Excuse Validation: A Study in Rule-Breaking

John Turri and Peter Blouw

> We do not call anything wrong unless we mean to imply that a person ought to be punished in some way or other for doing it.
>
> – *John Stuart Mill*

Abstract: Can judging that an agent blamelessly broke a rule lead us to claim, paradoxically, that no rule was broken at all? Surprisingly, it can. Across seven experiments, we document and explain the phenomenon of excuse validation. We found when an agent blamelessly breaks a rule, it significantly distorts people's description of the agent's conduct. Roughly half of people deny that a rule was broken. The results suggest that people engage in excuse validation in order to avoid indirectly blaming others for blameless transgressions. Excuse validation has implications for recent debates in normative ethics, epistemology and the philosophy of language. These debates have featured thought experiments perfectly designed to trigger excuse validation, inhibiting progress in these areas.

Keywords: blame; excuses; rules; speech acts; moral psychology

Introduction

Blame is a familiar fact of social life and philosophers and psychologists have studied blame extensively (e.g. Aristotle 350 BCE; Mill 1861; Strawson 1962; Feinberg 1970; Fischer & Ravizza 1998; Knobe 2010; Hindriks 2008; Guglielmo & Malle 2010, Nadelhoffer 2006; Cushman 2008; Gray & Wegner 2010; Fast & Tiedens 2009; Pizarro et al. 2003; Inbar et al.

2012; Edmondson 2004). One intriguing line of research by Mark Alicke and collaborators has shown that when we view someone negatively and want to blame them, we interpret facts in a way that justifies our negative reaction, a phenomenon known as *blame validation* (Alicke 1992; Alicke 2000; Alicke, Buckingham, Zell & Davis 2008; see also Alicke & Rose 2010). The motivation to blame can distort people's judgments about causal processes. For example, we are more likely to identify someone as the primary cause of an accident if he was on his way to hide a vial of cocaine from his parents than if we was on his way to hide their anniversary present from them (Alicke 1992, Study 1). This discovery was an important contribution to our understanding of the way we ordinarily evaluate agents and their conduct.

Common as blame is, we don't always blame people for bad outcomes and transgressions. Often times we excuse them (Austin 1956; Brandt 1969; Mehlman & Snyder, 1985; Snyder & Higgins, 1988; Tyler & Feldman, 2007; Botterell 2009; Tetlock, Self & Singh 2010). If people interpret facts so that it validates their desire to blame, then perhaps they also interpret facts so that it validates their desire to excuse. Is blame validation only half the story? Might people also engage in *excuse validation* too? That is, might people also interpret the facts so that it validates their desire to excuse?

If the answer is "yes," then it complicates standard methodology in normative theorizing. Normative theorizing in philosophy relies heavily on considered judgments about cases (Goodman 1955; Rawls 1971; Daniels 1979). And there is a general preference for conservatism when choosing among theories (Harman 1986). Other things being equal, we should prefer the theory that coheres better with our pretheoretical beliefs or with what we would ordinarily say; and we should prefer the theory that better "saves the appearances." But if our judgments about cases, our pretheoretical beliefs, what we would ordinarily say, or the way things appear is

infected by excuse validation, then this pollutes the well of evidence that we rely on when evaluating normative theories. In particular, it makes it harder to be confident that certain purported counterexamples to philosophical theories really are counterexamples.

For example, consider a consequentialist theory of moral rightness which says that an agent's action is morally wrong if he could easily have done something that promoted a significantly better outcome. A critic then produces a purported counterexample with this structure: an agent performs an action and has very good evidence that he could not easily have done something significantly better; but it turns out that, despite the evidence, such an option was available to him; yet no one wants to say that the agent's action is immoral (compare Kagan 1998: 64ff). So, the critic concludes, the consequentialist theory is probably wrong, due to this counterintuitive implication. But if we are prone to excuse validation when considering cases with this structure, then the critic incurs an additional burden: she faces the very real possibility that her intuition is actually a misleading response to the fact that the agent has engaged in *blameless* wrongdoing, rather than, contra consequentialist theory, no wrongdoing at all. In short, she should explain why the supposed counterintuitiveness isn't merely excuse validation in action.

Similarly, consider the recent debate in epistemology and the philosophy of language over the norm of assertion. Theorists have vigorously debated whether assertion is governed by a factive norm or rule. A factive norm of assertion says that you should assert a proposition only if it is true. According to factive accounts, a false assertion breaks the rule and so should not be made (Unger 1975; Williamson 2000: ch. 11; Weiner 2005; Schaffer 2008; Turri 2010, 2011, 2013; Benton 2012; Buckwalter & Turri, 2013). Although there is considerable evidence in favor of a factive account, critics have strenuously objected to it (Douven 2006; Lackey 2007; Hill and

Schechter 2007; Kvanvig 2009). Critics advance the following sort of argument: a speaker who makes a false but well justified assertion is not thereby blameworthy; so the speaker has probably not broken a rule of assertion; so assertion probably doesn't have a factive norm. The argument assumes that blamelessness is a defeasibly good indication that no rule has been broken. More generally, critics of factive accounts claim that it is just counterintuitive to say that a false but well justified assertion *breaks the rule* of assertion, or is incorrect *qua* assertion. But, again, if we are prone to excuse validation when considering cases of this sort, then there is a ready diagnosis of the critic's dissatisfaction: excuse validation, wherein blameless transgressions of the norm of assertion are mistaken for instances in which no transgression has occurred.

This suggests a recipe for how excuse validation is relevant to normative theory in general. Whenever a theory proposes that a certain form of activity is governed by a standard or rule R, a certain sort of case will seem tempting as a counterexample. The tempting case will have three features:

- An agent reasonably believes that he is following R; but
- he isn't actually following R, due to factors outside of his awareness or control; and
- many people find it counterintuitive to say that the agent is breaking the rule or acting incorrectly according to the rule.

Of course, some such cases might be genuine counterexamples. The present point is not that a tendency toward excuse validation automatically disqualifies such counterexamples. (As they say, just because you're paranoid doesn't mean they aren't out to get you.) Rather, the point is that excuse validation complicates the situation. Critics incur a burden to explain why the supposed counterintuitiveness isn't just excuse validation in action. Conversely, proponents of

the theory inherit a tool that can be wielded to explain away the supposedly counterintuitive verdict.

