The authors report findings bearing on the reliability of measures used in industrial marketing research surveys to identify the structure of buying groups. Results obtained in a pilot study of the purchase of lithographic plates by small printing firms reveal a lack of consensus about purchase influence between pairs of informants from the same organizations. Further, the ratings appear to differentiate among roles but not stages in the decision process.

Measuring Influence in Organizational Purchase Decisions

Received conceptualizations of organizational buying behavior routinely emphasize joint decision making and the attendant notion of a "buying center" as being the relevant unit of analysis. However, this orientation is often not reflected in empirical research in this field, which continues to focus on a single individual within an organization rather than on a buying collectivity or group (Spekman and Calder 1978; Wind 1978). The persistence of several formidable methodological issues is seen as the principal barrier to further progress. Among the most basic problems to be resolved is that of establishing a workable method for delineating the composition of buying centers. As Bonoma, Zaltman, and Johnston (1977) recently observed, "The complex, vague, and often changing composition of the buying center makes it difficult to ascertain empirically just who is involved in organizational buying" (p. 80).

Examining studies concerned with identifying what Webster and Wind (1972) refer to as the "locus of buying responsibility," one finds that reliance on a single "key informant" is the most common procedure used to obtain such data. Typically a purchasing agent or someone else believed on a a priori grounds to be a major participant is sought out and asked to report which members of the organization take part in the purchasing decision under study, what functions they perform, and/or what degree of involvement/influence these individuals have in different stages of the process (Brand 1972; Buckner 1967; Platten 1950). Despite widespread application of this approach, a review of the literature reveals that the reliability and related properties of measurements obtained in this manner have yet to be established. We seek to add to the limited stock of knowledge by reporting some findings on the degree of consensus and discrimination exhibited by measures of buying process involvement obtained from informants occupying different positions within the same firm.

We first examine the research in industrial marketing on the psychometric properties of measures used in surveys to identify the composition and structure of organizational buying groups. Attention is focused on findings bearing on the reliability of reports of buying involvement as reflected by the level of intraorganizational agreement that has been observed for such measures. Design and data collection instrument issues are noted and briefly discussed. Next, we describe a study of the purchasing of lithographic plates by small commercial printing firms. Some unresolved questions iden-

---

1. The key informant method is used in other areas of marketing research such as in the study of strategic business planning and vertical marketing systems. See Phillips (1980) for a discussion of this work. In consumer research, husbands and wives are used as family informants (Davis 1971; Davis, Douglas, and Silk 1981).
tified in the literature review pertaining to the intraorganization consensus and discriminant ability of influence assessments are investigated. We then report the extent of between-informant consensus found for these measures as well as a multitrait-multimethod analysis of their ability to discriminate among different roles and stages in the decision process. Finally, our findings lead us to question the prospects of obtaining reliable and valid assessments of influence through the simple types of rating questions upon which much industrial marketing research relies.

RESEARCH ON CONSENSUS

The body of empirical evidence on the quality of measures of involvement in organizational purchasing is not very extensive. However, an examination of this work suggests this inattention is not due to the absence of unresolved issues. The few relevant studies uncovered in a review of the industrial marketing literature are listed in Table 1.

The basic type of reliability information reported in each of these studies is an assessment of the degree of consensus or agreement observed among reports of purchase involvement obtained from two or more participants within the same organization or firm. Investigations of convergence in judgments of the same phenomenon made by different respondents are sometimes referred to as assessments of "validity" rather than "reliability" (e.g., Kane and Lawler 1978). However, we use the latter term in preference to the former to emphasize the distinction made by Campbell and Fiske (1959, p. 83), who defined reliability as "the agreement between two efforts to measure the same trait through maximally similar methods" and validity as "the agreement between two attempts to measure the same trait through maximally different methods" (emphasis added).

Consensus was assessed in a variety of ways in the studies, some formal and some informal, and thus we summarize each result as indicating either "high" or "low" consensus, in accordance with the judgment of the original authors in reporting their findings. Table 1 indicates a conflicting set of results: whereas a high level of consensus about participation was observed in four of the six relevant investigations, consensus on influence was found to be low in all four of the studies addressing this construct. Given these disparate conclusions, we scrutinized the studies for possible sources of unreliability that might account for the differences in outcomes. Materials offering valuable insights into this class of measurement problems are to be found in two traditions of organization research: the informant technique used by sociologists (Seidler 1974) and the method of peer assessment employed by psychologists (Kane and Lawler 1978). We first discuss the methods employed in the studies summarized and classified in Table 1, then attempt to order and interpret the findings by proposing three major conclusions which appear consistent with the overall patterns of results.

Methodological Considerations

Types of decisions and organizations investigated. As indicated by the entries in the second column, the studies constitute a broad sampling of purchase decisions. Note that in most of the investigations purchase decisions for the same specific product were examined in a cross-section of firms, whereas McMillan (1973), Patchen (1974), and Spekman (1977) studied a cross-section of products, generally focusing on a different product in each organization. Information is not always reported in sufficient detail to permit the purchase decisions to be categorized according to a typology of buying situations such as that proposed by Robinson, Faris, and Wind (1967)—straight and modified rebuys and new tasks—but a reasonable conjecture is that the latter two types predominate. Considerable heterogeneity characterized each of the samples of firms studied in terms of basic industries covered. Large-scale organizations with established purchasing units and procedures appear to have been the typical setting for most of the work, but meaningful summaries of other organizational variables could not be extracted from the source materials.

Informant selection. Data were collected by means of personal interviews in all studies except those of McMillan (1973) and Choffray (1977), who used mail questionnaires. Spekman (1977) used a combination of personal interviews and self-administered questionnaires. An important element of these studies not covered in Table 1 is the manner in which respondents were selected. The basic approach in all of the investigations was a sociometric one wherein a key informant was contracted initially within each organization and asked to identify others "involved" in the decision, some or all of whom became the target respondent population for subsequent measurement efforts. However, the studies differ considerably in terms of the number of decision participants contacted in the follow-up stage and the amount of control over their selection. Such differences are partially reflected in the variability within and across studies of the "number of informants per buying group" shown in Table 1. For example, Patchen (1974) instructed interviewers to contact all persons mentioned by any respondent as having had "any part" in the purchase. He was able to obtain completed interviews with an average of 5.5 respondents per decision where the mean number of individuals mentioned as involved in any way was 15. Restrictions placed on interviewers' access to relevant personnel by the cooperating firms was cited as the major constraint on achieving complete coverage of the buying group membership. In contrast, McMillan (1973) mailed questionnaires to purchasing agents, asking them to select the "scientist" and "manager" they considered to be "most associated" with purchase decisions for the particular product under study. Of the maximum possible response of three respondents per firm (i.e., pur-
Table 1
SUMMARY OF STUDIES OF CONSENSUS IN PARTICIPATION AND INFLUENCE MEASUREMENT

| Reference          | Purchase decisions investigated | Sample size | No. of informants per buying group | No. of roles rated per informant | Measure | Participation | Level of consensus | Influence | Measure | Response scale | Level of consensus |
|--------------------|---------------------------------|-------------|-----------------------------------|---------------------------------|---------|---------------|-------------------|-----------|---------|-----------------|-------------------|
| Weigand (1966)     | Inexpensive fabricated part     | 55          | 2                                 | 1                               | Specific (6 product elements)   | 4-point scale | Low (self-informant) | n.a.     | —       | —               | —                 |
| McMillan (1973)    | 3 chemical products             | 75          | Varied (mean = 2)                 | 3                               | n.a.    | n.a.          | n.a.              | Global   | 5-point scale | Low (self-informant) | —                 |
| Grashoff & Thomas  | Scientific & technical info.     | 171         | Varied (mean = 1–2)               | 3                               | n.a.    | n.a.          | n.a.              | Specific (11 stages) | Rank order several positions | Low (self-informant) |
| Kelly (1974)       | Offset press                    | 18          | Varied (mean = 2–3)               | Varied                          | Specific (5 stages)             | Open-ended | High (between-informant) | Global   | Rank order decision participants | ?                 |
| Grönhaug (1977)    | Computer                        | 16          | n.a.                              | n.a.                            | Specific (several stages)       | Open-ended | High (between-informant) | Global   | Open-ended | Low (between-informant) | —                 |
| Patchen (1974)     | "Non-repetitive purchases"      | 33          | Varied (mean = 5.5)               | Varied                          | Specific (several stages)       | Open-ended | High (self-informant)    | Global   | —       | —               | —                 |
| Choffray (1977)    | Solar cooling system computer terminal | 12          | 2                                 | 8                               | Specific (5 stages)             | Dichotomized rating High (between-informant) | n.a.     | —       | —               | —                 |
| Spekman (1977)     | "Commodity classes" (mfg. mat'l. & office supplies) | 21          | Varied (mean = 6)                 | 1                               | Global | Multiple items rated on 5-point scale | Low (between-informant) | Global | Multiple items rated on 5-point scale | Low (between-informant) |

