Adults' use of latent vs direct causes for Intervention

Background

This study builds on a prior pre-registered experiment conducted in our lab (Study 1, June 2025: [link redacted for future double-blind review]) which investigated the hypothesis that adults use two representational spaces to organize mental, physical and bodily events. One space groups events based on underlying latent factors, and another space describes how specific events are directly causally connected to each other. Study 1 had two main findings regarding the structure of our representations of other agents' capacities: (1) adults represent 3 latent factors namely "the mind", "the mechanical body" and "the biological body" that organize mental events, actions, and physiological events, and (2) a distinct representation of how these events causally relate to each other.

An open question regarding the two representational spaces is what kind of content they articulate. One possibility is that these two spaces are intuitive causal-explanatory framework theories about how other people work. Prior work has shown that intuitive theories are common-sense, abstract, causal-explanatory frameworks that help us navigate the world (Gerstenberg & Tenenbaum, 2017). They help people reason about events from different domains: Intuitive psychology, for example, conceives of actions as causally connected to mental states, which allows people to explain, predict and intervene on other agents' minds and behaviors. Intuitive theories also help people reason about events within the same domain: children's intuitive understanding of the biological domain helps them explain, predict and plan interventions on bodily states to prevent states of illness and disease, and to infer that members of living species that share internal structure, not superficial perceptual features, will have similar properties (Degn et al., 2025; Gelman & Markman, 1987).

Here, we propose to test the hypothesis that the judgments that we measured in Study 1 constitute two distinct intuitive theories that help us make sense of other agents as mental beings, physical actors, and living systems in terms of (1) Latent Causes and (2) Direct Causes. If the representational spaces measured in Study 1 constitute intuitive theories, then even when not asked to, adults should recruit these beliefs for prediction, explanation, counterfactuals, and intervention. If these two spaces represent two distinct ways of reasoning causally about other agents, then we should be able to predict when people will use beliefs about Latent Causes vs Direct Causes.

In summary, in the current work we propose to test for the functional roles of the Latent Causes and Direct Causes frameworks as distinct intuitive causal theories in common-sense social cognition.

General Alternative Hypotheses

The two frameworks (Latent Causes and Direct Causes) support common-sense reasoning: We hypothesize that the two causal frameworks measured in Study 1 are the basis for our intuitive reasoning within and across the domains of mind, action and body. We predict that adults will spontaneously make use of these representations for inference, explanation, intervention, and counterfactual reasoning.

The two frameworks do not support common-sense reasoning: An alternative hypothesis is that while adults can report their beliefs about how mental events, actions, and physiological events are organized when explicitly asked to (like in Study 1), they do not make use of these beliefs during everyday social cognition.

Current Study

Prior work has argued that people represent causal relations in terms of interventions: A can cause B if changing A would lead to changing B (Goddu & Gopnik, 2024; Gopnik & Wellman, 2012; Sobel & Kushnir, 2006; Woodward, 2005). This study tests whether people can use both representations of direct causes (i.e. whether A can make B happen) and latent causes (i.e. what is the common cause of A and B) to design interventions on other agents. By hypothesis, direct causes and latent causes are relevant for designing different kinds of interventions. Thus, we designed a task with two sorts of interventions, one where latent causes are more relevant and one more direct causes are more relevant. In one intervention (*Ordinary Intervention*), participants consider which of two possible actions (i.e. an event which is a **cause** of the target event and another which is **similar** owing to sharing a common cause with the target event) they would do to make a hypothetical person do a **target** activity. In the other intervention *Magical Intervention*) people consider which of two possible spells (i.e. the same pair of causal and similar items as in the ordinary intervention condition) they would cast to give a creature a **target** ability; they are told that they cannot directly grant the creature this ability, and to choose the spell that would most likely accomplish the same result.

