Resituating the Relevance of Alternatives for Causal Attributions[1]

Justin Sytsma

Abstract: Phillips et al. (2015) provide what looks like compelling evidence in favor of explaining the impact of broadly moral evaluations on causal attributions in terms of the relevance of alternative possibilities. As part of a series of manipulation studies, they found that asking participants to describe what an agent could have done differently in a neutral case (a case in which information about broadly moral considerations was removed) showed a similar effect on causal attributions to varying the morality of the agent's action. Phillips and colleagues take this to show that broadly moral evaluations impact the alternative possibilities people see as relevant, which in turn impact their attributions. This study leaves open the possibility that the manipulation impacts people's broadly moral evaluations which in turn impact their attributions, however, rather than directly impacting their attributions. But this alternative model conflicts with Phillips et al.'s account, while being compatible with competing explanations. These two models are tested using the same manipulation method. The results support the alternative model, indicating that the mechanism proposed by Phillips et al. cannot be presumed.

Broadly moral considerations have been repeatedly shown to have a notable impact on causal attributions.[2] Perhaps the most discussed example is the Pen Case from Knobe and Fraser (2008):

> The receptionist in the philosophy department keeps her desk stocked with pens. The administrative assistants are allowed to take the pens, but faculty members are supposed to buy their own.
>
> The administrative assistants typically do take the pens. Unfortunately, so do the faculty members. The receptionist has repeatedly e-mailed them reminders that only administrative assistants are allowed to take the pens.

[1] Forthcoming in *Oxford Studies in Experimental Philosophy, Volume 5*. I want to thank David Rose and two anonymous reviewers for *Oxford Studies in Experimental Philosophy* for their extremely helpful comments on previous versions of this paper. I'd also like to thank David for his expert guidance on the Greedy Equivalence Searches reported.

[2] For recent surveys of work in this area, see Willemsen and Kirfel (2019) and Henne (forthcoming). Empirical work on norms and causal attributions includes Alicke (1992, 2000), Knobe and Fraser (2008), Hitchcock and Knobe (2009), Alicke et al. (2011), Sytsma et al. (2012), Reuter et al. (2014), Kominsky et al. (2015), Clarke et al. (2015), Willemsen and Reuter (2016), Willemsen (2016), Samland and Waldmann (2016), Henne et al. (2017), Icard et al. (2017), Livengood et al. (2017), Rose (2017), Kominsky and Phillips (2019), Gerstenberg and Icard (2020), Livengood and Sytsma (2020), Henne et al. (2021), and Kirfel and Lagnado (2021), among many others.

> On Monday morning, one of the administrative assistants encounters Professor Smith walking past the receptionist's desk. Both take pens. Later that day, the receptionist needs to take an important message… but she has a problem. There are no pens left on her desk.

This scenario describes a situation in which two agents jointly bring about a bad outcome by performing actions that are symmetric outside of Professor Smith violating an injunctive norm (faculty members are prohibited from taking pens). Despite this, when participants were asked to rate agreement with a causal attribution for each agent, ratings were very notably higher for Professor Smith. How are we to explain findings like this?

A number of authors have developed explanations that center on the role of counterfactuals in causal cognition, although they differ in their technical details.[3] Here, I'll focus on the version of this type of view given by Phillips et al. (2015). They argue that a key part of the explanation of the impact of broadly moral considerations on a variety of judgments, including causal attributions, is how they influence the perceived relevance of alternative possibilities. The basic idea is that norm violations make the alternative possibility on which the norm-violation does not occur seem more relevant to people, such that they are more likely to consider it. And, in scenarios like the Pen Case, if the agent did not violate the norm (e.g., if Professor Smith did not take a pen), then the outcome would not have occurred. It is held that this then leads people to judge that the norm-violating agent is more causal since that alternative possibility highlights that but for that agent's action, the outcome would not have occurred.