n.a. = not available.
Research on the informant technique (Seidler 1974) and peer assessment (Kane and Lawler 1978) draws attention to the importance of taking account of such differences in the number and selection of participants in interpreting between-participant consensus. In examining specific estimates of agreement between a pair of involvement reports, one must keep track of the identities of the rater and ratee. In some studies listed in Table 1, consensus assessments are based on comparisons of respondents' self-reports and informants' reports of the same behavior, whereas in other studies the comparisons are between two or more different informants' reports of a third party's behavior. For example, a purchasing agent could provide a self-report of his/her own involvement in a particular decision, and the chief engineer and/or treasurer in the same organization each might also serve as informants to report on the purchasing agent's involvement in that decision. Accordingly, for each study listed in Table 1, we characterize the agreement reported as either "self-informant" or "between informant" although in a few instances the actual consensus reported may mix these two types of data.

Differentiating between respondent or self-report data and informant data is of consequence because they are likely to have different sources of bias and measurement error. Self-reports may inflate the respondents' role whereas a position bias may be present in informants' reports of others' behavior as a result of limited or selective access to information and opportunities to observe the buying process being investigated. Such methods factors can affect the level of convergence observed among different measures, and an understanding of them is crucial in evaluating the merits of competing procedures.

The relevant unit of analysis in the studies is the buying group; pertinent sample size information is reported in Table 1. For six of the eight studies, a single purchase decision was investigated in each of the sample firms and hence the number of buying groups equals the number of organizations. The former quantity exceeds the latter in the studies of Patchen (1974) and Spekman (1977) where several different purchase decisions were examined within some sample firms. The sample size information for respondents is simply the product of the mean number of informants per buying group and the number of buying groups.

Constructs measured. Conceptually, we can distinguish between two constructs measured in the studies: participation and influence. Operationally, participation was generally defined by answers to questions about "who was involved" in one facet or another of some overall decision process whereas influence was typically measured by responses to questions about "how much say" or effect someone had in the course or outcome of the decision process. Though the questions asked differed among studies, the wording of items reported clearly suggests one construct or the other. Weigand (1966), however, asked respondents to indicate on a four-point scale "how concerned" purchasing agents where with six "elements" of a purchase decision (e.g., "naming the manufacturer"). Weigand recognized the ambiguity of the measure and noted: "Concern" may be somewhat different than to ask about 'buying authority' or 'buying responsibility,' but concern is presumably an accompanying characteristic" (p. 82). Rather than disregarding this oft-cited study, we categorize it in Table 1, somewhat arbitrarily, as measuring participation.

Measurement instruments. Measures of participation and influence are categorized as "specific" or "global" depending on whether respondents were asked detailed questions about particular phases of the purchase decision process or a general question referring to the final outcome. The specific questions were usually phrased in terms of a set of decision process stages similar to the "buyphase" scheme proposed by Robinson, Faris, and Wind (1967). The importance of distinguishing between specific and global measures is suggested by Patchen (1963), who found that reports of influence in specific areas were more reliable than global measures.

The final feature of the measures selected for attention is whether the question elicited an assessment of participation or influence in the form of (1) a dichotomous judgment, (2) a rating on a prespecified scale, or (3) a rank ordering of persons. Most investigators determined participation through the use of open-ended questions of the general form "who did what?" and hence the response involved a dichotomous judgment in the sense that someone was either explicitly identified as "involved" or implicitly designated as "not involved." In contrast, all of the reported influence measures involved rating or ranking tasks. Kane and Lawler (1978, p. 557) point out these three types of questions require very different levels of discrimination from respondents.

Note from Table 1 that the number of positions evaluated by each informant was a fixed number in five studies but varied in the investigations of Kelly (1974), Grönhaug (1977), and Patchen (1974). In the latter three cases, each informant was asked questions about participation and/or influence that related to the entire buying group membership and, of course, the size of these groups varied across the decisions investigated. For all the remaining studies, the particular positions or job titles informants were asked to assess were prespecified by the investigators and fixed across all organizations. The studies of Weigand (1966) and Spekman (1977) concentrated on a single role, that of the purchasing agent.

Summary of Findings

1. Respondents tend to attribute more participation and influence to themselves and/or the positions they hold than other informants attribute to them and/or the same positions.
In three of the studies listed in Table 1 (Grashof and Thomas 1976; McMillan 1973; Weigand 1966), the same persistent tendency for self-assessments to be inflated in relation to assessments by other informants is noted. However, the comparisons reported in these three studies were all made on an aggregate basis, by contrasting the marginal distributions (or some measure of central tendency thereof) of “self-ratings” for a particular position with those for others’ ratings of the same position. None of these three studies reported a measure of association between ratings of the same position obtained from pairs of informants within the same organization. Comparison of marginal distributions for two variables does not yield any information about their covariance and hence these studies do not provide estimates of intraorganization consensus. A similar issue has arisen in research on family decision making as pointed out by Davis (1971), who cites examples of husbands’ and wives’ responses being similar when compared on an aggregate basis but dissimilar on a within-family basis. Nonetheless, these studies are valuable in having identified self-aggrandizement as an important source of method variance. The same type of position bias has been observed in other studies of industrial purchasing (e.g., Cooley, Jackson, and Ostrom 1978).

The evidence of high between-informant consensus reported by Choffray (1977; also see Choffray and Lilien 1980) is based on data collected in separate mail surveys undertaken to assess the market potential for new products. The informants’ task was to indicate the pattern of participation that would or would not be involved. In his analysis of consensus and discrimination, Choffray aggregated ratings across the several job titles informants rated, and hence the quality of the participation measures for specific job titles or positions was not addressed.

Kelly (1974; Patchen 1974) are similar in that they involved intensive investigations of past purchase decisions and employed unstructured measures of specific participation. Kelly (1974, p. 424) noted that “little disagreement was detected as to who had performed the five major functions, and members of the decision team were easily identified.” Grönhaug (1977, p. 440) remarked similarly that “In most organizations it was found quite easy to trace the persons involved in the buying process. Furthermore, information from various members of the organization showed consistency, indicating face validity of the observations.” Neither Kelly nor Grönhaug reported any specific data bearing on consensus. A marked degree of agreement about participation was also observed by Patchen (1974), who found that of the average of 5.5 persons identified as decision participants and subsequently interviewed, 4.8 confirmed that “they had some responsibility for or were consulted about either the ‘buy’ decision or the decision about the specific product to buy” (p. 203).

One can argue that specific questions are likely to have an advantage over global ones with respect to reliability or between-informant consensus if the difference between the two types of questions is interpreted in terms of a pair of variables which Sudman and Bradburn (1974) identify as sources of response effects in surveys: structure of the respondents’ task and complexity of the event or behavior to which a question refers. Those authors hypothesize that the less structured the respondents’ task and the more complex the behavior to be reported, the greater the relative response effects will be. These factors seem particularly applicable to organizational purchase decision processes which typically involve numerous interactions over extended periods of time. Specific questions asked with reference to particular stages or elements in the decision process help informants cope with this complexity by structuring their task. Global questions, in contrast, place the burden of decomposing the task on the informant and thus may introduce measurement error into responses because of ambiguity and varying interpretations of the task.

