Journal of Applied Psychology
2004, Vol. 89, No. 1, 36–51

Copyright 2004 by the American Psychological Association, Inc.
0021-9010/04/$12.00 DOI: 10.1037/0021-9010.89.1.36

The Forgotten Ones? The Validity of Consideration and Initiating Structure in Leadership Research

Timothy A. Judge, Ronald F. Piccolo, and Remus Ilies
University of Florida

This study provided a meta-analysis of the relationship of the Ohio State leadership behaviors—Consideration and Initiating Structure—with leadership. Overall, 163 independent correlations for Consideration and 159 correlations for Initiating Structure were analyzed. Results revealed that both Consideration (.48) and Initiating Structure (.29) have moderately strong, nonzero relations with leadership outcomes. Consideration was more strongly related to follower satisfaction (leader satisfaction, job satisfaction), motivation, and leader effectiveness, and Initiating Structure was slightly more strongly related to leader job performance and group–organization performance. Validities did vary by leadership measure, but in most cases validities generalized regardless of the measure used. Overall, the results provide important support for the validity of Initiating Structure and Consideration in leadership research.

The 1940s was a period of great importance to the field of leadership. Disappointed by the yield from studies investigating the trait theory of leadership, a group of researchers at Ohio State University—led by R. Stogdill, C. Shartle, and J. Hemphill—sought to uncover the behavioral indicators of effective leadership (see Stogdill, 1950). Although at various periods many behaviors were studied, the Ohio State studies isolated two factors: Consideration and Initiating Structure (or Structure). Consideration is the degree to which a leader shows concern and respect for followers, looks out for their welfare, and expresses appreciation and support (Bass, 1990). Initiating Structure is the degree to which a leader defines and organizes his role and the roles of followers, is oriented toward goal attainment, and establishes well-defined patterns and channels of communication (Fleishman, 1973). Until the advent of transformational leadership theory beginning in the late 1970s (Bass, 1985; Burns, 1978; House, 1977), these two dimensions dominated leadership research. As Fleishman (1995) noted, "Consideration and Initiating Structure have proven to be among the most robust of leadership concepts" (p. 51).[1]

In the more than half century since the discovery of Consideration and Initiating Structure, much has been learned about these concepts. At the same time, upon reflecting on this literature, one cannot help but be impressed by the mysteries surrounding Consideration and Structure, as well as how quickly they fell out of favor in leadership research. Questions were raised about the generality of the validities and the nature of the measures themselves; many may feel that these questions were never answered satisfactorily. Fleishman (1995), for example, argued that the validities of Consideration and Initiating Structure are curvilinear, such that there are diminishing returns to the increased use of consideration and structure on the part of the leader. However, with the exception of the highly cited study by Fleishman and Harris (1962), this hypothesis has not been tested further (Fleishman, 1998). Similarly, research was often criticized for its reliance on common source data in which the leadership behavior ratings and criteria were collected from the same source (Kerr & Schriesheim, 1974), although in fairness, many subsequent studies did use independent data sources (e.g., Ilgen & Fuji, 1976; Sheridan & Vredenburgh, 1978b). Another controversy is how Consideration and Structure should be measured. Several measures of the constructs have been developed, each has been criticized on various grounds (Schriesheim & Kerr, 1974), and yet each seems to have its adherents. As we note later, whether these measures in fact assess the same construct, and which measure is superior, is not

Timothy A. Judge, Ronald F. Piccolo, and Remus Ilies, Department of Management, Warrington College of Business, University of Florida.

Remus Ilies is now at the Department of Management, Michigan State University.

We thank Ed Fleishman and Henry Tosi for their assistance in locating some of the studies included in the meta-analysis.

Correspondence concerning this article should be addressed to Timothy A. Judge, Department of Management, Warrington College of Business, University of Florida, Gainesville, Florida 32611. E-mail: timothy.judge@cba.ufl.edu

[1] Although we make primary reference to the work of the Ohio State leadership studies, we note that researchers at the University of Michigan carried out a relatively concurrent and in some ways comparable leadership research program (e.g., Likert, 1961). Although the contributions of the Michigan studies should be acknowledged, the Michigan dimensions of employee-centered leadership and production-centered leadership were much less frequently studied in empirical research. Perhaps the most focal contribution of the Michigan studies was the Blake and Mouton managerial grid (Blake, Mouton, & Bidwell, 1962), which has been criticized in the literature on various grounds (Bernardin & Alvares, 1976). Similarly, at roughly the same time, Bales and associates at Harvard studied person- and task-oriented behaviors (Bales, 1954) although again, these contributions have been less recognized in the literature. As noted by Chemers (1997), "The most extensive program of research on the relative effectiveness of leadership behavior patterns was carried out by the very productive group of researchers at Ohio State" (p. 23). In accordance, we focused on the Ohio State dimensions of Consideration and Initiating Structure.

universally accepted (see Bass, 1990, and Fleishman, 1995, for reviews).

On a conceptual level, House's (1977) theory of charismatic leadership shifted the focus of leadership research to higher end constructs. Further, with the emergence of implicit leadership theories, researchers started to question the internal validity of behavioral leadership survey measures in general, and measures of Consideration and Structure were no exception (Rush, Thomas, & Lord, 1977). In addition, from an implicit theory perspective, the shift toward higher end leadership constructs perhaps has made researchers realize that constructs such as charisma and vision are more central to their own implicit theories of leadership than constructs such as Consideration and Structure, and this realization was then reflected in the leadership constructs and processes investigated in their research.

In sum, research on Consideration and Structure has been criticized on both methodological and conceptual grounds. Perhaps as a result of these criticisms, Consideration and Initiating Structure are widely believed to be of limited validity and little direct utility in contemporary leadership thinking and research. To be sure, these constructs are recognized as an important part of the lineage of present day leadership research (House & Podsakoff, 1994). However, in most comprehensive reviews of the leadership literature, one is left with the impression that these constructs are archaic. The Ohio State studies, contemporary thinking argues, are to be remembered only to better understand where we are today; Consideration and Structure are historical concepts that, due to various methodological and conceptual deficiencies and failed efforts at validation, led to other advancements that are of greater current relevance. One can readily find quotations in the literature that reflect this viewpoint:

> The results show a predominance of low to moderate correlations . . . there is as yet almost no evidence on the predictive validity of "Consideration" and "Initiating Structure" (Korman, 1966, p. 360).
>
> Researchers have not been able to establish a consistent link between task and relationship behaviors and outcomes such as morale, job satisfaction, and productivity (Northouse, 1997, p. 45).
>
> The results have been weak and inconsistent for most criteria of leadership effectiveness (Yukl, 1998, p. 49).
>
> Overall, the research based on a two-factor conceptualization of leadership behavior has added little to our knowledge about effective leadership (Yukl & Van Fleet, 1992, p. 156).
>
> Unfortunately, there was no pattern of leader behavior which was found to be consistently associated with subordinates' satisfaction or any criteria of supervisor or manager effectiveness (House & Aditya, 1997).

Although not everyone is as pessimistic about the contemporary relevance of the Consideration and Structure factors (notably, Fleishman, 1998), the above quotations suggest that most researchers consider these leadership behaviors to be outmoded.

Given the inconsistencies in the validity of Consideration and Structure that were concluded to exist, researchers have argued that a way to advance the literature is to investigate situational moderators (Korman, 1966). Yukl and Van Fleet (1992) effectively summarized this thinking by commenting, "The relative importance of specific forms of [Consideration and Initiating Structure] varies from situation to situation" (p. 157). Kerr and Schriesheim (1974) documented the situational moderators that had been studied at that time, which include subordinate factors (e.g., follower job knowledge), supervisor factors (e.g., upward influence), and task factors (e.g., autonomy). Given the assumed lack of support for the main effects of Consideration and Structure, in the 1970s many researchers turned to the investigation of situational factors that might moderate their effects. One of the most popular theories of leadership in the 1970s and 1980s—path–goal theory (House, 1971)—was based on the premise that factors moderate and mediate the effect of Consideration and Structure on outcomes.

Although we do not deny that situational moderators may exist and may prove productive in future research, we would note two points. First, theories that were explicitly developed to test moderators of Consideration and Initiating Structure have not received clear support in the literature. For example, support for path–goal theory (House, 1971) was labeled as "mixed" by the creator of the theory himself (House & Podsakoff, 1994, p. 52); Wofford and Liska (1993), on the basis of a meta-analytic review, described the support as "inconsistent" and "marginal." Similarly, support for the "neutralizers" in Kerr and Jermier's (1978) substitutes for leadership theory, which sought to explain why the leadership behaviors are of modest validity, was described as "not . . . all that more supportive than the tests of path–goal moderators" (House & Podsakoff, 1994, p. 53). As with path–goal theory, this is a conclusion borne out by meta-analytic evidence in which the substitutes main effects appear to exert more substantial influences than the moderating effects as predicted by the theory (Podsakoff, MacKenzie, & Bommer, 1996). Second, before one concludes that there are situational moderators based on the premise of inconsistent results, it is important to first document that these inconsistencies do in fact exist.

The purpose of the present study is to provide a comprehensive examination of the validity of Consideration and Initiating Structure in leadership research. We examine every available study and perform a meta-analytic review that (a) estimates the overall validity of Consideration and Initiating Structure in predicting both subjective (e.g., satisfaction) and objective (e.g., performance) leadership criteria, (b) estimates the correlation between Consideration and Initiating Structure, and (c) determines whether validities vary by measure and, if so, by how much. Given the limitations of narrative reviews (potentially misleading conclusions, over- or underweighting of particular studies, failure to consider study artifacts and measurement error; Hunter & Schmidt, 1990), it is worthwhile to consider the actual effect sizes once sampling and measurement error have been taken into account. At the very least, this review will provide a proper historical perspective on these constructs. More fundamentally, it could alter understanding of these constructs and suggest new areas for research. In the next section of the article, we present hypotheses regarding Consideration and Initiating Structure.

Review of the Literature and Hypothesized Relations

Although there may be considerable pessimism regarding the validity of Consideration and Initiating Structure, one conclusion that has often been drawn from an examination of the empirical

data is that Consideration correlates more strongly with follower satisfaction and Initiating Structure correlates more strongly with performance or effectiveness. Both Bass (1990) and Yukl (1998), for example, noted that the clearest set of results regarding the validity of the two behaviors is the correlation of Consideration with satisfaction. This pattern of associations fits well with the conceptual nature of the constructs. As noted by Halpin (1957b), one would expect leaders high on Initiating Structure to be more effective at meeting role expectations, whereas one would expect followers to prefer (and thus be more satisfied by) leaders who are considerate. Considerate leaders are empathetic (Fleishman & Salter, 1963), and thus should be skilled at sensing and subsequently satisfying the needs of their followers. Because the orientation of structuring leaders is toward the task (Bass, 1990), they should be more effective at producing performance outcomes. Although the strongest support can be offered for the expectation that Consideration correlates more strongly with follower satisfaction, whereas Initiating Structure correlates more strongly with performance and leader effectiveness (Hypothesis 3), in order to test their effects across all criteria, we hypothesized main effects for Consideration (Hypothesis 1) and Structure (Hypothesis 2) for all the criteria.

Hypothesis 1: Consideration will be positively related to (a) follower satisfaction (leader satisfaction, job satisfaction) and (b) leader performance or effectiveness (leader job performance, group–organization performance, leader effectiveness).