One of the most compelling pieces of evidence for this general type of account comes from the manipulation experiments performed by Phillips et al. (2015). Focusing on the study targeting causal attributions, they gave participants a neutral version of the Pen Case vignette in which details about the norm violation were removed. Phillips and colleagues then varied

[3] See, for example, Hitchcock and Knobe (2009), Halpern and Hitchcock (2015), Kominsky et al. (2015), Phillips et al. (2015), Icard et al. (2017), Kominsky and Phillips (2019), Morris et al. (2019), Henne et al. (2019), Quillien (2020), Gerstenberg et al. (2021), Quillien and Barlev (2022), and Henne and O'Neill (2022).

whether participants were asked to briefly describe what Professor Smith could have done differently or simply to describe the scenario. Phillips et al.'s idea was that writing about what the agent could have done differently would increase the perceived relevance of the alternative actions the agent might have taken, inducing the same effect found previously when information about broadly moral considerations was included. They argue that inducing such an effect by directly manipulating the relevance of alternative possibilities would therefore provide evidence for model **[1]** shown in Figure 1. And Phillips et al. interpret their findings as doing exactly this.

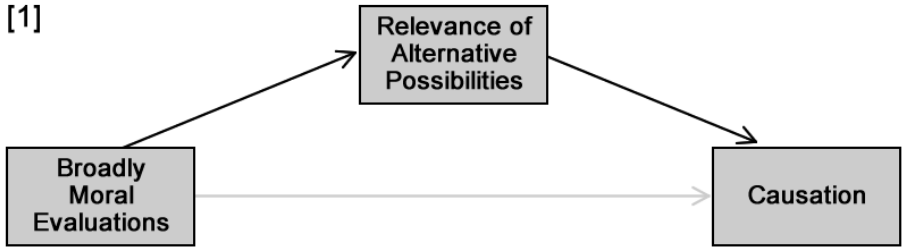

Figure 1: Model proposed by the alternative possibilities account for causal attributions, adapted from Figure 1 in Phillips et al. (2015).

There is a potential problem, however: Phillips and colleagues did not test participants' broadly moral evaluations. As such, their study leaves open the possibility that the alternative possibilities that participants considered relevant impacted their broadly moral evaluations and that these evaluations in turn impacted their causal attributions. In other words, model **[2]** in Figure 2 is also compatible with the results. But this model runs counter to the alternative possibilities account, positing that the impact of broadly moral evaluations on causal attributions primarily works directly, not through the alternative possibilities that people consider relevant.

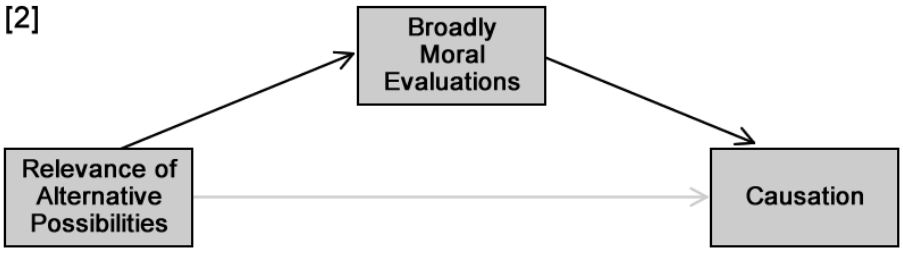

Figure 2: Proposed alternative model.

While model **[2]** runs counter to the alternative possibilities account, it is compatible with a number of competing explanations of the impact of norms on causal attributions in the literature. This includes Sytsma and Livengood's *responsibility account*, Samland and Waldmann's *pragmatic account*, and Alicke's *bias account*. The responsibility account contends that broadly moral evaluations matter for causal attributions because the ordinary concept of causation at play in such attributions is a broadly moral concept, akin to the concepts of responsibility and accountability (e.g., Sytsma et al. 2012, 2019; Livengood et al. 2017; Sytsma 2022a). As such, this account denies that the impact of broadly moral evaluations on causal attributions works primarily through the alternative possibilities people consider. Since the responsibility account is silent on how people arrive at their broadly moral evaluations, however, it allows that the alternative possibilities we consider relevant might play a role in those evaluations. As such, this account is compatible with the alternative explanation of Phillips et al.'s findings. And similarly, for the pragmatic and bias accounts.[4]

