

Attributional Analysis of the Discrepancy Between Raters in Job Performance Ratings.

DIAB M. AL-BADAYNEH
DEPARTMENT OF SOCIOLOGY
MU'TAH UNIVERSITY
MU'TAH , AL-KARAK
JORDAN

Badayneh@mutah.edu.jo

ABSTRACT

This study aimed to investigate possible effects of raters' ratings {nurses' (self); supervisors' (administrator) and patients' (client)} and hospitals (government and private hospitals) on nurses' performance. This study consisted of 303 nurses, 400 patients and 60 supervisors. An instrument originally developed by Zammuto, London, & Rowland, (1982). was adapted and translated into Arabic for this study.

Significant effects were found at alpha .05 level of raters' ratings of the performance of nurses. Paired comparison, using Scheffe F-test showed significant differences between all pairs of comparisons. Differences in ratings between ratters were interpreted in the light of the attribution theory.

Introduction

A causal attribution theory suggests an explanation to the differences in attributions made by actors and observers. Attribution theory was applied in several areas like Academic success and failure (AL-Badayneh, 1992), self-estimation and depression (Hadad, 1990) and assignment of responsibility (AL-Badayneh, 1993). Observers are likely to attribute the low performance of their subordinates more to internal than to external causes, while actors tend to attribute high performance to their effort and ability (Jones & Nisbet, 1972). Jones and Nisbett hypothesized that actors attribute their actions to situational requirements, whereas observers attribute the same action to demographic dispositions. Monson and Snyder (1977) modified this assumption as follows:

Actors should make more situational attributions than should observers about behavior acts that are under situational control; by contrast, actors' perceptions of behavior that are under dispositional control ought to be more dispositional than the perceptions of observers (p. 96).

AL-Badayneh (1993) examined the differences in the attributional styles in assigning responsibility for car accidents. The findings showed significant differences between raters in assigning responsibility.

Differences between and among different raters have been explained by Harris and Schaubroeck (1988) based on their review of the literature:

1. Egocentric bias. The underlying assumption is that observer ratings, or actor ratings of performance are biased in some fashion, while similar raters (e.g., peers and supervisors) share a set of common perceptions.
2. Differences in organizational level. Zammuto has asserted that raters at different levels weigh performance dimensions differently. Raters at different levels (e.g., peers and supervisors) disagree on the overall rating, but they would agree on dimensional ratings (Zammuto, London, & Rowland, (1982).). Raters at different levels define and measure performance differently (Landy, Farr, Saal, & Freytag, 1976).
3. Observational opportunities. This suggests that peers have more opportunities to observe the ratee and at more revealing times than do supervisors (Latham & Wexley, 1982). (Harris & Schaubroeck 1988:44-47).

Herman, Dunham, & Hulin, (1975) argued that employees who held similar positions and ranks in the organizational structure express similar attitudes with the work and pay, experienced the same level of motivation, and agreed on contingencies for interpersonal behavior; and employees at the same level agreed in their description of their supervision. Herman, Dunham, & Hulin, (1975) concluded the following:

If organizational-structure characteristics are more highly related to organizational behavior than are demographic characteristics in a variety of different organizational settings, then the effect must be related to employees' ability and willingness to adapt to their work environments (p. 230).

Organizational differences are likely to affect the ability of managers to accurately interpret and compare performance ratings. The way meanings are shared, assigned, and interpreted in an organization is believed to be one of the factors which cause differences in performance ratings. Kane and Lawler (1979) found that social characteristics of an organization may impact performance appraisal.

Porter and Lawler (1965) reviewed the literature on organizational structure and employee behavior and concluded that organizational structure and levels are strongly related to both attitudes and behavior of employees. Zammuto, London, & Rowland, (1982) stated the following

Organizational differences are likely to affect the ability of managers to accurately interpret and compare performance ratings across organizational settings. Such efforts are likely to be complicated because contextual differences cause raters to place different emphases on specific performance criteria (p. 644).

