

# Not as Bad as Painted?

## Legal Expertise, Intentionality Ascription, and Outcome Effects Revisited

**Karolina Prochownik (Karolina.Prochownik@ruhr-uni-bochum.de)**

Faculty of Law, Ruhr University Bochum, Universitätsstrasse 150, 44801 Bochum, Germany

**Melina Krebs (Melina.Krebs@ruhr-uni-bochum.de)**

Institute for Philosophy II, Ruhr University Bochum, Universitätsstrasse 150, 44801 Bochum, Germany

**Alex Wiegmann (Alexander.Wiegmann@ruhr-uni-bochum.de)**

Institute for Philosophy II, Ruhr University Bochum, Universitätsstrasse 150, 44801 Bochum, Germany

**Joachim Horvath (Joachim.Horvath@ruhr-uni-bochum.de)**

Institute for Philosophy II, Ruhr University Bochum, Universitätsstrasse 150, 44801 Bochum, Germany

### Abstract

Previous research by Kneer and Bourgeois-Gironde (2017) suggests that legal experts are susceptible to the “severity effect” – they ascribe a higher level of intentionality for actions if they lead to very bad side-effects than when they have somewhat bad side-effects. These results are potentially problematic for the legal system because ascriptions of intentionality in the law explicitly depend on the evaluation of mental states of the agent (*mens rea*), not on the badness of the outcomes she caused. In this paper, we provide and test an alternative explanation of the “severity effect” that has no troubling implications for the law. We suggest that it may be a subtype of a more general “side-effect effect” (Knobe, 2003), which is compatible with certain legal criteria of ascribing intentionality.

**Keywords:** severity effect; intentionality; *mens rea*; criminal law; legal expertise

### Introduction

The “expertise defense” is a claim that intuitions of experts are more reliable than those of non-experts, as far as matters of their professional expertise are concerned. On these grounds, empirical research on lay people’s intuitions triggered in response to hypothetical philosophical cases had been questioned as insufficient to provide insights about the reliability of philosophers’ intuitions (e.g., Ludwig, 2007; Horvath, 2010; Williamson, 2011). However, empirical findings do not provide evidence that professional philosophers’ intuitions are more reliable than ordinary people’s intuitions, and therefore speak against the view that the “expertise defense” works for philosophers (e.g., Feltz & Cokely, 2009; Schwitzgebel & Cushman, 2012; Horvath & Wiegmann, 2016; Wiegmann, Horvath & Meyer, 2020).

“Experimental jurisprudence” is a new research field that examines intuitions and psychological mechanisms underpinning legal or legally relevant concepts (e.g., consent, causation, ownership, reasonableness, intentionality; Sommers, in press; Macleod, 2019; Nancekivell, Millar, Summers & Friedman, 2016; Tobia, 2018). Researchers in this field often examine whether the

intuitions of the folk are congruent with the equivalent concepts in the legal system (e.g., Robinson & Darley, 1995; Kobick & Knobe, 2009). Some authors also argued that legal concepts should be adapted to better accommodate their folk counterparts (e.g., Kobick, 2010). However, an important question that arises for this type of research is whether data from lay people can be generalized to people with legal expertise.

One potential criticism to experimental jurisprudence could be an equivalent to the “expertise defense” argument in philosophy: that legal experts’ intuitions are more reliable in the matters of the law than folk intuitions, and that one cannot, therefore, undermine the reliability of legal intuitions based merely on the observation that folk intuitions are unreliable (i.e., in this context, “incongruent with the law”).

In particular, based on this “legal expertise defense”, it could be argued that potential challenges arising from the discrepancies between folk concepts and their equivalents in the law do not arise, providing that legal experts’ concepts are congruent with the concepts inherent to the legal system. For instance, even though lay people’s intuitions on some legal matters vary and depart from the law, justice could still be served, providing that legal professionals’ intuitions are stable and congruent with the law. In fact, lack of systematicity in folk intuitions could provide support for the departure from these intuitions in favor of adopting technical criteria in the law.

Despite its clear practical and societal importance, research on whether intuitions of legal experts are more reliable than folk intuitions have been scarce so far (e.g., Kneer & Bourgeois-Gironde, 2017; Donelson & Hannikainen, 2020; Prochownik, 2019; Tobia, in preparation). In this paper, we aim to make a modest contribution to this existing research by further investigating whether the ascriptions of intentionality of legal experts and lay people depend on the degree of negativity of the side-effects of evaluated actions, and therefore are inconsistent with the concept of *mens rea* in the criminal law (Kneer & Bourgeois-Gironde, 2017).

