Running Head: MUDDY WATERS

Welcome Clarity for Muddy Waters

Alexander Weiss and Timothy C. Bates

The University of Edinburgh

Abstract

In his article Lee advocates using directed acyclic graphs (**DAGs**) to test causal theory in differential psychology. We applaud Lee's efforts to introduce graph theory and causal hypotheses to differential psychology. We agree that these methods lead to better understanding of the mechanisms underlying behavior. We thus join in advocating the use of these methods that, while they require more creative effort, honestly confront and overcome the problems of assigning causality which plague differential psychology and public policy.

Welcome Clarity in Muddy Waters

Testing causal hypotheses against observational data is central to progress in differential psychology and the social sciences more generally. As noted by Lee, the lack of causal progress in the social sciences flows not from a lack of statistical tools, but rather from which tools are used and how. Lee's most important message is that, while social scientists are taught that "*correlation does not imply causation*", they are not taught what correlation **does** imply, namely an unresolved causal structure (Shipley, 2000). Lee lays out the requirements for causal research, namely how causal hypotheses must be expressed as DAGs or equivalent strucutres, and how theories thus expressed can then be objectively compared.

Since works such as Kerlinger's (1964) "*Foundations of Behavioral Research,*" students have been trained to "control" variables. This method has become embedded in the paradigm of psychological science, with the assumption that control leads to conclusions. In fact, however, it has long been recognized that inappropriate controls mischaracterize cause and effect (e.g. Meehl, 1992). In particular, apparently innocuous control of a variable influenced by traits under study can induce false associations between these traits of interest, which if interpreted as real can have harmful consequences (Bingham, Heywood, & White, 1991; Figueredo, Hetherington, & Sechrest, 1992). Beyond this, statistical control cannot, even in principle, test causal assumptions (Pearl, 2000). Knowledge about just how critical a limit to causal interpretations of most epidemiological models is far from wide-spread (Davey Smith, 2010), and hopefully Lee's article will stimulate this spread. The calculus of causal theory, in proving the consequences of lack of control and inappropriate statistical control, as well as the solutions to these problems, places causal theorizing on a firm

mathematical and logical footing. There simply can no longer be any place for theories not expressed, whether in words or figures, as DAGs making these effects explicit.

Confidence in the Progress of Theory

One important consequence of a logical framework for contrasting causal hypotheses is that the crisis of confidence that found its expression in the postmodern proposition that, as social constructions, all theories have equivalent value is set aside. While scientists still cannot know if they have found the truth simply by dint of applying DAGs, they can determine which of two competing models is closer too that truth. If the idea that a mechanism exists to objectively and iteratively select causal models which are not simply different, but which are more complete in an objective sense, this may be the biggest impact of all the changes that Lee's paper lays out.

Theory Generation

If statistical tools allow us to test causal hypotheses, they also highlight the requirement for researchers to generate theories. Importantly, when a model finds itself containing a correlation, as with rain and mud, one must "do(mud)" and measure the effects or lack thereof of this treatment on the likelihood of rain. We hope that this expression spreads widely, and that readers come to expect articles to be expressed in this fashion, compelling researchers to make clear the causal process they are predicting, be it do(school) or do(genetic polymorphism) or do(neuroticism). Lest other fields feel smug, such errors remain common in areas such as medicine that are more used by now to thinking causally. Variables merely associated with a disease may become proposed targets for intervention, sometimes to humorous effect (Cohen et al., 2000). A practical consideration when translating causal theory into practice is that

targeting outcomes rather than processes leading to outcomes. For instance, targeting school grades instead of effective teaching can have perverse consequences as factors not included in causal models come into play. These factors may include teachers invalidating tests as indicators of knowledge by teaching to the test or, worse, purchasing the answers to exams (Vasagar, 2011). Such problems should be expected in the study of any system that has multiple causes (see the First Law of Ecology in Hardin, 1963).

What are Common Factors?

