12-1993

Conducting interorganizational research using key informants

Nirmalya KUMAR
Singapore Management University, nirmalyak@smu.edu.sg

Louis W. STERN

James C. ANDERSON

Published in Academy of Management Journal, Volume: 36 Issue: 6 Pages: 1633-1651.
http://doi.org/10.2307/256824

CONDUCTING INTERORGANIZATIONAL RESEARCH USING KEY INFORMANTS

NIRMALYA KUMAR
Pennsylvania State University
LOUIS W. STERN
JAMES C. ANDERSON
Northwestern University

In this article, we examine the use of the key informant methodology by researchers investigating interorganizational relationships. Authors have advocated the use of multiple informants to increase the reliability and validity of informant reports. However, interorganizational research still tends to rely on single informants. We investigated informant selection and obtaining perceptual agreement among multiple informants, two problems that may have inhibited widespread use of multiple informants. We suggest procedures for dealing with those problems and provide an illustrative application of our proposals.

It is now widely recognized that, to survive and thrive in competitive environments, firms must seek cooperative relationships with other firms. Correspondingly, an increasing number of empirical studies on interorganizational relationships have appeared in the management and marketing literature over the past decade (e.g., Anderson & Narus, 1990; Cusumano & Takeishi, 1991; Heide & Miner, 1992; Provan & Skinner, 1989; Van de Ven & Walker, 1984). The authors of many of these studies have adopted an interaction perspective on interorganizational relationships, emphasizing how interorganizational trust, conflict, and cooperation emerge within "the context of a specific relationship . . . through ongoing interaction" (Heide & Miner, 1992: 266).

Researchers who want to conduct quantitative, large-scale, empirical investigations of interorganizational relationships must frequently confront a lack of archival data on organization- or relationship-level constructs of interest, such as commitment and power.[1] Thus, they must frequently rely

Grants from the Marketing Science Institute and Pennsylvania State University's Institute for the Study of Business Markets supported this research. We gratefully acknowledge the valuable contributions of Ravi S. Achrol, Hans Baumgartner, Christophe Van den Bulte, Barbara Gray, Brent Johnson, Gary Lilien, Edward Zajac, and the two anonymous reviewers for this journal.

[1] This situation is certainly not unique to interorganizational relationship research; strategy researchers have also noted the unavailability of archival data (e.g., Venkatraman & Ramanujam, 1987).

on reports of key informants. Relying on key informant accounts is appropriate when the content of inquiry is such that complete or in-depth information cannot be expected from representative survey respondents. Respondents describe "their personal feelings, opinions, and behaviors" (Seidler, 1974: 817), but informants generalize "about patterns of behavior, after summarizing either observed (actual) or expected (prescribed) organizational relations" (Seidler, 1974: 817). Researchers do not select informants to be representative of the members of a studied organization in any statistical sense. Rather, they are chosen because they are supposedly knowledgeable about the issues being researched and able and willing to communicate about them.

Like other research techniques, the key informant methodology has some significant drawbacks. Both informant bias and random error can taint informant reports. The former results from differences related to informants' varying organizational roles (Phillips, 1981; Seidler, 1974). For example, the views of CEOs may systematically vary from those of second-level executives because the organizational roles of both influence their interpretations of events (Golden, 1992; Hambrick, 1981). In addition, other more idiosyncratic sources of error may contaminate informant reports, especially retrospective accounts (Golden, 1992; Huber & Power, 1985; Schwenk, 1985). Informant reports suffer from individuals' memory failure, or inaccurate recalling of past events (Golden, 1992), as well as from memory distortion (Nutt, 1986). The latter can result from hindsight bias, attributional bias, subconscious attempts to maintain self-esteem, or impression management (Huber & Power, 1985; Salancik & Meindl, 1984). Thus, there may be little correspondence between informant reports and actual events.

Several researchers have advocated querying multiple informants to increase the reliability and validity of reports (Bagozzi, Yi, & Phillips, 1991; Golden, 1992; Phillips, 1981; Schwenk, 1985; Seidler, 1974). The theoretical and methodological benefits of multiple-informant studies are well documented; it is not our purpose here to restate them. Rather, because most interorganizational relationship research still relies on single informants, we examine two practical problems in using multiple informant methodologies and suggest procedures to help overcome them or, at a minimum, reduce their impact.[2]

The first problem, hereafter called the selection problem, is the challenge of identifying two or more informants competent to report on a particular dyadic relationship. The second, the perceptual agreement problem, is the frequent dissimilarity of the reports of competent multiple informants. Not surprisingly, constructing an organizational response out of divergent informant reports becomes problematic. As noted below, both extant ap-

[2] A longer version of this article that includes a table of the informant procedures used by past empirical investigations of interorganizational relationships is available from the first author.

proaches—aggregating multiple informants' responses and submitting them to structural equation modeling—suffer from significant drawbacks. A consensual approach to reconciling multiple informant reports has been suggested (Glick, Huber, Miller, Doty, & Sutcliffe, 1990) but never applied in interorganizational relationship research. We empirically explored a hybrid consensus-averaging method of constructing organizational responses from multiple informant survey data. We also demonstrated that self-assessments of informant competency are not related to perceptual disagreement between knowledgeable informants.

