Judgment Before Emotion: People Access Moral Evaluations Faster than Affective States

Corey Cusimano (cusimano@sas.upenn.edu)
Department of Psychology, University of Pennsylvania,
3720 Walnut St Philadelphia, PA 19104 USA

Stuti Thapa Magar (stuti_thapa_magar@brown.edu) and Bertram F. Malle (bfmalle@brown.edu)

Department of Cognitive, Linguistic, and Psychological Sciences, Brown University,
190 Thayer Street, Providence, RI 02912 USA

Abstract

Theories about the role of emotions in moral cognition make different predictions about the relative speed of moral and affective judgments: those that argue that felt emotions are causal inputs to moral judgments predict that recognition of affective states should precede moral judgments; theories that posit emotional states as the output of moral judgment predict the opposite. Across four studies, using a speeded reaction time task, we found that self-reports of felt emotion were delayed relative to reports of event-directed moral judgments (e.g. badness) and were no faster than person-directed moral judgments (e.g. blame). These results pose a challenge to prominent theories arguing that moral judgments are made on the basis of reflecting on affective states.

Keywords: affect, emotion, moral judgment, reaction time

Introduction

There is broad agreement that affective phenomena play an important role in moral cognition; there is widespread disagreement, however, over the particular role that affect plays. Many theories suggest that emotion acts as an input to moral judgment: that affective states help distinguish moral from non-moral events (Nichols, 2002), indicate the severity of the transgression (Haidt, 2001), or bias downstream cognitive processes (Alicke, 2000). In contrast to these "emotion-as-input" (em-in) models, "emotion-as-output" (em-out) models argue that, very often, considerations of rules, norms, risk or caused harm, and causal and mental information, guide moral judgments without any necessary causal precedence of affect (Huebner, Dwyer, & Hauser, 2009; Mikhail, 2011). According to these models, emotions are typically connected to moral judgments because as they motivate and scale our social responses.

The focus of our paper is on a prominent subset of the *em-in* theories, which claim that emotions precede and influence moral judgment through *felt affect*. For example, Schnall Haidt, Clore, and Jordan (2008) claim, "When making evaluative judgments, people *attend to their own feelings*, as if asking themselves: How do I feel about it?" (p. 1097, emphasis added). Similarly, Miller et al (2012) argue that "the likelihood of judging an action wrong is determined… by how upsetting you consider the action itself to be" (p 574). That is, moral judgments of some event are formed by recognizing and reporting one's emotional response to that event. This is why, according to these theories, the experience of negative affect on its own (i.e. absent appraisals of harm or risk) can yield negative moral evaluations (e.g. Haidt, 2001). The primary source of evidence for this comes from affect misattribution experiments in which inducing feelings of disgust (unrelated to the stimulus) both amplified the perceived wrongness of target behaviors (Schnall, et al., 2008; Cheng, Ottati, & Price, 2013) and appeared to cause ordinarily permissible behaviors to be judged as wrong (Wheatley & Haidt, 2005).

Though initially promising, many of the findings in favor of the *em-in* model of moral judgment have been called into question. First, it appears as though many moral judgments can be made absent any affective experience (Niedenthal, Rohmann, & Dalle, 2003) and, conversely, many strong emotional reactions occur without any corresponding moral judgment (Royzman, Goodwin, & Leeman, 2011). Additionally, the primary source of evidence for the causal role of felt affect has been called into question: a recent meta-analysis reports that, across dozens of experiments, there is not reliable effect of incidental disgust on moral judgment (Landy & Goodwin, 2015).

However, even though these findings are consistent with *em-out* models, a major challenge in assessing any of the theories regarding the role of emotion in moral judgment is the dearth of experimental paradigms that get at the heart of the causal primacy question—whether the routine causal sequence is, according to one set of theories, *event → emotion → moral judgment* or, according to the other set of theories, *event → moral judgment → emotion*. What is needed are independent and time-locked measurements of the relevant moral and affective processes as they emerge in response to a range of different moral violations.

