

Formative Versus Reflective Indicators in Organizational Measure Development: A Comparison and Empirical Illustration

Adamantios Diamantopoulos and Judy A. Siguaw*

Institute of Business Administration, University of Vienna, Bruenner Straße 72, A-1210, Vienna, Austria and

*Cornell-Nanyang Institute of Hospitality Management, Nanyang Technological University, S3-01B-49

Nanyang Avenue, Singapore 639792, Republic of Singapore

Email: adamantios.diamantopoulos@univie.ac.at [Diamantopoulos]; judysiguaw@ntu.edu.sg [Siguaw]

A comparison is undertaken between scale development and index construction procedures to trace the implications of adopting a reflective versus formative perspective when creating multi-item measures for organizational research. Focusing on export coordination as an illustrative construct of interest, the results show that the choice of measurement perspective impacts on the content, parsimony and criterion validity of the derived coordination measures. Implications for practising researchers seeking to develop multi-item measures of organizational constructs are considered.

Latent variables are widely utilized by organizational researchers in studies of intra- and inter-organizational relationships (James and James, 1989; Scandura and Williams, 2000; Stone-Romero, Weaver and Glenar, 1995). In nearly all cases, these latent variables are measured using *reflective* (effect) indicators (e.g. Hogan and Martell, 1987; James and Jones, 1980; Morrison, 2002; Ramamoorthy and Flood, 2004; Sarros *et al.*, 2001; Schaubroeck and Lam, 2002; Subramani and Venkatraman, 2003; Tihanyi *et al.*, 2003). Thus, according to prevailing convention, indicators are seen as functions of the latent variable, whereby changes in the latent variable are reflected (i.e. manifested) in changes in the observable indicators. However, as MacCallum and Browne point out, 'in many cases, indicators could be viewed as causing rather than being caused by the latent variable measured by the indicators' (MacCallum and Browne, 1993, p. 533). In these instances, the indicators are known as *formative* (or causal); it is changes in the indicators that determine changes in the value of the latent variable rather than the other way

round (Jarvis, Mackenzie and Podsakoff, 2003). Formally, if η is a latent variable and $x_1, x_2, \dots, x_n$ a set of observable indicators, the reflective specification implies that $x_i = \lambda_i \eta + \varepsilon_i$, where λ_i is the expected effect of η on x_i and ε_i is the measurement error for the i th indicator ($i = 1, 2, \dots, n$). It is assumed that $\text{COV}(\eta, \varepsilon_i) = 0$, and $\text{COV}(\varepsilon_i, \varepsilon_j) = 0$, for $i \neq j$ and $E(\varepsilon_i) = 0$. In contrast, the formative specification implies that $\eta = \gamma_1 x_1 + \gamma_2 x_2 + \dots + \gamma_n x_n + \zeta$, where γ_i is the expected effect of x_i on η and ζ is a disturbance term, with $\text{COV}(x_i, \zeta) = 0$ and $E(\zeta) = 0$. For more details, see Bollen and Lennox (1991), Fornell, Rhee and Yi (1991) and Fornell and Cha (1994).

With few exceptions (e.g. Law and Wong, 1999; Law, Wong and Mobley, 1998), formative measures have been a somewhat ignored topic within the area of organizational research. Indeed, nearly all of the work that exists in the area of formative measurement has stemmed from researchers housed in sociology or psychology (e.g. Bollen, 1984; Bollen and Lennox, 1991; Bollen and Ting, 2000; Fayers and Hand, 1997; Fayers *et al.*, 1997; MacCallum and Browne,

1993), marketing (e.g., Diamantopoulos and Winklhofer, 2001; Fornell and Bookstein, 1982; Jarvis, Mackenzie and Podsakoff, 2003; Rossiter, 2002) and strategy (e.g. Fornell, Lorange and Roos, 1990; Hulland, 1999; Johansson and Yip, 1994; Venaik, Midgley and Devinney, 2004, 2005). This situation is unfortunate given that, in many cases, work utilizing formative measures may better inform organization theory, as illustrated herein.

The current study seeks to extend previous methodological work by Bollen and Lennox (1991), Law and Wong (1999), and Diamantopoulos and Winklhofer (2001) by tracing the *practical* implications of adopting a formative versus reflective measurement perspective when *developing* a multi-item organizational measure from a pool of items.¹ More specifically, we explore whether conventional scale development procedures (e.g. see Churchill, 1979; DeVellis, 2003; Netemeyer, Bearden and Sharma, 2003; Spector, 1992) and index construction approaches (e.g. Diamantopoulos and Winklhofer, 2001; Law and Wong, 1999) as applied in organizational coordination research lead to materially different multi-item measures in terms of (a) *content* (as captured by the number/proportion of common items included in the measures), (b) *parsimony* (as captured by the total number of items comprising the respective measures), and (c) *criterion validity* (as captured by the ability of a reflective scale versus that of a formative index to predict an external criterion, i.e. some outcome variable).²

None of these issues has been systematically addressed in previous methodological research. With regards to content, previous comparisons of reflective and formative measures have implicitly assumed that *exactly the same* set of indicators can be used to operationalize the construct involved

(e.g. Bollen and Ting, 2000; Law and Wong, 1999; MacCallum and Browne, 1993). For example, the recent Monte Carlo simulation of measurement model misspecification in marketing by Jarvis, Mackenzie and Podsakoff (2003) was based on a simple reversal of the directionality of the paths between constructs and their indicators. This study assumes that the *only* difference resulting from applying a formative versus reflective measurement approach relates to the causal priority between the construct and its indicators. However, this is an untested and, most likely, unwarranted assumption as it implies that despite their *very* different nature (see next section), scale development and index construction strategies will result in measures that contain an *identical* set of indicators. With regard to parsimony, a natural extension of the assumption made with regards to measure content is that formative indexes and reflective scales are equally parsimonious (if both types of measures are assumed to be comprised of the exactly same items, then the number of items *must* be the same in both cases). Again, this is a questionable assumption as it implies that the measure purification procedures associated with scale development and index construction respectively will result in the exclusion (viz. inclusion) of exactly the same *number* of items (although the *specific* items dropped from the measures need not be the same). Lastly, with regards to criterion validity, no previous study has *empirically* examined whether multi-item measures generated by scale development (reflective) and index construction (formative) approaches respectively, perform similarly in terms of their ability in predicting some outcome variable. While considerations of validity have featured in previous discussions of measurement model specification, such discussions have been purely of a conceptual nature (e.g. Bagozzi and Fornell, 1982; Diamantopoulos, 1999).³

In the following section, we provide some conceptual background to the problem of developing multi-item measures in organizational research and contrast scale development and index construction procedures in the specific context of organizational coordination. Next, we apply these procedures to empirical data and

¹Throughout this paper we use the (generic) term 'measure' to refer to a multi-item operationalization of a construct, and the terms 'index' and 'scale' to distinguished between measures comprised of formative and reflective items respectively. The terms 'items' and 'indicators' are used interchangeably.

²Criterion (or criterion-related) validity 'concerns the correlation between the measure and some criterion variable of interest' (Zeller and Carmines, 1980, p. 79). It is also known as 'empirical' (e.g. Nachmias and Nachmias, 1976) 'pragmatic' (Oppenheim, 1992) and 'predictive' validity (e.g. Nunnally and Bernstein, 1994).

³For a conceptual discussion of the nature of validity as well as the analytic tools that can be used to aid its assessment, see Carmines and Zeller (1979).

generate multi-item measures for the same focal construct (export coordination) using a reflective and a formative measurement perspective, respectively. We follow this by a comparison of the derived coordination measures in terms of content, parsimony and criterion validity and conclude the article with some thoughts aimed at assisting organizational researchers in the selection of their measure development strategies.

Developing multi-item measures: scales versus indices

Consider a scenario where a researcher wishes to develop a multi-item measure for a particular organizational construct, η for subsequent use in empirical research. In tackling this task, the researcher can follow one of two strategies, depending upon his/her conceptualization of the focal construct: (s)he can either treat the (unobservable) construct as *giving rise* to its (observable) indicators (Fornell and Bookstein, 1982), or view the indicators as *defining characteristics* of the construct (Rossiter, 2002). In the former case, measurement items would be viewed as reflective indicators of η and conventional scale development guidelines (e.g. Churchill, 1979; DeVellis, 2003; Netemeyer, Bearden and Sharma, 2003; Spector, 1992) would be followed to generate a multi-item measure. In the latter case, measurement items would be seen as formative indicators of η and index construction strategies (Diamantopoulos and Winklhofer, 2001) would be applicable. Note, in this context, that the measure development procedures associated with the two approaches are very different. Scale development places major emphasis on the intercorrelations among the items, focuses on common variance, and emphasizes unidimensionality and internal consistency (e.g. see Anderson and Gerbing, 1982; Churchill, 1979; DeVellis, 2003; Nunnally and Bernstein, 1994; Spector, 1992; Steenkamp and van Trijp, 1991). Index construction, on the other hand, focuses on explaining abstract (unobserved) variance, considers multicollinearity among the indicators and emphasizes the role of indicators as predictor rather than predicted variables (e.g. see Bollen, 1984; Diamantopoulos and Winklhofer, 2001; Law and Wong, 1999; MacCallum and Browne, 1993).

