aim at a region of the darts board in which he would probably make [fewer/more] than 11 points. As he sets up his shot, however, he loses his balance, the dart slips out of his hand and wobbles toward the board, and... the dart lands in the [6-point/14-point] segment." In the skill condition "He aims the dart at a region of the board in which he would probably make [fewer/more] than 11 points. He sets up his shot and... the dart lands in the [6-point/14-point] segment.

Participants first answered the intentionality question, "Did George hit the [6-point/14-point] segment intentionally?", on a 4-point scale from 0 (Certainly no), 1 (Probably no), 2 (Probably yes) to 3 (Certainly yes). The scale was presented as continuous to make the judgment somewhat easier for participants, but for analysis purposes, the two no answers and the two yes answers were combined. (A continuous analysis leads to identical results.) Participants also answered an evaluation question: "Do you approve or disapprove of George's action?" on a 4-point scale from −2 (Very much disapprove), −1 (Somewhat disapprove), +1 (Somewhat approve), to +2 (Very much approve).

Finally, we devised an intuitively meaningful manipulation check for the skill factor that also served as a check for the equality of the actions' difficulty: "Their cousin Frank feels that George's shot was lucky. He challenges George to hit the [14-point/6-point] segment again. How likely is it that George will hit the [14/6] again?" The response scale ranged from 0 (Extremely unlikely) to 5 (Extremely likely). The skill manipulation should result in reliable differences on this likelihood rating (with the skilled agent being more likely than the unskilled agent to successfully perform the action again). However, the difficulty of the negative and

---

[4] The cutoff of 11 was chosen because it splits the 20 possible segments in two halves and makes the over/under appear equally difficult. Strictly speaking, in many games of darts the 11 would not be the midpoint of the *points* one can achieve (because double and triple scores increase the possible points all the way to 60), but we refrained from explaining this complication to the participants. We ran a variant of this study in which the over/under cutoff was 20 points, which is close to the median of the entire possible point range, and we found the same results.

positive actions should be equivalent, so the likelihood ratings should not differ across valence.

4.2. Results

The valence manipulation check showed that people approved of the benevolent action more ($M = 1.2$) than of the mean action ($M = 0.3$), $F(1, 114) = 25.5$, $p < .001$, $d = .94$. By contrast, valence did not have an impact on the agent's likelihood of repeating the same shot ($F = 1$, $d = .18$), suggesting that people viewed the positive and the negative act as equally difficult. The skill manipulation, however, had the expected effect on the likelihood rating, with the unskilled agent being less likely to repeat the same shot ($M = 1.5$) than the skilled agent ($M = 2.2$), $F(1, 116) = 16.5$, $p < .001$, $d = .73$.

The pattern of intentionality judgments contradicted the skill neglect claim and favored the standard model of intentionality. People more often judged the action intentional when the agent had skill (85%) than when he did not (27%). Importantly, the effect of skill on intentionality judgments was as strong for the positive action (83% vs. 30%, $d = .61$) as for the negative action (87% vs. 23%, $d = .84$). In a loglinear analysis predicting intentionality judgments from skill and valence, the main effect of skill was substantial and significant, $z = 5.9$, $p < .001$, $d = .76$, whereas the main effect of valence was not, $z = 0.1$, $d = .01$, nor was the valence × skill interaction, $z = 0.63$, $d = .08$ (see Fig. 3).

4.3. Discussion

The design of Study 3 assured that there was no basic action performed intentionally (akin to pulling the trigger) that could count as the broader action in question, and all actions were equated for difficulty. Thus, the only relevant difference between the positive and the negative action in Study 3 was moral valence. Yet no undue influence of moral evaluation on intentionality judgments emerged. Intentionality varied systematically as a function of skill for both positive and negative actions.

One possible concern is that the degree of disapproval elicited by the negative valence condition was not overwhelming, as only 40% of participants in this condition gave George's action an explicit disapproval rating. It is therefore possible that the negative act was not blameworthy enough

Fig. 3. Intentionality judgments are sensitive to skill but not to moral valence.

to elicit any skill neglect. To address this possibility, we repeated all analyses on only those participants in the negative valence condition who assigned a negative rating to George on the (dis)approval scale ($N = 24$ out of 60). Within this subsample, the results were just as strong: 77% saw the negative action as intentional when performed with skill, 9% saw it as intentional when performed without skill.