## Outcome Effects on Intentionality Ascriptions

Kneer and Bourgeois-Gironde (2017; henceforth: K&B), presented 32 professional judges in France with the following vignette describing the agent's action as leading to a "somewhat bad" or "very bad" side-effect (experimental manipulation in square brackets):

The mayor of a small beach town is approached by his advisor who says: "We could build a new highway connection. This would make car traffic much more efficient. However, there would be [minor/severe] adverse effects on the environment. During construction, the animals in the construction zone will [be disturbed/die]. This is [only temporary/not a temporary condition], [everything goes/things will not go] back to normal once construction is finished." The mayor responds: "I don't care at all about the environment. All I care about is making car traffic as efficient as possible. Let's build the new highway connection." They build the new highway connection. The animals in the zone are [temporarily disturbed/die]. [Everything goes/Things do not go] back to normal after construction is finished.

To what extent do you agree/disagree with the following statement: The mayor intentionally harmed the environment.

The researchers found that judges assigned a significantly higher degree of intentionality in the "very bad" condition than in the "somewhat bad" condition. They described this finding as the "severity effect".

The "severity effect" seems very problematic for the law. It violates a fundamental principle of many criminal law systems: subjective and objective components of a crime should be evaluated independently. In particular, more severe outcomes should not – *ceteris paribus* – lead to higher intentionality ascriptions. The corresponding legal concept, called *mens rea*, does not explicitly rely on the degree of negativity of an outcome. According to K&B, their findings provide evidence that legal experts ascribe intentionality in a way that is incompatible with the technical *mens rea* concept of the law (which might speak against the expertise defense for lawyers, cf. K&B, p. 141).

However, K&B's conclusion might be premature. Let us have a closer look at the "beach town" vignette. While the "very bad" version of this vignette undoubtedly describes severe harm to the environment (the animals in the construction zone die, and things do not go back to normal after construction is finished), it is far from clear whether the outcome in the "somewhat bad" condition was perceived as harmful (the animals are only temporarily disturbed, and everything goes back to normal after construction is finished).<sup>1</sup> It is then possible that the alleged "severity

<sup>1</sup> Note that there are potentially further problems with the vignette used by K&B. As indicated by anonymous reviewers of this paper, although the side-effect of the construction specifically concerns the animals in the construction zone, the intentionality question is in terms of harm to the environment (which might

effect" found by K&B might be an instance of the well-known "side-effect effect" (or "Knobe effect").<sup>2</sup> Broadly put, the "side-effect effect" consists in ascribing (more) intentionality to (morally) bad side-effects of action, but no (or less) intentionality to (morally) good or neutral side-effects (e.g., Knobe, 2003; Wright & Bengson, 2009). This effect was also observed in judges (see K&B, Experiment 1). However, the "side-effect effect" does not necessarily pose a challenge to the legal system (see Kobick & Knobe 2009; Kobick, 2010, for different views).

Legal experts are trained to ascribe intentionality in case of actions and their side-effects that are highly socially *negative*, not to those that are positive or neutral. In particular, they are expected to apply specifically legal standards of *mens rea* from criminal law to, for instance, the case of an agent's actions that harm the environment. When the agent helps or does not harm the environment, there is no expertise-specific expectation that legal professionals should apply different criteria of intentionality ascriptions than lay people. In other words, applying legal standards of intentionality only for cases when harm occurs is consistent with the concept of *mens rea*. And, when harm occurs, there are certain forms of *mens rea* in criminal law that may allow assigning intentionality to the agent who does not desire but foresees the negative side-effects of his actions (e.g., *dolus directus 2* or *dolus eventualis* in Germany and legal systems influenced by it; Taylor, 2004).

To sum up, unlike the "side-effect effect", the "severity effect"—i.e., legal experts (*ceteris paribus*) assigning more intentionality to actions with a very bad outcome than to actions with a somewhat bad outcome – poses a severe challenge to the legal practice. In particular, it seems to undermine the view that legal professionals are correctly using a technical concept of the law when assigning intentionality. However, the challenge raised by K&B rests on the assumption that harm really occurred in the "somewhat bad" version of the vignette. Their findings only cause trouble for the practice of law if this assumption is correct.

