



University of Navarra

MEASURES OF COLLABORATIVE  
VENTURE PERFORMANCE:  
AN ANALYSIS OF CONSTRUCT VALIDITY

Africa Ariño\*

*RESEARCH PAPER No 386*  
*April, 1999*

\* Professor of General Management, IESE

**Research Division**  
**IESE**  
University of Navarra  
Av. Pearson, 21  
08034 Barcelona - Spain

Copyright © 1999, IESE  
Do not quote or reproduce without permission

## **MEASURES OF COLLABORATIVE VENTURE PERFORMANCE: AN ANALYSIS OF CONSTRUCT VALIDITY**

### **Abstract**

This study presents a methodology to evaluate the construct validity of measures of collaborative venture (CV) performance and applies it to a sample of 69 CVs. After discussing content validity of the different measures, empirical results show that subjective measures of CV performance are reliable and present convergent validity. Results also support their discriminant validity with respect to other measures such as contractual changes, but not to longevity; no conclusive evidence is found with respect to survival.

---

I would like to thank Peter Bentler, José de la Torre, Dick Goodman, Elaine Mosakowski, and Bill Ouchi for their helpful comments on earlier drafts of this paper. Funding for this research was provided by IESE, International Graduate School of Management, University of Navarra (Barcelona, Spain).

## **MEASURES OF COLLABORATIVE VENTURE PERFORMANCE: AN ANALYSIS OF CONSTRUCT VALIDITY**

### **Introduction**

Construct validation research has been underemphasized in strategic management (Venkatraman and Grant, 1986). In particular, this is true of research on inter-firm collaborative ventures (1) (CVs) and, more specifically, of research on measures of CV performance (Parkhe, 1993 c). Whereas some work in the international management field has focused on the measurement of CV performance (Geringer and Hebert, 1991; Baird, Lyles and Reger, 1993), the validity of the underlying measures used in the literature –financial, objective, and subjective measures (Geringer and Hebert, 1991)– is still questionable.

Content validity, reliability, and convergent, discriminant, and criterion-related validity, are the main aspects of construct validity. Baird, Lyles and Reger's (1993) review of methods for evaluating CV performance may be considered as a first step in assessing the content validity of existing measures. However, a clear description of the dimensions that define this concept is still missing. Furthermore, no attempt has been made to estimate the empirical validity of such measures (Parkhe, 1993 c).

Therefore, assessing the construct validity of the measures of CV performance is important. Such a study is necessary for at least three reasons: first, to evaluate the soundness of inferences drawn from the relationships among CV performance and other theoretical constructs –inferences which are based upon the relationships among the measures of the constructs; second, to assist future researchers in the selection of their measures of CV performance; and, finally, to make valid recommendations to managers, based on the results from substantive research.

This paper begins with a brief review of the measures of CV performance used in the literature, followed by a presentation of the concept of construct validity and its various aspects. The definition of CV performance proposed herein allows an assessment of the content validity of the various measures of CV performance. We suggest that the subjective measures capture the concept better than the “objective” ones. Next, we present a methodology to assess the empirical aspects of construct validity. We follow this methodology to assess the empirical validity of the CV performance measures using a sample of CVs involving companies based in Spain. After discussing the results, we conclude the paper by indicating its limitations and its contributions.

---

(1) We define an inter-firm collaborative venture as an explicit agreement between two (or more) firms to collaborate in a limited aspect of their activity for a relatively long term, and which may or may not result in a separate organizational entity.

## Existing Measures of Collaborative Venture Performance

Geringer and Hebert (1991) classify the measures of CV performance used in the literature into three main groups: financial, objective, and subjective measures. *Financial measures* include various measures of profitability (Good, 1972; Renforth, 1974; Dang, 1977), growth (Good, 1972; Dang, 1977), and cost position (Renforth, 1974). Frequently used *objective measures* include stability measures such as longevity of the venture (Harrigan, 1988 b; Kogut, 1989), ownership or contract stability (Franko, 1971; Gomes-Casseres, 1987; Blodgett, 1992), and survival (Franko, 1971; Raveed, 1976; Harrigan, 1988 b). The most commonly used *subjective measure* is an overall assessment of the firm's satisfaction with CV performance (Killing, 1983; Schaan, 1983; Beamish, 1984; Geringer and Hebert, 1991; Parkhe, 1993 a, b). Other subjective measures used include the degree of fulfillment of strategic goals that the venture was aimed at covering (Parkhe, 1993 a, b), and the net spillover effects of the venture on other activities of the firm (Parkhe, 1993 b).

## Construct Validity

Construct validity represents “the correspondence between a construct (conceptual definition of a variable) and the operational procedure to measure [...] that construct” (Schwab, 1980: 5-6). Relationships among constructs are inferred from the observed relationships among their measures. Thus, the conclusiveness of these inferences depends upon the correspondence between constructs and measures (Schwab, 1980). Construct validity includes *content validity*, *reliability*, and *convergent*, *discriminant*, and *criterion-related validity* of the measures of the construct.

*Content validity* is “a qualitative type of validity where the domain of a concept is made clear and the analyst judges whether the measures fully represent that domain” (Bollen, 1989: 185). The domain of a concept is bounded by its theoretical definition, which should reflect the meanings associated with the concept in prior research (Bollen, 1989) and make its dimensions clear (Schwab, 1980; Bollen, 1989).

*Reliability* is defined traditionally as the consistency of measurement. An alternative definition, appropriate from a structural perspective, is that the reliability of a measure is the magnitude of the direct relations that all variables (except the error terms) have on that measure (Bollen, 1989).

*Convergent validity* is “the degree to which two or more attempts to measure the same concept [...] are in agreement” (Bagozzi and Phillips, 1982: 468). Measures of the same construct should present a large common variance. Convergence is especially important in areas in which there is a proliferation of measures presumably assessing the same construct (Schwab, 1980).

*Discriminant validity* is “the degree to which measures of distinct concepts differ” (Bagozzi and Phillips, 1982: 469). Measures of different constructs should share little variance. The analysis of discriminant validity may help to corroborate issues of content validity when it is suspected that some measures actually correspond to another concept.

*Criterion-related validity* is “the degree of correspondence between a measure and a criterion variable” (Bollen, 1989: 186). To evaluate it, a variable that is a standard against which to compare the measure in question is needed. Criterion validity may be *concurrent* –when the criterion exists at the same time as the measure– or *predictive* –when the criterion occurs in the future (Bollen, 1989).

## **Collaborative Venture Performance and Content Validity of its Measures**

Before assessing the content validity of CV performance measures, we need a definition of this concept. This definition has to meet two conditions. First, it should capture the meaning given to the concept in previous related studies. Second, it should specify the dimensions that mark the boundaries of the construct under study (Bollen, 1989). We will now review the interpretations of the term that appear in the literature, and suggest a definition of CV performance. Next, we will assess the content validity of those measures in the light of this definition.