[4] The pragmatic account holds that ordinary concept of causation is non-normative, but that for pragmatic reasons participants read the questions in probes like that given by Knobe and Fraser (2008) to instead be asking about a normative concept like responsibility or accountability (Samland and Waldmann 2016, Samland et al. 2016). Their account is also silent about how we arrive at our broadly moral evaluations, however, and as such is compatible with those evaluations being influenced by alternative possibilities. The bias account holds that people's judgments tend to be biased by their desire to blame or praise the agent (Alicke 1992, 2000; Alicke et al. 2011; Rose 2017; Rogers et al. 2019). And this account is compatible with considerations of alternative possibilities playing a role in our broadly moral evaluations and those evaluations then impacting blame and praise judgments.

Thus, determining whether the effect of Phillips et al.'s manipulation reflects the direct or the indirect impact of the relevance of alternative possibilities on causal attributions is important for determining whether their results support the alternative possibilities account over competitors. And similar points could be made for accounts of the impact of broadly moral considerations on other judgments, such as judgments of freedom (e.g., Phillips and Knobe 2009), doing versus allowing (e.g., Cushman et al. 2008), and intentionality (e.g., Knobe 2003). To test the mechanism behind the effect shown by Phillips et al., I replicated their manipulation study for causal attributions, adding two further questions—one about a broadly moral attribution and one about a responsibility attribution. The results suggest that the impact of the relevance of alternative possibilities on causal attributions primarily works through people's broadly moral evaluations.

Methods

Participants were recruited through advertising for a free personality test on Google. In addition to answering the questions reported below, participants were asked basic demographic questions and after the test questions were given a 10-item Big Five personality inventory. Responses were collected from 210 native English-speakers, age 16 or older.[5] The sample size was selected to give a power greater than 90% to detect an effect of the size reported by Phillips et al. (d=0.42) using a one-tailed test.

Each participant read the vignette for the neutral version of the Pen Case from Phillips et al.'s second study:

> The receptionist in the philosophy department keeps her desk stocked with pens. Both the administrative assistants and the faculty members are allowed to take the pens, and both the administrative assistants and the faculty members typically do take the pens. The

[5] 63.3% women (three non-binary), average age 41.4 years, ranging from 16 to 80.

> receptionist has repeatedly e-mailed them reminders that both administrators and professors are allowed to take the pens.
>
> On Monday morning, one of the administrative assistants encounters Professor Smith walking past the receptionist's desk. Both take pens. Later, that day, the receptionist needs to take an important message... but she has a problem. There are no pens left on her desk.

Participants were then either asked to describe the events of the story (control) or to write about what other decision Professor Smith could have made (alternatives). Finally, they rated their agreement or disagreement with each of three claims on a scale from 1 ("completely disagree") to 7 ("completely agree"):

> (Wrong) It was wrong for Professor Smith to take a pen.
> (Responsible) Professor Smith is responsible for the problem.
> (Caused) Professor Smith caused the problem.

The three claims were presented in random order.

Results

Results are shown in Figure 3 and the data is available at https://osf.io/sgu62. The original study successfully replicated. Further, analyzing Phillips et al.'s results together with those for the present study, an ANOVA looking at ratings for Caused with *study* and *condition* as between-participant factors showed no significant effects for *study*, but did show a significant main effect for *condition*, $F(1, 440)=15.0$, $p<0.001$, $\eta^2=0.033$. This suggests that including the two additional questions did not have a notable effect on responses to Caused.

