Received: 1 May 2018 | Revised: 29 July 2018 | Accepted: 21 September 2018

DOI: 10.1111/phc3.12562

ARTICLE

WILEY

Recent empirical work on the relationship between causal judgements and norms

Pascale Willemsen[1] | Lara Kirfel[2]

[1] Ruhr-University Bochum

[2] University College London

Correspondence
Pascale Willemsen, Ruhr-University Bochum, Institute for Philosophy II, Universitaetsstr. 150, 44801 Bochum, Germany.
Email: pascale.willemsen@rub.de

Funding information
Germany's Excellence Initiative, Grant/Award Number: DFG GSC 98/3

Abstract

It has recently been argued that normative considerations play an important role in causal cognition. For instance, when an agent violates a moral rule and thereby produces a negative outcome, she will be judged to be much more of a cause of the outcome, compared to someone who performed the same action but did not violate a norm. While there is a substantial amount of evidence reporting these effects, it is still a matter of debate how this evidence is to be interpreted. In this paper, we engage with the three most influential classes of explanations, namely, (a) the Norm-Sensitive Cognitive Process View, (b) the Normative Concept View, and (c) the Pragmatics View. We will outline how these theories explain the empirical results and in what ways they differ. We conclude with a reflection on how well these strategies do overall and what questions they still leave unanswered.

1 | INTRODUCTION

Judgements about causation play an important role in explaining and predicting the world. When catastrophes like the nuclear accident in Fukushima happen, we immediately ask *why* these things happened—what its causes are. And very often, the aim of gaining such causal knowledge is not only to better understand an event's history but also to be able to predict and prevent similar events in the future. For the same reasons, we try to understand what caused a certain epidemic to spread or young adults to slide into delinquency. When we determine that A caused B, we make a causal judgement.

Causal judgements play an important role in our social life and the social practices we have established. For instance, to determine who is legally responsible for a car accident, we need to settle the question of who caused it. Moral questions are equally closely tied to causal questions: We only know who to blame for the red stain on the carpet if we can determine who spilled his or her wine and thereby caused the stain.

Thus, it comes as no surprise that causal cognition is flourishing as an interdisciplinary research project whose results have gained significant attention. In this paper, we do not even try to give a complete overview of the research done in the last couple of years.[1] Instead, the focus of this paper will be on the alleged effect that norms have on judgements about singular causation, such as "Tom playing with his phone while driving caused the accident on Bedford Way this morning," in contrast to the general causal claim "Playing with one's phone while driving causes accidents." We will start by introducing the notion of norms that has been referred to in the literature, and we will describe a selection of effects that have been reported. Several explanations of these effects have been suggested. We outline those explanations to allow for a critical comparison of their respective explanatory power. The paper concludes with some general remarks and open questions.

2 | NORMS AND THEIR EFFECTS ON CAUSAL JUDGEMENTS

The literature on the role of norms in causal cognition has distinguished between at least two different kinds of norms[2]: First, there are *statistical norms* which describe the frequency or likelihood of a certain action or event (Bear & Knobe, 2017; Sytsma, Livengood, & Rose, 2012; Willemsen, 2016). For instance, it is normally the case that men are taller than women or that children favour chocolate over spinach. An action can be statistically abnormal in two ways: It can either be abnormal with respect to the population of which the agent is part and thus violate a population-level statistical norm. Alternatively, an action can be abnormal given the agent's typical behaviour. In that case, an action violates an agent-level statistical norm. While it might be abnormal for children to favour spinach, this might be very typical for an individual agent—she always chooses spinach when given a choice between spinach and chocolate. Alternatively, a child might usually prefer spinach over chocolate but this time chooses chocolate instead of spinach.

Second, norms can also be *injunctive* (Alicke, 1992; Knobe & Fraser, 2008; Kominsky, Phillips, Gerstenberg, Lagnado, & Knobe, 2015; Reuter, Kirfel, van Riel, & Barlassina, 2014; Sytsma et al., 2012; Willemsen & Reuter, 2016). Injunctive norms include both *prescriptive* norms, which tell people what they *should do*, and *proscriptive* norms, telling people what they *should not do*.[3] Moral norms are a prototypical example of injunctive norms, for example, "keep your promises", "tell the truth" (prescriptive norms), as well as "do not steal" and "do not harm others" (proscriptive norms). However, there is a variety of other injunctive, yet non-moral norms that have similar action-guiding functions, such as those found in the legal system. In addition, there are injunctive norms concerning table manners, dress codes, and norms of politeness.

In contrast to statistical norms, violating an injunctive norm is often followed by negative social feedback. Violating a moral norm against harming others is followed by blame, and in more severe cases, the legal system will even inflict some form of punishment. Violating a social norm by wearing a swimsuit in a church will be met with disapproval. In contrast, violating a statistical norm, for instance when a woman is taller than a man, might cause confusion or surprise but is not typically met with any form of social condemnation.

The most prominent evidence that norms can have an impact on causal judgements comes from a study by Knobe and Fraser (2008). In their experiment, Knobe and Fraser presented their participants with a short vignette in which a professor and an administrative assistant both take a pen from the departmental office. As a result of the lack of pens, the receptionist cannot take notes during an important phone call. In their study, Knobe and Fraser varied whether both agents or only the administrative assistant was officially allowed to take pens. When only the administrative assistant was allowed to take pens but the professor was not, people gave higher causal ratings for the outcome to the professor, that is, the agent who violated an injunctive norm.