### *The Concept of Collaborative Venture Performance*

An agreed-upon definition of CV performance is missing in the literature (Geringer and Hebert, 1991). Nonetheless, the view of goal accomplishment underlies most interpretations (Beamish, 1988; Hamel, Doz and Prahalad, 1989; Anderson, 1990; Baird, Lyles and Reger, 1993; Beamish and Delios, 1997). Thus, defining CV performance as the degree of accomplishment of CV goals is the first step in capturing the meaning given to this concept by previous researchers. If we are able to specify the goals of the parties involved in a CV, we will have identified the dimensions relevant to the concept. These dimensions have to be definite and yet flexible, so that the definition may be applied across various kinds of CVs, and to CVs at different stages of development.

#### *Levels of performance*

Drawing from the strategy literature, we may recognize three levels of performance that depend on the goals under consideration: *financial performance*, *operational performance*, and *organizational effectiveness* (Venkatraman and Ramanujam, 1986).

*Financial performance* is the relevant level of CV performance in those cases in which profitability is an explicit goal of a CV.

*Operational performance* “focuses on those key operational success factors that might lead to financial performance” (Venkatraman and Ramanujam, 1986: 804, emphasis as in the original). Operational performance can be measured by indicators of such key success factors (1).

*Organizational effectiveness* refers to the fulfillment of the organizational goals, taking into account the interests of multiple constituencies. Its measurement is the subject of open debate in the literature (Steers, 1975; Cameron and Whetten, 1983). The existence of several parties to the CV calls for a multiple-constituencies approach in evaluating organizational effectiveness. Baird, Lyles and Reger (1993) identify the following as potential key constituency groups likely to have interests in a CV: each of the partners to the CV as an independent organization, the CV management team, and the host government/community.

---

(1) Note that we depart here from Venkatraman and Ramanujam's (1986) assertion that the operational success factors of interest might lead to financial performance. Given that, in general, financial performance is not always relevant to CVs, we consider it sufficient to say that the operational success factors might lead to success understood as effectiveness –as we will show soon, effectiveness is the relevant level of CV performance.

### *Goals of the Multiple Constituencies Interested in a Collaborative Venture*

*The partners' goals.* With regard to the partners' goals, two issues have to be considered. First, each partner may have goals for the CV which are not shared by the other partner. CVs are organizational arrangements in which two or more sovereign organizations collaborate by combining their resources to pursue *shared* interests (Borys and Jemison, 1989). However, since each partner is indeed a sovereign organization, they are sure to have their own agenda (Habib, 1983; Buckley and Casson, 1988; Doz, 1988) with different goals for the CV. The shared interests are the *common* goals of the CV, while the goals that each firm has for the CV and does not share with its partner are the *private* goals (Ariño, 1995).

A second issue to consider with regard to the partners' goals is that both the shared and the non-shared goals may change over time (Moxon, Roehl and Truitt, 1988), yielding *emergent* goals that differ from the *initial* –common or private– ones. The partners' goals may change as a consequence of a closer interaction stemming from the start-up of a CV, as a consequence of shifts in their respective strategies (Franko, 1971; Harrigan, 1986, 1988a; Ariño and de la Torre, 1998) or as a consequence of environmental changes (Harrigan, 1985, 1986, 1988a; Ariño and de la Torre, 1998) that affect both firms, either equally or asymmetrically.

*The CV management's goals.* The management of the CV –if there is one– will in general embrace the partners' common goals when these are explicit; otherwise, it will embrace some other ideal deemed useful as a means to settle potential conflicts between the partners (Schaan and Beamish, 1988) and which, therefore, reflects the CV management's interpretation of the partners' common and private goals. The CV management may also develop *independent* goals –such as survival of the CV, or avoiding conflict with the partners– which may or may not be in conflict with the partners' goals (Schaan and Beamish, 1988). If these independent goals do not conflict with the partners' goals, pursuing them does not harm the performance of the CV. If they do, they may be considered as impediments that make the fulfillment of the partners' goals more difficult. In this case, the existence of independent goals may affect CV performance negatively. However, fulfillment of the CV management's independent goals is not an integral part of CV performance, given that a CV is not set up to accomplish these independent goals.

*The community's goals.* The host government or the community may have goals for the CV as well. These societal goals may be imposed on the CV (Harrigan, 1986; Baird et al., 1993) or, on the contrary, they may concur with the CV's aim. If societal goals act as a constraint, they will be reflected in the likewise constrained scope of the goals that the partners have for the CV; if they constitute a part of the CV's goals, they will be included in the partners' goals. Therefore, we do not need to consider them as a separate category when evaluating CV performance (1).

In sum, the interests that the different constituencies may have in a CV are integrated in the goals of the partners. The goals of a CV as an independent organization are the partners' common goals. The goals of the CV management are these same partners' common goals, or at least the management's interpretation of them. If the management develops some independent goals, these may affect the performance of the CV, but they are not an integral part of CV performance because a CV's *raison d'être* does not include fulfillment of these goals. Finally,

---

(1) This argument assumes that the community's goals are explicit. If not, then they will be difficult to ascertain, and also it will be difficult to evaluate their accomplishment.

the goals that the community may have for a CV are incorporated into the partners' goals, either as limits to these goals or as an integral part of them.

Thus, we may conclude from this discussion that the partners' common and private goals, whether they be initial or emergent, are the relevant ones when assessing CV performance from a strategic management perspective. Thus, we may define CV performance as *the degree of accomplishment of the partners' common and private, initial and emergent goals*. This definition suggests criteria to evaluate CV performance which are more parsimonious and operational than those proposed by Baird, Lyles and Reger (1993), and, also, flexible enough to be applied to different kinds of CVs at any of their various developmental stages.

### ***Assessing the Content Validity of the Various Measures of Collaborative Venture Performance***

Having established a clear definition of CV performance, we may next evaluate the content validity of financial, objective, and subjective measures of CV performance. We do so by judging the extent to which these measures reflect the degree of accomplishment of the partners' common and private, initial and emergent goals for a CV.

#### *Financial measures*

Financial measures of CV performance measure the degree of fulfillment of its financial goals. A CV is likely to have goals other than just financial ones (Anderson, 1990; Lorange and Roos, 1992; Baird et al., 1993). Thus, it seems that *financial indicators* are adequate measures of CV performance only in those cases in which financial performance is a prominent goal. While this might have been the case with early CVs, it does not seem to be so in more recent ones. The researcher cannot assume that financial goals are prominent and use financial indicators to draw comparisons of performance across different kinds of CVs. As a matter of fact, the use of financial indicators as measures of CV performance was more frequent in earlier times than it is nowadays (see Geringer and Hebert's [1991] review of financial measures of CV performance).

