6

Causation and the Silly Norm Effect

Levin Güver and Markus Kneer

6.1 Introduction

6.1.1 The Correspondence Assumption

Whereas in certain domains, the law relies on terms of art (e.g., "injunction," "double jeopardy," "punitive damages," and "bankruptcy"), in others—in particular in criminal law—it invokes, or takes itself to invoke, the plain everyday meaning of the expressions used. This is unsurprising: citizens, standardly not equipped with a law degree, must understand what the law says in order to adhere to it. Furthermore, in common law jurisdictions, nonexperts help decide court cases in jury trials. And when it comes to disputes in statutory interpretation among judges, turning to ordinary meaning is one, if not—as some scholars and practitioners argue[1]—*the* evident strategy to resolve them.[2]

According to what we term the correspondence assumption, certain central legal expressions are taken to refer to the same concepts as their corresponding ordinary language analogues (at least within designated spheres of the law). Candidate concepts for the correspondence assumption are plentiful. *Consent* is one. At a recent sexual misconduct trial in the United States, the judge refused to provide conceptual classification and stated that "the jury will decide what consent means to them" (Puente, Sloan, & Deerwester, 2018; for empirical work on the notion of consent, see Sommers, 2020). The expression "reasonable" and the concept it denotes constitute another example. As Gardner (2015) writes, the reasonable person standard "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" (p. 36). Naturally, for a maneuver of this sort to even begin to make sense, it must be assumed that the lay person's concept of reasonableness fits the law's demands.[3]

In many jurisdictions, the central mens rea concepts, such as intention, are subject to the correspondence assumption—which is perhaps one of the key reasons why, very frequently, they are left partially or entirely uncodified.[4] The English courts have made this explicit stating that "the legal meaning of the word 'intention' is the ordinary meaning of the word" (Herring, 2012, p. 135). In *R v. Moloney* [1985], Lord Bridge put it as follows:

The golden rule should be that, when directing a jury on the mental element necessary in a crime of specific intent, the *judge should avoid any elaboration or paraphrase of what is meant by intent, and leave it to the jury's good sense to decide whether the accused acted with the necessary intent*, unless the judge is convinced that, on the facts and having regard to the way the case has been presented to the jury in evidence and argument, some further explanation or elaboration is strictly necessary to avoid misunderstanding.[5]

Indeed, the courts have been very reluctant to provide the jury with further directions on intention, doing so only in "very rare"[6] or "very exceptional"[7] cases.[8]

6.1.2 Correspondence Trouble

Where correspondence is assumed, a complication can arise: although the law takes a certain legal expression *E* to mirror ordinary language, its application in daily life differs from what the law presumes. The divergence can be due to one of two reasons: misalignment in *application only* or misalignment in *semantics*. In the former case, the assumption of correspondence holds good in so far as the legal expression *E* and its ordinary language equivalent are semantically on a par—they mean the same. However, the application of the ordinary language expression sometimes differs radically from what the law assumes due to either pragmatics or bias.[9] Consider the expression "intention" and its cognates, for which English law assumes correspondence. Problematically, lay attribution of intentionality is sensitive to outcome valence (good v. bad, the Knobe Effect)[10] and outcome severity (the Severity Effect).[11] The Knobe Effect threatens to undermine a meaningful distinction between the mentes reae knowledge and intention for bad outcomes.[12] Both the Knobe Effect and the Severity Effect put pressure on the conceptual and procedural independence of actus reus and mens rea, since features of the former (outcome valence or severity) influence the attribution of the latter. Importantly, the problem is not limited to the judgments of lay juries. Legal professionals, including judges, also manifest the Knobe Effect[13] and the Severity Effect[14] for intentionality attributions. Differently put, even if the folk (as well as experts) would reflectively endorse certain implicit and explicit constraints the law imposes on the concept of intention (such as the possibility of a hard distinction between knowingly and intentionally committing a crime), its standard *application* can still be inconsistent with these assumptions.

The *second* complication that can arise with respect to the correspondence assumption runs deeper than the first. It goes beyond application, pragmatics, and potential bias and instead regards the very semantics of the expression at stake. In such a case the folk expression *E* does not actually mean what the law takes it to mean, and this explains why folk applications of the concept designated by *E* defy legal expectations. Differently put, the folk use or application might be perfectly adequate—what is off are the legal hypotheses as to what the folk concept *E* and its apparently corresponding legal equivalent actually mean. The aforementioned concept of reasonableness is a good example, since it might defy legal expectations in two orthogonal ways: *normativity* and *outcome dependence*.[15]

There is an extensive debate among legal scholars as to whether "reasonable" is best understood as a *descriptive* expression (such that what is reasonable is what the *ordinary* person would do), a prescriptive or normative expression (such that what is reasonable is what the *responsible, prudent,* or perhaps somewhat *ideal* citizen would do), or possibly a *hybrid* expression (or what philosophers refer to as a "thick" expression, having both descriptive and normative components).[16] Despite outsourcing the meaning of the expression to the folk, the law *does* speculate about, and thus constrain, what it can mean. Consider, for instance, the staggering variety of explanations of the reasonable person standard for negligence across jury instructions in US sates. Some of those are more in tune with a descriptive standard (focusing on "ordinary" conduct, such as Texas), and others—explicated in terms of the "reasonably careful" person (e.g., Illinois or Florida)—suggest a normative standard.[17] As Tobia's (2018) empirical work shows, however, the folk concept of *reasonable* seems to be hybrid. If so, its semantics is inconsistent with legal constraints that explicate it in purely descriptive or purely normative terms.

Perhaps an even more glaring divergence arises as regards the law's insistence on the outcome independence of what is reasonable. In evaluating criminal negligence, we must consider the defendant's conduct in light of "the circumstances known to him [or her]" so as to assess whether his or her conduct "involves a gross deviation from the standard of care that a reasonable person would observe *in the actor's situation.*"[18] What matters are the agent's epistemic circumstances ex ante, not what one might come to learn about the action's consequences ex post.[19] The folk concept of reasonableness, however, seems to be strongly sensitive to outcome information: decisions and actions undertaken from the same epistemic point of view are judged more or less reasonable depending on whether the outcome is good or bad. This is not just a matter of a possibly biased, outcome-sensitive *application* of the expression "reasonable." Even when the effect of the hindsight bias is corrected for, the folk seem to insist that outcome information matters to judgments of reasonableness (see the findings in Kneer, 2021).

6.1.3 The Correspondence Assumption with Respect to Causation

So far, we have explained what we call the correspondence assumption, provided a few examples, and examined two distinct types of problems that can arise in the wake of assumptions of this sort. With the basic conceptual framework in place, we will now turn to the concept of causation, which constitutes the topic proper of this chapter. Here too, we take it that there are at least decent grounds to hold that the correspondence assumption is in place for certain jurisdictions.

Causation lies at the heart of both tort law and criminal law. The actus reus (the "guilty act") is one of the two central requirements for criminal culpability besides mens rea (the "guilty mind"). In the rather rare cases of strict liability, the actus reus by itself can suffice. There's considerable evidence that common law jurisdictions, which overwhelmingly task lay juries with the process of determining causation, endorse the correspondence assumption (see Summers, 2018). Hart and Honoré's (1959) contention

that the legal notion of causation should be that of the "plain man" (p. 1) has been echoed many times by British and American courts. In a landmark English case, Lord Wright argued that "[c]ausation is to be understood as the man in the street, and not as the scientist or the metaphysician, would understand it."[20] A Scottish court under Lord Thomson highlighted that they would rather follow "the practical experience of the reasonable man" than "the theoretical speculations of the philosopher."[21] The US Supreme Court, in the much-cited *Burrage v. United States*, stated that courts should rely on "the common understanding of causation" and explicate causal relations with reference to what it "is natural to say."[22] It thus comes as no surprise that Knobe and Shapiro's (2021) analysis of a multitude of US cases concludes that "judges who invoke the doctrine of proximate causation [. . .] are doing what the rules tell them to do, namely, to engage in *ordinary causal reasoning*" (p. 235, emphasis added).[23]

Assumed correspondence between a certain legal concept and its folk analogue does not mean that the law defers to the folk, whatever their concept might be. Even when explicit definitions are lacking, partial clarifications (e.g., in the case of "reasonable" discussed earlier) or legal procedure constrain the concept of interest and its application. A question of fundamental importance is thus whether a particular folk concept *C*, to which the law wants to avail itself, is broadly consistent with the constraints it takes to govern said concept. To make some progress in this regard as concerns the concept of causation, we will proceed as follows: in Section 6.2, we examine the legal notion of causation in the United States. Section 6.3 surveys several accounts of the folk notion of causation and discusses ways in which they could correspond with the American legal analogue (or at least certain scholarly interpretations thereof). In the remainder of the chapter, we report a series of studies that casts doubt on the suggestion that the law should invoke the "ordinary man's" concept of causation.

6.2 Causation in the Law

Common law jurisdictions have converged on a two-layer model of causation for both criminal law and the law of torts, distinguishing between factual cause and legal cause. In a first step, the courts determine whether the action in question was the factual cause of the outcome. A factual cause is determined by employing the *but-for* test: an action is deemed the cause of an event *X* if, *but-for* the action, *X* would not have come about. Simply put, if *X* is a factual cause of consequence *Y*, *X* is a necessary condition for *Y*'s occurrence.[24] Factual causation is, however, unable to capture all constellations with which the courts are confronted in their day-to-day activities.[25] It is thus in a second step—that of *legal causation*—that the courts distinguish legally *relevant* causal factors from irrelevant ones, reducing the extensive class of factual causes to those that are of import for the determination of legal responsibility.

There is no "clear [and] crisp definition" (Moore, 2019, Section 6.2.3) of proximate causation in the United States, though we can sort the multitude of formulas employed by the courts into two overarching clusters. The first takes proximate causation to be a reflection of *actual* causal relations *in the world*, whereas the second cluster employs "policy-based" (Posner, 1986, p. 181) tests, that is, tests that take normative factors,

such as considerations of justice and social interests, into account (see Moore, 2019, Section 6.5.3). A prominent example of the former cluster is the test of *directness* and an example of the latter is the test of *reasonable foreseeability*.[26]

According to the test of directness, proximate causation is established if the causal connection between an action and outcome is sufficiently direct and there is no intervening factual cause that supersedes the defendant's action (i.e., there is no further cause that stands between the defendant's action and the harmful outcome).[27] Consider the following example: in a moment of inattentiveness, *A* swings her golf club and hits *B* in the face, breaking his nose. *B* requires medical attention. On his way to the hospital, *B* is hit by a bolt of lightning and dies instantly. Is *A*'s action the proximate cause of *B*'s death? Undeniably, her action was a factual cause: had *A* not hit *B* in the face, *B* would not have been struck by lightning on his way to the hospital. Nevertheless, the lightning supersedes *A*'s doing, it severs the causal chain between the injury and the death. If, on the other hand, we were to modify the example so that *A*'s golf swing kills *B* on the spot, the causal relation would be sufficiently *direct* to consider *A* the proximate cause of *B*'s death.

In applying the test of foreseeability, courts probe whether the defendant could, at the time of her action, have reasonably foreseen the resulting harm.[28] The underlying rationale is that it is unfair to hold someone legally accountable for an unforeseeable outcome, as this would largely constitute an instance of bad luck. Since what is *reasonably* foreseeable may be subject to a wide range of value judgments (for reasonableness, see the references in Section 6.1.2), the test of foreseeability can plausibly be taken to carry normative import. To illustrate, consider the following situation: *A* is speeding past a busy town square, just ahead of which *B* is crossing the road. *A*, who is unable to react in time, collides with and fatally injures *B*. Given that *A* could, at the time of driving, have reasonably foreseen that speeding past a well-frequented area is a recipe for disaster, her doing is regarded the proximate cause of *B*'s death. However, if *A* is not speeding but instead driving attentively, and *B*—in an unpredictable manner—runs onto the road, the legal assessment would change: the accident is not judged reasonably foreseeable, and *A* is absolved of legal liability.[29]

There is a long-standing legal dispute concerning proximate causation in the law.[30] Two camps can be distinguished. Legal formalists treat proximate causation as a descriptive enterprise. On their view, causation is taken to be something in the world, and when the courts select a proximate cause, they simply single out a special class of factual causes that are sufficient in causal strength to be considered *the* legal cause of a certain outcome.[31] Legal realists disagree. They claim that when the courts speak of proximate causation, they do not take themselves to be pointing out a state of affairs in the world. Instead, courts employ the veiling language of proximate causation to make normative ascriptions of responsibility—judgments that are based to a considerable extent on moral and policy considerations.[32]

The dispute itself has a descriptive and a prescriptive dimension. On the one hand, it concerns the question as to *what the courts are really doing*, or the *practice* of the law. What are the psychological mechanisms by virtue of which judges come to reach a verdict? Formalists contend that it is *via* the deductive application of certain rules and tests (Schauer, 1988), examples of which we have already seen. According to legal

Table 6.1 Formalism and Realism

		Practice	
		Descriptive	**Normative**
Nature	**Descriptive**	Formalism	Weak realism
	Normative	–	Strong realism

realists, however, judges construe the causal query as "post hoc justification for the moral judgment [they have] already made" (Knobe & Shapiro, 2021, p. 171), deciding ultimately "with their sense of justice and social utility" (p. 176).[33] On the other hand, there is disagreement as to how courts *ought* to assess proximate causation or what the *nature* of the law demands. Should the law exclusively rely on judgments free of normative considerations to establish proximate causation? Or do such normative factors have their proper place in such decisions?

The positions discussed can be plotted in a matrix distinguishing what the courts should do according to the nature of the law, and what they actually do in practice. Formalists will contend that proximate causation is descriptive in both nature and practice; *weak* realists argue that there is a mismatch between descriptive doctrine and normative legal reality, while *strong* realists take proximate causation to be rightfully normative in both nature and practice (see Table 6.1).[34]

Let's take stock. Causation is commonly assessed in two layers. The first layer, that of *factual* causation, refers to an entirely descriptive, counterfactual notion of causation. The second layer—that of proximate causation—is established through a multitude of tests, of which we have exemplarily assessed the test of directness and that of foreseeability in the context of both criminal and tortious liability. We then briefly reconstructed the long-standing debate between legal formalists and legal realists along two dimensions—those of proximate causation's nature and practice— and distinguished three distinct positions, that is, formalism, weak realism, and strong realism.

