A Semantics for Causing, Enabling, and Preventing Verbs Grounded in Structural Causal Models

Anonymous CogSci submission

Abstract

When choosing how to describe what happened, we have a number of causal verbs at our disposal. In this paper, we develop a model-theoretic formal semantics for nine causal verbs that span the categories of CAUSE, ENABLE, and PREVENT. We use structural causal models (SCMs) to represent participants' mental construction of a scene when assessing the correctness of causal expressions within the presented context. Furthermore, SCMs enable us to model events relating both the physical world as well as agents' mental states. In experimental evaluations, we find that the proposed semantics exhibits a closer alignment with human evaluations in comparison to prior accounts of the verb families.

Keywords: causality; language; structural causal models; semantics; psycholinguistics.

Introduction

Causal cognition is ubiquitous and foundational for reasoning about both the physical and the social world (Gerstenberg & Tenenbaum, 2017; Waldmann, 2017). How can we best capture people's causal knowledge about the world? Structural causal models (SCMs) (Pearl, 2009; Spirtes, Glymour, & Scheines, 2000) are a generic formalism where a set of variables can represent both the mental states and actions of agents, and the state of the physical world at various levels of detail. In this paper, we use structural causal models to develop a proposal about the semantics of a variety of causal verbs and experimentally evaluate the novel predictions our framework makes.

Our objects of study are English verbs of causing (*cause*, *get*, *make*), enabling (*enable*, *let*, *allow*), and preventing (*prevent*, *stop*, *block*). We investigate the meaning of these nine verbs when used in linguistic constructions of the form

$$X \begin{Bmatrix} \textit{caused} & \textit{enabled} \\ \textit{got} & \textit{allowed} \end{Bmatrix} \alpha \textit{ to } Z \qquad X \begin{Bmatrix} \textit{made} \\ \textit{let} \end{Bmatrix} \alpha\, Z$$

$$X \begin{Bmatrix} \textit{prevented} \\ \textit{stopped} \\ \textit{blocked} \end{Bmatrix} \alpha \textit{ from } Z\textit{ing}$$

where the subject X is an event, the object α is an agent, and Z is an event. Our choice to use these nine verbs was motivated by previous work on these verb families (Cao, Williamson, & Choi, 2022; Klettke & Wolff, 2003; Wolff, Klettke, Ventura, & Song, 2005). While each of these verbs undoubtedly has its own subtle meaning, our proposal is that each verb family will at least entail a "core" meaning of CAUSE, ENABLE, and PREVENT, respectively (see also Wolff, 2007). We propose that **causing verbs** entail that the event X causes the event Z with actions of α mediating, **enabling verbs** entail that X makes α able to bring about Z, and **preventing verbs** entail that X makes α unable to bring about Z.

The proposal that these verb families each entail a respective "core" meaning makes good on the insights from the psychological study of causal language, where periphrastic causatives (verbs that denote indirect causal relationships) have been organized into CAUSE, ENABLE, and PREVENT verb families (Beller, Bennett, & Gerstenberg, 2020; Cheng & Novick, 1991; Sloman, Barbey, & Hotaling, 2009; Wolff, 2007; Wolff et al., 2005; Wolff & Song, 2003; Wolff & Zettergren, 2002). As depicted in Figure 1, Wolff (2007) defines these three categories in terms of affector and patient forces, and how they combine to align with the endstate. The representational use of "forces" emphasizes the physical aspect of causal relationships, and thus anticipates agents' internal desires to manifest as a force. From another point of view, Cheng and Novick (1991) and Cheng (1997) differentiate causing, enabling, and preventing by measuring the covariation between potential factors and the effect over a set of contextually relevant events. Yet another view uses the framework of mental model theory in which different causal verbs are analyzed in terms of the logical possibilities that they imply (Goldvarg & Johnson-Laird, 2001).

More recent efforts have aimed to capture the differences between CAUSE, ENABLE, and PREVENT using SCMs. For example, Sloman et al. (2009) argues that the use of the verbs

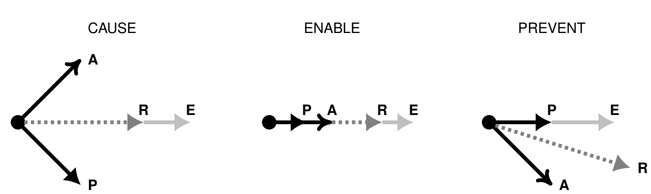

Figure 1: Representation of CAUSE, ENABLE, and PREVENT from Wolff (2007), where forces associated with the affector (A), forces associated with the patient (P) combine to form the resultant force (R) that may or may not be directed towards the endstate (E).

cause, *enable*, and *prevent* expresses beliefs about type and token cause-effect relationships in a causal model. Generally, previous experimental work on causal language such as Beller et al. (2020), Klettke and Wolff (2003), and Wolff (2003) have used non-agentive patients when exploring structured causal relationships between *affectors* and *patients* (in the sense of Wolff, 2007; Wolff et al., 2005). Our work explores agentive patients with observable preferences which provides a new perspective on these verb families.