#### *Objective measures*

Objective measures of CV performance measure its operational performance or stability. They may postulate content validity as a measure of CV performance, insofar as each one of the stability conditions –longevity, no ownership or contractual changes, survival– is necessary and sufficient to ensure the good performance of the CV. Next, we will analyze whether this is the case.

To assess whether *longevity* is a necessary and sufficient condition for success we need to bear in mind that different goals are fulfilled in different time frames –for instance, CVs set up to access knowledge are temporary in nature, while those seeking the benefits from scale are more enduring (Porter and Fuller, 1986).

In the case of CVs meant to endure, a prolonged existence is a necessary though not sufficient condition for their success. The problem is that some CVs may last a long time, even if unsuccessful, due to the existence of some barrier to exit, or to organizational inertia. Thus, the fact that they last a long time does not mean that they are successful. However, if

they last a period of time shorter than necessary for the fulfillment of goals, we may say with certainty that they were unsuccessful.

In the case of CVs meant to be temporary, they may actually last a short time, either because they are successful or because they fail to meet the management challenge characteristic of CVs (Devlin and Bleackley, 1988). Additionally, if they last a long time, this could be due again to the existence of some barrier to exit, or to organizational inertia or, on the contrary, to the fact that success leads the partners to extend the duration of their collaboration.

We may conclude that longevity does not present content validity as a measure of CV performance. The exception is the case of CVs meant to endure and actually having an existence shorter than necessary, which can be said to have failed in fulfilling their goals.

*Ownership or contractual changes* may respond to a number of causes (Gomes-Casseres, 1987). First, the CV may not have been set up adequately, and the partners may decide to change its structure in an attempt to compensate for its failure. Second, the CV may have been designed as an intermediate step, in which case a change in structure would have been anticipated and would indicate the CV's success (see also Kogut, 1991). Finally, environmental changes may modify the conditions for CV success, and a change in the basic CV contract may reflect a successful adaptation to the new conditions. Thus, ownership or contractual changes do not tell us anything conclusive about CV success, and we may conclude that they do not present content validity as a measure of CV performance.

Finally, the discussion on *survival* can be referred to the discussion on the two previous measures. If the CV ends by dissolution, the discussion on *longevity* applies: *survival* does not present content validity as a measure of CV performance, with the exception of CVs meant to endure and actually having an existence shorter than necessary, i.e., those which can be said to have failed in fulfilling their goals. If the CV is acquired by one of the partners, the discussion on *ownership changes* applies: *survival* does not present content validity as a measure of CV performance.

In sum, the objective measures can be related to performance only in particular instances. These instances require some knowledge about the temporal nature of CV goals. Using them as measures of performance without this previous knowledge would be misleading.

### *Subjective measures*

Subjective measures of CV performance measure the degree of fulfillment of several goals from the perspective of one of the partners. A caveat of these measures is that they do not capture the degree of fulfillment of the other partner's private goals. Depending on the purpose of using a measure, both partners' evaluations might be necessary. However, this is no impediment to the assessment of content validity. Furthermore, Geringer and Hebert (1991) found that the use of a single respondent for collecting performance data is appropriate.

A partner's *satisfaction with the overall performance* of the CV is one of the most frequently used measures of CV performance (Killing, 1983; Schaan, 1983; Beamish, 1984; Geringer and Hebert, 1991; Parkhe, 1993 a, b). Being an overall assessment, this measure evaluates the fulfillment of a partner's current common and private goals –which may be the initial ones or others that emerged during the course of the CV.

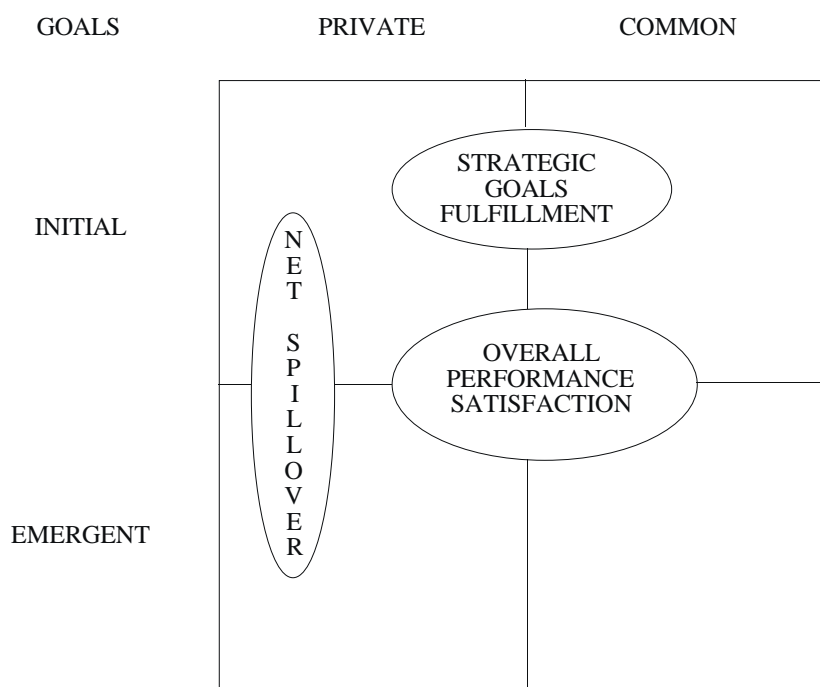


Parkhe (1993 b) measured CV performance as the *degree of fulfillment of strategic goals*. This measure allows respondents to rate the importance of the particular items suggested by the researcher for their CV. In particular, Parkhe developed his list from a broader list generated by Contractor and Lorange (1988). The formation of a CV responds to some strategic goal which constitutes its initial purpose. Thus, this measure captures the degree of fulfillment of the initial goals, be these common or private.

Finally, another measure is the *net spillover effects* of the CV on other activities of the company. It is defined as the difference between positive spillover effects –such as the application of know-how gained from CV activities to non-CV operations– and negative spillover effects –such as competition between the CV and other parent firm operations (Parkhe, 1993 a). Spillover effects pertain to, by definition, private goals, either initial or emergent. Thus, this measure captures the degree of fulfillment of a partner's private goals for the CV.

To summarize, the various subjective measures of CV performance present different degrees of content validity. A partner's satisfaction with the CV's overall performance evaluates the degree of fulfillment of this partner's goals –common and private, initial and emergent–, presenting the highest content validity among the subjective measures. The assessment of the accomplishment of strategic goals evaluates the degree of fulfillment of initial goals –common and private. Finally, the estimation of spillover effects evaluates the degree of fulfillment of private goals –initial and emergent (see Figure 1).