Hypothesis 2: Initiating Structure will be positively related to (a) follower satisfaction (leader satisfaction, job satisfaction) and (b) leader performance or effectiveness (leader job performance, group–organization performance, leader effectiveness).

Hypothesis 3: Compared with Initiating Structure, Consideration will be more strongly related to follower satisfaction (leader satisfaction, job satisfaction), whereas, compared with Consideration, Initiating Structure will be more strongly related to leader performance or effectiveness (leader job performance, group–organization performance, leader effectiveness).

In the literature, four measures of Consideration and Initiating Structure have been widely used: The Leader Behavior Description Questionnaire (LBDQ; Halpin, 1957a), the LBDQ, Form XII (LBDQ-XII; Stogdill, 1963), the Supervisory Behavior Description Questionnaire (SBDQ; Fleishman, 1989b), and the Leader Opinion Questionnaire (LOQ; Fleishman, 1989a). The LOQ is the most unique of these measures in that it asks leaders to indicate how often they believe they *should* (vs. actually *do*) engage in considerate and structuring behaviors. A common theme in the literature is that the specific measures correlate differently with outcomes (House & Aditya, 1997). Schriesheim and Kerr (1974) concluded that the LBDQ-XII is the best measure of Consideration and Initiating Structure. Fleishman (1995) disagreed, arguing that the SBDQ and LOQ were better measures. Irrespective of which measure is superior, in light of past research we expect validities to vary by measure.

Hypothesis 4: Validities of (a) Consideration and (b) Initiating Structure will vary by measure.

The correlation between Consideration and Initiating Structure has been the subject of much debate. The concern with the independence of these dimensions can be traced to two sources. First, orthogonality of the dimensions was often claimed in the literature; orthogonality suggests that the dimensions are wholly independent. Second, perhaps the most popular practical application of the leader behavior approach—the managerial grid—is based on the assumption of orthogonality. Weissenberg and Kavanagh (1972) reviewed the literature on the relationship between measures of Consideration and Structure and concluded that the two dimensions "are not always empirically independent as stated and implied" (p. 127). Bass (1990) agreed, noting, "Initiation and Consideration should be independent, but such is not the case" (p. 515). Weissenberg and Kavanagh (1972) further argued that the relationship between Consideration and Initiating Structure varied depending on the measure used. Fleishman (1995) also noted that the Consideration–Structure correlation could be expected to vary by measure, with the LOQ and SBDQ displaying lower intercorrelations. Thus, overall, we believe that Consideration and Initiating Structure will be positively related, but we also believe that the specific measure used will explain variability in these correlations across studies.

Hypothesis 5: There will be a significant (nonzero) correlation between Consideration and Initiating Structure.

Hypothesis 6: The intercorrelation of Consideration and Initiating Structure will vary by measure.

Method

Literature Search

To identify all possible studies of the relationships between Consideration, Initiating Structure, and relevant organizational criteria, we searched the PsycINFO database (1887–2001) for studies (articles, book chapters, dissertations, and unpublished reports) that referenced two general keyword categories. First, we searched for studies that referenced the terms *Consideration* or *Initiating Structure*, as well as related terms such as *employee orientation*, *production-centered*, or the names of the measures used to assess Consideration and Initiating Structure (e.g., *LBDQ*). We combined these keyword searches with search results for leadership criteria, including keywords such as *leader effectiveness*, *job satisfaction*, *group performance*, and *motivation*. In addition to the electronic searches, we examined the reference lists of comprehensive reviews of the literature (e.g., Bass, 1990; Fleishman, 1998; Fleishman, Mumford, Zaccaro, & Levin, 1991; Kerr & Schriesheim, 1974; Schreisheim, House, & Kerr, 1976; Schriesheim & Kerr, 1974) and included articles that appeared relevant.

Our search efforts resulted in the identification of 18 articles referenced in literature reviews or meta-analyses on relevant topics, and 1,180 abstracts identified by means of electronic searches (878 journal articles and 302 dissertations). In reviewing the abstracts, we eliminated studies that clearly did not include primary data (such as qualitative studies or reviews) and studies that did not appear to measure leadership. Further, we eliminated studies that did not appear to measure a relevant criterion such as leader job performance or motivation.

For the remaining 165 articles and 36 doctoral dissertations, we examined each study to determine whether it contained the information needed

to calculate validities. Several studies were excluded because they reported percentages or proportions or means with no standard deviations, or because they reported other measures of association that could not be converted to correlations. In total, 130 studies met the criteria for inclusion in the database (117 journal articles and 13 dissertations). These studies reported a total of 593 correlations computed from 457 independent samples. We formed three distinct databases in order to estimate the meta-analytic correlations between (1) Consideration and organizational criteria, (2) Initiating Structure and organizational criteria, (3) Consideration and Initiating Structure. Table 1 shows the number of primary correlations included in each meta-analytical database together with the number of studies that reported these correlations and the total number of independent samples that were used to compute the correlations.

Meta-Analytic Procedures

Using the methods of Hunter and Schmidt (1990), we conducted three main meta-analyses to estimate the correlations among Consideration, Initiating Structure, and organizational criteria. We corrected each primary correlation for attenuation because of unreliability in both the predictor and the criterion, and then we computed the sample-weighted mean of these corrected correlations. To estimate parameters describing the variability of the meta-analytical estimates and the confidence in these estimates, the variance of the observed individual estimates was corrected for the effects of both sampling and differential measurement error. When authors of original studies reported the internal consistency reliability for the measures used to compute the primary correlations—as was the case for the majority of correlations—we used this value to correct the observed correlation for attenuation. When reliabilities for predictor and criteria measures were not reported in the original studies, we averaged the reliabilities reported in the studies that did provide such estimates for similar measures and used these mean reliability values to correct the primary correlations.[2]

In addition to reporting point estimates for corrected correlations, it is also important to describe variability in these estimates. In accordance, we report 80% credibility intervals and 90% confidence intervals around the estimated population correlations. We believe it is important to report both confidence and credibility intervals because each tells us different things about the nature of the estimates. Confidence intervals provide an estimate of the variability around the estimated mean corrected correlation that is due to sampling error: a 90% confidence interval around a positive point estimate that excludes zero indicates that if the estimation procedures were repeated a large number of times, the point estimate would be larger than zero in 95% of the cases (the other 5% would be zero or negative). Credibility intervals provide an estimate of the variability of individual correlations across studies: an 80% credibility interval excluding zero indicates that 90% of the individual correlations in the meta-analysis excluded zero (for positive correlations, less than 10% are zero or negative, and 10% lie at or beyond the upper bound of the interval). Thus, confidence intervals estimate variability in the mean correlation, whereas credibility intervals estimate variability in the individual correlations across the studies.

Table 1
Number of Studies, Samples, and Primary Correlations Included in the Meta-Analyses

Relationship	Studies	Samples	Correlations
Consideration–Criteria	103	154	209
Initiating Structure–Criteria	99	151	203
Consideration–Initiating Structure	78	166	181

Note. In some cases, the number of correlations appearing in this table is greater than the number of correlations appearing in subsequent tables. To avoid violating the independence assumption, whereby only one correlation can be analyzed from each sample (Hunter & Schmidt, 1990), if a study reported correlations of Consideration with motivation and with leader satisfaction, only one of these correlations could be used in the overall analysis reported in Table 2. In such a case, the criterion was selected at random for the overall analysis.

Moderator Analyses

We divided the primary estimates into categories according to the expected moderator variables. We conducted separate meta-analyses for each of the categories to estimate the true correlations for the categories delimited by moderator variables. Meta-analytical evidence for the presence of moderators requires that (a) true estimates are different in the categories formed by the potential moderator variable, and (b) the mean corrected standard deviation within categories is smaller than the corrected standard deviation computed for combined categories. To test for the presence of moderator effects, as recommended by Sagie and Koslowsky (1993), we report the Q statistic (Hunter & Schmidt, 1990, p. 151), which tests for homogeneity in the true correlations across studies. A significant Q statistic (which is approximately distributed as a chi-square, χ^2), indicates the likelihood that moderators explain variability in the correlations across studies. If a significant Q statistic across moderator categories becomes nonsignificant within a moderator category, it suggests that the moderator explains a significant amount of the variability in the correlations across the moderator categories.

Results

Table 2 provides the results of the overall meta-analyses, which estimate the validity of Consideration and Initiating Structure, collapsed across all criteria and measures. The meta-analysis results, on the basis of 163 correlations for Consideration and 159 correlations for Initiating Structure, revealed that both Consideration and Structure evince nonzero correlations across the leadership criteria. Specifically, because the 90% confidence intervals excluded zero, we can be confident that the mean Consideration correlation and the mean Initiating Structure correlation are nonzero. Furthermore, because the 80% credibility intervals exclude zero, more than 90% of the individual correlations for Consideration and Structure are greater than zero (a maximum of 10% lie at or beyond the upper bound of the interval). Despite their significance, the average correlation for Consideration ($\hat{\rho} = .48$) is appreciably larger than the average correlation for Initiating Structure ($\hat{\rho} = .29$).

[2] As it was pointed out by an anonymous reviewer, compared with Hunter and Schmidt (1990), the procedure developed by Raju, Burke, Normand, and Langlois (1991) estimates the standard deviation of the true-score correlations more accurately (because it takes into account the sampling errors associated with sample-specific estimates of the reliabilities of the scores on the predictor and criterion measures when computing the sampling variance of the corrected correlations). However, for our specific analyses, the choice of method (Raju et al., 1991, vs. Hunter & Schmidt, 1990) made little difference in the analyses. For the overall analyses of the correlation of Consideration and Initiating Structure with criteria, for example, the standard deviations of the true score correlations differed by only .0099 (.2549 for Raju et al., 1991, vs. .2648 for Hunter & Schmidt, 1990) and .0112 (.2062 for Raju et al., 1991, vs. .2175 for Hunter & Schmidt, 1990), respectively.

Table 2
Relationship of Consideration and Initiating Structure to Leadership

Leader behavior	k	N	$\bar{r}$	$\hat{\rho}$	$SD_{\hat{\rho}}$	80% CV lower	80% CV upper	90% CI lower	90% CI upper	Q
Consideration	163	20,963	.41	.48	.26	.14	.82	.44	.52	1,500.83**
Initiating Structure	159	20,431	.24	.29	.21	.02	.56	.26	.33	775.09**

Note. k = number of correlations; N = combined sample size; $\bar{r}$ = mean observed correlation; $\hat{\rho}$ = estimated true score correlation; $SD_{\hat{\rho}}$ = standard deviation of true score correlation; CV = credibility interval; CI = confidence interval; Q = test for homogeneity in the true correlations across studies.
** $p < .01$.

