Causal and Counterfactual Reasoning about Gradual and Abrupt Events

Vanessa Cheung[1], **Cristina Leone**[1], **David Lagnado**[1], **& Samantha Kleinberg**[2]
vanessa.cheung.14@ucl.ac.uk, cristina.leone.19@ucl.ac.uk, d.lagnado@ucl.ac.uk, skleinbe@stevens.edu
[1] Department of Experimental Psychology, University College London, London WC1H 0AP, UK
[2] Department of Computer Science, Stevens Institute of Technology, Hoboken, NJ 07030, USA

Abstract

Determining what caused an event is common in everyday life, yet little is known about what aspects of real-world events affect causal attribution. Causes may unfold at multiple timescales, with gradual events (e.g., steady weight loss) and abrupt ones (e.g., acute illness) contributing to an outcome. We investigated causal attribution in real-world contexts (e.g., finance, health) where both types of events contribute to a positive or negative outcome. In Study 1, participants gave higher causal ratings to abrupt causes in negative scenarios about the environment or finance. Conversely, we found higher causal ratings for *gradual* causes of physical health regardless of outcome valence, and no significant differences in mental health scenarios. Further, participants' counterfactual responses were mostly consistent with their causal attributions. Study 2 suggests that the preference for abrupt causes may be explained by their temporal proximity to the outcome. We discuss explanations for these findings and their implications.
Keywords: counterfactual reasoning, causal attribution, time

Introduction

In real-world situations, many different causes can jointly contribute to an outcome. These causes can be categorized by a variety of different attributes, chief among which is timing. Some causes immediately precede their effects, such as a ruptured pipeline leading to an oil spill, while others take place long before the incident itself, such as a mistake in the construction of the pipeline contributing to the oil spill years later. In many cases, a more temporally proximate cause is also a sudden one with a brief duration – a burst pipeline, a lightning strike causing a fire – whereas causes that occur farther away may have a cumulative impact, such as one's investments steadily decreasing in value over time.

The relationship between suddenness and recency is not always one-to-one. One patient could be diagnosed with liver failure due to a lifetime of heavy drinking (a gradual cause). Another patient could receive the same diagnosis due to an drug-induced liver injury two decades ago, a high risk factor for liver disease later in life (a sudden cause that predates the effect). In many cases, however, it often occurs that acute, sudden causes have some immediate impact on the outcome. For example, a burst water pipe very suddenly and immediately leads to a flood in an abandoned house. Past research has often focused on how people attribute causality to events that are more temporally proximal to an outcome, but few studies have investigated these kinds of abrupt events.

In this article, we compare these two types of causes: *abrupt* causes, which most often refer to a self-contained event that unfolds instantaneously or in a short time frame and has a more immediate impact on an outcome, and *gradual* causes, which unfold incrementally over time and have a cumulative impact on the outcome.[1] How do people attribute causality to abrupt versus gradual events when they jointly cause an outcome? For instance, which event do people consider to have contributed more to flooding an abandoned house: a water pipe bursting, or a different water pipe that has been leaking for years?

Past work has indicated a tendency for people to judge the most recent causes as more important in cases of joint causation (e.g., Hart & Honoré, 1985; Lagnado & Channon, 2008; Reuter et al., 2014). This is possibly because the perceived strength of the connection between a cause and an outcome decreases as the time difference between them increases (Shanks et al., 1989). This is in line with how causal judgments of an action can be mitigated by reducing its physical and temporal proximity to the outcome (Cheung et al., 2025). How people might perceive the strength of a gradual cause, however, is a more complex question: while the threshold for triggering the outcome may occur relatively close in time to the outcome itself, the gradual event began at a point farther away in time and continued for a sustained period. This has not often been investigated in prior research, particularly in cases where multiple types of causes (i.e., both abrupt and gradual) contribute to an outcome, despite this often happening in everyday life.

Further, the reason why people might attribute more causality to an abrupt event may involve an overlap between two concepts. Some might argue that it is not only because the abrupt cause is more proximal to the outcome, but also because its sudden nature is regarded as more abnormal or atypical compared to a gradual cause. In the normality literature, an event is abnormal if it violates a prescriptive (e.g., moral) norm or a statistical norm (i.e., it rarely occurs); people typically consider both when assessing normality (Bear & Knobe, 2017). Following this, one may consider an abrupt event to be abnormal as it is more salient as a difference-making event due to how much it contrasts with the prior state. Conversely, gradual events could be perceived as less abnormal, as they usually occur for a longer time and cause a sustained, cumulative change. In addition, abrupt causes may be considered more statistically abnormal if they are perceived to be less common than gradual events. Both the abnormality (e.g., Icard et al., 2017) and the temporal proximity (e.g., Lagnado & Channon, 2008) of a cause have been found to increase its salience in cases of joint causation.

