

## **UC Merced**

# **Proceedings of the Annual Meeting of the Cognitive Science Society**

### **Title**

Blame Blocking and Expertise Effects Revisited

### **Permalink**

<https://escholarship.org/uc/item/9wp4647h>

### **Journal**

Proceedings of the Annual Meeting of the Cognitive Science Society, 43(43)

### **ISSN**

1069-7977

### **Authors**

Prochownik, Karolina

Wiegmann, Alex

Horvath, Joachim

### **Publication Date**

2021

Peer reviewed

# Blame Blocking and Expertise Effects Revisited

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

Faculty of Law, 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

This paper examines whether advanced law students are resistant to the blame blocking effect—a tendency to assign higher punishments for failed attempts than failed attempts with independent causal chains leading to intended harm (Cushman, 2008). This effect goes against the criminal law principle that intentionally acting towards committing a crime, and not accidental outcomes, determine liability for attempts. To further investigate whether advanced students of law are judging blame blocking scenarios correctly (in line with their legal expertise), in two experiments, we compared their punishment responses with four populations: beginning law students, advanced philosophy students, advanced natural science students, and laypeople with no academic background. We did not observe the blame blocking effect in either of the four student populations, and it was only partially present in the lay population. We discuss the implications of these findings for research on legal expertise and the blame blocking effect in general.

**Keywords:** blame blocking; punishment; expertise defense; experimental jurisprudence

## Introduction

An important area of research within experimental philosophy is the “expertise defense”—an empirical claim that philosophers’ intuitions regarding philosophical matters are less susceptible to irrelevant factors (e.g., order effects, personality effects, framing effects) than lay people’s intuitions. This defense has been generally shown not to hold for professional philosophers (e.g., Feltz & Cokely, 2009; Schwitzgebel & Cushman, 2012; Horvath & Wiegmann, 2016; Wiegmann, Horvath & Meyer, 2020; Horvath & Wiegmann, in press). However, it is an open question whether other types of expertise may impact people’s judgments and make them more resistant to irrelevant factors. This question has recently caught some attention in experimental jurisprudence (e.g., Kneer & Bourgeois-Gironde, 2017; Donelson & Hannikainen, 2020; Prochownik, Krebs, Wiegmann, & Horvath, 2020; Tobia, in preparation), which largely investigates lay people’s (or legal expert’s) concepts and their relationship to the corresponding (technical) concepts of the law and jurisprudence (e.g., Tobia,

2018; Sommers, 2020; Knobe & Shapiro, 2021; Tobia, forthcoming). The expertise defense in this context could take the form of an argument that legal professionals’ concepts reflect (technical) concepts of the law and jurisprudence, and it would thus predict that lawyers’ judgments about legally relevant cases are more reliable than lay people’s judgments (see also Prochownik et al., 2020). This claim is therefore an important question for empirical research.

Testing whether legal experts’ judgments are resistant to legally irrelevant factors seems especially relevant in cases where ordinary judgments depart from legal concepts and doctrines. This is the case for at least two recently studied effects—the *severity effect* (Kneer & Bourgeois-Gironde, 2017) and the *blame blocking effect* (Cushman, 2008). The severity effect occurs when people assign higher levels of intentionality for severely bad versus moderately bad outcomes. This goes against the legal concept of intentionality or *mens rea*, which by default depends on the mental states of the agent, and not on the level of badness of actual outcomes. Although initial research suggests that legal experts’ ascriptions of intentionality are sensitive to severity of outcome (Kneer & Bourgeois-Gironde, 2017), this has recently been questioned by other researchers (Prochownik et al., 2020; Tobia, in preparation).

The blame blocking effect occurs when people assign higher blame and punishment for cases of failed attempts that do not lead to harm than for cases of failed attempts with independently occurring harms (Cushman, 2008). As such, this effect is incompatible with the criminal law standard that assessing liability for attempts depends on whether an agent, through his actions, intentionally aimed at committing a crime (therefore, it does not rely on whether the harm occurred independently of the agent’s intentional actions)<sup>1</sup>. Recent research suggests that people with legal education are not susceptible to the blame blocking effect, especially when they have sufficient exposure to legal training (Prochownik & Unterhuber, 2018).