The choice of measurement perspective (and, hence, the use of reflective versus formative indicators) should be *theoretically* driven, that is, it should be based on the 'auxiliary theory' (Blalock, 1968; Costner, 1969) specifying 'the nature and direction of the relationship between constructs and measures' (Edwards and Bagozzi, 2000, p. 156). In many cases, this choice will be straightforward as the causal priority between the construct and the indicators is very clear. For example, 'constructs such as "personality" or "attitude" are typically viewed as underlying factors that *give rise* to something that is observed. Their indicators tend to be realized, then as reflective' (Fornell and Bookstein, 1982, p. 292, emphasis in the original). Similarly, constructs such as socio-economic status are typically conceived as *combinations* of education, income and occupation (Hauser, 1971, 1973) and, thus, their indicators should be formative; after all, 'people have high socio-economic status because they are wealthy and/or educated; they do not become wealthy or educated because they are of high socio-economic status' (Nunnally and Bernstein, 1994, p. 449).

In other instances, however, choosing correctly between a reflective and a formative measurement perspective may not be as easy as 'the directionality of the relationship is far from obvious' (Fayers *et al.*, 1997, p. 393). As Hulland observes, 'the choice between using formative or reflective indicators for a particular construct can at times be a difficult one to make' (1999, p. 201). In this context, several methodological studies (Cohen *et al.*, 1990; Diamantopoulos and Winklhofer, 2001; Jarvis, Mackenzie and Podsakoff, 2003; Rossiter, 2002) have provided examples of many constructs which have previously (and erroneously) been operationalized by means of reflective multi-item scales, although a formative perspective would have been theoretically appropriate. Similarly, Edwards and Bagozzi (2000) demonstrate that determining the nature and directionality of the relationship between a construct and a set of indicators can often be far from simple, while Bollen and Ting state that 'establishing the causal priority between a latent variable and its indicators can be difficult' (2000, p. 4). In short, it *cannot* be taken for granted that researchers will *always* make the correct choice when operationalizing constructs in organizational research efforts. This raises the question: What happens if they get it wrong?

		'Correct' Auxiliary Theory	
		Reflective	Formative
Researcher's Choice of Measurement Perspective	Reflective	Correct Decision	Type I Error
	Formative	Type II Error	Correct Decision

Figure 1. Choosing a measurement perspective

Errors in choosing a measurement perspective

Theoretically, we can distinguish between four possible outcomes when contemplating the choice of measurement perspective (Figure 1). Two of these are desirable and indicate that, *given the conceptualization of the construct in question*, the correct perspective has been adopted in its operationalization. The other two are clearly undesirable as they indicate that a wrong choice has been made. Specifically, a Type I error occurs when a reflective approach has been adopted by the researcher (and hence scale development procedures employed) although given the nature of the construct in question, the correct operationalization should have been formative (thus calling for index construction procedures). In contrast, a Type II error occurs when a formative specification has been chosen by the investigator, although a reflective approach would have been theoretically appropriate for the particular construct concerned.⁴

Looking at Figure 1, what is the likelihood of committing a Type I versus Type II error? Within the stream of organizational research, there is little doubt that a reflective perspective is by far the dominant approach (Law and Wong, 1999). In fact, there are very few instances in which the choice of a measurement perspective is *explicitly* defended in empirical studies. Indeed, some authors speak of an 'almost automatic acceptance of reflective indicators in the minds of researchers' (Diamantopoulos and Winklhofer, 2001, p. 274); moreover, according to a recent

meta-analysis, 'most of the errors in measurement model specification resulted from the use of a reflective measurement model for constructs that should have been formatively modelled' (Jarvis, Mackenzie and Podsakoff, 2003, p. 207).

In light of the above, we suspect that the probability of *erroneously* selecting a reflective perspective (and thus committing a Type I error) is currently much higher than the corresponding probability of erroneously opting for a formative perspective (Type II error). From a practical point of view, the question then becomes: given that a formative perspective *should* have been employed when developing the measure of interest, what are the implications of committing a Type I error?

To approach this problem conceptually, it is helpful to distinguish between three key stages in measure development: item generation, measure purification and measure validation. With regard to item generation, an obvious question is whether the *initial* pool of items generated under a reflective perspective would be very different from what would have been generated had a formative perspective been (correctly) adopted instead. Surprisingly, and despite the fundamental nature of this question, whether item *generation* under a reflective versus a formative perspective is likely to result in substantially different item pools has not as yet been explicitly addressed in past research, either conceptually or empirically.⁵

Insights into this issue, however, can be gained by looking at existing methodological guidelines for scale development (e.g. DeVellis, 2003; Spector, 1992) and index construction respectively (e.g. Bollen and Lennox, 1991; Diamantopoulos and Winklhofer, 2001). Such guidelines turn out to be remarkably similar, as *both* approaches strongly emphasize the need to be *comprehensive* and *inclusive* at the item generation stage (albeit for different reasons).⁶ Specifi-

⁴Needless to say, that the terms 'Type I' and 'Type II' errors as used here have nothing to do with the use of these terms, in the context of conventional significance testing procedures (e.g. see Henkel, 1976).

⁵Indeed, as already mentioned in previous sections, past studies have invariably used an identical set of indicators to specify both reflective and formative models.

⁶The emphasis on comprehensiveness in scale development is driven by (a) the need to capture all dimensions of the construct, and (b) the need for redundancy among items, which enhances internal consistency (Churchill, 1979; DeVellis, 2003; Spector, 1992). In contrast, the emphasis on comprehensiveness in index construction is because 'failure to consider all facets of the construct will lead to an exclusion of relevant indicators (and thus

cally, in generating reflective items for scale development purposes, organizational researchers are told that

the content of each item should primarily reflect the construct of interest ... Although the items should not venture beyond the bounds of the defining construct, they should exhaust the possibilities for types of items within those bounds ... at this stage of the scale development process, it is better to be overinclusive, ... It would not be unusual to begin with a pool of items that is three or four times as large as the final scale ... In general, the larger the item pool, the better (DeVellis, 2003, pp. 63–66).

Similarly, in generating formative indicators for potential inclusion in an index, researchers are advised that ‘the items used as indicators must cover the entire scope of the latent variable ... [and] be sufficiently inclusive in order to fully capture the construct’s domain of interest’ (Diamantopoulos and Winklhofer, 2001, pp. 271–272) as well that ‘breadth of definition is extremely important to causal indicators’ (Nunnally and Bernstein, 1994, p. 484).

In short, according to the extant literature, there appears to be no compelling reason as to why the *initial* item pool would differ *purely* because of the choice of measurement perspective. Assuming that literature guidelines on comprehensiveness and inclusiveness are diligently followed, item generation under each perspective would not be expected to result in widely divergent item pools. Therefore, the impact of a Type I error would not appear to have particularly adverse consequences at the item generation stage.

Unfortunately, the same cannot be said when it comes to measure *purification*, whereby the final set of items is selected for inclusion in the measure. Here, the effects of a Type I error can be profound because scale development and index construction procedures use fundamentally different criteria for retaining (excluding) indicators. Specifically, high intercorrelations among items are desirable in scale development since high intercorrelations enhance internal consistency (Bollen and Lennox, 1991); conversely, low

intercorrelations among indicators are typically interpreted as signals of problematic items which then become candidates for exclusion during scale purification (DeVellis, 2003). In contrast to scale development, under index construction, items are retained as long as they have a *distinct* influence on the latent variable (Bollen, 1989); indeed as Fayers and Hand observe, ‘we do not in general expect high correlations amongst the causal indicators, since this could indicate that there is unnecessary redundancy’ (Fayers and Hand, 1997, p. 147; see also Rossiter, 2002). In fact, if intercorrelations among items are very high, this is likely to result in multicollinearity problems (Bollen and Lennox, 1991; Law and Wong, 1999) that are resolved by excluding one or more items that exhibit strong linear dependencies with other items included in the measure (Diamantopoulos and Winklhofer, 2001).

Given the above differences, the *final* set of items comprising the reflective measure is likely to differ considerably from that comprising the formative measure. Thus, committing a Type I error can be expected to have major implications at the item purification stage and it *cannot* be automatically assumed that exactly the same set of items will be included in both the reflective and formative specifications (as previous studies have tended to assume – see Bollen and Ting, 2000; Fornell and Bookstein, 1982; Jarvis, Mackenzie and Podsakoff, 2003; Law and Wong, 1999; MacCallum and Browne, 1993).

As far as *criterion validity* is concerned, insights into the potential impact of a Type I error can be gained by considering the correlation between the measure developed using a reflective perspective and an external criterion (i.e. outcome variable) and comparing its magnitude to that between the criterion and the measure that would have resulted from following a formative perspective. Any significant differences between the correlations would be indicative of over- or under-estimation of criterion validity as a result of having committed a Type I error. Specifically, if the correlation between the reflective scale and the chosen criterion is significantly higher (lower) than the correlation between the formative index and the criterion, then Type I error would have resulted in an overestimation (underestimation) of criterion validity.

exclude part of the construct itself’ (Diamantopoulos and Winklhofer, 2001, p. 271).