Because causality implies temporal precedence, we reasoned that if attended emotions cause moral evaluations, then participants ought to experience (and be able to report) certain emotions (such as feeling angry or upset) before being able to judge the moral status of that behavior. In contrast, if moral judgments guide affect or emotions based on perceived norm violations, causal and mental information, and so on, then felt emotions should follow moral judgments.

Experimental Paradigm

To examine questions of causal primary, we conducted a series of reaction time experiments to test the relative speed of moral, non-moral, and affective reactions to value-laden events. We relied on a variant of the *simultaneous inference paradigm* (*SIP*, Smith & Miller, 1983; Malle & Holbrook, 2012) to measure the speed at which people make different judgments in response to short descriptions of moral transgressions. In the SIP, participants learn to associate a question with a short cue (or, hereafter, "probe") which is then used to elicit responses in a speeded-judgment task. These probes minimize the latency between the presented question and the participant's comprehension, as well as differences in the length and complexity of full questions.

Prior research using the SIP trained participants on dichotomous Yes-No judgments (e.g. "Did the behavior reveal a certain goal the actor has?"), which required modification for two reasons: First, many moral and affective reactions are graded: stabbing someone is worse than keying their car, which is worse than stealing their pencil. A simple Yes-No judgment does not indicate that someone is sensitive to these differences. Second, and relatedly, a prediction of *em-in* models is that the extremity of the affective reaction predicts the perceived severity of the transgression (e.g. Miller et al, 2012), which makes the best test of these models one in which both moral and affective judgments require reporting this more specific, nuanced information. To do this, we presented each probe along with a 7-point rating scale from which participants selected their response as quickly as they could.

We also varied the type of moral scenario participants would react to. Different kinds of events reliably lead to different moral and affective reactions (e.g. people blame transgressors more for intentional harms relative to unintentional ones), and these different outputs are thought to reflect different underlying cognitive processes (Cushman, 2013; Malle et al, 2014). Furthermore, variation in encountered behavior better reflects the experience of encountering random morally relevant behaviors in the real world and generates variation that requires participants' attention. To this end, studies 1-2 mixed intentional and unintentional violations, while studies 3 and 4 mixed intentional, unintentional, and non-agent caused events.

Across four experiments, we measured reaction times for four response types: (1) non-moral judgments (e.g "Intentional?"), (2) moral evaluations of the event (e.g., "Bad?" or "Good?"), (3) moral judgments of the person (e.g., "Blame?"), and (4) reports of one's own affective state (e.g., "Angry?"). As argued above, the *em-in* models predict that people's responses to the affective probes should be faster than the responses to the moral probes, whereas the *em-out* models predict the opposite. For the purpose of the current report, we will focus on these *a priori* contrasts of response times, setting aside the speed of other probes and the specific ratings people provided.

Study 1

Methods

Participants. 241 people (130 self-reported as female, mean age = 35) recruited from Amazon's Mechanical Turk (AMT) participated in this experiment.

Stimuli. We constructed 24 short descriptions of an agent causing harm either intentionally or unintentionally (e.g. "*When she walked by a homeless man asking for money, Lisa spit on the ground in front of him"*). Intentionality was verified through pretesting (mean intentionality ratings for intentional and unintentional descriptions were 8.17 and 2.35 respectively on a 1-9 scale). The 12 intentional and unintentional sentences were matched on length (15.3 and 15.8 words for intentional and unintentional conditions, respectively) and varied in moral severity (valence ratings -1.1 to -4.0, M = -2.45, for intentional transgressions, and -0.65 to -3.88, M = -2.00 for unintentional transgressions on a -4 to +4 scale).

