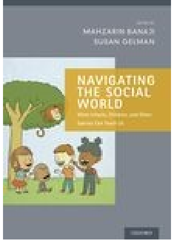

Navigating the Social World: What Infants, Children, and Other Species Can Teach Us
Mahzarin R. Banaji (ed.), Susan A. Gelman (ed.)

https://doi.org/10.1093/acprof:oso/9780199890712.001.0001

Published: 2013 **Online ISBN:** 9780199332779 **Print ISBN:** 9780199890712

CHAPTER

3.12 How Causal Learning Helps Us Understand Other People and How Other People Help Us Learn About Causes: Probabilistic Models and the Development of Social Cognition

Elizabeth Seiver, Alison Gopnik, Daphna Buchsbaum

https://doi.org/10.1093/acprof:oso/9780199890712.003.0034 Pages 186–190
Published: April 2013

Abstract

This chapter reviews studies on children's causal learning. These studies suggest a two-way interaction between causal learning and social knowledge. Children clearly use causal inference to draw important conclusions about the social world around them. At the same time their knowledge of the social world may itself shape the kinds of inferences they make. This back and forth between what we already know about people and what we learn about them means progress in understanding the complexities of social life. The new computational tools of probabilistic models and Bayesian inference can let us understand this learning in a deeper way.

Keywords: causal learning, social cognition, social knowledge, children, causal infence causal learning, social cognition, social knowledge, children, causal infence

Subject: Neuropsychology, Cognitive Psychology

Since the emergence of Theory of Mind research some 25 years ago, it has been a commonplace that understanding how other people function is crucially important for social life. Figuring out how perceptions lead to beliefs, or how intentions cause actions, can help us to make predictions about what people will do and to design ways to get them to do what we want. Developmental psychologists have charted the changes in children's causal understanding of the mind, and they have begun to extend this work to children's causal understanding of the social world more generally—their understanding of intelligence and effort, in-groups and out-groups, race and gender, and even good and evil.

Many researchers have suggested that our adult social knowledge is like a set of intuitive everyday theories—coherent, abstract causal representations that, much like scientific theories, allow us to make new predictions about the world around us and to act to change that world. Children develop and revise these theories in much the same way that scientists do. This "theory theory" has been articulated both in developmental and social psychology, though the similar ideas in the two fields have rather surprisingly only come together recently (see, e.g., Ames et al., 2001; Carey, 1985; Chiu, Hong, & Dweck, 1997; Gopnik & Meltzoff, 1997; Gopnik & Wellman, 1992; Wellman & Gelman, 1992).

Typically, research on the development of these intuitive social theories has examined when particular kinds of social knowledge emerge. But there is a deeper question to ask, not just when but why and how? Why do these changes take place? How is it possible for children to develop an increasingly accurate view of the social world, especially since the social world is even more subject to cultural and historical variation than the physical world? And then there is an even deeper question. The fundamental assumption of cognitive science is that we can understand the brain as a kind of computer designed by evolution and programmed by experience—though a computer far more powerful than any we know of now. What kinds of computations could underlie the remarkable human capacity for social learning specifically?

In the past we did not have the right tools for a computational account of intuitive theories in adults or their development in children. This reflected a deeper theoretical tension at the heart of developmental cognitive science. Children—even infants—have abstract, structured representations of the world, including intuitive theories of the mind and other people more generally. At the same time, children learn in prodigious amounts. They transform their representations of other people based on concrete experiences—the contingent, probabilistic evidence they see about what people do. How do children infer abstract psychological and social structure from the concrete contingencies of human action?

Until recently, the theoretical and computational possibilities on offer have not given a satisfying answer to this question. Connectionist theories allowed learning but denied that there are abstract representations; nativist theories allowed representation but denied that there is substantive learning. Although Piaget famously suggested "constructivism" as a middle path, he offered little detail about how constructivist
p. 187 processes might work. ↳

This problem is particularly vivid and relevant for social cognition. One might be able to provide an associationist account of relatively concrete kinds of knowledge like our ability to recognize objects or perform simple motor tasks. Or you might be able to justify a nativist account of such universally similar cognitive abilities as our ability to perceive three-dimensional spatial structure or to use syntax. But social cognition is one of the best examples of knowledge that is at once highly abstract and structured and highly variable across contexts and cultures.

