Developing Explicit Understanding of Probabilistic Causation: Patterns and Variation in Children's Reasoning

Tina Grotzer
Harvard University

EPLIP Lecture Series
April 5, 2012




Acknowledgements

We would like to express our appreciation to the teachers and students who allowed us to collect data on their reasoning patterns.

Thank you to S. Lynneth Solis, Samantha Marengell, Evelyn Chen, and Reuben Posner for blind coding the data and assessing reliability between the two coders.

This work is supported by the National Science Foundation, Grant No. REC-0845632 All opinions, findings, conclusions or recommendations expressed here are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Project Team

- S. Lynneth Solis
- M. Shane Tutwiler
- Evelyn Chen
- Leslie Duhaylongsod
- Reuben Posner
- Molly Levitt
- David Jeong
- Megan Powell
- Samantha Marangell
- Erika Spangler
- Adi Flesher

Probabilistic Causality

-When contigencies between causes and effects are not one to one; they are in some respects stochastic

Studies of Cognition in relation to Probabilistic Causation must account for instances that are…

-inherent in the way that the world behaves…

-inherent in our perception of the world; the information that we take in and how we respond to it

Why Study Probabilistic Causation?

- The appearance of stochastic events—how do we reason about them?
- A key mechanism in causal induction?
- Arguments that we are causal determinists
- Interacts with how students reason in science class, about evidence over time, and about statistical regularities
- It interacts with how we notice and reason about complexity

In every day causality…

- Calling a friend sometimes gets a response.
- Pushing a button on a flashlight usually makes it light.
- If we plant seeds, many of them, but not all, may grow.
- Pushing a button on a game, might make something happen.

…but not always. The contingency patterns vary and so does our attention span!

→Events give the appearance of being stochastic!

A Key Mechanism in Causal Induction?

- Causal Bayes Net (CBN) Models assume that we sum across probabilistic instances of causation to discern causal relationships.
- Gopnik and colleagues have shown that even preschoolers behave in ways that fit with CBN Models.
- However, there is some possibility that this is about tracking (Schultz & Mendelson, 1975; Siegler, 1976; Siegler & Liebert, 1974) rather than acceptance.

Child Development, March/April 2006, Volume 77, Number 2, Pages 427–442

God Does Not Play Dice: Causal Determinism and Preschoolers' Causal Inferences

Laura E. Schulz
Massachusetts Institute of Technology

Jessica Sommerville
University of Washington

Three studies investigated children's belief in causal determinism. If children are determinists, they should infer unobserved causes whenever observed causes appear to act stochastically. In Experiment 1, 4-year-olds saw a stochastic generative cause and inferred the existence of an unobserved inhibitory cause. Children traded off inferences about the presence of unobserved inhibitory causes and the absence of unobserved generative causes. In Experiment 2, 4-year-olds used the pattern of indeterminacy to decide whether unobserved variables were generative or inhibitory. Experiment 3 suggested that children (4 years old) resist believing that direct causes can act stochastically, although they accept that events can be stochastically associated. Children's deterministic assumptions seem to support inferences not obtainable from other cues.

Many researchers have proposed that children's knowledge about the world can take the form of causal theories, in which *unobserved* causes play a central role (Carey, 1985; Gopnik, 1988; Gopnik & Meltzoff, 1997; Keil, 1989; Perner, 1991; Wellman, 1990). Children invoke unobserved mental states to explain human behavior (see, e.g., Wellman, 1990), invisible forces to explain physical events (Shultz, 1982), and invisible, internal mechanisms to explain biological events (Gelman, Coley, & Gottfried, 1994).