Judgments. Our non-moral, social judgment probed intentionality (*Cue:* INTENTIONAL? *Full:* Was the main character's behavior INTENTIONAL?) on a [1] definitely not intentional to [7] definitely intentional scale. Our event-directed moral evaluation probed "badness" (*Cue:* BAD? *Full:* How BAD was the thing that happened? *Scale:* [1] not at all bad – [7] the most bad possible), while our person-directed moral judgment probed judgments of blameworthiness (How much BLAME does the main character deserve? *Scale*: [1] no blame at all – [7] the most blame possible). Finally, to assess participants' affective states, we used a general feeling probe (*Cue:* FEEL? *Full:* How much did the story make you FEEL something? Scale: [1] no feeling at all – [7] the most feeling possible).

Design. The study crossed two within-subject factors: behavior type (intentional vs unintentional) and judgment type (INTENTIONAL, BAD, BLAME, FEEL). The 24 experimental stimuli and four judgments types were distributed over participants such that they were probed for each judgment type 6 times, (three for intentional behaviors, three for unintentional behaviors). We used a Latin-square design to pair each of the four judgments with each of the 24 stories across four lists. The order of stimuli and probes was randomized for each participant within each list.

Procedure. The entire experiment was conducted through the participant's web browser. At the beginning of the experiment, participants received instructions, including a description of the cues and their associated meanings, as well as the fact that they would be doing a speeded-judgment task and so would have limited time to read and respond to the vignettes. They then completed a training session in which they were taught the single-word cues for the associated judgments (e.g. "BAD?" for "How bad was the thing that happened?").

During the experiment, for each trial, a one-sentence description of a transgression was displayed in the center of

the screen. It remained for 4.5s and was replaced by the probe and a seven-point rating scale (participants did not know which probe would be displayed for any trial). Participants were instructed to place their fingers on the number row of the keyboard and press the corresponding number to respond. Once they indicated their response, the cue and scale disappeared and the next trial automatically started. At the end of the experiment, participants filled out a brief demographics questionnaire indicating their gender, age, and language background.

Results

We removed all trials with reaction times greater than 10 seconds (0.7% data loss). Otherwise, no other trials or participants were removed from data analysis.

Following previous studies (Malle & Holbrook, 2012), we conducted simple effects tests comparing RTs for the affect probe with RTs for other judgment types separately within intentional and within unintentional behaviors. Our primary question was whether reaction times to the affective judgment probe were slower than to other judgment probes. We used linear mixed-effect models (LMEM) to regress RTs on judgment type, which was dummy coded with affective judgment (here, FEEL) as the baseline. Finally, due to the within-subject design, each model included random intercepts and slopes for each participant as well as a random intercept for each scenario.

When judging intentional behaviors, reaction times for FEEL judgments (M = 2563, SD = 1246) were significantly slower compared to INTENTIONAL (M = 2113, SD = 985, b = -452.52, SE = 62.9, t = -7.19, p < 0.001), BAD (M = 2332, SD = 1204, b = -233.04, SE = 57.49, t = -4.054, p < 0.001), and BLAME (M = 2355, SD = 1211, b = -214.44, SE = 55.964, t = -3.832, p < 0.001). However, when judging unintentional transgressions, we observed no significant difference between FEEL (M = 2640, SD = 1267) and INTENTIONAL (M = 2540, SD = 1240, b = -98.71, SE = 59.936, t = -1.647, p = 0.1), BAD (M = 2545, SD = 1244, b = -96.17, SE = 59.09, t = -1.63, p = 0.104), or BLAME (M = 2647, SD = 1339, b = 13.88, SE = 60.93, t = 0.23, p = 0.82).

Discussion

Study 1 provides preliminary support against *em-in* models in favor of *em-out* models. Reaction times for the emotion probe FEEL were slower than those for judgments of INTENTIONALITY, BAD, and BLAME, at least when considering intentional norm violations. There were no comparable RT differences between the affect probe and the remaining probes in response to unintentional violations, perhaps because relevant moral rules are more difficult to access, harm more difficult to calculate, or responsibility more complicated to assess (e.g. Malle, et al., 2014).