Fortunately, over the past 10 years, researchers in philosophy of science and machine learning have started to explain how theories, including theories of other people, can be represented and changed through probabilistic models and Bayesian learning techniques (Glymour, 2001; Griffiths & Tenenbaum, 2007; Griffiths, Chater, Kemp, Perfors, & Tenenbaum, 2010; Pearl, 2000, Spirtes, Glymour, & Scheines, 1993). This framework promises a computationally precise developmental cognitive science that can integrate structure and learning.

Probabilistic causal models and Bayesian learning formalize the processes we are all familiar with as practicing scientists. As scientists we analyze the probabilistic, statistical patterns in the evidence; we perform experiments; and we use the data from these two sources to infer what causes what and to test and revise our theories. These ideas began in the philosophy of science as a way of saying mathematically how we can best make causal inferences from these kinds of evidence. But they can also help us understand how the mind actually does make these inferences.

In fact, over the past 10 years developmental and cognitive researchers have shown that adults and children actually can solve the problem of causal learning, in general, using implicit probabilistic models (e.g., Gopnik et al., 2004; Gopnik & Schulz, 2007; Gopnik & Tenenbaum, 2007; Lu et al., 2008; Sloman, 2005 for recent reviews see Gopnik 2012; Gopnik & Wellman, 2012). They use statistical evidence and experiments to infer causal structure in a Bayesian way. But almost all of these experiments have focused on physical causality. Researchers have only recently started to apply these ideas to psychological and social causality.

In our lab we have begun to use these ideas to explore some basic questions in social psychology. First, where do our explanations of human action come from? In particular, when do we think that a person's actions are the result of her long-lasting individual personality traits, and when do we think, instead, that those actions are caused by her immediate situation? Forty years of research in social psychology have shown that these attributions have important consequences for how we act and how we treat other people (e.g., Levy & Dweck, 1998). In fact, these attributions can literally be a matter of life and death. For example, did the Abu Ghraib guards torture because they were particularly sadistic people or because of isolation, stress, and a toxic institutional structure?

Western adults show a strong bias, sometimes called "the fundamental attribution error," to answer questions like these by referring to traits rather than situations (Jones & Harris, 1967; Kelley, 1967; Ross, 1977). People from other cultures show much less of a trait bias (Morris & Peng, 1994). Although this bias has been the focus of research in adult social psychology, we know very little about how or why trait attributions and biases develop in the first place. We know something about the "when"—the general consensus in the literature has been that children do not spontaneously explain actions in terms of traits or use those traits to make new predictions until they are about 7 years old—but we do not know why they start to do so (Alvarez, Ruble, & Bolger, 2001; Peevers & Secord, 1973; Rholes & Ruble, 1984; Shimizu, 2000).

An old idea is that we might infer traits based on consistent statistical patterns of behavior (Kelley, 1967). For example, if we see that Josie consistently takes risks across many situations while Anna is consistently risk averse, we may conclude that Josie and Anna's actions are caused by some internal personality trait rather than an external situation.

Probabilistic causal models let us formalize these inferences; they tell us just what sorts of causal inferences we should draw from statistical patterns like these (see Gopnik et al., 2004, Gopnik & Schulz, 2007). Bayesian versions of probabilistic models have an additional advantage: They can explain how our prior knowledge, the knowledge we have accumulated in the past, can interact with the new evidence we observe at a particular time to shape our inferences (Griffiths & Tenenbaum, 2007). We wanted to explore whether these causal learning processes might explain the emergence of trait attributions, in general, and the Western reliance on traits in particular (Seiver, Gopnik, & Goodman, 2012).

We showed 4-year-olds, who do not yet use trait explanations spontaneously, different patterns of
p. 188 contingency between people, situations, ↳ and actions. Anna and Josie are little dolls that can play on a miniature trampoline and bicycle. We showed half the children that Anna happily goes on the trampoline and leaps on the bicycle three out of four times, but Josie can only bring herself to get on the trampoline and bicycle one out of four times. We showed the other half of the children that Anna and Josie both bounce on the trampoline three out of four times but only dare approach the bicycle one out of four times. The events are the same, but the statistical patterns are different. If children are making causal inferences in a Bayesian way, they should conclude that the actions are more likely to be caused by traits in the first condition, and by situations in the second.

