APPLIED PSYCHOLOGY: AN INTERNATIONAL REVIEW, 1998, *47* (3), 397–420

Transformational Leadership and Objective Performance in Banks

Alois L.J. Geyer and Johannes M. Steyrer

University of Economics and Business Administration Vienna, Austria

Dans cette recherche, on analyse les relations existant entre les leaderships de transaction et de transformation et des indices de performance de vingt banques, cela à partir d'un échantillon de quelques 1500 observations. Le comportement de leadership est mesuré grâce au Questionnaire Multifactoriel de Leadership (MLQ) qui est appliqué ici pour la première fois à des germanophones. La valeur psychométrique du MLQ allemand a été analysée et jugée pas satisfaisante. Nous avons mis au point et contre-validé un MLQ modifié à quatre facteurs. Il s'agit de la première étude empirique qui rapporte le leadership à des indices objectifs de performance, échappant à toute influence qualitative ou subjective. Les mesures objectives de performance prennent en considération les charactéristiques des clients et du marché local et indiquent si des objectifs raisonnables ne sont pas atteints ou au contraire dépassés. A partir des évaluations du modèle de Lisrel, nous confirmons l'hypothèse selon laquelle le leadership de transformation favorise davantage la performance que le leadership de transaction. La distinction entre performance à long terme et performance à court-terme est une autre nouveauté de cette étude. Il apparaît que les dimensions centrales du leadership de transformation sont plus fortement corrélées avec la performance à long terme. La considération portée à l'individue présente une corrélation positive avec la performance à court-terme, mais est négativement corrélée avec la performance à long terme.

This investigation examines the relations between transformational/ transactional leadership and performance indicators of 20 different banks, using a sample of some 1500 observations. Leadership behaviour is measured by the Multifactor Leadership Questionnaire (MLQ), which is used for the first time in the German-speaking area. The psychometric quality of the German MLQ is analysed and found to be inadequate. We therefore derive and cross-validate a modified four-factor MLQ. This is the first empirical study

Requests for reprints should be sent to Dr. Alois L.J. Geyer, University of Economics and Business Administration Vienna, Augasse 2-6, A-1090 Wien, Austria. E-mail: geyer@wu-wien.ac.at

We would like to thank two anonymous referees and the editor for many helpful comments and suggestions.

that relates leadership behaviour to objective performance indicators without any qualitative or subjective influence. Objective performance measures take into account customer features and local market conditions, and show whether reasonable targets are over- or underperformed. Based on LISREL model estimates we find support for the hypothesis that transformational leadership affects performance over and above transactional leadership. Another new aspect of this study is the distinction between long-term and short-term performance. We find that core dimensions of transformational leadership are more strongly related to long-term than to short-term performance. Individualised Consideration is positively related to short-term but negatively related to long-term performance.

INTRODUCTION

Bass (1985) introduced a theory of "transformational" leadership that was based on Burns' (1978) classification of "transactional" and "transformational" political leaders. Bass argued that leadership is generally conceptualised as a transactional exchange process. A transactional leader motivates subordinates by exchanging rewards for services rendered. The concern of such leaders is to clarify their subordinates' goals and to arrange contingent rewards as inducements towards goal attainment. Transformational leaders, on the other hand, motivate subordinates to perform "beyond expectations" by developing, intellectually stimulating, and inspiring followers to transcend their own self-interests for a higher collective purpose, mission, or vision.

In order to measure transformational/transactional leadership, Bass and his colleagues developed the "Multifactor Leadership Questionnaire" (MLQ). Factor-analytic studies (Bass, 1985; Bass & Avolio, 1990a; Bass, Avolio, & Goodheim, 1987; Hater & Bass, 1988) confirmed the hypothesis about the underlying composition of the two types of leadership. Four of seven scales were identified and defined as characteristic of transformational leadership. They are as follows:

Charisma involves gaining respect for, trust, and confidence in others and transmitting a strong sense of mission to them. Inspirational Motivation means communicating a vision with fluency and confidence, increasing optimism, team spirit, and enthusiasm. Intellectual Stimulation means actively encouraging others to look at old methods in new ways, fostering innovation, and stressing the use of intelligence. Individualised Consideration gives personal attention to all individuals, making each individual feel valued, and delegates tasks as a means of developing followers.

Two scales were identified and defined as characteristic of transactional leadership: Contingent Reward leadership occurs when the leader rewards or disciplines the follower depending on the adequacy of the follower's performance. Management by Exception avoids giving directions if

established procedures are working and allows followers to continue doing their jobs as always if performance goals are met.

The seventh scale is called Laissez-Faire. It relates to the avoidance or absence of leadership and is, by definition, the most inactive type of leadership. As opposed to transactional leadership, Laissez-Faire represents a non-transaction.

Review of Previous Research

The MLQ has been examined in over 75 research studies (Lowe & Kroeck, 1995), in several countries, and in a variety of organisational contexts, for leaders with high and low levels of responsibility. These studies have generally reported statistically significant relations between leader effectiveness and the transformational scales. The transactional scale Contingent Reward is also positively related to outcome measures, but the association is weaker than with transformational scales. Correlations for Management by Exception are low and often negative, if significant at all. Laissez-Faire leadership is significantly negative related to the outcome measures (e.g. Avolio & Yammarino, 1990; Bass, 1985; Bass & Avolio, 1989; Hater & Bass, 1988; Seltzer & Bass, 1990; Seltzer, Numerof, & Bass, 1989; Singer, 1985; Yammarino & Bass, 1990).

The effectiveness indicators can be separated into two groups: in the first group the rating of leadership style and the rating of leader effectiveness are completed by the subordinate, using the effectiveness criterion measures included in the MLQ (extra effort, satisfaction with the leader, and perceived effectiveness of the leader). The second group of criterion measures are quasi-institutional measures of the leader's effectiveness including either hard measures (e.g. percentage of goals met) or soft measures such as supervisor's ratings of work groups. Large and significant differences have been found between subordinate ratings of effectiveness and quasi-organisational measures of effectiveness for all MLQ scales. Significantly stronger relationships have been found for the former than for the latter measures. Two meta-analytic reviews aggregating about 35 published and unpublished studies (Coleman et al., 1995; Lowe & Kroeck, 1995) have confirmed this tendency.

Results using subordinate ratings of effectiveness can be criticised because it is difficult to assess the connection between leadership and performance when respondents are the source of information about both leadership behaviour and effectiveness. This might cause artificially inflated correlations by the proclivity of respondents to answer questions in consistent patterns ("common method variance"). Using a single source to collect predictor and criterion data may enhance the impact of the rater's "implicit theories" about the relation between leadership behaviour and

performance, thus inflating the observed correlation between the two measures (Binning, Zaba, & Whattam, 1986; Eden & Leviatan, 1975; Larson, Lingle, & Scerbo, 1984; Lord, Binning, Rush, & Thomas, 1978; Phillips & Lord, 1982; Rush, Thomas, & Lord, 1977).

