

PERSON CHARACTERISTIC VERSUS ROLE CONGRUENCY EXPLANATIONS FOR ASSESSMENT CENTER RATINGS

CRAIG J. RUSSELL

Rutgers, The State University of New Jersey

Ratings from assessment centers can consistently predict such subsequent performance criteria as rate of future salary increases, promotions, and supervisory ratings of performance and potential for entry-level management (Thornton & Byham, 1982). In assessment centers, exercises and simulations are used to generate examples of candidates' behaviors, and assessors base dimensional and overall ratings on those behaviors. Unfortunately, little is known about why ratings from assessment centers have predictive value or what it is that they measure (Klimoski & Strickland, 1977; Neidig & Neidig, 1984; Sackett & Dreher, 1984). The literature on assessment centers offers two competing explanations. In the *person characteristic* model, assessment centers are seen as generating candidate behaviors that correspond directly to underlying, causal personal characteristics, skills, and abilities. According to those who design assessment centers, conceptual definitions of assessment dimensions represent a candidate's specific characteristics (Byham, 1970; Holmes, 1977).

In their formulation of the *role congruency* explanation, Sackett and Dreher (1982, 1984) hypothesized that assessors' ratings reflect the degree to which candidates' behavior is congruent with managerial role expectations within a particular exercise. Instead of reflecting a candidate's characteristics, ratings indicate how well a candidate fits expectations of what a hypothetical good manager might do in a given exercise. The question then becomes: Do assessment centers measure (1) a person's characteristics or (2) how well a person behaves like a manager in the eyes of an assessor?

The two explanations have different implications for development of a theory of managerial selection and lead to competing predictions. If personal characteristics constitute the major source of variation in dimensional ratings, ratings of the same dimension across different exercises should be correlated, and ratings made on different dimensions within the same exercise should not be correlated. If the major source of variation is how well a candidate plays the part of a manager in different exercises, ratings within an exercise across dimensions should be more highly correlated than ratings of the same

I would like to thank Douglas Bray, Richard Beatty, Charles Fay, Martin Ford, Dan Morrison, and Kevin Murphy for their comments on an earlier version of this paper. Deborah Good assisted with data analysis.

dimension across exercises. A number of investigators have reported evidence supporting a role congruency explanation (Archambeau, 1979; Neidig, Martin, & Yates, 1978, 1979; Sackett & Dreher, 1982; Sackett & Harris, 1983; Silverman, Dalessio, Woods, & Johnson, 1986; Turnage & Muchinsky, 1982).

Though assessment center architects still adhere to the person characteristic model (Neidig & Neidig, 1984), no researcher has reported results supporting it. However, there have been challenges to results supporting the role congruency model. Neidig and Neidig (1984) argued that convergence of different dimensional ratings within a single exercise can be interpreted as support for an exercise-specific characteristic like skill at handling a simulated in-basket or interview. The only way to test this alternative explanation is to introduce a third measure of either (1) role congruent behavior or (2) person-specific, rather than exercise-specific, skills and abilities. Convergence of a role congruent behavioral measure with exercises requiring that particular type of behavior would support the role congruency explanation. Convergence of a person-specific skill or ability measure with ratings made on individual dimensions across exercises would support the person characteristics explanation.

The current study contains two sources of information: (1) six dimensional ratings made after two exercises and (2) self-ratings of effective interpersonal behaviors for 75 candidates being considered for entry-level management positions. The self-ratings represent a measure of behaviors congruent with the role of manager. The measure is useful in this application because a large literature exists indicating that no single personal characteristic, skill, or ability underlies the self-reported behaviors it assesses (Ford & Tisak, 1983). If a role congruency explanation fits, the measure of effective interpersonal behavior should be most highly correlated with postexercise dimensional ratings that come from an exercise requiring interpersonal interaction, such as an interview simulation. If a person characteristic explanation fits, the effective behavior measure should be more highly correlated with ratings on those dimensions that correspond most closely with interpersonal behavior, such as sensitivity, regardless of the exercise on which the rating was based.

METHODS

Participants and Procedures

Assessment center ratings were obtained from a *Fortune* 500 company in the summer and fall of 1984. Current employees of divisions located all across the continental United States either nominated themselves or were nominated by their immediate supervisors to participate in an assessment center. There were 75 participants, 35 percent of whom were women; their average age was 28 years and their average job tenure was 4 years.