Because the overall correlations reported in Table 2 are fairly strong and the credibility intervals exclude zero, the overall validities for Consideration and Initiating Structure are interpretable. However, for both Consideration and Structure, the Q statistic is significant, meaning that there is significant variability in the correlations even after taking measurement and sampling error into account. Thus, moderators of the overall relationships likely exist. In accordance, we specified analysis of variance (ANOVA) models in which the corrected correlation ($\hat{\rho}$) was the dependent variable and the study-level moderators were the explanatory variables. The study-level moderators were as follows: Criteria (which of the six criteria was used in the study), Measure (which of the five measures [LBDQ, LBDQ-XII, LOQ, SBDQ, other] was used in the study), Research Design (cross-sectional or longitudinal), Independence of Data Sources (same source or different source), Study Setting (business, college, military, public sector), and Level of the Leader (supervisory or mid- to upper-level leaders). For Consideration, the ANOVA results revealed that two variables—Criteria and Measure—explained a significant amount of variance in the correlations across studies, with Criteria accounting for 14.7% unique variance ($p < .01$) and Measure accounting for 6.2% unique variance ($p < .05$). As a set, the moderators explained 32.1% of the variance in the Consideration correlations across samples. For Initiating Structure, only Measure accounted for a significant amount of variance. It explained 18.8% ($p < .01$) unique variance in the Initiating Structure correlations across samples. Overall, the moderators explained 24.7% of the variance in the Structure correlations across samples. Thus, Hypothesis 4 was supported by the results—validities for both Consideration and Structure varied significantly by measure.

Given the potential importance of the moderators, especially Criteria and Measure, we conducted separate meta-analyses within each moderator variable category. Table 3 provides the meta-analysis results for Consideration and Initiating Structure, broken down by the six criteria: follower job satisfaction, follower satisfaction with the leader, follower motivation, leader job performance, group–organization performance, and leader effectiveness. As the table shows, both leadership behaviors have nonzero relations with the criteria. The only exceptions are the correlation of Consideration with leader job performance and the correlation of Initiating Structure with follower job satisfaction. In both of these cases, the average correlation is distinguishable from zero (the 90% confidence interval excludes zero), but the 80% credibility interval includes zero, indicating that more than 10% of the correlations are zero or negative. In general, however, the behaviors are significantly related to the criteria, supporting Hypotheses 1 and 2.

In Hypothesis 3, we predicted that leader Consideration will be more strongly related to criteria that reflect follower satisfaction (leader satisfaction, job satisfaction) whereas Initiating Structure will be more strongly related to criteria that reflect leader performance or effectiveness (leader job performance, group–

Table 3
Relationship of Consideration and Initiating Structure to Leadership Criteria

Criterion	Consideration k	Consideration N	Consideration $\bar{r}$	Consideration $\hat{\rho}$	Initiating Structure k	Initiating Structure N	Initiating Structure $\bar{r}$	Initiating Structure $\hat{\rho}$	Z
Follower job satisfaction	76	11,374	.40	.46[a,b]	72	10,317	.19	.22[a]	20.49*
Follower satisfaction with leader	49	7,871	.68	.78[a,b]	49	8,070	.27	.33[a,b]	44.15*
Follower motivation	11	1,067	.36	.50[a,b]	12	1,041	.26	.40[a,b]	2.94*
Leader job performance	25	2,330	.18	.25[a]	22	2,085	.19	.24[a,b]	.36
Group–organization performance	27	2,008	.23	.28[a,b]	27	2,079	.23	.30[a,b]	.73
Leader effectiveness	20	1,605	.39	.52[a,b]	20	1,960	.28	.39[a,b]	4.11*

Note. Because of space limitations, the Q statistic is not reported but was significant in all cases except for the correlations of Consideration and Initiating Structure with follower motivation. k = number of correlations; N = combined sample size; $\bar{r}$ = mean observed correlation; $\hat{\rho}$ = estimated true score correlation; Z = Steiger (1980) statistic for differences in correlations of Consideration and Initiating Structure with the criteria.
[a] 90% confidence interval excluded zero.
[b] 80% credibility interval excluded zero.
* $p < .05$.

organization performance, leader effectiveness). To test this hypothesis, we conducted a significance test on the correlations to determine if differences were statistically significant. Because the correlations reported in this particular moderator analysis are not independent, our tests for equality must take the dependency of the correlations into account. For example, the correlation between Consideration and follower job satisfaction is not independent of the correlation between Initiating Structure and follower job satisfaction because these correlations are computed from the same population. Thus, to test for equality of dependent correlations, we used the technique described by Steiger (1980) to test the hypotheses, H_0: $\hat{\rho}_{y-c} = \hat{\rho}_{y-is}$. The Steiger test takes dependency into account and allows us to calculate a test statistic with a Student's t distribution.

Consistent with Hypothesis 3, Consideration is more strongly related to follower satisfaction (and motivation as well), whereas Initiating Structure is slightly more strongly related to criteria that reflect leader performance. Contrary to Hypothesis 3, Consideration was more strongly related to leader effectiveness, although both correlations were moderately strong in magnitude. Although most of the Q statistics are still statistically significant, their size is greatly reduced, suggesting the criterion does reduce much of the unexplained variability in the correlations.

Table 4 provides the meta-analysis results for both behaviors, broken down by measure. As revealed in the table, with a sole exception, the validities of the specific measures generalize across the population of primary studies. Specifically, in all cases except the SBDQ measure of Initiating Structure, both the 90% confidence intervals and the 80% credibility intervals exclude zero. There were two cases in which the validities did vary appreciably. The LOQ-Consideration measure and the SBDQ-Initiating Structure measure both displayed lower correlations than the other measures. In the other cases, validities did not appear to vary much by measure (excepting these two cases, the average deviation in validity was only .01 for Consideration and .04 for Initiating Structure).

One might wonder whether validities should be nested hierarchically, specifically Measure within Criteria. Though there were sufficient sample sizes, in many cases, to perform these nested analyses, there was little reason to believe that validities were interactive. In the ANOVA models explaining variance in Consideration and Initiating Structure, a Criteria × Measure interaction was specified. In neither case did this interaction explain a significant amount of variance in the validities. Furthermore, when separate meta-analyses were conducted by Measure within each criterion group, the results did not appear to vary any more than they did in the separate moderator analyses.

One hierarchical moderator analysis we performed was to investigate whether research design (cross-sectional vs. longitudinal) or method variance (common methods vs. no common methods) explained differences in the validities by measure (e.g., whether the average validity of LOQ measures of Consideration is lower than the other validities because studies using this measure employed more longitudinal designs that minimized common method variance). Specifically, we performed a hierarchical moderator analysis that computed the validity of each measure within each of the four moderator categories (cross-sectional, longitudinal, common methods, and no common methods). The results of this hierarchical moderator analysis are provided in Table 5. In general, the results show that the effects of research design and method variance do not vary substantially by measure. With few exceptions, when a research design effect or method effect was found for one measure, it was found for the others. For example, for all four measures (LBDQ, LBDQ-XII, LOQ, SBDQ), the Consideration validities were significantly higher when leadership behaviors and outcomes were measured with common methods than when there were no common methods. Similarly, for Initiating Structure, none of the cross-sectional versus longitudinal correlations were significantly different for any of the measures. Thus, overall, it does not seem that research design or method variance substantially explains the different validities obtained by the measures of Consideration and Structure.

Although less powerful than criterion type and measure, the results for the other moderators are provided in Table 6. With a single exception (Initiating Structure with longitudinal designs), the validities generalized across studies, and in all cases, the average validities were distinguishable from zero. To determine whether validities varied significantly across moderator categories,

Table 4
Validities of Consideration and Initiating Structure by Measure

Measure	Consideration					Initiating Structure				
	k	N	$\bar{r}$	$\hat{\rho}$	Q	k	N	$\bar{r}$	$\hat{\rho}$	Q
LBDQ	44	4,401	.43	.51[a,b]	277.79**	43	4,360	.29	.37[a,b]	211.21**
LBDQ-XII	86	13,110	.47	.54[a,b]	1,168.84**	86	12,945	.27	.32[a,b]	372.38**
LOQ	14	772	.24	.34[a,b]	31.04**	14	772	.27	.40[a,b]	41.26**
SBDQ	38	5,349	.46	.54[a,b]	536.50**	38	5,202	.05	.07	181.03**
Other–unidentified measure	27	2,655	.43	.52[a,b]	191.06**	22	2,305	.29	.37[a,b]	111.47**

Note. LBDQ = Leader Behavior Description Questionnaire; LBDQ-XII = Leader Behavior Description Questionnaire, Form XII; LOQ = Leader Opinion Questionnaire; SBDQ = Supervisory Behavior Description Questionnaire; k = number of correlations; N = combined sample size; $\bar{r}$ = mean observed correlation; $\hat{\rho}$ = estimated true score correlation; Q = test for homogeneity in the true correlations across studies.

[a] 90% confidence interval excluded zero.

[b] 80% credibility interval excluded zero.

** $p < .01$.

Table 5
Effects of Research Design and Common Methods on Validities by Measure

	Cross-sectional (CS) vs. longitudinal (L)		Common methods (CM) vs. no common methods (NCM)	
Measure	$\hat{\rho}_{CS}$	$\hat{\rho}_{L}$	$\hat{\rho}_{CM}$	$\hat{\rho}_{NCM}$
Consideration				
LBDQ	.55	.36	.57	.37
LBDQ-XII	.49	.39	.50	.31
LOQ	.44	.26	.50	.20
SBDQ	.62	.53	.63	.31
Other	.58	.16	.60	.35
Structure				
LBDQ	.38	.35	.41	.30
LBDQ-XII	.34	.23	.34	.20
LOQ	.45	.33	.52	.27
SBDQ	.01	.17	.04	.14
Other	.39	.34	.25	.51

Note. $\hat{\rho}$ = estimated true score correlation; LBDQ = Leader Behavior Description Questionnaire; LBDQ-XII = Leader Behavior Description Questionnaire, Form XII; LOQ = Leader Opinion Questionnaire; SBDQ = Supervisory Behavior Description Questionnaire.

we used the Quiñones, Ford, and Teachout (1995) Z test.[3] For Consideration, the average correlation from cross-sectional research designs was significantly higher than the average correlation from longitudinal designs ($Z = 2.23$, $p < .05$), but for Initiating Structure, the difference between the two designs was not significant ($Z = 0.78$, *ns*). When both Consideration and the leadership criterion were measured by the same source, the average correlation was significantly higher than when measured by different sources ($Z = 3.73$, $p < .01$). The independence of data sources, however, did not affect the average correlation for Initiating Structure. With the exception of a difference between the average correlation for Consideration in business versus in public sector settings ($Z = -2.16$, $p < .05$), no other study setting moderated the average validity of Consideration or Initiating Structure. Further, there did not exist any significant differences in the average validity for Consideration or Initiating Structure based on the level of the leader within the organization studied.

Finally, although not provided in the tables, we also investigated whether validities for Consideration and Structure varied by whether the correlation came from a published or unpublished (dissertation) source. Overall, the validities were as follows: Consideration-published, $\hat{\rho}_{p} = .47$ ($k = 149$); Consideration-unpublished, $\hat{\rho}_{u} = .54$ ($k = 14$); Structure-published, $\hat{\rho}_{p} = .30$ ($k = 145$); Structure-unpublished, $\hat{\rho}_{u} = .30$ ($k = 14$). Using the Quiñones et al. (1995) Z test, the published–unpublished differences in validity were not significantly different (Consideration: $\hat{\rho}_{p}$ vs. $\hat{\rho}_{u}$, $Z = -1.16$, *ns*; Structure: $\hat{\rho}_{p}$ vs. $\hat{\rho}_{u}$, $Z = -.01$, *ns*), suggesting that the overall validities do not differ between published and unpublished sources.