Causal language has also been studied extensively in linguistics (Dowty, 1979; Levin & Hovav, 1994; Siegal & Boneh, 2020). Like psychologists, linguists have been interested in the conceptual reasoning behind speakers who choose to use one causal expression over another. Recently, there has been a growing interest in using structural causal models to define natural language semantics (Baglini & Siegal, 2020, 2021; Lassiter, 2018; Lauer & Nadathur, 2020; Schulz, 2011). For example, Lauer and Nadathur (2018, 2020) argue that causal necessity and sufficiency differentiate lexical uses of *make* and *force*, while Baglini and Siegal (2020) uses SCMs to explain the asymmetric entailment relation between *cause* and lexical causatives (e.g., *kill*).

In this paper, we build on the linguistic and psychological work by integrating SCMs into a semantics for causal language. We then experimentally test the predictions that our semantics make that are in conflict with existing accounts and find support for the proposed semantics.

The paper is organized as follows. We first define time-indexed causal models with agents for jointly representing social and physical dynamics. Then, we use this formalism to define "core meaning" concepts CAUSE, ENABLE, PREVENT, and propose that verbs in the cause, enable, and prevent families entail their respective concept. We experimentally support three predictions made by our model that conflict with existing accounts. In the experiment, we asked participants to watch videos of a simple grid world and evaluate whether English sentences are an accurate description. We model participants as constructing some time-indexed causal models with agents of the grid world that is used to evaluate the truth of the English sentences, and give two examples of such models. We close by discussing our results, which support the proposed semantics, and highlighting directions for future work building on these results.

Time-Indexed SCMs with Agents

In this section, we first define causal models in the sense of Pearl (2009). Using the logic of structural causal models (SCMs), we define a model-theoretic semantics for the concepts of CAUSE, ENABLE, and PREVENT.

For the purposes of our paper, SCMs can simulate the mechanics and entities of a particular world. Causal models carve up a phenomenon into a set of variables with a causal structure that connects them and causal mechanisms that determine their value. We additionally privilege certain subsets of variables that represent the mental states and actions of agents.

Definition 1. Models. We define a **time-indexed causal model with agents** $\mathcal{M}$ to consist of:

- **Variables** where each variable X_t, indexed by a timestep $t \in \{0, 1, 2, \ldots\}$, has an associated set of **values** it can take on $\mathsf{Val}(X_t)$.
- **Causal Structure** represented by arrows running from "parent" variables to "child" variables. We require that all parents immediately precede their children. Equivalently, if P_t is a parent of $C_{t'}$, then $t' = t + 1$.
- **Causal Mechanisms** that determine a node's value based on the value of its parents.
- **Agents** where each agent α has associated sets of variables encoding mental states $\mathbf{M}^{\alpha}$ and actions $\mathbf{A}^{\alpha}$. We require that the children of mental state variables be mental state variables or action variables of the same agent.

Definition 2. Partial and Total Settings. A **setting** assigns some number of variables values. **Total settings** assign every variable a value, while **partial settings** assign values to some subset of variables.

The variables at timestep zero have causal mechanisms that output constant values, which, in turn, determine the values for variables at timestep one, which determine the values for timestep two, and so on. Think of this total setting as capturing what actually happens.

Definition 3. Events. We define an event $\mathbf{E} = \mathbf{e}$ to be a partial setting $\mathbf{e}$ of a set of variables $\mathbf{E}$. An event *happens* in a model $\mathcal{M}$, written $\mathbf{E} = \mathbf{e}$, when the total setting that satisfies the mechanisms of $\mathcal{M}$ projected onto the variables $\mathbf{E}$ results in the partial setting $\mathbf{e}$.

The fundamental operation on a causal model is an intervention that fixes the values of some variables, which in turn may have downstream changes on other variables. Interventions can be understood as a function that takes in a causal model and outputs a new causal model where the intervened-on variables have their causal mechanisms fixed to be functions mapping to constant values.

Definition 4. Interventions. An intervention $\mathbf{I} \leftarrow \mathbf{i}$ is a partial setting $\mathbf{i}$ of variables $\mathbf{I}$. A proposition ϕ is true under an intervention, written $\langle \mathbf{I} \leftarrow \mathbf{i} \rangle \varphi$, if φ is true in the model identical to $\mathcal{M}$ except where the causal mechanisms of $\mathbf{I}$ are set to be constant functions mapping to the values in $\mathbf{i}$.

We include a list of agents that are associated with mental state variables and action variables, which allows us to define the *dynamic modality* of agents, that is, what an agent is and isn't able to do.

Definition 5. Action Sequences. We define an action sequence $\mathbf{a}^{\alpha}_{t:t'}$ to be a partial setting that fixes only the action variables of an agent α from time t to time t', inclusive.