However, these definitions of (ab)normality do not clearly

[1] In this article, we use abrupt and gradual 'events' and 'causes' interchangeably.

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society*

map onto abrupt versus gradual events: while some abrupt events may be more surprising due to their sudden impact (and are therefore perceived to be more abnormal), gradual events may also be abnormal. For instance, both a leaky pipe and a burst pipe may be considered abnormal, as both have violated a functional norm (see Kominsky and Phillips, 2019). Similarly, while a gradual cause (e.g., a leaky pipe) may occur more frequently, the point by which it reaches a certain threshold to impact the outcome (e.g., enough to cause a flood) is also an infrequently occurring event – though it is less clear at what point this threshold is reached.

Prior studies manipulating normality often describe a scenario where two causes contribute to an outcome: one that frequently occurs, and one that rarely occurs. In most of these scenarios, the two competing causes are usually both abrupt events with similar timing that together led to the outcome (e.g., a sprinkler and rain contributing to a flood; Icard et al., 2017). Studies testing for normality effects have not typically included gradual causes in the scenarios, despite their prevalence in real-world contexts. This makes it an open question as to how timing and suddenness may factor into judgments.

According to past research, abrupt events might be considered more causal than gradual events. However, is this always the case across different real-world contexts, and when the events lead to good or bad outcomes? It remains an open question whether one should expect differences between real-world domains (e.g., health or finance). Causal inferences may be guided by domain-general mechanisms, but there may also be domain differences in causal attributions due to differences in experiences or familiarity (e.g., gradual causes may be more salient than abrupt causes when improving health outcomes, such as improving diet and exercise, compared to in other domains). Further, past research on causal attributions often focus on negative outcomes. Given that outcome valence can influence causal attributions (e.g., Alicke, 2000), we aimed to include both positive nad negative outcomes in this study.

Counterfactual Reasoning

How do people attribute causation in situations where multiple causes contribute to an outcome? One of the most prominent explanations is that they engage in counterfactual reasoning – they consider what would have happened if events had unfolded differently from how they actually did (Gerstenberg, Lagnado, et al., 2014; Gerstenberg et al., 2021; Pearl et al., 2000). For instance, when two events (X and Y) are present, we can consider whether Y would still be present if X had been absent. If removing X prevents Y from occurring, then we know that X is the cause of Y (Hagmayer et al., 2007). However, when imagining counterfactuals, people often do not simply remove a cause: they follow a more flexible version of counterfactual dependence (Chockler & Halpern, 2004), for example by imagining the contribution that a cause *would have* made if the situation had been slightly different (Gerstenberg & Lagnado, 2010). If changing X changes how Y occurs, then X is a cause of Y (Gerstenberg et al., 2021).

The Counterfactual Simulation Model (Gerstenberg et al., 2021) offers a more unified account of how people make causal judgments about physical events by simulating counterfactual possibilities. More specifically, it makes predictions about the extent to which different candidate causes in physical settings are perceived to have caused an outcome. However, this model mainly focused on causal attribution in physical events involving objects, and it remains unclear how well it can generalize to more complex, real-world situations involving sentient agents and events that involve changes in probability, timing, and intensity (e.g., Merck & Kleinberg, 2016). When people engage in counterfactual reasoning about more realistic situations with different types of causes, what do they change from the world?

In more naturalistic contexts, the normality of the event (Icard et al., 2017; Kirfel & Lagnado, 2021), expectations about what will happen (Gerstenberg et al., 2021), and temporal factors (Henne et al., 2021), all play a role in what kind of counterfactuals people consider. As people might have different intuitions about these factors for gradual and abrupt causes, this may affect the kinds of counterfactuals they consider and how they attribute causality.

The Present Study

We explored 1) whether people attribute more causality to gradual or abrupt causes in naturalistic, everyday situations, 2) if attributions depend on real-world domain and whether the events contribute to a positive or negative outcome, and 3) whether people are more likely to consider counterfactuals for gradual or abrupt events. To do so, we conducted two studies with realistic scenarios and investigated how participants attribute causality to abrupt and gradual causes and what counterfactual alternatives they consider.