TWO PROBLEMS IN USING MULTIPLE INFORMANTS

The constructs usually of interest to interorganizational relationship researchers are unobservable, theoretical, and accessible only as shared constructions about what a focal organization is and does; a firm's cooperation with another firm is an example. Thus, researchers must develop organization-level constructs out of multiple informant reports on observed measures (Anderson, 1987; Lincoln & Zeitz, 1980).[3]

The Selection Problem

Response errors are likely to be higher for informants whose roles are not closely associated with the phenomena under study. Key informants are usually chosen on the basis of their formal roles in an organization; frequently, they are companies' owners or general managers. Usually researchers obtain no explicit verification of the competency of the informants but merely assert that the selected person is uniquely qualified to answer questions about the issues under investigation (Cusumano and Takeishi [1991], Heide and John [1990], and Heide and Miner [1992] are exceptions). Not surprisingly, in studies in which informants' claims of competency have been examined, researchers have eliminated some reports because of informants' inadequate knowledge (e.g., Heide & John, 1990; Heide & Miner, 1992).[4]

The issue of informant competence seems to be intimately tied to the decision to forgo using multiple informants. A variety of scholars have explained that their decisions not to use multiple informants were based on a lack of qualified people (e.g., Heide & John, 1990; Provan & Skinner, 1989). Yet relatively few investigators report formally evaluating the unavailability of multiple informants. Given the enormous amount of effort it frequently takes to obtain even a single organizational informant to discuss often delicate matters, like interfirm dependence and conflict, researchers may feel

---

[3] As one reviewer noted, we are assuming the presence of a single underlying organizational trait or construct rather than multiple individual realities.

[4] In contrast, when informant competency is evaluated by examining the mean scores for informants as a whole (e.g., Phillips, 1981) poor informants cannot be discerned and eliminated.

that the work involved in formally assessing competency or ferreting out additional knowledgeable informants is excessive.

Two approaches have been adopted in the past to formally evaluate the competency of informants to report on interorganizational relationships. Some researchers have relied on global measures, or overall evaluations of informant competency. Examples include using the length of an informant's tenure with a firm (e.g., Phillips, 1981), the informant's knowledge of or involvement with the other firm or firms engaged in relationships with the focal firm (e.g., Heide & Miner, 1992), or the length of time the informant has observed or interacted with the other firms (e.g., Phillips, 1982). Because informants are usually asked to report on their firms' relationships with other firms, logically it would seem that the tenure question should focus on how long informants have interacted with the others, not on the length of their current employment.

Alternatively, researchers (e.g., Cusumano & Takeishi, 1991) have used specific measures that assess the level of an informant's knowledge of each major issue included in a study ("How knowledgeable are you about the level of trust in your firm's relationship with firm X," for example). Because specific measures allow assessment of differences in an informant's ability to report on the various issues, they would appear preferable to global measures.

The Perceptual Agreement Problem

When data have been collected from multiple informants in interorganizational relationship research, the reports have often failed to demonstrate high levels of perceptual agreement (e.g., Anderson & Narus, 1990; Molnar & Rogers, 1979). Disagreements may arise because of differences in both knowledge and perceptions (Golden, 1992; Hambrick, 1981; Schwenk, 1985). If the reason for disagreement is some informants' inadequate knowledge about the issues under investigation, the solution is relatively straightforward: the reports of such informants should be ignored. But knowledgeable informants may disagree because they hold different organizational positions and thus different perspectives on the same organizational phenomena. In other words, informant bias taints their reports.

How to combine the discrepant responses of multiple informants into an organizational response is an unresolved issue. Past researchers have applied three approaches. The latent trait approach involves modeling the independent reports of the multiple informants as reflective indicators of a latent construct using structural equation techniques (e.g., Anderson & Narus, 1990; Bagozzi et al., 1991). It allows partitioning the sources of the variance of a measure into trait (construct of interest), method (informant bias and measure specificity), and random error. However, inadequate perceptual agreement among multiple informants frequently creates problems in obtaining acceptable models using the initial set of items, resulting in ill-fitting models (e.g., Anderson & Narus, 1990; Bagozzi et al., 1991; Marsh & Bailey, 1991).

The aggregation approach pools the responses of the multiple infor-

mants to create organization-level indicators, usually via a simple unweighted average (e.g., Chatterjee, Lubatkin, Schweiger, & Weber, 1992; Johnson, Sakano, & Onzo, 1990; Van de Ven & Walker, 1984). However, as James (1982) pointed out, perceptual agreement between informants must be demonstrated before measurements taken from them can be aggregated. Researchers employing this approach have either demonstrated perceptual agreement (e.g., Johnson et al., 1990), discarded observations and measures for which such agreement did not exist (e.g., Chatterjee et al., 1992), or ignored James's recommendation and simply pooled all the multiple informant data.

The consensual approach requires that multiple informants develop a shared position on the items on which they initially disagree (e.g., Eisenhardt & Bourgeois, 1988; Nutt, 1986). Resolving discrepancies among reports can enhance the validity of data (Glick et al., 1990), since objectivity often "results from the heated, intense, and biased confrontation between the somewhat biased ideas of somewhat biased individuals" (Mitroff, 1972: 615). Furthermore, examining the different organization members' views of the same phenomenon and discussing differences may improve a researcher's understanding of the organizational phenomena under investigation (Schwenk, 1985). However, consensual responses may suffer from differences in power among informants and conformity pressures.

