

# Reasonableness on the Clapham Omnibus: Exploring the outcome-sensitive folk concept of *reasonable*

Markus Kneer  
University of Zurich  
markus.kneer@uzh.ch

To appear in Bystranowski, P, Janik, B. & Prochnicki, M. (2021) *Judicial Decision-Making: Integrating Empirical and Theoretical Perspectives*. Springer Nature. Please cite the final version.

## 1 Introduction<sup>1</sup>

This paper presents a series of studies (total N=579) which demonstrate that folk judgments concerning the reasonableness of decisions and actions depend strongly on whether they engender positive or negative consequences. A particular decision is deemed more reasonable in retrospect when it produces beneficial consequences than when it produces harmful consequences, even if the situation in which the decision was taken and the epistemic circumstances of the agent are held fixed across conditions. This finding is worrisome for the law, where the reasonable person standard plays a prominent role. The legal concept of reasonableness is outcome-insensitive: whether the defendant acted in a reasonable fashion or not depends exclusively on her context of action, no matter how things play out. Folk judgments of reasonableness are thus inconsistent with the legal concept of reasonableness. Problematically, in common law jurisdictions, the decision whether a defendant's behavior was reasonable or not is frequently (though not necessarily) delegated to a lay jury. What could, and should, be done about this problem depends on what explains the outcome-sensitivity of folk judgments of reasonableness.

*Prima facie*, the outcome-sensitivity of reasonableness judgments allows for different interpretations. On a bias interpretation, perceived reasonableness *post hoc* is inappropriately distorted by outcome information. This view comes in two flavours: The *unmediated bias* view, according to which the effect of outcome information on reasonableness judgments is (predominantly) direct. Alternatively, there is the *mediated bias* view, according to which the distorting effect of outcome information is mediated at least partially by another factor. What kind of factor? Perceived probability is, perhaps, the most plausible candidate. On this view, a decision is deemed more or less reasonable because the (beneficial or harmful) consequences of said decisions are deemed more or less *probable post hoc* depending on what actually happened. Or, in a nutshell, the outcome bias in reasonableness judgments could be reduced, at least in parts, to the hindsight bias.

---

<sup>1</sup> I would like to thank Teneille Brown, Morris Hoffman, Joshua Knobe, James Macleod, Kevin Tobia, the members of the Guilty Minds Lab Zurich and the Editors for very helpful comments and suggestions. Research on this paper was supported by an SNSF Ambizione grant for the project *Reading Guilty Minds* (PZ00P1\_179912).

Importantly, however, the fact that outcome impacts reasonableness judgments need not constitute a bias in the first place. The folk might deem it *appropriate* to assess the reasonableness of decisions and actions *inter alia* on the basis of the consequences these decisions and actions produce. In such a case, it is more adequate to treat the folk concept of *reasonable* as inherently outcome-dependent, and the outcome-dependence of reasonableness judgments would not only be *consistent with* but *due to* a correct deployment of the concept.

In this paper, I will explore which interpretation of the outcome-sensitive folk judgments of reasonableness is correct and what this means for the law, in particular in common law systems where lay jurors (i.e. the folk) play a key role in adjudicating whether a defendant acted in a reasonable fashion. To do so, I will begin with *theoretical preliminaries*, focusing on the scope of the reasonable person standard, the legal concept of *reasonable*, and general problems that can arise when the law relies on jurors. A brief survey of previous pertinent *empirical work* motivates the experiments. The bulk of the paper reports three studies that suggest that the folk concept of reasonableness is outcome-dependent and explores what that means for the law – in particular in countries where questions pertaining to reasonableness are frequently decided by lay juries.

### ***1.1 The Reasonable Person Standard***

Many of the law’s demands are tied to what it is *reasonable* to believe, feel, judge and do.<sup>2</sup> Negligence, for instance, is attributed if a reasonable person would have been aware of a substantial risk in the defendant’s circumstances.<sup>3</sup> Self-defense can be claimed if a reasonable person would have justifiably feared for their life in the defendant’s situation.<sup>4</sup> But reasonableness also plays a prominent role in cases concerning, for instance, excessive fees, rates or prices,<sup>5</sup> statutory rape,<sup>6</sup> deceptive advertising,<sup>7</sup> restraints on trade,<sup>8</sup> delivery on contractually agreed terms,<sup>9</sup> or the public dissemination of private information.<sup>10</sup> Consequently, the reasonable person is an

---

<sup>2</sup> For discussion, see e.g. Fletcher (1984), Gardner (2001, 2015), Moran (2003, 2010), Westen (2008), Miller & Perry (2012), Zipursky (2014), Tobia (2018).

<sup>3</sup> Model Penal Code §2.02. (d)

<sup>4</sup> UK Criminal Law Act, 1967, c. 58, § 85 sch. 3(1) (Eng. & Wales) (“A person may use such force as is reasonable in the circumstances in the prevention of crime [ . . . ] .”); United States v. Peterson 483 F 2d 1222 (1973); NY Penal Code Art. 35; see also Baron (2011).

<sup>5</sup> See, e.g., 16 U.S.C. § 824e(a) (2012) (stating that the Federal Energy Regulatory Commission must set “the just and reasonable rate”); 16 U.S.C. § 2621(18)(B) (2012) (stating that states should ensure “a reasonable rate” of return when authorizing utility companies. More generally, see Berger (1977).

<sup>6</sup> Model Penal Code § 213.6(1)

<sup>7</sup> See, e.g., Fink v. Time Warner Cable, 714 F.3d 739, 741 (2d Cir. 2013); Elias v. Hewlett-Packard Co., 950 F. Supp. 2d 1123, 1131 (N.D. Cal. 2013).

<sup>8</sup> See e.g. Leegin Creative Leather Prod., Inc. v. PSKS, Inc., 551 U.S. 877, 885 (2007), (“[T]he Court has repeated time and again that §1 ‘outlaw[s] only unreasonable restraints.’”); Standard Oil Co. of N.J. v. United States, 221 U.S. 1, 60 (1911) (stating that “the standard of reason” regards anti-competitive trade restraints under the Sherman Act).

<sup>9</sup> See e.g. Mayfield v. Koroghli, 184 P.3d 362, 366 (Nev. 2008) (stating that parties must deliver on contracts within a “reasonable time”).

<sup>10</sup> Restatement (Second) of Torts §652D (1977).

authority on everything from the appropriate degree of emotional distress when under attack to which kinds of information that are deemed too embarrassing to have disclosed in public.

The reasonable person standard, Gardner argues, “exists to allow the law *to pass the buck*, to help itself *pro tempore* to standards of justification that are not themselves set by the law, and which therefore are only as good as the standards of justification used by the person or people to whom the buck is passed.” Recurrence to the reasonable person allows the law to significantly expand its scope without the introduction of countless specific and controversial rules. Instead the buck is passed to the fact-finder, frequently a jury in common law countries, “who is invited to use extra-legal standards to bridge the gap from legal rule to legal ruling.”<sup>11</sup>

### ***1.2 Two General Threats to Trial by Jury***

When the law “passes the buck” to lay juries, they need the relevant competences in the tasks assigned to them. One competence regards the grasp of the particular concept at stake as defined by the law. Jurors must know (often following instructions) how expressions such as “reasonable”, but also “intention”, “cause”, or “probability” are understood in the legal context. Another competence regards their application: Jurors must assess the reasonableness of the action, and other relevant features of the case at issue in responsible and unbiased fashion. In trials where jurors are involved, justice is thus under potential threat from at least two sources: *unnoticed conceptual divergence* between the legal and the folk concepts at stake, and *cognitive bias*.

Naturally, once identified, each problem requires a different remedy. If the law takes itself to operate with a folk concept of e.g. reasonableness, causation, or intention, and this assumption turns out mistaken (the actual folk concept differs from the presumed folk concept), we must work towards *conceptual alignment*. On the other hand, if conceptual alignment exists, yet cognitive biases distort the assessment of reasonableness, causation, *mens rea* and the like, strategies to *debias such judgments* must be devised. Given the vast reach of the reasonable person standard and serious concerns as to whether the folk concept of, e.g. intention or causation fit the ones in use in the law,<sup>12</sup> it is of interest to test whether folk judgments of what is reasonable are consistent with the legal concept of reasonableness. In order to do this, we will begin with a closer examination of the legal meaning of “reasonable”.