The relation between transformational/transactional leadership and quasi-institutional measures has been analysed much less frequently in empirical studies using the MLQ. Various indicators of effectiveness are used: (1) supervisory appraised performance (Bass, 1985); (2) performance appraisal evaluation, various measures of satisfaction (Waldman, Bass, & Einstein, 1987); (3) supervisors' ratings of officers and work group (Hater & Bass, 1988); (4) effectiveness rated by subordinates, supervisory performance evaluation frequency, promotion recommendation (Waldman, Bass, & Yammarino, 1990); (5) early promotion (Bass & Yammarino, 1991); (6) project quality, budget schedule efficiency (Keller, 1992); (7) percentage of goals met (Howell & Avolio, 1993); (8) sales performance rating (Yammarino & Dubinsky, 1994).

These empirical studies have the following limitations: (1) All of these results refer to leaders of only one organisational context. So far there exists no study that is based on data from different organisations. (2) Three of the studies mentioned were carried out in military organisations (Bass & Yammarino, 1991; Hater & Bass, 1988; Waldman et al., 1990), so that the transferability of conclusions to business organisations is questionable. (3) Except for the study by Howell and Avolio (1993), who used measures that were at least 80% objective and no more than 20% subjective, all other studies were based on qualitative measures. They were either subordinate-determined or based on soft criteria such as supervisory performance appraisals. As these qualitative evaluations are based on the same social context, it is possible that attributional processes in the sense of performance cues are involved. Furthermore, halo effects (Murphy & Reynolds, 1988) and similarity effects (Turban & Jones, 1988) may influence individual assessments of perceived performance.

All these effects should be less influential when analysing objective performance outcomes instead of subordinate-based or quasi-institutional measures with a qualitative and subjective influence. To our knowledge this is the first empirical study that uses objective indicators where no such effects are active.

Purpose of the Current Study and Hypotheses

The objectives of the present paper are as follows:

(a) To use the MLQ in the German-speaking area for the first time and scrutinise its psychometric quality.

(b) To analyse how transformational and transactional leadership dimensions are related to subordinate ratings of effectiveness and objective performance indicators of branches of banking organisations.

(c) To use data from several comparable but independent organisations.

The empirical work on the relation between leadership and effectiveness discussed earlier suggests the following working hypotheses:

H1: Transformational leadership has an effect on performance over and above the effect of transactional leadership. This is referred to as the "augmentation hypothesis" in the transformational leadership theory (Bass, 1985).

H2: Within the transactional leadership realm, Contingent Reward leadership is positively associated with performance, and Management by Exception is not related to performance.

H3: Transformational/transactional leadership scales are more strongly related to subordinate ratings of effectiveness than to objective measures.

As regards the distinction between long-term and short-term performance the present paper should be viewed as a potential source for generating new hypotheses.

METHOD

Sample

Our sample consisted of 1456 direct reports of branch managers of 116 branches who were employed by 20 different Austrian banks. The survey was organised as follows: each subordinate was informed about the objectives of the study and was obliged to participate, although no formal check was carried out. An anonymous treatment of questionnaires was guaranteed. Each questionnaire was mailed directly and individually to us. About 90% of the subordinates actually participated.

Employees without leadership responsibilities evaluated branch managers or division managers who were their direct superiors. The branch managers and division managers in turn rated either one or both of their immediate directors. The numbers of questionnaires, raters, and leaders are presented in Table 1.

Organisational Context

It is important to note that the organisations studied are almost identical with respect to several basic business conditions. For example, they have a common range of products; common training, payment, and promotion

TABLE 1
Number of Questionnaires for Different Types of Leaders

Type of Leader	*Number of Leaders*	*Number of Questionnaires*	*Average Number of Questionnaires per Leader*
Director	45	760	17
Division manager	54	253	5
Branch manager	116	443	4
Total	215	1456	7

schedules; coordinated marketing policies; and a far-reaching common organisational system. Nevertheless, the management of each bank acts completely independently and has the sole responsibility for its performance.

However, the performance of individual branches is influenced by the local business environment, i.e. customer characteristics and local market conditions. We have attempted to account for the effect of these factors in the computation of our objective performance indicators.

Leadership Measures

Leadership is measured on the basis of the seven scales of the Multifactor Leadership Questionnaire (MLQ; form 5R). Each dimension is represented by several items, using a 5-step scale ranging from "not at all" to "frequently, if not always".

To our knowledge this is the first study that is based on a German translation of the MLQ. Each item was first translated from English to German and then—independently—retranslated back to English. This sequence was repeated in order to refine item phrasing and to obtain an accurate translation. However the question can be raised whether the translated questionnaire can be considered an adequate instrument, and whether it still represents the constructs conceptualised by its originators. Despite careful checking for translation accuracy, some of the subtleties of item phrasing may have got lost during the translation process. Therefore one objective of our data analyses was to investigate whether the factor structure obtained by Bass and Avolio (1990a) can still be supported when a German translation of the questionnaire is used.

Performance Indicators

The objective performance indicators were developed in close collaboration with bank managers to provide them with suitable guidelines for planning purposes, target assignments, and performance comparisons among

branches. As the usual performance measures such as earnings or profits were not available, we started to derive performance indicators on the basis of ratios that relate the volume of production in each of nine major product categories (various savings products, loans, insurance products, other invisible services and securities) to (a) the number of checking accounts maintained in a branch and (b) the number of people living in appropriately defined areas which represent the potential number of customers of a branch.[1]

However, these ratios are not immediately suitable for performance comparisons, as branches operate in various (mainly rural) areas under rather different conditions. These differences can be attributed to characteristics of existing customers (age, income, and job distribution)[2] and market conditions (number of competitors, purchasing power and age distribution of population, the number of years a branch has been operating). Therefore, a particular branch may perform better (i.e. have higher ratios) than another branch because it operates in an area with a higher level of customers' income, or faces less competition. Thus, performance comparisons among branches have to take into account differences in customer features and market conditions.

This calls for a procedure that results in relative performance measures. Therefore, in the next step we carried out multiple regression analyses for each of the nine product categories. The ratios were used as the dependent variable and the data on customer profiles and market conditions were entered as explanatory variables in these analyses.[3] This procedure is comparable to the PIMS approach which attempts to find data-based laws of the marketplace for strategic planning purposes (see Schoeffler, Buzzell, & Heany, 1974).

