POST-PRINT AUTHOR COPY
ACCEPTED IN *MEASUREMENT: INTERDISCIPLINARY RESEARCH AND PERSPECTIVES* ON 16 AUGUST 2016

PRIDE AND PREJUDICE, AND CAUSAL INDICATORS

Nick Lee[1]
Professor of Marketing
Marketing Group
Warwick Business School
Warwick University
Coventry CV4 7AL UK
+44 24 765 28674
nick.lee@wbs.ac.uk

Laura Chamberlain
Lecturer in Marketing
Marketing and Strategy Group
Aston Business School
Aston University
Birmingham B4 7ET UK
+44 121 204 3155
l.m.chamberlain1@aston.ac.uk

[1]Corresponding author.

PRIDE AND PREJUDICE, AND CAUSAL INDICATORS

Aguirre-Urreta, Ronkko, and Marakas' (2016) paper in *Measurement: Interdisciplinary Research and Perspectives* (hereafter referred to as ARM2016) is an important and timely piece of scholarship, in that it provides strong analytic support to the growing theoretical literature that questions the underlying ideas behind causal and formative indicators (e.g. Cadogan and Lee, 2012; Edwards, 2011, Hardin, et al., 2011; Howell, Brievik, and Wilcox, 2007; Lee, Cadogan, and Chamberlain, 2014; 2013; Rhemtulla, Bork, and Borsboom, 2015). Such literature provides in our view compelling reasoning to avoid, or at best be extremely cautious in using, formative / causal indicators. However, the theoretical arguments presented in such work seem to have had little impact on either the common use of causal / formative indicators in practice, nor the continuing proliferation of methodological articles defending their use (e.g. Bollen, 2007, 2011; Bollen and Bauldry, 2011; Bollen and Diamantopoulos, 2015; Diamantopoulos, 2011; Diamantopolous, Riefler, and Roth, 2008). It should be no surprise that we hope that the approach used in ARM2016 proves more convincing evidence to scholars that there are significant dangers in applying the causal / formative approach to measurement.

In this commentary, we hope to supplement and add clarity to a small number of areas of ARM2016. In doing so, we hope both to add support to the main conclusions of ARM2016, but also to open up the potential for causal / formative indicators to provide some useful function in future work, rather than the rather confused and contradictory place they seem to occupy at present. Specifically, we explain that – while we support ARM2016 strongly – there really should have been no need for such a demonstration, because basic understanding of the principles of measurement leads to exactly the same conclusions. In doing so, we first explain what ‘measurement’ actually means, and demonstrate where literature on formative / causal indicators makes important missteps, leading to erroneous conclusions. We are hardly the first to point this out (e.g. Borsboom, 2005), yet such lessons continue to go unheeded. Second, we diverge from ARM2016 in recommending a distinct nomenclature that distinguishes between formative and causal indicators, which follows from our earlier work (e.g. Lee, 2010; Lee, Cadogan, and Chamberlain, 2013), and again remains generally unheeded in current literature. Finally, we briefly suggest how separating formative from causal indicators allows each to have their distinct uses in empirical research, even though neither are measurement models.

Just what is ‘measurement’?

There is much confusion over the meaning of the term ‘measurement’, and space limitations preclude us from doing more than touching on the issues herein. At its most basic level, measurement can be defined as “the assignment of numbers…to entities and events to represent their properties and relations” (Savage and Ehrlich, 1992, p. 3). Similarly, Hand (2004, p. 12, emphasis in original) defines measurements as “mappings from objects in the universe being studied to a numerical representation called a *variable*”. However, such formal definitions are at best ontologically agnostic, and at worst almost meaninglessly broad, and leave many important questions unaddressed, the most serious of which concerns exactly *what* can be measured. In addressing this question we follow the logic compellingly expressed by Guildford (1954, p. 3) in saying “whatever exists in some amount can be measured”. Of course, this implies the opposite: that which does not exist cannot be measured. The question is: what exactly does it mean for something to exist? In answering this question, we draw heavily from Borsboom (2005, see also Markus and Borsboom, 2013). In essence, Borsboom (2005) shows that the only plausible notion of measurement is a realist one, which assumes the existence of real attributes that cause variation in measurement devices. Markus and Borsboom (2013) term this a *causal theory of measurement* (CTM).