It is thus a question of first importance to normative philosophical theorizing whether excuse validation is a robust human tendency. In light of recent developments in social psychology, in particular Alicke's work on blame validation noted above, we set out to answer this question. To test whether people engage in excuse validation, we conducted seven experiments that asked participants to evaluate cases involving transgressions. In Experiments 1 and 2, participants read a story about a blameless transgressor and were then asked whether the protagonist had acted incorrectly or broken the rules. In Experiment 3, participants read a similar story that lacked clear exculpatory information. In Experiments 4 and 5, participants again read a story about a blameless transgressor, but this time they were provided with more discriminating options for characterizing the transgression. In Experiment 6, participants evaluated transgressions that caused dramatically different amounts of societal harm. In Experiment 7, participants were asked to evaluate a third-party's assertion about a blameless transgression. We found that participant judgment about rule-breaking is distorted when exculpatory information is present but not when it is absent. We also found that participants improve at identifying rule-breaking when they are allowed to identify it as "unintentional." Increasing a transgression's harmfulness did not affect the rate at which excuse validation occurred. Moreover, excuse validation occurred at similar rates whether participants were asked to directly judge the transgression or to evaluate a third-party's assertion about it. We propose an explanation of excuse validation in terms of the pragmatics of indirect speech.

Experiment 1

Method

Participants (N = 153, 70 female, aged 18–66 years, mean age = 30 years) were recruited and tested using Amazon Mechanical Turk + Qualtrics and compensated $0.30 for approximately 2 minutes of their time. Participation was restricted to U.S residents and 97% listed English as their native language. We excluded data from 44 participants who failed comprehension questions designed to ensure that they were reading the scenario carefully enough. But including these participants yielded the same pattern of results described below. We used Amazon Mechanical Turk to recruit and compensate participants in the same way for all subsequent experiments. Repeat participant was not allowed and participants who had taken previous surveys were excluded by their AMT Worker ID. As an added precaution, we manually screened for repeat IP addresses.

Participants were randomly assigned to read one of five stories: Baking, Farming, Asserting, Driving, and Chess. We used multiple cover stories about a wide range of activities to rule out explanations of the results in terms of superficial features of any one story or common presuppositions about any one activity. The stories share a common structure: the protagonist has good evidence to think that they are following the relevant rule and correctly engaging in the activity, but something outside of their awareness foils them and they end up breaking the rule anyway. For example, in the Driving story, Doreen just had her car serviced and is driving home from the mechanic's shop. She wants to get home without unnecessary delay, but she does not want to break any traffic laws. The traffic law sets the speed limit at 55 miles per hour, so she looks down to see how fast she is going. The speedometer says that she is going 55 miles per hour. But the speedometer malfunctions and underestimates the car's speed. Doreen is not aware of this and is actually driving sixty miles per hour.

Following the story, participants answered four comprehension questions, then a question about whether the protagonist should be criticized, then the key test question: did the protagonist act incorrectly? For example, here are the questions for the Driving story:

1. How fast does Doreen think she is driving? [55/60]
2. Is it reasonable for Doreen to think that she is driving that fast? [Yes/No]
3. What is the speed limit? [55/60]
4. How fast is Doreen driving? [60/55]
5. Should Doreen be criticized for driving that speed? [Yes/No]
6. Is there a sense in which it is incorrect for Doreen to drive that speed? [Yes/No]

The questions were always asked in the same order and presented on a single screen with the story on top. Response options were presented in random order. After testing, participants filled out a brief demographic questionnaire. All the experiments reported in this paper used these same basic procedures. All the stories and questions are included in Appendix A.

Results and discussion

The protagonist in the story breaks a relevant and salient rule. Moreover, participants answered comprehension questions whose correct answers together straightforwardly entail that the protagonist broke the rule. Thus one might reasonably expect, at the very least, a very strong majority of participants to agree that there is a sense in which the protagonist's behaviour is incorrect (i.e. answer "yes" to the test question). For example, in the Driving story Doreen is driving 60 mph in a 55 mph zone. It follows unmistakeably that she is breaking the speeding law. In the context of the story, this establishes a clear and relevant sense in which her driving speed is incorrect.

However, the protagonist is justified in thinking that they're following the rule and they end up breaking the rule in light of misleading evidence. Thus we should expect most participants to deny that the protagonist should be criticized for their behaviour (i.e. answer "no" to the criticism question). For instance, the speedometer on Doreen's well maintained car tells her that she's not breaking the speed limit, and Doreen is motivated to obey the traffic laws. Thus it seems inappropriate to criticize or blame her for breaking the speed limit.

Cases like this provide an excellent test for whether excuse validation actually occurs. If excuse validation does occur, then many participants will deny that the protagonist's behaviour is incorrect. But if excuse validation doesn't occur, then it's unlikely that many people will deny that the protagonist's behaviour is incorrect. Instead, we'd expect almost everyone to agree that there is a sense in which the protagonist's behaviour is incorrect, which is the objectively correct answer.

Two preliminary notes about the analyses that follow: (i) as expected, cover story had no effect on response to the test question (i.e. is the protagonist's conduct incorrect?), 41–56% answering "yes," $\chi^2(df = 4, N = 153) = 2.202$, $p = .699$, n.s., all tests two-tailed unless otherwise noted; (ii) participant gender had no effect on response to the test question (M/F: 54/44% answering "yes"), Fisher's exact test, $p = .257$, n.s., so we collapse across gender. The same is true for all the experiments reported in this paper.

The overwhelming majority (92%) said that the protagonist should not be criticized, binomial test, $p < 0.001$, and cover story didn't affect participant response on this matter, 85-97%, $\chi^2(df = 4, N = 153) = 3.85$, $p = .427$, n.s. By contrast, only half of participants (49.9%) agreed that the protagonist's behaviour was in some sense incorrect. In other words, participants

were no better than chance at correctly answering the test question, binomial test, $p = 1$. (See Table 1.)

This is a very surprising result that provides strong initial support for the hypothesis that people engage in excuse validation. Consider the Chess story. Nicolas is playing in a chess tournament where "the rook rule" is in effect: players are not allowed to take more than 30 seconds to move the rook. But Nicolas takes 45 seconds to move his rook. Participants were eliminated from the analysis if they didn't acknowledge both of those facts, which together trivially entail that Nicolas's move is incorrect in at least this sense: he is breaking the tournament rules. But only 41% of participants agree that there is a sense in which Nicolas's move is incorrect. Participants who read the other stories agreed at similar rates (Table 1).

Table 1: Experiment 1: Percentage of participants agreeing that the protagonist's action is incorrect.

Baking	**Farming**	**Asserting**	**Driving**	**Chess**
56%	55%	53%	45%	41%

One deflationary explanation for this result is that participants interpreted the questions from the protagonist's perspective: Doreen isn't aware that she is breaking the speed limit, so from her perspective there is nothing wrong with her speed.[1] But this is unlikely because if participants were interpreting the questions from the protagonist's perspective, then they would have failed the comprehension questions: Doreen is equally unaware of the fact that she is driving 60 mph, but all participants correctly answered that she is driving 60 mph. Moreover, as

[1] Thanks to Laurie Rudman for proposing the explanation.

observed in Experiments 4-7 below, almost all participants are happy to answer that the protagonist "unintentionally" broke the rules, even though, from the protagonist's perspective, they are no more unintentionally breaking the rules than they are breaking them.