To assess this assumption, we present data from two preregistered experiments with law students in Germany (because in the German criminal law community, there are

---

encompass a much broader set of consequences). Besides, the fact that the mayor aimed at a positive result (making traffic more efficient) might have been weighed against the negative side-effects by the participants. This positive result could, in their view, have outweighed the side-effect of temporarily disturbing the animals, but not of causing their death (Prochownik, 2019). Although we do not focus on these alternative hypotheses here, we address them indirectly in Experiment 2, which investigates the "severity effect" across a new set of scenarios and questions.

<sup>2</sup> Our proposal is, therefore, the reverse of the hypothesis of K&B (2017) that the "Knobe effect" might be a subtype of a broader "severity effect" (p. 143). However, note that in this paper, we focus on examining scenarios intended to involve *some degree* of harm (whereas the original study on the "Knobe effect" involves contrasting positive and negative side-effects, e.g., Knobe, 2003).

relatively clear and generally accepted criteria of particular cases of *mens rea*, which makes it easier to test whether people depart from them). To conduct well-powered experiments, and because professional judges are not easily accessible in great numbers, we recruited advanced law students instead (from the 3<sup>rd</sup> semester on). Since the “severity effect” goes against a very fundamental principle of the legal system that is routinely covered in introductory lectures in criminal law at German universities, we assumed that this level of legal education should already make students sufficiently resistant to outcome biases (if at all). We also examined lay people in Germany, as the “severity effect” was only tested in legal experts by K&B, and it was not clear whether it applies to lay people too.

## Experiment 1

In Experiment 1, we tested whether the “severity effect” replicates in advanced law students and lay people in Germany. For this purpose, we translated the original stimuli from French to German. To conduct an exact replication of the K&B study, we kept the vignettes and questions as close to the French original as possible. Besides, we asked new questions about harm to test if some participants in the “somewhat bad” case thought no harm occurred at all compared to those in the “very bad” case.

## Methods

160 participants (80 law subjects and 80 lay subjects)<sup>3</sup> were assigned either to the “somewhat bad” or the “very bad” outcome condition of the original “beach town” scenario. After reading the story, participants indicated whether they disagree or agree with the statement, “The mayor intentionally harmed the environment” (from “1 – completely disagree”, to “7 – completely agree”). Participants who disagreed with the intentionality statement (and chose responses from “1 – completely disagree”, to “3 – rather disagree”) were presented with a follow-up question of understanding, asking them why they disagreed. Participants could choose one of the following two options: “because the environment was not harmed at all” or “the mayor harmed the environment, but he didn’t do it intentionally”. Finally, participants were asked two questions about harm (whether they agreed or disagreed with the statements “The mayor harmed the environment” and “The environment was harmed by the construction”), and whether they thought the mayor committed a crime (“yes” or “no”).<sup>4</sup> Two additional questions about harm were included to further explore whether some participants perceive the “somewhat bad” condition as involving no harm at all in comparison to the “very bad” condition.

<sup>3</sup> The sample size was determined based on the “small telescope” principle, according to which the sample size of the replication study should be 2.5 times the original sample size (Simonsohn, 2015). In K&B (2017), there were 32 participants.

<sup>4</sup> See Supplementary Materials online for the full set of questions: <https://osf.io/n9h2b/>.

Sample sizes, exclusion criteria, and main analyses were preregistered: <https://osf.io/w3gt8> and <https://osf.io/ncy9b>. Study materials and data are also available online: <https://osf.io/n9h2b> (see folder: Experiment 1).

## Results

**Law Participants** We analyzed 80 valid advanced law students’ responses.<sup>5</sup> We replicated the original finding of K&B: there was a significant difference in intentionality ascriptions between the “somewhat bad” ( $M = 4.20$ ,  $SD = 1.847$ ) and the “very bad” condition ( $M = 4.95$ ,  $SD = 1.932$ ),  $t(78) = -1.784$ ,  $p = .039$  (one-tailed;  $p = .078$  if two-tailed)<sup>6</sup>,  $d = 0.397$ . Excluding five participants who indicated, in the understanding question, that they did not assign intentionality because the environment was not harmed (12.2% out of 41 participants in the “somewhat bad” condition)<sup>7</sup> decreased the “severity effect”. The average intentionality assigned in the “somewhat bad” condition ( $M = 4.50$ ,  $SD = 1.732$ ) did not significantly differ anymore from the average judgment assigned in the “very bad” condition ( $M = 4.95$ ,  $SD = 1.932$ ),  $t(73) = -1.056$ ,  $p = .147$  (one-tailed;  $p = .295$  if two-tailed),  $d = 0.245$  (Figure 1).