The "pen case" study was soon followed by similar experiments to demonstrate the influence of injunctive norms on causal judgements. When two agents log in to a computer (Knobe & Fraser, 2008) or enter a room with a motion detector (Icard, Kominsky, & Knobe, 2017; Kominsky et al., 2015), but only one agent is allowed to perform this action, the norm-violating agent is judged to be more of a cause for the subsequent outcome of their actions. It

has, thus, been argued that actions or events that violate injunctive norms, that is, the way people or things generally ought to behave, receive higher causal attribution for an outcome than actions or events that adhere to injunctive norms. This phenomenon has often been dubbed "abnormal inflation."

In addition to the effect that causal responsibility ratings are increased for norm-violating agents, Kominsky and colleagues demonstrated that the effect of norms goes beyond the causal attributions to the norm-violating factor. If two agents or events jointly cause an outcome, people not only attribute higher causality to the factor that is norm-violating but also reduce their causal attribution to the other, norm-conforming factor (Kominsky et al., 2015). This phenomenon has been described as "causal supersession," expressing the idea that the perceived causal contribution of a cause can be superseded by the causal contribution of another causal factor that is norm-violating (Kominsky et al., 2015).

Similar evidence has been demonstrated for the influence of statistical norms. If an agent does something atypical, for example, driving a different route home (Kahneman & Miller, 1986), or produces an unlikely outcome, for example, rolling a number higher than 11 with two dice (Icard et al., 2017; Kominsky et al., 2015), this behaviour is assigned higher causal contribution compared to a behaviour that is statistically normal.

It might be argued that the fact that both injunctive and statistical norm violations elicit the same effects might depend on an important connection between both kinds of norms. When an agent violates an injunctive norm, for instance by harming someone, this action also violates a statistical norm. Typically, people do not harm each other. Although injunctive norm violations often covary with statistical norm violation, it has been shown that the influence of injunctive norms remains unchanged once statistical norm-conformity is controlled for (Knobe & Fraser, 2008). A study by Sytsma et al. (2012) suggest that the statistical normality of an action can also indicate whether an injunctive norm was violated. In some cases, people believe that if you violate a statistical norm, you have also violated an injunctive norm. As an example, imagine that the department's secretary usually sends around reminder emails about important guest speakers and important departmental events. It is not her job to send these emails, she just always does it because she knows that these things are easily forgotten about. One day, she does not send an email and, as a consequence, half of the staff forgets about an important talk. While the secretary only violated a statistical norm, one might feel that because of her routine of sending these emails and her well-known reliability, she was also supposed to send the email.

As Henne, Pinillos, and De Brigard (2016) as well as Willemsen (2016) demonstrate, norms do affect causal judgements not only about actions but also about omissions: An omission that violated an injunctive norm is considered much more causally relevant for the outcome, compared to an omission that did not violate an injunctive norm. For example, when an agent promised to water his neighbour's plants, his not doing so was considered the cause of the plants' death. In contrast, if he did not promise to do so, people denied that the agent's omission caused the plants' death. An effect in the same direction was found for statistical norms. When an agent deviates from her agent-specific norm of regularly attending a gym class, her not attending is judged to be more of a cause for the other participants to believe that the agent was sick, compared to when she did not violate such a norm by not attending (Willemsen, 2016).

Given the evidence for the similar influence of both injunctive and statistical norms on causal judgements about actions as well as omissions, the general claim for the influence of norms on causal judgements has been made. It has been argued that norm-violating actions or events receive higher causal attribution for an outcome than actions or events that adhere to norms (Halpern & Hitchcock, 2015; Hitchcock & Knobe, 2009; Icard et al., 2017; Knobe & Fraser, 2008).

3 | THREE EXPLANATIONS

In the literature, several explanations of the influence of injunctive norms on causal judgements have been proposed. In general, these explanations differ along two dimensions. On the one hand, some authors have proposed that causal judgements are influenced by norms because the cognitive process leading to those causal judgements is responsive

to norms. Call this the *Norm-Sensitive Cognitive Process View*. Whenever agents make causal judgements, they will consider norms as an important guide. On such a view, people start with a purely descriptive concept of causation. "Descriptive" in this context means that a causal judgement is made true by the world—A caused B if A and B stand in a certain relation to one another. Whether such a relation exists between A and B is a matter of facts, not norms. However, in the process of reaching a causal judgement, the application of this descriptive concept is affected by norms.

In contrast, others have argued that the concept of causation is itself normative. Call this the *Normative Concept View*. Thus, the cognitive process of causal judgement making starts off from a concept of causation that already has built-in normative content. The difference between theories that argue for a norm-sensitive causal process and those who argue for an inherently normative concept of causation is that according to the latter, no causal judgement can ever, by definition, be purely descriptive. Every causal judgement always includes information about injunctive or statistical norms that are relevant to the specific context. The former kind of theories does not rule out that in some situations, causal judgements are entirely descriptive and do not rely in any way on norms.

