Cognitive Psychology 65 (2012) 381–413

Contents lists available at SciVerse ScienceDirect

Cognitive Psychology

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

Causal imprinting in causal structure learning

Eric G. Taylor *, Woo-kyoung Ahn

Yale University, Psychology Department, 2 Hillhouse Avenue, New Haven, CT 06520, United States

A R T I C L E I N F O

Article history:
Accepted 1 July 2012
Available online 1 August 2012

Keywords:
Causal structure learning
Causal reasoning
Belief revision
Cognitive biases
Bayesian modeling

A B S T R A C T

Suppose one observes a correlation between two events, B and C, and infers that B causes C. Later one discovers that event A explains away the correlation between B and C. Normatively, one should now dismiss or weaken the belief that B causes C. Nonetheless, participants in the current study who observed a positive contingency between B and C followed by evidence that B and C were independent given A, persisted in believing that B causes C. The authors term this difficulty in revising initially learned causal structures "causal imprinting." Throughout four experiments, causal imprinting was obtained using multiple dependent measures and control conditions. A Bayesian analysis showed that causal imprinting may be normative under some conditions, but causal imprinting also occurred in the current study when it was clearly non-normative. It is suggested that causal imprinting occurs due to the influence of prior knowledge on how reasoners interpret later evidence. Consistent with this view, when participants first viewed the evidence showing that B and C are independent given A, later evidence with only B and C did not lead to the belief that B causes C.

1. Introduction

Imagine you are a scientist investigating the causes of myopia (near-sightedness). You discover a correlation between using a nightlight as a child and developing myopia as an adult, and based on this correlation, you infer that nightlights are a cause of myopia. To support this inference, you also develop a theory that explains how light exposure during the evening could alter the eye's physiology (Quinn, Shin, Maguire, & Stone, 1999). A year later, however, new evidence emerges that a third factor—whether the child's parents have myopia—explains away the original correlation (Gwiazda, Ong,

* Corresponding author. Fax: +1 203 432 7172.
E-mail address: eric.taylor@yale.edu (E.G. Taylor).

0010-0285/$ - see front matter
http://dx.doi.org/10.1016/j.cogpsych.2012.07.001

Held, & Thorn, 2000). Parents who have myopia both tend to use a nightlight in their child's room and tend to pass along myopia to their children. For children whose parents do not have myopia, the original correlation disappears, undermining the belief that nightlights are a cause of myopia.

Examples like the one above (taken from a series of articles published in the journal *Nature*) are commonplace in science and in everyday causal reasoning. Presumably due to their abundance, students are repeatedly told in research methods and statistics courses that correlation does not imply causation and that they should look out for a hidden common cause that accounts for the observed correlation. Yet, as noted in the above examples, people frequently neglect such warnings. The main question we address in the current work is: Given that people frequently do infer causality from correlations without considering potential common causes, what happens to those beliefs when a hidden common cause is revealed?

One possibility is that people reconsider the initial evidence in light of the common cause, which leads them to discard their initial belief. For example, as the scientists did in the myopia scenario, reasoners might use their knowledge of parent myopia to discard their belief that nightlights cause child myopia. We refer to this behavior as *belief revision*. Another possibility is that reasoners do not reconsider the initial evidence, but instead maintain their initial belief and merely add the common cause relations. We refer to this behavior as *causal imprinting*, based on the idea that the initial evidence imprints a belief into the reasoner's mind, making it difficult to dispel despite the later evidence.

In what follows, we first review previous studies on causal structure learning and discuss the reasons to expect either belief revision or causal imprinting. Then, we present our empirical framework for distinguishing between these two outcomes, including a Bayesian analysis of the data we present to participants. Our Bayesian analysis shows under what conditions causal imprinting is normative, and the current study empirically tests these conditions.

1.1. Causal structure learning

Many previous studies have examined how people use covariation evidence to judge whether two events are causally related (Buehner, Cheng, & Clifford, 2003; Cheng, 1997). For example, one can use the data from a number of medical patients to judge whether taking a pill causes or prevents a headache. When learning only a single causal relation, people tend to give approximately normative judgments, consistent with statistical models of causal inference (Buehner et al., 2003; Griffiths & Tenenbaum, 2005, 2009; Rottman, Ahn, & Luhmann, 2011).

However, many causal judgments involve situations where there are more than just two events, and in these tasks people behave less normatively (Lagnado & Sloman, 2004, 2006; Steyvers, Tenenbaum, Wagenmakers, & Blum, 2003). For example, Steyvers et al. (2003) found that fewer than half of their participants could distinguish a common cause structure (A causes B and C) from a common effect structure (B and C cause A) based on observations alone. Inferring the causal relations between multiple events is difficult because it requires verifying which events are merely correlated, and which are genuinely causally related (Pearl, 2000). As in the myopia example, a set of three events may all be correlated with each other because one is the common cause of the other two. In this case, one may perform numerous computations to verify the true causal structure: first, the overall contingency between each pair of events, and then, the conditional contingency of each pair of events, given the value of the third event (Scheines, Spirtes, Glymour, & Meek, 1994).

Despite these apparent difficulties, people do have knowledge of complex, real world causal systems, such as economic trends, weather patterns, and social hierarchies. Such learning may be possible because people assemble their causal models piece by piece, learning one relation at a time, rather than attempting to learn them all at once (Ahn & Dennis, 2000; Fernbach & Sloman, 2009).

If the learning of causal structures often occurs incrementally, then it is crucial to know how people update their initial causal beliefs when faced with new evidence. This is especially true when the new evidence reveals previously hidden causal factors that lead to different interpretations of the initial evidence. Though some work has addressed how people reason about hidden causes (Rottman et al., 2011), the current studies are the first to examine how learners update their beliefs when these hidden causes are revealed.