Moreover, a descriptively higher proportion (24.4%) of participants indicated that they disagreed with the statement that the mayor harmed the environment in the “somewhat bad” case, as compared to 12.8% in the “very bad” case.<sup>8</sup> Similarly, 14.6% indicated that the environment was not harmed by the construction in the “somewhat bad” condition, while only 2.6% did so in the “very bad” condition. Finally, there was a tendency to judge the mayor’s action as more criminal in the “very bad” condition (35.9% thought he committed a crime in this condition, in comparison to 17.1% in the “somewhat bad” condition).

**Non-Law Participants** We analyzed 80 valid responses from lay subjects in the same fashion.<sup>9</sup> Again, we replicated the “severity effect”: participants perceived the agent’s action as more intentional in the “very bad” case ( $M = 6$ ,  $SD = 1.320$ ) in comparison to the “somewhat bad” case ( $M = 4.48$ ,  $SD = 1.754$ ),  $t(78) = -4.393$ ,  $p < .001$ ,  $d = 0.98$ ).

<sup>5</sup> After preregistered exclusion criteria were applied. Only the first 80 valid responses were included.

<sup>6</sup> We assumed that K&B’s “severity effect” hypothesis is directional, and so we took one-tailed  $p$  values as decisive, but we report the two-tailed values as well for Experiment 1.

<sup>7</sup> Nobody chose this response in the “very bad” condition.

<sup>8</sup> We do not report inferential statistics for the harm variables for two reasons. First, the main focus of this experiment is the replication of the “severity effect”, and the sample size was chosen to achieve sufficient power to replicate this effect, not in terms of being able to detect potentially small differences regarding the perceived harm in the “somewhat bad” versus the “very bad” case. Second, we used several questions about harm to explore if some participants thought that no harm occurred in the “somewhat bad” case (we did not have specific hypotheses for these questions).

<sup>9</sup> We excluded participants who didn’t satisfy our preregistered criteria. Only the first 80 valid responses were included.

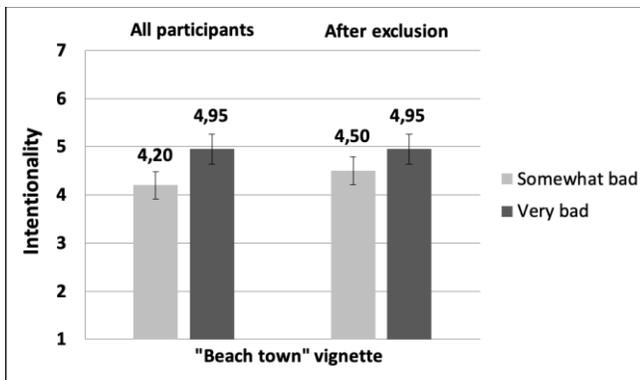


Figure 1: Intentionality judgments of law participants in two severity conditions (all participants, and after the exclusion of those who indicated that no harm occurred). Error bars represent standard errors.

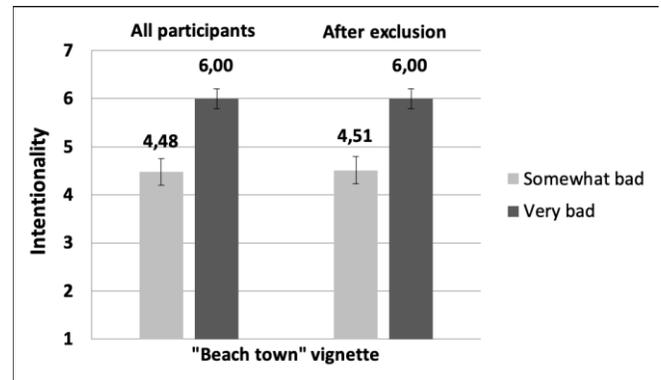


Figure 2: Intentionality judgments of non-law participants in two severity conditions (all participants, and after the exclusion of one person indicating that no harm occurred). Error bars represent standard errors.

Excluding one participant (in the “somewhat bad” condition) who, in the follow-up question, indicated that no harm occurred did not change this finding (Figure 2). No person indicated that no harm occurred in the “very bad” condition.