Two other concerns about the results naturally suggest themselves. First, perhaps asking whether the protagonist "should be criticized" encouraged many participants to interpret "incorrect" in the test question to mean "incorrect in a way that merits criticism." This could be amplified by the fact that asking "is there a sense in which it is incorrect …?" is somewhat unusual, which could encourage participants to search for contextual cues about how to interpret it. Second, perhaps something in the comprehension questions drove the results. For instance, perhaps a series of questions about a blameless transgression led some participants to react protectively when answering the test question. Experiment 2 addresses both of these concerns.

Experiment 2

This experiment addresses the two worries raised about Experiment 1. In this experiment, we ask participants no comprehension questions and change the wording of the test question: we simply asked them whether the protagonist broke the rules.

Method

Participants (N = 150, 75 female, aged 18-65, mean age = 31 years; 96% reporting English as a native language) were randomly assigned to read one of the same five stories used in Experiment 1. Following the story, participants were asked a single question: did the protagonist break the rules? Participants were then asked to rate how confident they were in their answer, on a scale of

1-10 (low to high). We added this question to determine whether people who answer "no" were less confident in their answer.

Results and discussion

If the concerns raised about Experiment 1 are correct, then switching to a single, straightforward question about "breaking the rules" should result in a very different outcome. More specifically, participants should perform much better than chance when answering the test question. By contrast, if excuse validation is a robust phenomenon, then the results will probably be similar to Experiment 1.

The results were very similar to Experiment 1. Only 56% of participants agreed that the protagonist broke the rules, which doesn't differ from what could be expected by chance, binomial test, p = .165, or from response to the test question in Experiment 1, binomial test, p = .159, test proportion = .499. Moreover, participants who answered "no" to the test question were no less confident than participants who answered "yes," M = 8.27/8.60, SD = 2.10/1.77, independent-samples t-test, $t(148) = -1.02$, p = .310, n.s. Thus the results replicate the main finding from Experiment 1 while also addressing the two concerns raised about it.

Considering Experiment 2 on its own, it's natural to wonder whether the effect is being driven by carelessness. After all, there are no comprehension questions. But this concern is met by considering the results in light of Experiment 1. Experiment 1 included four comprehension questions, but the observed results were substantially similar. Moreover, Experiments 6 and 7 below include comprehension questions, and the observed results are again very similar.

Experiment 3

Although we've seen that people are surprisingly unwilling to classify blameless rule-breaking as rule-breaking, the results thus far don't establish that this has anything special to do with *blameless* rule-breaking. Consistent with the findings thus far, people might display a *general* reluctance to describe others as breaking rules or acting incorrectly, whether blameless or not. Or perhaps all that matters is that the protagonist *believes* or *sincerely believes* that she or he is following the rules, or that she or he has been *tricked* into breaking the rules. This experiment tests these possibilities.

Method

Participants (N = 95, 25 female, aged 18-60 years, mean age = 27 years; 95% reporting English as a native language; five participants excluded who had taken a prior survey) were randomly assigned to one of four conditions: Minimal, Believes, Sincerely Believes, and Tricked. Participants in each condition read a variation of the Chess story. Participants in the Minimal condition read this story:

> Nicolas is participating in a chess tournament. Tournament rules say that players are not allowed to take more than 30 seconds to make a move. Nicolas just took 45 seconds to make a move.

Participants in the Believes, Sincerely Believes, and Tricked conditions read a very similar story, except that they were told, "Nicolas believes/sincerely believes/was tricked into believing that the rules allow players to take 60 seconds to make a move. As a result, Nicolas just took 45 seconds to make a move." Notice that the story for the Tricked condition is simply an abstract description of the original Chess story, omitting the details of *how* Nicolas is tricked.

After reading the story, all participants were asked a single yes/no question, "Did Nicolas break the rules?"

Results and discussion

One explanation of our earlier results is that many people are simply unwilling in general to describe behaviour as incorrect or as transgressions. Or, at least, people are unwilling to describe behaviour as incorrect or as a transgression when the agent believes, sincerely believes, or was tricked into believing that the behaviour is correct and permissible. Thus it's possible that the observed results, surprising though they may be, aren't due to the perception of *blameless* rule-breaking specifically. If this objection is on the right track, then we should observe similar results in the present experiment. That is, in at least one of the four conditions, participants should engage in excuse validation, with roughly half answering that Nicolas did not break the rules.

There was no effect of condition on participant response, χ^2 (df = 3, N = 95) = 2.694, p = .441, n.s. Participants in all four conditions overwhelmingly agreed that Nicolas broke the rules, at rates far exceeding chance, 95-100%, binomial test, all ps < .001. Indeed, only 2 out of 95 participants answered that Nicolas did not break the rules. This rules out the proposed explanations of the earlier results.

At this point we think it is reasonable to conclude that many people engage in excuse validation. Moreover, although participants engage in excuse validation when they are given specific exculpatory information, participants did not do so when given general information that suggests the protagonist should be excused (e.g. having been "tricked"). This suggests that

excuse validation is not merely a matter of verbal generosity on the part of some participants.[2] The next two studies test one explanation for why excuse validation happens.

Experiment 4

Why do people engage in excuse validation? We find the following hypothesis plausible. Saying that someone "broke a rule" or acted "incorrectly" is often a way of *indirectly* criticizing or blaming them (for the theory of indirect speech acts, see Searle 1979: ch. 2). But participants don't want to (appear to) unfairly criticize the protagonist. So instead of agreeing that the protagonist "broke the rules," which would be true but misleading, many participants disagree. In other words, many participants say that the protagonist did not break the rules because it's the only available option that won't implicate them in unwarranted criticism.

According to this hypothesis, many participants say something literally false in order to communicate something relevant (Grice 1989). They literally answer "no" to the question "did the protagonist break the rules?" but we should understand that "no" to mean, roughly, "No, I'm not going to blame the protagonist."

If this hypothesis is correct, then a slight change to the wording of the test question could dramatically change the results. For instance, suppose we instead ask participants to decide which statement better describes the case: the protagonist "unintentionally broke the rules" or the protagonist "did not break the rules." We often say that an outcome was "unintentional" in order to excuse or at least mitigate blame. We predicted that this small change would lead to a dramatically different pattern of response. This experiment tests this prediction.

---

[2] Thanks to Ori Friedman for helpful discussion on this point.

Method

Participants (N = 58, 24 female, aged 18-53 years, mean age = 26 years; 98% reporting English as a native language) read one of two of the stories used in Experiment 1, Baking and Asserting. After reading the story, participants were asked to pick which of two options best describes the case: *the protagonist unintentionally broke the rules* and *the protagonist did not break the rules*.

Results and discussion

We hypothesized that excuse validation occurs because people don't want to suggest that a blameless transgressor should be unfairly blamed. If our hypothesis is correct, then participants in this experiment will not engage in excuse validation because agreeing that the protagonist unintentionally broke the rules does not similarly suggest blame. That is, this experiment doesn't force participants to choose between answering truthfully and avoiding unfairness. Instead, by agreeing that the protagonist unintentionally broke the rules, participants can correctly identify it as a transgression while indirectly excusing it.