Despite its strengths, the consensus method has never, to the best of our knowledge, been employed in interorganizational relationship research, probably because of the operational demands it places on both researchers and informants. To encourage greater use of this approach, some means must be found to reduce the amount of effort required. Therefore, we propose a hybrid approach in which consensual judgments are collected only when there is substantial disagreement between knowledgeable multiple informants on an item. Remaining minor differences can be resolved simply through averaging reports.

METHODS

Research Setting

We assessed informant competency and used a hybrid approach for resolving discrepant reports in a study of reseller performance from a supplier perspective. The data were collected from a major vehicle rental company (the supplier) with a network of over 5,000 dealers (the resellers) in the United States and Canada. The dealers were mostly small, independently owned businesses in which vehicle rental was one of multiple products.

Seven performance facets were assessed: contribution to sales, contribution to profits, reseller competence, reseller compliance, contribution to growth, reseller adaptability, and customer satisfaction. To collect informant data from the focal supplier on each of the facets, we developed items appropriate for a survey questionnaire. On the basis of confirmatory factor analysis, we used three items to measure each facet of reseller performance.

The facet scales demonstrated adequate substantive validity, content validity, unidimensionality, reliability, convergent validity, and discriminant validity. Responses were on seven-point Likert-type scales. Complete details on these procedures, the items used, and the theoretical basis of the study are available elsewhere (Kumar, Stern, & Achrol, 1992). The Appendix contains the definition and a sample item for each facet scale as well as the response format.

Informants and Data

Given the focus of the study, the appropriate informants were those in the supplier's organization who had adequate knowledge of and information on the dealers' performance. Interviews with officers from the supplier's organization led to the identification of two positions, sales manager and fleet manager, whose occupants should theoretically have met the competency requirement. Both the sales and fleet managers were responsible for managing the supplier's relationships with its dealers. We identified 150 pairs of informants, each pair consisting of one sales and one fleet manager. To obtain independence of measurement, we asked each informant to report on only a single dealer; for each pair of informants, we randomly chose a focal dealer from the dealers assigned to them.

Prior to the collection of any substantive data, we mailed an informant competency questionnaire to all identified informants. Two global items assessed (1) the tenure of informants with their firm and (2) how long they had interacted with the dealers on whom they would report. In addition, specific items elicited the level of information and knowledge that the informants had about the focal dealers' performance on each of the seven facets. Five-point Likert-type scales were used for this purpose.

Once we had qualified informants regarding their competency to evaluate specific dealers, we mailed each a questionnaire including the 21 performance assessment items. The informants were instructed to complete this questionnaire individually.

Following the recommended procedure for dealing with the perceptual agreement problem, we employed both a consensual and a statistical approach to resolve discrepancies in informant reports. When the two informants reporting on a dealer differed substantially on any item, we asked them to resolve their differences through discussion. We considered a difference of two points or more on the seven-point scales used in this study to be a substantial difference. We gave the pairs feedback on the items on which there was substantial disagreement after receiving the completed reports. Informants were instructed to contact each other and come up with a consensual rating for each item. This consensual response was then considered the organizational response. For the items on which an informant pair did not substantially disagree, we considered the simple average of their responses to be the organizational response.

All three questionnaires (informant competency, substantive seven-facet, and consensus) were returned from 98 pairs of informants, amounting

to a 65.3-percent response rate. Tests to assess nonresponse bias indicated that although respondents were generally more knowledgeable than nonrespondents, results of only 5 of the 23 tests conducted were significant; thus, nonresponse bias did not appear to be considerable.

RESULTS

Evaluating Informant Competency

Table 1 presents the means for both the sales and the fleet managers on each of three types of indicators of informant competency. Overall, the average sales manager was more experienced than the average fleet manager, had interacted more with the focal dealer, and was more confident of his or her ability to evaluate the dealer on each of the facets. Irrespective of those differences, on the basis of the mean scores for each type of informant alone (cf. Phillips, 1981, 1982), both the average sales manager and the average fleet manager appeared to be adequately qualified.

Because overall mean scores are compensatory, an investigation of individual informant responses to competency questions may indicate that the reports of some informants should be excluded because of insufficient knowledge. We were very conservative in rejecting potential informants, eliminating only those who indicated they lacked adequate information or knowledge (a "1" response) for all seven performance facets listed on the competency questionnaire. Under this rule, two informants were removed from the study, resulting in two pairs of informants being dropped from the analysis.[5]

Table 1 reports the correlations between the three types of indicators used to measure informant competency. Tenure with the firm has significant, though not particularly large, correlations with the number of months an informant had worked with a focal dealer (.33 and .31 for sales and fleet managers, respectively). The correlations between these two global items and the informants' assessments of their ability to evaluate dealer performance on the facets are even lower, ranging from .23 to −.16, with only one of them significant at the .05 level. This pattern suggests that these global items are inadequate indicators of competency because they do not reflect an ability to report on the specific issues of interest to the researcher.

Assessment of Informant Bias

We expected the sales managers and the fleet managers to have different perspectives about dealer performance. To assess whether significant informant bias existed, we estimated four structural equation models: The null model, model 1, represents the hypothesis that the variation in the measures

[5] We explored use of more stringent rules that eliminated more pairs, such as eliminating informants who responded with a "1" to three or more of the seven facets. This criterion would have eliminated an additional five pairs, but the results were not significantly altered.