Other authors have objected that neither is the process of reaching a causal judgement sensitive to norms nor is the concept of causation normative. Norms have no effect on causal judgements either way. Instead, it is argued that when people make causal judgements in an experimental setting, the context in which they are asked for their causal evaluation triggers an alternative interpretation of the test question. Call this the *Pragmatics View*. According to this view, in many of the studies taken to support the *Norm-Sensitive Cognitive Process* or the *Normative Concept View*, participants did not answer a question about *causation* but about a related notion such as accountability. According to such an explanation of the empirical results, these results do not shed light on causal cognition at all.

In the following, we will have a closer look at each of these views in more detail. In Section 4, we will compare these views with each other to evaluate their conclusiveness and explanatory power.

3.1 | Norm-Sensitive Cognitive Process View

There are two dominant explanations of the effect of norms on causal judgements that both focus on the norm-sensitivity of the cognitive process leading to causal judgements. In the following, we refer to them as the *Point of Intervention Account* and the *Bias Account*.

Many researchers have provided evidence that people rely heavily on counterfactual reasoning when making causal judgement (Gerstenberg, Goodman, Lagnado, & Tenenbaum, 2014; Lagnado, Gerstenberg, & Zultan, 2013; Petrocelli, Percy, Sherman, & Tormala, 2011). The Point of Intervention Account builds on the idea that counterfactual reasoning is key to causal judgements. When people make a causal judgement, they primarily reach this judgement by asking "what if?" and imagining what would have happened, had the situation been slightly different.

But if people reason about causation counterfactually, why do people choose that when both the administrative staff and the professor take pens, it is the professor who caused the problem? It is correct that had the professor not taken pens, there would not have been a problem, as there would be one pen left (keeping the rest of the story identical). However, it seems equally true that had the administrative staff not taken pens, there would not have been a problem either (again, keeping everything else the same). Both counterfactuals lead to the outcome being avoided. Consequently, people should say that both the administrative staff and the professor caused the problem. The reason for this is that one agent violated a norm, while the other one did not.

The Point of Intervention Account suggests that the main purpose of all causal judgements is to identify the best point of intervention, so the factor with which we need to fiddle to prevent similar effects in the future, or to bring about the same effect in the future again, if the outcome is desirable. When selecting this best point of intervention, norms provide a helpful guide. How do they do that?

Hitchcock & Knobe (2009) argue that norms influence counterfactual reasoning. Imagine that you observe a traffic accident. According to Knobe and colleagues, when you now determine who caused the accident, you engage in counterfactual reasoning to determine the best point of intervention. Kominsky and colleagues as well as Icard and colleagues. When doing so, not all the possible counterfactual scenarios are equally likely to cross

you mind. Imagine you know that one of the two drivers crossed a green light, the other one a red light. The first thing you will think about is a counterfactual scenario in which the second driver does not cross a red light. Other counterfactuals are less likely to come to mind. For instance, would you ever consider a case in which a powerful alien casts a spell that turns cars into soft, puffy little clouds or opens a portal that allows the two cars safe passage? If the world had exploded just before the cars hit each other, there would have been no car accident. While this could have prevented the accident as well, these counterfactuals are so abnormal that they would not even cross your mind. Norms, on the other hand, make it very likely that a counterfactual scenario will be considered.

The Point of Intervention Account suggest that people have a purely descriptive concept of causation that depends on counterfactual reasoning.

According to the Bias Account, norms enter the picture in an importantly different way. The Bias Account, as suggested by Alicke (1992, 2000, 2008), agrees with the Point of Intervention Account that when people make a causal judgement, they start off with a purely descriptive concept of causation that does not include any normative elements. When confronted with a negative outcome that needs causal explanation, two cognitive processes are initiated. First, there is an objective process evaluating the agent's causal involvement, and there is a second cognitive process which spontaneously and unreflectively evaluates the agent's moral responsibility. If this second process generates a sufficiently strong blame (or praise, for that matter) response, the initially descriptive, and value-free causal judgement is re-evaluated in a blame (or praise) validation mode. The aim of this re-evaluation is to vindicate the quick and unreflected moral judgement that the agent is to blame or to praise.

The Point of Intervention and the Bias Account, thus, differ in the cognitive process they postulate, and they further differ in whether they consider a normatively influenced causal judgement to be competent or biased. The Bias Account considers the impact of norms on causal judgements as systematic performance error (Alicke, 2000). When people's causal judgements are affected by their moral intuitions, their causal judgement get distorted in the way that the agent's causal contribution is overstated. The Point of Intervention Account does not believe that the effect of norms on causal judgements constitutes a bias. Instead, the effect tracks an important feature of the function of causal judgements, namely, to identify points of intervention (Knobe, 2010; Kominsky et al., 2015). Both accounts yet share that they consider the concept of causation to be purely descriptive and norms to affect the application of this concept.[4]

3.2 | Normative Concept View

Recently, Sytsma and Livengood (2018) have argued against the claim that people's concept of causation is descriptive. According to their view, which they refer to as the *Responsibility View*, the ordinary concept of causation has some built-in normative content and is generally similar to the concept of responsibility. This means that when people make a causal judgement, they do not make a judgement about the metaphysical relationship in the world between two causal relata that philosophers are concerned with; rather, they make a judgement about who or what is responsible for a certain event to occur. Challenging the Norm-Sensitive Cognitive Process View, Sytsma and Livengood doubt that people first apply a purely descriptive concept of causation that is affected by norms along the way to making a causal judgement. Instead, they argue that the concept of causation is already pretty much indistinguishable from the concept of responsibility. Sytsma and Livengood take the existent empirical evidence to support the Normative Concept View. First of all, the Normative Concept View considers the valence of the outcome relevant for people's causal judgements. Sytsma and Livengood claim that people's causal verdicts will be affected by the valence of the outcome if the valence of the outcome affects people's responsibility judgements. People are more likely to ascribe responsibility for a bad than for a good outcome. Assuming that causal attributions are relevantly similar to the attribution of responsibility, the valence of the outcome will often make a difference for causal judgements.