A simple example should clarify our procedure and the meaning of resulting indicators. Consider a branch that operates in a low-income area with younger customers than average. Suppose its ratio for life-insurances per checking account is .03 and the average ratio for all branches is .05. The regression analysis shows that for that particular branch a ratio of .02 is an appropriate number. .02 is the fitted value from the regression that can be considered a sales target for the branch that takes into account local market conditions. Comparing the actual ratio .03 to the average .05 and concluding

[1] These survey measures were obtained during the last three months of the same year for which product data was collected.

[2] In order to obtain customer profiles, data on age, income, and employment has been collected from a sample of some 300,000 customers holding checking accounts in the regions where the banks are located. Again this data was collected during the last three months of the same year for which product data was obtained.

[3] The variables that best represented customer and market factors were selected by a stepwise regression procedure.

a bad performance is inappropriate because specific conditions of the branch are ignored. However, the difference between the current ratio .03 and the fitted ratio .02—the regression residual—can serve as a suitable performance indicator for this branch. It shows whether a branch over- or underperforms reasonable targets. We use the standardised residuals of these regressions as a measure of performance for each branch. This measure is free from any effects induced by customer features or market conditions.

The residuals are relative performance indicators that are assumed to be affected by the leadership behaviour of branch managers and directors. We are aware that factors other than leadership may also be reflected in these indicators. However, during several workshops with branch managers and directors we found that the relative performance measured with these indicators conforms well to the expectations and subjective rankings of leaders.

Finally the nine performance indicators were split into two groups: the first group contains indicators for bank products where the underlying ratios are based on stock data (e.g. the volume of savings and loans; six items). Therefore, these performance measures describe the effectiveness of the branch as a result of the staff members' and leaders' activity over a longer period of time. The second group of indicators is based on flow data (e.g. the annual production of insurance products and subsidised housing loans; three items). These performance measures describe the effectiveness of the branch as a result of the staff members' and leaders' activity during one year. Moreover, special emphasis was put on promoting and enforcing sales of this product group during the period when data had been collected.

Subordinate-based performance is measured by three extra effort items contained in the MLQ. According to Bass (1985), extra effort is a central outcome of transformational leadership and shows how a leader motivates subordinates to perform beyond expectations.

Analyses

Structural Equation Models. We use structural equation models (LISREL)—which have become a standard research tool in social sciences—to carry out confirmatory factor analyses, to estimate correlations among leadership scales, to obtain measures of reliability for both leadership and performance indicators, and to estimate structural relations between leadership and performance. Structural equation models contain measurement equations that relate observed indicators and latent variables, and structural equations that describe the relationship among latent variables. Confirmatory factor analyses are based on measurement equations alone.

The structural relations between latent performance and leadership variables have the form of a regression equation. The direct path coefficients in this equation are estimated for each of the three performance variables in turn. In order to test the augmentation effect (H1) an hierarchical regression analysis was carried out. In each step of this analysis, leadership variables were entered cumulatively in a prespecified sequence to the regression equation. The increase in the coefficient of determination R^2 from one step to the next was used to test for the augmentation effect.

In order to measure the goodness-of-fit obtained in confirmatory factor analyses and structural equation models we use the usual χ^2 statistic, Akaike's Information Criterion (AIC), and the Root Mean Square Residual (RMSR). In addition we present some commonly used incremental fit indices (see Hoyle & Panter, 1995, p.166): the Tucker-Lewis Index (TLI), the Normed Fit Index (NFI), and the Comparative Fit Index (CFI).

Reliability. Reliability of leadership and performance scales is assessed in several ways. Cronbach's coefficient alpha is provided as a standard measure of reliability. The results from estimating structural equation models are used to calculate two reliability measures suggested by Fornell and Larcker (1981); the construct reliability and the average variance extracted.

In addition we measure the degree of consensus among subordinates' ratings by calculating the interrater reliability according to James, Demaree, and Wolf (1984). This measure accounts for potential response biases generated by central tendency, or social desirability. For each potential bias a corresponding null distribution of responses is constructed. The resulting interrater reliabilities measure the degree of consensus relative to the corresponding prior distribution. In this reliability analysis we included only leaders for which ratings of at least five subordinates were available.

Sample Size. Confirmatory factor analyses of the MLQ are based on all available ratings. Because of missing values only 1161 observations were used in this analysis (listwise deletion). For the purpose of testing the relationship between leadership and performance, only ratings of branch managers were used. This is because there is only one set of performance indicators for each branch that cannot be meaningfully assigned to directors and division managers who are responsible for several branches. Therefore questionnaires for these leaders had to be dropped from the analysis. When more than one questionnaire was available for a branch, the data on performance measures was expanded accordingly. A total of 376 observations is used in this analysis.

RESULTS

Confirmatory Factor Analyses of the MLQ

As we were using a German translation of the MLQ, we first investigated whether the factor structure obtained by Bass and Avolio (1990a) could still be supported. For that purpose we carried out a confirmatory factor analysis, where each item is associated with only one latent leadership variable according to the assignments of Bass and Avolio.

Reliability of MLQ Scales. Alpha coefficients for each scale indicate acceptable reliability, although alpha values for Management by Exception and Laissez-Faire are somewhat lower (see Table 2). The alpha values we obtain are even higher than those reported by Bass and Avolio (1990a). Construct reliability derived from results of confirmatory factor analysis are almost as high as the alpha coefficients. The average variance extracted reveals a worse picture, however. The values for transformational scales and Contingent Reward are low, but just acceptable. However, average variance extracted for Management by Exception and Laissez-Faire scales is clearly below the .5 cutoff.[4] The distinctly lower reliabilities of these scales are probably due to raters' difficulties in articulating socially undesirable behaviour.

This explanation is not supported by the interrater reliabilities, however. First, interrater reliabilities for Management by Exception and Laissez-Faire are almost equal to those of other scales. Second, assuming a "central tendency" null response distribution—corresponding to subordinates' tendency to avoid extreme, socially undesirable ratings—indicates lower reliability similar to other scales, and still at acceptable levels.

Another explanation is obtained when considering items with the worst reliabilities[5] (e.g. "focuses attention on irregularities, mistakes, exceptions, and deviations from what is expected of me"; "avoids telling me how to perform my job"; "makes me feel that whatever I do is OK with him/her"). These items refer to situations where leaders may behave differently depending on the rater's performance. Leaders may react to substandard performance of subordinates e.g. by closer supervision (McFillen, 1978), by more directive-structuring and critical-punitive behaviour (Szilagyi, 1980) or more autocrative behaviour (Barrow, 1976). Thus, varying performance

[4] It should be noted that no items were deleted before the average variance extracted was calculated—for instance, because estimated loadings were low or insignificant.