Accepting such a framework of measurement unavoidably places significant constraints on formal measurement definitions such as those given above. In particular, it places an ontological restriction on what can be measured – in that if something does not exist as an attribute, it cannot be measured. Questions regarding existence of attributes are challenging, and our prior work attempted to provide a framework for such thinking, to varying degrees of success (e.g. Cadogan, Lee, and Chamberlain, 2013; Lee, 2010). Specifically, it seems to us that attributes that are somehow 'created' by researchers, as combinations or composites of other attributes (for example, Socioeconomic Status, or SES), are problematic in this regard. Perhaps even more importantly, the clear causal asymmetry in CTM mitigates against causal indicators being considered measures. This chimes with basic logic and intuition, in that it simply seems illogical to consider changes in the value of a measure to cause changes in the value of an attribute. Causal indicators may have many uses, but measurement is not one of them (Rhemtulla, Bork, and Borsboom, 2015).

The Importance of Distinguishing Causal from Formative

ARM2016 settles on the use of the term 'causal' to refer to what different authors have on various occasions also termed 'formative'. Bollen and Bauldry (2011) correctly pointed out the inconsistency of terminology, as did we (e.g. Lee, 2010; Lee, Cadogan, and Chamberlain, 2013). However, it seems that the key implications of distinguishing 'causal' from 'formative' in this literature remain unappreciated. First, a ‘causal’ indicator must be exactly that, causal. In other words, it must have a causal relationship to the attribute represented by a latent variable. Few scholars who write about formative / causal indicators address exactly what they mean when they use the term ‘cause’. However, the term is not neutral, and implies a number of important features, which are somewhat intertwined, and which we also attempted to unpick in prior work (e.g. Lee, 2010; Lee, Cadogan, and Chamberlain, 2014). In particular, the notion of ‘cause’ is complex, and definitions differ across various philosophical points of view. However, it seems there are at least two features, relevant to our present discussion, that seem reasonably consistent. First, in order to cause something, a thing has to be real. This seems almost so obvious as to be trivial, but it does raise issues regarding our commitments to reality of our unobservables, and how we justify these claims to reality, touched upon above. Second, as Markus and Borsboom (2013, p. 92) note, “the notion (or notions) of causation that underlies typical behavioral science research assumes that the cause is an entity distinct from the effect”. This follows the Humean (2008) tradition, which in turn follows Aristotle’s (1984) concept of ‘efficient causation’ (see Lee, Cadogan, and Chamberlain, 2014). In this case, a ‘causal’ indicator is one which has a causal effect on another attribute. This seems to match well with early discussion of causal indicators in the literature, such as Blalock’s (1963) ‘exposure to discrimination’ variable, which has ‘race’ as a causal indicator. While this model has flaws touched on later, it does at least maintain the separation between cause and effect.

Like ARM2016, we consider the confusion to have begun to creep into the literature with Bollen and Lennox (1991), who (clearly shown by the quotes in ARM2016) conflated the terms ‘cause’ and ‘form’, and introduced the idea that causal indicators somehow compose the latent variable, and thus assign its meaning. As such, Bollen and Lennox (1991) suggest that omitting a causal indicator changes the meaning of the latent variable. However, if the cause and effect are separate, this makes no sense at all. In Blalock's (1963) model for example, 'race' does not somehow define the meaning of 'exposure to discrimination', it is a cause of it. This situation is completely different from what should be called 'formative', where the indicators are actually the defining components of a formed composite – as is clearly the case with archetypal examples such as SES. In such a case, SES is nothing more than the combined scores on income, education, and occupation. Change the components, and you change what SES is. They do not *cause* SES, they *are* SES.

Of course, as we point out in Lee, Cadogan, and Chamberlain (2014), Aristotle's (1984) concept of 'material causation', suggests that a cause can be considered something of which the final consequence is made. Such a definition of cause could indeed encompass the formative composite model. However, while this is termed 'causation' by Aristotle, it is a completely different concept to the Humean notion of cause, and thus could not be used to refer to the notion of cause encoded within Blalock's 'causal indicator' model, and those which draw from it. As such, there remains a problematic conflation of ideas, even though the term 'cause' could refer to either, depending on how one defines it.