Sytsma and Livengood (2018) say that in many circumstances, their account makes very similar predictions as the Norm-Sensitive Cognitive Process View, such as for the pen or the computer case (Knobe & Fraser, 2008, Kominsky

et al., 2015). The reason is that in these cases, norm violation coincides with the moral evaluation of the agent and as a consequence with responsibility judgements. However, trolley cases provide a way of de-confounding these two aspects. Take the switch variant in which five people are standing on the tracks and are about to be killed by a run-away trolley. The agent could divert the train onto another track where there is only one person standing. By flipping the switch, the five would be saved at the costs of the one person. If the agent flips the switch, so Sytsma and Livengood argue, the agent is responsible for the one person's death—as she would not have died but for the agent flipping the switch. However, as experimental studies have repeatedly shown, the vast majority of participants believe that flipping is the morally right thing to do (Cushman & Young, 2009, 2011).

Given this moral evaluation, the Norm-Sensitive Cognitive Process View predicts that the agent violated a norm when he does not flip the switch, so Sytsma and Livengood claim. As a consequence, causal judgements for the outcome (in this case, the one person's death) should be inflated, compared to causal judgements for the outcome when the agent does flip the switch (in that case, the death of the five)—either because people want to vindicate their moral disapproval or because of the norm violation. The Normative Concept Account makes the opposite prediction and claims that causal judgements will be similar to responsibility judgements for the death. As a consequence, Sytsma and Livengood predict that agent will be considered more causally relevant when he flips the switch, compared to when he does not.

In a series of experiments, Sytsma and Livengood (2018) find empirical support for their view. People's responsibility judgements reliably predict their causal judgements. Thus, conforming to a moral norm and doing what participants considered to be the right thing to do increased causal responsibility ratings. They take this as strong evidence in favour of their own and against both the Point of Intervention and the Bias Account.

It should be noted that Sytsma and Livengood remain very vague in spelling out what this new notion of responsibility is. Responsibility, as used in the philosophical and experimental literature, is often accompanied by a qualification such as “moral” or “causal” or “legal” responsibility. Responsibility, as used in Sytsma and Livengood's account, seems to be none of these things. It seems related and yet different from blame judgements, as blame and responsibility judgements can come apart. But what exactly the similarities and differences are is unclear. Further, there are many open questions concerning the difference between responsibility and accountability. Until we have a clearer understanding of what responsibility is, Sytsma and Livengood's account might be an interesting way to think about the connection between norms and causation, but it is hard to assess the theory's predictive and explanatory power.

While Sytsma and Livengood repeatedly contrasted their own view with Joshua Knobe's one by interpreting Knobe as a defender of the Norm-Sensitive Cognitive Process View, it is not clear whether this interpretation of Knobe's account is adequate. On the one hand, Knobe uses formulations that suggest that the concept of causation is itself descriptive but enriched by information about norms: “We have seen that people's *ordinary application* of a variety of different concepts can be influenced by moral considerations” (Knobe, 2010, p. 320, emphasis added); “What we have here is a model of the *competence underlying people's use* of one particular concept” (Knobe, 2010, p. 328, emphasis added); “Even *the processes* that look most ‘scientific’ actually take moral considerations into account” (Knobe, 2010, p. 328, emphasis added).

In contrast, Knobe sometimes chooses formulations that seem more in line with the Normative Concept View and, thus, very much in line with the suggestions made by Sytsma and Livengood:

> *There is now good reason to believe* there are no concepts *anywhere in folk psychology that enable one to describe an agent's attitudes in a way that is entirely independent of moral considerations. The impact of moral judgments, we suspect, is utterly pervasive. (Pettit & Knobe, 2009, p. 602)*

Indication that Knobe does in fact subscribe to a normative concept of causation is further provided by the strong emphasis on the role of counterfactuals, given that many theories hold that the meaning of causal claims can be explained in terms of counterfactual conditionals (see, for instance, Hitchcock & Knobe, 2009; Knobe, 2010; Kominsky et al., 2015; Icard et al., 2017).

3.3 | The Pragmatics View

Both the Norm-Sensitivity and Normative Concept Views share the assumption that when researchers ask participants for their causal intuitions about vignettes, their responses actually express participants' causal judgements. According to these views, when we find that whether the professor is allowed to take pens or not has an effect on her causal evaluation, this effect reveals something about how people make causal judgements. However, not all researchers are convinced that research on causal cognition necessarily tells us anything about causal cognition. An alternative view has been suggested in a recent set of studies conducted by Samland and Waldmann (2014, 2015, 2016) and Samland, Josephs, Waldmann, and Rakoczy (2015). The authors argue that the interpretation of the causal test question given to participants is systematically ambiguous. When presented with a text-based scenario in which an agent violates an injunctive norm and, thereby, causes a certain outcome, the experimental query asking for "the cause" or the extent to which the agent caused the outcome can be understood in two different ways.