However, we still have a puzzling finding to address. We have argued that most people in Knobe's (2003b) original study viewed the immoral action (killing the aunt) as intentional because there was skillful and intentional basic action—pulling the trigger—that counted as the broader action of killing. But in the neutral condition, the agent performed the same basic action ("He raises the rifle, gets the bull's-eye in the sights, and presses the trigger...The bullet lands directly on the bull's-eye. Jake wins the contest."), yet just 28% judged the action of hitting the bull's-eye as intentional. If intentionally pressing the trigger counts as intentionally killing the aunt in the negative condition, then why does it not count as intentionally hitting the bull's-eye in the neutral condition? Here the moral bias claim comes into play. Setting aside the basic action of pulling the trigger (which was constant across negative and neutral actions in Knobe, 2003b), the moral bias claim posits that people saw unskilled neutral behaviors as less intentional than unskilled negative behaviors *because of* the latter's moral valence.

By contrast, we propose that, in addition to their moral valence, the two conditions in Knobe (2003b) also differed in the difficulty of the action in question. Whereas "hitting the bull's-eye" specifies a single precise condition for success (having the bullet land directly on the small bull's-eye), "killing the aunt" offers a greater number of success conditions—the agent could have hit many different body parts or used many different methods of killing his aunt. Therefore, we hypothesize that the act of killing was seen as relatively easy (even for the unskilled protagonist) because it allowed for many lethal ways of accomplishing it, whereas the act of hitting the bull's-eye was seen as relatively difficult because it allowed for only one appropriate way of accomplishing it. If so, then we might be able to explain the puzzling pattern of intentionality judgments by recognizing that the low-skill agent actually had sufficient skill to intentionally perform the easy negative action but not enough skill to intentionally perform the difficult neutral action.

5. Study 4

Study 4 directly tested the hypothesis that, in the typical skill challenge findings, killing is judged as easier than hitting the bull's-eye. In addition, Study 4 attempted to vary the actions' difficulty. This aim was met by asking people to make two intentionality judgments: the standard one about the act of *killing* (which we assume to be general and easy) and a new one about the act of *hitting the aunt in the heart* (which we assume to be specific and difficult). Knobe (2003b) observed an intentionality asymmetry when comparing a general negative act (killing) with a specific neutral one (hitting the bull's-eye). Study 4 additionally compares
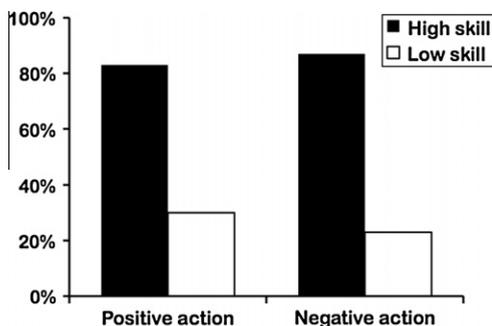
the specific neutral action with a specific negative action (hitting the aunt's heart) and examines whether any intentionality asymmetry remains once the specificity of the actions is held constant (i.e., comparing the act of *hitting* the heart to that of *hitting* the bull's-eye). According to the moral bias claim, the negative specific act should more often be judged intentional than the neutral specific act.

5.1. Method

Participants were 120 undergraduate students who completed a larger computer-presented questionnaire and received partial course credit in return. Two participants had missing values and were excluded. Each read either the low-skill negative or low-skill neutral vignettes from Knobe's (2003b) original study, then answered a difficulty question ("How challenging was the shot that James took?") and a skill question ("How much skill do you think James had when he took the shot?"), both on a 0–6 scale. Participants in the negative condition then answered a specific intentionality question ("Did James hit his aunt in the heart intentionally?") and the standard general intentionality question ("Did James kill his aunt intentionally?"). To keep the conditions as similar as possible, participants in the neutral condition also answered a specific intentionality question ("Did James hit the bull's-eye intentionally?") followed by a more general one ("Did James win the contest intentionally?").

5.2. Results

Skill judgments, as expected, were quite low overall and did not differ notably between the neutral ($M = 1.57$) and negative conditions ($M = 1.22$), $t(118) = 1.59$, $p > .10$, $d = .30$. Difficulty, however, varied by valence. As expected, people viewed taking the shot to kill as easier ($M = 2.8$) than the shot to hit the bull's-eye ($M = 3.8$), $t(116) = 2.79$, $p < .01$, $d = .51$.