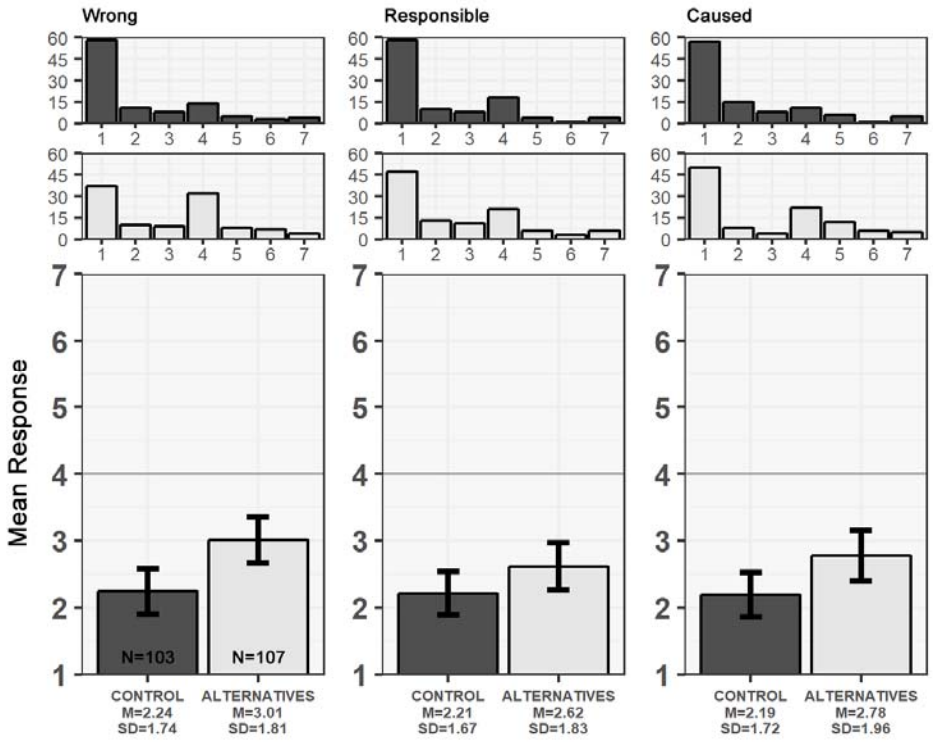

Figure 3: Results showing 95% confidence intervals; histograms above.

In line with the effect reported by Phillips et al., the mean rating for Caused in the alternatives condition was significantly higher than the mean rating in the control condition, $t(206.18)=2.29$, $p=0.01$, $d=0.32$.[6] A similar effect was found for Responsible, $t(207.5)=1.67$, $p=0.048$, $d=0.23$, and Wrong, $t(208)=3.13$, $p=0.0010$, $d=0.43$. In line with previous results (e.g., Sytsma 2021, 2022b; Sytsma and Livengood 2021), there was a strong correlation between Responsible and Caused, $r=0.76$, $t(208)=17.1$, $p<0.001$. Further, no significant difference was found between them for either the control condition, $t(102)=0.18$, $p=0.85$ (two-tailed), $d=0.011$, or the alternatives condition, $t(106)=1.17$, $p=0.24$ (two-tailed), $d=0.08$. Somewhat weaker

[6] Welch's t-tests are used throughout. Since the effect at issue is directional—mean causal rating being higher in the alternatives condition—I use one-tailed tests unless specified otherwise.

correlations were found between Wrong and Responsible, r=0.54, t(208)=9.28, p<0.001, and between Wrong and Caused, r=0.47, t(208)=7.61, p<0.001.

To test the two mechanisms noted above, I used the Greedy Equivalence Search (GES) algorithm in TETRAD 6.4.0.[7] Since the alternative possibilities account does not make a specific prediction about responsibility attributions, I began by excluding this variable. The search returned model **[A]** in Figure 4, which was a good fit for the data, df=1, χ^2=0.95, p=0.33, *CFI*=1.00, *RMSEA*=0.00, *BIC*=-12.1. This model corresponds with model **[2]**—the alternative to Phillips et al.'s model shown in Figure 2—with the effect of condition on Caused working through Wrong, rather than directly as predicted by the alternative possibilities account. To test for alternative models, I ran a second GES specifying that condition has a direct effect on Caused. The search returned model **[B]** in Figure 4, df=1, χ^2=5.67, p=0.017, *CFI*=0.95, *RMSEA*=0.15, *BIC*=0.33. Following the rules of thumb given by Schermelleh-Engel et al. (2003), **[B]** is an "acceptable fit" based on p value (>0.01) and *CFI* (>0.95) but is not an acceptable fit based on χ^2 (>3df) or *RMSEA* (>0.08). In contrast, **[A]** is a "good fit" on all four measures (p value > 0.05, *CFI* > 0.97, χ^2 > 3df, *RMSEA* < 0.05).[8]