### ***1.3 The Legal Concept of Reasonableness***

What is the legal concept of reasonableness, and what should it be? This question has received considerable attention over the last two decades. One debate explores the conceptual category or nature of reasonableness: Does “reasonable” denote a *descriptive* or a *normative* concept? More precisely, must we understand that which is reasonable descriptively as what an *average* or *ordinary* person would do? Or must we rather understand it as what an *upright* or *ideal* person

---

<sup>11</sup> Gardner, 2015, p. 36.

<sup>12</sup> See e.g. Alicke (1992, 2000), Nadelhoffer (2006), Macleod (2019), Kneer & Bourgeois-Gironde (2017).

would do, in which case the concept of reasonableness would be normative?<sup>13</sup> Perhaps neither is correct *tout court*, and the expression “reasonable” is ambiguous and takes *different* meanings in different legal contexts (Zipursky, 2014). It could even have a *hybrid* meaning with both descriptive and normative features (Tobia, 2018). A distinct debate examines whether it is appropriate to hold everyone, no matter their particular individual or “subjective” features to account with respect to a single “objective” standard.<sup>14</sup>

The descriptive/normative and the subjective/objective debates are fascinating and of considerable practical importance to the law. I will set them aside in this paper, and turn to a third question which I will label the potential *outcome-sensitivity* of the folk concept of reasonableness.<sup>15</sup> In recent work on practical rationality broadly conceived, I found judgments of what is reasonable, what is rational, and what should be done, to depend on outcome valence. In two situations A (negative outcome) and B (positive outcome), in which two agents have the same information and act with the same frame of mind, A’s decision is judged significantly less reasonable than B’s (Kneer, in prep. a, see also Baron & Hershey, 1988). These results are at odds with the legal concept of reasonableness (or at least the correct application thereof): Whether or not a decision, belief, emotional reaction or action is reasonable does not depend on the consequences, but the situation in which the agent finds herself. Or, to put it somewhat more technically, the semantics of the legal expression “reasonable” is sensitive to the *context of decision or action*, not the *ex post context of assessment* where the consequences are known (Kneer in prep. b, for the semantic framework, see MacFarlane, 2014). The Model Penal Code, for instance, is perfectly explicit about this with respect to criminal negligence.<sup>16</sup> The short paragraph emphasizes that what matters is *the agent’s* epistemic perspective and circumstances and not some context *ex post*, and this is stressed not only once, but twice (see italics):

A person acts negligently with respect to a material element of an offense when he should be aware of a substantial and unjustifiable risk that the material element exists or will result from his conduct. The risk must be of such a nature and degree that the actor's failure to perceive it, considering the nature and purpose of his conduct and *the circumstances known*

---

<sup>13</sup> For discussion, see Ripstein (1998), Gardner (2001, 2013), Moran (2003), Westen (2008), Miller & Perry, 2012, Zipursky (2014), Tilley (2016), Tobia (2018).

<sup>14</sup> For discussion, see. e.g. Wells (1982), Kelman (1991), Armour (1994), Moran (2003, 2010), Lee (2007), Simons (2008), Nourse (2008).

<sup>15</sup> This question, whose practical importance might even exceed the one of the other two debates has received next to no attention thus far. Two exceptions are the excellent empirical papers by Baron & Hershey (1988) and Kamin & Rachlinski (1995). The latter is limited to probability judgments in tort negligence, the former explicitly includes “reasonable” as a dependent variable, though mentions a plethora of other expressions in the same scale. Both studies have very low power, and neither of the two employs a clean within-subjects design, which – as will be argued below – is essential to a proper understanding of the folk concept of reasonableness.

<sup>16</sup> The same emphasis on the agent’s circumstances is found in the definition of recklessness (see MPC 2.02 (c)), where the “reasonable person” is substituted with “the law-abiding person”. Many commentators, however, treat negligence and recklessness as very similar, and the “law-abiding person” as a close cousin of the “reasonable person”. Both negligence and recklessness are a matter of *running an unreasonable risk*, the difference lying in whether the defendant did so knowingly or not. For discussion, see e.g. Dressler (1995), Simons (2002).

*to him*, involves a gross deviation from the standard of care that a reasonable person would observe *in the actor's situation*. (MPC, 2.02. (d), italics added)

Section 220.3 (regarding manslaughter) of the Model Penal Code further elucidates how “reasonableness” must be understood. Criminal homicide is a case of manslaughter if the killing “committed under the influence of extreme mental or emotional disturbance for which there is reasonable explanation or excuse. The reasonableness of such explanation or excuse shall be determined *from the viewpoint of a person in the actor’s situation under the circumstances as he believes them to be*.”<sup>17</sup>

The Restatement (3<sup>rd</sup>) of Torts on Negligence §3 further demonstrates the point:

A person acts negligently if the person does not exercise reasonable care under all the circumstances. Primary factors to consider in ascertaining whether the person's conduct lacks *reasonable care are the foreseeable likelihood* that the person's conduct will result in harm, the *foreseeable severity of any harm* that may ensue, and the burden of precautions to eliminate or reduce the risk of harm. (italics added)

What matters for the evaluation of reasonable care are the *foreseeable* likelihood of harm, and the *foreseeable* severity of harm – in short, the *foreseeable* risk. The law’s insistence on the foreseeability of risk thus limits the factors relevant for the assessment of reasonable care or negligence to features of the defendant’s circumstances *ex ante*. What actually happens, and what becomes known *ex post*, is irrelevant.<sup>18</sup>

In short, the legal meaning of “reasonable” is outcome-independent (for further discussion, see also the excellent Kamin & Rachlinski, 1995 and Rachlinski, 1998). Imagine, for a moment, that it were not. Then the vast area of the law in which judgments of reasonableness take centre stage would be a matter of pure outcome-luck.<sup>19</sup> Under conditions of risk and uncertainty, any action – no matter how rigorously thought through and carefully executed – can produce harmful

---

<sup>17</sup> For discussion, see e.g. Low 1987 and Dressler 1993.

<sup>18</sup> Given that the law explicitly addresses the *severity* of the foreseeable consequences, one might think that the legal concept of reasonable care (and negligence) *is* outcome-sensitive. And in some sense, of course, it is: Holding probability fixed, it is considerably more unreasonable to embark on an action where a human life is at risk than an action that might lead to a broken finger. Importantly, however, the outcome-sensitivity thus understood regards *possible* outcomes. In the main text, when I write about outcome-dependence, I am interested in the influence of *actual* outcomes – whether or not the possible bad outcome does come to pass – on the assessment of the defendant’s behaviour. Differently put, when I write about the “outcome-dependence of reasonableness”, what is meant is not the *ex ante*, or epistemic, reading thereof, but the *ex post*, and thus partially metaphysical understanding of “outcome-dependence”. Thanks a lot to Kevin Tobia and Morris Hoffman for pointing out the need for clarification in this regard. Thanks to Kevin Tobia for drawing my attention to the passage from Tort law.

<sup>19</sup> On moral luck in general, see Nagel, 1979, Williams, 1993, Hartman 2017. With a view to the law, see Enoch, 2007, 2010; Enoch & Marmor, 2007, for empirical work, see Martin & Cushman, 2016 and Kneer & Machery, 2019.

consequences. But if reasonableness, on its legal understanding, were outcome-dependent, then the reasonable person standard would just be strict liability in fanciful disguise: No matter how reasonable the decision *ex ante*, if the defendant is unlucky and the consequences dire, her decision or action is likely to be judged unreasonable *ex post*.

#### **1.4 Identifying the Problem**

The legal concept of reasonableness is outcome-insensitive – what is, and is not reasonable, must be evaluated with regards to the context of decision or action, not post hoc, at a later context when the consequences are known. Preliminary empirical findings, however, suggest that folk judgments of reasonableness are inconsistent with the legal concept of reasonableness. People’s assessments of what is reasonable depend strongly on outcomes. This, we said, might be the consequence of one of two more general potential problems: On the one hand, there might be conceptual alignment between the legal and the folk understanding of reasonableness, yet folk judgments are distorted by an outcome bias. To put it differently: As presumed by lawmakers and legal practitioners, the legal meaning of “reasonable” is the same as the ordinary language meaning of the expression. But despite conceptual alignment, people might find it difficult to disregard outcomes. An effect of this sort is also found, for instance, when it comes to the ascription of causation and *mens rea* in between-subjects experiments (Alicke, 1992, 2000; Alicke & Rose, 2010; Nadelhoffer, 2006; Kneer & Bourgeois-Gironde, 2017, Bourgeois-Gironde & Kneer, 2017). Hence, if reasonableness judgments were distorted, we should work towards a better understanding of their underlying “mechanics”, so as to devise means to alleviate the bias.