TABLE 1
Results of the Informant Competency Check

	Sales Manager			Fleet Manager		
	Mean	Correlation with Months Employed	Correlation with Months of Interaction with Reseller	Mean	Correlation with Months Employed	Correlation with Months of Interaction with Reseller
Months employed	64	1.0		44[a]	1.0	
Months of interaction with reseller	24	.33*	1.0	21	.31*	1.0
Information/knowledge adequate to evaluate the reseller's performance on						
Contribution to sales	4.80	−.11	.10	3.85[a]	−.15	−.07
Contribution to profits	4.11	−.02	.23*	3.62[a]	−.13	−.14
Reseller competence	4.52	−.16	.03	4.47	.06	−.10
Reseller compliance	4.48	−.02	.11	4.20[a]	.11	−.03
Contribution to growth	4.15	.08	.15	3.55[a]	.18	.00
Reseller adaptation	4.03	−.04	.16	3.54[a]	.20	.06
Customer satisfaction	4.28	−.06	.10	4.03[a]	.08	.00

[a] Significantly different from mean of sales managers. For example, 44 months is significantly different from 64 months.

* $p < .05$

is explained only by random error and no common factors exist (χ^2_{91} = 777.22, $p < .001$). In the trait-only model (model 2), the variation in the measures is posited to be explained completely by correlated traits plus random error (χ^2_{56} = 189.37, $p < .001$). According to the method-only model (model 3), correlated methods plus random error completely explain the variation (χ^2_{76} = 202.05, $p < .001$). Finally, in the trait-method model (model 4), correlated traits, uncorrelated methods, and random error account for the variation (χ^2_{42} = 49.09, $p = .210$).

The data used to estimate these models were summed facet scales from the individual informant reports. Comparing model 1 to model 3 and model 2 to model 4 provides a test of method variance (Bagozzi et al., 1991). The results of the chi-square difference test indicate that the introduction of method variables to both the null model (model 3 versus model 1: χ^2_{15} = 575.17) and the trait-only model (model 4 versus model 2: χ^2_{14} = 140.28) produced significant ($p < .001$) improvements. Thus, significant informant bias does exist.[6]

To gain further understanding of perceptual agreement, following Jones, Johnson, Butler, and Main (1983), we conducted two alternate assessments of the perceptual agreement between the two informant positions. First, we examined the absolute deviation between the two informants' scores on each facet (i.e., $|sales\ manager_{facet\ i} - fleet\ manager_{facet\ i}|$). This procedure produced seven mean absolute deviations, one for each facet. We tested each of these mean absolute deviations against the null hypothesis that there was no difference between the reports of the two informants on a particular facet and rejected all seven null hypotheses; *t*-values ranged from 11.08 to 13.66 ($p < .001$). The correlation between the reports of the two informant positions on each of the facets provided the second assessment of perceptual agreement. As can be seen from Table 2 (column 1), although these correlations are significant for six of the seven facets, they also are considerably lower than 1. Thus, the results from these alternate assessments indicate that there are significant differences between the reports of the two informant positions on each of the seven facets.

Informant Competency and Perceptual Agreement

Phillips (1981) argued that because the informants in his study were adequately qualified, failure to observe perceptual agreement was not a knowledge deficiency artifact. We decided to examine whether the extent of perceptual agreement between the reports of the two informants in a pair was related to self-assessments of competency. The measure of perceptual agreement we used was the absolute deviation for a pair of informants summed across all 21 performance items for that pair. To explore the relationship of this measure with informant competency, we classified the sales

[6] Furthermore, in model 4, the average variance levels accounted for by trait factors, informant bias, and random error were 38, 22, and 40 percent, respectively.

TABLE 2
Convergence Between Reports[a]

Variables	Sales and Fleet Managers	Fleet Managers and Consensus-Averaging	Sales Managers and Consensus-Averaging			Averaging and Consensus-Averaging
			r	*Z*	*p*	
Contribution to sales	.62	.77	.93	5.31	.001	.95
Contribution to profits	.19[b]	.55	.75	2.46	.01	.83
Reseller competence	.29	.65	.81	2.56	.01	.91
Reseller compliance	.31	.62	.82	3.14	.001	.90
Contribution to growth	.47	.66	.90	5.26	.001	.91
Reseller adaptation	.41	.68	.88	4.20	.001	.93
Customer satisfaction	.40	.76	.83	1.49	.068	.96

[a] Statistics are correlations. In addition, we computed *Z*- and *p*-values (one-tailed test) to test whether the correlation between the sales managers' reports and the consensus-averaging process differed significantly from the correlation between the fleet managers' reports and the consensus-averaging process on each of the facets. Thus, for example, .93 is significantly greater than .77 for the contribution-to-sales facet. We used the formula to test the significance of difference between dependent *r*'s (Meng, Rosenthal, & Rubin, 1992).

[b] This correlation is not significant at the .05 level. All other correlations are significant at the .01 level.

managers into two groups, highly competent and less competent to report on the reseller, using each variable measuring competency—months employed by the supplier, months the informant had known the dealer, and the specific competence measures—independently to categorize the sales managers into the two groups. For the first two measures, we used the median to split the sales managers into high- and low-competence groups. On the specific measure, sales managers were classified as highly competent only if they had responded with a "4" or "5" to each of the competency questions regarding the seven facets. Similar procedures were employed to split the fleet managers into highly competent and less competent sets.

