University of Navarra

Working Paper

MEASURING THE PERFORMANCE OF INTERNATIONAL LOGISTICS OUTSOURCING PARTNERSHIPS: A DYADIC PERSPECTIVE ANALYSIS

Joan Jané*
Alejandro Lago**
Africa Ariño***

---

*   Hewlett-Packard
**   Professor of Production, Technology and Operations Management, IESE
*** Professor of General Management, IESE

MEASURING THE PERFORMANCE OF INTERNATIONAL LOGISTICS OUTSOURCING PARTNERSHIPS: A DYADIC PERSPECTIVE ANALYSIS

Abstract

We analyze the validity of five performance measures of international logistics outsourcing partnerships, using information from both partners. Each partner's assessment of performance is captured by a single construct, which underlies four of the measures. This construct, however, is different for each party. Consequently, we examine a focal partner's perceptions of the other partner's performance assessment, and show that these *inter-party perceptions* are a poor measure of the latter's actual performance assessment.

MEASURING THE PERFORMANCE OF INTERNATIONAL LOGISTICS OUTSOURCING PARTNERSHIPS: A DYADIC PERSPECTIVE ANALYSIS*

Introduction

The measurement of strategic alliance performance has become an object of study for international management researchers (Geringer and Hebert, 1991; Yan and Zeng, 1999; Ariño, 2003). Despite the various studies on the topic, however, evidence on the validity of the available measures of alliance performance is inconclusive. Doubts as to the most appropriate methodological approach to deal with both partners' assessments of alliance performance persist. This study advances knowledge in this area by analyzing the construct validity of performance measures for a particular type of alliance – international logistics outsourcing partnerships – using information from both sides of the dyad.

Logistics outsourcing partnerships are a type of strategic alliance that has become common in today's international business environment. Analysis of their performance has received increasing attention in the literature (Lieb and Randall, 1996; Sink and Langley, 1997; Murphy and Poist, 2000). Logistics partnerships are a particular class of long-term arrangement in which there is an asymmetric relationship between the partners (i.e., a buyer vs. seller arrangement). Also, they are based on a service, rather than on goods. Hence, it is important to extend our empirical knowledge of alliances with these characteristics (i.e., partner asymmetry and service objectives).

The strategic alliance literature has acknowledged the multi-dimensional nature of performance (e.g., Geringer and Hebert, 1991). Similarly, case- and interview-based research has revealed that the partners in a logistics partnership may pursue a wide variety of goals (Lambert et al., 1999). Yet, to date, survey-based research on logistics alliances has measured performance only in the aggregate, e.g., using a single question about partnership satisfaction (e.g., Lieb and Randall, 1996; Sink and Langley, 1997; Murphy and Poist, 2000). Hence, there are benefits to be gained by analyzing the validity of different performance measures in more detail, so as to complement existing evidence in the field of strategic alliances, and asymmetric partnerships in particular.

More importantly, the empirical analysis of strategic alliance performance is limited because information is often obtained from one of the partners only (see Parkhe, 1993). Consequently, it is often assumed – either explicitly or implicitly – that performance measures and evaluation are common to both sides of the partnership. Evidence that compares information from both parties is very scarce (Geringer and Hebert, 1991) and inconclusive. Empirical work on the performance of logistics partnerships is not exempt from this same problem.

In this paper, we address these limitations. We surveyed and analyzed a sample of 95 international logistics partnerships, asking *both partners* – logistics provider and customer – about their assessment of partnership performance along different dimensions. We succeeded in getting answers from both parties in 73 cases. This paper thus offers three main contributions. First, we determine which measures are appropriate for evaluating the performance of international logistics outsourcing partnerships. As empirical evidence in the field is not very broad, it is relevant to draw from similar studies in the field of strategic alliances (Geringer and Hebert, 1991; Ariño, 2003) and build on them with new and more complete data. Second, because our database includes information from both sides of the dyad, we are able to test whether performance measures are common to both partners, whether the partners' performance assessments are correlated, and to what extent a focal partner's perceptions about the other partner's performance assessment –what we call *inter-party perceptions*– may be a reliable measure of the latter's actual performance assessment. To the best of our knowledge, this is the first empirical study in the field of asymmetric partnerships to consider the perspective of both partners,[1] and one of the very few to have done so in the broader field of strategic alliances. Finally, responding to past calls to fill an existing gap (Yan and Zeng, 1999; Ariño, 2003), we develop and test the validity of a measure of process performance.

The paper is organized as follows. We start by discussing the nature of logistics partnerships and the content validity of measures of their performance. We do so in the context of the literature on strategic alliance performance measurement. The following section contains methodological details of our survey of international logistics outsourcing partnerships and presents the operational measures evaluated. Then, we analyze the empirical validity of these measures of performance for each partner. Next, we complete the empirical analysis by assessing the validity of *inter-party perceptions* as measures of actual performance assessments. We conclude by discussing implications for future research.

Logistics outsourcing partnerships and their performance

In this section, we define what we understand by a logistics outsourcing partnership, considered as a particular class of strategic alliance, and review key literature on measures of strategic alliance performance. Then, we turn to discuss the content validity of these measures in our case.

A strategic alliance is defined as a formal agreement between two or more business organizations to pursue a set of private and/or common interests through the sharing of resources in contexts involving uncertainty about outcomes (see Ariño, 2003). As such, a logistics outsourcing partnership may be considered one type of strategic alliance: a partnership which involves a logistics provider, often referred to as a third party logistics

---

[1] Murphy and Poist (2000) analyzed differences in satisfaction levels between third party logistics providers and customers and found that differences were non-significant. The study, however, was not based on matched pairs or *dyads*, i.e., logistics providers' and customers' answers did not necessarily represent the same partnership.

provider or 3PL, offering a wide array of customized services – including transportation, warehousing, inventory control, distribution and other value-added activities – to a customer firm on a long-term basis and aimed at achieving specific objectives and mutual benefits (Bagchi and Virum, 1996; Cooper and Gardner, 1993; Murphy and Poist, 2000). Hence, it is appropriate to approach the analysis of international logistics partnership performance measures using available knowledge on the measurement of the performance of strategic alliances in general.