Interestingly, however, this dichotomy between descriptive and normative theories of causation is not limited to the legal sphere and extends to the psychological literature. Given that courts tend to postulate correspondence between the legal and the folk concepts of causation, we will turn to the psychological literature next.

6.3 Theories of the Folk Concepts of Causation

6.3.1 The Norm Effect

We have argued that there is evidence that in common law jurisdictions like the United States and the United Kingdom, a correspondence assumption is in place for causation. For correspondence to hold, the folk concept must be consistent with the constraints the law imposes on the legal concept of causation. Having briefly sketched how causation is established in the law in the previous section, we will turn to recent

empirical work on the folk concept of causation. More precisely, we will focus on one aspect of this large body of work, according to which perceived causation is sensitive to salient norm violations (the *norm effect*).[35]

Consider the following situation (henceforth *Rollerblading*), which is based on a Swiss Federal Court case:[36] Mark is rollerblading on a footpath, and Lauren is walking ahead of him. Suddenly, a cat jumps out of the brush. In an attempt to evade it, Lauren steps into the lane of Mark. Mark crashes into Lauren. Who or what caused the accident? It seems natural to deem the cat as the cause of the accident. This intuition is consistent with recent findings concerning *normality* in the ascription of causation, highlighting that people tend to elevate the causal contribution of *abnormal* events—here, the suddenly appearing cat—in jointly causal structures.[37]

Now consider a variation of the scenario, in which everything is held fixed, except that it is legally *prohibited* to skate on the footpath. This is an example of an injunctive norm, as it expresses not what typically happens (a statistical norm), but what *ought* or *ought not to* be done. Despite the prohibition, Mark is rollerblading on the same footpath that Lauren is walking on. Lauren sidesteps the cat, walks into Mark's lane and the two collide. Who caused the accident? In this case, our response might differ from the original case, or so a series of empirical studies on the folk concept of causation suggests.[38] When two agents—one of them in violation of an injunctive norm—jointly bring about an outcome, the norm-violating agent is deemed more causal. This effect, standardly known as the norm effect, extends to scenarios where an outcome is brought about by a single agent, once in a norm-conforming and once in a norm-violating manner (Livengood, Sytsma, & Rose, 2017; Sytsma, Livengood, & Rose, 2012).

6.3.2 The Folk Concept of Causation

6.3.2.1 *The Counterfactual and the Pragmatic View*

There are at least four families of accounts in the literature, which purport to explain the norm effect, the first of which is known as the Counterfactual View. According to its proponents, norm violations—be they of the prescriptive or descriptive kind—motivate people to reason about counterfactual scenarios in which the agent adhered to the norm in question.[39] For Hitchcock and Knobe (2009), this is mainly the case for *abnormal* causes, which give rise to this kind of counterfactual reasoning to a significantly higher degree than normal causes. Such counterfactual reasoning, they hold, renders the abnormal factor more salient, and thus increases perceived causal contribution.

Proponents of the Pragmatic View, by contrast, hold that the locution "*A* caused *B*" can be read in one of two ways (Samland & Waldmann, 2014, 2015, 2016). Under the narrow reading, it refers to the descriptive causal processes linking events *A* and *B*. Under the broad reading, it refers to an assessment of *accountability*, a notion which extends beyond the descriptive into the normative realm. Judgments of causation in this sense are sensitive to considerations like the agent's foresight of the outcome, their desire to bring it about, and of course also norms and whether the agent was

aware of them (see Samland & Waldman, 2016, p. 165). Different contexts trigger different uses of "cause" which is, according to the Pragmatic View, what explains the norm effect.

6.3.2.2 *The Bias View and the Responsibility View*

In this chapter, we will not have much to say about the Counterfactual and Pragmatic Views. We'll predominantly focus on the Bias View and the Responsibility View, whose import for legal causation is (perhaps) clearer and more pronounced. On the Bias View (by and large Alicke's Culpable Control Model), ordinary people have a descriptive *concept* of causation, yet in *attributing* causation they fall victim to a pervasive bias (Alicke, 1992, 2000; see also Lagnado & Channon, 2008). When an agent breaks a norm, people blame her for doing so. In an implicit act of backward rationalization, their desire to blame the norm-violating agent triggers attributions of causality, even though people would agree in reflective judgment that causation does not depend on normative factors (see also Rose, 2017, p. 1327). This effect is not limited to the violation of norms. Rather, any factor that is able to elicit a desire to blame the agent— such as an agent's bad general character, wicked motives, race, gender, status, and ideology—can have downstream consequences and distort laypeople's attribution of causation.[40] Common to these factors is that people view the agent in a negative light and then project factors that justify their desire to blame them post hoc (Alicke, Rose, & Bloom, 2011, p. 670).

Take the following schematic illustration of the Bias View (Figure 6.1a): there are factors that affect blame independently of causation—whether appropriately so (e.g., mens rea—solid black arrow) or not (e.g., race, gender, and status—dashed black arrow). Furthermore, there are descriptive factors pertinent to causation (e.g., how directly the agent was involved in bringing about the outcome) that do and should affect causation, and therefore blame (solid gray arrows). There is, however, also a variety of factors that affect causation *via* blame (dashed black arrow). Although they should *not* matter for causation, an increase in perceived blame is inadequately justified post hoc by means of an increase of the ascribed causal contribution of the agent. Note that on this view, there are two ways the bias can arise: naturally, factors that inappropriately influence blame might distort perceived causation (dashed black paths). However, factors that appropriately influence blame (e.g., mens rea), yet that are conceptually independent from causation proper, might too (solid black and dashed black paths). Whereas in such a case "blame amateurs," as Alicke calls them, get it wrong only once, in the former case, they get it wrong twice (Alicke, 2008, p. 179).
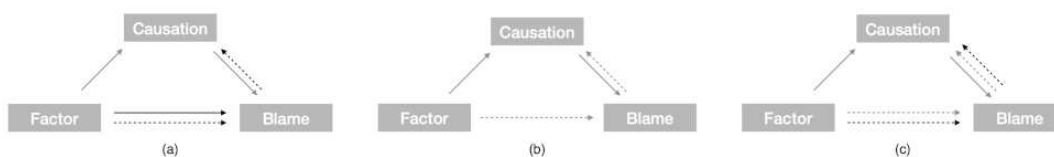
Figure 6.1 Schematic illustrations of the Bias View (6.1a), the Anything-Goes View (6.1b), and the Responsibility View (6.1c). © Levin Güver and Markus Kneer.

The Responsibility View, by contrast, holds that the influence of norms on causal judgment is not a *bug* but a *feature* (Livengood, Sytsma, & Rose, 2017, p. 284). The folk are not systematically biased in the application of a descriptive concept of causation—rather, the ordinary concept of causation is inherently normative.[41] The meaning of "X caused Y" is, in Sytsma's (2022) terminology, "quite similar" to "X is *responsible* for Y" (p. 6). Responsibility, in turn, is taken to encompass "broadly *moral* evaluations" (Sytsma, 2020, p. 21), though the notion is not further specified. Perceived responsibility can be increased not only through norm violations but also due to factors relevant to the mental state of the agent, such as her *foresight* or *desire* of the outcome (see Sytsma, 2019 and for interested related findings, Kirfel & Lagnado, 2021).

When we attempt to schematize the Responsibility View, it is not entirely clear what, exactly, it entails. The most permissive extrapolation of causal and moral responsibility being "quite similar" is that *any* factor that affects the one, can (though need not necessarily) affect the other (Figure 6.1b). Uncontroversially, descriptive features that affect causation can have an impact on blame (solid gray arrows). On a maximally permissive account, the folk concept of causation would be such that *any* factor that has an impact on *perceived* blameworthiness can have an impact on causation (dashed gray arrows). This *Anything-Goes View*—primarily discussed for didactic reasons here—is surely not what Sytsma and Livengood have in mind. True, the view makes room, for instance, for norm infractions to influence blame and therefore perceived causation, a point Sytsma defends at length. But it overgenerates: if an agent's gender impacts perceived blame in misogynistic ways, then—on this view—it would be fine to wind up with a difference in causal attribution across gender. An account of this sort, needless to say, cannot helpfully be contrasted with the Bias View, since it rules out the possibility of bias from the get-go.

What, exactly, is Sytsma's view? Following Alicke, Sytsma acknowledges the distinction between features that are "peripheral" to moral responsibility—such as, for example, "the actor's or victim's race and character" (Alicke, Rose, & Bloom, 2011, p. 674)—and those that are not (Sytsma, 2019, p. 4, 2022, pp. 11–12). Differently put, Sytsma agrees with Alicke that there are factors that *appropriately* influence moral responsibility and blame and those that do so *inappropriately*. But once *actual moral responsibility* and *perceived moral responsibility* can come apart (in contrast to the Anything-Goes View), the possibility of bias is back, and the differences between the two accounts of causation can be stated clearly. On the Responsibility View, we take the following to hold good (Figure 6.1c): uncontroversially, factors that have a direct influence on causation can have downstream normative consequences on blame (solid gray arrows). However—and this is the distinguishing feature of the account—factors that *appropriately* influence blame can also have an appropriate impact on perceived causation (dashed gray arrows). Differently put, certain factors, such as salient norms, that prima facie have no clear connection to causation can impact it nonetheless in virtue of their justified impact on perceived moral responsibility or blame. However—and this prevents the account from collapsing into an unpalatable Anything-Goes View—not just any factor that has an influence on perceived blame has a *valid* impact on causation: factors—like, for example, race, gender, or general character—that *bias*

blame are not considered appropriate influences on perceived causation (dashed black arrows).

6.3.2.3 *Recent Support for the Responsibility View*

The Responsibility and Bias Views make similar predictions for the *Rollerblading* case stated earlier. They both hypothesize that the violation of a reasonable and pertinent norm will affect blame (or moral responsibility) and thus—on one account adequately and on the other inadequately—attributed causality. Differently put, the predictions of the two views are identical with regard to all and only those factors that justly bear on moral responsibility. The two views do, however, come apart as concerns factors that should *not* bear on—or are "peripheral to"—moral responsibility or blame. According to the Bias View, such peripheral factors, which *inappropriately* influence perceived blame, will increase perceived causality just like nonperipheral ones. The Responsibility View, however, predicts that they will not—which is what prevents it to collapse into the Anything-Goes View.

One peripheral feature already briefly mentioned earlier may be the agent's *general character*. Assume that two agents *A* (a good person) and *B* (a bad person) do the exact same thing with the same state of mind, and their actions lead to a harmful outcome. Whether or not the agent is a good person should not matter for the assessment of their moral responsibility for the harmful outcome. The Responsibility View thus predicts the perceived causal contribution of the two agents to be the same. The Bias View, however, hypothesizes that factors normatively irrelevant or "peripheral" to moral responsibility, like general character, might well have an impact on blame, and—in an attempt of post hoc justification thereof—on perceived causality.

In a famous experiment, Alicke (1992)—the main proponent of the Bias View—tested the prediction. He designed a vignette where a speeding driver collides with another car. In one version, he was speeding to hide an anniversary present for his parents (good character); in the other, the driver was speeding to hide a vial of cocaine from his parents (bad character). Participants deemed the driver significantly more causal in the latter version. On Alicke's view, this is because our desire to blame the bad driver more than the good driver makes us exaggerate his causal contribution. Sytsma (2019) disagrees, hypothesizing that the two vignettes trigger not only different inferences as to the agents' general character, but also as concerns their driving ability, a feature which is relevant to causal assessment. And indeed, Sytsma shows, if driving ability is held fixed across scenarios, the effect of character on causation disappears.

In further studies with a different scenario (*Lauren Alone*, first used in Livengood, Sytsma, & Rose, 2017), Sytsma shows that manipulating character only affects causality if it also affects the attribution of inculpating states of mind (in particular, knowledge). In the scenario, Lauren works for a company that has an unstable mainframe. The company does not know that the mainframe will crash if anyone logs into it. One day, Lauren logs into the mainframe, and the system crashes. Following the crash, the company institutes a policy that forbids its employees from logging into the mainframe. In one study, Sytsma manipulates the agent's character (*not specified* v. *bad*) and her mental states concerning the system crash (*not specified* v. *specified as absent*). He

finds that character has an effect on causal judgment when knowledge and desire are left unspecified. When it is explicitly stated that the agent lacks knowledge or desire of the bad outcome, the effect disappears. What this suggests is that the participants draw an inference from bad character to an inculpating attitude toward the outcome, which then influences causal judgment because it does—and should—influence moral responsibility. In further studies, Sytsma finds that participants' causal judgments are most sensitive to the agent's knowledge of the outcome (i.e., the system crash) and, to a lesser extent, to her desire to bring it about (see also Kirfel & Lagnado, 2021).

In a nutshell, then, Sytsma shows that what really drives Alicke's astonishing results is not general character (a feature peripheral to both moral responsibility and causation), but other features (ability, mens rea) which can covary with the former, but which are not peripheral to moral responsibility (and thus, on Sytsma's view, causation).

6.4 Matching Legal and Psychological Accounts

The legal and psychological accounts discussed have prescriptive and descriptive features: they take position as to the *nature* of causation and its *actual attribution*, be it in court or our day-to-day lives. Formalists argue that the legal concept of causation is descriptive and that's how it is applied (i.e., solid gray arrows only in any of our graphs). Weak realists also hold that the legal concept of causation is descriptive, though its application has certain normative facets. Those who are vocal in their critique of the normative application of what is ultimately a descriptive concept presumably agree— by and large—with Alicke's account (Figure 6.1a). Strong realists, by contrast, argue that there is no genuine mismatch between the application of the legal concept of causation and its nature: the concept is sensitive to normative factors, so its application can be, too. This seems—at least prima facie—a good fit for Sytsma's Responsibility View. Naturally, if Sytsma's account as to what the folk concept is were correct, then some strong realist account of legal causation fits the Folk View of Causation (at least broadly). We would have actual correspondence between the legal expression (and concept) on the one hand and the folk accounts thereof.

