



Neglecting Negligence? The Role of Outcome Bias and Negligence on Moral Luck

Niklas Fritsche, Claudia Gerhardt

Hochschule Fresenius, University of Applied Sciences

Datum der Einreichung: 14.07.2023 | Datum der Annahme: 16.01.2024

ABSTRACT

This study examines the influences of outcome bias and negligence on moral judgments within the framework of moral luck. Through a conceptual replication using a novel vignette and a between-subjects design, this research explores how different outcomes and levels of perceived negligence affect judgments of punishment, blame, and wrongness. The study engaged 363 participants, manipulating outcome and negligence variables to assess their impact on moral evaluations.

The findings indicate that negative outcomes intensify moral judgments, underscoring the robust influence of outcome bias. While negligence did not significantly moderate the severity of punishment judgments, it substantially mediated the effects of outcomes on perceptions of blame and wrongness. Notably, and contrary to prior findings, this research reveals significant differences in moral evaluations between neutral and negligent conditions, highlighting the contextual variability of negligence perception.

These results suggest that negligence significantly shapes moral evaluations and that outcome biases are a pervasive influence. This study contributes to the understanding of moral luck and informs practical applications in legal and professional settings, advocating for more nuanced considerations of negligence in moral judgments.

Schlüsselbegriffe: outcome bias, negligence, moral luck, moral judgements

1 Neglecting Negligence? The Role of Outcome Bias and Negligence on Moral Luck

Moral luck is a complex phenomenon that emerges when individuals are subject to moral judgment despite certain aspects of their evaluation being dependent on factors beyond their control. This concept challenges the traditional Kantian understanding of moral judgments by highlighting the inherent tension between intuition regarding moral responsibility and luck's role in shaping moral outcomes. The concept of moral luck suggests that moral judgments can be influenced not only by the intentions and actions of individuals but also by external factors that determine the outcomes of those actions, such as luck (Levy, 2016; Nagel, 1979; Williams, 1981). Understanding this phenomenon is crucial to our daily lives as it challenges our intuitive belief that moral judgments should be solely based on actions and intentions that one can control (Kant, 1785), highlighting the complex interplay between luck and morality.

Initially, morality and luck may seem contradictory (Statman, 1993), but both philosophy and psychology provide evidence of moral luck, in which random factors affect

moral judgment (Gino et al., 2009; Martin & Cushman, 2016; Nobes & Martin, 2022). To illustrate, imagine a driver who agrees to drive through a large leaf pile at the request of the younger brother. Unbeknownst to them, there are hidden sticks in the pile. If the car passes through the pile unharmed, the driver is hardly judged. However, if the hidden objects turn out to be two children who are killed by the car, the driver is judged disproportionately harshly. The morality of the driver's action is judged based on the unforeseeable outcome of the situation, which is beyond their control (Martin & Cushman, 2016).

However, judging others based on factors beyond their control while believing oneself is only responsible for controllable factors, conflicts with the control principle (Kant, 1785), which essentially states that one ought not to judge the morality of an agent if the outcome of the action is influenced by factors outside the agent's control. The Kantian idea of the moral judgment process is still propagated today (Lench et al., 2015; Young et al., 2010) and contradicts findings concerning moral luck, prompting researchers and philosophers to investigate its influence on moral evaluation.

Given the lack of consensus on the definition of luck, this study adopts Nagel's (1979) framework to distinguish between resultant, situational, constitutive, and causal luck. Aligned with the methodology of previous research in this domain (Kneer & Machery, 2019; Levy, 2016; Nobes & Martin, 2022), this study uses the definition of resultant luck, which affects the focal outcome of a given situation. Hence, resultant luck is used interchangeably with "moral luck."

During the research on this erratic phenomenon, one cognitive distortion, outcome bias, materialised and was continuously used to explain the occurrence of moral luck. In their seminal paper, Baron and Hershey (1988) coined the term "outcome bias" to describe the phenomenon where the assessment of an action and the actor is distorted by the perceived success or failure of the action, even when all other information is the same. This bias is believed to occur implicitly and outside of conscious cognition, creating a gap between normative theories of decision-making (Brown et al., 1974) and real-world practice. Outcome bias is a widely observed and replicable phenomenon that occurs across various domains (Broude & Levy, 2019; Gruppen et al., 1994), indicating external validity (Cole et al., 2012; Lefgren et al., 2014). Explanation approaches for outcome bias include the overgeneralization heuristic (Baron & Hershey, 1988), attributing clairvoyant ability to the agent (Gino et al., 2009), and underestimating the causal role of external factors (Savani & King, 2015), to name a few.

Additionally, outcome bias is frequently examined as an influencing factor in moral luck in economic and medical settings (Gino et al., 2009; and most recently Nobes & Martin, 2022). Here, the outcome of the situation determines the valence of the evaluations (e.g. blame, punishment, and wrongness). Martin and Cushman (2016) developed a nuanced framework to explain how such judgments arise. In their two-process model, Martin and Cushman (2016) proposed two distinct routes that affect moral judgments differently. Regarding perceived wrongness and permissibility, the sole indicator for condemning the agent is the good or bad intention attributed to them, whereas punishment and blame are influenced by the combination of this route and the agent's outcome (good or bad) (Martin & Cushman, 2016). Despite evidence of the validity of the model (Martin & Cushman, 2016; Margoni et al., 2019), the results of a more recent study point to the fact that the model is incomplete and lacks at least one other factor (Nobes & Martin, 2022).

In many legal systems, judges depend on the assessment of negligence when inferring the punishment and condemnation of victims. In German law, negligence (Fahrlässigkeit) refers to the failure to exercise reasonable care or diligence, resulting in harm or damage to another person or property (§ 276 Abs. 2 Verantwortlichkeit des Schuldners v. 26.11.2001, BGBL. I, S. 3138). Several studies have shown the importance of negligence in moral judgments (Kneer & Machery, 2019; Laurent et al., 2016; Rizzo et al., 2019) and have indicated that even young children (to some extent) base their decisions on it (Mulvey et al., 2020; Nobes et al., 2017). Thus, it is reasonable to assume that negligence influences moral luck. Fol-

lowing the same approach, Nobes and Martin (2022) investigated the effects of negligence on moral luck.

They hypothesised that the impact of the outcome of unintentional accidents on moral luck is smaller than previously assumed and that a larger part of the occurrence of moral luck can be explained by the negligence of the agent. To demonstrate this, the authors conducted two experiments in which they slightly altered the vignette used by Martin and Cushman (2016) to include negligence as an additional independent variable. (The vignette describes a more detailed version of the girl driving through a leaf pile).