[5] Reliability for individual items is assessed using $R_i^2 = 1 - \widehat{\mathrm{var}}[\epsilon_i]/\widehat{\mathrm{var}}[x_i]$ (see Bollen, 1989, pp.221, 288) and measures the amount of variance of indicator i explained by the corresponding latent variable. $\widehat{\mathrm{var}}[\epsilon_i]$ and $\widehat{\mathrm{var}}[x_i]$ denote the estimated variance of residuals and indicators, respectively.

TABLE 2
Reliability of MLQ Scales

	Alpha	Construct Reliability	Average Variance Extracted	Interrater Reliability		
				Uniform	Central	Skew
Charisma	.93	.93	.59	.88	.77	.65
Inspirational Motivation	.88	.86	.47	.88	.77	.65
Intellectual Stimulation	.93	.91	.52	.92	.82	.67
Individualised Consideration	.92	.91	.49	.90	.76	.65
Contingent Reward	.92	.91	.54	.88	.77	.61
Management by Exception	.77	.71	.26	.91	.82	.67
Laissez-Faire	.71	.71	.22	.88	.77	.61

Uniform, central, and skew refer to the prior response distribution that is used in the calculation of the respective interrater reliability. The "uniform" entries are based on the usual assumption of a uniform prior. "Central" accounts for response biases due to central tendency. "Skew" is based on a prior response distribution that is skewed, possibly because of social desirability.

of raters may introduce more variation in these items and therefore lower reliability.

This interpretation is supported by the fact that other items of Management by Exception and Laissez-Faire scales that refer to leaders' interference in exceptional cases have higher reliability (e.g. "is content to let me do my job the same way as I've always done it, unless changes seem necessary"; "avoids intervening except when I fail to meet objectives"). The same applies to items that refer to behaviour which seems to be relatively independent of the performance level of the staff member (e.g. "avoids making decisions"; "is hard to find when a problem arises").

Model Fit. We found a rather poor fit of the measurement model indicated by the high χ^2 statistic and the goodness-of-fit indices which are all below acceptable levels (see Table 3). We explain the weak support for the MLQ by the fact that the MLQ factor structure has been mainly derived on the basis of exploratory factor analyses. As confirmatory factor analyses impose much stronger restrictions than exploratory analyses, the poor fit is no surprise. Nevertheless, it remains an open question whether this result is specific to the present data set, the German translation of the MLQ, or a general weakness of the factor structure of the MLQ.

Correlations Among MLQ Scales. We obtained extremely high correlations among transformational scales (see Table 3). This result agrees with many other studies (e.g. Avolio, Waldman, & Einstein, 1988; Bass, 1985, 1990; Bass & Avolio, 1990b, 1993; Hater & Bass, 1988; Yammarino & Bass, 1990). In these studies, correlations range from .70 to .91 for subordinates' reports and from .49 to .73 for superiors' reports. We found

TABLE 3
Correlation Among MLQ Scales

	1	2	3	4	5	6	7
1. Charisma	1.00	.85	.80	.81	.73	−.37	−.60
2. Inspirational Motivation	.98	1.00	.85	.69	.68	−.43	−.57
3. Intellectual Stimulation	.89	.97	1.00	.69	.67	−.39	−.55
4. Individualised Consideration	.88	.84	.77	1.00	.80	−.26	−.53
5. Contingent Reward	.81	.83	.76	.89	1.00	−.28	−.50
6. Management by Exception	−.20	−.32	−.29	−.11	−.23	1.00	.54
7. Laissez-Faire	−.76	−.77	−.72	−.64	−.61	.59	1.00
Number of items	10	7	10	10	9	10	10
Mean	3.24	3.18	2.98	3.07	2.68	3.03	2.62
Standard deviation	1.01	0.90	0.85	0.91	0.92	0.62	0.63

Correlations above the main diagonal as well as mean and standard deviation are calculated from observed scores. Correlations below the main diagonal are estimated by LISREL.

Fit indices for LISREL model: χ^2: 16082, df: 2058, $P < .0001$, AIC: 11966, RMSR: .1138, TLI: .7149, NFI: .6989, CFI: .7264, $N = 1161$.

high positive correlation between transformational scales and Contingent Reward. Other studies also show that transformational and transactional leadership scales are positively related, but the correlations are distinctly lower. On average, the transformational factors' correlation with Contingent Reward is about .55. The negative correlations we obtained for Management by Exception and Laissez-Faire leadership are in the expected range. Other studies find an average correlation of −.2 of transformational scales with Management by Exception and values ranging from −.47 to −.57 for Laissez-Faire leadership (Bass, 1985, 1990; Bass & Avolio, 1990b, 1993; Kuhnert & Lewis, 1987; Waldman et al. 1990; Yammarino & Bass, 1990).

The results obtained from the confirmatory factor analysis are not directly comparable to other studies, however. Correlations are frequently based on observed scales, whereas we estimated correlations among latent constructs. For comparison we also provide the correlations obtained from observed scales in Table 3. These correlations are very close to those obtained in the studies mentioned earlier.

The high correlation among transformational scales is not a completely unexpected result. We point out that the original factor analyses upon which the MLQ was developed used two samples of a total of 196 US Army colonels. A higher-order factor analysis disclosed two clusters of factors: an active higher-order dimension included Charisma, Inspirational Motivation, Intellectual Stimulation, and Individualised Consideration. Contingent Reward, Management by Exception, and Laissez-Faire leadership were identified as the passive dimension of leadership (Bass, 1985).

Despite the fact that later analyses showed that the overall factor

structure continues to provide a meaningful framework, the original differentiation into seven scales could never again be verified in the same form. Hater and Bass (1988) and Waldmann et al. (1987) were able to identify five out of the seven factors. Bass and Avolio (1993) refer to several unpublished studies where one single transformational factor instead of four emerged. High correlations—approaching or exceeding .9—among transformational scales were also found in a study by Podsakoff, MacKenzie, Moorman, and Fetter (1990). Other research led to a subdivision of Contingent Reward into two factors (Howell & Avolio, 1993) and to a division of Management by Exception into two parts, respectively (Bass, 1990; Hater & Bass, 1988). Bass (1995) also reports that none of the later analyses led to a differentiation between Charisma and Inspirational Motivation.

Modified MLQ Scales. The results obtained so far indicate (a) that the factor structure of the MLQ cannot be maintained and (b) the need to construct new latent variables out of the highly intercorrelated subscales. For that purpose we split the sample randomly into two parts.[6] One half of the data was used as a calibration sample for which a new factor structure was derived on the basis of exploratory and confirmatory factor analyses. The second half of the sample was used for cross-validation.

