

Faculty of Informatics

Faculty of Informatics - Papers

University of Wollongong

Year 2008

Formative versus reflective measurement
models: Two applications of formative
measurement

T. Coltman* T. M. Devinney†
D. F. Midgley‡ S. Venaik**

*University of Wollongong, tcoltman@uow.edu.au

†University of New South Wales

‡INSEAD, France

**University of Queensland

This article was originally published as Coltman, T, Devinney, TM, Midgley, DF & Venaik, S, Formative versus reflective measurement models: Two applications of formative measurement, *Journal of Business Research*, 61(12), 2008, 1250-1262. Original journal article available here

This paper is posted at Research Online.

<http://ro.uow.edu.au/infopapers/689>

Formative versus Reflective Measurement Models:

Two Applications of Formative Measurement

Tim Coltman, University of Wollongong

Timothy M. Devinney, The University of New South Wales

David F. Midgley, INSEAD

Sunil Venaik, The University of Queensland

December 2007

The views here are solely those of the authors, who appear in alphabetical order. The authors thank the anonymous reviewers and the editor of the special issue for their many helpful comments. Send correspondence to Tim Coltman, Centre for Business Services Science, University of Wollongong, Wollongong, Australia 2522 (tcoltman@uow.edu.au). Timothy M. Devinney, Australian School of Business, The University of New South Wales, Sydney, Australia 2052 (T.Devinney@agsm.edu.au). David F. Midgley, Marketing Area, INSEAD, Boulevard de Constance, 77305 Fontainebleau, France (david.midgley@insead.edu). Sunil Venaik, UQ Business School, The University of Queensland, Brisbane, Australia 4072 (svenaik@business.uq.edu.au).

Abstract

This paper presents an organizing framework that assists researchers in the design and validation of formative and reflective measurement models. The framework draws from the extant literature, includes both theoretical and empirical considerations, and is illustrated through two important examples, one from international business and one from marketing. Both examples concern constructs that are fundamental to theory-building in these disciplines, and constructs that most scholars measure reflectively. In contrast, application of the framework to these examples suggests that a formative measurement model may be more appropriate. These results reinforce the need for all researchers to justify, both theoretically and empirically, their choice of measurement model for their constructs. Utilization of an incorrect measurement model undermines the content validity of the constructs, misrepresents the structural relationships within which these constructs are embedded, and ultimately lowers the usefulness of management theories for business researchers and practitioners. The main contribution of this paper is to question the unthinking assumption of reflective measurement seen in much of the business literature.

Keywords: Formative, reflective, international business, integration-responsiveness, marketing, market orientation

Formative versus Reflective Measurement Models:

Two Examples of Formative Measurement

1. Introduction

Management scholars often identify structural relationships among latent, unobserved constructs by statistically relating covariation between the latent constructs and the observed variables or indicators of the latent constructs (Borsboom, Mellenbergh, and Heerden, 2003, 2004). This allows scholars to argue that if variation in an indicator *X* is associated with variation in a latent construct *Y*, then exogenous interventions that change *Y* can be detected in the indicator *X*. Most scholars assume this relationship between construct and indicator is reflective. In other words, the change in *X* reflects the change in the latent construct *Y*. With reflective (or effect) measurement models, causality flows from the latent construct to the indicator.

However, not all latent constructs are entities that are measurable with a battery of positively correlated items (Bollen and Lennox, 1991; Edwards and Bagozzi, 2000; Fornell, 1982). A less common, but equally plausible approach is to combine a number of indicators to form a construct without any assumptions as to the patterns of inter-correlation between these items. A formative or causal index results (Blalock, 1964; Diamantopoulos and Winklhofer, 2001; Edwards and Bagozzi, 2000) where causality flows in the opposite direction, from the indicator to the construct. Although the reflective view dominates the psychological and management sciences, the formative view is common in economics and sociology.

The distinction between formative and reflective measures is important because proper specification of a measurement model is necessary to assign meaningful relationships in the

structural model (Anderson and Gerbing, 1988). Theoretical work in construct validity (Blalock, 1982; DeVellis, 1991; Edwards and Bagozzi, 2000) and structural equation modeling (Baumgartner and Homburg, 1996; Chin and Todd, 1995; Shook, Ketchen, Hult, and Kacmar, 2004) enhances our understanding, however, considerable debate still exists regarding the procedures a working researcher should follow to achieve construct validity (e.g., Diamantopoulos, 2005; Finn and Kayande, 2005; Rossiter, 2005). This paper is not to repeat or continue this debate. Rather, the authors take the middle ground, building on the work of both those who stress theoretical justifications for constructs and those who argue for empirical validation as part of measure development.

This paper presents an organizing framework for construct measurement that begins with theoretical justification to define the nature of the focal constructs, and then employs a series of empirical tests to support the causal direction between constructs and their measures. The framework builds on the work of Jarvis, Mackenzie, and Podsakoff (2003) who provide a set of decision rules for deciding whether the measurement model should be formative or reflective. However, the framework here differs from Jarvis et al.'s decision rules in several respects, most importantly in the procedures proposed and the attention to measurement error.

The major contribution of this paper is to question the common assumption of a reflective measurement model seen in much of the empirical business literature. The validity of this assumption is measured by applying the proposed framework to two widely used constructs in the business literature, integration responsiveness (from the discipline of international business) and market orientation (from the discipline of marketing). These two empirical examples are chosen: (1) because of the predominance of the reflective modeling approach for these constructs, even though a formative model can be theoretically more appropriate, and (2) due to the

criticality of the underlying phenomena to the development of the disciplines of international business and marketing.

In the case of the integration responsiveness framework, the diverse measures of each of the integration and responsiveness pressures are unlikely to be highly intercorrelated as a reflective structure requires. A priori, a formative approach to measurement would seem worthy of consideration, yet most of the work in this area takes the reflective stance to measurement, often without any consideration of alternatives (Venaik, Midgley, and Devinney, 2004). Similarly, most research on market orientation defines it as a one-dimensional construct measured through a multi-item reflective scale. Yet, the main scales that measure market orientation—MARKOR (Kohli and Jaworski, 1990) and MORTN (Deshpande and Farley, 1998)—are conceptualized as a set of activities that make up the attribute (see Narver and Slater, 1990, p. 21), implying a formative model. Furthermore, the substantive inconsistencies in the market orientation literature (Langerak, 2003) raise many questions about the dimensionality (Siguaw and Diamantopoulos, 1995) and measurement (Narver, Slater, and MacLachlan, 2004) of the market orientation construct. These examples serve to illustrate a problem in the international business and marketing literature, where insufficient attention is paid to the measurement of constructs.

The paper is organized as follows. The next section presents the organizing framework for designing and validating reflective and formative models using both theoretical and empirical considerations. Then our framework is applied to the two illustrative and important examples taken, respectively, from international business and marketing. The purpose here is to examine whether reflective or formative measurement models are more or less appropriate, not to debate the content validity of the measures that various scholars adopt.

2. An organizing framework for designing and validating reflective and formative models

In recent years, scholars have begun to challenge the blind adherence to Churchill's (1979) procedure with its strict emphasis on exploratory factor analysis (Spearman, 1904), internal consistency (Cronbach, 1951) and the domain sampling model (Nunnally and Bernstein, 1994). In psychology, Borsboom et al. (2003, 2004) use basic logic and measurement theory to argue that the choice of model is dependent upon the ontology invoked by the latent construct. In marketing, Rossiter (2002) provides a general procedure for scale development which extends "accepted" practice by reemphasizing the importance of theoretical considerations. Borsboom and Rossiter both argue that scholars should focus only on theoretical considerations and resist the temptation to conduct empirical tests.

Alternatively, Diamantopoulos (2005) and Finn and Kayande (2005) argue that both theoretical and empirical criteria are necessary to design and validate measurement models. Empirical analyses provide an important foundation for content validity, especially to detect errors and misspecifications or wrongly conceived theories. For example, finding a negative relationship when theory and common sense suggest a positive relationship would be a concern for researchers.

