IESE Business School
Universidad de Navarra

Centro Anselmo Rubiralta
de Globalización y Estrategia

Working Paper

WP No 529

November, 2003

TESTING ASSUMPTIONS ABOUT EVALUATING STRATEGIC ALLIANCE PERFORMANCE

Paul Olk *
Africa Ariño **

* Professor, Daniels College of Business, University of Denver
** Professor, General Management, IESE

IESE Business School - Universidad de Navarra
Avda. Pearson, 21 - 08034 Barcelona. Tel.: (+34) 93 253 42 00 Fax: (+34) 93 253 43 43
Camino del Cerro del Águila, 3 (Ctra. de Castilla, km. 5,180) - 28023 Madrid. Tel.: (+34) 91 357 08 09 Fax: (+34) 91 357 29 13

The Anselmo Rubiralta Center for Globalization and Strategy aims to be an international benchmark for companies, universities, business schools and governments. It promotes the exchange of ideas in the area of globalization and international strategy.

The Center's main objectives are to:

- Develop a conceptual structure that will help companies to understand and manage the impact of globalization and develop strategies of internationalization;
- Generate and disseminate new knowledge in this field;
- Develop new, high-quality teaching materials.

www.iese.edu/globalcenter

TESTING ASSUMPTIONS ABOUT EVALUATING STRATEGIC ALLIANCE PERFORMANCE

Abstract

Researchers have used a variety of measures to evaluate strategic alliance performance. In this paper we use data collected on performance of R&D consortia in the U.S. and of Spain-based equity and non-equity dyadic alliances to investigate empirically five basic assumptions made by strategic alliance researchers. We find that while several assumptions are supported, others are not. Results are consistent across samples. We conclude with recommendations for how to evaluate performance in future research into strategic alliances.

Key words: Strategic Alliance, Performance Evaluation, Equity Alliances, Consortia.

TESTING ASSUMPTIONS ABOUT EVALUATING STRATEGIC ALLIANCE PERFORMANCE

Introduction

The proliferation of strategic alliances over the last few decades has been accompanied by a similar growth in strategic alliance research. As part of an effort to understand what leads to alliance effectiveness, researchers have examined a host of factors to describe, and often to develop normative statements about, when and how to build, develop and perhaps terminate a strategic alliance. Typically, research examines different components of one aspect of an alliance – e.g., partnering structure (Parkhe, 1993b), learning (Inkpen, 2002), or trust (Zaheer *et al.*, 1998) – and relates them to performance. Central to an understanding of the determinants of alliance effectiveness is the type of performance measures used. What emerges from reviewing a broad number of strategic alliance studies is that there is no single, agreed upon measure of performance. This has led to research identifying different types of performance measures (see for example Gray, 2000; Olk, 2002) or examining comparability between specific measures (e.g., Ariño, 2003; Geringer, 1998; Glaister and Buckley, 1998; Geringer and Hebert, 1991), but generally concluding that the measures used are often not equivalent.

That researchers have not reached a consensus about strategic alliance performance measures is not unusual. Numerous areas of strategic management research have benefited from a periodic examination of measurement. Hoskisson *et al.* (1993), for example, called attention to the similarity of an objective, entropy measure of diversification strategy to a subjective measure. Likewise, Bergh and Fairbank (2002) examined alternative approaches to measuring strategic change. They concluded that the most common approach may produce flawed results, and that researchers need to recognize and address the complexities of evaluating change in order to improve the empirical support for theory. Similarly, Nath and Gruca (1997) examined three alternatives to measuring strategic groups, finding consistency between archival measures with direct and with perceptual measures, but not between direct and perceptual measures. For each of these strategic management topics, these studies reflect an important stage in knowledge development. They consolidated findings and gave researchers an evaluation of which measures to use and when and where measures are appropriate. In general these studies are part of normal science, as the researchers examine prior research for construct coherence and operationalization in their role as "validity police" (Hirsch and Levin, 1998). For strategic alliance research, the challenge of developing comparable measures of performance is compounded by researchers often not justifying their measure selection. This practice results in implicit assumptions being made about measurement development.

We gratefully acknowledge the financial support of the Anselmo Rubiralta Center for Research on Globalization and Strategy at IESE.

Our paper calls attention to these assumptions and evaluates whether they have support. These assumptions –found in research using a variety of theoretical lenses, including transaction cost, resource dependence, resource based view, learning, and an options approach, and in research examining a broad array of alliance types– focus on the comparability of various measures and respondents, the preference for select measures, and recommendations for how to combine measures. As such, they affect the ability to achieve consistency in findings and to develop a deeper understanding of what makes for an effective alliance.

The contributions of our paper to an understanding of how to build effective alliances serve as the basis for its structure. We begin by describing five assumptions researchers have made in measuring alliance effectiveness. From these, we develop hypotheses about relationships between different measures that permit us to test whether there is support for the assumptions. We then test these hypotheses using data from two different samples. The first consists of data collected from U.S. member companies of U.S. –based research and development consortia and the second is comprised of equity and non-equity dyadic alliances based in Spain. Testing our hypotheses using these different databases adds to the robustness of our findings. Finally, we discuss the implications of our findings for researchers trying to reconcile different findings related to strategic alliance effectiveness as well as for those seeking to operationalize performance in future research.

Assumptions about performance evaluation

The assumptions were identified from reviewing a wide variety of strategic alliance articles and represent a range of different performance evaluation issues. These assumptions cover issues of how best to evaluate alliance performance, as well as the relatedness of different performance indicators. Specifically, the assumptions focus on the relationship between continued membership and company satisfaction; on which performance criteria are most useful; and on the evaluation of performance, in terms of relatedness among different measures, how to combine different performance measures, and the reliability of a partner's evaluation for representing other partners. As noted, rarely do researchers make explicit such an assumption –and we will note where researchers have tested or noted disagreement with an assumption– nor do they test the impact of the assumption on research findings. In this section we identify and discuss these assumptions and, for each, we develop hypotheses that will allow us to test for support.

Relationship Between Stability or Termination and Company Satisfaction

Assumption 1: Measures of stability or termination are good indicators of a company's overall satisfaction with, or satisfaction in achieving its strategic interests in, the alliance.

Many researchers examining stability or termination of an alliance assume that longevity indicates good performance and that when performance is poor, the alliance will be terminated or restructured (e.g., Lampel and Shamsie, 2000; Marios and Abdessemed, 1996; Meschi, 1997; Spekman *et al.*, 1996). Several reasons have been offered for this relationship. First, the lack of a change reflects some degree of success in creating the arrangement, either in terms of creating some value or in terms of conflict reduction. The alliance is likely to be more productive when there is an absence of conflict (e.g., Steensma and Lyles, 2000). As

Cropper (1996) noted, stability indicates an absence of opportunistic behavior, or at least the partners have ex post safeguarded their investments in the alliance, and perhaps that the partners have made a commitment to future persistence, continuity and continuing viability of the relationship.

Support for this assumption comes from a generally positive correlation found between longevity and managerial satisfaction. Geringer and Hebert (1991), using data from two samples, found a strong agreement between stability and company satisfaction. Drawing from this, many researchers (e.g., Lee and Beamish, 1995) have argued that measures of stability may reflect alliance satisfaction.

This relationship has not always received strong support. Glaister and Buckley (1998), attempting to replicate Geringer and Herbert (1991), found only weak support for a positive relationship between measures of alliance survival, stability and duration, and satisfaction measures (i.e., company's satisfaction with alliance, company's perception of partner's satisfaction, company's perception of the alliance manager's satisfaction, and company's evaluation of the costs and benefits of the alliance). In their sample, the satisfaction measures were positively related to termination, but showed virtually no significant relationship with stability and duration. Likewise, Hatfield *et al.* (1998) concluded from their analysis that while alliance survival, duration and company satisfaction are related indicators of performance, they reflect different phenomena and are not equivalent.