An empirical illustration

To explore the potential consequences of a Type I error empirically, we focus on the construct of intraorganizational 'export coordination' for illustrative purposes and use data drawn from a survey of exporting practices of US firms ($n = 206$; see Appendix 1 for data-collection details). Both intra- and interorganizational coordination have been of interest to management academicians and practitioners for decades (e.g. Al-Dosary and Garba, 1998; Cheng, 1984; Faraj and Sproull, 2000; Grandori, 1997; Gittell, 2002; Harris and Raviv, 2002; Montoya-Weiss, Massey and Song, 2001; Sahin and Robinson, 2002; Tsai, 2002). Indeed, the concept of coordination lies at the heart of organizational systems theory. Social organizations, as one of the more complex classifications in the systems paradigm, for example, are defined as groups of individuals acting in concert (Pondy and Mitroff, 1979). Similarly, the role-system perspective espoused by Katz and Kahn (1978) views human organizations as a set of complementary work roles that together create a meaningful composition. In other words, to achieve system effectiveness, the members of the organization(s) must coordinate their functions and activities (Cheng, 1983; Lawrence and Lorsch, 1967). Consequently, coordination is a critical element of organizational and interorganizational structure (e.g. Harris and Raviv, 2002), and relationships, including those developed within supply channels (e.g. Chen, Federgruen and Zheng, 2001; Taylor, 2001, 2002), strategic alliances, networks, and hybrids (Grandori, 1997; Nault and Tyagi, 2001). Coordination has been linked to a number of outcome measures, including, but not limited to, output quantity, team effectiveness, organizational performance, innovation, knowledge sharing, total cost and service level (e.g. Cheng, 1983, 1984; Faraj and Sproull, 2000; Persaud, Kumar and Kumar, 2002; Zhao, Xie and Zhang, 2002).

In light of the sweeping effect of organizational coordination on important outcomes, as well as its widespread prevalence in the management literature, it is important to pursue a thorough understanding of the measurement specification of this construct. However, the choice of measurement perspective to operationalize coordination has not been explicitly defended in prior

research. This situation is regrettable because, 'wrongly treating variables as effect rather than causal indicators can bias parameter estimation and lead to incorrect assessments of the relationships between variables' (Bollen and Ting, 2000, p. 4). Thus, an incorrect operationalization of coordination as a result of a Type I error may have important consequences in the context of substantive research, particularly as far as the link between coordination and organizational outcomes is concerned.

Bearing the above in mind and returning to our illustrative construct of export coordination, this construct captures organizational coordination within the export department and between the export department and other functional areas. Its conceptual domain has been described as consisting of 'several interrelated themes: communication and common understanding; organizational culture emphasizing responsibility, cooperation and assistance; a lack of dysfunctional conflict; and common work-oriented goals' (Cadogan, Diamantopoulos and Pahud de Mortanges, 1999, p. 692). Based on this conceptualization, Cadogan, Diamantopoulos and Pahud de Mortanges (1999) used samples of British and Dutch exporters to develop a scale of export coordination (out of an initial item pool containing 30 items based on an extensive literature review, in-depth interviews with export managers and two pilot studies). Their operationalization strategy follows conventional scale development procedures and is thus based on the assumption that sound export coordination in a firm would be *reflected* in sound communication *and* responsibility/cooperation/assistance *and* lack of dysfunctional conflict, etc. (since it is the construct that is assumed to bring about variation in the indicators rather than the other way round). However, a more credible case could be made that coordination is *made up* of good communication, sharing of responsibility, cooperation, etc., since an increase (decrease) in *any* of these constituents would positively (negatively) impact the degree of export coordination in the firm. Having more (less) teamwork, communication, cooperation and so on would be expected to result in greater (lesser) coordination and there is no compelling reason why scoring high, on, say, communication *necessarily* implies also scoring high on, for example, teamwork. Thus (export) coordination is a good example of an organiza-

tional construct where a Type I error might have been committed during its operationalization.⁷

We now trace empirically the implications of this error along the lines described in the previous section. Specifically, we use the original item pool created by Cadogan, Diamantopoulos and Pahud de Mortanges (1999)⁸ and apply it to our own sample of US exporters to develop (a) a reflective scale of export coordination using conventional scale development procedures, and (b) a formative index of export coordination following the guidelines recently proposed by Diamantopoulos and Winklhofer (2001). We subsequently compare the derived measures in terms of content, parsimony and criterion validity thus illustrating potential consequences of Type I error on the export coordination measure itself and the assessment of the relationship between coordination and organizational (export) performance.

Scale development

Following the methodological guidelines of Steenkamp and van Trijp (1991) and Gerbing and Hamilton (1996), a two-stage approach was used to investigate the dimensionality of the export coordination items. First, the original pool of 30 items (shown in Appendix 2 as X1 to X30) was subjected to an exploratory factor analysis (EFA) with Varimax rotation to gain initial insights as to item dimensionality (Comrey, 1988; Hattie, 1985). Two primary factors were extracted together accounting for 53.4% of overall variance and, following inspection of the items loading on each factor, these were labelled 'collaboration' and 'conflict' (absence of) respectively. However, five items (X5, X18, X20, X23 and X30) displayed high cross-loadings and a further item (X4) failed to load highly on either factor; consequently, these six items were dropped from further analysis. In the second

stage, the remaining 24 items were entered into a confirmatory factor analysis (CFA) (conducted with the LISREL8 program), with the number of factors constrained to two and items allocated to each factor on the basis of the EFA results (see Gerbing and Hamilton, 1996). The resulting two-factor, CFA model's fit ($\chi^2 = 535.97$, $df = 251$; RMSEA = 0.083; GFI = 0.79; NNFI = 0.88; CFI = 0.89) was then compared to the fit of a one-factor model ($\chi^2 = 1110.2$, $df = 252$; RMSEA = 0.23; GFI = 0.45; NNFI = 0.64; CFI = 0.67) by means of a chi-square difference (D^2) test. The difference in fit was highly significant ($D^2 = 424.77$, $df = 1$, $p < 0.001$), indicating that two factors captured the covariation among the 24 items much better than a single common factor. Inspection of modification indices, however, revealed that the fit of the two-factor model could be further improved by eliminating several cross-loading items (X7, X10, X12, X19, X21, X22, X26 and X27). Accordingly, cross-loading items were dropped one at a time, starting with the one displaying the highest modification index. A significant improvement in fit was noted ($\chi^2 = 188.75$, $df = 103$; RMSEA = 0.07; GFI = 0.87; NNFI = 0.94; CFI = 0.94), however, further examination of modification indices revealed highly correlated measurement errors between some items. As the introduction of correlated measurement errors could not be *theoretically* defended in the present study (Danes and Mann, 1984) and bearing in mind that correlated errors 'are fall-back options nearly always detracting from the theoretical elegance and empirical interpretability of the study' (Bagozzi, 1983, p. 450), it was decided to eliminate some of the items concerned (starting with the highest modification index and eliminating the item with the lower squared multiple correlation). In all, five items were dropped (X2, X3, X15, X24, and X26). The final CFA model contained a total of 11 items, four capturing collaboration (X6, X8, X11, and X13) and seven capturing conflict (X1, X9, X14, X16, X17, X28, and X29) with all items loading significantly on their respective factors, no cross-loadings and no correlated measurement errors. The model's overall fit was very satisfactory ($\chi^2 = 58.17$, $df = 43$; RMSEA = 0.04; GFI = 0.94; NNFI = 0.98; CFI = 0.98). Construct (composite) reliability calculations also revealed high internal consistency for both subscales (collaboration = 0.87,

⁷It is not our intention here to engage in a substantive critique of Cadogan, Diamantopoulos and Pahud de Mortanges' (1999) measure of export coordination; we merely focus on export coordination as an *illustrative* construct to trace the implications of following a reflective versus formative approach for measure development purposes.

⁸We would like to thank John Cadogan for kindly making the initial item pool available for purposes of the present study.

conflict = 0.90). Lastly, average variance extracted (AVE) for collaboration was 0.62 and for conflict was 0.56; AVE scores greater than 0.50 indicate that a higher amount of variance in the indicators is captured by the construct compared to that accounted for by measurement error (Fornell and Larcker, 1981).

In summary, the application of conventional scale development procedures resulted in a reflective measure of export coordination comprising two unidimensional and highly reliable subscales and displaying excellent overall model fit. Despite these attractive properties, however, it should be recalled that this measure of export coordination is based on the *wrong* measurement perspective, given the nature of the construct. The measure is thus statistically sound (in that it easily satisfies conventional criteria for evaluating reflective multi-item measures) but theoretically questionable. The measurement perspective that *ought* to be applied is the formative perspective, necessitating the application of index construction (as opposed to scale development) procedures. This is the subject of the next section.

Index construction

As with the development of the reflective scale of export coordination, the 30-item pool served as the starting point for the construction of a formative index for the construct. First, multicollinearity among the items was assessed; high levels of multicollinearity in a formative measure can be problematic because the influence of each indicator on the latent construct cannot be distinctly determined (Bollen, 1989; Law and Wong, 1999). Using a 0.30 tolerance level as the cut-off criterion,⁹ 16 items were eliminated (X2, X5, X6, X8, X11, X12, X19, X20, X22, X23, X24, X25, X27, X28, X29 and X30). Subsequently, the

remaining 14 items were correlated with the 'global' statement 'activities and individuals are coordinated in our firm'. This step follows Diamantopoulos and Winklhofer's suggestion to correlate each indicator with 'a global item that summarizes the essence of the construct that the index purports to measure' (2001, p. 272, see also Fayers *et al.*, 1997). All correlation coefficients were found to be positive and significant ($p < 0.001$) and, thus, the 14 items were retained for further analysis. Next, a Multiple Indicators Multiple Causes (MIMIC) model (Jöreskog and Goldberger, 1975) was estimated with the 14 items as 'proximal antecedents' (Rossiter, 2002, p. 34) of export coordination and two further items (see Y1 and Y2 in Appendix 2) as reflective indicators of the construct (necessary for model identification). The two reflective items captured the effectiveness of the organization in disseminating export knowledge throughout the firm (and, clearly, coordination can be expected to positively impact upon dissemination effectiveness).¹⁰ Although the initial estimation of the MIMIC model yielded a good fit ($\chi^2 = 16.46$, $df = 13$; RMSEA = 0.035; GFI = 0.99; NNFI = 0.97; CFI = 1.00), several of the t-values associated with the parameters of the 14 items were not significant. To obtain a more parsimonious model, items with non-significant parameters were excluded from the model in an iterative process that deleted one item at a time, starting with the lowest t-value (Jöreskog and Sörbom, 1989); as a result, items X1, X3, X9, X13, X14, X15, X18, X21, and X26 were eliminated. The final MIMIC model included five formative indicators and also displayed good fit ($\chi^2 = 0.095$, $df = 4$; RMSEA = 0.0; GFI = 1.00; NNFI = 1.06; CFI = 1.00). The initial (14-item) and final (five-item) MIMIC models were then compared and no significant deterioration in fit was noted ($D^2 = 15.51$, $df = 9$, $p > 0.05$). Thus the five items (X4, X7, X10, X16, and X17) from the MIMIC analysis were retained to form an index of export coordination. Note that no dimensionality or reliability tests were performed on the