As predicted, participants overwhelmingly answered that the protagonist unintentionally broke the rule, 91%, binomial test, $p < .001$. This provides good initial support for our hypothesis that excuse validation is caused by a preoccupation to avoid unfairly blaming the protagonist. Experiment 5 further tests the hypothesis.

Experiment 5

Method

Participants (N = 58, 33 female, aged 18-69 years, mean age = 30 years; 93% reporting English as a native language) read one of the two stories used in Experiment 4. After reading the story,

participants were asked, "Which of the following statements accurately describes Brenda/Robert? Check all that apply." The options were (1) the protagonist broke the rules, (2) the protagonist unintentionally broke the rules, and (3) the protagonist did not break the rules. The options were rotated randomly. Participants were then taken to a different screen and asked the same test question as in Experiment 2, namely, whether the protagonist broke the rules (Yes/No). Participants could not go back and change their previous answer.

Results and discussion

We expected that when asked to "check all that apply," participants would tend to select that the protagonist unintentionally broke the rules (option 2) and tend to avoid selecting that the protagonist broke the rules (option 1). Provided this expectation was met, we predicted that participants would feel entitled to then agree that the protagonist broke the rules when asked a direct Yes/No question. We predicted this because we expected that giving participants the opportunity to first note that the rule-breaking was unintentional would eliminate or at least significantly reduce any implication that the rule-breaking was blameworthy.

The expectations were met and the prediction was true. Cover story didn't affect the rate at which participants selected any of the three "check all that apply" options: option 1, 17/24%, Fisher's exact test, $p = .747$; option 2, 97/97%, Fisher's, $p = 1$; option 3, 3/7%, Fisher's, $p = 1$. Participants selected option 1 ("broke the rule") at rates far below chance, 21%, binomial test, $p < .001$, and participants selected option 2 ("unintentionally broke the rule") at rates far exceeding chance, 97%, binomial test, $p < .001$. Almost no one selected option 3 ("did not break the rule"), 5%, binomial test, $p < .001$. Cover story didn't affect participant response to the subsequent dichotomous test question (did the protagonist break the rules?), 83/83%, Fisher's exact test, p =

1. Participants overwhelmingly agreed that the protagonist broke the rules (83%), significantly exceeding the rate of agreement in Experiment 2, binomial test, $p < .001$, test proportion = .56.

These results further support the hypothesis that people engage in excuse validation in order to avoid indirectly blaming a blameless transgressor.

Experiment 6

We have seen that people are surprisingly unwilling to classify blameless rule-breaking as rule-breaking, and we have seen that allowing people the option to classify it as "unintentional" blocks this tendency. But further questions remain about the scope of the effect. For instance, will increasing the societal harm caused by the blameless transgression also block excuse validation? And will excuse validation persist when people are asked not to classify the transgression themselves, but rather to evaluate a third-party's classification of it? The remaining two experiments pursue these questions in turn.

Method

Participants (N = 71, 27 female, aged 18-66 years, mean age = 29 years; 96% reporting English as a native language) were randomly assigned to read one of two stories: Significant Harm and Insignificant Harm. We excluded data from 9 participants who failed comprehension questions designed to ensure that they were reading the scenario carefully enough. But including these participants yielded the same pattern of results described below.

The stories featured Andrew, a technician who works on a vessel. When the vessel performs a certain exercise ("exercise Omega"), it's Andrew's job to send a certain signal ("signal M100") to notify interested parties, according to the ship's rules. In the story for

Significant Harm, Andrew works on a military vessel and if the right signal isn't sent, it could be interpreted as *an act of war* by adversaries. In the story for Insignificant Harm, Andrew works on a commercial vessel and if the right signal isn't sent, *a database will have to be updated manually* by corporate headquarters. In each story, Andrew sends the signal listed in his copy of the ship's rule manual, but his copy contains a misprint and, as a result, he sends the wrong signal ("signal N001"). In the Significant Harm scenario, war ensues; in the Insignificant Harm scenario, headquarters updates its database manually.

After reading the story, participants answered two comprehension questions about which signal the rules require Andrew to send and which signal Andrew actually sends, followed by the test question, "Did Andrew break the rules?" (Yes/No). Participants were then taken to a different screen — they could not go back and change an answer — and asked two more yes/no questions: "Did Andrew unintentionally break the rules?" and "Should Andrew be criticized for breaking the rules?" We asked the criticism question to ensure that participants viewed Andrew as blameless. We asked the "unintentionally broke" question as a within-subjects factor to see whether participants who had just said that Andrew *did not break the rules* were then willing to agree that Andrew *did break the rules unintentionally*. The stories and all the questions are included in Appendix B.

Results and discussion

We are interested in whether increased harm caused by a transgression blocks excuse validation. If it does block excuse validation, then participants in the Significant Harm condition will probably agree at significantly higher rates that Andrew broke the rules. But if it doesn't, then

participants in the Significant Harm condition probably will not agree at significantly higher rates.

There was no effect of condition on whether participants agreed that Andrew broke the rules: 41% answered "yes" in Insignificant Harm compared to 49% in Significant Harm, Fisher's exact test, $p = .635$, n.s. This suggests that dramatically increasing the transgression's harmfulness does not block excuse validation. Overall 45% of participants answered "yes," which doesn't differ from what could be expected by chance, binomial test, $p = .477$, thus replicating the main finding from Experiments 1 and 2.

There was no effect of condition on whether participants subsequently agreed that Andrew unintentionally broke the rules: 68% answered "yes" in Insignificant Harm compared to 82% in Significant Harm, Fisher's exact test, $p = .276$, n.s. Overall 75% of participants agreed that Andrew unintentionally broke the rules, which is much higher than could be expected by chance, binomial test, $p < .001$. Participants were significantly more likely to agree that Andrew unintentionally broke the rules than that Andrew broke the rules, McNemar's test, $p < .001$. Of the 39 participants who said that Andrew did not break the rules, 29 (74%) answered that Andrew unintentionally broke the rules, which exceeds chance rates, binomial test, $p = .003$. Indeed, the most frequent combination of answers to the two rule-breaking questions was the paradoxical, "No, Andrew did not break the rules" but "Yes, Andrew unintentionally broke the rules." Forty-one percent of participants (29/71) selected this particular combination. (See Fig. 1.)