An alternative to the bias view is the following: The law might wrongly assume conceptual alignment. Whereas the law operates with an outcome-insensitive and dynamically invariant concept of reasonableness, the folk concept might simply differ – without the outcome-sensitivity of “reasonable” constituting a bias. The folk concept of punishment, for instance, is strongly sensitive to outcomes (Cushman, 2008), without this clearly suggesting that anything is amiss: People tend to judge deserved punishment differently across outcomes *even* in case in which they assess blame and wrongness identically (Kneer & Machery, 2019). If this possibility were born out, the law, with its perpetual insistence that what is reasonable just is what the “man on the Clapham omnibus” or “the average Joe” takes it to be, would be completely mistaken, because the women and men on the Clapham bus in fact have a different concept of reasonableness than what the law *takes* them to have. The law, that is to say, would unwittingly operate with a *technical* concept of reasonableness, yet assume that the legal concept is the non-technical folk concept.

#### **1.5 Structure of the Paper**

The paper proceeds as follows: In section 2 I will say a few more words about extant research regarding outcome effects in empirical legal scholarship and experimental jurisprudence.<sup>20</sup> Section

---

<sup>20</sup> Experimental jurisprudence is a young discipline which explores philosophical topics regarding the law by empirical means. See e.g. Kneer & Bourgeois-Gironde (2017), Bourgeois-Gironde & Kneer (2018), Prochownik et

3-5 report three experiments based on the well-known *TJ Hooper* torts negligence case, as well as a medical malpractice scenario. The results suggest that the folk concept of reasonableness is outcome-sensitive, dynamically variant across *ex ante* and *ex post* contexts of assessment and that these effects are not due to the hindsight bias. Section 6 concludes by arguing that, out of the two threats to trial by jury we are faced with conceptual misalignment. It is thus of utmost practical importance for the law to come to understand that it is operating with a concept of reasonableness that differs radically from the folk concept and to take precautions against the likely injustice this generates on a daily basis.

## 2. Outcome Effects

### 2.1 Intention and Knowledge

Culpability in Western criminal law requires both an *actus reus* (a “guilty act”) and a *mens rea* (a “guilty mind”). Since neither a bad intention nor a harmful accident, by themselves, engender criminal responsibility (Coke’s law), the *actus reus* and the *mens rea* must be established independently and treated as conceptually distinct. Findings from psychology and experimental philosophy raise a worry for this neat picture of two independent constitutive judgments feeding into culpability assessment: Outcome is known to influence all-things-considered judgments such as blame, culpability and punishment. It might thus be the case, that certain features of the *actus reus* – such as outcome valence and severity – affect judgments as to whether a defendant acted with a guilty mind.

There is, in fact, ample evidence that this is precisely the case at the more inculpatory spectrum of the *mens rea* catalogue. Outcome *valence* (good v. bad) has an impact on attributions of intentionality (the Knobe effect, see Knobe, 2003a, 2003b; for reviews, see Feltz, 2007, Cova, 2016) as well as knowledge and belief (the epistemic side-effect effect, Beebe & Buckwalter, 2010; Beebe & Jensen, 2012). Furthermore, there is considerable evidence that the *severity* of harm correlates positively with *mens rea* ascription. For instance, the more severe an undesired, yet knowingly occasioned harm, the higher the likelihood that people will judge it intentional (Kneer & Bourgeois-Gironde, 2017; Garcia-Olier & Kneer, in prep, Prochownik et al. 2020, Tobia, 2020) or that they will judge it as known or foreseen (Kneer et al., in prep).

These findings are of importance for common law countries, where the *mens rea* question in criminal trials is frequently decided by lay juries. The situation might be similarly problematic in countries relying exclusively (or near-exclusively) on professional judges: The Knobe effect and severity effects have also been found to arise among professional judges, lawyers, and law students across several countries (Kneer & Bourgeois-Gironde, 2017 & Bourgeois-Gironde & Kneer, 2017

---

al. (2020) and Tobia (2020a) on *mens rea* attribution, MacLeod (2019) and Knobe & Shapiro (2020) on causation, Lidén et al. (2019) on the confirmation bias in guilt assessments, Strohmaier et al. (2020) on the hindsight bias in the law, Donelson & Hannikainen (2020) on Fuller’s theory of law, Struchiner et al. (2020) on the Fuller/Hart debate, or Tobia (2020b) on ordinary meaning and legal interpretation to name but a few.

for French judges and lawyers respectively; Prochownik et al. 2020 for German law students; Kneer et al., in prep., for professionals in the UK, Poland, the Netherlands and Colombia, the Knobe effect – though not severity effect – for US professionals, Tobia, 2020).

## ***2.2 Recklessness, Negligence and the Reasonable Person Standard***

The reasonable person standard features prominently in the lower echelons of the *mens rea* catalogue. An agent is criminally liable when grossly deviating from standards of reasonable care, be it knowingly (thus recklessly) or unwittingly (thus negligently).<sup>21</sup> Although the topic has not been explored much in the empirical literature, there is at least some evidence that negligence attributions are outcome-sensitive (Kamin & Rachlinski, 1995, Kneer & Machery, 2019; Frisch et al. in prep). In between-subjects designs, people are more inclined to ascribe negligence to an unlucky, well-intentioned agent whose actions lead to a bad outcome, than to a lucky agent, whose otherwise identical actions do not. Interestingly, the outcome effect on negligence is mediated by perceived probability: People deem a negative outcome more likely in retrospect than a neutral one and thus tend to be more willing to ascribe negligence in the former than in the latter case. This finding dovetails with Fischhoff's (1975) hindsight bias, i.e. the change in perceived probabilities of events as they unfold. Naturally, this is bad news for the law (cf. Rachlinski, 1998, Hoffman, ms.; Strohmaier et al. 2020). As the above-quoted passage from the Model Penal Code (2.02 (d)) makes clear, the law cares about the probabilities, and hence the risks, at the context of action, not some later context when the consequences are known. But if perceptions of probabilities and risks are outcome-dependent, it comes as no surprise that evaluations of negligence and recklessness – in contrast to what the law requires – are, too.

The experiments presented below aspire to break new ground in several ways: First, and in contrast to Kamin & Rachlinski (1995) and Kneer & Machery (2019), the dependent variable of Experiment 1 (and the follow-up experiments) is *reasonableness per se*, not merely a related variable, negligence or probability. The implications of the results are thus considerably broader in scope, since the reasonable person standard is not limited to negligence cases. Second, Experiment 2 explores whether the outcome-dependence of reasonableness judgments is a fundamental feature of the folk-concept of *reasonable*, or a bias, as is standardly simply assumed. Third, Experiment 3 investigates whether the outcome-dependence of reasonableness judgments is but a consequence of the hindsight bias. This is done by holding perceived probability fixed across outcomes, so as to explore whether a direct effect of outcome on reasonableness remains under well-controlled circumstances. In sum, the paper aspires to make considerable progress not only as to whether outcome influences attributions of reasonableness, but what explains the outcome dependence, how it is to be interpreted and what all this means for the law.

---

<sup>21</sup> Cf. MPC 2.02 (c,d), and Dressler, 2015 §10.04.D, see also Homicide Act 1957 s3 and Gardner (2015). For discussion of the reasonable person standard with respect to negligence, see e.g. Fletcher (1970), Hurd & Moore (2002), Moore & Hurd (2011), Alexander, Ferzan & Morse (2009) and Stark (2016).



### 3. Experiment 1

The first experiment employs a between-subjects design to explore whether people deem a decision more reasonable *ex post* if the consequences are positive than when they are negative (holding all else fixed across conditions). The scenario is based on a famous negligence case, *The TJ Hooper*.<sup>22</sup> In that case, the plaintiff lost two barges and their cargo in a storm, claiming that this was due to the fact that the defendant's tugboats – the *TJ Hooper* and the *Montrose* – did not have reliable radio equipment, which could have warned them of the storm.