In addition to empirical reservations, several researchers have raised conceptual concerns about the measure for representing a company's strategic interests. One is that termination combines natural and untimely deaths (Gulati, 1998). With a simple measure, alliances designed for limited duration cannot be distinguished from instances where the termination is not planned. A second is that the measure assumes stability is in the company's interest and tells us little about the company's reasons for terminating (Reuer and Koza, 2000). For example, it combines situations where the company may have considered the alliance to be performing well, but it was terminated for other reasons (e.g., the partner decided to withdraw; the company's strategic interests changed, making the alliance no longer relevant even though it was performing well; or the company decided to internalize the activities of the alliance and no longer wanted a separate entity), with situations where the alliance was considered not to be performing well. Alternatively, Inkpen and Ross (2001) reported evidence that some alliances persist even when the expected results are not being achieved. Finally, Yan and Zeng (1999) argued, in addition to the above arguments, that research on instability suffered from a static focus on stability as a destination rather than as a process, and that it did not provide specific recommendations for management. They concluded that instability is both a neutral concept (reflecting neither good nor poor performance) and a dynamic one (it may be the result of many developments, and can represent one of several different changes in the alliance). These criticisms suggest the need to consider the strategic interests of the company when examining whether stability is related to overall satisfaction with an alliance.

One of the sources of these problems is that most studies have examined two-party strategic alliances. In studying these alliances, particularly from an outsider's perspective, researchers cannot always distinguish between bilateral decisions leading to venture instability and one member's decision to depart. In addition to our data on two-party alliances based in Spain, by including in our analysis R&D consortia, which in this study all have more than two parties, we can evaluate when a member company departs without this leading to alliance dissolution. Therefore, in this sample we differentiate sample venture instability from membership change and provide a more refined test of the relationship between satisfaction and the decision to depart than previous studies.

Drawing on these arguments, we propose two hypotheses to evaluate the assumption of a positive relationship between stability or termination and company satisfaction. The first examines the general condition of a positive relationship between the propensity to remain in an alliance and company general satisfaction. The second refines this relationship by considering the strategic interests of the company and evaluates whether there is a positive relationship between propensity to remain in an alliance and satisfaction in achieving strategic objectives.

Hypothesis 1: Propensity of a company to stay in an alliance is positively related to the company's overall satisfaction with the alliance.

Hypothesis 2: Propensity of a company to stay in an alliance is positively related to the company's satisfaction with its strategic objectives for the alliance.

Performance Criteria Usefulness

Assumption 2: Goal-related criteria of performance are more useful than non-goal criteria, in general, and criteria that reflect a company's specific goals are more useful than criteria that reflect other possible goals.

A second assumption made by many strategic alliance researchers occurs when asking managers to evaluate the performance of the alliance against the company's goals for entering the alliance. Alongside termination, this is likely the most common measure of strategic alliance performance and is what Gray (2000) refers to as a "problem-focused" approach. The advantage of this approach is that it accesses the strategic reasons for the company's participation in the alliance. Directly asking managers is considered a valid indicator of the company's strategic intent and is said to be a better measure of success, as defined by the decision makers, than longevity. The goals researchers have used include self-defined goal accomplishment (Saxton, 1997), objectives met (Brockhoff and Teichert, 1995; Lorange and Roos, 1992), and initial expectations (Geringer and Hebert, 1991), as well as more specific goals such as learning (Lane and Lubatkin, 1998; Kale *et al.*, 2000).

What is not tested, in many of these studies, is whether managers find goal-related criteria to be more important than other types of criteria. Although some researchers have included criteria such as emergent or unintended benefits, benefits to other partners or network externalities, or level of cooperation between the partners, there has been no testing of whether a goal-related criterion is more useful than non-goal criteria. We test this assumption by developing two hypotheses. In the general form of the hypothesis (H3), we propose that all goal-related criteria are more relevant than non-goal-related criteria. Specifically, we suggest that managers will find performance criteria that reflect motives for partnership more useful than criteria that do not reflect possible goals. In a more specific form of the hypothesis (H4), we propose that a company's primary goal for partnership will be more important in determining useful performance criteria than will other motives. Researchers have identified a variety of reasons why a company might join an alliance (e.g., Contractor and Lorange, 1988). Criteria that reflect a company's primary reason for joining will likely be more useful to the company than criteria that reflect other motives. Stated more formally, we anticipate that criteria that reflect a company's primary goals will be rated more useful to managers than criteria that reflect other goals.

Hypothesis 3: Performance criteria that reflect goals for alliance partnership will be evaluated as significantly more useful than criteria that do not reflect possible goals.

Hypothesis 4: Performance criteria that reflect a company's primary alliance goal will be evaluated as significantly more useful than criteria that reflect other goals.

Performance Evaluation

Assumption 3: Goal-related evaluations of performance are related to other evaluations of performance.

Turning away from concerns about performance criteria and towards actual performance, a third assumption can be found in the few studies that use more than one type of measure to indicate performance. Researchers have had to make decisions about how to analyze the different measures. Some have conducted separate analyses on each performance indicator (e.g., Luo and Chen, 1997; Lyles and Salk, 1997). Others have combined the different measures into a single evaluation, typically combining meeting a company's objectives with indicators of company satisfaction and some non-goal-related measure. Examples of such efforts include measures of company satisfaction and joint venture economic performance (Cullen et al., 1995; Osland and Cavusgil, 1996), and company satisfaction, meeting company objectives and company profit (Mjoen and Tallman, 1997).

In combining the measures, researchers assumed, but did not show evidence, that these evaluations are related to one another and can be combined into a composite measure. If these variables are pooled, the variables need to reflect an underlying construct of performance, and therefore should be positively related[1]. This practice raises some concerns. Researchers have argued that at times a company will have multiple and sometimes conflicting motives for alliance partnership (e.g., Beamish and Delios, 1997). To test the assumption that these different measures can be combined, we examine two approaches found in the literature: combining indications of attaining strategic interests with non-goal performance, and combining indications of attaining strategic interests with company satisfaction. We argue that there should be a positive relationship between a company's goal-related performance and its non-goal performance and that there should be a positive relationship between goal-related performance and company satisfaction.

Hypothesis 5: Goal-related and non-goal performance evaluations of an alliance are positively related to one another.

Hypothesis 6: Goal-related performance evaluations and overall satisfaction with an alliance are positively related to one another.

Assumption 4: When using multiple measures of performance, it does not matter whether the various measures are added together or weighted for each item's importance.

Extending the issue raised in Assumption 3 is a concern about how to combine the variables. When using several measures of performance, researchers have had to make decisions about whether to add them equally or to weight them based upon the importance of

the criterion. For some, the measures have been added together (e.g., Steensma and Corley, 2000). For others, the indicators have been weighted based upon the criterion's relative importance (e.g., Parhke, 1993a). In these studies, by not directly comparing the different approaches, there is an implicit supposition that it does not make any difference how the variables are combined. To test this, we propose a hypothesis stating that the indication of performance derived by adding together different indicators will not be significantly different from the indication of performance developed from first weighting the indicators. Formally stated, there will be a positive relationship between a composite performance evaluation derived by simply adding together the factors and a composite performance evaluation derived by weighting the factors.

Hypothesis 7: Alliance performance derived by adding together non-weighted measures will be positively related to performance derived by adding together measures first weighted by importance.

Assumption 5: One strategic alliance member's evaluation of performance is significantly related to its partner's performance evaluation.

A fifth assumption one finds in the literature centers on the representativeness of a focal company's performance assessment for the partner's assessment. Although some researchers have noted that partners often have different and even competing goals for an alliance (e.g., Hamel, 1991) and others argue that success should be evaluated in terms of both partners' achieving their goals (e.g., Yan and Gray, 1994), the costs and difficulties of collecting data from more than one partner has led some researchers to argue that it is often unfeasible to collect data from both partners when evaluating performance (e.g., Parkhe, 1993a). Further, Geringer and Hebert (1991) reported a strong correlation between one partner's assessment of an alliance and the other's. Subsequently, a number of researchers have assumed that partners of an alliance will evaluate performance similarly. We test this assumption in a hypothesis that states that there will be no significant difference among the performance evaluations made by partners in an alliance.

Hypothesis 8: There will be no significant differences in performance evaluations by partners of the same alliance.

METHODS

We tested our hypotheses with two distinct data sets, collected without knowledge of one another. In Study 1, we used data collected from a sample of U.S.-based R&D consortia representing a range of industries (e.g., chemicals, energy, automotive, composite materials) but emphasizing computers and telecommunications. These consortia were all equity agreements around pre-competitive research activities, and involved more than two member companies, almost all of them from the U.S. In Study 2, we worked with data gathered from a sample of companies operating in Spain involved in strategic alliances, also from a broad range of industries (e.g., energy, chemicals, machinery, electronic equipment, transportation), but emphasizing financial services. The alliances were both domestic and international, equity and non-equity agreements around competitive activities, and involved only two partners. The general differences between the two data sets are such that consistent results between them will enhance the robustness of our arguments.