Although this line of research seems relevant for the expertise defense as applied to people with legal training, it could be argued that the mere fact that people in this group are not susceptible to certain effects to the same extent as lay

<sup>1</sup> For instance, see Fletcher (1998, pp. 171-187) for a comparative legal perspective on attempts; see Bohlander (2009, pp. 137-152) for the German criminal law doctrine of attempts.

people is not sufficient evidence for the special status of legal expertise. In particular, it could be that higher (academic) education or capability for reflective reasoning in general is what makes people's judgments less susceptible to irrelevant factors. If this is so, these effects could not be attributed to legal expertise specifically.<sup>2</sup> Therefore, it seems that claims regarding the applicability of the expertise defense to legal professionals should be based on systematic comparative studies involving people with legal and other types of expertise, just as much as lay people.

This paper aims to advance previous research by taking this broader approach to the assessment of the legal expertise defense. In particular, we aim to further explore whether people with legal training are resistant to the blame blocking effect, and whether this can be attributed to their sufficient exposure to legal education. In our experiments, we therefore examine and compare ascriptions of punishment for failed attempts with no harm done and independently caused harms in five populations: beginning law students; advanced law students (3<sup>rd</sup> or higher semester); advanced philosophy students (3<sup>rd</sup> or higher semester); advanced natural science students (3<sup>rd</sup> or higher semester)<sup>3</sup> (Experiment 1); and lay people (with no university education) (Experiment 2). In particular, we investigate whether advanced students of law are less prone to the blame blocking effect in comparison to other examined populations.

## Experiment 1

In the first experiment we investigated whether the blame blocking effect occurs across four groups in Germany: beginning law students (1st semester), advanced law students (3rd or higher semester), advanced philosophy students (3rd or higher semester), and advanced natural science students (3rd or higher semester). Based on the standard curriculum of legal studies' programs in Germany, we assumed that

---

<sup>2</sup> See Tobia (in preparation) for a similar but slightly different approach. He examines the "severity effect" in lay people, people with legal training, and people with other expert training, but assumes that similarity in judgments of both expert groups is compatible with a certain account of legal expertise—that certain kinds of people (e.g., intelligent, reflective) choose to study law.

<sup>3</sup> This includes, for instance, biology, chemistry, physics, astronomy, and geoscience students.

<sup>4</sup> In particular, for the purposes of the current study, we treated 3<sup>rd</sup> or higher-semester students of law as "advanced", because, typically, they will already have completed at least one or more courses in criminal law at this point. This should have sufficiently exposed them to the basic principles of legal responsibility for attempted crimes that are standardly covered during such classes at German universities (and should be attested by passed exams). In line with this assumption, only 2 people (less than 1%) in our advanced law students sample ( $n = 252$ ) indicated that they had not completed any lecture or seminar in penal law (16% indicated that they had completed one lecture, 49% two to three lectures, and 35% more than three lectures).

<sup>5</sup> In total, we collected 1260 valid responses. Since we preregistered a sample size of 1008, we only included the 252 first recorded responses in each group. Data analyses for the complete

students in the third and higher semesters should have already completed a number of lectures in criminal law. This should have given them sufficient training in legal principles that govern liability assessment for failed attempts.<sup>4</sup> Therefore, we predicted that they will not be susceptible to the blame blocking effect. In contrast, beginning law students who have only started their exposure to criminal law should still be susceptible to the blame blocking effect. Similarly, we expected that advanced students of natural science and philosophy should manifest this effect because their professional training does not cover the principles of criminal law.

## Methods

1008 participants (252 in each of the four groups) were randomly assigned to a *harm* and *no-harm condition*.<sup>5</sup> In each condition, they read two stories of attempted murders (adapted from Cushman, 2008; Prochownik, 2017; Prochownik & Unterhuber, 2018)<sup>6</sup>. In the no-harm variant of the *allergy vignette*, two runners compete in a championship race. One of the runners wants to kill his rival by sprinkling hazelnuts on his salad, thinking he is deadly allergic to them. However, this is incorrect and the plan fails. In the harm variant of this story, the runner performs the same action, but his supposed victim is actually allergic to peanuts and dies because they were tragically included in his salad. In the no-harm *construction scenario*, two engineers compete for a position. One wants to kill the other by executing an electric shock, but the electrical wiring system is faulty and the plan fails. In the harm variant, the story is the same, except that a heavy beam caused by the wind to fall actually kills the supposed victim (unlike in the no-harm version). Below each story, participants had to specify the amount of punishment that, in their opinion, the agent deserved (from 0 to 25 years in prison).<sup>7</sup>

set of responses is available under "Main and Supplementary Data Analyses" on the OSF webpage of this project.