However, little is known about how children infer unobserved causes. Until recently, developmental psychologists have looked primarily at children's ability to infer causal structure from spatiotemporal cues (Cheng & Novick, 1992; Leslie & Keeble, 1987) and information about substantive, domain-specific mechanisms (Ahn, Gelman, Amsterlaw, Hohenstein, & Kalish, 2000; Bullock, Gelman, & Baillargeon, 1982; Carey & Spelke, 1994; Shultz, 1982; Spelke, Breinlinger, Macomber, & Jacobson, 1992). In adult cognitive psychology, by contrast, researchers have focused primarily on domain-general causal learning from the strength of association (Shanks, 1985; Shanks & Dickinson, 1987; Spellman, 1996) and patterns of covariation (Cheng, 1997, 2000) among events.

However, we can sometimes have causal knowledge even without knowing much about underlying mechanisms. If increasing serotonin levels relieve depression, we may conclude that low serotonin levels cause depression even if we do not know how. On the other hand, understanding causation seems to involve more than recognizing patterns of correlation. Lack of exercise is correlated with depression and we could imagine a plausible mechanism connecting the two (e.g., metabolic changes associated with exercise might regulate emotional arousal). However, if manipulating serotonin levels affects depression and manipulating physical activity does not, we will conclude that serotonin plays a causal role in depression and exercise does not.

Recently, psychologists, philosophers of science, and statisticians have suggested that the crucial piece missing from both mechanism and covariation accounts of causal inference is the notion of intervention (Gopnik et al., 2004; Gopnik & Schulz, in press; Pearl, 2000; Spirtes, Glymour, & Scheines, 1993; Woodward, 2003). Intuitively, if X is causally related to Y, then (all else being equal) there will be something we can do to change the value of X that will change the value of Y; that is, intervening

This research was supported by a National Science Foundation Graduate Fellowship Award (99–136), an American Association of University Women American Dissertation Fellowship, and a McDonnell Foundation Causal Learning Collaborative Initiative grant to Laura E. Schulz. We are very grateful to Alison Gopnik for significant contributions to this work. We are grateful also to Catherine Griffith, Carla Hudson, Anna Jenkins, Tamar Kushnir, Rebecca Saxe, and Christine Schaeffer, for their help and suggestions. Finally, we appreciate the support of the children, families, and staff at Fruit and Flower, Northeast Community Child Development Center, Helen Gordon Child Development Center, the Goddard School, and the Harold T. Jones Child Study Center.

Correspondence concerning this article should be addressed to Laura E. Schulz, MIT Department of Brain and Cognitive Sciences, 46-4011, 77 Massachusetts Avenue, Cambridge, MA 02139-4307. Electronic mail may be sent to lschulz@mit.edu.

Are Children Causal Determinists?

- Schulz and Sommerville (2006) found that children resist probabilistic causes and search for nonobvious causes when a cause is not apparent.
- They studied novel mechanical devices; gave training instructions about one thing making another thing happen; and in some studies, they studied one to one effects (4x) followed by one to none effects (4x) and vice versa.

How do children reason in contexts that invite probabilistic causation?

- games; biology; social instances?
- In familiar contexts?
- With scaffolds designed to help them?
- Could we encourage an explicit understanding of probabilistic causation?

→Why would we care?

In the face of stochastic evidence…

- …inferring non-obvious causes can be an adaptive stance.
- …however, an implicit assumption of determinism may lead to missing evidence of causal relationships that are unreliable or that fall beyond our attentional span.
- And an explicit assumption of determinism can lead to rejecting probabilistic causal relationships when reasoning about evidence.

- But the issue is bigger when we turn it inside out.

In a complex world, we need to be sensitive to possible causal connections despite…

- Noise in the environment that make them hard to detect.
- Attentional patterns that make us short-sighted.
- Causal patterns that significantly depart from one to one correspondence (accumulation patterns; triggering effects; increases in tendency/statistical summing across) for instance, increases in hurricanes, flooding, and snow amounts due to climate change.