Table 7 provides results for the meta-analysis on the relationship between Consideration and Initiating Structure across the 166 samples where we were able to locate such a correlation. As is revealed in the table, the average corrected correlation is relatively weak ($\hat{\rho} = .17$). Although this average corrected correlation is distinguishable from zero, the 80% credibility interval included zero, indicating that in more than 10% of the samples, the Consideration–Structure correlation was zero or negative. Indeed, an examination of the correlations in the database revealed that the correlation was zero or negative in 35% of the samples. Thus, Hypothesis 5 (a significant Consideration–Structure correlation) received mixed support. Furthermore, the very large and significant Q statistic suggests that there are factors that moderate this relationship. Hypothesis 6 predicted that the measure used would moderate the Consideration–Structure relationship. In accordance, as with previous analyses, we specified an ANOVA model in which measure was used to explain variability in the Consideration–Structure correlations across samples. The ANOVA results revealed that the measure of Consideration and Structure explained a large and significant amount—52.7% ($p < .01$)—of the variance in the correlations.

Given that the measure used in studies explained much of the variability in the Consideration–Structure correlation across samples, we estimated separate meta-analyses by the measure used in the study. The results of these meta-analyses broken down by measure are provided in Table 8. As revealed in the table, the Consideration–Structure correlations varied substantially by measure. In two cases (LBDQ and LBDQ-XII), the average correlations are moderately to strongly positive, in two cases (LOQ and SBDQ), the average correlations are weakly negative, and for other–unidentified measures, the average correlation is moderately positive. Using the Quiñones et al. (1995) test, the average Consideration–Structure correlation based on the LBDQ measure

[3] The difference between the Steiger (1980) test and the Quiñones et al. (1995) Z test is that the former tests for differences in dependent correlations (a comparison of correlations from the same sample based on the same participants), whereas the latter tests for differences in independent correlations (a comparison of correlations from two independent samples or two distinct moderator categories).

Table 6
Other Moderator Analysis Results

Moderator	Consideration				Initiating Structure			
	k	N	$\bar{r}$	$\hat{\rho}$	k	N	$\bar{r}$	$\hat{\rho}$
Research design								
Cross-sectional	104	13,419	.40	.46[a,b]	104	13,346	.25	.31[a,b]
Longitudinal	43	3,615	.28	.36[a,b]	38	3,135	.20	.26[a]
Independence of data sources								
Same	98	13,260	.41	.47[a,b]	92	12,625	.24	.29[a,b]
Different	52	4,099	.25	.32[a,b]	53	4,181	.23	.30[a,b]
Study setting								
Business	79	10,305	.37	.43[a,b]	78	10,062	.24	.29[a,b]
College	11	1,331	.46	.56[a,b]	10	1,297	.19	.25[a,b]
Military	11	649	.30	.40[a,b]	12	678	.31	.46[a,b]
Public sector	51	6,616	.46	.53[a,b]	47	6,311	.23	.28[a,b]
Level of leader								
Supervisory	124	16,574	.40	.47[a,b]	121	16,215	.23	.28[a,b]
Mid or upper	25	2,244	.45	.54[a,b]	23	2,050	.25	.33[a,b]

Note. k = number of correlations; N = combined sample size; $\bar{r}$ = mean observed correlations; $\hat{\rho}$ = estimated true score correlation.
[a] 90% confidence interval excluded zero.
[b] 80% credibility interval excluded zero.

is significantly stronger than the average correlation based on the LOQ or SBDQ measures ($Z = 10.33$, $p < .01$, and $Z = 7.22$, $p < .01$, respectively). Similarly, the average correlation based on the LBDQ-XII is significantly stronger than the average correlation based on the LOQ or SBDQ measures ($Z = 13.56$, $p < .01$, and $Z = 8.26$, $p < .01$, respectively). Thus, Hypothesis 6 is supported. The correlation between Consideration and Initiating Structure does vary depending on the measure used to assess these constructs.

Discussion

Despite their predominance in the literature a generation ago, Consideration and Initiating Structure have long been dismissed as important influences on leadership effectiveness. House and Podsakoff (1994) labeled the body of findings as "relatively disappointing" (p. 50). Yukl (1998) labeled the results as "weak and inconsistent" (p. 49). Northouse (1997) remarked that the "preponderance of the research in this area was inconclusive" (p. 38). Yukl and Van Fleet (1992) concluded, "The long fixation on consideration and initiating structure appears to have come to an end" (p. 159). Even models that sought to delineate the conditions under which Consideration and Initiating Structure were relevant—path–goal theory and substitutes for leadership—have not received strong support in the literature (Dionne, Yammarino, Atwater, & James, 2002; Wofford & Liska, 1993). Since 1980, there have been only a handful of empirical journal articles on Consideration or Initiating Structure, and there have been none since 1987. A recent 500-page book on leadership revealed only a single, passing citation to the Ohio State studies (Mathieu, 2001, p. 454). Similarly, Consideration and Structure were completely absent from a recent 53-chapter edited book on leadership (Hickman, 1998). These behaviors seem to be in danger of being viewed as

Table 7
Meta-Analysis of Correlation Between Consideration and Initiating Structure

Statistic	Value
Number of correlations (k)	181
Combined sample size (N)	26,295
Average uncorrected correlation ($\bar{r}$)	.14
True score correlation ($\hat{\rho}$)	.17
Standard deviation of true score correlation ($SD_{\hat{\rho}}$)	.33
80% credibility value lower	−.25
80% credibility value upper	.59
90% confidence interval lower	.12
90% confidence interval upper	.22
Q	1,826.28**

Note. Q = test for homogeneity in the true correlations across studies.
** $p < .01$.

Table 8
Consideration–Initiating Structure Correlation by Measure

Measure	k	N	$\bar{r}$	$\hat{\rho}$	Q
LBDQ	37	5,138	.36	.44[a,b]	315.50**
LBDQ-XII	32	5,806	.37	.46[a,b]	153.27**
LOQ	78	10,051	−.06	−.08[a]	242.95**
SBDQ	18	2,858	−.07	−.08	120.02**
Other–unidentified measure	16	2,442	.22	.28[a]	130.66**

Note. LBDQ = Leader Behavior Description Questionnaire; LBDQ-XII = Leader Behavior Description Questionnaire, Form XII; LOQ = Leader Opinion Questionnaire; SBDQ = Supervisory Behavior Description Questionnaire; k = number of correlations; N = combined sample size; $\bar{r}$ = mean observed correlation; $\hat{\rho}$ = estimated true score correlation; Q = test for homogeneity in the true correlations across studies.
[a] 90% confidence interval excluded zero.
[b] 80% credibility interval excluded zero.
** $p < .01$.

historical artifacts in leadership research—important artifacts—but artifacts of little contemporary relevance nonetheless.

This denouement for the Ohio State leadership behaviors, however, may be premature. The results of the present quantitative review revealed that both Consideration and Initiating Structure have important main effects on numerous criteria that most would argue are fundamental indicators of effective leadership. It is striking how the validities for each behavior generalized—across criteria, across measures, and even over time and across sources. Of course, these behaviors are not all there is to solving the mysteries of leadership effectiveness. However, just as surely, the results do suggest that these behaviors—Consideration and Initiating Structure—are important pieces in the leadership puzzle.

In interpreting the meaning of this meta-analytic review, one should note that these results are not the result of "fantastic" meta-analytic corrections. We simply corrected the leader behavior and criterion results on the basis of coefficient alpha reliability estimates, which, in most cases, were at acceptable levels (the average reliability, across all behaviors and criteria, was $\alpha = .81$). To be sure, there is an argument that interrater reliability estimates would be the more appropriate corrections (see Murphy & DeShon, 2000, and Schmidt, Viswesvaran, & Ones, 2000). Furthermore, one could even argue that the correlations should be corrected for range restriction because it is likely that individuals are selected to leadership positions on the basis of the degree to which they are considerate and structuring. It is likely that performing these two corrections (using interrater reliability and correcting estimates for range restriction) would have made the corrected correlations considerably higher. The point here is that these results are not news because our corrections transformed weak correlations into moderately strong ones. The correlations were always there, which makes it unfortunate that they were overlooked.

Why were the correlations overlooked? How could the conclusions of reviewers of this literature have been so wrong? One means of reconciling the pessimistic conclusions of past research with the more positive results of this review is to examine the history of other areas. In the area of personality and job performance, prior to Barrick and Mount's (1991) meta-analysis, researchers had cautioned against the use of personality tests (Mount & Barrick, 1998). Similarly, meta-analysis dramatically altered the conclusions in the area of intelligence and integrity tests (see Schmidt & Hunter, 1998, for a review). Although there is nothing wrong with qualitative reviews per se, they are subject to various errors, including the fact that subjective accounting of study results often leads to inaccurate conclusions. As a specific case in point, Kerr and Schriesheim (1974) concluded, "LOQ studies have typically yielded much less significant relationships between predictors and criteria than have those using the LBDQ or SBDQ" (p. 561). Yet, our results show that although for Consideration this conclusion is accurate, for Initiating Structure this conclusion clearly is inaccurate. Perhaps the conclusions in past qualitative reviews were not tested meta-analytically because the literature had already been pronounced as dead before the advent of meta-analysis in industrial–organizational psychology. Why meta-analyze a literature, researchers may have asked, that had long been obsolete?

That Consideration and Initiating Structure are far from obsolete—that they have significant main effects—does not deny the possibility that these effects are moderated by other variables. Indeed, our moderator analyses were successful in revealing several factors that appear to affect the validity of Consideration and Initiating Structure. There may be other moderators. Furthermore, our results do not fully address concerns that have been expressed about the measures of Consideration and Initiating Structure. For example, concerns have been raised about skew in responses to the items, and there are concerns about leniency, social desirability, and halo effects (see Schriesheim & Kerr, 1974). That the measures correlate with "hard" performance and effectiveness criteria ameliorates some, but certainly not all, of these concerns. We further note that measures of transformational leadership likely suffer from many of the same limitations. Indeed, a meta-analysis of the transformational leadership literature (Lowe, Kroeck, & Sivasubramaniam, 1996) appears to be dominated by correlations suffering from common-source limitations.

The measure used in leadership studies did moderate the validity of both Consideration and Initiating Structure. Although Schriesheim and Kerr (1974) favored the LBDQ-XII, and Fleishman (1995) preferred the LOQ or SBDQ, the original LBDQ and the LBDQ-XXII have the highest validities averaged across Consideration and Structure. That is, for Consideration, the LOQ was less valid than the other three measures and, for Initiating Structure, the SBDQ was less valid than the other three. On the other hand, in only one of eight possible relationships involving the four measures of the two behaviors was the validity not generalizable (the LOQ measure of Initiating Structure). Thus, the measure of Consideration and Initiating Structure does matter, but only in one case does it matter so dramatically as to destroy the validity of the behavior. Overall, the pattern of correlations is more consistent than has been depicted in previous reviews (e.g., Yukl, 1998).

In general, Consideration exhibited stronger relationships with the criteria than did Initiating Structure. This was especially true with respect to follower satisfaction (follower job satisfaction, follower satisfaction with the leader). Initiating Structure did have slightly stronger relations with group–organization performance. Thus, to some extent, our hypothesis that followers prefer considerate leaders but will perform more effectively for structuring leaders, was supported. On the other hand, Consideration was linked to leader job performance and group–organization performance, and Initiating Structure was linked to leader satisfaction. Both behaviors also were linked to follower motivation and leader effectiveness, with Consideration being somewhat more important.