Mean differences between raters were studied by Holzbach (1978), who studied supervisors, self, and peer performance ratings of 107 managerial and 76 professional employees in a medium-sized manufacturing location. Holzbach reported that the mean of nurse-ratings was greater than the mean of supervisors and greater than the mean of peer ratings. The mean of peer ratings was greater than the mean of supervisors ratings. Shore and Thornton III (1986). Findings showed that subordinates' self-ratings were higher than their supervisors' ratings.

A Meta-analysis was conducted by Harris and Schaubroeck (1988) on the findings based on reviews of self-supervisor, self-peer, and peer-supervisor ratings studies. The results indicated a high relationship between peer and supervisor ratings ($r = .62$) but only a moderate correlation between self-supervisors ($r = .35$) and nurse-peer ratings ($r = .36$).

Borman (1974) reported less supervisor-peer rating agreement than was found within either type of rater group. Klimoski and London (1974) reported that supervisor ratings showed a strong correlation between effort and performance ratings, whereas peer ratings and self-ratings differentiated between effort and performance.

Borman (1974) suggested that different raters have different perspectives on performance, and Blood (1974) noted that these differences may provide valuable information for the diagnosis of organizational problems.

The effect of actual performance level of the ratee on ratings of that performance was examined by Bigoness (1976), who found that actual performance had the greatest effect on performance ratings. Other studies (Hamner, Kim, Baird, & Bigoness, 1974) found that actual performance accounted for the largest percentage of variance in performance ratings (30%), the sex and race of the ratees and raters accounted for an additional 23 percent of the rating variance.

Shore and Thornton III (1986) concluded, based on the review of the literature, that the gender of supervisors and subordinates may affect levels of agreement between self- and supervisory ratings because (1) men tend to rate their performance more favourably than women rate theirs, and (2) women tend to rate others more favourably than men rate others.

Two factors were found to be important in the study of rater-ratee interaction: (1) frequency of contact and (2) relevancy of the contact. A study by Rothaus, Morton, and Hanson, (1965) showed that increased psychological distance of the rater tended to result in ratings that were more critical and negative. Landy and Guion (1970) reported that raters with daily contact with ratees had a median inter-rater reliability of 0.24 in contrast to a median reliability of .62 for those raters with more relevant contacts with the ratees.

In their review, Landy and Farr (1980: 73) concluded that the literature in the area of performance ratings is fragmented. Some of the research is on different rating formats, whereas others examine characteristics of raters and ratees. They argued that the rating instrument and the characteristics of raters and ratees are only parts of the larger system. Based on their review of literature on performance, they proposed a model of performance rating which was composed of five aspects: (1) the roles (rater and ratee), (2) the vehicle (the rating instrument), (3) the rating context (the type of organization, and the purpose of ratings), (4) the rating process (administrative constraints, individual rater strategies, etc.), (5) and the results of rating (raw and transformed performance information, and actions based on that information, actions based on that information).

Review of the related literature showed that performance ratings would be affected by the following factors: (1) raters, and raters' characteristics (i.e., age, sex), (2) ratees' and ratees' characteristics (i.e., age, sex), (3) rating instrument and (4) organizational characteristics (size, formality, complexity, position).

Research Problem

The measurement of performance has a great importance for administrators as well as for social scientists. Rating is the most common method of performance evaluation. Performance evaluation affects personnel decisions (Landy & Farr, 1980). The effect of the rater provides valuable information for the diagnosis of the hospital problems. The present study aimed to examine the effect of rater {nurses (self); supervisors (administrators) and patients (clients)} and hospital type (government and private hospitals) on nurses' performance ratings.

Research Questions

1. To what extent are the nurses' ratings congruent with supervisory ratings and with patients' ratings of nurses' performance?
2. To what extent are the government hospitals ratings congruent with private hospitals in nurses' performance

ratings?

3. To what extent are the variance ratings of raters congruent with variance ratings in hospitals in nurses' performance?