Perhaps greater between-informant consensus about participation is achieved by asking questions requiring nominations or dichotomous judgments rather than more elaborate ratings or rankings because the former task requires the informant to make fewer fine discriminations than the latter (Kane and Lawler 1978). Choffray (1977; Choffray and Lilien 1980) investigated the discriminant ability of his specific, dichotomous measure of participation using a variant of Campbell and Fiske’s (1959) method for analyzing discriminant validity. The results confirmed the ability of informants to differentiate between phases of the decision process in anticipating who would or would not be involved. In his analysis of consensus and discrimination, Choffray aggregated ratings across the several job titles informants rated, and hence the quality of the participation measures for specific job titles or positions was not addressed.
Three of the four studies listed in Table 1 which report low levels of consensus for influence (McMillan 1973; Patchen 1974; Spekman 1976) utilized a global measure that required some type of ordinal assessments. Patchen (1974) administered two global influence questions which asked informants "who was the most influential" in making designated generic and specific product purchase decisions. Comparisons of the nominations received from different informants indicated that "the number of persons named as most influential increases almost as fast as the number of informants increases," leading Patchen to conclude (p. 206) that "the people involved in each decision do not agree very much about who had ‘most influence’." McMillan (1973) also employed a global influence scale and found significant mean differences between the aggregate distributions of self and informant ratings, but did not report any measure of intraorganization covariance or agreement for the ratings. The nature of consensus in influence measures was not pursued by Kelly (1974), but he did mention that "a majority opinion decision rule was used to determine the final rankings of the importance of each team member" (p. 425), thereby acknowledging the presence of some inconsistencies in the global influence rankings.

Another piece of evidence that warrants further discussion is found in Spekman’s (1976) dissertation. Most of the details mentioned here are available in published accounts of the original study (Spekman 1978; Spekman and Stern 1979). Using global measures of purchasing agents' relative participation and influence, Spekman found the level of consensus among buying group members to be generally low for the 52 decisions he investigated. For each buying group, Kendall's coefficient of concordance was computed as an index of agreement among informants who in the average buying group numbered about six. The median value of the 52 coefficients was .314 for the participation measure and .483 for the influence measure. Both measures were total scores obtained by summing responses across multiple items, each of which informants had rated on a five-point scale. The items in the participation scale were general whereas those in the influence scale were specific in the sense of relating to eight phases of the decision process. Given that the measures were summed scores, they are identified as "global" in Table 1. Both scales have a sufficient degree of internal consistency to warrant treating the items as comprising a scale, the value of Cronbach’s (1951) α statistic being .70 for the five-item participation scale and .78 for the eight-item influence scale. Both values appear statistically significant at the .01 level as judged by Feldt’s (1965) approximate test. Spekman found consensus to be similarly limited for measures of several other constructs relating to the structure of buying groups.

The rather disheartening implication of Spekman’s findings is that for global measures of relative participation and influence, the use of internally consistent multiple-item scales (rather than the usual practice of relying on a single item) does not seem to lead to high consensus among informants. Two matters not discussed by Spekman, but of considerable interest in the present context, are (1) to what extent informants differentiated among decision process stages in their ratings of the purchasing agents’ influence and (2) to what extent consensus varied according to specific stages in the decision process. Spekman focused on consensus with respect to the influence total score, not the items composing it relating to specific stages. Other studies have shown that purchasing agents are perceived to be more influential at some stages than others (Bellozzi and Walter 1980; Cooley, Jackson, and Ostrom 1977). If this were the case, influence ratings for different stages would not be single factored. Note that this possibility is not precluded by the fact that Spekman found substantial internal consistency among the eight specific items as reflected by the value of .78 for Cronbach’s α, inasmuch as α estimates the proportion of total score variance due to all common factors present in the items and is an upper bound for the principal factor running through the items (Cronbach 1951).

Unresolved Issues

Before we discuss some unresolved issues, it is helpful to summarize in the following matrix the pattern of results indicated by our analyses of previous studies.

| MEASURE | Global | Participation | Influence |
|---------|--------|---------------|-----------|
| Specific |        | high consensus (proposition 2) |          |
|         |        | focus of present study          | low consensus (proposition 3) |

Thus, measures of participation have been found to exhibit a satisfactory level of reliability as reflected in high between-informant consensus. For measures of influence, however, the evidence, though not extensive, is consistently negative. A possible explanation for these divergent results relates to differences in the types of questions used to measure participation and influence. The high degree of consensus reported for participation has been achieved almost exclusively through the use of the specific questions which refer to particular stages or aspects of the decision process and which were found to be capable of differentiating among such stages on the one occasion when the matter was investigated. In contrast, with a single exception, all studies which found low consensus on influence relied on global measures. The only study in which the instrument measuring influence distinguished among specific stages of the de-
cision process was that of Grashof and Thomas (1976), who found systematic differences between self and informant ratings. However, those authors examined consensus only on an aggregate basis rather than within organizations or buying groups. No assessment of the discriminant ability of specific measures of influence has appeared in any of the literature on organization purchasing of which we are aware.

As argued heretofore, specific questions are likely to have less measurement error than global ones because the former, by structuring the informant's task, may effect a reduction in its complexity. Given the widespread acceptance of Robinson, Faris, and Wind’s “buyphase” conceptualization of industrial purchasing, it is surprising to find that the psychometric properties of specific measures of buying influence have received so little attention—especially in light of the poor performance of global influence measures and the more favorable record of specific participation measures. Other authors have commented on the limitations of global measures of influence. Patchen (1974, p. 208), for example, suggested that the divergence among informants’ reports of global influence observed in his study of organizational buying was probably due to the diffuse nature of the decision process wherein “‘various people play different kinds of roles in the decision, often at different stages.’” Interestingly, in an earlier study of decisions relating to compensation and work assignments in organizations, Patchen (1963) found that reports of influence in specific areas showed better reliability (between-informant agreement) than a global influence measure.

Returning to the measure-construct matrix, we should note that past studies do not shed any light on intraorganizational consensus attainable with either global questions on participation (top left quadrant) or measures of influence addressing specific aspects of a decision (bottom right quadrant). In the former case, task complexity in providing global judgments would lead us to posit that the level of between-informant consensus is likely to be less than that with judgments of participation pertaining to specific aspects of a decision.

The key issue addressed in our study is whether specific assessments of influence obtained from informants can be shown to have a satisfactory level of intraorganization consensus and the ability to discriminate among stages in the decision process. In the next section we describe a small-scale study that investigated this question.

METHOD

Background

A pilot study of the purchase of lithographic plates was undertaken among a purposive sample of 25 printing firms in the greater metropolitan Boston area. All firms included in the sample were general commercial printers rather than specialized or in-house printing operations. In terms of Robinson, Faris, and Wind’s (1967) typology of buying situations, the purchase of lithographic plates by the types of firms studied is somewhere between the straight and modified rebuy categories of purchase decisions. Lithographic plates are a supply routinely used in printing operations and purchased at regular intervals. Although the plates account for only a small fraction of printing costs, their performance has a critical bearing on the quality of the final printing. The principal suppliers are three very large manufacturers and they direct much competitive marketing effort at printing firms. As a result, printers frequently have occasion to re-evaluate their brand selections and switching occurs.

A distinctive feature of our work is that the organizations investigated were small and did not employ purchasing specialists. Twelve of the 25 firms had fewer than 30 employees and only six had more than 100. In all but two cases, the owners were involved in the everyday management of the firms’ operations. Though much industrial marketing activity involves relatively undifferentiated organizations with few employees, most available research on organizational purchasing behavior is concerned with highly structured, large-scale organizations that have professional purchasing staffs. Phillips (1980) has suggested that as organization size increases, it becomes more difficult for an informant to observe relevant processes and events; hence variability among informants’ reports may increase with increasing organization size. We therefore might expect to find greater between-informant agreement in the relatively small organizations we investigated than has been reported in the studies reviewed. Role specialization is also likely to be related to organization size. To the extent that purchasing responsibilities and procedures are more formalized in large than in small organizations, the effect of increasing organization size on the difficulty of an informant’s task may be less direct.