The participants were all candidates for entry-level managerial positions. The procedure required four assessors to make ratings on 9 to 12 dimensions across four or five exercises. The type and number of exercises and dimen-

sions depended on the specific job for which an individual was a candidate. For instance, the assessments for sales manager and manufacturing supervisor differed on one exercise and two dimensions. Assessors individually observed candidates' behaviors, rated them on all 9 to 12 dimensions after each exercise, and arrived at an overall rating for the exercise. Thus, four exercises and 9 dimensions would yield 36 postexercise dimensional ratings per assessor, in addition to one overall rating for each of the four exercises. Discussion to achieve consensus and overall dimensional ratings followed each exercise. Finally, assessors arrived at their own overall assessment ratings for each candidate and then discussed each candidate to arrive at a final, consensual overall assessment rating.

Analyses are limited to postexercise dimensional ratings made on in-basket and interview simulations because they were the only two exercises used consistently across all centers. Six dimensions were rated in common across the two exercises: oral communication skills, sensitivity, planning and organization, management control, information gathering, and decision making. Appendix A presents conceptual definitions of the dimensions as well as a brief description of the in-basket and interview simulation exercises.

Effective Interpersonal Behavior

A number of research efforts directed at measuring a personal-skill-and-ability construct labeled social intelligence occurred 50 years ago, in conjunction with ongoing developments in the area of intelligence testing (Broom, 1928, 1930; McClatchy, 1929; Thorndike, 1920, 1936). These efforts were noteworthy for their failure (Ford & Tisak, 1983). More recently, Ford and his colleagues (M. E. Ford, 1983, 1985, in press; Ford, Burt, & Bergin, 1984; Ford & Tisak, 1983) have developed measures of the interpersonal processes or behaviors that people engage in to achieve various social goals. Because achieving organizational goals through others is a primary managerial role, I viewed effective interpersonal behavior as a role requirement for effective managers. The scale used in the present study was a modified form of the self-rating version used in previous research (Ford, 1985; Ford et al., 1984). Candidates were asked to describe how well they do in 24 different interpersonal situations with co-workers. Responses could range from 1 = not at all good to 6 = extremely good.

A common factor analysis yielded seven factors with eigenvalues greater than 1. As scree test and visual interpretation of the interfactor correlations indicated that a three-factor solution might be more parsimonious. The three-factor solution, which employed a loading rule of .40 on the major factor and .30 at most on all others, generated clean factor loadings in the factor pattern matrix. The average loadings on major and minor factors were .65 and .18. Table 1 contains the factor pattern matrix, and Appendix B describes the items and their development.

The factors are almost identical to those that Ford and his colleagues reported (M. E. Ford, 1985, in press; Ford et al., 1984). They used the same items phrased to reflect a nonwork social setting. Ford (1983, 1985) saw the

TABLE 1
Factor Pattern Matrix
for the Effective Interpersonal Behavior Scale

Items ^a	Factors		
	1	2	3
Morality-related			
1	.83	.29	.21
2	.78	.26	.28
3	.92	.19	.27
4	.61	.20	.28
5	.55	.21	.29
6	.41	.29	.20
Independence-related			
1	.28	.75	.25
2	.27	.55	.29
3	.23	.54	.27
4	.26	.49	.22
5	.28	.60	.25
6	.28	.60	.25
Friendship-related			
1	.22	.26	.66
2	.24	.28	.78
3	.22	.24	.83
4	.23	.26	.85
5	.16	.28	.74
Eigenvalues	7.2	4.3	3.8

^a Appendix B describes the numbered items in each scale.

three factors—designated prosocial, self-efficacy, and instrumental—as reflections of the implicit frameworks people use to categorize social behavior. I labeled the three factors morality-related behavior, independence-related behavior, and friendship-related behavior. The items loading on the three factors had standardized alpha coefficients of .83, .83, and .76, respectively. For the entire effective interpersonal behavior scale, $\alpha = .92$. The average interfactor correlation was .30, and the highest correlation was between the morality and friendship factors ($r = .35$).

Participants took the scale before going to an assessment center. People who were not associated with the local assessment center administered and collected the questionnaire, which was titled the “University Study of Managerial Skills.”

A final measurement issue was the possibility of leniency bias in candidates’ self-ratings on the behavioral scales. The centers were not used as a diagnostic tool to identify shortcomings but as a selection device. Candidates who received unfavorable evaluations in assessment were rarely considered further for promotion. With limited opportunity for reassessment, it was to the candidates’ advantage to make sure they were at their best.

