Formative versus Reflective Measurement Models: Two Applications of Erroneous Measurement

Tim R. Coltman[a], Timothy M. Devinney[b], David F. Midgley [c,]*, Sunil Venaik[d]

[a]Centre for Business Services Science, University of Wollongong, Australia
[b]Australian Graduate School of Management, University of NSW, Australia
[c]INSEAD, France
[d]School of Business, University of Queensland, Australia

***Corresponding author:**

David F. Midgley, Marketing Area, INSEAD, Boulevard de Constance, 77305 Fontainebleau, France. +33 1 60 72 42 78 (phone) +33 1 60 74 55 00 (fax)

david.midgley@insead.edu

The views expressed here are solely those of the authors, who are listed in alphabetical order.

ABSTRACT

An organizing framework is presented which assists researchers in the design and validation of formative and reflective measurement models. This framework is drawn from the extant literature and includes both theoretical and empirical considerations. The framework is then applied to 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 which most scholars have previously measured *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 Erroneous Measurement

1. Introduction

Management scholars often identify structural relationships among latent constructs by statistically relating covariation between the latent constructs and the observed variables or indicators used to measure these latent, unobserved constructs (Borsboom et al. 2003). 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 commonly this relationship between construct and indicator is assumed to be *reflective.* That is, the change in X is a reflection of (determined by) the change in the latent construct Y. With reflective (or *effect*) measurement models causality flows from the latent construct to the indicators.

However, not all latent constructs are entities that can be measured with a battery of positively correlated items, as is typically assumed with *reflective* indicators (Bollen and Lennox 1991; Edwards and Bagozzi 2000; Fornell 1982). It is equally plausible to define a construct as being determined by (or *formed)* from a number of indicators without any assumptions as to the patterns of inter-correlation between these items. This is termed a *formative* or *causal index* (Blalock 1964; Diamantopoulos and Winklhofer 2001; Edwards and Bagozzi 2000). Here causality flows in the opposite direction, namely from the indicator to the construct. Although the reflective view dominates in 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 before meaning can be assigned to the relationships implied in the structural model (Anderson and Gerbing 1988). Although, theoretical work in construct validity (Blalock 1982; DeVillis 1991; Edwards and Bagozzi 2000) and structural equation modeling (Baumgartner and Homberg 1996; Chin and Todd 1995; Shook et al. 2004) has enhanced our understanding of complex constructs, there is still considerable debate regarding the procedures a working

researcher should follow to achieve construct validity (see, for example, Rossiter 2005; Diamantopoulos 2005; Finn and Kayande 2005). It is not the purpose of this paper 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 proposed causal direction between constructs and their measures. The framework also builds on the work of Jarvis et al (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 specific empirical 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 the assumption of reflective measurement is tested 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 examples are chosen: (1) because of the predominance of the reflective modeling approach for these constructs even though a formative model can be theoretically argued to be 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, there is little reason to believe that the diverse measures of each of the integration and responsiveness pressures will be highly inter-correlated, as is required by a reflective structure. A priori, a formative approach to measurement would seem worthy of consideration, yet most of the published work in this area takes the reflective stance to measurement, often without any consideration of alternatives (Venaik et al. 2004). Similarly the vast majority of research on market orientation has defined it as a one-dimensional construct measured through a multi-

item reflective scale. Yet, the main scales used to measure market orientation—MARKOR (Kohli and Jaworski 1990) and MORTN (Deshpande and Farley 1998)—are conceptualized as a set of activities that compose the attribute (see Narver and Slater 1990, p.21), implying a formative model. Furthermore, the substantive inconsistencies reported in the market orientation literature (Langerak 2003) raise many questions about the dimensionality (Siguaw and Diamantopoulos 1995) and measurement (Narver et al. 2004) of the market orientation construct. These examples serve to illustrate a problem in the international business and marketing literature where insufficient attention has been 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 based on both theoretical and empirical considerations. Next, this framework is applied to the two illustrative and important examples taken, respectively, from international business and marketing. It should be noted that the purpose here is to examine whether reflective or formative measurement models are more or less appropriate in these fields, not to debate the content validity of the measures adopted by various scholars.