STUDY 1

Sample and procedures

Surveys were distributed to 110 multi-party research and development consortia registered with the U.S. Department of Justice under the National Cooperative Research Act of 1984, listed in the Federal Register during the period January 1985 through January 1992. We obtained addresses for almost all of the U.S. member organizations, representing 915 memberships in the consortia. Telephone calls made to each member organization identified the person considered the primary liaison between the consortium and the member organization. After the initial mailing and follow-up phone calls, we received 253 responses, and the actual number of questionnaires usable for the purposes of this research question was 168 (18.3%), representing 77 of the 110 consortia (70%). Of these responses, 134 came from 43 consortia in which at least one other member responded.

We tested for a non-response bias in two ways. First, we compared the respondents to non-respondents in terms of reliably known data, including company sales, number of employees, industry SIC code, consortium size, and number of consortia in which a company was a member. We found no variation in terms of the industries and consortia represented, consortium size, and number of consortia in which a company was a member. A t-test did indicate that responding companies were slightly, yet significantly, larger than non-respondents. Second, we compared early and late respondents across all the consortium variables used (Armstrong and Overton, 1977), and we found that they diverged significantly only on the age of the consortium. Early responses tended to come from members of slightly older consortia. These analyses suggest that, while this sample appears overall to be representative, the findings may be somewhat biased toward larger members and toward members of older consortia.

We also tested for the possibility of bias derived from reliance upon a single key informant. Consistent with research on evaluating the reliability of a single informant (Kumar *et al*, 1993), we examined respondents' knowledge on multiple global indicators. First, we examined whether there was any influence due to inadequate knowledge about the company's membership in the consortium. To test this, we reran the analyses using only those informants who had been involved in the consortium for more than a year. The findings are consistent with those we report. Second, we examined whether respondents' position in the consortium or the company might influence the knowledge about membership. We separated respondents based on their self-reported position in their own company and on their affiliation with the consortium. The positions were collapsed into upper management, middle management and lower management. Most respondents (78%) reported coming from middle management in their company, but held an upper management position in the consortium (81%). Comparisons across these three levels on the variables used in this study found no significant differences. Since neither test provides evidence of unreliability, we concluded that our reliance upon key informants has not introduced any substantial bias to the study.

Measures

The measures used to test the hypotheses required deriving three types of variables: membership and satisfaction variables, usefulness of performance criteria variables, and performance evaluation. Appendix 1 provides a detailed list of the questions used to create the variables.

Membership and Satisfaction Variables

Propensity to Stay. This dummy variable indicates the informants' assessment of their company's intention to continue to belong to this consortium. Tests confirmed that this measure reliably indicated subsequent consortia membership changes.

Overall Satisfaction. This variable is a 3-item, 7-point scale measuring the informants' assessment of the extent to which their company was satisfied with the consortium (Cronbach's alpha=.80).

Usefulness of Performance Criteria Variables

Respondents were asked to evaluate 31 different criteria in terms of the usefulness of each to their company, whether or not it was actually used, for evaluating the consortium (7-point scale). The criteria were derived from the literatures on R&D management and on strategic alliances (see Table 1). These 31 items were grouped into goal-related and non-goal criteria based upon informant-supplied information on the company's motives. Respondents were asked to evaluate the importance of nine potential reasons for membership in the consortium (7-point scale). A factor analysis of the nine items revealed three factors: learning motive, profit motive and efficiency motive. Two items did not load on to these factors and were dropped from the analysis. The highest rated motive reported by the respondents was considered the primary goal of the member. The various performance criteria included 10 that reflected these three motives: five for learning, two for profit and three for efficiency. These 10 items were averaged to derive the measure *Usefulness of Goal-Related Criteria.* Averaging the subset items resulted in measures of *Usefulness of Learning Criteria, Usefulness of Efficiency Criteria, and Usefulness of Profit Criteria.* The 21 other criteria not representing a company's strategic interest in the consortium were averaged to derive the *Usefulness of Non-goal Criteria measure.*

Performance Evaluation Variables

In addition to asking informants which criteria were important, we also asked respondents their evaluation on each criterion of the performance over the previous 12 months. For subjective criteria this was a 7-point scale. For the objective evaluations (e.g., patents, return on investment) we asked for the actual value. In order to combine the subjective and objective evaluations into a single composite measure, we normalized the objective measures on a 7-point scale (i.e., lowest response became a 1 and the highest a 7). From this we derived the following variables:

Non-Weighted Goal-Related Performance. Averaging together all the subjective and objective evaluations of goal-related performance derived this summary indicator of performance. This was also done for each of the goal-related performance measures: *Non-Weighted Learning Performance, Non-Weighted Profit Performance and Non-Weighted Efficiency Performance.*

*Weighted Goal-Related Performance.*This second summary indicator of performance was derived by first multiplying the evaluation of the performance on a specific criterion by the reported usefulness of the criterion. The weighted scores for the various indicators of performance were then averaged together and then rescaled (1=lowest and 7=highest) to develop a weighted performance indicator. This was also done for each of the goal-related

Table 1. Performance Criteria

Criteria	*Goal Represented by Criterion**
Percentage Changes in Consortium's Annual Budget	
Annual Percentage of Consortium Research Project Completed	
Annual Number of Papers Consortium Researchers Present at Conferences or Publish	
Annual Number of New Products Developed by Consortium	
Annual Number of Patents From Consortium Research	
Annual Number of Licensing Arrangements From Consortium Patents	
Annual Profitability of Consortium – Return on Investment	
Planned Versus Actual Costs of Consortium Research	
Annual Number of Member Organizations Retained	
Annual Number of New Member Organizations Attracted	
Annual Number of Contacts Between Consortium and Your Organization	Learning
Annual Amount of Time Consortium Research Saves Your Organization	Efficiency
Annual Return on Your Organization's Investment	Profit
Annual Percentage of Consortium Research Results Adopted by Your Organization	Learning
Number of Spin-off Companies Created From Consortium Research	
Probability of Commercial Success of Consortium Research Findings	
Technical Quality of Consortium Research	
External Reputation of Consortium	
Responsiveness of Consortium to Your Organization's Needs	
Improvement in Consortium Researchers' Skills and Knowledge	Learning
Consortium's Contribution to Improvement in Skills and Knowledge of Your Organization's Researchers	Learning
Consortium's Contribution to Improvement in Your Organization's Competitive Position	Profit
Level of Cooperation Among Member Organizations	
Value of Information Obtained on the Internal Research of Other Member Organizations	Learning
Perceived Benefits to U.S. Competitiveness	
Perceived Benefits to Member Organizations Versus Non-members	
Equitable Distribution of Benefits Among Member Organizations	
Planned Versus Actual Progress of Consortium Research	
Planned Versus Actual Outcomes From Consortium Research	
Value of Consortium Research in Comparison to Your Organization's Internal Research Efficiency	
Value of Consortium Research In Comparison to Your Organization's Other Cooperative Research Activities	Efficiency

* All items not representing a member-level goal were categorized as other or non-goal. + Respondents saw this list twice. The first time they rated each item in terms of how useful it was as an indicator of performance (1=Not at all important, 7=Extremely important). The second time, they evaluated the consortium's performance over the previous 12 months on the item. The former responses were used to develop measures of usefulness of performance criteria, the latter to develop measures of performance.

performance measures, resulting in *Weighted Learning Performance, Weighted Profit Performance and Weighted Efficiency Performance.*

Non-Goal Performance. The performance evaluations on criteria not reflecting the goals were averaged together to create a measure of non-goal performance.

STUDY 2

Sample and procedures

Surveys were mailed to a sample of 189 Spanish firms that appeared in Funk and Scott's Countries Index - Europe (1986-1992) as having announced their engagement in collaborative activities, beginning with Spain's adhesion to the European Community (1986) and concluding with the establishment of the Single European Market (1992). This is a period that can be expected, a priori, to include high venturing activity, especially if we take into account that the international exposure of Spanish firms was low at that time. Only alliances involving two partners were considered. 91 (48%) of the questionnaires were returned, and they were all usable for this study.

In order to examine potential non-response bias, we assessed possible differences in alliances' industries (Chi-square test) and in firm size, measured by the number of employees (one-way ANOVA tests), between early and late respondents (Armstrong and Overton, 1977). These tests provided no indication of response bias. Examination of the variables included in the study revealed that there were significantly more survivors among late than early respondents (90% vs. 65%).

