EDITORIAL

Causal Learning Beyond Causal Judgment: An Overview

José C. Perales[1] and David R. Shanks[2]

[1]*Faculty of Psychology, Department of Experimental Psychology, University of Granada, Granada, Spain*

[2]*Division of Psychology and Language Sciences, University College London, London, England*

Most research articles studying how people learn to detect causal relationships in their environments commence with some sort of example to illustrate the relevance of causality in our daily lives. These examples allude to routine problems faced by doctors, economists, social psychologists, and others and emphasize the importance of deepening our understanding of causal reasoning. But despite these frequent applied examples, it is somewhat surprising that research on causal learning has only had a modest impact in applied disciplines. After three decades or so of intense study, it is probably time to wonder why this is the case. Plainly, we do not want causal learning to become a super-specialized topic, perfectly constructed but unable to generate useful knowledge of wider relevance.

In our view, cross-boundary work to fulfill this ambition is being undertaken, but to make a full impact it requires a reformulation of the implicit paradigm for causal learning research. The core of this paradigm is simple and can be summarized in two principles: first, causal knowledge can be assessed by means of verbal or numerical judgments of causal strength, and second, a single mental algorithm is sufficient to account for how environmental conditions (including covariation between cues and outcomes, time delays, and statistical interactions) map onto judgments. This research program has produced an impressive corpus of data (see [1] for a recent review), but also a current feeling that wider progress and impact is not being achieved.

This special issue is a joint attempt to present a vision of how research on causal learning might develop in the future, and to push that process forward. With regard to the first principle, it is important to acknowledge that judgments are not the only way to assess causal knowledge. Judgments reflect causal beliefs, and causal beliefs are probably the basis for other responses, such as decisions or interventions. But it is not possible to predict decisions or interventions on the basis of judgments alone. It would be naïve to think that causal beliefs reflected in simple causal judgments are the sole input to decision-making and intervention processes. Much effort is necessary to ascertain how causal knowledge is employed in all of these competencies, so we can build bridges between what we have discovered in recent decades and other aspects of behavior.

With regard to the second principle, we argue that a reconsideration of how theory needs to develop in the future is also necessary. Most researchers now accept that people interpret the world as causal, and build mental representations of their environment in which events are causally related. Still, these causal models must be constructed from some sort of evidence, and that evidence is provided by basic coding mechanisms capturing regularities in the environment. In other words, causal learning not only serves to detect and code statistical regularities, but also to uncover the hidden causal structure that generates those regularities. For example, the correlation between smoking and lung cancer has been known for a long time. However, some scientists and tobacco manufacturers denied the existence of a causal link between the two variables, because they believed that some other causal factor was responsible for the co-occurrence (for example, populations from certain social origins could be more likely both to smoke and suffer cancer). Obviously, if smoking were not a direct cause of cancer, it would be useless to recommend people to quit smoking.

In theoretical terms, we need some basic coding mechanism(s) to capture statistical regularities, and some other mechanism(s) to infer causal structures from them. Most psychologists and neuroscientists would accept that the brain is a hugely sophisticated form of connectionist net, capable of building quite reliable models of the regularities in our experiences and interactions with the world (see Morís, Cobos, & Luque's paper in this volume). Miller's comparator model [2], and Allan's [3] recent work, emphasize the idea that basic coding processes, either associative or episodic, can generate representations of the world in which, with sufficient attentional resources, most relevant events, their conjunctions, their relations of time and order, and their statistical dependencies, are conserved.

Additionally, a number of algorithms have been postulated in artificial intelligence that are capable of using the sort of output generated by these basic coding mechanisms to build structural causal representations [4, 5]. The limits of bounded rationality and actual research indicate, however, that the use of such algorithms requires the management of quantities of information that are beyond human processing limits. So, causal induction is also a learning problem: certain second-order cues (abstract features of interrelations among cues) can indicate what is a cause and what is not. We, and most of the authors in this volume, support this cues-to-causality approach (see the papers by Lagnado & Speekenbrink, and Hagmayer et al. in this volume, and [6], for more detailed discussions). Much research is needed to ascertain how we learn to manage these cues, the quantitative

and qualitative role of cue information in the formation of causal beliefs and in open behavior, and how cues interact when more than one are simultaneously present[1].

ORGANIZATION OF THE VOLUME

This volume is organized in three sections. In the first, the authors deal with *interactions between causality and low-level cognitive processes*. **Fugelsang** and **Roser**'s paper is aimed at describing the conditions under which perceptual causality occurs, and reviews the evidence showing that perceptual causality depends on automatic low-level perceptive processes not implying any kind of rational inference. **Morís**, **Cobos**, and **Luque**'s paper, on the other hand, considers the relationship between low-level cognition and causality in the opposite direction: how automatic associative activation diffusion processes (in a connectionist net) determine learning in contingency judgment tasks. The data they present indicate that pairings between cues and outcomes in contingency learning tasks automatically generate associations which, in turn, can account for priming effects. These priming effects seem to reflect competitive interactions between cues in the formation of associations. The question whether these competition effects are incorporated into a connectionist model which, in turn, serves as the basis for causal inference remains open, thus raising important issues about the possible multiple locus of cue interaction effects.

The second section is concerned with *the interplay between causality, reasoning, and decision making*, or the interaction between causality and forms of reasoning and behavior beyond causal judgment.

Oaksford and Chater's article stresses the importance of structural relations (including causal relations in net-like structures) in determining how people understand and endorse conditionals and draw inferences from them. Given that those structures are necessarily part of reasoners' knowledge, the main thesis is that causal learning is at the core of human language and reasoning.