First, in a narrow sense, the term "cause" can refer to a causal relation, that is, the causal link, between what the agent did and the outcome of their action. Second, in a broader sense, the term "cause" can refer to an accountability relation, that is, how accountable or responsible the agent is for the outcome in a moral sense. According to Samland and Waldmann's (2016) *Accountability Hypothesis*, participants form a hypothesis about the meaning of the term "cause" in the experiment's causal test question. Pragmatic contextual features influence whether participants interpret the causal response measure as a request to assess causality or a request to assess accountability. In context of a vignette like the pen case, story features like human agents and norm violations make it likely for participants to answer the causal test question in terms of accountability. The accountability hypothesis argues that the influence of norms on causal judgement can be explained by the pragmatic context of the experimental situation, and thus, the effects are not effects on participants' causal judgements at all.

Samland and Waldmann published a compelling series of studies (Samland & Waldmann, 2014, 2015, 2016; Samland et al., 2015) that support their view. When using unambiguous causal strength measures or pragmatic contexts that suggest a narrow causality interpretation of the causal test question, the influence of norms on causal judgements disappears (Samland & Waldmann, 2016). Moreover, Samland and Waldmann manipulated which feature of the situation was highlighted in the causal query—the agent or the inanimate objects that was used by the agent. They found that injunctive norms only affect causal judgements when participants were asked about the causal contributions of the agents. In contrast, there was no such effect when they were asked about the causal contributions of the objects with which the outcome was caused (Samland & Waldmann, 2016). Samland and Waldmann take this as strong support for the claim that the pragmatic context affects people's interpretation of the causal query, such that highlighting the causal mechanism rather than human agents provides a context that eliminates norm effects.

The Accountability Hypothesis takes a *Pragmatics View* on the influence of norms on causal judgements—it explains norm effects as the result of experimental pragmatics. Rather than assuming the influence of norms at the stage of the availability or sampling of counterfactuals (Hitchcock & Knobe, 2009; Icard et al., 2017) or a re-evaluation of the causal concept to match blame attribution (Alicke, 1992), the Accountability Hypothesis argues that in case of norm effects, people's judgement go beyond a mere causal evaluation. Contextual features influence whether people access their causal concept only or the related accountability concept when making a judgement. Although Samland and Waldmann argue that a causal judgement is necessary to make an accountability judgement in the first place, that is, a causal link between agent and outcome has to be established, the response given in these experiments is a response to considerations of accountability, rather than causality alone. As a consequence, the authors argue that the effects found in the literature reveal nothing about causal cognition.

4 | WHAT ARE THE RESPECTIVE MERITS AND PROBLEMS?

Three classes of theories are currently on the table in order to explain the effect of norms on causal judgements. So which one is most successful in offering a complete explanation of norm effects?

Let us start with the two Norm-Sensitive Cognitive Process Views. Both the Intervention and Bias Accounts have been developed in direct response to the empirical findings the authors themselves presented and, for many situations, they make quite similar predictions as to how people's causal judgements will respond to norm violations. The subtle differences, however, become crucial when the two accounts are directly compared with each other.

For instance, the Intervention Account holds that it is norms and norms alone that affect causal judgements by changing the salience of the counterfactual conditionals relevant for making a causal judgement. The Bias View puts less weight on the norms themselves and argues that the valence of the outcome of an action has a crucial effect as well. People are sensitive to bad outcomes, as they provoke an immediate affective reaction which then needs to be validated by a post hoc modification of the initially made causal judgement. As a consequence of the different roles that the Intervention and Bias Accounts assign to norms and outcomes, they make different empirical predictions concerning the conditions under which causal judgements will be affected. For instance, while a harmful outcome alone can be sufficient to attribute causality to an agent according to the Bias Account, the Interventionist Account requires that, in addition, the agent also has to have clearly violated a norm.

By using an altered version of the "computer vignette" (Hitchcock & Knobe, 2009), Reuter et al. (2014) wanted to test whether moral judgements influence causal selection in a scenario in which the agents did not intend or foresee the outcome. Against the predictions of the Point of Intervention Account, their results revealed that the nature of the effect, that is, whether the outcome was good, bad, or neutral, is a significant predictor of judgements of actual causation, independent of whether a norm had been violated or not. The Point of Intervention Account "makes no mention of any sort of moral judgements regarding the effect" (Hitchcock & Knobe, 2009) and therefore predicts no effect of the valence of the outcome. However, speaking against the Bias Account, they also found that when the effect of an action was not foreseen or intended, people will attribute less rather than more causality to the agent when the effect is negative, compared to when the effect is neutral. Since the Bias Account states that the desire to blame the agent increases causal responsibility, the exact opposite effect should have occurred. In favour of both accounts, Reuter et al. also found that, as soon as an agent violated a norm, people attribute causality for the outcome to the agent. The fact that causal attributions are no longer moderated by the valence of the outcome once a norm violation occurs supports the Interventionist Account. In consequence, the authors conclude that neither the Point of Interventionist Account nor the Blame Account can account for causal attributions in the complex interplay of foreseeability, the valence of the effect, and the violation of a norm (Reuter et al., 2014).