We constructed a stimulus set based on the items and results of Study 1. The set consists of 15 triads of items, each triad consisting of a **target** item (such as “get tired”) and a choice set of two items: an event that was most **similar** (i.e. close by in the Sorting Task RDM, but far away in the Causal Task RDM) to the target, and an event that was most **causally relevant** (i.e. close by in the Causal Task RDM but far away in the Sorting Task RDM). Thus, we chose these items based on their joint causal and similarity distance from the target: similar items had lowest similarity distance and largest causal distance to the target, whereas causal items had lowest causal distance and highest similarity distance to target. We constrained the similar option to be from the same domain as the target (this was already true for 13/15 target items), and the causal option to be from a different domain (this was already true for 14/15 target items). The full set of 15 item triads and their respective distances are listed in Figure 1.

Target and Choice Items			
Domain	Target Item	Causal Choice (Causal Distance, Freesort Distance)	Similar Choice (Causal Distance, Freesort Distance)
mind	see something	take a walk (0.2, 0.58)	hear something (0.42, 0.18)
	hear something	take a walk (0.25, 0.58)	see something (0.44, 0.18)
	choose what to do	experience pain (0.2, 0.57)	remember something (0.15, 0.36)
	remember something	take a walk (0.24, 0.61)	think about something (0.07, 0.15)
	think about something	get sick (0.25, 0.65)	remember something (0.08, 0.15)
action	reach for something	become hungry (0.21, 0.56)	kick something (0.54, 0.4)
	sit down	experience pain (0.13, 0.58)	jump up and down (0.54, 0.3)
	jump up and down	see something (0.29, 0.6)	sit down (0.83, 0.3)
	kick something	think about something (0.37, 0.6)	sit down (0.7, 0.42)
	take a walk	think about something (0.26, 0.58)	sit down (0.77, 0.33)
body	get tired	jump up and down (0.1, 0.54)	feel scared (0.42, 0.4)
	become hungry	take a walk (0.18, 0.55)	feel scared (0.67, 0.44)
	feel scared	see something (0.12, 0.54)	get tired (0.51, 0.4)
	experience pain	kick something (0.13, 0.51)	get tired (0.43, 0.41)
	get sick	see something (0.38, 0.62)	get tired (0.39, 0.34)

Figure 1: Stimuli for the current study based on group-averaged responses from two separate samples (total N = 151). The first two columns list the target events and their domains. The remaining two columns list the two options (the "causal choice" vs "similar choice") associated with that target, and their similarity and causal distance from the target.

Hypothesis

We hypothesize that when people design interventions they can rely both on representations of direct causes (i.e. whether A can make B happen), and representations of latent causes (i.e. whether A and B are both caused by a common variable). Specifically, when people are asked what they would do to intervene (e.g. target = get tired; causal choice = jump up and down; similar choice = feel scared), they will select the direct cause to intervene in the ordinary situation (e.g. make the person jump up and down) and the item with a shared latent cause to intervene in the magical situation (e.g. give the creature the capacity to feel scared).

We predict that people will be (1) more likely to select the causal item in the Ordinary than Magical Intervention condition, and that they will be (2) more likely than chance to select the causal item in the Ordinary Intervention condition, and (3) more likely than chance to select the similar item in the Magical Intervention condition.

Dependent Variable

On each trial, the choice between the **causal** and **similar** item.

Independent Variable

Whether the intervention is **ordinary** or **magical.**

Procedure

The experiment will have two between-subjects conditions. Subjects in each condition will be presented with the following instruction page:

> **Ordinary Intervention condition:** Welcome! In this game you will be asked to imagine that you are trying to make another person do something. On each trial, there will be an intended outcome (for example “choose what to do”), and two possible actions (for example “see something” and “jump up and down”). Please read carefully and respond with the action you would choose in the given scenario.
>
> **Magical Intervention condition**: Welcome! In this game you will be asked to imagine that you are a wizard who can cast spells to give creatures the ability to do things. On each trial, there will be an outcome you want to produce (for example, giving a creature the ability to “choose what to do”), and two possible spells (for example “see something” and “jump up and down”). Please read carefully and respond with the magic spell you would use in the given scenario.