Next, we placed the pair of informants reporting on a particular dealer into one of three categories—high-, mixed-, and low-competence. If both managers reporting on a dealer were highly (or less) competent, that pair of informants was placed in the high (or low) category. If one member of a pair was highly competent and the other was less competent, that pair of informants was placed in the mixed category. We performed this process for each of the three competency variables.

We employed Helmert contrasts (Finn, 1974) to explore whether more

competent pairs of informants demonstrated greater perceptual agreement, first examining differences in perceptual agreement between pairs of informants placed in mixed and low categories. As Table 3 shows, there were no significant differences on each of the three measures, and thus it was acceptable to pool the informant pairs in the mixed and low categories and contrast them with those in the high category. However, after this procedure had been completed, none of the contrasts were significant. Thus, differences in self-assessments of knowledge—informant competency—do not explain the obtained disagreements between the paired informants.

Individual and Consensual Informant Reports

A relatively unexplored issue in the literature is the degree to which such consensus ratings will be more aligned with one informant's report than with another's. On the basis of logic alone, it would seem that (1) knowledge about the issues under investigation, (2) place in the hierarchy of an organization, and (3) tenure in the organization would serve as strong predictors. Because the sales managers studied had greater expert (Table 1) and legitimate power than the fleet managers (the former were higher in the hierarchy), we expected that the organizational reports constructed using our recommended hybrid procedure would be closer to the sales managers' independent reports.

To test whether the ratings derived through our hybrid averaging-consensus procedure were more aligned with the sales managers' than the fleet managers' original ratings, we computed the correlations between the original ratings of both types of manager and our hybrid procedure. The middle four columns of Table 2 report the results. The correlations between the consensus responses and the individual informant reports are high and significant for all facets. Furthermore, as expected, the correlations between the sales managers' responses and the consensus responses (.93–.75) are higher than the correlations between the fleet managers' responses and the consensus responses (.77–.55) for each facet. As Table 2 indicates, these

TABLE 3
Effects on Informant Competence on Perceptual Agreement[a]

Competence Categories	Months Employed	Months of Interaction with Reseller	Conjunctive Competence Measure
High	25.35	23.08	23.35
Mixed	23.76	25.03	24.82
Low	24.09	24.03	24.08
Low versus mixed	0.03	0.29	0.17
High versus average of low and mixed	0.54	0.66	0.34

[a] In the top three rows, larger values indicate lower perceptual agreement. Statistics in the last two rows are *F*s. All *F*s were insignificant.

differences between the correlations are significant ($p < .05$) for all facets except customer satisfaction.

It was critical to determine whether the hybrid procedure we adopted yields results different from other, less cumbersome and time-consuming ways of constructing organization-level indicators. We considered two frequently used alternatives: (1) using single informants and (2) merely averaging all multiple informant reports, an aggregation approach. If we had applied a single-informant method in this study, we would have considered only reports from the sales managers, given their higher position and greater expertise.

To examine the relationship between the sales managers' reports and the organizational reports, we constructed a measure of the agreement between the sales managers' reports on a particular facet and those obtained using our averaging-consensus approach for each facet. This measure was the absolute deviation between the sales manager facet score and the hybrid facet score ($|sales\ manager_{facet\ i} - hybrid_{facet\ i}|$). This procedure resulted in seven mean facet deviations, one for each facet. We tested each deviation against the null hypothesis that there was no difference between the two methods on a particular facet. We rejected all seven null hypotheses on the basis of t-values ranging from 8.49 to 10.74 ($p < .001$). As an additional analysis, we computed the correlations between the sales managers' reports and the hybrid consensus-averaging reports on each of the facets (Table 2, column 3). Although the resulting correlations are large, all of them are still significantly different from a correlation of 1 ($t < .05$). Thus, both alternate analyses indicated significant differences between the reports of sales managers as single informants and organizational indicators constructed using our hybrid procedure.

To examine the relationship between using the hybrid averaging-consensus procedure and merely averaging all the reports of both informants (an aggregation approach), we constructed a measure of the agreement between the facet scores obtained using our hybrid approach and aggregation for each facet. This measure was the absolute deviation between the consensus-averaging score on a particular facet and simply averaging the two informants' facet scores ($|average_{facet\ i} - hybrid_{facet\ i}|$). This procedure resulted in seven mean facet deviations, one for each facet. We tested each of these scores against the null hypothesis that there was no difference between these two methods on a particular facet. We rejected all seven null hypotheses on the basis of t-values ranging from 5.60 to 9.21 ($p < .001$). Once again, we computed the correlations between the facet scores resulting from our hybrid procedure and the aggregation approach for an additional analysis (Table 2, last column). Although the resulting correlations are large, they are all significantly different from 1. Furthermore, one reason the correlations are so large is that by restricting collection of consensual responses to only those items on which there was a substantial difference between a pair of informants, on average only 15 percent of the responses had to be referred for

consensus.[7] Overall, the results from these alternate analyses suggest that there are significant differences between organizational indicators constructed using our suggested hybrid approach and those resulting from a strict aggregation approach.

CONCLUSIONS

Our research provided four primary results. First, a variety of measures of informant competency were found to not converge with each other. Second, significant informant bias was observed between informants. Third, this lack of perceptual agreement between informants was not a function of informants' self-assessments of their knowledge. And fourth, the organizational reports generated using our hybrid consensus-averaging approach were found to significantly differ from organizational reports constructed using either a single-informant or an averaging approach.