⁹We chose this level as current guidelines suggest that tolerance values greater than 0.35 may cause multicollinearity problems (see, for example, *InStat Guide to Choosing and Interpreting Statistical Tests*, available at <http://www.graphpad.com/instat3/instat.htm>). Note, in this context, that 'the default values for excluding collinear variables allow for an extremely high degree of multicollinearity. For example, the default tolerance value in SPSS for excluding a variable is .0001, which means that ... more than 99.99 percent of variance is predicted by the other independent variables' (Hair *et al.*, 1998, p. 193).

¹⁰For a theoretical discussion of this relationship as well as empirical evidence, see Cadogan *et al.* (2001). Note that the two items (Y1 and Y2 in Appendix 2) were *not* part of the item pool for the export coordinating construct and were simply used to enable the estimation of the MIMIC model.

formative index; this is because ‘factorial unity in factor analysis and internal consistency, as indicated by coefficient alpha, are not relevant’ (Rossiter, 2002, p. 315) under the formative measurement perspective (see also Bagozzi, 1994; Bollen and Lennox 1991; Jarvis, Mackenzie and Podsakoff, 2003). However, the items in the final index were checked to ensure that they still exhibited sufficient breadth of content to capture the domain of the coordination construct. This was deemed necessary because, as Diamantopoulos and Winklhofer point out, ‘indicator elimination – by whatever means – should not be divorced from conceptual considerations when a formative measurement model is involved’ (2001, p. 273).

In summary, the application of index construction procedures resulted in a formative measure of export coordination satisfying the criteria established by Diamantopoulos and Winklhofer (2001) for the evaluation of formative measures. Importantly, and unlike the reflective measure of export coordination discussed in the previous section, the statistical soundness of the measure is accompanied by theoretical appropriateness. In terms of Figure 1, for the *specific* construct under consideration, the formative index is located in the lower right quadrant (indicating a correct decision), whereas the reflective scale falls in the upper right quadrant (indicating a Type I error). The next section assesses the implications of a Type I error by explicitly comparing the two export coordination measures in terms of content, parsimony and criterion validity.

Measure comparison

Content

Only two items (X16, X17) were found to be common to both measures derived by the procedures described above; thus scale development and index construction efforts resulted in markedly different multi-item measures. More specifically, out of the 14 items *discarded* during scale purification, three (X4, X7, and X10) were subsequently *included* in the formative index. This lends empirical support to Bollen’s (1984) speculation that ‘it seems quite possible that a number of items (or indicators) have not been used in research because of their low or negative correlation with other indicators designed to

measure the same concept. If some of these are cause-indicators, researchers may have unknowingly removed valid measures’ (Bollen, 1984, p. 383). Thus our findings confirm the point made previously that scale development and index construction (as distinct approaches approaches to deriving multi-item measures) are likely to produce substantially different operationalizations of the same organizational construct *even* if the initial pool of items is the same: committing a Type I error (see Figure 1 earlier), impacts upon the content of the derived measure. Having said that, in this particular instance, the two measures are positively intercorrelated ($r = 0.782$, $p < 0.001$) thus apparently providing a consistent picture of export coordination.¹¹

Parsimony

With five items, the formative index provides a more concise operationalization of export coordination than the reflective scale (which contains twice as many items), indicating that a Type I error may also affect the parsimony of the measure. The practical implication is that less onerous demands would be placed upon respondents during data-collection efforts in subsequent studies (e.g. when a measure of export coordination is included in a research questionnaire). The explanation for the difference in length between the two measures most probably lies in the fact that index construction procedures tend to *eliminate* highly intercorrelated items (in order to minimize multicollinearity), whereas scale development procedures tend to *retain* highly intercorrelated items (in order to maximize internal consistency). Put differently, index construction discourages redundancy among indicators, whereas scale development encourages redundancy.¹²

Criterion validity

Export performance was used as the (external) criterion to assess the validity of the reflective and

¹¹The correlations between the formative index and the two reflective subscales (collaboration and conflict) came to 0.669 and 0.686 respectively; both are significant at $p < 0.001$.

¹²As one colleague aptly put it, ‘one model’s collinearity is another model’s reliability’.

Table 1. Export coordination and export performance

<i>Sales performance</i>	<i>Reflective Scale (R^2)</i>	<i>Formative Index (R^2)</i>	<i>z-value^a</i>
Export sales intensity (% of sales from export)	0.054**	0.091***	-2.242*
Export sales growth (3-year growth rate)	0.027	0.016	-1.208
Export sales per employee	0.029	0.049*	-1.382
<i>Profit performance</i>			
Export profit intensity (% of profit from export)	0.064**	0.111***	-2.671**
Relative profitability (export versus domestic market)	0.048*	0.004	4.695***
<i>Competitive performance</i>			
Sales versus major competitor	0.041*	0.053**	-0.877
Profit versus major competitor	0.035	0.023	0.974
Market share versus major competitor	0.053*	0.067**	-0.916
New market entry versus major competitor	0.024	0.038*	-1.160
<i>Satisfaction with performance</i>			
Sales	0.014	0.055***	-3.736***
Profit	0.024	0.010	1.727
Market share	0.025	0.069***	-3.425**
New market entry	0.016	0.049**	-3.034**
<i>Overall performance</i>	0.134***	0.137***	-0.149

Notes: ^az-test for differences in R^2 -values;

* $p < 0.05$; ** $p < 0.01$; *** $p < .001$.

formative measures of export coordination. In this context, any measure of a firm's export coordination would be expected to be positively correlated with measures of the firm's export performance (see Cadogan, Diamantopoulos and Pahud de Mortanges, 1999; Cadogan, Diamantopoulos and Siguaw, 2002). Accordingly, we used regression analysis to examine the magnitude and significance of the relationships between the derived measures and different aspects of export performance and formally tested any differences in R^2 -values (Kleinbaum *et al.*, 1998; Olkin, 1967). More specifically, for the reflective specification multiple regression analysis was used with the two subscales of collaboration and conflict as predictors, whereas a simple regression was performed on the formative measure. Given the multi-faceted nature of export performance (see Katsikeas, Leonidou and Morgan, 2000), a total of 14 export performance measures were employed in the assessment of criterion validity capturing sales performance, profit performance, competitive performance and satisfaction with performance (see Appendix 3 for measurement details).¹³

¹³The export performance measures selected have been widely used in previous empirical research and were chosen following consultation of the relevant literature on the conceptualization and measurement of export performance (e.g. Al-Khalifa and Morgan, 1995; Katsi-

Three points become apparent from Table 1. First, there is only a single instance where a significant link is observed with the reflective scale but a non-significant relationship with its formative counterpart (relative profitability). In contrast, there are five instances (export sales per employee, new market entry performance versus major competitor, satisfaction with sales, satisfaction with market share, and satisfaction with new market entry) in which the formative measure returns a significant result but the reflective measure does not. Lastly, in most instances in which significant relationships with export performance are observed for both the reflective *and* the formative measures, the magnitudes of the relationships associated with the latter are significantly higher (see results of z-test for differences in R^2 -values in Table 1).

From the above, it is evident that the consequences of a Type I error can also be manifested in the criterion validity of the derived measure. In our example, the results in Table 1 indicate that the adoption (erroneous) of a reflective perspective would have resulted in an underestimation of the links between export coordination and export performance. More worryingly, for some performance indicators, different *substantive* conclusions would have been

keas, Leonidou and Morgan, 2000; Matthyssens and Pauwels, 1996; Shoham, 1991; Thack and Axinn, 1994).

drawn as a result of having committed a Type I error; for example, export coordination would have been deemed to have no effect on sales export per employee (i.e. export productivity) or new market entry relative to competition (as the relevant R^2 -values under the reflective specification are non-significant). Thus, it is clear that a Type I error can have (potentially serious) implications in terms of the extent to which the derived measure relates to other measures of interest. With particular reference to the large body on organizational research that has been largely based on reflective coordination measures, it is thus possible that the true effects of coordination as both the 'glue' that ensures organizational effectiveness and as a determinant of performance outcomes may have been underestimated as a result of a Type I error.

At this point, it is important to note that the impact of Type I error on criterion validity was not simply the product of a reversal of directionality of the links between the construct and its indicators. In fact, given that aggregate (sub) scale scores (i.e. simple linear combinations) were used in the regression analysis in Table 1, the links between the construct and its indicators were not *explicitly* modelled; it was only the procedures (i.e. scale development versus index construction) used to *select* the items for inclusion in the measures that differed between the reflective and formative specifications. Thus the observed effect on criterion validity (underestimation in our example) was because (mostly) *different* items were finally included in the reflective and formative measures and not as a result of simply re-specifying the *same* set of items from reflective to formative (as done for example, by Fornell and Bookstein, 1982; Jarvis, MacKenzie and Podsakoff, 2003; Law and Wong, 1999; MacCallum and Browne, 1993).