The pattern of intentionality judgments showed a striking pattern. Within the negative condition, the general question about killing elicited almost perfect intentionality rates (98%) whereas the specific question about hitting the aunt in the heart elicited very low intentionality rates (27%), $z(59) = 6.4$, $p < .001$, $d = .83$.[5] Thus, the identical action elicited remarkably different intentionality judgments depending on whether it was described as "killing" or "hitting in the heart."

Within the neutral condition, the question about winning elicited moderate intentionality judgments (43%), and the question about hitting the bull's-eye did not differ from this rate (38%), $p > .10$. This is understandable since the story itself explicitly said that there was only one way of winning the contest—by hitting the bull's-eye—so the two actions were of comparable specificity.

Finally, we compared intentionality rates for the two specific (and comparably difficult) acts—the neutral one of hitting the bull's-eye (38%) and the negative one of hitting the aunt's heart (27%)—and they did not differ from each other, $p > .10$, $d = .17$. If anything, the negative act elicited slightly lower intentionality rates.

5.3. Discussion

The results suggest that Knobe (2003b) and other researchers confounded difficulty with valence when they used the verb phrases *killing the aunt* and *hitting the bull's-eye*. With this choice of verbs, people see the negative action of killing as easier to accomplish than the neutral one of hitting the bull's-eye. For an agent with low skill, an easier action has a greater chance of being performed intentionally because even a low level of skill may be sufficient for its successful performance. Thus, the standard negative condition in past studies yielded more intentionality judgments than the neutral condition because the former was easier, not because it was morally significant. When we correct for this confound and equate the specificity and difficulty of the actions at issue—comparing hitting the aunt's heart with hitting the bull's-eye—intentionality judgments for the negative action drop substantially. Only 27% of participants deemed the negative action of hitting the aunt's heart intentional, as low a rate as the neutral condition in Knobe's (2003b) original demonstration and its various replications (Nadelhoffer, 2004, 2005). This specificity effect is consistent with Nadelhoffer's (2004) findings, which showed that when an agent rolled a six on a die to detonate a bomb and kill his enemy, people viewed the general act of killing more often as intentional than the specific act of rolling a six.

Thus far, our results have revealed two important patterns. Studies 1 through 3 showed that, even for immoral actions, an agent's skill informs judgments about whether the agent's action was intentional. Specifically, when a basic action (such as pulling the trigger) is performed unskillfully, intentionality judgments about the negative outcome (killing, hitting the body) drop markedly. Moreover, Study 4 showed that the difficulty or specificity of an action is critical for intentionality—whereas people may view very general negative actions (e.g., killing) as intentional, they are far less inclined to view specific negative actions (e.g., hitting the heart) as intentional.

Combining these insights leads to a final prediction that counters the skill challenge. If the agent's action is truly unskilled (e.g., failing to press the trigger) *and* if people are asked about a specific action (e.g. hitting the heart), intentionality judgments should become essentially nonexistent, even when the agent remains highly blameworthy. Study 5 tested this prediction.

6. Study 5

6.1. Method

Participants were 58 undergraduate students who completed the study as the initial task in a larger study on

---

[5] These within-subject dichotomous data can be analyzed in a variety of ways, from nonparametric tests to repeated-measures ANOVA (assuming that underlying the dichotomous intentionality judgment is a continuous state of confidence). All analyses lead to the same results, and we are reporting here the more conservative Wilcoxon's signed rank test and its corresponding effect-size estimate.

moral cognition. Each received partial course credit in return. All participants read the *low skill* vignette from Study 1, in which the agent's hand "slipped on the barrel of the gun and a shot went off." However, participants answered intentionality questions about both the general action ("Did Jake kill his aunt intentionally?") and the specific action ("Did Jake shoot his aunt in the heart intentionally?"). Participants then rated the difficulty of each action ("How challenging was it for Jake to kill his aunt [shoot his aunt in the heart]?" on a 0–5 scale).[6] Finally, they rated blame ("How much blame does Jake deserve?" on a 0–5 scale).