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) used 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. His five-step C-OAR-SE procedure offers an important extension to "accepted" practice by reemphasizing the importance of theoretical considerations. Common to both Borsboom and Rossiter, is the suggestion that scholars resist the temptation to conduct empirical tests and focus on theoretical considerations only.

Alternatively, Diamantopoulos (2005) and Finn and Kayande (2005) argue that both theoretical and empirical criteria are necessary to design and validate measurement models. According to these scholars it is important to conduct empirical analyses even if one has followed a rigorous assessment of content validity, especially to detect any errors, misspecifications or wrongly conceived theories. In their view researchers should be concerned if, for example, a negative relationship is found when theory and common sense suggest a positive relationship. As Judd et al. (1991 pp. 56–57) state "validity is demonstrated when the empirical relationship observed with a measure match the theoretically postulated nomological net of the construct."

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

==========

INSERT TABLE 1 ABOUT HERE

==========

2.1 Theoretical considerations

There are three broad theoretical considerations in deciding whether the measurement model should be formative or reflective. These are: (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 of 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 used (Borsboom et al. 2004; Rossiter 2002). Typical

examples of reflective scenarios include measures of attitudes and personality. These are traits that exist in individuals, which are traditionally measured by eliciting responses to indicators that are evoked by these constructs. Practically all published scales in business and related methodological texts on scale development (Bearden and Netmeyer 1999; Bruner II et al. 2001; Netmeyer et al. 2003; Spector 1992) are based on reflective measurement. For example, inspection of papers published in the *Journal of International Business Studies* and the *Journal of Marketing* in 2006 reveals that nearly 95% 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 based on three dimensions of human development: 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 the causality flows from the construct to the indicators; whereas in the case of formative models, it is the reverse, the causality flows from the indicators to the construct. Hence, in reflective models a change in the construct is assumed to cause 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 shown 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 how these models are estimated; topics discussed in section 2.2.

==========

INSERT FIGURE 1 ABOUT HERE

==========

Consideration 3. Characteristics of indicators: there are significant differences in the characteristics of the indicators used to 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 can be interchanged with other indicators that are likely to be elicited by the same construct. Due to this interchangeability, it is adequate 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 represented by the construct is sensitive to the number and types of indicators selected to represent the construct. Hence, adding or removing an indicator can change the conceptual domain of the construct. However, as pointed out by Rossiter (2002), this does not mean that we need a census of indicators as suggested by Bollen and Lennox (1991). As long as the indicators selected 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 there are also three empirical considerations that should inform our 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 of 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. This is unlike the formative model, where the indicators do not necessarily share the same theme and hence have no preconceived

pattern of intercorrelation. Indicators can theoretically possess no intercorrelation or high or low intercorrelation.

That being said, researchers should check that indicator intercorrelations are as they expect. Such checks can be conducted as part of the various preliminary analyses that are necessary whenever constructs are measured through 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 as hypothesized by the researcher (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 inter-correlation 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. Stronger tests are needed for that.

Since reflective indicators are expected to have positive intercorrelations, their individual and composite reliabilities can also be assessed empirically with the help of measures such as factor loading and communality, Cronbach alpha, average variance extracted and internal consistency (Trochim, 2006). However, as these measures of reliability are based on the assumption of internal consistency—i.e., high intercorrelations among the indicators in question—they are inappropriate for formative indicators, where there is no theoretical assumption made about inter-item correlation. This is one of the key operational issues in the use of formative indicators: *there are no simple, easy and universally accepted criteria for assessing the reliability of formative indicators.*

Consideration 5. Indicator relationships with construct antecedents and consequences: in the case of reflective models, it is required that the indicators have a similar (positive/negative, significant/non-