Bearing the above in mind, one might argue that, instead of a regression analysis, structural equation modeling techniques should have been employed to assess the relationship between export coordination and export performance, as this would have enabled *explicit* modeling of the epistemic relationships between constructs and indicators as reflective or formative. While we acknowledge the merits of this approach and while we fully recognize that 'linear composites of indicators are not the same as the latent variable with which they are associated' (Bollen and

Lennox, 1991, p. 312), we opted for a simpler, regression-based approach using summed scores for three reasons.

First, our main concern was with the way in which Type I error can impact the selection of items when generating a multi-item measure rather than with the effects of measurement mis-specification on structural relations between constructs. The latter issue has already been addressed in previous studies by Law and Wong (1999), MacCallum and Browne (1993) and, most recently, by Jarvis, Mackenzie and Podsakoff (2003).

Second, aggregate summated multi-item measures are widely used in empirical research; indeed, 'the summated rating scale is one of the most frequently used tools in the social sciences' (Spector, 1992, p. 1). While researchers often use SEM approaches (e.g. CFA) when *developing* their measures, once items have been selected following dimensionality and reliability tests, it is common practice to combine them to generate overall (i.e. aggregate) measures of the construct(s) of interest. Note, in this context, that the methodological literature also encourages the use of aggregate measures; for example, it has been argued that a 'well-developed summated rating scale can have good reliability and validity' (Spector, 1992, p. 2).

Third, a SEM approach would have made the direct comparison of criterion validity under reflective and formative measurement conditions highly problematic, since the resulting structural equation models could not have been compared by means of nested model tests (Bagozzi and Yi, 1988). A model, say M1, is said to be nested within another model, say M2, if 'M1 can be obtained from M2 by constraining one or more of the free parameters in M2 to be fixed or equal to other parameters. Thus, M1 can be thought of as a special case of M2' (Long, 1983, p. 65). In our case, given that the reflective and formative specifications contain different (unique) indicators plus the fact that the links between the latent variable and the indicators have different directionality, the resulting measurement models would not be nested: it is simply not possible to get from the reflective to the formative specification simply by constraining (or relaxing) model parameters. Moreover, as noted earlier, the reflective specification comprises a two-factor representation of export coordination, which means that *two* structural paths impact upon

export performance, as opposed to one in the formative specification. Thus the only criterion validity comparison that could reasonably be made between the two models would be in terms of the difference in the coefficient of determination (R^2) on the criterion (export performance) variable. This assessment is not that different from comparing R^2 -values along the lines of Table 1. In this context, it is highly doubtful whether the additional effort involved of specifying SEM models, repeating this several times (to cover all aspects of export performance), only to again compare R^2 -values at the end would have resulted in substantially better insights on the criterion validity of the measures.¹⁴

Conclusion

The purpose of this article was to inform organizational theory by systematically comparing conventional scale development procedures based on a reflective measurement perspective with a formative measure approach involving index construction, and tracing empirically the consequences of committing a Type I error (i.e. erroneously following a reflective measurement perspective when a formative perspective should have been adopted). Our study, conducted within the context of organizational coordination, is thus fully consistent with calls in the organizational research literature that 'more empirical studies directed towards the comparison of these two views for various management constructs should ... be conducted' (Law and Wong, 1999, p. 156). According to our findings, the two approaches result in materially different coordination measures in terms of content, parsimony and criterion validity; thus, the choice of measurement perspective (i.e. reflective versus formative) and the resultant choice of procedure (i.e. scale development versus index construction) *does* matter from a practical point of view.

¹⁴Moreover, identification problems with the formative specification would have to be solved under a SEM approach, because the formative model 'in isolation is statistically underidentified' (Bollen and Lennox, 1991, p. 312). Unfortunately, there is still no consensus in the literature as to the best way of dealing with this problem (for different views see Bollen and Davis, 1994; Diamantopoulos and Winklhofer, 2001; Jarvis, Mackenzie and Podsakoff, 2003; MacCallum and Browne, 1993).

For practising organizational researchers, the implications are five-fold. First, explicit consideration at the construct definition stage of the likely causal priority between the latent variable and its indicators is essential so as to avoid 'obvious' errors in the choice of measurement perspective (Edwards and Bagozzi, 2000). Unfortunately, and despite warnings that 'researchers should not automatically confine themselves to the unidimensional classical test model' (Bollen and Lennox, 1991, p. 312), within the domain of organizational research (and certainly within organizational coordination studies), a reflective perspective is by far the dominant approach. However, making the wrong choice and committing a Type I error is not without costs since, as our illustrative example showed, an inaccurate assessment of the relationship between the focal construct (coordination) and important outcomes (performance) may result. Thus we strongly urge organizational researchers to consider the potential applicability of a formative measurement perspective *before* setting out to develop multi-item measures for their construct(s) of interest (see also Law and Wong, 1999). Consultation of the comprehensive set of guidelines recently offered by Jarvis, Mackenzie and Podsakoff (2003) for choosing between reflective and formative specifications should help considerably with this task.

Second, assuming that a reflective perspective has been chosen on *theoretical* grounds (which implies that a formative perspective *has* been carefully considered but ultimately rejected), it is *not* acceptable to 'change one's mind' based on the results obtained during scale development. Any tendency to use the formative measurement model as 'a handy excuse for low internal consistency' (Bollen and Lennox, 1991, p. 312) must be firmly resisted. Switching from scale development to index construction purely on the basis of data-driven considerations (e.g. because the dimensionality and/or reliability of the scale was unsatisfactory) amounts to nothing less than *knowingly* committing a Type II error (see Figure 1)! Clearly, this is as undesirable as committing a Type I error.¹⁵ The choice of measurement

¹⁵In addition to involving 'cheating', there is also no guarantee that a set of items that failed to generate a good reflective scale will succeed in producing a sound formative index.

perspective should always be based on *theoretical* considerations regarding the direction of the links between the construct and its indicators (Blalock 1968; Costner, 1969; Edwards and Bagozzi, 2000). While organizational researchers may sometimes get this wrong, and thus *unknowingly* commit a Type I or Type II error, choosing a measurement perspective purely by observing whether scale development 'works better' than index construction on a particular dataset is bound to capitalize on chance; this will limit the empirical replicability of the derived measure and will certainly not improve its conceptual soundness.

Third, as Edwards and Bagozzi point out, 'if measures are specified as formative, their validity must still be established. It is bad practice to ... claim that one's measures are formative, and do nothing more' (Edwards and Bagozzi, 2000, p. 171). Unfortunately, in the few instances where formative measures have been intentionally employed in previous studies (e.g. Ennew, Reed and Binks, 1993; Homburg, Workman and Krohmer, 1999; Johansson and Yip, 1994), no attempt was made to assess their quality. However, this most probably reflects the fact that it was not until very recently that *concrete* guidelines for constructing multi-item indexes with formative indicators were made available in the mainstream literature (Diamantopoulos and Winklhofer, 2001; Jarvis, Mackenzie and Podsakoff, 2003). Future efforts by organizational researchers to model constructs with formative indicators should, therefore, ensure that such guidelines are indeed followed so as to provide a basis for judging the validity of the derived indexes. The practical implication of this is that provision should be made at the study *design* stage for the incorporation of *additional* items (external to the index) to enable the specification of MIMIC models, assess external validity, etc (see Diamantopoulos and Winklhofer, 2001). Failure to make such provision is bound to cause serious problems at the index construction stage not least because, on its own, a formative indicator measurement model is statistically underidentified (Bollen, 1989; Bollen and Lennox, 1991). To estimate the model it is necessary to introduce some reflective indicators (as in a MIMIC specification) or relate it to other constructs operationalized by means of reflective measures (Bollen and Davis, 1994; Jarvis, Mackenzie and Podsakoff, 2003; Law and Wong, 1999;

MacCallum and Browne, 1993; Williams, Edwards and Vandenberg, 2003).

Fourth, it is sometimes argued that reflective measures (and use of covariance structure analysis) are better suited for theory development and testing purposes, whereas formative measures (accompanied by partial least squares (PLS) estimation) are better for prediction (Anderson and Gerbing, 1988). For example, Fornell and Bookstein state that 'should the study intend to account for observed variances ... reflective indicators are most suitable. If the objective is explanation of abstract or 'unobserved' variance, formative indicators ... would give greater explanatory power' (1982, p. 292).¹⁶ We are somewhat uncomfortable with these suggestions because (a) they are based on the (implicit) assumption that the multi-item measure of the organizational construct in question will comprise exactly the same set of indicators irrespective of whether a reflective or a formative perspective is applied (and this *cannot* be taken for granted as the results of the present study clearly show), and (b) they imply that the epistemic relations between the organizational construct and its measurements can be 'manipulated' to suit the objective of the investigator. We feel that such potential for manipulation is inconsistent with the principles of sound 'auxiliary theory', whereby the nature and direction of relationships between constructs and their measures are clearly specified *prior* to examining substantive relationships between theoretical constructs (Blalock, 1968; Costner, 1969).¹⁷ As

¹⁶Given the *same* set of indicators, 'the reflective formulation can never account for more variance in the dependent variable than the formative specification' (Fornell, Rhee and Yi, 1991, p. 317). The reason for this can be traced to the fact that the variance in the true score is smaller than the variance in the observed variables (i.e. the indicators) under a reflective specification; in contrast, under a formative specification, the variance in the true score is larger than the variance in the indicators (see Fornell and Cha, 1994; Namboodiri, Carter and Blalock, 1975). However, as the present study has shown, it *cannot* be assumed that the same set of indicators will indeed be included in a measure irrespective of whether scale development or index construction strategies have been followed.