As an indication of the competence of key informants, over 63 percent of the respondents had participated directly in the negotiation of the alliance in question. On average they had been involved in the alliance for 4.9 years, with 91 percent having been involved at least since the alliance formation, and 4.5 percent during its first year of operation. Validity checks with secondary data suggested no evidence of unreliability[2]. Preliminary versions of the questionnaire were reviewed by business scholars to ensure face validity. The survey was then translated into Spanish and reviewed by two Spanish-speaking researchers. The translated survey was pre-tested with six Spanish executives experienced in managing alliances, and several changes were made after the pre-testing stage. The final Spanish version was reverse translated into English by a person unfamiliar with the study, and there was a high degree of correspondence between the Spanish and English versions.

For analytical purposes, we split the sample into two sub-samples. The first included 39 equity alliances, and the second 52 contractual alliances. These two types of alliances are likely to result in diverse dynamics that have different performance implications. Thus, the sample split allows us to assess the robustness of our results across these alliance types.

---

[2] We used available secondary data for which we had the exact survey items. Specifically, we examined whether or not the respondent firm was state-owned and whether the partner firm was a Spanish company, a subsidiary, or a foreign company. Matches were present for 98 and 96 percent of the cases, respectively.

Measures

Appendix 2 contains the questionnaire items from which the variables were derived. Information on performance criteria variables was not collected in this study, as it had not been designed for this purpose initially.

Membership and satisfaction variables

Alliance Survival. This dummy variable indicates whether the alliance was still operating at the time of data collection.

Overall Satisfaction. This variable is a single-item, 5-point scale measuring the informants' assessment of the extent to which their firm was satisfied with the overall performance of the alliance.

Partner's Overall Satisfaction. In a similar fashion, we asked informants their assessment of the extent to which their partner was satisfied with the overall performance of the alliance (5-point scale).

Performance Evaluation Variables

Weighted Strategic Goal Fulfillment. This variable is a composite measure of the importance multiplied by the degree of *fulfillment* of the specific strategic goals an alliance may be meeting (Parkhe, 1993b). *Importance* is a 5-point scale measuring the informants' assessment of the extent to which each of the possible goals embraced by the firm for an alliance was important to their firm. *Fulfillment* is a 5-point scale measuring the informants' assessment of the extent to which each of the identified strategic goals for the alliance was fulfilled. A list of eight strategic goals was adapted from the literature (Contractor and Lorange, 1988; Parkhe, 1993a, 1993b). We also provided the category "not applicable" as an option.

Non-Weighted Strategic Goal Fulfillment. We summed the degree of fulfillment of specific strategic goals an alliance may be meeting as just defined.

Partner's Weighted Strategic Goal Fulfillment. Similarly to strategic goal fulfillment, we constructed this composite measure by multiplying the firm's assessments of importance by the degree of *fulfillment* of the partner's specific strategic goals an alliance may be meeting.

Net Spillover Effects. We asked informants to assess on a 5-point scale the difference between positive and negative effects of the alliance on other activities of their firm (Parkhe, 1993a).

RESULTS

Table 2 summarizes the expected relationships between the variables in each of the two studies. Means, standard deviations and correlations among the variables appear in Table 3 (Study 1) and Table 4 (Study 2).

Table 2. Expected Relationships between Variables in Study 1 and Study 2

Hyp	Expected Relationship	Variables in Study 1		Variables in Study 2	
H1	+ correlation	Propensity to stay	Overall satisfaction	Alliance survival	Overall satisfaction
	+ correlation			Alliance survival	Partner's overall satisfaction
H2	+ correlation	Propensity to stay	Non-Weighted Learning performance, if learning motive primary	n.a.	n.a.
	+ correlation	Propensity to stay	Non-Weighted Efficiency performance, if efficiency motive primary		
	+ correlation	Propensity to stay	Non-Weighted Profit performance, if profit motive primary		
H3	sign.difference	Usefulness of Goal Related Criteria	Usefulness of Non-Goal Criteria	n.a.	n.a.
H4	sign.difference	Usefulness of Learning criteria if learning motive primary	Usefulness of Efficiency criteria if learning motive primary	n.a.	n.a.
	sign.difference	Usefulness of Learning criteria if learning motive primary	Usefulness of Profit criteria if learning motive primary		
	sign.difference	Usefulness of efficiency criteria if efficiency motive primary	Usefulness of Learning criteria if efficiency motive primary		
	sign.difference	Usefulness of efficiency criteria if efficiency motive primary	Usefulness of Profit criteria if efficiency motive primary		
	sign.difference	Usefulness of profit criteria if profit motive primary	Usefulness of Learning criteria if profit motive primary		
	sign.difference	Usefulness of profit criteria if profit motive primary	Usefulness of Efficiency criteria if profit motive primary		
H5	+ correlation	Non-Weighted Learning performance	Non-goal performance	Weighted strategic goal fulfillment	Net spillover effects
	+ correlation	Non-Weighted Efficiency performance	Non-goal performance		
	+ correlation	Non-Weighted Profit performance	Non-goal performance		
	+ correlation	Non-Weighted Learning performance, if learning motive primary	Non-goal performance		
	+ correlation	Non-Weighted Efficiency performance, if efficiency motive primary	Non-goal performance		
	+ correlation	Non-Weighted Profit performance, if profit motive primary	Non-goal performance		

Table 2. Expected Relationships between Variables in Study 1 and Study 2 (continued)

H6	+ correlation	Non-Weighted Learning performance, if learning motive primary	Overall satisfaction	Weighted strategic goal fulfillment	Overall satisfaction
	+ correlation	Non-Weighted Efficiency performance, if efficiency motive primary	Overall satisfaction		
	+ correlation	Non-Weighted Profit performance, if profit motive primary	Overall satisfaction		
H7	+ correlation	Weighted performance	Non-weighted performance	Weighted strategic goal fulfillment	Non-weighted strategic goal fulfillment
	+ correlation	Weighted Learning performance, if learning motive primary	Non-Weighted Learning performance, if learning motive primary		
	+ correlation	Weighted Efficiency performance, if efficiency motive primary	Non-Weighted Efficiency performance, if efficiency motive primary		
	+ correlation	Weighted Profit performance, if profit motive primary	Non-Weighted Profit performance, if profit motive primary		
H8	n.s.difference	Multiple partners' Propensity to stay		Overall satisfaction	Partner's overall satisfaction
	n.s.difference	Multiple partners' Overall satisfaction		Weighted strategic goal fulfillment	Partner's overall satisfaction
	n.s.difference	Multiple partners' Learning performance			
	n.s.difference	Multiple partners' Efficiency performance			
	n.s.difference	Multiple partners' Profit performance			
	n.s.difference	Multiple partners' Unweighted Goal related performance			
	n.s.difference	Multiple partners' Weighted Goal related performance			
	n.s.difference	Multiple partners' non-goal performance			

Table 3. Means, Standard Deviations and Correlations* for Variables in Study 1

	Mean	Std.Dev.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1. Propensity to Stay	1.82	0.39	100															
2. Overall Satisfaction	4.60	1.25	27	100														
3. Usefulness of Goal-Related Criteria	5.06	0.90	-29	-5	100													
4. Usefulness of Non-Goal Criteria	4.67	0.59	-17	8	47	100												
5. Usefulness of Learning Criteria	4.92	0.95	-28	-5	86	48	100											
6. Usefulness of Profit Criteria	5.44	1.22	-31	-6	77	41	52	100										
7. Usefulness of Efficiency Criteria	5.05	1.24	-14	-2	81	26	46	54	100									
8. Non-Weighted Learning Performance	3.86	1.06	6	35	17	26	23	3	10	100								
9. Non-Weighted Profit Performance	3.73	1.57	18	46	2	9	-4	2	7	54	100							
10. Non-Weighted Efficiency Performance	4.11	1.42	22	58	10	27	2	6	17	45	58	100						
11. Weighted Learning Performance	3.44	0.93	-8	28	46	36	60	22	21	**85**	45	38	100					
12. Weighted Profit Performance	3.71	1.42	8	40	21	16	10	30	17	52	**92**	52	49	100				
13. Weighted Efficiency Performance	3.48	1.23	9	39	46	28	20	27	66	39	49	**78**	40	50	100			
14. Non-Goal Performance	4.40	0.88	5	66	5	19	-1	6	7	**46**	**51**	**55**	35	48	43	100		
15. Non-Weighted Performance	4.22	0.86	12	65	6	25	3	4	8	67	66	71	53	61	55	93	100	
16. Weighted Performance	3.21	0.85	1	62	31	41	26	23	25	67	62	69	64	64	65	88	**93**	100