6.2. Results

As predicted, the act of shooting the aunt's heart was seen as more difficult ($M = 2.45$) than the act of killing ($M = 1.40$), $t(57) = 4.14$, $p < .001$, $d = .57$. Difficulty ratings were associated with intentionality in the expected direction—the greater the perceived difficulty, the less the action was deemed intentional—but this pattern was not significant, $r = -.13$, $p > .10$.

Most important, whereas just over half of participants judged the general act of killing as intentional (62%), hardly anyone judged the specific act of hitting the aunt's heart as intentional (10%), $z = 5.0$, $p < .001$, $d = .66$. This result contradicts the skill neglect claim. In addition, blame ratings were high overall ($M = 4.3$), and, although blame ratings correlated strongly with intentionality judgments about killing ($r = .54$, $p < .01$), they correlated only weakly with intentionality judgments about hitting the heart ($r = .17$, $p > .10$). Even among the participants who assigned a great deal of blame (i.e., ratings of 4 or 5), only 13% viewed the specific act of hitting the heart as intentional. This result speaks strongly against the moral bias claim, for on that account, judging an action as blameworthy should lead people to judge it intentional.

Comparing Study 4 and Study 5, we now see that within each level of action specificity people are sensitive to performance-based skill manipulations, contradicting the hypothesis of skill neglect (see Fig. 4). Consider first the general act of *killing*. When the agent performed the skilled action of pulling the trigger (Study 4), 98% of people deemed his killing intentional; when no such skilled action was described (Study 5), 62% of people deemed the killing intentional, $\chi^2 = 24.7$, $p < .001$. (This drop was even greater in Study 1—from 93% to 42%.) By contrast, the specific act of hitting the heart was considered intentional by far fewer people. But even this action was more often deemed intentional when the agent pulled the trigger (27% in Study 4) than when no such action was mentioned (10% in Study 5), $\chi^2 = 5.18$, $p < .05$.

Fig. 4. Sensitivity to skill within each level of action specificity (combining Studies 4 and 5).

7. General discussion

7.1. The challenge

Knobe (2003b) challenged traditional theories of intentionality and moral judgment by suggesting that people judge morally significant behaviors as intentional even if a component of intentionality (such as skill) is missing—quite in contrast to neutral behaviors, which are judged unintentional when missing a component. The moral valence (and especially blameworthiness) of a behavior, he proposed, guides perceptions of intentionality—not the other way around, as was traditionally assumed. He provided evidence for this hypothesis by showing that intentionality judgments for an unskilled act of killing are frequent whereas intentionality judgments for an unskilled act of hitting the bull's-eye are infrequent. Similar intentionality discrepancies between valenced and neutral actions were reported by Nadelhoffer (2004, 2005).

These provocative "skill challenge" findings appeared to support two related claims. First, when assessing the intentionality of morally significant (e.g., blameworthy) actions, people neglect information about skill, which is typically a necessary condition of intentionality (Malle & Knobe, 1997). Second, at a given level of skill, people judge such actions as more intentional as a direct consequence of regarding them as blameworthy.

7.2. Answering the challenge

Our studies have examined these two claims derived from the skill challenge, and, in doing so, have identified two nonmoral mechanisms that account for the original findings. The first is that broader actions (such as killing) can be accomplished by basic actions (such as pulling the trigger), and performing a basic action intentionally can count as performing the broader action intentionally. Studies 1–3 showed that when this basic action is performed unskillfully or is absent, intentionality judgments drop substantially. The second mechanism is that more general actions are less difficult and therefore require less skill to be performed intentionally. Study 4 showed that Knobe's (2003b) original actions (killing and hitting a bull's-eye) differed not just in moral valence but also in specificity

---

[6] The order of the two intentionality questions, along with the order of the two difficulty questions, was randomized for each participant. There was a marginally significant order effect on intentionality judgments about the specific action, such that people were slightly more likely to say that Jake intentionally shot his aunt's heart when this question was asked first (19%) than when it was asked second (3%), $\chi^2(1) = 3.64$, $p < .10$. There was no order effect on the general intentionality question, the difficulty questions, or the blame question.
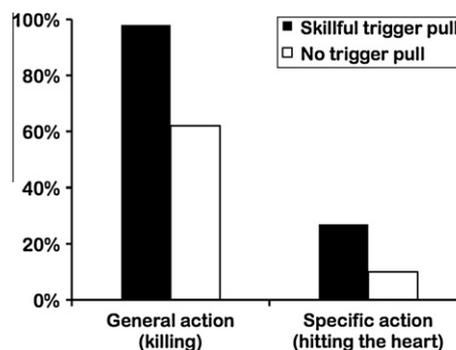
and hence difficulty. When we held specificity constant, people deemed morally negative actions intentional just as often as morally neutral ones. Study 5 combined these two mechanisms, showing that when the skillful basic action was absent *and* people were asked about the specific act of hitting the heart, intentionality judgments were essentially non-existent, even though people assigned a great deal of blame.