Their results showed an outcome effect concerning the agent's moral judgment, meaning that a negative outcome led to more negative judgment (punishment, wrongness, and blame). Interestingly, this influence was moderated (punishment) or mediated (wrongness and blame) by negligence and perceived negligence, respectively. However, an even more significant finding of the study was the non-existent distinction between neutral and negligent agents regarding perceived negligence. In other words, in a situation without cues about the negligence of an agent, the actor is considered as negligent as a person who acts negligently.

To put it more clearly, if someone (judge, line manager, etc.) were to evaluate an agent's negligence and the context would not provide sufficient or any information about the level of negligence, the evaluator would perceive the agent as equally negligent compared to a situation in which he knows that the agent acted negligently which would subsequently shape his assessment. One can easily see how this could have remarkable repercussions in daily settings in the courts of law, business, or interpersonal situations. Moreover, this finding has important implications for (prior) research investigating moral luck, as well as practical settings.

First, it may indicate that, in earlier studies, the outcome effects observed in situations that did not reveal any information about the agent's negligence may have occurred because of the intuitive attribution of negligence by the participants, which, in turn, influenced their judgment. Second, this could mean that in real-life settings, the search for clues of negligence is substantial in one's moral evaluation. This could imply that the absence of these cues could lead to an unfair and disproportionately harsh judgment of the agent.

This concerns social interactions and possible damage for one's social reputation. However, even more worrying could be the grave detrimental effects on a person's life when judged in high-stakes situations, for example, in the court of law or in a corporate environment. Here, insufficient evidence about an agent's negligence may decide between substantial imprisonment and probation, or between being laid off and receiving a warning.

These myriad situations highlight the importance of gaining a deeper understanding of the underlying theory that explains moral luck. In this context, this study aims to replicate the findings of Nobes and Martin (2022) concep-

tually, as it emerges as a more fitting and potent approach, offering robust theoretical and practical contributions and functions as an important component of the scientific method which is often neglected in the social sciences (Schmidt, 2009). This assertion rests on the inherent intricacies of the subject matter, characterised by the multifaceted nature of social phenomena and the intricate interplay of contextual, and individual variables. Given the dynamic and context-dependent nature of human behaviour, the rigid constraints of exact replication may inadvertently stifle the exploration of underlying theoretical constructs (Crandall & Sherman, 2016).

Conversely, a conceptual replication allows for necessary adaptability, permitting variations in operationalisations, research designs, and participant demographics, aligning with the real-world complexity of social phenomena. Such flexibility is imperative for illuminating the breadth and depth of a theory's applicability and transcending specific experimental conditions in order to identify generalisable principles and theoretical nuances (Lynch et al. 2015). Conceptual replications, therefore, offer a superior platform for discerning the broader practical implications of a theory as they facilitate the identification of boundary conditions, exploration of contextual variations, and revelation of potential moderating or mediating factors, thereby enhancing theoretical understanding (Crandall & Sherman, 2016; Lynch et al., 2015).

Therefore, we decided to use a vignette that was different from the original study. This should serve as a tool to scrutinise the proposed underlying mechanisms that evoke moral luck and detect whether these uphold across multiple contextual settings. More specifically, it inspects the adequacy and robustness of the two-process model as well as the newly identified influence of negligence in a novel context.

To test this, we propose the following hypotheses: The extensive literature that shows that participants judge agents more harshly (punishment, blame, wrongness) when their actions lead to a negative outcome compared to a positive outcome leads us to hypothesise that:

H1 The agent is judged more harshly (punishment, blame, wrongness) when their actions entail a negative outcome compared to a positive outcome.

Regarding the negligence hypothesis, our second hypothesis is as follows.

H2 The outcome effect on the degree of punishment is positively moderated by the perceived negligence of the agent.

This was based on the research of Nobes and Martin (2022), who demonstrated this effect in their study design. They, as well as Kneer and Machery (2019), also found that outcome effect on blame and wrongness judgment were largely explained by a link through the agent's perceived negligence. To test these findings, we hypothesised the following.

H3 The outcome effect on blame and wrongness judgments is mediated by the perceived negligence of the agent.

Results from Kneer and Machery (2019) and Nobes and Martin (2022) showed that the outcome of the situation better explained how strongly people punished the agent compared to how blameworthy and wrong they perceived the agent to be. In other words, the influence of the outcome was stronger on punishment judgments than on blame or wrongness judgments. As punishment judgments intuitively seem to have the most impactful and diverse consequences for an individual's life, we are especially interested in determining if these findings also emerge in this conceptual replication.

Finally, we investigate the weighty findings concerning the nonexistent differences in perceived negligence between an agent that, to an evaluator, obviously acts negligently compared to an agent that acts in a neutral fashion (Nobes & Martin, 2022). That is, that there are no salient cues for the assessor to infer negligent or non-negligent behaviour. Thus, we examine whether this pattern is also evident in our study design and consequently critically analyse the influence of negligence in the realm of moral luck and judgments.

These hypotheses and research questions aim to replicate and probe previous findings in the realm of moral luck and particularly focus on the recent findings of Nobes and Martin (2022) regarding the influence of negligence on moral evaluation and punishment. To replicate these findings conceptually, the research design was adapted, but an explicitly different vignette was used. Further details are provided in the methods section.

2 Method

As this study aims to reproduce and scrutinise the results found by Nobes and Martin (2022), many parallels and methodological procedures will be adopted to ensure optimal comparability while also allowing the possibility of novel findings.

Sample

Similar to the original study, the majority of participants were recruited through subject pool platforms (e.g. SurveyCircle and SurveySwap), psychology student panels, and social media. In total, 398 respondents participated in the study, but due to inconsistent answers (31), failing an attention check (3), or young age (here, it was questionable whether the participant would have sufficient understanding of the constructs at hand (e.g. wrongness), some participants were removed from the pool, leading to an effective sample size of $N = 363$. An attention check was included for quality assurance because of the risk that participants may answer the questions inattentively when recruiting through subject-pool platforms (Curran, 2017). This sample consisted of 218 female (60%) and 144 male (40%) participants, with one person identifying as diverse. The age of the participants ranged from 17 to 79 years ($M = 29.56$, $SD = 12.99$). The size and distribution of sex and age did not significantly differ from Nobes and

Martin's (2022) sample (Study 1: female (70%), age ($M = 28.7$, $SD = 12.1$; Study 2: female (59%), age ($M = 39.1$, $SD = 17.5$)).

An a priori power analysis was performed using GPower 3.1 (Rasch et al., 2014). Based on the effect sizes and power of the original study, the power set at .95 with $\alpha = .05$, determined a minimum sample size of $N = 251$ (that is, at least 42 per group) for the ANOVA to detect a medium effect size of $f = 0.25$ ($\eta_p^2 = .06$) (Cohen, 1992). For the linear regression, a minimum sample size of $N = 107$ was found to identify a medium effect of $f = 0.15$ (Cohen, 1992) with a power set at .95 and $\alpha = .05$.