One possible reason for the slow unfolding of affective reactions is that participants found it difficult to respond to a vague probe such as "feel". To address this possibility, we conducted another experiment using a more concrete easily identifiable and morally relevant emotion probe: *angry*.

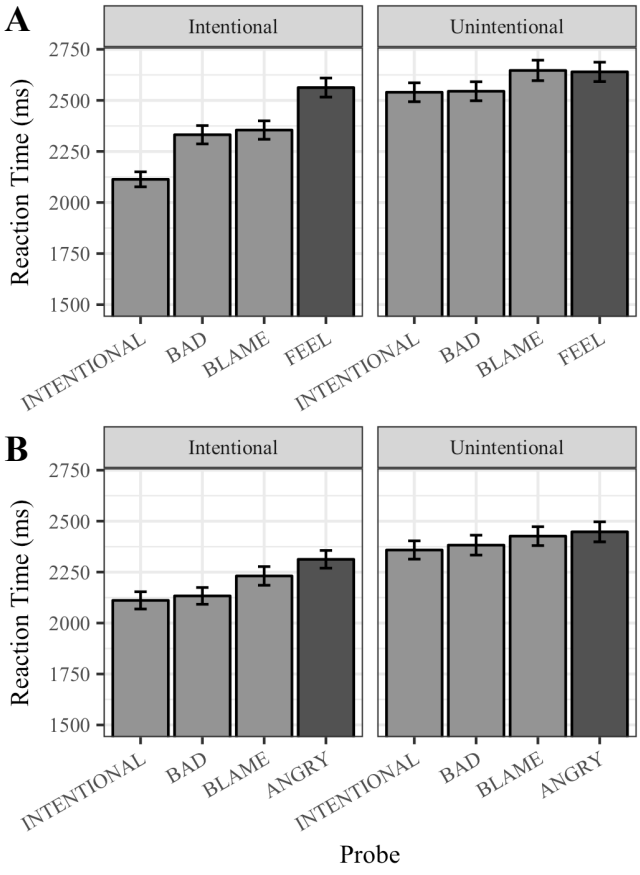

Figure 1: Reaction time (and standard error) for participants' responses to probes in study 1 (A) and 2 (B)

Study 2

Methods

Participants, Materials, & Procedure. 237 people (134 self-reported as female, mean age = 35) recruited from Amazon's Mechanical Turk (AMT) participated in this experiment. Study 2 was identical to Study 1 except that the affective judgment probe assessed anger (*Cue*: ANGRY? *Full:* How ANGRY are you at the main character? *Scale*: [1] not at all angry to [7] the most angry possible).

Results

Reaction times from Study 2 were analyzed identically to Study 1. Before conducting analyses, we removed all trials with RTs greater than 10s (2.8% data loss). No other data were removed.

Replicating Study 1, we found that, in the intentional condition, INTENTIONAL (M = 2111, SD = 1119) and BAD (M = 2133, SD = 1091) judgments were both significantly faster than ANGER judgments (M = 2313, SD = 1150; INTENTIONAL: b = -203.90, SE = 52.32, t = -3.90, p < 0.001; BAD: -183.18, SE = 52.90, t = -3.46, p = 0.001). We did not observe significant differences between ANGER (M = 2448, SD = 1300) and other judgments for unintentional violations

(INTENTIONAL: M = 2358, SD = 1189, -88.43, SE = 56.65, t = -1.561, p = 0.119; BAD: M = 2382, SD = 1291, b = -61.22, SE = 59.91, t = -1.02, p = 0.307). Finally, we detected no significant difference between ANGER and BLAME for either intentional (M = 2231, SD = 1214, b = -80.41, SE = 52.94, t = -1.52, p = 0.129) or unintentional (M = 2426, SD = 1229, b = -18.10, SE = 55.28, t = -0.33, p = 0.743) behaviors (see Figure 1).

