Formative versus Reflective Measurement Implications for Explaining Innovation in Marketing Partnerships

Tania Bucic, University of New South Wales
Siegfried P Gudergan, University of Technology, Sydney

Abstract

While standard tests for assessing the reliability of reflective scales exist, the marketing literature places less emphasis on approaches that assist in evaluating formative scales. In this paper we apply the confirmatory tetrads test for the evaluation of formative measurement scales. We do this in the context of a model explaining innovation in marketing partnerships, which includes constructs that are measured using either formative or reflective scales. The findings illustrate the problems associated with the misspecification of measurement scales and related effects on structural model estimations.

Keywords: measurement, formative scales, reflective scales, tetrad test

Introduction

Marketing models often include relationships among sets of latent constructs. Related measurement scales for such latent constructs in SEM's are either formative or reflective in nature. Contrary to the LISREL, EQS and AMOS applications of the covariance-based method, PLS-Graph and MPlus applications of the variance-based approach can handle both types of scales, since formative item weights and reflective item loadings as well as structural model parameters can be estimated simultaneously. Despite PLS applications becoming more common in marketing, there is limited work concerning the assessment of formative scales. Not only are reflective scales often inappropriate for the measurement of certain constructs, formative indicators have better predictive power attributable to their minimization of traces of residual variance in the structural model. While standard tests for assessing the reliability of reflective scales exist, the marketing literature places less emphasis on approaches that assist in evaluating formative scales.

Formative scales are used when a construct is viewed as an explanatory combination of its indicators (Fornell and Bookstein, 1982; Fornell, 1987). In this case, the construct is defined as a total weighted score across all the items, where each item represents an independent dimension in its own right. Increase in the value on one indicator translates into a higher score for the overall scale, regardless of the value on the other indicators. The final score for the construct for any given response is the sum of the weighted scores on all the items. A good formative scale is one that exhausts the entire domain of the construct completely, meaning that the items should collectively represent all the relevant aspects of the construct of interest. In contrast, in a reflective scale all observed indicators are viewed as being caused by some underlying common dimension or construct (Bagozzi, 1982; Bagozzi and Fornell, 1982; Fornell and Bookstein, 1982). Unlike items used in a formative scale, each item in a reflective scale is assumed to share a common core, which is the underlying construct of interest. An increase in the value of the construct then translates into an increase in the value for all the items representing the construct. Typically in marketing studies, formative indicators are tested for their validity using a theoretic rationale and expert opinion (e.g., Rossiter, 2002).

Reflective measures are examined using a range of techniques of scale construction and measurement assessment including factor analysis (Spearman, 1904) and classical test theory (Lord and Novick, 1968) as well as parametric item response theory employing analyses such as Rasch analysis (Rasch, 1960) and non-parametric item response theory using analyses such as Mokken analysis (Mokken, 1971). Thus, while reflective scales are tested concerning unidimensionality, this should not be the case for formative scales (Bollen and Lennox, 1991).

Much of the problem surrounding the absence of formative indicator testing is attributable to construct misspecification. As explained by Jarvis, MacKenzie and Podsakoff (2003), few studies use formative measurement scales, even though such measurement scales would be appropriate in these studies. With the aim of determining the extent of misspecification in marketing studies, these authors reviewed contributions in JMR, JM, JCR and MS over the past 24 years and concluded that only 71 per cent of all measurement scales have been specified correctly; with 68 per cent being reflective ones and 3 per cent being formative ones. Thus, misspecification of the appropriate measurement approach is pervasive (Jarvis et al., 2003).

Misspecification of measurement scales can bias structural parameter estimation and lead to incorrect assessments of structural relationships in SEM's. Both, structural model and measurement scale specification and estimation techniques are relevant in this context; estimation approaches for structural models that account for multiple indicators and multiple causes are, for example, MIMIC (Hauser and Goldberger, 1971) and PLS (e.g., Fornell and Cha, 1994; Fornell and Larcker, 1981). Precluding misspecification of measurement scales assists in avoiding incorrect parameter estimations of structural relationships as suggested in Anderson and Gerbing's *Two-Step Approach* for SEM (Anderson and Gerbing, 1988). However, while this approach focuses on reflective measurement scales and, thus, advocates testing for unidimensionality as a *posterior* practice, employing an *a-priori* technique such as the COARSE procedure (Rossiter, 2002), which provides a framework for *a-priori* measurement scale development, is more appropriate as it accounts for both reflective and formative scales and provides a structured approach prior to scale specification and model estimation. While an *a-priori* practice is required, *posterior* testing for assessing the use of reflective versus formative scales is desirable. However, the marketing literature does not stress techniques for assessing statistically the application of reflective versus formative measurement scales in structural models. In the strategic management literature, Gudergan, Devinney and Ellis (2003) have applied the vanishing tetrad test for evaluating formative and reflective scales. In the next sections of this paper, we will first explain the vanishing tetrad test. We will then apply it to a study of innovation in marketing partnerships. We will follow with conclusions about reflective and formative measurement approaches and testing in marketing.