Then we asked the children to explain why Anna and Josie acted the way they did and to predict what they would do in a new situation. In the first condition, children explained the character's actions in terms of traits (e.g., "She's brave" or "She knows how to ride a bike"), and they predicted that she would continue to be brave in new situations—she would go off the diving board, too. The second group said the children acted that way because of situations—the trampoline was safe and the bicycle was dangerous, and this was also reflected in their predictions. So the pattern of probabilistic covariation led the children to prefer one type of explanation and prediction to the other.

Bayesian inferences depend on two factors, however: the evidence you see and your prior belief in the probability of the hypotheses under consideration. The 4-year-olds in our experiment weighed the evidence equally in the two conditions, which suggested that they initially thought that trait and situation attributions were equally likely. But 6-year-olds already showed a bias toward traits. They did take the evidence into account, but they consistently thought trait explanations were more likely than situation explanations, even when the immediate evidence did not support this. This would make sense, however, if between 4 and 6, children found consistently more support for trait than situation hypotheses, both directly in their observations of what people do and indirectly in the information they receive from others. Such a pattern of evidence would lead to a stronger prior probability for trait rather than situation hypotheses. To explore this possibility, we are testing children in China, who may have less strong evidence for traits than American children.

The attribution literature concerns high-level explanations for human action. But before we can explain actions, we have to identify them. Although this might seem simple, in fact, inferring intentional goal-directed actions from a continuous stream of movements is challenging (Baldwin, Andersson, Saffran, & Meyer, 2008). This difficulty is apparent in the literature on imitation; sometimes children "overimitate," including all the details of a complex adult action (Lyons, Young, & Keil, 2007), and sometimes they only imitate the parts of that action that are relevant to the particular outcome (Gergely, Bekkering, & Király, 2002; Williamson, Meltzoff, & Markman, 2008). Why and how do children choose one analysis of action over another?

In another series of studies, we have shown that children can use statistical information to analyze and imitate actions as well as to explain them (Buchsbaum, Gopnik, Griffiths, & Shafto, 2011). We showed 4-year-old children five sequences of three actions each on a toy (e.g., the experimenter would squish it, then shake it, and then roll it). Some of these sequences, but not others, led to the toy playing music. Children could either imitate all the actions they saw or just choose a subset of the most effective actions. A Bayesian model worked out the probability that different sequences would cause the machine to go, given the patterns of evidence. The children's behavior fit this model well. When the evidence suggested that an action sequence was more likely to produce the effect, they were more likely to produce that sequence, even if they had never seen it in isolation. So children's imitation looked like a rational attempt to use the evidence to make sense of the actions they saw. Imitation, a major engine of social development, reflects a kind of probabilistic causal inference.

A last twist to this study makes an important point, however. In one condition the experimenter acted inept, as if she did not know anything about how the machine worked. In that case the children relied on the evidence. But in another condition the experimenter told the children that she was showing them how the machine worked—and then showed them exactly the same sequence of actions. In that pedagogical condition children were much more likely to overimitate, mimicking everything the experimenter did. So children made different causal inferences depending on the social context.

These studies suggest that there is a two-way interaction between causal learning and social knowledge. Children clearly use causal inference to draw important conclusions about the social world around them. At
p. 189 the same time their ↳ knowledge of the social world may itself shape the kinds of inferences they make, as in the contrast between the pedagogical teacher and the inept demonstrator. In a way that Piaget would have appreciated, this back and forth between what we already know about people and what we learn about them lets us slowly but surely make progress in understanding the complexities of social life. The new computational tools of probabilistic models and Bayesian inference can let us understand this learning in a deeper way.

Acknowledgments

This research was supported by the J.S. McDonnell Causal Learning Collaborative. We thank Adrienne Wente for invaluable assistance.