Besides, only a small proportion of non-law participants responded that the mayor did not harm the environment (12.5% in the “somewhat bad” case versus 2.5% in the “very bad” case). Similarly, 10% answered the environment was not harmed by the construction in the “somewhat bad” case versus 0% in the “very bad” case. Finally, participants’ judgment on whether the mayor’s action was a crime differed depending on the condition: 52.5% perceived it to be criminal in the “very bad” condition compared to 12.5% in the “somewhat bad” condition.<sup>10</sup>

## Discussion

We replicated the “severity effect” in lay people and advanced law students: intentionality ascriptions were higher when the outcome was more severe. However, our results also provide initial evidence for an alternative interpretation of our replication and the previous findings. The “severity effect” in law students and legal professionals (as previously examined by K&B) might be based on the fact that a higher proportion of participants in the “somewhat bad” condition, as compared to the “very bad” condition, think that no harm occurred in the “beach town” vignette.

If this hypothesis is correct, the findings of K&B (2017) might not be troubling for the practice of law, because ascribing intentionality in the criminal law takes place with respect to socially negative (harmful) actions and outcomes. The findings would be troubling only if people with legal expertise differentiated the level of intentionality between

two *clear* harm cases with different degrees of negativity, which would explicitly go against the legal principle of judging subjective elements of the crime (i.e., an agent’s mental states) independently of the evaluation of outcomes. The aim of the second experiment is to test this hypothesis directly.<sup>11</sup>

## Experiment 2

The main goal of Experiment 2 was to test whether the “severity effect” would occur in an experimental setting that rules out interpretations of the effect that are not troubling for the practice of law. In particular, we aimed to make it clear that genuine harm occurred in the “somewhat bad” condition (while still being evidently less harm than in the “very bad” condition). For this reason, we created four new vignettes describing different types of harm, and proportioned the degree of harm adequately to each of the two “severity” conditions. To make sure that our manipulation worked as intended, we presented participants with a question about the respective degree of harm after they answered the main question about intentionality.

## Methods

We used a 2 x 4 mixed design with *outcome severity* (somewhat bad vs. very bad) as a between-subject factor and *type of vignette* as a within-subject factor. 394 participants (194 law and 200 non-law subjects) were assigned to four scenarios (presented in random order) either in the “somewhat bad” or “very bad” variant. In the instructions, participants were asked to imagine that they

<sup>10</sup> We also found the effect of severity of outcome on ascriptions of blame, for both law participants,  $t(78) = -4.128, p < .001$  (one- and two-tailed),  $d = 0.922$ , and non-law participants,  $t(78) = -2.486, p = .0075$  (one-tailed; if two-tailed  $p = .015$ ),  $d = 0.556$ .

<sup>11</sup> Note that we do not analyze non-law subjects in detail here, or compare them with law students because we examined the former primarily to investigate whether the “severity effect” also occurs in lay people (which was not tested by K&B). However, because of the reviewers’ suggestions, we conducted exploratory analyses on potential differences between law and lay participants (see Supplementary Analyses online: <https://osf.io/n9h2b>).

were a judge deciding the cases, and to answer a couple of questions about these law-related cases.

For instance, in the “environment” scenario (a modified version of K&B’s “beach town” vignette) we specified that construction would cause air pollution *and* some animals’ health would be harmed (in the “somewhat bad” condition), or some animals would die (in the “very bad” condition). In the “hospital” vignette, we stated that patients would lose 10% of their hearing capacity (in the “somewhat bad” condition), or that they would completely lose their hearing capacity (in the “very bad” condition) as a side-effect of a chief physician’s decision to test a new medication.<sup>12</sup>

After reading each scenario, participants indicated how much they agreed with the statement: “The [agent] harmed the [victim] intentionally” (on a scale from “1 – completely disagree to “7 – completely agree”). They also evaluated the degree of severity of harm (“In your opinion how big was the harm caused by the [agent]”; on a scale from “1 – very little” to “7 – very big”).<sup>13</sup>

Main analyses and dependent variables, sample sizes, and exclusion criteria were preregistered: <https://osf.io/ynzb5> and <https://osf.io/r9fvz>. Full study stimuli and data are also available at <https://osf.io/n9h2b> (see folder: Experiment 2).

## Results

**Law Participants** We included 194 valid responses of advanced law students,<sup>14</sup> and conducted a mixed ANOVA with the severity of outcome as a between-subject factor and vignette as a within-subject factor. We found no significant main effect of outcome severity on judgments of intentionality,  $F(1, 192) = .975, p = .325, \eta_p^2 = .005$  (Figure 3).<sup>15</sup> In contrast, such a main effect was observed for judgments of the degree of harm,  $F(1,192) = 6.545, p = .011, \eta_p^2 = .033$  (Figure 4).<sup>16</sup> Although law students perceived the harm in the “somewhat bad” and the “very bad” condition as significantly different, they did not assign different intentionality levels.