significant) relationship with the antecedents and consequences of the construct being measured. This is not the case for formative indicators as they do not necessarily share a common theme and, therefore, cannot be expected to have the same types of linkages with the antecedents and consequences of the construct being created. This is a significant issue in the use of formative models, particularly as it has implications about the appropriate level of aggregation of formative indicators. Whereas 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. Such a single-item criterion approach is also applied by Diamantopoulos and Siguaw (2006) in comparing formative and reflective models of 'export coordination.' This single-item approach is used for the second example in this paper (market orientation). Second, one applies a Multiple Indicators and Multiple Causes (MIMIC) model, where the construct is measured through both formative and reflective indicators. Third, one applies a structural model linking the formatively measured construct with another construct with which it is theoretically expected to relate and which is measured with reflective items. This approach establishes criterion and nomological validity and is the one followed in the first example in this paper (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, measurement error for each indicator can be identified and eliminated using common factor analysis because the factor score contains only that part of the indicator that it shares 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 the data is collected. Diamantopoulos (2006) suggests two possible ways to eliminate the error term (thus ensure that $\zeta \cong 0$) through design: (1) capture all possible causes on the construct, and (2) specify the focal construct in such a way as to ensure that the full set of indicators is captured. In both cases the error term has been legitimately excluded. Given our discussion it is clear that unlike the reflective model, there is no simple way to empirically assess the impact of measurement error in a formative model.

However, Bollen and Ting (2000) have suggested that the vanishing tetrad test—a procedure based on covariance algebra—can provide some assistance for the assessment of measurement error in formative models. Derived from the work by Spearman and Holzinger (1924) the test is based on 'nested' vanishing tetrads that are implied by comparing two theoretical measurement models. More specifically, a tetrad refers to the difference between the products of two pairs of error covariances. In the case of a reflective model the null hypothesis is that the set of non-overlapping tetrads vanishes. More simply, 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 vanishing tetrad test confirms whether or not this is true.

The tetrad test is a confirmatory procedure that was never meant to be a standalone criterion for distinguishing formative from reflective models. For example, if the hypothesis that the errors are uncorrelated is rejected, it can be for 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. It is possible,

although unlikely in practice, that a formative model is correct but that the indicator error structures are uncorrelated. Thus the results from the vanishing tetrad test do not provide definitive proof as to the correct measurement model but serve as an important pointer as to whether the measurement error meets the criteria required by the assumed model.

Another measurement issue that needs to be checked in the case of a formative model 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 regressing the indicators on this variable and computing standard diagnostics such as the condition index.

In the next section, this framework, with its three sets of theoretical criteria and three sets of empirical criteria, is 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) has been 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. In an extensive review of the literature, it was found that nearly all researchers have taken the reflective route and only a handful used a formative model—but more critically that there has been little published debate, justification or validation for the choice each researcher has made (Venaik et al. 2004). 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, diversity and complexity. Researchers in this field have characterized them as global integration pressures—global competition, the need to reduce costs, and the pressures of technological change and complexity, etc.—and local responsiveness pressures—diversity of market infrastructure, country based regulation, local customer heterogeneity, etc. Conceptually, this categorization of pressures has been developed and defined by international business scholars, with an emerging consensus in the field on how they should be labeled. Given this, it is difficult to think of them as being innate characteristics of the business environment that cause either global integration or local responsiveness.

Consideration 2: direction of causality. It is logical to view the diverse facets of the environment as forming the IR pressures rather than the other way around. Indeed, the very word 'pressures' itself 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 are defined as the pressures, rather than the pressures causing the measures to change in concert. 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 standard 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,' particularly across such diverse industries as software, capital

equipment, financial services and food products. 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, a number of empirical tests are applied to corroborate the suitability of a formative model for this application.

3.2. Empirical considerations

Based on a comprehensive survey of the IR literature, 23 indicators of IR pressures were identified and administered via a questionnaire to a sample of 163 managers belonging to the subsidiaries of multinational firms in 35 countries (Venaik et al. 2004). Three statistical procedures that follow from the empirical considerations set out earlier were used to corroborate the appropriateness of a formative model.

Consideration 4: indicator intercorrelation. As discussed above a range of preliminary analyses were conducted 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 the data used here, 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 five pressures, the directionality and strength of the indicators also fits 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.

==========

INSERT TABLE 2 ABOUT HERE

==========

Consideration 5: indicator relationships with construct antecedents and consequences. The five formatively measured pressures were used to predict the independent reflectively measured construct of subsidiary *Autonomy* (which was demonstrated to be one-dimensional and measured with a composite reliability of 0.90 and average variance extracted of 61%). This additional construct of *Autonomy* was used as the criterion construct to identify 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. A series of control variables were also included to provide greater confidence that any observed effects were not spurious results of industry and firm heterogeneity. The technique of partial least squares (PLS) was used for this analysis (Chin 1998). Figure 2 shows the results obtained. These results add further support to the formative model as the five pressures predict *Autonomy* reasonably 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*.

