Journal of Business Research 67 (2014) 403–406

Contents lists available at ScienceDirect

Journal of Business Research

Construct measurement in management research: The importance of match between levels of theory and measurement☆

Bo Bernhard Nielsen

Copenhagen Business School, Department of Strategic Management and Globalization, Kilen 14, 2nd floor, 2000 Frederiksberg, Denmark
University of Technology Sydney, School of Marketing, Australia

ARTICLE INFO

Article history:
Received 1 October 2012
Received in revised form 1 December 2012
Accepted 1 December 2012
Available online 5 January 2013

Keywords:
Construct measurement
Management research
JV relatedness
Multilevel research
SIC code
Alliance termination

ABSTRACT

Far too often do management scholars resort to crude and often inappropriate measures of fundamental constructs in their research; an approach which calls in question the interpretation and validity of their findings. Scholars often legitimize poor choices in measurement with a lack of availability of better measures and/or that they are simply following existing research in adopting previously published measures without critically assessing the validity, appropriateness, and applicability of such measures in terms of the focal study. Motivated by a recent dialog in *Journal of Business Research*, this research note raises important questions about the use of proxies in management research and argues for greater care in operationalizing constructs with particular attention to matching levels of theory and measurement.

1. Introduction

Scholars seeking to investigate important research questions and develop new theories often face challenges with respect to the measurement of the constructs that are pivotal to their frameworks. With expectation to test empirically novel theoretical frameworks, researchers employ a variety of data collection approaches that may compromise construct validity. One frequently used approach involves approximation, which enables the researcher to investigate important research questions at the expense of precise measures. It is often argued that approximation is reasonable given the data constraints and that issues of lack of precise measures should not prevent us from investigating important research questions (e.g., Cui & Kumar, 2012b). In this commentary, I argue that data limitations should prevent us from analyzing research questions (testing theory) for which we do not have appropriate data or measures; rather this limitation than running the risk of providing "evidence" of key relationships that may not hold up to closer empirical scrutiny.

Cui and Kumar (2012a) follow previous research and use a 2-digit SIC code as proxy for JV relatedness in order to investigate to what extent alliance termination differs among related and unrelated joint ventures. Essentially, their approach is a contingency approach which specifies a set of factors at different levels that may influence the termination rate of JVs. Chief among these variables is JV relatedness and they advance the argument that the impact of various factors (at multiple theoretical levels) on termination vary between related and unrelated JVs. The underlying theoretical logic is that the evolution (and thus probability of termination) of related and unrelated JVs is likely to differ because of important differences in the motives of the firms to form such JVs.

As a commentary (Nielsen, 2012) to Cui and Kumar (2012a) notes, however, industry-level SIC codes say very little about the underlying motives behind alliance formation, nor do they approximate well firm-level resources and capabilities. Yet, the authors argue in their rejoinder (Cui & Kumar, 2012b) to the commentary that "SIC codes provide a good proxy for firm resources and capabilities because… businesses in the same industry tend to have similar assets, operations, as well as similar intangible resources…." However, using a 2-digit industry level SIC code to proxy firm intangible resources and capabilities is highly problematic as such industry overlap says absolutely nothing about firm level operations, efficiencies, procedures, knowledge, experience, let alone assets or motivation behind JV formation. Moreover, within industries, there is likely to be variance between firms in terms of resources, capabilities and strategic motives for JV formation. This example, which is by no means unique in management research (see Boyd, Gove, & Hitt, 2005), raises a number of fundamental questions about measurement in management research, the importance of matching the levels of theory and measurement, and the utility of previously established measures.

☆ The author thanks Jacob Lyngsie, Copenhagen Business School, and Stewart Miller, University of Texas at San Antonio for their valuable comments and suggestions. Any and all errors are the sole responsibility of the author.

E-mail address: bn.smg@cbs.dk.

0148-2963/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jbusres.2012.12.020

2. Construct measurement in management research

Unobserved constructs (such as capabilities or managerial processes) lie at the core of management phenomena, which puts a premium on the researcher's ability to develop sound strategies for operationalizing and testing constructs that are unobservable (Godfrey & Hill, 1995). More fundamentally, I suggest that management scholars should strive harder to ensure construct validity of their measures and seek to acquire direct measures rather than relying on poor proxies when examining complex latent constructs. Proxies for variables are often seemingly selected without concern for their reliability or validity and future researchers must employ measurement strategies, such as cross-validation, in order to establish that a model's constructs are appropriately measured.

The use of dichotomous variables is an extreme example of how simple measures are used to proxy complex relationships in management research. For instance, prior experience with the partner is often coded as a dummy based on any type of previous relationships between parties and theorized to reflect relational quality between partners. Yet, such binary variables are imperfect measures of the nature (type of transaction) and quality (positive or negative) of interaction between JV partners and thus fail to capture the underlying theoretical construct. Thus, using dichotomous variables as crude representations of more complex relations allow for little variance which likely biases results.