Few studies (Geringer and Hebert, 1991; Yan and Zeng, 1999; Ariño, 2003) have focused on the conceptual underpinnings and empirical validity of measures of alliance performance. At the conceptual level, Yan and Zeng (1999) elaborate on the complex relationship between alliance instability and performance, and advocate a new research focus on how alliance changes and evolutionary processes influence their performance. Ariño (2003) – building on work in the strategic management field (Venkatraman and Ramanujam, 1986) – recognizes three different levels of performance, depending on the goals to be achieved: financial performance, operational performance and organizational effectiveness. She argues that organizational effectiveness – understood as the fulfillment of the partnership goals, taking into account the interests of the multiple constituencies involved – is the most comprehensive of the three. Following Yan and Zeng (1999), she suggests that the concept of strategic alliance performance be broadened to encompass both outcome and process performance.

At the empirical level, Geringer and Hebert (1991) compare a range of objective and subjective measures of alliance performance. They analyze a sample of 69 dyadic joint ventures with one U.S.-partner. They find correlations among objective and subjective measures to be generally positive and significant, the correlations being stronger among measures of overall performance than among measures based on more specific individual dimensions. Using a sample of 16 dyadic and triadic joint ventures from Canada, they also found a focal partner's evaluation of the other partner's satisfaction with the joint venture's performance to be correlated with the latter's actual reported satisfaction. Hence, the authors conclude that collecting performance data from a single participant might be appropriate, although more so for overall performance information than for detailed, specific aspects of performance.

Ariño (2003) evaluates the construct validity of measures of alliance performance in two samples, one consisting of 34 equity alliances and the other of 45 purely contractual ones. All of the alliances were dyadic and had one partner operating from Spain. Based on her conceptual discussion of content validity, she focuses on subjective organizational effectiveness measures, and analyzes them using a structural analysis approach. She finds that a specific measure such as the fulfillment of strategic goals reflects an underlying construct different from other general measures such as overall performance satisfaction or net spillover effects. These two measures had convergent and discriminant validity relative to objective measures such as contractual changes and survival, but not relative to longevity. As mentioned earlier, she argues the need to develop measures of process performance, as current measures do not capture it properly.

Based on insights from this research, we turn now to discuss the content validity of different existing measures of organizational effectiveness as applied to the case of international logistics partnerships. The most common measure of effectiveness is each partner's *satisfaction with overall performance* (see Lieb and Randall, 1996; Sink and Langley, 1997; Murphy and Poist, 2000 for logistics partnerships; and Parkhe, 1993 for other types of strategic alliances). From a conceptual point of view, it is also the most complete one, as it entails satisfaction with all possible goals – initial and emergent.

Alternatively, one could directly measure *overall satisfaction with the fulfillment* of *expectations* – i.e., a partner's *initial goals,* private or otherwise.

Performance can also be measured more specifically by the degree of *fulfillment of strategic goals*. These normally refer to some objectives initially set for the partnership (see Parkhe, 1993). Usual objectives sought in logistics partnerships include: cost reduction through specialization, enhanced customer service, reduced risk and uncertainty, and strategic flexibility (Frankel at al., 1996). Most of these goals could be considered private goals for each of the partners. Not surprisingly, 3PLs and customers pursue different strategic goals. 3PLs seek profit, profit stability and competitive advantage, while often customers look for improvements in customer service levels, in costs, and in strategic flexibility (Laarhoven and Sharman; 1996; Lambert et al., 1999). Underlying these objectives are the partners' particular views of how to gain competitive advantage (Bowersox, 1990).

Besides these goals, which are pursued decisively, additional benefits are often achieved, such as improvement in decision making, or size reduction of the logistics department (Laarhoven and Sharman, 1996). These emergent objectives – not necessarily expected at the beginning of the partnership – are referred to as *net spillover effects*.

Finally, monitoring of logistics limited to the physical activities and their interrelated administrative functions seems insufficient. It has been suggested that process performance may also be a valid performance measure for logistics alliances, including aspects of how well inter-organizational relations are working (AT Kearney, 1995; Bagchi and Virum, 1996).

In summary, we may consider at least five measures of logistics partnership performance that present content validity. Next, we evaluate empirically if they are reliable measures, and the extent to which they all represent a single latent construct in the case of both 3PLs and customers.

Methods

Sample

The target population for our study was logistics partnerships in which the 3PL company was among the top 75 in Europe. Because of reliability considerations, we included in the sample only partnerships for which the target respondent in the 3PL company could be clearly identified. Each selected manager from 49 3PLs was asked to choose a minimum of three logistics partnerships with a European scope – one considered as successful, another considered as unsuccessful, and a third one selected randomly – and to fill in one questionnaire per partnership. They were also asked to invite the customer in each contract to *independently* evaluate the partnership using a questionnaire that we had adapted for that purpose (for more details on the survey instrument, see below). The 49 companies selected included the top 10 3PLs in the European market, as well as 17 of the top 35 3PLs worldwide. Forty-two were European companies and seven were from elsewhere, but all had a relevant presence in the European market.

Questionnaires were sent electronically to all 49 managers, but 12 of them declined to participate in the study. We received responses regarding 95 partnerships. These include 95 responses from 3PLs, and 73 matched pairs from the customer side (50 percent response rate at the dyad level). We attribute this high response rate to the care taken in identifying

the appropriate respondent and to the follow-up procedure used (Dillman, 1978), which included supplemental phone calls. To further encourage responses, respondents were assured confidentiality and access to the study findings. We had asked respondents to choose partnerships with clients in certain suggested areas of activities. Table 1 outlines the characteristics of the sample in terms of firm size as measured by gross revenue and number of employees of each partner company, and in terms of customers' industry.