Thus, researchers desiring to conduct survey research on interorganizational relationships must address a number of critical issues regarding the appropriate use of key informants. Some of these questions concern assessing informant competency and using multiple informant reports to construct organization-level indicators. Given our results, we offer the following suggestions.

Assessing Informant Competency

Advocating tests for informant competency is not likely to meet with much resistance or argument, yet it is surprising how seldom researchers formally conduct such tests, however basic. We employed both global and specific measures to evaluate informant competency. In general, the correlations between these measures were low. However, each type of scale has advantages and disadvantages. The specific scale allowed for more precise measurement by evaluating each informant's ability on each major issue of interest. This process permits elimination of informants on some elements but not on others, if such a distinction is desired. When one informant does not dominate other informants on all the issues of interest (e.g., Walker & Weber, 1987), a specific scale may be more useful than a global scale.

Global items are easier to administer and less reactive, that is, they have less chance of conditioning later responses of informants than specific competency questions; further, the needed data may be available from archival sources. A researcher who is operating under severe constraints on questionnaire length or is concerned about reactivity may favor global measures. If a global measure is to be used, the length of time an informant has interacted with the other firm of interest and overall knowledge items are pref-

[7] On average, across all informants the responses of the two informants were identical on 31 percent of the items, and they differed by only one point on 54 percent of the items.

erable to tenure with the focal organization. Furthermore, we strongly recommend the use of multiple global items to increase reliability.

Regardless of the measures used to assess informant competency, researchers should scrutinize individual informant reports to exclude informants who are not adequately qualified to report on the issues under investigation. Moreover, in the interests of full information, the rule used to eliminate informant reports should be reported, as well as the number of informants whose reports were excluded from the analysis.

Using Multiple Informant Reports

As have previous researchers, we observed considerable disagreement between the reports of the multiple informants we studied. However, we found that variance in self-reported knowledgeability was not related to differences in informants' perceptions. This finding makes us reasonably confident that the perceptual differences were due to significant informant bias. In other words, just because two informants are knowledgeable does not necessarily mean they will completely agree with each other.

Since knowledgeable multiple informants often disagree, researchers need to adopt a procedure for combining their reports into an organizational concept. Three methods—aggregation, consensus, and latent trait modeling—have been employed in the past. In contrast, our suggested hybrid approach potentially uses all three methods while overcoming some of the problems that researchers encounter when using any of them in isolation. Our approach compares multiple informant reports on individual items, pooling them when there is acceptable agreement and requiring consensus on the remaining items.[8] By requiring consensus only when differences were substantial, on average we only had to obtain agreement on 15 percent of the items for each pair of informants. On the other hand, a consensual process raises several questions having to do with differences in hierarchical power among informants, coalition formation, conformity pressures, and groupthink (Janis, 1972), all of which may lead the responses from such a procedure to differ significantly from independent ratings (Schwab & Heneman, 1986). The degree to which those issues affect consensual responses is a question for future research. However, we advise researchers with strong concern about such influences to use the informant competency measures

[8] The impact of such a procedure on measurement error should be noted. Although averaging reduces random error, systematic error (or informant bias) is only diminished if the method factors are uncorrelated (Rousseau, 1985). We expect this pattern to hold true for consensual judgments as well. Furthermore, when data are collected from only two informants, as is predominantly the case, distinguishing correlated informant bias from perceptual agreement is difficult (Anderson, 1987). Thus, if we assume that informant bias factors are uncorrelated, as others have (Anderson, 1987; Anderson & Narus, 1990; Marsh & Bailey, 1991), both systematic and random error will be reduced with the averaging-consensus approach. However, one reviewer suggested that allowing discrepancies to be resolved in the fashion proposed does not enhance the validity of the data since various sources of error and biases may creep into the measurement process.

drawn from a specific scale to determine which informant's responses be included in their analysis.

In proposing our hybrid approach, we do not mean to suggest that researchers abandon the multitrait-multimethod (MTMM) approaches to model informant bias Bagozzi and colleagues (1991) outlined. Rather, at a cost of a modest amount of additional data collection, our recommended procedure provides insurance against some of the problems with estimating these complicated models on empirical interorganizational relationship data. Furthermore, these estimation difficulties may reflect more general problems with MTMM models (Marsh & Bailey, 1991). It is also important to note that the organizational responses resulting from our hybrid procedure can subsequently be used to estimate structural equation models for the assessment of convergent, discriminant, and nomological validity (see Kumar, Stern, and Achrol, [1992] for such an application). Since method factors are obviated, the occurrence of improper solutions would be greatly reduced. Our hybrid approach incorporates all three previously used methods to varying degrees for the purposes of constructing organizational concepts from informant reports.

Final Comments

Some limitations to this study should be mentioned. First, we have provided an illustrative application of our proposed suggestions, and thus, questions of generalizability remain. The nearly ideal conditions surrounding our application, including independence of observation, standardization, and lack of cultural barriers between informants, made obtaining perceptual agreement and consensual responses much easier than it might be under other conditions. In contexts with greater noise, the correlations between the reports of different informants are likely to be lower and informant bias higher. Second, since the informants who were higher in the organization's hierarchy (the sales managers) were also more competent to report the information of interest here, we were unable to separate the impacts of hierarchy and competence on the direction of the consensual response. Third, more stringent criteria for excluding informants should probably be used, especially if a larger pool of potential informants is available. However, an exploration of our data did not yield significant differences based on the exclusion rule used. Perhaps the care we took in selecting informants at the outset ensured that all our informants were reasonably knowledgeable.