#### 3.1 Participants

I recruited 179 participants on Amazon Mechanical Turk. The IP address was restricted to the United States. As preregistered,<sup>23</sup> I excluded participants failing an attention check and those who responded to the first question (which required reading the scenario) in less than 15 seconds. 159 participants remained (70 female, age M=43 years, SD=14 years).

#### 3.2 Methods and Materials

Participants were presented with one of two conditions of the following scenario. The two conditions differed only in terms of outcome (positive v. negative; labels in bold omitted):

Mr Garcia runs a flourishing logistics business with two dozen cargo jets. Until now, he has exclusively focused on air transport. He decides to expand to shipping and buys five cargo ships.

For ships, there is a large variety of security systems available. They track other ships, make detailed weather predictions, and minimize risks for ships in distress. Garcia decides to get a package deal from a provider who installs several different types of maritime security systems on his cargo freighters.

The provider offers different plans for servicing the security systems and updating the software. There are four options: A monthly check, a bimonthly check, a check every six months, and a plan where the systems are only checked and updated once a year. The yearly plan is by far the cheapest, the more regular the checks get, the more expensive the plan.

Out of the four options, Garcia decides to take the second-cheapest. The security systems on his boats are thus checked and updated every six months.

**Positive outcome:** Garcia's cargo shipping business has many clients. 15 months after commencing operations, there has not been a single security incident on any of his cargo freighters.

---

<sup>22</sup> The T.J. Hooper, 53 F.2d 107 (S.D.N.Y. 1931), for discussion see e.g. Epstein (1992).

<sup>23</sup> <https://aspredicted.org/blind.php?x=ke2ia5>

**Negative outcome:** Garcia's cargo shipping business has many clients. 15 months after commencing operations, one of Garcia's cargo ships crashes into a rock due to the malfunctioning of one of its security systems. The ship sinks. The 40 tons of crude oil it was transporting are spilled into the ocean. The damage to the environment is very serious.

Participants were asked to what extent they agreed with the claim "It was reasonable for Garcia to choose the six-month service plan." Responses were collected on a 7-point Likert scale anchored at 1 (completely disagree) and 7 (completely agree). Next, participants were asked "[h]ow much praise or blame, if any, does Garcia deserve for choosing the six-month servicing plan?". The 7-point Likert scale was anchored at 1 (lots of praise) and 7 (lots of blame), the midpoint 4 being labelled "neither praise nor blame".

### 3.3. Results

There was a pronounced effect of outcome valence on judgments of reasonableness *ex post*. In the positive outcome condition, where there has not been a single security incident, mean reasonableness ( $M=5.24$ ,  $SD=1.43$ ) was significantly higher than in the negative outcome condition, in which one of the ships crashes ( $M=3.96$ ,  $SD=1.88$ ),  $t(1,149)=4.70$ ,  $p<.001$ , Cohen's  $d=.76$ , a medium-large effect). As regards blame and praise, the situation was reversed. Mean blame/praise ascriptions in the negative outcome condition ( $M=4.92$ ,  $SD=1.36$ ) significantly exceeded mean blame/praise in the positive outcome condition ( $M=3.69$ ,  $SD=1.05$ ),  $t(1,149)=-6.24$ ,  $p<.001$ , Cohen's  $d=1.01$ , a large effect). Figure 1 represents the results graphically.

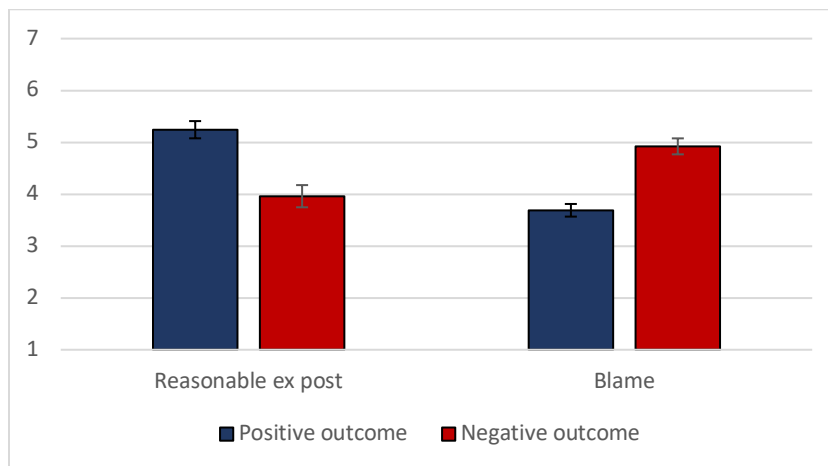


Figure 1: Mean reasonableness *ex post* and mean blame/praise attributions across conditions (positive v. negative outcome). Error bars denote standard error of the mean.

### 3.4 Discussion

Consistent with previous findings (Baron & Hershey, 1988, Kneer, in prep. a), outcome had a significant and pronounced effect on judgments of reasonableness and blame. The results raise a

serious worry: The law employs an outcome-insensitive concept of reasonableness, but people's *post hoc* evaluations of whether an agent's decisions *were* reasonable are strongly sensitive to outcomes. Given these results, the question arises which of the two threats to trial by jury discussed in the introduction is at stake. Are we facing a *performance error* in the application of a folk concept of reasonableness that is identical to the legal one in so far as it is outcome-insensitive (if correctly employed)? Or might there be *tacit conceptual divergence*, i.e. might it be the case that the folk concept of reasonableness just *is* outcome-sensitive, and thus differs from the legal concept?

Just like our first experiment, the little work that touches upon the reasonable person standard in existence cannot answer this question. Kamin and Rachlinski (who test for probability and negligence rather than reasonableness more generally) ran a between-subjects design, Baron & Hershey (1986) presented participants with a whole battery of scenarios with good and bad outcomes. In order to better understand what drives the outcome effect in assessments of reasonableness *ex post*, and thus to identify which of the two threats to jury trial is at stake, a different design is required: We should explore whether people, when assessing a decision or action *ex ante* – i.e. before the consequences are known – are willing to *change* their evaluations of reasonableness *ex post* once the consequences have been revealed. If people, in the novel design, assess reasonableness by and large the same *ex ante* and *ex post*, this would indicate that their concept of reasonableness is, like the legal one, outcome-independent. The fact that, in between-subjects designs, we find a pronounced outcome effect could thus be interpreted as a performance error or bias. If, on the other hand, people change their views of whether an action was reasonable *ex post* in contrast to what they judged *ex ante* on the basis of the outcome, then the folk concept is simply not what the law presumes it to be. In that case, the folk concept of reasonableness (*qua* folk concept) is *appropriately* sensitive to outcome, and the outcome effect in Experiment 1 the consequence of its proper application.

#### **4. Experiment 2**

The goal of Experiment 2 is to investigate whether the outcome-sensitivity of reasonableness judgments *ex post* constitutes a performance error or else, whether the folk concept of reasonableness is simply outcome-sensitive (in contrast to the legal concept). If people judge reasonableness *ex ante* and *ex post* the same, we would have evidence in favour of the former hypothesis. If people's *ex post* judgments of reasonableness diverge from their *ex ante* judgments depending on outcome, we would have evidence in favour of the latter hypothesis.

##### **4.1 Participants**

I recruited 181 participants on Amazon Mechanical Turk. The IP address was restricted to the United States. As preregistered,<sup>24</sup> I excluded participants failing an attention check and those who

---

<sup>24</sup> <https://aspredicted.org/blind.php?x=hp8zx4>

responded to the first question (which required reading the scenario) in less than 15 seconds. 148 participants remained (71 female, age  $M=43$  years,  $SD=13$  years).

#### **4.2 Methods & Materials**

The materials were similar to those in Experiment 1. The main difference was that, once Garcia's general situation has been explained, participants are merely told that Garcia *considers* choosing the second-least expensive security plan, according to which the boats are serviced every six months. They then had to state whether they agree or disagree with the claim that "[i]t is reasonable for Garcia to choose the six-month service plan" on a 7-point Likert scale. Differently put, participants rated reasonableness at the context of decision, i.e. reasonableness *ex ante*. On the next screen, participants were told that Garcia in fact chose the six-month plan, and they were presented with either the positive outcome (smooth sailing) or negative outcome (one of the boats crashes). They were then asked to what extent they agree or disagree with the claim that it was reasonable for Garcia to choose that plan (7-point Likert scale), and to what extent they would attribute praise or blame to him for deciding as he did (7-point Likert scale, for details, see Appendix). The mixed design experiment thus allows us to contrast perceived reasonableness *ex ante* and *ex post* for each of the two conditions, and reasonableness *ex post* and blame across outcomes.