**Figure 1**  
**Mapping of subjective measures of collaborative venture performance**



## Empirical Validity of Collaborative Venture Performance Measures

While assessment of content validity relies on conceptual arguments, evaluation of the other aspects of construct validity has an empirical component. We follow a structural approach, similar to the one Hoskisson, Hitt, Johnson, and Moesel (1993) used in their study of an entropy measure of diversification strategy. This procedure requires that we specify a measurement model. The conceptual discussion on content validity may be useful for this purpose.

Our main conclusions in the previous discussion were three. One conclusion was that satisfaction with the overall performance of the CV is the measure with the strongest content validity. Hence, we should expect this measure to show the highest reliability among the other measures:

*Hypothesis 1:* Overall performance satisfaction is the measure of CV performance that presents the highest reliability.

A second conclusion was that subjective measures have content validity as a measure of CV performance. Thus, we should expect them to measure one same unobservable phenomenon; therefore:

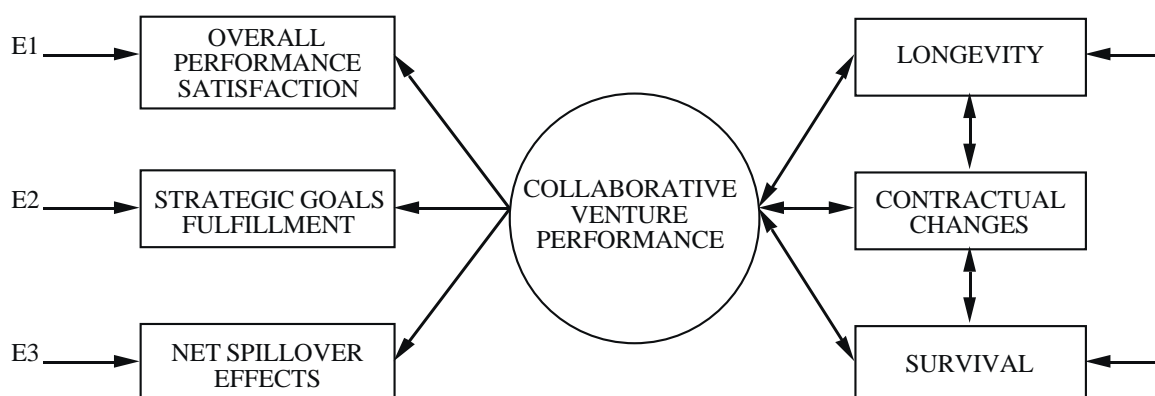
*Hypothesis 2:* Subjective measures present convergent validity as a measure of CV performance.

Finally, we also concluded that objective measures do not have content validity as measures of CV performance. Consequently, we should expect them to measure something different than the subjective measures do; hence:

*Hypothesis 3:* Subjective measures of CV performance have discriminant validity with respect to objective measures.

From the last two hypotheses –and consistent with the first one– it follows that the measurement model to be tested is one that incorporates subjective measures as indicators of CV performance, and objective measures as constructs that are allowed to freely correlate with CV performance. Figure 2 represents this model in a graphic form. Reliability, convergent validity, and discriminant validity of the subjective measures will be assessed on the grounds of this measurement model.

**Figure 2**  
Measurement model of collaborative venture performance



Assessment of the criterion-related validity of a measure requires the existence of a criterion variable that can be used as a standard against which to compare that measure. As stated by Bollen (1989: 188), “for many measures no criterion is available.” We believe this is the case for CV performance. Usually, CV performance is treated in the literature as an outcome that needs to be explained, and it has been associated with a good number of factors that have an incidence on CV performance. For instance, Parkhe (1993 a) shows that perception of opportunistic behavior and commitment of non-recoverable investments in the CV have an influence on CV performance. He also shows that this is associated with the pattern of payoffs and the shadow of the future. Mohr and Spekman (1994) find that primary characteristics of CV success are: partnership attributes of commitment, coordination and trust; communication quality and participation; and the conflict resolution technique of joint problem-solving. Dussage and Garrette (1995) identify the following as having an impact on CV performance: partner asymmetries, distribution of ownership and control in the case of joint ventures, scope and breadth of purpose, industry structure and competitive context. On the one hand, the variety of factors associated with CV performance and the lack of a model of CV success make it difficult to select one criterion variable that would suggest concurrent validity. Besides, today's performance is the consequence of past conditions, and it is not necessarily associated with today's conditions. On the other hand, the general use of CV performance as a dependent variable to be explained precludes us from identifying a variable that is generally accepted to be a consequence of CV performance; thus, predictive validity cannot be assessed either.

## Methods

### Sample

The sample for this study was drawn from Spanish firms that appeared in the Funk and Scott's (F&S) Countries Index - Europe (1986-1992) as having announced their engagement in venturing activities. The selected time period begins with Spain's accession to the European Community (1986), and concludes with the establishment of the Single European Market (1992), a period that can be expected *a priori* to present a high venturing activity. We selected as target industries those with a higher number of CVs (see Table 1).

**Table 1. Industries and responses**

Industry description	No. of responses	% of responses	Questionnaires mailed	% of total mailed
Energy (petroleum and electricity)	6	6.6	19	10.1
Chemicals	14	15.4	15	7.9
Machinery except electrical	5	5.5	7	3.7
Electronic equipment	4	4.4	7	3.7
Transportation equipment	4	4.4	5	2.6
Transportation	6	6.6	8	4.2
Communications	0	0.0	2	1.1
Financial services	37	40.6	95	50.3
Other services	15	16.5	31	16.4
TOTAL	91	100.0	189	100.0

The target informant was the person in each firm most directly related to the CV. Sacrificing quantity for quality, we sent out questionnaires only to those firms in which we were able to identify this person. Of the 189 mailed questionnaires 91 (48%) were returned. We attribute this rather high response rate to the care taken in identifying the target respondent and in the follow-up process. For this process we followed the procedure Dillman [1978] suggests, supplemented with phone calls. More than 63% of the informants had participated in the negotiation of their firm's CV, and they had been involved with the CV for 4.9 years on average.

For the purpose of this study, eighteen of the returned questionnaires were incomplete. In four cases, we received the answers from each side of the CV dyad. To insure independent data points we dropped one of the parties to each CV, selecting it randomly by the flip of a coin. This left a final sample of 69 questionnaires for this study.

## **Measures**

### *Subjective measures of CV performance*

*Overall performance satisfaction* is a five-point scale measuring the informants' assessment of the extent to which their firm is satisfied with the overall performance of the CV (Harrigan, 1988 b; Geringer and Hebert, 1991; Parkhe, 1993 c).

*Strategic goals fulfillment* is a composite measure of the *importance* multiplied by the degree of *fulfillment* of specific strategic goals a CV may be meeting (Parkhe, 1993 a, c). *Importance* is a five-point scale measuring the informants' assessment of the extent to which each of the possible goals embraced by the firm for a CV was important to their firm. *Fulfillment* is a five-point scale measuring the informants' assessment of the extent to which each of the identified strategic goals for the CV was fulfilled. The list of strategic goals was adapted from the literature (Contractor and Lorange, 1988; Parkhe, 1993 a, c; Ring and Van de Ven, 1994). We added an open-ended item so that the informants might consider additional goals not listed. We also provided the category "not applicable" as an option.