Data Collection

Preliminary investigation revealed that occupants of some or all of four positions constituted the main sphere of potentially relevant decision participants. Though the exact titles differed among organizations, for ease of reference we employ the following set of role designations: (1) general manager (chief operating executive of the firm), (2) foreman (supervisor in charge of all printing operations), (3) platemaker (responsible for carrying out the processes required to prepare the plate used in the printing process), (4) pressman (skilled craftsman who operates the press which produces the printed copy).

Personal interviews were conducted with two members of each firm, one with either a general manager or foreman and the other with a platemaker or pressman. The former group we identify as “managers” and the latter as “users.” Both statuses were represented within the pairs of informants selected from each firm to minimize the possibility that shared positional bias might operate as a
methods factor to inflate the level of between-informant consensus (Campbell and Fiske 1959). In 16 of the firms studied, the “managers” interviewed were “general managers” and in the other nine they were “foremen.” The “users” were “platemakers” in 16 firms and “pressmen” in the remaining nine. The firms were initially contacted by telephone to solicit their cooperation, identify the manager to be interviewed, and set up a convenient appointment. The managers, in turn, provided the assistance needed to arrange the user interviews.

Both managers and users were asked the following set of three questions for each of the four roles.

1. If you had to make a change in lithoplate brands because of problems which were occurring with your usual brand, how likely would each of the following individuals be to suggest the need for a change?
2. How much influence would you expect each of the following individuals to have in evaluating alternative brands?
3. How much influence would you expect each of the following individuals to have in making the final decision as to what brand to change to?

These three questions enabled respondents to differentiate their assessments according to particular stages in the decision process. Descriptions of how switches in suppliers actually came about, obtained in a preliminary investigation of other similar firms, suggested that these three questions tapped key discernible phases of the decision process. Note that the questions ask about expectations rather than past events. Such questions are routinely asked in industrial marketing research studies where information is sought about the composition of buying groups to support various marketing mix decisions.

Respondents were provided with a five-point rating scale for each question. For the first question pertaining to the likelihood of initiating the change, the response categories were labeled “unlikely,” “somewhat likely,” “likely,” “very likely,” and “extremely likely.” For the two influence questions, Patchen’s (1963) scale descriptors were used: “little or no influence,” “some influence,” “quite a bit of influence,” “great influence,” and “very great influence.” Responses were scored on a one to five scale, the higher scores indicating greater involvement or influence.

RESULTS

Aggregate Comparisons

Separate frequency distributions of the managers’ and users’ responses were tabulated for the 12 questions asked—four positions rated with respect to each of three decision functions or stages. Table 2 reports the medians and modes for the two sets of marginal distributions. Judged by either measure of central tendency, the managers’ and users’ ratings appear very similar. For seven of the 12 questions, the medians of the two distributions are identical and only in one case do the medians differ by more than one scale point. An examination of the modes indicates only slightly less consistent aggregate similarity in the ratings provided by managers and users.

Though the summary statistics in Table 2 are derived from a small sample of 25 observations and do not vary greatly across positions rated or decision functions, an interpretation of purchasing involvement can be drawn from them. The foreman appears to be a key decision participant in all three stages. Not surprisingly, general managers were judged as being less likely to initiate a change than any of the other three types of personnel, all of whom are directly involved in using the plates. However, the general manager was acknowledged to be influential in evaluating alternatives and making a final decision. The role of platemakers and pressmen was foremost in initiating changes but diminished in evaluating alternatives and making a final decision. Opinion prevalent in the printing industry suggests greater differentiation between the platemaker and pressman roles than our data indicate.

Position Bias

A series of analyses were undertaken to investigate the possible presence of position bias which had been observed in several of the studies reviewed. First, the relative frequency of agreement and disagreement between informants was found to be independent of the specific role or position held by each member of the pair of “manager” (general manager or foreman) and “user” (platemaker or pressman) selected from each firm. The sample of 25 sets of two informants consisted of nine platemaker-general manager pairs, seven platemaker-foreman pairs, seven pressman-general manager pairs, and two pressman-foreman pairs. When the relative in-

| Stage/position rated | Median Managers | Median Users | Mode Managers | Mode Users |
|----------------------|-----------------|--------------|---------------|------------|
| Initiate change      |                 |              |               |            |
| Pressman             | 5               | 4            | 5             | 4          |
| Platemaker           | 4               | 4            | 5             | 4          |
| Foreman              | 5               | 4            | 5             | 5          |
| General manager      | 3               | 3            | 5             | 1          |
| Evaluate alternatives|                 |              |               |            |
| Pressman             | 4               | 2            | 5             | 2          |
| Platemaker           | 4               | 3            | 4             | 2,4        |
| Foreman              | 4               | 4            | 5             | 5          |
| General manager      | 4               | 4            | 5             | 5          |
| Final decision       |                 |              |               |            |
| Pressman             | 3               | 3            | 3             | 2          |
| Platemaker           | 2               | 3            | 2             | 2          |
| Foreman              | 4               | 4            | 4             | 4          |
| General manager      | 4               | 4            | 5             | 5          |
idence of between-informant agreement and disagreement—distinguishing between whether the manager-informant or user-informant gave a higher rating to a particular role/stage in the entire set of influence ratings (12 per firm)—was compared for these four types of informant pairs, virtually no association could be detected between the two classifications. The value of Goodman and Kruskal’s (1954) index of predictive association \( \lambda_p \) was found to be zero, indicating that information about the particular positions occupied by a pair of manager and user informants would not reduce the probability of error in predicting whether a pair of informants from the same firm agreed or disagreed (or the direction of their disagreement) in attributing influence to the same position.

A second aspect of position bias relates to the distinction between “self-informant” and “between-informant” comparisons. Given that both informants from each firm rated the same four positions, a particular respondent was providing a “self-report” when rating the position he occupied and an “informant report” when rating any of the other three positions. Thus, the pairs of ratings obtained for the 25 firms studied are a mixture of “self plus informant” reports and “two-informant” reports. If the kind of self-aggrandizement position bias discussed before were operating, we would expect agreement to be relatively less frequent in self-informant comparisons than in between-informant comparisons. For each of the 12 ratings, a test was conducted of the null hypothesis of no differences between the self-informant and between-informant comparisons in terms of the relative frequency of agreement versus disagreement. In no case could the null hypothesis be rejected at the 0.05 level. Furthermore, the incidences of disagreement for each of the 12 ratings revealed that the proportion of cases where a management informant attributed more (rather than less) influence to a position than his user counterpart did not differ significantly for the self-informant and between-informant comparisons.

The foregoing results indicated that the ratings obtained from the four types of informant pairs could be pooled without systematically distorting any underlying patterns of agreement or disagreement in the disaggregated data. Thus in the remaining analyses we treat general managers and foremen as “manager” informants and platemakers and pressmen as “user” informants. Table 3 is a summary of the results obtained when the ratings supplied by the manager and user in the same firm were compared. For each of the 12 influence assessments, Table 3 shows the frequency of agreement and, in the cases of disagreement, the frequency with which one or the other informant’s rating was higher. A moderate tendency for managers to give higher ratings than users is evident. For each rating, a Wilcoxon matched-pairs ranked-sign test (Siegel 1956, p. 75–83) was carried out to test the null hypothesis of no difference between the ratings of the pairs of manager and user informants. The two-tail probabilities of obtaining values of the test statistic as large as those observed under the null hypothesis are also reported in Table 3. For only two of the 12 ratings do the Wilcoxon test results indicate the presence of between-informant differences that are significant at the .05 level. The tendency sometimes reported in previous studies for ratings to be biased in a direction related to the informant’s own position is not evident in these data to any marked degree. On the contrary, managers were more inclined than users to attribute involvement to pressmen and platemakers for two of the three decision functions rated.