References

Alvarez, J. M., Ruble, D. N., & Bolger, N. (2001). Trait understanding or evaluative reasoning? An analysis of children's behavioral predictions. *Child Development,* 72 (5), 1409–1425. 10.1111/1467-8624.00356
Google Scholar WorldCat Crossref

Ames, D. R., Knowles, E. D., Rosati, A. D., Morris, M. W., Kalish, W. C., & Gopnik, A. (2001). The social folk theorist: Insights from social and cultural psychology on the contents and contexts of folk theorizing. In B. Malle, L. Moses, & D. Baldwin (Eds.), *Intentions and intentionality: Foundations of social cognition* (pp. 287–307). Cambridge, MA : MIT Press.
Google Scholar Google Preview WorldCat COPAC

Baldwin, D., Andersson, A., Saffran, J., & Meyer, M. (2008). Segmenting dynamic human action via statistical structure. *Cognition,* 106 (3), 1382–1407. 10.1016/j.cognition.2007.07.005
Google Scholar WorldCat Crossref

Buchsbaum, D., Gopnik, A., Griffiths, T. L., & Shafto, P. (2011). Children's imitation of causal action sequences is influenced by statistical and pedagogical evidence. *Cognition,* 120(3), 331–340. 10.1016/j.cognition.2010.12.001
Google Scholar WorldCat Crossref

Carey, S. (1985). *Conceptual change in childhood.* Cambridge, MA : MIT Press. Press/Bradford Books.
Google Scholar Google Preview WorldCat COPAC

Chiu, C. Y., Hong, Y. Y., & Dweck, C. S. (1997). Lay dispositionism and implicit theories of personality. *Journal of Personality and Social Psychology,* 73 (1), 19–30. 10.1037/0022-3514.73.1.19
Google Scholar WorldCat Crossref

Gergely, G., Bekkering, H., & Király, I. (2002). Rational imitation in preverbal infants. *Nature,* 415 (6873), 755. 10.1038/415755a
Google Scholar WorldCat Crossref

Glymour, C. N. (2001). *The mind's arrows: Bayes nets and graphical causal models in psychology.* Cambridge, MA: MIT Press.
Google Scholar Google Preview WorldCat COPAC

Gopnik, A. (2012). Scientific thinking in very young children: Theoretical advances, empirical discoveries and policy implications. *Science,* 28 September 2012, 1623–1627. 10.1126/science.1223416
Google Scholar WorldCat Crossref

Gopnik, A., Glymour, C., Sobel, D. M., Schulz, L. E., Kushnir, T., & Danks, D. (2004). A theory of causal learning in children: causal maps and Bayes nets. *Psychological Review,* 111 (1), 3–32. 10.1037/0033-295X.111.1.3
Google Scholar WorldCat Crossref

Gopnik, A., & Meltzoff, A. (1997). *Words, thoughts and theories.* Cambridge, MA : MIT Press.
Google Scholar Google Preview WorldCat COPAC

Gopnik, A., & Schulz, L. (Eds.) (2007). *Causal learning: Psychology, philosophy, and computation.* Oxford University Press. 10.1093/acprof:oso/9780195176803.001.0001
Google Scholar Google Preview WorldCat COPAC Crossref

Gopnik, A., & Tenenbaum, J. (2007). Bayesian networks, Bayesian learning and cognitive development. *Developmental Science (Special section on Bayesian and Bayes-Net approaches to development),* 10 (3), 281–287.
Google Scholar WorldCat

Gopnik, A., & Wellman, H. M. (1992). Why the child's theory of mind really is a theory. *Mind and Language,* 7 (1–2), 145–171. 10.1111/j.1468-0017.1992.tb00202.x
Google Scholar WorldCat Crossref

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

Griffiths, T. L., Chater, N., Kemp, C., Perfors, A., & Tenenbaum, J. B. (2010). Probabilistic models of cognition: Exploring representations and inductive biases. *Trends in Cognitive Sciences,* 14 (8), 357–364. 10.1016/j.tics.2010.05.004
Google Scholar WorldCat Crossref

