Journal of Business Research 67 (2014) 2821–2826

Contents lists available at ScienceDirect

Journal of Business Research

Formative vs. reflective measures: Facets of variation

Adam Finn [a,*], Luming Wang [b]

[a] Alberta School of Business, University of Alberta, Edmonton, AB, Canada T6G 2R6
[b] Asper School of Business, University of Manitoba, Winnipeg, MB, Canada R3T 5V4

ARTICLE INFO

Article history:
Received 1 September 2010
Received in revised form 1 November 2010
Accepted 1 July 2012
Available online 5 September 2012

Keywords:
Multi-facet data
Formative constructs
Reflective constructs
Covariance components

ABSTRACT

Research that treats observed measures of a business construct as a function of a latent true score plus random error is making two strong assumptions. The first, the assumed direction of causality, has generated the burgeoning formative measurement literature to which Cadogan and Lee (2013) contribute. The second is overlooked. Classical test theory assumes there is only one legitimate source of variance. Academics validating measures of business constructs invariably assume that this source is their respondents, who can be managers, employees, or customers, depending on the context (i.e. business discipline). The invoked causality accounts for covariance at the respondent level, ignoring whether it also applies to other sources of variance—such facets as brands, companies, departments, locations, service providers or work groups. Researchers need to be clear about a business construct's conceptual domain, about the sources of variance that are focal to the theoretical relationships being investigated, and whether the construct's relationship with its indicators is formative or reflective for each facet.

1. Introduction

Valid measures are a necessary condition for business theory development and testing (Peter, 1981, p. 133). Validity assessment of business constructs began to be taken seriously in the classical test theory era, making the assumption classical test theory makes about the relationship between a construct and its indicators, namely that observed scores for measures of a construct are a function of a latent true score, plus random error. Variation in the underlying latent construct causes variation in the observed measures (Bagozzi, 1980, 1984; Bollen, 1989; Nunnally, 1978).

However, classical test theory makes two assumptions that can both compromise an assessment of construct validity. The first, the assumed direction of causality, has generated the burgeoning literature to which Cadogan & Lee (2013) contribute. A reflective causal flow from the latent construct to its measures may not always be appropriate. It sometimes makes more sense to view causality as flowing from the measures to the construct (Blalock, 1964; Fornell & Bookstein, 1982). This distinction between formative and reflective measurement can substantially affect theory testing. There are two sources of bias in studies that mis-specify formative constructs as reflective (Diamantopoulos, Riefler, & Roth, 2008). First, an incorrect reflective specification for a set of formative indicators of an exogenous construct leads to over estimation of the structural parameters for its effects. If it is for an endogenous construct, the structural parameters are under estimated. Second, applying reflective scale development principles to a formative construct can delete valid items and thereby compromise the content validity of the construct.

The second assumption has received far less attention. Classical test theory assumes that there is only one legitimate source of variance. Academics, who are validating measures of business constructs, invariably assume this source is their respondents, who can be managers, employees, or customers, depending on the context (i.e. business discipline). Whether indicators of constructs are assumed to be formative or reflective, to validate constructs and test their theories, business academics invoke the causality to account for covariance at the respondent level. This means that they are ignoring the existence of other systematic sources of variance, such as brands, companies, departments, locations, service providers or work groups. However, if other sources of construct variance are of focal interest, construct validity must be established for each source of variance.

The objective of this paper is to establish that rigorous assessment of construct validity has to move beyond the formative versus reflective question and address both of these assumptions. Researchers need to be clear about a business construct's conceptual domain, about the sources of variance that are the facet of differentiation involved in the theoretical relationships being investigated, and whether the constructs relationship with its indicators is formative or reflective for each facet. Only the latter choice between formative and reflective models, which substantially affects estimation procedures, has received adequate attention in the literature.

* Corresponding author. Tel.: + 1 780 492 5369; fax: +1 780 492 3325.
E-mail addresses: adam.finn@ualberta.ca (A. Finn), luming.wang@ad.umanitoba.ca (L. Wang).

0148-2963/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.jbusres.2012.08.001

2. Formative or reflective measures

Three approaches have been taken to this issue in the business literature.

2.1. Philosophy of science perspective

The simplest approach to the formative or reflective question treats it as one of philosophy of science. As summarized by Borsboom, Mellenbergh, and Van Heerden (2003, p. 209), "The realist interpretation of a latent variable implies a reflective model, whereas constructivist, operationalist, or instrumentalist interpretations are more compatible with a formative model." From a realist perspective all constructs are reflective, and the strength of attachment to this realist position is evident in two attempts to preserve the reflective model in the face of the formative challenge.

First, the reverse causal flow that is central to formative measurement can be achieved by modeling each formative measure as an antecedent construct, underlying its own reflective measure. This respecification can avoid what Edwards (2010) calls the fallacy of formative measurement. Edwards and Bagozzi (2000) modify the formative measurement model, so that formative measures x1, x2, and x3 of construct Z are replaced by corresponding constructs X1, X2, and X3 as antecedents of Z. Thus x1, x2, and x3 become reflective measures of these three constructs, each with its own uniqueness and random error, d1, d2, and d3. Latent constructs X1, X2, and X3 rather than the observed measures x1, x2, and x3 are the causes of Z, thereby avoiding the position that measures cause constructs.