Table 1. Sample characteristics

Third Party Logistics Providers[a]				
Revenue (million €)		# employees		
Category	Distribution	Category	Distribution	
<500	20.0%	<2,500	8.7%	
>500 and <1,000	10.5%	>2,500 and <7,500	23.9%	
>1,000 and <5,000	28.4%	>7,500 and < 15,000	18.5%	
>5,000 and <10,000	21.1%	>15,000 and <30,000	14.1%	
>10,000 and <20,000	6.3%	>30,000 and < 50,000	12,0%	
>20,000	13.7%	>50,000	22.8%	

Customers[b]					
Revenue (million €)		# employees		Industry	
Category	Distribution	Category	Distribution	Category	Distribution
<500	18.8%	<2,500	25.8%	Technology	42.1%
>5,000 and <10,000	18.8%	>2,500 and <7,500	10.6%	Automotive	16.8%
>10,000 and <30,000	10.9%	>7,500 and < 15,000	12.1%	Fast moving consumer goods	10.5%
>30,000 and <60,000	14.1%	>15,000 and <30,000	24.2%	Pharma/Healthcare	9.5%
>60,000 and <100,000	17.2%	>30,000 and < 50,000	7.6%	Retail	9.5%
>100,000	20.3%	>50,000	19.7%	Indsutrial equippment	6.3%
				Services	3.2%
				Chemical	2.1%

[a]N = 95. [b]N = 73.

The target informants were the senior managers most directly related to these partnerships. For 3PLs, these were typically European VPs in global accounts or business development, while for customers they were European logistics or supply chain directors. For 3PLs, respondents' average years of experience in the company were 7.3, with 3.4 years in the last position and 2.6 years involved in the logistics partnership. For customers, these figures were 9.7, 3.7, and 3.3 years, respectively. In addition, 70% of the 3PLs' and 75% of the customers' respondents had participated in the negotiation of that specific logistics partnership.

In order to provide some external validity, we checked 3PL size in terms of gross income and number of employees against available secondary data. Matches were present for 97% and 96% of the cases, respectively. We also checked information provided by both 3PLs and customers on partnership sector and type of logistics services considered. In this case, matches were found for 98% and 90% of the cases, respectively.

The questionnaires corresponding to the 73 partnerships for which we had responses from both sides contained all of the necessary information. These 73 alliances

involved thirty-three 3PLs. At each 3PL, questionnaires were answered by different people. In order to examine potential nonresponse bias, we assessed possible differences in customers' industries and in 3PL size, measured by gross income, between early and late respondents, on the assumption that late respondents are more similar to non-respondents than early respondents are to non-respondents (Armstrong & Overton, 1977). An analysis comparing the sectoral distribution of alliances for early and late respondents, and a one-way ANOVA for firm size across these groups, gave insignificant results, providing no indication of response bias. Additional descriptive statistics on the sample appear in the results section.

Survey Instrument

Preliminary versions of the questionnaire were reviewed by business scholars and logistics managers with experience in this type of partnership to ensure face validity. We sought to address the possibility of consistency artifacts and common method bias. First, we arranged the questionnaire items (see Appendix) so that the subjective items appeared prior to objective data (Salancik & Pfeffer, 1977). Second, we used Harman's (1967) single-factor test to examine whether a significant amount of common method variance exists in the data. If so, a factor analysis of all of the variables will generate a single factor or a general factor that accounts for most of the variance in the data (e.g., Podsakoff & Organ, 1986). Unrotated factor analysis using the eigenvalue-greater-than-one criterion revealed four factors, and the first factor explained only 21.7 percent of the variance in the data. Thus, we concluded that the analysis was not subject to common method bias.

Measures of performance

Following the discussion in the previous section, we considered the following five measures of logistics partnership performance (see Appendix for details):

- *Overall expectations fulfillment* is a seven-point scale measuring the informants' assessment of the extent to which their firm is satisfied with the fulfillment of expectations with respect to the partnership.

- *Global performance satisfaction* is a seven-point scale measuring the informants' assessment of the extent to which their firm is satisfied with the global performance of the partnership.

- *Strategic goal fulfillment* is a composite measure obtained as the average of the level of fulfillment of specific strategic goals weighted by the importance assigned to each of these goals, as assessed by the informants. The specific goals considered were different for 3PLs and for customers (see questions 1P and 1C in the Appendix). These goals (six for the 3PLs and eight for the customers) were identified through a literature review, and through prior fieldwork, including interviews with twenty companies involved in logistics partnerships (twelve 3PLs selected among the European top 20 and representing around 38% of the European market share in 2004, and eight customer firms that were leaders in different industries in Europe).

- *Net spillover effects* captures the difference between the positive and negative effects of the partnership on other company activities. Based on the measure developed by Parkhe (1993), this variable is the cumulative sum of the informants' assessment of the extent to which five possible effects of the

partnership are present (seven-point scale; Cronbach's alpha values were .72 for the 3PLs sub-sample, and .81 for the customer one). By means of the prior interviews referred to above, five possible such effects were identified for the 3PLs, and five others for the customers (see questions 5P and 5C in the Appendix).

- *Process performance* is also a composite measure obtained as the average of the informants' assessment of the extent to which their firm is satisfied along five dimensions of the partnership process (seven-point scale; Cronbach's alpha values of .86 and .84 were obtained for the 3PL and customer sub-samples, respectively) (see question 6P in the Appendix). This measure is a novel one and was built in response to Ariño's (2003) call for the development of measures of strategic alliance process performance.

Empirical validity of measures of logistics partnership performance for 3PLs and customers

In this section we evaluate the empirical validity of measures of logistics partnership performance. First, we assess the empirical validity of these measures in the 3PL and in the customer sub-samples independently. Second, we analyze whether 3PLs and customers are evaluating one and the same performance construct or whether their assessments capture different performance evaluations. In other words, we evaluate the *convergent validity* of assessments of partnership performance made by 3PLs and by customers, and the *discriminant validity* among such assessments.[2]

Reliability

Table 2 shows the descriptive statistics and correlations of the variables in the study for the 3PL and customer sub-samples. All correlations are high and significant at least at the .01 level, and suggest the existence of a single latent construct.