Study Design

- Microgenetic studies, 4 students in each grade (K, 2, 4, 6) (n = 16) across the school year.
- High density interviews at points when students' classroom and study experiences suggested that change was likely—as in Opfer and Siegler (2004).
- Interviews proceeded from open-ended to increasingly structured to assess how students frame the concepts and the accessibility of concepts.
- Scaffolds used familiar examples and compared analogous causal forms in different problem contexts through "mutual alignment" (e.g. Kurtz, Miao, & Gentner, 2001) incorporated in the form of design studies (Brown, 1992; Collins, 1999).

Why Microgenetic Studies?

- Summing across cases helps us to learn trends.
- However, in teaching and learning, outliers are an important source of information.
- They teach us what we can do to help other students—they offer leverage points.

Subjects:

Students were from Boston and Cambridge and are primarily Black and Latino with less than 1% Caucasian and 78% on “Free or Reduced Lunch.”

Tasks:

- Four domains: biological, mechanical, social, and games (focus on authentic tasks)
- Included seed planting, hatching chicks, bubble gum machines, videotapes of brief social interactions, and a set of games
- Tasks that children might be familiar with from their everyday worlds were intentionally chosen to elicit their expectations and existing knowledge.

Tasks With Stochastic Effects

Scoring and Analysis

Sessions were intensively analyzed. Using ATLAS.TI, we coded:

1) transcripts and videos of students responses etic categories of probabilistic versus deterministic statements and;
2) emergent analysis of patterns in students' reasoning.

Independent coders coded the interpretive aspects and agreement levels were assessed with refinements made until there was at least 85% agreement.

Emergent codes independently generated by two of more coders are reported here.

Narratives were then developed for each student.

Student	Document	Reliability Round #	Final Reliability Round	Deterministic Agreement	Deterministic Detection Range	Deterministic Agreement over Lowest Detection	Deterministic Agreement over Mean Detection	Probabilistic Agreement	Probabilistic Detection Range	Probabilistic Agreement over Lowest Detection	Probabilistic Agreement over Mean Detection	Contradictions (Disagreements)	Overall Agreement over Lowest Detection	Overall Agreement over Mean Detection	Overall Agreement over Overlaps	Overall Detection
J	3.26 Bear	1		22	22-31	100%, (22/22)	83%, (22/26.5)	0	0-9	-	0%, (0/4.5)	9	100%	41.5%	71%, (22/31)	100%, (31/31)
		2	X	30	31-31	97%, (30/31)	97%, (30/31)	5	6-6	83%, (5/6)	83%, (5/6)	2	90%	90%	95%, (35/37)	100%, 37/37
K	2.12 Bunny	1	X	5	5-6	100%, (5/5)	91%, (5/5.5)	11	11-13	100%, (11/11)	92%, (11/12)	0	100%	91.5%	100%, (16/16)	100%, (16/16)
	4.29 Plants	1	X	7	7-7	100%, (7/7)	100%, (7/7)	2	2-2	100%, (2/2)	100%, (2/2)	0	100%	100%	100%, (9/9)	100%, (9/9)
Ta	2.6 Bunny	1	X	7	7-8	100%, (7/7)	93%, (7/7.5)	0	0-0	-	-	0	100%	93%	100%, (7/7)	87.5%, (7/8)
	5.6 Bear	1	X	16	16-16	100%, (16/16)	100%, (16/16)	8	8-9	100%, (8/8)	94%, (8/8.5)	0	100%	97%	100%, (24/24)	96%, (24/25)
Ty	12.9 Bunny	1	X	27	27-32	100%, (27/27)	92%, (27/29.5)	1	1-3	100%, (1/1)	50%, (1/2)	2	100%	71%	93%, 28/30	91%, (30/33)

COLUMN DESCRIPTIONS

Deterministic Agreement: Indicates the number of instances in which both coders agreed that a statement was Deterministic Thinking.

Deterministic Detection Range: Indicates the number of statements that each coder indipendently scored (detected) as Determnistic Thinking.

Deterministic Agreement over Lowest Detection: Indicates the percentage of Deterministic Thinking agreements out of the lowest number of Deterministic Thinking detections.