However, even with these limitations in mind, it is important that interorganizational researchers consider our suggestions, because most will be compelled to rely on informant reports. This situation is in part a result of the limitations of archival data for measuring the constructs of interest to organization theorists. Besides the problem of their availability, archival data are frequently deficient and do not converge with informant data (Pennings, 1973). Usually, only a single indicator is obtained for a construct. If, as in this study, the facets are conceptualized as abstract concepts, a single item will not be an adequate sampling of its domain (have adequate content

validity). Performance facets that at first glance appear to be well represented by a single item may on further reflection be closer to abstract constructs. And even when archival data are available, they usually have to be supplemented with informant reports to aid interpretation (Golden, 1992; Schwenk, 1985).

Researchers are still grappling with issues related to the appropriate use of informants in both organizational and interorganizational research. Some of our conclusions with respect to assessment of informant competency and constructing organization-level responses out of multiple informant reports are also relevant for organizational research. For example, researchers examining relationships between divisions or business units within the same firm should use multiple informants and may wish to apply our method to the selection and perceptual agreement problems that they will face. Failure to account for informant bias in both organizational and interorganizational research may lower the degree of correspondence between informant reports and the organizational concepts they are intended to represent, thereby jeopardizing the validity of any substantive findings.

REFERENCES

Anderson, J. C. 1987. An approach for confirmatory measurement and structural equation modeling of organizational properties. ***Management Science,*** 33: 525–541.

Anderson, J. C., & Narus, J. A. 1990. A model of distributor firm and manufacturer firm working relationships. ***Journal of Marketing,*** 54: 42–58.

Bagozzi, R. P., Yi, Y., & Phillips, L. W. 1991. Assessing construct validity in organizational research. ***Administrative Science Quarterly,*** 36: 421–458.

Chatterjee, S., Lubatkin, M. H., Schweiger, D. M., & Weber, Y. 1992. Cultural differences and shareholder value in related mergers: Linking equity and human capital. ***Strategic Management Journal,*** 13: 319–334.

Cusumano, M. A., & Takeishi, A. 1991. Supplier relations and management: A survey of Japanese, Japanese-transplant, and U.S. auto plants. ***Strategic Management Journal,*** 12: 563–588.

Eisenhardt, K. M., & Bourgeois, L. J. III. 1988. Politics of strategic decision making in high-velocity environments: Toward a midrange theory. ***Academy of Management Journal,*** 31: 737–770.

Finn, J. D. 1974. ***A general model for multivariate analysis.*** New York: Holt, Rinehart & Winston.

Glick, W. H., Huber, G. P., Miller, C. C., Doty, D. H., & Sutcliffe, K. M. 1990. Studying changes in organizational design and effectiveness: Retrospective event histories and periodic assessments. ***Organization Science,*** 1: 293–312.

Golden, B. R. 1992. The past is the past—or is it? The use of retrospective accounts as indicators of past strategy. ***Academy of Management Journal,*** 35: 848–860.

Hambrick, D. C. 1981. Strategic awareness within top management teams. ***Strategic Management Journal,*** 2: 263–279.

Heide, J. B., & John, G. 1990. Alliances in industrial purchasing: The determinants of joint action in buyer-seller relationships. ***Journal of Marketing Research,*** 27: 24–36.

Heide, J. B., & Miner, A. S. 1992. The shadow of the future: Effects of anticipated interaction and frequency of contact on buyer-seller cooperation. ***Academy of Management Journal,*** 35: 265–291.

Huber, G. P., & Power, D. J. 1985. Retrospective reports of strategic-level managers: Guidelines for increasing their accuracy. ***Strategic Management Journal,*** 6: 171–180.

James, L. R. 1982. Aggregation bias in estimates of perceptual agreement. ***Journal of Applied Psychology,*** 67: 219–229.

Janis, I. L. 1972. ***Victims of groupthink.*** Boston: Houghton-Mifflin.

Johnson, J. L., Sakano, T., & Onzo, N. 1990. Behavioral relations in across-culture distribution systems: Influence, control, and conflict in U.S.-Japanese marketing channels. ***Journal of International Business Studies,*** 21: 639–655.

Jones, A. P., Johnson, L. A., Butler, M. C., & Main, D. S. 1983. Apples and oranges: An empirical comparison of commonly used indices of interrater agreement. ***Academy of Management Journal,*** 26: 506–519.

Kumar, N., Stern, L. W., & Achrol, R. S. 1992. Assessing reseller performance from the perspective of the supplier. ***Journal of Marketing Research,*** 29: 238–253.

Lincoln, J., & Zeitz, G. 1980. Organizational properties from aggregate data: Separating individual and structural effects. ***American Sociological Review,*** 45: 391–408.

Marsh, H. W., & Bailey, M. 1991. Confirmatory factor analyses of multitrait-multimethod data: A comparison of alternative models. ***Applied Psychological Measurement,*** 15: 47–70.

Meng, X., Rosenthal, R., & Rubin, D. 1992. Comparing correlated correlation coefficients. ***Psychological Bulletin,*** 111: 172–175.

Mitroff, I. I. 1972. The myth of objectivity or why science needs a new psychology of science. ***Management Science,*** 18: B613–618.

Molnar, J. J., & Rogers, D. L. 1979. A comparative model of interorganizational conflict. ***Administrative Science Quarterly,*** 24: 405–425.