Table 2. Descriptive statistics and correlations by sub-sample

Variable	Third Party Logistics Providers[a]							Customers[b]						
	Mean	s.d.	1	2	3	4	5	Mean	s.d.	1	2	3	4	5
(1) Overall expectations fulfillment	5.29	1.17	1.00					5.19	1.09	1.00				
(2) Global performance satisfaction	5.06	1.33	0.69	1.00				5.07	1.16	0.87	1.00			
(3) Strategic goal fulfillment	5.20	0.92	0.55	0.50	1.00			5.05	0.83	0.76	0.72	1.00		
(4) Net spillover effects	22.11	4.96	0.48	0.40	0.48	1.00		21.07	6.09	0.39	0.44	0.42	1.00	
(5) Process performance	4.86	1.09	0.67	0.68	0.65	0.38	1.00	4.83	1.02	0.72	0.72	0.69	0.37	1.00

[a]N=95. Correlations above r=0.38 are significant at P<0.01.
[b]N=73. Correlations above r=0.37 are significant at P<0.01.

---

[2] Although a complete analysis of construct validity would require assessment of criterion-related validity, we believe no criterion is available in this case (see Ariño, 2003).

Following Ariño (2003), we performed two tests to assess the reliability of measures of logistics partnership performance (see Table 3):

Table 3 Reliability tests for measures of logistics partnership performance

Variable	Third Party Logistics Providers[a]		Customers[b]	
	Reliability coefficient	Alpha if item deleted	Reliability coefficient	Alpha if item deleted
Overall expectations fulfillment	0.74	0.52	0.85	0.43
Global performance satisfaction	0.69	0.52	0.84	0.41
Strategic goal fulfillment	0.62	0.56	0.76	0.46
Net spillover effects	0.42	0.86	0.32	0.92
Process performance	0.73	0.54	0.73	0.45
(Alpha)		0.62		0.52

[a]N = 95. [b]N = 73.

Test 1: reliability coefficient. We performed a 1-factor exploratory analysis and obtained the commonalities for the five measures of performance for 3PLs and for customers separately, as shown in Table 3. Commonalities indicate the percentage of variance explained by the single factor – i.e., they are equivalent to the reliability coefficient. Note that the reliability coefficients are substantially high, except for *net spillover effects*.

Test 2: Cronbach's alpha. Cronbach's alpha for the five performance measures as a set took a value of .61 for the 3PLs sub-sample, and .52 for the customers sub-sample (below the recommended .60 limit). When the *net spillover effects* measure was dropped out, this coefficient increased to .86 and .96, respectively.

The results from these two tests suggest that single constructs of partnership performance as assessed by 3PLs and as assessed by customers may exist. Nevertheless, the *net spillover effects* measure may be partially explained by another latent factor. This is more evident in the customer sub-sample.

Since the measures are reliable, it is pertinent to assess their convergent and discriminant validity across sub-samples. Hence, we turn now to analyze whether joint valid measures of partnership performance can be developed that might be applied to evaluate performance assessments of 3PLs and customers indistinctly. That is, we want to know if 3PLs' performance assessments coincide with those of the customers. A preliminary analysis not reported here showed that most correlations between performance measures evaluated by 3PLs and the equivalent measures evaluated by customers are significant.

Convergent validity

Convergent validity of measures of logistics partnership performance assessed by 3PLs and by customers is examined through two tests carried out in measurement Model 1 (see Figure 1), which excludes *net spillovers effects*, given its low reliability relative to the other measures. For the sake of completeness, we analyzed two other models: one in which *net spillover effects* was considered as measuring a concept other than partnership performance; and another in which *net spillover effects* was considered as a valid measure of partnership performance. Both performed worse than Model 1 presented here. The results of the tests are shown in Figure 1 and Table 4.

Figure 1. Measurement Model 1 with standardized coefficients
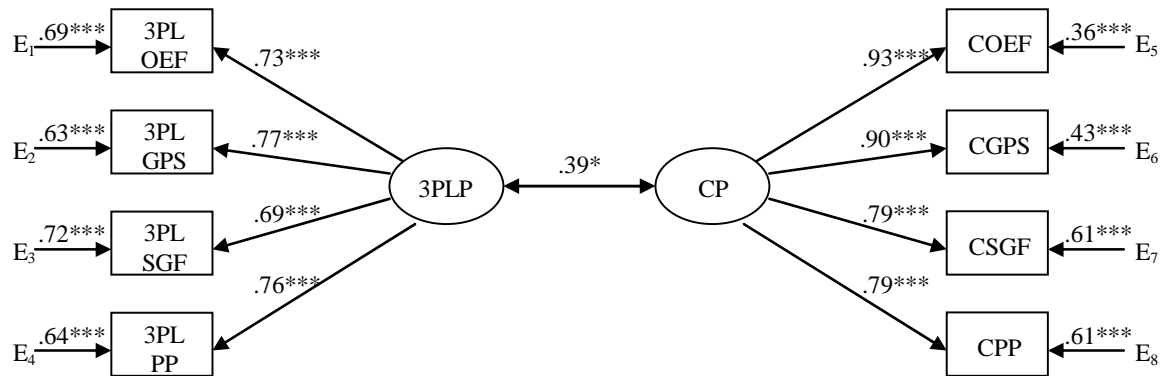
N=73.

Significance levels obtained from non-standardized solutions:

*** significant at p<.005

** significant at p<.01

* significant at p<.05

3PLP = logistics provider's (3PL's) performance; 3PLOEF = 3PL's overall expectations fulfillment; 3PLGPS = 3PL's global performance satisfaction; 3PLSGF = 3PL's strategic goal fulfillment; 3PLPP = 3PL's process performance.

CP = customer's performance; COEF = customer's overall expectations fulfillment; CGPS = customer's global performance satisfaction; CSGF = customer's strategic goal fulfillment; CPP = customer's process performance.

E_i: error terms.

Table 4. Tests of measurement models with goodness-of-fit indices

Models	X^2	d.f.	P	CFI	NNFI	X^2 difference	d.f. change	P
(1) Baseline model	21.68	19	0.30	0.99	0.99	--	--	--
(1b) Spillover effects in	42.17	34	0.16	0.97	0.96	20.49	15	0.15
(1c) Performance: 3PL=Customer	98.29	18	0.00	0.73	0.62	76.61	1	0.00

N=73.