### **3.3. Results**

#### **3.3.1 Within-subjects Results**

In this experiment, people rated the reasonableness of Garcia's decision both before the outcome was revealed, and once again once outcomes had been presented. The purpose of the experiment was to explore whether, on average, people change their assessment of the reasonableness of a decision in light of its consequences once the latter are revealed. This was indeed the case for both outcomes. In the positive outcome condition, people deemed it significantly more reasonable for Garcia to choose the six-month plan *ex post* ( $M=5.40$ ,  $SD=1.53$ ) than they did *ex ante* ( $M=5.12$ ,  $SD=1.62$ ), paired-samples t-test,  $t(1,77)=-2.58$ ,  $p=.012$ , Cohen's  $d=.29$ , a small effect, see Figure 2. In the negative outcome condition, people deemed it significantly less reasonable to choose the six-month plan *ex post* ( $M=4.09$ ,  $SD=1.99$ ) than they did *ex ante* ( $M=5.43$ ,  $SD=1.48$ ),  $t(1,69)=6.38$ ,  $p<.001$ , Cohen's  $d=.76$ , a medium-large effect).

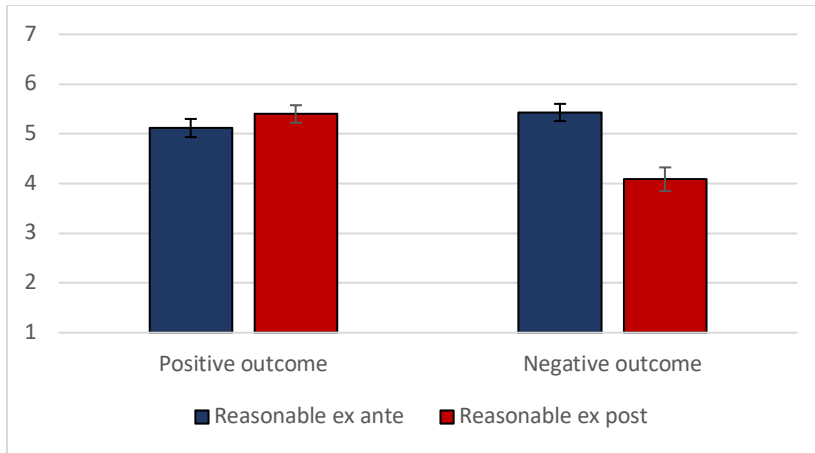


Figure 2: Mean reasonableness *ex ante* and *ex post* across outcomes. Error bars denote standard error of the mean.

### 3.3.2 Between-subjects Results

The first part of the scenario (at which the outcome is still withheld) was identical across conditions. Consequently, assessments of reasonableness *ex ante* should not differ, which is indeed what we find (independent samples t-test,  $t(1,146)=-1.22, p=.223$ , two-tailed).

Participants across the different conditions then viewed the distinct outcomes on a second screen, and had to assess reasonableness *ex post*. The impact of outcome on reasonableness was significant,  $t(1,146)=4.52, p<.001$ , Cohen's  $d=.74$ , a medium-large effect, see Figure 3. The same held for blame/praise,  $t(1,146)=-6.43, p<.001$ , Cohen's  $d=1.06$ ).

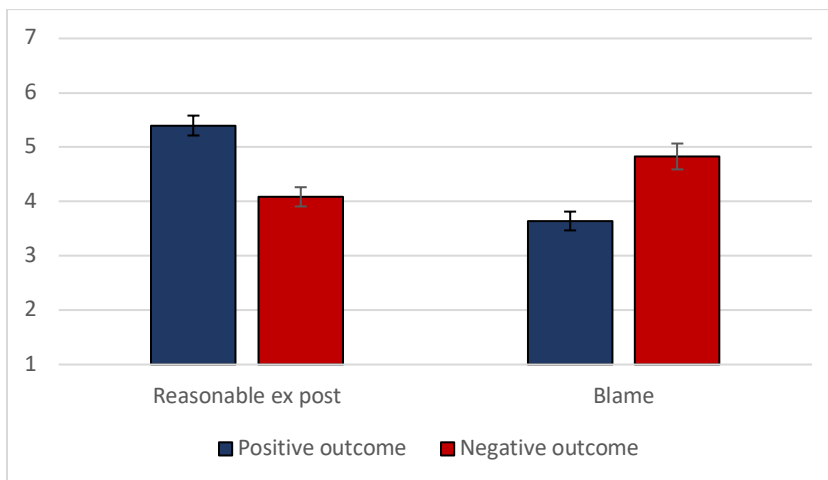


Figure 3: Mean reasonableness *ex post* and mean blame/praise attributions across conditions (positive v. negative outcome). Error bars denote standard error of the mean.

### 3.3.3 Comparison across Designs

The data collected so far allows a further analysis: We can explore whether reasonableness assessments *ex post* differ across subjects who first rated reasonableness *ex ante* (a sort of priming

technique for debiasing purposes, results from Experiment 2) and subjects exclusively confronted with questions of reasonableness *ex post* (Experiment 1). The same analysis can be conducted for blame. A 2 (outcome: positive v. negative) x 2 (design: *ex ante* assessment or priming v. no *ex ante* assessment or no priming) ANOVA reveals the expected significant main effect for outcome ( $F(1,295)=42.43, p<.001$ ). Importantly, though, there was neither a significant main effect for experiment design ( $F(1,295)=.49, p=.484$ ) nor a significant outcome\*design interaction ( $F(1,295)=.005, p=.941$ ). For blame, the pattern of results was the same: a significant main effect for outcome ( $F(1,295)=79.94, p<.001$ ), a nonsignificant main effect for design ( $F(1,295)=.27, p=.601$ ) as well as a nonsignificant interaction ( $F(1,295)=.01, p=.941$ ). This is quite remarkable: Priming people by aid of an *ex ante* assessment of the reasonableness of Garcia's decision – and thus potentially immunizing them against an outcome effect or hindsight bias – had *no effect whatsoever*. Note that the probability of finding a mid-size effect of  $f=.25$  with the recruited sample size of  $N=299$  across the two experiments and an alpha of .05 exceeds 99% to find such an effect, if there is one to be found (power analysis conducted with G\*Power 3.1).

#### 4.4 Discussion

Consistent with the results of Experiment 1, and broadly in tune with those of Baron & Hershey (1988) and Kamin & Rachlinski (1995), we found that ascriptions of reasonableness (or related variables) are outcome-dependent. None of these previous studies, however, address the central question of whether the outcome effect testifies to a biased ascription of an outcome-insensitive concept of reasonableness, or whether the folk-concept of reasonableness simply is outcome-dependent (and thus correctly applied). The design of Experiment 2 sheds some light on the matter, because participants judge Garcia's decision *both ex ante and ex post*. If the outcome effect on reasonableness were a bias, then people would presumably not change their assessment once the consequences have been revealed. Interestingly, they do: People judge a decision that leads to negative consequences as less reasonable *ex post* than *ex ante*, and a decision that leads to positive consequences as more reasonable *ex post* than *ex ante*. Furthermore, the two-step procedure (first *ex ante*, then *ex post* assessment) does not mitigate the outcome effect on reasonableness *ex post* at all, as our comparison between Experiment 1 (only *ex post* assessments) and Experiment 2 (assessment both *ex ante* and *ex post*) reveal. Together, the results suggest that the folk concept of reasonableness is genuinely outcome-dependent, and thus diverges radically from the legal concept, and from *what the law assumes about the folk concept*.