<sup>6</sup> We adjusted the previous vignettes to make sure that the harm and no-harm versions are as similar as possible in terms of length and information provided. To the no-harm construction vignette in particular, we added the information that the wind caused a beam to fall which, however, narrowly missed the engineer (while in Prochownik & Unterhuber, 2018, only the harm-vignette included this information, and indicated that the engineer was killed by the beam). In the allergy vignette we made the (potentially) food-allergy triggering ingredients as similar as possible (i.e., "peanuts" and "hazelnuts"), while these ingredients belonged to different food categories in previous vignettes (e.g., "poppy seeds" and "hazelnuts" in Cushman, 2008). The full text of the vignettes (German original and English translation) is available in the OSF webpage of the project under "Supplementary Materials".

<sup>7</sup> The scale was designed to include possible liability thresholds for attempted murders (like those examined in our study) according to the German criminal code (from 3 to 15 years prison time). Also, our response format required that participants type in their punishment responses themselves. In this respect, it differs from the scales used in previous research (e.g., Cushman, 2008, used a scale anchored with particular prison sentences: "None", "6 months", "1

Sample sizes, exclusion criteria<sup>8</sup>, study questions, main hypotheses and analyses were preregistered at: [osf.io/hk7p8](https://osf.io/hk7p8). Data and study vignettes are also available online on the OSF webpage of this project: [osf.io/wfpzc](https://osf.io/wfpzc).

## Results

We first conducted an overall mixed ANOVA with Story (allergy, construction) as within-subject factor, Harm (harm vs. no-harm) and Population (four student groups) as between-subjects factors ( $N=1008$ ). We found no significant main effect of Harm,  $F(1, 1000) = 1.103, p = .294, \eta^2 = .001$ , and no significant interaction between Harm and Student Population,  $F(3, 1000) = .553, p = .646, \eta^2 = .002$ . We observed significant main effects of Population,  $F(3, 1000) = 17.933, p < .001, \eta^2 = .051$ , and Story,  $F(1, 1000) = 93.823, p < .001, \eta^2 = .086$ . Finally, the interaction between Story and Harm showed a small significant effect,  $F(1, 1000) = 17.041, p < .001, \eta^2 = .017$ .

We now revisit the preregistered hypotheses and analyses. First, we hypothesized there would be no blame blocking effect for advanced law students ( $n = 252$ ). This hypothesis was supported by a mixed ANOVA with Story as a within-subject factor and Harm as between-subjects factor: there was no main effect of Harm in this population,  $F(1, 250) = 0.085,$

$p = .771, \eta^2 < .001$ . Secondly, we predicted that the blame blocking effect would be present in beginning law students ( $n = 252$ ). This was not supported by the same mixed ANOVA conducted for this sample,  $F(1, 250) = 1.190, p = .276, \eta^2 = .005$ . Contrary to our predictions, we also did not find punishment ratings to be higher in the no-harm than in the harm condition among advanced students of philosophy ( $n = 252$ ),  $F(1, 250) = 0.96, p = .757, \eta^2 < .001$ , and natural science ( $n = 252$ ),  $F(1, 250) = 1.309, p = .254, \eta^2 = .005$ .

Our further hypotheses concerned specific comparisons between the four groups. We predicted that the blame blocking effect would be lower in advanced law students (i.e., expert subjects) than in the other three non-expert populations. We ran a set of three ANOVAs with Story as within-subject factor, Harm and Population as between-subjects factors (comparing respective populations only).<sup>9</sup> Contrary to our predictions, we did not find any significant interaction between Student Population and Harm in any of the conducted analyses: comparing advanced law and philosophy students,  $F(1, 500) = 0.177, p = .674, \eta^2 < .001$ ; comparing advanced law and natural science students,  $F(1, 500) = 0.603, p = .438, \eta^2 = .001$ ; and comparing advanced and beginning law students,  $F(1, 500) = 0.286, p = .593, \eta^2 = .001$ .<sup>10</sup> See Figure 1 for the summary of the main findings.