Despite the prima facie room for convergence just discussed, a lot depends on the details. Take the factor of mens rea as an example. The law draws a strict conceptual and procedural distinction between mens rea on the one hand and the actus reus (the "guilty act") on the other. Culpable are only those who fulfill both requirements (except in cases of strict liability). Whereas Sytsma's Responsibility View might make room for a legitimate impact of mens rea on causation via responsibility (solid gray arrows, Figure 6.2), an account of this sort breaks with the hard distinction between mens rea and actus reus. According to Western criminal law and torts, the fact that a certain factor, like mens rea, appropriately increases perceived moral responsibility does not warrant an inference as to heightened causal contribution (solid gray and dashed black arrows, dashed black indicating an error/bias).

Let's take an example: suppose that we face a *many hands problem*, meaning we cannot clearly attribute causal responsibility for a harmful consequence to any of the many agents involved. Now it turns out that one agent, François, acted with knowledge
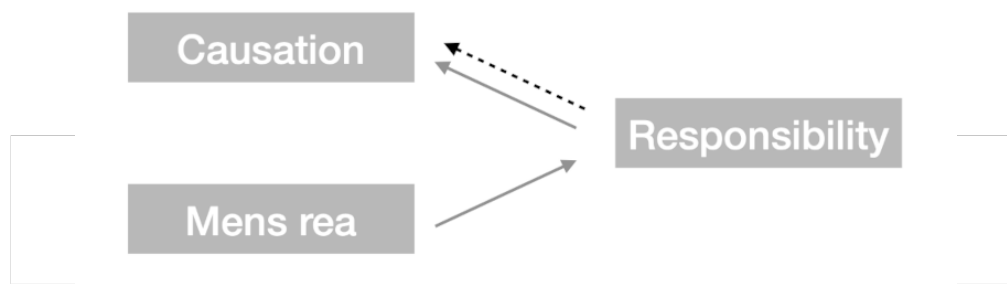
Figure 6.2 The implications of the Responsibility View for the law. © Levin Güver and Markus Kneer.

(i.e., was practically certain the harm would occur), whereas all others were merely aware of a substantial risk (i.e., acted recklessly). On the moral scoreboard, François' standing is naturally somewhat worse in this case, but does this mean that he is more *causally* responsible? According to most legal accounts of causation (and interpretations thereof), the answer is negative. Sytsma's view, however, can—and *does* (see Sytsma, 2019, pp. 5–6)—make room for such an inference. What this shows, in short, is that from the legal point of view not just any normatively relevant factor warrants an inference to causation. On the Responsibility View of Causation, however—or at least this is what can be gleaned from Sytsma's papers—any factor that *appropriately* influences perceived moral responsibility (and what can be more paradigmatic than mens rea?) can justly exert some influence on perceived causation.

Taking stock: certain versions of strong realism map onto Sytsma's Responsibility View. Causation can legitimately be influenced by normative features, be it the infraction of a pertinent rule or other factors that appropriately influence moral responsibility. Whether the latter can include mens rea depends on the particular type of strong realism at stake. Naturally, to maintain a firm distinction between mens rea and actus reus, it is not appropriate to count all factors which, like mens rea, legitimately impact responsibility as adequate desiderata for attributed causal responsibility. In contrast to strong realists, weak realists might sympathize with Alicke's Bias View: causation, on this account, is a robustly descriptive phenomenon, but in "blame validation mode" its attribution is frequently marred by normative factors.

The details of this admittedly rough matching of legal and psychological accounts of causation might seem to matter a lot—except if it could be *shown* that the folk attribution of causation fits the predictions of the Bias View. This is what we will attempt to do in the following. If our findings are on the right track, the correspondence assumption as regards causation (no matter potential restrictions of scope) is problematic: the law might be well advised to distance itself from the folk concept of causation and should have a close look over the latter's shoulders in juror trials.

6.5 Setting the Stage for the Experiments

According to the Responsibility View, certain situational features, such as the character of the agent should not have a *direct* impact on perceived causation. As Sytsma has

shown, in fact it doesn't: the influence of character on perceived causation is mediated by perceived mens rea, that is, the knowledge and desire to bring about a harmful outcome. Since mental states do play a legitimate role in the assessment of moral responsibility, it is only reasonable, on his view, that they also influence causality attributions. Just like mens rea, the violation of pertinent, contextually salient norms also does—and should—influence perceived causation.

The Responsibility View and the Bias View, we said, make identical predictions concerning perceived causation when an agent violates a norm pertinent to a harmful consequence of the agent's action. The predictions of the two views come apart as regards features "peripheral" to moral responsibility (such as character, race, status, gender, etc.). Alicke expects them to influence causation just the same; Sytsma does not (at least as long as they do not have an impact on a factor that legitimately influences moral responsibility). One such peripheral factor might be norms whose infraction is *nonpertinent* to the harmful outcome. Contrast two versions of the *Rollerblading* scenario: in one, skaters are not allowed on the footpath. In the other, they must wear a helmet—a rule that is aimed at their own protection. In the first case, where Mark is not supposed to skate on the path, he might legitimately be considered morally responsible for the accident with Lauren. However, in the second case, his moral responsibility should not be sensitive to the fact that he violates a norm. The rule to wear a helmet is supposed to protect *him*, and it simply isn't pertinent to the moral assessment or causal structure of the accident. This thought can be dramatized by invoking a patently silly norm: assume that people are only allowed to skate on the path if they like pizza, own a pet, or wear a gray T-shirt. On any account of moral responsibility worth its salt, moral responsibility should not be sensitive to the infraction of norms of this sort. On Sytsma's view, causation should thus not be sensitive to them either. Here we will present two experiments that explore whether they are.

6.6 Experiment 1

In our first experiment, we set out to test whether the effect of increased causality attribution is limited to pertinent norms or whether it extends to norms not pertinent to the consequences and even to outright silly norms. For the scenario, we used the *Rollerblading* vignette introduced in Section 6.3.1.[42]

6.6.1 Participants

Responses were collected from 278 participants on Amazon Mechanical Turk. The IP address was restricted to the United States. As preregistered,[43] participants were excluded if they failed an attention check, spent less than ten seconds reading the vignette, failed the comprehension question, or were not native English speakers. In total, 220 participants remained (female: 44%; mean age: 43 years, $SD = 13$ years, range: 22–74 years).

6.6.2 Methods and Materials

Participants were shown a vignette (*Rollerblading*) in which Mark was rollerblading on the same footpath that Lauren was walking on. It read as follows (conditions in square brackets):

> One recent summer afternoon, Mark is rollerblading outside. The path Mark is on is commonly used by cyclists, rollerbladers and pedestrians. [However, there is a sign stating that it is forbidden to be on the path as a cyclist or rollerblader. Cyclists and rollerbladers are fined $100 if they use the path.] / [However, it is forbidden to be on the path as a cyclist or rollerblader unless one wears a helmet. Mark is not wearing a helmet. He is thus not allowed to be on the path.] / [However, it is forbidden to be on the path as a cyclist or rollerblader unless one wears a gray t-shirt. Mark is not wearing a gray t-shirt. He is wearing a blue t-shirt. He is thus not allowed to be on the path.]
>
> One of these pedestrians is Lauren, who is walking ahead of Mark.
>
> Suddenly a cat jumps onto the path right in front of Lauren. Lauren is startled and steps to the left to evade it.
>
> Mark, who is approaching speedily on rollerblades from behind, collides with Lauren. The collision sweeps her off her feet and knocks her to the ground. Lauren sustains bruises all over.

Participants were randomly assigned to one of four conditions. In the *no norm* condition (displayed above without the addition of brackets), no norms as to the usage of the path were specified. In the *norm* condition, rollerbladers and cyclists were not allowed to use the path (first bracketed phrase). In the *nonpertinent norm* condition, rollerbladers and cyclists were only allowed to use the path if they wore a helmet, which Mark didn't do (second bracketed phrase). In the *silly norm* condition, everybody on the path was required to wear a gray T-shirt, and Mark's shirt was blue (third bracketed phrase).

Having read the scenario, participants had to answer a binary True/False comprehension question to confirm that they had read the vignette attentively and were aware both of Mark's action and its norm status. Participants were then asked questions about the causal contribution of Mark and the cat toward the accident. On a 7-point Likert scale, they had to report their agreement or disagreement with the following claims (labels in bold omitted):

> **Causation Mark:** "Mark caused the accident." (1 = completely disagree; 7 = completely agree)
> **Causation Cat:** "The cat caused the accident." (1 = completely disagree; 7 = completely agree)

Next, we tested two types of mental state ascriptions to Mark: knowledge and desire. As discussed earlier, Sytsma (2019) has shown that even when causality attributions *seem*

to be influenced by peripheral features (character in Alicke's cases, nonpertinent or silly norms in our case), the latter might actually impact features that *are* pertinent to moral responsibility—and on Sytsma's view, therefore causal responsibility. In Sytsma's replications of Alicke's famous cases, the impact of character on perceived causation was mediated by knowledge and desire attributions which are (at least on Sytsma's view) nonperipheral to the causation question.[44] The questions asked to what extent people agreed or disagreed with the following claims (labels in bold omitted):

Knowledge: "Mark knew that the accident would occur." (1 = completely disagree; 7 = completely agree)
Desire: "Mark desired the accident." (1 = completely disagree; 7 = completely agree)

Finally, we tested three types of moral judgment: blame, moral responsibility, and deserved punishment,[45] to see how they behave with respect to different types of norm violations (labels in bold omitted):

Blame: To what extent do you think that Mark is blameworthy, if at all, for the accident? (1 = not at all blameworthy; 7 = totally blameworthy)
Responsibility: To what extent do you think that Mark is *morally* responsible, if at all, for the accident? (1 = not at all morally responsible; 7 = totally morally responsible)
Punishment: How much punishment, if any, does Mark deserve for the accident? (1 = no punishment at all; 7 = severe punishment)

6.6.3 Results

We ran one-way ANOVAs to test the impact of norms (no norm, pertinent norm, nonpertinent norm, and silly norm) on all dependent variables (Table 6.2). Figure 6.3 provides an overview of the most important findings. We found that norm type had a significant effect on causation and moral judgment (all ps < .001). The effect size for Mark being the cause was large (η^2 = .218) and the same held for all three moral variables (all

Table 6.2 One-Way ANOVAs Exploring the Influence of Norms on Causality Ascriptions, Mental States, and Moral Judgments

	df	*F*	*p*	η^2
Causation Mark	3	20.03	<.001	.218
Causation Cat	3	8.26	<.001	.103
Knowledge	3	3.92	.009	.052
Desire	3	1.23	.298	.017
Blame	3	40.39	<.001	.359
Responsibility	3	36.45	<.001	.336
Punishment	3	35.68	<.001	.331

© Levin Güver and Markus Kneer.
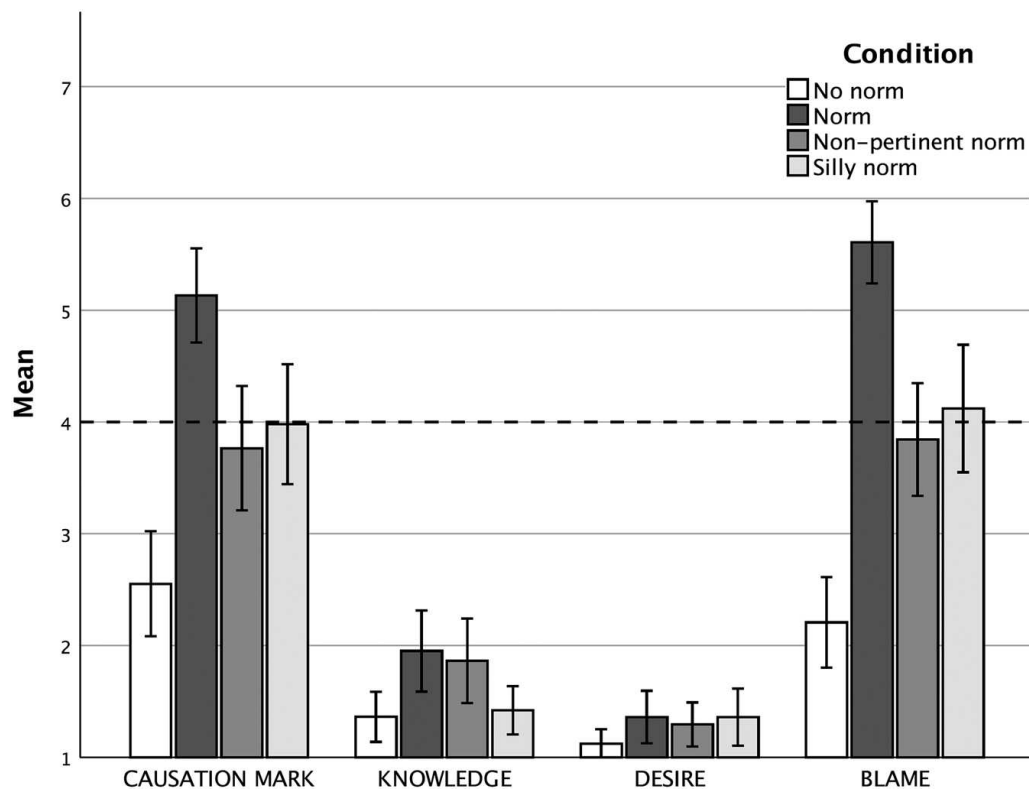
Figure 6.3 Comparison of means across all four conditions. Error bars denote 95 percent confidence intervals. © Levin Güver and Markus Kneer.

η^2s > .330). The effect of norm type on desire was nonsignificant (p = .298) and, although it reached significance for knowledge (p = .009), here the effect size was small (η^2 = .052).

According to Sytsma's view, perceived causality should covary with perceived moral responsibility (or moral blame). Across all four norm-type conditions, Mark's causal contribution correlated strongly with moral responsibility (r = .77) and blame (r = .84), in line with Sytsma's hypothesis. We also ran a mixed ANOVA (within-subject factor: judgment type—causation v. responsibility; between-subject factor: norm type—no norm v. norm v. nonpertinent norm v. silly norm). Again confirming Sytsma's view, we found that, aggregating across the four norm-type conditions, participants' causality judgments did not differ significantly from their judgments of responsibility (F(1,216) = .001, p = .972, η_p^2 = .000). In a similar mixed ANOVA with causation v. blame as the within-subject factor, we also found no significant difference in the attribution of these two DVs (F(1,216) = 1.25, p = .265, η_p^2 = .01).