### Consensus

To determine the degree of consensus between managers and users within the same firm on their perceptions of purchase influence, we cross-tabulated the two sets of responses and computed a measure of agreement for each of the 12 ratings. The coefficient used is that proposed by Cohen (1960) and discussed by Bishop, Fienberg, and Holland (1975, p. 395–7). Fleiss (1975) examined the numerous measures of agreement for

---

**Table 3**

| Stage/position evaluated | Frequency | Wilcoxon test probability |
|--------------------------|-----------|---------------------------|
|                         | M = U     | M > U | M < U | (2-tail test) |
| **Initiate change**      |           |       |       |              |
| Pressman                 | 7         | 10    | 8     | .866         |
| Platemaker               | 7         | 11    | 7     | .442         |
| Foreman                  | 7         | 14    | 4     | .001         |
| General manager          | 8         | 11    | 6     | .306         |
| **Evaluate alternatives**|           |       |       |              |
| Pressman                 | 6         | 13    | 6     | .032         |
| Platemaker               | 6         | 13    | 6     | .156         |
| Foreman                  | 10        | 9     | 6     | .678         |
| General manager          | 7         | 7     | 11    | 1.000        |
| **Final decision**       |           |       |       |              |
| Pressman                 | 5         | 9     | 11    | .522         |
| Platemaker               | 10        | 8     | 7     | .890         |
| Foreman                  | 8         | 11    | 6     | .548         |
| General manager          | 9         | 10    | 8     | .640         |

M = Managers’ ratings.  
U = Users’ ratings.  
*For each row in the table, the sum of the three column frequencies is 25.*

---

2 The Wilcoxon test makes use of information about the sign and magnitude of the difference between each pair of informants’ ratings. Disregarding the information about the magnitude of differences and focusing only on a directional comparison between pairs of ratings, we performed a sign test of the null hypothesis that managers’ ratings exceeded users’ ratings with no greater frequency than the reverse occurrence. The results were essentially similar to those indicated by the Wilcoxon test. The null hypothesis could be rejected at .10 level (two-tail test) for only one of the 12 ratings (the role of the foreman in initiating changes).
categorical data which have been proposed in the psychometric and statistics literature and concluded that $\kappa$ is one of only two measures "defensible both as chance-corrected measures and as intraclass correlation coefficients." More recently, Kraemer (1979) has shown how $\kappa$ relates to the classical psychometric model for reliability of interval data (the ratio of true score variance to observed score variance) and discussed its interpretation and use "to indicate the degree of loss of precision or power of statistical procedures" due to the unreliability of observations. Previous applications of $\kappa$ to marketing research problems similar to ours are reported by Davis, Douglas, and Silk (1981). As applied here, the value of the coefficient is given by:

$$\kappa = \frac{O - E}{1 - E},$$

where:

$O =$ observed proportion of cases within the sample where the manager and user from the same organization gave identical responses in rating a particular position with respect to a specific decision stage, and

$E =$ expected proportion of identical ratings under the assumption that the managers' and users' ratings are independent.

Thus this coefficient reflects the excess of observed over chance agreement, normalized by the maximum possible value of this difference given the marginal distributions of the observed responses. The coefficient is zero when the observed agreement is just equal to that expected by change and unity when the maximum possible excess of observed over chance agreement is obtained. Negative values of the coefficient indicate less observed agreement than expected by chance. Note that this coefficient is a measure of agreement in the sense that if depends only on the frequency of identical ratings—entries in the main diagonal—as distinct from an index of association which would also take account of the off-diagonal frequencies.

To illustrate the calculation of $\kappa$ consider the following contingency table, obtained by cross-tabulating the managers' and users' ratings of the "general manager's" influence in the "initiating change" stage.

|          | 1 | 2 | 3 | 4 | 5 | Total |
|----------|---|---|---|---|---|-------|
| Manager's ratings |   |   |   |   |   | 8     |
| User's ratings      | 1 | 2 | 1 | 2 | 1 | 4     |
|                     | 1 | 3 | 1 |   |   | 5     |
|                     | 1 |   | 1 |   | 3 | 5     |
|                     | 2 |   |   |   | 1 | 3     |
| Total               | 2 | 6 | 6 | 4 | 7 | 25    |

The figures in parentheses along the main diagonal are the expected frequencies computed by using the marginal distributions of the two sets of ratings. For these data we obtain:

$$O = \frac{1}{25} (2 + 1 + 3 + 1 + 1) = 0.32,$$

$$E = \frac{1}{25} (0.64 + 0.96 + 1.2 + 0.8 + 0.84) = 0.1776,$$

$$\kappa = \frac{0.32 - 0.1776}{1 - 0.1776} = + 0.173.$$

Two sets of agreement coefficients were computed. One was obtained by collapsing the five-point scale into two categories and treating the ratings as a dichotomous measure of influence. The two highest scale levels formed the "influential" category and the three lesser ratings were combined into a "noninfluential" grouping. The second set of coefficients was computed by using the full range of ratings from the original five-

---

Table 4
COEFFICIENTS OF AGREEMENT BETWEEN MANAGERS' AND USERS' INFLUENCE RATINGS OF DECISION PARTICIPANTS
BY STAGE: DICHHOTOMIZED VS. FIVE-POINT SCALE RESPONSES
(n = 25)
Table 5
INTERCORRELATIONS WITHIN MANAGERS’ INFLUENCE RATINGS OF DECISION PARTICIPANTS BY STAGE
(Q-values, n = 25)

|       | P     |       |       | Managers’ ratings |       |       |       |
|-------|-------|-------|-------|-------------------|-------|-------|-------|
|       | I     | E     | D     | I     |       |       |       |
| E     | (P)   | (E)   | (D)   | I     | E     | D     | I     | E     | D     |
| D     | (.71) | .63   | 1.0   | —     | —     | —     | —     | —     | —     |
| L     | .20   | (.81) | 1.0   | (.83) | —     | —     | —     | —     | —     |
| I     | .34   | .45   | (1.0) | (1.0) | —     | —     | —     | —     | —     |
| F     | (.71) | —     | 1.0   | (.86) | .22   | (1.0) | —     | —     | —     |
| E     | —     | (.87) | —     | (.33) | —     | —     | —     | —     | —     |
| D     | 0.0   | —     | (.10) | .11   | —     | (.10) | —     | —     | (.85) |
| I     | (.33) | .14   | —     | (.47) | —     | —     | (.10) | —     | —     |
| F     | —     | (.37) | (.29) | —     | —     | —     | —     | —     | —     |
| E     | —     | (.37) | —     | (.37) | —     | —     | —     | —     | —     |
| D     | (.45) | (.27) | (.55) | .86   | .05   | (.20) | —     | (.63) | (.10) |

Notation key:
- P = Pressman
- L = Platemaker
- I = Initiate change
- F = Foreman
- D = Final decision
- G = General manager

Discrimination Among Roles and Decision Stages

Some further analysis was undertaken to investigate the extent to which the measures discriminated among the positions or roles rated and stages in the decision process. Because an evaluation of the discriminating ability of measures is warranted only if they exhibit some meaningful level of convergence (Campbell and Fiske 1959), attention is confined here to a subset of role-stage influence ratings for which the level of between-informant consensus was sufficient to warrant further examination. Thus, the analysis is opportunistic and should be interpreted with caution. Intercorrelations among the ratings provided by managers and users for the various roles and stages were examined with reference to two criteria suggested by Campbell and Fiske’s (1959) multitrait-multimethod matrix approach to assessing discriminant validity.