---

year”, “2 years”, “4 years”, “8 years”, “16 years”, “32 years”, “Life”; and a 7-point scale anchored with abstract terms: “None at all”, “Some”, and “Very much” punishment).

<sup>8</sup> Noteworthy, we preregistered the exclusion of first-semester law students who had already completed a lecture in penal law. However, the great majority of participants in this group indicated they had already completed at least one lecture in penal law (i.e., only 5% of 252 participants who satisfied all remaining inclusion criteria answered that they had not; 57% indicated that they had completed one course, 14% two to three courses, and 24% more than three courses). Since first-semester students obviously have not yet completed their first semester (nor any courses), it seems likely that the students understood this question as asking whether they were currently *enrolled* in a lecture in penal law, and not whether they had already *completed* it (i.e., with an exam). Due to this likely misunderstanding, we decided not to exclude first semester-law students based on this criterion. Deviations from preregistered exclusion criteria are also reported online under “Main and Supplementary Data Analyses”.

<sup>9</sup> Note that numerous other comparisons between groups were possible here, but we will focus on the preregistered analyses

(concerning potential differences in blame blocking in advanced law students and three other populations).

<sup>10</sup> Additionally, for the sake of symmetry with the analyses reported for experiment 2, we conducted a set of eight exploratory chi-square tests comparing the frequencies of “no punishment” vs. “any punishment” responses in the harm and no-harm condition (Cushman, 2008). We conducted these comparisons separately for allergy and construction scenarios for each population. For this reason, we recoded all punishment responses dichotomously: all “0” years of punishment responses as “no punishment”, all responses from “1” to “25” years of punishment as “any punishment”. We did not find any significant differences in how frequently participants assigned “no punishment” vs. “any punishment” across harm and no-harm conditions in neither of the four groups and two scenarios examined. Only for the advanced philosophy students, we observed a tendency in the direction of the blame blocking effect in the construction scenario: 16% of participants assigned “no punishment” in the harm condition, while 7.9% did so in the no-harm condition,  $X^2(1, n = 252) = 3.966, p = 0.046$  (not significant after adjusting for multiple comparisons).