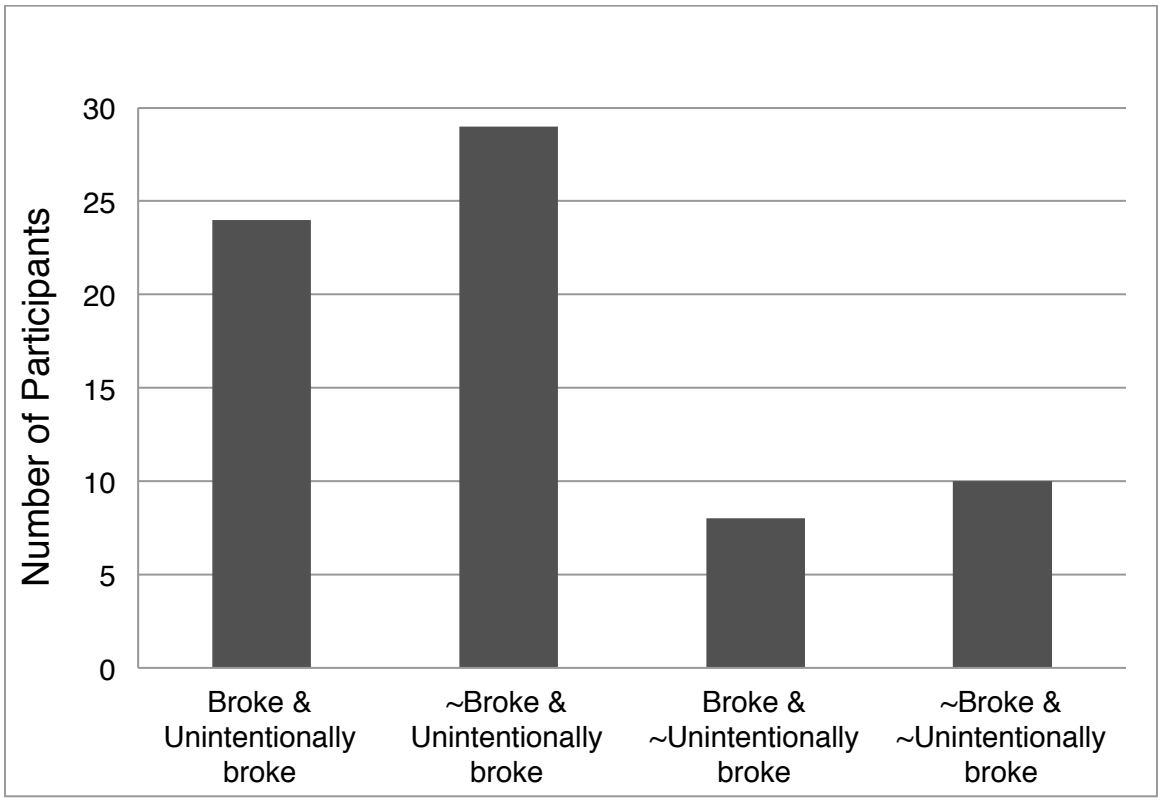

Fig. 1. Experiment 6. Number of participants who opted for the four possible combinations of answers to the two rule-breaking questions: "Did Andrew break the rules?" (Broke) and "Did Andrew unintentionally break the rules?" (Unintentionally broke).

The remaining 10 of the 39 participants who said that Andrew did not break the rules also went on to deny that Andrew unintentionally broke the rules. It is perhaps surprising that a nontrivial number of participants refuse to acknowledge that a rule has been broken even when it is possible to do so in way that emphasizes the rule-breaker's blamelessness. One explanation for this result is that some participants first deny rule-breaking in order to avoid conveying blame and then continue to deny rule-breaking in order to avoid making obviously inconsistent statements. Thus these participants successfully avoid the paradoxical combination of first denying and then affirming that a rule has been broken.

Finally, there was no effect of condition on answer to the "criticism" question, Fisher's exact test, $p = .434$, n.s. Participants in both conditions overwhelmingly answered that Andrew should not be criticized, 89-95%. Overall 92% of participants said that Andrew should not be criticized, which far exceeds what could be expected by chance, binomial test, $p < .001$, and matches exactly the 92% observed in Experiment 1.

Experiment 7

This experiment investigates whether the basic pattern of excuse validation observed above occurs when evaluating a third-party's judgment about blameless rule-breaking.

Method

Participants (N = 73, 28 female, aged 18-63 years, mean age = 31 years; 100% reporting English as a native language) were randomly assigned to read one of two stories: Broke and Unintentionally Broke. We excluded data from 7 participants who failed comprehension questions. Including these participants yielded the same pattern of results described below.

Participants in each condition read the same story used in the Significant Harm condition from Experiment 6. The only difference was an additional sentence at the very end, in which a third party tells Andrew that he broke the rules (unintentionally). Participants in the Broke condition read, "Soon after the vessel finishes its exercise, a war breaks out. The captain approaches Andrew and says, 'You broke the rules.'" The Unintentionally Broke scenario was exactly the same, except for a single word: the captain says to Andrew, "You broke the rules unintentionally."

Participants then answered the same two comprehension questions as in Experiment 6. The test question for the Broke scenario was, “Is the captain right that Andrew broke the rules?” The test question for the Unintentionally Broke scenario was, “Is the captain right that Andrew broke the rules unintentionally?”

Results and discussion

There was an effect of condition on response to the test question, Fisher’s exact test, $p < .002$. In the Broke condition, 54% of participants said the captain was right that Andrew “broke the rules,” which does not differ from what could be expected by chance, binomial test, $p = .743$, n.s., or from the 49% in the Significant Harm condition of Experiment 6 who agreed *in propria persona* that Andrew broke the rules, binomial test, $p = .652$, n.s., test proportion = .49. In the Unintentionally Broke condition, 89% of participants said the captain was right that Andrew “broke the rules unintentionally,” which far exceeds what could be expected by chance, binomial test, $p < .001$, but does not differ from the 82% in the Significant Harm condition of Experiment 6 who agreed *in propria persona* that Andrew broke the rules unintentionally, binomial test, $p = .398$, n.s., test proportion = .82.

These results suggest that excuse validation occurs whether participants are asked to directly judge blameless transgressions for themselves, or they are asked to evaluate a third-party’s judgment about such matters.

General Discussion

Some important philosophical debates in normative theory turn on judgments and intuitions about thought experiments involving blameless agents. Inspired by recent work in social

psychology that has demonstrated the distorting influence that the desire to blame can have ("blame validation"), we investigated whether the desire to excuse can have a similar distorting influence ("excuse validation"). In particular, we wanted to know whether certain debates in normative ethics, epistemology and the philosophy of language had been stalled partly because they have focused on thought experiments that are apt to trigger excuse validation in us when we consider them.

Seven experiments provided evidence that people engage in excuse validation and offered initial support for one explanation of why they do this. When presented with cases of blameless transgression and asked directly whether a transgression occurred, roughly half of participants contradicted the plain facts of the case by denying that the protagonist had transgressed (Experiments 1 and 2). This tendency is not due to a general unwillingness to identify others as transgressors, even when they sincerely believe that they are following the rules or when they have been "tricked" into breaking the rules (Experiment 3). Nor is it due to an inability to identify blameless transgression specifically, because the vast majority of participants identified the action as an "unintentional" transgression (Experiments 4–7). We saw similar response patterns using both between-subjects and within-subjects designs. Participants engaged in excuse validation whether the transgression caused significant or insignificant societal harm (Experiment 6). Participants also engaged in excuse validation when evaluating third-party assertions about blameless transgression (Experiment 7). All our studies were conducted on adults in the Unites States. But there is evidence that intuitions about blame and exculpation vary cross-culturally (Tetlock et al., 2010). Future work could investigate the extent to which excuse validation varies across cultures.