To explore the impact of norms in more detail, we ran independent samples t-tests for the contrasts between the norm, nonpertinent and silly norm conditions with the no norm condition respectively.

6.6.3.1 No Norm v. Norm

The findings of the *no norm* and *norm* conditions are visualized in Figure 6.4. Contrasting *no norm v. norm* results, we found that participants deemed Mark
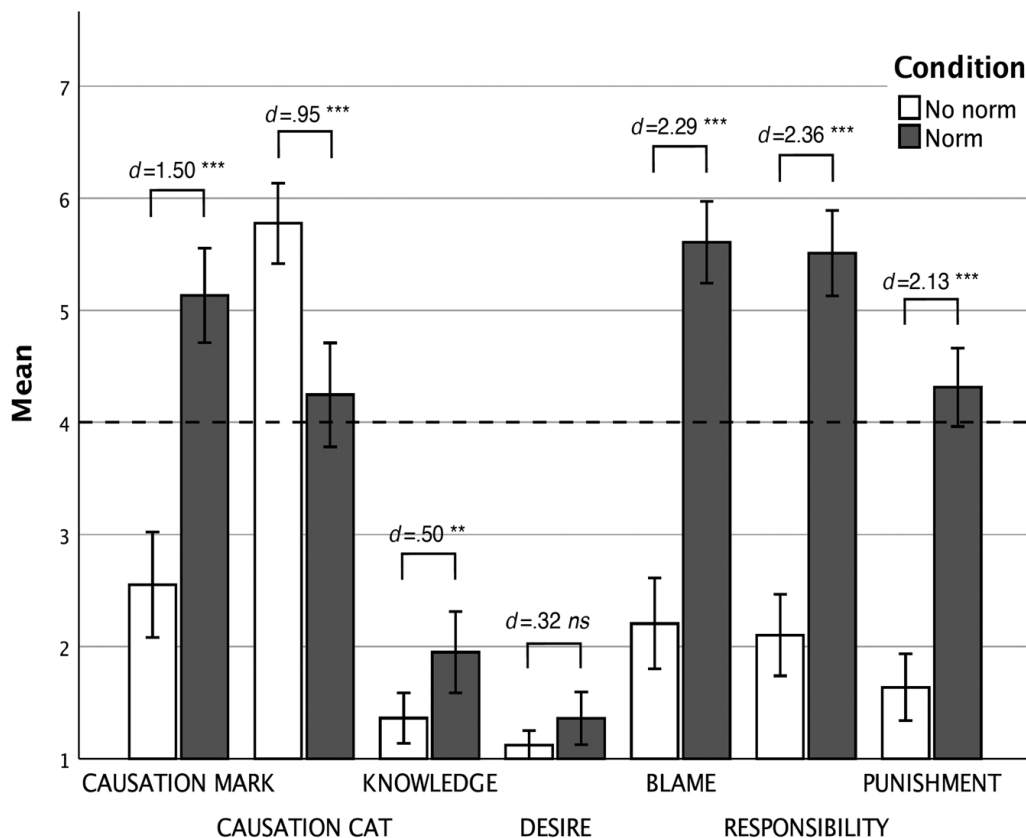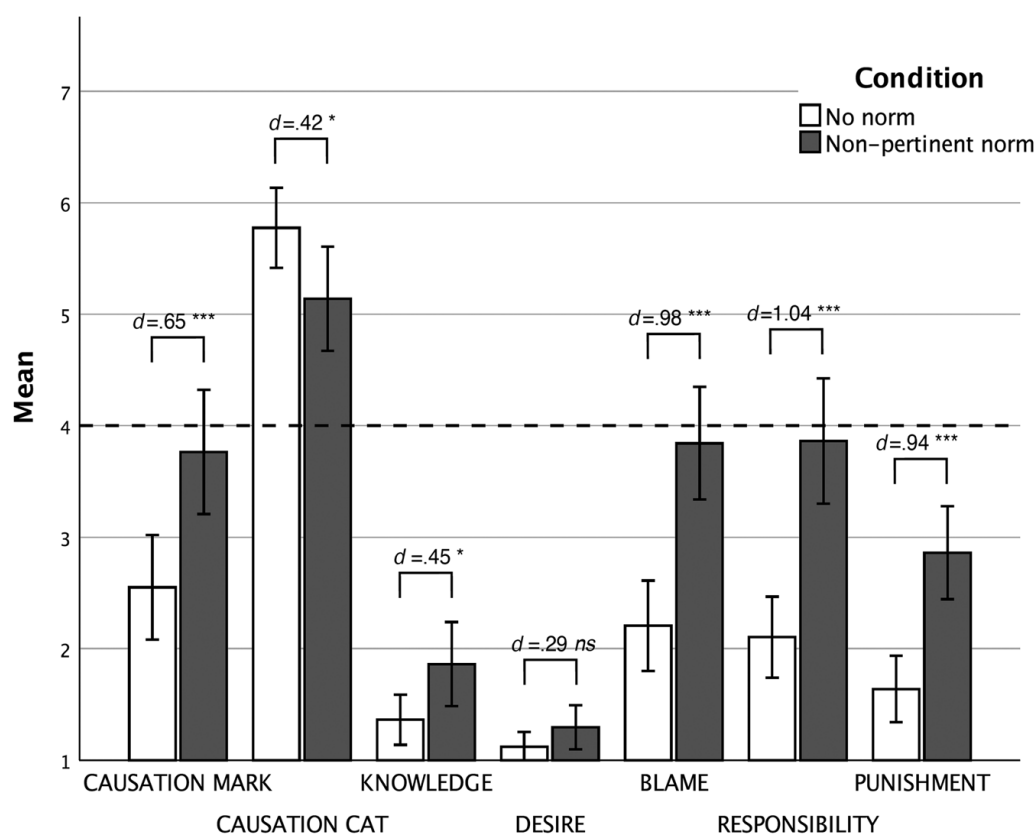
Figure 6.4 Comparison of means between the *no norm* and *norm* conditions. Effect sizes are given in terms of Cohen's *d*; * indicates *p* < .05, ** indicates *p* < .01, and *** indicates *p* < .001. Error bars denote 95 percent confidence intervals. © Levin Güver and Markus Kneer.

significantly more causal in the *norm* condition than the *no norm* condition ($p < .001$, $d = 1.50$, a large effect). This is consistent with previous findings (see note 35). There was also a significant and pronounced effect on the moral variables of blame, responsibility, and punishment (all $ps < .001$, all $ds > 2.12$, which are large effects). Additionally, participants considered Mark to have had significantly more foreknowledge of the accident ($p = .007$, $d = .50$, a medium-sized effect). There was no significant effect of norm status on perceived desire to cause an accident ($p = .080$).

6.6.3.2 No Norm v. Nonpertinent Norm

A comparison of the *no norm* and *nonpertinent norm* conditions revealed a similar effect as the one just discussed: in the nonpertinent norm conditions, participants gave significantly higher ratings for all DVs (all $ps < .031$) except the desire to cause an accident ($p = .136$), see Figure 6.5. Participants thus judged Mark significantly more causal in the *nonpertinent* norm condition than the *no norm* condition and the effect size was considerable ($d = .65$), despite the fact that Mark violated a norm that was peripheral to the outcome and (we take it) to his moral responsibility. As the data shows, however, the folk disagree with this assessment (for the moral variables all $ds > .97$).
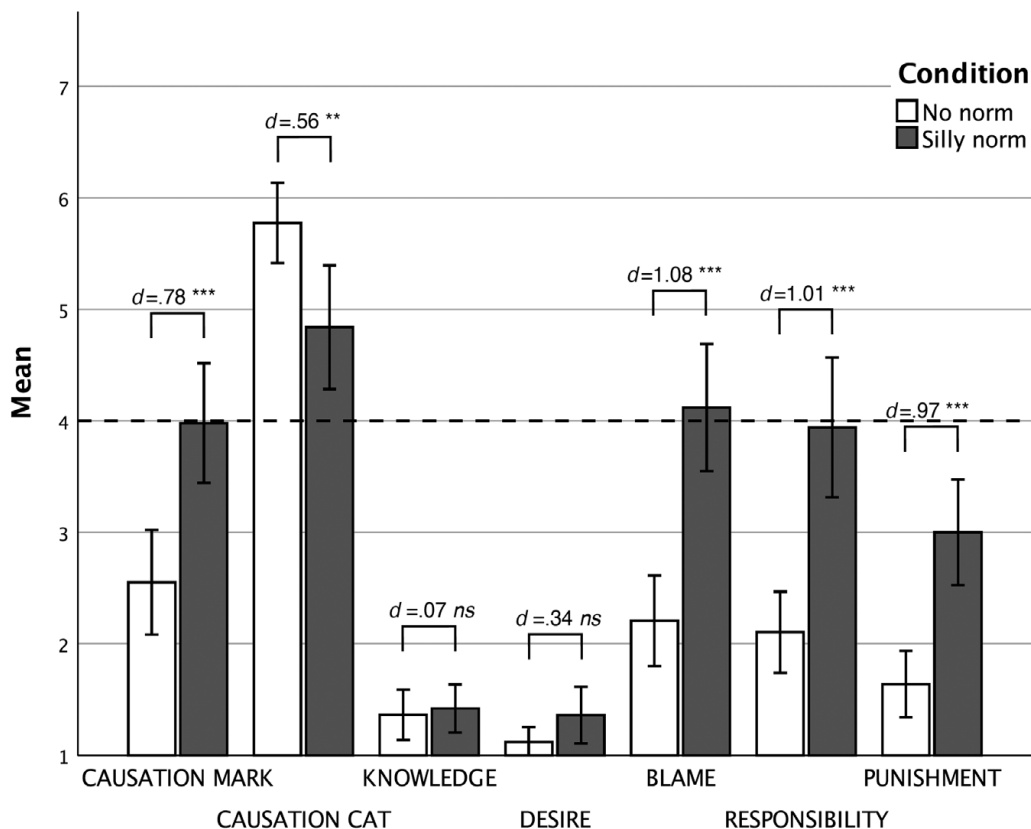
Figure 6.5 Comparison of means between the *no norm* and *nonpertinent norm* conditions. Effect sizes are given in terms of Cohen's *d*; * indicates $p < .05$, ** indicates $p < .01$, and *** indicates $p < .001$. Error bars denote 95 percent confidence intervals. © Levin Güver and Markus Kneer.

6.6.3.3 *No Norm v. Silly Norm*

Comparing the no norm and silly norm conditions, we found a significant difference for causality attributions and the moral variables (all ps $< .001$; see Figure 6.6), though we did not find a significant difference in knowledge or desire attributions (ps $> .098$). Again, the impact of a norm—albeit a silly one in this case—on causation was close to large in size ($d = .78$).

6.6.4 Discussion

Our experiment replicated previous findings according to which the violation of a norm pertinent to the moral assessment of an action influences perceived moral responsibility and—in line with the Responsibility View of Causation—the perceived causal contribution of the agent. Two conditions, in which the norm was either not pertinent to the consequences that ensued, or else patently silly, however, cast doubt on the plausibility of Sytsma's view. Since they are peripheral to moral responsibility, neither the nonpertinent or silly norm violations *should* influence responsibility or blame and hence causation. However, they do. This is in line with

Figure 6.6 Comparison of means between the *no norm* and *silly norm* conditions. Effect sizes are given in terms of Cohen's *d*; * indicates $p < .05$, ** indicates $p < .01$, and *** indicates $p < .001$. Error bars denote 95 percent confidence intervals. © Levin Güver and Markus Kneer.

Alicke's Culpable Causation Model: justified or not, bad outcomes frequently trigger blame, and when they do, people tend to rationalize their inclination to "stick it" to the agent either by exaggerated attributions of mens rea or causal contribution. Where attempts of post hoc blame justification via mens rea seem implausible (as in our scenario: all means for knowledge < 2.00, all means for desire < 1.50, no significant differences for either in the silly norm case), people seem to resort to causation.

The results are robust: we have replicated them in two further preregistered studies, each of which used a different scenario. Consistent with the findings here reported, we found a significant and pronounced effect of nonpertinent and silly norms on blame, moral responsibility, and causation. Their effect on attributed desire and knowledge was nonsignificant.

A proponent of the Responsibility View, we take it, could respond in one of two ways: *first*, they might argue that what actually matters is not *warranted* moral responsibility or blame but *ascribed* moral responsibility or blame. And indeed, the correlations between causation on the one hand and ascribed responsibility and blame on the other hand are strong in our study (across conditions, *r*s > .76, in all individual conditions *r*s > .60, see Supplemental Materials for details in note 43). But

on such an interpretation, the Responsibility View collapses into the aforementioned Anything-Goes View (Section 6.3.2.2). Many factors peripheral to moral responsibility proper—such as race, gender, character, status, and, as it turns out, the breaching of silly norms—can influence perceived blame. Since such biased moral assessments are inadequate, it is not clear why their post hoc justifications of exaggerated causation attributions should be any better.

Sytsma would agree with this assessment, we take it: after all he goes through considerable efforts to show that the impact of the "morally peripheral" feature of general character in Alicke's (1992) experiments is driven by a confound (driving ability). He further shows that, when no such confound is present, the effect of general character on causation unfolds via mens rea attribution, and mens rea is certainly relevant for moral responsibility. As discussed, we do not find an effect of the silly norm on mens rea (neither do we find one in the replications). Hence, the silly norm effect on causation is not easily explained by reference to attributed knowledge or desire. But this is where the *second* possible and certainly more plausible objection to our experiment might arise: the mens rea questions we ran following Sytsma's studies might be inadequate for the specific case at hand. In our scenario, one might argue, it simply makes little sense to attribute *foresight* (or knowledge) of an accident, so it is unsurprising that we could not detect a significant difference across conditions. However, other types of mens rea could well be relevant. The most plausible candidate is reasonable foreseeability of an accident and thus carelessness (i.e., the legal category of negligence). This is indeed a promising consideration: norm violators of any sort might be deemed careless rascals, and an increase in perceived moral responsibility, blame, and causal contribution might thus be traced back to an increase in negligence.