==========

INSERT FIGURE 2 ABOUT HERE

==========

However, it is difficult to judge a structural equation model in isolation. Hence, for this paper the five pressures were also analyzed as <u>reflectively</u> measured, by rerunning the PLS analysis with indicator directionality reversed. This 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 performed much worse than the formative

one. The reflectively measured pressures explained 17% of the variance in *Autonomy* compared with 23% for the formative model. (The total variance explained, including firm and industry controls, was 22% and 26% respectively). Examination of the item coefficients showed that this difference in performance was not due to poor measurement—for the reflective model all the item loadings and t-values were high, for the formative model all pressures had an adequate number of significant weights. Instead, the difference was attributable to the reflectively measured pressures not explaining the independent construct as well as did the formatively measured ones. Indeed, only <u>one</u> of the five reflectively measured pressures had a significant and meaningful path coefficient with *Autonomy* (where meaningful is $\beta > 0.20$, Meehl 1990), namely *global competiton* 3 ($\beta = -0.45$, $p < 0.01$), whereas <u>three</u> of the formatively measured pressures had 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. Many scholars would 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 specification and measurement model commonly accepted in the literature: e.g., 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% (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 empirical results support the theoretical arguments that formative measurement may be more appropriate.

Consideration 6: measurement error and collinearity. To corroborate the above discussion the vanishing tetrad test was applied to each construct. This test rejected 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 did 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).

==========

INSERT TABLE 3 ABOUT HERE

==========

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 there is no *prima facie* rationale 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 each other. The second three empirical considerations and

statistical analyses, together with tetrad tests, lend further support to the formative measurement model. Next, the same six considerations are applied 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. It stipulates that organizations should allocate resources to the systematic gathering and analysis of customer and competitor information and to make use of this knowledge in guiding 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 academic literature has suffered from diverse and inconsistent measures (Mason and Harris 2005).

A closer look at the empirical evidence indicates that the power of market orientation to predict advantage or performance is still an open question (Langerak 2003). For example, Agarwal et al. (2003) report no direct relationship, while Grewal and Tansuhaj (2001) show mixed results. These inconsistent findings imply that either the theory underpinning market orientation does not hold generally or that the measurement model used to operationalize the construct is incorrect. This paper will suggest the latter is arguably one cause of these inconsistent results.

4.1. Theoretical considerations

Consideration 1: nature of the construct. Market orientation has been conceptualized based on either cultural or behavioral criteria. According to the cultural perspective, market orientation is a potential source of competitive advantage because it creates a deeply rooted customer value system among all employees (Hunt and Morgan 1995). Others have suggested that market orientation is a behavioral concept that is largely a matter of choice and resource allocation (Ruekert 1992). Therefore, from an ontological standpoint, market orientation could be measured reflectively (cultural perspective) or

formatively (behavioral perspective). The literature on market orientation has uncritically assumed the reflective view.

It is also important to note that although both cultural and behavioral definitions of market orientation have been used in the literature, the measures of market orientation have largely been couched in terms of behaviors. For example, Narver and Slater (1990, p.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 co-ordination (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 has sought to measure 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 they are based on activities or behaviors that compose 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 component of market orientation (Hamel and Prahalad 1994). Indeed, Narver et al. (2004) have argued that much of the criticism surrounding market orientation has resulted from confusion surrounding the meaning of the term and, consequently, the way it is measured. The solution they argue is to divide market orientation into reactive and proactive components. Others have expressed related concerns with

the way market orientation is measured and recommended examination of its dimensionality (Siguaw and Diamantopoulos 1995) or encouraged 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 a number of empirical tests are used to further assess the suitability of a formative model.

4.2. Empirical considerations

To address the issue of whether market orientation is more validly measured through formative or reflective models, responses from a survey of senior executives were analyzed. This is a different sample to the first application and numbers 90 respondents. Eight indicators of market orientation drawn from a literature review of reactive and proactive market oriented scales were included in the questionnaire.