Nutt, P. C. 1986. The tactics of implementation. ***Academy of Management Journal,*** 29: 230–261.

Pennings, J. 1973. Measures of organizational structure: A methodological note. ***American Journal of Sociology,*** 79: 686–704.

Phillips, L. W. 1981. Assessing measurement error in key informant reports: A methodological note on organizational analysis in marketing. ***Journal of Marketing Research,*** 18: 395–415.

Phillips, L. W. 1982. Explaining control losses in corporate marketing channels: An organizational analysis. ***Journal of Marketing Research,*** 19: 525–549.

Provan, K. G., & Skinner, S. J. 1989. Interorganizational dependence and control as predictors of opportunism in dealer-supplier relations. ***Academy of Management Journal,*** 32: 202–212.

Rousseau, D. M. 1985. Issues of level in organizational research: Multi-level and cross-level perspectives. In B. M. Staw & L. L. Cummings (Eds.), ***Research in organizational behavior,*** vol. 7: 1–37. Greenwich, CT: JAI Press.

Salancik, G., & Meindl, J. R. 1984. Corporate attributions as strategic illusions of management control. ***Administrative Science Quarterly,*** 29: 238–254.

Schwab, D. P., & Heneman, H. G., III. 1986. Assessment of a consensus-based multiple information source job evaluation system. ***Journal of Applied Psychology,*** 71: 354–356.

Schwenk, C. R. 1985. The use of participant recollection in the modeling of organizational decision processes. ***Academy of Management Review,*** 10: 496–503.

Seidler, J. 1974. On using informants: A technique for collecting quantitative data and controlling for measurement error in organizational analysis. ***American Sociological Review,*** 39: 816–831.

Van de Ven, A. H., & Walker, G. 1984. The dynamics of interorganizational cooperation. ***Administrative Science Quarterly,*** 29: 598–621.

Venkatraman, N., & Ramanujam, V. 1987. Measurement of business economic performance: An examination of method convergence. ***Journal of Management,*** 13: 109–122.

Walker, G., & Weber, D. 1987. Supplier competition, uncertainty, and make-or-buy decisions. ***Academy of Management Journal,*** 30: 589–596.

APPENDIX
Measures

Responses for all items were on the following Likert-type scale: 1, strongly disagree; 2, disagree; 3, mildly disagree; 4, neither agree nor disagree; 5, mildly agree; 6, agree; 7, strongly agree. Reliability values are for the sales managers', fleet managers', and organizational responses, respectively. Items appear verbatim from the questionnaire; however, "the supplier" replaces the name of the specific supplier. Definitions did not appear in questionnaire.

Contribution to Sales (αs = .81, .76, .82)

Definition: It includes the revenues that the reseller has generated on behalf of the supplier relative to the market potential as well as the penetration that the reseller has achieved for the supplier in his market area.

"Over the past year, the dealer has been successful in generating high rental revenues for [the supplier], given the level of competition and economic growth in his market area."

Contribution to Profits (αs = .61, .57, .68)

Definition: The level of financial returns (profits, cash, etc.) that the supplier realizes through its association with the reseller compared to the supplier's effort and investment in the reseller.

"[The supplier] made inadequate profits from this dealer over the past year because of the amount of time, effort, and energy which [the supplier] had to devote to assisting him."

Reseller Competence (αs = .69, .70, .77)

Definition: The value of the human resources of the reseller's organization to the supplier. It encompasses the reseller's experience and knowledge of the suppliers and its competitors offerings.

"The dealer has amassed a great deal of knowledge about the features and attributes of [the supplier's] products and services."

Reseller Compliance (αs = .67, .74, .72)

Definition: The reseller's participation in the supplier's programs and its compliance with the supplier's policies.

"The dealer has frequently violated stipulations contained in his contract with [the supplier]."

Reseller Adaptation (αs = .78, .71, .79)

Definition: The reseller's ability to accept and initiate new ideas on behalf of the supplier as well as market the supplier's products in an innovative manner.

"The dealer is very innovative in his marketing of [the supplier's] products and services in his neighborhood."

Contribution to Growth (αs = .62, .73, .69)

Definition: The increases in revenues that the supplier has generated in the past and can expect to generate in the future through the reseller.

"In the past [the supplier's] business with the dealer has grown steadily."

Customer Satisfaction (αs = .84, .81, .82)

Definition: The level and quality of the services that the reseller provides actual and potential end-users of the supplier's products and services.

"The dealer provides customers with good assistance in the solution of any problems involving [the supplier's] products and services."

Nirmalya Kumar is an assistant professor of marketing at Smeal College of Business Administration, Pennsylvania State University. He received his Ph.D. degree in marketing from Northwestern University. His current research interests include interorganizational relationships in marketing channels, joint ventures, and marketing strategy.

Louis W. Stern is the John D. Gray Distinguished Professor of marketing at Northwestern University's J. L. Kellogg Graduate School of Management. He received his Ph.D. degree in marketing from Northwestern University. His research interests focus on the design and management of marketing channels and on antitrust issues in marketing.

James C. Anderson is the William L. Ford Distinguished Professor of Marketing and Wholesale Distribution and a professor of behavioral science in management at the J. L. Kellogg Graduate School of Management, Northwestern University. He received his Ph.D. degree in psychology from Michigan State University. His research interests are working relationships between firms in business markets, and measurement techniques.