Methodology

a. Subjects

Nurses

All nurses who volunteered to participate (238 registered nurses employed in the government hospitals and 65 registered nurses employed in the private hospitals) made up the nurse sample. The majority of nurses in the sample were females (86.1%) and most of them between the ages of 18 and 39 years (95.7%). More than half of nurses were Muslims (71.3%), and unmarried (57.1%).

Supervisors

Supervisor sample was consisted of 65 individuals (45 supervisors employed in the top five large government hospitals and 15 supervisors employed in the top five large private hospitals). The ratio of supervisors to nurses was .189 in government and .231 in private hospitals. Most of them were females (75%); between the ages of 18 and 39 years (81.7%); Muslims (60%), and were married (58.3%).

Table 1
Percentage Distribution for Respondents From Registered Nurses,
Supervisors, and Patients

	Supervisors N =60		R. Nurses N =303		Patients N=400	
Variable	Number	%	Number	%	Number	%

Religion

Muslim	36	60	216	71.3	366	91.5
Christian	24	40	87	28.7	34	8.5
Sex						
Females	45	75	261	86.1	196	49
Males	15	25	42	13.9	204	51
Age						
18-29	24	40	216	71.3	202	50.5
30-39	25	41.7	74	24.4	90	22.5
40-49	06	10	11	3.6	65	16.3
50 & over	05	08.3	2	0.7	43	10

Patients

The patients' sample consisted of 400 adult inpatients receiving treatment and care in the ten selected hospitals (40 patients from each hospital). Anonymity of subjects was maintained by the use of a code number which consisted of the number assigned to the hospital. A detailed description of the selection was discussed in the procedures employed in the study.

Half of the patient sample were males (51%); between the ages of 18 and 39 years of age (63.5%); Muslims (91.5%) and were married (69%). In terms of employment status, 20% of the patients were employed by government, 27% by private sector, 10% have their own business, 36.5% were unemployed, and 6.5% were not in the labour market. More than half the patients have been in hospitals prior to their current admission (57.8%). Of those who had been in hospitals before, 36% had been admitted four times to different hospitals; and 37.8% had been hospitalized for the same condition, 42% had been in the same hospital before; 32% had been patients more than three times in the same hospital. Only 37.5% had made outpatient visits to the hospital. Of those who made outpatient visits, 18% made three visits. And, finally, 49.5% of the patients paid their hospital bills, 30.5% were paid by government, 12% by private sector, and 8% by other sources.

b. Instrument Development

An instrument originally developed by Zammuto, London, & Rowland, (1982). was adapted for use in this study in order to measure nurses' performance. The instrument was translated into the Arabic with some modifications by the author. If the items were vague and hard to understand, Arabic synonyms were included and added in parentheses. Using

page 9

self-report-paper-pen, closed-ended answer questionnaire techniques was used. The items consisted of 18 characteristics derived from performance evaluation instruments used in four hospitals in California, and the 19th item was a rating of overall performance. The Performance Rating Index consisted of 19 items. The 19 items were: (1) technical competence, (2) ability to organise and schedule workloads, (3) skills in planning nursing care, (4) acceptability of completed work, (5) attendance and promptness, (6) observance of rest and lunch periods, (7) amount of work performed, (8) completion of work on schedule, (9) adaptability in emergencies, (10) quality of work, (11) dependability, (12) willingness to perform duties, (13) observance of rules and regulations, (14) effort applied, (15) accepting responsibility for own behavior, (16) making a high impression on visitors, (17) personal appearance, (18) skills in communications, and (19) overall performance. The second part consisted of demographic data including age, sex, education, income, and so forth.

Cronbach's alpha was computed for nurses' performance index. The alpha coefficients for nurses are ranged in value between 0.91 to 0.92. Cronbach alpha for the Job Description Index Scale was 0.91.

c. Procedures

All subjects were asked to rate nurses' preference on a single 20-point scale. A Performance Rating Questionnaire (Zammuto et al's instrument) was given to all subjects (nurses, supervisors and Patients) in the ten government and private hospitals included in this study. The instrument consisted of 19 performance items. All subjects asked to indicate on 20-point scale (1 = lowest to 20 = highest) how they would rate nurses' job performance on that day. The researcher collected these forms the next day.