Materials and Procedure

The participants joined the study via a link, leading them to a Unipark questionnaire. At the beginning of the questionnaire, the participants were informed of the duration and use of the data. Subsequently, their age and sex were determined. Using the random generator of the Unipark program, participants were assigned to one of six groups without their knowledge. A vignette was then presented, which differed slightly among the groups. A new vignette was created to validate the results of Nobes and Martin (2022). This expands the knowledge base of moral luck by examining its effects in a novel situation. It also tests the robustness of the results and can rule out that the results are scenario specific.

The general framework of the vignette presented here describes how an au pair goes out to eat with a child of the host family. A gender-neutral name was chosen for the au pair (Noah) in order to prevent possible gender effects. This seemed to be successful, since in the open-ended survey, both the pronouns "he" and "she" were used. Participants assigned to the positive outcome group experienced the following:

After both had eaten, they went home satisfied.

The following observations were made for the negative outcome group:

After both had eaten their first servings, the child froze, had an allergic reaction, and died in the restaurant.

The conditions for negligence were manipulated as follows. In the non-negligent condition, Noah explicitly asks the waiter if the food contains allergenic ingredients that are dangerous to the child. After the waiter confirms with the kitchen that there are none, the food is ordered. This demonstrates that Noah performed everything in their power to prevent harm to the child. However, in the negligent condition, Noah forgets to ask about allergenic ingredients and orders the same food, although it is stated that the parents have informed them about the allergies of the child. As in the negligent condition of Nobes and Martins (2022), the actor knows about the possible factors that can lead to a negative outcome (there: children who frequently play in the pile of leaves) but forgets about them. This is also in line with the German judiciary's definition of negligence; that is, the expectation that Noah, based on his knowledge of the allergy, could and should

have acted attentively and carefully (see § 276 Abs. 2 Verantwortlichkeit des Schuldners v. 26.11.2001, BGBL. I, S. 3138). This neutral condition was comparable to that of the vignette described by Martin and Cushman (2016). Here, no information is given about negligence, but only about how both go out to eat together and the child either dies or they go home together (see Appendix B for an overview of the vignette).

Afterwards, participants responded by answering how strongly they would punish Noah's behaviour. The measurement was carried out using a 7-point Likert scale (not at all – very strong) and additionally through an open-text field where participants could explicitly state their punishment measures. The other evaluation questions (negligence, guilt, and wrongness) followed and were measured using a 7-point Likert scale (not at all – very much).

The questions were translated into English as literally as possible for maximum comparability. The construct of wrongness was translated here as moral decency (German: moralische Anständigkeit) because the literal translation of wrongness (German: Falschheit) left too much room for misinterpretation. For the sake of conciseness and readability, the term wrongness will be used in the rest of this paper when referring to this item. The entire structure of the questionnaire with the specific questions can be found in the appendix.

Design

Similar to the Nobes and Martin study, a 3×2 design was applied with two independent variables (negligence and outcome) that were manipulated in their levels (non-negligent, neutral [no negligence information], negligent; positive [no harm], negative [child dies]). Consequently, six different experimental groups were identified. Each group was tested for four quantitative dependent variables: blame, deserved punishment, perceived negligence, and wrongness, which were measured on 7-point Likert scales. In addition, one open question aimed to identify the concrete level of punishment that each participant assigned to the agent. This item was included to gain a more accurate understanding of what abstract differentiations on the quantitative scale translate to on a concrete level. Considering that a "5" on the punishment scale may be viewed as a fine for one person and probation for another.

3 Results

The following section presents the results of the statistical analysis of the hypotheses. Data analysis was performed using SPSS software (Version 29.0.0.0 (241)) for Mac OS. Moderation and mediation analyses were conducted using the Process plug-in (Version 4.2) by Andrew Hayes (2022) for SPSS. The significance level for all calculations was set at $\alpha = .05$.

To test H1, which postulated that a negative outcome leads to greater punishment, greater blame, and greater perceived wrongness of the agent, we examined whether there were significant mean differences between groups with positive and negative outcomes with respect to the

dependent variables mentioned (DV). As the groups showed unequal variances with respect to punishment, a one-way ANOVA with Welch's test was performed (Urdu, 2005); the other DVs were analyzed using t-tests. Participants punished the actor less severely in the positive outcome condition ($M = 1.90$, $SD = 1.45$) than in the negative outcome condition ($M = 2.78$, $SD = 1.85$). This difference was statistically significant according to the Welch test, $F(1, 358.59) = 25.85$, $p < .001$, with a medium effect size of $f = 0.26$ ($\eta_p^2 = .06$).

There was also a difference between the positive ($M = 2.78$, $SD = 2.11$) and negative ($M = 3.36$, $SD = 1.99$) groups with respect to the actor's perceived blame. This result was statistically significant ($t(361) = -2.70$, $p = .004$). Additionally, the difference between the groups in their perception of the agent's wrongness was significant, $t(361) = -2.83$, $p = .002$ (positive: $M = 2.45$, $SD = 1.85$; negative: $M = 3.01$, $SD = 1.84$). The effect sizes according to Cohen's d were $d_{\text{blame}} = 0.28$ and $d_{\text{wrongness}} = 0.30$. Therefore, H1 is confirmed.

To identify the interaction effect of negligence and the focal outcome, a moderation analysis was conducted. The overall model of the moderation analysis for predicting the level of punishment through the interaction with the outcome of the action with moderator negligence explained 60.31% of the variance and was significant, $F(3, 359) = 121.45$, $p < .001$. However, a significant moderating effect of the two variables was not found, $\Delta R^2 = .002$, $F(3, 359) = 1.21$, $p = .272$, 95% CI [-0.150, 0.490]. However, the calculation of a two-factor ANOVA with both independent variables of the design and punishment as the dependent variable showed that the level of punishment was primarily influenced by the manipulation of the negligence variable, $F(2, 363) = 211.33$, $p < .001$, $\eta_p^2 = .36$, and less by the outcome of the action, $F(1, 363) = 42.63$, $p < .001$, $\eta_p^2 = .13$. The interaction effect was not significant ($F(2, 363) = 0.99$, $p = .371$, $\eta_p^2 = .006$).