We will finish with a discussion of Lee's statement regarding common factors. Lee "*allows a factor to play the role of cause or effect in graphs depicting the relations among high-level emergent entities*." The meaning ascribed to a common factor such as extraversion is basic to psychology and we appreciated the nuanced claim that a common factor may "*play the role*" of a cause. Lee is, here as elsewhere, taking causal reasoning seriously. Latent variable on SEM diagrams represent emergent properties of their indicators, but, like correlations, they also represent as-yet unresolved causal structures and must be explained by mechanisms. In psychology, explanations of constructs such as in-group favoritism are often given in what in Lee's terms would be long-form labels for the emergent property, or even as additional indicators. Lee's conceptualization thus refocuses our attention on the need to hypothesize causes for latent variables, not simply generate labels or additional indicators for them. Just as in physics the emergent properties of water are accounted for by non-wet, non-liquid causes, so too personality domains such as Extraversion must be accounted for by layers of mechanisms, from biology through typical characteristics to the objective biography of behavior (McCrae, 1996), rendering a set of objectively-specified and

parameterized mechanisms generating behavior on the fly as these systems are run in real environments (Lewis & Bates, 2011).

Causal hypothesizing and testing is, of course, no "*royal road*" to knowledge: while the means of testing causal mechanisms are established, causal hypotheses cannot be generated automatically. As Gödel (1962) demonstrated, steps towards completeness require creative mental effort which is not automatable. The power of modern differential psychology, then, depends on specifying theory in DAGs permitting causal inference, and we commend Lee's article to as broad a readership as possible.

References

Bingham, R. D., Heywood, J. S., & White, S. B. (1991). Evaluating schools and teachers based on student performance: Testing an alternative methodology. *Evaluation Review, 15*, 191-218.

Cohen, D., Spear, S., Scribner, R., Kissinger, P., Mason, K., & Wildgen, J. (2000). "Broken windows" and the risk of gonorrhea. *American Journal of Public Health, 90*, 230-236.

Davey Smith, G. (2010). Mendelian Randomization for Strengthening Causal Inference in Observational Studies: Application to Gene × Environment Interactions. *Perspectives on Psychological Science*, 527-545. doi: 10.1177/1745691610383505

Figueredo, A. J., Hetherington, J., & Sechrest, L. (1992). Water under the bridge: A response to Bingham, Heywood, and White. *Evaluation Review, 16*, 40-62.

Gödel, K. (1962). *On formally undecidable propositions of Principia mathematica and related systems* (First English ed.). Edinburgh: Oliver and Boyd.

Hardin, G. (1963). The cybernetics of competition: A biologist's view of society. *Perspectives in Biology and Medicine, 7*, 58-84.

Kerlinger, F. N. (1964). *Foundations of Behavioral Research*. New York, NY: Holt, Rinehart and Winston, Inc.

Lewis, G. J., & Bates, T. C. (2011). From Left to Right: How the personality system allows basic traits to influence politics via characteristic moral adaptations. *British Journal of Psychology, 102*, 546-558.

McCrae, R. R. (1996). Integrating the levels of personality. *Psychological Inquiry, 7*, 353-356.

Meehl, P. E. (1992). A funny thing happened to us on the way to the latent entities. In E. I. Megargee & C. D. Spielberger (Eds.), *Personality assessment in America: A retrospective on the occasion of the fiftieth anniversary of the Society for Personality Assessment* (pp. 113-125). Hillsdale, NJ: Lawrence Erlbaum Associates.

Pearl, J. (2000). *Causality : models, reasoning, and inference*. Cambridge, U.K. ; New York: Cambridge University Press.

Shipley, B. (2000). *Cause and correlation in biology : A user's guide to path analysis, structural equations, and causal inference*. Cambridge, UK: Cambridge University Press.

Vasagar, J. (2011, Thursday 8 December 2011). Exam boards scandal: the economic pressures that broke the system, *The Guardian*. Retrieved from http://www.guardian.co.uk/education/2011/dec/08/exam-boards-scandal-analysis

Author Note

Alexander Weiss and Timothy C. Bates, Department of Psychology, School of Philosophy, Psychology and Language Sciences, The University of Edinburgh, Edinburgh, United Kingdom.

Correspondence concerning this article should be addressed to Alexander Weiss or Timothy Bates, Department of Psychology, School of Philosophy, Psychology and Language Sciences, The University of Edinburgh, 7 George Square, Edinburgh EH8 9JZ, United Kingdom. Email: alex.weiss@ed.ac.uk