Deterministic Agreement over Mean Detection: Indicates the percentage of Deterministic Thinking agreements out of the average (mean) number of Deterministic Thinking detections.

Probabilistic Agreement: Indicates the number of instances in which both coders agreed that a statement was Probabilistic Thinking.

Probabilistic Detection Range: Indicates the number of statements that each coder indipendently scored (detected) as Probabilistic Thinking.

Probabilistic Agreement over Lowest Detection: Indicates the percentage of Probabilistic Thinking agreements out of the lowest number of Probabilistic Thinking detections.

Probabilistic Agreement over Mean Detection: Indicates the percentage of Probabilistic Thinking agreements out of the average (mean) number of Probabilistic Thinking detections.

Contradictions (Disagreements): Indicates the number of instances in which both coders scored (detected) a statement but disagreed on the code (e.g., one coder used Deterministic Thinking and the other used Probabilistic Thinking).

Overall Agreement over Lowest Detection: Indicates the average percentage of the Deterministic and Probabilistic Agreements on Lowest Detection; calculated by adding the Deterministic and Probabilistic Agreement over Lowest Detection percentages and dividing them by 2.

Overall Agreement over Mean Detection: Indicates the average percentage of the Deterministic and Probabilistic Agreements on Mean Detection; calculated by adding the Deterministic and Probabilistic Agreement over Mean Detection percentages and dividing them by 2.

Overall Agreement over Overlaps: Indicates the agreement in scoring (i.e., Deterministic Agreement + Probabilistic Agreement) out of the total number of statements which both coders coincided in scoring (detecting) (i.e., Deterministic Agreement + Probabilistic Agreement + Contradictions/Disagreement).

Overall Detection: Indicates the total number of statements which both coders coincided in scoring (detecting) (i.e., Deterministic Agreement + Probabilistic Agreement + Contradictions/Disagreement) out of the total number of statements scored (detected) by the first coder.*

*NOTE: Kindergarten data was scored 100% by one coder and 25% by a second coder.

Siegler's Overlapping Waves Theory of Microgenetic Change

- Path—*how the child sequenced his/her behaviors to get to the change.*
- Rate—*how quickly and with what supports the child moved from the realization of the new concept to its consistent application.*
- Breadth—*how narrowly or broadly the child gained the concept.*
- Variability—*the difference between children on the dimensions above.*

Most students were initially quite deterministic on the games tasks.

- This was expected for machines that are designed to work in a certain way, but not necessarily for games. *[On the bunny game: "You pick a card and then if you get one and you count like one two three and if you get a card and if you get the bunny on the circle, you turn the carrot and then the rabbit will fall." (ln. 10)]*

- Most of the students persistently pursued a pattern so that they could "be sure."

- For most students, this deterministic stance did not shift despite multiple opportunities to play the games and designed opportunities to make certain features of the game salient.

Some Students Maintained a Deterministic Stance Throughout…

- Even prior experiences in biology such as planting seeds were interpreted deterministically and explained in a reductionist manner.
- When seeking patterns, Andre referred to “finding evidence in science” and that if you look for the pattern, you’ll find evidence.

Jordan and Carter (K) had never planted before and they each predicted a one to one correspondence between seed planting and the number of resulting plants (16-16; 4-4 and 3-3; 21-21, and 5-5, respectively).

However, Carter purposely chose not to plant a cracked bean in the package because he said, “it wouldn’t grow.”

→Similar patterns were found in grade 2 despite a seed planting activity.

However, a few students at each grade level seemed to allow for probabilistic responses at the outset…

A few students approached the tasks in a more open manner—allowing for the possibility of probabilistic or deterministic responses.

Table 1. Comparing Response Patterns: Elena and Andre

		Elena		Andre	
Session	Task Domain	Deterministic	Probabilistic	Deterministic	Probabilistic
1	Mechanical/Game (Funny Bunny)	63%	36%	94%	5%
3	Mechanical/Game (Uno Attack)	60%	40%	90%	10%
5	Connections Across Domains	43%	57%	57%	43%

Maia (K, Session 1) engages notions of subjective uncertainty and objective uncertainty. She attends to the patterns in the game but holds a stance of uncertainty about what will happen.