In the exploratory factor analysis a preliminary classification of items led to four factors. Loadings had to be greater than .50 and to differ at least .20 to the loading of another factor. In a confirmatory factor analysis this preliminary structure was further refined on the basis of residual analysis and modification indices. Some items were dropped or reassigned to other factors. In addition, we introduced some parameters to account for the covariance among error terms of measurement relations. Such was only the case for similarly phrased items associated with the same factor.

The cross-validation of the modified factor structure was based on the second half of the sample. In the validation model all parameters (loadings, factor correlations, error variances, and error covariances) were set equal to the estimated parameters of the calibration model. Table 4 provides fit indices for the two models. One can see acceptable indices for the calibration sample that change only slightly in the validation model. Therefore we conclude that we were able to cross-validate the modified MLQ across the two samples.

The modified four-factor structure—details of which are available from the authors—still maintains the basic distinction into transformational and

[6] The split was based on the sequence number of the questionnaires. The two samples were found to be comparable in terms of the relative frequency of each type of leader and the composition of branches and banks.

TABLE 4
Fit Indices for the Modified MLQ from the Calibration and Validation Samples

	χ^2	*df*	*P<*	*AIC*	*RMSR*	*TLI*	*NFI*	*CFI*	*N*
Calibration	1701	546	.0001	609	.0588	.9062	.8786	.9139	608
Validation	2090	595	.0001	900	.0733	.8883	.8505	.8883	618

transactional leadership. The interpretation of the new scales agrees to a large extent with the corresponding original scales. In the modified MLQ the four original transformational scales are collapsed into two scales and the Laissez-Faire scale is dropped. Out of the remaining 35 items 83% are still associated with their original MLQ scales. In the remaining cases the new assignments do not strongly affect the original interpretation of scales. The original scales' labels are therefore maintained as far as possible. The modified four-factor structure can be described as follows:

The factor "Core Transformational Leadership" (11 items) includes indicators from the original scales Charisma (2 items), Inspirational Motivation (3 items), and Intellectual Stimulation (5 items). In addition it contains one item from the original Individualised Consideration scale which may be interpreted as intellectually stimulating ("gets me to look at problems as learning opportunities").

The second factor "Individualised Consideration (mod)" (8 items) is mainly constituted by the original Individualised Consideration scale (5 items) and contains three items from the original Charisma scale that emphasise relational aspects (e.g. "makes me feel good when I am around him/her"; "I have complete faith in him/her"). They express whether the leader creates an environment of emotional support, warmth, friendliness, and trust. This behaviour is consistent with leaders who provide coaching and teaching, and treat each follower as a respected individual.

The third factor "Contingent Reward (mod)" (10 items) mainly contains items from the original Contingent Reward scale (7 items) and is enhanced by three items of the original Individualised Consideration scale which express verbal feedback (e.g. "finds out what I want and helps me to get it"; "lets me know how I am doing"; "expresses appreciation when I do a good job"). Thus the contents of Contingent Reward in the sense of a positively reinforcing interaction is not altered.

The fourth factor is called "Management by Exception passive" (6 items). Hater and Bass (1988) were the first to distinguish among active and passive Management by Exception. In its active form the leader arranges to actively monitor deviances from standards, mistakes, and errors in the follower's assignments and to take corrective action as necessary. The passive form implies waiting passively for deviances and then taking corrective action. As four of the items of this new scale refer to passive leader behaviour (e.g. "avoids intervening except when I fail to meet objectives") and as it contains two items of the original Laissez-Faire scale (e.g. "avoids getting involved in our work") the term "Management by Exception passive" seems appropriate.

The modified MLQ consists of only 35 items. Many of the eliminated items can be viewed as redundant as there exist other items which refer to the same behavioural aspect.[7] We further note that all attitude-based items (e.g. "has my respect") were eliminated during the calibration procedure. As the MLQ has been frequently criticised because it contains a mixture of behaviour-based questions and attitude-based questions (e.g. Hunt, 1991) the modified MLQ represents an improvement in this respect. Only two of the Laissez-Faire items are still used. In the majority of studies using the MLQ this scale is not used at all (Lowe & Kroeck, 1995). Moreover, this scale was not considered in the explorative factor analyses that lead to the MLQ (Bass, 1985) and had only subsequently been included (Bass & Avolio 1990a). Thus, ignoring this scale is not considered a significant loss of information.

The correlations obtained on the basis of the modified scales for the whole sample are presented in Table 5. It turns out that the highest correlations there are considerably lower than among the original scales. Likelihood ratio tests show that each correlation is significantly different from 1.0, which indicates an improvement in discriminant validity. Further support for the modified factor structure is provided by the goodness-of-fit indices. Finally, Table 5 indicates that reliability indices have generally improved, in particular for Management by Exception passive. Summing up, we consider the modified MLQ as preferable to the original MLQ for our purposes.

Leadership and Performance

In this section we provide empirical evidence on the relation between leadership and performance. As the analysis of the relation between leadership and performance is based on branch data only, we first conducted a confirmatory factor analysis for that subset.

Reliability and Correlations. The reliability of leadership scales for branch data turns out to be better than their reliability for the whole dataset (see Table 6). This can be easily explained by the greater homogeneity in this

[7] We mention some examples for these redundancies. The item "praises me when I do a good job" has been eliminated but the leader behaviour it refers to is also expressed in "commends me when I do good work" or "expresses appreciation when I do a good job". Another example is the eliminated item "sets high standards" which is also covered by "communicates expectations of high performance to me".

TABLE 5
Correlation Among and Reliability of Modified MLQ Scales

	1	2	3	4
1. Core Transformational Leadership	1.00	.58	.60	−.16
2. Individualised Consideration (mod)	.68	1.00	.70	.06
3. Contingent Reward (mod)	.70	.80	1.00	−.13
4. Management by Exception passive	−.23	.02	−.21	1.00
Number of items	11	8	10	6
Mean	3.15	3.29	2.61	3.46
Standard deviation	.91	.98	.92	.88
Alpha	.93	.92	.93	.85
Construct reliability	.91	.90	.92	.83
Average variance extracted	.50	.54	.53	.45
Interrater reliability—uniform	.91	.87	.89	.83
Interrater reliability—central	.82	.73	.78	.73
Interrater reliability—skew	.68	.63	.61	.58

Correlations above the main diagonal as well as mean and standard deviation are calculated from observed scores. Correlations below the main diagonal are estimated by LISREL.