Whereas the validities of Consideration and Initiating Structure (in correlating with the criteria) do not vary strongly across measures (as a rule), the intercorrelation between the two behaviors does vary considerably across measures. Fleishman (1995) commented that the LOQ and SBDQ Consideration and Initiating Structure intercorrelations are quite small, and our results confirm this observation. The correlations between the LBDQ and LBDQ-XII Consideration and Initiating Structure scales were appreciably larger but not so large as to render the two concepts redundant. For example, the correlations among the transformational leadership dimensions (average $r = .78$), and even between transformational leadership and transactional leadership in the form of contingent reward (average $r = .68$), are far higher than the LBDQ Consid-

eration and Structure correlation (see Lowe et al., 1996, p. 421). The implications of the Consideration–Initiating Structure correlation, whether the LOQ and SBDQ are to be preferred due to their lower Consideration–Structure intercorrelation, and implications of this association for the incremental validity of the scales, are issues that need further analysis in future research.

The validity of Consideration was significantly stronger in cross-sectional designs than in longitudinal designs and when variables were measured by the same source. Perhaps there is a higher social desirably component to Consideration, which would explain why same-source and cross-sectional validities were substantially higher and, perhaps in part, why Consideration correlates so strongly with leader satisfaction. It is interesting to note that these methodological features did not affect the validity of Initiating Structure.

Future Research

Given the validity estimates reported here, further study of Consideration and Initiating Structure in leadership research is warranted. There are two general directions for future research. First, research should address some questions that were unanswered (perhaps due to the virtual abandonment of the topic) in past research. Second, Consideration and Structure have virtually disappeared in contemporary research; these concepts should be integrated with more recent theorizing in leadership research specifically and organizational behavior research more generally. Thus, beyond the areas already suggested, we provide below an agenda for future research.

Mediators of consideration and structure effects. Future research should explore mediators that explain why Consideration and Structure have their effects on the outcomes. As noted by Northouse (1997), past research "has not adequately shown how leaders' styles are associated with performance outcomes" (p. 45). Although many potential mediators could be investigated, an especially promising mediating framework is organizational justice. Justice research has suggested that there are at least three organizational justice dimensions: distributive, procedural, and interactional (Cropanzano, Byrne, Bobocel, & Rupp, 2001). It seems likely that leaders high on Initiating Structure are likely to foster distributive justice. As discussed by Greenberg and Lind (2000), the equity principle is that outcomes should be distributed according to contributions. Leaders high on Initiating Structure define rules regarding work duties, maintain standards, and determine the consequences of goal attainment. It seems likely that followers of such leaders should feel that their rewards are distributed according to the equity rule. Interactional justice is the degree to which individuals are treated with dignity, respect, and sensitivity by authority figures. Because Consideration is "the extent to which a leader exhibits concern for the welfare of the other members of the group" (Bass, 1990, p. 511), leaders high on Consideration should be better at fostering interactional justice. Finally, it appears that both Consideration and Structure are relevant to procedural justice, although in different ways. Considerate leaders should provide higher procedural justice in that they appear to follow several of Leventhal's (1980) rules in that they "put subordinates' suggestions into operation" and "obtain subordinates' approval on important matters before going ahead" (Bass, 1990, p. 511). As for Structure, on the one hand, structuring leaders should foster procedural justice in that they clearly communicate expectations and set clear standards across people. On the other hand, their directive nature may cause followers to feel they are without a voice in how the rules and procedures are established. Clearly, the integration of the Ohio State factors with justice theory is an important area for future research.

Further mediating mechanisms and integration with other literatures. Although formulations of leader–member exchange (LMX) theory suggest that leaders behave differently toward different followers, affect or liking is closely linked to LMX operationalizations (Liden, Sparrowe, & Wayne, 1997). Are considerate leaders more effective, in part, because they are better liked? Liking has been speculated to be a contaminant in the measurement of the behavioral dimensions (Schriesheim & Kerr, 1974). Perhaps it is, instead, a mediating variable that explains why considerate leaders are more satisfying to followers. As for mediators of Initiating Structure effects, perhaps structuring leaders are more effective because they are more likely to set hard and specific goals for their followers. The motivational effects of goal setting are well documented (Locke, 1997), and Kirkpatrick, Locke, and Latham (1996) integrated the goal-setting and leadership literatures. This effort did not, however, include Initiating Structure. Thus, it seems possible that integrating Consideration and Structure with the concepts in LMX and goal-setting theories, respectively, might prove worthwhile. Another mechanism that might mediate the relationship between Consideration and Structure and follower performance concerns followers' beliefs about their performance abilities. That is, because goals are thought to influence self-efficacy (e.g., Locke, 1997), Initiating Structure may contribute to follower feelings of increased efficacy which, in turn, lead to performance (Bandura, 1986). In addition, supportive leader behavior is linked to follower self-confidence in House's (1996) revised theory of path–goal leadership, which suggests that Consideration should also lead to increased follower self-efficacy and performance. Future research should investigate such mediating mechanisms.

Integration with transformational leadership. Transformational–charismatic leadership is the most popular contemporary leadership theory, and one that also is supported by meta-analytic evidence (Lowe et al., 1996). Although transformational leadership is seen as an advance over the Consideration and Initiating Structure dimensions, there has been surprisingly little discussion of the relationship of the Ohio State leader behaviors to transformational leadership. Bass (1999) pointedly argued that a facet of transformational leadership—individualized consideration—was to be distinguished from the Ohio State Consideration factor. However, Seltzer and Bass's (1990) study revealed moderately strong correlations of transformational leadership with Consideration and Initiating Structure. Mindful of Bass's (1999) argument that Consideration and individualized consideration are distinct, we are aware of no empirical research that has fully addressed this issue. How the two frameworks can be integrated, given their validity, is deserving of attention.

Causal ordering. As Bass (1990) noted, a limitation of past research is the inability to ascertain whether "leadership is a cause, a consequence, or a coincidence of group effectiveness, satisfaction, or other valued outcomes" (p. 542). This is a long-standing

criticism of this literature (Korman, 1966), yet with few exceptions there has been little effort to study the causal relationship between Consideration, Structure, and outcomes. Moreover, because the characteristics *understanding*, *caring*, and *concerned*, as well as *decisive*, *directive*, and *organized* are endorsed by individuals as attributes of leaders (Lord, Foti, & De Vader, 1984), it seems possible that implicit theories of leadership may explain the validities of Consideration and Structure. Specifically, individuals may attribute effective leadership by perceiving such leaders as considerate and structuring, irrespective of whether those behaviors actually led to effective leadership.

Construct validity of measures. Do different measures of Consideration and, in particular, Initiating Structure, reflect the same construct and, if not, what differing aspects of leadership do they tap? Is one more construct-valid than the others? There appears to be a trade-off in that the most valid measures across both factors (LBDQ, LBDQ-XII) also have the strongest intercorrelations between the two factors. It seems that the higher overall validity of the LBDQ and LBDQ-XII measures of Consideration and Structure is paid in the coin of higher intercorrelations between the measures. The measurement of the dimensions was a prominent theme in the literature (e.g., Schriesheim & Kerr, 1974) but one that never reached a resolution. Given the results observed in this review, more construct validity work is needed.

Causes of Consideration and Structure. What causes a leader to display considerate and structuring behaviors? It seems possible that agreeable individuals are more considerate, whereas conscientious (and perhaps less open) individuals are more structuring. Fleishman (1957) found that leaders who had high scores on Consideration also had high scores on benevolence, a trait that seems closely aligned to agreeableness. Similarly, Bass (1990, p. 522) reported the results of a study showing that charm was related to Consideration; charming is a characteristic that could easily be associated with extraverts, along with related traits such as witty, flamboyant, and vivacious (Goldberg, 1990). Bass also reported that ascendancy, a trait associated with conscientiousness, is related to Initiating Structure (p. 523). Although these studies predate the popularity of the five-factor model, they do suggest that future research linking personality to Consideration and Structure would be worthwhile.

Nonlinear relations. Fleishman (1995, 1998) has repeatedly noted that the field has practically ignored potential nonlinearities in the validity of Consideration and Structure. In 1998, for example, Fleishman commented, "It is difficult to find other work stimulated by our findings of curvilinearity in leadership–work group performance relationships" (p. 831). Nonlinearities involving the Ohio State factors could take several forms. First, there may be interactions between Consideration and Structure. Fleishman and Harris (1962) found that relatively considerate leaders could increase structure with little or no negative consequences, whereas the same was not true for less considerate leaders. Second, the effects of Consideration and Structure may be curvilinear, as Fleishman and Harris (1962) also found in their study. Finally, although moderating effects (e.g., pressure for production, climate) have been hypothesized and tested (see Fleishman, 1995), the results are piecemeal and replication attempts have been rare. We note that because we found linear relations of Consideration and Structure with the outcomes in this review does not deny the existence or importance of nonlinearities. Linear and nonlinear effects can coexist.

Additional outcomes. It seems possible that the validity of Consideration and Structure might be even more impressive if additional criteria were studied. Perhaps past researchers have not studied the criteria that these measures are most likely to predict. Specifically, it would seem that structuring leaders might result in greater team efficacy and performance, and considerate leaders might produce greater team cohesiveness. Perhaps considerate leaders, by facilitating trust, generate more integrative conflict resolution strategies and greater intra- and interorganizational coordination. Undoubtedly, there are other criteria that are conceptually close to Consideration and Structure to which these behaviors could be related.

Levels of analysis. Finally, although nearly all of the studies included in our meta-analysis were at the individual leader level of analysis, in future research this may not be the case. Most studies were at the individual level of analysis where a follower evaluated the leadership of his or her leader, and these evaluations were related to a criterion corresponding to that leader, such as follower satisfaction with the leader or an independent rating of leader job performance (e.g., House, Filley, & Kerr, 1971). In a few studies (e.g., Parker, 1963), Consideration and Structure were measured with follower ratings averaged across followers. Such designs implicitly assume that the leader effects are equivalent across followers, which may or may not be a valid assumption. Given that followers are nested within leaders and that there may be multiple levels of leaders, in the future it would benefit researchers to use a multilevel framework in investigating the effects of Consideration and Structure.

Limitations

Although the findings reported here make a contribution to the leadership literature, several limitations must be acknowledged. First, from a conceptual perspective, our moderator analyses were limited by the nature of the primary studies in that we could not investigate the effects of theoretically relevant moderators such as those proposed in House's (1996) revised theory of path–goal leadership (e.g., uncertainty, predictability, interdependence, etc.). Second, a methodological limitation stems from the fact that the number of correlations for some of the moderator analyses is relatively small. However, because meta-analytic estimates based on relatively small number of studies are not biased, and only their variability is affected (Schmidt, Hunter, Pearlman, & Hirsh, 1985), such meta-analytic results remain the best available estimates in the literature. Finally, our results, because they are a summary of the extant literature, reflect more correlations relating the behaviors to some criteria (satisfaction) than to others (commitment, performance, etc.). However, we should note that, in general, the behaviors do appear to be related to performance and other objective criteria, even though the number of correlations is relatively small.