In a recent corpus analysis, Sytsma, Bluhm, Willemsen, and Reuter (n.d.) also attempted to decide between the Point of Intervention and the Bias Account. They found that causal language is most frequently used in contexts with a highly negative connotation, such as accidents, deaths, problems, and so forth. If the outcome was, as Knobe and collaborators claim, irrelevant for causal evaluation, we should expect a much more balanced result. It might be objected that the fact that causal language is used when talking about negative events does not show that the outcome has an impact on causal judgements themselves. Alternatively, people might just talk about negative things more often than about positive things. But even if this objection is correct, the Intervention Account still owes an explanation of why we are so much more interested in causally explaining negative events.

Sytsma and colleagues consider these results to support the Normative Concept View. The Normative Concept View shares with the Bias Account that the outcome plays an important role in people's causal judgements. It is yet to be decided which theory makes more reliable predictions. The crucial challenge for the Normative Concept View is to spell out what the folk concept of causation is and to what extent it resembles and differs from responsibility. Both views still need to provide a more general explanation of norm effects that goes beyond moral contexts.

How about the Point of Intervention View and the Pragmatics View? As we described, Samland and Waldmann challenge Knobe and colleagues' account of norms and causal judgements. In a direct reply to Samland and Waldmann (2016), Phillips and Kominsky (2017) have developed a set of studies that aims to further test the accountability hypothesis against the counterfactual view. They argue that the absence of norm effects on causal judgement about inanimate objects—as demonstrated by Samland and Waldmann (2016)—might also be a result of people representing counterfactual scenarios in a more granular way than suggested by the Accountability

Hypothesis. Rather than envisaging a counterfactual situation in which both agent and object are absent, Phillips and Kominsky argue that people might imagine a counterfactual in which the agent but not the object is absent because injunctive norms usually refer to human agents only. In a follow-up study to Samland and Waldmann (2016), they show that injunctive norms have no effect on both causal and also counterfactual relevance ratings about inanimate objects. In contrast, they show that norms that specifically refer to inanimate objects, like norms of proper function, still influence both causal and counterfactual relevance judgements about them. However, because the relationship between the causal judgement and counterfactual reasoning that Phillips and Kominsky demonstrate in their studies is correlational only, further studies testing the pragmatic view against counterfactual theories have yet to be developed.

To what extent does the Pragmatics View differ from Sytsma and Livengood's Normative Concept View? Both views seem to operate with rather under-determined notions of accountability and responsibility. The main difference lies in how the relation between causation and accountability or responsibility respectively is defined. According to Samland and Waldmann, accountability is a concept that is different from causation and yet often verbally expressed in the same way. Causal questions are, thus, ambiguous. Sytsma and Livengood do not claim that causal queries are *ambiguous*, but norms are built in to the concept of causation. However, there is no independent evidence to support these claims, and it is not clear whether such claims are even empirically testable. Finally, as a result of these differences, Sytsma and Livengood consider the empirical evidence discussed before to be evidence about causal cognition. Samland and Waldmann explicitly deny this.

As the comparison of these accounts and their explanatory power demonstrates, much more research is in order before we can be confident to fully understand norm effects. For now, it seems that the Point of Intervention Account by Knobe and colleagues is the most extensive account regarding norms. While Phillips and Kominsky did not completely rule out all the scepticism against the Point of Intervention Account raised by Samland and Waldmann, the Point of Intervention Account has some advantages over its competitors. First, this account convinces by its sparsity of the postulated explanans, that is, the influence norms or normality, and by the variety of causal attribution effects it can explain by that. Moreover, it presents a neat extension of already existing counterfactual and sampling accounts of causal attribution (Lewis, 1973; Halpern & Pearl, 2005; Gerstenberg et al., 2014). However, the exclusive focus on norms makes it hard to account for factors that influence causal attributions yet go beyond norm violations or abnormality, such as the epistemic states of agents (Lagnado & Channon, 2008) or the valence of the effect (Reuter et al., 2014). Further, the Point of Intervention Account makes very general claims about the interaction of norms and causal judgements and does not differentiate between different types of causal cognition.

Danks et al. (2013) point out we need to distinguish at least between causal learning and causal reasoning. The current evidence focuses exclusively on causal reasoning, that is, making a causal judgement based on information that participants were previously given. While one might argue that causal reasoning is affected by norms, we still need evidence that the same effects occur for causal learning. As Danks and colleagues argue, there is a variety of processes through which we can learn causal information. Their evidence suggest that whether causal learning is affected by norms and the valence of the outcome strongly depends on the learning process involved. They do not affect people's judgements when they learned information from experience rather than from texts (which is the dominant way in which participants learned causal information in the experiments discussed in this paper). As a consequence, even if we believe that the Point of Intervention Account is already on the right track, we should be hesitant to overstate what the experimental evidence shows. To provide a more holistic picture of causal cognition, we need to shift our focus away from causal reasoning and towards causal learning and expand our methodological toolkit.