This paper follows the stance of Diamantopoulos, and Finn and Kayande but takes a different perspective on empirical measurement and the role that measures play in the choice of a formative or reflective measurement model. To comprehensively capture the necessary theoretical and empirical aspects, the paper presents an organizing framework for designing and validating formative and reflective models (see Table 1). As shown in the table, three theoretical considerations and three empirical considerations distinguish formative models from reflective ones. The following sections briefly discuss each of these considerations.

Table 1: A Framework For Assessing Reflective and Formative Models: Theoretical and Empirical Considerations

Considerations	Reflective model	Formative model	Relevant literature
<i>Theoretical Considerations</i>			
1. Nature of construct	Latent construct is existing ➤ Latent construct exists independent of the measures used	Latent construct is formed ➤ Latent constructs is determined as a combination of its indicators	Borsboom et al. (2003, 2004)
2. Direction of causality between items and latent construct	Causality from construct to items ➤ Variation in the construct causes variation in the item measures ➤ Variation in item measures does not cause variation in the construct	Causality from items to construct ➤ Variation in the construct does not cause variation in the item measures ➤ Variation in item measures causes variation in the construct	Bollen and Lennox (1991); Edwards and Bagozzi (2000); Rossiter (2002); Jarvis et al. (2003)
3. Characteristics of items used to measure the construct	Items are manifested by the construct ➤ Items share a common theme ➤ Items are interchangeable ➤ Adding or dropping an item does not change the conceptual domain of the construct	Items define the construct ➤ Items need not share a common theme ➤ Items are not interchangeable ➤ Adding or dropping an item may change the conceptual domain of the construct	Rossiter (2002) ; Jarvis et al. (2003)
<i>Empirical Considerations</i>			
4. Item intercorrelation	Items should have high positive intercorrelations ➤ Empirical test: internal consistency and reliability assessed via Cronbach alpha, average variance extracted, and factor loadings (e.g., from common or confirmatory factor analysis)	Items can have any pattern of intercorrelation but should possess the same directional relationship ➤ Empirical test: indicator reliability cannot be assessed empirically; various preliminary analyses are useful to check directionality between items and construct	Cronbach (1951); Nunnally and Bernstein (1994); Churchill (1979); Diamantopoulos and Siguaw (2006)
5. Item relationships with construct antecedents and consequences	Items have similar sign and significance of relationships with the antecedents/consequences as the construct ➤ Empirical test: content validity is established based on theoretical considerations, and assessed empirically via convergent and discriminant validity	Items may not have similar significance of relationships with the antecedents/consequences as the construct ➤ Empirical test: nomological validity can be assessed empirically using a MIMIC model, and/or structural linkage with another criterion variable	Bollen and Lennox (1991); Diamantopoulos and Winklhofer (2001); Diamantopoulos and Siguaw (2006)
6. Measurement error and collinearity	Error term in items can be identified ➤ Empirical test: common factor analysis can be used to identify and extract out measurement error	Error term cannot be identified if the formative measurement model is estimated in isolation ➤ Empirical test: vanishing tetrad test can be used to determine if the formative items behave as predicted ➤ Collinearity should be ruled out by standard diagnostics such as the condition index	Bollen and Ting (2000); Diamantopoulos (2006)

2.1 Theoretical considerations

Three broad theoretical considerations are important in deciding whether the measurement model is formative or reflective. These considerations include: (1) the nature of the construct, (2) the direction of causality between the indicators and the latent construct, and (3) the characteristics of the indicators used to measure the construct [numbering relates to the rows in Table 1].

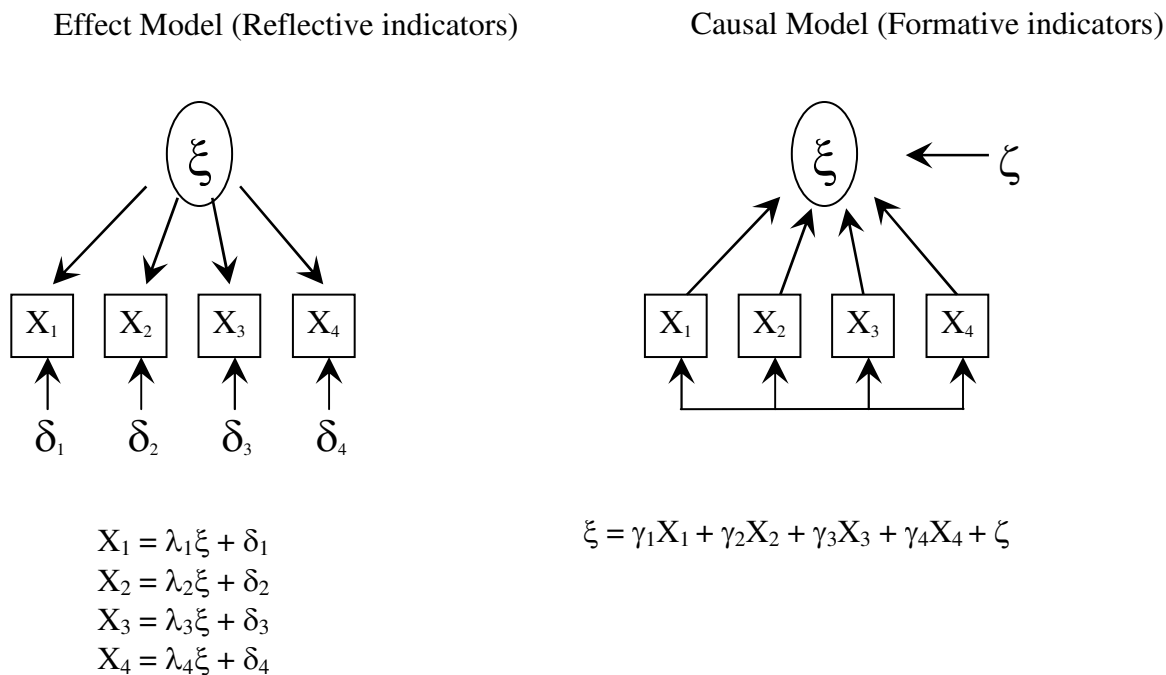
Consideration 1: The nature of the construct. In a reflective model, the latent construct exists (in an absolute sense) independent of the measures (Borsboom et al., 2004; Rossiter, 2002). Typical examples of reflective scenarios include measures of attitudes and personality that are measured by eliciting responses to indicators. Practically all scales in business and related methodological texts on scale development (Bearden and Netmeyer, 1999; Bruner II, James, and Hensel, 2001; Netmeyer, Bearden, and Sharma, 2003; Spector, 1992) use a reflective approach to measurement. For example, examination of papers in the *Journal of International Business Studies* and the *Journal of Marketing* in 2006 reveals that nearly 95 percent of the constructs measured with multiple items use a reflective structure without consideration of an alternative formulation.

In contrast, in a formative model, the latent construct is dependent upon a constructivist, operationalist or instrumentalist interpretation by the scholar (Borsboom et al., 2003). For example, the human development index (HDI) does not exist as an independent entity. Rather, it is a composite measure of human development that includes: health, education and income (UNDP, 2006). Any change in one or more of these components is likely to cause a change in a country's HDI score. In contrast to the reflective model, few examples of formative models are seen in the business literature.

Consideration 2: Direction of causality. The second key theoretical consideration in deciding whether the measurement model is reflective or formative is the direction of causality between the construct and the indicators. As shown in Figure 1, reflective models assume that causality flows from the construct to the indicators. In the case of formative models, the reverse is the case, causality flows

from the indicators to the construct. Hence, in reflective models, a change in the construct causes a change in the indicators. In the case of formative models, it is the other way around; a change in the indicators results in a change in the construct under study. Thus, the two models in Figure 1 are different, both psychometrically and conceptually (Bollen and Lennox, 1991). The difference in causal direction has profound implications both for measurement error (Diamantopoulos, 2006) and model estimation; topics discussed in section 2.2.