In addition, because the overall difference in the harm ratings was not very strong ( $M = 5.441, SD = 0.829$  in the “somewhat bad” conditions,  $M = 5.745, SD = 0.824$  in the “very bad” conditions), we performed a one-way ANOVA for the scenario for which the difference in harm-ratings between the two conditions was the highest (the “data”

scenario with  $M = 5.92, SD = 1.093$  in the “very bad” condition and  $M = 5.25, SD = 1.412$  in the “somewhat bad” condition,  $d = 0.53$ ). Even in this case, we found no main effect of outcome severity on judgments of intentionality,  $F(1, 192) = .816, p = .368, \eta_p^2 = .004$ .

**Non-Law Participants** We included 200 valid lay subjects’ responses to the analysis,<sup>17</sup> and performed a mixed ANOVA with severity of outcome as a between-subject factor and vignette as a within-subject factor. There was no main effect of outcome severity on intentionality ascriptions,  $F(1, 198) = 1.758, p = .186, \eta_p^2 = .009$  (Figure 5),<sup>18</sup> while such an effect was observed for the ratings of harm,  $F(1,198) = 17.884, p < .001, \eta_p^2 = .083$  (Figure 6). Altogether, there was no “severity effect”, although people perceived the degree of harm to be different in the two outcome conditions.

Additionally, because the difference in harm ratings was not very strong overall (in the “somewhat bad” conditions  $M = 5.537, SD = 0.723$ , in the “very bad” conditions  $M = 5.972, SD = 0.726$ ), we performed a separate analysis only for the scenario with the most pronounced difference in ratings of harm in the two conditions (the “data” scenario with  $M = 5.45, SD = 1.220$  in the “somewhat bad” condition, and  $M = 5.98, SD = 1.116$  in the “very bad” condition,  $d = 0.453$ ). A one-way ANOVA with severity of outcome as fixed factor and intentionality as dependent variable was not significant,  $F(1, 198) = 1.977, p = .161, \eta_p^2 = .010$ .

## Discussion

The results of our second experiment indicate that the “severity effect” in law participants and non-law participants disappears if it is made clear that harm also occurred in the “somewhat bad” condition. Although participants in both groups indicated that harm occurred in the “somewhat bad” and in the “very bad” condition, and perceived the harm as significantly larger in the “very bad” condition, they did not choose significantly higher intentionality ratings in the “very bad” case in comparison to the “somewhat bad” case.

However, there is a potential concern about this interpretation. Although both groups of participants rated the relevant harm to be significantly larger in the “very bad” than in the “somewhat bad” condition, and the difference is intuitively substantial, the absolute difference in harm ratings was actually not very large. Hence, the possibility remains that if the manipulation is made even stronger (while still ensuring that harm is perceived to happen in the “somewhat bad” condition), a “severity effect” would eventually occur – which would be troubling for the practice of law.

<sup>12</sup> The full text of all four study vignettes and questions (in German and English) is provided online: <https://osf.io/n9h2b>.

<sup>13</sup> See online Supplementary Materials for all study questions.

<sup>14</sup> We included participants who satisfied our preregistered criteria, and that took part before 15.08.2019.

<sup>15</sup> There was also no significant main effect of outcome severity on judgments of blame,  $F(1, 192) = 0.392, p = .532, \eta_p^2 = .002$ .

<sup>16</sup> To address some reviewers’ comments that the “environment” vignette could be problematic (for, in this case, law participants perceived the harm to be bigger in the “somewhat bad” condition than in the “very bad” condition; Figure 4), we repeated the main analyses only for the three remaining scenarios. This had no impact on the main findings (see Supplementary Analyses at <https://osf.io/n9h2b>).

<sup>17</sup> We excluded participants who did not satisfy the preregistered criteria. Only the 200 first valid responses were included.

<sup>18</sup> We also did not find a significant effect of outcome severity on blame ascriptions,  $F(1, 198) = 3.174, p = .076, \eta_p^2 = .016$ .