To minimize that potential bias, the instructions indicated that the purpose of the questionnaire was to help future candidates determine whether they were ready for assessment or whether they should postpone it until they had an opportunity to address shortcomings. The instructions also emphasized that their responses to this questionnaire would not be used by any assessor or in any administrative decisions concerning their own employment status. The range of average responses to the scale items was 3.62 to 4.29. The range of standard deviations for item responses was 1.82 to 2.44. It appeared that, although a constant positive bias might characterize responses to items, there was adequate variance to warrant examination.

ANALYSES AND RESULTS

Table 2 contains the means and standard deviations of all variables as well as all intercorrelations. A, B, and C represent a multitrait, multimethod (MTMM) correlation matrix comparable to those reported in previous studies. The average correlation among dimensions evaluated by the in-basket exercise (heterotrait-monomethod in A) was .53. The average correlation among dimensions evaluated by the interview simulation (heterotrait-monomethod in C) was .52. The average monotrait-heteromethod correlation was .25, and the average heterotrait-heteromethod correlation was .19 (B). These findings replicate those reported by Sackett and Dreher (1982) and Sackett and Harris (1983), though the absence of heterotrait-heteromethod convergence could be due to a context effect. Of particular interest are the correlations between the effective interpersonal behavior scales and the postexercise dimensional ratings. Although only 2 of the 28 correlations with in-basket exercise ratings achieve significance, 12 of the 28 correlations with interview simulation ratings achieve significance.

DISCUSSION

The findings presented here extend previous findings in support of the role congruency model. No support emerged for Neidig and Neidig's (1984) alternative explanation of an exercise-specific skill or ability. The convergence of the behavioral scales with dimensional ratings based on an interview simulation, which required candidates to demonstrate interpersonal behaviors, and an in-basket exercise, which did not require interpersonal behaviors, provides support for the role congruency explanation of assessment center ratings. Postexercise dimensional ratings failed to demonstrate convergent and discriminant validity in the MTMM analysis. As reported in prior research, dimensional ratings converge within exercises and discriminate between exercises.

It is also of interest that all but three of the correlations between the behavioral scales and interview simulation ratings are negative. Self-ratings of effective interpersonal behavior with peers are negatively related to assessors' ratings of behavior in the interview simulation. If both sets of ratings are construct-valid measures of interpersonal role requirements, the

findings suggest that assessors may view behavior indicating good peer relations at nonmanagerial levels as detrimental to good superior-subordinate relations. Although the finding is of obvious practical and theoretical interest in the area of managerial development, further research examining relationships between self-, peers', and superiors' perceptions of effective interpersonal behavior is needed.

Assessment center architects have been premature in concluding that postexercise dimensional ratings reflect underlying personal characteristics that cause behaviors observed in assessment centers. Ten years ago Klimoski and Strickland (1977) raised concerns about the construct validity of assessment center ratings, asking whether they predict performance or how well candidates conform to a given set of role expectations held by assessors. This study adds to the considerable body of research findings indicating the latter.

REFERENCES

- Archambeau, D. J. 1979. Relationships among skill ratings assigned in an assessment center. *Journal of Assessment Center Technology*, 2: 7-20.
- Broom, M. E. 1928. A note on the validity of a test of social intelligence. *Journal of Applied Psychology*, 12: 426-428.
- Broom, M. E. 1930. A further study of the validity of a test of social intelligence. *Journal of Educational Research*, 22: 403-405.
- Byham, W. C. 1970. Assessment centers for spotting future managers. *Harvard Business Review*, 48 (4): 150-160.
- Ford, D. H. In press. *Humans as self-constructing living systems*. Hillsdale, N. J.: Lawrence Erlbaum Associates.
- Ford, M. E. 1983. *The nature of social intelligence: Processes and outcomes*. Paper presented at the 91st Annual Meeting of the American Psychological Association, Anaheim, California.
- Ford, M. E. 1985. The concept of competence: Themes and variations. In H. A. Marlowe, Jr. & R. B. Weinberg (Eds.), *Competence development*: 210-250. Springfield, Ill.: Charles C. Thomas.
- Ford, M. E. In press. A living systems conceptualization of social intelligence: Outcomes, processes, and developmental change. In R. Steinberg (Ed.), *Advances in psychology of human intelligence*. Hillsdale, N. J.: Lawrence Erlbaum Associates.
- Ford, M. E., Burt, R., & Bergin, C. 1984. *The role of goal setting and goal importance in adolescence social competence*. Paper presented at the American Education Research Association Meetings, New Orleans.
- Ford, M. E., & Miura, I. 1983. *Childrens' and adults' conceptions of social competence*. Paper presented at the 91st Annual Meeting of the American Psychological Association, Anaheim, California.
- Ford, M. E., & Tisak, M. S. 1983. A further search for social intelligence. *Journal of Educational Research*, 75: 196-206.
- Holmes, D. S. 1977. How and why assessment works. In J. L. Moses & W. C. Byham (Eds.), *Applying the assessment center method*: 127-142. New York: Pergamon Press.
- Klimoski, R. J., & Strickland, W. J. 1977. Assessment centers: Valid or merely prescient? *Personnel Psychology*, 30: 353-363.