In short, Sytsma might hypothesize that the violation of a nonpertinent or silly norm triggers justified inferences regarding mens rea (negligence), and since these are relevant for moral responsibility these can have justified effects on perceived causation. Interestingly, the law makes room for similar considerations pertaining to the actus reus: as we have seen in Section 6.2, both criminal law and the law of torts employ tests of foreseeability in their assessment of legal causation. By testing not Mark's foresight of the accident, but its foreseeability, we can thus make headway on multiple fronts. We set out to test these hypotheses in the following experiment.

6.7 Experiment 2

6.7.1 Participants

We collected responses from 315 participants on Amazon Mechanical Turk. Their IP address was restricted to the United States. As preregistered,[46] we excluded participants who failed an attention check, spent less than ten seconds reading the vignette, or were not native English speakers. In total, 284 participants remained (female: 52%; mean age: 41 years, *SD* = 12 years, range: 20–78 years).

6.7.2 Methods and Materials

Participants were presented with the *Rollerblading* vignette from Experiment 1, though it was split into two parts. In the first step, participants were told that Mark was rollerblading on the path, that Lauren was walking ahead of him, and what type of norm applied (if any)—there being again four conditions: in the *no norm* condition, no further information was specified. In the *norm* condition, participants were told that Mark was not allowed to rollerblade on the path. In the *nonpertinent norm* condition, they were told that rollerbladers were required to wear a helmet, and Mark was not wearing one. In the *silly norm* condition, participants were told that everyone on the path was required to wear a gray T-shirt, whereas Mark was wearing a blue one.

Having read the first part of the vignette, participants were then asked to make an ex ante judgment as to the *foreseeability* of an accident.[47] The question read as follows (label in bold omitted):

> **Foreseeability:** To what extent do you agree or disagree with the following statement: "Mark could have reasonably foreseen the occurrence of an accident."
> (1 = completely disagree; 7 = completely agree)

Afterward, participants were shown the second part of the vignette, which detailed the appearance of the cat, Lauren's stepping into Mark's lane, and the ensuing collision. They were then asked to rate Mark's causal contribution toward the accident, the cat's causal contribution, the extent to which Mark is to be blamed and morally responsible for the accident, and how much punishment he deserves. The questions were phrased exactly as in Experiment 1 (see Section 6.6.2).

6.7.3 Results

We ran one-way ANOVAs to explore the influence of the four norm-type conditions on the dependent variables (see Table 6.3). Figure 6.7 provides an overview of the most important findings. We found a nonsignificant difference in participants' assessments of foreseeability across the four conditions ($p = .059$, $\eta^2 = .026$). Nevertheless, the effect of norm type on Mark's causal contribution and all moral variables was significant ($ps < .001$) and large in size for all DVs (η^2s > .194).

Table 6.3 One-Way ANOVAs Exploring the Influence of Norms on Foreseeability, Causality, and Moral Judgments

	df	*F*	*p*	η^2
Foreseeability	3	2.51	.059	.026
Causation Mark	3	22.57	<.001	.195
Causation Cat	3	7.22	<.001	.072
Blame	3	33.99	<.001	.267
Responsibility	3	42.10	<.001	.311
Punishment	3	31.89	<.001	.255
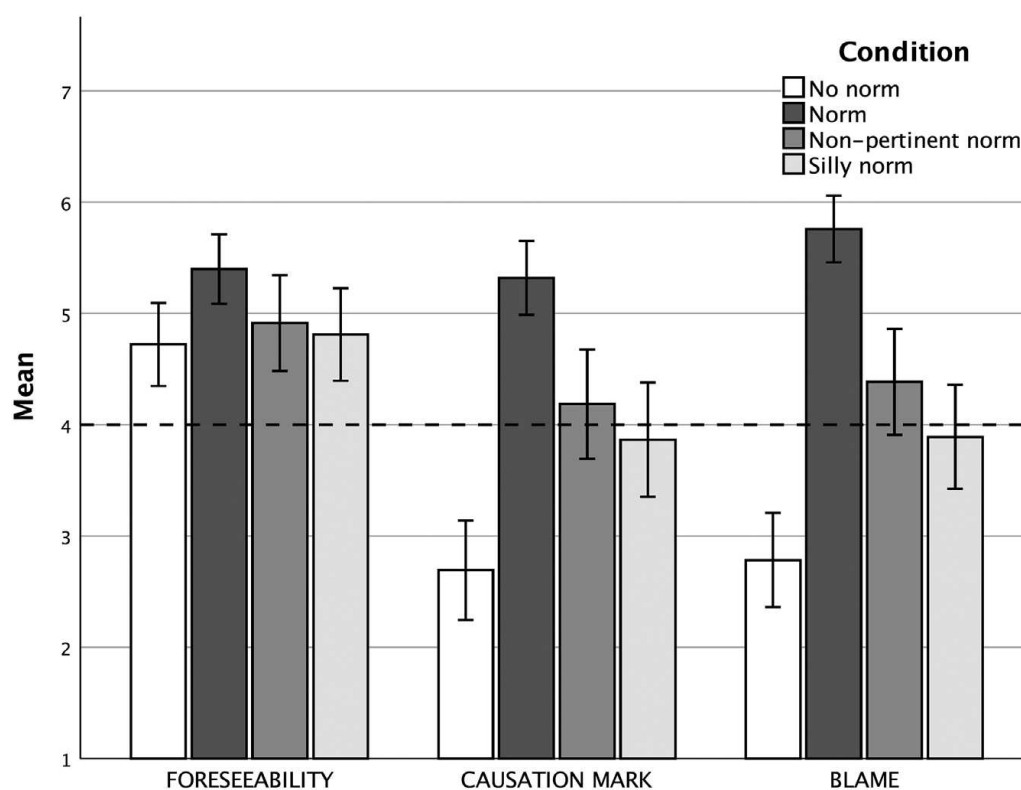
Figure 6.7 Comparison of means across all four conditions. Error bars denote 95 percent confidence intervals. © Levin Güver and Markus Kneer.

6.7.3.1 No Norm v. Norm

A comparison of the *no norm* and *norm* conditions revealed a significant difference for judgments of foreseeability ($p = .006$, $d = .47$, a medium-sized effect), see Figure 6.8. Participants also deemed Mark significantly more causal in the *norm* condition than the *no norm* condition ($p < .001$, $d = 1.62$, a very large effect), replicating the results of Section 6.6.3.1. There was also a significant and pronounced effect on the moral variables of blame, responsibility, and deserved punishment (all ps $< .001$, all ds > 1.79, which are very large effects).

6.7.3.2 No Norm v. Nonpertinent Norm

In comparing the *no norm* and *nonpertinent norm* conditions, we found no significant difference in judgments of foreseeability ($p = .507$, $d = .12$). Nevertheless, norm type had a pronounced impact on Mark's causal contribution ($p < .001$, $d=.77$, close to a large effect) and the moral variables (ps $< .001$, ds $> .85$, large effects), see Figure 6.9.

6.7.3.3 No Norm v. Silly Norm

Comparing the *no norm* and *silly norm* conditions, we found no significant difference in judgments of foreseeability ($p = .757$, $d = .05$). There was, however,
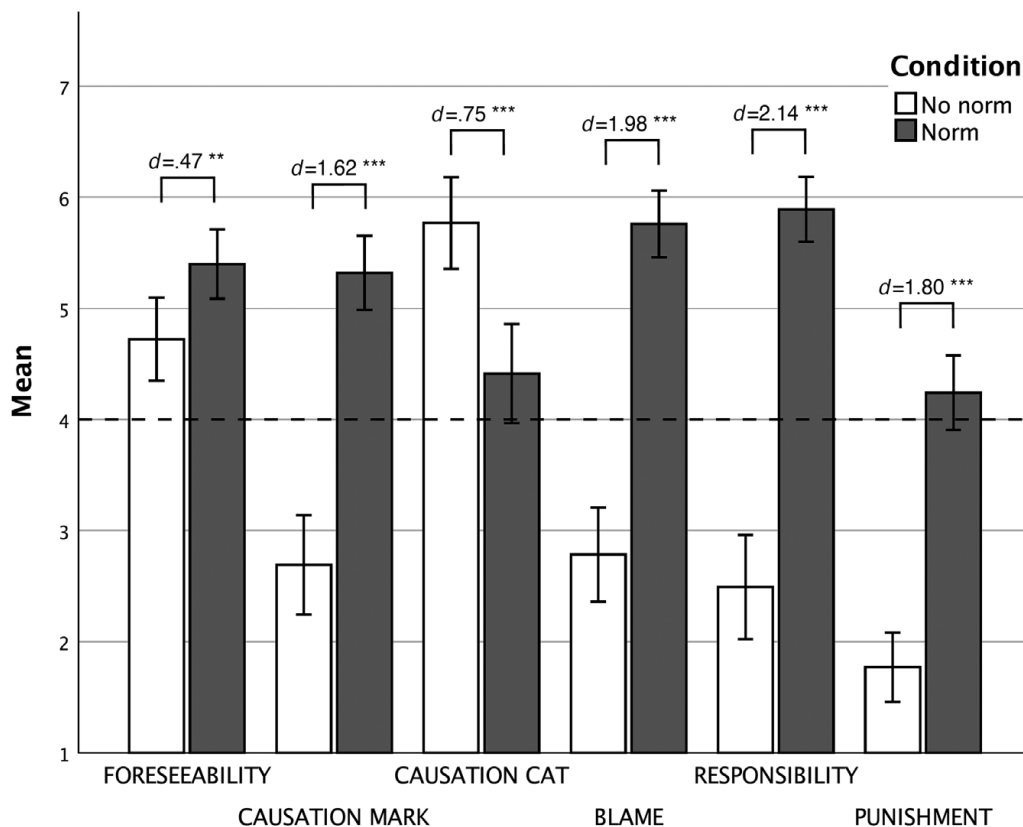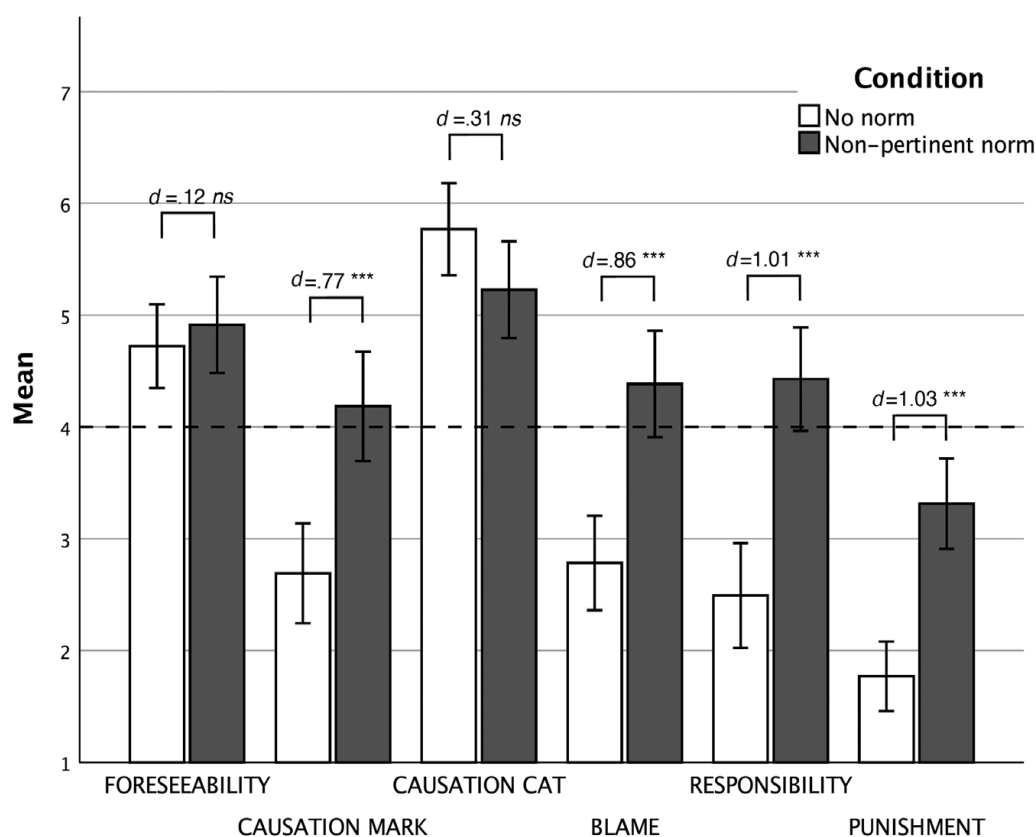
Figure 6.8 Comparison of means between the *no norm* and *norm* conditions. Effect sizes are given in terms of Cohen's *d*; * indicates $p < .05$, ** indicates $p < .01$, and *** indicates $p < .001$. Error bars denote 95 percent confidence intervals. © Levin Güver and Markus Kneer.

a significant effect of norm type on Mark's causal contribution ($p = .001$, $d = .58$, a medium-sized effect) and on blame, responsibility, and punishment ($ps < .002$, $ds > .58$), see Figure 6.10.