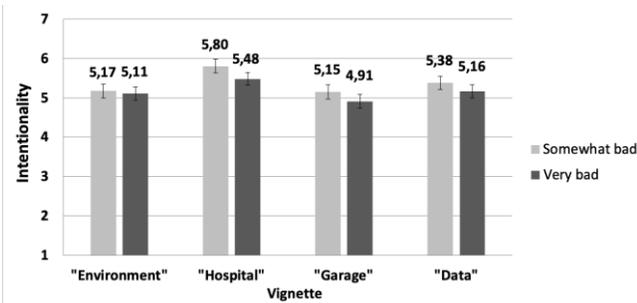


Figure 3: Intentionality judgments of law participants in two severity conditions for four scenarios. Error bars represent standard errors.

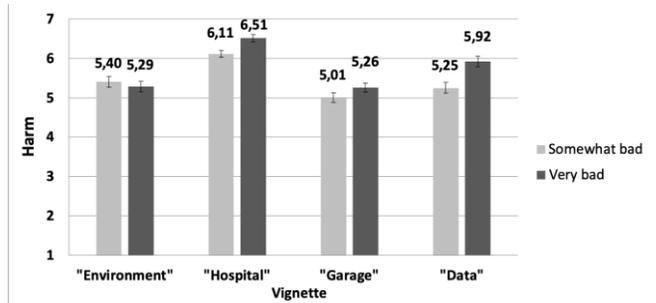


Figure 4: Harm ratings of law participants in two severity conditions for four scenarios. Error bars represent standard errors.

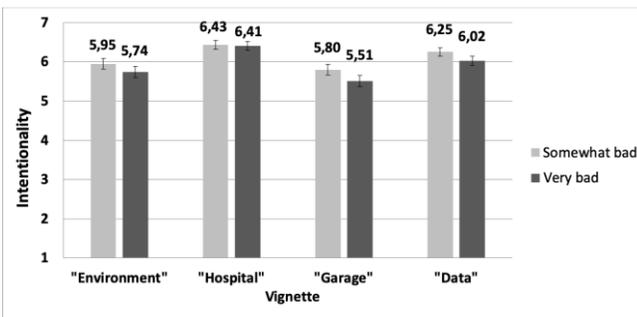


Figure 5: Intentionality judgments of non-law participants in two severity conditions for four scenarios. Error bars represent standard errors.

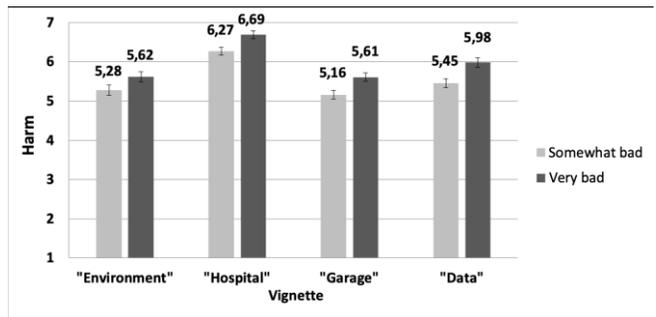


Figure 6: Harm ratings of non-law participants in two severity conditions for four scenarios. Error bars represent standard errors.

### General Discussion

Our two experiments cast doubt on K&B's conclusion that legal experts ascribe intentionality in a way that is problematic for legal practice. In Experiment 1, we replicated the "severity effect", using the original "beach town" vignette with law and non-law participants (K&B, 2017). We found that both groups ascribed higher levels of intentionality to the agent who caused a "very bad" outcome in comparison to a "somewhat bad" outcome. However, we also found initial evidence that this effect in law participants might be due to a higher proportion of participants in the "somewhat bad" condition, as compared to the "very bad" condition, who think that no harm occurred. If this explanation were correct, the findings would be compatible with the legal concept(s) of intentionality in many criminal law systems, and therefore less problematic than previously suggested.

In Experiment 2, we thus created a new set of scenarios, making sure that both the "somewhat bad" and "very bad" outcomes were perceived as clearly harmful. This time, we did not find the "severity effect" either in law or in non-law participants, although both groups perceived both outcomes to be harmful, and significantly different in degree. This finding again suggests that intentionality ascriptions of people with legal education might be less problematic than previously assumed (or not at all). However, we would also like to note a potential limitation of our second experiment.

The difference in the perceived degree of harm between the two outcome conditions was not very large overall. Therefore, we analyzed the data only for the scenario with the largest difference in these ratings, where the "severity effect" was most likely to show up. But even in this case, the effect was not present in either of the two tested groups. Still, the possibility remains that if the manipulation were made even stronger (while outcomes still being perceived as harmful in both conditions), an effect of severity of harm would occur. Thus, one potential future research direction could be to investigate an even stronger contrast between the two "bad outcome" conditions, such that the difference in harm ratings becomes even larger. This could provide further and more decisive evidence about whether there really is a troubling "severity effect" for the practice of law.