Fit indices for LISREL model: χ^2: 3054, df: 546, $P < .0001$, AIC: 1962, RMSR: .0608, TLI: .8978, NFI: .8883, CFI: .9062, $N = 1226$.

data-subset where only ratings for branch managers are considered, whereas ratings of three different types of directors and managers were used in the earlier analyses. The reliability of the three performance scales is generally high. They are only weakly correlated among each other, which indicates that we are using three rather different aspects of performance. The fit indices are somewhat lower than for the whole dataset.

Correlations among leadership scales obtained from the branch subset are similar to or larger in absolute terms than those from the complete data set. Transformational scales and Contingent Reward are positively correlated with performance, but Individualised Consideration is unrelated to long-term performance. Management by Exception passive is negatively correlated with all performance variables. Finally, we note support for hypothesis H3: leadership variables correlate much more strongly with extra effort than with objective performance.

Structural Relations. We estimated structural relations between latent performance and leadership variables on the basis of an hierarchical regression analysis. In the first step, the two transactional scales Contingent Reward and Management by Exception passive were put at once into the structural equation. In step two, Individualised Consideration was entered. Adding Core transformational Leadership in the final step provides a complete model for analysing the relation between performance and leadership. Table 7 contains the standardised path coefficients obtained in

TABLE 6
Correlations Between Modified MLQ Scales and Performance for Branch Data

	Modified MLQ				Performance		
	1	2	3	4	5	6	7
1. Core Transformational Leadership	1.00	.59	.76	−.34	.84	.28	.26
2. Individualised Consideration (mod)	.65	1.00	.72	−.07	.69	.04	.16
3. Contingent Reward (mod)	.84	.79	1.00	−.32	.82	.17	.21
4. Management by Exception passive	−.40	−.14	−.41	1.00	−.30	−.13	−.19
5. Extra effort	.92	.77	.88	−.36	1.00	.19	.22
6. Objective performance (long-term)	.33	.07	.21	−.19	.24	1.00	.44
7. Objective performance (short-term)	.24	.21	.20	−.22	.25	.40	1.00
Mean	3.15	3.51	3.00	3.16	2.96	.22	.05
Standard deviation	.97	.89	.96	1.02	1.12	.88	.89
Alpha	.95	.91	.94	.89	.92	.89	.83
Construct reliability	.95	.90	.94	.90	.92	.90	.87
Average variance extracted	.62	.53	.60	.60	.79	.61	.71

Correlations above the main diagonal as well as mean and standard deviation are calculated from observed scores. Correlations below the main diagonal are estimated by LISREL.

Fit indices for LISREL model: χ^2: 2840, df: 998, $P < .0001$, AIC: 844.4, RMSR: .0683, TLI: .8753, NFI: .8338, CFI: .8849, $N = 376$.

each of these steps together with R^2, incremental R^2, and F-statistics for testing the significance of R^2 increments. The goodness-of-fit measures for the complete model are very close to those obtained in the confirmatory factor analyses (see Tables 5 and 6).

The main conclusion we draw from Table 7 is a confirmation of the augmentation hypothesis (H1): transformational scales have an effect on performance over and above the effect of transactional scales. For all performance aspects we find significant increases in R^2 when transformational scales are added. There is some need for differentiation, however. For extra effort and long-term performance, Core Transformational Leadership contributes more to the total augmentation effect than Individualised Consideration. For short-term performance, their contribution is almost equally strong. This conclusion is supported when the sequence of entry is changed and Core Transformational Leadership is entered before Individualised Consideration. In this case the R^2s for the three-variable models from step two are .889, .123, and .081.

When interpreting path coefficients of the complete four-variable models one has to take into account that some of the correlations among leadership variables are still substantial. For extra effort, for instance, multicollinearity appears to reduce the coefficient of Contingent Reward as Individualised Consideration and Core Transformational Leadership are entered. The results for Management by Exception passive seem to be rather unaffected by multicollinearity: it is unrelated to extra effort and long-term

TABLE 7
Standardised Path Coefficients from Hierarchical Regression Analysis Relating Leadership Variables to Performance

	extra effort				*long-term*				*short-term*			
step	*1*	*2*	*3*		*1*	*2*	*3*		*1*	*2*	*3*	
Core Transformational Leadership			.619**				.478**				.283**	
Individualised Consideration (mod)		.208**	.233**			-.228*	-.216*			.212*	.240*	
Contingent Reward (mod)	.887**	.701**	.174*		.176**	.375**	-.037		.138*	-.048	-.310*	
Management by Exception passive	-.001	-.047	-.006		-.111	-.058	-.041		-.159**	-.216**	-.204**	
R^2	.787	.798	.905		.059	.074	.138		.062	.081	.106	
Incremental R^2		.011	.107			.015	.064			.019	.025	
F-test for incremental R^2		20.3**	419**			6.04*	27.6**			7.71**	10.4**	

Fit indices for the LISREL model of the final step: χ^2: 2842, df: 997, $P < .0001$, AIC: 847.6, RMSR: .0682, TLI: .8750, NFI: .8337, CFI: .8848, $N = 376$.
* $P < .05$; ** $P < .01$

performance but has a significant negative effect on short-term performance.

For short-term performance the sequence of path coefficients—as additional variables are entered—lends to the conclusion that Contingent Reward may have a suppressing effect on both transformational scales.[8] Finally, Table 7 shows that leadership variables account for almost all of the variance in extra effort. The picture is quite different for objective performance, where in the complete models only about 12% of the variance is accounted for by leadership variables.

DISCUSSION

The results of the present study allow conclusions regarding (a) the factor structure of the German MLQ, and (b) the relation between leadership and performance. The inadequate discriminant validity and other deficiencies of the original MLQ required the development of a new factor structure. The newly derived and cross-validated modified MLQ still allows for the distinction into the two basic dimensions of the theory: transformational and transactional leadership. A differentiation into seven subdimensions—as derived by Bass and Avolio (1990a)—cannot be maintained, however.

As regards the relations between leadership and performance we find support for the augmentation hypothesis (H1) that transformational leadership has an effect on performance over and above the effect of transactional leadership. The following findings may deserve special attention for generating new hypotheses: Core Transformational Leadership appears to be more strongly related to long-term than to short-term performance. Individualised Consideration is positively related to short-term performance but negatively related to long-term performance. How can these differences be explained?