¹⁷To be fair, Fornell and Bookstein themselves appear to recognize this problem when they state that 'one may wish to minimize residual variance in the structural portion of the model, which suggests use of formative indicators, even though the constructs are conceptua-

Venaik, Midgley and Devinney state ‘the choice between formative and reflective models must be driven fundamentally by theory, not empirical testing’ (2004, p. 42). We therefore strongly endorse Law and Wong’s view that ‘the conceptualization of the constructs should be theory-driven’ (Law and Wong, 1999, p. 156) as well as Bollen’s recommendation that ‘the researcher should decide in advance which are effect- and which are cause-indicators’ (Bollen, 1984, p. 383; see also Jarvis, Mackenzie and Podsakoff, 2003). In this context, a particularly important consideration applicable to both reflective and formative indicators has to do with the *content adequacy* of a measure. Specifically, when constructing a measure, one has to reconcile the theory-driven conceptualization of the measure with the desired statistical properties of the items comprising the measure as revealed by empirical testing. For example, blindly eliminating items to improve reliability (in a reflective scale) or to reduce multicollinearity (in a formative index) may well have adverse consequences for the content validity of the derived measure. While no hard and fast rules can be offered on how to balance content adequacy and statistical considerations, it has to be emphatically stated that exclusive focus on the latter is unlikely to result in robust and replicable measures.

Fifth, although the focus of the present study was on a comparison of *first-order* reflective and formative measurement models,¹⁸ it should be noted that these are not the *only* options available to researchers. In some instances, the conceptualization of the construct may be such as to necessitate the specification of a higher-order model (Edwards, 2001). As Jarvis, Mackenzie and Podsakoff observe, ‘conceptual definitions of constructs are often specified at a more abstract level, which sometimes include multiple formative and/or reflective first-order dimensions’ (2003,

p. 204). In such cases, more complex measurement models may be invoked in which either a reflective or, a formative specification is applied at all levels (e.g. first-order reflective, second-order reflective)¹⁹ or, alternatively using a *combination* of reflective and formative specifications (e.g. first-order reflective, second-order formative).²⁰ Researchers should be aware of such options for operationalizing complex constructs and should carefully consider their potential applicability when deciding on their measurement strategy.

In conclusion, we would like to emphasize that we do not by any means consider a formative perspective to be *inherently* superior to a reflective approach (or vice versa) in coordination studies, or more broadly, organizational research. Although our particular illustration of the consequences of a Type I error resulted in a formative coordination measure that was more parsimonious and a stronger correlate of an external criterion than its reflective counterpart, other applications with different datasets and/or focal constructs may well produce different results (e.g. Type I error may result in overestimation of external validity). Our purpose was *not* to derive generalizations of the conditions under which one approach outperforms the other in terms of parsimony and/or criterion validity; indeed, it is extremely doubtful whether such general conditions do, in fact, exist given that item pools for different constructs in different intra- and interorganizational studies are likely to vary substantially in terms of the patterns of intercorrelations among the items. Instead, our much more modest aim was to demonstrate that the choice of measurement perspective matters from a practical point of view and, more specifically, that committing a Type I error is not inconsequential as far as the properties (content, parsimony, criterion validity) of the derived measure are concerned. Hopefully, the insights provided by our analysis will help future organizational researchers make their choice of measurement perspective more wisely.

lised as giving rise to the observations (which suggests use of reflective indicators)’ (Fornell and Bookstein, 1982, p. 294).

¹⁸In Edwards and Bagozzi’s (2000) terminology, the measurement models considered in this paper were the ‘direct reflective’ and ‘direct formative’ models respectively. These are by far the most common measurement models used in substantive research, hence the focus of the present study on them.

¹⁹This is probably the most widely used form of a second-order model, i.e. the second-order factor model (e.g. see Rindskopf and Rose, 1988).

²⁰For a taxonomy of second-order measurement models, see Edwards (2001) or Jarvis, Mackenzie and Podsakoff (2003).

Appendix 1. Data collection

Data were collected from US exporters using the *Export Yellow Pages* (available from the US Department of Commerce) as a sampling frame. A sample of 2036 firms was randomly selected and a questionnaire, together with a cover letter, were mailed to the listed contact person. The titles of the latter indicated that all were upper executive level (e.g. national account manager, sales manager, general manager, vice-president, president). Shortly after the initial mailing, a second mailing, including questionnaire and cover letter, was undertaken in order to increase the response rate; in total, 206 responses were obtained.

We approached the problem of ineligibility by telephone, contacting a randomly selected sample of 100 non-respondents to *directly* determine reasons for non-response (Lesley, 1972). Based on this information, we were able to calculate a 95% confidence interval for the number of ineligibles in the sample (Wiseman and Billington, 1984). Ineligibles included instances where (a) the firm no longer existed, (b) the firm was not involved in exporting, (c) the contact name no longer worked for the company or (d) the wrong address was included in the sampling frame. Adjusted for ineligibility, the 206 responses constitute an effective response rate between 22% and 34% (reflecting respectively lower and upper 95% confidence limits of the ineligibility estimates). To further investigate potential non-response bias, a comparison of early and late respondents was undertaken (Armstrong and Overton, 1977). Early respondents were defined as the usable questionnaires returned within the first three days of returns and late respondents were those who responded after the follow-up. At the 5% level of significance, no significant differences were observed, thus indicating that response bias was unlikely to be a major problem in the present study.

Appendix 2. Export coordination: item pool

X1. In our company, departments/individuals compete with each other to achieve their own goals rather than working together to achieve common objectives.

- X2. Key players from other functional areas (e.g. production, finance) hinder the export related activities of this firm.
- X3. In our company, the objectives pursued by export personnel do not match those pursued by members of the manufacturing or R&D departments.
- X4. In our company, if the export unit does well, the reward system is designed so that everyone within the firm benefits.
- X5. Key players from other functional areas (e.g. production, finance) are supportive of those involved in the firm's export operations.
- X6. Export personnel build strong working relationship with other people in our company.
- X7. Salespeople coordinate very closely with other company employees to handle post-sales problems and services in our export markets.
- X8. In this firm, when conflicts between functional areas occur (e.g. between export personnel and manufacturing), we reach mutually satisfying agreements.
- X9. Those involved in the firm's export operations have to compete for scarce company resources with other functional areas (e.g. domestic sales team, marketing, R&D).
- X10. Export personnel work together as a team.
- X11. Employees within the export unit and those in other functional areas (e.g. engineering) help each other out.
- X12. Departments in our company work together as a team in relation to our export business.
- X13. Those employees involved in our firm's export operations look out for each other as well as for themselves.
- X14. Export personnel work independently from other functional groups within our company.
- X15. Other than export personnel, it could be stated that few people in this organization contribute to the success of the firm's export activities.
- X16. Certain key players in our firm attach little importance to our export activities.
- X17. The export activities of this company are disrupted by the behaviour of managers from other departments (e.g. manufacturing).

- X18. In this company, there is a sense of teamwork going right down to the shop floor.
- X19. There is a strong collaborative working relationship between export personnel and production.
- X20. Functional areas in this firm pull together in the same direction.
- X21. Competition for scarce resources reduces cooperation between functional areas in our firm (e.g. between export unit and R&D).
- X22. The activities of our business functions (e.g. marketing/sales, manufacturing, R&D, finance/accounting, etc.) are integrated in pursuing a common goal.
- X23. Our managers understand how everyone in our business can contribute to creating value for export customers.
- X24. In this company, export staff share programmes and resources with other business functions.
- X25. We resolve issues and conflicts through communication and group problem-solving.
- X26. In our company, it is considered that the less interaction export personnel have with other functional areas the better.
- X27. People from different functional areas in our firm discuss their problems openly and constructively.
- X28. There are tensions among functional areas (e.g. export personnel and manufacturing) that interfere with the company's export activities.
- X29. In our firm there is interdepartmental conflict.
- X30. In this firm, our business functions (e.g., export, manufacturing) are integrated in serving the needs of our export markets.

Dissemination items (for MIMIC model)

- Y1. We are effective at disseminating export information throughout our company.
- Y2. Data on export customer satisfaction are disseminated at all levels in this company on a regular basis.

Appendix 3. Export performance measures

Sales Performance

1. *Export Sales Intensity*: approximately what percentage of your total sales turnover is derived from exports?
2. *Export Sales Growth*: please indicate the approximate growth/decline rate of your export sales over the last three years?
3. *Export Sales per Employee*: total export sales turnover (in \$) divided by total number of employees.