[7] GES searches over equivalence classes of models by assigning an information score (in this case BIC) to each. The algorithm starts with the null graph, then adds the edge that most improves the score and applies the edge-orientation rules in Meek (1997). This is iterated until no additions further improve the score. The algorithm then considers deletions, removing edges while doing so increases the score, and again applying Meek's orientation rules. As Danks (2016, 467) notes, GES "is now the most widely used [score-based search algorithm] since it has proven to be the most reliable." For discussion see Chickering (2002) and Danks (2016). For applications in experimental philosophy, see Rose et al. (2011), Rose and Nichols (2013), Rose (2017).

[8] As expected, bootstrap mediation analyses (Preacher and Hayes 2008) showed comparable results. First, I tested whether Wrong mediates the effect of condition on Caused using 5000 resamples. There was a significant indirect effect (95% CIs [0.13, 0.63]). I then tested whether Caused mediates Wrong using 5000 resamples. Again, there was a significant indirect effect (95% CIs [0.04, 0.50]). While both indirect effects are significant, Wrong mediates a notably larger proportion of the effect of condition on Caused (0.60) than Caused does on Wrong (0.33), and while the direct effect of condition on Caused when controlling for Wrong is not significant (p=0.34), the direct effect of condition on Wrong when controlling for Caused is significant (p=0.014).

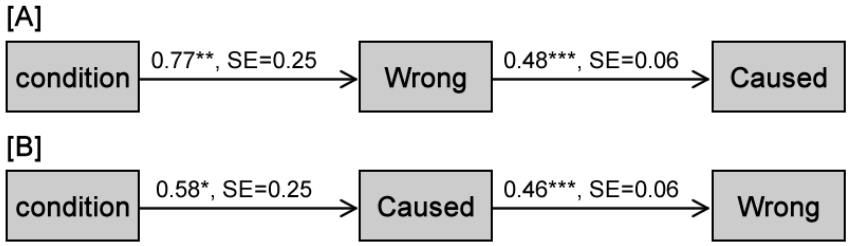

Figure 4: GES outputs from TETRAD excluding Responsible with standardized path coefficients and standard errors. Model **[A]** shows search without forcing a direct path from condition to Caused, **[B]** forcing a direct path. Asterisks indicate significance, **p*<0.05, ***p*<0.01, ****p*<0 .001.