- M: (22:30) *Gets a Carrot*
- I: Maia before you turn that, what do you think it’s going to make happen?
- M: Either one of the other holes are going to open, or one of the rabbits is going to fall down
- I: How do you know which?
- M: (22:51) I don’t know which one. I cannot.
- I: Cannot know… uh huh… do you know why you cannot?
- M: No
- I: (28:24) Do you have a prediction about what hole will open?
- M: I don’t know which one

Some students could generate their own examples of stochastic effects:

For instance, Layla (Gr. 2) says:

> "My shower, because sometimes when you turn the thing nothing comes out. And I'm like mom the shower stopped working! And she just tells me to get in the shower, so I go in the shower and all of a sudden water starts popping out and its cold and sometimes it will be piping hot. Oh and another thing, my baby brother does that. I'll be looking for him and he'll be standing in one place. And I'll be like Tyrone don't move, that's my brother's name, and then I'll go over there, and I'll come back and he's nowhere in sight. I'll go all the way around the house. Then I go in the hall way and I open the door and then boo out of nowhere. He pops out like the card machine and my shower sometimes doesn't work."

Those students responded in a more nuanced way to task features.

Table 2. Elena's Responses Across Domains

Session	Task Domain	Deterministic	Probabilistic
1	Mechanical/Game (Funny Bunny)	63%	36%
2	Mechanical (Bubble Gum Machine)	38%	61%
3	Mechanical/Game (Uno Attack)	60%	40%
4	Social (cheating; calling mom, pestering sister)	27%	73%
5	Connections Across Domains	43%	57%

In supported contexts, when contrasting domains, students were more likely to offer both kinds of statements.

Table 1. Comparing Response Patterns: Elena and Andre

		Elena		Andre	
Session	Task Domain	Deterministic	Probabilistic	Deterministic	Probabilistic
1	Mechanical/Game (Funny Bunny)	63%	36%	94%	5%
3	Mechanical/Game (Uno Attack)	60%	40%	90%	10%
5	Connections Across Domains	43%	57%	57%	43%

Certain contexts were more likely than others to elicit probabilistic responses…

- Gumball Machine
- Uno
- Social
- Seeds

Gumball Machine:

-Violation of Expectation

-Salience of Experience (Anchoring)

-One to One Correspondence

-Immediate Response (within attention set)

Kaylee (Gr. 2) appeared to make a breakthrough went she recounted in great detail the indignance of a gumball machine that did not deliver.

K: One time we went to a store. They had a gumball machine. My brother put one quarter in and he got NOTHING.

In unsupported contexts, Some students began to make a shift towards recognizing stochastic causes saying that they could not predict the outcome in every case, "just most of the time." They predicted what a best guess would be even if it "would not always be right."

By the fifth session, we see a dramatic increase in Rajon's use of the term risk. He gets a one-hop card. He moves his back bunny one space. He does not move his bunny on 20 to lower space 21.]

R: I'll risk moving this bunny [refers to back bunny] but I'm not risking that bunny [referring to bunny on 20].

R: It's gonna be on one of these low ones. [He sweeps his fingers across the holes on the top part of the game.] You can't predict so good.

Remaining issues….

- Learning opportunities over many instances is confounded with exploring different cases and contexts.
- It is possible that certain contexts invite students to map the features of the inherent causality better.
- It is possible that the varied cases allowed for learning.
- Supported learning opportunities resulted in more balanced responses.

Remaining Questions…

- A deterministic stance may be adaptive; might it reflect a sense of empowerment?
- How does an explicit notion of probabilistic causality impact reasoning about evidence in science?
- How does it impact the likelihood that we will detect causal patterns in the world that are difficult to detect and appear probabilistic? Can it?