7.2.1. An integrated explanation

The two mechanisms we have identified can be integrated within a framework of an action's *scope*, which refers to the ease and number of ways in which the action can be accomplished (cf. Goldman, 1970; Wegner & Vallacher, 1986). An action with wide scope (e.g., work, talk) has two features: First, it can be accomplished by a variety of basic actions that count as performing that wide-scope action. Second, among the basic actions that constitute a wide-scope action, some basic actions will be relatively easy, so the wide-scope action also becomes easier and can be performed intentionally even with a low level of skill. By contrast, an action with narrow scope (e.g., simmer the sauce, break the 100-meter record) can be accomplished only by very few basic actions and requires considerable skill to do so.

In each of Knobe's (2003b) studies, a morally significant action (killing, saving lives) was compared with a morally neutral action (hitting the bull's-eye), but both morally significant actions had wide scope whereas the neutral action had narrow scope. As a result, killing and saving lives could be accomplished by a relatively easy basic action (pressing the trigger), which allowed the agent to perform these actions intentionally even at low levels of skill. Hitting the bull's-eye, by contrast, was not sufficiently constituted by pulling the trigger and was seen as more difficult, thus preventing the agent with low levels of skill to perform the action intentionally. The low-skill gunman in Knobe's (2003b) studies did not have enough skill to intentionally hit the bull's-eye, but he had enough skill to intentionally kill.

The present studies therefore disconfirm the two claims of the skill challenge—that people neglect skill when assessing the intentionality of morally significant behavior and that people's moral evaluations bias their intentionality judgments. But beyond that, they teach us several lessons about both intentionality judgments and moral judgment.

7.3. Lessons learned

7.3.1. Skill and difficulty

Our results document not only that people are sensitive to levels of skill when judging intentionality but that they integrate the agent's skill with the action's difficulty. The amount of skill necessary for an action to be intentional is not some absolute amount but is lower for easy actions and higher for more difficult actions. This interplay between skill and difficulty was first described by Heider (1958) as the concept of *can*, which he defined as follows:

$$can = f \text{ (person's ability – difficulty of environmental factors).}$$

We may thus sharpen Malle and Knobe's (1997) model of the folk concept of intentionality by replacing the *skill* component with the *can* component. What is needed for a behavior to be seen as intentional is that the agent *can* perform the behavior (in a reliable and controlled manner), so she must have enough skill to overcome the given action's difficulty.

7.3.2. Blame and intentionality

Judgments of blame are not as strongly tied to intentionality judgments as was previously claimed by the dominant skill neglect and moral bias interpretations (Alicke, 2008; Knobe, 2003b). In studies 2 and 5 we probed both judgments, and blame correlated with intentionality of the general act of killing on average at $r = .37$. In Study 5 we also asked about the specific act of hitting the aunt's heart, which was even more weakly correlated with blame ($r = .17$). Indeed, most people assigned a great deal of blame but hardly anyone judged the action intentional. This is because people can incur substantial blame even for unintentional actions, particularly when they intend to cause negative outcomes or when they knowingly (and perhaps willingly) allow such outcomes to occur (Guglielmo & Malle, in press; Guglielmo et al., 2009). Admittedly, in the current studies, people typically examined actions of extreme valence (murdering or injuring another person), which restrict the range of blame judgments. However, Study 3 examined actions of more moderate valence and showed that blame and intentionality only correlated at $r = .19$. In fact, blame was most sensitive to variations in moral valence whereas intentionality was most sensitive to variations in skill. Of course, when we generally compare blame for behaviors that are clearly intentional with blame for behaviors that are clearly unintentional, the two judgments will be highly correlated. However, the present studies suggest that such a correlation holds because intentionality guides blame, not the other way around.