Consideration 4: indicator intercorrelation. A range of preliminary analyses were performed 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 fits expectations. However, the relationship of one indicator from the literature—'working with lead users'—is unclear; it related fairly equally with both dimensions in all analyses. Diagnostics for the common factor model, although better than for the first application, were again not high enough to provide support for the reflective model. Overall, the initial analyses support the theoretical considerations by tentatively suggesting two constructs measured formatively rather than one measured reflectively.

==========

INSERT TABLE 4 ABOUT HERE

==========

Consideration 5: indicator relationships with construct antecedents and consequences. A PLS model was used 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. This independent construct was measured 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 (Rossiter 2002; Bergvist and Rossiter 2007). The data were reverse scored for the analysis, where 5 = "far better." Second, a high pro-active market orientation should correlate significantly with success at generating revenue from new products. This was measured with a similar question: "Compared to the highest performing business in your industry, success with generating revenue from new products." In contrast there should be no significant correlation between the reactive construct and the proactive criterion or the proactive construct and the reactive criterion, providing a stronger test of the measurement model. Controls were added to the analysis to control for heterogeneity. These controls were for firm and respondent characteristics (but not industry given the question comparison is with a firm in the same industry). The results assuming a formative measurement model are as shown in Figure 3. Only one control is significant, that for firm size. Firm size has a negative coefficient and explains 3% of the variance on the reactive criteria and 2% on the proactive. Excluding this control, the market orientations themselves explain 16% of the reactive criterion and 22% of the proactive criterion.

==========

INSERT FIGURE 3 ABOUT HERE

==========

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 that 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 while three indicators for the reactive construct and two for the proactive 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. There is no measurement problem here but the prediction of the reactive criterion is worse, with an R-square of 10% (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. There is also little difference in performance on the proactive criterion where the reflective model has an R-square of 19% and a similar path magnitude to 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% of the reactive criterion and 17% of the proactive criterion (excluding controls). Both path magnitudes are significant ($\beta_{reactive} = 0.23$, $p < 0.02$ and $\beta_{proactive} = 0.41$, $p < 0.01$). Again there is a fall in prediction of the reactive criterion compared with the formative model, but at this sample size the difference is not significant. Although not changing the results materially, the diagnostics on this construct measure can be improved by pruning two items (raising average variance extracted to 52%).

Overall, the empirical results here are inconclusive and point towards 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, further assessment of the error structures was undertaken using the vanishing tetrad test. These test results are reported in Table 5 and reject the reflective model for both dimensions of market orientation (reactive at the 1% and proactive at the 10% level). Further investigation using bootstrapping shows that the 10%

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 both sets of indicators are 14.6 and 13.1 respectively, both 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).

==========

INSERT TABLE 5 ABOUT HERE

==========

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 Five 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 lend further support to those arguing for two separate constructs of market orientation measured with formative indicators.

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. In this paper, the authors synthesize previous work and present an organizing framework for designing and testing measurement models based on both theoretical and empirical considerations derived from extant literature. They agree with Borsboom et al. (2004) and Rossiter (2002) that measurement models must be designed on theoretical and conceptual considerations. However, the authors are also

aligned with the work of Bollen and Ting (2000), Diamantopoulos and Winklhofer (2001) and others who emphasize that empirical examination is also required. As the authors show, once the data is 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. But there is also a possibility, that the theory underlying the measurement model is incorrect. Since empirical validation is accepted as a norm to validate structural model hypotheses, the authors see little reason why the same should not apply to test the hypotheses about measurement models.

Next, the proposed framework was illustrated through its application to two important concepts in management, integration-responsiveness pressures and market orientation. In both cases there is justification for the belief that a formative model is more appropriate than a reflective one. In some cases, a reflective model is obvious, e.g., in personality and attitude measurement. In others, a formative model is understandable, e.g., in human development index or an index of economic freedom for countries. However, it is not uncommon to encounter situations in social science 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* may be measured formatively when viewed by the researcher as a composite comprising adaptation of the various elements of the marketing mix. However, the construct *propensity to adapt the marketing mix* can be measured reflectively as it drives the degree to which the various elements of the marketing mix will be adapted by a firm. Depending on the interpretation given to "mix adaptation" by the researcher either measurement model can be appropriate. In the case of the two applications in this paper, both theoretical and empirical considerations suggest that formative models might be 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 paper is that in choosing indicators from the literature for their data collection the authors have used questionnaire items developed in 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. That suggests one area where further research is needed: namely, better procedures for the design of formative indicators. Another is the development of statistical techniques for assessing the appropriateness of formative versus reflective models. 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.