An additional analysis of the qualitative results showed that participants in the non-negligent and neutral groups consistently did not demand punishment when the outcome was positive. The vast majority of participants showed the same pattern of judgment, even when the outcome was negative. The only deviation was that in the non-negligent condition ($n = 72$), 5% demanded a fine, 11% demanded dismissal, and just under 3% demanded imprisonment. In the neutral condition with negative outcomes ($n = 67$), 13% demanded dismissal, 6% each demanded a fine or probation, under 3% demanded a warning or reprimand, 4% demanded community service, and under 3% imprisonment; the rest demanded no punishment. In the negligent condition with a positive outcome ($n = 68$), the majority (54%) suggested serious conversation as punishment and 22% demanded dismissal. The remainder did not demand punishment. In the negligent condition with a negative outcome ($n = 70$), 27% of the participants suggested probation as a punishment, 18% suggested imprisonment, and 17% suggested a fine. The rest varied among dismissal, reprimand, and community service. Put simply, these numbers indicate that an increasing level of negligence leads to a more severe level of punishment, suggesting a possible underlying positive

moderation of punishment through negligence. However, this rather unconventional analysis must be treated with caution and will be discussed in more detail.

Mediation analyses were conducted to test whether the effect of the outcome on blame and wrongness was mediated by the perceived negligence of the agent, as stated in H3. In the blame mediation model, the outcome had a significant total effect ($c = 0.584$, $p < .001$, CI [0.159, 1.010]). However, this effect was accounted for more by its indirect effect through negligence ($ab = 0.851$, 95% CI [0.486, 1.213]) than by its direct effect ($c' = -0.267$, $p = .022$). The model for wrongness showed a total effect of $c = 0.555$, $p < .001$, CI [0.169, 0.914]. After adding the mediator to the model, the outcome effect was wholly accounted for by the indirect effect of perceived negligence ($ab = 0.737$, 95% CI [0.428, 1.069]), with a direct effect of $c' = -0.182$, $p = .105$. The results indicate that the outcome's effect on blame and wrongness is primarily mediated by perceived negligence, as the indirect effects through negligence were significant, whereas the direct effects were not consistently significant.

Previous studies (Kneer & Machery, 2019; Martin & Cushman, 2016; Nobes & Martin, 2022) have suggested that the outcome of an action has a stronger influence or explanatory power on the level of punishment than on the level of blame and wrongness. To determine if these results uphold in a novel context, we calculated linear regressions of these three variables with the outcome as the explanatory factor. The linear regression models showed that, in all three models, the outcome of the action significantly predicted the level of the respective variables. The results also indicate that the outcome predicts the severity of punishment to a greater extent than blame and wrongness. The results of these calculations, including the coefficients, can be found in Table 1. Furthermore, the outcome of the action explained a significant portion of the variance in the levels of punishment ($R^2 = .06$, $F(1, 361) = 24.04$, $p < .001$), blame ($R^2 = .02$, $F(1, 361) = 7.28$, $p = .007$), and wrongness ($R^2 = .02$, $F(1, 361) = 8.02$, $p = .005$). These results corroborate the findings of the aforementioned studies is also evident in our study.

To test if the pattern of Nobes and Martin's (2022) study concerning the nonexistent differences between the negligent and neutral conditions regarding perceived negligence would also be evident in this study, a one-way ANOVA was conducted. Due to unequal variances across groups, a Welch test was computed (Urdu, 2005), and the Games-Howell post-hoc test was used to detect differences between groups due to unequal group sizes (Toothaker, 1993). Welch's test revealed a significant difference between the groups for each criterion (Table 2). The Games-Howell post-hoc test showed a significant difference ($p < .001$) in punishment scores between the negligent and neutral groups ($\Delta M = 1.78$, 95% CI [1.31, 2.24]). The difference was also significant for blame ($\Delta M = 2.64$, $p < .001$, 95% CI [2.16, 3.12]), wrongness ($\Delta M = 2.20$, $p < .001$, 95% CI [1.73, 2.67]), and perceived negligence ($\Delta M = 2.86$, $p < .001$, 95% CI [2.35, 3.37]). The results showed significant differences between the groups for the tested variables. To conclude, we cannot confirm that the result of nonexistent differences in evalu-

ation between neutral and negligent agents, as found in Nobes and Martin's (2022) results, replicated in our de-

sign; we found a clear difference between both groups.

Table 1: The Effect of Outcome on Moral Judgments

	Dependent Variable		
	Punishment	Blame	Wrongness
Outcome	0.879* (0.179)	0.584* (0.217)	0.555* (0.196)
95% CI	[0.526, 1.232]	[0.159, 1.010]	[0.169, 0.941]

Note. This table reports the linear regression estimates of the negative outcome on the agent's judgment. Standard errors are shown in parentheses.

*Significant at the 1 percent level.

Table 2: Welch-test to Test for Equality of Means

		Statistic ^a	df ₁	df ₂	p
Punishment	Welch	72.622	2	232.887	<.001
Perceived negligence	Welch	204.747	2	225.642	<.001
Blame	Welch	183.618	2	229.138	<.001
Wrongness	Welch	128.455	2	229.358	<.001

Note. ^aAsymptotically F-distributed.

4 Discussion

This study aimed to conceptually replicate the recent findings of Nobes and Martin (2022) and generate original data on the topic of moral luck to gain a better understanding of the phenomenon, its occurrence, and outcome bias as an explanatory factor. Moreover, it scrutinizes the underlying theoretical explanations of earlier studies (Martin & Cushman, 2016; Nobes & Martin, 2022) and examines the generalizability and suitability of these frameworks in a novel context. The results of the first hypothesis provide clear evidence for the occurrence of outcome bias, that is, the influence of the evaluation of the agent. Participants tended to punish the agent more severely, ascribed more blame, and considered them more wrong if their action resulted in a negative outcome rather than a positive outcome (the death of the child). Thus, as many previous studies have shown (Cushman, 2008; Gino et al., 2009; Lench et al., 2015; Kneer & Machery, 2019; Martin & Cushman, 2016; Nobes & Martin, 2022), an outcome bias distorts the evaluation of an agent. This study shows that these effects are also observed in the German samples.

This provides further evidence for the robustness of outcome effects in the realm of moral discussions and situations but is particularly interesting because of the vignette used in this study. In previous studies on the topic of moral luck, only situations that already gave reason to anticipate an accident or mishap were described (to the best of our knowledge from literature research), at least to a greater extent than in this study design. Examples of such situations include doctors approving a new, possibly risky medication (Gino et al., 2009), governments dis-

cussing the use of flood-proof tents (Gino et al., 2009), parents leaving their toddlers unattended in the bathroom next to a full bathtub (Kneer & Machery, 2019), or the example of the girl driving a car through a pile of leaves beside the road (Martin & Cushman, 2016).