Figure 1: Reflective and Formative Measures



Consideration 3: Characteristics of indicators. Significant differences are present in the characteristics of the indicators that measure the latent constructs under reflective and formative scenarios. In a reflective model, change in the latent variable must precede variation in the indicator(s). Thus, the indicators all share a common theme and are interchangeable. This indicator

interchangeability, enables researchers to measure the construct by sampling a few relevant indicators underlying the domain of the construct (Churchill, 1979; Nunnally and Bernstein, 1994). Inclusion or exclusion of one or more indicators from the domain does not materially alter the content validity of the construct.

However, the situation is different in the case of formative models. Since the indicators define the construct, the domain of the construct is sensitive to the number and types of indicators representing the construct. Hence, adding or removing an indicator can change the conceptual domain of the construct. However, as Rossiter (2002) points out, this does not mean that we need a census of indicators as Bollen and Lennox (1991) suggest. As long as the indicators conceptually represent the domain of interest, they may be considered adequate from the standpoint of empirical prediction.

2.2 Empirical considerations

Paralleling the three theoretical considerations above, are three empirical considerations that inform understanding of the measurement model: (4) indicator intercorrelation, (5) indicator relationships with construct antecedents and consequences, and (6) measurement error and collinearity [numbering relates to the rows in Table 1].

Consideration 4: Indicator intercorrelation. In a reflective model, the indicators are evoked by the underlying construct and have positive and, desirably, high intercorrelations. In a formative model, the indicators do not necessarily share the same theme and hence have no preconceived pattern of intercorrelation. Indicators in a formative model can theoretically possess no intercorrelation or high or low intercorrelation.

Regardless, researchers should check that indicator intercorrelations are as they expect. Such checks are a necessary part of the various preliminary analyses for questionnaire items administered to samples of respondents. These preliminary analyses include checking for the presence of outliers (e.g., using distances in factor spaces for reflective measurement models or regression influence diagnostics

for formative models); checking that the dimensionality of the construct is consistent with a researcher's hypothesis (e.g., using common factor models or principal components analysis); establishing that the correlations between items and constructs have the expected directionality and strength (e.g., through bivariate correlations, factor or regression analysis); reliability statistics (in the case of the reflective measurement model); and, where several constructs are part of a theoretical structure, showing that common method bias is not an issue (e.g., by the absence of one common factor). Some of these preliminary analyses (and the diagnostics that go with them) shed useful light on issues of indicator intercorrelation and inferentially suggest whether one measurement model or another might be preferred. However, in themselves, they cannot either support or disconfirm theoretical expectations as to the nature of the measurement model. For that, researchers require stronger tests.

Since reflective indicators have positive intercorrelations, measures such as factor loading and communality, Cronbach alpha, average variance extracted and internal consistency are used to empirically assess the individual and composite reliabilities of the indicators (Trochim, 2007). However, as these measures of reliability assume internal consistency—that is, high intercorrelations among the indicators in question—they are inappropriate for formative indicators, where no theoretical assumption is made about inter-item correlation. One of the key operational issues in the use of formative indicators is that no simple, easy and universally accepted criteria exists for assessing the reliability of formative indicators.

Consideration 5: Indicator relationships with construct antecedents and consequences. In the case of reflective models, the indicators have a similar (positive/negative, significant/non-significant) relationship with the antecedents and consequences of the construct. The requirement for interrelated indicators is not the case for formative indicators as they do not necessarily share a common theme and, therefore, do not have the same types of linkages with the antecedents and

consequences of the construct. This requirement is a significant issue when using formative models, particularly as it has implications about the appropriate level of aggregation of formative indicators. While aggregating indicators to create a construct achieves the objective of model parsimony, it may come at a significant cost in terms of the loss of the rich, diverse and unique information embedded in the individual indicators underlying the theoretical model. Edwards (2001) makes a similar point for second and higher order dimensions.

In the case of formative measurement, Diamantopoulos and Winklhofer (2001) suggest three possible approaches. First, one relates the indicators to some simple overall index variable, such as a summary or overall rating—this approach is taken for the second example (market orientation). Second, one applies a Multiple Indicators and Multiple Causes (MIMIC) model, where both formative and reflective indicators measure the construct. Third, one applies a structural model linking the formatively measured construct with another construct that is theoretically related and measured with reflective items. This approach establishes criterion and nomological validity, and is the approach taken in the first example (i.e., integration-responsiveness pressures).

Consideration 6: Measurement error and collinearity. A key difference between formative and reflective models is the treatment of measurement error. As shown in Figure 1, an important assumption underlying the reflective measurement model is that all error terms (δ_i of Figure 1) are associated with the observed scores (x_i) and, therefore, represent measurement error in the latent variable. Such a correlational structure is not assumed in the case of a formative model. The disturbance term (ζ) is not associated with the individual indicator or the set of indicators as a whole, and therefore does not represent measurement error (Diamantopoulos, 2006).

In the case of reflective models, researchers can identify and eliminate measurement error for each indicator using common factor analysis because the factor score contains only that part of the indicator that is shared with other indicators, and excludes the error in the items used to compute the

scale score (Spearman, 1904). However, in the case of formative models, the only way to overcome measurement error is to design it out of the study before collecting the data. Diamantopoulos (2006) suggests two possible ways to eliminate the error term: (1) capture all possible causes on the construct, and (2) specify the focal construct in such a way as to capture the full set of indicators. Both approaches legitimately exclude the error term ($\zeta=0$). In the light of the above, it is clear that unlike the reflective model, no simple way exists to empirically assess the impact of measurement error in a formative model.

However, Bollen and Ting (2000) suggest that the tetrad test can provide some assistance for the assessment of measurement error in formative models. A “tetrad” refers to the difference between the products of two pairs of error covariances. Derived from Spearman and Holzinger (1924), the test is based on nested vanishing tetrads that are implied by comparing two theoretical measurement models. In the case of a reflective model, the null hypothesis is that the set of non-overlapping tetrads vanishes. In simpler terms, when the intercorrelations between pairs of errors are compared, they should tend to zero. Referring back to Figure 1, the assumption underlying the reflective model is that the correlations between the δ_i are zero. The tetrad test confirms whether or not this is true.

The tetrad test is a confirmatory procedure that should not be used as a stand-alone criterion for distinguishing formative from reflective models. Specifically, if the hypothesis that the errors are uncorrelated is rejected, it can be for one of two alternative reasons. One is that the construct is better measured formatively, not reflectively. The other is that reflective measurement is more appropriate but the error structure is contaminated. One possible source of contamination is common method error. Similarly, if the hypothesis that the errors are uncorrelated is accepted, this could still be a mistake. A possibility, although unlikely in practice, is that a formative model is correct but that the indicator error structures are uncorrelated. Thus, while serving as an important pointer, the results from the tetrad test do not provide definitive proof as to the correct measurement model.

Another measurement issue that researchers need to check in formative models is collinearity. The presence of highly correlated indicators will make estimation of their weights in the formative model difficult and result in imprecise values for these weights. Given a criterion variable, as above, an estimate of the impact of collinearity can be made by regressing the indicators on this variable and computing standard diagnostics such as the condition index.

In the next section, the three sets of theoretical criteria and three sets of empirical criteria, are applied to two key constructs in international business and marketing, integration-responsiveness and marketing orientation.

3. Application one: measuring international business pressures

The Integration Responsiveness (IR) framework of Prahalad and Doz (1987) is widely used in the international business literature to characterize the environmental pressures confronting firms as they expand worldwide. According to this framework, firms come under countervailing pressures to simultaneously coordinate the activities and strategies of their local business units to attain global competitive advantage (global integration) while adapting these activities and strategies to the unique circumstances of the countries in which they operate (local responsiveness).