- McClatchy, V. R. 1929. A theoretical and statistical critique of the concept of social intelligence and of attempts to measure such a process. *Journal of Abnormal and Social Psychology*, 24: 217–220.
- Neidig, R. D., Martin, J. C., & Yates, R. E. 1978. *The FBI's management aptitude program assessment center: Research report No. 1 (TM-78-3)*. Washington, D.C.: U.S. Civil Service Commission (NTIS No. PB-286 181).
- Neidig, R. D., Martin, J. C., & Yates, R. E. 1979. The contribution of exercise skill ratings to final assessment center evaluations. *Journal of Assessment Center Technology*, 2: 21–23.
- Neidig, R. D., & Neidig, P. J. 1984. Multiple assessment center exercises and job relatedness. *Journal of Applied Psychology*, 69: 182–186.
- Sackett, P. R., & Dreher, G. F. 1982. Constructs and assessment center dimensions: Some troubling findings. *Journal of Applied Psychology*, 67: 401–410.
- Sackett, P. R., & Dreher, G. F. 1984. Situation specificity of behavior and assessment center validation strategies: A rejoinder to Neidig and Neidig. *Journal of Applied Psychology*, 69: 187–190.
- Sackett, P. R., & Harris, M. 1983. *A further examination of the constructs underlying assessment center ratings*. Paper presented at the 91st Annual Meeting of the American Psychological Association, Anaheim, California.
- Silverman, W. H., Dalessio, A., Woods, S. B., & Johnson, R. L., Jr. 1986. Influence of assessment center methods on assessors' ratings. *Personnel Psychology*, 39: 565–578.
- Thorndike, E. L. 1920. Intelligence and its uses. *Harper's Magazine*, 140: 227–235.
- Thorndike, E. L. 1936. Factor analysis of social and abstract intelligence. *Journal of Educational Psychology*, 27: 231–233.
- Thornton, G. C. III, & Byham, W. C. 1982. *Assessment centers and managerial performance*. New York: Academic Press.
- Turnage, J. J., & Muchinsky, P. M. 1982. Transsituation variability in human performance with assessment centers. *Organizational Behavior and Human Performance*, 30: 174–200.

APPENDIX A

Definitions of Assessment Dimensions

The dimensions rated by assessors examined in this study included the following:

1. Oral communication skills – ability to effectively present information orally.
2. Sensitivity – ability to perceive subtle cues in behaviors of others.
3. Planning and organization – ability to allocate or assign, sequence, and coordinate resources to achieve a goal or objective.
4. Control – ability to develop and maintain sources, procedures, and mechanisms to become and remain informed of important matters.
5. Information gathering – ability to interact with others in ways necessary to gather important information.
6. Decision making – ability to make high quality decisions.

The exercises in which participants engaged were:

1. In-basket – Candidates receive a package of background materials and problems, specific to the type of job for which they are being considered, with predetermined relationships, decision requirements, and priorities. Candidates have a specific time period within which to assume the role of a responsible administrator.
2. Interview simulation – Candidates play a role structured around an interview situation that might occur on the target job. For a position in sales management, it involved a face-to-face interaction with an unhappy customer; for the position of manufacturing supervisor, it involved discussing a performance problem with a subordinate.