Of course, even in these situations, it cannot be assumed that participants immediately expected an accident to occur. However, it seems reasonable to assume that these situations represent a more noticeable danger than having two people eat dinner together. Thus, it is important to acknowledge that moral luck and its associated outcome effects also occur in situations that could be considered safe or inconspicuous, and not only in situations that presumably tend to stimulate semantic networks to perceive possible accident scenarios.

Consistent with previous studies (Kneer & Machery, 2019; Martin & Cushman, 2016; Nobes & Martin, 2022), our results show that the influence of the outcome is particularly strong for punishment judgments. That is, the outcome explained the severity of the punishment ascribed to the agent to a greater extent than the degree of moral condemnation (blame and wrongness). The pattern found in these three studies was also evident, indicating the robustness of the findings. Moreover, we could replicate the findings that the outcome effects on blame and wrongness are partially or wholly mediated, respectively, when adding perceived negligence as an explanatory factor.

These results provide limiting arguments to the Two-Process Model of Martin and Cushman (2016). According to the model, the attributed blame of an agent is influ-

enced by the consequences of the action as well as their intention. Given the present situation, it would be unreasonable for someone to assume that Noah had a bad intent; it is more likely the opposite, as a shared meal is presumably associated with good intent. Therefore, this factor is eliminated. Accordingly, the outcome remains the determining factor for blame judgment in the model. However, this was largely mediated by perceived negligence.

The same pattern is observed when examining wrongness. As stated in the Two-Process Model, this is explained through a single path that evaluates the intentions and beliefs of the actor; however, this is most likely to be omitted based on the above reasoning. The outcome of the situation is an unintentional accident, and according to the Two-Process Model, should not result in Noah being judged as morally wrong, regardless of the outcome of the situation.

Despite that, Noah is judged differently depending on the outcome of the situation; yet, this effect is explained by the perceived negligence, which suggests that the model of Martin and Cushman (2016) does not provide an exhaustive explanation on how moral judgments form and needs to be extended for factors such as negligence.

While the moderating effect of negligence on punishment judgments ceased to be found in classical statistical moderation analysis, a separate analysis of the qualitative answers of the participants gave rise to the notion that there is an underlying interaction between perceived negligence and the severity of punishment. Participants ascribed more severe punishments (imprisonment, probation, and fines) only when the outcome was negative and Noah acted negligently. In the case of a positive outcome, the vast majority mentioned that Noah should not receive punishment; however, the frequency of more severe punishments increased when moving from the non-negligent group to the negligent group. That is, individuals seemed to punish an agent more severely the more negligent he acted. In addition to the intuitively plausible nature of this effect, this is also in line with the results obtained by Nobes and Martin (2022).

However, very importantly, this finding should be treated with caution as the interaction analysis using Hayes (2008) linear regression model did not yield a moderation effect. A possible explanation for why this effect was not found in the analytical moderation analysis may be the measurement of punishment level. The Likert scale differentiates between no punishment - somewhat strong - very strong (1-7). However, these levels did not specify an unequivocal meaning (except for no punishment). That is, a level of five may mean Noah's dismissal for one person, while for another, it may mean a six-month prison sentence. This distorts the results and analysis and may have led to the absence of this effect in the moderation analysis. This might also explain why Nobes and Martin found an interaction effect in their design, as they assigned concrete punishments to different levels of their Likert scale, making the different levels unambiguous.

In addition, despite indicating a moderating effect of negligence on the punishment of the agent, the qualitative analysis lacks a rigorous, systematic scaffold from a statistical moderation analysis, thus questioning the significance of the conclusion. This is very important to keep in mind when engaging with this finding, as it should be noted that it provides suggestive evidence of a possible moderating effect; however, it is by no means conclusive. We propose that it should serve as a nudge for further research to employ a more systematic approach to identifying the interaction effect of negligence on punishment, as this might hold substantial implications for decision-making, spanning from interpersonal to judiciary to corporate contexts.

Supposedly, the most relevant findings of this study relate to the results of Nobes and Martin (2022), which dealt with the differences in the perception of the agent's negligence between the neutral and negligent groups. Nobes and Martin (2022) found no significant differences between these groups in their study design, leading them to assume that people automatically assume negligence in the absence of clues to the agent's negligence.

Considering the results that illustrate the effects of perceived negligence on the punishment and evaluation of an agent, this could mean that the evaluation of actors in situations without information about negligence is disproportionately negatively influenced.

Since the negligence of an agent cannot always be easily identified, this could affect interpersonal situations where the moral appearance or social status of a person is impaired and legal judgments where the stakes are usually much higher. Moreover, this plays a crucial role in an organizational context, as blame judgments can have a substantial impact on organizational citizenship behavior (Costa & Neves, 2017) and may be related to moral distress, burnout, compassion fatigue, lateral violence, and second-victim syndrome (for a review see Davidson et al., 2015).

Consequently, the investigation of this result was of substantial significance. However, our findings showed significant differences between the groups. That is, our study showed no evidence of the pattern found by Nobes and Martin (2022). Not only with regard to the perceived negligence of the agent, but also in terms of their evaluation. That is, the more negligent the behavior, the more negligent the person was perceived, and the more strongly they were condemned, showing that these findings could not be replicated.

5 Limitations

One limitation of this study was the composition of the participant pool. The majority of participants were relatively young (75% between 17 and 29 years old) and were primarily recruited through student channels. Therefore, inferences regarding these effects on the general population must be treated with caution (Bortz & Döring, 2006). However, as most other studies on moral luck relied on student-based sample sizes, it grants adequate

comparability, especially when compared to the original study by Nobes and Martin (2022).

Another limitation is the validity of the negligence variable. Owing to the time frame and scope, no pre-test could be conducted to test the validity of the variable. However, the high correlation between IV negligence and the perceived negligence variable, $r_{(361)} = .67, p < .001$, provided strong evidence for the validity of the variable.

As already mentioned, this study used an unusual vignette compared to the vignettes used in similar studies, in the sense that it displayed a much more inconspicuous situation. This deviance might shed light on why Nobes and Martin's (2022) findings regarding the nonexistent differences in perceived negligence between the neutral and negligent groups could not be replicated. This is because of the lower salience of possible accidents/dangers in our design compared to Nobes and Martin's (2022), which could have led to a greater perceptual discrepancy between the neutral and negligent conditions.

This means that our participants may not have thought about a potential accident danger in this trivial situation without a cue for negligence, thereby resulting in a significant difference between the neutral and negligent conditions in our design. Presumably, this discrepancy was not as clear in the situation with the girl driving on the road, which may have led to non-significant effects because participants were more prone to judge her actions as negligent regardless of whether the situation offered information about her negligence due to the nature of the situation. To gain a better understanding on how this effect may come into existence and how situational cues influence it, this aspect should be further investigated, particularly in situations that pose a more salient risk.