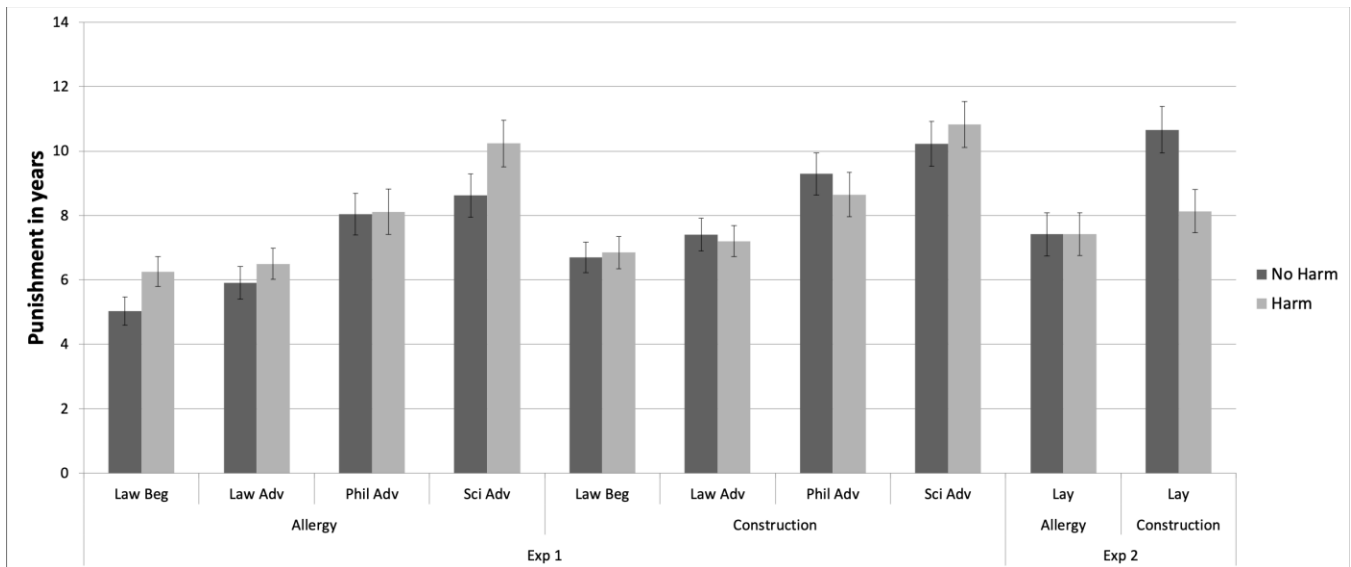


Figure 1: Results of Experiments 1 and 2. Punishment in years as a function of story, subject, and harm. Error bars represent standard error of means.

## Discussion

In the first experiment, we investigated whether the blame blocking effect occurs in four student populations in Germany, and whether legal expertise has an attenuating impact on it. In line with our predictions, we did not find this effect among advanced law students. Contrary to our other predictions, we also did not find the effect in the three other populations with minimal or zero legal expertise (beginning law students, advanced philosophy students, and advanced natural science students).

Let us consider two potential explanations of these findings. First, the four examined groups could share certain common feature(s) that reduce their sensitivity to irrelevant factors as they were found to affect punishment ascriptions in blame blocking scenarios in previous research. The most obvious feature would be that our participants were all university students. Being a university student (and educational achievements in general) are typically associated with a number of capacities, such as critical reasoning skills or higher intelligence (e.g., Deary, Strand, Smith, & Fernandes, 2007), which may attenuate people's susceptibility to the blame blocking effect (and, possibly, other effects). Regardless which of these capacities may be most relevant here, this explanation seems to undermine claims about the special status of legal expertise with respect to such effects (but see Tobia, in preparation). Namely, it could be that it is being a university student in general, and not legal education specifically, that makes people's judgments "immune" to the irrelevant factors in blame blocking scenarios.<sup>11</sup> Second, our failure to replicate the

blame blocking effect raises the possibility that the effect is not very robust or appears only under certain circumstances (e.g., not for our German populations).

In our second experiment, we mainly investigated the first possibility. If it is being a university student that is responsible for the absence of a blame blocking effect in the four student populations we examined, then people who are not university students or who have not had any university education should manifest the blame blocking effect. This would also support the hypothesis that a general selection effect related to entering a university education, and not legal education specifically, attenuates the blame blocking effect.

## Experiment 2

In the second study, we investigated whether the blame blocking effect occurs in a German population of lay people with no university education whatsoever.

## Methods

We recruited 252 lay participants who indicated that they do not have a university degree, that they never studied at a university, and that they are not currently studying at a university. Study design, vignettes, and main questions were the same as in our first experiment. Sample size, exclusion criteria, and main analyses were preregistered at: [osf.io/ues95](https://osf.io/ues95).

Full text of the study vignettes and questions as well as the data are also available online at: [osf.io/wfpzc](https://osf.io/wfpzc).

<sup>11</sup> Whether this manifests itself specifically in punishment judgments (and in the context of blame blocking in particular), or a broader set of judgments, is a question for further research.

## Results

All reported analyses were preregistered. We first conducted a mixed ANOVA with Story as a within-subject factor and Harm as a between-subjects factor ( $n = 252$ ). Over all vignettes, we found no significant main effect of Harm in the lay population,  $F(1, 250) = 1.907, p = .169, \eta^2 = .008$ . Secondly, we conducted two chi-square tests for two stories comparing the frequency of “no punishment” vs. “any punishment” responses in the harm and no-harm condition (see Cushman, 2008). For the sake of these tests, we recoded all no punishment responses to “0” and all punishment responses (from 1 to 25) to “1”. In the allergy story, there was no difference in the frequencies of “no punishment” vs “any punishment” across two conditions: 20.5% assigned “no punishment” in the harm condition comparing to 16.9% in the no-harm condition,  $\chi^2(1, n = 252) = .528, p = 0.467$ . In the construction story, we found a significant blame blocking effect: 18.9% assigned “no punishment” in the harm condition comparing to 6.9% in the no-harm condition,  $\chi^2(1, n = 252) = 8.079, p = .004$ .<sup>12</sup> See Figure 1 for the summary of the main results.

