Absences and Alternatives: Reasoning about Causal Exceptions Resembles, but is Not Explained by, Reasoning about Generics

Elizabeth Lapidow (elapidow@uwaterloo.ca) & Stephanie Denison
Department of Psychology, University of Waterloo

Abstract

All causal relations are subject to exceptions. Theories of how causal cognition handles exceptions focus primarily on their *frequency* or *probability,* but does the *nature* of exceptions also influence reasoning? In two experiments, we presented adults with information about two candidate causes. Both produced the same effect in four of six cases. In the remaining two cases, one candidate produced no discernible effect while the other produced a similar but distinct effect. Participants were sensitive to this difference: despite the two candidates' identical probabilities of producing the majority effect, participants had strong preferences about which candidate was acceptable to cite in causal claims (Experiment 1) and employ in causal interventions (Experiment 2). These preferences held even in the absence of generic language and appear to be sensitive to complex and specific intuitions. This study offers novel insights into how causal reasoning understands and navigates exceptions in a rich, conceptual way.

Keywords: causal reasoning; generics; probability; determinism; generalization; domain differences

Introduction

Exceptions abound within causal reasoning. We readily accept, for example, '*smoking causes lung cancer*,' even while knowing that not every smoker, or even the majority of smokers, develops the disease. Theories of causal cognition account for this in a number of ways, for example, by highlighting the role of background conditions or inhibitors (e.g., Pearl, 2000), positing concepts of 'stability,' 'exportability,' and so on (e.g., Lombrozo, 2010; Woodward, 1997). Several theories of causal cognition propose that causal knowledge is fundamentally a 'probabilistic contrast' drawn from observed frequencies (e.g., Cheng, 1997, 2000) or takes the form of probabilistic models where relationships hold with different weights (e.g., Gopnik et al., 2004; Griffiths & Tenenbaum, 2009).

However, these accounts largely speak to how our causal cognition interprets the *frequency* of exceptions (i.e., how often X fails to bring about Y), rather than on what *meaning* or interpretation reasoners *place on the fact* that X sometimes fails to bring about Y. A great deal of how causal reasoners understand and interpret exceptions remains unknown. For example, consider the side effects of two experimental medications: 60% of patients taking Medication A experience headaches, and 40% experience no side effects; while 60% of patients taking Medication B experience headaches, and 40% experience stomach aches. Both medications bring about headaches with the same frequency, but to say "*Medication A causes headaches*" feels obviously true in a way that "*Medication B causes headaches*" does not. Here, we report results of an empirical examination into this aspect of causal exceptions—that is, whether the *nature* of exceptions, rather than their frequency, influences causal representations and reasoning.

Specifically, we ask whether and how causal reasoning about exceptions parallels reasoning about exceptions to *generics*. Generic statements, such as '*ducks lay eggs*,' convey information about kinds and are found across human language (Gelman, 2009). Furthermore, like causal claims, generics are subject to exceptions (e.g., not every duck, or even the majority of ducks, lays eggs). Connecting these two areas of reasoning may initially seem trivial; after all, causal claims and generics capture and convey *generalizations*, and useful generalizations have exceptions (e.g., Cartwright, 1983). However, our intuitive reasoning about exceptions to generic statements provides an intriguing starting point for investigating this underexplored aspect of causal reasoning. In particular, exceptions to generics are inevitable, but only some exceptions are tolerated. Compare, for instance, the apparent truth of '*ducks lay eggs*' to the apparent untruth of '*ducks are female.*' Leslie (2007) suggests that this difference lies, at least in part, in the nature of the statements' exceptions (see also Cohen, 2004). Generics accepted as true have exceptions that are *negative*, absences of the cited property (i.e., ducks that do not lay eggs), rather than exceptions that have *positive*, alternative properties (i.e., ducks that are male). Recently, Simmons and Gelman (2025) conducted the first empirical test of this theory of generics. They find that, indeed, both children and adults intuitively expect novel generics (e.g., "*Wugs have blue horns*") to refer to categories with negative exceptions (e.g., novel animals with either blue horns or no horns), not positive exceptions (e.g., novel animals with either blue horns or red horns).