The solution to this is clear. The term *formative* should refer to composite based models, where the indicators are somehow combined to create a composite score, which is not representative of a real attribute that can exist distinct from the indicators. The term *causal* should refer to models where the indicators and the attribute they are supposed to relate to are distinct and independent of each other. Causal indicators do not therefore determine the meaning of the attribute, any more than an antecedent cause determines the meaning of its consequence, in a Humean causal framework at least. A thing cannot cause itself from this viewpoint.

Conclusions

In light of the delineation of the meaning of causal and formative above, a number of conclusions can be drawn regarding the findings of AMR2016. In terms of true 'causal' indicators, which are considered distinct causes of a separate real attribute, it becomes completely obvious that a census of indicators is not required. Indeed, why should it? The causal indicators cannot be considered measures, since the causal flow runs in the wrong direction (Markus and Borsboom, 2013). As such, they are simply antecedent causes, and cannot serve to define the meaning of the consequence. With this in mind, what can we consider the meaning of the consequence to be? Here, we advocate a return to Blalock's (e.g. 1963) nomenclature of *unmeasured variable,* which sums up exactly the problem we face, and the one Blalock was searching for a solution to. A model which sets up a latent variable *X*, with causal indicators and some downstream consequences, has no measure of *X*. It should be clear to all but those blinded to alternatives that the modelled *X* therefore has no stability across models and data sets. Such models may be acceptable – even necessary – in many situations of secondary data (e.g. trying to model 'exposure to discrimination' with only census demographic data), such as those common to sociology, where the model first emerged. However, setting out to collect primary data for such a model, if alternatives are available, is akin to tying one's own hands. The alternative of course, if one can define a sensible case that *X* is a real unobservable, is to measure it, by which we mean first set up a plausible measurement theory, involving causal influence from *X* to some measure (Markus and Borsboom, 2013). Note the causal asymmetry – cause must flow *from* the attribute *to* the measure for this to make sense. The alternative is simply not a plausible theory of measurement (Borsboom, 2005).

What then of formative indicators? Neither measure nor (Humean) cause, how should we think of them? Here, it seems a census of indicators is by definition essential. Indeed, since the choice of indicators creates the meaning of the composite, it seems almost impossible to *not* have a census. The conceptual trap is mistakenly thinking of the composite as a real thing. It is not, it is simply the combination of the components defined. The analytical trap is allowing data to estimate the relationships and combination, when in fact these should be part of the definition (Lee and Cadogan, 2013; Lee, Cadogan, and Chamberlain, 2013). Space restrictions preclude us from elaboration, and we refer interested readers to the cited works.

The ongoing controversy over formative and causal indicators is disturbing, because it seems to us that it contains an over-reliance on empirically-based argument, and an under-appreciation of the philosophical consequences of various theories. Put simply, no amount of identified models can make up for a theory which makes no sense. More specifically, we agree with Borsboom (2005, p. 9), who states that "the proper ground for the evaluation of conceptual frameworks like measurement models lies not in their empirical implications, but in their philosophical consequences. We may find such consequences plausible, or they may strike us as absurd…[in evaluating measurement models] plausibility and absurdity play roles analogous to the roles of truth and falsity in empirical research". It seems plain to us that the philosophical consequences of the theory of causal indicators as defined in the relevant literature, and summarized in ARM2016, are completely implausible. ARM2016's results are welcome support for these conceptual conclusions.

REFERENCES

Aguirre-Urreta, M. I, Ronkko, M., & Marakas G. M. (this issue). Omission of causal indicators: Consequences and implications for measurement. *Measurement: Interdisciplinary Research and Persepctives XX(X),* X-XX/

Aristotle (1984) The Complete Works: The Revised Oxford Translation, vol. 1, Barnes, J. (ed.). Princeton University Press: NJ

Blalock, H. M. (1963). Making causal inferences for unmeasured variables from correlations among indicators. American Journal of Sociology, 69(1), 53–62.