Finally, we conducted an exploratory analysis (mentioned in the preregistration) on the combined dataset including four student populations from Experiment 1 ( $n = 1008$ ) and the lay population from Experiment 2 ( $n = 252$ ). We conducted a mixed ANOVA with Story as a within-subject factor and Harm and Population as between-subjects factors ( $N = 1260$ ). Also for this model, we observed no significant main effect of Harm,  $F(1, 1250) = 0.057, p = .811, \eta^2 < .001$ : participants assigned similar punishments in the harm ( $M = 8.019$ ) and no-harm condition ( $M = 7.930$ ). There was a significant main effect of Population,  $F(4, 1250) = 13.120, p < .001, \eta^2 = .04010$ : beginning law students ( $M = 6.208$ ) and advanced law students ( $M = 6.755$ ) punished the least, lay people ( $M = 8.409$ ) and advanced philosophy students ( $M = 8.523$ ) were at a medium level, and advanced natural science students punished the most ( $M = 9.979$ ). We also observed a significant main effect of Story,  $F(1, 1250) = 132.456, p < .001, \eta^2 = .09611$ : punishments were generally higher in the construction ( $M = 8.61$ ) than in the allergy ( $M = 7.36$ ) vignette. Moreover, the following interactions were significant: Story and Harm,  $F(1, 1250) = 32.493, p < .001, \eta^2 = .025$ <sup>13</sup>; Story and Population,  $F(4, 1250) = 3.061, p = .016, \eta^2 = .010$ ; Story, Population and Harm,  $F(4, 1250) = 2.412, p = .047, \eta^2 = .008$ .

---

<sup>12</sup> This finding was further confirmed by exploratory ANOVAs (mentioned in the preregistration) conducted for separate stories on a full range of responses. In the analysis for the construction story, we found a significant blame blocking effect: Lay participants assigned higher punishments in the no-harm ( $M = 10.66, SD = 8.164$ ) than in the harm condition ( $M = 8.13, SD = 7.436$ ),  $F(1, 250) = 6.590, p = .011, \eta^2 = .026$ .

<sup>13</sup> Overall, participants tended to assign punishments slightly in the direction of blame blocking effect in the construction vignette. This response pattern was reversed in the allergy vignette.

<sup>14</sup> It should be noted that this earlier finding was also not robust, because the blame blocking effect occurred only in one, but not in

## Discussion

Our second experiment examined whether the blame blocking effect is present among German lay people. Overall, lay people did not assign higher punishments in the no-harm versus harm condition. The blame blocking effect was *partially* observed for one of the two vignettes. Therefore, the results of our second study do not support the hypothesis that general capacities due to a selection effect related to entering a university education attenuates the blame blocking effect.

Finally, the blame blocking effect was not present for the analysis that directly compared the five populations from our two experiments, which overall suggests at best weak evidence for this effect in our sample. Below we further investigate potential reasons behind our replication failure, and we discuss them in light of relevant previous findings.

### General Discussion

We conducted two experiments which investigated whether the blame blocking effect was present in five populations in Germany: beginning and advanced law students, advanced philosophy and natural science students, and lay people with no university education. We found very limited evidence for this effect in our sample: Only non-academic lay participants in only one of the two study vignettes assigned higher punishments in the no-harm than in the harm condition.

Our findings have several implications for past and future research. First, we did not replicate the previous finding that the blame blocking effect occurs for beginning law students but not for advanced law students (Prochownik & Unterhuber, 2018). In fact, we did not observe any blame blocking effect in law students at all (and neither for the two other student populations examined).<sup>14</sup> Also, contrary to other studies with lay people (in the US and Poland: Cushman, 2008; Prochownik & Cushman, 2019; Prochownik & Unterhuber, 2018), we found only partial and rather weak evidence for the occurrence of this effect in our non-academic lay sample (in Germany). Because we did not replicate the blame blocking effect in the first place, the issue of (legal) expertise strictly speaking does not even arise.<sup>15</sup> Moreover, our failure to replicate the blame blocking effect makes it a live hypothesis that it is not very robust, or that it only occurs under specific circumstances. Investigating these possibilities goes beyond the scope of this paper, but we will discuss potential factors that could account for our replication failure

the other sample of beginning law students (Prochownik & Unterhuber, 2018).