*Net spillover effects* is a five-point scale measuring the informants' assessment of the difference between positive and negative effects of the CV on other activities of their firm. This measure was developed by Parkhe (1993 c).

### *"Objective measures"*

*Longevity* measures the number of years elapsed between the moment the CV was formed and the moment it was terminated, or November 1994 (the time when data collection was carried out) if the CV was still operating then.

*Contractual changes* measures the number of changes in the ownership structure or in the CV contract.

*Survival* measures whether the CV was still operating in November 1994.

### *Control variables*

*Temporal horizon.* From our discussion on the content validity of “objective measures” we may expect their relationship to CV performance to vary across CVs involving different temporal horizons. A CV is considered to be temporary if there is a set date for its termination, and enduring if there is no such date.

*Equity ownership.* Equity CVs can entail very high exit costs (Gulati, 1995). The relationship between CV performance and the “objective measures” may not be univocal in equity CVs due to the exit barriers involved, but this may not hold in non-equity CVs. A CV is considered as an equity CV if it involves shared ownership or equity exchanges between the partners, and as a non-equity CV otherwise.

### *Estimation and Results*

We might expect that informants who were satisfied with their CV would be more prompt to respond to our questionnaire than those who were dissatisfied. It is not feasible to conduct a direct check for a non-response bias of this sort; thus we checked for differences between early and late respondents (Armstrong and Overton 1977). A one-way between-groups ANOVA was used for the continuous variables, and likelihood analysis for the dummy variables. A significant difference (.05 level) between early and late respondents was found for *survival*: there were significantly more survivors among late than early respondents. This means that although the number of non-survivors in the sample is low relative to survivors, our sample includes a significantly high proportion of non-survivors. This bias needs to be taken into consideration when interpreting the results.

Table 2 shows the descriptive statistics and correlations of the variables in the study for the full sample. We performed the tests for reliability, convergent validity, and discriminant validity with the maximum likelihood (ML) estimation method of the EQS structural equation package (Bentler, 1992a). We selected this method because it has been proved to reject models more frequently than the generalized least squares method (GLS) with small-sized samples (Hu, Bentler, and Kano, 1992), thus providing a more trustworthy test for our case. Given our earlier discussion, we may expect somewhat different results for temporary and enduring CVs, as well as for equity and non-equity ones. Thus, in addition to the tests for the full sample we performed tests for these two sets of sub-samples.

**Table 2. Descriptive statistics and correlations for the full sample**

Variables	Mean	s.d.	1	2	3	4	5
1. Overall performance satisfaction	3.29	1.07					
2. Strategic goals fulfillment	9.84	3.72	.52				
3. Net spillover effects	3.61	0.83	.71	.53			
4. Longevity	3.73	1.83	.31	.32	.29		
5. Contractual changes	0.41	0.69	-.01	-.02	-.15	.03	
6. Survival	0.74	0.44	.45	.31	.32	.43	-.06

Correlations above  $r = .24$  significant at  $p < .05$

Spearman rank correlations are reported where ordinal data are used

### Reliability

The reliability coefficient is the squared multiple correlation coefficient for the variable (Bollen, 1989:221). The reliability coefficients were calculated from the measurement model (Figure 2), except for the equity CVs sub-sample, for which the coefficients were calculated from the revised model (see origin of this model in the sub-section on convergent validity). In the full sample case, the coefficient was .71 for *overall performance satisfaction*, meaning that 71% of the variance of this variable is explained by the latent construct *CV performance*. The coefficient was .36 for *strategic goals fulfillment*, and .71 for net *spillover effects*. As was the case in the full sample, the reliability coefficient was lower for *strategic goals fulfillment* than for the other variables in all of the sub-samples (see Table 3).

**Table 3. Reliability coefficients of subjective measures of performance**

SAMPLE USED	N	Overall performance satisfaction	Strategic goals fulfillment	Net spillover effects
Full sample	69	.71	.36	.71
Temporary CVs	31	.66	.47	.70
Enduring CVs	38	.67	.30	.80
Equity CVs (revised model)	31	.58	.29	.90
Non-equity CVs	38	.72	.33	.69

### Convergent validity

Convergent validity is examined through three tests involving the measurement model, which is built in such a way that the factor representing CV performance is allowed to freely correlate with each of the “objective measures.” Table 4 shows the results of the tests for convergent validity performed on the full sample and on the four sub-samples.

*Test 1: Overall goodness-of-fit.* The goodness-of-fit is assessed with the chi-square test, and the Comparative Fit Index (CFI) developed by Bentler (1990 a). These tests compare the hypothesized model with the null model, where there is complete independence among all the variables employed. We chose the CFI among other indexes of fit, as the CFI takes into account sample size, and thus it is preferable (Bentler, 1990 b). A model chi-square that is statistically non-significant rejects the null hypothesis of complete independence, thus showing a good fit of the hypothesized relationships to the data. Magnitudes of .90 or greater for the CFI also evidence a good fit of the model to the data (Bentler, 1992 b).

The results with the full sample show that the model chi-square is statistically non-significant, thus rejecting the null hypothesis of complete independence. The CFI is above the required level. These results show a good fit of the hypothesized relationships to the data.

The results hold in the non-equity CVs sub-sample as well as in the temporary and enduring CVs. However, estimation of the model with the equity CVs sub-sample shows the model chi-square to be statistically significant and the CFI to be below the required level. These results suggest that the fit of the model to this sub-sample is not good enough. The subsequent tests of convergent validity will indicate whether model improvement may come from adding or from dropping some parameters in the model.

**Table 4. Results of convergent validity tests**

	Test 1		Test 2			Test 3						
	X2	CFI	Satisfaction	Goals fulfillment	Spillover	Satisfaction/Goals fulfillment		Satisfaction/Spillover		Goals fulfillment/Spillover		
			loading	loading	loading	$\Delta$ X2	$\Delta$ CFI	$\Delta$ X2	$\Delta$ CFI	$\Delta$ X2	$\Delta$ CFI	
Proof of convergent validity	p > .05	> .9	p < .05	p < .05	p < .05	p > .05	~0	p > .05	~0	p > .05	~0	
<b>SAMPLE USED</b>	<b>N</b>											
Full sample	69	p = .48	1.0	p < .001	p < .001	p < .001	p < .01	-.08	p < .2	-.01	p < .001	-.11
Temporary CVs	31	p = .63	1.0	p < .001	p < .001	p < .001	p < .03	-.01	p < .80	0	p < .01	-.12
Enduring CVs	38	p = .17	.95	p < .001	p < .001	p < .001	p < .10	-.05	p < .20	-.01	p < .05	-.06
Equity CVs	31	p = .01	.84	p < .01	p < .05	p < .001	p < .10	-.03	p < .30	0	p < .10	-.04
Equity CVs, revised model	31	p = .23	.97									
Non-equity CVs	38	p = .67	1.00	p < .01	p < .05	p < .01	p < .05	-.03	p < .50	0	p < .03	-.06

*Test 2: Factor loadings of the observed variables.* There is evidence of convergent validity if a z-test shows all of the observed variables measuring a construct to be significant at least at the .05 level (Bentler, 1992 a).