Findings

As can be seen from table (2), there was a significant effect of the interaction between raters' ratings of nurses' performance and the hospital type. Nurses and patients' ratings of nurses' performance in the private hospitals were higher than in the government hospitals, whereas supervisors ratings of nurses' performance was lower in the private hospitals than in the government hospitals. Figure 1 shows the interaction.

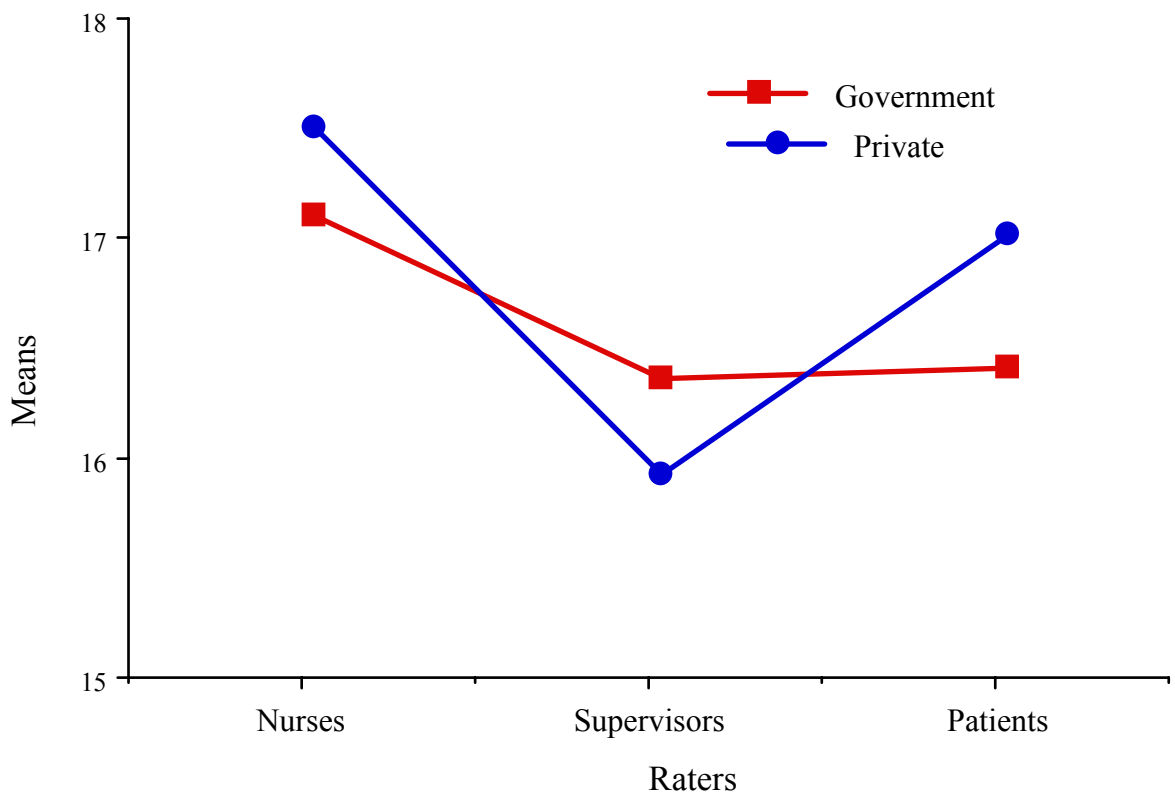


Figure 1
Mean difference between raters and hospitals.

The analysis of Two-Way ANOVA showed that there was a significant effect attributed to the raters (Nurses vs. patients vs. Supervisors) whereas no significant effect was found for the hospital type (government vs. private) on the ratings of nurses' performance (table 2). A positive strong relationship was found between the ratings of nurses and patients ($r=.886$), nurses and supervisors ($r=.928$), and patients and supervisors ($r=.949$).