Griffiths, T. L., & Tenenbaum, Joshua, B. (2007). Two proposals for causal crammar. In A. Gopnik & L. Schulz (Eds.), *Causal learning psychology, philosophy, and computation* (p. 323). New York : Oxford University Press. 10.1093/acprof:oso/9780195176803.003.0021
Google Scholar Google Preview WorldCat COPAC Crossref

Jones, E. E., & Harris, V. A. (1967). The attribution of attitudes. *Journal of Experimental Social Psychology,* 3 (1), 1–24. 10.1016/0022-1031(67)90034-0
Google Scholar WorldCat Crossref

Kelley, H. H. (1967). Attribution theory in social psychology. In D. Levine (Ed.), *Nebraska symposium on motivation* (Vol. 15, pp. 192–238). Lincoln : University of Nebraska.
Google Scholar Google Preview WorldCat COPAC

Levy, S. R., & Dweck, C. S. (1998). Trait- versus process-focused social judgment. *Social Cognition,* 16 (1), 151–172. 10.1521/soco.1998.16.1.151
Google Scholar WorldCat Crossref

Lu, H., Yuille, A. L., Liljeholm, M., Cheng, P. W., & Holyoak, K. J. (2008). Bayesian generic priors for causal learning. *Psychological Review,* 115 (4), 955–984. 10.1037/a0013256
Google Scholar WorldCat Crossref

Lyons, D. E., Young, A. G., & Keil, F. C. (2007). The hidden structure of overimitation. *Proceedings of the National Academy of Sciences USA, 104*, 19751–19756.
Google Scholar Google Preview WorldCat COPAC

Morris, M. W., & Peng, K. (1994). Culture and cause: American and Chinese attributions for social and physical events. *Journal of Personality and Social Psychology,* 67, 949–971. 10.1037/0022-3514.67.6.949
Google Scholar WorldCat Crossref

p. 190 Pearl, J. (2000). *Causality.* Cambridge, England : Cambridge University Press. ↳
Google Scholar Google Preview WorldCat COPAC

Peevers, B. H., & Secord, P. F. (1973). Developmental changes in attribution of descriptive concepts to persons. *Journal of Personality and Social Psychology,* 27, 120–128. 10.1037/h0034469
Google Scholar WorldCat Crossref

Rholes, W. S., & Ruble, D. N. (1984). Children's understanding of dispositional characteristics of others. *Child Development,* 55, 550–560. 10.2307/1129966
Google Scholar WorldCat Crossref

Ross, L. (1977). The intuitive psychologist and his shortcomings: Distortions in the attribution process. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 10, pp. 173–220). New York : Academic Press.
Google Scholar Google Preview WorldCat COPAC

Seiver, E., Gopnik, A,, & Goodman, N. (2012). Did she jump because she was the big sister or because the trampoline was safe? Causal inference and the development of social attribution. *Child Development, on-line publication* 24 SEP 2012 DOI: 10.1111/j.1467–8624.2012.01865.x 10.1111/j.1467–8624.2012.01865.x
Google Scholar WorldCat Crossref

Shimizu, Y. (2000). Development of trait inference: Do young children understand the causal relation of trait, motive, and behavior? *Japanese Journal of Educational Psychology,* 48 (3), 255–266.
Google Scholar WorldCat

Sloman, S. A. (2005). *Causal models: How people think about the world and its alternatives.* New York: Oxford University Press.
Google Scholar Google Preview WorldCat COPAC

Spirtes, P., Glymour, C., & Scheines, R. (1993). *Causation, prediction, and search.* (2nd ed.). New York : Springer-Verlag. 10.1007/978-1-4612-2748-9
Google Scholar Google Preview WorldCat COPAC Crossref

Wellman, H. M., & Gelman, S. A. (1992). Cognitive development: Foundational theories of core domains. *Annual Review of Psychology*, *43*, 337–375. 10.1146/annurev.ps.43.020192.002005
Google Scholar WorldCat Crossref

Williamson, R. A., Meltzoff, A. N., & Markman, E. M. (2008). Prior experiences and perceived efficacy influence 3-year-olds' imitation. *Developmental Psychology,* 44 (1), 275–285. 10.1037/0012-1649.44.1.275
Google Scholar WorldCat Crossref