Definition 6. Dynamic Modality. We define an agent α to be able to bring about an event $\mathbf{Z}=\mathbf{z}$ at time t, written $\textsc{Can}(\alpha, \mathbf{Z}=\mathbf{z}, t)$, if there is a time $t' > t$ and action sequences $\mathbf{a}^{\alpha}_{t:t'}$ and $\mathbf{b}^{\alpha}_{t:t'}$ such that

$$\langle \mathbf{A}^{\alpha}_{t:t'} \leftarrow \mathbf{a}^{\alpha}_{t:t'} \rangle \mathbf{Z}=\mathbf{z} \wedge \langle \mathbf{A}^{\alpha}_{t:t'} \leftarrow \mathbf{b}^{\alpha}_{t:t'} \rangle \mathbf{Z}=\mathbf{z}'.$$

Semantics for Causing, Enabling, and Preventing Verbs

Here, we take the definitions given by Pearl (2009) and use them to build a semantics for verbs of causing, enabling, and preventing. We take the fairly standard philosophical view that causation is a binary relation between *events* (Davidson, 1967; Lewis, 1973, 1986; cf. Gerstenberg, Goodman, Lagnado, & Tenenbaum, 2021). This means that in a sentence depicting a causal relationship of the structure [X CAUSE/ENABLE/PREVENT α Z], X and α Z are events and X implicitly or explicitly embeds both an agent and event led by the agent when the syntactic subject is agentive (Hitchcock, 2020). Others have argued that *facts* are the relata of causal relationships, since facts are better able to account for negative events or absences (Bennett, 1988; Mellor, 2004). For our purposes, we allow *event* to have a fairly wide domain and include previously debated phenomena such as states and events of omission (Beebee, 2004; Gerstenberg & Stephan, 2021; Henne, Pinillos, & De Brigard, 2017; McGrath, 2005).

Causing Verbs

We hypothesize that a verb from the cause family entails that X was a cause of α taking actions to bring about Z. Formally, we define $\textsc{Cause}(\mathbf{X}=\mathbf{x}, \alpha, \mathbf{Z}=\mathbf{z}, t)$ to be true when the following hold

1. The event $\mathbf{X}=\mathbf{x}$ happens.
2. The event $\mathbf{Z}=\mathbf{z}$ happens.
3. There exist a sequence of actions $\mathbf{a}^{\alpha}_{\geq t}$ such that
   (a) The event of agent α taking the actions $\mathbf{a}^{\alpha}_{\geq t}$ happens
   (b) The event $\mathbf{A}^{\alpha}_{\geq t}=\mathbf{a}^{\alpha}_{\geq t}$ causes the event $\mathbf{Z}=\mathbf{z}$ and this causal relationship is fully mediated[1] by the event $\mathbf{X}=\mathbf{x}$, meaning there exists $\mathbf{x}'$, $\mathbf{a}'$, and $\mathbf{z}'$ such that

$$\langle \mathbf{X} \leftarrow \mathbf{x}' \rangle (\mathbf{A}^{\alpha}_{\geq t}=\mathbf{a}' \wedge \mathbf{Z}=\mathbf{z}') \wedge \langle \mathbf{X} \leftarrow \mathbf{x}', \mathbf{A}^{\alpha}_{\geq t} \leftarrow \mathbf{a}^{\alpha}_{\geq t} \rangle \mathbf{Z}=\mathbf{z}.$$

Consider the following sentence as an example: "The deer running across the street caused Josie to slam on the breaks." Condition 1 tells us that the event of *the deer running across the street* actually occurring logically follows. Condition 2 tells us that the event of *Josie slamming on the breaks* actually happens as well. Finally, Condition 3 tells us that Josie took a (sequence of) actions that fully mediates the *the deer running across the street* causing *Josie slamming on the breaks*, such as taking her foot off the gas and pushing on the break.

[1] *Mediation* in the sense that the indirect effect is transmitted to the outcome via the mediator (Pearl, 2014).

Enabling Verbs

We hypothesize a verb from the enable family entails that X was a cause of α having available actions that bring about the event Z. Formally, we define $\textsc{Enable}(\mathbf{X}=\mathbf{x}, \alpha, \mathbf{Z}=\mathbf{z}, t)$ to be true when the following hold

1. The event $\mathbf{X}=\mathbf{x}$ happens.
2. The agent α is able to bring about the event $\mathbf{Z}=\mathbf{z}$
$$\textsc{Can}(\alpha, \mathbf{Z}=\mathbf{z}, t).$$
3. The event $\mathbf{X}=\mathbf{x}$ causes the agent to be able to bring about the event, meaning there exists an $\mathbf{x}'$ such that
$$\langle \mathbf{X} \leftarrow \mathbf{x}' \rangle \neg \textsc{Can}(\alpha, \mathbf{Z}=\mathbf{z}, t).$$

Again, consider the following example sentence: "The ice freezing enabled Jin to skate on the lake." Condition 1 holds because the ice actually froze, Condition 2 holds because Jin has the ability to choose to skate on a lake given that the water has frozen, and finally, Condition 3 holds because if the ice hadn't frozen, then Jin wouldn't have been able to choose to skate on the lake.