Table 2
Tow-Way ANOVA Results Testing the effects of Raters

and Hospitals in Performance Ratings.

Source	df	Sum Squares	Mean Squares	F	P
Hospitals	1	6.228	6.228	1.199	.2737
Raters	2	137.007	68.504	13.192	.0001
Interaction	2	35.911	17.956	3.458	.0319
Error	1000	5192.772	5.193		

As shown in Table 3, the mean ratings of nurses on the performance was greater than the mean of supervisors' ratings and the ratings of patients. Nurses' ratings of their performance was higher than either patients and supervisors. Ratings of nurses performance by all raters were higher in the private hospitals than in the government hospitals. There is strong evidence to say that performance rating differed according to different raters. This findings supports the attributional assumption which asserts such differences. Therefore, raters occupy different angles and rate performance from these angles. Each rater has different interest in nurses' performance and focus on these aspects more than the others.

Table (3)

Mean Comparison (Raters by Hospitals)

Raters/ Hospitals	Government mean #	Private mean #	Total mean #
Nurses	17.066 238	17.466 65	17.151 303
Patients	16.396 200	16.981 200	16.689 400
Supervisors	16.323 238	15.891 65	16.231 303
Total	16.606 676	16.862 330	16.69 1006

Further analysis using Scheffe F-test, to examine the possible differences between each pair of raters showed that differences between all pairs were significant. As can be seen from table 2, mean differences between nurses and patients between nurses and, between supervisors and patients and supervisors, were all significant at .05 (table 4).

Table (4)

paired comparison of raters

Scheffe-F-test

Raters	Mean-Differences	p-value
Nurses vs. Supervisors	0.921	0.0001* s
Nurses vs. Patients	0.463	0.0289* s
Patients vs. Supervisors	0.458	0.0311* s

Discussion

The data analysis showed that actors (nurses) rated themselves significantly higher than observers (patients and supervisors). Moreover, patients rated nurses significantly higher than supervisors. This effect is not independent of the hospital type (government and private). Nurses in the government hospitals rated themselves significantly lower than nurses in the private hospitals. Patients in the private hospitals rated nurses significantly greater than patients in the government hospitals, whereas supervisors rated nurses in private hospitals significantly lower than nurses in the government hospitals. This finding is consistent with other research findings for different occupations. Similar findings were reported for clerical employees (Parker, Taylor, Barrett, & Martens, 1959), technical employees (Kirchner, 1966), nurses (Klimoski & London, 1974), first level supervisors (Walden & Thornton, 1979), and executives (Holzbach, 1978). Interpretation of these findings can be attributed to the fact that nurses observe different dimensions of their performance and have different definitions of performance dimensions and, consequently, arrive at different assessments of their performance compared to supervisors and patients. Nurses self-rating was more lenient than either supervisory or client ratings. These findings supported the argument of attribution theory made by Monson and Snyder (1977) and Jones & Nisbet, (1972) and supported by other researchers (Borman, 1974; Landy & Farr, 1980).

Nurses, supervisors, and patients occupy different levels in the hospitals, consequently they disagreed in rating the nurse's performance. This can be explained by differences in the organizational levels occupied by raters. Differences between nurses, supervisors and patients' ratings of nurses' performance was due to the type of relationship they held in regard to the nurses. This finding supports the argument (Borman, 1974; Landy et al., 1976) that raters at different levels disagree on both dimensions and overall performance ratings. Findings of this study support Zammuto, London, & Rowland, (1982) argument that raters at different levels (nurses, supervisors, & patients) rate the overall performance differently, but they would agree on some dimensional ratings. All explanations (i.e., Zammuto, London, & Rowland, (1982) ; Borman, 1974; & Landy et al., 1976) suggest that raters who occupy the same level would provide similar ratings. These interpretations are consistent with Merton's

(1968) and Kane and Lawler's (1979) analyses that organizational structural characteristics influence employee behavior and attitudes.