Study 1

Study 1 explored whether people attribute more causality to the gradual or the abrupt cause across different real-world contexts. We also used a naturalistic qualitative method to test what counterfactual alternatives were more salient.

Method

Participants This study received IRB approval from Stevens Institute of Technology. We conducted an *a priori* power analysis for a mixed ANOVA. To achieve 80% power for detecting a medium effect size at an alpha-level of 0.05, 104 participants were required. We recruited 110 participants from a U.S. sample via Prolific. We excluded one participant due to incomplete responses. Our final sample size was $N = 109$ (55 male, 53 female, 1 did not indicate gender; $M_{age} = 33.3, SD_{age} = 10.4$).

Design and Materials Participants saw eight vignettes in the domains of environment, finance, physical health, and mental health. We showed all participants two different scenarios for each of the four domains. As a between-subjects manipulation, some participants were randomized to see a

positive version of a vignette (in which the events lead to a good outcome), and some a negative version (in which the events lead to a bad outcome). Each vignette described a scenario where both a gradual and an abrupt cause contribute to an outcome. We include an example vignette from the physical health domain below:

> Barbara is a relatively healthy woman in her forties. For some time, she has been incorporating some running into her day-to-day routine: building her aerobic base, increasing her mileage, and trying out routes with different types of terrain. One day, her friends convince her to join them for a week-long trip hiking a portion of the Appalachian trail. She doesn't have time to do any specific preparation, and is suddenly thrown into hiking 12-15 miles each day. The abrupt increase in activity is certainly challenging, but Barbara manages it without injury. Later, she visits the doctor and finds that her VO2 max has significantly improved compared to her checkup last year.

In this vignette, the gradual cause of Barbara's improved VO2 max is her gradually increasing fitness due to routine running, and the abrupt cause is going on the hiking trip. Importantly, we counterbalanced the order of the two causes presented in each vignette to eliminate potential order effects.

Participants saw a random selection of eight of 32 possible vignettes that varied in valence (positive, negative) with counterbalanced order (i.e., reading about the gradual event first vs. abrupt event first) (2 Valence $\times$ 2 Position $\times$ 4 Domain $\times$ 2 Vignettes per domain). Each participant saw both positive and negative vignettes in both event orders throughout the experiment, but only one version of each scenario.

In the negative version of the example vignette, Barbara's VO2 max decreased significantly compared to last year due to her skipping runs in her fitness routine (the gradual cause) and having contracted a respiratory illness which impacted her lung capacity (the abrupt cause).

For each vignette, participants completed two questions in which they rated the extent to which the gradual and abrupt events caused the outcome. Following the previous example, participants were asked to indicate their agreement with the following statements on a scale of 1 ("Strongly disagree") to 7 ("Strongly agree"): "Barbara's increased day-to-day fitness routine caused her VO2 max to improve significantly compared to last year." and "Barbara's hiking trip caused her VO2 max to improve significantly compared to last year." The order of these two statements was randomized depending on which cause was shown first in the vignette.

To see what counterfactual alternatives were more salient to participants, we also asked them an open question similar to what was used in Kirfel and Lagnado (2021):

> In the text box below, please complete the sentence by indicating what could have gone differently so that the outcome (Barbara's VO2 max increased significantly from last year) would not have occurred.
>
> *If [blank], Barbara's VO2 max levels would not have increased significantly compared to last year.*
>
> This response does not need to fit into the format of the example sentence and can be as long or as short as needed.

Data Analysis All data and analysis code are available at `https://osf.io/wah5b/`. We conducted a mixed model analysis using the `mixed` function in the `afex` package (Singmann et al., 2022) in R. In the overall analysis, we used the following variables as fixed effects: Domain (environment, finance, mental health, physical health), Valence (positive and negative), Position (reading about gradual first vs. abrupt first), Cause Type (gradual cause and abrupt cause), and Vignette (the eight different scenarios). We were primarily interested in the interaction between Domain, Valence, and Cause Type. The Position variable was included solely to counterbalance the order of the materials. We included participants as random effects.[2]

We then analyzed the results within each domain with separate mixed models using a similar procedure as the overall analysis. As we were mainly interested in the interaction between Valence and Cause Type, we conducted pairwise comparisons using `emmeans` (Lenth, 2022) with Tukey adjusted p-values.