Preventing Verbs

We hypothesize that their use entails that X was a cause of α having no available actions that can bring about the event Z. Formally, we define $\textsc{Prevent}(\mathbf{X}=\mathbf{x}, \alpha, \mathbf{Z}=\mathbf{z}, t)$ to be true when the following hold

1. The event $\mathbf{X}=\mathbf{x}$ happens.
2. The agent α is unable to bring about the event $\mathbf{Z}=\mathbf{z}$
$$\neg \textsc{Can}(\alpha, \mathbf{Z}=\mathbf{z}, t).$$
3. The event $\mathbf{X}=\mathbf{x}$ causes the agent to be unable to bring about the event, meaning there exists an $\mathbf{x}'$ such that
$$\langle \mathbf{X} \leftarrow \mathbf{x}' \rangle \textsc{Can}(\alpha, \mathbf{Z}=\mathbf{z}, t).$$

Consider the following example of a preventing verb: "The storm warning being issued prevented Juan from visiting his family over the holidays." This statement tells us that the storm warning was actually issued (Condition 1), Juan is unable to travel to his family (Condition 2) and if the storm warning hadn't been issued, then Juan would have been able to visit his family (Condition 3).

Novel Predictions

Our proposal is that the verb families of cause, enable, and prevent verbs have meanings that logically entail the concepts CAUSE, ENABLE, and PREVENT. We experimentally test three predictions made by our proposal that are in conflict with previous accounts.

> **H1**. X may be an event of omission for cause, enable, or prevent verbs.

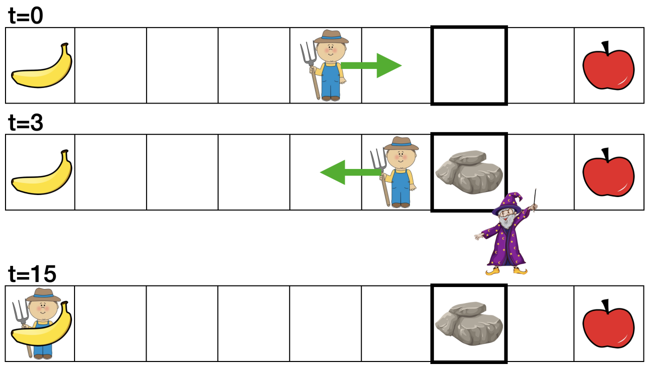

Figure 2: Three frames from the short videos shown to participants. The bolded cell starts empty and the farmer moves towards the apple. Then, the wizard places a rock which blocks the farmer's way to his (apparently) preferred fruit. In the end, the farmer reaches the banana instead. The green arrows indicate in which direction the farmer moves during the video and are added only for demonstration. The number of cells was reduced due to space constraints.

The hypothesis **H1** is in conflict with the accounts of Dowe (2004) and Salmon (1998), who predict that X *may not* be an event of omission.

> **H2**. Enabling verbs do not entail that Z happened.

The hypothesis **H2** is in conflict with the account of Beller et al. (2020) and Wolff et al. (2005) who predict that enabling verbs *do* entail that Z actually happened.

> **H3**. Preventing verbs do not entail that Z would have happened if not for X.

The hypothesis **H3** is in conflict with the account of Beller et al. (2020) and Wolff et al. (2005), who predict that preventing verbs *do* entail that Z would have happened if not for X.

Experiment

We tested our semantics by presenting participants with short animated videos including scenes such as shown in Figure 2, and asking them to select whether multiple expressions accurately or inaccurately describe the scene.

Methods

Materials The videos used in our experiment can be found here: `tinyurl.com/yc6e5dzm` . We created 7 different videos each less than 10 seconds long. Figure 2 shows the general structure of each video. In each video, there is a **wizard** and hallway with a **farmer** in the middle, an **apple** on the far right, a **banana** on the far left, and a **bolded cell** between the farmer and the apple which could be empty or contain a rock. Across the videos, we varied (1) whether a rock is present in the bolded cell at the beginning of the video, (2) whether the wizard casts a spell that either removes or places the rock, and (3) whether the farmer prefers the apple or banana.[2] In Figure 2, we show three frames for the video where the bolded cell starts empty, the farmer walks toward the apple, but then the wizard places a rock stopping the farmer who ends up going to the banana.

Language Stimuli For each video, we constructed nine sentences of the form

The NP *of the rock* verbed *the farmer (to/from) reach(ing) the apple.*

where *NP* is either *appearance*, *disappearance*, *presence*, or *absence* depending on which event happened.

Participants. 80 native English-speaking participants (*age*: Mean = 40, SD = 12; *gender*: 36 female, *nationality*: US) were recruited over Prolific. Each participant was provided with an introduction to the study and had to pass a simple comprehension question to continue. Failing the comprehension check brought the participants back to the introductory instructions, after which they could re-attempt the comprehension question. Participants took on average 7.32 minutes (SD = 4.68) to complete the task and were compensated at a rate of 12.57 USD per hour.

