Australasian Marketing Journal 23 (2015) 246–258

Contents lists available at ScienceDirect

Australasian Marketing Journal

journal homepage: www.elsevier.com/locate/amj

Visualizing·matching·generalizing: Case identification hypotheses and case-level data analysis

Arch G. Woodside *

Department of Marketing, Carroll School of Management, Boston College, 140 Commonwealth Avenue, Chestnut Hill, MA 02467, USA

ARTICLE INFO

Article history:
Received 8 June 2015
Revised 20 June 2015
Accepted 10 July 2015
Available online 14 August 2015

Keywords:
Behavior
Case
Empiricism
Management
Theory
Variable

ABSTRACT

The traditional and still dominant logic among nearly all empirical positivist researchers in schools of management is to write symmetric (two-directional) variable hypotheses (SVH) even though the same researchers formulate their behavioral theories at the case (typology) identification level. The behavioral theory of the firm, theories of buyer behavior, and Miles and Snow's typology of organization's strategy configurations (e.g., "prospectors, analyzers, and defenders") are iconic examples of formulating theory at the case identification level. When testing such theories, most researchers automatically, unconsciously, switch from building theory of beliefs, attitudes, and behavior at the case identification level to empirically testing of two-directional relationships and additive net-effect influences of variables. Formulating theory focusing on creating case identification hypotheses (CIH) to describe, explain, and predict behavior and then empirically testing at SVH is a mismatch and results in shallow data analysis and frequently inaccurate contributions to theory. This paper describes the mismatch and resulting unattractive outcomes as well as the pervasive practice of examining only fit validity in empirical studies using symmetric tests. The paper reviews studies in the literature showing how matching both case-based theory and empirical positivist research of CIH is possible and produces findings that advance useful theory and critical thinking by executives and researchers.

CHINESE ABSTRACT

在各个管理学流派的几乎所有经验实证主义研究者当中，传统上以及当前依然占主要地位的逻辑是撰写对称（双向）变量假设（SVH），即便这些研究者都是在案例（类型学）识别层面用公式表示他们的行为理论。《企业行为理论》（Cyert & March，1963年）、《购买者行为理论》（Howard & Sheth，1969）以及迈尔斯和斯诺（1978年）的组织策略构形类型学（例如"探索者、分析者和防守者"）是在案例识别层面用公式表示理论的形象例子。在检验该等理论时，多数研究者都在不经意间自动地从在案例识别层面对信念、态度和行为理论的建设转向了对双向关系和各变量的附加净有效影响的经验主义检验。以设立案例识别假设（CIH）为焦点用公式表示理论从而对行为进行描述、说明和预测，然后以对称（双向）变量假设进行经验主义检验，这是一个错误的配对，导致数据分析肤浅，以及对理论提供的[信息]经常不准确。本论文描述了这一错误的配对及其所导致的没有吸引力的结果以及在经验主义研究中使用对称性试验只对合适的有效性进行考察的普遍做法。基于案例的理论与对案例识别假设进行的经验实证主义研究有多大配对可能性，本论文回顾了对之进行阐释的文献的研究并提供了以高管和研究人员推动有用理论和批判性思维向前发展的研究结果。

1. Introduction

The present paper describes the mismatch between case-based identification theory creation and symmetric variable (two directional) hypotheses (SVH) and describes workable solutions to the mismatch. The solutions include case-based group-level analysis within symmetric testing (Bass et al., 1968), resorting to simple algorithms to replace relying on symmetric tests (McClelland, 1998), and using asymmetric Boolean-algebra based indexes instead of symmetric tests (Ragin, 2008). (Asymmetric tests are one-directional tests to indicate whether or not high scores in a complex statement of X associate consistently with high scores in Y, an outcome condition. A following section explains asymmetric tests in some detail.) This paper builds from a foundational premise, "Scientists' tools are not neutral" (Gigerenzer, 1991). Thus, if you test SVH via

The author is grateful for the comments by anonymous reviewers, Barry Babin (Louisiana Technical University), and Carol M. Megehee (Coastal Carolina University), on prior drafts of this article.
* Tel.: 1 617 552 3069; fax: 1 617 552 6677.
E-mail address: arch.woodside@bc.edu.

http://dx.doi.org/10.1016/j.ausmj.2015.07.002
1441-3582/

symmetric tests, you are not testing case-level causal conditions of complex wholes; you are testing a theory of net-effects directional hypotheses. SVH theory and tests include severe limitations beyond the mismatch between case-based theory and SVH. This paper reviews these severe limitations (cf. Fiss, 2007).

Most empirical research in behavioral science and business research journals present a mismatch theory and data analysis. Most researchers develop theory from the perspective of the individual firm or consumer but formulate their hypotheses from the perspective of the net effects influence of individual variables on a dependent variable. The researchers' shift in thinking is seemingly subtle and likely an unconscious one. Their prior training in data analyses focuses on how to do symmetric tests such as analysis of variance, multiple regression analysis (MRA), and structural equation modeling (SEM); they apply this training to extract information from their data without recognizing the transformation in their focus from case (e.g., firm, consumer, organization, nation) identification into a symmetrical variable hypotheses (SVH) perspective. Symmetric tests of variable hypotheses examine two-way directional hypotheses such as high versus low X (independent) scores associated with high versus low Y scores. For example, Hofstede's (2001) cultural value theory proposes that each nation is a complex whole of a combination of distinct cultural values (e.g., collectivism/individualism, masculinity, uncertainty avoidance, and power distance); a vast number of studies examine Hofstede's theory in many different contexts but almost all of these studies examine the net effects of each cultural value using symmetric tests. Almost none of these studies examine the influence of culture values from the perspective of culture as complex wholes (i.e., recipes), including the studies by Hofstede and colleagues; for exceptions, see Hsu et al. (2013) and Woodside, Hsu, and Marshall (2011).