* All correlations 15 and above are significant at least at the .05 level. Numbers in bold indicate results relevant for hypotheses testing.

n=168

Table 4. Means, Standard Deviations and Correlations* for Variables in Study 2

Equity SAs Sub-Sample

			Correlations						
	Mean	Std Dev.	1	2	3	4	5	6	7
1. Alliance Survival	0.77	0.43	100						
2. Overall Satisfaction	3.49	1.12	48	100					
3. Partner's Overall Satisfaction	3.64	0.96	42	-75	100				
4. Weighted Strategic Goal Fulfillment	54.54	25.99	-25	67	53	100			
5. Non-Weighted Strategic Goal Fulfillment	16.85	8.06	-17	68	57	92	100		
6. Weighted Partner's Strategic Goal Fulfillment	64.05	30.21	-4	44	48	62	69	100	
7. Net Spillover Effects	3.56	0.88	43	70	65	42	50	26	100

* All correlations 42 and above are significant at the .05 level. Numbers in bold indicate results relevant for hypotheses testing.
n=39

Non-Equity SAs Sub-Sample

			Correlations						
	Mean	Std Dev.	1	2	3	4	5	6	7
1. Alliance Survival	0.75	0.44	100						
2. Overall Satisfaction	3.25	0.97	43	100					
3. Partner's Overall Satisfaction	3.63	0.84	31	69	100				
4. Weighted Strategic Goal Fulfillment	53.58	23.87	0	42	22	100			
5. Non-Weighted Strategic Goal Fulfillment	17.53	0.07	20	34	18	82	100		
6. Weighted Partner's Strategic Goal Fulfillment	55.63	27.86	15	28	24	61	64	100	
7. Net Spillover Effects	3.67	0.74	35	64	53	37	34	17	100

* All correlations 28 and above are significant at the .05 level. Numbers in bold indicate results relevant for hypotheses testing.
n=52

To test hypothesis 1, which argued that decisions regarding propensity to stay were related to overall satisfaction, we correlated these two measures in the sample from Study 1. The results, as can be seen in Table 3, reveal a significant and positive relationship between these two variables (r=.27, p<.05). In the sample from Study 2 (see Table 4), we correlated alliance survival with overall satisfaction, yielding positive and significant results both in the equity and non-equity sub-samples (r=.48, p<.05 and .43, p<.05 respectively). Alliance survival is also positively and significantly related to the partner's overall satisfaction (r= .42, p<.05 and .31, p<.05 for each sub-sample). Thus, the decision to stay in an alliance is positively and significantly correlated with overall satisfaction, supporting hypothesis 1.

The test for hypothesis 2 –as well as for hypotheses 3 and 4– could be done with the sample from Study 1 only. It involved assessing the correlation between propensity to stay and performance along each of the primary motives for partnership. These appear in Table 5. The three separate correlations provide mixed support for the hypothesis. While efficiency performance was significantly related to remaining in the consortium when a company's primary motive for being in the consortium was efficiency (r=.24, p<.05), the relationships of

the decision to stay with learning performance (r=-.01) and profit performance (r=.54) were not significant when the primary motives were learning and profit, respectively. It appears that the decision to stay in an alliance represents a general satisfaction but not necessarily satisfaction with the strategic interests in the alliance.

Table 5. Correlations for sub-samples based upon primary motives of Learning, Efficiency and Profit, Study 1

Learning Motive Primary (n=55)

	Mean	Std Dev.	1	2	3	4	5	6	7	8	9
Propensity to Stay	1.71	0.42	100								
Overall Satisfaction	4.36	1.14	34	100							
Non-Weighted Learning Performance	3.90	1.07	**-1**	**26**	100						
Non-Weighted Profit Performance	3.61	1.50	7	26	44	100					
Non-Weighted Efficiency Performance	3.70	1.33	6	50	44	59	100				
Non-goal Performance	4.25	0.78	-4	44	**50**	47	61	100			
Usefulness of Learning Criteria	**5.01**	0.86	-45	-18	21	1	-1	22	100		
Usefulness of Profit Criteria	**5.25**	1.26	-39	-21	-23	-5	-7	1	55	100	
Usefulness of Efficiency Criteria	**4.97**	1.27	-11	-4	-10	15	26	8	26	33	100

* All correlations 27 and above are significant at the .05 level. Numbers in bold indicate results relevant for hypotheses testing.

Efficiency Motive Primary (n=109)

	Mean	Std Dev.	1	2	3	4	5	6	7	8	9
Propensity to Stay	1.87	0.34	100								
Overall Satisfaction	4.74	1.24	25	100							
Non-Weighted Learning Performance	3.84	1.05	10	41	100						
Non-Weighted Profit Performance	3.81	1.61	22	54	57	100					
Non-Weighted Efficiency Performance	4.33	1.45	**24**	**62**	46	56	100				
Non-goal Performance	4.48	0.91	9	71	45	52	**52**	100			
Usefulness of Learning Criteria	**4.89**	1.00	-17	-1	26	-5	8	-8	100		
Usefulness of Profit Criteria	**5.54**	1.13	-28	1	24	9	16	10	53	100	
Usefulness of Efficiency Criteria	**5.08**	1.29	-16	-2	21	5	16	7	54	65	100

* All correlations 20 and above are significant at the .05 level. Numbers in bold indicate results relevant for hypotheses testing.

Profit Motive Primary (n=5)

	Mean	Std Dev.	1	2	3	4	5	6	7	8	9
Propensity to Stay	1.75	0.50	100								
Overall Satisfaction	3.83	2.06	-59	100							
Non-Weighted Learning Performance	3.75	1.20	42	47	100						
Non-Weighted Profit Performance	3.13	2.02	**54**	**29**	96	100					
Non-Weighted Efficiency Performance	4.00	2.31	58	28	88	79	100				
Non-goal Performance	4.27	1.47	-33	96	71	**54**	53	100			
Usefulness of Learning Criteria	**4.40**	0.71	-94	33	-60	-63	-81	6	100		
Usefulness of Profit Criteria	**5.38**	1.89	-57	-25	-83	-72	-99	-49	82	100	
Usefulness of Efficiency Criteria	**5.33**	0.94	-94	31	-66	-70	-82	2	99	81	100

* All correlations 95 and above are significant at the .05 level. Numbers in bold indicate results relevant for hypotheses testing.

Hypothesis 3 proposed that criteria that reflect goals for why a company would be in an alliance would be rated more useful than criteria that were not goal-related. A t-test comparing the mean of the importance of goal-related criteria (X_{mn}=5.06) with non-goal criteria (X_{mn}=4.67) reveals that this difference is significant (t=6.36, p<.01). The results support the assumption that goal-related criteria are more useful than non-goal criteria.

The fourth hypothesis refined the test of the importance of goal-related criteria by proposing that criteria reflecting a company's primary goal for partnership will be rated as more useful than criteria reflecting other goals. To test this hypothesis, we compared the means for the three "usefulness of goal criteria" variables when each of the three identified goals (learning, efficiency, and profit) was the primary motive for joining the consortium (see Table 5). A t-test analysis of the means for each type of motive revealed that when learning is the highest-rated motive (n=55), learning criteria (X_{mn} =5.01) are not significantly different in usefulness than efficiency criteria (X_{mn} =4.97, p>.05) nor profit criteria (X_{mn} =5.25, p>.05), and efficiency and profit criteria are not significantly different from one another. When efficiency is the highest-rated motive (n=109), t-test analysis revealed that efficiency criteria (X_{mn} =5.08) are evaluated significantly less useful than profit criteria (Xmn =5.54, p<.05), and similarly to learning criteria (X_{mn} =4.89, p>.05), while profit criteria are rated higher than learning criteria (p<.05). In the few instances when profit is the highest-rated motive (n=5), profit criteria (X_{mn} =5.38) are not rated significantly more useful than either learning (X_{mn} =4.40, p>.05) or efficiency criteria (X_{mn} =5.33, p>.05), while efficiency criteria are significantly less useful than learning criteria (p<.05), although these findings are likely affected by the small sample size. These three tests do not support the hypothesis. In each instance, the criteria related to the primary goal were not significantly more useful than other goal criteria, and for efficiency, even lower. These findings indicate that there is not a direct relationship between the goals when entering an alliance and the preferred criteria for evaluation[3].