<sup>15</sup> However, our findings do suggest some legal expertise effects for punishment ascriptions: law students assigned the lowest punishments in comparison to three other populations. This might be due the fact that according to German criminal law, liability for cases of failed attempts like those described in our vignettes would be relatively low (between 3 and 15 years prison time). Moreover, since beginning and advanced students ascribed punishment similarly, our findings may suggest *fast* effects of legal training on this category of judgments.

below, by taking into account several limitations of our current study.

We can think of a few reasons why we did not manage to replicate the blame blocking effect, although it was found by previous researchers (e.g., Cushman, 2008; Prochownik & Cushman, 2019). In particular, it may be that we failed to reproduce particular *circumstances* that were present in previous research, and under which this effect occurs. First, the most obvious possibility is that we departed from the previous methods. The two most significant departures were that: a) we adapted our harm and no-harm vignettes to be as similar as possible in terms of length and information provided; b) we used a new punishment question format in which participants had to specify the amount of punishment themselves instead of choosing it on a scale with explicitly given options.

Regarding the first possibility (a), it may be that by adding to the no-harm vignettes information about causal chains similar to those described in the harm vignettes, but not leading to harm, we made the two conditions too similar to each other, resulting in our study participants not differentiating punishment across them. This interpretation of our results is in line with an alternative account which points out that, in previous research, harm vignettes were more complex and richer in information than no-harm vignettes, and that this might have distracted participants in the harm vignettes from evaluating the agent, leading to the blame blocking effect (Prochownik & Cushman, 2019). Therefore, it may be that after we made our harm and no-harm vignettes equally “distracting”, we did not find a significant blame blocking effect.<sup>16</sup> However, it should be noted that it is actually good methodological practice to make experimental vignettes as similar as possible, apart from the to-be-tested difference, in order to avoid “noise” and possible confounds. Moreover, it also speaks against this “distraction account” that a recent preregistered study with a wider set of scenarios found a significant blame blocking effect even after making the harm and no-harm conditions more symmetrical (Prochownik & Cushman, in preparation). Therefore, more research is needed regarding the particular features of the vignettes under which the blame blocking effect does and does not occur. In particular, both current and previous studies consistently report that the blame blocking effect is stronger for some of the tested vignettes than others (Cushman & Prochownik, 2019; Prochownik, 2017; Prochownik & Unterhuber, 2018)

The second possibility is that we failed to replicate the blame blocking effect because we changed the format of the main question for punishment responses (b). Instead of giving participants a scale with pre-given punishment options (Cushman, 2008; Prochownik & Cushman, 2019), or a slider scale with a range of choices to select from (Prochownik & Unterhuber, 2018), we asked them to type in the amount of punishment the agent deserved in an open text-entry by themselves (we did indicate, though,

this should be within the range of 0 and 25 years in prison). It is possible that, in this way, we encouraged participants to reflect more on the sentences the agents deserved. For instance, sentencing somebody to a certain amount of imprisonment *oneself* may be associated with a greater sense of responsibility than clicking on one of the pre-given punishment choices. This “reflective thinking account” may be compatible with the previous findings that support the blame blocking effect. In particular, it may be that this phenomenon is limited to highly intuitive moral judgments, and that it disappears even after a minimal amount of reflection has been exercised (in contrast to related findings about intuitive expertise where reflection did not make a difference, such as Schwitzgebel & Cushman, 2015). Therefore, if the way we asked the punishment question sufficiently stimulated participants’ reflective reasoning, this could explain our failure to replicate this effect.

Future research should examine whether the blame blocking effect only occurs under conditions of intuitive vs. reflective moral reasoning. One possible direction of research could involve replicating our current study, but rephrasing the punishment question in a way that triggers purely intuitive responses. For instance, participants could be asked to choose between only two options: punishment vs. no punishment—which seems to be the most “intuitive” response format available in the literature (Prochownik & Cushman, in preparation), and which also comes closest to the dichotomous data analysis in the original study reporting the blame blocking effect (Cushman, 2008, Experiment 4).