The majority of studies in scholarly behavioral science and business-related journals on many different topics exhibit this theory–analysis mismatch. Most of the resulting published articles report low levels of explained variance (R^2) in their dependent variables in the findings section and struggle in their discussion sections to show how net effects of individual variables are relevant for complex wholes of the firm, person, or organization. While most of the relevant literature fails to acknowledge this mismatch specifically – except for McClelland (1998) and Fiss (2007) – a few researchers (Armstrong, 2012; Bass et al., 1968; Gigerenzer and Brighton, 2009; Ordanini et al., 2014; Ragin, 2008) do describe problems with the still current dominant logic of reporting findings using symmetric tests; these authors offer helpful solutions to overcome these problems. A bit of headway is now occurring in doing what McClelland advocated in the 1990s – taking steps to overcome the limitations of using regression analysis (symmetric tests) and the mismatching of theory and data analysis. Possibly a tipping point is appearing in the literature due to the subsequent work of Meier and Donzé (2012), Fiss (2007, 2011), Fiss et al. (2013), Ordanini et al. (2014), Ragin (2008), Woodside, Hsu, and Marshall (2011), Woodside (2013), and the studies by researchers who are members of COMPASSS.ORG.

Following this introductory section, section 2 focuses directly on overcoming the limitations of symmetric tests (regression analysis) and indirectly on the mismatch between case identification hypotheses (CIH) and SVH; section 2 also describes one of the "illusions in regression analysis" (Armstrong, 2012). Section 3 describes McClelland's algorithm procedure for overcoming the mismatch without completely moving beyond symmetric testing. Section 4 describes moving completely beyond symmetric tests and SVH to the use of asymmetric theory construction and tests of CIH. Section 4 describes a set of data to practice SVH versus CIH. Section 5 briefly describes three examples of matching asymmetric case-based theory construction with asymmetric CIH testing. Section 6 includes a visual and general discussion of case-based theory construction and CIH testing versus the convention (and dominant logic) of box and arrow presentation of SVH representations of case-based theory. Section 7 concludes with a call for finance, management, and marketing scholars to recognize the current pervasive mismatch between case-based (typology) theories with the use of symmetric variable hypotheses testing – a mismatch that does not need to continue.

2. Symmetric testing of configural outcomes to overcome regression analysis limitations

Most empirically-based studies include symmetric (two-directional) variable hypotheses (SVH). The general form of these hypotheses includes the following expressions. (1) Increases (decreases) in X (independent variable) associates with increases (decreases) in Y (dependent variable). (2) Increases in W (a second independent variables) associates with increases (decreases) in Y. (3) The impact of changes in X on Y is greater than the impact of changes in W on Y. (4) An interaction effect (X by W) occurs whereby the joint impact of increases in W and X is greater than the separate impacts of the levels of W and X on the level of Y. Researchers use symmetric tests such as analysis of variance and regression analysis to confirm or reject these expressions for variables in a given data set. Thus, the hypotheses are written in the following formats (with ß representing a standardized partial regression coefficient of influence of change in Y due to a change in the independent variable):

(1) $Y = ß_1X$, with $|ß_1| > 0$; estimating the net effect of level of X on level of Y
(2) $Y = ß_2W$, with $|ß_2| > 0$; estimating the net effect of level of W on level of Y
(3) $Y = ß_1X + ß_2W$, where $|ß1| > |ß2|$; comparing the relative sizes of net effects
(4) $Y = ß_1X + ß_2W + ß_3(X{\cdot}W)$

with all three $|ß_i| \geq 0$; testing for an interaction effect beyond the net effects influence. Examining interaction effects is a step toward recognizing the need to consider recipes but only to a limited extent. Interaction effects often compete with main effects in interpretations of regression models; interaction effects become difficult to interpret especially among 3+ independent variables especially. The use of CHAID (chi-squared automatic interaction detection, see Kass, 1980; Magidson, 1994) takes the researcher an additional step toward recognizing that the levels of any one individual variable is usually insufficient in accurately predicting the level of an outcome condition; however, CHAID falls short in comparison to the capabilities of more recent software in identifying surprising algorithms that are accurate consistently in identifying cases having high scores in an outcome condition of interest. This present study describes this more recent software below.

In an article receiving a substantial number of citations (n = 200+) since its publication but widely ignored in practice, Bass et al. (1968) point to the severe limitations of using testing by symmetric metrics only. Their study, "Market Segmentation: Group versus Individual Behavior" points out, "The evidence is overwhelming that R^2 (explained variance) is low when individual household purchase rates are related to socioeconomic variables. The intuitive conclusion, perhaps, suggested by the evidence is that market segmentation based on socioeconomic measurement is infeasible. This is the conclusion of Twedt (1964), Frank (1967) and others" (Bass et al., 1968, p. 264). Bass et al. (1968) go on to quote a few highly-cited papers that include this inaccurate conclusion.