Although this framework has been applied for over a decade, the issue of relevance here is whether the formative or reflective measurement model is appropriate for these pressures. Venaik et al.'s (2004) review of the literature demonstrates that nearly all researchers have taken the reflective route and only a handful the formative. More critically, their study shows that little published debate, justification or validation can be found to justify the route that each researcher took. Hence, it is important to apply the theoretical and empirical considerations enunciated in Table 1.

3.1. Theoretical considerations

Consideration 1: nature of the construct. The environmental pressures facing a multinational enterprise cover a domain of enormous breadth and diversity. Researchers in international business

have characterized these pressures as global integration pressures—global competition, the need to reduce costs, and the pressures of technological change, etc.—and local responsiveness pressures—diversity of market infrastructure, country based regulation, local customer heterogeneity, etc. It is difficult to think of these pressures as being innate characteristics of the business environment that actually cause overall global integration or local responsiveness pressures.

Consideration 2: direction of causality. A more logical approach is to view the diverse facets of the environment as forming IR pressures rather than the other way around. Indeed, the very word “pressures” implies this view (from the Latin *pressura*—the action of pressing, Webster’s Dictionary). Thus, the direction of causality is from the various aspects of the international business environment to what the researcher defines as the pressures through the measures chosen, rather than “latent” pressures being reflected in correlated measures. Therefore, a formative model is likely to be a more appropriate structure for testing the IR framework.

Consideration 3: characteristics of indicators. Additionally, it is not clear that the individual items in this domain—be they questionnaire items or variables from economic databases—share a common theme in the way required by the reflective approach. For example, any number of integration pressures may underlie the firm’s need to integrate its activities worldwide—these could include “the importance of multinational customers”, “investment intensity”, etc. (Prahalad and Doz, 1987; pp. 18–19). There is little reason to believe that all these pressures are sampled from a common domain and are interchangeable, as is required when applying a reflective approach. Why, for example, would an item designed to measure the “importance of multinational customers” necessarily be related with one designed to measure “investment intensity.” Similarly, country infrastructure is a different aspect of local responsiveness pressures than say, subsidiary country regulations, even though both force firms to design their strategies on a country-by-county basis. Indeed, the diversity of phenomena that needs to

be considered under the heading of IR pressures suggests at least a *prima facie* case for the formative viewpoint.

Based on the three theoretical considerations in the proposed framework—the nature of the construct, the direction of causality, and the characteristics of the items used to represent the construct—the IR framework is best conceptualized and measured using a formative model. Next, we apply a number of empirical tests to corroborate the suitability of this model.

3.2. Empirical considerations

Based on a comprehensive survey of the IR literature, Venaik et al. (2004) administered 23 indicators of IR pressures to a sample of 163 managers from the subsidiaries of multinational firms in 35 countries. These data form the basis for the empirical tests discussed below.

Consideration 4: indicator intercorrelation. As discussed above, Venaik et al. conducted a range of preliminary analyses on these data (including outlier detection, bivariate correlation analysis, principal component analysis and common factor analysis). The major conclusion from these analyses relevant to this paper is that more than two integration-responsiveness pressures are needed to adequately represent the domain of the 23 indicators. At least for these data, five pressures are needed to represent what much of the literature has forced into two. Table 2 shows the association between the 23 items and these five pressures of government influence, quality of local infrastructure, global competition, technological change and resource sharing, as shown by preliminary analyses. The five pressures are largely independent of one another, as demonstrated by low intercorrelations in oblique rotations. Given these five pressures, the directionality and strength of the indicators also fit expectations. However, diagnostics for the common factor model were poor, raising concerns as to whether the reflective model was appropriate. Overall, these initial analyses support the theoretical considerations above by tentatively suggesting five formatively measured pressures rather than two reflectively measured ones.

Table 2: IR Pressures. Dimensionality and Association between Constructs and Indicators Suggested by Preliminary Analyses

No. Indicators	Constructs				
	Government influence	Quality of local infrastructure	Global competition	Technological change	Resource sharing
1 Product decisions influenced by government	√				
2 Price decisions influenced by government	√				
3 Advertising decisions influenced by government	√				
4 Promotion decisions influenced by government	√				
5 Sourcing decisions influenced by government	√				
6 R&D decisions influenced by government	√				
7 Quality of local infrastructure: logistics		√			
8 Quality of local infrastructure: channels		√			
9 Quality of local infrastructure: advertising		√			
10 Quality of local infrastructure: personnel		√			
11 Quality of local infrastructure: suppliers		√			
12 Competitors are mostly global			√		
13 Competitors sell globally standardized products			√		
14 The nature of competition is global			√		
15 Co-ordination of production is global			√		
16 Co-ordination of procurement is global			√		
17 Rate of product innovation				√	
18 Rate of process innovation				√	
19 Technological complexity				√	
20 Rate of technological change				√	
21 Sharing of production resources					√
22 Sharing of R&D resources					√
23 Sharing of management services					√

Source: Adapted from Venaik et al. (2004).

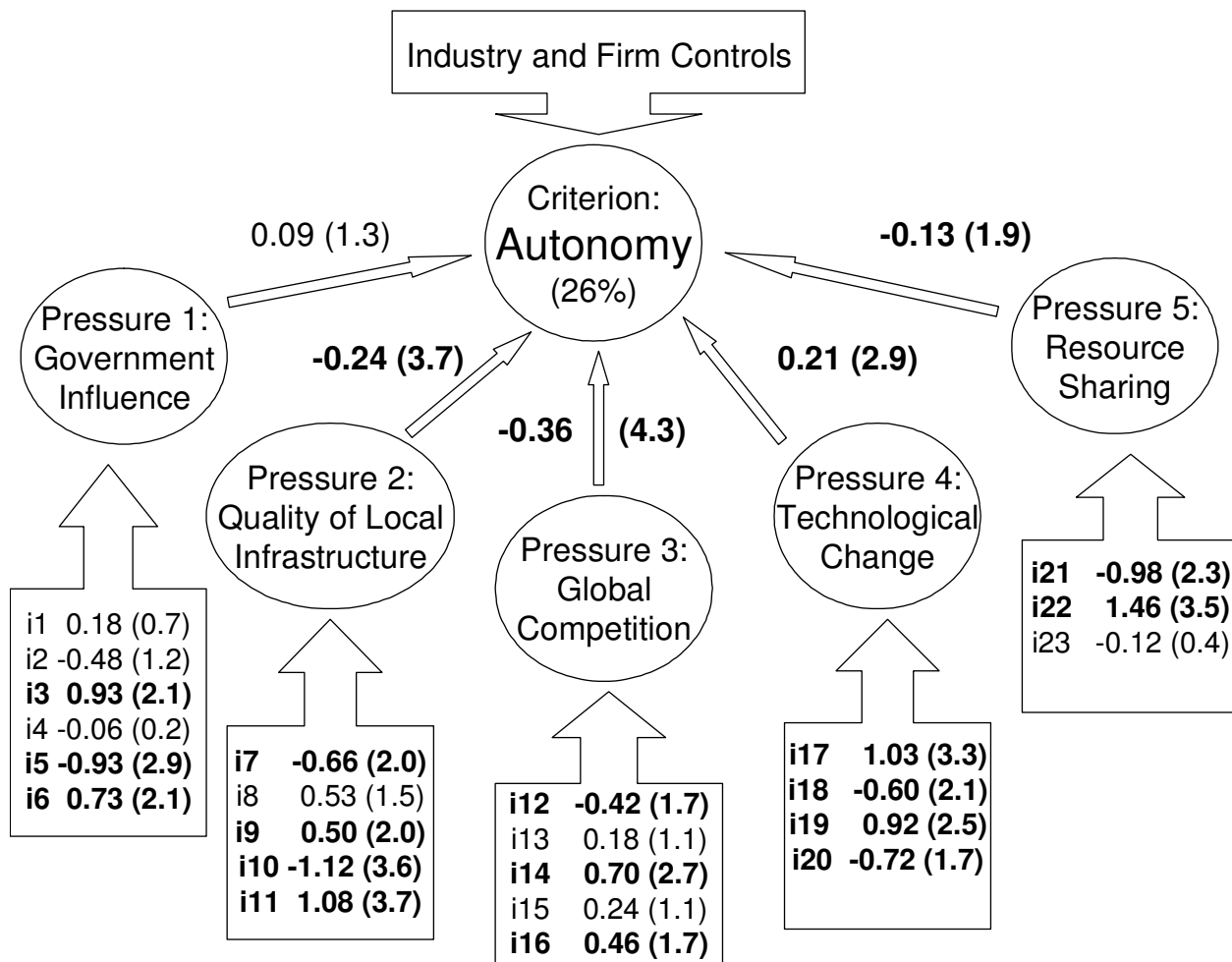
Consideration 5: indicator relationships with construct antecedents and consequences.