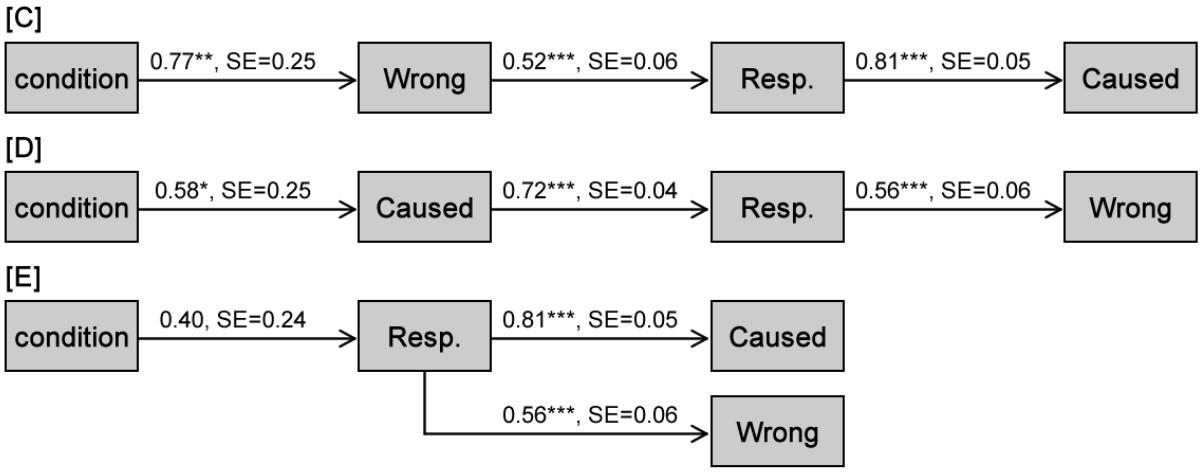

Figure 5: GES outputs from TETRAD including Responsible with standardized path coefficients and standard errors. Model **[C]** shows search without forcing a direct path, **[D]** forcing a direct path from condition to Caused, and **[E]** forcing a direct path to Responsible. Asterisks indicate significance, **p*<0.05, ***p*<0.01, ****p*<0 .001.

I then ran a GES including Responsible. The search returned model **[C]** in Figure 5, which is again a good fit for the data based on each of the four measures, *df*=3, χ^2=3.99, *p*=0.26, *CFI*=1.00, *RMSEA*=0.040, *BIC*=-12.1. This model also has the effect of condition on Caused working through Wrong, although in addition it now runs through Responsible. To test for alternative models, I first ran a GES specifying that condition has a direct effect on Caused. The search returned model **[D]** in Figure 5, *df*=3, χ^2=8.95, *p*=0.030, *CFI*=1.00, *RMSEA*=0.097, *BIC*=-

7.09. This model is a good fit based on *CFI* and an acceptable fit based on *p* value and χ^2 but is not an acceptable fit based on *RMSEA*. Finally, I ran a GES specifying that condition has a direct effect on Responsible. The search returned model **[E]** in Figure 5, df=3, χ^2=10.3, p=0.016, *CFI*=0.99, *RMSEA*=0.11, *BIC*=-5.78. This model is a good fit based on *CFI* and an acceptable fit based on *p* value but is not an acceptable fit based on χ^2 or *RMSEA*.

Discussion

The manipulation studies conducted by Phillips et al. (2015) show that asking participants to write about what an agent could have done differently impacts ratings for a range of attributions, including causal attributions, and does so in a way similar to that previously found for including information bearing directly on broadly moral considerations. They take these studies to provide direct empirical evidence for a link between the relevance of alternative possibilities and attributions that is independent from broadly moral evaluations. But these studies did not test whether the manipulation *also* impacts people's broadly moral evaluations, leaving open the possibility that the effect on attributions primarily works through such evaluations.

I tested these competing models for causal attributions, following Phillips et al.'s procedure except for also having participants rate a broadly moral evaluation and a responsibility attribution. The predicted effect was again found for causal attributions, replicating Phillips et al.'s finding, but importantly was also seen for the other two ratings. This indicates that the mechanism suggested by Phillips et al. cannot be presumed. Further, causal search indicates that the best fitting model has the impact of the manipulation on causal attributions working through broadly moral attributions, and in turn responsibility attributions. Indeed, this model was a notably better fit than the best alternative model with the manipulation directly affecting causal

attributions. This model runs counter to that proposed by Phillips et al., but is in line with alternative explanations of the impact of broadly moral considerations on causal attributions, including that given by the responsibility account. As such, these findings put pressure on Phillips and colleagues' alternative possibilities account. At the same time, it should be emphasized that this is just one piece of evidence in an ongoing debate and that the results are for just one vignette. Further work is needed to extend these findings to further scenarios and to include further measures, such as a direct measure of the perceived relevance of alternative possibilities.[9]

Finally, it is worth noting that the preferred model produced by TETRAD for the present study differs from that previously suggested for the responsibility account (Sytsma 2021), where broadly moral evaluations were modeled as a common cause for both causal attributions and responsibility attributions. Model **[C]** instead indicates a chain, with broadly moral evaluations impacting responsibility attributions, which in turn impact causal attributions. This model remains consistent with the key insight of the responsibility account, however, which is that the ordinary concept of causation at play in causal attributions is a normative concept, just as responsibility and accountability are normative concepts. Further, the strong correlation between causal attributions and responsibility attributions found in the present study lends further credence to this view.