There is a final worry. Kamin & Rachlinski (1995) as well as Kneer & Machery (2018) have argued that the outcome effect on negligence is due to an outcome effect on perceived probability *post hoc*. In other words, the outcome effect on negligence is mediated by perceived probability *post hoc*. A similar hypothesis is possible for reasonableness more broadly conceived: The fact that the probability of harm *post hoc* is judged higher when harm actually does occur than when it does not might lead people to evaluate a decision that leads to harm as less reasonable than one that does not. Whereas an outcome effect on the perceived probability of the sort invoked here is

standardly interpreted as a *bias*,<sup>25</sup> a *rationalist* interpretation is also possible.<sup>26</sup> Especially in cases where probabilities are fuzzy and hard to assess, it makes sense to use outcomes as an indicator of an event's probability: Whether a decision or action *did lead* to a particular outcome or not constitutes useful information for the assessment (and potential *reassessment*) of whether it was *likely that it would lead* to that sort of outcome. While this is an interesting question to pursue in further detail, our final experiment focuses on an even more fundamental one: Is there, at least sometimes, a *direct* or *unmediated* effect of outcome on reasonableness *ex post*? In other words, can an outcome effect on reasonableness *ex post* be observed even if we were to hold perceived probability *ex post* constant across conditions? If so, this would show that the outcome-sensitivity of reasonableness cannot be reduced to a hindsight effect – whether the latter is best interpreted as a bias or not.

### 5. Experiment 3

The goal of the final experiment consists in exploring the effect of outcome on reasonableness in situations where the risk can be clearly quantified *ex ante*. The vignette explicitly specifies the probability of a harmful outcome, which, paired with a manipulation check focusing on probability, allows us to minimize the influence of the hindsight effect on reasonableness attributions. That way, we can explore whether the outcome-effect on reasonableness is tied to perceived probability, or whether outcome can influence perceived reasonableness directly.

The vignette itself is based on a duty of care case – a type of case where reasonableness also plays a central role. Manufacturers, for instance, must exercise reasonable care in providing consumers with safe products, doctors are under the obligation to prevent reasonably foreseeable harm to their patients. In cases of this sort, it is once again up to the fact-finder to decide what, exactly, constitutes a reasonable measure of care and what does not.

#### 5.1 Participants

I recruited 219 participants on Amazon Mechanical Turk. The IP address was restricted to the United States. As preregistered,<sup>27</sup> I excluded participants failing an attention check and those who responded to the first question (which required reading the scenario) in less than 15 seconds. As also preregistered (and detailed below), I also excluded all participants who, after having learned the outcome, did not report the precise probability of a harmful consequence, which was specified explicitly in the scenario. This ensured that the outcome effect on reasonableness *ex post* was *not*

---

<sup>25</sup> For instance by Kamin & Rachlinski, 1995, as well as nearly all the papers on the “hindsight bias” surveyed by Rachlinski, 1998. Note how rare it is to read about a hindsight *effect* rather than a hindsight *bias*. The problem is discussed in more detail by Kneer & Machery (2018), who explore the rationalistic interpretation of the hindsight effect at length.

<sup>26</sup> I'd like to thank Joshua Knobe for pressing me on this important point.

<sup>27</sup> <https://aspredicted.org/blind.php?x=ng28n4>

driven by the hindsight bias (i.e. a diverging assessment of probability *ex ante* v. *ex post*). 125 participants remained (55 female, age M=41 years, SD=13 years).

## **5.2 Methods and Materials**

Like Experiment 2, Experiment 3 also took a 2 (between-subjects factor *outcome*: positive v. negative) x 2 (within-subjects factor *reasonableness*: *ex ante* v. *ex post*) design. All participants were presented with the following first part of the vignette and had to rate the doctor's decision to prescribe a certain drug on a 7-point Likert scale:

Ms Woods goes to the hospital since she has been experiencing severe abdominal pain for several days. Dr Miller examines her carefully and concludes that she has an inflammation of the intestines.

Dr Miller is thinking about prescribing a new drug called Fluxis, which has proven very effective with the kinds of inflammation Ms Woods has. As most drugs, it has some possible side-effects. According to recent studies, there's a 3% probability that Fluxis will cause damage to the liver.

Q: To what extent do you agree or disagree with the following claim:

"It is reasonable for Miller to prescribe Fluxis." (1: completely disagree – 7: completely agree)

Having responded to the reasonableness *ex ante* question, participants were randomly assigned to one of the two outcomes (labels in square brackets omitted):

**[Good outcome]** Dr Miller prescribes the drug Fluxis to Ms Woods. The drug works perfectly. The inflammation is cured within days, and Ms Woods has no side-effects whatsoever.

**[Bad outcome]** Dr Miller prescribes the drug Fluxis to Ms Woods. The drug does not work as hoped. The inflammation persists and Fluxis causes serious damage to Ms Woods' liver.

Q1: To what extent do you agree or disagree with the following claim:

"It was reasonable for Miller to prescribe Fluxis." (1: completely disagree – 7: completely agree)

Q2: How much praise or blame, if any, does Miller deserve for prescribing Fluxis to Ms Woods? (1: lots of praise, 4: neither praise nor blame, 7: lots of blame).



Besides responding to the two main questions, participants also had to answer the following question regarding probability: “On a scale of 0% (completely unlikely) to 100% (certain), how high was the probability that Fluxis would cause damage to the liver?”. As explained above, this question served as a manipulation check. In order to explore the impact of outcome on reasonableness *independently* of a hindsight effect, those participants who did not respond that the probability of liver damage was 3% (as specified in the vignette) were excluded.

### 5.3 Results

#### 5.3.1 Within-subjects Results

In the positive outcome condition, people deemed it significantly more reasonable to prescribe the drug *ex post* ( $M=5.98$ ,  $SD=1.38$ ) than they did *ex ante* ( $M=5.45$ ,  $SD=1.40$ ), paired-samples t-test,  $t(1,61)=-4.40$ ,  $p<.001$ , Cohen’s  $d=.56$ , a medium-sized effect). In the negative outcome condition, people deemed it significantly less reasonable to prescribe the drug *ex post* ( $M=4.51$ ,  $SD=1.76$ ) than they did *ex ante* ( $M=5.48$ ,  $SD=1.27$ ),  $t(1,62)=7.04$ ,  $p<.001$ , Cohen’s  $d=.89$ , a large effect), see Figure 4.

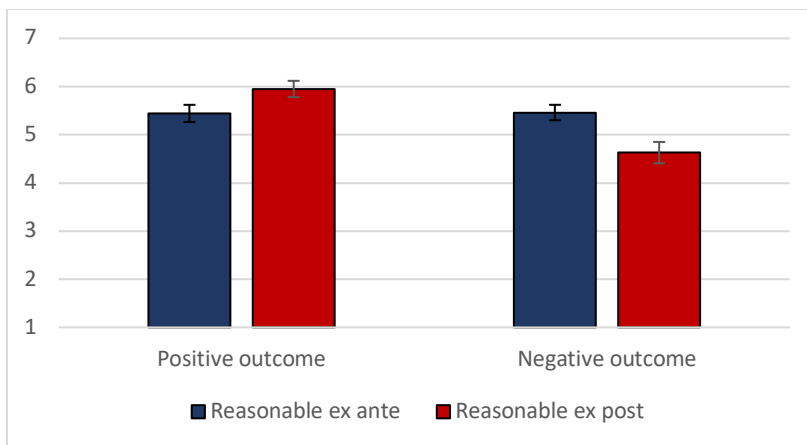


Figure 4: Mean reasonableness *ex ante* and *ex post* across outcomes. Error bars denote standard error of the mean.

#### 5.3.2 Between-subjects Results

Across conditions, outcome had a significant impact on reasonableness *ex post*, independent-samples t-test  $t(1,123)=5.28$ ,  $p<.001$ , Cohen’s  $d=.94$ , a large effect.<sup>28</sup> The same held true for blame,  $t(1, 123)=-6.06$ ,  $p<.001$ , Cohen’s  $d=1.08$ , a large effect), see Figure 5.<sup>29</sup>

<sup>28</sup> Unsurprisingly, there is no significant difference across conditions for reasonableness *ex ante* ( $t(1,123)=-.103$ ,  $p=.918$ ), since the two conditions are – at that point – still identical (outcomes have not yet been revealed).

<sup>29</sup> As expected, all contrasts remain strongly significant when one leaves those participants who did not respond with the exact probability of a harmful outcome in the dataset ( $N=157$ ). There was no significant effect of outcome on perceived probability *post hoc*,  $t(1,155)=-.715$ ,  $p=.476$ ). This suggest that, in explicitly specifying the probability in the scenario, we managed to limit its potential effect even without excluding all those who participants who did not report it correctly.