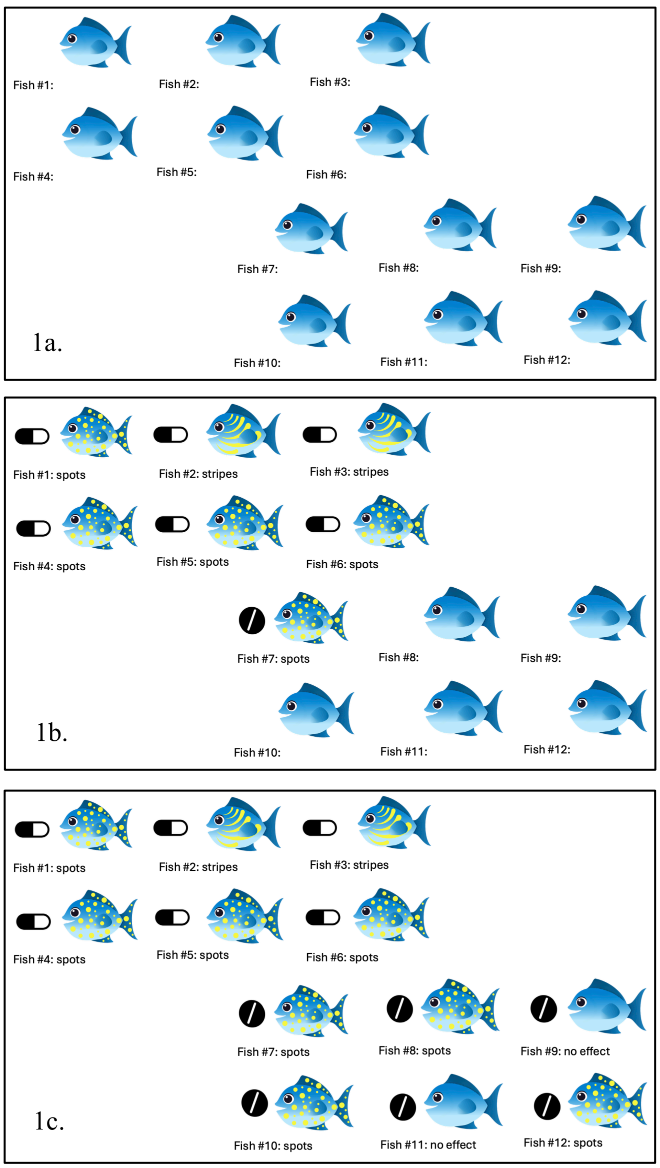
Does the same kind of content-focused evaluation of exceptions also influence causal reasoning? If so, it would shed light on how causal cognition navigates exceptions beyond accounts of how frequency leads to impressions or appearances of probabilistic causes. However, there is an immediate potential objection to this suggestion: Causality is very commonly presented *using* generic statements. The acceptability of "*Medication A causes headaches*" versus "*Medication B causes headaches*" might not be a reflection of causal intuitions, but instead an instance of intuitions about

In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana and T. Torrent (Eds.), *Proceedings of the 48th Annual Conference of the Cognitive Science Society*

generic language applied to causality. This argument is plausible and has been suggested previously. Frosch and Johnson-Laird (2011) propose that the intuitive lay concept of causality is deterministic and does not, despite appearances, allow for exceptions. The apparent tolerance of exceptions to causal claims such as "*smoking causes lung cancer*" derives, they argue, not from the fact that the claim expresses a causal relation, but from the fact that the claim expresses that relation generically, without an explicit universal qualifier (e.g., '*all'*).

This objection resonates with the close interrelationships among causality, explanation, and generalization in everyday reasoning (e.g., Lombrozo & Vasilyeva, 2017). However, to our knowledge, it has not been tested empirically. Furthermore, our reasoning about exceptions to causal claims has implications that do not extend to the broader class of generic statements. In addition to influencing organization of kind and category knowledge, causality also critically and uniquely implies and guides *action*. Indeed, according to some accounts, the knowledge captured by the statement 'X is a cause of Y' is that 'an intervention to change X will also change Y' (e.g., Woodward, 2010). The current study will therefore examine whether sensitivity to exceptions is merely an effect of generic language or extends beyond evaluating claims to how causal information is used for intervention.

Figure 1. Example animation used in the *Biological* trial.



We designed a task modelled on Simmons and Gelman (2025) to determine whether causal reasoners are sensitive to not only the probability or frequency of exceptions but also their conceptual form (i.e., *absence* and *alternative*). Participants learned about two candidate causes, both of which produce the same outcome (*target effect*) in the majority of cases (Figure 1). Candidates differ in the minority cases. One has *alternative exceptions* and produces an outcome similar to, but clearly distinct from, the *target effect*. The other has *absence exceptions* and produces no discernible outcome. In Experiment 1, we then presented participants with a claim that an unnamed candidate causes the *target effect* and asked them to choose which of the two was being referred to. In Experiment 2, participants were instead asked to select one of the candidates as the better intervention for bringing about the *target effect*. Critically, the statistical probability of each candidate producing that outcome was always identical. Therefore, any preference in selecting between them would indicate that causal reasoners are sensitive to the conceptual nature of exceptions, rather than just their frequency.