When constructs are complex and multi-faceted, multi-dimensional construct measurement is typically warranted in order to improve validity and fit with theory. A single measure provides no structure on which to evaluate construct validity. That is not to say that single measures cannot be valid, however, one must establish such validity by showing that results obtained from a single measures would be the same if other measures within the domain were used, or show that indeed a single measure captures the underlying construct better than a host of ambiguous items (for an example of trust in alliances, see Nielsen & Nielsen, 2009). Notwithstanding, combining several measures provides greater construct validity and, importantly, improves generalizability relative to using a single measure.

Construct measurement is of pivotal importance if we are to advance management research and scholars must (at a minimum) demonstrate that (1) measures employed plausibly capture the theoretical constructs and (2) theoretical and empirical levels of analysis for the proposed construct match (Lawrence, 1997).

3. Levels of theory and measurement

Researchers need to be much more watchful of potential mismatches between theory and measurement when operationalizing difficult-to-measure management phenomena; levels-of-analysis ambiguity may seriously misrepresent the relationships a researcher would have found if data had been collected and analyzed at the same level as the theory (Klein, Dansereau, & Hall, 1994; Lawrence, 1997). Level of theory refers to the focal unit or target (e.g., firm or alliance) that a researcher aims to explain; "it is the level to which generalizations are made" (Rousseau, 1985: 4). The focal unit, in turn, determines the *appropriate* level associated with the key constructs of interests in a study (e.g., JV relatedness is at the JV level). Explicit attention must be paid to the level of theory (and measurement) because neglecting to do so increases the risks of falsely attributing effects at one level (e.g., industry level) to another (e.g., firm or JV level) and thus committing "fallacies of the wrong level" (Rousseau, 1985: 5). Related, level of measurement refers to the actual source of the data from which inferences are made; in other words the entities from which the data are drawn or to which the data is attached (e.g., firms, dyads and industries). The level of measurement should correspond to the level of constructs specified in the focal unit in order to increase the variability predicted by the theory.

For instance, if the theory specifies within-group homogeneity (e.g., firms within industries are similar in terms of motives for forming JVs), data collection should be conducted at the firm level in order to assess whether such homogeneity exists. However, theories like RBV assert that firms *differ* in terms of resources and capabilities, and that this heterogeneity determines variation in firm strategy and performance — even within the same industry. Such firm-level heterogeneity, in turn, is likely to influence the motivational intend underlying JV formation and, as a result, the probability of JV termination. In other words, measuring JV-firm relatedness at the industry level assumes no firm heterogeneity within the same 2-digit SIC code, yet such heterogeneity is clearly possible (indeed likely according to theory) as not all firms with similar 2-digit SIC codes are similar (homogenous) in terms of resources, capabilities or motivational intend. Testing such theories is best accomplished by (a) using measures that (like the theory) highlight the position of each individual firm relative to the JV and by (b) maximizing within-group variability (Nielsen, 2010).

Before higher-level (e.g., industry) data are used to measure lower-level (e.g., JV or firm) phenomena, psychometric evidence of the suitability of using the data in this way must be demonstrated. Different types of validity evidence exist depending on the nature of the constructs in question (see Chan, 1998; Chen, Mathieu, and Bliese, 2004). More generally, however, researchers need to be explicit about how data collected at one level of analysis is related to constructs at a lower/higher level of analysis (see also Klein & Kozlowski, 2000). Here it should be noted, that simply arguing that more appropriate measures were difficult to obtain and that this reflects "the constraints that a researcher typically faces when conducting research" (Cui & Kumar, 2012b) is a poor "*proxy*" for rigorous and valid research. If indeed theory (and data) warrants multilevel treatment, then researchers need to account for this in adequate ways both theoretically and empirically. Building multilevel theories and measuring and testing them accordingly is difficult but necessary if one wishes to answer challenging and interesting research questions. Conceptualizing multilevel yet building and testing a single-level model without paying due attention to the levels of theory and measurement in the study design may mislead the reader and lead to erroneous conclusions due to level-related confounds (Klein & Kozlowski, 2000; Nielsen, 2010).

4. The utility of existing measures

Moreover, as is often the practice, it can be dangerous to follow existing research in their operationalization of measures since legitimizing the use of a potentially poor measure by referencing its previous use in the literature does not ensure its appropriateness or indeed validity. Poor measurement oftentimes appears to be the result of relying on previously published work, regardless of the quality and purpose of that measurement approach, and may be responsible for the persistence and pervasive reliance on poor measures in a particular field (Boyd et al., 2005). To be sure, adopting existing measures can be very valuable; however, doing so without assessing the applicability and meaning of such measures may lead to poor match between theory and measurement. Meaningfulness depends upon context and thus researchers must take extra steps to ensure validity when utilizing existing measures in different contextual settings. Hence, although it is normally thought to lend credibility and legitimacy to a study that measures (and particularly proxies) have been established previously in the literature, we, as researchers, have to be more critical of the way we adopt and apply such measures in our work.

First of all, be careful in adopting existing measures at face value without evaluating to what extent the original study was within the same domain theoretically and empirically (e.g., levels of theory and measurement). For instance, while a number of scholars have used industry similarity measured by SIC codes to infer business relatedness, the fact remains that industry overlap (in terms of 2 or even