Conclusion

In discussing the results of a meta-analytic study, the goal should be to answer the question, "'Where are we now that this

meta-analysis has been conducted?'" (Rosenthal, 1995, p. 190). Where we are now, in terms of the validity of Consideration and Initiating Structure, is at a considerably more optimistic place than past reviews have placed us. We are not arguing that all of the criticisms of past research have been wrong or that the present investigation answers all of the unanswered questions regarding the Ohio State leadership behaviors. For example, our results cannot speak to nonlinearities in the relations, a possibility that has been repeatedly raised (Fleishman, 1995, 1998) but seldom investigated. Furthermore, the paradox of differing results by measure, although being informed by this review, is essentially left unresolved. What we are arguing is that it is inadvisable, at this point, to abandon Consideration and Initiating Structure in leadership research. Some of the more prominent organizational psychologists in the history of the field—Stogdill, Shartle, and Fleishman, among others—spent the better part of their careers researching these concepts. It appears their investigations were more productive than previously thought, and that the fruit born of these investigations could prove useful once again in leadership research.

References

References marked with an asterisk indicate studies included in the meta-analysis.

*Abdel-Halim, A. A. (1981). Personality and task moderators of subordinate responses to perceived leader behavior. *Human Relations, 34,* 73–88.

*Aldag, R. J., & Brief, A. P. (1977). Relationships between leader behavior variability indices and subordinate responses. *Personnel Psychology, 30,* 419–426.

*Anderson, L. R. (1966). Leader behavior, member attitudes, and task performance of intercultural discussion groups. *Journal of Social Psychology, 69,* 305–319.

*Badin, I. J. (1974). Some moderator influences on relationships between consideration, initiating structure and organizational criteria. *Journal of Applied Psychology, 59,* 380–382.

Bales, R. F. (1954). In conference. *Harvard Business Review, 32,* 44–50.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice Hall.

Barrick, M. R., & Mount, M. K. (1991). The big five personality dimensions and job performance: A meta-analysis. *Personnel Psychology, 44,* 1–26.

*Bartol, K. M., & Wortman, M. S. (1975). Male versus female leaders: Effects on perceived leader behavior and satisfaction in a hospital. *Personnel Psychology, 28,* 533–547.

*Bartolo, K., & Furlonger, B. (2000). Leadership and job satisfaction among aviation fire fighters in Australia. *Journal of Managerial Psychology, 15,* 87–97.

*Bass, B. M. (1956). Leadership opinions as forecasts of supervisory success. *Journal of Applied Psychology, 40,* 345–346.

*Bass, B. M. (1958). Leadership opinions as forecasts of supervisory success: A replication. *Personnel Psychology, 11,* 515–518.

Bass, B. M. (1985). *Leadership and performance beyond expectations.* New York: Free Press.

Bass, B. M. (1990). *Bass and Stogdill's handbook of leadership.* New York: Free Press.

Bass, B. M. (1999). Two decades of research and development in transformational leadership. *European Journal of Work and Organizational Psychology, 8,* 9–32.

*Beatty, R. W. (1974). Supervisory behavior related to job success of hard-core unemployed over a two-year period. *Journal of Applied Psychology, 59,* 38–42.

*Begum, H. A., & Anwar, M. N. (1978). Consideration and structure dimensions of leadership behaviour and supervisory effectiveness as determined by managers' ratings and group productivity. *Asian Journal of Psychology & Education, 3,* 40–47.

*Bernardin, H. J. (1976). The influence of reinforcement orientation on the relationship between supervisory style and effectiveness criteria. *Dissertation Abstracts International, 37,* 1018.

*Bernardin, H. J. (1987). Effect of reciprocal leniency on the relation between consideration scores from the Leader Behavior Description Questionnaire and performance ratings. *Psychological Reports, 60,* 479–487.

Bernardin, H. J., & Alvares, K. M. (1976). The managerial grid as a predictor of conflict resolution method and managerial effectiveness. *Administrative Science Quarterly, 21,* 84–92.

*Besco, R. O., & Lawshe, C. H. (1959). Foreman leadership as perceived by superiors and subordinates. *Personnel Psychology, 12,* 573–582.

*Bird, A. M. (1977). Development of a model for predicting team performance. *Research Quarterly, 48,* 24–32.

Blake, R. R., Mouton, J. S., & Bidwell, A. C. (1962). Managerial grid. *Advanced Management–Office Executive, 1,* 12–15.

*Bledsoe, J. C. (1978). Morale of curriculum directors as related to perceptions of leadership of superintendents. *Psychological Reports, 43,* 1283–1288.

*Bowers, D. G., & Seashore, S. E. (1966). Predicting organizational effectiveness with a four-factor theory of leadership. *Administrative Science Quarterly, 11,* 238–263.

*Brown, S. E., & Bledsoe, J. C. (1978). Job satisfaction of school superintendents as related to perceptions of leader's behavior. *Psychological Reports, 42,* 171–174.

*Bryman, A., Bresnen, M., Ford, J., & Beardsworth, A. (1987). Leader orientation and organizational transience: An investigation using Fiedler's LPC Scale. *Journal of Occupational Psychology, 60,* 13–19.

Burns, J. M. (1978). *Leadership.* New York: Harper & Row.

Chemers, M. M. (1997). *An integrative theory of leadership.* Mahwah, NJ: Erlbaum.

*Conger, J. A., & Kanungo, R. N. (1992). Perceived behavioral attributes of charismatic leadership. *Canadian Journal of Behavioral Science, 24,* 86–102.

*Conger, J. A., & Kanungo, R. N. (1994). Charismatic leadership in organizations: Perceived behavioral attributes and their measurement. *Journal of Organizational Behavior, 15,* 439–452.

Cropanzano, R., Byrne, Z. S., Bobocel, D. R., & Rupp, D. E. (2001). Moral virtues, fairness heuristics, social entities, and other denizens of organizational justice. *Journal of Vocational Behavior, 58,* 164–209.

*Cummins, R. C. (1971). Relationship of initiating structure and job performance as moderated by consideration. *Journal of Applied Psychology, 55,* 489–490.

*Cummins, R. C. (1972). Leader-member relations as a moderator of the effects of leader behavior and attitude. *Personnel Psychology, 25,* 655–660.

*Dawson, J. E., Messe, L. A., & Phillips, J. L. (1972). Effect of instructor–leader behavior on student performance. *Journal of Applied Psychology, 56,* 369–376.

*Dessler, G., & Valenzi, E. R. (1977). Initiation of structure and subordinate satisfaction: A path analysis of path–goal theory. *Academy of Management Journal, 20,* 251–259.

Dionne, S. D., Yammarino, F. J., Atwater, L. E., & James, L. R. (2002). Neutralizing substitutes for leadership theory: Leadership effects and common-source bias. *Journal of Applied Psychology, 87,* 454–464.

*Distefano, M. K., & Pryer, M. W. (1973). Comparisons of leader and

subordinate descriptions of leadership behavior. *Perceptual & Motor Skills, 37,* 714.

*Downey, H. K., Sheridan, J. E., & Slocum, J. W. (1975). Analysis of relationships among leader behavior, subordinate job performance and satisfaction: A path–goal approach. *Academy of Management Journal, 18,* 253–262.

*Drory, A., & Ben-Porat, A. (1980). Leadership style and leniency bias in evaluation of employees' performance. *Psychological Reports, 46,* 735–739.

*Ehrlich, S. B., Meindl, J. R., & Viellieu, B. (1990). The charismatic appeal of a transformational leader: An empirical case study of a small, high-technology contractor. *Leadership Quarterly, 1,* 229–247.

*Evans, M. G. (1970). The effects of supervisory behavior on the path–goal relationship. *Organizational Behavior & Human Decision Processes, 5,* 277–298.

*Evans, M. G. (1974). Extensions of a path–goal theory of motivation. *Journal of Applied Psychology, 59,* 172–178.

*Fleishman, E. A. (1953). The description of supervisory behavior. *Journal of Applied Psychology, 37,* 1–6.

*Fleishman, E. A. (1957). A leader behavior description for industry. In R. M. Stogdill & A. E. Coons (Eds.), *Leader behavior: Its description and measurement* (pp. 103–119). Columbus, OH: Bureau of Business Research.

*Fleishman, E. A. (1957). The Leader Opinion Questionnaire. In R. M. Stogdill & A. E. Coons (Eds.), *Leader behavior: Its description and measurement* (pp. 120–133). Columbus, OH: Bureau of Business Research.

Fleishman, E. A. (1973). Twenty years of consideration and structure. In E. A. Fleishman & J. G. Hunt (Eds.), *Current developments in the study of leadership* (pp. 1–40). Carbondale, IL: Southern Illinois University Press.

*Fleishman, E. A. (1989a). *Leadership Opinion Questionnaire (LOQ) examiner's manual.* Park Ridge, IL: Science Research Associates.

*Fleishman, E. A. (1989b). *Supervisory Behavior Description Questionnaire (SBD) examiner's manual.* Park Ridge, IL: Science Research Associates.

Fleishman, E. A. (1995). Consideration and structure: Another look at their role in leadership research. In F. Dansereau & F. J. Yammarino (Eds.), *Leadership: The multiple-level approaches* (pp. 51–60). Stamford, CT: JAI Press.

Fleishman, E. A. (1998). Patterns of leadership behavior related to employee grievances and turnover: Some post hoc reflections. *Personnel Psychology, 51,* 825–834.

*Fleishman, E. A., & Harris, E. F. (1962). Patterns of leadership behavior related to employee grievances and turnover. *Personnel Psychology, 15,* 43–56.

Fleishman, E. A., Mumford, M. D., Zaccaro, S. J., & Levin, K. Y. (1991). Taxonomic efforts in the description of leader behavior: A synthesis and functional interpretation. *Leadership Quarterly, 2,* 245–287.

*Fleishman, E. A., & Peters, D. R. (1962). Interpersonal values, leadership attitudes and managerial "success." *Personnel Psychology, 15,* 127–143.

*Fleishman, E. A., & Salter, J. A. (1961). Humanizing relationships in a small business. *Advanced Management Journal, 26,* 18–20.

*Fleishman, E. A., & Salter, J. A. (1963). Relationship between the leader's behavior and his empathy toward subordinates. *Journal of Industrial Psychology, 1,* 79–84.

*Fleishman, E. A., & Simmons, J. (1970). Relationship between leadership patterns and effectiveness ratings among Israeli foremen. *Personnel Psychology, 23,* 169–172.

Goldberg, L. R. (1990). An alternative "description of personality": The Big-Five factor structure. *Journal of Personality and Social Psychology, 59,* 1216–1229.

Greenberg, J., & Lind, E. A. (2000). The pursuit of organizational justice: From conceptualization to implication to application. In C. L. Cooper & E. A. Locke (Eds.), *Industrial and organizational psychology: Linking theory with practice* (pp. 72–108). Oxford, United Kingdom: Blackwell.

*Greene, C. N. (1975). The reciprocal nature of influence between leader and subordinate. *Journal of Applied Psychology, 60,* 187–193.

*Gruenfeld, L. W., & Weissenberg, P. (1966). Supervisory characteristics and attitudes toward performance appraisals. *Personnel Psychology, 19,* 143–151.

*Halpin, A. W. (1954). The leadership behavior and combat performance of airplane commanders. *Journal of Abnormal & Social Psychology, 49,* 19–22.

*Halpin, A. W. (1955). The leader behavior and leadership ideology of educational administrators and aircraft commanders. *Harvard Educational Review, 25,* 18–32.