Test 1: Overall goodness-of-fit. Non-significant values for a model X^2, and magnitudes of .90 or greater for the comparative fit index (CFI) and non-normed fit index (NNFI) provide evidence of an acceptable fit between model and data (Bentler, 1992). Model 1 showed an acceptable goodness-of-fit (see Table 4).

Test 2: Significance of the factor loadings and factor correlations. A z-test showing all of the observed variables measuring a construct to be significant provides evidence of convergent validity (Bentler, 1992). In our case, all of the factor loadings were significant at the .005 level (see Figure 1). The correlation between these latent constructs capturing partnership performance as assessed by 3PLs and by customers is also significant at least at the .05 level.

Discriminant validity

We tested two alternative models to determine the discriminant validity of our measurement model. First, we tested a model in which *net spillover effects* was correlated with the latent construct *partnership performance* (Model 1b). Secondly, by constraining the correlation between the two latent constructs in Model 1 to be equal to 1, we tested whether there could be a single *partnership performance* latent construct that was common to 3PLs and customers (Model 1c). We use a *Chi-square difference test* to assess discriminant validity. This test compares a baseline model (in our case, Model 1) with a more restricted model in which the correlation between the two constructs under examination is constrained to equal 1.0 (Joreskog, 1971). Evidence of discriminant validity is provided by a significantly higher chi-square for the model in which the correlation is restricted, as this would indicate a non-perfect correlation between the constructs (Bagozzi and Phillips, 1982).

The results are shown in Table 4. Regarding Model 1b, the goodness of fit is worse than in Model 1, as assessed by the X^2, CFI and NNFI statistics. This result, together with the low convergent validity of the spillover measure obtained above, suggests that Model 1 is preferable to Model 1b. Notwithstanding, Model 1b shows a non-significant increase in the value of X^2 relative to Model 1. Hence, we have to be cautious about totally rejecting the hypothesis that *net spillover effects* may also be a good indicator of the latent construct *performance*, captured by the other four constructs. On the other hand, Model 1c shows a non-significant increase in the value of X^2. Furthermore, the restricted model has a significant X^2 at least at the .001 level, which is evidence of very poor goodness of fit. This allows us to reject the hypothesis that a common latent perception of performance exists for both parties.

Interpretation of results

Our results suggest some interesting conclusions. First, the different measures of partnership performance reasonably measure the same construct, with the exception of *net spillover effects*. This result holds both for 3PLs' and for customers' assessments of performance. However, it deviates from those obtained by Ariño (2003). She found strategic goal fulfillment to capture a different construct than the other examined measures. The differences were justified on the basis that the strategic goal fulfillment measure captures outcome performance only, while other measures –overall satisfaction and net spillover effects– represent both process and outcome performance. Conversely, logistics partnerships may be considered a type of long-term relationship where specific customer goals are very well defined a priori in the form of certain service delivery requirements. Hence, performance can be strictly tied to those goals – at least to those of the customer. On the other hand, additional benefits –i.e., spillover effects– are valued but are seldom considered as part of partnership performance. Hence, we suggest that in *asymmetric* business partnerships (e.g., with one partner as customer and the other as provider), performance may indeed be clearly tied to some specific customer goals – which may or may not be included in the contractual agreement. Additionally, achievement of these goals results in a smoother process.

Second, partnership performance assessments differ across partners. The latent constructs that represent each party's performance assessments are correlated, but they do not represent one single performance assessment that is evaluated equally by 3PLs and by customers. This is important, as it indicates that researchers need to ask both parties about their assessments of partnership performance before drawing any conclusions about whether the partnership was successful or not. When this is not possible, researchers usually ask the only available focal partner about the other partner's performance assessment (what we call *inter-party perceptions*). As long as these perceptions about the counterpart's assessment of partnership performance accurately capture the latter's actual assessment, collecting data

from only one side of the partnership will be enough. In the next section, we evaluate the extent to which a focal partner's perception of the other partner's assessment of performance is a reliable measure of the latter's actual assessment

A focal partner's perception of the other partner's performance assessment

We asked each focal partner –both 3PLs and customers– about their perception of the other partner's assessment of the different dimensions of partnership performance. Specifically, respondents from 3PLs were asked about their perception of their customer's assessment of (1) *overall expectations fulfillment*, (2) *global performance satisfaction*, (3) *fulfillment of specific strategic goals*, and (4) *process performance* (see Appendix). Respondents from customer partners were asked about their perception of the 3PL's assessments of the first three variables. For the sake of brevity and given that the survey had a broader purpose, customers were not asked about perceptions of *process performance*. For the same reasons, and because net spillover effects are unlikely to be observed by external parties, neither 3PLs nor customers were asked about perceptions of such effects. Based on these responses, we built similar measures to those presented in the section on measures of performance, and we called them *inter-party perceptions*.

Reliability of inter-party perceptions

Table 5 shows the correlations of 3PLs' (customers') perceptions of the customers' (3PLs') performance assessments and the customers' (3PLs') actual assessments. All of the correlations are high and significant at least at the .01 level. We completed the reliability tests reported earlier (i.e., the reliability coefficient and Cronbach's alpha tests) to assess if the *inter-party perceptions* regarding the different measures of performance could be explained by a single construct of "inter-party perceived performance assessment." The results in Table 6 suggest that each set of *inter-party perceptions* –that is, those of 3PLs and those of customers– is reliable enough and could represent one and the same underlying construct, with the exception of the perception of *strategic goal fulfillment*. We speculate that this is due to the fact that this measure may actually capture private information. It may be difficult for a focal partner to assess the extent to which the other partner's goals have been fulfilled, and hence this perception might not consistently match the other performance assessments. Consequently, we dropped this measure from our analysis.

Table 5. Correlations of inter-party perceptions of performance

Variable	1		2		3		4
(1) Overall expectations fulfillment	.54	.17	.12		.12		n.a.
(2) Global performance satisfaction	.46		.44	.19	.12		n.a.
(3) Strategic goal fulfillment	.53		.50		.35	.28	n.a.
(4) Process performance	.43		.44		.35		.52	n.a.