Five formatively measured pressures are used to predict the independent reflectively measured construct of subsidiary Autonomy (a one-dimensional construct with composite reliability of 0.90 and average variance extracted of 61%). This additional construct of Autonomy is the criterion construct which identifies the formative model (Diamantopoulos and Winklhofer, 2001). Autonomy is of

theoretical relevance as it is considered in the literature to be one of the most important consequences of global pressures on firms. Control variables are included to provide greater confidence that any observed effects are not spurious results of industry and firm heterogeneity. The technique of partial least squares (PLS) is used for this analysis (Chin, 1998) and Figure 2 shows the results obtained. These results add further support to the formative model as the five pressures predict Autonomy well and the majority of outer item coefficients and inner path coefficients have the right signs and adequate t-statistics. The exception is government influence, which, although the formative model seems appropriate from the individual indicator perspective, does not predict Autonomy.

[white space intentionally left blank]

Figure 2: Test of Criterion Validity for IR Pressures Measured Formatively



Note: Boxes contain outer indicator coefficients; inner path coefficients are next to arrows (with the absolute values of the bootstrap t-statistic in parentheses). All significant values are shown in bold type ($p < 0.05$). The percentage under the independent, reflective construct of *Autonomy* is the R^2 . The indicator numbers i1, i2, etc. refer to the measures listed in Table 2.

Source: Adapted from Venaik et al. (2004).

However, it is difficult to judge a structural equation model in isolation and the five pressures are measured reflectively measured, by rerunning the PLS analysis with indicator directionality reversed. This additional analysis provides a clear comparison between reflective and formative measurement models (Diamantopoulos and Siguaw, 2006).

Although noting that reflective models always explain less variance than formative models (which are optimized for prediction), the reflective measurement model performs much worse than the formative one. The reflectively measured pressures explain 17 percent of the variance in Autonomy compared with 23 percent for the formative model. (The total variance explained, including firm and industry controls, is 22% and 26%, respectively). Examination of the item coefficients shows that this difference in performance is not due to poor measurement—for the reflective model, all the item loadings and t-values are high, and for the formative model all pressures have an adequate number of significant weights. Instead, the difference is attributable to the reflectively measured pressures not explaining the independent construct as well as the formatively measured ones do. Indeed, only one of the five reflectively measured pressures has a significant and meaningful path coefficient with Autonomy (where meaningful is $\beta > 0.20$, Meehl, 1990), namely global competition ($\beta = -0.45$, $p < 0.01$), whereas three of the formatively measured pressures have a significant and meaningful path coefficient (quality of local infrastructure, global competition and technological change, $\beta = -0.24$, -0.36 and 0.21 , respectively; all with $p < 0.01$). For the international business literature, this is an important finding. Most scholars expect IR pressures to impact on the degree of subsidiary autonomy (e.g., Dunning, 1988) and thus, above and beyond its demonstrated empirical superiority, would consider the formative model more theoretically valid.

The other model comparison that is relevant is with the measurement model commonly accepted in the literature: for example, a two-dimensional model where the pressures of Global Integration (dimensions 3 through 5 in Figure 2) and Local Responsiveness (dimensions 1 and 2) are measured reflectively. For these data, this model is neither theoretically nor empirically compelling. Although the R-square is adequate at 17 percent (excluding 4% of variance explained by controls), only the path from Global Integration to Autonomy is significant ($\beta = -0.42$, $p < 0.02$). The path from Local Responsiveness to Autonomy is not significant ($\beta = -0.14$, $p > 0.15$). The latter should be of concern

to IR theorists. Furthermore, as might be expected when several dimensions are collapsed into two, the reflective measurement diagnostics are not strong, especially average variance extracted (which is less than 30% in both cases). If the measures are pruned in the traditional manner, these diagnostics can be improved, but only at the expense of prediction and meaning. Global Integration becomes defined solely as global competition and all the other pressures disappear from the model. For IR pressures, these results support the theoretical arguments that formative measurement may be more appropriate.

Consideration 6: measurement error and collinearity. We also apply the vanishing tetrad test to each construct. This test rejects the reflective model for four of the five constructs, lending added support to the formative view taken here. However, with the fifth construct, the pressure of resource sharing, the test does not reject the reflective model (see Table 3). As noted above, this can be because this construct is truly better measured reflectively or because indicators in the formative construct are uncorrelated. Here the correlations between the sharing of production, R&D and management service resources are modest but not zero. Re-running the PLS analysis switching resource sharing from a formative to a reflective measurement model results in a non-significant impact of this construct on Autonomy. Although it is not possible to reach a definitive conclusion with these data, it does suggest that resource sharing might also be better conceptualized formatively, as theory indicates.

Collinearity is not an issue in these results as the largest condition indices from regressions of the five sets of indicators range from 7.1 to 13.8, all of which are less than 15 (the accepted heuristic for the point at which some concerns of collinearity start to emerge) and well below 30 (the accepted heuristic for clear collinearity problems).

Table 3: IR Pressures: Tetrad Test Results for Formative Indicators

Constructs	Number of indicators	χ^2 (Df)	Df	Significance	Implication
Government influence	6	22.4	9	< 0.01	Formative
Quality of local infrastructure	5	19.9	5	< 0.01	Formative
Global competition	5	20.5	5	< 0.01	Formative
Technological Change	4	9.8	2	< 0.01	Formative
Resource sharing	3*	1.0	2	0.59	Reflective

*As this construct had three indicators, a fourth—unrelated—indicator was added to the test. This follows the advice of Bollen and Ting (2000).

To sum up, much of the extant research uncritically assumes a reflective measurement model when empirically representing the integration-responsiveness pressures confronting multinational firms. However, both theoretical and empirical analysis shows that this assumption is debatable. The first three theoretical considerations clearly indicate that no prima facie rationale exists for the large set of measures that represent the broad, diverse and complex domain of integration-responsiveness pressures to share a common theme or be related to one another. The second three empirical considerations and statistical analyses, together with tetrad tests, lend further support to the formative measurement model. Next, we apply the same six considerations to another important construct, market orientation.

4. Application two: measuring market orientation

The concept of market orientation has long been a cornerstone in marketing strategy. The literature in marketing stipulates that organizations should allocate resources to the systematic gathering and analysis of customer and competitor information and to make use of customer knowledge to guide a customer linking strategy (Hunt and Morgan, 1995; p. 11). The emphasis placed on market orientation as a driver of competitive advantage and business performance in marketing is not surprising. The main tenets of this view—that is, customer-oriented thinking, customer analysis and

understanding—are fundamental to the beliefs of the discipline. However, despite the concept's apparent credibility, the literature suffers from inconsistent measures (Mason and Harris, 2005).

The empirical evidence also indicates that the power of market orientation to predict advantage or performance is still an open question (Langerak, 2003). For example, Agarwal and Erramilli (2003) report no direct relationship, while Grewal and Tansuhaj (2001) show mixed results. These inconsistencies imply that either the theory underpinning market orientation does not hold or that the measurement model used to operationalize the construct is incorrect. This paper aims to demonstrate that the latter is arguably one cause of these inconsistent results.