Finally, our results are also relevant for discussions about the reliability of expert intuitions in experimental philosophy and jurisprudence. Our findings do not support the view that the legal experts' concept of intentionality is incompatible with the technical concept of *mens rea* in the law (as it is construed in the German legal system). Results of our first experiment suggest that intentionality ascriptions of law participants were less affected by the severity of outcome than those of non-law participants. Overall, this suggests that the expertise defense in the area of law might be more successful than in philosophy. This hypothesis should be subject to future empirical research.

## Acknowledgments

We would like to thank Markus Kneer for sharing the original vignettes and questions of the French study (K&B, 2017), and his valuable feedback on a previous version of this manuscript. We would also like to thank four anonymous reviewers for their very helpful comments and suggestions.

## References

- Donelson, R., & Hannikainen, I. (2020). Fuller and the folk: The inner morality of law revisited. In T. Lombrozo, J. Knobe, & S. Nichols (Eds.), *Oxford studies in experimental philosophy* (Vol. 3). Oxford: Oxford University Press.
- Feltz, A. & Cokely E. T. (2009). Do judgments about freedom and responsibility depend on who you are? Personality differences in intuitions about compatibilism and incompatibilism. *Consciousness and Cognition*, 18(1), 342–350.
- Horvath, J. (2010). How (not) to react to experimental philosophy. *Philosophical Psychology*, 23(4), 447–480.
- Horvath, J., & Wiegmann, A. (2016). Intuitive expertise and intuitions about knowledge. *Philosophical Studies*, 173(10), 2701–2726.
- Kneer, M. & Bourgeois-Gironde, S. (2017). Mens rea ascription, expertise and outcome effects. Professional judges surveyed. *Cognition*, 169, 139–146.
- Knobe, J. (2003). Intentional action and side effects in ordinary language. *Analysis*, 63, 190-193.
- Kobick, J. (2010). Discriminatory intent reconsidered: Folk concepts of intentionality and equal protection jurisprudence. *Harvard Civil Rights-Civil Liberties Law Review*, 45, 517-562.
- Kobick, J., & Knobe, J. (2009). Interpreting intent: How research on folk judgments of intentionality can inform statutory analysis. *Brooklyn Law Review*, 75, 409-431.
- Ludwig, K. (2007). The epistemology of thought experiments: First person versus third person approaches. *Midwest Studies in Philosophy*, 31(1), 128–159.
- Macleod, J. (2019). Ordinary causation: A study in experimental statutory interpretation. *Indiana Law Journal*, 94 (3), 956-1029.
- Nancekivell, S., Millar, J., Summers, P., & Friedman, O. (2016). Ownership rights. In J. Sytsma & W. Buckwalter (Eds.), *A companion to experimental philosophy*. Oxford: Wiley-Blackwell.
- Prochownik, K. (2019) Conditions of criminal responsibility from the perspective of experimental legal philosophy (Unpublished doctoral dissertation). Jagiellonian University, Krakow.
- Robinson, P. H., & Darley, J. M. (1995). Justice, liability, and blame: Community views and the criminal law. Boulder: Westview Press.
- Simonsohn, U. (2015). Small telescopes: Detectability and the evaluation of replication results. *Psychological Science*, 26(5), 559-569.
- 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.
- Sommers, R. (in press). Commonsense consent. *Yale Law Journal*. Retrieved from <https://ssrn.com/abstract=2761801>
- Taylor, G. (2004). Concepts of intention in German criminal law. *Oxford Journal of Legal Studies*, 24(1), 99-127.
- Tobia, K. P. (2018). How people judge what is reasonable. *Alabama Law Review*, 70, 293.
- Tobia, K. P. (2020, April 6). *Legal concepts and legal expertise*. Manuscript in preparation. Retrieved from <https://ssrn.com/abstract=3536564>
- Wiegmann, A., Horvath, J. & Meyer, K. (2020). Intuitive expertise and irrelevant options. In T. Lombrozo, J. Knobe, S. Nichols (Eds.), *Oxford studies in experimental philosophy* (Vol. 3). Oxford: Oxford University Press.
- Williamson, T. (2011). Philosophical expertise and the burden of proof. *Metaphilosophy*, 42(3), 215–229.
- Wright, J. C., & Bengson, J. (2009). Asymmetries in judgments of responsibility and intentional action. *Mind & Language*, 24(1), 24-50.