6.7.4 Discussion

Our experiment produced several findings. *First*, we replicated the results from Experiment 1 and the literature more generally as regards the comparison between the no norm v. norm conditions: the presence of a pertinent norm has a significant and large effect on perceived causation ($d = 1.62$) and the moral variables (all *d*s > 1.79). Note, however, that it is unlikely that this effect can be fully accounted for by foreseeability. Here, too, we found a significant norm effect, though its size is comparatively small ($d = .47$). We do not want to suggest that this needs to be problematic for either Sytsma's Folk View of Causation or certain accounts of legal causation. For instance, the findings do not pose a problem for *strong* realist readings of proximate causation, as they might concede from the get-go that a plethora of broadly normative factors can play into one's causal judgment—a position consistent with Sytsma's Responsibility View.
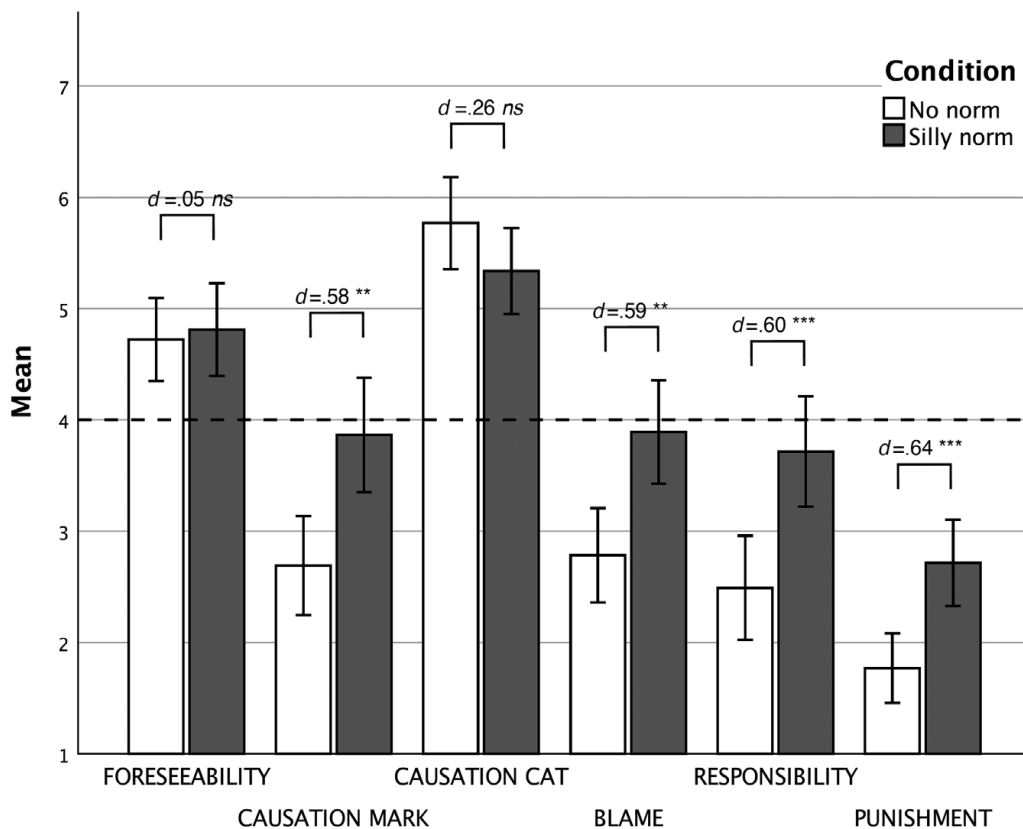
Figure 6.9 Comparison of means between the *no norm* and *nonpertinent norm* conditions. Effect sizes are given in terms of Cohen's *d*; * indicates $p < .05$, ** indicates $p < .01$, and *** indicates $p < .001$. Error bars denote 95 percent confidence intervals. © Levin Güver and Markus Kneer.

What *is* problematic for most accounts is our *second* set of findings. According to the Responsibility View and by and large any account of legal causation, nonpertinent or silly norms should not influence causation directly. They certainly also shouldn't influence causation via reasonable foreseeability, since what is *reasonably* foreseeable simply doesn't depend on what kinds of nonpertinent or silly norms happen to be in place. And in fact, the folk concurs here: contrasts of no norm v. nonpertinent norm as well as no norm v. silly norm revealed no significant effect of norm type on foreseeability ($ps > .506$). Problematically, however, both for the nonpertinent norm and the silly norm conditions Mark's causal contribution was judged as significantly more pronounced than in the no norm condition ($ps < .002$, $ds > .57$). Overall, then, our findings suggest that features peripheral to causation according to the law and any plausible version of the Responsibility View nonetheless do influence perceived causation, and that this *cannot* be explained with reference to foreseeability.

6.8 General Discussion

In certain domains, the law assumes that a legal expression *E* corresponds to its natural language analogue. Explicit clarification, established procedures, or case law provide

Figure 6.10 Comparison of means between the *no norm* and *silly norm* conditions. Effect sizes are given in terms of Cohen's *d*; * indicates *p* < .05, ** indicates *p* < .01, and *** indicates *p* < .001. Error bars denote 95 percent confidence intervals. © Levin Güver and Markus Kneer.

constraints on what such expressions can mean in the legal context—whether or not the law's take on the ordinary meaning of *E* is in fact correct. For some expressions the constraints are minimal (e.g., as concerns "reasonable" in the United States and the United Kingdom) or entirely absent (e.g., "intention," which is left uncodified in many jurisdictions, see note 4). For others, like causation, where at least in the United States and the United Kingdom, a correspondence assumption is arguably in place, the constraints are rather ample. Naturally, *presumed* and *actual* correspondence and the potential divergence that can arise are of particular importance in common law jurisdictions, where jurors are tasked with evaluating causation.

One important facet of correspondence concerns the question whether an expression (and denoted concept) is descriptive, normative, or hybrid (i.e., "thick" in philosophical jargon). The expression "reasonable" raises this question and so does the expression "cause" and cognates. Formalists argue that the doctrine of proximate causation is stated descriptively and—by and large—applied in such a fashion (naturally, the occasional slip does not mean that there are *systematic* mistakes). The law's explicit correspondence assumption (see, e.g., *Burrage*) stands and falls with what the folk expression of "cause" actually means and what the concept it denotes actually is. If, like Alicke, we hold that its sensitivity to nondescriptive factors such

as manifested by the norm effect is a bias, the possibility of correspondence in the *semantics* across folk and legal concept is possible. The fact that the folk *application* of the concept is systematically distorted by normative factors would mean that, in court, the law should be vigilant that lay jurors don't make mistakes. Suppose, on the other hand, that Sytsma were correct, such that the normatively inflected folk attributions of causation are in line with an unobjectionably normative folk concept of causation. From a formalist point of view, correspondence is thus no longer tenable: the *semantics* of the legal expression and that of the folk expression differ radically. In such a case, the law would be well advised to reign in its speculations as to correspondence and explicitly instruct jurors that in court, "cause" means something quite distinct from what it means on the street. So, in a nutshell, on formalist premises, the norm effect either poses a *threat* to the adequate *application* of the concept of cause in court by jurors (Bias View of the Folk Concept) or testifies to a difference in *semantics*, which means assumptions as to correspondence must be retired (Responsibility View of the Folk Concept).

Let's turn to *weak* realism, which holds that the letter of the law operates with a descriptive concept of causation, yet its application in court tends to be systematically inflected. If this is seen as problematic, and if we agreed with Alicke's account of folk causation, an eerie correspondence is in place. Semantically, the legal and the folk expression "cause" are on a par, yet *in and outside court* people are prone to systematic bias. If, by contrast, we agreed with Sytsma, then correspondence would once again be under pressure: the folk concept makes room for normative factors; the legal concept does not. And even if this were not to matter much given that the application of either concept is, in fact, frequently normative—that is, there's correspondence in application, though not in semantics—this would be quite a formidable mess.

What about *strong* realism? The classical norm effect can be accounted for by the strong realist position. However, on this view, too, the influence of nonpertinent and silly norms nevertheless spells trouble. Even a strong realist account of legal causation, we take it, does not amount to an Anything-Goes View, according to which any factor that might influence *perceived* moral responsibility can also legitimately influence causation. Consequently, even strong realists should, and presumably would, be alarmed by the silly norm effect on causal judgment—at least if they assume correspondence between the legal and the folk concept of causation.

So much for the possible *legal* implications of the norm effect and the silly norm effect. In our experiments, we have also tried to make progress regarding the question as to what, exactly, the folk concept of causation really is. Everyone, we take it, agrees that violations of silly norms should *not* influence perceived causation—at least if they do not impact the foreseeability of consequences or other morally *relevant* mediators. However, we found a substantial effect of silly and nonpertinent norms on causation, and we ruled out potential confounds due to foresight, desire, and foreseeability. This is problematic for Sytsma's view. As discussed at length, the fact that nonpertinent and silly norms affect perceived moral responsibility and that the latter correlates strongly with causation is of little help: the violation of silly norms, just as an agent's gender or race, should affect *neither* moral responsibility *nor* causation.

Our findings concerning the inadequate influence of silly norms on causation allow one of two interpretations: *first*, one might take them to support a *winner-takes-it-all* victory of Alicke's view. The silly norm effect suggests that people's desire to blame the agent led them to project the necessary causal prerequisites post hoc. But if this is so for morally peripheral factors, there's little reason to assume that the *process* of judgment for nonperipheral factors is much different—and it is this process that Alicke's account is about. On this interpretation, then, the legal implications we traced out under the premise of Alicke's account being correct would hold good.

Alternatively, one might opt, *second*, for a more limited conclusion. According to the latter, we still cannot adequately say whether Alicke or Sytsma are right as regards nonperipheral normative factors such as the (nonsilly) norm effect. Only as regards clearly peripheral factors—like the violation of silly norms—Alicke has a point. On a modified account of Sytsma's view, there are thus factors that appropriately and inappropriately influence folk attributions of causation. For this to be convincing, what's needed is of course some explanation why the processing of morally peripheral and nonperipheral factors should invoke different *psychological* mechanisms. We doubt that an explanation of this sort is easy to come by. What is obvious is that the earlier-expressed recommendations of caution and care as regards the possibly biased *application* of the concept of causation in court are very much in order. One might take these warnings to be restricted in scope to *juror* trials. But we have limited trust in legal expertise when deep-seated patterns of judgment distortion are at stake. Given that legal experts are just as sensitive to the Knobe Effect and the Severity Effect on mens rea attribution (Kneer & Bourgeois-Gironde, 2017; Kneer et al., ms), even when the mode of presentation is the exact same as in court (Kneer & Bublitz, ms), we doubt that all is gas and gaiters when it comes to causation. Given the powerful impact of morally peripheral normative factors on causation among laypeople, future research should address whether experts do any better in this regard.

6.9 Conclusion

Is the folk concept of causation suited for legal purposes? Does it make sense for the law to rely on "the practical experience of the reasonable man" rather than "the theoretical speculations of the philosopher" in this regard?[48] The response to these questions depends in part on what the folk concept is, and in part on the legal constraints it needs to live up to. In this chapter, we have provided an example of how an inquiry of this sort can proceed, focusing on just one facet of the empirical literature about causation, namely, the norm effect.

The norm effect demonstrates that folk *attributions* of causality are sensitive to normative factors. Whether this shows that the folk *concept* of causation is inherently normative, however, is a matter of debate. On Sytsma's Responsibility View the question is answered in the affirmative. According to Alicke's Culpable Control Model, the norm effect constitutes a bias. One can thus draw a very rough analogy between these folk psychological views on the one hand, and strong and weak realism about causation in the law on the other.

In our experiments, we have shown that the violation of morally nonpertinent and silly norms *also* has a powerful effect on causality ascriptions. Furthermore, we found that these effects cannot be explained by a potentially legitimate difference in the foreseeability of possible consequences. Effects of this sort thus constitute a bias, we take it, *both* on Sytsma's *and* Alicke's account. Do these findings suggest that the standard (pertinent) norm effect familiar from the literature—as well as other normative factors—must *also* be treated as a bias? This question requires further research. Whether or not the answer is "yes," it is evident, however, that the law should be cautious about reliance on the folk concept of causation and its application, as the latter might not be what the law *takes* them to be. If the folk concept of causation is normative, it might be unsuited for legal purposes, at least if we share formalist or weak realist premises. Even on strong realist assumptions, however, it is hard to make sense of a concept of causation that is susceptible to factors like character, gender, or silly norms. Courts must thus strive to limit inappropriate normative influences on causation judgments, in particular in juror trials, no matter how broad the class of legitimate normative influences is defined.

Notes

1 *E.g.*, Brannon (2015); Kavanaugh (2016); Solan and Gales (2016).

2 For further reflections on the relation between ordinary and legal language, see, for example, Jiménez (2021); Knobe and Shapiro (2021); Prochownik (2022); Tobia (2020, 2022).

3 See, for example, Westen (2008); Moran (2010); Gardner (2015); Zipursky (2015); Mangini (2018).

4 Both intention and knowledge are left uncodified in, for example, England, Germany (see the glaring lack in § 15 German Penal Code), France, the Netherlands, and Spain.

5 *R v Moloney* [1985] AC 905, 926 (italics added).

6 See *R v McNamara (Richard)* [2009] EWCA Crim 2530.

7 See *R v Allen* [2005] Crim LR 698; *R v Phillips* [2007] EWCA Crim 1042.

8 Herring (2012) makes the plausible assumption that the courts have not elaborated further on what the ordinary meaning of intention *actually* is because they think "it is obvious" (p. 135).

9 Note that for many puzzling phenomena (e.g., the Knobe Effect concerning intention and the norm effect concerning causation) for which one type of account is available (say, a pragmatic explanation) and the other one is standardly available also (a bias account).

10 Knobe (2003, 2006), for reviews see Cova, Lantian, and Boudesseul (2016); Feltz (2007), for early discussion in the legal context see Nadelhoffer (2006) and Kobick and Knobe (2009).

11 According to this effect, more severe outcomes correlate positively with the willingness to attribute intentionality. See Kneer and Bourgeois-Gironde (2017); Olier and Kneer (ms), and for a large cross-cultural replication of the effect with samples from over a dozen countries from the Americas, Asia, and Europe, see

Kneer et al. (ms). For a review of outcome effects broadly conceived, see Robbennolt (2000).

12 Whereas the Knobe effect blurs the boundary between knowledge and intention, the boundary between knowledge and *recklessness*, too, seems to be susceptible to influences not anticipated by lawgivers. In an experiment conducted by Severance, Goodman, and Loftus (1992), lay participants were presented with the MPC definitions of the four mentes reae and asked to apply the terms to different legally relevant scenarios. Surprisingly, participants were unable to distinguish knowledge from purpose, recklessness, or negligence—indeed, the only distinction they could reliably make was between intention and negligence (p. 115). Further studies have confirmed the boundary between knowledge and recklessness to be especially opaque, see Shen, Hoffman, Jones, Greene, and Marois (2011, pp. 1337, 1343); similarly, Levinson (2005), though *contra* Robinson and Darley (1995); Vilares et al. (2017).