4.1. Theoretical considerations

Consideration 1: nature of the construct. The conceptualization of market orientation builds from either cultural or behavioral criteria. According to the cultural perspective, market orientation creates a deeply rooted customer value system among all employees and is a potential source of competitive advantage (Hunt and Morgan, 1995). Others suggest that market orientation is a behavioral concept that is largely a matter of choice and resource allocation (Ruekert, 1992). Therefore, from an ontological standpoint, researchers can measure market orientation reflectively (cultural perspective) or formatively (behavioral perspective). The market orientation literature uncritically assumes the reflective view.

Although both cultural and behavioral definitions of market orientation are used in the literature, the measures of market orientation are largely couched in terms of behaviors. For example, Narver and Slater (1990, pp. 20–21) define market orientation as “the business culture that most effectively and efficiently creates superior value for customers.” Yet, they measure market orientation through behavioral items relating to customer orientation, competitor orientation and inter-functional coordination (Langerak, 2003). Arguably, adding or removing any of these components would change the conceptual interpretation of the construct, again implying a formative model is more appropriate.

Consideration 2: direction of causality. Virtually all the published literature in marketing, measures market orientation through three highly cited scales that have subsequently been synthesized into the MORTN summary scale (Deshpande and Farley, 1998). Close examination of the items contained in these scales reveals that the items are based on activities or behaviors that make up the construct. Hence, conceptual justification would imply that the direction of causality is from the indicator to the construct and not the other way around.

Consideration 3: characteristics of indicators. Lastly, all ten indicators in the MORTN scale are concerned with a customer's expressed needs, implying that the construct is one-dimensional and conceived as a reaction to these needs. Yet, no attention is given to intelligence-related items that support a proactive market orientation. The lack of emphasis currently given to proactive market orientation is problematic, given the growing evidence that industry and customer foresight are probably the most important components of market orientation (Hamel and Prahalad, 1994). Indeed, Narver et al. (2004) argue that much of the criticism surrounding market orientation is due to confusion surrounding the meaning of the term and, consequently, the way market orientation is measured. The solution, they argue, is to divide market orientation into reactive and proactive components. Others express related concerns about the way market orientation is measured and recommend examination of the construct's dimensionality (Siguaw and Diamantopoulos, 1995) or encourage modifications to the published scales (Rossiter, 2002).

Hence, based on the three theoretical considerations in the proposed framework—the nature of the construct, the direction of causality and the characteristics of the items used to represent the construct—it appears that market orientation is best conceptualized and measured using a formative model. To support this conclusion, we conduct a number of empirical tests.

4.2. Empirical considerations

To address the issue of whether market orientation is more validly measured through formative or reflective models, we analyze responses from a survey of senior executives. This sample is different to the first application and comprises 90 respondents. The questionnaire includes eight indicators of market orientation drawn from a literature review of reactive and proactive market-oriented scales.

Consideration 4: indicator intercorrelation. We perform a range of preliminary analyses, as for the first application. These analyses identify two separate dimensions for reactive and proactive market orientation, supporting Narver et al. (2004). These dimensions are largely independent of each other, as demonstrated by low intercorrelations in oblique rotations. The association between the indicators and these two dimensions is shown in Table 4. Given two constructs, the directionality and strength of these indicators largely fit expectations. However, the relationship of one indicator from the literature—“working with lead users”—is unclear; it relates fairly equally with both dimensions in all analyses. Diagnostics for the common factor model, although better than for the first application, are again not high enough to provide support for the reflective model. Overall, these analyses support the theoretical considerations by suggesting two constructs measured formatively rather than one measured reflectively.

[white space intentionally left blank]

Table 4: Market Orientation. Dimensionality and Association between Constructs and Indicators Suggested by Preliminary Analyses

No. Indicators	Constructs	
	Reactive Orientation	Proactive Orientation
1 Responsiveness to individual customer needs relative to competitors	√	
2 Ease to do business with relative to competitors	√	
3 Share customer experience across business relative to competitors	√	
4 Business driven by customer satisfaction relative to competitors		√
5 Predicting new market developments relative to competitors		√
6 Discovery of latent needs relative to competitors		√
7 Brainstorm customer usage relative to competitors		√
8 Work closely with lead users relative to competitors		unclear

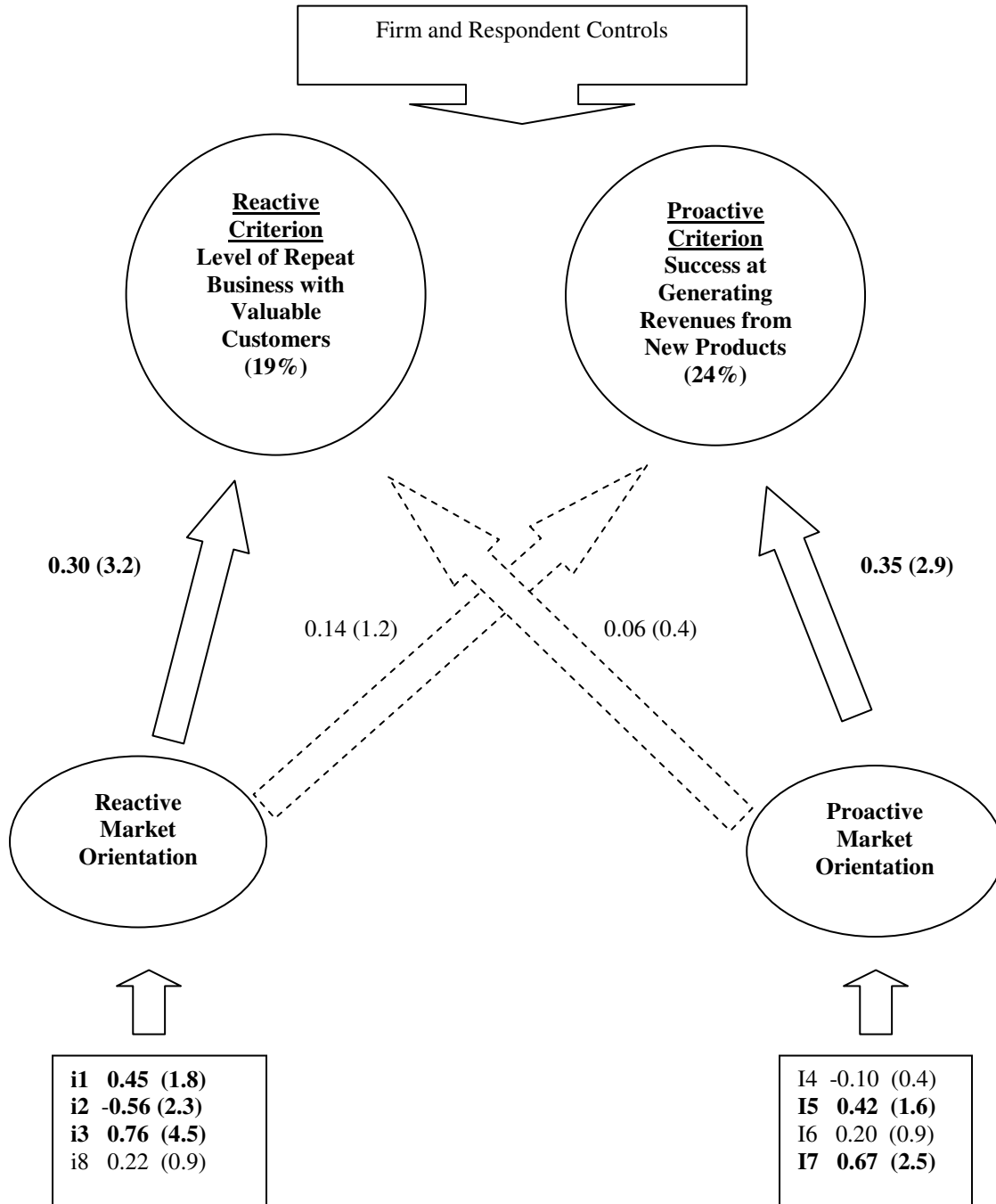
Consideration 5: indicator relationships with construct antecedents and consequences. We use a PLS model to assess criterion validity against two theoretically relevant and independent single-item constructs (see Figure 3). First, a high reactive market orientation should correlate significantly with the level of repeat business with valuable customers. We measure this independent construct through a single item on a 5-point Likert scale: “Compared to the highest performing business in your industry, the level of repeat business with valuable customers is far better to much worse.” This question is worded to ensure that respondents perceive it as a concrete, singular object. Hence, a single-item measure is entirely appropriate (Bergvist and Rossiter, 2007; Rossiter, 2002). The data are reverse scored for the analysis, where 5 = “far better”. Second, a high proactive market orientation should correlate significantly with success at generating revenue from new products. The study measures this revenue generating success with a similar question: “Compared to the highest performing business in your industry, the level of success generating revenue from new products is far better to much worse.”