Profit Performance

4. *Export Profit Intensity*: approximately what percentage of your total profits is derived from exports?
5. *Relative Profitability*: overall, how profitable are your export sales in relation to sales in the domestic market? (1 to 7 scale, where 1 = much less profitable and 7 = much more profitable)

Competitive Performance

In relation to your major competitors, how would you rate your export performance along the following dimensions (1 to 7 scale, where 1 = much worse and 7 = much better):

6. Export Sales
7. Export Profits
8. Export Market Share
9. Rate of New Market Entry

Export Satisfaction

Overall, how satisfied are you with your performance along the following dimensions (1 to 7 scale, where 1 = very unsatisfied and 7 = very satisfied):

10. Export Sales
11. Export Profitability
12. Export Market Share
13. Rate of New Market Entry

Overall Performance

14. Overall, how would you rate your firm's export performance? (1 to 7 scale, where 1 = poor and 7 = outstanding).

References

- Al-Dosary, A. and S. B. Garba (1998). 'Inter-organizational coordination: A case study of the education and training system in Saudi Arabian manpower planning', *International Journal of Management*, **15**(2), pp. 161–174.
- Al-Khalifa, A. and N. A. Morgan (1995). 'Export performance measurement: A review and suggested directions', *Marketing Theory and Applications*, **6**, pp. 313–318.
- Anderson, J. C. and D. W. Gerbing (1982). 'Some methods for respecifying measurement models to obtain unidimensional construct measurement', *Journal of Marketing Research*, **19**, pp. 453–460.
- Anderson, J. C. and D. W. Gerbing (1988). 'Structural equation modeling in practice: A review and recommended two-step approach', *Psychological Bulletin*, **103**, pp. 411–423.
- Armstrong, J. B. and T. S. Overton (1977). 'Estimating non-response bias in mail surveys', *Journal of Marketing Research*, **14**(3), pp. 396–402.
- Bagozzi, R. P. (1983). 'Issues in the application of covariance structure analysis: A further comment', *Journal of Consumer Research*, **9**, pp. 449–450.
- Bagozzi, R. P. (1994). 'Structural equation models in marketing research: Basic principles'. In: R. Bagozzi (ed.), *Principles of marketing research*, pp. 317–385. Blackwell Publishing, Oxford.
- Bagozzi, R. P. and C. Fornell (1982). 'Theoretical concepts, measurements, and meaning'. In: C. Fornell (ed.), *A second generation of multivariate analysis*, **1**, pp. 24–38. Praeger, New York.
- Bagozzi, R. P. and Y. Yi (1988). 'On the evaluation of structural equation models', *Journal of the Academy of Marketing Science*, **16**(1), pp. 74–94.
- Blalock, H. M. (1968). 'The measurement problem: A gap between the languages of theory and research'. In: H. M. Blalock and A. B. Blalock (eds), *Methodology in social research*. McGraw-Hill, New York.
- Bollen, K. A. (1984). 'Multiple indicators: Internal consistency or no necessary relationship?', *Quality and Quantity*, **18**, pp. 377–385.
- Bollen, K. A. (1989). *Structural equations with latent variables*. John Wiley & Sons, New York.
- Bollen, K. A. and W. R. Davis (1994) Causal indicator models: Identification, estimation, and testing. Paper presented at the American Sociological Association Convention, Miami.
- Bollen, K. A. and R. Lennox (1991). 'Conventional wisdom on measurement: A structural equation perspective', *Psychological Bulletin*, **110**, pp. 305–314.
- Bollen, K. A. and K. F. Ting (2000). 'A tetrad test for causal indicators', *Psychological Methods*, **5**(1), pp. 3–22.
- Carmines, E. G. and R. A. Zeller (1979). *Reliability and validity assessment*. Sage Publications, Beverly Hills.
- Cadogan, J. W., A. Diamantopoulos and C. Pahud de Mortanges (1999). 'A measure of export market orientation: Scale development and cross-cultural validation', *Journal of International Business Studies*, **30**, pp. 689–707.
- Cadogan, J. W., A. Diamantopoulos and J. A. Siguaw (2002). 'Export market orientated activities: Their antecedents, performance consequences, and moderating influences', *Journal of International Business Studies*, **33**(3), pp. 615–626.
- Cadogan, J. W., N. J. Paul, R. T. Salminen, K. Puumalainen and S. Sundqvist (2001). 'Key antecedents to "export" market-oriented behaviors: A cross-national empirical examination', *International Journal of Research in Marketing*, **18**, pp. 261–282.
- Chen, F., A. Federgruen and Y. Zheng (2001). 'Coordination mechanisms for a distribution system with one supplier and multiple retailers', *Management Science*, **47**(5), pp. 693–708.
- Cheng, J. L. C. (1983). 'Interdependence and coordination in organizations: A role-system analysis', *Academy of Management Journal*, **26**(1), pp. 156–162.
- Cheng, J. L. C. (1984). 'Organizational coordination, uncertainty, and performance: An integrative study', *Human Relations*, **37**, pp. 829–851.
- Churchill, G. A. (1979). 'A paradigm for developing better measures of marketing constructs', *Journal of Marketing Research*, **16**, pp. 64–73.
- Cohen, P., J. Cohen, J. Teresi, M. Marchi and N. Velez (1990). 'Problems in the measurement latent variables in structural equations causal models', *Applied Psychological Measurement*, **14**, pp. 183–196.
- Comrey, A. L. (1988). 'Factor-analytic methods of scale development in personality and clinical psychology', *Journal of Consulting and Clinical Psychology*, **56**, pp. 754–761.
- Costner, H. L. (1969). 'Theory, deduction, and the rules of correspondence', *American Journal of Sociology*, **75**, pp. 245–263.
- Danes, J. E. and O. K. Mann (1984). 'Unidimensional measurement and structural equation models with latent variables', *Journal of Business Research*, **12**, pp. 337–351.
- DeVellis, R. F. (2003). *Scale development: Theory and applications*. (2nd edn.). Sage Publications, Newbury Park, CA.
- Diamantopoulos, A. (1999). 'Export performance measurement: Reflective versus formative indicators', *International Marketing Review*, **16**, pp. 444–457.
- Diamantopoulos, A. and H. Winklhofer (2001). 'Index construction with formative indicators: An alternative to scale development', *Journal of Marketing Research*, **37**, pp. 269–277.
- Edwards, J. R. (2001). 'Multidimensional Constructs in Organizational Behavior Research: An Integrative Analytical Framework', *Organizational Research Methods*, **4**(2), pp. 144–192.
- Edwards, J. R. and R. P. Bagozzi (2000). 'On the nature and direction of relationships between constructs and measures', *Psychological Methods*, **5**, pp. 155–174.
- Ennew, C. T., G. V. Reed and M. R. Binks (1993). 'Importance-performance analysis and the measurement of service quality', *European Journal of Marketing*, **27**(2), pp. 59–70.
- Faraj, S. and L. Sproull (2000). 'Coordinating expertise in software development teams', *Management Science*, **46**, pp. 1554–1568.
- Fayers, P. M. and D. J. Hand (1997). 'Factor analysis, causal indicators and quality of life', *Quality of Life Research*, **6**, pp. 139–150.
- Fayers, P. M., D. J. Hand, K. Bjordal and M. Groenvold (1997). 'Causal indicators in quality of life research', *Quality of Life Research*, **6**, pp. 393–406.