APPENDIX B

Effective Interpersonal Behavior Scale

The work of Ford and his colleagues was the basis for the development of the scales used in this study. Ford and Miura (1983) took an implicit approach, asking people to describe the most socially competent person they knew. Cluster analysis of subsequent codings of those descriptors yielded four major components underlying peoples' conceptions of social competency: (1) prosocial behavior – sensitive to the feelings of others, dependable, socially responsible, respectful of other people's rights and viewpoints, interested in other people; (2) instrumental behavior – knows how to get things done, likes to set goals, communicates well; (3) social ease – opens up to people, adapts socially, is involved and enjoys social activities; and (4) self-efficacy – has own values and identity, a good outlook and self-concept.

D. H. Ford (in press) also took an explicit approach based on the social motivation literature and "living systems theory." He identified eight categories of self-assertive and integrative behavioral competencies in social settings. The integrative competencies dealt with behaviors maintaining and promoting the well-being of other individuals and groups. However, using peer nominations, teacher ratings, and self-evaluations with high school students, M. E. Ford (in press, 1985) and Ford and his colleagues (1984) found convergent evidence of the three implicit categories of prosocial, instrumental, and self-efficacy behavior. The findings reported in Table 1 support these implicit categories. The items for the three subscales are listed below:

Morality-related behavior:

1. Some people are good at following through when they say they'll do something for a co-worker, but other people are not so good at this.
2. Some people are good at treating co-workers equally, even those with very different beliefs or characteristics, but others are not so good.
3. Some people are good at responding to the needs of their co-workers, but other people are not so good.
4. Some people are good at keeping their commitments to their co-workers, but others are not so good at this.
5. Some people are good at giving everyone an equal opportunity to state their opinion when making a group decision, but other people are not so good at this.
6. Some people are good at sticking to agreements they have made with their co-workers, but other people are not so good at this.

Independence-related behavior:

1. Some people are good at resisting group pressure when making choices, but other people are not so good at this.
2. Some people are good at developing their own way of doing things, but other people are not so good at this.
3. Some people are good at doing what they want even when their co-workers want them to do something else.
4. Some people are good at coming out ahead in most comparisons with their peers, but other people are not so good at this.
5. Some people are good at maintaining their own identity within a group, but other people are not so good at this.
6. Some people are good at forming their own individual viewpoints, but other people are not so good at this.

Friendship-related behavior:

1. Some people are good at helping their co-workers get through hard times, but other people are not so good at this.
2. Some people are good at finding a co-worker who can make them feel better when they are lonely or depressed, but other people are not so good at this.
3. Some people are good at finding a co-worker who will stand by them when they have a problem, but other people are not so good at this.

4. Some people are good at helping their co-workers feel good about themselves, but other people are not so good at this.
5. Some people are good at getting their co-workers to help them with their work, but other people are not so good at this.

Craig J. Russell earned his Ph.D. degree in business administration from the University of Iowa; he currently is an assistant professor at the Institute of Management and Labor Relations, Rutgers, The State University of New Jersey. His research interests include management selection, compensation, turnover, and job choice.

⁵ *Academy of Management Journal*
1987, Vol. 30, No. 4, 826-835.

USING MULTIPLE ATTRIBUTES TO ASSESS SOURCES OF PERFORMANCE FEEDBACK

DAVID M. HEROLD

ROBERT C. LIDEN

Georgia Institute of Technology

MARYA L. LEATHERWOOD

University of Illinois

Feedback on employees' performance in work organizations has long been recognized as important for the learning and maintaining of work-related behaviors (Ashford & Cummings, 1983; Ilgen, Fisher, & Taylor, 1979; Taylor, Fisher, & Ilgen, 1984). However, as Ilgen and colleagues noted, "Generalizations about the effects of feedback on individuals are few" (1979: 349). In their review, those investigators concluded that one major reason for the scarcity of generalizable conclusions is that researchers have paid little attention to differences in sources of feedback.

Previous attempts to address such differences have focused on identifying salient sources of feedback in work environments and on the amount and nature of the information that each source provided. Greller and Herold (1975) and Hanser and Muchinsky (1978) found differences in the perceived informativeness of five sources of performance feedback; the sources were formal organizations, co-workers, supervisors, tasks themselves, and focal individuals' own feelings and ideas. Using factor analysis, Hanser and Muchinsky (1978) and Herold and Greller (1977) demonstrated the existence of separate factors representing different sources of feedback.

Although the amount of information provided by feedback sources is important, other attributes of sources are also relevant (Guetzkow, 1965; Porter & Roberts, 1976). Examining such other attributes, Hanser and Muchinsky (1978) reported differences in the perceived reliability of information from different sources, and Greller (1980) reported differences in the perceived usefulness of information from different sources.