All of the factor loadings for the indicators of CV performance are significant below the .001 level when we estimate the model with the full sample. When we estimate the model with the sub-samples, all of the factor loadings are significant at least at the .05 level. As this holds in the case of equity CVs, we cannot expect that dropping any of the CV performance indicators will improve the fit of the model to the data in this sub-sample.

*Test 3: Comparison of baseline model to more restricted models.* The baseline model is compared to other models in which factor loadings of measures of the same construct are restricted to equal each other (Hoskisson et al., 1993). We compared the baseline model to other more restricted models in which the factor loadings of measures of performance are constrained to be equal two at a time (*overall performance satisfaction* and *strategic goals fulfillment*; *overall performance satisfaction* and *net spillover effects*; and *strategic goals fulfillment* and *net spillover effects*). A non-significant increase in chi-square as well as invariance in the CFI provide evidence for convergent validity (Byrne, 1994).

The results obtained with the full sample show a non-significant chi-square increase in the model where the factor loadings for *overall performance satisfaction* and *net spillover effects* were forced to be equal. The increase in chi-square was significant in the other two models. The CFI variation was not substantial in any case.

The same results hold for the temporary CVs and the non-equity CVs sub-samples. In the enduring CVs sub-sample only the model where the factor loadings for strategic goals fulfillment and net spillover effects were forced to be equal shows a significant increase in chi-square. In the case of the equity CVs sub-sample, none of the models has a significant chi-square increase. None of the models shows important CFI variations in any of the sub-samples. The results of this test for the equity-based sub-sample confirm that model improvement will not come from dropping any of the CV performance indicators.

These results suggest that the measurement model shows a very good fit to the data—evidence of convergent validity—except in the case of the equity-based CVs sub-sample, calling for a model respecification in this case. Tests 2 and 3 suggest that dropping any of the CV performance indicators will not improve the model fit. The EQS output suggested that the model would improve by not allowing *longevity* and *contractual changes* to correlate, and by allowing *survival* to freely correlate with the error term associated to *net spillover effects*. The revised model incorporates these changes. This model is statistically non-significant and the CFI is above the acceptance level (see Table 4). Thus, we may accept the revised model as one in which convergent validity of the indicators of CV performance can be substantiated for equity CVs. The test for discriminant validity for this sub-sample is done using the revised model as the baseline model.

#### *Discriminant validity*

Discriminant validity is assessed with the chi-square difference test, comparing the baseline model with a more restricted model in which the correlation between the two constructs under examination is constrained to equal 1.0 (Joreskog, 1971). A significantly higher chi-square for the model in which the correlation is restricted would indicate a non-perfect correlation between the constructs, which is evidence of discriminant validity



(Bagozzi and Phillips, 1982). This evidence is also provided by an important variation in the goodness-of-fit measures (Byrne, 1994). This test is done for one pair of constructs at a time (Anderson and Gerbing, 1988).

Table 5 reports the results of comparing the measurement model to those in which the correlations between *CV performance*, and *longevity*, *contractual changes*, and *survival* are constrained to equal 1.0 one at a time. When the estimation is done with the full sample, the restricted model in which the correlation between *CV performance* and *longevity* is set to 1 shows a non-significant increase in chi-square and no variation in CFI. When the restriction is imposed on the correlation between *CV performance* and *contractual changes*, the increase in chi-square is significant and the CFI shows a substantial decrease. The last model that restricts the correlation between *CV performance* and *survival* did not converge after 90 iterations; model underidentification –a potential cause of non-convergence– was not detected.

**Table 5. Results of discriminant validity test**

Proof of discriminant validity	Longevity		Contractual changes		Survival	
	$\Delta X^2$	$\Delta CFI$	$\Delta X^2$	$\Delta CFI$	$\Delta X^2$	$\Delta CFI$
	p < .05	>>0	p < .05	>>0	p < .05	>>0
<b>SAMPLE USED</b>	<b>N</b>					
Full sample	69	p < .30	0	p < .001	-.70	n.c.
Temporary CVs	31	p < .20	0	p < .001	-.46	n.c.
Enduring CVs	38	p < .95	.01	p < .001	-.82	n.c.
Equity CVs, revised model	31	P < .90	.01	P < .01	.03	n.c.
Non-equity CVs	38	p < .10	0	p < .001	-1.00	n.c.

n.c. non-convergence

Similar results hold for all four sub-samples. One exception is the model that constrains the correlation between *CV performance* and *contractual changes* when estimated with the equity CVs sub-sample. In this case, although the increase in chi-square is significant, the CFI variation is not important.

These results indicate that the subjective measures of CV performance have discriminant validity with respect to *contractual changes* and survival, but not with respect to longevity.

## Discussion

This study has three main findings. First, that *overall performance satisfaction* and *net spillover effects* are more reliable measures of CV performance than *strategic goals fulfillment*. Second, that taken together the subjective measures of CV performance are good measures of this concept, although *overall performance satisfaction* and *net spillover effects* are more similar to each other than any of them is to *strategic goals fulfillment*. Third, that subjective measures of CV performance show discriminant validity with respect to *contractual changes* and *survival*, but not with respect to *longevity*.

The results obtained from the sub-samples are similar to those obtained from the full sample. The only substantial difference is that subjective measures show a greater convergence in the case of equity CVs and enduring CVs than in the other cases, meaning that the measures have a greater common variance in these cases than in the other sub-samples.

We had anticipated (Hypothesis 1) that *overall performance satisfaction* would be the measure of CV performance with the highest reliability. In fact, *net spillover effects* is as reliable as *overall performance satisfaction*. The result does not invalidate our hypothesis but rather expands it, as both measures display an equal degree of high reliability.

The convergent validity tests confirm that subjective measures have this kind of validity (Hypothesis 2). The tests also indicate that *overall performance satisfaction* and *net spillover effects* are better measures of CV performance than *strategic goals fulfillment*, the first two having a greater common variance than either of them does with *strategic goals fulfillment*. This could be due to the fact that the former two measures have a more global nature, while the latter is more detailed and composite in nature. However, it may also have to do with the different content validity of these measures. From our initial discussion on content validity, we concluded that *overall performance satisfaction* evaluates the degree of fulfillment of common and private, initial and emergent goals; *strategic goals fulfillment* evaluates the degree of fulfillment of initial goals, common or private; and *net spillover effects*, that of private goals, initial or emergent.