5 | CONCLUSION

In this paper, we provided a systematic overview over the recent developments in causal cognition, with special attention to the effects that norms have for causal judgements. We compared three types of explanations for these

effects. The Norm-Sensitive Cognitive Process View claims that the cognitive process underlying causal judgements is sensitive to norms. In contrast the Normative Concept View holds that norms are an inseparable part of the concept of causation all along, such that all empirical findings are to be explained as the result of the application of this norm-laden concept. And, finally, the Pragmatics View denies that norms have an effect on causal judgements—they rather trigger a reinterpretation of causal language in favour of a related notion of accountability. We outlined various limitations of each account and open questions that we hope will be answered by future research.

ENDNOTES

[1] See, among others, Lombrozo (2010), Woodward (2011, 2012), Danks, Rose, and Machery (2013), Halpern and Hitchcock (2015), Waldmann and Hagmayer (2013), and Sloman and Lagnado (2015).

[2] Some researchers believe there to be a third kind, namely, norms of proper functioning (Hitchcock & Knobe, 2009; Phillips & Kominsky, 2017). For instance, a washing machine functions properly if it cleans and spins the laundry. In this paper, we will focus on statistical and injunctive norms.

[3] Please note that the term "prescriptive" norms is sometimes used as an umbrella term for both prescriptive and proscriptive norms, so basically synonymous to "injunctive."

[4] Whether the effect of norms on non-moral judgements demonstrates a conceptual competence or a bias or performance error is not only debated in context of causation. In the experimental philosophy of action, it seems that moral intuitions affect people's application of the concept of intentionality, an effect called the side-effect effect or Knobe effect (see Knobe, 2003; Knobe & Burra, 2006; Leslie, Knobe, & Cohen, 2006; Nadelhoffer, 2006). Machery (2008) raised some serious concern about whether the competence versus performance error debate could ever be experimentally decided. The same worries apply for the disagreement between the Point of Intervention and the Bias View as well.

ORCID

Pascale Willemsen http://orcid.org/0000-0002-4563-1397

REFERENCES

Alicke, M. D. (1992). Culpable causation. *Journal of Personality and Social Psychology*, *63*(3), 368–378. https://doi.org/10.1037/0022-3514.63.3.368

Alicke, M. D. (2000). Culpable control and the psychology of blame. *Psychological Bulletin*, *126*(4), 556–574. https://doi.org/10.1037/0033-2909.126.4.556

Alicke, M. D. (2008). Blaming badly. *Journal of Cognition and Culture*, *8*(1), 179–186. https://doi.org/10.1163/156770908X289279

Bear, A., & Knobe, J. (2017). Normality: Part descriptive, part prescriptive. *Cognition*, *167*, 25–37. https://doi.org/10.1016/j.cognition.2016.10.024

Cushman, F., & Young, L. (2009). The psychology of dilemmas and the philosophy of morality. *Ethical Theory and Moral Practice*, *12*(1), 9–24. https://doi.org/10.1007/s10677-008-9145-3

Cushman, F., & Young, L. (2011). Patterns of moral judgment derive from nonmoral psychological representations. *Cognitive Science*, *35*(6), 1052–1075. https://doi.org/10.1111/j.1551-6709.2010.01167.x

Danks, D., Rose, D., & Machery, E. (2013). Demoralizing causation. *Philosophical Studies*, *17*(2), 251–277. https://doi.org/10.1007/s11098-013-0266-8

Gerstenberg, T., Goodman, N. D., Lagnado, D. A., & Tenenbaum, J. B. (2014). From counterfactual simulation to causal judgment. In *Proceedings of the 36th Annual Meeting of the Cognitive Science Society*, 523–528.

Halpern, J. Y., & Hitchcock, C. (2015). Graded causation and defaults. *British Journal for the Philosophy of Science*, *106*(11), 587–612. https://doi.org/10.1093/bjps/axt050

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), 843–887. https://doi.org/10.1093/bjps/axi147

Henne, P., Pinillos, Á., & De Brigard, F. (2016). Cause by omission and norm: Not watering plants. *Australasian Journal of Philosophy*, *95*, 1–14. https://doi.org/10.1080/00048402.2016.1182567.

Hitchcock, C., & Knobe, J. (2009). Cause and norm. *Journal of Philosophy*, *106*(11), 587–612. https://doi.org/10.5840/jphil20091061128

Icard, T. F., Kominsky, J. F., & Knobe, J. (2017). Normality and actual causal strength. *Cognition*, *161*, 80–93. https://doi.org/10.1016/j.cognition.2017.01.010

Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review*, *93*(2), 136–153. https://doi.org/10.1037/0033-295X.93.2.136

Knobe, J. (2003). Intentional action and side effects in ordinary language. *Analysis*, *63*(279), 190–194. https://doi.org/10.1111/1467-8284.00419

Knobe, J. (2010). Person as scientist, person as moralist. *Behavioral and Brain Sciences*, *33*(04), 315–329. https://doi.org/10.1017/S0140525X10000907

Knobe, J., & Burra, A. (2006). The folk concepts of intention and intentional action: A cross-cultural study. *Journal of Cognition and Culture*, *6*(1), 113–132. https://doi.org/10.1163/156853706776931222

Knobe, J., & Fraser, B. (2008). Causal judgment and moral judgment: Two experiments. *Moral Psychology*, *2*(1992), 441–448.