Another limitation was the perceived causality of the agent. While the results strongly confirm the effects of outcome bias, negligence, and moral luck, this study did not measure how causally responsible the agent was perceived. Causality is also a frequently studied factor in the field of moral luck (Cushman, 2008; Martin & Cushman, 2016; Nobes & Martin, 2022), which influences the evaluation of the agent, albeit mediated by negligence (Nobes & Martin, 2022).

Accordingly, perceived causality cannot be excluded as another influencing factor, and there is a need for a similar design that includes a variable to measure the agent's perceived causality. This could be particularly interesting in this or similar vignettes, as one could argue that the agent's direct perceived causality is somewhat less salient. The rationale for this is that the effective reason for the accident - the food brought - is not directly connected to the au pair as it is brought by the waiter.

In other vignettes, the direct influence of the agent, and thus the perceived causality, may be more salient; for example, the self-executed action of the girl driving through the pile of leaves, the mother knowingly leaving her child alone next to the full bathtub, or the doctor who is directly responsible for issuing the medication. Consequently, it would be interesting to determine the role of

causality in situations where the accident cause is not immediately connected to the agent.

6 Conclusion

The present study investigated the influence of negligence on moral judgments and evaluations in the context of moral luck. Using a vignette-based approach, we conceptually replicated the recent study by Nobes and Martin (2022) and demonstrated the robustness of moral luck and the influence of outcome effects in a novel context. Our results provide further evidence for the significant role of negligence in moral decision making and evaluations, while also limiting some of the findings.

In particular, this study contributes to the literature by highlighting the potential impact of negligence as a factor in understanding moral luck and outcome bias. Our findings suggest that negligence plays a substantial role in shaping moral judgments, particularly in between-subjects designs where direct comparisons or alternative outcomes are not readily available. The observed outcome bias was largely influenced by perceived negligence, emphasizing the need to consider this contextual factor in various real-life contexts such as legal, private, or professional evaluations.

Furthermore, our study aligns with previous research that has shown divergent results regarding the occurrence of moral luck and outcome bias in within-subject designs. By incorporating negligence as an independent variable, future within-subjects studies can shed light on the conditions under which negligence influences moral luck and outcome effects. This avenue of research deepens our understanding of the interplay among negligence, moral judgment, and contextual factors.

On the topic of contextual factors, we showed that the presumably highly impactful findings regarding non-different perceptions of an agent's negligence did not hold up in our setting. Thus, it does not seem to be the case that this phenomenon is pervasive across situations and cannot be generalized. Our findings suggest that there is more nuance to this evaluation which should be further explored through subsequent studies, especially in different situations.

From a practical standpoint, these findings have important implications for decision makers in legal and professional settings. Given the prevalence of between-subject designs and the limited opportunities for direct comparisons, considering the role of negligence is crucial. As our findings did not replicate the nonexistent differences between neutral and negligent agents, we strongly urge decision makers to seek out any clues that provide information about the negligence of an actor to come to a more substantiated assessment of an agent's negligence. This could make the difference between an individual being fired or getting off with a warning.

Presenting alternative outcomes and encouraging evaluators to assess the extent to which an actor's behavior could have led to different outcomes may help contextualize assessments and reduce outcome effects (Bazerman et

al., 2011; Gino et al., 2009; Kneer & Machery, 2019) and might also help to pigeonhole the negligence of an agent. This recommendation underscores the importance of incorporating negligence manipulation as an independent variable in further investigations.

In conclusion, this study adds to the growing body of literature on moral luck by conceptually replicating previ-

ous findings through a different vignette, emphasizing the significance of negligence in moral decision making, and highlights the need for future research to explore the interrelationships among negligence, moral luck, and outcome bias. These findings offer valuable insights for understanding moral judgments and have practical implications for the decision-making contexts in which moral judgments occur.

7 References

- Baron, J. & Hershey, J. C. (1988). Outcome Bias in Decision Evaluation. *Journal of Personality and Social Psychology*, 54(4), 569-579. <https://doi.org/10.1037/0022-3514.54.4.569>
- Bazerman, M. H., Gino, F., Shu, L. L. & Tsay, C. (2011). Joint evaluation as a real world tool for managing emotional assessment of morality. *Emotion Review*, 3(3), 290-292. <https://doi.org/10.1177/1754073911402370>
- Bortz, J. & Döring, N. (2006). *Forschungsmethoden und Evaluation: für Human- und Sozialwissenschaftler*. Springer.
- Broude, T. & Levy, I. (2019). Outcome Bias and Expertise in Investigations under International Humanitarian Law. *European Journal of International Law*, 30(4), 1303-1318. <https://doi.org/10.1093/ejil/chaa005>
- Brown, R. V., Kahr, A. S. & Peterson, C. (1974). *Decision analysis for the manager*. Holt, Rinehart and Winston.
- Cohen, J. (1992). A power primer. *Psychological bulletin*, 112(1), 155-159. <https://doi.org/10.1037/0033-2909.112.1.155>
- Cole, S., Healy, A. & Werker, E. (2012). Do voters demand responsive governments? Evidence from Indian disaster relief. *Journal of Development Economics*, 97(2), 167-181. <https://doi.org/10.1016/j.jdeveco.2011.05.005>
- Costa, S., & Neves, P. (2017). Job insecurity and work outcomes: The role of psychological contract breach and positive psychological capital. *Work & Stress*, 31(4), 375-394. <https://doi.org/10.1080/02678373.2017.1330781>
- Crandall, C. S., & Sherman, J.W. (2016). On the scientific superiority of conceptual replications for scientific progress. *Journal of Experimental Social Psychology*, 66, 93-99. <https://doi.org/10.1016/j.jesp.2015.10.002>
- Curran, P. G. (2017). Methods for the detection of carelessly invalid responses in survey data. *Journal of Experimental Social Psychology*, 66, 4-19. <https://doi.org/10.1016/j.jesp.2015.07.006>
- Cushman, F. (2008). Crime and punishment: Distinguishing the roles of causal and intentional analyses in moral judgment. *Cognition*, 108(2), 353-380. <https://doi.org/10.1016/j.cognition.2008.03.006>
- Cushman, F. & Martin, J. W. (2016). The adaptive logic of moral luck. In J. Systema & W. Buckwalter (Eds.), *A Companion to Experimental Philosophy* (S. 190-205). Wiley-Blackwell.
- Davidson, J. E., Agan, D. L., Chakedis, S., & Skrobik, Y. (2015). Workplace Blame and Related Concepts: An Analysis of Three Case Studies. *Chest*, 148(2), 543-549. <https://doi.org/10.1378/chest.15-0332x>
- Davidson, R. & MacKinnon, J. G. (1993). *Estimation and Inference in Econometrics*. Oxford University Press.
- Gino, F., Moore, D. A. & Bazerman, M. H. (2009). No harm, no foul: outcome bias in ethical judgments. *Harvard Business School NOM Working Paper*, 08-080.
- Gruppen, L. D., Margolin, J., Wisdom, K. & Grum, C. M. (1994). Outcome bias and cognitive dissonance in evaluating treatment decisions. *Academic Medicine*, 69(10), 57-59. [10.1097/00001888-199410000-00042](https://doi.org/10.1097/00001888-199410000-00042)
- Kant, I. (1785). *Die Grundlegung zur Metaphysik der Sitten*. J.F. Hartknoch.
- Kneer, M. & Machery, E. (2019). No luck for moral luck. *Cognition*, 182, 331-348. <https://doi.org/10.1016/j.cognition.2018.09.003>
- Laurent, S. M., Nuñez, N. L. & Schweitzer, K. A. (2016). Unintended, but still blameworthy: The roles of awareness, desire, and anger in negligence, restitution, and punishment. *Cognition and Emotion*, 30(7), 1271-1288. <https://doi.org/10.1080/02699931.2015.1058242>
- Lefgren, L., Platt, B. & Price, J. (2014). Sticking with What (Barely) Worked: A Test of Outcome Bias. *Management Science*, 61(5), 1121-1136. <https://doi.org/10.1287/mnsc.2014.1966>
- Lench, H. J., Domsky, D., Smallman, R. & Darbor, K. E. (2015). Beliefs in moral luck: When and why blame hinges on luck. *British Journal of Psychology*, 106(2), 272-287. <https://doi.org/10.1111/bjop.12072>
- Levy, N. (2016). Dissolving the Puzzle of Resultant Moral Luck. *Review of Philosophy and Psychology*, 7, 127-139. <https://doi.org/10.1007/s13164-015-0249-0>
- Margoni, F., Geipel, J., Hadjichristidis, C. & Surian, L. (2019). The influence of agents' negligence in shaping younger and older adults' moral judgment.