Observational opportunities offer another explanation for the lack of convergence between nurses, supervisors, and patients' ratings of nurses' performance. It is assumed that patients have more opportunities to observe nurses' performance at more revealing times than do supervisors. This explanation implies that supervisors disagree with nurses and with patients because they have fewer opportunities to observe the nurses' performance. Findings of this study supports this contention.

Future research is needed to examine the individual differences in performance ratings that can be explained by personal and situational attributions. Also, research is needed to study the effects of raters and rater's characteristics on performance ratings.

Applications

Attribution theory can be applied to different areas in the complex organizations. Findings of this study showed that Performance ratings can be used as effective developmental tool (especially self-ratings). Since observers (i.e., supervisors and patients) tend to make personal-oriented performance attribution whereas actors (i.e., nurses) tend to make situational-oriented performance attributions especially for poor performance, then a comprehensive picture of employees' performance can be achieved.

Minimising the discrepancy between raters and ratees can be achieved by improving the channels of communications, increasing the similarity, and familiarity and determining the a common performance criteria. Inadequate observations can cause supervisors' misunderstanding of nurses' performance.

Performance rating can be used as a criteria for promotion. It can be used to improve performance, and as a tool in diagnosing organizational pathologies. The ideal performance rating system would combine information from multiple sources, to form and integrate a tool that can be agreed upon from all relevant raters.

References

- Al-Badayneh, Diab (1993) "Azo masoleiat Hawadeth Al seer: derash tajrebeh ' (in Arabic). Mu'tah Lil-Bohooth Wa Al-Dirasat Vol. 8 No. 3, 1993 9-42.
- Al-Badayneh, Diab (1992) "Fahes faratheat natheret al azo fe al orden" Majalet shoon Ejtemaeh. Vol. 35, 1992, 55-79
- Bigoness, N. J. (1976). Effects of applicant's sex, race and performance on employees' performance ratings: Some additional findings. Journal of Applied Psychology, 61, 80-84.
- Blood, M. R. (1974). Spin-offs from behavioral expectation scale procedures. Journal of Applied Psychology, 59, 513-515.
- Borman, W. C. (1974). The rating of individuals in organizations: An alternative approach. Organizational Behavior and Human Performance, 12, 105-124.
- Cascio, W. F., & Phillips, N. F. (1979). Performance testing: A rose among thorns? Personnel Psychology, 32, 751-766.
- Cascio, W.F., & Valenzi, E. R. (1977). Behaviorally anchored rating scales: Effects of education, and job experience of raters and ratees. Journal of Applied Psychology, 62, 278-282.
- Farr, J. L., O'Leary, B. S., & Bartlett, C. J. (1971). Ethnic group membership as a moderator of the prediction of job performance. Personnel Psychology, 24, 609-636.
- Feldman, M. J. (1981). Beyond attribution theory: Cognitive processes in performance appraisal. Journal of Applied Psychology, 66 (2), 127-148.
- Ferguson, L.W. (1949). The value of acquaintance ratings in criteria research. Personnel Psychology, 2, 93-102.
- Hadad, Yasmeeen "Asaleeb alazoo wa takder althat wal ekteab: ertebatatahom al motabadaleh wa alakatahom bel momarasat al waledeah" (in Arabic) Majelet derasat, 17, n0. 3, pp. 33-66, 1990
- Hamner, W. C., Kim, J. S., Baird, L., & Bigoness, W. J. (1974). Race and sex as determinants of ratings by potential employers in a simulated work sampling task. Journal of Applied Psychology, 59, 705-711.
- Harris, M., M., & Schaubroeck, M. (1988). A meta- analysis of self-supervisor, self-peer, and peer-supervisor ratings. Personnel Psychology, 41, 43-61.
- Herman, J. B., Dunham, R. B., & Hulin, C. L. (1975). Organizational structure, demographic characteristics, and employee responses. Organizational Behavior and Human Performance, 13, 206-232.
- Holzbach, L. R. (1978). Rater bias in performance ratings: Superior, nurse, and peer ratings. Journal of Applied Psychology, 63 (5), 579-588.
- Hunter, J. E., & Hunter, R. F. (1984). Validity and utility of alternative predictors of job performance. Psychological Bulletin, 96, 72-98.