When participants were questioned in a way that did not force them to choose between (i) accurately identifying the transgression and (ii) avoiding unfairly suggesting that the transgressor should be blamed, almost all participants correctly identified that the transgression had taken place. That is, when offered an option that combined both fairness and accuracy, nearly all participants answered accurately. By contrast, when people were forced to choose between accuracy and fairness, just as many chose fairness as chose accuracy. Overall, the findings are well explained by the hypothesis that people engage in excuse validation in order to avoid blaming, either directly or indirectly, a blameless transgressor.

Excuse validation differs from blame validation in one important respect. According to Alicke, blame validation has certain clear limitations. It affects judgments of causality primarily when things are complicated and messy, particularly when there are multiple independent causal factors contributing to an outcome, none of which "obviously necessitate[s] the outcome more than the others" (Alicke 1992: 377). More generally, blame validation "is limited by reality constraints," so that even an extreme bigot probably won't blame a member of the targeted group "for an offense in which he had absolutely no causal involvement" (Alicke 2008: 183–4). In short, potent negative affect can "slant" perceptions but cannot "skew" them beyond all agreement with plain facts. However, we observed that excuse validation not only slants but can radically skew. In Experiments 1 and 2, roughly half of participants implicitly contradicted themselves and the plain facts of the case. For instance, in the Chess story, Nicolas took forty-five seconds to move his rook, and the rules prohibit anything more than thirty seconds, but a majority of participants said that Nicolas did not break the rules. And in Experiment 6, 40% of participants explicitly contradicted themselves by agreeing *both* that Andrew unintentionally broke the rules *and* that Andrew did not break the rules.

We don't suggest that people, especially theoreticians, faced with blameless wrongdoing reason explicitly as follows: "This person clearly broke a rule. But I shouldn't say that he broke a rule. So he didn't actually break a rule." Presumably that is too transparently inconsistent to be a plausible account of the underlying psychological process. The correct account will likely be subtler than that. When faced with blameless wrongdoing, humans implicitly sense that they shouldn't say that the person broke a rule, but they don't necessarily have immediate unproblematic access to the reason why; such evaluations are often unconscious and automatic, even if deliberate reflection can alter them (Fazio 2001; Cunningham & Zelazo 2007; Cushman, Young & Hauser 2006). We're left to make sense of our reaction. This is not always easy.[3] One natural explanation for why we shouldn't say that he broke a rule is that it's false that he broke the rule. We shouldn't expect theorists to be immune to this temptation.[4]

Despite how illogical excuse validation can initially seem, it makes sense that we would be disposed to engage in excuse validation, both from the perspective of individual assessors and as a matter of punitive convention. Consider these in turn. First, criticism and blame aren't taken lightly. Unwarranted criticism can hurt. Even the perception of unwarranted criticism or too strong a tendency toward being "judgmental" can damage relationships. Better to avoid the appearance of unwarranted criticism than to state the truth and suffer negative repercussions.

---

[3] Hauser, Cushman, Young, Jin & Mikhail (2007: 15) report that a large and diverse group of subjects tested on versions of trolley cases "generally failed to provide justifications that could account for the pattern of their judgments."

[4] Tobia, Buckwalter & Stich (2013) report findings which suggest that professional philosophers are susceptible to framing effects when making moral judgments. Schwitzgebel and Cushman (2012) report that professional philosophers, even those specializing in ethics, are susceptible to order effects, and they speculate that a philosopher's "skill at moral reasoning" is primarily a tool for "post-hoc rationalization." Indeed, Schwitzgebel and Cushman observed that non-philosophers were *less* susceptible to ordering effects.

Second, criticism and blame are mild forms of punishment. Punishment is one main way that normative systems are maintained and punitive sentiments are a deep feature of human psychology (Sripada 2005). But policing a norm through punishment can be costly to the enforcer. This creates an incentive to opt out of enforcement, which in turn creates a need for "punishment stabilization mechanisms" that disincentivize opting out (Boyd & Richerson 1992). But if the transgressor's heart is in the right place — that is, if she is already clearly committed to following the norm — then it might well be dysfunctional to punish. Punishing such an individual does not well serve the primary point of punitive practices, which is to provide incentive and motivate compliance. At the same time, it would be bad to encourage an overly lax attitude toward enforcement, since that would also be dysfunctional. Excuse validation is the perfect compromise: we omit blaming her because she "didn't break the rules." One thus avoids sending the disadvantageous signal to observers that one is a bad partner, unwilling to bear one's share of the punitive burden needed to maintain social order (Gintis, Smith & Bowles, 2001).

Acknowledgments — For helpful feedback and conversations, we thank Mark Alfano, Wesley Buckwalter, Gerry Callaghan, Fiery Cushman, Michael Devitt, Mathieu Doucet, Chris Eliasmith, Ori Friedman, Joshua Knobe, Bertram Malle, David Rose, Laurie Rudman, Paul Thagard, and Angelo Turri. Thanks also to audiences at the University of Waterloo (September 2012), the 39th meeting of the Society for Philosophy and Psychology (Brown University, June 2013), and the 6th Rocky Mountain Ethics Congress (University of Colorado, Boulder, August 2013). This research was kindly supported by the Social Sciences and Humanities Research Council of Canada and an Early Researcher Award from the Ontario Ministry of Economic Development and Innovation.

Appendix A: Experiments 1-5

(Baking) Brenda just entered a natural baking contest. She is in the kitchen where the contest is being held, and is starting to prepare her dish. The rules of the contest say that the only sweetener contestants are allowed to use is natural sugar, so Brenda was careful to buy only sweetener clearly labeled "natural sugar." But the label on the package is wrong, because there was a mix-up at the factory: an artificial sweetener that looks just like sugar, SweeTooth, was accidentally packed in a package labeled "natural sugar," without anybody noticing. Brenda isn't aware that this happened. As a result, Brenda is actually using SweeTooth.

1. What kind of sweetener does Brenda think she is using? [Sugar/SweeTooth]
2. Is it reasonable for Brenda to think that she is using that kind of sweetener? [Yes/No]
3. What kind of sweetener is allowed? [Sugar/SweeTooth]
4. What kind of sweetener is Brenda using? [Sugar/SweeTooth]
5. Should Brenda be criticized for using that sweetener? [Yes/No]
6. Is there a sense in which it is incorrect for Brenda to use that sweetener? [Yes/No]

(Farming) Ori just entered the organic farming business. Today Ori received a letter from the agriculture department explaining the rules for organic farming. It says, "Farmers are allowed to use the fertilizer GreenFeed, but not CleanFeed." But Ori's neighbour Carl is also an organic farmer, and he doesn't want Ori competing with him. So Carl sabotaged Ori by intercepting Ori's letter and changing the text. The actual official organic farming rules say, "Farmers are allowed to use the fertilizer CleanFeed, but not GreenFeed." Ori isn't aware that Carl did this. As a result, Ori fertilizes his crops with GreenFeed.