Experiment 1

The design and analysis plan for Experiment 1 was preregistered: https://aspredicted.org/vgrh-hnyg.pdf.

Participants

Data was collected from 124 adults, recruited through the University of Waterloo undergraduate participant pool. Our preregistered sample size, based on Simmons & Gelman (2025), was 120. We slightly exceeded this number as a consequence of reopening the participation pool following exclusions. An additional 15 individuals were excluded and replaced following preregistered criteria: noncompletion (n = 7), incorrect responses on one or both control trials (n = 5), or two or more of three attention check questions (n = 3).

Materials

The task used five sets of stimuli (three test and two control). For each set, we created a cover story to fit causal claims in the form of "[Object] causes [subject] to [target effect]" and animated GIFs to accompany them. Animations began by showing two groups of six identical 'subjects' (Fig. 1a), each with a label (e.g., "Subject #1"). During the animation, one of two 'objects' would appear next to each subject (Fig. 1b), and one of three outcomes would occur: the *target effect* (a change to the subject, cited in the causal claim), the *alternative exception* (a similar, but different change from the *target*), or the *absence exception* (no change to the subject).

We created three different cover stories for the test trials, one each for the *Psychological, Biological,* and *Mechanical* domains. Simmons and Gelman (2025) found that preference for *absence* over *alternative* exceptions could not be explained by differences in salience; however, the stimuli were selected such that the *'target'* and *'alternative'* effects would be clearly distinct, but without obvious differences in valence and/or value. In *Biological* (Figure 1), 'subjects' (blue fish) either develop a yellow pattern—either spots (*target*) or stripes (*alternative*)—or remain unchanged (*absence*). In *Psychological*, the 'objects' (two different designs of puppet) either caused 'subjects' (children, represented by neutral expression emoticons) to laugh (*target*, emoticon changes to laughing face), to clap (*alternative*, emoticon changes to smiling face with clapping hands), or had no effect (*absence*, neutral expression emoticon does not change). In *Mechanical*, the 'objects' (two different types of power sources) either caused 'subjects' (propellers, still images of fan) to spin clockwise (*target*, image spins to the right and a right facing arrow appears), spin counterclockwise (*alternative*, image spins left and a left facing arrow appears), or had no effect (*absence*, still fan image does not change).

In the three test trials, the *target effect* occurred in four of the six subjects, and either the *alternative* or *absence exception* occurred in the other two (Figure 1c). In the control trials, non-target outcomes were either all *alternative exceptions* (e.g., dough rising or collapsing) or all *absence exceptions* (e.g., plants growing or not), but the *target effect* occurred in five out of six subjects for one object and in only one out of six subjects for the other object. Since one of the two objects was an objectively stronger cause than the other, these trials both helped obscure the experiment's focus from participants and served as a comprehension check. Participants who did not select the stronger cause on both trials ($n = 5$) were excluded and replaced.

Procedure

Testing was online and asynchronous. Participants were initially told that they would be asked to read scenarios and make judgments about them. There were three test trials (*Psychological, Biological, Mechanical*) and two control trials. Trials were presented in a counterbalanced order, with test trials always presented first, third, and fifth, and control trials second and fourth.

Each trial began with written text about researchers studying the effects of two objects from the same category on particular subjects. For example, the *Biological* trial opened with, "A research team is studying the effects of different vitamin varieties on fish. They are currently conducting tests of two different varieties: B-3X and HV-5. Click 'Next' to see what happened when each was tested on six different fish." When the participant clicked the next button, a GIF appeared, initially showing a still image (Figure 1a). After a short delay, the GIF began to play, showing the animations for each subject (e.g., fish) in turn (Figure 1b). One of two objects (e.g., a circular or an oblong pill) would appear next to the subject, and either the *target effect*, *alternative exception*, or *absence exception* would occur.