13 For French legal experts, see Kneer and Bourgeois-Gironde (2017) and Bourgeois-Gironde and Kneer (2018). For US experts, see Tobia (2020).

14 For French legal experts, see Kneer and Bourgeois-Gironde (2017); for conflicting results, see Prochownik, Krebs, Wiegmann, and Horvath (2020); Tobia (2020). Although the evidence seems mixed, a large cross-cultural replication with experts from the UK, Brazil, Poland, and the Netherlands finds a robust effect for each sample, see Kneer et al. (ms).

15 Naturally, if one considers the outcome sensitivity of the expression "intentional" to be part of its ordinary language *semantics*, as some have argued, this expression, too, falls under the second category of correspondence trouble.

16 See, for example, Gardner (2001, 2015); Tobia (2018); Westen (2008); Zipursky (2015).

17 Note that "careful" has a normative ring to it, which the related noun "care" (in contrast to "carefulness") does not. The fact that the latter is frequently invoked—for example, in formulations of "ordinary care"—does not necessarily import normativity. One can effect an action with "ordinary care" and yet fail to satisfy standards of "reasonable carefulness."

18 Section 2.02(d) MPC (italics added); for another example, see Section 210.3 MPC regarding manslaughter.

19 The same point holds for Torts (Third Restatement) §3, according to which negligence turns on the "*foreseeable* likelihood" of harm and the "*foreseeability* severity of any harm," since foreseeability is tied clearly to the ex ante circumstances of the agent. More on this in footnote 47.

20 *Yorkshire Dale Steamship Co Ltd v Minister of War Transport* [1942] AC 691 (HL) 706. In a similar vein, Lord Salmon argued in *Alphacell Ltd v Woodward* [1972] A.C. 824, 847, that "[w]hat or who caused an event to occur is essentially a practical question of fact which can best be answered by ordinary common sense than abstract metaphysical theory."

21 *Blaikie v British Transport Commission* [1961] SC 44, 49, reaffirmed in *Kane v HM Advocate* [2009] HCJAC 8.

22 *Burrage v. United States*, 571 US 204 (2014). Representative in the realm of factual causation are, for example, Dressler (2012, p. 160) and Solan and Darley (2001, pp. 271–2). For a plethora of further references, see Macleod (2019, pp. 982–5), Tobia (2021, pp. 91–2), and Summers (2018, pp. 3–5).

23 See also Lagnado and Gerstenberg (2017), who argue that "legal concepts of causation are closely related to everyday causal reasoning" (p. 565).

24   See, for example, Fletcher (1998, 2000); Dressler (2015, § 14); Herring (2018, pp. 80–96); and Section 2.03(1)(a) Model Penal Code. Some authors have argued for modified forms of the test for factual causation, see, for example, Harpwood (2009) and Elliott and Quinn (2017).

25   The *but-for* test is both over- and under-inclusive in certain respects, as it faces well-known complications with situations of *overdetermination* and *preemption*, see Stuckenberg (2014) for a review, and see also the recent empirical findings by MacLeod (2019).

26   In recent years, the test of directness has fallen out of favor and was largely replaced by the criterion of foreseeability in the assessment of tortious conduct (see Owen, 2009; Goldberg & Zipursky, 2010; but cf. Keeton, 1963), whereas both criteria are used conjunctively in criminal law (Dressler, 2015, pp. 189–90). As to their general interrelation, cf. Grady (2002, pp. 9–10).

27   See, for example, Dressler (2015, p. 190); *State v. Dunn* 850 P.2d 1201, 1215. For the law of torts, cf. Epstein and Sharkey (2020); Harpwood (2000, pp. 158–9).

28   Cf. the slightly different formulations of the foreseeability test depending on whether causation or, for example, questions of tortious duty (Harpwood, 2000, pp. 31–2) or its breach (Simons, 2002, pp. 291–4) are at stake. Courts oftentimes fail to conceptually hold these two layers—those concerning breach and causation—apart, cf. Harpwood (2000, p. 27); see also Brown (ms).

29   The courts emphasis on the foreseeability of an outcome is backed by recent empirical work on causal cognition, see Kirfel and Lagnado (2021). A third test of proximate causation which Hart and Honoré (1959) were early to allude to—namely, one that probes the atypicality or abnormality of the causal chain—has also received strong support from the literature on causal cognition, see, for example, Halpern and Hitchcock (2015); Hitchcock and Knobe (2009); Icard, Kominsky, and Knobe (2017).

30   Thus, Grady (2002, p. 2) writes that "[n]o common law doctrine is more puzzling than the proximate cause limitation on negligence liability," and Swisher (2002, p. 351) reiterates: "In all of Anglo-American law, there is no concept that has been [. . .] so pervasive—and yet so elusive—as the causation requirement."

31   See, for example, Beale (1920); Kadish (1985); Schauer (1988); Hart and Honoré (2002); for an overview Moore (2009).

32   See, for example, Green (1929); Keeton (1963), Prosser and Keeton (1984); for a comprehensive overview Leiter (2005).

33   For instance, comprehensive analysis of case law in common law jurisdictions gives reason to believe that key elements of proximate causation (e.g., "directness" for the United States and "operativeness" and "substantiveness" for the United Kingdom) are judged highly inconsistently—judgments that can only be made sense of after taking into account the court's "sense of justice" or "public policy considerations," Dressler (2015, p. 189). For a similar analysis concerning Swiss case law, see Frei (2010).

34   Increasing support for the strong realist position can be found throughout the Restatement of Torts, which, in the Second Restatement, rebranded proximate causation to "legal causation," and in the Third Restatement, had the duty analysis usurp the concept entirely, see American Law Institute (1985). Arguably, both Harper, James, and Gray (1986) and Prosser, Keeton, Dobbs, Keeton, and Owen (1984) can be considered *strong* realists as well.

35   For example, Knobe and Fraser (2008); Gerstenberg and Icard (2020); Hitchcock and Knobe (2009); Knobe (2009); Kirfel and Lagnado (2017); Kominsky, Phillips,

Gerstenberg, Lagnado, and Knobe (2015); Livengood, Sytsma, and Rose (2017); Samland and Waldmann (2015, 2016); Sytsma (2019); Schwenkler and Sytsma (ms); for a review see Willemsen and Kirfel (2019).

36   Swiss Federal Court verdict of April 18, 2011, 6B_974/2010.

37   For example, Halpern and Hitchcock (2015); Hitchcock and Knobe (2009); Icard, Kominsky, and Knobe (2017); Kominsky et al. (2015); Kahneman and Miller (1986). The origins of such normality-based accounts can be traced back to at least Hart and Honoré (1959, p. 10).

38   For example, Alicke, (1992, 2000); Alicke, Rose, and Bloom (2011); Hitchcock and Knobe (2009); Knobe (2006); Knobe and Fraser (2008); Samland and Waldmann (2016); Sytsma (2019); Sytsma, Livengood, and Rose (2012).

39   Halpern and Hitchcock (2015); Henne, Kulesza, Perez, and Houcek (2021a); Henne, O'Neill, Bello, Khemlani, and De Brigard (2021b); Hitchcock and Knobe (2009); Icard, Kominsky, and Knobe (2017); Kominsky et al. (2015); cf. also Gerstenberg, Goodman, Lagnado, and Tenenbaum (2015); Halpern and Pearl (2005); Lagnado, Gerstenberg, and Zultan (2013); Woodward (2008).

40   Alicke (1992, 2000); Alicke, Rose, and Bloom (2011); on blame more generally, see Malle, Guglielmo, and Monroe (2014).

41   Livengood, Sytsma, and Rose (2017); Sytsma (2020); Sytsma, Livengood, and Rose (2012). Differently put, "we are simply dealing with the judgments that result from the correct application of a normative concept akin to responsibility or accountability. [. . .] The explanation of the norm effect is simply that we ordinarily use the lemma 'cause' in a normative way" (Sytsma, 2022, p. 28).

42   *Rollerblading* is adapted from a judgment of the Swiss Federal Court, 6B_974/2010 from April 18, 2011.

43   Available under https://aspredicted.org/ve2p4.pdf. Supplemental Materials and data for this and all further experiments can be found on the project's OSF site at https://osf.io/8meca/.

44   For interesting similar findings, see Kirfel and Lagnado (2021). Kirfel and Lagnado defend an account according to which folk causation depends on *foreseeability* rather than moral responsibility.

45   We distinguished between blame and moral responsibility mainly because Alicke invokes the former and Sytsma the latter in their accounts, though neither gives a definition of either. We included deserved punishment as a further measure since it is of direct relevance to the law on the one hand, and since there is an ongoing debate as to whether blame and punishment judgments draw on the same process of judgment. Cushman (2008) argues that this is the case, and Kneer and Machery (2019) and Frisch, Kneer, Krueger, and Ullrich (2021) challenge the view.

46   Available under https://aspredicted.org/j3sx5.pdf.

47   Note that the law explicitly highlights that reasonable foreseeability is foreseeability ex ante, not ex post (for the law of torts: Goldberg & Zipursky, 2010; cf. also Owen, 2009, pp. 1281–2, 1294; for criminal law: Dressler, 2015, pp. 189–90). So as to avoid serious worries regarding a potential hindsight bias when it comes to the assessment of negligence (see, e.g., Kamin & Rachlinski, 1995; Kneer, 2021; Kneer & Skoczen, forthcoming), the question concerning foreseeability was presented to participants before the outcome (i.e., the accident) was described.

48   As expressed by Lord Thomson in *Blaikie v British Transport Commission* [1961] SC 44, 49.

References

Alicke, M. D. (1992). Culpable causation. *Journal of Personality and Social Psychology*, *63*(3), 368–78.

Alicke, M. D. (2000). Culpable control and the psychology of blame. *Psychological Bulletin*, *126*(4), 556–74.

Alicke, M. D. (2008). Blaming badly. *Journal of Cognition and Culture*, *8*(1–2), 179–86.

Alicke, M. D., Rose, D., & Bloom, D. (2011). Causation, norm violation, and culpable control. *The Journal of Philosophy*, *108*(12), 670–96.

American Law Institute. (1985). *Model Penal Code and Commentaries (Official Draft and Revised Commentaries): With text of Model Penal Code as adopted at the 1962 annual meeting of the American Law Institute at Washington, D.C., May 24, 1962*. American Law Institute.

Beale, J. H. (1920). The proximate consequences of an act. *Harvard Law Review*, *33*(5), 633–58.

Bourgeois-Gironde, S., & Kneer, M. (2018). Intention, cause, et responabilité: Mens rea et effet Knobe. In S. Ferey & F. G'Sell (Eds.), *Causalité, Responsabilité et Contribution à la Dette* (pp. 117–44). Bruylant.

Brannon, V. C. (2015). Statutory interpretation: Theories, tools, and trends, *Congressional Research Service Reports*, *R45153*, 1–64.

Brown, T. R. (forthcoming). Minding accidents. *University of Colorado Law Review*.

Cova, F., Lantian, A., & Boudesseul, J. (2016). Can the Knobe effect be explained away? Methodological controversies in the study of the relationship between intentionality and morality. *Personality & Social Psychology Bulletin*, *42*(10), 1295–308.

Cushman, F. (2008). Crime and punishment: Distinguishing the roles of causal and intentional analyses in moral judgment. *Cognition*, *108*(2), 353–80.

Dressler, J. (2012). *Understanding criminal law* (6th ed.). LexisNexis.

Dressler, J. (2015). *Understanding criminal law* (7th ed.). LexisNexis.

Elliott, C., & Quinn, F. (2017). *Tort law* (11th ed.). Pearson.

Epstein, R. A., & Sharkey, C. M. (2020). *Cases and materials on torts* (12th ed.). Wolters Kluwer.

Feltz, A. (2007). The Knobe effect: A brief overview. *The Journal of Mind and Behavior*, *28*(3/4), 265–77.

Fletcher, G. P. (1998). *Basic concepts of criminal law*. Oxford University Press.

Fletcher, G. P. (2000). *Rethinking criminal law*. Oxford University Press.

Frei, M. A. (2010). *Der rechtlich relevante Kausalzusammenhang im Strafrecht im Vergleich mit dem Zivilrecht: Adäquate Kausalität und Voraussehbarkeit, Gefahrschaffung, Risikoverringerung, erlaubtes Risiko, Vertrauensgrundsatz, rechtmässiges Alternativverhalten, Schutzzweck der Norm, eigenverantwortliche Selbstgefährdung, Handeln auf eigene Gefahr, allgemeines Lebensrisiko und Sozialadäquanz* (Vol. 56). Schulthess.

Frisch, L. K., Kneer, M., Krueger, J. I., & Ullrich, J. (2021). The effect of outcome severity on moral judgment and interpersonal goals of perpetrators, victims, and bystanders. *European Journal of Social Psychology*. Online first, available under https://doi.org/10.1002/ejsp.2805.

Gardner, J. (2001). The mysterious case of the reasonable person. *The University of Toronto Law Journal*, *51*(3), 273–308.

Gardner, J. (2015). The many faces of the reasonable person. *Law Quarterly Review*, *131*(1), 563–84.

Gerstenberg, T., Goodman, N. D., Lagnado, D. A., & Tenenbaum, J. (2015). How, whether, why: Causal judgments as counterfactual contrasts. In D. C. Noelle, R. Dale, A. S. Warlaumont, J. Yoshimi, T. Matlock, C. D. Jennings, & P. P. Maglio (Eds.), *Proceedings of the 37th annual conference of the cognitive science society* (pp. 782–7). Cognitive Science Society.

Gerstenberg, T., & Icard, T. (2020). Expectations affect physical causation judgments. *Journal of Experimental Psychology: General*, *149*(3), 599–607.

Goldberg, J. C. P., & Zipursky, B. C. (2010). *Torts*. Oxford University Press.

Grady, M. F. (2002). *Proximate cause decoded*. UCLA *Law Review*, *50*, 293–335).

Green, L. (1929). Are there dependable rules of causation? *University of Pennsylvania Law Review*, *77*(5), 601–28.