7.3.3. Levels of action

People simultaneously assess intentionality for basic and for broader actions and integrate the two as a function of the action's scope. In the case of wide-scope actions, the intentionality of any suitable basic action transfers to the broader action, whereas in the case of narrow-scope actions, such a suitable basic action is hard to find.

The notion that people respond differentially to behaviors with narrow or wide scope finds kinship with two related literatures. Action identification theory (e.g., Vallacher & Wegner, 1987) posits that any given action can be identified in multiple ways, from its low-level performance features to its high-level meaning. Studies have shown that easier actions are conceptualized at a higher level whereas more difficult actions are conceptualized at a lower level (Vallacher, Wegner, & Somoza, 1989), which parallels the inverse relation between difficulty and scope in the present studies. Similarly, the linguistic category model (Semin, 2008) suggests that speakers use more or

fewer abstract verbs to influence others' impressions of an individual or group. Descriptive action verbs such as " stare" or "push" denote specific actions of shorter temporal duration that are easily verifiable, whereas interpretive action verbs such as "deride" or "harm" denote broader classes of actions that are more difficult to verify, more open to interpretation (Semin & Fiedler, 1988). In line with this perspective, we suggest that narrow-scope descriptions use descriptive action verbs (e.g., hitting the bull's-eye) to refer to an action's concrete performance that is temporally circumscribed and easily verified—as a result, criteria of intentionality apply precisely. Wide-scope descriptions use interpretive action verbs (e.g., kill) to refer to the action's broader meaning, which is more difficult to verify and leaves more room for interpretation of properties such as intentionality.

The wide scope of killing may help explain why 40–60% of people (in Studies 1 and 5) still viewed this action as intentional, even when the pulling of the trigger was not explicitly mentioned. Killing may be so general that the manner in which one achieves the goal may be nearly irrelevant for intentionality judgments. Simply committing to an action plan and exhibiting a minimal amount of skill in its implementation (e.g., by aiming a gun and putting one's finger over the trigger) may be sufficient for accomplishing the wide-scope act of killing. Thus, considerations of morality may play a role in people's judgments of intentionality, not by directly changing perceptions of intentionality, but by inviting people to describe (im)moral actions with wide-scope verbs (e.g., kill, harm, benefit). Such verbs portray actions as more flexible in their manner of fulfillment, and people are more likely to regard such actions as intentional.

7.3.4. Causal deviance

The present findings help clarify the role of causal deviance in moral judgment (e.g., Chisholm, 1966). Previous studies have shown that judgments of both blame and praise are weakened when an agent's behavior is causally deviant—when the agent's action outcome unfolds in a way that was not intended (Pizarro, Uhlmann, & Bloom, 2003). Study 1 demonstrated that causal deviance sometimes leaves both intentionality and blame untouched, namely, when the deviance occurs after intentionally performing a basic action that counts as the broader action at issue. However, if causal deviance interferes with the basic action and indicates a lack of control, the broader action is less likely to be deemed intentional, even if it results in a severely negative outcome (see Mele & Cushman, 2007 for similar findings). Moreover, only if people view the causally deviant action as unintentional will they reduce assignments of blame.

7.3.5. Jury decision making

Knobe's (2003a, 2003b) challenges to the standard intentionality model threaten the good sense of asking people to make mens rea judgments in legal proceedings (Nadelhoffer, 2006b). We should be wary of jurors' assessments of a defendant's intent if the negative valence of the presumed criminal act biases them toward "seeing" such intent. However, as the present findings meet the skill challenge, and those reported elsewhere meet the side-effect challenge (Guglielmo & Malle, in press), we can maintain our confidence in jurors' ability to handle mens rea judgments (Malle, 2006; Malle & Nelson, 2003).

Nonetheless, extant research highlights the conditions under which such judgments can go awry, since, sadly, some of the methodological problems in Knobe's and other researchers' studies are mirrored by procedural problems in the courtroom. When cases are presented with vague, wide-scope action verbs, generally unfavorable background information about the defendant (see Alicke, 2000), and limited judgment categories (e.g., intentionality—yes or no?; see Guglielmo & Malle, in press), jurors will not be at their best in assessing the intentionality of a defendant's action. Conversely, such assessments will be more reliable with precise formulations of the actions in question, careful selection of what background is relevant (i.e., information that helps infer the mental states at hand, not general character), and permission to use multiple conceptual categories to describe the action in question (e.g., willingly, knowingly, with luck).