Overall, 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 opens 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, Homberg 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 concept of validity. Psychological Review 2004;111(4):1061–1071.

Borsboom D, Mellenbergh GJ, Heerden JV. The theoretical status of latent variables. Psychological Review 2003;110(2):203–219.

Bruner II GCB, James KE, Hensel PJ. Marketing Scales Handbook. Chicago: American Marketing Association, 2001.

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.

Chin WW. Issues and opinion on structural equation modeling. MIS Quarterly 1998;22(1):vii–xvi.

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, 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. 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. 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: and 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.

Judd CM, Smith ER, Kidder LH. Research Methods in Social Relations. Ft Worth, Texas: Holt, Rinehart and Winston, 1991.

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, MacLachlan DL. Responsive and proactive market orientation and new product success. The Journal of Product Innovation Management 2004;21(5):334–347.

Narver JC, Slater SF. The effect of a market orientation on business profitability. Journal of Marketing 1990;54(10):20–35.

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. Reminder: a horse is a horse. International Journal of Research in Marketing 2005;22(1):23–25.

Rossiter JR. The C-OAR-SE procedure for scale development in marketing. International Journal of Research in Marketing 2002;19(4):1–31

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, Holzinger K. The sampling error in the theory of two factors. British Journal of Psychology 1924;15:17–19.

Spearman C. General intelligence objectively determined and measured. American Journal of Psychology 1904;15(2):201–292.

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.

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

Considerations	Reflective model	Formative model	Relevant literature
Theoretical Considerations			
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)
Empirical Considerations			
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	Bollen and Ting (2000); Diamantopoulos (2006)

- Collinearity should be ruled out by standard diagnostics such as the condition index

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

		Constructs				
No.	Indicators	Government influence	Quality of local infra-structure	Global competition	Techno-logical 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)

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

Constructs	Number of indicators	χ^2(Df)	Df	Significance	Implication
Government influence	6	22.39	9	0.008	Formative
Quality of local infrastructure	5	19.90	5	0.001	Formative
Global competition	5	20.46	5	0.001	Formative
Technological Change	4	9.84	2	0.007	Formative
Resource sharing	3*	1.04	2	0.593	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).

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

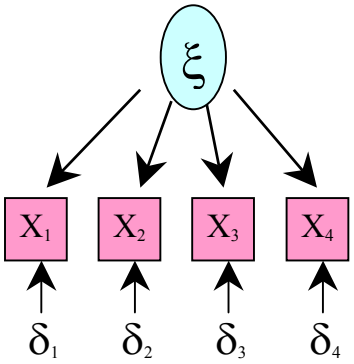
No.	Indicators	Constructs: Reactive Orientation	Constructs: Proactive Orientation
1	Responsiveness to individual customer needs relative to competitors	√	
2	Ease to do business with relative to competitors	√	
3	Share customer experience across business relative to competitors	√	
4	Business driven by 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	

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

Constructs	Number of Items	χ^2(Df)	Df	Significance	Implication
Reactive Orientation	4	8.08	2	0.018	Formative
Proactive Orientation	4	4.66	2	0.097	Formative

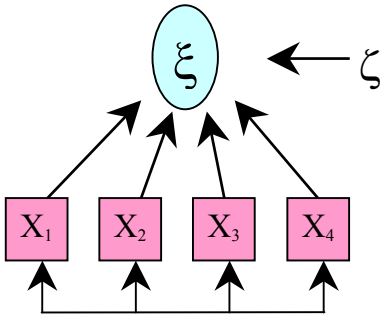
Figure 1: Reflective and Formative Measures

Effect Model (Reflective indicators)