Confirmatory Tetrad Test for Formative Scale Evaluation

Bollen and Ting (2000) suggest employing the *vanishing tetrad test* for assessing formative measurement scales. This test provides an empirical assessment of whether a formative or reflective scale specification is appropriate. In the next paragraphs we will explain the principles for determining and interpreting vanishing tetrads for measurement scales types. Tetrads, in our context, refer to the difference between the product of a pair covariances and the product of another pair among four measurement scale items[1]. For a minimum of four

[1] Bollen and Ting (1993) provide test statistics that allow undertaking the test of vanishing tetrads for both correlation and covariance data.

items, six covariances can be calculated and three tetrads computed. A vanishing tetrad means that the tetrad equals zero (Spearman, 1904). Bollen and Ting (1993) proposed a confirmatory tetrad test where within the context of measurement scales, index models are specified in advance. These are thus suitable for *posterior* testing regarding the relevance of reflective versus formative scales.

In reflective measurement scales the difference between the covariance products is, by definition, equal to zero; i.e., $\tau_{abcd} = 0$ *for* $a \neq b \neq c \neq d$ (assuming four measurement items). With formative measurement scales this is, by definition, not true unless the individual covariances themselves are zero. Based on these particulars a test can be structured that examines the extent to which a vector of non-redundant tetrads is different from zero. A non-redundant tetrad is one that is linearly related to another; e.g., $\tau_{abcd} = \sigma_{ab}\sigma_{cd} - \sigma_{ac}\sigma_{bd}$ *and* $\tau_{abcd} = \sigma_{ac}\sigma_{db} - \sigma_{ab}\sigma_{cd}$ are redundant and one is excluded from the test. The test statistic utilized is: $T = N\, t'\, \Sigma_{tt}^{-1}\, t$; where N is the sample size, t the vector of independent sample tetrad differences and Σ_{tt}^{-1} the inverse of the covariance matrix of the limiting distribution of t as N goes to infinity, T is asymptotically distributed χ^2 with d degrees of freedom equal to the number of non-redundant tetrads considered (Bollen and Ting, 1993, 2000). The null hypothesis, in our context, is that the measurement scale is reflective, i.e., that $t = 0$ (Bollen and Ting, 2000)[2].

The underlying foundations of tetrad analysis can be linked the works of Wright (1921, 1923) in which causal structures have been modelled statistically. His approach, *path analysis*, combines directed graphs, representing causal hypotheses, and linear statistical models, standing for systems of linear regression equations and statistical constraints. Wright, and others after him, recognise that the causal structure of this approach (the directed graph) determines statistical predictions that can be tested without undertaking experiments. For example, if a causes b directly and b causes c directly, as long as the strengths of the causal effects is not zero, a and c are correlated, but the partial correlation of a and c controlling for b vanishes. With a, b and c being observable items, these causal hypotheses can be tested without undertaking a controlled experiment. If a and c are significantly correlated and not significantly correlated when controlled for b, then the causal hypotheses are supported; if not, then the causal claims are questionable.

Assuming that the causal relationships are captured entirely by the directed graph, the statistical model provides the basis for analysing vanishing correlations and vanishing partial correlations as means of examining statistical constraints as well as statistical consequences of causal relationships (e.g., Blalock, 1961, 1971). Pearl and his colleagues (see Pearl, 1988) use directed acyclic graphs to determine probabilities and the conditional independence relations among them—the d-separation algorithm allows computing the conditional independence relations entailed by the graphs[3]. Spirtes (1994) showed that d-separation correctly calculates conditional independence relations captured by directed graphs interpreted as linear statistical models and Richardson (1994) developed a procedure to determine when linear models are d-separation equivalent. Spirtes and co-authors (1996) showed that d-separation is workable for linear

---

[2] Gudergan, Devinney and Ellis (2003) have previously applied this confirmatory tetrad test and examined this hypothesis within a non-equity alliance context in which alliance performance has been modelled using formative and reflective measurement scales; Venaik, Midgley and Devinney (2004) demonstrate the validity of tetrad tests in management research.