Tables 5–7 report the matrix of intercorrelations among the 24 separate influence ratings (2 informants × 4 roles × 3 stages). The measure of association is Yule’s $Q$, computed for the $2 \times 2$ contingency tables formed by cross-tabulating pairs of ratings, dichotomized in the manner explained previously in connection with the coefficient of agreement. As noted in the review of previous research, there is reason to believe that the reliability and discriminant ability of dichotomous judgments may be greater than those of scaled ratings. In the present context, $Q$ reflects how much better than chance one could predict one assessment of influence (or noninfluence) by always predicting the same outcome indicated by a second assessment (see Davis 1971, Chapter 2). $Q$ may range from $-1$ to $+1$ and is zero where the two assessments are independent. A positive $Q$-value indicates a tendency for the two assessments of influence vs. noninfluence to be similar whereas a negative value
Table 6
INTERCORRELATIONS BETWEEN MANAGERS’ AND USERS’ INFLUENCE RATINGS OF DECISION PARTICIPANTS BY STAGE*  
(Q-values, n = 25)

|       | P            | L            | F            | G            |
|-------|--------------|--------------|--------------|--------------|
|       | I    | E   | D   | I    | E   | D   | I    | E   | D   | I    | E   | D   |
| I     | .73  | (.67)| (.40)| [.23]| -.03| -.08| [.48]| -.23| .33  | [.08]| .76 | -.35|
| P E   | (.45)| .54 | (.93)| .35 | .54 | .81 | 1.0  | [.05]| 1.0  | .29 | (.18)| -.18|
| D     | (.56)| (.86)| .90 | .46 | .86 | [.71]| -.30| -.28 | [.11]| .14 | .01 | (.01)|
| I     | [.33]| -.14| .19 | -.22| (.20)| (0.0)| [.22]| .14  | .53 | [.64]| .71 | .50 |
| L E   | .71 | (.56)| 1.0 | .86 | .78 | (.88)| 1.0  | [.01]| 10   | -.14| (.05)| -.05|
| Users’ | D   | .71 | 1.0 | [.95]| .66 | (1.0)| .73  | -.48| -.20 | [.33]| .08 | -.03| (.03)|
| ratings| I    | .01 | -.37| [.09]| -.64| -.47| .30  | [.09]| (.11)| .82 | .33 | .01 |
|        | F E  | -.14| [.05]| -.16| -.33| [.05]| -.20| (.10)| .01  | (.60)| .45 | (.05)| .37 |
|        | D    | 0.0 | -.13| [.19]| -.56| -.45| [.33]| (.22)| (.14)| .00 | .64 | .45 | (.20)|
|        | I    | -.33| -.50| -.19| -.14| -.20| 0.0  | [.22]| .22  | .00 | .61 | (.20)| (.45)|
|        | G E  | -.45| [.27]| -.16| -.33| [.05]| -.20| (.10)| -.33| -.60| (.71)| .05 | (.37)|
|        | D    | -.26| -.40| [.05]| -.11| -.09| [.07]| .04  | -.11| -.20| (.56)| (.09)| .23 |

*Notation key:  
P = Pressman  
I = Initiate change  
L = Platemaker  
E = Evaluate alternatives  
F = Foreman  
D = Final decision  
G = General manager

indicates a tendency toward dissimilar assessments.

Tables 5–7 correspond to the main elements of Campbell and Fiske’s (1959) multitrait-multimethod matrix. The two informants (manager and user) are treated as different “methods” whereas the various combinations of roles and decision stages rated constitute different “traits.” Tables 5 and 7 contain the “within-informant” correlations for the managers and users, respectively, and are analogous to Campbell and Fiske’s mono-method-heterotrait triangles. The “between-informant” correlations are reported in Table 6 and are treated as two “heteromethod-heterotrait” triangles separated by the diagonal of convergent coefficients (underscored) which represent the correlations between managers’ and

Table 7
INTERCORRELATIONS WITHIN USERS’ INFLUENCE RATINGS OF DECISION PARTICIPANTS BY STAGE*  
(Q-values, n = 25)

|       | P            | L            | F            | G            |
|-------|--------------|--------------|--------------|--------------|
|       | I    | E   | D   | I    | E   | D   | I    | E   | D   | I    | E   | D   |
| I     | 1    |     |     |     |     |     |     |     |     |     |     |     |
| P E   | (.11)|     |     |     |     |     |     |     |     |     |     |     |
| D     | (.23)|     |     |     |     |     |     |     |     |     |     |     |
| L E   | (.73)| .08 | .14 |     |     |     |     |     |     |     |     |     |
| Users’ | D   | .02 | .65 | .94 | (.33)| (.88)|     |     |     |     |     |     |
| ratings| I    | (.55)| -.05| -.58| (.71)| -.35| -.55  |     |     |     |     |     |
|        | F E  | -.03| [.18]| -.01| .20  | [.05]| .03  | (.63)|     |     |     |     |
|        | D    | .45 | .45 | [.14]| .61  | .14  | [.45] | (.88)| (.88)|     |     |     |
|        | G E  | -.03| [.18]| -.01| .20  | [.37]| .03  | .55  | [.93]| .50  | (.10)|     |
|        | D    | .25 | -.06| [.23]| .07  | -.23| [.59]| .54  | .89  | [.83]| (1.0)| (.97)|

*Notation key:  
P = Pressman  
L = Platemaker  
E = Evaluate alternatives  
F = Foreman  
D = Final decision  
G = General manager
users' ratings of the same roles and decision stages. Within each of the four hetero-role/stage triangles in Tables 5–7, three different types of correlations can be distinguished: (1) correlations between ratings of the same role for two different stages of the decision process (enclosed in parentheses), (2) correlations between ratings of two different roles for the same stage of the decision process (enclosed in brackets), and (3) correlations between ratings of two different roles for two different stages of the decision process (not enclosed in either parentheses or brackets).

Recall from the previous discussion that consensus as measured by Cohen's coefficient of agreement (Table 4) appeared strongest for five of the six measures relating to the influence of the pressman and platemaker. The Q-values for these measures in the main diagonal of Table 6 reflect the degree of association or convergence between two measures of the same trait and are the counterparts of the coefficients of agreement reported previously as indices of consensus. Note that "agreement" is a special case of "association" (Bishop, Fienberg, and Holland 1975, p. 393–4). Fisher's exact test was computed for the 2 × 2 contingency tables underlying each of the five congruent Q-coefficients reported in Table 8 whose magnitude ranged from .54 to .90. The one-tail probabilities for Fisher's test are also shown in Table 8 and indicate that the null hypothesis of no association between managers' and users' ratings could be rejected at the .10 level (or less) for four of the Q-coefficients but at only the .20 level for the fifth coefficient.

As evidence of discriminant validity, Campbell and Fiske suggest that a measure should correlate more highly with another measure of the same trait than it does with other measures having neither trait nor method in common. The mechanics of performing this test involve comparing the Q-values for the relevant convergent correlations in the main diagonal of Table 6 with each of the other 12 Q-values representing the "between-informant" correlations in the same row and column of the two hetero-role/stage triangles in Table 6. Counts of the number of directional confirmations found for this criterion are summarized in Table 8. The convergent coefficients exceed the "same stage-different role" and "different stage-different role" correlations in more than 80% of the comparisons, thereby providing a favorable indication of discrimination. However, only half of the comparisons involving "same role-different stage" correlation are confirmatory. The pattern of these results suggests that the influence ratings reflect some differentiation among roles but not decision stages.