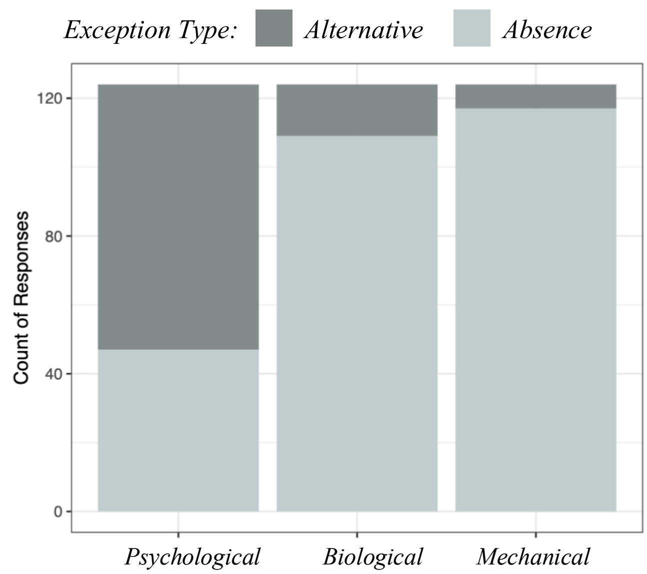
After the animation, the final frame remained on screen (Figure 1c), and the test question appeared. This question presented a causal claim about the *target effect* and asked participants to select which object it referred to. For example, in the *Biological* trial, the text said, "The next day, the research team reports: 'B-3X causes fish to develop spots.' Which one is B-3X?" Under this text were images of the oblong and circular pills that participants could select by clicking. After submitting their selection, participants were taken to the start of the next trial. After five trials, participants answered three attention-check questions in which they had to select which of four statements they had seen during the task. Participants who answered two or more of these incorrectly were excluded and replaced ($n = 3$).

Results

Figure 2 shows participants who selected the *alternative* versus the *absence exception* option as the referent of the causal claims. We find a strong preference for *absence exceptions* over *alternative exceptions* in both *Biological* (87.9%) and *Mechanical* (94.35%), but the opposite preference in *Psychological* (37.9%).

To determine whether there was a significant overall effect of exception type on causal judgments, we used generalized logistic regression[1] to predict, from average performance across participants, which type (*absence* or *alternative*) was selected on a trial. This model revealed that judgments were

Figure 2. Causes selected in Experiment 1 by domain.



[1] Following our preregistered analysis plan, random effect of participant was initially included in this model, but had to be removed as it led to a singular fit error.

indeed sensitive to exception type, with an overall preference for causes with *absence exceptions* (0.73) over causes with *alternative exceptions* (0.27, $p < 0.001$).

At the same time, the results also revealed that this sensitivity was not uniform. We conducted a generalized logistic mixed-effects regression to predict the choice of *alternative exception* options, with a fixed effect of domain (*Psychological, Biological, Mechanical*) and random effect of participant. *Psychological* served as the reference level, and the intercept was significantly different from zero ($\beta = 0.63$, $p < 0.01$), indicating that the choice of *alternative exceptions* on this trial differed from chance (estimated probability = 65.17%). In comparison to *Psychological*, selection of *alternative exception* causes was significantly lower in both *Biological* ($\beta = -2.99$, $p < 0.001$) and *Mechanical* ($\beta = -3.87$, $p < 0.001$). The estimated probability of selecting the *alternative exception* in these two domains was 4.81% for *Biological* and 2.03% for *Mechanical*.

Discussion

The results of Experiment 1 establish that causal reasoners are sensitive to the nature, and not just the frequency, of observed exceptions and that this sensitivity influences their inferences. Even though participants observed both options produce the target outcome with the same probability (66%), they did not judge them equally likely candidates for a causal claim. Instead, their intuitions were determined by the conceptual form of the exception events and the domain in which they occurred.

This finding echoes and extends the effect of positive versus negative exceptions posited to underlie the evaluation of generic statements (Leslie, 2009; Simmons & Gelman, 2025). Participants in the current study were similarly willing to accept causal claims, despite observing exceptions, but this depended on whether those exceptions were *absences* or *alternatives*. In most trials, the direction of sensitivity was also similar: participants felt strongly that "X causes Y" was appropriate for a cause that was only observed bringing about Y, even though it did not always do so, rather than a cause that sometimes brought about effects other than Y.

At the same time, the results of Experiment 1 suggest that causal reasoners are not merely applying their intuitions about generics to causal claims. If they were, the preference for *absence* exceptions should have been observed across all test trials, since all included a generic causal claim. However, participants' reaction to exception-type in *Biological* and *Mechanical* was reversed in the *Psychological* domain. This pattern of sensitivity cannot be explained as merely a consequence of stating causality in generic terms (e.g., Frosch & Johnson-Laird, 2011). It suggests instead that reasoning about exceptions in generics and causality stems from distinct, if similar, intuitions, and that what underlies the behaviour observed here is potentially specific to reasoning about causal relationships.