- Jones, E. E., & Davis, K. E. (1965). From acts to dispositions: The attribution processes in person perception. In L. Berkowitz, (ed.), Advances in experimental social psychology, Vol. 2. Orlando, FL: Academic Press.
- Jones, E. E., & Nisbett, R. E. (1972). The actor and the observer: Divergent perceptions of the causes of behavior. In E. E. Jones, D. Kanouse, H. H. Kelly, R. E. Nisbett, S. Valins, & B. Weiner. (Eds.). Attributions: Perceiving the causes of behavior (pp. 79-94). New York: General Press.
- Kane, J. S., & Lawler, E., E., III. (1978). Methods of peer assessment. Psychological Bulletin, 85, 553-586.
- Kane, J. S., & Lawler, E.E., III. (1979). Performance appraisal effectiveness: Its assessment, and determinants. In B. M. Staw (Ed.), Research in organizational behavior, (pp. 425-478). Greenwich, CT: JAI Press.
- Kanouse, D. E., & Hanson, L.R. (1972). Negativity in evaluation. In E. E. Jones, D. Kanouse, H. H. Kelly, R. E. Nisbett, S. Valins, & B. Weiner (Eds.). Attributions: Perceiving the causes of behavior (pp. 47-62). New York: General Press.
- Keeley, M. (1974). Subjective performance evaluation and person-role conflict under conditions of uncertainty. Academy of Management Journal, 20, 301-314.
- Kelley, H. H. (1972). Attribution in social interaction. In E. E. Jones, D. Kanouse, H. H. Kelly, R. E. Nisbett, S. Valins, & B. Weiner (Eds.), Attributions: Perceiving the causes of behavior (pp 1-26). New York: General Press.
- Kelly, E. E. (1967). Attribution theory in social psychology. Nebraska Symposium on Motivation, 14, 192-241.
- Kelly, L. (1985). Dimensions of professional nursing. New York: Macmillan.
- Kirchner, W. K., & Reisberg, D.J. (1962). Differences between better and less effective supervisors in appraisal of subordinates. Personnel Psychology, 15, 295-302.
- Kirchner, W. K. (1966). Relationships between supervisory and subordinate ratings for technical personnel. Journal of Industrial Psychology, 3, 57-60.
- Klimoski, R. J., & London, M. (1974). Role of rater in performance appraisal. Journal of Applied Psychology, 59, 455-451.
- Klores, M. S. (1966). Rater bias in forced-distribution ratings. Personnel Psychology, 19, 411-421.
- Landy, F. J. (1978). An opponent process theory of job satisfaction. Journal of Applied Psychology, 63 (5), 533-547.
- Landy, F. J. (1985). Psychology of work behavior. Homewood, IL. Dorsey Press
- Landy, F. J., Barnes, J., & Murphy, K. (1978). Correlates of perceived fairness, and accuracy in performance appraisal. Journal of Applied Psychology, 63, 751-754.
- Landy, F. J., & Farr, J. L. (1976). Police performance appraisal. JSAS Catalog of Selected Documents in Psychology, 6, 83 (Ms. No. 1315).
- Landy, F. J., & Farr J. L. (1980). Performance rating. Psychological Bulletin, 87 (1), 72-107.
- Landy, F. J., & Farr, J. L. (1983). The measurement of work performance: Methods, theory and applications. New York: Academic Press.
- Landy, F. J., Farr, J. L., Sael F. E., Freytag, W. R. (1976). Behaviorally anchored scales for rating the performance of police officers. Journal of Applied Psychology, 61, 458-557.
- Landy, F. J., & Guion, R. M. (1970). Development of scales for the measurement of work motivation. Organizational Behavior and Human Performance, 5, 93-103.
- Latham, G. P. & Wexley, K. N. (1982). Increasing productivity through performance appraisal. Reading, MA: Addison-Wesley.
- Merton, K. R. (1968). Social theory and social structure. New York: The Free Press.
- Monson, T.C. & Snyder, M. (1977). Actors, observers, and the attribution process: Toward a reconceptualization. Journal of Experimental Social Psychology, 13, 89-111.