Procedure. Each participant completed 7 trials. Each trial contained one of the short videos paired with four randomly sampled sentences of interest and a trivial attention check question about the video. Participants were asked to select whether each sentence was "accurate" or "inaccurate"[3]. 8 participants were excluded for failing at least one of the seven attention checks.

Defining Mental Causal Models

We assume that participants watching the video stimuli mentally construct a representation that is a time-indexed causal model and then use it to evaluate the truth of natural language sentences. A crucial benefit of semantics grounded in SCMs is that we can remain agnostic about the details of the participants' mental model.

A participant might have a detailed, low-level causal model of the grid world's mechanistic updates at each time-step with variables representing the values of each cell, or a high-level causal model with variables representing the occurrence of major events and aggregated timesteps. In both cases, the model will be compatible with the proposed semantics.

There are two agents, **Wizard** and **Farmer**. The variables are defined to be $\mathcal{V} = \mathsf{Grid} \cup \mathbf{A}_{\text{Farmer}} \cup \mathbf{A}_{\text{Wizard}}$ where

$$\mathsf{Grid} = \{G_t^j : 0 \leq j \leq 24 \quad \wedge \quad t \in \mathbb{N}\}$$
$$\mathbf{A}_{\text{Farmer}} = \{A_t^F : t \in \mathbb{N} \setminus \{3\}\} \quad \mathbf{A}_{\text{Wizard}} = \{A_3^W\}$$

The values of these variables are defined to be

$$\mathsf{Val}(G_j^t) = \{\textbf{Blank}, \textbf{Farmer}, \textbf{Wizard}, \textbf{Rock}, \textbf{Banana}, \textbf{Apple}\}$$
$$\mathsf{Val}(A_t^F) = \{\rightarrow, \leftarrow\} \quad \mathsf{Val}(A_3^W) = \{\textbf{Cast}, \textbf{Don't Cast}, \textbf{Remove}\}$$

---

[2] This results in seven videos, because when the rock is present and the wizard casts no spell, the video is the same regardless of the farmer's preference.

[3] We also conducted an experiment where we allowed for responses on a continuous slider scale instead of a binary choice setup. This lead to similar results which are omitted here.