Before speculating about this domain difference and what underlies sensitivity to exception-type in causal reasoning (we return to both in the *General Discussion*), we seek to conduct a more direct test of the finding's independence from generic language. As mentioned above, both generic statements and causal claims convey generalizations, but causality is uniquely tied to action. If sensitivity to *absence* versus *alternative* exceptions is truly a reflection of causal reasoning, rather than merely a by-product of generics, then it should also appear in reasoning about actions. We test this hypothesis in Experiment 2 by presenting the same cover stories and stimuli as Experiment 1, but asking participants to select which option is the better causal intervention.

If our causal representations capture exceptions in a rich, more nuanced way than by tracking frequency alone, then choice behaviour should resemble Experiment 1. However, it might easily be that exception type is irrelevant to reasoners when evaluating actions based on causal evidence. If an agent needs or wants to produce a specific outcome, then an alternative and an absence are both equally *not* that target. In such a situation, the probability of producing the target outcome may be overridingly salient, with differences in exception type not influencing consideration. If this were the case, then we would expect participants to choose between the options in Experiment 2 at chance and to find no differences between domains.

Experiment 2

The design and analysis plan for Experiment 2 was preregistered: https://aspredicted.org/d4ji9n.pdf.

Participants

Participants were 84 adults recruited from the University of Waterloo undergraduate participant pool in exchange for academic credit and randomly assigned to one of two conditions (Evaluation: N = 40, Intervention: N = 44). Following our preregistration, our sample size was determined by collecting a pilot sample of 20 participants and using the smir package to conduct a power analysis. This analysis indicated a sample size of 40 per condition would have between 83.45% and 98.75% power to find an effect.

An additional 15 individuals were tested but excluded and replaced due to one of the three preregistered criteria: not completing the task (n = 4), incorrectly responding on one or both of the control trials (n = 2), or on more than one of the three attention check questions (n = 9).

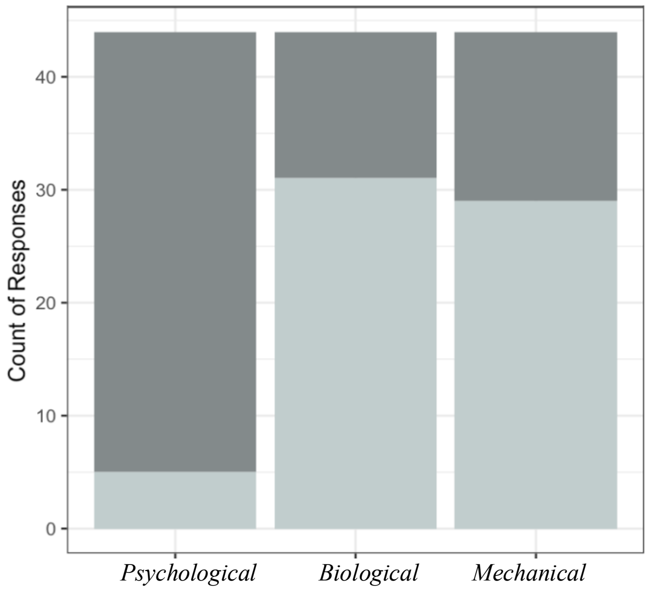
Method

The stimuli and procedure of Experiment 2 were identical to those of Experiment 1. In the *Evaluation* condition, the question was identical, providing a comparison group for the current experiment and a test of the reproducibility of Experiment 1's results. In the *Intervention* condition, participants were not presented with causal claims; instead, they were asked, "Choose the best [object] for making [target effect] occur in [subjects]." For example, "Choose the best variety for making fish develop spots", in *Biological*, again with a forced-choice between the two objects.

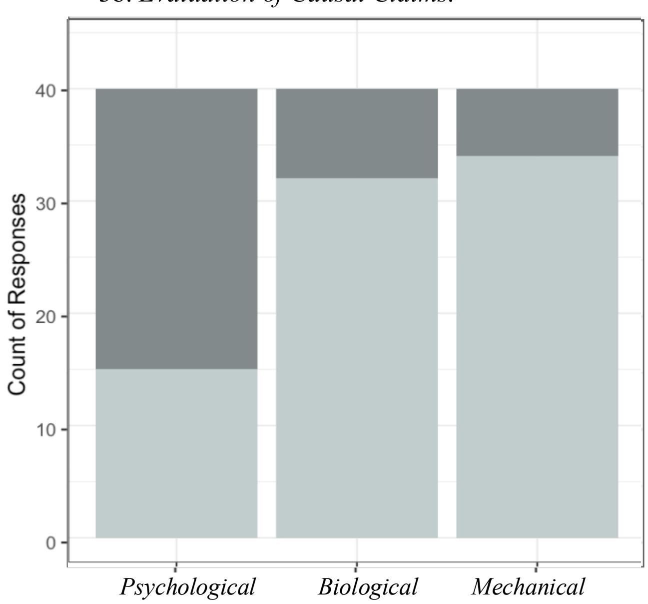
Figure 3. Causes selected in Experiment 2.