For the counterfactual responses, the first author and a research assistant (who was blind to the research questions) independently coded qualitative responses to categorize whether it corresponded to the abrupt cause, the gradual cause, or whether it was irrelevant or ambiguous. Responses could correspond to more than one category (e.g., mentioned both causes). The inter-coder reliability for each category ranged from 85% to 94% ($M = 92\%$). The two coders then discussed any disagreements and agreed on a final decision. When analyzing the data, we excluded responses categorized as irrelevant or ambiguous. We then conducted a logistic regression using `glm` for each domain.[3]

Results

In the overall analysis of causal attributions, we found a significant interaction between Valence and Cause Type when averaging across the different domains, $F(1, 1593.88) = 7.92, p = .005$ (see Figure 1). We also found a significant interaction between Domain, Valence, and Cause Type, $F(3, 1579.99) = 5.22, p = .001$, which we examine more closely in the following sections.

Environment In the environment domain, overall, participants attributed more causality to the abrupt cause over the gradual cause, $F(1, 263.66) = 20.62, p < .001$. We found

[2] Starting from the maximal model, we reduced the terms until the model converged. Given that we only had eight vignettes (two per domain), the study was not sufficiently powered to include Vignette as a random effect. For the overall analysis, our final model was: scores $\sim$ domain $*$ framing $*$ cause type $*$ position $+$ vignette $+$ (cause type $+$ position $||$ pID).

[3] We used the following model: code $\sim$ cause type $\times$ valence $\times$ position $+$ vignette.

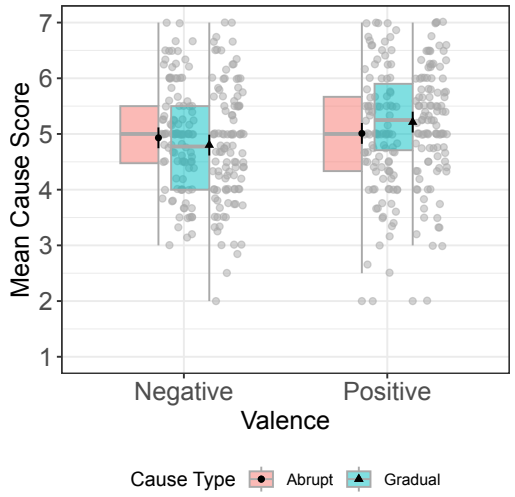

Figure 1: Overall mean causal judgments for gradual versus abrupt causes for positive and negative scenarios averaged across different domains. Error bars indicate 95% CI.

a significant interaction between Cause Type and Valence, $F(1,263.67) = 7.56, p = .006$. As shown in Figure 2, when the scenario led to a negative outcome, participants preferred the abrupt cause ($M = 5.45$, 95% CI [5.15, 5.75]) over the gradual cause, ($M = 4.53$, 95% CI [4.23, 4.83]), $t(260) = 5.18, p < .001$. There was no significant difference in causal attribution scores between the abrupt cause and the gradual cause for positive scenarios, $t(260) = 1.26, p = .209$. When looking at the individual vignettes, the preference for the abrupt cause in negative scenarios was stronger in one vignette than the other, but the interaction between Cause Type, Valence, and Vignette was not significant, $F(1,263.66) = 0.51, p = .474$.

As shown in Figure 3, we found a similar pattern of results in participants' counterfactual responses: they were more likely to consider counterfactuals for the abrupt cause in negative scenarios, $z = 4.74$, OR $= 7.55$, 95% CI [3.27, 17.44], $p < .001$, but there was no significant difference in positive scenarios, $z = 0.90$, OR $= 1.41$, 95% CI [0.66, 3.02], $p = .370$.

Finance We found a similar result in the finance domain, where, overall, participants attributed more causality to the abrupt cause, $F(1,102.15) = 11.55, p < .001$. We found no significant interaction between Cause Type and Valence, $F(1,292.24) = 0.26, p = .614$. As shown in Figure 2, participants attributed more causality to the abrupt cause ($M = 5.31$, 95% CI [5.01, 5.61]) than the gradual cause ($M = 4.90$, 95% CI [4.60, 5.20]) in negative scenarios, $t(180) = 2.09, p = .04$. They also attributed more causality to the abrupt cause ($M = 5.25$, 95% CI [4.95, 5.55]) than the gradual cause ($M = 4.70$, 95% CI [4.60, 5.20]) in positive scenarios, $t(180) = 2.79, p = .006$. This preference was consistent in individual vignettes.