Discussion

Study 2 largely replicated the findings from Study 1. However, both studies are limited in several respects. First, the behaviors were always negative, and the agent in every description always a causer of harm. It is possible under these conditions that moral norms become more salient and accessible or that expectations of high causal agency sped up moral evaluations. Additionally, anger is typically directed at persons, and a different emotion term may be more appropriate for affective reactions to events. Finally, the studies did not limit participants' response time. While most responses occurred within several seconds, it is nevertheless possible that judgments are consciously accessible before then without an incentive to reveal these judgments as soon as they are accessible. The next two studies were designed to address these shortcomings.

Studies 3 and 4

Methods

Participants. 111 people (58 self-reported as female, mean age = 37.4) in Study 3 and 193 people (90 self-reported as female, mean age = 33.4) in Study 4, recruited from AMT participated in this experiment.

Stimuli. We constructed 28 single-sentence descriptions, 14 featuring a good event and 14 featuring a bad event. For each valence condition, we constructed four stimuli that had an agent with no causal role (*Non-causal*) in the good or bad event, four in which an agent unintentionally did a good or bad thing (*Unintentional*), and a final six in which an agent intentionally did something good or bad (*Intentional*). Pretesting ensured that these items were matched on intentionality and length across valence, and that all agency conditions were comparable (see Table 1).

Judgments. We modified the full questions associated with each cue to accommodate the greater variety of stimulus events. Similar to Studies 1-2, we included measures of (1) intentionality (*Cue*: INTENTIONAL? *Full:* "Was it intentional (what the character did)?" *Scale*: [1] definitely not intentional - [7] definitely intentional), and (2) the badness of the event (*Cue*: BAD? *Full:* "How bad was it (what happened)?" *Scale*: [1] not at all bad - [7] extremely bad). We also included a measure of (3) the goodness of the event in order to accommodate the positive valence items (*Cue*: GOOD? *Full:* "How good was it (what happened)?" *Scale*: [1] not good at all - [7] extremely good).

Table 1: Pretest ratings for stimuli in Studies 3 & 4

Behavior Types		**Pretest Values**		
Valence	*Agency*	*Intentionality*	*Valence*	*Words*
Negative	Intentional	8.41	-2.58	16.83
	Unintentional	2.40	-2.57	16.00
	Non-causal		-2.70	13.75
Positive	Intentional	8.22	2.57	15.17
	Unintentional	2.03	2.64	18.75
	Non-causal		2.79	12.50

Studies 3 and 4 were identical except for the affective judgment probe. Study 3 measured anger (*Cue*: ANGRY? *Full:* "How angry were you (about what happened)?" *Scale*: [1] not at all angry – [7] extremely angry), whereas Study 4 measured "upset" (*Cue*: UPSET? *Full*: "How upset were you (about what happened)?" *Scale*: [1] not upset at all – [7] extremely angry).

Design. Studies 3 and 4 crossed two within-subject factors: event type (non-causal, intentional, and unintentional) and judgment type (INTENTIONAL, BAD, GOOD, and UPSET or ANGRY). The 28 experimental stimuli and four judgment types were distributed over participants in the following pattern: across the 28 items, participants responded to eight INTENTIONAL probes, four for intentional behavior conditions, two for unintentional behavior conditions, and two in the non-caused behavior conditions. This distribution meant that roughly half the probes would results in low intentionality ratings and the other half would result in high intentionality ratings. Affect probes were also distributed this way: four for intentional behavior stimuli, two for unintentional behavior stimuli, and two for non-caused. These probes were evenly divided between valence conditions. Finally, participants saw six BAD probes and six GOOD probes, which were matched to valence. That is, participants made BAD judgments only following negatively-valenced stimuli and GOOD judgments only following positively-valenced stimuli. The twelve moral evaluation probes (6 BAD and 6 GOOD) were evenly divided between event type conditions.

Probes were distributed across four stimulus lists according to a Latin-square design. At the beginning of the experiment, participants were randomly assigned to one of the four lists and, during the experiment, the order of the stimulus sentences and probes was randomized.