Similarly, **Meder**, **Gerstenberg, Hagmayer**, and **Waldmann**'s paper assigns a central role to causal structures in interventional decisions. When faced with complex causal systems, where should the individual (for example a politician who wants to improve the results of a health system) intervene? The authors provide a heuristic for interventional decisions based on the observed probability of the effect given each known cause, which operates on the basis of a skeletal causal structure that consists of the desired effect and its direct and indirect causes.

García-Retamero, **Hoffrage**, **Müller**, and **Maldonado**'s interest also focuses on the role of causal learning on decision making, but from a different perspective. When deciding which of two companies is going to do best in the stock market, an investor will not search all the cues indicative of which company will do better, but will tend to select those she thinks are causally related to the criterion. In addition, when assessing the relationship between each of the cues and the criterion on the basis of the feedback received for her choices, causal knowledge will determine how such relationships are perceived, and thus will determine future decisions.

The three previous papers in this section are mostly about how causal knowledge determines reasoning and decision making. The importance of statistical dependencies in the updating of causal knowledge is well-appreciated, but where else does causal knowledge come from? As noted above, most research on causation has focused on the question of how statistical dependencies map onto judgments. **Hagmayer**, **Meder**, **Osman**, **Mangold**, and **Lagnado** consider the intriguing hypothesis that people interacting with and trying to control a dynamic causal system spontaneously discover important structural features of the system they are acting upon. In other words, active manipulation provides cues to causality that cannot be accounted for by mere instrumental learning. This closes an incremental loop in which causal knowledge alters the pattern of interventions, and intervention generates cues for enriching the causal structures that supported intervention in the first place.

Finally, the third section presents *new theoretical developments* in the mapping between environmental conditions and causal beliefs. More particulalrly, it reviews some classic controversies in the field and reformulates them in integrative ways. **Ramos** and **Catena** present a model for causal learning that attempts to go beyond the traditional dialogue between rule-based and associative theories. In their model, raw cue-outcome and cue-cue statistical dependencies are computed without interaction. Cue interaction effects will occur – or not – depending on a number of factors, including features of the experimental procedure and the prior knowledge of the learner, together with knowledge about potential causal structures supporting the observed dependencies. Importantly, the mechanism underlying cue interaction is not conditionalization, as assumed in most inferential theories, but relativization, a more powerful mechanism based on the same assumptions as linear regression in statistical methods.

Perales, **Shanks**, and **Lagnado** insist on the centrality of causality, and try to make it independent of the mechanism/covariation controversy. Their proposal is that, once knowledge about a relation between events reaches a causal status, their connection is irreducible, either to mechanisms (mechanisms are perceptually unreachable) or covariations (covariation does not imply causation), and acquires functional properties independent of the origin of the evidence on which that belief is based. Causal knowledge can be acquired either from information about potential mechanisms connecting cause and effect, or from observed statistical dependencies. However, the origin is irrelevant for further updating processes.

[1]This approach shares some features with the heuristic-toolbox approach to decision making (see García-Retamero et al's paper in this volume), in the sense that cues to causality can be interpreted as heuristics to decide whether a perceived relationship is causal or not. It also shares with Gopnik's modular theory [7] the idea that causal knowledge requires some inference at unveiling causal links behind statistical and temporal relations, although the modular theory assumes the existence in our brains of a rationality-mimicking and hard-wired "causal learning" module responsible for such inference processes.

1874-3501/10 2010 Bentham Open

Finally, **Lagnado** and **Speekenbrink** focus on the effect of delays between candidate causes and effects in causal inference. This is one of the seminal effects in causal learning research and one of the few that gave rise to the original causal learning program. The authors revisit this effect and qualify it. First, they demonstrate that not only delay, but also delay variability, hinders causal judgments. And second, delays imply higher probabilities of other events intervening between the candidate cause and the effect; the probability of these intervening effects can be manipulated orthogonally to delay, and easily overcomes its effect. In summary, whatever the origin of these effects (associative or inferential), they make rational sense as they maximize the probability of discovering causal relations where they really exist.

REFERENCES

[1] Perales JC, Shanks DR. Models of covariation-based causal judgment: a review and synthesis. Psychonomic Bull Rev 2007; 14: 577-96.
[2] Stout SC, Miller RR. Sometimes-competing retrieval (SOCR): a formalization of the comparator hypothesis. Psychol Rev 2007; 114: 759-83.
[3] Allan LG, Siegel S, Tangen JM. A signal detection analysis of contingency data. Learn Behav 2005; 33: 250-63.
[4] Pearl J. Causality: Models, reasoning, and inference. Cambridge: Cambridge University Press 2000.
[5] Glymour C. Learning causes: psychological explanations of causal explanation. Mind Mach 1998; 8: 39-60.
[6] Perales JC, Catena A. Human causal induction: a glimpse at the whole picture. Eur J Cogn Psychol 2006; 18: 277-320.
[7] Gopnik A, Glymour C, Sobel D, Schulz L, Kushnir T, Danks, D. A theory of causal learning in children: causal maps and Bayes nets. Psychol Rev 2004; 111: 3-32.

Jose Perales
Faculty of Psychology, Department of Experimental Psychology, University of Granada, Granada
Spain
Tel: +34 958 249609
Fax: +34 958 246239
E-mail: jcesar@ugr.es

David R. Shanks
Division of Psychology and Language Sciences, University College London, Gower St., London WC1E 6BT
England
Tel: (020) 7679 7588
Fax: (020) 7436 4276
E-mail: d.shanks@ucl.ac.uk