In light of these conclusions, and taking into account that *net spillover effects* is as reliable a measure as *overall performance satisfaction*, the empirical results suggest that the initial goals lose importance as new goals develop from a closer partner interaction. This interpretation is a plausible explanation of why *strategic goals fulfillment* is less similar to the other measures. It is also consistent with a recent research stream that shows the importance of evolutionary processes in CVs, processes that lead to a constant re-evaluation and adaptation of the goals and terms of CVs (Ring and Van de Ven, 1994; Doz, 1996; Ariño and de la Torre, 1998).

The discriminant validity tests show that *contractual changes* and *survival* measure something other than CV performance. This result does not hold for *longevity*. Thus, discriminant validity of the subjective measures (Hypothesis 3) receives mixed support. One explanation for the unexpected result for *longevity* may be that, empirically, only successful alliances are long-lived, and that short-lived CVs are either failed CVs that have been dissolved or CVs still in their first years of operation which have not yet had enough time to be considered as successful. This explanation is consistent with the fact that survival is not a good proxy for CV performance: there may be CVs in operation not yet considered successful because they have not passed the test of time.

However, we cannot overlook the bias towards non-survivors in the sample. Consistent with our results, Geringer and Hebert (1991) found that subjective measures were strongly correlated with *longevity*, and only weakly with *contractual changes*. Contrary to our results, they found that the correlation between subjective measures and *survival* was strong. Therefore, we cannot be conclusive regarding the discriminant validity of subjective measures with respect to *survival*.

## Conclusion

The main contribution of this study is to provide a methodology for assessing construct validity of CV performance measures. The results demonstrate that subjective measures of CV performance show convergent validity, although *overall performance satisfaction* and *net spillover effects* are more reliable than *strategic goals fulfillment*. The results also show that while *longevity* may be safely used as a proxy for CV performance when access to subjective measures is restricted, *contractual changes* may not. Regarding the appropriateness of using *survival* as such a proxy, we hold our judgement because of the sample bias in this variable.

We have to acknowledge that this study is limited in that it is confined to the sample used. Replication of the study with a different sample, ideally a larger one, is necessary to achieve generalizability of the results. One caveat is that the study is based upon the assumption that performance refers to goals accomplishment, an assumption that is common among researchers embracing the strategic management perspective, but that may not be shared by those from other disciplines. In particular, some theorists may adhere to the assumption that an organization's main goal is survival. In interpreting the results of this study, we need to keep in mind the assumptions upon which it is based.

Despite its limitations, this study provides guidelines for researchers engaged in understanding CV performance. Its results help us interpret past research on the subject, and will assist future researchers in their selection of measures of CV performance. □

## References

- Anderson, Erin (1990) 'Two firms, one frontier: On assessing joint venture performance', *Sloan Management Review*, 32, Winter, pp. 19-30.
- Anderson, James C. and David W. Gerbing (1982). 'Structural equation modeling in practice: A review and recommended two-step approach', *Psychological Bulletin*, 103, pp. 411-423.
- Ariño, Africa (1995). *Inter-firm Collaborative Ventures: Performance and Cooperative Behavior*. Unpublished doctoral dissertation, University of California, Los Angeles.
- Ariño, Africa and José de la Torre (1998). 'Learning from failure: Towards an evolutionary model of collaborative ventures', *Organization Science*, 9(3), pp. 306-325.
- Armstrong, J. Scott and Terry S. Overton (1977). 'Estimating nonresponse bias in mail surveys', *Journal of Marketing Research*, 14, pp. 396-403.
- Bagozzi, Richard P. and Lynn W. Phillips (1982) 'Representing and Testing Organizational Theories: A Holistic Construal,' *Administrative Science Quarterly*, 27, pp. 459-489.
- Baird, Inga S., Marjorie A. Lyles, and Rhonda K. Reger (1993) 'Evaluation of cooperative alliances: Integration and future directions', paper presented at the Academy of Management Annual Meeting, Atlanta, GA.
- Beamish, Paul W. (1984). *Joint Venture Performance in Developing Countries*. Unpublished doctoral dissertation, University of Western Ontario.
- Beamish, Paul W. (1988). *Multinational Joint Ventures in Developing Countries*. Routledge, New York.
- Beamish, Paul B. and Andrew Delios. 1997. 'Improving joint venture performance through congruent measures of success.' In P.W. Beamish and J.P. Killing (eds.), *Cooperative Strategies: European Perspectives*. The New Lexington Press, San Francisco, CA, pp. 103-127.
- Bentler, Peter M. (1990 a). 'Comparative fit indexes in structural models', *Psychological Bulletin*, 107, pp. 238-246.
- Bentler, Peter M. (1990 b). 'Fit indexes, Lagrange multipliers, constraints changes, and incomplete data in structural models', *Multivariate Behavioral Research*, 25, pp. 163-172.
- Bentler, Peter M. (1992 a). *EQS: Structural Equations Program Manual*. BMDP Statistical Software, Los Angeles.
- Bentler, Peter M. (1992 b). 'On the fit of models to covariances and methodology to the *Bulletin*.' *Psychological Bulletin*, 112, pp. 400-404.
- Blodgett, Linda Longfellow (1992). 'Factors in the instability of international joint ventures: An event history analysis', *Strategic Management Journal*, 13, pp. 475-481.