- Cognitive Development*, 49, 116-126. <https://doi.org/10.1016/j.cogdev.2018.12.002>
- Mulvey, K. L., Gönültaş, S. & Richardson, C. B. (2020). Who is to blame? Children's and adults' moral judgments regarding victim and transgressor negligence. *Cognitive Science*, 44(4), Article e12833. <https://doi.org/10.1111/cogs.12833>
- Nagel, T. (1979). *Mortal Questions*. Cambridge University Press.
- Nobes, G. & Martin, J. W. (2022). They should have known better: The roles of negligence and outcome in moral judgments of accidental actions. *British Journal of Psychology*, 113(2), 370-395. <https://doi.org/10.1111/bjop.12536>
- Nobes, G., Panagiotaki, G. & Engelhardt, P. (2017). The development of intention-based morality: The influence of intention salience and recency, negligence, and outcome on children's and adults' judgments. *Developmental Psychology*, 53, 1895-1911. <https://doi.org/10.1037/dev0000380>
- Rasch, B., Friese, M., Hofmann, W. & Naumann, E. (2014). *Quantitative Methoden 2 – Einführung in die Statistik für Psychologen und Sozialwissenschaftler*. Heidelberg: Springer.
- Rizzo, M. T., Li, L., Burkholder, A. R. & Killen, M. (2019). Lying, negligence, or lack of knowledge? Children's intention-based moral reasoning about resource claims. *Developmental Psychology*, 55, 247-258. <https://doi.org/10.1037/dev0000635>
- Savani, K. & King, D. (2015). Perceiving outcomes as determined by external forces: The role of event construal in attenuating the outcome bias. *Organizational Behavior and Human Decision Processes*, 130, 136-146. <https://doi.org/10.1016/j.obhdp.2015.05.002>
- Statman, D. (1993). *Moral Luck*. State University of New York Press.
- Toothaker, L. E. (1993). *Multiple comparison procedures*. Sage Publications.
- Urdan, T. C. (2005). *Statistics in Plain English*. Mahwah: Lawrence Erlbaum Associates, Inc.
- Williams, B. (1981). *Moral Luck*. Cambridge University Press.
- Young, L., Nichols, S. & Saxe, R. (2010). Investigating the Neural and Cognitive Basis of Moral Luck: It's Not What You Do but What You Know. *Review of Philosophy and Psychology*, 1(3), 333-349. <https://doi.org/10.1007/s13164-010-0027-y>

Appendix A: Original Vignette Used in Study Design

Nicht fahrlässig bei positivem Handlungsausgang

Noah ist Au-pair bei einer Familie mit einem Kind (5 Jahre). Heute sind die beiden zusammen in einem Restaurant essen. Das Kind bestellt sein Essen. Nachdem sie ihre Bestellungen aufgegeben haben, fragt Noah den Kellner, ob in den Gerichten bestimmte Nahrungsmittel enthalten sind, da das Kind darauf allergisch reagiert. Der Kellner hält Rücksprache mit dem Koch und verneint. Noah ist beruhigt und die Bestellung geht in die Küche. Nachdem beide gegessen haben, gehen sie gesättigt nach Hause.

Nicht fahrlässig bei negativem Handlungsausgang

Noah ist Au-pair bei einer Familie mit einem Kind (5 Jahre). Heute sind die beiden in einem Restaurant essen. Das Kind bestellt sein Essen. Nachdem sie ihre Bestellungen aufgegeben haben, fragt Noah den Kellner, ob in den Gerichten bestimmte Nahrungsmittel enthalten sind, da das Kind darauf allergisch reagiert. Der Kellner hält Rücksprache mit dem Koch und verneint. Noah ist beruhigt und die Bestellung geht in die Küche. Nachdem beide ihre ersten Portionen gegessen haben, erstarrt das Kind, bekommt einen allergischen Schock und verstirbt noch in dem Restaurant.

Keine Fahrlässigkeitsinformation bei positivem Handlungsausgang

Noah ist Au-pair bei einer Familie mit einem Kind (5 Jahre). Heute sind die beiden zusammen in einem Restaurant

essen. Das Kind bestellt sein Essen. Der Kellner nimmt die Bestellung auf und bringt sie in die Küche. Nachdem beide gegessen haben, gehen sie gesättigt nach Hause.

Keine Fahrlässigkeitsinformation bei negativem Handlungsausgang

Noah ist Au-pair bei einer Familie mit einem Kind (5 Jahre). Heute sind die beiden zusammen in einem Restaurant essen. Das Kind bestellt sein Essen. Der Kellner nimmt die Bestellung auf und bringt sie in die Küche. Nachdem beide ihre ersten Portionen gegessen haben, erstarren das Kind, bekommt einen allergischen Schock und verstirbt noch in dem Restaurant.