References

Alicke, M. (1992). "Culpable causation." *Journal of Personality and Social Psychology*, 63: 368–378.

Alicke, M. (2000). "Culpable Control and the Psychology of Blame." *Psychological Bulletin*, 126(4): 556–574.

[9] I want to thank an anonymous reviewer for this suggestion.

Alicke, M., D. Rose, and D. Bloom (2011). “Causation, Norm Violation, and Culpable Control.” *Journal of Philosophy*, 108: 670–696.

Chickering, D. (2002). “Optimal structure identification with greedy search.” *Journal of Machine Learning Research*, 3: 507–554.

Clarke, R., J. Shepherd, J. Stigall, R. Waller, and C. Zarpentine (2015). “Causation, norms, and omissions: A study of causal judgments.” *Philosophical Psychology*, 28(2): 279–293.

Cushman, F., J. Knobe, and W. Sinnott-Armstrong (2008). “Moral appraisals affect doing/allowing judgments.” *Cognition*, 108: 281–289.

Danks, D. (2016). “Causal Search, Causal Modeling, and the Folk.” In J. Sytsma and W. Buckwalter (eds.), *A Companion to Experimental Philosophy*, pp. 463–471, West Sussex: Wiley Blackwell.

Gerstenberg, T., N. Goodman, D. Lagnado, and J. Tenenbaum (2021). “A counterfactual simulation model of causal judgments for physical events.” *Psychological Review*, 128(5): 936–975.

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

Halpern, J. and C. Hitchcock (2015). “Graded Causation and Defaults.” *British Journal for the Philosophy of Science*, 66: 413–457.

Henne, P. (forthcoming). “Experimental Metaphysics: Causation.” In A. Bauer and S. Kornmesser (Eds.), *The Compact Compendium of Experimental Philosophy*. De Gruyter.

Henne, P. and K. O’Neill (2022). “Double Prevention, Causal Judgments, and Counterfactuals.” *Cognitive Science*, 46(5): e13127.

Henne, P., L. Niemi, Á. Pinillos, F. De Brigard, and J. Knobe (2019). “A counterfactual explanation for the action effect in causal judgment.” *Cognition*, 190: 157–164.

Henne, P., K. O’Neill, P. Bello, S. Khemlani, and F. De Brigard (2021). “Norms Affect Prospective Causal Judgments.” *Cognitive Science*, 45(1): e12931.

Henne, P., Á. Pinillos, and F. Brigard (2017). “Cause by Omission and Norm: Not Watering Plants.” *Australasian Journal of Philosophy*, 95(2): 270–283.

Hitchcock, C. and J. Knobe (2009). “Cause and Norm.” *The Journal of Philosophy*, 106: 587–612.

Icard, T., J. Kominsky, and J. Knobe (2017). “Normality and Actual Causal Strength.” *Cognition*, 161: 80–93.

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

Knobe, J. (2003). “Intentional action and side effects in ordinary language.” *Analysis*, 63: 190–193.

Knobe, J. and B. Fraser (2008). “Causal judgments and moral judgment: Two experiments.” In W. Sinnott-Armstrong (ed.), *Moral Psychology, Volume 2: The Cognitive Science of Morality*, pp. 441–447, Cambridge: MIT Press.

Kominsky, J., J. Phillips, T. Gerstenberg, D. Lagnado, and J. Knobe (2015). “Causal superseding.” *Cognition*, 137: 196–209.

Kominsky, J. and J. Phillips (2019). “Immoral Professors and Malfunctioning Tools: Counterfactual Relevance Accounts Explain the Effect of Norm Violations on Causal Selection.” *Cognitive Science*, 43(11): e12792.

Livengood, J. and J. Sytsma (2020). “Actual Causation and Compositionality.” *Philosophy of Science*, 87(1): 43–69.