7.3.6. Limitations and future directions

One limitation of our studies is the reliance on especially negative behaviors, usually murder or severe injury. Given that such scenarios were the source of the skill challenge (Knobe, 2003b; Nadelhoffer, 2004, 2005), we had to make them our primary focus in order to fully account for the phenomenon. However, Study 3 hints at the worthwhile aim for future research to explore more fully the relationship between difficulty, skill, intentionality, and blame/praise across a range of morally negative and positive behavior.

Second, whereas our findings counter the claim that judgments of moral valence guide intentionality judgments, we did not assess the temporal properties of these judgments. For highly negative behaviors, blame and intentionality will often co-occur, and disentangling their temporal order and causal direction remains challenging. One research strategy for such disentangling is to examine people's reaction times of making the various judgments in response to an observed behavior. We are currently conducting studies that use such a reaction time approach and offer promise for further understanding the interplay between blame and intentionality (Guglielmo & Malle, 2008).

The literature will benefit from further study of people's spontaneous conceptualizations and descriptions of morally significant behaviors. As we have shown, intentionality judgments are highly sensitive to the precise way in which a given action is described. But how do people themselves describe these actions? A negative behavior is less likely to be seen as intentional when framed as having narrow (as opposed to broad) scope. Speakers may therefore opt to describe such behaviors as narrow or broad depending on whether they want to convince their audience of the action's intentionality or unintentionality. We suspect that, in the courtroom, attorneys for the prosecution will favor broad descriptions to heighten the impression of intentionality, whereas attorneys of the defense will favor narrow descriptions to weaken the impression of intentionality.

Previous studies (Knobe, 2003b; Nadelhoffer, 2004, 2005) mimicked the prosecution. We hope that our studies have given justice to both sides and have demonstrated the subtlety with which ordinary people think about intentionality and blame.

References

Adams, F. (1986). Intention and intentional action: The simple view. *Mind and Language, 1*, 281–301.

Alicke, M. D. (2000). Culpable control and the psychology of blame. *Psychological Bulletin, 126*, 556–574.

Alicke, M. D. (2008). Blaming badly. *Journal of Cognition and Culture, 8*, 179–186.

Chisholm, R. (1966). Freedom and action. In K. Lehrer (Ed.), *Freedom and determinism*. New York: Random House.

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

Cushman, F., & Mele, A. R. (2008). Intentional action: Two-and-a-half folk concepts? In J. Knobe & S. Nichols (Eds.), *Experimental philosophy*. New York: Oxford University Press.

Goldman, A. (1970). *A theory of human action*. Englewood Cliffs, NJ: Prentice-Hall.

Guglielmo, S. & Malle, B. F. (in press). Can unintended side effects be intentional? Resolving a controversy over intentionality and morality. *Personality and Social Psychology Bulletin*.

Guglielmo, S., & Malle, B. F. (2008). *The priority of intentionality or blame: Evidence from a reaction time experiment*. Albuquerque, NM: Poster presented at the meeting of the Society for Personality and Social Psychology.

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

Hamilton, V. L. (1978). Who is responsible? Towards a social psychology of responsibility attribution. *Social Psychology, 41*, 316–328.

Heider, F. (1958). *The psychology of interpersonal relations*. New York: Wiley.

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

Knobe, J. (2003b). Intentional action in folk psychology: An experimental investigation. *Philosophical Psychology, 16*, 309–324.

Knobe, J. (in press). Person as scientist, person as moralist. *Behavioral and Brain Sciences*.

Malle, B. F. (2006). Intentionality, morality, and their relationship in human judgment. *Journal of Cognition and Culture, 6*, 61–86.

Malle, B. F., & Knobe, J. (1997). The folk concept of intentionality. *Journal of Experimental Social Psychology, 33*, 101–121.

Malle, B. F., & Knobe, J. (2001). The distinction between desire and intention: A folk-conceptual analysis. In B. F. Malle, L. J. Moses, & D. A. Baldwin (Eds.), *Intentions and intentionality: Foundations of social cognition* (pp. 45–67). Cambridge, MA: MIT Press.