Hypotheses 5-8 examined managers' evaluations of performance. For hypothesis 5, we examined whether goal-related performance and non-goal performance are significantly and positively related to one another. In Study 1, the results were conducted first for the entire sample and then for each of the primary motives for membership. Each of the three goal-related performance variables in Table 3 were significantly related to non-goal performance [non-goal and learning performance (r=.46), non-goal and efficiency performance (r=.51) and non-goal and profit performance (r=.55), p<.05]. Comparing the measure of primary motive performance with the single measure of non-goal performance indicators, in Table 5, we found a correlation of .50, or above, for each motive [non-goal and learning performance (r=.50); non-goal and efficiency performance (r=.52); non-goal and profit performance (r=.54)]. Because of its small size, for the profit sample the correlation is not significant, although it is significant for each of the other two motives (p<.05). In Study 2, we correlated our measures of non-weighted strategic goal fulfillment and net spillover effects, and found this correlation to be positive and significant (r=.50 p<.05 and .34, p<.05 for the equity and non-equity sub-samples respectively) (see Table 4). This finding provides support for hypothesis 5, that goal-related and non-goal related performance evaluations are positively related. In general, it appears that the assumption by researchers that these measures are consistent with one another and may be combined appears to have support.

---

[3] Because of sample size limitations, we were unable to examine all the combinations where respondents reported more than one motive above a theoretical mean suggesting importance (e.g., coding both efficiency and learning motives greater than 4 on a scale of 1=not important, 7=very important). Where we were able to conduct these analyses, we found support that goal-related criteria were significantly more important than non-goal-related criteria.

The next analysis on the performance evaluation examined whether goal-related performance is positively related to overall satisfaction, hypothesis 6. Examining the relationship between the overall satisfaction measure and performance for a company's primary motive revealed generally consistent support. As reported in Table 5, for the Study 1 sub-sample for which learning is the primary motive, learning performance is marginally related to satisfaction measures ($r=.26$, $p<.10$). For the sub-sample where efficiency is the primary motive, efficiency performance is related to satisfaction ($r=.62$, $p<.05$). For the sub-sample for which profit is the primary motive, however, profit performance is not significantly related to satisfaction ($r=.29$, $p>.05$), although we attribute this finding to the small sub-sample size. As for Study 2, we tested this hypothesis by correlating non-weighted strategic goal fulfillment and overall satisfaction. The results (Table 4) are positive and significant for both the equity ($r=.68$, $p<.05$) and the non-equity ($r=.34$, $p<.05$) sub-samples. Overall, it appears that satisfaction is associated with goal-related performance and there is support for hypothesis 6.

Hypothesis 7 proposed that it made no difference whether a composite performance evaluation measure was derived by adding together the unweighted individual items or whether they were first weighted to reflect their relative importance. Results reported in Tables 3 and 4 support the hypothesis. In Study 1, the overall simple average and the weighted average evaluations of performance were highly related ($r=.93$, $p<.05$)[4]. Also, for each of three motive-related performance measures, there was a very high correlation ($r=.85$ for learning, $r=.78$ for efficiency and $r=.92$ for profit) between the weighted and the unweighted performance measures. As for Study 2, the correlation between weighted and non-weighted measures of strategic goal fulfillment was also positive and significant in both sub-samples ($r=.92$, $p<.05$ and $.82$, $p<.05$ respectively). Thus, the overall results support hypothesis 7 that such two kinds of measures would be so related[5].

The final test evaluated hypothesis 8, and the assumption that one partner's response is similar to other partners. To test this assumption in Study 1, we conducted an analysis of variance of the various performance measures used for those alliances in which we had more than one respondent. The results are mixed, and appear in Table 6. While there is support for partners in an R&D consortium having some similar evaluations of performance (e.g., propensity to stay, overall satisfaction, weighted learning performance and weighted profit performance), for other measures of performance (e.g., non-weighted learning performance, non-weighted efficiency performance, non-weighted profit performance, weighted efficiency performance, non-goal performance, non-weighted and weighted performance), there are differences. This suggests that measuring one company's evaluation may not always reflect

---

[4] In a refined test, we eliminated low-weighted criteria from the analysis. Including unimportant measures would incorporate unnecessary information into the composite measure. Consequently, we eliminated those criteria for which the respondent did not rate above the theoretical mean (i.e., 4 on the 7-point scale). We multiplied the performance evaluation by the importance assigned to the criterion (5, 6 or 7) and calculated the mean of these scores and divided by 7, to rescale the measure from 1 to 7. This variable had a mean of 3.79 and a standard deviation of .73, and is highly correlated with both the non-weighted and the weighted composite measures.

[5] An alternative approach to evaluating the importance of weighting performance variables is to test hypotheses 5 and 6 using weighted measures of performance. In examining the correlations reported in Tables 3, 4 and 5, as well as in additional analyses not reported, one finds the same relationships in 12 of the 13 correlations using weighted measures as were reported with non-weighted measures. The one exception was found in the sub-sample learning motives, in Study 1. The correlation between learning performance and satisfaction, which is marginally significant using non-weighted measures, is non-significant using weighted measure of learning performance. This consistency supports the finding of comparability between weighted and non-weighted measures.

the partners' evaluation, and that researchers will need to confirm the similar ratings. In Study 2 we did not have data from the partner's side. As a proxy, we used the informant's assessment of the partner's performance. In this way, we ran an analysis of variance between measures of performance of the focal firm and measures of this firm's perceptions of its partner's performance. The results appear in Table 7 and show a significant difference between the means of overall satisfaction and the partner's overall satisfaction, as well as between the means of strategic goal fulfillment and the partner's strategic goal fulfillment. These results hold for both the equity and the non-equity sub-samples in Study 2, and –though not straight tests for hypothesis 8 that partners' performance evaluations are similar– are inconsistent with this hypothesis.

Table 6. Results from ANOVA of Within Alliance Consensus on Performance Evaluations, Study 1

Measures of Performance	F-Score
Propensity to Stay	1.60*
Overall Satisfaction	1.77*
Non-Weighted Learning Performance	1.39
Non-Weighted Profit Performance	1.43
Non-Weighted Efficiency Performance	1.11
Weighted Learning Performance	1.70*
Weighted Profit Performance	1.53*
Weighted Efficiency Performance	0.99
Non-goal Performance	1.39
Non-weighted Performance	1.44
Weighted Performance	1.48

* Significant at p<.05

Table 7. Results from ANOVA of Within Alliance Consensus on Performance Evaluations, Study 2

Measures of Performance	F-Score	
	Equity SAs Sub-Sample	Non- Equity SAs Sub-Sample
Overall Satisfaction	12.59 ***	9.81 ***
Strategic Goal Fulfillment	3.35 +	9.84 ***

+ Significant at p<.10
*** Significant at p<.001

One of the strengths of our analysis is that we have used data from two samples and used different measures of our constructs. While the dataset from Study 2 did not permit us to test hypotheses 2, 3, and 4, we were able to test the remaining hypotheses. The results from both sub-samples in Study 2 are consistent with those from Study 1, supporting hypotheses 1, 5, 6, and 7, and not supporting hypothesis 8. This consistency endorses the robustness of our conclusions.

DISCUSSION

As the practice of alliances continues, evaluation of effectiveness will remain a central issue for managers and researchers. Critical in developing knowledge of how to build an effective alliance is an understanding of how to evaluate effectiveness. As reviews have recognized (e.g., Gray, 2000; Olk, 2002), there is currently no single way to evaluate a strategic alliance. In this paper we built on the few studies that have investigated comparability of measures and provided an empirical testing of five common assumptions. We found only two of the assumptions received complete support.

The first assumption tested was whether measures of alliance partnership continuity are related to company satisfaction. A general version of the relationship (H1) was supported in both samples, suggesting that the relationship is not affected by differences in the number of partners, the activities of the alliance, or whether the alliance is equity or non-equity. A more specific version, (H2), which considered satisfaction in achieving strategic objectives and continuity, was not supported. Only companies engaged in the alliance primarily for efficiency reasons had the expected positive relationship between the decision to remain and partner satisfaction on efficiency criteria. It appears that researchers can use a company's decision to withdraw as an indication of its general satisfaction with an alliance but not necessarily satisfaction with its strategic interests. This finding suggests that research might want to distinguish between operational satisfaction and strategic objective satisfaction with an alliance. The former may refer to satisfaction for an alliance that has acceptable levels of conflict. When these levels are exceeded, or there is the presence of opportunistic behaviors, one may expect to find the associated relationship of satisfaction with alliance termination (e.g., Park and Ungson, 1997). Strategic objective satisfaction, however, refers to the impact on the strategic interests of the company. As we argued, companies may withdraw because they are satisfied, after reaching their strategic goals for the alliance, or they may stay in because they have not yet satisfied their strategic objectives. The determinants of each of these types of satisfaction may not be the same and research should continue to explore their differences.