The next pages will present 15 trials in each condition, randomly interspersed. Across both conditions, trials will consist of a test question and two alternatives in a forced-choice paradigm. The trials will read as follows:

> **Ordinary Intervention condition**: Choose the action that you think will best lead to the outcome - the one most likely to succeed. Suppose that you wanted to make someone else {*target ability*}. Which of these two actions would most likely accomplish your goal? Would you …
>
> **Magical Intervention condition**: Choose the spell that you think will best lead to the outcome - the one most likely to succeed. Suppose you wanted to give a creature the ability to {*target ability*}. However, you ran out of that specific spell. Which of these two other spells would most likely accomplish your goal? Granting the creature the ability to…

Each of the 15 targets will be shown once per condition. Below the test question will be an option between a similar choice and a causal choice. The left-right positions of the choices will be randomised across trials.

There will be two randomly interspersed attention checks, and participants will be excluded if they failed at least one attention check.

After the 30 trials, subjects will be presented with a debrief followed by a demographics questionnaire.

Analysis Plan

To test for our 3 predictions—namely, that people will be (1) more likely to (1) select the causal item in the Ordinary than Magical Intervention condition (2) more likely than chance to select the causal item in the Ordinary Intervention condition, and (3) more likely than chance to select the similar item in the Magical Intervention condition—we will use a mixed effects logistic regression using the lme4 package (Bates et al., 2015) in R. The model specification will be: **response ~ condition + (1 | subject_id), family = binomial(link = 'logit')**. We will set the reference group to be the Magical Intervention condition, thus the "intercept" will represent the log-odds of selecting the causally relevant item in the Magical Intervention condition, the slope coefficient will capture the log-odds difference between the Magical Intervention and Ordinary Intervention conditions, and the sum of the intercept and the condition coefficient will give the log-odds of selecting the causally relevant item in the Ordinary Intervention condition. We will confirm our first prediction if the slope coefficient is significant, and we will confirm our second and third predictions by running 'allEffects()' from the *effects*' package, which will generate confidence intervals that show whether each condition mean is different from chance.

Our threshold for statistical significance will be p = .05, two-tailed, and we will use the check_model() function from the performance package (Lüdecke et al., 2021) to conduct quality assurance.

Sample Size and Stopping Rule

Power analysis for Confirmatory Analysis

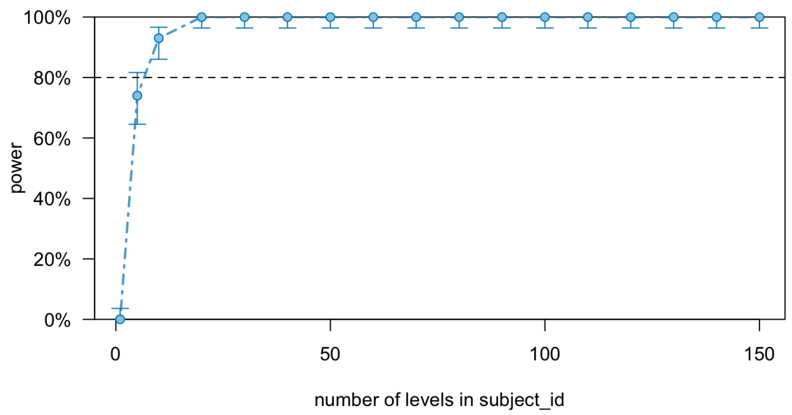
We used the R package *simr* to estimate the power curves for replicating all 3 predicted results. For power analysis purposes, will run three separate models targeting each prediction separately: the original mixed effects regression for detecting a condition difference, a separate Ordinary Intervention-only model for detecting if its causal response mean is above chance, and a separate Magical Intervention-only model for detecting if its similar response mean is above chance.

Based on a pilot study (N = 100), the sample size needed to achieve 80% power to replicate the slope coefficient for the combined condition model is <10; the sample size to replicate the Ordinary Intervention effect is N<20 and the sample size to replicate the Magical Intervention is N<20 (Figure 2).