Bollen, K. A. (2007). Interpretational confounding is due to misspecification, not to type of indicator: Comment on Howell, Breivik, and Wilcox (2007). *Psychological Methods, 12*(2), 219–228.

Bollen, K. A. (2011). Evaluating effect, composite, and causal indicators in structural equation models. *MIS Quarterly, 35*(2), 359–372.

Bollen, K. A., & Bauldry, S. (2011). Three Cs in measurement models: Causal indicators, composite indicators, and covariates. *Psychological Methods, 16*(3), 265–284.

Bollen, K. A., & Diamantopoulos, A. (2015, September 21). In Defense of Causal–Formative Indicators: A Minority Report. *Psychological Methods*. Advance online publication. http://dx.doi.org/10.1037/met0000056

Bollen, K. A., & Lennox, R. (1991). Conventional wisdom on measurement: A structural equation perspective. *Psychological Bulletin, 110*(2), 305–314.

Borsboom, D. (2005). *Measuring the mind*. Cambridge, UK: Cambridge University Press.

Cadogan, J., & Lee, N. (2012) Improper Use of Formative Endogenous Variables. *Journal of Business Research*. Vol 66: 233-241.

Cadogan, J., Lee, N., & Chamberlain, L. (2013) Formative Variables are Unreal Variables: Why the Formative MIMIC Model is Invalid. *Academy of Marketing Science Review*. Vol 3 (1): 38-49.

Diamantopoulos, A. (2011). Incorporating formative measures into covariance-based structural equation models. *MIS Quarterly, 35*(2), 335–358.

Diamantopoulos, A., Riefler, P., & Roth, K. P. (2008). Advancing formative measurement models. *Journal of Business Research*, *61*(12), 1203–1218.

Edwards, J. R. (2011). The fallacy of formative measurement. *Organizational Research Methods*, *14*(2), 370-388.

Guildford, J. P. (1954) *Psychometric methods* (rev. ed.). New York: McGraw-Hill.

Hand, D. J. (2004). *Measurement: Theory and Practice*. London, UK: Arnold.

Hardin, A. M., Chang, J. C.-J., Fuller, M. A., & Torkzadeh, G. (2011). Formative measurement and academic research: In search of measurement theory. *Educational and Psychological Measurement, 71*(2), 281–305.

Howell, R. D., Breivik, E., & Wilcox, J. B. (2007). Reconsidering formative measurement. *Psychological Methods*, *12*(2), 205–218.

Hume, D. (2008) An Enquiry Concerning Human Understanding, Oxford University Press: Oxford, England.

Lee, N. (2010) On a Distinction Between Causal and Formative Variable Models: Theoretical Issues and Practical Implications for Organizational Researchers. *Aston University Working Paper Series* 2010 ISBN 978-1-85449-765-9

Lee, N., Cadogan, J. W. & Chamberlain, L. (2014). Material and Efficient Cause Interpretations of the Formative Model: Resolving Misunderstandings and Clarifying Conceptual Language. *Academy of Marketing Science Review*. Vol. 4: DOI 10.1007/s13162-013-0058-5

Lee, N., Cadogan, J. W., & Chamberlain, L. (2013) The MIMIC Model and Formative Variables: Problems and Solutions. *Academy of Marketing Science Review*. Vol 3 (1): 3-17.

Markus, K. A. & Borsboom, D. (2013). *Frontiers of test validity theory*. New York: Routledge.

Rhemtulla, M., Bork, R. van, & Borsboom, D. (2015). Calling models with causal indicators “measurement models” implies more than they can deliver. *Measurement: Interdisciplinary Research and Perspectives, 13*(1), 59–62. http://doi.org/10.1080/15366367.2015.1016343

Savage, C. W. & Ehrlich, P. (1992) A brief introduction to measurement theory and to the essays. In Savage, C. W. & Ehrlich, P. (Eds), *Philosophical and Foundational Issues in Measurement Theory*. (pp. 1-15). Hillsdale, NJ: Lawrence Erlbaum Associates.