The second assumption tested was whether managers valued goal-related measures over other types of measures. Many researchers have evaluated alliances from a company's perspective, based upon whether the company achieved its goals for partnership. Using a general form of the hypothesis (H3), we found that managers found goal-related criteria more useful than non-goal criteria. When we examined a more specific form of the hypothesis (H4), that managers would prefer measures that reflect their primary motive to measures that reflect other motives, we found no support. The assumed linkage between motives and criteria did not hold for any of the motives. While the findings for profit may be a consequence of the small sample size, the findings for learning and efficiency reveal that managers may use a broader array of criteria than is indicated by the motive. For example, managers may say that they are involved primarily for learning, but also be concerned with performance in terms of efficiency and profit. This indicates that research which measures performance only in terms of a company's stated or specific goals (e.g., Brockhoff and Teichert, 1995; Lin and German, 1998; Makino and Delios, 1996) may not capture the strategic benefits beyond the specific motives.

We followed this test with others that examined the assumption made by some researchers that various evaluations of performance were distinct but related, and could be combined. The findings from both data sets supported the assumption. We found that goal-related performance was similar to both satisfaction measures and to non-goal performance, and there was support for combining each of these measures with goal-related performance. While we only considered a subset of the variables that have been used, it appears that

researchers using multiple indicators of performance may combine performance indicators based upon goal-related performance with satisfaction and with non-goal performance.

An assumption on how to combine different measures was evaluated by whether the measures should be added together or weighted based upon the criterion's importance. Findings from our two analyses indicated that these measures are highly related and that this assumption was supported. Researchers may not have to worry about whether performance measures are simply added together or first weighted before combining. Future research should consider exploring more complex types of combinations. As noted, we used cross-sectional data that weighted current measures. Over time, however, the importance of some measures may change in importance, increasing the impact of weighting. For our data, however, we found no effect.

Our final assumption tested was whether one company's performance evaluations were comparable to its partner's evaluations. Our findings suggest the relationship varied by the measure used for sample 1 and was not supported for sample 2. Unfortunately, there is no clear relationship between the type of performance measurement and similarity in evaluations in sample 1. It appears, though, that researchers should not assume that one company's perspective extends to others. Future research might extend these findings by considering how the level of consensus between or among companies varies by the alliance's and the partners' context as well as the types of measures used. There may be situations where the criteria of one company do in fact reflect the partner's. Until these situations are known, we can conclude that researchers will need to confirm one partner's evaluation reflects another's for each dataset.

Although this work provides a test of these assumptions and furthers our understanding about how to evaluate an alliance, and an important strength of our study is that we tested most of our hypothesis with two data sets and found similar results, we recognize that one study cannot evaluate the variety of different performance issues and that there are limitations to it. We only collected data from one informant in each company. While we conducted reliability checks, we were not able to examine the fairly common assumption that these informants are reliable indicators of their company's strategy. In addition, profit represented only a small percentage of sample 1 company's primary motive. In other samples, where companies have more profit interests, the relationship between motives and performance criteria may differ from the ones we found. Besides these suggestions, another area in which to extend these findings is to examine predictors of performance. There is some evidence (e.g., Erden, 1997) that even though different indicators of alliance performance are related, they are not equivalent and have different predictors. Understanding what each variable captures, and if or when different variables are equivalent, will help researchers understand how findings from different studies compared. Finally, future research should continue to explore the relationship among different types of performance measures. The present study addressed two common assumptions found in studies using multiple measures: evaluating whether the select measures are significantly related (#3) and if they can be added together to create a composite score (#4). Other researchers have argued that some performance indicators may be interim measures for subsequent measures of performance (Hatfield *et al.*, 1998). For example, meeting one's strategic objectives may precede an overall evaluation of satisfaction. Because each of our datasets was a cross-sectional survey, we were not able to examine temporal relationships among the measures. Exploring in the future such relationships will add needed insights into strategic alliance evaluation.

These findings offer several specific implications for theories about strategic alliances. First, theories that are based on specific strategic interests of the company may

have to pay special attention to the criteria they use in evaluating performance and to recognize that similar types of criteria may not be used for each strategic interest. For efficiency-interested theories (e.g., transaction cost), it appears that focusing on efficiency criteria is not adequate. Likewise for learning approaches, learning criteria do not appear to capture adequately all the performance interests of the partner. Second, research that examines bilateral issues –such as partner trust, negotiations, decision-making– and relates these to performance cannot rely upon a single partner's evaluation. Researchers will have to collect data on both partners. Finally, there appears to be significant room for researchers to contribute to our alliance understanding by examining multiple measures of performance, particularly those that take a process and a dynamic perspective to alliance performance. We know very little about these measures and there is an abundance of research that can be conducted that will shed light on the development of an alliance.

Our research also has implications for managers. First, our finding on the relative importance of strategic interests suggests that managers should focus on evaluating performance against attaining the strategic reasons for participation. While this may seem obvious, we have witnessed alliances where not all managers in the company are aware of the strategic reasons for participation. And even if they do know the reasons for participation, they often do not know their performance on these criteria. In the absence of this information, we have noticed that they tend to focus on other measures, such as satisfaction or stability, which as we found may not reflect strategic performance. Second, our finding on the difference among partners with respect to their performance indicates that managers need to realize that their company's performance may not be the same as their partner's, and to be aware of the consequences of these differences. Again, we have witnessed alliances where the partners were not cognizant of each other's interests or performance in the alliance. This lack of attention to the partner's interests led to problems as these alliances developed. Finally, our finding on how to combine different performance measures suggests that managers may not need complex metrics to evaluate an alliance's performance. Measures that are simpler to collect, and likely simpler to understand, may provide similar evaluations. In summary, we set out to investigate five common assumptions found in the strategic alliance literature. We found that while there is support for some of these assumptions, there are areas in which strategic alliance researchers need to be careful in how they evaluate an alliance. Building from our assessment will provide additional guidance on how to evaluate a strategic alliance, and continuing to investigate this issue will improve the practice and theory of how to develop an effective strategic alliance.□

REFERENCES

Ariño A. 2003. Measures of strategic alliance performance: an analysis of construct validity. *Journal of International Business Studies* **34**: 66-79.

Armstrong JS, Overton T. 1977. Estimating nonresponse bias in mail surveys. *Journal of Marketing Research,* **14**: 396-402.

Beamish P, Delios A. 1997. Improving joint venture performance through congruent measures of success. *In Cooperative Strategies: European Perspectives, Beamish PW,* Killing JP (eds.). The New Lexington Press: San Francisco; 103-127.

Bergh D, Fairbank J. 2002. Measuring and testing change in strategic management research. *Strategic Management Journal* **23**: 359-365.

Brockoff K, Teichert T. 1995. Cooperative R&D and partners' measures of success. *International Journal of Technology Management* **10**: 111-123.

Contractor F, Lorange P. 1988. Why should firms cooperate? In *Cooperative Strategies in International Business*, Contractor F, Lorange P (eds.). Lexington Books: Lexington, MA; 3-28.

Cropper S. 1996. Collaboration in practice: key issues. In *Creating Collaborative Advantage,* Huxham C (ed.). Sage Publication: London; 80-100.

Cullen J, Johnson J, Sakano T. 1995. Japanese and local partner commitments to IJVs: psychological consequences of outcomes and investments in the IJV relationship. J*ournal of International Business Studies* **26**: 91-116.

Erden D. 1997. Stability and satisfaction in cooperative FDI: partnerships in Turkey. In *Cooperative Strategies: European Perspectives.* Beamish PW, Killing JP (eds.). The New Lexington Press: San Francisco; 158-183.

Geringer, JM. 1998. Assessing replication and extension. A commentary on Glaister and Buckley: measures of performance in UK international alliances. *Organization Studies* **19**: 119-138.

Geringer JM, Herbert L. 1991 Measuring performance of international joint ventures. *Journal of International Business Studies* **22**: 249-263.

Glaister KW, Buckley PJ. 1998. Measures of performance in UK international alliances. *Organization Studies* **19**: 89-118.