Finally, there are other options that could explain why we did not replicate the blame blocking effect in addition to the two main ones discussed above. In particular, it is possible that this effect is specific to certain cultures and languages. Our replication was the first study on the blame blocking effect in Germany and in the German language. Perhaps we failed to replicate it because it does not generalize to German culture and language. We can think of no clear reason why this should be the case, but more cross-cultural research would be needed to clearly rule out this possibility.

Although many questions remain unanswered, we considered three possibilities that may shed light on our failure to replicate the blame blocking effect—that it may be limited to a certain type of *vignette*, to a certain type of *cognition*, or a certain type of *culture or language*. Future research should investigate them more thoroughly and try to specify the exact conditions under which blame blocking does or does not occur.

<sup>16</sup> Moreover, since most substantial changes considered our “construction” vignette, and this was the vignette driving the blame

blocking effect in previous findings (Prochownik, 2017), this could explain why we observed only a limited effect for this vignette.

## Acknowledgments

We thank four anonymous reviewers for their very helpful comments and suggestions.

Joachim Horvath and Alex Wiegmann's work on this paper was supported by an Emmy Noether grant of the German Research Foundation (DFG), project number 391304769. Karolina Prochownik's work was supported by the DFG-CAPES research grant "Experimental Legal Philosophy: The Concept of Law Revisited", project number 434400506.

## References

- Bohlander, M. (2009). *Principles of German criminal law*. Oxford, Portland, OR: Hart Publishing.
- Cushman, F. (2008). Crime and punishment: distinguishing the roles of causal and intentional analyses in moral judgment. *Cognition*, 108(2), 353-380.
- Deary, I. J., Strand, S., Smith, P., & Fernandes, C. (2007). Intelligence and educational achievement. *Intelligence*, 35(1), 13-21.
- 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.
- Fletcher, G. P. (1998). *Basic concepts of criminal law*. Oxford: Oxford University Press.
- Horvath, J., & Wiegmann, A. (2016). Intuitive expertise and intuitions about knowledge. *Philosophical Studies*, 173(10), 2701-2726.
- Horvath, J., & Wiegmann, A. (in press). Intuitive expertise in moral judgements. *Australasian Journal of Philosophy*.
- Kneer, M., & Bourgeois-Gironde, S. (2017). Mens rea ascription, expertise and outcome effects. Professional judges surveyed. *Cognition*, 169, 139-146.
- Knobe, J., & Shapiro, S. J. (2021). Proximate cause explained: An essay in experimental jurisprudence. 88 *University of Chicago Law Review*, 165-236.
- Prochownik, K. (2017). Do people with a legal background dually process? The role of causation, intentionality and pragmatic linguistic considerations in judgments of criminal responsibility. In J. Stelmach, B. Brożek & Ł. Kurek (Eds.), *The province of jurisprudence naturalized*. Warsaw: Wolters Kluwer.
- Prochownik, K., & Unterhuber, M. (2018). Does the blame blocking effect for assignments of punishment generalize to legal experts? *Proceedings of the 40<sup>th</sup> annual conference of cognitive science society* (pp. 2285-2290). Austin, TX: Cognitive Science Society.
- Prochownik, K., & Cushman, F. A. (2019). Outcomes speak louder than actions? Testing a challenge to the two-process model of moral judgment. *Proceedings of the 41<sup>st</sup> annual conference of the cognitive science society* (pp. 2606-2612). Montreal, QB: Cognitive Science Society.
- Prochownik, K., & Cushman, F. (2020). *Testing challenges to blame blocking*. Manuscript in preparation.
- Prochownik, K., Krebs, M., Wiegmann, A., & Horvath, J. (2020). Not as bad as painted? Legal expertise, intentionality ascription, and outcome effects revisited. In *Proceedings of the 42<sup>nd</sup> annual conference of the cognitive science society* (pp. 1930-1936). Cognitive Science Society.
- 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.
- Schwitzgebel, E., & Cushman, F. (2015). Philosophers' biased judgments persist despite training, expertise and reflection. *Cognition*, 141, 127-137.
- Sommers, R. (2020). Commonsense consent. *Yale Law Journal*, 129.
- 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>
- Tobia, K. P. (forthcoming). Experimental jurisprudence. 89 *University of Chicago Law Review*. Retrieved from: <https://ssrn.com/abstract=3680107>
- 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.