This pattern was reflected in their counterfactual responses (Figure 3): they were more likely to consider counterfactuals for the abrupt cause in both negative ($z = 3.02$, OR $= 2.99$, 95% CI [1.47, 6.08], $p = .003$) and positive scenarios ($z = 4.87$, OR $= 4.78$, 95% CI [2.55, 8.96], $p < .001$).

Physical Health We found the opposite pattern of results in the physical health domain. Overall, participants assigned more causality to gradual events than abrupt events $F(1,308.95) = 78.65, p < .001$. We also found a significant interaction between Cause Type and Valence, $F(1,308.95) = 22.45, p < .001$. As shown in Figure 2, participants attributed more causality to the gradual cause ($M = 4.91$, 95% CI [4.63, 5.19]) than the abrupt cause ($M = 4.36$, 95% CI [4.08, 4.65]) for negative scenarios, $t(310) = 2.93, p = .004$. They also attributed more causality to the gradual cause ($M = 6.00$, 95% CI [5.72, 6.29]) than the abrupt cause ($M = 4.21$, 95% CI [3.92, 4.49]) for positive scenarios, $t(310) = 9.58, p < .001$. However, when looking at the individual vignettes, we did not always observe a strong preference for the gradual cause in negative scenarios, but this preference was consistent for positive scenarios (this interaction between Cause Type, Valence, and Vignette was significant, $F(1,308.95) = 22.33, p < .001$).

Participants tended to consider counterfactuals for the gradual cause in both negative ($z = -4.43$, OR $= 0.25$, 95% CI [0.14, 0.47], $p < .001$) and positive scenarios ($z = -6.80$, OR $= 0.07$, 95% CI [0.03, 0.15], $p < .001$).

Mental Health We found no significant difference in how participants attributed causality to abrupt and gradual events in the mental health domain, $F(1,101.14) = 0.00, p = .973$. We found no significant interaction between Cause Type and Valence, $F(1,312.09) = 0.83, p = .364$. Participants did not significantly prefer one cause type when considering counterfactuals in negative scenarios, $z = -1.92$, OR $= 0.50$, 95% CI [0.24, 1.01], $p = .055$, but this difference was significant in positive scenarios with more mentions of counterfactuals for the gradual cause, $z = -2.00$, OR $= 0.53$, 95% CI [0.28, 0.99], $p = .045$ (see Figure 3).

Discussion

Overall, participants attributed more causality to the abrupt cause in scenarios about the environment or finance with a bad outcome. Conversely, they attributed more causality to the *gradual* cause in the physical health scenarios regardless of valence, and we found no difference in causal judgments in the mental health scenarios. However, in all vignettes, the abrupt event occurred closer in time to the outcome than the gradual event, raising the question of whether participants' causal attributions are due to differences in temporal proximity between abrupt and gradual causes.

Study 2

In Study 1, we found a difference in causal attributions for gradual and abrupt events in some domains. However, this may have been confounded by the timing of events, as gradual ones tended to occur earlier. To investigate this, Study 2 added temporal order as a manipulation (i.e., the abrupt cause can either occur more or less recently than the gradual cause).