Exception Type: *Alternative* *Absence*

3a. *Choice of Causal Interventions*.



3b. *Evaluation of Causal Claims*.



Results

Figure 3 shows the number of participants who selected the *alternative* or *absence exception* options in Experiment 2. The *Evaluation* condition (Figure 3b) replicated the results of Experiment 1, with participants preferring *absence exceptions* in *Biological* (80%) and *Mechanical* (85%), but the opposite preference in *Psychological* (37.5%). Responses in the *Intervention* condition (Figure 3a) follow the same pattern, with the majority of participants selecting *absence exceptions* in *Biological* (70.45%) and *Mechanical* (65.91%), but the *alternative exception* in *Psychological* (88.64%).

To determine whether exception type influences the selection of causal interventions as well as the evaluation of causal claims, we conducted a generalized logistic mixed-effects model with domain (within-subjects), condition (between-subjects), and their interaction as fixed effects, and participant as a random effect.

None of the interaction effects were significant (*ps* > 0.1), but the model revealed significant effects of both condition and domain. There was a significant negative effect of condition ($\beta = -1.85$, $p < 0.01$), indicating that, when collapsing across test trials, participants in the *Evaluation* condition selected *alternative exception* options less often than those in the *Intervention* condition (estimated prob ability of 13.6% for *Evaluation*, 86.4% for *Intervention*). At the same time, there was also a significant effect of domain, collapsing across the two conditions. As in Experiment 1, participants selected *alternative exception* options significantly less often in *Biological* ($\beta = -3.62$, $p < 0.001$) and *Mechanical* ($\beta = -3.35$, $p < 0.001$) as compared to *Psychological*. Again, *Psychological* (the reference level) was significantly greater than chance ($\beta = 2.5$, $p < 0.001$), with an estimated 92.45% probability of selecting the *alternative-exception* compared to 2.61% for *Biological* and 3.38% for *Mechanical*.

This analysis provides important validation of the results of Experiment 1. Following this, we examined how reasoning about causal claims compared to reasoning about causal interventions using separate models for each condition[2], with the domain as a fixed-effect in both. In *Evaluation*, the preference for *alternative-exceptions* in *Psychological* did not reach significance (estimated probability = 62.5%, $p = 0.12$), but was still significantly increased compared to the *Biological* (estimated probability of selecting *alternative-exceptions* = 13.04%, $p < 0.001$) and *Mechanical* (estimated probability = 9.57%, $p < 0.001$) domains. In *Intervention*, the preference for *alternative-exceptions* was significantly above chance in *Psychological* (estimated probability = 95.32%, $p < 0.001$) and significantly lower in the *Biological* (1.23%, $p < 0.001$) and *Mechanical* (1.69%, $p < 0.001$).

Discussion

The results of Experiment 2 establish that sensitivity to the nature of exceptions in causal reasoning is not merely a byproduct of describing causality using generic language. Even when no claims were presented, and participants were asked to use the evidence to choose an action to bring about a particular outcome, they considered more than just the observed probability of that outcome. At the same time,

[2] Following our preregistration, the model for the Intervention condition includes a random effect of participant. This random effect was also initially included in the model for the *Evaluation* condition, but this structure resulted in a convergence failure. We also deviate from the preregistered plan to report intercept only models, as these analyses are uninformative, and potentially misleading, due to the different directions of preference depending on domain.

participants clearly differentiated between *Evaluation* and *Intervention*, as the preference for *absence exceptions* (in *Mechanical* and *Biological*) was more moderate in the latter. This difference is reasonable and consistent with our expectations: when the goal is to produce a specific outcome, differences in how causes might fail to produce that outcome matter less than the likelihood of failure. Notably, however, this was not true for the *Psychological* trials. Participants not only had the opposite preference, selecting the *alternative exception* over the *absence exception*, but this preference was also markedly stronger in selecting interventions than in evaluating causal claims. Below, we consider what this difference, and those in Experiment 1, suggest about the underlying nature of causal reasoning about exceptions.