Fahrlässig bei positivem Handlungsausgang

Noah ist Au-pair bei einer Familie mit einem Kind (5 Jahre). Heute sind die beiden zusammen in einem Restaurant essen. Beide bestellen ihr Essen. Die Eltern des Kindes hatten Noah zwar darüber unterrichtet, dass ihr Kind Nussallergie hat, jedoch vergisst Noah dies und fragt nicht, ob diese Stoffe in dem Essen verarbeitet wurden. Der Kellner bringt die Bestellung in die Küche. Nachdem beide gegessen haben, gehen sie gesättigt nach Hause.

Fahrlässig bei negativem Handlungsausgang

Noah ist Au-pair bei einer Familie mit einem Kind (5 Jahre). Heute sind die beiden zusammen in einem Restaurant essen. Beide bestellen ihr Essen. Die Eltern des Kindes hatten Noah zwar darüber unterrichtet, dass ihr Kind Nussallergie hat, jedoch vergisst Noah dies und fragt nicht, ob diese Stoffe in dem Essen verarbeitet wurden.

Der Kellner bringt die Bestellung in die Küche. Nachdem die beiden ihre ersten Portionen gegessen haben, erstarrt das Kind, bekommt einen allergischen Schock und ver stirbt noch in dem Restaurant.

Appendix B: Translated Original Vignette Used in Study Design

Non-negligent with positive outcome

Noah is an au pair with a family with a child (5 years old). Today, two of them eat together at a restaurant. The child orders his food. After they have placed their orders, Noah asks the waiter if there are certain foods in the dishes, because the child is allergic to them. The waiter consults the cook and answers in the negative. Noah is reassured and the order goes to the kitchen. After both have eaten, they go home satiated.

Non-negligent with negative outcome

Noah is an au pair with a family with a child (5 years old). Today, two of them eat together at a restaurant. The child orders his food. After they have placed their orders, Noah asks the waiter if there are certain foods in the dishes, because the child is allergic to them. The waiter consults the cook and answers in the negative. Noah is reassured and the order goes to the kitchen. After both have eaten their first portions, the child freezes, experiences an allergic shock, and dies while still in the restaurant.

No negligence information with positive outcome

Noah is an au pair with a family with a child (5 years old). Today, two of them eat together at a restaurant. The child

orders his food. The waiter takes the order and brings it to the kitchen. After both have eaten, they go home satiated.

No negligence information with negative outcome

Noah is an au pair with a family with a child (5 years old). Today, two of them eat together in a restaurant. The child orders his food. The waiter takes the order and brings it to the kitchen. After both have eaten their first portions, the child freezes, experiences an allergic shock, and dies while still in the restaurant.

Negligent with positive outcome

Noah is an au pair with a family with a child (5 years old). Today, two of them eat together in a restaurant. They both order their food. The parents of the child had informed Noah that their child has nut allergies, but Noah forgets this and does not ask whether these substances were used in the food. The waiter takes the order to the kitchen. After both have eaten, they go home satiated.

Negligent with negative outcome

Noah is an au pair with a family with a child (5 years old). Today, two of them eat together in a restaurant. They both order their food. The parents of the child had informed Noah that their child has nut allergies, but Noah forgets this and does not ask whether these substances were used in the food. The waiter takes the order to the kitchen. After the two have eaten their first portions, the child freezes, experiences allergic shock, and dies in the restaurant.

Appendix C: Study Questionnaire

Table 3: Original questions and response rating scales

Question	Response
1. Was ist in dem eben gezeigten Szenario passiert?	1 = Zwei Personen waren zusammen essen; 2 = Zwei Personen waren zusammen sportmachen; 3 = Zwei Personen haben zusammen Hausaufgaben erledigt
2. Falls Sie Noahs Verhalten bestrafen würden, wie stark würden Sie das tun?	1 = gar nicht; 4 = mittelstark; 7 = sehr stark
3. Geben Sie bitte an, was für eine Bestrafung Sie als angemessen betrachten (Bspw. könnten Sie unterscheiden zwischen Entlassung, Gefängnisstrafe, Geldstrafe, Bewährungsstrafe. Sie können natürlich auch gar keine, mildere oder stärkere Bestrafungen als angemessen erachten; es gibt keine richtigen oder falschen Antworten. Geben Sie einfach das Maß an Bestrafung an, dass Sie als angemessen empfinden)	Offene Textangabe

4. Für wie fahrlässig halten sie das Verhalten von Noah?	1 = gar nicht fahrlässig; 4 = einigermaßen fahrlässig; 7 = sehr fahrlässig
5. Wie schuldfähig ist Noah für den Ausgang der Situation?	1 = gar nicht schuldfähig; 4 = einigermaßen schuldfähig; 7 = sehr schuldfähig
6. Für wie unanständig/moralisch inkorrekt halten Sie das Verhalten von Noah?	1 = gar nicht unanständig/inkorrekt; 4 = einigermaßen unanständig/inkorrekt; 7 = sehr unanständig/inkorrekt

Appendix D: Translated Study Questionnaire

Table 4: Translated questions and response rating scales

Question	Response
1. What happened in the displayed scenario?	1 = Two persons went out to eat together; 2 = Two persons exercised together; 3 = Two persons did homework together
2. In case you would punish Noah's actions, to what extent would you punish them?	1 = not at all; 4 = somewhat; 7 = very harshly
3. Please indicate what kind of punishment you consider appropriate (e.g., you could distinguish between dismissal, imprisonment, fine, suspended sentence). Of course, you can also consider no punishment at all, milder or stronger punishment as appropriate; there are no right or wrong answers. Simply indicate the level of punishment that you feel is appropriate.)	Open response field
4. To what extent were Noah's actions negligent?	1 = not negligent at all; 4 = somewhat negligent; 7 = very negligent
5. To what extent was Noah blameworthy for the way things turned out?	1 = not blameworthy at all; 4 = somewhat blameworthy; 7 = very blameworthy
6. How wrong were Noah's actions?	1 = not wrong at all; 4 = somewhat wrong; 7 = very wrong

Correspondence addresses:

Niklas Fritsche
 9 Vanbern House
 83-85 Prince of Wales Road
 NW5 3LY London, UK
 n.c.fritsche@lse.ac.uk

Prof. Dr. Claudia Gerhardt
 Hochschule Fresenius
 Alte Rabenstraße 1
 20148 Hamburg
 DEUTSCHLAND
 gerhardt@hs-fresenius.de