In contrast, no significant correlation should be found between the reactive construct and the proactive criterion or between the proactive construct and the reactive criterion. This pattern of expected correlations between constructs and criterion questions provides a stronger test of the measurement model.

The results when the control variables are added—based on a formative measurement model—are shown in Figure 3. Only one control is significant, that for firm size. Firm size has a negative coefficient and explains 3 percent of the variance on the reactive criterion and 2 percent on the proactive criterion. Excluding this control, the market orientations themselves explain 16 percent of the reactive criterion and 22 percent of the proactive criterion.

[white space intentionally left blank]

Figure 3: Test of Criterion Validity for Market Orientation Measured Formatively



Note 1: Boxes contain outer indicator coefficients; inner path coefficients are next to arrows (with the absolute values of the bootstrap t-statistic in parentheses). All significant values are shown in bold type ($p < 0.05$). The percentage under each of the independent criteria is the R^2 . Indicator numbers, i1, i2, etc. refer to the measures shown in Table 4.

The results are as theoretically expected. The path from the reactive construct to the reactive criterion is both significant and meaningful ($\beta = 0.30$; $p < 0.01$), as is the path from the proactive construct to the proactive criterion ($\beta = 0.35$; $p < 0.01$). Also as expected, the crossover paths from each construct to the criterion for the other are not significant. However, three of the eight measurement indicators taken from the literature have insignificant weights. Thus a small number of indicators essentially drive the performance on the two criteria.

If the analysis is rerun assuming a reflective measurement model, the loadings on all eight indicators have significant t-statistics. Measurement error is not a problem here but the prediction of the reactive criterion is worse, with an R-square of 10 percent (excluding the variance explained by the single control) and a weaker path ($\beta = 0.24$ $p < 0.02$). However, the difference between the magnitude of this path coefficient in the reflective and formative models is not statistically significant. For the proactive criterion, the reflective and formative models have a similar performance, with the reflective model having an R-square of 19 percent and a similar path magnitude to that of the formative model ($\beta = 0.33$ $p < 0.01$).

The other comparison of relevance is with the one-dimensional, reflective model common in the literature. This results in a reflective measure with reasonable diagnostics (composite reliability of 0.86 and average variance extracted of 43%) that explains 9 percent of the reactive criterion and 17 percent of the proactive criterion (excluding controls). Both path magnitudes are significant ($\beta_{\text{reactive}} = 0.23$, $p < 0.02$ and $\beta_{\text{proactive}} = 0.41$, $p < 0.01$). Again, a fall in the predictive power of the reactive criterion is evident when compared with the formative model, but at this sample size, the difference is not significant.

Overall, the empirical results here are inconclusive and point toward the need for additional tests to support or reject a formative model structure. Unlike the first example of integration-

responsiveness pressures, both the formative and reflective models of market orientation are reasonably aligned with theoretical predictions on these specific tests.

Consideration 6: measurement error and collinearity. As in the first application, we undertake further assessment of the error structures using the vanishing tetrad test. The test results (see Table 5) reject the reflective model for both dimensions of market orientation (reactive at the 2% and proactive at the 10% level). Further investigation using bootstrapping shows that the 10 percent level for proactive market orientation is more likely a result of sample size limitations on the chi square test than the incorrect rejection of a reflective model. These results therefore imply that a formative model may be a better way of measuring both reactive and proactive market orientation. Again, collinearity is not an issue in these results as the largest condition indices from regressions of the two sets of indicators are 14.6 and 13.1, respectively.

Table 5: Market Orientation: Tetrad Test Results for Causal Indicators

Constructs	Number of Items	χ^2 (Df)	Df	Significance	Implication
Reactive Orientation	4	8.1	2	< 0.02	Formative
Proactive Orientation	4	4.7	2	< 0.10	Formative

The weight of evidence (both theoretical and empirical) largely supports the finding that market orientation is best represented by a two-dimensional construct measured formatively. The only qualification to this support is for consideration 5 where the formative and reflective measurement models both fit theoretical predictions for the criteria chosen here. The support for a two-dimensional construct measured formatively has important intellectual implications because virtually all the work conducted in marketing has viewed market orientation as a one-dimensional, reflectively measured construct. Both the theoretical and empirical work presented here indicate that current scales based on

one-dimensional reflective measures may not be completely valid, and also lend further support to those arguing for two separate constructs.

5. Discussion and Conclusions

Most researchers in the management sciences assume that the correct measurement model is a reflective one, whereas there are many instances in which this assumption may not be theoretically or empirically justified. This paper synthesizes previous work and presents an organizing framework for designing and testing measurement models based on both theoretical and empirical considerations derived from extant literature. The authors agree with Borsboom et al. (2004) and Rossiter (2002) that measurement models must be designed on theoretical considerations. However, we are also in agreement with the work of Bollen and Ting (2000), Diamantopoulos and Winklhofer (2001) and others who emphasize that empirical examination is required. As shown in the paper, once the data are collected, it is often useful to know if the assumptions underlying the measurement model hold empirically or not. Of course, it is possible that the reasons for empirical disconfirmation may be due to incorrect instrument design or mistaken responses by the respondents. Another possibility is that the theory underlying the measurement model is incorrect. Since empirical validation is accepted as a norm to validate structural model hypotheses, the same should apply to test the hypotheses about measurement models.

Next, the proposed framework is illustrated through its application to two important concepts in management, integration-responsiveness pressures and market orientation. In both cases, the appropriateness of a formative model over a reflective one is justified. In some cases, in personality and attitude measurement, for example, a reflective model is obvious. In other cases, a formative model is understandable, for example, in a human development index or an index of economic freedom for countries. However, it is not uncommon to encounter situations in social sciences where individual interpretation can lead to ambiguous results, especially when the construct definition and/or

nomenclature are inconsistent. For example, the construct of marketing mix adaptation is measured formatively when viewed by the researcher as a composite comprising adaptation of the various elements of the marketing mix. However, the construct of propensity to adapt the marketing mix is measured reflectively as it drives the degree to which the various elements of the marketing mix are adapted by a firm. Depending on the interpretation given to “mix adaptation” by the researcher, either measurement model is appropriate. In the case of the two applications in this paper, both theoretical and empirical considerations suggest that formative models are more plausible than reflective ones. This claim is not definitive, but simply offers an alternative lens for viewing and operationalizing these two important constructs.

A potential limitation of this study is that the indicators chosen from the literature for our questionnaire items are based on the reflective tradition. However, a counter-argument is that such items represent a conservative test of the proposition that formative measurement is worth considering. Indicators developed especially for formative measurement ought to perform better than those used here. This suggests one area where further research is needed: namely, better procedures for the design of formative indicators. Another area for research is the development of statistical techniques for assessing the appropriateness of formative versus reflective models. The tetrad test aside, the academic world is split between covariance and partial least squares model testing, each of which has strengths but few complementarities that help researchers to apply the empirical tests suggested here.