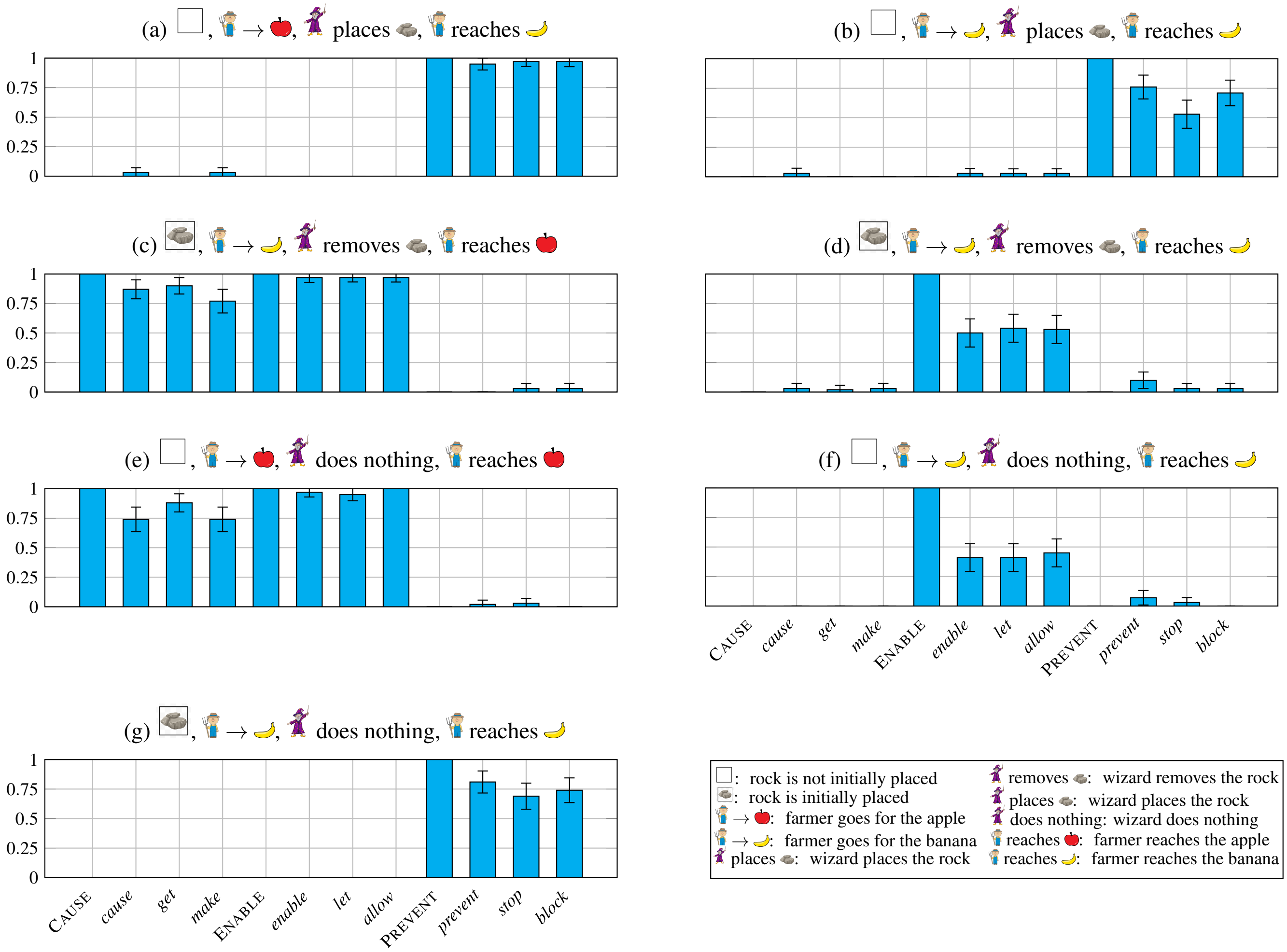

Figure 3: Results from our binary choice experiment with 95% confidence intervals for each of our seven video stimuli ($n = 72$). The bars for CAUSE, ENABLE, and PREVENT indicate whether those logical formulas are true for the shown video and have no error bars because they are theoretical quantities. The sentences shown had the form *The* NP *of the rock* verbed *the farmer (to/from) reach(ing) the apple*, where *NP* is either *appearance*, *disappearance*, *presence*, or *absence* depending on which event happened.

for $0 \leq j \leq 24$ and $t \neq 3$.

The causal mechanisms of the grid variable on the first timestep are constant functions that set the scene, with the rock only appearing in certain experimental conditions.

$$\mathcal{F}_{G_0^j} = \begin{cases} \textbf{Farmer} & j = 12 \\ \textbf{Banana} & j = 2 \\ \textbf{Rock} \text{ or } \textbf{Blank} & j = 18 \\ \textbf{Apple} & j = 22 \\ \textbf{Blank} & \textit{otherwise} \end{cases}$$

For the timestep $t = 4$, the causal mechanisms of the grid variables determine the values of each cell based on the value of the cell on the previous timestep and any action taken by the wizard on the previous timestep.

$$\mathcal{F}_{G_4^j}(g_3^j, a_3^W) = \begin{cases} \textbf{Rock} & a_3^W = \textbf{Cast} \\ & \text{and } g_{t-1}^j = \textbf{Blank} \\ \textbf{Blank} & a_3^W = \textbf{Cast} \\ & \text{and } g_{t-1}^j = \textbf{Rock} \\ \textbf{Rock} & a_3^W = \textbf{Don't Cast} \\ & \text{and } g_{t-1}^j = \textbf{Rock} \\ \textbf{Blank} & a_3^W = \textbf{Don't Cast} \\ & \text{and } g_{t-1}^j = \textbf{Blank} \\ g_2^j & \textit{otherwise} \end{cases}$$

For all other timesteps $t \notin \{0, 4\}$, unless the Wizard takes the action to Remove an existing Rock, the causal mechanisms of the grid variables determine the values of each cell based on the value of the cell on the previous timestep and any action taken by the farmer on the previous timestep.

$$\mathcal{F}_{G_t^j}(g_{t-1}^{j-1}, g_{t-1}^{j}, g_{t-1}^{j+1}, a_{t-1}^F) = \begin{cases} \textbf{Farmer} & a_{t-1}^F = \rightarrow \text{ and } g_{t-1}^{j-1} = \textbf{Farmer} \\ \textbf{Farmer} & a_{t-1}^F = \leftarrow \text{ and } g_{t-1}^{j+1} = \textbf{Farmer} \\ \textbf{Blank} & a_{i<t}^W = \textbf{Remove} \\ g_{t-1}^j & otherwise \end{cases}$$

Using this low-level mental model of the gridworld, participants would be able to record events such as the *Farmer's initial direction* (i.e., $\leftarrow$ *or* $\rightarrow$), *actions taken by the Wizard* (i.e. **Place Rock** and/or **Lift Rock**), as well as the dynamic constraints of the gridworld as given described in its Causal Mechanisms.

Alternatively, participants may represent their causal models at a higher level of abstraction. The videos can be understood as a sequence of four event variables: (1) R_0, the rock is **present** or **absent**, (2) F_1, the farmer moves **left** or **right**, (3) W_2, the wizard **casts** or **doesn't cast**, and (4) F_3, the farmer moves **left** or **right**. Like in the lower-level model, there are two agents, **Wizard** and **Farmer**. These variables have binary domains and their causal mechanisms directly encode the contrasting conditions in our experiments. The constant mechanism $\mathcal{F}_{R_0} = \textbf{present}$ or **absent** encodes one of two starting positions, the mechanisms for farmer movement $\mathcal{F}_{F_1}(r_0)$ and $\mathcal{F}_{F_3}(w_2)$ encodes one of two fruit preferences (apple or banana), the mechanism for wizard action $\mathcal{F}_{W_2}(f_1)$ encodes one of two wizard mindsets (helpful and unhelpful).