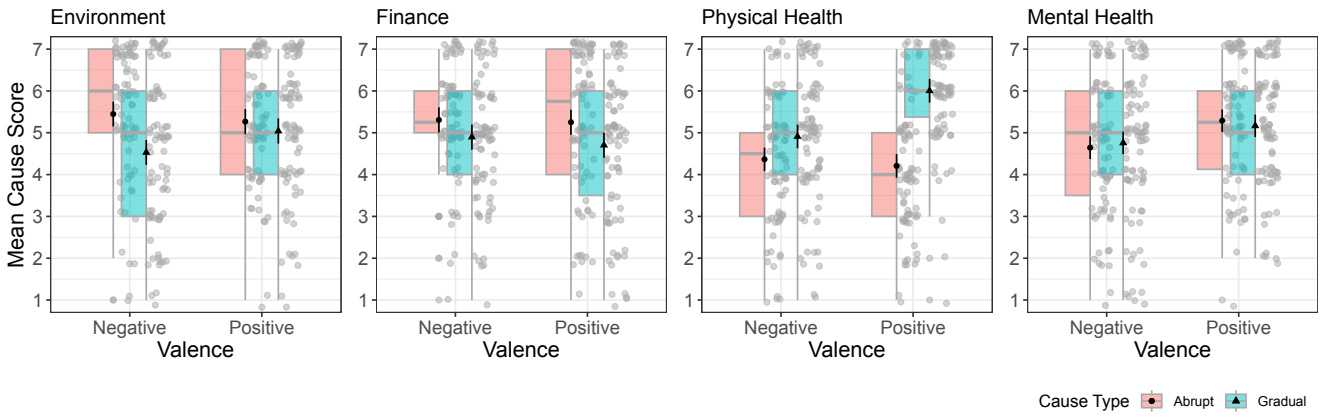

Figure 2: Participants attributed causality differently between gradual and abrupt causes depending on domain in Study 1. Error bars indicate 95% CI.

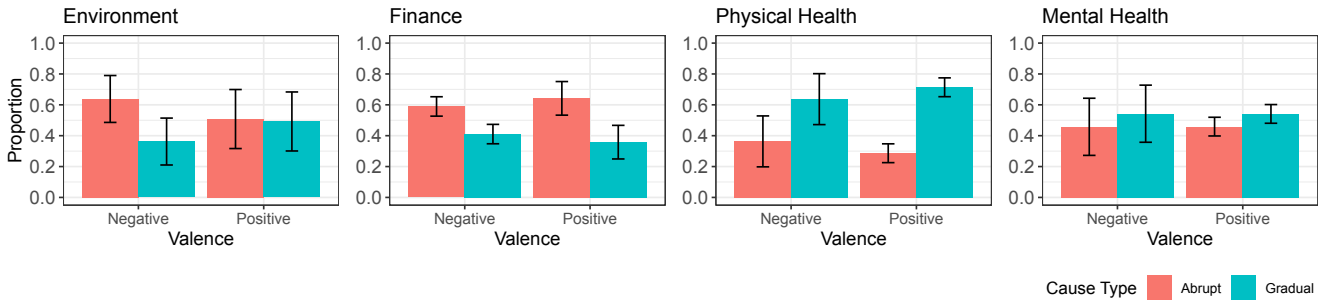

Figure 3: Participants mentioned counterfactual alternatives about gradual versus abrupt causes differently depending on domain in Study 1. Error bars indicate 95% CI.

Method

Participants We recruited 165 participants on Prolific. We excluded one participant from analysis due to incomplete responses. Our final sample size was $N = 164$ (84 male, 80 female; $M_{age} = 38.4$, $SD_{age} = 10.9$).

Design and Materials In this pilot study, we only included the finance vignettes from Study 1, revised such that the abrupt cause happened either more or less recently than the gradual cause. We used a 2 (Valence) $\times$ 2 (Position) $\times$ 2 (Temporal Order) $\times$ 2 (Vignette) mixed design. Participants saw one of eight versions of each vignette. As in Study 1, we analyzed effects averaging across Position and Vignette.

Data Analysis We used the same data analysis method as in Study 1 for both causal attribution scores and counterfactual responses, except the latter was only coded by the first author.

Results

We found no significant interaction between Cause Type, Temporal Order, and Valence, $F(1, 416.35) = 0.41, p = .524$. As shown in the left panel of Figure 4, which averages across positive and negative scenarios, we found a significant interaction effect between Cause Type and Temporal Order, $F(1, 391.30) = 5.95, p = .015$. However, pairwise compar-

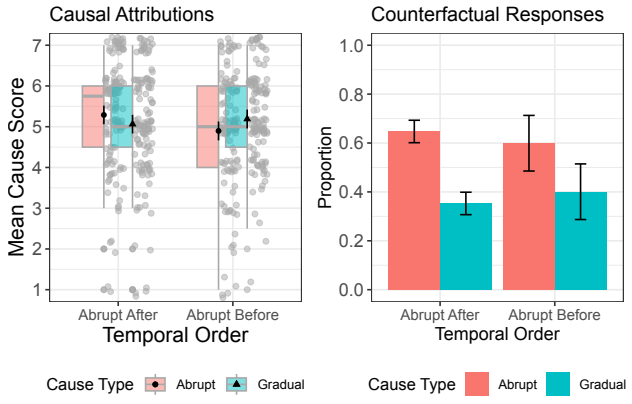

Figure 4: Temporal order affected how participants attributed causality to gradual versus abrupt causes (left panel) but not the salience of counterfactual alternatives (right panel) in Study 2. Error bars indicate 95% CI.

isons showed no statistically significant differences (abrupt after gradual: $t(260) = 1.88, p = .061$; abrupt before gradual: $t(256) = 1.48, p = .140$).