[3] As noted by Pearl (1988) causal independence, also known as Reichenbach's Principle or no correlation without causation, is rooted in the principle of no action of a distance (Arntzenius, 1990).

statistical models with correlated errors; thus providing support for the use of vanishing tetrads and the vanishing tetrad test.

An Application of the Confirmatory Tetrad Test

In this paper, we apply the confirmatory tetrads test for the evaluation of formative measurement scales within the context of a model explaining innovation in marketing partnerships which includes latent constructs of which some are measured using reflective scales and others using formative scales. Leaning on Kenny's (1979) work we examine the measurement scale properties based on internal consistency/validity and consistency of the epistemic relationships/correlations. To examine the internal consistency/validity we apply the confirmatory tetrad analysis using the CTA-SAS routine (Ting, 1995) and PLS estimations of two structural models for (*i*) theory-based based measurement scales (i.e., reflective and formative measurement scales that have been developed using the COARSE procedure) and (*ii*) misspecified measurement scales (i.e., measurement scales that have been constrained to reflective structures using factor and reliability analyses; excluding formative measurement structures). Table 1 represents the results from carrying out reliability analyses and the confirmatory tetrad tests and Table 2 from undertaking the PLS estimations.

Table1: Reliability and Confirmatory Tetrad Test Results

Constructs & Measurement Scales	**Specified Model**				**Misspecified Model**[4]			
	Cronbach Alpha	_2(Df)	Df	p-value	Cronbach Alpha	_2(Df)	Df	p-value
Partnership Creativity	0.89				0.89			
Partnership Learning Process	-----	*51.23*	*20*	*0.0001*	*0.55*	*n.a.*	*n.a.*	*n.a.*
Partnership Knowledge Stock	-----	*53.07*	*20*	*0.0001*	*0.73*	*3*	*2*	*0.2229*
Partnership Innovation	-----	*106.13*	*44*	*0.0000*	*0.82*	*15.72*	*2*	*0.0004*
Partnership Absorptive Capacity	0.82				0.82			
Intrinsic Motivation	0.78				0.78			
Critical Thinking	0.82				0.82			
Extrinsic Motivation	0.79				0.79			
Diversity	-----	*93.18*	*35*	*0.0000*	*0.85*	*57.09*	*9*	*0.0000*
Job Autonomy	0.81				0.81			
Communicative Interaction	-----	*61.01*	*35*	*0.0047*	*0.84*	*26.01*	*2*	*0.0000*
Structure – Formality	0.75				0.75			
Structure – Centralisation	0.69				0.69			
Culture – Risk Orientation	0.75				0.75			
Culture – Collectivism	0.81				0.81			

The results reported in Table 1 show that those scales that have been conceptualised as formative scales should indeed be measured accordingly, with the possible exception of Learning Process. However, based on the initial *a-priori* procedure and related confirmatory

[4] Misspecification of the model is based on employing a reflective scale instead of a formative scale for the measurement of relevant constructs. To accomplish fit statistics consistent with the use of reflective measurement properties some items have been deleted, as they did not relate to an underlying one-dimensional core. The latter is, of course, to be expected given the apriori development of the scales. The reflective measurement scale for Partnership Learning has resulted in keeping three items only in the scale; the confirmatory tetrad test requires at least four items and, thus, could not be run for this scale.

tetrad test, Learning Process should be measured using a formative structure. The results reported in Table 2 show how misspecification of measurement models affects the estimation of the structural model (i.e. changes in significance levels for structural parameters and the overall explanatory power as suggested by the r-squares). Thus, using the confirmatory tetrad test and estimation of the structural model using PLS allows us to examine more thoroughly measurement scale issues by looking at internal consistency/validity and consistency of the epistemic relationships/correlations.