Kominsky, J., Phillips, J., Gerstenberg, T., Lagnado, D., & Knobe, J. (2015). Causal superseding. *Cognition*, *137*, 196–209. https://doi.org/10.1016/j.cognition.2015.01.013

Lagnado, D., & Channon, S. (2008). Judgments of cause and blame: The effects of intentionality and foreseeability. *Cognition*, *108*(3), 754–770. https://doi.org/10.1016/j.cognition.2008.06.009

Lagnado, D., Gerstenberg, T., & Zultan, R. (2013). Causal responsibility and counterfactuals. *Cognitive Science*, *37*(6), 1036–1073. https://doi.org/10.1111/cogs.12054

Leslie, A. M., Knobe, J., & Cohen, A. (2006). Acting intentionally and the side-effect effect: Theory of mind and moral judgment. *Psychological Science*, *17*(5), 21–27. https://doi.org/10.1111/j.1467-9280.2006.01722.x

Lewis, D. (1973). Causation. *The Journal of Philosophy*, *70*(17), 556–567. https://doi.org/10.2307/2025310

Lombrozo, T. (2010). Causal–explanatory pluralism: How intentions, functions, and mechanisms influence causal ascriptions. *Cognitive Psychology*, *61*(4), 303–332. https://doi.org/10.1016/j.cogpsych.2010.05.002

Machery, E. (2008). The folk concept of intentional action: Philosophical and experimental issues. *Mind & Language*, *23*(2), 165–189. https://doi.org/10.1111/j.1468-0017.2007.00336.x

Nadelhoffer, T. (2006). Desire, foresight, intentions, and intentional actions: Probing folk intuitions. *Journal of Cognition and Culture*, *6*(1), 133–157. https://doi.org/10.1163/156853706776931259

Petrocelli, J., Percy, E., Sherman, S., & Tormala, Z. (2011). Counterfactual potency. *Journal of Personality and Social Psychology*, *100*(1), 30–46. https://doi.org/10.1037/a0021523

Pettit, D., & Knobe, J. (2009). The pervasive impact of moral judgment. *Mind & Language*, *24*(5), 586–604. https://doi.org/10.1111/j.1468-0017.2009.01375.x

Phillips, J., & Kominsky, J. (2017). Causation and norms or proper functioning. Counterfactuals are (still) relevant. Proceedings of the 39th annual conference of the cognitive science society.

Reuter, K., Kirfel, L., van Riel, R., & Barlassina, L. (2014). The good, the bad, and the timely: How temporal order and moral judgment influence causal selection. *Frontiers in Psychology*, *5*(1336), 1–10. https://doi.org/10.3389/fpsyg.2014.01336

Samland, J., Josephs, M., Waldmann, M. R., & Rakoczy, H. (2015). The role of prescriptive norms and knowledge in children's and adults' causal selection. *Journal of Experimental Psychology. General*, *145*, 125–130. https://doi.org/10.1037/xge0000138

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. 1359–1364).

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*.

Samland, J., & Waldmann, M. R. (2016). How prescriptive norms influence causal inferences. *Cognition*, *156*, 164–176. https://doi.org/10.1016/j.cognition.2016.07.007

Sloman, S. A., & Lagnado, D. (2015). Causality in thought. *Annual Review of Psychology*, *66*, 223–247. https://doi.org/10.1146/annurev-psych-010814-015135

Sytsma, J., Bluhm, R., Willemsen, P., Reuter, K. (n.d.). Causal attributions and corpus analysis. http://philsci-archive.pitt.edu/1

Sytsma, J., & J. Livengood (2018), 'Intervention, bias, responsibility … and the trolley problem'. http://philsci-archive.pitt.edu/14549/

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), 814–820. https://doi.org/10.1016/j.shpsc.2012.05.009

Waldmann, M. R., & Hagmayer, Y. (2013). Causal reasoning. In D. Reisberg (Ed.), *Oxford library of psychology. The Oxford handbook of cognitive psychology* (pp. 733–752).

Willemsen, P. (2016). Omissions and expectations: A new approach to the things we failed to do. *Synthese* https://doi.org/10.1007/s11229-016-1284-9, *195*, 1587–1614.

Willemsen, P., & Reuter, K. (2016). Is there really an omission effect? *Philosophical Psychology*, *29*, 1–18. https://doi.org/10.1080/09515089.2016.1225194.

Woodward, J. (2011). Psychological studies of causal and counterfactual reasoning. In H. Hoerl, T. McCormack, & S. R. Beck (Eds.), *Understanding counterfactuals, understanding causation* (pp. 16–53). Oxford University Press.

Woodward, J. (2012). Causation, interactions between philosophical theories and psychological research. *Philosophy of Science*, *79*, 961–972. https://doi.org/10.1086/667850

AUTHOR BIOGRAPHIES

Pascale Willemsen is assistant professor at the Institute for Philosophy II at Ruhr-University Bochum. Before that, she was a postdoctoral researcher at the Department of Experimental Psychology at University College London, working at the Causal Cognition. Lab Pascale wrote her dissertation on the causal and moral relevance of omissions.

Lara Kirfel is a PhD student at the Department of Experimental Psychology at University College London. She holds an MA in philosophy from King's College London and an MSc in psychology from University College London.

How to cite this article: Willemsen P, Kirfel L. Recent empirical work on the relationship between causal judgements and norms. *Philosophy Compass*. 2019;14:e12562. https://doi.org/10.1111/phc3.12562