N=73.
Figures at lower diagonal and below: correlations among Third Party Logistics Providers' (3PLs') perceptions of customers' performance assessment and customers' actual assessments. All correlations are significant at $P<0.01$.
Figures at upper diagonal and above: correlations among Customers' perceptions of 3PLs' performance assessment and 3PLs' actual assessments. None of the correlations is significant, except for $r_{3,3}$ ($P<0.05$).

Table 6. Reliability test for measures of inter-party perceptions of performance

Variable	Third Party Logistics Provider's perception of customer's performance		Customer's perception of Third Party Logistics Provider's performance	
	Reliability coefficient	Alpha if item deleted	Reliability coefficient	Alpha if item deleted
Overall expectations fulfillment	0.86	0.80	0.80	0.63
Global performance satisfaction	0.79	0.83	0.85	0.57
Strategic goal fulfillment	0.60	0.90	0.49	0.88
Process performance	0.75	0.85	--	--
(Alpha)		0.89		0.79

N=73.

Correspondence between inter-party perceptions and actual performance assessments

Our next task was to assess whether a focal partner's perceptions about the other party's assessments of partnership performance are a valid indicator of the latter. We expected those focal partner's perceptions to be strongly influenced by the other partner's actual performance assessments (as revealed in the survey), but also by the focal partner's own performance assessments. Hence, we tested structural measurement Model 2, which captures these expected effects (see Figure 2).

Figure 2. Measurement Model 2 with standardized coefficients
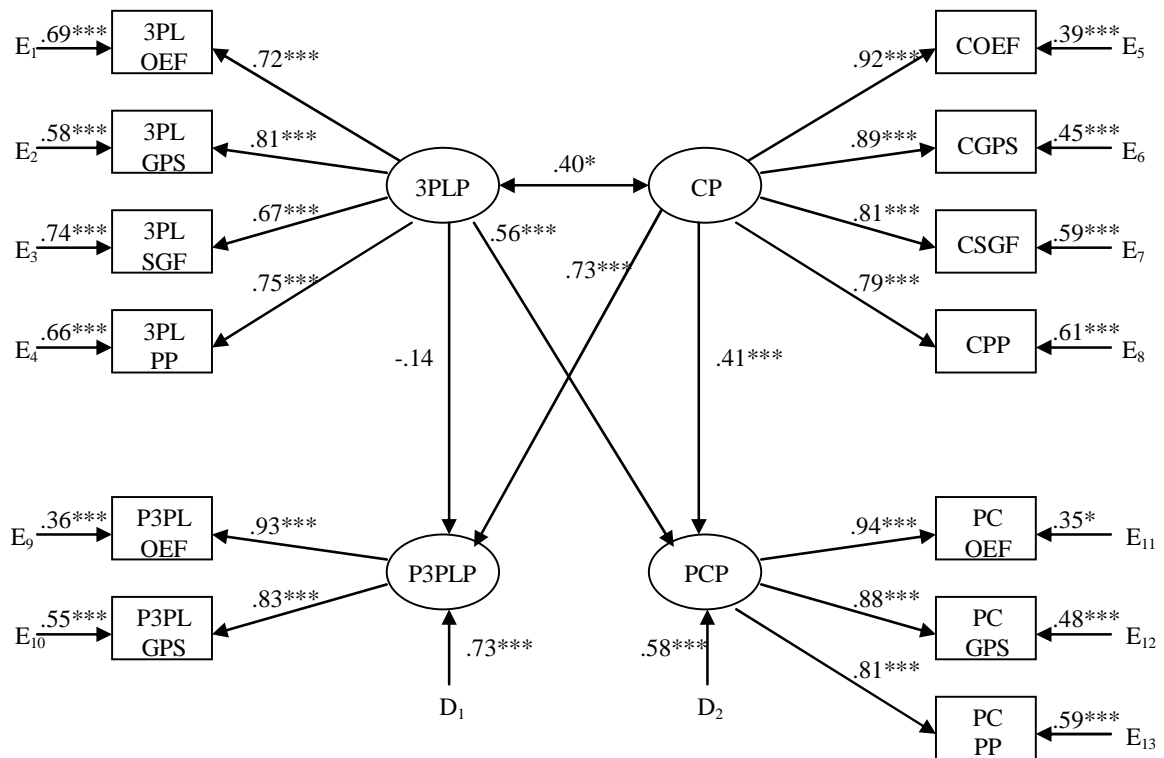
N=73. Model X^2=82.68 (60 d.f.; p<.03); CFI=0.96; NNFI=0.95.

Significance levels obtained from non-standardized solutions:
*** significant at p<.005
** significant at p<.01
* significant at p<.05

3PLP = logistics provider's (3PL's) performance; 3PLOEF = 3PL's overall expectations fulfillment; 3PLGPS = 3PL's global performance satisfaction; 3PLSGF = 3PL's strategic goal fulfillment; 3PLPP = 3PL's process performance.

P3PLOEF = perceived (by the customer) 3PL's overall expectations fulfillment; P3PLGPS = perceived (by the customer) 3PL's global performance satisfaction.

CP = customer's performance; COEF = customer's overall expectations fulfillment; CGPS = customer's global performance satisfaction; CSGF = customer's strategic goal fulfillment; CPP = customer's process performance.

PCOEF = perceived (by the 3PL) customer's overall expectations fulfillment; PCGPS = perceived (by the 3PL) customer's global performance satisfaction; PCPP = perceived (by the 3PL) customer's process performance.

E_i, D_i: error terms.

In other words, we tested the convergent validity of inter-party perceptions relative to the measures of actual performance as assessed by the other partner. To do so, we carried out the two tests put forward earlier:

Test 1: Overall goodness-of-fit. This test yields mixed results: both the values of CFI and NNFI are well above .90, but the X^2 is statistically significant at the .03 level. This last result makes it difficult to reject the null hypothesis of total independence.