Halpin, A. W. (1957a). *Manual for the Leader Behavior Description Questionnaire.* Columbus, OH: Bureau of Business Research, Ohio State University.

*Halpin, A. W. (1957b). The leader behavior and effectiveness of aircraft commanders. In R. M. Stogdill & A. E. Coons (Eds.), *Leader behavior: Its description and measurement* (pp. 52–64). Columbus, OH: Bureau of Business Research.

*Halpin, A. W. (1957c). The observed leader behavior and ideal leader behavior of aircraft commanders and school superintendents. In R. M. Stogdill & A. E. Coons (Eds.), *Leader behavior: Its description and measurement* (pp. 65–68). Columbus, OH: Bureau of Business Research.

*Halpin, A. W., & Winer, B. J. (1957). A factorial study of the leader behavior descriptions. In R. M. Stogdill & A. E. Coons (Eds.), *Leader behavior: Its description and measurement* (pp. 39–51). Columbus, OH: Bureau of Business Research.

*Hand, H. H., & Slocum, J. W. (1972). A longitudinal study of the effects of a human relations training program on managerial effectiveness. *Journal of Applied Psychology, 56,* 412–417.

*Hemphill, J. K. (1955). Leadership behavior associated with the administrative reputation of college departments. *Journal of Educational Psychology, 46,* 385–401.

*Hemphill, J. K. (1957). Leader behavior associated with the administrative reputations of college departments. In R. M. Stogdill & A. E. Coons (Eds.), *Leader behavior: Its description and measurement* (pp. 74–85). Columbus, OH: Bureau of Business Research.

*Hemphill, J. K., & Coons, A. E. (1957). Development of the Leader Behavior Description Questionnaire. In R. M. Stogdill & A. E. Coons (Eds.), *Leader behavior: Its description and measurement* (pp. 6–38). Columbus, OH: Bureau of Business Research.

*Herold, D. M. (1974). Interaction of subordinate and leader characteristics in moderating the consideration–satisfaction relationship. *Journal of Applied Psychology, 59,* 649–651.

Hickman, G. R. (Ed.). (1998). *Leading organizations: Perspectives for a new era.* Thousand Oaks, CA: Sage.

*House, R. J. (1971). A path–goal theory of leader effectiveness. *Administrative Science Quarterly, 16,* 321–338.

House, R. J. (1977). A 1976 theory of charismatic leadership. In J. G. Hunt & L. L. Larsen (Eds.), *Leadership: The cutting edge.* Carbondale, IL: Southern Illinois University Press.

House, R. J. (1996). Path–goal theory of leadership: Lessons, legacy and a reformulated theory. *Leadership Quarterly, 7,* 323–352.

House, R. J., & Aditya, R. N. (1997). The social scientific study of leadership: Quo vadis? *Journal of Management, 23,* 409–473.

*House, R. J., Filley, A. C., & Gujarati, D. N. (1971). Leadership style, hierarchical influence, and the satisfaction of subordinate role expectations: A test of Likert's influence proposition. *Journal of Applied Psychology, 55,* 422–432.

*House, R. J., Filley, A. C., & Kerr, S. (1971). Relation of leader consid-

eration and initiating structure to R and D subordinates' satisfaction. *Administrative Science Quarterly, 16,* 19–30.

*House, R. J., & Kerr, S. (1973). Organizational independence, leader behavior, and managerial practices: A replicated study. *Journal of Applied Psychology, 58,* 173–180.

House, R. J., & Podsakoff, P. M. (1994). Leadership effectiveness. In J. Greenberg (Ed.), *Organizational behavior: The state of the science* (pp. 45–82). Hillsdale, NJ: Erlbaum.

*Hunt, J. G., & Liebscher, V. K. (1973). Leadership preference, leadership behavior, and employee satisfaction. *Organizational Behavior & Human Decision Processes, 9,* 59–77.

Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis.* Newbury Park, CA: Sage.

*Ilgen, D. R., & Fuji, D. S. (1976). An investigation of the validity of leader behavior descriptions obtained from subordinates. *Journal of Applied Psychology, 61,* 642–651.

*Johns, G. (1978). Task moderators of the relationship between leadership style and subordinate responses. *Academy of Management Journal, 21,* 319–325.

*Johnson, P. O., & Bledsoe, J. C. (1973). Morale as related to perceptions of leader behavior. *Personnel Psychology, 26,* 581–592.

*Katerberg, R., & Hom, P. W. (1981). Effects of within-group and between-groups variation in leadership. *Journal of Applied Psychology, 66,* 218–223.

*Katz, R. (1977). The influence of group conflict on leadership effectiveness. *Organizational Behavior & Human Decision Processes, 20,* 265–286.

*Kenis, I. (1978). Leadership behavior, subordinate personality, and satisfaction with supervision. *Journal of Psychology, 98,* 99–107.

Kerr, S. J., & Jermier, J. M. (1978). Substitutes for leadership: Their meaning and measurement. *Organizational Behavior and Human Performance, 22,* 375–403.

Kerr, S. J., & Schriesheim, C. (1974). Consideration, initiating structure, and organizational criteria—An update of Korman's 1966 review. *Personnel Psychology, 27,* 555–568.

Kirkpatrick, S. A., Locke, E. A., & Latham, G. P. (1996). Implementing the vision: How is it done? *Polish Psychological Bulletin, 27,* 93–106.

*Kohli, A. K. (1989). Effects of supervisory behavior: The role of individual differences among salespeople. *Journal of Marketing, 53,* 40–50.

Korman, A. K. (1966). "Consideration," "initiating structure," and organizational criteria—A review. *Personnel Psychology, 19,* 349–361.

*Korsgaard, M. A., Schweiger, D. M., & Sapienza, H. J. (1995). Building commitment, attachment, and trust in strategic decision-making teams: The role of procedural justice. *Academy of Management Journal, 38,* 60–84.

*Lahat-Mandelbaum, B.-S., & Kipnis, D. (1973). Leader behavior dimensions related to students' evaluation of teaching effectiveness. *Journal of Applied Psychology, 58,* 250–253.

*Larson, L. L., Hunt, J. G., & Osborn, R. N. (1976). The great hi-hi leader behavior myth: A lesson from Occams razor. *Academy of Management Journal, 19,* 628–641.

*Lee, C., & Schuler, R. S. (1980). Goal specificity and difficulty and leader initiating structure as strategies for managing role stress. *Journal of Management, 6,* 177–187.

Leventhal, G. S. (1980). What should be done with equity theory? New approaches to the study of fairness in social relationship. In K. J. Gergen, M. S. Greenberg, & R. H. Willis (Eds.), *Social exchange: Advances in theory and research* (pp. 27–55). New York: Plenum.

Liden, R. C., Sparrowe, R. T., & Wayne, S. J. (1997). Leader–member exchange theory: The past and potential for the future. *Research in Personnel and Human Resources Management, 15,* 47–119.

Likert, R. (1961). *New patterns of management.* New York: McGraw-Hill.

Locke, E. A. (1997). The motivation to work: What we know. *Advances in Motivation and Achievement, 10,* 375–412.

Lord, R. G., Foti, R. J., & De Vader, C. L. (1984). A test of leadership categorization theory: Internal structure, information processing, and leadership perceptions. *Organizational Behavior and Human Performance, 34,* 343–378.

Lowe, K. B., Kroeck, K. G., & Sivasubramaniam, N. (1996). Effectiveness correlates of transformation and transactional leadership: A meta-analytic review of the MLQ literature. *Leadership Quarterly, 7,* 385–425.

*Lowin, A., Hrapchak, W. J., & Kavanagh, M. J. (1969). Consideration and initiating structure: An experimental investigation of leadership traits. *Administrative Science Quarterly, 14,* 238–253.

Mathieu, J. E. (2001). "Into the box" thinking about leadership research. In S. J. Zaccaro & R. J. Klimoski (Eds.), *The nature of organizational leadership* (pp. 437–463). San Francisco: Jossey-Bass.

*Matsui, T., Osawa, T., & Terai, T. (1975). Relations between supervisory motivation and the consideration and structure aspects of supervisory behavior. *Journal of Applied Psychology, 60,* 451–454.

*McCabe, J. B. (1986). Leadership behavior and job satisfaction of medical technologists: A path–goal approach. *Dissertation Abstracts International, 46,* 3789.

*Michaelsen, L. K. (1973). Leader orientation, leader behavior, group effectiveness and situational favorability: An empirical extension of the contingency model. *Organizational Behavior & Human Decision Processes, 9,* 226–245.

*Miles, R. H., & Petty, M. M. (1977). Leader effectiveness in small bureaucracies. *Academy of Management Journal, 20,* 238–250.

Mount, M. K., & Barrick, M. R. (1998). Five reasons why the "Big Five" article has been frequently cited. *Personnel Psychology, 51,* 849–857.

Murphy, K. R., & DeShon, R. (2000). Interrater correlations do not estimate the reliability of job performance ratings. *Personnel Psychology, 53,* 873–900.

*Nealy, S. M., & Blood, M. R. (1968). Leadership performance of nursing supervisors at two organizational levels. *Journal of Applied Psychology, 52,* 414–422.

Northouse, P. G. (1997). *Leadership: Theory and practice.* Thousand Oaks, CA: Sage.

*Nystrom, P. C. (1978). Managers and the hi-hi leader myth. *Academy of Management Journal, 21,* 325–331.

*Oaklander, H., & Fleishman, E. A. (1964). Patterns of leadership related to organizational stress in hospital settings. *Administrative Science Quarterly, 8,* 520–532.

*Oldham, G. R. (1976). The motivational strategies used by supervisors: Relationships to effectiveness indicators. *Organizational Behavior & Human Decision Processes, 15,* 66–86.

*Osborn, R. N., & Hunt, J. G. (1975). Relations between leadership, size, and subordinate satisfaction in a voluntary organization. *Journal of Applied Psychology, 60,* 730–735.

*Osborn, R. N., & Vicars, W. M. (1976). Sex stereotypes: An artifact in leader behavior and subordinate satisfaction analysis? *Academy of Management Journal, 19,* 439–444.

*Parker, T. C. (1963). Relationships among measures of supervisory behavior, group behavior, and situational characteristics. *Personnel Psychology, 16,* 319–334.

*Pennington, P. W. (1998). Principal leadership and teacher motivation in secondary schools. *Dissertation Abstracts International Section A: Humanities & Social Sciences, 58,* 2943.

*Pettman, P. J. (1977). Consideration, initiating structure, field dependence and student evaluation of teacher behavior. *Dissertation Abstracts International, 37,* 5418.

*Petty, M. M., & Bruning, N. S. (1980). A comparison of the relationships between subordinates' perceptions of supervisory behavior and mea-

sures of subordinates' job satisfaction for male and female leaders. *Academy of Management Journal, 23,* 717–725.

*Petty, M. M., & Lee, G. K. (1975). Moderating effects of sex of supervisor and subordinate on relationships between supervisory behavior and subordinate satisfaction. *Journal of Applied Psychology, 60,* 624–628.

Podsakoff, P. M., MacKenzie, S. B., & Bommer, W. H. (1996). Meta-analysis of the relationships between Kerr and Jermier's substitutes for leadership and employee job attitudes, role perceptions, and performance. *Journal of Applied Psychology, 81,* 380–399.