Referring to the right panel of Figure 4, this pattern was not observed for participants' counterfactual responses: they were significantly more likely to consider counterfactuals for the abrupt cause in both scenarios: where the abrupt cause came after the gradual cause ($z = 5.58$, OR $= 3.87$, 95% CI $[2.40, 6.23]$, $p < .001$), and where the abrupt cause came before the gradual cause ($z = 2.61$, OR $= 2.09$, 95% CI $[1.20, 3.64]$, $p = .009$).

Discussion

When we manipulated temporal proximity, participants attributed causality differently to the abrupt versus gradual cause depending on the temporal order of events. Participants attributed more causality to whichever cause was closer in time to the outcome. However, as in Study 1, participants were still more likely to mention counterfactual alternatives for the abrupt cause regardless of temporal order, suggesting that counterfactual considerations may not entirely explain causal attribution.

General Discussion

Across two exploratory studies, we found that participants attributed causality to abrupt and gradual events differently across domains, and that temporal proximity changed how people attributed causation to gradual and abrupt events: they tended to report higher causal attribution scores to the abrupt cause only when it occurred closer in time to the outcome.

The general finding that temporal proximity plays an important role in cases of joint causality is in line with previous literature (Hart & Honoré, 1985; Lagnado & Channon, 2008; Reuter et al., 2014). Our results suggest that, at least in some contexts, the key characteristic of the abrupt cause that affects people's causal attributions is that it often has an immediate impact on the outcome. Nonetheless, the effect of temporal order does not fully explain why we found a different pattern of results depending on domain in Study 1, given that the abrupt cause occurred more recently than the gradual cause in almost all cases.

One explanation for this interaction between domains could be participants' expectations. For example, in the physical health domain, people may expect that causes of positive health (e.g., a healthy diet, increasing exercise) are gradual. However, both abrupt and gradual causes often contribute to a negative health outcome (e.g., an injury or a condition that worsens over time), which could explain the smaller preference for gradual causes in this scenario.

Notably, we found some differences across domains and between causal and counterfactual judgments. For example, people's counterfactual responses in finance scenarios focused on abrupt rather than gradual causes – even when they did not attribute cause differently (Study 2). This suggests a possible dissociation between people's causal judgments and the counterfactuals they bring to mind (Mandel, 2003). This difference has been explained in various ways, such as people focusing more on events that are abnormal (Kahneman & Miller, 1986) or events they could have changed (Girotto et al., 1991). Our studies suggest a new alternative – that a single abrupt event is easier to bring to mind (and thus to mentally change) compared to multiple gradual events. For example, it is easier to imagine a single large gain or loss of money compared with many small changes.

Limitations and Further Directions

While we examined four different domains, more work is needed to explain the interaction between domain and cause type. One straightforward improvement is to replicate our findings with more vignettes to ensure robustness of the observed effects. In addition, as Study 2 was only a pilot study, it was not sufficiently powered to detect the interaction between Valence, Cause Type, and Temporal Order ($\sim$ 65% power).

Moreover, we aim to further examine interactions between familiarity and valence. For example, in the physical health domain, one negative vignette described a person being diagnosed with Type 2 diabetes due to an unhealthy diet and being infected with the SARS-CoV-2 virus. The positive version described a person with Type 2 diabetes who reduced their blood sugar levels due to sticking to a healthy diet and having weight-loss surgery. Surgery and infection are both abrupt, but their perceived causal strength may differ: participants may be less familiar with the link between Covid-19 and Type 2 diabetes, which has only been studied recently (e.g., Ssentongo et al., 2022). Future work should control for these factors more tightly, as well as causal structure (see Henne et al., 2021), which could have interacted with domain.

Additionally, we aim to ask more targeted counterfactual questions that correspond more directly with existing frameworks such as the Counterfactual Simulation Model. For example, to investigate whether participants are reasoning about the necessity and robustness of each cause, we could ask to what extent they think the outcome would still have occurred if one of the causes had been removed or changed. Together, these extensions could build upon this exploratory work to investigate what factors are driving this effect.