- Mullins, C. J., & Force, R. C. (1962). Rater accuracy as a generalized ability. Journal of Applied Psychology, 46, 191-193.
- Parker, J. W., Taylor, E. K., Barrett, R. S., & Martens, L. (1959). Rating scale content: Relationship between supervisory and nurse-ratings. Personnel Psychology, 12, 49-63.
- Perrow, C. (1965). Hospitals: Technology, structure and goals. In March G. James Handbook of organizations (pp. 910-971). Chicago: Rand McNally.
- Porter, L. W., & Lawler, E. E. III. (1965). Properties of organization structure in relation to job attitudes and job behavior. Psychological Bulletin, 64, (1), 23-51.
- Pfiffner, J. M., & Sherwood, F. P. (1960). Administrative organization. Englewood Cliffs, NJ: Prentice-Hall.
- Rothaus, P., Morton, R. B., & Hanson, P. G. (1965). Performance appraisal, and psychological distance. Journal of Applied Psychology, 49, 48-54.
- Rothbart, M. (1968). Effects of motivation equity and compliance on the use of reward, and punishment. Journal of Personality and Social Psychology, 9, 35-362.
- Schroeder, L. L., Sjoquist L. D., & Stephan E. P. (1989). Understanding regression analysis: An introductory guide. Newbury Park, CA: Sage.
- Thornton, G. C., (1968). The relationship between supervisor-and nurse-appraisal of executive performance. Personnel Psychology, 21, 441-445.
- Shore, M. L., & Thornton, G. III., (1986). Effects of gender on self-and supervisory ratings. Academy of Management Journal, 29 (1), 115-129.
- Vance, J. R., Winne, S. P., & Wright, S.E. (1985). A longitudinal examination of rater and ratee effects in performance ratings. Personnel Psychology, 36, 609-619.
- Walden, O. A., & Thornton, G. G. (1979). A comparison of supervisors' self-appraisals and their administrators' appraisals. Medical Group Management, 22, 24, 58.
- Waldman, D. A., & Avolio, B. J. (1986). A meta-analysis of age differences in job performance. Journal of Applied Psychology, 71, 33-38.
- Wexley, K. N. & Klimoski, R. (1984). Performance appraisal: An update. In K. M. Rowland & Q. D. Ferris (Eds), Research in personnel and human resources management (pp. 35-79). Greenwich, CT: JAI Press.
- Wexley, K. N., & Nemeroff, W. F. (1977). The effects of racial prejudice, race of applicant, and biographical similarity on interviewer evaluation of job applicants. Journal of Social and Behavioral Science, 20, 66-76.
- Wexley, K. N., Alexander, R. A., Greenawalt, J. P., & Couch, M. A. (1980). Attitudinal congruence and similarity as related to interpersonal evaluations in manager-subordinate dyads. Academy of Management Journal, 23, 320-330.
- Wexley, K.N., & Pulakos, E.D. (1982). Sex effects on performance ratings on managers-subordinate dyads: A field study. Journal of Applied Psychology, 67, 433-439.
- Zammuto, F. R., London, M., & Rowland, M. K. (1982). Organization and rater differences in performance appraisals. Personnel Psychology, 35, 643-657.
- Zedeck, S., & Baker, H. T. (1972). Nursing performance as measured by behavioral expectation scales: A multitrait-multirater analysis. Organizational Behavior and Human Performance, 7, 457-466.