Test 2: Significance of the factor loadings and factor correlations. All of the factor loadings and the coefficients that relate inter-party perceptions to actual assessments of performance are statistically significant, with the exception of the one that measures the influence of the 3PLs' actual assessment on the customers' perception of this assessment (see Figure 2). The correlation between the latent constructs capturing partnership performance as assessed by 3PLs and by customers is also significant at the .05 level.

Some important conclusions may be drawn from these results. First, the degree of correspondence between actual performance assessments and the inter-party perception of those assessments is generally low, as shown by the standardized (reliability) coefficient between these two constructs. This is true both for 3PLs and customers. More importantly, customers' perceptions of 3PL performance assessment are not significantly affected by the 3PLs' actual assessments, which evidences that customers tend to ignore or misjudge 3PLs' performance assessments.

Second, 3PLs' perceptions about customers tend to be biased by the 3PLs' own partnership performance assessment. The magnitude of this bias (in terms of standard deviations) is comparable to the influence of the customers' actual performance assessment (.56 vs. .41).

Finally, variations in perceptions may be also affected by other aspects not explicitly considered in Model 2 (note that the variance of the errors associated with the inter-party perceptions constructs, i.e., percentage of non-explained variance, are still high).

Conclusions

This study is limited in that it focuses on just one type of alliance, and the extent to which its conclusions may be generalizable to other alliance types merits further study. Despite this limitation, the study has important implications for research. First, our analysis shows that the examined measures of international logistics partnership performance actually capture one common underlying construct – maybe with the exception of net spillover effects, which may also capture some other construct. This is in contrast to results in Ariño (2003), which showed strategic goal fulfillment to capture a different underlying construct than net spillover effects and overall performance satisfaction. Based on her results, Ariño traced a distinction between process performance and outcome performance. We propose a reliable measure of process performance that is used for the first time, and our results could be read as if that distinction were not significant. However, support for this statement or for the opposite one would require analysis of panel data that allowed us to dig into the interactions between process and outcome performance.

Second, this study shows that participants in a logistics outsourcing partnership assess its performance differently. This finding offers more robust support than the existing literature about the important differences that may exist in the perceptions of providers and customers (Pisharodi and Langley, 1990). This may have to do both with the asymmetric relationship between 3PLs and their customers, and with the subjectivity of each party's assessment of the services, especially if the parties have different strategic goals. This result may be extended to other types of alliances as long as partners differ in their goals for the alliance. That may be the case in other service outsourcing partnerships, such as information systems outsourcing, for instance. Also, our results are more likely to hold in complementary alliances than in pooling alliances (Nohria and García-Pont, 1991), as partner asymmetry is more prominent in the former, while in the latter achieving the benefits of scale tends to be a predominant goal shared by all partners.

Finally, our findings suggest that *inter-party perceptions* –a focal partner's perception of the other partner's performance assessments– are a poor measure of the other partner's actual performance assessment. This result is in contrast to the findings in Geringer and Hebert (1991). One possible explanation has to do with the different nature of the alliances studied: (contractual) international logistics outsourcing partnerships vs. equity joint ventures. The likely closer interaction among partners in the second case might be one justification for the results in Geringer and Hebert (1991). However, our results are in agreement with the mainstream business-to-consumer service literature, which acknowledges a pervasive *service gap,* i.e., the inability of the service provider to adequately measure and understand how the service is perceived by the customer (Zeithaml et al., 1990). The extent to which such a gap exists in business-to-business partnerships over the long term is unknown. Our results suggest, however, that it may be widespread. Nevertheless, from a methodological point of view, our study casts doubt on the validity of surveying one partner only to infer the global performance of a partnership and its assessment by both partners, and more so in the case of purely contractual alliances. Given the difficulties in collecting data from both sides of the dyad, researchers may need to choose whom to survey. Our study suggests that in supplier-customer relationships the surveyed party should be the supplier as –consistent with our results– they are more likely to be aware of their customer's satisfaction with the partnership than the other way around.

References

A.T. Kearney (1995) 'Strategic Partnerships in Logistics: A Key to Competitive Advantage', EPCA Distribution Meeting; 23 October 1995; Monaco.

Ariño, A. (2003) 'Measures of Strategic Alliance Performance: An Analysis of Construct Validity', *Journal of International Business Studies*, **34**(1): 66-79.

Armstrong, J.S. and Overton, T.S. (1977) 'Estimating Nonresponse Bias in Mail Surveys', *Journal of Marketing Research,* **14**: 396-403.

Bagchi, P.K. and Virum, H. (1996) 'European Logistics Alliances: A Management Model', *The International Journal of Logistics Management*, **7**(1): 93-107.

Bagozzi, R.P. and Phillips, L.W. (1982) 'Representing and Testing Organizational Theories: A Holistic Construal', *Administrative Science Quarterly*, **27**: 459-489.

Bentler, P.M. (1992) *EQS: Structural Equations Program Manual*, Los Angeles: BMDP Statistical Software.

Bowersox, D.J. (1990) 'The Strategic Benefits of Logistics Alliances', *Harvard Business Review*, **68**(4): 36-45.

Cooper, M.C. and Gardner J.T. (1993) 'Building Good Business Relationships: More than just Partnering or Strategic Alliances?', *International Journal of Physical Distribution and Logistics Management*, **23**(6): 14-26.

Dillman, D.A. (1978) *Mail and Telephone Surveys: The Total Design Method,* New York: John Wiley.

Frankel, R., Schmitz, J. and Frayer D.J. (1996) 'Formal versus Informal Contracts: Achieving Alliance Success', *International Journal of Physical Distribution & Logistics Management*, **26**(3): 47-63.

Geringer, J.M. and Hebert, L. (1991) 'Measuring the Performance of International Joint Ventures', *Journal of International Business Studies*, **22**(2): 249-264.

Harman, H.H. (1967) *Modern Factor Analysis*. Chicago, IL: University of Chicago Press.

Joreskog, K.G. (1971) 'Simultaneous Factor Analysis in Several Populations', *Psychometrika*, **36**: 409-426.