Procedure. The training and overall experiment procedures were the same as in Studies 1 and 2, with one exceptions: For each trial, the judgment screen containing the cue (e.g. "BAD?") and the rating scale disappeared after five seconds after being displayed. If no response was offered before then, no response was recorded for that trial. Participants were informed of the time restriction in the instructions. Prior to the experiment, participants conducted five practice trials to get accustomed to the procedure.

Results

We removed all trials in which the participant did not provide an answer within the time constraint (1.6% data loss in Study 3, 3% data loss in Study 4). No other data were removed. Similar to Studies 1 and 2, we conducted separate analyses on the *Intentional*, *Unintentional*, and *Non-Causal* behaviors, using the same mixed-effect regression models and adding the Valence (*positive* vs. *negative*) term, which predicted changes in the dummy variable (affect) as a function of positive or negative behaviors. Across all models in both studies, valence was not significant and did not improve model fit, and so was removed as a predictor.

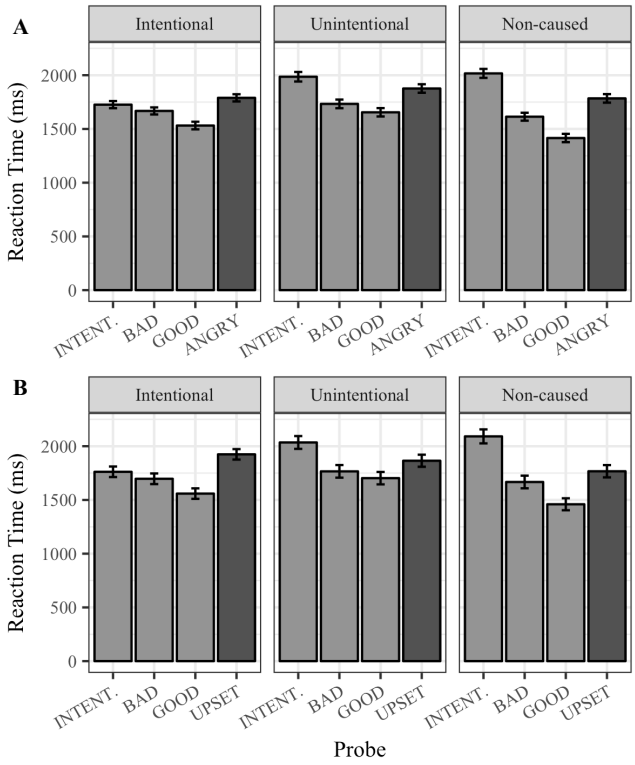

Figure 2: Reaction time (and standard error) for participants' responses to probes in study 3 (A) and 4 (B)

Intentional Behaviors. In Study 3, moral evaluations were faster than ANGER (M = 1746, SD = 739; BAD: M = 1601, SD = 647, b = -178.46, SE = 37.35, t = -4.78, p < 0.001; GOOD: M = 1438, SD = 645 , b = -294.32, SE = 38.32, t = -7.68, p < 0.001), while INTENTIONAL ratings were not (M = 1687, SD = 739, b = -57.65, SE = 36.01, t = -1.60, p = 0.109). Similarly, in Study 4, INTENTIONAL (M = 1656, SD = 675), BAD (M = 1560, SD = 607), and GOOD (M = 1456, SD = 642) judgments were significantly faster than UPSET (M = 1822, SD = 682; INTENTIONAL: b = -163.51, SE = 42.7, t = -3.83, p < 0.001; BAD: b = -273.68, SE = 48.50, t = -5.64, p < 0.001; GOOD: b = -366.28, SE = 48.19, t = -7.60, p < 0.001).