Results and Discussion

We predicted that the causing, enabling, and preventing verbs would have meanings that entail the logical formulas CAUSE, ENABLE, and PREVENT, respectively. The results support this hypothesis – whenever a logical formula is not true, the verbs in the corresponding family are near zero. Figure 3 depicts the proportion of participants selecting "accurate" for each verb.

H1. Our hypothesis that X may be an event of omission is supported in Figures 3e, 3f, and 3g where participants found sentences accurate even when the wizard took no action. In all these scenarios, the rock either remained absent (Figures 3e and 3f) or present (Figure 3g), but a significant number of participants still found causing, enabling, and preventing verbs acceptable for describing the scene ($p < 0.001$ in all cases). It has been long observed that people ordinarily judge omissions to be causes (see, e.g., Gerstenberg & Stephan, 2021; Henne, Pinillos, & Brigard, 2017; Walsh & Sloman, 2011). Our experiment extends this result to enabling and preventing verbs.

H2. Our hypothesis that the effect Z is not entailed by enabling verbs is supported in Figures 3d and 3f where a significant portion of participants ($p < 0.001$ in all cases) found sentences with enabling verbs accurate even when the farmer never reached the apple.

H3. Our hypothesis that preventing verbs do not entail that Z would have happened if not for X is supported in Figures 3b and 3g where a significant proportion of participants ($p < 0.001$ in all cases) found sentences with preventing verbs accurate when it is clear that the farmer wouldn't reach the apple even if he had been able to (because he prefers the banana anyways).

It is worth noting, however, that a portion of participants disagreed with **H2** and **H3**. Potential explanations include sub-populations having different underlying semantics, different thresholds for when the term *accurate* is appropriate, and variations in pragmatic reasoning about implications. Regardless, a semantics and pragmatics for enabling and preventing verbs should explain the disagreement among participants. The proposed semantics has the potential to model participant disagreement with an appropriate pragmatic account because it is compatible with **H1**, **H2** and **H3**, and a pragmatic account can weight preferences according to complex contextual constraints. We hope that these semantics can form the basis for the exploration of holistic models of pragmatic causal language production and comprehension.

In sum, these empirical results present challenges to three common assumptions in accounts of causal language but which are predicted by the presented structural causal models.

Limitations and Future Directions

While our proposal better predicts participants' judgements than previous work in the presented experiment, our semantics is limited in that it cannot make graded predictions. For example, consider that our notion of *bringing about* is binary – either an agent is able to bring about the effect, or it isn't.

We are also interested in understanding the level of granularity at which participants mentally model causal scenarios. For instance, as discussed in the Experiment section, participants may internally reason using the low-level model or the high-level model. What are the implications of using one or the other for natural language judgements? A theory of causal abstraction provides a rich avenue for future work (Geiger, Potts, & Icard, 2023).

Conclusion

This paper proposes a model-theoretic semantics for the verbs *cause, get, make, enable, let, allow, prevent, stop, block* using the logic of structural causal models (SCMs). SCMs enable us to not only model affector and patients' mental states, but also allow us to represent participants' construal of a presented video at different levels of detail. In an experiment that asked participants to rate descriptions of a context, we found that the results aligned better with the proposed semantics than with previous accounts of the causing, enabling, and preventing verb families. This suggests that an SCM account of causal language provides a valuable new perspective to understanding causal event language and judgments.

References

Baglini, R., & Siegal, E. A. B.-A. (2020). Direct causation: A new approach to an old question. *University of Pennsylvania Working Papers in Linguistics*, *26*, 19–28.

Baglini, R., & Siegal, E. A. B.-A. (2021). Modelling linguistic causation. *manuscript, Aarhus University and Hebrew University of Jerusalem*.

Beebee, H. (2004). Causing and nothingness. In J. Collins, N. Hall, & L. A. Paul (Eds.), *Causation and counterfactuals* (pp. 291–308). MA: MIT Press Cambridge.

Beller, A., Bennett, E., & Gerstenberg, T. (2020). The language of causation. *Proceedings of the 42nd Annual Conference of the Cognitive Science Society*.

Bennett, J. (1988). *Events and their names*. Oxford University Press UK.

Cao, A., Williamson, G., & Choi, J. (2022, June). A cognitive approach to annotating causal constructions in a cross-genre corpus. In *Proceedings of the 16th linguistic annotation workshop (law) at lrec* (pp. 151–159). Online: European Language Resources Association.

Cheng, P. W. (1997). From covariation to causation: A causal power theory. *Psychological Review*, *104*(2), 367–405.

Cheng, P. W., & Novick, L. R. (1991). Causes versus enabling conditions. *Cognition*, *40*(1), 83-120.

Davidson, D. (1967). Causal relations. *Journal of Philosophy*, *64*(21), 691–703.