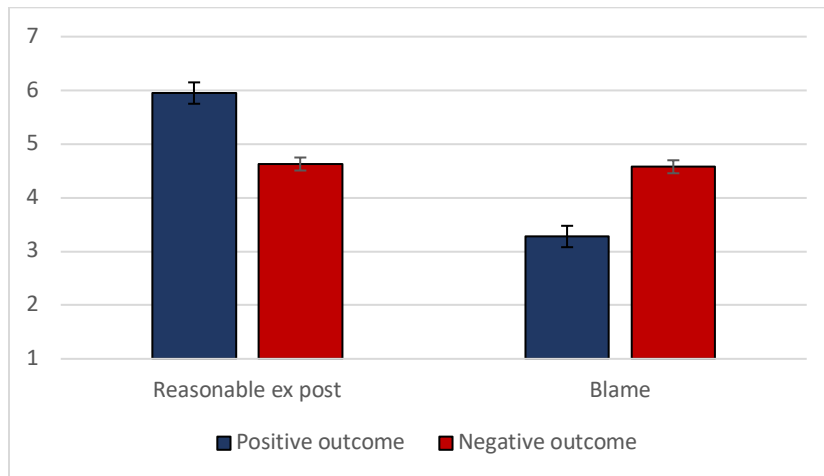


Figure 5: Mean reasonableness *ex post* and mean blame/praise attributions across conditions (positive v. negative outcome). Error bars denote standard error of the mean.

#### 5.4 Discussion

Replicating the between-subjects results from Experiments 1 and 2, we found that decisions leading to a negative consequence are judged less reasonable *post hoc* than those that lead to positive consequences. Replicating the within-subjects results from Experiment 2, people judge the reasonableness of a decision producing harm as less reasonable *ex post* than *ex ante* (before the outcome was known), and *vice versa* for positive outcomes. These findings suggest that the folk *application* of the concept of reasonableness is not subject to a performance error, but that the folk work with a concept of reasonableness which *is* outcome-dependent. Is this outcome-dependence driven by a hindsight effect, as some scholars like Kamin and Rachlinski (1995) suggest? Perhaps, in certain situations, this might *also* be the case. However, even when one carefully controls for potential hindsight bias, we find a strong *direct* effect of outcome on reasonableness.

#### 6. Conclusion

In their capacity as fact-finders determining what is, and is not, reasonable, jurors, we said, must have two types of competences: They must have *concept* competence, i.e. understand and use the legal concept of reasonableness. And they must possess *application* competence, that is, they must not fall prey to cognitive biases when deciding what is reasonable and what is not.

The legal concept of reasonableness is *outcome-insensitive* and *dynamically invariant*. In other words, whether a decision produces a positively or negatively valenced outcome must not have an impact on whether it is judged reasonable, and a decision that is reasonable *ex ante* remains reasonable *ex post*, even if it produces undesirable consequences. In line with previous, vaguely related studies, however, all three experiments suggest that folk attributions of reasonableness are strongly sensitive to outcome. In principle, this could be due to one of two reasons: *First*, if – as

assumed by the law – the legal concept of reasonableness just *is* the folk concept of reasonableness, then the potential jurors fall prey to an outcome bias, as they tend to also do in the attribution of *mens rea*. Or else, *second*, folk ascriptions might actually be true to the folk concept of reasonableness – a concept that is very different from the legal concept of reasonableness and from what the law takes the folk concept to be. Experiments 2 and 3 suggest that this is in fact the case: Even in within-subjects designs, people find it perfectly perfectly appropriate to revise their *ex ante* assessments of reasonableness *ex post*. It thus appears that outcome-sensitivity is built into the ordinary language expression “reasonable” and the concept it denotes. The law seems to wrongly assume conceptual alignment and unwittingly operate with a technical (outcome-insensitive) concept of reasonableness.

Hardnosed advocates of the bias interpretation could, broadly in line with Kamin and Rachlinski (1995), object to our preliminary conclusion in the following way: What is, and is not reasonable is (at least frequently) tied to the notion of *risk*, i.e. the product of *probability* and severity of outcome. Given that probability assessments are subject to the exceptionally robust and pervasive hindsight bias, the variation of reasonableness assessed *ex ante* vs. *ex post* might simply be a consequence of distorted probability judgments. The findings from Experiment 3, however, suggest that this objection is on the wrong track.<sup>30</sup> Even if we hold perceived probability fixed across the *ex ante* context of decision and the *ex post* context of assessment on the one hand, as well as across differently valenced outcomes on the other, our results from Experiments 1 and 2 replicate perfectly. Naturally, this leaves open the possibility that the outcome sensitivity and dynamic variance of reasonableness is *exacerbated* by the hindsight bias in the many situations where probabilities are unknown or hard to quantify. However, it demonstrates that the outcome-sensitivity and dynamic variance of reasonableness does not constitutively depend on perceived probability. The folk concept seems to be directly sensitive to outcomes.

Over the last two decades, considerable effort has been devoted to a better understanding of what the legal concept of reasonableness does and should mean, and how potential pitfalls in the application of the reasonable person standard can be avoided. Perhaps the subjective/objective and the normative/descriptive debates have received the most attention, and they are, without a doubt, of great practical importance. The potential outcome-sensitivity of reasonableness ascriptions, by contrast, has received next to no attention, and there is no literature on the dynamic sensitivity of reasonableness at all. This constitutes a serious shortcoming because there is *no* legal case invoking the reasonable person standard where these two features do *not* matter. It is thus hard to overstate the practical importance of the present findings and more research – both theoretical and empirical – is required to explore how conceptual alignment can be achieved so as to prevent further legal injustice due to the law’s misunderstanding of the folk concept of reasonableness.

---

<sup>30</sup> It does not matter whether the effect of hindsight on perceived probability is viewed as a *bias* or as a *rational heuristic* (as discussed in Section 4.4). The point here is simply that the effect of outcome on reasonableness is, at least in parts, independent of probability judgments (no matter how they are interpreted).

## Bibliography

- American Law Institute (1982). Model Penal Code, annotated.
- Adams, F. (2015). The Knobe-effect and the law. *Method-Analytic Perspectives*, 4(6)
- Armour, J. D. (1994). Race ipsa loquitur: Of reasonable racists, intelligent Bayesians, and involuntary Negrophobes. *Stanford Law Review*, 781-816.
- Alexander, L., Ferzan, K. K., & Morse, S. J. (2009). *Crime and culpability: A theory of criminal law*. Cambridge University Press.
- Alicke, M. D. (1992). Culpable causation. *Journal of personality and social psychology*, 63(3), 368.
- Alicke, M. D. (2000). Culpable control and the psychology of blame. *Psychological bulletin*, 126(4), 556.
- Alicke, M., & Rose, D. (2010). Culpable control or moral concepts?. *Behavioral and Brain Sciences*, 33(4), 330.
- Baron, Marcia (2011) "The Standard of the Reasonable Person in Criminal Law". in R.A. Duff, Lindsay Farmer, S.E. Marshall, Massimo Renzo, and Victor Tadros (eds.) *The Structures of Criminal Law* Oxford University Press, pp.11-36.
- Baron, J., & Hershey, J. C. (1988). Outcome bias in decision evaluation. *Journal of personality and social psychology*, 54(4), 569.
- Bourgeois-Gironde S. & M. Kneer (2018). Intention, cause, et responsabilité: *Mens Rea* et effet Knobe. In S. Ferey & G'Sell F. (ed.), *Causalité, responsabilité et Contribution à la dette de réparation. Une Perspective interdisciplinaire*. Bruylant.
- Beebe, J. R., & Buckwalter, W. (2010). The epistemic side-effect effect. *Mind & Language*, 25(4), 474-498.
- Beebe, J. R., & Jensen, M. (2012). Surprising connections between knowledge and action: The robustness of the epistemic side-effect effect. *Philosophical Psychology*, 25(5), 689-715.
- Berger, S. R. (1977). Court Awarded Attorneys' Fees: What is Reasonable. *U. Pa. L. Rev.*, 126, 281.
- Cova, F. (2016). The folk concept of intentional action: Empirical approaches.
- Cushman, F. (2008). Crime and punishment: Distinguishing the roles of causal and intentional analyses in moral judgment. *Cognition*, 108(2), 353-380.
- Donelson, R. and Ivar R. Hannikainen (2020) Fuller and the Folk: The Inner Morality of Law Revisited. In (Lombrozo, T., Knobe, J. and Nichols, S.) *Oxford Studies in Experimental*. Oxford University Press (2020).
- Dresser, R. (1993). Culpability and Other Minds. *S. Cal. Interdisc. LJ*, 2, 41.
- Dressler, J. (1995). *Understanding criminal law*. Matthew Bender.
- Duff, A. (2015). Intention, intentional action and the law. *Method-Analytic Perspectives*, 4(6), 136-146.
- Enoch, D. (2007). Luck between morality, law, and justice. *Theoretical Inquiries in Law*, 9(1), 23-59.
- Enoch, D. (2010). Moral luck and the law. *Philosophy Compass*, 5(1), 42-54.
- Enoch, D., & Marmor, A. (2007). The case against moral luck. *Law and Philosophy*, 26(4), 405-436.
- Epstein, R. A. (1992). The Path to "The TJ Hooper": The Theory and History of Custom in the Law of Tort. *The Journal of Legal Studies*, 21(1), 1-38.