Second, Treiblmaier, Bentler, and Mair (2011) preserve the realist position by arguing that a "closely matched common factor" can represent any formative latent. Their two-step procedure uses the measures of the formative construct to create two or more composites, and then models these composites as reflective indicators of a common factor that closely approximates the formative latent and can be incorporated into any larger SEM model. They claim that this is "a compromise in the formative/reflective controversy and conceptually speaking, a formative construct can be scientifically meaningful, but practically and operationally, a thoughtfully developed reflective measurement approach is the most appropriate way to implement it." (p. 19).

However, this view that a construct must be conceptually reflective and empirically unidimensional is often inconsistent with the way constructs are defined in the field. Whether a construct is viewed as unidimensional or can be multidimensional may depend on the level of abstraction used to define the construct (Jarvis, MacKenzie, & Podsakoff, 2003). For example, job satisfaction is frequently defined as being composed of several different aspects, including satisfaction with one's pay, coworkers, supervisor, opportunities for advancement, and so forth. Although one can look at each aspect as a separate construct, at a more abstract level, they are all integral parts of a person's job satisfaction. Indeed, we think that this kind of abstract multidimensional construct definition is quite common in the business literature. Thus the formative versus reflective measures issue does not only exist between measures and first-order latent constructs. If constructs are conceptualized at a more abstract level, they can have formative and/or reflective relationships at each level (Jarvis et al., 2003). For example, a single multidimensional construct might have one type of measurement model relating its measures to its first-order components and a different measurement model relating its components to the underlying second-order factor.

2.2. A priori assessment of causal directionality

An increasingly popular position holds that particular constructs are inherently formative or inherently reflective. Rossiter (2002) takes this position, presented first for marketing and now all the social sciences (Rossiter, 2011). His C-OAR-SE scale development paradigm classifies attributes of objects (i.e., constructs) as concrete (singular) or abstract, in which case they have multiple first or second-order components. Abstract attributes can be formed (responses to the components cause the attribute), or eliciting (the attribute causes the responses to the components). Thus, C-OAR-SE asserts that constructs are a priori identifiable as formative or reflective by expert judgment. Service quality is his prime example of a formed second order attribute (formative), whereas the trait of need for cognition (Cacioppo & Petty, 1982) is an example of an eliciting attribute (reflective).

Jarvis et al. (2003, Table 1) present a set of four decision rules for determining whether a construct should be modeled as having formative or reflective indicators. These are (1) the direction of causality from construct to measure implied by the conceptual definition, (2) interchangeability of the indicators/items, (3) covariation among the indicators, and (4) nomological net of the construct indicators. They report that their rules could reliably classify 98 percent of the 1192 constructs modeled using confirmatory factor analysis in the four leading marketing journals between 1977 and 2000. As a result, Jarvis et al. (2003, Table 3) list 31 examples of marketing constructs their criteria identify as first or second order formative. Similarly, MacKenzie, Podsakoff, and Jarvis (2005) identify job performance, transformational leadership, and procedural justice as formative rather than the previously assumed reflective constructs from the organizational and behavioral research literature. Diamantopoulos et al. (2008, Table 1) identify additional examples, including some from the management and information technology literatures.

2.3. Empirical assessment of causal directionality

A causal structure proposed for the measures of a particular construct is only a hypothesis, and it remains a hypothesis until the structure is supported for the data provided by respondents. When a construct is hypothesized to be formative rather than reflective, it has implications for the covariance between the construct's measures.

The most commonly used reflective specification is the congeneric factor model, where all measures are a linear function of one underlying construct plus random error. Thus, reflective indicators of a latent construct must be unidimensional, and any two measures that are equally reliable are interchangeable. Moreover, the construct remains unchanged when the number of indicators included in the measurement model changes (Bollen & Lennox, 1991).

In contrast, the set of formative measures for a construct should not be unidimensional. The formative measurement model does not even assume or require the measures to be correlated. Indeed, formative indicators can be completely uncorrelated. This is almost evidently the case where a formative construct is formed by a set of mutually exclusive types of behavior. For the formative model the consequences of omitting a measure are quite serious. Since the error term captures aspects of the construct's domain that the set of indicators neglect, Diamantopolous (2006) proposes using the variance of the error term as a measure of construct validity. The lower the error variance, the more valid the construct (see also Williams, Edwards, & Vandenberg, 2003).

If a researcher specifies a measurement model as reflective or formative to match a particular conceptualization, "this does not guarantee that this is the correct causal structure nor does it override the possibility that empirical characteristics of the data might imply a different block structure" (Grace & Bollen, 2008, p. 202). Bollen and Ting (2000) provide confirmatory tetrad analysis tests to distinguish empirically between formative and reflective and even mixed measurement structure models. If these tests show that the proposed reflective measures of a construct are not unidimensional, a reflective first-order specification can only be preserved by postulating a more