General Discussion

The current study offers insight into an underexplored aspect of causal cognition. It shows that probabilistic evidence influences causal reasoning beyond inferring relationships of different strengths (Cheng, 1997), weights (Griffiths & Tenenbaum, 2009), stability (Woodward, 1997), and so on. We find that causal reasoners make inferences from the nature, and not just the frequency, of probabilistic evidence. These results highlight the complexity with which causal reasoners consider exceptions, a consideration that resembles, but is not explained by, intuitions about exceptions to generics.

Having established sensitivity to exceptions within causal cognition, several avenues for further research suggest themselves. First and foremost, what is the nature of the intuition underlying this sensitivity? Are relationships with alternative exceptions flatly unacceptable, striking us as untrue in much the same way as '*ducks are male,*' or simply less acceptable in comparison to absence? In our force-choice task, causes were presented in the context of both types of exceptions. Participants' consideration was necessarily relative, and their responses were restricted to a binary selection. Follow-up studies asking reasoners to provide independent, continuous ratings of acceptability for different types of exceptions are currently underway in our lab.

Another clear avenue for further consideration is the apparent evidence of domain differences. In both experiments, the preference for absence over alternative exceptions in reasoners' judgements of causal relationships in the *Mechanical* and *Biological* trials was strikingly reversed in the *Psychological* trials. Of course, as we only tested one vignette per domain, it is entirely possible this was merely a stimuli-specific quirk, rather than a reflection of causal intuitions. That is an empirical question that requires further testing. However, if we, at least for now, take the significant effect of trial at face value, what does it suggest about causal cognition?

One possibility is that the psychological domain is distinct within causal reasoning. Human reasoners have, by default, an intimate awareness of the complexity of human behaviour that we lack in other domains. This awareness might conceivably lead to particular, possibly unique intuitions about causal relationships involving human agents. Indeed, prior research finds intuitive expectations that psychological events have more, and more complex, causes than physical ones (Strickland et al., 2017) and that causal relationships have a capacity to vary in ways that other domains do not allow (Lapidow & Denison, 2025; Nichols & Knobe, 2007). While intriguing and plausible, the possibility that causal reasoners have a higher tolerance for exceptions in the psychological domain cannot account for the current results. A higher tolerance for exceptions might predict that exception type would simply have less impact on participants' judgments, not that the direction of that impact would reverse, or be more pronounced in Experiment 2 than in Experiment 1.

An intriguing alternative that would be consistent with these points is the possibility that there is a domain-general intuition underlying reasoning about causal exceptions: a preference for causes that don't 'do something else.' Recently, Stephan and colleagues (2023) proposed and demonstrated that causal reasoners have a default expectation of *dilution*: that the greater the number of effects produced by a single cause, the weaker it is at producing any of those effects. Might the difference between *Psychological* and the other trials be due not to differences in domain intuitions but in whether the *alternative* struck participants as an additional, and therefore diluting, effect? Recall that the *Psychological target* outcome was a child laughing in response to a puppet. It could be that the *alternative* (clapping) struck participants not as 'similar, but distinct,' but the *same* effect. Rather than 'causing laughing' and 'causing clapping', both options can be seen as 'causing a positive audience reaction.' Such equivalence is not equally apparent in the other conditions. While both the *target* and *alternative* in *Mechanical* might be seen as 'causing movement,' participants' prior knowledge would likely suggest little difference in consequences if a puppet intended to cause laughter produces applause, whereas a part unpredictably rotating in opposite directions might easily have disastrous consequences for a mechanical system. An experiment is currently underway in our lab to test this possibility by presenting reasoners with *alternatives* that are more and less similar to the *targets*.

Finally, we are not alone in exploring the connection between causality and generics. Recent research has also explored how both are subject to scope (Ritchie et al., 2025; Vasil et al., 2025). Whether and how expectations about scope and form interact to produce intuitions about exceptions will be important to examine. In addition, adults often overgeneralize tolerance of exceptions in generics to universal statements (e.g., "*all ducks lay eggs*") (e.g., Khemlani et al., 2007; Leslie et al., 2011). While it seems unlikely that reasoners would judge "*smoking always causes lung cancer*" as true, this overgeneralization effect is robust, and the possibility warrants empirical examination.

Beyond all else, these results highlight that reasoning about probabilistic causal evidence involves a richness and complexity yet to be examined, and that offers a path towards a better understanding of the realities of causal cognition.

Acknowledgments

Special thanks to Leanne Schutt, without whose efforts in stimuli development, survey construction, and study troubleshooting, this study would not have been completed. This work was supported by the Jacobs Foundation Young Fellows Award to Elizabeth Lapidow and a Natural Sciences and Engineering Research Council of Canada Discovery Grant to Stephanie Denison.