Conclusion

We examined how people attribute causality to gradual versus abrupt events across different real-world domains. We found a preference for abrupt over gradual causes in some domains, but the opposite pattern in others. Moreover, we found some evidence that the preference for abrupt events was driven by the temporal proximity of the cause to the outcome. Further, the counterfactuals that people generated were not always consistent with their causal judgments. We suggest that this might be due to the ease of imagining one large change over multiple small changes in some contexts. Future research is needed to further explore the relationship between causal attribution and counterfactual thinking in real-world scenarios.

Acknowledgements

We would like to thank Daisy Li for coding the qualitative responses for Study 1.

References

Alicke, M. D. (2000). Culpable control and the psychology of blame. *Psychological Bulletin*, *126*(4), 556–574.

Bear, A., & Knobe, J. (2017). Normality: Part descriptive, part prescriptive. *cognition*, *167*, 25–37.

Cheung, V., Qiao, M. H., Brown, T., & Lagnado, D. A. (2025). Attribution of responsibility between agents in a causal chain. https://osf.io/preprints/psyarxiv/863eq

Chockler, H., & Halpern, J. Y. (2004). Responsibility and blame: A structural-model approach. *Journal of Artificial Intelligence Research*, *22*, 93–115.

Gerstenberg, T., Goodman, N. D., Lagnado, D. A., & Tenenbaum, J. B. (2021). A counterfactual simulation model of causal judgments for physical events. *Psychological review*, *128*(5), 936.

Gerstenberg, T., & Lagnado, D. A. (2010). Spreading the blame: The allocation of responsibility amongst multiple agents. *Cognition*, *115*(1), 166–171.

Gerstenberg, T., Lagnado, D. A., et al. (2014). Actual and counterfactual worlds. *Oxford Studies in Experimental Philosophy, Volume 1*.

Girotto, V., Legrenzi, P., & Rizzo, A. (1991). Event controllability in counterfactual thinking. *Acta Psychologica*, *78*(1-3), 111–133.

Hagmayer, Y., Sloman, S. A., Lagnado, D. A., & Waldmann, M. R. (2007). Causal reasoning through intervention. *Causal learning: Psychology, philosophy, and computation*, 86–100.

Hart, H. L. A., & Honoré, T. (1985). *Causation in the law*. OUP Oxford.

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

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

Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review*, *93*(2), 136–153.

Kirfel, L., & Lagnado, D. A. (2021). Changing minds — epistemic interventions in causal reasoning. *Cognition*, *212*, 104721. https://doi.org/10.1016/j.cognition.2021.104721

Kominsky, J. F., & Phillips, J. (2019). Immoral professors and malfunctioning tools: Counterfactual relevance accounts explain the effect of norm violations on causal selection. *Cognitive Science*, *43*(11), e12792.

Lagnado, D. A., & 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

Lenth, R. V. (2022). *Emmeans: Estimated marginal means, aka least-squares means* [R package version 1.8.1-1]. https://CRAN.R-project.org/package=emmeans

Mandel, D. R. (2003). Judgment dissociation theory: An analysis of differences in causal, counterfactual and covariational reasoning. *Journal of Experimental Psychology: General*, *132*(3), 419.

Merck, C., & Kleinberg, S. (2016). Causal explanation under indeterminism: A sampling approach. *Proceedings of the AAAI Conference on Artificial Intelligence*, *30*(1).

Pearl, J., et al. (2000). Models, reasoning and inference. *Cambridge, UK: Cambridge University Press*, *19*(2).

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.

Shanks, D. R., Pearson, S. M., & Dickinson, A. (1989). Temporal contiguity and the judgement of causality by human subjects. *The Quarterly Journal of Experimental Psychology*, *41*(2), 139–159.

Singmann, H., Bolker, B., Westfall, J., Aust, F., & Ben-Shachar, M. S. (2022). *Afex: Analysis of factorial experiments* [R package version 1.1-1]. https://CRAN.R-project.org/package=afex

Ssentongo, P., Zhang, Y., Witmer, L., Chinchilli, V. M., & Ba, D. M. (2022). Association of covid-19 with diabetes: A systematic review and meta-analysis. *Scientific Reports*, *12*(1), 20191.