Cognition 137 (2015) 137–153

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

On the acquisition of abstract knowledge: Structural alignment and explication in learning causal system categories

Micah B. Goldwater[a,*], Dedre Gentner[b,1]

[a] University of Sydney, School of Psychology, Brennan MacCallum Building (A18), Camperdown, NSW 2006, Australia
[b] Northwestern University, Department of Psychology, 2029 Sheridan Road, Evanston, IL 60208, United States

ARTICLE INFO

Article history:
Received 31 March 2014
Revised 1 December 2014
Accepted 2 December 2014

Keywords:
Analogy
Causal learning
Categorization
Education

ABSTRACT

This research studies a relatively unexplored aspect of expertise – the ability to detect causal relational patterns in multiple contexts – and demonstrates learning processes that foster this ability. Using the Ambiguous Sorting Task (AST), in which domain information competes with causal patterns, we previously found that science experts spontaneously noticed and sorted by causal patterns such as positive feedback, while novices sorted primarily by content domain. We investigated two kinds of learning experiences that we claim are needed to achieve high fluency in detecting key cross-domain patterns. We found that *direct explication* of example phenomena increased people's accuracy in depicting the examples, but did not increase sensitivity to the causal patterns in new examples. However, *analogical comparison* between parallel examples did lead to greater propensity to detect the causal patterns across diverse examples. Combining within-example explication with between-example alignment led to the greatest gains in generalized sensitivity to causal patterns.

An important and understudied aspect of expertise is the ability to spontaneously notice key relational patterns in the flow of experience. For example, the Swiss inventor George de Mestral developed the idea of Velcro when he and his dog returned from an Alpen hike with masses of burrs in their clothes and fur. During the tedious process of removing the burrs, he began to focus on their extraordinary clinging power. After examining the burrs with a microscope, he came up with the idea for a reversible fastener, with stiff hooks (like the burrs) on one side and soft loops (like fur or fabric) on the other side. This ability to see beyond the routine irritation of dealing with burrs to a valuable insight suggests a creative mind at work. But it also underlines the importance of a *prepared* mind. De Mestral was a trained engineer, working in the machine shop of an engineering company and pursuing his own inventions on the side. (He received a patent for a toy airplane at the age of 12). His amassed experience in how things work gave him a rich internal vocabulary with which to interpret causal patterns that on the surface bear little resemblance to his everyday work.

Our question here is how people acquire generalized sensitivity to key causal patterns. We focus on causal patterns because of the importance and pervasiveness of causality in human cogition (Ahn, Kim, Lassaline, & Dennis, 2000; Mackie, 1980; Sloman, 2005). Having an abstract understanding of causal structure allows for deep connections to be made across domains. For example, both the melting of polar icecaps and the growth of economic pricing bubbles are governed by a *positive feedback* causal structure. An understanding of causal structure is critical for explanation and prediction (e.g., Lombrozo & Carey, 2006; Sloman, 2005; Thagard, 1989). It influences category

* Corresponding author. Tel.: +61 2 9351 5453; fax: +61 2 9351 2603.
E-mail addresses: micah.goldwater@sydney.edu.au (M.B. Goldwater), genter@northwestern.edu (D. Gentner).
[1] Tel.: +1 (847) 467 2035; fax: +1 (847) 491 7859.

http://dx.doi.org/10.1016/j.cognition.2014.12.001

organization (Ahn et al., 2000; Rehder & Burnett, 2005; Sloman, Love, & Ahn, 1998) and is reflected in linguistic structure (Fillmore, 1978; Jackendoff, 1983; Kuehne & Forbus, 2002; McCawley, 1968; Wolff & Song, 2003). For this reason, causal knowledge and reasoning has been a focus of cognitive science from the outset (de Kleer & Brown, 1981; Forbus, 1984; Hayes, 1979). There has been extensive research on how causal knowledge is represented, using formalisms such as qualitative process models (Forbus, 1985; Forbus, Nielsen, & Faltings, 1991) or causal Bayesian networks (e.g., Gopnik et al., 2004; Pearl, 2000; Waldmann, Hagmayer, & Blaisdell, 2006).