Table 2: PLS Analysis Results

Effects	path coefficients		t-values		significance levels	
	Spec. M.	Misspec. M.	Spec. M.	Misspec. M.	Spec. M.	Misspec. M.
Effects on Partnership Innovation						
(Spec. M. *R_=0.308;* Misspec. M. *R_=0.164)*						
Partnership Creativity	0.2410	0.2320	4.6973	4.6357	****	****
Partnership Learning Process	0.2920	0.2890	4.1323	4.7993	****	****
Partnership Knowledge Stock	0.1700	-0.0120	4.6835	-0.2118	****	n.s.
Effects on Partnership Creativity						
(Spec. M.*R_=0.516;* Misspec. M. *R_=0.492)*						
Intrinsic Motivation	0.1990	0.2130	3.5219	4.5343	****	****
Critical Thinking	0.2510	0.2640	6.3289	6.0761	****	****
Extrinsic Motivation	-0.0370	-0.0350	-1.8421	-0.9536	**	n.s.
Diversity	0.1160	-0.0100	3.5552	-0.1762	****	n.s.
Job Autonomy	0.0850	0.0970	-0.4805	2.4275	n.s.	***
Communicative Interaction	0.1420	0.0900	3.3657	2.2618	****	***
Structure – Formality	0.0020	0.0300	-1.0158	0.4151	n.s.	n.s.
Structure – Centralisation	-0.1120	-0.1200	-2.3590	-2.3504	***	***
Culture – Risk Orientation	0.2900	0.3070	9.6911	7.4765	****	****
Effects on Partnership Learning						
(Spec. M. *R_=0.247;* Missp. M.						
R_=0.151)	0.0820	0.0280	2.9266	0.3548	***	n.s.
Partnership Creativity	0.1770	0.1270	4.7102	2.2139	****	***
Intrinsic Motivation	0.1810	0.2030	2.2407	3.2004	**	***
Critical Thinking	0.0100	0.1100	1.4616	1.9799	*	*
Extrinsic Motivation	0.0330	-0.0340	0.1929	-0.3264	n.s.	n.s.
Diversity	0.2040	0.1160	6.0014	2.1487	****	**
Communicative Interaction	-0.1690	0.0290	-2.3528	0.3069	***	n.s.
Structure – Formality	-0.0850	0.0200	-2.9692	0.2666	****	n.s.
Structure – Centralisation	-0.0510	-0.0080	-1.1130	-0.1279	n.s	n.s
Culture – Collectivism						
Effects on Partnership Knowledge Stock						
(Spec. M. *R_=0.453;* Missp. M.*R_=0.213)*						
Partnership Learning	0.6350	0.2550	11.721	4.3828	****	****
Partnership Capacity	0.0530	0.3010	2.2154	5.3294	**	****

Conclusions

While standard tests for assessing the reliability of reflective scales exist, the marketing literature places less emphasis on approaches that assist in evaluating formative scales. In this paper, we apply the confirmatory tetrads test for the evaluation of formative measurement scales within the context of a model explaining innovation in marketing partnerships which includes latent constructs of which some are measured using reflective scales and others employing formative scales. The findings illustrate the problems associated with misspecification of measurement scales and related effects on structural model estimations.

References

Anderson, J. C., and Gerbing, D. W., 1988. Structural Equation Modeling in Practice: A Review and Recommended Two Step Approach. *Psychological Bulletin.* 103 (May). 411-23.

Arntzenius, F., 1990. Physics and Common Causes. *Synthese.* 82, 77-96.

Bagozzi, R. P., 1982. A Field Investigation of Causal Relationships Among Cognitions, Affect, Intentions, and Behaviour. Journal of Marketing Research. 19, 562-584.

Bagozzi, R. P., and Fornell, C., 1982. Theoretical Concepts, Measurements and Meaning. In A Second Generation of Multivariate Analysis, ed. C. Fornell. New York: Praeger.

Blalock, H., 1961. *Causal Inferences in Non-experimental Research.* University of North Carolina Press, Chapel Hill.

Blalock, H., 1971. *Causal Models in the Social Sciences.* Aldine-Atherton, Chicago.

Bollen, K.A. and Lennox, R. 1991. Conventional Wisdom on Measurement: A Structural Equation Perspective. *Psychological Bulletin.* 110: 305-314.

Bollen, K.A. and Ting, K., 1993. Confirmatory Tetrad Analysis. In P.M. Marsden (ed.) Sociological Methodology: p147-175. Washington, D.C.: American Sociological Association.

Bollen, K.A. and Ting, K. 2000. A Tetrad Test for Causal Indicators. *Psychological Methods.* 5, 3-22.

Fornell, C., 1987. A Second Generation of Multivariate Analysis: Classification of Methods and Implications for Marketing Research. In M. Houston (Ed.), Review of Marketing, American Marketing Association, Chicago, IL.

Fornell, C., and Bookstein, F., 1982. Two Structural Equation Models: LISREL and PLS Applied to Consumer Exit-Voice Theory. *Journal of Marketing Research.* 19, 440-452.

Fornell, C. and Cha, J., 1994. Partial Least Squares. *Advanced Methods of Marketing Research. I*n R. P. Bagozzi. Oxford, Blackwell Publishers.