Halpern, J. Y., & Hitchcock, C. (2015). Graded causation and defaults. *The British Journal for the Philosophy of Science*, *66*(2), 413–57.

Halpern, J. Y., & Pearl, J. (2005). Causes and explanations: A structural-model approach. Part I: Causes. *The British Journal for the Philosophy of Science*, *56*(4), 843–87.

Harper, F., James, F. Jr., & Gray, O. S. (1986). *The law of torts* (2nd ed., Vol. 4). Little, Brown and Company.

Harpwood, V. (2000). *Principles of tort law*. Cavendish.

Harpwood, V. (2009). *Modern tort law*. Routledge–Cavendish.

Hart, H. L. A., & Honoré, T. (1959). *Causation in the law*. Clarendon Press.

Hart, H. L. A., & Honoré, T. (2002). *Causation in the law* (2nd ed. 1985, repr.). Clarendon Press.

Henne, P., Kulesza, A., Perez, K., & Houcek, A. (2021a). Counterfactual thinking and recency effects in causal judgment. *Cognition*, *212*, 104708.

Henne, P., O'Neill, K., Bello, P., Khemlani, S., & De Brigard, F. (2021b). Norms affect prospective causal judgments. *Cognitive Science*, *45*(1), e12931.

Herring, J. (2012). *Criminal law: Text, cases, and materials* (5th ed.). Oxford University Press.

Herring, J. (2018). *Criminal law: Text, cases, and materials* (8th ed.). Oxford University Press.

Hitchcock, C., & Knobe, J. (2009). Cause and norm. *The Journal of Philosophy*, *106*(11), 587–612.

Icard, T. F., Kominsky, J. F., & Knobe, J. (2017). Normality and actual causal strength. *Cognition*, *161*, 80–93.

Jiménez, F. (2021). Some doubts about folk jurisprudence: The case of proximate cause. *The University of Chicago Law Review Online*. Available under https://doi.org/10.2139/ssrn.3815405.

Kadish, S. H. (1985). Complicity, cause and blame: A study in the interpretation of doctrine. *California Law Review*, *73*(2), 323–410.

Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review*, *93*(2), 136–53.

Kamin, K. A., & Rachlinski, J. J. (1995). Ex post ≠ ex ante: Determining liability in hindsight. *Law and Human Behavior*, *19*(1), 89–104.

Kavanaugh, B. M. (2016). Fixing statutory interpretation: Review of 'Judging Statutes' by Robert A. Katzmann. *Harvard Law Review*, *129*(8), 2118–63.

Keeton, R. E. (1963). *Legal cause in the law of torts*. Ohio State University Press.

Kirfel, L., & Lagnado, D. (2017). 'Oops, I did it again.' The impact of frequent behaviour on causal judgement. In R. Granger, U. Hahn, & R. Sutton (Eds.), *Proceedings of the 39th annual meeting of the cognitive science society* (pp. 2420–5). Cognitive Science Society.

Kirfel, L., & Lagnado, D. (2021). Causal judgments about atypical actions are influenced by agents' epistemic states. *Cognition*, *212*, 104721.

Kneer, M. (2022). Reasonableness on the Clapham Omnibus: Exploring the outcome-sensitive folk concept of reasonable. In P. Bystranowski, J. Bartosz, & P. Maciej (Eds.), *Judicial decision-making: Integrating empirical and theoretical perspectives* (pp. 25–48). Springer Nature.

Kneer, M., & Bourgeois-Gironde, S. (2017). Mens rea ascription, expertise and outcome effects: Professional judges surveyed. *Cognition*, *169*, 139–46.

Kneer, M., & Bublitz, J. C. (ms). Outcome effects on mens rea attribution: A comparative study with German legal experts, law students and laypeople (in prep.).

Kneer, M., Hannikainen, I., Zehnder, M.-A., Almeida, G., Aguiar, F., Bystranowski, P., … Struchiner, N. (ms). Outcome effects on mental state ascriptions across cultures (in prep.).

Kneer, M., & Machery, E. (2019). No luck for moral luck. *Cognition*, *182*, 331–48.

Kneer, M., & Skoczen, I. (forthcoming). Outcome effects, moral luck and the hindsight bias. *Cognition*. Available under http://dx.doi.org/10.2139/ssrn.3810220.

Knobe, J. (2003). Intentional action and side effects in ordinary language. *Analysis*, *63*(279), 190–4.

Knobe, J. (2006). *Folk psychology, folk morality* (Dissertation). Princeton University.

Knobe, J. (2009). Folk judgments of causation. *Studies in History and Philosophy of Science Part A*, *40*(2), 238–42.

Knobe, J., & Fraser, B. (2008). Causal judgment and moral judgment: Two experiments. In W. Sinnott-Armstrong (Ed.), *Moral Psychology* (Vol. 2, pp. 441–7). MIT Press.

Knobe, J., & Shapiro, S. J. (2021). Proximate cause explained: An essay in experimental jurisprudence. *University of Chicago Law Review*, *88*, 165–236.

Kobick, J., & Knobe, J. (2009). Interpreting intent: How research on folk judgments of intentionality can inform statutory analysis. *Brooklyn Law Review*, *75*(2), 409–31.

Kominsky, J. F., Phillips, J., Gerstenberg, T., Lagnado, D., & Knobe, J. (2015). Causal superseding. *Cognition*, *137*, 196–209.

Lagnado, D. A., & Channon, S. (2008). Judgments of cause and blame: The effects of intentionality and foreseeability. *Cognition*, *108*(3), 754–70.

Lagnado, D. A., & Gerstenberg, T. (2017). Causation in legal and moral reasoning. In M. R. Waldmann (Ed.), *The Oxford handbook of causal reasoning* (pp. 565–601). Oxford University Press.

Lagnado, D. A., Gerstenberg, T., & Zultan, R. (2013). Causal responsibility and counterfactuals. *Cognitive Science*, *37*(6), 1036–73.

Leiter, B. (2005). American legal realism. In M. P. Golding & W. A. Edmundson (Eds.), *The blackwell guide to the philosophy of law and legal theory* (pp. 50–66). John Wiley & Sons, Ltd.

Levinson, J. D. (2005). Mentally misguided: How state of mind inquiries ignore psychological reality and overlook cultural differences. *Howard Law Journal*, *49*(1), 1–29.

Livengood, J., Sytsma, J., & Rose, D. (2017). Following the FAD: Folk attributions and theories of actual causation. *Review of Philosophy and Psychology*, *8*(2), 273–94.

Macleod, J. (2019). Ordinary causation: A study in experimental statutory interpretation. *Indiana Law Journal*, *93*(3), 957–1030.

Malle, B. F., Guglielmo, S., & Monroe, A. E. (2014). A theory of blame. *Psychological Inquiry*, *25*(2), 147–86.

Mangini, M. (2018). Toward a theory of reasonableness. *Ratio Juris*, *31*(2), 208–30.

Moore, M. S. (2009). *Causation and responsibility: An essay in law, morals, and metaphysics*. Oxford University Press.

Moore, M. S. (2019). Causation in the law. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Metaphysics Research Lab, Stanford University. Available under https:// plato.stanford.edu/archives/win2019/entries/causation-law/.

Moran, M. (2010). The reasonable person: A conceptual biography in comparative perspective. *Lewis & Clark Law Review*, *14*(4), 1233–83.

Nadelhoffer, T. (2006). Bad acts, blameworthy agents, and intentional actions: Some problems for juror impartiality. *Philosophical Explorations*, *9*(2), 203–19.

Olier, J. G., & Kneer, M. (ms). The Knobe effect as an instance of a Severity effect (under review).

Owen, D. (2009). Figuring foreseeability. *Wake Forest Law Review*, *44*(5), 1277–308.

Posner, R. A. (1986). Legal formalism, legal realism, and the interpretation of statutes and the constitution. *Case Western Reserve Law Review*, *37*(2), 179–217.

Prochownik, K. (2022). Causation in the law, and experimental philosophy. In P. Willemsen & A. Wiegmann (Eds.), *Advances in Experimental Philosophy of Causation* (pp. 165–88). Bloomsbury Publishing.

Prochownik, K., Krebs, M., Wiegmann, A., & Horvath, J. (2020). Not as bad as painted? Legal expertise, intentionality ascription, and outcome effects revisited. In S. Denison, M. Mack, Y. Xu, & B. C. Armstrong (Eds.), *Proceedings of the 42nd annual conference of the cognitive science society* (pp. 1930–6). Cognitive Science Society.

Prosser, W. L., Keeton, W. P., Dobbs, D. B., Keeton, R. E., & Owen, D. G. (1984). *Prosser and Keeton on the law of torts* (5th ed.). West Publishing Co.

Puente, M., Sloan, G., & Deerwester, J. (2018, April 25). Bill Cosby retrial, day 13: Jury Adjourns for night after seeking definition of consent. *USA Today*. https://www .usatoday.com/story/life /2018/04/25/bill-cosby-retrial-day-13-jury-begins-deliberati ons/548593002 [https:// perma.cc/2ADC-BTZH].

Robbennolt, J. K. (2000). Outcome severity and judgments of 'Responsibility': A meta-analytic review. *Journal of Applied Social Psychology*, *30*(12), 2575–609.

Robinson, P., & Darley, J. (1995). *Justice, liability, and blame: Community views and the criminal law*. Westview Press.

Rose, D. (2017). Folk intuitions of actual causation: A two-pronged debunking explanation. *Philosophical Studies*, *174*(5), 1323–61.

Samland, J., & Waldmann, M. R. (2014). Do social norms influence causal inferences. In P. Bello, M. Guarini, M. McShane, & B. Scassellati (Eds.), *Proceedings of the 36th annual conference of the cognitive science society* (pp. 1359–64). Cognitive Science Society.

Samland, J., & Waldmann, M. R. (2015). Highlighting the causal meaning of causal test questions in contexts of norm violations. In D. C. Noelle, R. Dale, A. S. Warlaumont, J. Yoshimi, T. Matlock, C. D. Jennings, & P. P. Maglio (Eds.), *Proceedings of the 37th annual conference of the cognitive science society* (pp. 2092–7). Cognitive Science Society.

Samland, J., & Waldmann, M. R. (2016). How prescriptive norms influence causal inferences. *Cognition*, *156*, 164–76.

Schauer, F. (1988). Formalism. *Yale Law Journal*, *97*(4), 509–48.

Schwenkler, J., & Sytsma, J. (ms). Reversing the norm effect on causal attributions (preprint). Available under http://philsci-archive.pitt.edu/18220/.

Severance, L. J., Goodman, J., & Loftus, E. F. (1992). Inferring the criminal mind: Toward a bridge between legal doctrine and psychological understanding. *Journal of Criminal Justice*, *20*(2), 107–20.

Shen, F., Hoffman, M., Jones, O., Greene, J., & Marois, R. (2011). Sorting guilty minds. *New York University Law Review*, *86*, 1306–60.

Simons, K. W. (2002). Dimensions of negligence in criminal and tort law. *Theoretical Inquiries in Law*, *3*(2), 1–57.

Solan, L. M., & Darley, J. M. (2001). Causation, contribution, and legal liability: An empirical study. *Law and Contemporary Problems*, *64*(4), 265–98.

Solan, L. M., & Gales, T. (2016). Finding ordinary meaning in law: The judge, the dictionary or the corpus? *International Journal of Legal Discourse*, *1*(2), 253–76.

Sommers, R. (2020). Commonsense consent. *Yale Law Journal*, *129*(8), 2232–324.

Stuckenberg, C.-F. (2014). Causation. In M. D. Dubber & T. Hörnle (Eds.), *The Oxford handbook of criminal law*. Oxford University Press.

Summers, A. (2018). Common-sense causation in the law. *Oxford Journal of Legal Studies*, *38*(4), 793–821.

Swisher, P. N. (2002). Insurance causation issues: The Legacy of Bird v. St. Paul Fire & Marine Ins. Co. *Nevada Law Journal*, *2*, 351–85.

Sytsma, J. (2019). The character of causation: Investigating the impact of character, knowledge, and desire on causal attributions (preprint). Available under http://philsci-archive.pitt.edu/16739/.

Sytsma, J. (2020). Causation, responsibility, and typicality. *Review of Philosophy and Psychology*. Online first, available under https://doi.org/10.1007/s13164-020-00498-2.

Sytsma, J. (2022). The responsibility account. In P. Willemsen & A. Wiegmann (Eds.), *Advances in experimental philosophy of causation* (pp. 145–64). Bloomsbury Publishing.

Sytsma, J., Livengood, J., & Rose, D. (2012). Two types of typicality: Rethinking the role of statistical typicality in ordinary causal attributions. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, *43*(4), 814–20.

Tobia, K. (2018). How people judge what is reasonable. *Alabama Law Review*, *70*(2), 293–359.

Tobia, K. (2020). Legal concepts and legal expertise (working paper). Available under https://doi.org/10.2139/ssrn.3536564.

Tobia, K. (2021). Law and the cognitive science of ordinary concepts. In B. Brozek, J. Hage, & N. Vincent, *Law and mind: A survey of law and the cognitive sciences* (pp. 86–96). Cambridge University Press.

Tobia, K. (2022). Experimental jurisprudence. *University of Chicago Law Review*, *89*. Available under https://doi.org/10.2139/ssrn.3680107.

Vilares, I., Wesley, M. J., Ahn, W.-Y., Bonnie, R. J., Hoffman, M., Jones, O. D., … Montague, P. R. (2017). Predicting the knowledge–recklessness distinction in the human brain. *Proceedings of the National Academy of Sciences*, *114*(12), 3222–7.

Westen, P. (2008). Individualizing the reasonable person in criminal law. *Criminal Law and Philosophy*, *2*(2), 137–62.

Willemsen, P., & Kirfel, L. (2019). Recent empirical work on the relationship between causal judgements and norms. *Philosophy Compass*, *14*(1), e12562.

Woodward, J. (2008). Psychological studies of causal and counterfactual reasoning. In C. Hoerl, T. McCormack, & S. R. Beck (Eds.), *Understanding counterfactuals, understanding causation* (pp. 16–53). Oxford University Press.

Zipursky, B. (2015). Reasonableness in and out of negligence law. *University of Pennsylvania Law Review*, *163*, 2131–70.