Table 9

| Comparison coefficients | Within manager | Within user |
|-------------------------|----------------|------------|
| Role-stage Q-value | Same role-diff. stage | Same stage-diff. role | Same role-diff. role | Same stage-diff. role | Diff. role |
| PI | .73 | 0/2 | 2/3 | 6/6 | 2/2 | 6/6 |
| PE | .54 | 0/2 | 2/3 | 5/6 | 1/2 | 5/6 |
| PD | .90 | 0/2 | 1/3 | 5/6 | 1/2 | 6/6 |
| LE | .78 | 0/2 | 2/3 | 5/6 | 1/2 | 6/6 |
| LD | .73 | 0/2 | 1/3 | 5/6 | 1/2 | 6/6 |
| Totals | 0/10 | 8/15 | 24/30 | 6/10 | 11/15 | 29/30 |

(no. of confirmations/no. of comparisons)
A second criterion of discriminant validity put forth by Campbell and Fiske is that a measure should correlate more highly with a different measure of the same trait than with measures of other traits employing the same method. In the present context, this calls for comparisons between a measure's convergent correlation and its correlation with the 11 other role-stage ratings in each of the "within-manager" and "within-user" triangles in Tables 5 and 7, respectively. Table 9 summarizes the outcome of the comparisons. The results indicate that the discriminant quality of users' ratings was somewhat better than that of the managers' ratings but, as in the previous test, the absence of differentiation between stages in the decision process is apparent. Poor discrimination with respect to roles as well as stages is evident for the managers' ratings where only about half of the comparisons involving "same stage-different role" are in a confirmatory direction.

As a final indication of the discriminating qualities of the ratings, the median values of the Q-coefficients within each of the four hetero-role/stage triangles were computed for the three types of role-stage correlations. The median Q-values are reported in Table 10. The median of the "same role-different stage" correlations is positive for all four triangles and generally appears elevated in comparison with the median Q's for the "same stage-different role" and "different stage-different role" correlations which fluctuate near zero.

In summary, we find some modest evidence of role differentiation but little indication of discrimination among stages in the decision process. The influence attributed to a role at some particular decision stage tends to be essentially independent of assessments of other roles for the same stage. However, how influential a role was judged to be at one stage is generally positively related to the judgment made of that role for other stages. We are aware of no a priori reason for expecting the latter type of correlation and hence the lack of discrimination among stages is perplexing. Whether the limitations of the measure's discriminating qualities are due to method factors or real interdependencies among the roles and stages cannot be clearly ascertained from the analyses. With a more extensive data base consisting of additional continuous measures and a larger sample, more powerful structural equation methods for analyzing multitrait-multimethod matrices developed by Joreskog (1970, 1971) could be used to address this matter. Phillips (1980) and Phillips and Bagozzi (1980) have recently performed this kind of analysis on informant data.

Certain limitations of the foregoing analyses warrant mention. First, the number of ratings (12 per informant) and correlations (276) examined is very large in relation to the number of observations underlying the analysis (25 pairs of informants). Second, with a small sample, the magnitude of Yule's Q is sensitive to shifts in the raw ratings and no test-retest reliabilities or other estimates of the magnitude of within-informant random measurement error are available for the individual ratings. Hence, our results should be regarded as tentative.

**DISCUSSION**

The low level of consensus and limited degree of purchase stage discrimination exhibited by the measures employed, along with the results reviewed from previous studies, afford a rather discouraging picture of the quality of influence assessments obtained through the use of simple rating procedures commonly administered in industrial marketing research surveys. However, the research that has accumulated to date, including ours, has not progressed very far in establishing knowledge about the sources and magnitude of systematic and random measurement errors nor has it offered much direction about how to improve measurement in this area. The formidable nature of the latter task is underscored by
experience in other research on influence measurement and the key informant method. For example, even in experimentally created groups, attributions of influence by informants have been found to be relatively independent of measures of opinion change (March 1956). Somewhat more encouraging results were obtained by Phillips (1980), who collected reports on several organizational variables including power-dependence relations with customers and suppliers from multiple informants in each of 506 wholesale distribution companies. Phillips found that if an unmeasured methods factor was associated with each set of informants, tests for convergent and discriminant validity could generally be satisfied by informants' reports. However, methods and random error variance typically accounted for more than 50% of the observed variance in the informants' data, leading Phillips to conclude that the measures "must be viewed as questionable." Further analysis of these data by Phillips and Bagozzi (1980) showed that structural relationships among constructs were frequently attenuated by the presence of these substantial components of measurement error in the informant measurements.

Virtually all of the research on influence we examined has relied on questionnaire items that depict the underlying decision process as a progression through a series of phases or stages, beginning with the recognition of a problem or need and ending with the selection of suppliers and/or evaluation of their performance. Even when such items refer to several specific stages, as in our study, the measures obtained are very global or general in the sense that an informant is asked to provide an overall assessment about the outcome for a particular stage, where the stage itself would ordinarily encompass a myriad of events and interactions distributed over time. The presumption that, somehow, global ratings about relative influence by decision phase ought to be reliable and valid needs to be reexamined and appears to be the appropriate point of departure for future efforts to improve the quality of measures of organizational purchasing influence.

Our results indicating that influence ratings discriminated among roles but not decision stages might be interpreted as suggesting that the task informants were asked to perform was excessive. In his discussion of the use of the informant technique in sociological research on organizations, Seidler (1974, p. 817) observed that informants "are often asked, at least implicitly, to perform calculations otherwise left for the computer." Referring to their work on response effects in survey interviews, Cannell, Oksenberg, and Converse (1977) concluded, "The demands placed on the respondent by many survey questions are greater than generally has been realized, and the respondent's inability or unwillingness to meet these demands is a major source of invalidity" (p. 309). Some appreciation for the burden of influence questions placed on respondents can be obtained by examining the nature of the influence processes described in intensive descriptive studies of purchasing decision making like those done by Cyert, Simon, and Trow (1956) and Pettigrew (1975) and imagining the difficulties informants in those organizations would have in making judgments about the relative influence of different participants in various stages. Those studies reveal that influence is exerted in a variety of complicated ways, including through search and control of information, power relations, and negotiation. Patchen (1974) uncovered similar phenomena and emphasized the view of organizational purchase decision making as a "group process involving some kind of accommodation among individuals and units" (p. 155). In light of such studies, the kinds of influence processes that surround purchase decisions apparently are not easily related to hierarchical or multistage notions like the "buyphase" scheme of Robinson, Faris, and Wind (1967), which is the conceptualization underlying most of the empirical research on influence measurement.

To assess influence, it seems preferable to ask questions that relate to accommodation and other processes through which influence is manifested. Microtheoretical notions about social influence like those discussed by Bagozzi (1978) offer a starting point in conceptualizing the kinds of processes relevant here. However, descriptive studies involving intensive qualitative exploration of purchasing decisions (Calder 1977) and/or observational and unobtrusive methods (Cyert, Simon, and Trow 1956; Pettigrew 1975) would be required to develop instruments suitable for applications in survey research.

CONCLUSIONS

Three major conclusions can be drawn about methods used in marketing research to study the structure and composition of organizational buying groups.

1. Measures of purchase decision participation having a satisfactory level of between-informant reliability may be obtained by asking informants questions which focus on specific stages or elements of the overall decision process. Instruments which elicit nominations or dichotomous judgments about the identity of decision participants appear preferable to those which require informants to make finer discrimination in the form of ratings or rankings.

2. The reliability and validity of methods currently used to measure the influence of different participants in organizational purchase decisions have yet to be established. Self-reports have been shown to be prone to varying amounts of position bias. Although the body of relevant evidence is limited, studies of between-informant consensus have consistently found low levels of agreement for global types of influence assessments. These negative results should discourage the common practice of basing influence measurements on single questions or items which implicitly assume that informants are willing and able to provide episodic summary judgments about complex decision processes. Steps to structure the informant's task, such as asking for reports of events and interactions rather than for inferences about social phenomena, are
likely to reduce measurement error. Methods of measuring influence are needed which go beyond the conceptualization of organization purchasing as phased or sequential decision processes and which are related to the nature of processes through which influence is exerted.

3. As has been previously advocated (Wind 1978), assessments of purchase participation and influence should be obtained from more than one informant within an organization to enable the researcher to detect instances of disagreement and avoid being unknowingly led to accept a misleading view of buying center composition and structure. Controls need to be exercised in selecting informants to ensure familiarity with the processes being investigated and to cope with position biases. Analyses of the convergent and discriminant qualities of informant measures should be a routine part of buying center studies. Particularly useful for such purpose are the structural equation methods which Phillips (1980) and Phillips and Bagozzi (1980) have recently applied to analyze the psychometric properties of informant data.