The main contribution of this paper is to show the need for researchers to explicitly justify their choice of reflective or formative measurement models by providing the supporting theoretical arguments and empirical corroboration. Uncritical and universal application of a reflective structure to oversimplify the measurement of broad, diverse and complex real-world constructs such as integration-responsiveness pressures and market orientation exposes scholars to the risk of reducing the rigor of business theory and research and its relevance for managerial decision making.

References

- Agarwal S. Erramilli KM. Dev CD. Market orientation and performance in service firms: role of innovation. *Journal of Services Marketing* 2003; 17(1): 68-82.
- Anderson JC. Gerbing DW. Structural equation modeling in practice: a review and recommended two-step approach. *Psychological Bulletin* 1988; 103(3): 411-423.
- Baumgartner H. Homburg C. Applications of structural equation modeling in marketing and consumer research: a review. *International Journal of Research in Marketing* 1996; 13(2): 139-161.
- Bearden WO. Netmeyer RG. *Handbook of Marketing Scales*. Thousand Oaks: Sage, 1999.
- Bergvist L. Rossiter JR. The predictive validity of multi-item versus single-item measures of the same constructs. *Journal of Marketing Research* 2007; 44(22): 175-184.
- Blalock HM. *Causal Inferences in Nonexperimental Research*. Chapel Hill, NC: University of North Carolina Press, 1964.
- Blalock HM. *Conceptualization and Measurement in the Social Sciences*. Beverly Hills, California: Sage, 1982.
- Bollen KA. Lennox R. Conventional wisdom in measurement: a structural equation perspective. *Psychological Bulletin* 1991; 110(2): 305-314.
- Bollen KA. Ting KF. A tetrad test for causal indicators. *Psychological Methods* 2000; 5(1): 3-22.
- Borsboom D. Mellenbergh GJ. Heerden JV. The theoretical status of latent variables. *Psychological Review* 2003; 110(2): 203-219.
- Borsboom D. Mellenbergh GJ. Heerden JV. The concept of validity. *Psychological Review* 2004; 111(4):1061-1071.
- Bruner II GCB. James KE. Hensel PJ. *Marketing Scales Handbook*. Chicago: American Marketing Association, 2001.
- Chin WW. Issues and opinion on structural equation modeling. *MIS Quarterly* 1998; 22(1): vii-xvi.

- Chin WW. Todd PA. On the use, usefulness, and ease of use of structural equation modeling in MIS research: a note of caution. *MIS Quarterly* 1995; 26(2/4): 237-246.
- Churchill GA. A paradigm for developing better measures of marketing constructs. *Journal of Marketing Research* 1979; 16(Feb):64-73.
- Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951; 16(3): 297-334.
- Deshpande R. Farley J. Measuring market orientation: generalization and synthesis. *Journal of Market-Focused Management* 1998; 2(3): 213-232.
- DeVillis RF. *Scale Development: Theories and Applications*. Newbury Park, California: Sage, 1991.
- Diamantopoulos A. The C-OAR-SE procedure for scale development in marketing: a comment. *International Journal of Research in Marketing* 2005; 22(1): 1-9.
- Diamantopoulos A. The error term in formative measurement models: interpretation and modeling implications. *Journal of Modeling in Management* 2006; 1(1):7-17.
- Diamantopoulos A. Siguaw JA. Formative versus reflective indicators in organizational measure development: a comparison and empirical illustration. *British Journal of Management* 2006; 17(4): 263-282.
- Diamantopoulos A. Winklhofer HM. Index construction with formative indicators: an alternative to scale development. *Journal of Marketing Research* 2001; 38(5): 269-277.
- Dunning JH. *Explaining International Production*. London: Unwin Hyman, 1988.
- Edwards JR. Multidimensional constructs in organizational behavior research: an integrative analytical framework. *Organizational Research Methods* 2001; 4(2): 144-192.
- Edwards J. Bagozzi R. On the nature and direction of relationships between constructs and measures. *Psychological Methods* 2000; 5(2): 155-174.
- Finn A. Kayande U. How fine is C-OAR-SE? A generalizability theory perspective on Rossiter's procedure. *International Journal of Research in Marketing* 2005; 22(1): 11-21.

- Fornell C. *A Second Generation of Multivariate Analysis*. New York: Praeger, 1982.
- Grewal R. Tansuhaj P. Building organizational capabilities for managing economic crisis: the role of market orientation and strategic flexibility. *Journal of Marketing* 2001; 65(4): 67-80.
- Hamel G. Prahalad CK. *Competing for the Future*. Boston, MA: Harvard Business School, 1994.
- Hunt SD. Morgan RM. The comparative advantage theory of competition. *Journal of Marketing* 1995; 59(2): 1-15.
- Jarvis CB. Mackenzie SB. Podsakoff PM. A critical review of construct indicators and measurement model misspecification in marketing and consumer research. *Journal of Consumer Research* 2003; 30(3): 199-218.
- Kohli AK. Jaworski BJ. Market orientation: the construct, research propositions, and managerial implications. *Journal of Marketing* 1990; 54(4): 1-8.
- Langerak F. An appraisal of research on the predictive power of market orientation. *European Management Journal* 2003; 21(4): 447-464.
- Mason K. Harris LC. Pitfalls in evaluating market orientation: an exploration of executives' interpretations. *Long Range Planning* 2005; 38(4): 373-391.
- Meehl PE. Why summaries of research on psychological theories are often uninterpretable. *Psychological Reports* 1990; 66: 195-244.
- Narver JC. Slater SF. The effect of a market orientation on business profitability. *Journal of Marketing* 1990; 54(10): 20-35.
- Narver JC. Slater SF. MacLachlan DL. Responsive and proactive market orientation and new product success. *The Journal of Product Innovation Management* 2004; 21(5): 334-347.
- Netmeyer RG. Bearden WO. Sharma S. *Scaling Procedures: Issues and Applications*. Thousand Oaks: Sage, 2003.
- Nunnally JC. Bernstein IH. *Psychometric Theory*. New York: McGraw-Hill, 1994.

- Prahalad CK. Doz Y. *The Multinational Mission: Balancing Local Demand and Global Vision*. New York: Free Press, 1987.
- Rossiter JR. The C-OAR-SE procedure for scale development in marketing. *International Journal of Research in Marketing* 2002; 19(4): 1-31.
- Rossiter JR. Reminder: a horse is a horse. *International Journal of Research in Marketing* 2005; 22(1): 23-25.
- Ruekert RW. Developing a market orientation: an organizational strategy perspective. *International Journal of Research in Marketing* 1992; 9(3): 225-245.
- Shook CL. Ketchen DJ Jr. Hult TMG. Kacmar MK. An assessment of the use of structural equation modeling in strategic management research. *Strategic Management Journal* 2004; 25(4): 397-404.
- Siguaw JA. Diamantopoulos A. Measuring market orientation: some evidence on Narver and Slater's three-component scale. *Journal of Strategic Marketing* 1995; 3: 77-88.
- Spearman C. General intelligence objectively determined and measured. *American Journal of Psychology* 1904; 15(2): 201-292.
- Spearman C. Holzinger K. The sampling error in the theory of two factors. *British Journal of Psychology* 1924; 15: 17-19.
- Spector PE. *Summated Rating Scale Construction*. Newbury Park: Sage, 1992.
- Trochim WMK. <http://www.socialresearchmethods.net/kb/index.php>. Visited on 19th July 2007.
- UNDP. *Human Development Report*. New York: Palgrave, 2006.
- Venaik S. Midgley DF. Devinney TM. A new perspective on the integration-responsiveness pressures confronting multinational firms. *Management International Review* 2004; 44(1, Special Issue): 15-48.