Unintentional Behaviors. In Study 3, moral evaluations were significantly faster than the ANGER ratings (M = 1837, SD = **714;** BAD: M = 1682, SD = 694**,** b = -174.2, SE = 43.07, t = -4.05, p < 0.001; GOOD: M = 1601, SD = 647**,** b = -193.3, SE = 43.86, t = -4.41, p < 0.001), while INTENTIONAL was slower (M = 1932, SD = 775**,** b = 115.11, SE = 41.88, t = 2.75, p = 0.006). In Study 4, UPSET (M = 1760, SD = 617) was significantly slower than GOOD (M = 1603, SD = 624**,** b = -157.62, SE = 53.89, t = -2.93, p = 0.003), and significantly faster than INTENTIONAL (M = 1972, SD = 769**,** b = 212.92, SE = 53.78, t = 3.96, p < 0.001), but not reliably different from BAD (M = 1673, SD = 686, b = -82.75, SE = 53.86, t = -1.54, p = 0.124).

Non-Caused Behaviors. Moral evaluations of uncaused good and bad events were significantly faster than the ANGER judgments (M = 1728 , SD = 653; BAD: M = 1561, SD = 611, b = -163.31, SE = 38.24, t = -4.27, p < 0.001; GOOD: M = 1335, SD = 541, b = -390.61, SE = 37.89, t = -10.31, p < 0.001), while INTENTIONAL ratings were slower (M = 1969, SD = 749, b = 241.79, SE = 39.55, t = 6.11, p < 0.001). In Study 4, UPSET (M = 1659, SD = 616) was significantly slower than GOOD (M = 1340, SD = 505, b = -309.67, SE = 50.41, t = -6.14, p < 0.001) and BAD (M = 1556, SD = 639, b = -103.24, SE = 50.39, t = -2.05, p = 0.04), but significantly faster than INTENTIONAL (M = 1947 , SD = 723, b = 292.20, SE = 50.56, t = 5.78, p < 0.001).

Discussion

Results from Studies 3 and 4 replicated our previous findings (see Figure 2): Moral evaluations of intentional violations were reliably faster than reports of felt anger and upsetness. For unintentional violations, reporting feeling upset was not significantly slower than negative moral evaluations, but anger was. Perhaps upset feelings are more globally sensitive to any unfortunate outcome and therefore converge with (but on average do not precede) badness judgments. Lastly, intentionality judgments were slowed in response to unintentional and uncaused events—which is not entirely surprising given that those events are clearly *not* intentional; the detection of negation may take time.

General Discussion

Across four experiments, participants were reliably slower at reporting their emotional states in response to norm violations compared to reporting their moral judgments. More specifically, the results from these studies showed a clear speed advantage for *event-directed* judgments of badness and, often, intentionality judgments. These results fit both with theoretical models of moral judgment arguing that moral appraisals precede emotion, as well as prior work showing that intentionality and norm violation detection can occur extremely quickly (Malle & Holbrook, 2012; Van Berkum et al, 2009). These findings did not extend to person-directed moral judgments (blame), consistent with theories that blame is more complex than event-directed evaluations (e.g. Cushman, 2013; Malle et al, 2014).

One important limitation of these studies comes from the observation that, while we are interested in characterizing

the cognitive processes underlying moral judgment when people are exposed to a morally relevant stimulus, we measured people's reaction times to probes that were displayed *after* the stimulus had been shown. It is possible that participants attended to their affective reactions when they were first exposed to the stimulus, which resulted in a moral judgment, which they later more quickly retrieved during the post-stimulus probe. However, while we cannot rule this out, it is not clear why affect would be more difficult to retrieve post-stimulus as opposed to during-stimulus (when one's attention is presumably directed outward toward reading the stimulus). Additionally, this account does not explain why the speed of retrieving affective information would change as a function of the behavior (Study 4). Finally, prior work using a simultaneous inference paradigm found that post-stimulus reaction times directly recapitulated online measures (Malle & Holbrook, 2012).

Second, affect may have been slower relative to moral and social judgments because of an attention switching cost: the non-affect judgments targeted the stimulus while the affect judgment targeted oneself. Because *em-in* models explicitly predict a shift in attention from the behavior to one's affective state, a delay in reporting due to switching attention is not, in principle, a confound for our test. That said, this cost may have been exacerbated by the relative balance of event-directed (75%) versus self-directed (25%) probes, and future studies should use an even balance of affect and moral judgments.