Dowe, P. (2004). Causes are physically connected to their effects: Why preventers and omissions are not causes. In C. Hitchcock (Ed.), *Contemporary debates in philosophy of science* (pp. 189–196). Blackwell.

Dowty, D. R. (1979). *Word meaning and Montague grammar: the semantics of verbs and times in generative semantics and in Montague's PTQ* (No. v. 7). Dordrecht ; Boston: D. Reidel Pub. Co.

Geiger, A., Potts, C., & Icard, T. (2023). *Causal abstraction for faithful interpretation of ai models.* (arXiv:2106.02997)

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*(6), 936–975.

Gerstenberg, T., & Stephan, S. (2021). A counterfactual simulation model of causation by omission. *Cognition*, *216*, 104842.

Gerstenberg, T., & Tenenbaum, J. B. (2017). Intuitive theories. In M. Waldmann (Ed.), *Oxford handbook of causal reasoning* (pp. 515–548). Oxford University Press.

Goldvarg, E., & Johnson-Laird, P. N. (2001). Naive causality: A mental model theory of causal meaning and reasoning. *Cognitive Science*, *25*(4), 565–610.

Henne, P., Pinillos, Á., & Brigard, F. D. (2017). Cause by omission and norm: Not watering plants. *Australasian Journal of Philosophy*, *95*(2), 270-283.

Henne, P., Pinillos, Á., & De Brigard, F. (2017). Cause by omission and norm: Not watering plants. *Australasian Journal of Philosophy*, *95*(2), 270–283.

Hitchcock, C. (2020, 07). Communicating causal structure. In (p. 53-71).

Klettke, B., & Wolff, P. (2003). Differences in how english and german speakers talk and reason about cause. In *Proceedings of the annual meeting of the cognitive science society* (Vol. 25).

Lassiter, D. (2018). Causation and probability in indicative and counterfactual conditionals. *Unpublished manuscript*, 1–27.

Lauer, S., & Nadathur, P. (2018). *Sufficiency causatives.* (Unpublished manuscript)

Lauer, S., & Nadathur, P. (2020). Causal necessity, causal sufficiency, and the implications of causative verbs. *Glossa: a journal of general linguistics*, *5*, 49–105.

Levin, B., & Hovav, M. R. (1994). A preliminary analysis of causative verbs in english. *Lingua*, *92*, 35-77.

Lewis, D. (1973). Causation. *Journal of Philosophy*, *70*(17), 556–567.

Lewis, D. (1986). Events. In D. Lewis (Ed.), *Philosophical papers vol. ii* (pp. 241–269). Oxford University Press.

McGrath, S. (2005). Causation by omission: A dilemma. *Philosophical Studies*, *123*(1), 125–148.

Mellor, D. H. (2004). For facts as causes and effects. In N. Hall, L. A. Paul, & J. Collins (Eds.), *Causation and counterfactuals* (pp. 309–23). Cambridge: Mass.: Mit Press.

Pearl, J. (2009). *Causality: Models, reasoning and inference* (2nd ed.). Cambridge University Press.

Pearl, J. (2014, 06). Interpretation and identification of causal mediation. *Psychological methods*, *19*.

Salmon, W. C. (1998). *Causality and Explanation*. Oxford University Press.

Schulz, K. (2011). "if you'd wiggled a, then b would've changed": Causality and counterfactual conditionals. *Synthese*, *179*(2), 239–251.

Siegal, E. B.-A., & Boneh, N. (2020). Causation: From metaphysics to semantics and back. In *Perspectives on causation: Selected papers from the jerusalem 2017 workshop* (p. 3-51). Cham: Springer.

Sloman, S., Barbey, A., & Hotaling, J. (2009, January). A causal model theory of the meaning of cause, enable, and prevent. *Cognitive Science*, *33*(1), 21–50.

Spirtes, P., Glymour, C., & Scheines, R. (2000). *Causation, prediction, and search* (2nd ed.). MIT press.

Waldmann, M. R. (Ed.). (2017). *The oxford handbook of causal reasoning*. Oxford University Press.

Walsh, C. R., & Sloman, S. A. (2011). The meaning of cause and prevent: The role of causal mechanism. *Mind & Language*, *26*(1), 21-52.

Wolff, P. (2003). Direct causation in the linguistic coding and individuation of causal events. *Cognition*, *88*(1), 1-48.

Wolff, P. (2007). Representing causation. *Journal of Experi-*

mental Psychology: General, *136*(1), 82–111.
Wolff, P., Klettke, B., Ventura, T., & Song, G. (2005). Expressing causation in english and other languages. *Categorization inside and outside the laboratory: Essays in honor of Douglas L. Medin*.
Wolff, P., & Song, G. (2003). Models of causation and causal verbs. *Cognitive Psychology*, *47*, 276–332.
Wolff, P., & Zettergren, M. (2002). A vector model of causal meaning. In *Proceedings of the twenty-fifth annual conference of the cognitive science society.* Erlbaum.