Malle, B. F., Moses, L. J., & Baldwin, D. A. (2001). *Intentions and intentionality: Foundations of social cognition*. Cambridge, MA: MIT Press.

Malle, B. F., & Nelson, S. E. (2003). Judging *mens rea*: The tension between folk concepts and legal concepts of intentionality. *Behavioral Sciences and the Law, 21*, 563–580.

McCann, H. J. (2005). Intentional action and intending: Recent empirical studies. *Philosophical Psychology, 18*, 737–748.

Mele, A. R. (1992). *Springs of action: Understanding intentional behavior*. New York: Oxford University Press.

Mele, A. R. (2003). Intentional action: Controversies, data, and core hypotheses. *Philosophical Psychology, 16*, 325–340.

Mele, A. R., & Cushman, F. (2007). Intentional action, folk judgments, and stories: Sorting things out. *Midwest Studies in Philosophy, 31*, 184–201.

Mele, A. R., & Moser, P. K. (1994). Intentional action. *2, 28*, 39–68.

Mele, A. R., & Sverdlik, S. (1996). Intention, intentional action, and moral responsibility. *Philosophical Studies, 82*, 265–287.

Nadelhoffer, T. (2004). The Butler problem revisited. *Analysis, 64*, 277–284.

Nadelhoffer, T. (2005). Skill, luck, control, and intentional action. *Philosophical Psychology, 18*, 341–352.

Nadelhoffer, T. (2006a). Desire, foresight, intentions, and intentional actions: Probing folk intuitions. *Journal of Cognition and Culture, 6*, 133–157.

Nadelhoffer, T. (2006b). Bad acts, blameworthy agents, and intentional actions: Some problems for jury impartiality. *Philosophical Explorations, 9*, 203–220.

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

Nichols, S., & Ulatowski, J. (2007). Intuitions and individual differences: The Knobe effect revisited. *Mind and Language, 22*(4), 346–365.

Ohtsubo, Y. (2007). Perceived intentionality intensifies blameworthiness of negative behaviors: Blame–praise asymmetry in intensification effect. *Japanese Psychological Research, 49*, 100–110.

Pettit, D., & Knobe, J. (2009). The pervasive impact of moral judgment. *Mind and Language, 24*, 586–604.

Pizarro, D. A., Uhlmann, E., & Bloom, P. (2003). Causal deviance and the attribution of moral responsibility. *Journal of Experimental Social Psychology, 39*, 653–660.

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

Searle, J. R. (1983). *Intentionality: An essay in the philosophy of mind*. Cambridge, England: Cambridge University Press.

Semin, G. R. (2008). Language puzzles: A prospective retrospective on the linguistic category model. *Journal of Language and Social Psychology, 27*, 197–209.

Semin, G. R., & Fiedler, K. (1988). The cognitive functions of linguistic categories in describing persons: Social cognition and language. *Journal of Personality and Social Psychology, 54*, 558–568.

Shaver, K. G. (1985). *The attribution of blame: Causality, responsibility, and blameworthiness*. New York: Springer.

Sousa, P., & Holbrook, C. (2010). Folk concepts of intentional action in the contexts of amoral and immoral luck. *Review of Philosophy and Psychology, 1*, 351–370.

Tedeschi, J. T., & Reiss, M. (1981). Verbal strategies as impression management. In C. Antaki (Ed.), *The psychology of ordinary social behaviour* (pp. 271–309). London: Academic Press.

Thompson, S. C., Armstrong, W., & Thomas, C. (1998). Illusions of control, underestimations, and accuracy: A control heuristic explanation. *Psychological Bulletin, 123*, 143–161.

Vallacher, R. R., & Wegner, D. M. (1987). What do people think they're doing? Action identification and human behavior. *Psychological Review, 94*, 3–15.

Vallacher, R. R., Wegner, D. M., & Somoza, M. P. (1989). That's easy for you to say: Action identification and speech fluency. *Journal of Personality and Social Psychology, 56*, 199–208.

Wegner, D. M., & Vallacher, R. R. (1986). Action identification. In R. M. Sorrentino & E. T. Higgins (Eds.), *Handbook of motivation and cognition: Foundations of social behavior* (pp. 550–582). New York: Guilford.

Weiner, B. (1995). *Judgments of responsibility: A foundation for a theory of social conduct*. New York: Guilford.