- Bollen, Kenneth A. (1989). *Structural Equations with Latent Variables*. Wiley, New York.
- Borys, Brian and David B. Jemison (1989). 'Hybrid arrangements as strategic alliances: Theoretical issues in organizational combinations', *Academy of Management Review*, 14(2), pp. 234-249.
- Buckley, Peter J. and Mark Casson (1988). 'A theory of cooperation in international business.' In Contractor, F. J. and P. Lorange (eds.), *Cooperative Strategies in International Business*. Lexington Books, Lexington, MA, pp. 31-53.
- Byrne, Barbara M. (1994). *Structural Equation Modeling with EQS and EQS/Windows*. Sage, Thousand Oaks, CA.
- Cameron, K. S. and D. A. Whetten (1983). 'Organizational effectiveness: One model or several?' In K. S. Cameron and D. A. Whetten (eds.), *Organizational Effectiveness: A Comparison of Multiple Methods*. Academic Press, New York, pp. 1-24.
- Contractor, Farok J., and Peter Lorange (1988). 'Why should firms cooperate? The strategy and economic basis for cooperative ventures.' In Contractor, F. J. and P. Lorange (eds.), *Cooperative Strategies in International Business*. Lexington Books, Lexington, MA, pp.3-30.
- Dang, Tran Thanh (1977). *Ownership, Control and Performance of the Multinational Corporation: A Study of U.S. Wholly-Owned Subsidiaries and Joint Ventures in the Philippines and Taiwan*. Unpublished doctoral dissertation, University of California, Los Angeles.
- Devlin, Godfrey and Mark Bleackley (1988). 'Strategic alliances – Guidelines for success', *Long Range Planning*, 21, pp. 18-23.
- Dillman, D. A. (1978). *Mail and Telephone Surveys: The Total Design Method*. John Wiley, New York.
- Doz, Yves L. (1988). 'Technology partnerships between larger and smaller firms', *International Studies of Management and Organization*, 17(4), pp. 31-57.
- Doz, Yves L. (1996). 'The evolution of cooperation in strategic alliances: Initial conditions or learning processes?', *Strategic Management Journal*, 17, special summer issue, pp. 55-84.
- Dussage, Pierre and Bernard Garrette (1995). 'Determinants of success in international strategic alliances: Evidence from the global aerospace industry', *Journal of International Business Studies*, 26(3), pp. 505-530.
- Franco, Lawrence G. (1971). *Joint Venture Survival in Multinational Corporations*. Praeger, New York.
- Geringer, J. Michael and Louis Hebert (1991). 'Measuring performance of international joint ventures', *Journal of International Business Studies*, 22, pp. 249-264.
- Gomes-Casseres, Benjamin (1987). 'Joint venture instability: Is it a problem?', *Columbia Journal of World Business*, Summer, pp. 97-107.

- Good, Loretta (1972). *United States Joint Ventures and Manufacturing Firms in Monterrey, Mexico: Comparative Styles of Management*. Unpublished doctoral dissertation, Cornell University.
- Gulati, Ranjay (1995). 'Does familiarity breed trust? The implications of repeated ties for contractual choice in alliances', *Academy of Management Journal*, 38, pp. 85-112.
- Habib, Ghazi M. (1983). *Conflict Measurement in the Distribution Channel of Joint Ventures: An Empirical Investigation*. Unpublished doctoral dissertation, Texas Tech University, Lubbock, Texas.
- Hamel, Gary, Yves L. Doz, and C. K. Prahalad (1989). 'Collaborate with Your Competitors – and Win,' *Harvard Business Review*, January-February, pp. 133-139.
- Harrigan, Kathryn Rudie (1985). *Strategies for Joint Ventures*. Lexington Books, Lexington, MA.
- Harrigan, Kathryn Rudie (1986). *Managing for Joint Venture Success*. Lexington Books, Lexington, MA.
- Harrigan, Kathryn Rudie (1988 a). 'Joint ventures and competitive strategies', *Strategic Management Journal*, 9, pp. 141-158.
- Harrigan, Kathryn Rudie (1988 b). 'Strategic alliances and partner asymmetries'. In Contractor, F. J. and P. Lorange (eds.), *Cooperative Strategies in International Business*. Lexington Books, Lexington, MA.
- Hoskisson, Robert E., Michael A. Hitt, Richard A. Johnson, and Douglas D. Moesel (1993). 'Construct validity of an objective (entropy) categorical measure of diversification strategy', *Strategic Management Journal*, 14, pp. 215-235.
- Hu, Li-tze, P. M. Bentler and Yutaka Kano (1992). 'Can test statistics in covariance structure analysis be trusted?' *Psychological Bulletin*, 112(2), pp. 351-362.
- Joreskog, K. G. (1971). 'Simultaneous factor analysis in several populations', *Psychometrika*, 36, pp. 409-426.
- Killing, J. Peter (1983). *Strategies for Joint Venture Success*. Praeger, New York.
- Kogut, Bruce (1989). 'The stability of joint ventures: Reciprocity and competitive rivalry', *Journal of Industrial Economics*, 38, pp. 183-198.
- Kogut, Bruce (1991). 'Joint ventures and the option to expand and acquire', *Management Science*, 37, pp. 19-33.
- Lorange, Peter and Johan Roos (1992). *Strategic Alliances. Formation, Implementation, and Evolution*. Blackwell, Cambridge, MA.
- Mohr, Jakki and Robert Spekman (1994). 'Characteristics of partnership success: Partnership attributes, communication behavior, and conflict resolution techniques', *Strategic Management Journal*, 15(2), pp. 135-152.

- Moxon, Richard W., Thomas W. Roehl, and J. Frederick Truitt (1988). 'International cooperative ventures in the commercial aircraft industry: Gains, sure, but what's my share?' In Contractor, F. J. and P. Lorange (eds.), *Cooperative Strategies in International Business*. Lexington Books, Lexington, MA.
- Parkhe, Arvind (1993 a). 'The structuring of strategic alliances: A game-theoretic and transaction-cost examination of interfirm cooperation', *Academy of Management Journal*, 36, pp. 794-829.
- Parkhe, Arvind (1993 b). ' "Messy" research, methodological predispositions, and theory development in international joint ventures', *Academy of Management Review*, 18, pp. 227-268.
- Parkhe, Arvind (1993 c). 'Partner nationality and the structure-performance relationship in strategic alliances', *Organization Science*, 4, pp. 301-324.
- Porter, Michael E. and Mark B. Fuller (1986). 'Coalitions and global strategy.' In M.E. Porter (ed.), *Competition in Global Industries*. Harvard Business School Press, Cambridge, MA.
- Raveed, Sion (1976). *Joint Ventures Between U.S. Multinational Firms and Host Governments in Selected Developing Countries: A Case Study of Costa Rica, Trinidad and Venezuela*. Unpublished doctoral dissertation, Indiana University.
- Renforth, William E. (1974). *A Comparative Study of Joint International Business Ventures with Family Firm or Non-Family Firm partners: The Caribbean Community Experience*. Unpublished doctoral dissertation, Indiana University.
- Ring, Peter Smith and Andrew H. Van de Ven (1994). 'Developmental processes of cooperative interorganizational relationships', *Academy of Management Review*, 19, pp. 90-118.
- Schaan, Jean-Louis (1983). *Parent Control and Joint Venture Success: The Case of Mexico*. Unpublished doctoral dissertation, University of Western Ontario.
- Schaan, Jean-Louis and Paul W. Beamish (1988). 'Joint venture general managers in LDCs.' In Contractor, F. J. and P. Lorange (eds.), *Cooperative Strategies in International Business*. Lexington Books, Lexington, MA.
- Schwab, Donald P. (1980). 'Construct validity in organizational behavior', *Research in Organizational Behavior*, 2, pp. 3-43.
- Steers, R. M. (1975). 'Problems in the measurement of organizational effectiveness', *Administrative Science Quarterly*, 20, pp. 546-558.
- Venkatraman, N. and Vasudevan Ramanujam (1986). 'Measurement of business performance in strategy research: A comparison of approaches', *Academy of Management Review*, 11, pp. 801-814.