Gray B. 2000. Assessing inter-organizational collaboration: multiple conceptions and multiple methods. In *Cooperative Strategies: Economic Business and Organizational Issues*. Faulkner D, de Rond M (eds.). Oxford University Press: Oxford; 243-260.

Gulati R. 1998. Alliances and networks. *Strategic Management Journal* **19**: 293-317.

Hamel G. 1991. Competition for competence and inter-partner learning within international strategic alliances. *Strategic Management Journal* **12**, 83-103.

Hatfield L, Pearce JA, Sleeth R, Pitts M. 1998. Toward validation of partner goal achievement as a measure of joint venture performance. *Journal of Managerial Issues* **10**: 355-372.

Hirsch P, Levin D. 1999. Umbrella advocates versus validity police: a life-cycle model. *Organization Science* **10**: 199-212.

Hoskisson R, Hitt M, Johnson R, Moesel D. 1993. Construct validity of an objective (entropy) categorical measure of diversification strategy. *Strategic Management Journal* **14**, 215-234.

Inkpen A. 2002. Learning, knowledge management, and strategic alliances: so many studies, so many unanswered questions. In *Cooperative Strategies and Alliances*, Contractor F, Lorange P (eds.). Pergamon: Amsterdam; 267-289.

Inkpen A, Ross J. 2001. Why do some strategic alliances persist beyond their useful life? *California Management Review* **44**: 132-148.

Kale P, Singh H, Perlmutter H. 2000. Learning and protection of proprietary assets in strategic alliances: building relational capital. *Strategic Management Journal* **21**: 217-238.

Kumar N, Stern L, Anderson J. 1993. Conducting interorganizational research using key informants. *Academy of Management Journal* **36**:1633-1651.

Lampel J, Shamsie J. 2000 Probing the unobtrusive link: dominant logic and the design of joint ventures in General Electric. *Strategic Management Journal* **21**: 593-602.

Lane P, Lubatkin M. 1998. Relative absorptive capacity and interorganizational learning. *Strategic Management Journal* **19**: 461-478.

Lee C, Beamish P. 1995. The characteristics and performance of Korean joint ventures in LDCs. *Journal of International Business Studies* **26**: 637-654.

Lin X, Germain R. 1998. Sustaining satisfactory joint venture relationships: the role of conflict resolution strategy. *Journal of International Business Studies* **29**: 179-196.

Lorange P, Roos J. 1992. *Strategic Alliances: Formation, Implementation and Evolution.* Blackwell Publishers: Cambridge, MA.

Luo Y, Chen M. 1997. Business strategy investment strategy and performance of international joint ventures: the case of China. In *Cooperative Strategies: Asian Perspectives*, Beamish PW, Killing JP (eds.). The New Lexington Press: San Francisco; 341-374.

Lyles M, Salk J. 1997. Knowledge acquisition from foreign parents in international joint ventures: an empirical examination in the Hungarian context. In *Cooperative Strategies: European Perspectives*, Beamish PW, Killing JP (eds.). The New Lexington Press: San Francisco; 325-355.

Makino S, Delios A. 1997. Local knowledge transfer and performance: implications for alliance formation in Asia. *Journal of International Business Studies* **27**: 905-927.

Marois B, Abdessemed T. 1996. Cross-border alliances in the French banking sector. *International Studies of Management and Organization* **26**: 35-58.

Meschi P-X. 1997. Longevity and cultural differences of international joint ventures: toward time-based cultural management. *Human Relations* **50**: 211-228

Mjoen H, Tallman S. 1997. Control and performance in international joint ventures. *Organization Science* **8**: 257-274.

Nath D, Gruca T. 1997. Convergence across alternative methods for forming strategic groups. *Strategic Management Journal* **14**, 745-760.

Olk P. 2002. Evaluating strategic alliance performance. In *Cooperative Strategies and Alliances*, Contractor F, Lorange P (eds.). Pergamon: Amsterdam; 119-143.

Osland G, Cavusgil S. 1996. Performance issues in U.S.-China joint ventures. *California Management Review* **38**: 106-130.

Park S, Ungson G. 1997. The effect of national culture, organizational complementarity, and economic motivation on joint venture dissolution. *Academy of Management Journal* **40**: 279-307.

Parkhe A. 1993a. Partner nationality and the structure-performance relationship in strategic alliances. *Organization Science* **4**: 301-314.

Parkhe A. 1993b. Strategic alliance structuring: a game theoretic and transaction cost examination of interfirm cooperation. *Academy of Management Journal* **36**: 794-829.

Reuer J, Koza M. 2000. International joint venture instability and corporate strategy. In *Cooperative Strategies: Economic Business and Organizational Issues*, Faulkner D, de Rond M (eds.). Oxford University Press: Oxford; 261-280.

Saxton T. 1997. The effects of partner and relationship characteristics on alliance outcomes. *Academy of Management Journal* **40**: 443-462.

Spekman R, Isabella L, MacAvoy T, Forbes T. 1996. Creating strategic alliances which endure. *Long Range Planning* **29**: 346-357.

Steensma K, Corley K. 2000. On the performance of technology-sourcing partnerships: the interaction between partner interdependence and technology attributes. *Academy of Management Journal* **43**: 1045-1067.

Steensma K, Lyles M. 2000. Explaining IJV survival in a transitional economy through social exchange and knowledge-based perspectives. *Strategic Management Journal* **21**: 831-852.

Yan A, Gray B. 1994. Bargaining power management control and performance in United States - China joint ventures: a comparative case study. *Academy of Management Journal* **37**: 1478-1517.

Yan A, Zeng M. 1999. International joint venture instability: a critique of previous research, a reconceptualization, and directions for future research. *Journal of International Business Studies* **30**: 397-414.

Zaheer A, McEvily B, Perrone V. 1998. Does trust matter? Exploring the effects of interorganizational and interpersonal trust on performance. *Organization Science* **9**: 123-141.

Appendix 1

Variables and questions used to derive variables, Study 1

Propensity to Stay (reverse coded)

Will your organization continue to belong to this consortium? 1=Yes, 2=No

Overall Satisfaction

Average of responses to three questions (#3 is reverse coded)

1. Overall, how effective is this consortium? (1=not at all, 7=extremely)

2. To what extent has this consortium met your expectations for its research and development activities? (1=not at all, 7=completely)

3. If you had your way, to what extent would you change the manner in which this consortium conducts its activities? (1=not at all, 7=completely)

Goals

How important is each of the following as a reason for your organization's membership in this consortium? (1=not at all, 7=very)

Learning motive was represented by the following items

– Learn about or develop a new technology
– Improve your organization's capabilities
– Train researchers

Profit motive was represented by the following items

– Deter entry of other organizations into this research area
– Make a profit

Efficiency motive was represented by the following items

– Share risks of uncertain research
– Avoid unnecessary duplication of research

Appendix 2

Variables and questions used to derive variables, Study 2

Alliance Survival

Is the venture still operating? 1=Yes, 0=No

Overall satisfaction / Partner's overall satisfaction

Overall, to what extent do you think YOUR FIRM / YOUR PARTNER is satisfied with the global results of this venture? (1=very unsatisfied, 5= very satisfied)

Strategic Goal Fulfillment / Partner's Strategic Goal Fulfillment

Product of responses to two questions

1. Collaborative ventures can be aimed at different strategic goals. How would you describe the importance for YOUR FIRM / YOUR PARTNER of each of the following strategic goals when the venture agreement WAS SIGNED?

   (1=minimal, 5= vital; 6=n.a.)

   a. Reducing costs/obtaining scale economies
   b. Gaining access to a market in the same industry
   c. Gaining access to a market in another industry
   d. Developing new technologies

   e. Blocking the competition
   f. Meeting government requirements
   g. Developing new skills
   h. Reducing risks
   i. Other (specify)

2. How do you think each of YOUR FIRM'S / YOUR PARTNER'S strategic goals in relation to this venture has been met?

   (1=very poorly, 5= very well; 6=n.a.)

   (Same list of goals as above)

Net Spillover Effects

Many collaborative ventures result in SIDE EFFECTS for their parent firms. For example, there are POSITIVE side effects when the skills that are being developed through the venture can be applied profitably to other operations within the company. There are NEGATIVE side effects if the collaboration has damaging repercussions on other activities in the company. In this venture, the NET side effects for YOUR FIRM have been: (1=strongly negative, 5=very positive).