1. What kind of fertilizer does Ori think is allowed? [Green-Feed/CleanFeed]

2. Is it reasonable for Ori to think that? [Yes/No]
3. What kind of fertilizer is allowed? [CleanFeed/GreenFeed]
4. What kind of fertilizer does Ori use? [GreenFeed/CleanFeed]
5. Should Ori be criticized for using that fertilizer? [Yes/No]
6. Is there a sense in which it is incorrect for Ori to use that fertilizer? [Yes/No]

(Asserting) Recently Robert started collecting coins. Today Robert made a purchase for an 1804 US silver dollar at a local coin shop, and put the coin in a display in his dining room. He is having dinner guests over tonight, and he plans to tell them about his new acquisition. But the coin dealer cheated Robert: the coin Robert purchased is actually a 1904 US silver dollar that has been tampered with so that it looks like it says '1804' on it. Robert isn't aware that the dealer did this. As a result, Robert tells his guests, "There is an 1804 US silver dollar in this display."

1. What kind of coin does Robert think is in the display? [1804/1904]
2. Is it reasonable for Robert to think that? [Yes/No]
3. Robert tells his guests that a _____ coin is in the display. [1804/1904]
4. Robert's statement to his guests is _____. [False/True]
5. Should Robert be criticized for making that statement? [Yes/No]
6. Is there a sense in which it is incorrect for Robert to make that statement? [Yes/No]

(Driving) Doreen just had her car serviced and is driving home from the mechanic's shop. She wants to get home without unnecessary delay, but she does not want to break any traffic laws. The traffic law sets the speed limit at 55 miles per hour, so she looks down to see how fast she is going. The speedometer says that she is going 55 miles per hour. But the mechanic intentionally tampered with a setting on the speedometer, so that it often cannot correctly

measure the car's speed. Doreen isn't aware that the mechanic did that. As a result, Doreen is actually driving sixty miles per hour.

1. How fast does Doreen think she is driving? [50/55/60]
2. Is it reasonable for Doreen to think that she is driving that fast? [Yes/No]
3. What is the speed limit? [55/60]
4. How fast is Doreen driving? [60/55]
5. Should Doreen be criticized for driving that speed? [Yes/No]
6. Is there a sense in which it is incorrect for Doreen to drive that speed? [Yes/No]

(Chess) Nicolas just accepted an invitation to participate in a special chess tournament. The invitation explains the tournament's special rule, known as "the rook rule." The invitation says, "A player is not allowed to take more than 60 seconds when moving the rook." But Nicolas's fierce competitor, Vladimir, is also in the tournament and wants to sabotage Nicolas. Vladimir intercepted Nicolas's invitation and tricked him by changing the text. The actual official rook rule says: "A player is not allowed to take more than 30 seconds when moving the rook." Nicolas isn't aware that Vladimir did this. As a result, during the first match, Nicolas takes forty-five seconds to move the rook.

1. How much time does Nicolas think the rook rule allows? [60/30 seconds]
2. Is Nicolas's belief about the rook rule reasonable? [Yes/No]
3. How much time does the rook rule actually allow? [60/30 seconds]
4. How much time does Nicolas take to move his rook? [45/30 seconds]
5. Should Nicolas be criticized for taking that much time? [Yes/No]
6. Is there a sense in which it is incorrect for Nicolas to take that much time? [Yes/No]

Appendix B: Experiments 6-7

(Vessel: Significant/Insignificant Harm) Andrew is a technician on a [military/commercial] vessel. When the vessel performs certain exercises, he is expected to notify [nearby adversaries/corporate headquarters] by sending certain signals, as detailed in the vessel's rule manual. If the vessel performs an exercise without sending the right signal, the consequences [could/will] be [significant/insignificant]: [it could be interpreted as an act of war/headquarters will have to update its database manually]. ¶[5] Today the vessel's captain orders the crew to perform exercise Omega, which they had never done before. Andrew consults the copy of the vessel's rule manual that he was given, and it says, "If the vessel performs exercise Omega, send signal N001." But Andrew's copy of the manual contains a misprint. The actual official rule says, "If the vessel performs exercise Omega, send signal M100." Andrew isn't aware of the misprint. As a result, Andrew sends signal N001. ¶ Soon after the vessel finishes its exercise, headquarters updates its database manually. [The captain approaches Andrew and says, "You broke the rules [unintentionally]."]

1. The rules officially require Andrew to send signal _____. [M100/N001]
2. Andrew sent signal _____. [N001/M100]
3. Did Andrew break the rules [unintentionally]? [Y/N]

>>page break<< (the following two questions were also presented in Experiment 6)

4. Did Andrew unintentionally break the rules? [Y/N]
5. Should Andrew be criticized for breaking the rules? [Y/N]

References

[5] Indicates paragraph break on the participant's screen

Alicke, M. D. (1992). Culpable causation. Journal of Personality and Social Psychology, 63(3), 368–378.

Alicke, M. D. (2000). Culpable control and the psychology of blame. Psychological Bulletin, 126(4), 556. doi:10.1037/0033-2909.126.4.556

Alicke, M. D. (2008). Blaming badly. Journal of Cognition and Culture, 8(1), 179–186.

Alicke, M., & Rose, D. (2010). Culpable control or moral concepts? Behavioral and Brain Sciences, 33(04), 330–331. doi:10.1017/S0140525X10001664

Aristotle. (350 BCE/1941). Nichomachean ethics. (W. D. Ross, Trans., R. McKeon, Ed.).

Austin, J. L. (1956). A plea for excuses. Proceedings of the Aristotelian Society, 57, 1–30.

Benton, M. A. (2012). Knowledge norms: assertion, belief, and action. Rutgers University, New Brunswick.

Botterell, A. (2009). A Primer on the Distinction between Justification and Excuse. Philosophy Compass, 4(1), 172–196. doi:10.1111/j.1747-9991.2008.00200.x

Boyd, R., & Richerson, P. J. (1992). Punishment allows the evolution of cooperation (or anything else) in sizable groups. Ethology and Sociobiology, 13, 171–195.

Brandt, R. B. (1969). A utilitarian theory of excuses. The Philosophical Review, 78(3), 337–361.