- Fornell, C. and F. L. Bookstein (1982). 'A comparative analysis of two structural equation models: LISREL and PLS applied to market data'. In: C. Fornell (ed.), *A second generation of multivariate analysis*, 1, pp. 289–324. Praeger, New York.
- Fornell, C. and J. Cha (1994). 'Partial least squares'. In: R. P. Bagozzi (ed.), *Advanced methods of marketing research*, pp. 52–78. Blackwell, Oxford.
- Fornell, C. and D. F. Larcker (1981). 'Evaluating structural equation models with unobservable variables and measurement error', *Journal of Marketing Research*, 18, pp. 39–50.
- Fornell, C., P. Lorange and J. Roos (1990). 'The cooperative venture formation process: A latent variable structural modeling approach', *Management Science*, 36(10), pp. 1246–1255.
- Fornell, C., B. D. Rhee and Y. Yi (1991). 'Direct regression, reverse regression, and covariance structure analysis', *Marketing Letters*, 2, pp. 309–320.
- Gerbing, D. W. and J. G. Hamilton (1996). 'Viability of exploratory factor analysis as a precursor to confirmatory factor analysis', *Structural Equation Modeling*, 3(1), pp. 62–72.
- Gittell, J. H. (2002). 'Coordinating mechanisms in care provider groups: Relational mediator and input uncertainty as a moderator of performance effects', *Management Science*, 48, pp. 1408–1426.
- Grandori, A. (1997). 'An organizational assessment of interfirm coordination modes', *Organization Studies*, 18, pp. 897–925.
- Hair, J. F. Jr., R. E. Anderson, R. L. Tatham and W. C. Black (1998). *Multivariate data analysis*. (5th edn). Prentice Hall, New Jersey.
- Harris, M. and A. Raviv (2002). 'Organization design', *Management Science*, 48, pp. 852–865.
- Hattie, J. (1985). 'Methodology review: Assessing unidimensionality of tests and items', *Applied Psychological Measurement*, 9(June), pp. 139–164.
- Hauser, R. M. (1971). *Socioeconomic background and educational performance*. Rose Monograph Series American Sociological Association, Washington DC.
- Hauser, R. M. (1973). 'Disaggregating a social-Psychological model of educational attainment'. In: A. S. Goldberger and O. D. Duncan (eds), *Structural equation models in the social sciences*. Seminar Press, New York.
- Henkel, R. E. (1976). *Tests of significance*. Sage Publications, Beverly Hills, CA.
- Hogan, E. A. and D. A. Martell (1987). 'A confirmatory structural equations analysis of the job characteristics model', *Organizational Behavior and Human Performance*, 34, pp. 141–174.
- Homburg, C., J. P. Workman Jr. and H. Krohmer (1999). 'Marketing's influence within the firm', *Journal of Marketing*, 63(2), pp. 1–17.
- Hulland, J. (1999). 'Use of partial least squares (pls) in strategic management research: A review of four recent studies', *Strategic Management Journal*, 20(2), pp. 195–204.
- James, J. R. and L. A. James (1989). 'Causal modelling in organizational research', *International Review of Industrial and Organizational Psychology*, 11, pp. 371–404.
- James, L. R. and A. P. Jones (1980). 'Perceived job characteristics and job satisfaction: An examination of reciprocal causation', *Personnel Psychology*, 33, pp. 97–135.
- Jarvis, C. B., S. B. MacKenzie and P. M. Podsakoff (2003). 'A critical review of construct indicators and measurement model misspecification in marketing and consumer research', *Journal of Consumer Research*, 30(2), pp. 199–218.
- Johansson, J. K. and G. S. Yip (1994). 'Exploiting globalization potential: U.S. and Japanese strategies', *Strategic Management Journal*, 15(8), pp. 579–601.
- Jöreskog, K. G. and A. S. Goldberger (1975). 'Estimation of a model with multiple indicators and multiple causes of a single latent variable', *Journal of the American Statistical Association*, 10, pp. 631–639.
- Jöreskog, K. G. and D. Sörbom (1989). *LISREL 7: A guide to the program and applications*. SPSS Inc, Chicago.
- Katsikeas, C. S., L. C. Leonidou and N. A. Morgan (2000). 'Firm-level export performance assessment: review, evaluation, and development', *Journal of the Academy of Marketing Science*, 28, pp. 493–511.
- Katz, D. and R. L. Kahn (1978). *The social psychology of organizations*. (2nd edn). Wiley, New York.
- Kleinbaum, D. G., L. L. Kupper, K. E. Muller and N. Azhar (1998). *Applied regression analysis and other multivariate methods*. Duxbury Press, Pacific Grove, CA.
- Law, K. S. and C. Wong (1999). 'Multidimensional constructs in structural equation analysis: An illustration using the job perception and job satisfaction constructs', *Journal of Management*, 25, pp. 143–160.
- Law, K. S., C. Wong and W. H. Mobley (1998). 'Toward a taxonomy of multidimensional constructs', *Academy of Management Review*, 23, pp. 741–755.
- Lesley, L. L. (1972). 'Are high response rates essential for valid surveys?', *Social Science Research*, 1, pp. 323–334.
- Lawrence, P. R. and J. W. Lorsch (1967). 'Differentiation and integration in complex organizations', *Administrative Science Quarterly*, 12, pp. 1–47.
- Long, J. S. (1983). *Covariance structure models: An introduction to LISREL*. Sage Publications, Beverly Hills.
- MacCallum, R. C. and M. W. Browne (1993). 'The use of causal indicators in covariance structure models: Some practical issues', *Psychological Bulletin*, 114, pp. 533–541.
- Matthyssens, P. and P. Pauwels (1996). 'Assessing export performance measurement'. In: T. K. Madsen and S. T. Cavusgil (eds), *Advances in International Marketing*, 8, pp. 85–114. JAI Press, London.
- Montoya-Weiss, M. M., A. P. Massey and M. Song (2001). 'Getting it together: Temporal coordination and conflict management in global virtual teams', *Academy of Management Journal*, 44, pp. 1251–1262.
- Morrisson, E. W. (2002). 'Newcomers' relationships: The role of social network ties during socialization', *Academy of Management Journal*, 45, pp. 1149–1160.
- Nachmias, D. and C. Nachmias (1976). *Research methods in the social sciences*. St Martin's Press, London.
- Namboodiri, N. K., L. F. Carter and H. M. Blalock Jr. (1975). *Applied multivariate analysis and multivariate analysis and experimental designs*. McGraw-Hill, New York.
- Nault, B. R. and R. K. Tyagi (2001). 'Implementable mechanisms to coordinate horizontal alliances', *Management Science*, 47(6), pp. 787–799.

- Netemeyer, R. G., W. O. Bearden and S. Sharma (2003). *Scaling Procedures: Issues and Applications*. Sage Publications, London.
- Nunnally, J. C. and I. H. Bernstein (1994). *Psychometric theory*. (3rd edn.). McGraw-Hill, New York.
- Olkin, I. (1967). 'Correlation revisited'. In J. C. Stanley (ed.), *Improving experimental design and statistical analysis*. Rand McNally, Chicago.
- Oppenheim, A. N. (1992). *Questionnaire design and attitude measurement*. (2nd edn.). St. Martin's Press, London.
- Persaud, A., U. Kumar and V. Kumar (2002). 'Coordination structures and innovative performance in global R&D labs', *Canadian Journal of Administrative Sciences*, **19**(1), pp. 57–75.
- Pondy, L. and I. Mitroff (1979). 'Beyond open system models of organization'. In: B. Shaw (ed.), *Research in organizational behaviour*, pp. 3–39. JAI Press, Greenwich, CT.
- Ramamoorthy, N. and P. C. Flood (2004). 'Gender and Employee Attitudes: The Role of Organizational Justice Perceptions', *British Journal of Management*, **15**(3), pp. 247–258.
- Rindskopf, D. and T. Rose (1988). 'Some theory and applications of confirmatory second-order factor analysis', *Multivariate Behavioral Research*, **23**, pp. 51–67.
- Rossiter, J. R. (2002). 'The C-OAR-SE procedure for scale development in marketing', *International Journal of Research in Marketing*, **19**, pp. 305–335.
- Sahin, F. and E. P. Robinson (2002). 'Flow coordination and information sharing in supply chains: Review, implications and directions for future research', *Decision Sciences*, **33**, pp. 505–536.
- Sarros, J. C., G. A. Tanewski, R. P. Winter, J. C. Santora and I. L. Densten (2002). 'Work Alienation and Organizational Leadership', *British Journal of Management*, **13**(4), pp. 285–304.
- Scandura, T. and E. Williams (2000). 'Research methodology in management: Current practices, trends, and implications for future research', *Academy of Management Journal*, **43**, pp. 1248–1264.
- Schaubroeck, J. and S. S. K. Lam (2002). 'How similarity to peers and supervisor influences organizational advancement in different cultures', *Academy of Management Journal*, **45**, pp. 1120–1136.
- Shoham, A. (1991) 'Performance in exporting: A state-of-the-art literature review and synthesis and directions for future research'. Paper presented at the Academy of International Business Annual Meeting Miami, FL, 18–20 October.
- Spector, P. E. (1992). *Summated rating scales construction: An introduction (Quantitative Applications in the Social Sciences, Paper No. 82)*. Sage Publications, Newbury Park, CA.
- Steenkamp, J. B. E. M. and H. C. M. van Trijp (1991). 'The use of LISREL in validating marketing constructs', *International Journal of Research in Marketing*, **8**, pp. 283–299.
- Stone-Romero, E. F., A. E. Weaver and J. L. Glenar (1995). 'Trends in research design and data analytic strategies in organizational research', *Journal of Management*, **21**, pp. 141–157.
- Subramani, M. R. and N. Venkatraman (2003). 'Safeguarding investments in asymmetric interorganizational relationships: Theory and evidence', *Academy of Management Journal*, **46**, pp. 46–62.
- Taylor, T. A. (2001). 'Channel coordination under price protection midlife returns and end-of-life returns in dynamic markets', *Management Science*, **47**, pp. 12220–12234.
- Taylor, T. A. (2002). 'Supply chain coordination under channel rebates with sales effort effects', *Management Science*, **48**, pp. 992–1007.
- Tihanyi, L., R. A. Johnson, R. E. Hoskisson and M. A. Hitt (2003). 'Institutional ownership differences and international diversification: The effects of boards of directors and technological opportunity', *Academy of Management Journal*, **46**, pp. 195–211.
- Tsai, W. (2002). 'Social structure of "coopetition" within a multiunit organization: Coordination, competition and intra-organizational knowledge sharing', *Organization Science*, **13**(2), pp. 179–190.
- Venaik, S., D. F. Midgley and T. M. Devinney (2004). 'A New Perspective on the Integration-Responsiveness Pressures Confronting Multinational Firms', *Management International Review*, **44**(1), pp. 15–48.
- Venaik, S., D. F. Midgley and T. M. Devinney (2005). 'Dual paths to performance: the impact of global pressures on MNC subsidiary conduct and performance', *Journal of International Business Studies*, **36**(6), pp. 655–675.
- Williams, L. J., J. R. Edwards and R. J. Vandenberg (2003). 'Recent Advances in Causal Modeling Methods for Organizational and Management Research', *Journal of Management*, **29**(6), pp. 903–936.
- Wiseman, F. and M. Billington (1984). 'Comment on a standard definition of response rates', *Journal of Marketing Research*, **21**, pp. 336–338.
- Zeller, R. E. and E. G. Carmines (1980). *Measurement in the social sciences*. Cambridge University Press, Cambridge.
- Zhao, X., J. Xie and W. J. Zhang (2002). 'The impact of information sharing and ordering co-ordination on supply chain performance', *Supply Chain Management*, **7**(1), pp. 24–40.

Adamantios Diamantopoulos is Professor and Head of the Department of International Marketing at the University of Vienna, Austria. His main research interests are in international marketing, marketing research, and research methodology and he is the author of some 200 publications in these areas. His work has appeared, among others, in the *Journal of Marketing Research*, *Journal of International Business Studies*, *Journal of the Academy of Marketing Science*, *International Journal of Research in Marketing*, *International Journal of Forecasting*, *Journal of International Marketing* and *Journal of Business Research*.

Judy A. Siguaw is the Founding Dean of Cornell-Nanyang Institute of Hospitality Management. She is also a Professor of Marketing in the School of Hotel Administration at Cornell University and holds a J. Thomas Clark Chair in Entrepreneurship and Personal Enterprise. She has published 40 journal articles, including those appearing in the *Journal of Marketing Research* and *the Journal of Marketing*, among others. She is co-author of four books.