Finally, even if we accept that felt emotions do occur after explicit moral judgment, our data do not rule out the possibility that *pre-conscious* affective processes play a role in moral judgment formation (say, by interfering with cognitive processes, Alicke, 2000). It is also possible that conscious affect may play a causal role when judging more ambiguous situations, in which relevant harm or rule information is difficult to access. Consistent with this, badness and blame judgments were not reliably faster than emotion reports when judging accidental bad behavior. Thus, our results may only hold for relatively common or extreme, but not unusual or novel situations.

In summary, we found that people could report moral evaluations of norm-violating events more quickly than their emotional reactions to these events. These results pose a challenge to models claiming that felt affect plays a necessary role in forming moral judgment.

Acknowledgments

The authors would like to thank Fiery Cushman, John Voiklis, members of the SCSRL and anonymous reviewers for valuable feedback on this project. This research was funded in part by a grant from the Office of Naval Research (ONR), N00014-13-1-0269.

References

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

Cheng, J. S., Ottati, V. C., & Price, E. D. (2013). The arousal model of moral condemnation. *Journal of Experimental Social Psychology*, *49*, 1012-1018.

Cushman, F. (2013). Action, outcome, and value a dual-system framework for morality. *Personality and Social Psychology Review*, *17*, 273-292.

Haidt, J. (2001). The emotional dog and its rational tail: a social intuitionist approach to moral judgment. *Psychological Review*, *108*, 814.

Huebner, B., Dwyer, S., & Hauser, M. (2009). The role of emotion in moral psychology. *Trends in Cognitive Sciences, 13*, 1–6.

Landy, J. F., & Goodwin, G. P. (2015). Does incidental disgust amplify moral judgment? A meta-analytic review of experimental evidence. *Perspectives on Psychological Science*, *10*, 518-536.

Malle, B. F., Guglielmo, S., & Monroe, A. E. (2014). A theory of blame. *Psychological Inquiry*, *25*, 147-186.

Malle, B. F., & Holbrook, J. (2012). Is there a hierarchy of social inferences? The likelihood and speed of inferring intentionality, mind, and personality. *Journal of Personality and Social Psychology*, *102*, 661-684.

Mikhail, J. M. (2011). *Elements of moral cognition: Rawls' linguistic analogy and the cognitive science of moral and legal judgment*. New York, NY: Cambridge Uni. Press.

Miller, R. M., Hannikainen, I. A., & Cushman, F. A. (2014). Bad actions or bad outcomes? Differentiating affective contributions to the moral condemnation of harm. *Emotion*, *14*, 573-587.

Nichols, S. (2002). Norms with feeling: Towards a psychological account of moral judgment. *Cognition*, *84*, 221-236.

Niedenthal, P. M., Rohmann, A., & Dalle, N. (2003). What is primed by emotion concepts and emotion words. *The psychology of evaluation: Affective processes in cognition and emotion*, 307-333.

Paulhus, D. L., & Lim, D. T. (1994). Arousal and evaluative extremity in social judgments: A dynamic complexity model. *European J. of Social Psychology*, *24*, 89-99.

Royzman, E. B., Goodwin, G. P., & Leeman, R. F. (2011). When sentimental rules collide: "Norms with feelings" in the dilemmatic context. *Cognition*, *121*, 101-114.

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

Smith, E. R., & Miller, F. D. (1983). Mediation among attributional inferences and comprehension processes: Initial findings and a general method. *Journal of Personality and Social Psychology*, *44*, 492.

Van Berkum, J. J., Holleman, B., Nieuwland, M., Otten, M., & Murre, J. (2009). Right or wrong? The brain's fast response to morally objectionable statements. *Psychological Science*, *20*, 1092-1099.

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