Buckwalter, W., & Turri, J. (2013). Telling, showing and knowing: a unified theory of pedagogical norms. Analysis, 1–5. doi:10.1093/analys/ant092

Cunningham, W. A., & Zelazo, P. D. (2007). Attitudes and evaluations: a social cognitive neuroscience perspective. Trends in Cognitive Sciences, 11(3), 97–104. doi:10.1016/j.tics.2006.12.005

Cushman, F. (2008). Crime and punishment: Distinguishing the roles of causal and intentional analyses in moral judgment. Cognition, 108(2), 353–380. doi:10.1016/j.cognition.2008.03.006

Cushman, F., Young, L., & Hauser, M. (2006). The role of conscious reasoning and intuition in moral judgment testing three principles of harm. Psychological Science, 17(12), 1082–1089.

Daniels, N. (1979). Wide reflective equilibrium and theory acceptance in ethics. The Journal of Philosophy, 76(5), 256–282.

Douven, I. (2006). Assertion, Knowledge, and Rational Credibility. Philosophical Review, 115(4), 449–485. doi:10.1215/00318108-2006-010

Edmondson, A. C. (2004). Learning from Mistakes is Easier Said than Done: Group and Organizational Influences on the Detection and Correction of Human Error. Journal of Applied Behavioral Science, 40(1), 66–90. doi:10.1177/0021886304263849

Fast, N. J., & Tiedens, L. Z. (2010). Blame contagion: The automatic transmission of self-serving attributions. Journal of Experimental Social Psychology, 46(1), 97–106. doi:10.1016/j.jesp.2009.10.007

Fazio, R. H. (2001). On the automatic activation of associated evaluations: An overview. Cognition & Emotion, 15(2), 115–141.

Feinberg, J. (1970). Doing and deserving: essays in the theory of responsibility. Princeton: Princeton University Press.

Gintis, H., Smith, E. A., & Bowles, S. (2001). Costly Signaling and Cooperation. Journal of Theoretical Biology, 213(1), 103–119. doi:10.1006/jtbi.2001.2406

Goodman, N. (1955). Fact, fiction, and forecast. Cambridge, Mass: Harvard University Press.

Gray, K., & Wegner, D. M. (2011). To escape blame, don't be a hero—Be a victim. Journal of Experimental Social Psychology, 47(2), 516–519. doi:10.1016/j.jesp.2010.12.012

Grice, H. P. H. (1989). Studies in the way of words. Cambridge: Harvard University Press.

Guglielmo, S., & Malle, B. F. (2010). Can Unintended Side Effects Be Intentional? Resolving a Controversy Over Intentionality and Morality. Personality and Social Psychology Bulletin, 36(12), 1635–1647. doi:10.1177/0146167210386733

Harman, G. (1986). Change in View. Cambridge, Mass.: MIT Press.

Hauser, M., Cushman, F., Young, L., Kang Xing Jin, R., & Mikhail, J. (2007). A dissociation between moral judgments and justifications. Mind & Language, 22(1), 1–21.

Hill, C., & Schechter, J. (2007). Hawthorne's lottery puzzle and the nature of belief. Philosophical Issues, 17, 102–122.

Hindriks, F. (2008). Intentional action and the praise-blame asymmetry. The Philosophical Quarterly, 58(233), 630–641. doi:10.1111/j.1467-9213.2007.551.x

Inbar, Y., Pizarro, D. A., & Cushman, F. (2012). Benefiting From Misfortune: When Harmless Actions Are Judged to Be Morally Blameworthy. Personality and Social Psychology Bulletin, 38(1), 52–62. doi:10.1177/0146167211430232

Kagan, S. (1998). Normative ethics. Boulder, Colorado: Westview Press.

Knobe, J. (2010). Person as scientist, person as moralist. Behavioral and Brain Sciences, 33(04), 315–329. doi:10.1017/S0140525X10000907

Kvanvig, J. (2009). Assertion, knowledge, and lotteries. In D. Pritchard & P. Greenough (Eds.), Williamson on knowledge (pp. 140–160). Oxford: Oxford University Press.

Lackey, J. (2007). Norms of assertion. Nous, 41(4), 594–626.

Mehlman, R. C., & Snyder, C. R. (1985). Excuse theory: A test of the self-protective role of attributions. Journal of Personality and Social Psychology, 49(4), 994. doi:10.1037/0022-3514.49.4.994

Mill, J. S. (1861/1979). Utilitarianism. (G. Sher, Ed.). Indianapolis: Hackett.

Nadelhoffer, T. (2006). Bad acts, blameworthy agents, and intentional actions: Some problems for juror impartiality. Philosophical Explorations, 9(2), 203–219.

Pizarro, D., Uhlmann, E., & Salovey, P. (2003). Asymmetry in Judgments of Moral Blame and Praise: The Role of Perceived Metadesires. Psychological Science, 14(3), 267–272. doi:10.1111/1467-9280.03433

Rawls, J. (1971). A theory of justice. Harvard University Press.

Schaffer, J. (2008). Knowledge in the image of assertion. Philosophical Issues, 18(1), 1–19.

Schwitzgebel, E., & Cushman, F. (2012). Expertise in Moral Reasoning? Order Effects on Moral Judgment in Professional Philosophers and Non-Philosophers. Mind & Language, 27(2), 135–153. doi:10.1111/j.1468-0017.2012.01438.x

Searle, J. R. (1979). Expression and Meaning. Cambridge: Cambridge University Press.

Snyder, C. R., & Higgins, R. L. (1988). Excuses: Their effective role in the negotiation of reality. Psychological Bulletin, 104(1), 23. doi:10.1037/0033-2909.104.1.23

Sripada, C. S. (2005). Punishment and the Strategic Structure of Moral Systems. Biology & Philosophy, 20(4), 767–789. doi:10.1007/s10539-004-5155-2

Strawon, P. F. (1962). Freedom and resentment. Proceedings of the British Academy, 48, 1–25.

Tetlock, P. E., Self, W. T., & Singh, R. (2010). The punitiveness paradox: When is external pressure exculpatory – And when a signal just to spread blame? Journal of Experimental Social Psychology, 46(2), 388–395. doi:10.1016/j.jesp.2009.11.013

Tobia, K., Buckwalter, W., & Stich, S. (2013). Moral intuitions: Are philosophers experts? Philosophical Psychology, 26(5), 629–638. doi:10.1080/09515089.2012.696327

Turri, J. (2010). Prompting challenges. Analysis, 70(3), 456–462. doi:10.1093/analys/anq027

Turri, J. (2011). The Express Knowledge Account of Assertion. Australasian Journal of Philosophy, 89(1), 37–45. doi:10.1080/00048401003660333

Turri, J. (2013). The test of truth: An experimental investigation of the norm of assertion. Cognition, 129(2), 279–291. doi:10.1016/j.cognition.2013.06.012

Tyler, J. M., & Feldman, R. S. (2007). The double-edged sword of excuses: When do they help, when do they hurt. Journal of Social and Clinical Psychology, 26(6), 659–688.

Unger, P. (1975). Ignorance: a case for skepticism. Oxford: Oxford University Press.

Weiner, M. (2005). Must we know what we say? Philosophical Review, 114(2), 227–251.

Williamson, T. (2000). Knowledge and its limits. Oxford: Oxford University Press.