Livengood, J., J. Sytsma, and D. Rose (2017). “Following the FAD: Folk attributions and theories of actual causation.” *Review of Philosophy and Psychology*, 8(2): 274–294.

Meek, C. (1997). *Graphical Models: Selecting Causal and Statistical Models*. PhD Thesis, Carnegie Mellon University.

Morris, A., J. Phillips, T. Gerstenberg, and F. Cushman (2019). “Quantitative causal selection patterns in token causation.” *PLOS ONE*, 14(8): e0219704.

Phillips, J. and J. Knobe (2009). “Moral judgments and intuitions about freedom.” *Psychological Inquiry*, 20: 30–36.

Phillips, J., J. Luguri, and J. Knobe (2015). “Unifying Morality’s Influence on Non-moral Judgments: The Relevance of Alternative Possibilities.” *Cognition*, 145: 30–42.

Preacher, K. and A. Hayes (2008). “Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models.” *Behavior Research Methods*, 40: 879–891.

Quillien, T. (2020). “When do we think that X caused Y?” *Cognition*, 205: 104410.

Quillien, T. and M. Barlev (2022). “Causal judgment in the wild: evidence from the 2020 US presidential election.” *Cognitive Science*, 46(2): e13101.

Rogers, R., M. Alicke, S. Taylor, D. Rose, T. Davis, and D. Bloom (2019). “Causal deviance and the ascription of intent and blame.” *Philosophical Psychology*, 32(3): 404–427.

Rose, D. (2017). “Folk Intuitions of Actual Causation: A Two-pronged Debunking Explanation.” *Philosophical Studies*, 174(5): 1323–1361.

Rose, D., J. Livengood, J. Sytsma, and E. Machery (2011). “Deep trouble for the deep self.” *Philosophical Psychology*, 25: 629–646.

Rose, D. and S. Nichols (2013). “The lesson of bypassing.” *Review of Philosophy and Psychology*, 4: 599–619.

Samland, J., M. Josephs, M. Waldmann, and H. Rakoczy (2016). “The Role of Prescriptive Norms and Knowledge in Children’s and Adults’ Causal Selection.” *Journal of Experimental Psychology: General*, 145(2): 125–130.

Samland, J. and M. R. Waldmann (2016). “How prescriptive norms influence causal inferences.” *Cognition*, 156: 164–176.

Schermelleh-Engel, K. and H. Moosbrugger (2003). “Evaluating the Fit of Structural Equation Models: Tests of Significance and Descriptive Goodness-of-Fit Measures.” *Methods of Psychological Research Online*, 8(2): 23–74.

Sytsma, J. (2022a). “The Responsibility Account.” In P. Willemsen and A. Wiegmann (Eds.), *Advances in Experimental Philosophy of Causation*, 145–164, Bloomsbury Press.

Sytsma, J. (2022b). “Crossed Wires: Blaming Artifacts for Bad Outcomes.” *The Journal of Philosophy*, 119(9): 489–516.

Sytsma, J. (2021). “Causation, Responsibility, and Typicality.” *Review of Philosophy and Psychology*, 12: 699–719.

Sytsma, J., R. Bluhm, P. Willemsen, and K. Reuter (2019). “Causal Attributions and Corpus Analysis.” In E. Fischer and M. Curtis (Eds.), *Methodological Advances in Experimental Philosophy*, Bloomsbury, 209–238.

Sytsma, J. and J. Livengood (2021). “Causal Attributions and the Trolley Problem.” *Philosophical Psychology*, 34(8): 1167–1191.

Sytsma, J., J. Livengood, and D. Rose (2012). “Two types of typicality: Rethinking the role of statistical typicality in ordinary causal attributions.” *Studies in History and Philosophy of Biological and Biomedical Sciences*, 43: 814–820.

Willemsen, P. (2016). “Omissions and expectations: A new approach to the things we failed to do.” *Synthese*, 195: 1587–1614.

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

Willemsen, P. and K. Reuter (2016). “Is there really an omission effect?” *Philosophical Psychology*, 29: 1–18.