- Feltz, A. (2007). The Knobe effect: A brief overview. *The Journal of Mind and Behavior*, 265-277.
- Fletcher, G. P. (1970). Theory of criminal negligence: A comparative analysis. *U. Pa. L. Rev.*, 119, 401.
- Fletcher, G. P. (1985). The right and the reasonable. *Harvard Law Review*, 949-982.
- Frisch, L. K., Kneer, M., Krueger, J. I., & Ullrich, J. (2020). Do You Feel the Same? The Effect of Outcome Severity on Moral Judgment and Interpersonal Goals of Perpetrators, Victims, and Bystanders.
- García Olier, J. & Kneer, M. (in prep.) What drives the Knobe Effect: Valence or severity?
- Gardner, J. (2001). The mysterious case of the reasonable person. *The University of Toronto Law Journal*
- Gardner, J. (2015). The many faces of the reasonable person. *Law Quarterly Review*, 131(1), 563-584.
- Vol. 51, No. 3 (Summer, 2001), pp. 273-308
- Hartman, R. J. (2017). *In defense of moral luck: Why luck often affects praiseworthiness and blameworthiness* (Vol. 38). Taylor & Francis.
- Hoffman, M. (ms.). The psychology of the trial judge.
- Heller, K. J. (1998). Beyond the Reasonable Man-A Sympathetic But Critical Assessment of the Use of Subjective Standards of Reasonableness in Self-Defense and Provocation Cases. *Am. J. Crim. L.*, 26, 1.
- Hurd, H. M., & Moore, M. S. (2002). Negligence in the Air. *Theoretical Inq. L.*, 3, 333.
- Kamin, K. A., & Rachlinski, J. J. (1995). Ex post ≠ ex ante. *Law and Human Behavior*, 19(1), 89-104.
- Kneer (in prep. a). The folk concept of practical rationality.
- Kneer (in prep. b). “Reasonable” and “rational” are assessment-sensitive.
- Kneer, M., & Bourgeois-Gironde, S. (2017). *Mens rea* ascription, expertise and outcome effects: Professional judges surveyed. *Cognition*, 169, 139-146.
- Kneer, M., Hannikainen, I.R., Almeida, G., Aguiar, F., Bystranowski, P., Dranseika, V., Janik, B. M., Garcia Olier, J., Güver, L., Liefgreen, A., Tobia, K., Próchnicki, M., Rosas, A., Skoczeń, I., Strohmaier, N. & Struchiner, N. (in prep.) Outcome effects on mental state ascriptions across cultures
- Kneer, M., & Machery, E. (2019). No luck for moral luck. *Cognition*, 182, 331-348.
- Knobe, J. (2003a). Intentional action and side effects in ordinary language. *Analysis*, 63(3), 190-194.
- Knobe, J. (2003b). Intentional action in folk psychology: An experimental investigation. *Philosophical psychology*, 16(2), 309-324.
- Knobe, J. (2010). Person as scientist, person as moralist. *Behavioral and Brain Sciences*, 33(4), 315–329.
- Knobe, J., & Shapiro, S. J. (2020). Proximate Cause Explained: An Essay in Experimental Jurisprudence. *University of Chicago Law Review*, forthcoming.
- Kobick, J. (2010). Discriminatory intent reconsidered: Folk concepts of intentionality and equal protection jurisprudence. *Harvard Civil Rights-Civil Liberties Law Review*, 45, 517–562.
- Lee, C. K. Y. (2007). *Murder and the reasonable man: Passion and fear in the criminal courtroom* (Vol. 37). NYU Press.

- Lidén, M., Gräns, M., & Juslin, P. (2019). ‘Guilty, no doubt’: detention provoking confirmation bias in judges’ guilt assessments and debiasing techniques. *Psychology, Crime & Law*, 25(3), 219-247.
- Low, P. W. (1987). The model penal code, the common law, and mistakes of fact: recklessness, negligence, or strict liability. *Rutgers LJ*, 19, 539.
- MacFarlane, J. (2014). *Assessment sensitivity: Relative truth and its applications*. OUP Oxford.
- Macleod, J. A. (2019). Ordinary causation: a study in experimental statutory interpretation. *Ind. LJ*, 94, 957.
- Martin, J. W., & Cushman, F. (2016). The adaptive logic of moral luck. *The Blackwell companion to experimental philosophy*, 190-202.
- Miller, A. D., & Perry, R. (2012). The reasonable person. *NYUL Rev.*, 87, 323.
- Moore, M. S., & Hurd, H. M. (2011). Punishing the awkward, the stupid, the weak, and the selfish: The culpability of negligence. *Criminal Law and Philosophy*, 5(2), 147-198.
- Moran, M. (2010). The Reasonable Person: A Conceptual Biography in Comparative Perspective. *Lewis & Clark L. Rev.*, 14, 1233.
- Nagel, T. (1979). *Mortal Questions (Vol. 89, Vol. 3)*. Cambridge University Press.
- Nadelhoffer, T. (2006). Bad acts, blameworthy agents, and intentional actions: Some problems for juror impartiality. *Philosophical Explorations*, 9(2), 203–219.
- Nourse, V. (2008). After the Reasonable Man: Getting over the Subjectivity/Objectivity Question. *New Criminal Law Review*, 11(1), 33-50.
- Prochownik, K., Krebs, M., Wiegmann, A., & Horvath, J. (2020). Not as Bad as Painted? Legal Expertise, Intentionality Ascription, and Outcome Effects Revisited.
- Rachlinski, J. J. (1998). A positive psychological theory of judging in hindsight. *The University of Chicago Law Review*, 65(2), 571-625.
- Simons, K. W. (2002). Dimensions of negligence in criminal and tort law. *Theoretical Inquiries in Law*, 3(2).
- Simons, K. W. (2008). Self-Defense: Reasonable Beliefs or Reasonable Self-Control?. *New Criminal Law Review: An International and Interdisciplinary Journal*, 11(1), 51-90.
- Stark, F. (2016). *Culpable carelessness: Recklessness and negligence in the criminal law*. Cambridge University Press.
- Strohmaier, N., Pluut, H., van den Bos, K., Adriaanse, J., & Vriesendorp, R. Hindsight bias and outcome bias in judging directors’ liability and the role of free will beliefs. *Journal of Applied Social Psychology*.
- Struchiner, N., Hannikainen, I., & Almeida, G. (2020). An experimental guide to vehicles in the park. *Judgment and Decision Making*, 15(3).
- Tobia, K. P. (2018). How people judge what is reasonable. *Ala. L. Rev.*, 70, 293.
- Tobia, K. P. (2020a). Law and the cognitive science of ordinary concepts. *Handbook on Law and the Cognitive Sciences (Cambridge University Press, forthcoming)*.
- Tobia, K. P. (2020b). Testing ordinary meaning: An experimental assessment of what dictionary definitions and linguistic usage data tell legal interpreters. *Harvard Law Review*, 133.
- Wells, C. (1982). Swatting the subjectivist bug. *Criminal Law Review*, (APR), 209-220.
- Westen, P. (2008). Individualizing the reasonable person in criminal law. *Criminal Law and Philosophy*, 2(2), 137-162.
- Williams, B. (1993). Moral luck. In D. Statman (Ed.). *Moral Luck*. SUNY Press.
- Zipursky, B. C. (2014). Reasonableness in and out of Negligence Law. *U. Pa. L. Rev.*, 163, 2131.