Empirical work on causality has examined how people determine the causal structure of a particular domain or phenomenon. While people typically find it quite difficult to infer complex causal patterns purely from observation (e.g., Lagnado & Sloman, 2004; Steyvers, Tenenbaum, Wagenmakers, & Blum, 2003), this research has revealed kinds of experiences that lead people to infer causal relationships within a domain. These include being exposed to particular kinds of statistical relations among variables (Cheng, 2000), through direct causal interventions (e.g., Hagmayer, Sloman, Lagnado, & Waldmann, 2007) or to evidence of causal mechanisms (Ahn & Kalish, 2000; Ahn, Kalish, Medin, & Gelman, 1995; Lagnado, Waldmann, Hagmayer, & Sloman, 2007; Rehder & Hastie, 2001), and engaging in self-explanation (Chi, de Leeuw, Chiu, & LaVancher, 1994; Lombrozo, 2010) or counter-factual reasoning (e.g., Harris, German, & Mills, 1996).

Our focus here is on different. Instead of examining how people learn and use the causal relations that govern a particular phenomenon, we ask how people learn abstract categories of causal systems that apply across domains and phenomena. In our prior work we developed a sorting task aimed at assessing people's propensity to notice key causal patterns amidst competing information (Rottman, Gentner, & Goldwater, 2012). In this task, the *Ambiguous Sorting Task* (AST), subjects are asked to sort descriptions of causal phenomena into categories. They are given an array of five 'seed cards' to use in sorting (as well as an "Other" category). Each of the five seed cards depicts a different causal system and a different content domain from the others. This means that subjects are free to sort either by causal structure or by content domain. Likewise, the phenomena descriptions they are asked to sort vary both in their content domain (biology, economics, etc.) and in their causal structure (positive feedback, causal chain, etc.). (see Table 1 for the entire list.) The idea here is that people can achieve a successful sort simply by attending to the relatively obvious domain-level commonalities; thus, there is no need to seek some other sorting principle. However, if people spontaneously notice the causal commonalities, they may choose to use these instead.

To discover whether fluent knowledge of abstract causal patterns varies with expertise, we gave the AST to advanced physical science students and to social science students. There were two results of interest. First, students in social science and economics sorted primarily by content domain – evidence that the causal patterns were not obvious even to college students. Second, advanced physical science students – many of whom had taken courses in multiple science disciplines – sorted primarily by causal system – evidence that fluent perception of causal patterns increases with expertise (Rottman et al., 2012; and see Chi, Feltovich, & Glaser, 1981 for a similar result). To analogize to the de Mestral example, there was a shift from focusing on the domain level (removing burrs) to focusing on a cross-domain abstract principle (achieving an adhesive connection).

Our goal in the present research is to understand how people come to be sensitive to abstract causal patterns. We hypothesize that at least two kinds of learning experiences are needed for forming abstract causal categories. The first is experience and instruction on particular causal phenomena, as in much of the prior research on causal learning. Learning the causal structure of particular cases is important, but we suggest that by itself it is not enough. The second important contributor is a way to form abstract causal representations that apply across domains. We suggest that analogical comparison of cases in which the same causal system applies can achieve such abstractions.

These two contributors operate in different ways. The first contributor, encountering causal explanations for specific phenomena, seems likely to improve causal understanding of specific phenomena. But by itself it is unlikely to promote general causal abstractions. Rather, we hypothesize that comparing analogous phenomena from different domains is critical in promoting abstraction of the common causal structure. There is considerable prior evidence consistent with the idea that analogical comparison highlights common relations through a process of structural alignment (Falkenhainer, Forbus, & Gentner, 1989; Gentner & Markman, 1997), and that such an alignment permits learners to notice the common structure, which can then be applied more broadly. The idea that analogical comparison promotes transfer from specific learning contexts has received support from research with adults (Gick & Holyoak, 1983; Goldwater & Markman, 2011; Kurtz, Boukrina, & Gentner, 2013) and children (Christie & Gentner, 2010; Gentner, Anggoro, & Klibanoff, 2011) and by studies in educational contexts (Gentner, Loewenstein, & Thompson, 2003; Klahr & Chen, 2011; Rittle-Johnson & Star, 2009).

However, although these studies show that analogical comparison can improve transfer, in general it is not clear that it does so via abstracting the common structure. Analogical comparison can also improve the individual case representations, and this in itself could improve transfer (e.g. Gary, Wood, & Pillinger, 2012). Because our study includes a separate (and prior) assessment of the accuracy of the case representations, as described below, we can test whether there are effects of comparison on abstraction over and above its effects on representation accuracy.

Thus, we hypothesize that both causal explication and structural alignment will lead to increased sensitivity to causal structure, but for different reasons. Causal explication of the training examples will lead to better representation of the causal structure of the individual examples. Structural alignment of training examples will lead to abstraction of the causal pattern, and therefore to increased ability to perceive that pattern in other phenomena. Based on this reasoning, a further prediction is that