References

Cartwright, N. (1983). *How the Laws of Physics Lie* (1st ed.). Oxford University PressOxford. https://doi.org/10.1093/0198247044.001.0001

Cheng, P. W. (1997). From covariation to causation: A causal power theory. *Psychological Review*, *104*(2), 367–405. https://doi.org/10.1037/0033-295X.104.2.367

Cheng, P. W. (2000). Causality in the mind: Estimating contextual and conjunctive power. In *Explanation and cognition* (pp. 227–253). The MIT Press.

Cohen, A. (2004). Generics and Mental Representations. *Linguistics and Philosophy*, *27*(5), 529–556. https://doi.org/10.1023/B:LING.0000033851.25870.3e

Frosch, C. A., & Johnson-Laird, P. N. (2011). Is everyday causation deterministic or probabilistic? *Acta Psychologica*, *137*(3), 280–291. https://doi.org/10.1016/j.actpsy.2011.01.015

Gelman, S. A. (2009). 6 Generics as a Window onto Young Children's Concepts. In F. J. Pelletier (Ed.), *Kinds, Things, and Stuff: Mass Terms and Generics* (p. 0). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195382891.003.0006

Gopnik, A., Glymour, C., Sobel, D. M., Schulz, L. E., Kushnir, T., & Danks, D. (2004). A Theory of Causal Learning in Children: Causal Maps and Bayes Nets. *Psychological Review*, *111*(1), 3–32. https://doi.org/10.1037/0033-295X.111.1.3

Griffiths, T. L., & Tenenbaum, J. B. (2009). Theory-based causal induction. *Psychological Review*, *116*(4), 661–716. https://doi.org/10.1037/a0017201

Khemlani, S., Leslie, S.-J., Glucksberg, S., & Fernandez, P. R. (2007). Do Ducks Lay Eggs? How People Interpret Generic Assertions. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *29*(29). https://escholarship.org/uc/item/4tm460rm

Lapidow, E., & Denison, S. (2025). Maybe She'll Say Yes: How Young Learners Acquire and Apply Knowledge about Inconsistent Causal Relationships from Different Domains. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *47*(0). https://escholarship.org/uc/item/79j7p447

Leslie, S. (2007). Generics and the Structure of the Mind. *Philosophical Perspectives*, *21*(1), 375–403. https://doi.org/10.1111/j.1520-8583.2007.00138.x

Leslie, S.-J., Khemlani, S., & Glucksberg, S. (2011). Do all ducks lay eggs? The generic overgeneralization effect. *Journal of Memory and Language*, *65*(1), 15–31. https://doi.org/10.1016/j.jml.2010.12.005

Lombrozo, T. (2010). Causal–explanatory pluralism: How intentions, functions, and mechanisms influence causal ascriptions. *Cognitive Psychology*, *61*(4), 303–332.

Lombrozo, T., & Vasilyeva, N. (2017). Causal Explanation. In *The Oxford handbook of causal reasoning* (p. 415). Oxford University Press.

Nichols, S., & Knobe, J. (2007). Moral Responsibility and Determinism: The Cognitive Science of Folk Intuitions. *Noûs*, *41*(4), 663–685. https://doi.org/10.1111/j.1468-0068.2007.00666.x

Pearl, J. (2000). *Causality: Models, reasoning, and inference*. Cambridge University Press.

Ritchie, K., Curiel, A., Zhu, T. A., Pederneschi, A., & Vasil, N. (2025). The scope of generic generalizations: Developmental changes in judgments about contextually-restricted generics. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *47*(0). https://escholarship.org/uc/item/3c17j9r6

Simmons, E., & Gelman, S. A. (2025). The role of exceptions in children's and adults' judgments about generic statements. *Cognition*, *255*, 106016. https://doi.org/10.1016/j.cognition.2024.106016

Stephan, S., Engelmann, N., & Waldmann, M. R. (2023). The perceived dilution of causal strength. *Cognitive Psychology*, *140*, 101540. https://doi.org/10.1016/j.cogpsych.2022.101540

Strickland, B., Silver, I., & Keil, F. C. (2017). The texture of causal construals: Domain-specific biases shape causal inferences from discourse. *Memory and Cognition*. https://doi.org/10.3758/s13421-016-0668-x

Vasil, N., Campos, G., Rouzbehani, B., Curiel, A., Caoile, J., Khurelkhuu, S., & Ritchie, K. (2025). Contextual restriction and faultless disagreement about generics across development. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *47*(0). https://escholarship.org/uc/item/57r2n3tm

Woodward, J. (1997). Explanation, invariance, and intervention. *Philosophy of Science*, *64*, S26–S41.