Individualised Consideration behaviour (e.g. "provides advice to me when I need it") refers to situations where leader actions can have an immediate impact on what is done and how it is done. It therefore seems obvious that short-term performance (measured in terms of flow data) reacts strongly to Individualised Consideration. In the context of long-term performance (measured in terms of stock data) an explanation of the estimated negative relation may require one to take into account the experience of subordinates and the associated—inversely related—need for consideration. Branches with a high level of long-term performance may have well trained and experienced employees who require little

[8] The conditions for suppression—correlation patterns and sign reversals—are fulfilled (see Cohen & Cohen, 1983, p.94). In addition, the path coefficients of the two transformational leadership dimensions become smaller when Contingent Reward is omitted from the regression equations.

consideration. Conversely, a low performance level may be due to deficiencies in the qualification of employees. These require appropriate leader activities to compensate or eliminate such deficits, which implies that leaders may spend more time on coaching and teaching than in branches with high long-term performance. Therefore the conjecture of Bass and Avolio (1990b, p.242) that transformational leaders "gain greater levels of long-term performance" than transactional leaders may be too general and may only hold for core transformational dimensions but not for Individualised Consideration.

This result also bears some relation to results obtained in the context of the Ohio State leadership studies. As the modified Individualised Consideration scale now also contains items of the original Charisma scale (e.g. "makes me feel good when I'm around him or her") it is more similar to the Consideration scale of the Ohio State studies than the original MLQ scale. Hundreds of studies have been conducted on the effects of Consideration, but the results for most criteria were inconsistent or inconclusive (Bass, 1990, Kerr & Schriesheim, 1974; Yukl, 1971). Given our results, it appears reasonable that the time dimension has to be accounted for when the relation between (Individualised) Consideration and performance is analysed. This might resolve some of the inconsistencies found in previous studies.

Beyond that, this result could also be a consequence of the specific context of our study: in banking organisations we are primarily concerned with bureaucratic process cultures (Deal & Kennedy, 1982) which rely on a "Mediating Technology" (Thompson, 1967) or a "Routine-Technology" (Perrow, 1970). According to Deal and Kennedy such cultures are characterised by employees who do not consider it important "what they do" but "how they do something" (Deal & Kennedy, 1982, p.119). The values of such organisations focus primarily on technical perfection. Employees are only appreciated when they get the process and the details right. From these considerations and from our empirical findings it could be concluded that those leadership behaviours are most effective that are in a clear atmospheric contrast to the formalisation and instrumental aspects in such organisations. Such a contrast may be invoked by key aspects of core transformational behaviour: strengthening values, loyalty and concern for continuous improvement, communicating an inspiring vision, or arousing followers to think in new ways. As Core Transformational Leadership is mainly concerned with influencing corporate culture and strategies for attaining organisational objectives, it is not surprising to find a comparatively stronger impact of this behaviour on long-term than on short-term performance.

As regards transactional behaviour, hypothesis H2 states that Contingent Reward leadership is positively associated with performance. As it turns out,

this is only true for extra effort but not for objective performance. Contingent Reward is unrelated to long-term performance. Its significant negative relation with short-term performance may be due to suppressor effects which are difficult to explain and would require further study. However, our result agrees with Howell and Avolio's (1993) study, which used a rather objective performance measure. One of their explanations for the negative path coefficients between performance and Contingent Reward was that such leader behaviour also means supervising followers' behaviour rather than rewarding it: "If followers perceive leaders as restricting their freedom of action, then it is possible that followers' motivational levels might decline" (Howell & Avolio 1993, p.399). We conclude that within the scope of highly trusting and inspiring interaction relationships—which contrast bureaucratic process cultures—Contingent Reward behaviour can be ineffective or counterproductive.

Hypothesis H2 further states that Management by Exception is not related to performance. We find agreement with this hypothesis for extra effort and long-term performance, only. For short-term performance, the evidence contradicts hypothesis H2: Management by Exception passive has a significant negative effect on this performance aspect. This may be explained by the importance of supporting and advising behaviour in this context. In addition, the leader's passive waiting for deviances implies that subsequent corrective actions come too late in the short run.

We finally note support for the hypothesis that leadership scales are more strongly related to subordinate-based performance measures than to objective measures (H3). This result agrees with many other empirical studies and deserves no further discussion given the problems associated with subjective measures mentioned earlier.

Referring to MLQ research, Avolio, Yammarino, and Bass (1991, p.571) state that "in the social and behavioural sciences, data collection concerning multiple constructs from a single source (rater) is quite common" but the validity of such data presents "an unresolved sticky issue" for behavioural science researchers. A strength of the present study is its use of objective performance indicators that have been obtained independently from the raters. A central dilemma of research on leader behaviour that could not be resolved in this study is the tendency of respondents to attribute desirable behaviour to a leader who is perceived as effective, even though such behaviour is not actually observed (Lord & Maher, 1991). The perceived performance level may therefore contribute to systematic response patterns that inflate the significance of effects. Meindl (1990) refers to "hyperromanticism" and shows that a connection between romanticisation of leadership and the tendency to "see" more transformational qualities of leaders exists. This aspect should also be taken into account when interpreting our results.

From the results of this study two aspects for further research can be derived: first, the divergent results for short- and long-term performance suggest that one should explicitly take into account the time-frame when analysing objective performance and formulating hypotheses. Second, the relation between performance, transformational leadership, and information processing requires further analysis. We believe that future empirical research has to cope with the duality of transformational leadership both as a determinant of performance and as a perceptual phenomenon that unfolds over time.

Manuscript received April 1995
Revised manuscript received September 1996

REFERENCES

Avolio, B.J., Waldman, D.A., & Einstein, W.O. (1988). Transformational leadership in a management game simulation. *Group and Organization Studies*, *17*, 571–587.

Avolio, B.J., & Yammarino, F.J. (1990). Operationalizing charismatic leadership using a levels-of-analysis framework. *Leadership Quarterly*, *1*, 193–208.

Avolio, B.J., Yammarino, F.J., & Bass, B.M. (1991). Identifying common methods variance with data collected from a single source: An unresolved sticky issue. *Journal of Management*, *17*, 571–587.

Barrow, J.C. (1976). Worker performance and task complexity as causal determinants of leader behavior style and flexibility. *Journal of Applied Psychology*, *61*, 433–440.

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. (1995). Theory of transformational leadership redux. *Leadership Quarterly*, *6*, 463–478.

Bass, B.M., & Avolio, B.J. (1989). Potential biases in leadership measures: How protoypes, leniency, and general satisfaction relate to ratings and rankings of transformational and transactional leadership constructs. *Educational and Psychological Measurement*, *59*, 508–527.

Bass, B.M., & Avolio, B.J. (1990a). *Transformational leadership development: Manual for the multifactor leadership questionnaire*. Palo Alto, CA: Consulting Psychologists Press.

Bass, B.M., & Avolio, B.J. (1990b). The implications of transactional and transformational leadership for individual, team, and organizational development. *Research in Organizational Change and Development*, *4*, 231–272.