Laarhoven, P. and Sharman, G. (1996) 'Logistics Alliances: The European Experience', *The McKinsey Quarterly*, (1): 106-115.

Lambert, D.M., Emmelhainz, M.A. and Gardner, J.T. (1999) 'Building Successful Logistics Partnerships', *Journal of Business Logistics*, **20**(1): 165-181.

Lieb, R.C. and Randall H.L. (1996) 'A Comparison of the Use of Third Party Logistics Services by Large American Manufacturers', *Journal of Business Logistics*, **17**(1):305-320.

Murphy, P.R. and Poist R.F. (2000) 'Third Party Logistics: Some User versus Provider Perspectives', *Journal of Business Logistics*, **21**(1): 121-133.

Nohria, N. and Garcia-Pont, C. (1991) 'Global Strategic Linkages and Industry Structure', *Strategic Management Journal*, **12**(Summer special issue): 105-124.

Parkhe, A. (1993) 'The Structuring of Strategic Alliances: A Game-Theoretic and Transaction-Cost Examination of Interfirm Cooperation', *Academy of Management Journal*, **36**(4): 794-829.

Pisharodi, R.M. and Langley, C.J. Jr. (1990) 'A Perceptual Process of Customer Service Based on Cybernetic/Control Theory', *Journal of Business Logistics*, **11**(1): 26-48.

Podsakoff, P.M. and Organ, D.W. (1986) 'Self-Reports in Organizational Research: Problems and Prospects', *Journal of Management*, **12**: 531-544.

Salancik, G.R. and Pfeffer, J. (1977) 'An Examination of Need-Satisfaction Models of Job Attitudes', *Administrative Science Quarterly*, **22**: 427-456.

Sink, H.L. and Langley, C.J. Jr. (1997) 'A Managerial Framework for the Acquisition of Third Party Logistics Services', *Journal of Business Logistics*, **18**(2): 163-167.

Venkatraman, N. and Ramanujam, V. (1986) 'Measurement of Business Performance in Strategy Research: A Comparison of Approaches', *Academy of Management Review*, **11**(4): 801-814.

Yan, A. and 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**(2): 397-414.

Zeithaml, V.A., Parasuraman, A. and Berry, L. (1990) *Delivering Quality Service*, New York: The Free Press.

Appendix

Questionnaire items

Third Party Logistics Providers

1P. Logistics contractual relationships may be set up to fulfill different strategic objectives. When this contract WAS SIGNED, how important for YOUR COMPANY was each of the following strategic objectives? (1=very low, 7=very high, 8=n.a.): business growth; profitability (risk/reward balance); financial stability; efficiencies/economies of scale; market positioning; expertise in new business/sector.

2P. Overall, indicate the level of satisfaction of YOUR COMPANY with the fulfillment of EXPECTATIONS for this alliance? (1=very unsatisfied, 7=very satisfied).

3P. Overall, indicate the level of satisfaction of YOUR COMPANY with the GLOBAL PERFORMANCE of this alliance (1=very unsatisfied, 7=very satisfied).

4P. How much do you feel each of the strategic objectives of YOUR COMPANY are fulfilled with this alliance? (1=very little, 7=very much, 8=n.a.) (same objectives as in question 1P).

5P. Indicate the extent to which each of the following EFFECTS is present in this alliance: (1=very little; 7=very much): leverage infrastructure, people and systems into other operations; increase buying power; develop additional skills and expertise in certain logistics solutions; expand portfolio of logistics services offered to other customers; improve market credibility or become leader in a specific sector or segment.

6P. Indicate the level of satisfaction of YOUR COMPANY with each of the following PROCESSES in the relationship (1=very unsatisfied, 7=very satisfied): contract negotiation; information exchange; conflict resolution; pricing renegotiation; performance and strategy review.

7P. By the time this contract WAS SIGNED, how important for YOUR PARTNER do you think each of the following strategic objectives was? (1=very low, 7=very high, 8=n.a.): (same objectives as in question 1C).

8P. Overall, what do you think the level of satisfaction of YOUR PARTNER with the fulfillment of expectations for this alliance is? (1=very unsatisfied, 7=very satisfied).

9P. Overall, what do you think the level of satisfaction of YOUR PARTNER with the GLOBAL PERFORMANCE of this alliance is? (1=very unsatisfied, 7=very satisfied).

10P. How much do you feel each of the strategic objectives of YOUR PARTNER are fulfilled with this alliance? (1=very little, 7=very much, 8=n.a.) (same objectives as in question 1C).

11P. What do you think the level of satisfaction of YOUR PARTNER with each of the following PROCESSES in the relationship is? (1=very unsatisfied, 7=very satisfied) (same processes as in question 6P).

Appendix (continued)

Customers

1C. Logistics contractual relationships may be set up to fulfill different strategic objectives. When this contract WAS SIGNED, how important for YOUR COMPANY was each of the following strategic objectives? (1=very low, 7=very high, 8=n.a.) cost competitiveness; service and quality excellence; flexibility; risk minimization; outsourcing strategy; supply chain optimization; new market access; gaining expertise.

2C. (Same as question 2P).

3C. (Same as question 3P).

4C. How much do you feel each of the strategic objectives of YOUR COMPANY are fulfilled with this alliance? (1=very little, 7=very much, 8=n.a.) (same objectives as in question 1C).

5C. Indicate the extent to which each of the following EFFECTS is present in this alliance: (1=very little; 7=very much): improve market credibility and logistics leadership; gain expertise and logistics knowledge; increase information and control of the supply chain; expand portfolio of logistics services to our customers; identify additional logistics or outsourcing opportunities.

6C. (Same as question 6P).

7C. By the time this contract WAS SIGNED, how important for YOUR PARTNER do you think each of the following strategic objectives was? (1=very low, 7=very high, 8=n.a.) (same objectives as in question 1P).

8C. (Same as question 8P).

9C. (Same as question 9P).

10C. How much do you feel each of the strategic objectives of YOUR PARTNER are fulfilled with this alliance? (1=very little, 7=very much, 8=n.a.) (same objectives as in question 1P).