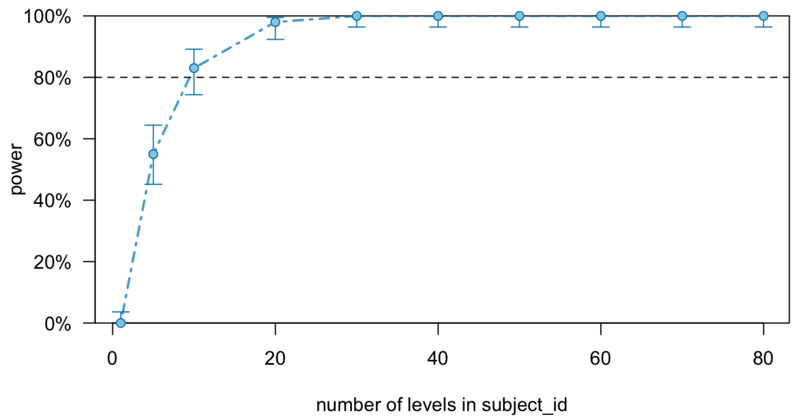
Stopping rule and exclusions

Considering the largest sample required to replicate all effects (N = 20), and that the effect size in the pilot could be inflated due to a small sample size, we conservatively estimate a target sample of 100 participants. To account for a <1% exclusion rate from the pilot study, we will conservatively collect data from 105 participants, prior to exclusions.

(a)



(b)



(c)

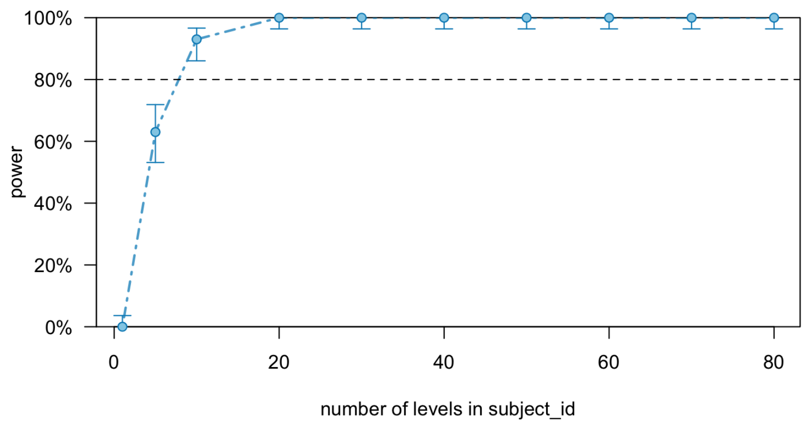


Figure 2(a-c): Power curve showing the power that would be achieved (y-axis) using different sample sizes (x-axis) to replicate the fixed effects for the responses ~ condition + (1 | subject_id) models with different datasets. (a) Shows power for the “combined data for both conditions” model, (b) shows power curve for the ordinary intervention-only model, and (c) is for the magical intervention-only model.

References

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models Usinglme4. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Degn, P., Fiber, Z., & Sullivan, J. (2025). Children's understanding of the causal mechanisms underlying disease prevention. *Cognitive Development*, *74*(101563), 101563. https://doi.org/10.1016/j.cogdev.2025.101563

Gelman, S. A., & Markman, E. M. (1987). Young children's inductions from natural kinds: The role of categories and appearances. *Child Development*, *58*(6), 1532. https://doi.org/10.2307/1130693

Gerstenberg, T., & Tenenbaum, J. B. (2017). *Intuitive Theories* (M. R. Waldmann (ed.)). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199399550.013.28

Goddu, M. K., & Gopnik, A. (2024). The development of human causal learning and reasoning. *Nature Reviews Psychology*, *3*(5), 319–339. https://doi.org/10.1038/s44159-024-00300-5

Gopnik, A., & Wellman, H. M. (2012). Reconstructing constructivism: causal models, Bayesian learning mechanisms, and the theory theory. *Psychological Bulletin*, *138*(6), 1085–1108. https://doi.org/10.1037/a0028044

Lüdecke, D., Ben-Shachar, M., Patil, I., Waggoner, P., & Makowski, D. (2021). Performance: An R package for assessment, comparison and testing of statistical models. *Journal of Open Source Software*, *6*(60), 3139. https://doi.org/10.21105/joss.03139

Sobel, D. M., & Kushnir, T. (2006). The importance of decision making in causal learning from interventions. *Memory & Cognition*, *34*(2), 411–419. https://doi.org/10.3758/bf03193418

Woodward, J. (2005). *Making things happen*. Oxford University Press.