*Pool, S. W. (1996). The path–goal theory of organizational leadership: An empirical analysis of leader behavior by integrating leader substitute theory. *Dissertation Abstracts International Section A: Humanities & Social Sciences, 57,* 1225.

Quiñones, M. A., Ford, J. K., & Teachout, M. S. (1995). The relationship between work experience and job performance: A conceptual and meta-analytic review. *Personnel Psychology, 48,* 887–910.

Raju, N. S., Burke, M. J., Normand, J., & Langlois, G. M. (1991). A new meta-analytic approach. *Journal of Applied Psychology, 76,* 432–446.

*Rambo, W. W. (1958). The construction and analysis of a leadership behavior rating form. *Journal of Applied Psychology, 42,* 409–415.

*Rider, L. H. (1975). Leader behavior, locus of control and consultation effectiveness of school psychologists. *Dissertation Abstracts International, 35,* 5134.

*Rooker, J. L. (1968). The relationship of need achievement and need affiliation to leader behavior. *Dissertation Abstracts International, 28,* 4426.

Rosenthal, R. (1995). Writing meta-analytic reviews. *Psychological Bulletin, 118,* 183–192.

*Rowland, K. M., & Scott, W. E. (1968). Psychological attributes of effective leadership in a formal organization. *Personnel Psychology, 21,* 365–377.

Rush, M. C., Thomas, J. C., & Lord, R. G. (1977). Implicit leadership theory: A potential threat to the internal validity of leader behavior questionnaires. *Organizational Behavior and Human Decision Processes, 20,* 93–110.

Sagie, A., & Koslowsky, M. (1993). Detecting moderators with meta-analysis: An evaluation and comparison of techniques. *Personnel Psychology, 46,* 629–640.

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin, 124,* 262–274.

Schmidt, F. L., Hunter, J. E., Pearlman, K., & Hirsh, H. R. (1985). Forty questions about validity generalization and meta-analysis. *Personnel Psychology, 38,* 697–798.

Schmidt, F. L., Viswesvaran, C., & Ones, D. S. (2000). Reliability is not validity and validity is not reliability. *Personnel Psychology, 53,* 901–912.

*Schriesheim, C. A. (1979). The similarity of individual directed and group directed leader behavior descriptions. *Academy of Management Journal, 22,* 345–355.

*Schriesheim, C. A. (1982). The great high Consideration–high Initiating Structure leadership myth: Evidence on its generalizability. *Journal of Social Psychology, 116,* 221–231.

*Schriesheim, C. A., Cogliser, C. C., & Neider, L. L. (1995). Is it "trustworthy?": A multiple-levels-of-analysis reexamination of an Ohio state leadership study, with implications for future research. *Leadership Quarterly, 6,* 111–145.

*Schriesheim, C. A., & DeNisi, A. S. (1981). Task dimensions as moderators of the effects of instrumental leadership: A two-sample replicated test of path–goal leadership theory. *Journal of Applied Psychology, 66,* 589–597.

*Schriesheim, C. A., House, R. J., & Kerr, S. (1976). Leader initiating structure: A reconciliation of discrepant research results and some empirical tests. *Organizational Behavior & Human Decision Processes, 15,* 297–321.

Schriesheim, C. A., & Kerr, S. (1974). Psychometric properties of the Ohio State leadership scales. *Psychological Bulletin, 81,* 756–765.

*Schriesheim, C. A., Kinicki, A. J., & Schriesheim, J. F. (1979). The effect of leniency on leader behavior descriptions. *Organizational Behavior & Human Decision Processes, 23,* 1–29.

*Schriesheim, C., & von Glinow, M. A. (1977). The path–goal theory of leadership: A theoretical and empirical analysis. *Academy of Management Journal, 20,* 398–405.

*Schuler, R. S. (1976). Participation with supervisor and subordinate authoritarianism: A path–goal theory reconciliation. *Administrative Science Quarterly, 21,* 320–325.

*Seeman, M. (1957). A comparison of general and specific leader behavior descriptions. In R. M. Stogdill & A. E. Coons (Eds.), *Leader behavior: Its description and measurement* (pp. 86–102). Columbus, OH: Bureau of Business Research.

*Seltzer, J., & Bass, B. M. (1990). Transformational leadership: Beyond initiation and consideration. *Journal of Management, 16,* 693–703.

*Shapiro, M. I. (1971). Initiating structure and consideration: A situationist's view of the efficacy of two styles of leadership. *Dissertation Abstracts International, 31,* 4382–4383.

*Sheih, C. S.-M. (1999). The effect of perceived leadership behavior on the job satisfaction of subordinates in Taiwan's university libraries. *Dissertation Abstracts International Section B: The Sciences & Engineering, 59,* 3749.

*Sheridan, J. E., & Vredenburgh, D. J. (1978a). Predicting leadership behavior in a hospital organization. *Academy of Management Journal, 21,* 679–689.

*Sheridan, J. E., & Vredenburgh, D. J. (1978b). Usefulness of leadership behavior and social power variables in predicting job tension, performance, and turnover of nursing employees. *Journal of Applied Psychology, 63,* 89–95.

*Sheridan, J. E., & Vredenburgh, D. J. (1979). Structural model of leadership influence in a hospital organization. *Academy of Management Journal, 22,* 6–21.

*Sheridan, J. E., Kerr, J. L., & Abelson, M. A. (1982). Leadership activation theory: An opponent process model of subordinate responses to leadership behavior. In J. G. Hunt, U. Sekaran, & C. A. Schriesheim (Eds.), *Leadership: Beyond establishment views* (pp. 122–141). Carbondale, IL: Southern Illinois University Press.

*Sims, H. P., & Szilagyi, A. D. (1975). Leader structure and subordinate satisfaction for two hospital administrative levels: A path analysis approach. *Journal of Applied Psychology, 60,* 194–197.

*Skinner, E. W. (1969). Relationships between leadership behavior patterns and organizational–situational variables. *Personnel Psychology, 22,* 489–494.

*Spitzer, M. E., & McNamara, W. J. (1964). A managerial selection study. *Personnel Psychology, 17,* 19–40.

*Stanley, D. M. (1975). The effects of consideration and initiating structure as moderated by incremental influence. *Dissertation Abstracts International, 35,* 3639.

Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin, 87,* 245–251.

*Stinson, J. E., & Johnson, T. W. (1975). The path–goal theory of leadership: A partial test and suggested refinement. *Academy of Management Journal, 18,* 242–252.

Stogdill, R. M. (1950). Leadership, membership and organization. *Psychological Bulletin, 47,* 1–14.

Stogdill, R. M. (1963). *Manual for the Leader Behavior Description Questionnaire, Form XII.* Columbus, OH: Bureau of Business Research, Ohio State University.

*Stogdill, R. M., & Shartle, C. L. (1956). *Methods in the study of admin-*

istrative leadership (Research Monograph No. 80). Columbus, OH: Bureau of Business Research.

*Stogdill, R. M., Goode, O. S., & Day, D. R. (1963a). The leader behavior of United States Senators. *Journal of Psychology, 56,* 3–8.

*Stogdill, R. M., Goode, O. S., & Day, D. R. (1963b). The leader behavior of corporation presidents. *Personnel Psychology, 16,* 127–132.

*Stogdill, R. M., Goode, O. S., & Day, D. R. (1964). The leader behavior of presidents of labor unions. *Personnel Psychology, 17,* 49–57.

*Swanson, R. G., & Johnson, D. A. (1975). Relation between peer perception of leader behavior and instructor-pilot performance. *Journal of Applied Psychology, 60,* 198–200.

*Szilagyi, A. D., & Keller, R. (1976). A comparative investigation of the supervisory behavior description questionnaire (SBDQ) and the revised leader behavior description questionnaire (LBDQ-Form XII). *Academy of Management Journal, 19,* 642–650.

*Szilagyi, A. D., & Sims, H. P. (1974). An exploration of the path–goal theory of leadership in a health care environment. *Academy of Management Journal, 17,* 622–634.

*Tscheulin, D. (1973). Leader behavior measurement in German industry. *Journal of Applied Psychology, 57,* 28–31.

*Tucker, J. H. (1977). Leadership in autonomous group environments, task/person orientation and interpersonal competence. *Dissertation Abstracts International, 37,* 4202–4203.

*Valenzi, E. R., & Dessler, G. (1978). Relationships of leader behavior, subordinate role ambiguity and subordinate job satisfaction. *Academy of Management Journal, 21,* 671–678.

*van Engen, M. L., van der Leeden, R., & Willemsen, T. (2001). Gender, context, leadership styles: A field study. *Journal of Occupational & Organizational Psychology, 74,* 581–598.

*Vecchio, R. P. (1980). Alternatives to the least preferred co-worker construct. *Journal of Social Psychology, 112,* 261–269.

*Vecchio, R. P. (1981). Situational and behavioral moderators of subordinate satisfaction with supervision. *Human Relations, 34,* 947–963.

Weissenberg, P., & Kavanagh, M. J. (1972). The independence of initiating structure and consideration: A review of the evidence. *Personnel Psychology, 25,* 119–130.

Wofford, J. C., & Liska, L. Z. (1993). Path–goal theory of leadership: A meta-analysis. *Journal of Management, 19,* 857–876.

*Yoho, S. K. (1995). Follower-centered leadership: An investigation of leader behavior, leader power, follower competency, and follower job performance in leader–follower relationships. *Dissertation Abstracts International Section A: Humanities & Social Sciences, 56,* 0629.

Yukl, G. (1998). *Leadership in organizations.* Upper Saddle River, NJ: Prentice Hall.

*Yukl, G., & Lepsinger, R. (1990). Preliminary report on validation of the Managerial Practices Survey. In K. E. Clark & M. B. Clark (Eds.), *Measures of leadership* (pp. 223–237). West Orange, NJ: Leadership Library of America.

Yukl, G., & Van Fleet, D. D. (1992). Theory and research on leadership in organizations. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (Vol. 2, pp. 147–197). Palo Alto, CA: Consulting Psychologists Press.

*Yunker, G. W., & Hunt, J. G. (1976). An empirical comparison of the Michigan Four-Factor and Ohio State LBDQ leadership scales. *Organizational Behavior & Human Decision Processes, 17,* 45–65.

Received September 3, 2002
Revision received April 7, 2003
Accepted May 2, 2003 ■

Low Publication Prices for APA Members and Affiliates

Keeping you up-to-date. All APA Fellows, Members, Associates, and Student Affiliates receive—as part of their annual dues—subscriptions to the *American Psychologist* and *APA Monitor.* High School Teacher and International Affiliates receive subscriptions to the *APA Monitor,* and they may subscribe to the *American Psychologist* at a significantly reduced rate. In addition, all Members and Student Affiliates are eligible for savings of up to 60% (plus a journal credit) on all other APA journals, as well as significant discounts on subscriptions from cooperating societies and publishers (e.g., the American Association for Counseling and Development, Academic Press, and Human Sciences Press).

Essential resources. APA members and affiliates receive special rates for purchases of APA books, including the *Publication Manual of the American Psychological Association,* and on dozens of new topical books each year.

Other benefits of membership. Membership in APA also provides eligibility for competitive insurance plans, continuing education programs, reduced APA convention fees, and specialty divisions.

More information. Write to American Psychological Association, Membership Services, 750 First Street, NE, Washington, DC 20002-4242.

Copyright of Journal of Applied Psychology is the property of American Psychological Association and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.