REFERENCES

Bagozzi, Richard P. (1978), “Exchange and Decision Processes in the Buying Center,” in Organizational Buying Behavior, Thomas V. Bonoma and Gerald Zaltman, eds. Chicago: American Marketing Association, 100–25.

Bellozzi, Joseph A. and C. K. Walter (1980), “Purchasing Agents’ Influence in the Buying Process,” Industrial Marketing Management, 9 (April), 137–41.

Bishop, Yvonne M. M., Stephen E. Fienberg, and Paul W. Holland (1975), Discrete Multivariate Analysis. Cambridge, MA: MIT Press.

Bonoma, Thomas V., Gerald Zaltman, and Wesley J. Johnstone (1977), “Industrial Buying Behavior,” Report No. 77–117, Marketing Science Institute, Cambridge, MA.

Brand, Gordon T. (1972), The Industrial Buying Decision. New York: John Wiley & Sons, Inc.

Buckner, Hugh (1967), How British Industry Buys. London: Hutchinson.

Calder, Bobby J. (1977), “Structural Role Analysis of Organizational Buying: A Preliminary Investigation,” in Consumer and Industrial Buying Behavior. Arch G. Woodside, Jagdish N. Sheth, and Peter D. Bennett, eds. New York: North-Holland, 193–9.

Campbell, Donald T. and Donald W. Fiske (1959), “Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix,” Psychological Bulletin, 52 (March), 81–105.

Cannell, Charles F., Lois Oksenberg, and Jean M. Converse (1977), “Striving for Response Accuracy: Experiments in New Interviewing Techniques,” Journal of Marketing Research, 14 (August), 306–15.

Choffray, Jean-Marie (1977), “A Methodology for Investigating the Nature of the Industrial Adoption Process and the Differences in Perceptions and Evaluation Criteria Among Decision Participants,” Ph.D. thesis, Massachusetts Institute of Technology.

——— and Gary Lilien (1980), “Industrial Market Segmentation by the Structure of the Buying Process,” Industrial Marketing Management, 9 (October), 331–42.

Cohen, Jacob (1960), “A Coefficient of Agreement for Nominal Scales,” Educational and Psychological Measurement, 20 (Spring), 37–46.

Cooley, James R., Donald W. Jackson, Jr., and Lonnie L. Ostrom (1977), “Analyzing the Relative Power of Participants in Industrial Buying Decisions,” Proceedings, 1977 Educators’ Conference. Chicago: American Marketing Association, Series No. 41, 243–6.

———, ———, and ——— (1978), “Relative Power in Industrial Buying Decisions,” Journal of Purchasing and Materials Management (Spring), 18–20.

Cronbach, Lee J. (1951), “Coefficient Alpha and the Internal Structure of Tests,” Psychometrika, 16 (September), 297–334.

Cyert, Richard M., Herbert A. Simon, and Donald B. Trow (1956), “Observation of a Business Decision,” Journal of Business, 29 (October), 237–48.

Davis, Harry L. (1971), “Measurement of Husband-Wife Influence in Consumer Purchase Decisions,” Journal of Marketing Research, 8 (August), 305–12.

———, Susan P. Douglas, and Alvin J. Silk (1981), “Measure Unreliability: A Hidden Threat to Cross-National Consumer Research?” Journal of Marketing, 45 (Spring), 98–109.

Davis, James A. (1971), Elementary Survey Analysis. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Feldt, Leonard S. (1965), “The Approximate Sampling Distribution of Kuder-Richardson Reliability Coefficient Twenty,” Psychometrika, 30 (September), 357–70.

Fleiss, Joseph L. (1975), “Measuring Agreement Between Two Judges on the Presence or Absence of a Trait,” Biometrics, 31 (September), 651–9.

Goodman, Leo A. and William H. Kruskal (1954), “Measures of Association for Cross Classifications,” Journal of the American Statistical Association, 49 (December), 732–64.

Grashof, John F. and Gloria P. Thomas (1976), “Industrial Buying Center Responsibilities: Self Versus Other Member Evaluations of Importance,” Proceedings, 1976 Educators’ Conference. Chicago: American Marketing Association, Series No. 39, 344–9.

Grönhaug, Kjell (1977), “Exploring a Complex Organizational Buying Decision,” Industrial Marketing Management, 6 (December), 439–44.

Jöreskog, K. G. (1970), “A General Method for Analysis of Covariance Structures,” Biometrika, 57, No. 2, 239–51.

Kelly, J. Patrick (1974), “Functions Performed in Industrial Purchasing Decisions with Implications for Marketing Strategy,” Journal of Business Research, 2 (October), 420–34.

Kraemer, Helen C. (1979), “Ramifications of a Population Model for K as a Coefficient of Reliability,” Psychometrika, 44 (December), 461–72.

March, James G. (1956), “Influence Measurement in Experimental and Semi-Experimental Groups, Sociometry, 19 (December), 260–71.

McMillan, James R. (1973), “Role Differentiation in Industrial Buying Decisions,” Proceedings, 1973 Educators’ Conference. Chicago: American Marketing Association, Series No. 35, 207–11.

Patchen, Martin, (1963), “Alternative Questionnaire Approaches to the Measurement of Influence in Organizations,” American Journal of Sociology, 69 (July), 41–52.

——— (1974), “The Locus and Basis of Influence in Orga-
nizational Decisions," *Organizational Behavior and Human Performance,* 11 (April), 195–221.

Pettigrew, Andrew M. (1975), "The Industrial Purchasing Decision as a Political Process," *European Journal of Marketing,* 9, 4–19.

Phillips, Lynn W. (1980), "Assessing Measurement Error in Key Informant Reports on Organizational Properties," Working Paper No. 554, Graduate School of Business, Stanford University.

--- and Richard P. Bagozzi (1980). "On Measuring Organizational Properties: Methodological Issues in the Use of Key Informants," Working Paper, Sloan School of Management, Massachusetts Institute of Technology.

Platten, J. H., Jr. (1950), *How Industry Buys.* New York: Scientific American.

Robinson, Patrick J., Charles W. Faris, and Yoram Wind (1967), *Industrial Buying and Creative Marketing,* Boston: Allyn and Bacon.

Seidler, John (1974), "On Using Informants: A Technique for Collecting Quantitative Data and Controlling Measurement Error in Organization Analysis," *American Sociological Review,* 39 (December), 816–31.

Siegel, Sidney (1956), *Nonparametric Statistics.* New York: McGraw-Hill Book Company.

Spekman, Robert E. (1977), "A Contingency Approach to Power Relationships Within the Industrial Buying Task Group," unpublished Ph.D. dissertation, Northwestern University.

--- (1978), "A Macro-Sociological Examination of the Industrial Buying Center: Promise or Problem," *Proceedings, 1978 Educators' Conference.* Chicago: American Marketing Association, Series No. 43, 111–15.

--- and Bobby J. Calder (1978), "Narrowing the Conceptual-Empirical Gap in Organizational Buying Behavior," in *Advances in Consumer Research,* Vol. 5, H. Keith Hunt, ed., Association for Consumer Research, 653–6.

--- and Louis W. Stern (1979), "Environmental Uncertainty and Buying Group Structure: An Empirical Investigation," *Journal of Marketing,* 43 (Spring), 54–64.

Sudman, Seymour and Norman M. Bradburn (1974), *Response Effects in Surveys.* Chicago: Aldine.

Webster, Frederick E., Jr. and Yoram Wind (1972), *Organization Buying Behavior.* Englewood Cliffs, NJ: Prentice-Hall, Inc.

Weigand, Robert E. (1966), "Indentifying Industrial Buying Responsibility," *Journal of Marketing Research,* 3 (February), 81–4.

Wind, Yoram (1978), "Organization Buying Center: A Research Agenda," in *Organizational Buying Behavior,* Thomas V. Bonoma and Gerald Zaltman, eds. Chicago: American Marketing Association, 67–76.