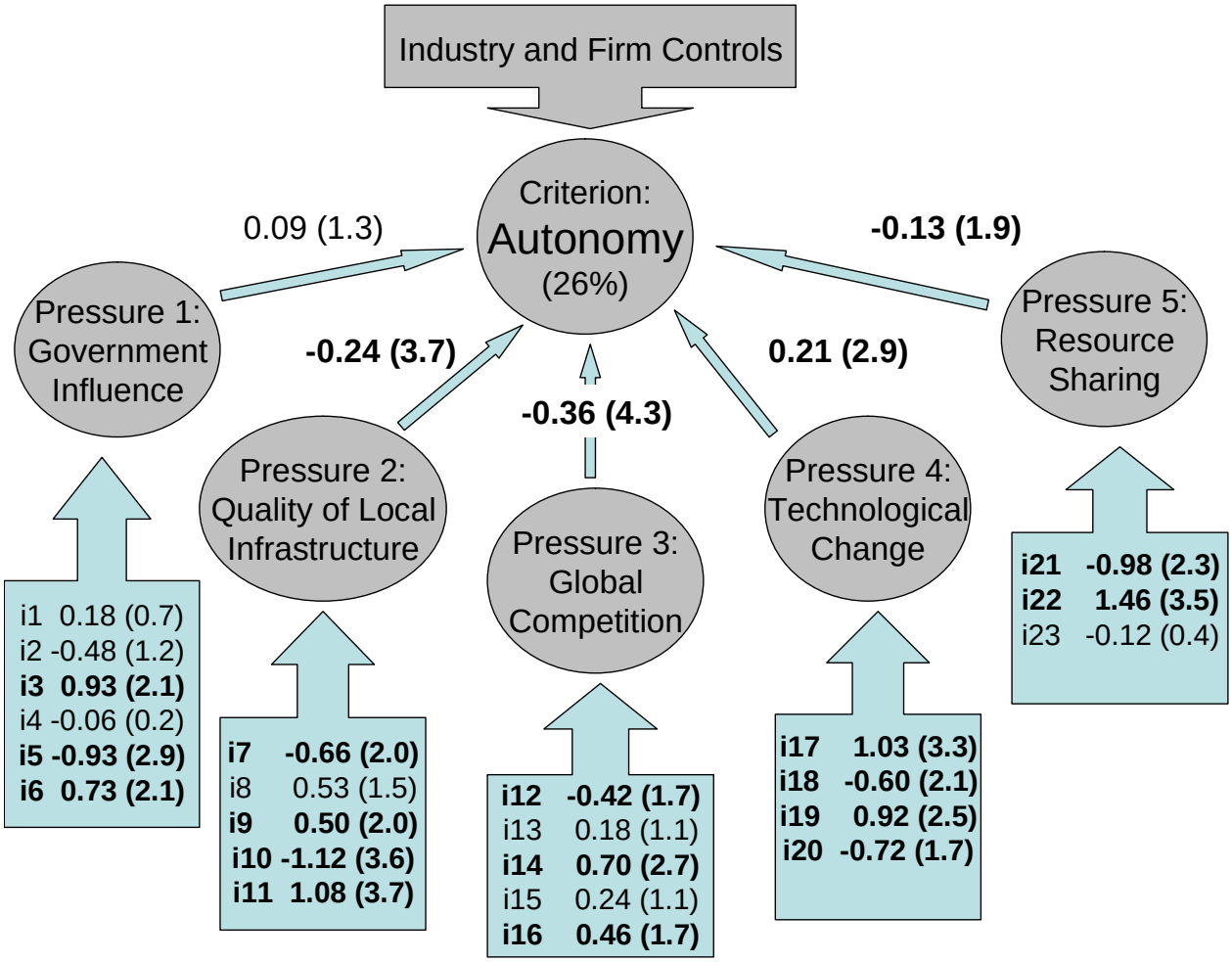
$$X_1 = \lambda_1\xi + \delta_1$$
$$X_2 = \lambda_2\xi + \delta_2$$
$$X_3 = \lambda_3\xi + \delta_3$$
$$X_4 = \lambda_4\xi + \delta_4$$