Fornell, C. and Larcker, V.F., 1981. Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. *Journal of Marketing Research.* 18, 39-50.

Gudergan, S., Devinney, T. and Ellis, S. R., 2003. A Competence-Innovation Framework of Non-Equity Alliance Performance. *ISBM Working Paper Series* at PennState.

Hauser, R.M. and Goldberger, A.S., 1971. The treatment of Unobservable Variables in Path Analysis. In H.L. Costner (Ed.), *Sociological Methodology*, 81-117. Jossey-Bass, San Francisco.

Jarvis, C., MacKenzie, S. and Podsakoff, P., 2003. A Critical Review of Construct Indicators and Measurement Model Misspecification in Marketing and Consumer Research. *Journal of Consumer Research.* 30 (September), 199-218.

Kenny, D.A., 1979. *Correlation and Causality.* John Wiley & Sons, New York.

Lord, F. M. and Novick, M.R., 1968. Statistical Theories of Mental Test Scores. Reading, MA: Addison-Wesley.

Mokken, R.J., 1971. *A Theory and Procedure of Scale Analysis with Applications in Political Research.* De Gruyter, New York.

Pearl, J., 1988. TETRAD and SEM. *Multivariate Behavioural Research.* 33(1), 119-128.

Rasch, G., 1960. *Probabilistic Models for Some Intelligence and Attainment Tests.* Nielsen and Lydiche, Copenhagen.

Richardson, T., 1994. *Properties of Cyclic Graphical Models.* MS Thesis. Carnegie Mellon University, Pittsburgh.

Rossiter, J., 2002. The C-OAR-SE Procedure for Scale Development in Marketing. *International Journal of Research in Marketing.* 19(4), 305-335.

Spearman, C., 1904. General Intelligence: Objectively Determined and Measured. *American Journal of Psychology.* 15, 201-293.

Spirtes, P., 1994. Conditional Independence in Directed Cyclic Graphical Models for Feedback. *Technical Report CMU-PHIL-54.* Department of Philosophy, Carnegie Mellon University, Pittsburgh.

Spirtes, P., Richardson, T., Meek, C., Scheines, R., and Glymour, C., 1996. Using d-Separation to Calculate Zero Partial Correlations in Linear Models with Correlated Errors. *Technical Report CMU-PHIL-72.* Department of Philosophy, Carnegie Mellon University, Pittsburgh.

Ting, K.-F., 1995. Confirmatory Tetrad Analysis in SAS. *Structural Equation Modelling.* 2(2), 163-171.

Venaik, S, Midgley, D.F. and Devinney, T.M., 2004. A New Perspective on the Integration-Responsiveness Pressures Confronting Multinational Firms. *Management International Review.*

Wright, S., 1921. Correlation and Causation. *Journal of Agricultural Research.* 20, 557-585.

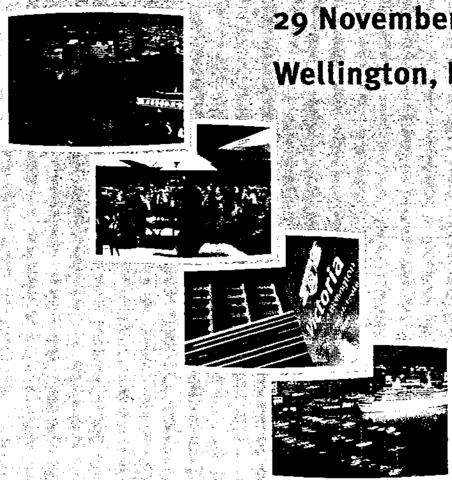
Wright, S., 1923. The Theory of Path Coefficients: A reply to Niles' criticism. Genetics. 8, 239-255.

ANZMAC 2004

Marketing Accountabilities and Responsibilities

29 November – 1 December 2004

Wellington, New Zealand



Conference Proceedings

System recommendations

Macintosh: Macintosh power PC, 32MB of RAM, Apple System Software verion 8.0 or later and CD ROM drive.

Windows: Pentium PC, 32MB of RAM, Microsoft Windows 95 or later and CD ROM drive.

To access: double-click on home.htm to start.

This CD ROM requires Acrobat Reader 4.0 or later.

ISBN Conference proceedings: CD 0-475-22215-1

Every reasonable endeavour has been undertaken to ensure that all materials contained in this CD ROM were correct at the time of creation – November 2004

www.vuw.ac.nz/anzmac