Bass, B.M., Avolio, B.J., & Goodheim, L. (1987). Biography and the assessment of transformational leadership at the world class level. *Journal of Management*, *13*, 7–19.

Bass, B.M., & Avolio, B.J. (1993). Transformational leadership: A response to critiques. In M.M. Chemers & R. Ayman (Eds.), *Leadership theory and research. Perspectives and directions*. San Diego, CA: Academic Press; Harcourt Brace Jovanovich.

Bass, B.M., & Yammarino, F.J. (1991). Congruence of self and others' leadership ratings of naval officers for understanding successful performance. *Applied Psychology: An International review*, *40*, 437–454.

Binning, J.F., Zaba, A.J., & Whattam, J.C. (1986). Explaining the biasing effects of performance cues in terms of cognitive categorization. *Academy of Management Journal*, *29*, 521–535.

Bollen, K.A. (1989). *Structural equations with latent variables*. New York: Wiley.

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

Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences* (2nd Edn). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Coleman, E., Patterson, P., Bryan Fuller J., Hester, K., & Stringer, D.Y. (1995). *A meta-analytic examination of leadership style and selected follower compliance outcomes*. Unpublished manuscript. Alabama.

Deal, T.E., & Kennedy, A.A. (1982). *Corporate cultures. The rites and rituals of corporate life*. Reading, MA: Addison-Wesley.

Eden, D., & Leviatan, U. (1975). Implicit leadership theory as a determinant of the factor structure underlying supervisory behavior scales. *Journal of Applied Psychology*, *60*, 736–741.

Fornell, C., & Larcker, D.F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, *18*, 39–50.

Hater, J.J., & Bass, B.M. (1988). Supervisors' evaluations and subordinates' perceptions of transformational and transactional leadership. *Journal of Applied Psychology*, *73*, 695–702.

Howell, J.M., & Avolio, B.J. (1993). Transformational leadership, transactional leadership, locus of control and support for innovation: Key predictors of consolidated business unit performance. *Journal of Applied Psychology*, *78*, 891–902.

Hoyle, R.H., & Panter, A.T. (1995). Writing about structural equation models. In R.H. Hoyle (Ed.), *Structural equation modeling*. Thousand Oaks, CA: Sage.

Hunt, J.G. (1991). *Leadership: A new synthesis*. London: Sage.

James, L.R., Demaree, R.G., & Wolf, G. (1984). Estimating within-group interrater reliability with and without response bias. *Journal of Applied Psychology*, *69*, 85–98.

Keller, R.T. (1992). Transformational leadership and the performance of research and development project groups. *Journal of Management*, *18*, 489–501.

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

Kuhnert, K.W., & Lewis, P. (1987). Transactional and transformational leadership: A constructive/developmental analysis. *Academy of Management Review*, *12*, 648–657.

Larson, J.R., Lingle, J.H., & Scerbo, M.M. (1984). The impact of performance cues on leader-behavior ratings. *Organizational and Human Performance*, *33*, 323–349.

Lord, R.G., Binning, J.F., Rush, M.C., & Thomas, J.C. (1978). The effect of performance cues and leader behavior on questionnaire ratings of leadership behavior. *Organizational Behavior and Human Performance*, *21*, 27–39.

Lord, R.G., & Maher, K.J. (1991). *Leadership and information processing. Linking perceptions and performance*. London, Sydney, Wellington: Harper Collins.

Lowe, K.B., & Kroeck, K.G. (1995). *Effectiveness correlates of transformational and transactional leadership: A meta-analytic review*. Unpublished manuscript. Binghamton/ Miami.

McFillen, J.M. (1978). Supervisory power as an influence in supervisor–subordinate relations. *Academy of Management Journal*, *21*, 419–433.

Meindl, E. (1990). On leadership: An alternative to the conventional wisdom. In B.M. Staw & L.L. Cummings (Eds.), *Research in organizational behavior*. London: Jai Press.

Murphy, K.R., & Reynolds, D.H. (1988). Does true halo affect observed halo? *Journal of Applied Psychology*, *73*, 235–238.

Perrow, C. (1970). *Organizational analysis: A sociological view*. London: Tavistock.

Phillips, J.S., & Lord, R.G. (1982). Schematic information processing and perceptions of leadership in problem-solving groups. *Journal of Applied Psychology*, *67*, 486–492.

Podsakoff, P.M., MacKenzie, S.B., Moorman, R.H., & Fetter, R. (1990). Transformational leader behaviors and their effects on followers' trust in leader, satisfaction, and organizational citizenship behaviors. *Leadership Quarterly*, *1*, 107–142.

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 Performance*, *20*, 93–110.

Schoeffler, S., Buzzell, R.D., & Heany, D.F. (1974). Impact of strategic planning on profit performance. *Harvard Business Review*, *52*, 137–145.

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

Seltzer, J., Numerof, R.E., & Bass, B.M. (1989). Transformational leadership: Is it a source of more burnout and stress? *Journal of Health and Human Resource Administration*, *12*, 174–185.

Singer, M.P. (1985). Transformational vs. transactional leadership: A study of New Zealand company managers. *Psychological Reports*, *57*, 143–146.

Szilagyi, A.D. (1980). Causal inferences between leader reward behavior and subordinate goal attainment, absenteeism, and work satisfaction. *Journal of Occupational Psychology*, *53*, 195–204.

Thompson, J.D. (1967). *Organizations in action.* New York: McGraw-Hill.

Turban, D.B., & Jones, A.P. (1988). Supervisor–subordinate similarity: Types, effects, and mechanisms. *Journal of Applied Psychology*, *73*, 228–234.

Waldman, D.A., Bass, B.M., & Einstein, W.O. (1987). Leadership and outcomes of performance appraisal processes. *Journal of Occupational Psychology*, *60*, 177–186.

Waldman, D.A., Bass, B.M., & Yammarino, F.J. (1990). Adding to contingent–reward behavior: The augmenting effects of charismatic leadership. *Group & Organization Studies*, *15*, 381–394.

Yammarino, F.J., & Bass, B.M. (1990). Long-term forecasting of transformational leadership and its effects among naval officers: Some preliminary findings. In K.E. Clark & M.B. Clark (Eds.), *Measures of leadership*. Greensboro, NC: Center for Creative Leadership.

Yammarino, F.J., & Dubinsky, A.J. (1994). Transformational leadership theory: Using levels of analysis to determine boundary conditions. *Personnel Psychology*, *47*, 787–811.

Yukl, G.A. (1971). Toward a behavioral theory of leadership. *Organizational Behavior and Human Performance*, *6*, 414–440.