Causal Model (Formative indicators)



$$\xi = \gamma_1X_1 + \gamma_2X_2 + \gamma_3X_3 + \gamma_4X_4 + \zeta$$

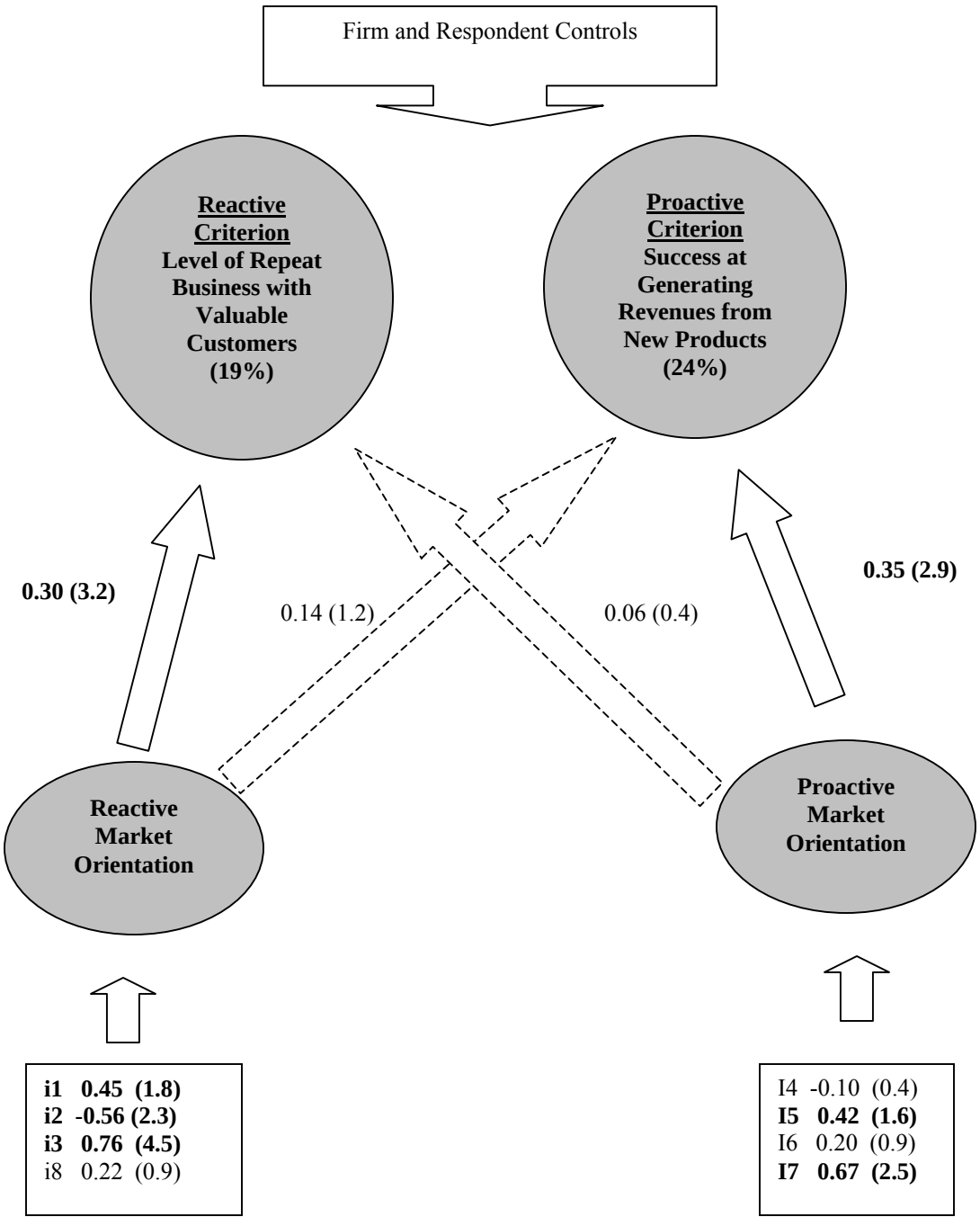
Figure 2: Test of Criterion Validity for IR Pressures Measured Formatively



Note: Boxes contain outer indicator coefficients; inner path coefficients are next to arrows (with the absolute values of the bootstrap t-statistic in parenthesis). 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)

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 parenthesis). 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.