

Assessment Centers: Research and Applications

A Special Issue of the
Journal of Social Behavior and Personality, 1997, Vol. 12, No. 5

Editors' Preface

Ronald E. Riggio and Bronston T. Mayes

ASSESSMENT CENTER RESEARCH

Section I

Assessment Centers: Past, Present, and Future

Insights into the History and Future of Assessment Centers: An Interview with Dr. Douglas W. Bray and Dr. William Byham
Bronston T. Mayes

A Reassessment of Assessment Centers: Challenges for the 21st Century
Ann Howard

The Assessment Center Process: New Directions
Phillip E. Lowry

Section II

Validity and Reliability Issues in Assessment Centers

Construct Validity of Assessment Centers: Appropriate Use of Confirmatory Factor Analysis and Suitable Construction Principles
Martin Kleinmann and Olaf Köller

Assessment Center Construct Validity and Behavioral Checklists: Some Additional Findings

Lisa M. Donahue, Donald M. Truxillo, John M. Cornwell, and Melanie J. Gerrity

Construct Validity of Assessment Center Judgments: Analyses of the Behavioral Reporting Method

George C. Thornton III, Aharon Tziner, Michal Dahan, James P. Clevenger, and Elchanan Meir

New Evidence on the Construct Validity of Diagnostic Assessment Centers: The Findings May Not Be So Troubling After All
Jeffrey D. Kudisch, Robert T. Ladd, and Gregory H. Dobbins

Congruence Between Self-Other Ratings and Assessment Center Performance
Kenneth M. Nowack

Section III

The Role of Individual Differences in Assessment Centers

A Person Perception Explanation for Validation Evidence from Assessment Centers
Robert G. Jones

Assessing the Assessor: The Relationship of Assessor Personality to Leniency in Assessment Center Ratings

Lynn K. Bartels and Dennis Doverspike

The Role of Gender in a Developmental Assessment Center

Ted H. Shore, Armen Tashchian, and Janet S. Adams

A Study of State, Trait, and Test Anxiety, and Their Relationship to Assessment Center Performance

Clive Fletcher, Catriona Lovatt, and Caroline Baldry

ASSESSMENT CENTER APPLICATIONS

Section IV

Operational Applications of Assessment Centers

Development and Validation of an Assessment Center During Organizational Change

Warren Bobrow and Janet S. Leonards

Undercover Agent Assessment Centers: Crafting Vice and Virtue for Impostors
Michel Girodo

Assessment Centers: Cross-Cultural and Cross-National Issues
Dennis R. Briscoe

Section V

Educational Applications of Assessment Centers

The Use of Assessment Center Methods for Student Outcome Assessment

Ronald E. Riggio, Monica Aguirre, Bronston T. Mayes, Chris Belloli, and Carolyn Kubiak

Using an Assessment Center as a Developmental Tool for Graduate Students: A Demonstration

Janet L. Kottke and Kenneth S. Shultz

Assessment Centers for Course Evaluations: A Demonstration

Bronston T. Mayes, Christopher A. Belloli, Ronald E. Riggio, and Monica Aguirre

Name Index

Subject Index

Congruence Between Self-Other Ratings and Assessment Center Performance

Kenneth M. Nowack

Organizational Performance Dimensions
137 Strand Street, Santa Monica, CA 90405

Multi-rater feedback is based upon the tenet that congruence between self and others is associated with managerial success and effectiveness (Tornow, 1993; Yammarino & Atwater, 1993). This study investigated the relationship between self and managerial rating congruence on two measures of assessment center performance (overall assessor ratings and in-basket scores) with 144 production supervisors. Results of hierarchical regression analyses indicated that manager ratings of supervisory effectiveness ($\Delta R^2 = .26, p < .01$) and self-ratings on 14 assessment center dimensions ($\Delta R^2 = .06, p < .01$) significantly contributed towards predictions of overall assessor ratings of performance (OAR). In-agreement/good raters and overestimators were rated significantly higher in assessment center exercises by assessors compared to underestimators (Tukey's HSD test, $p < .01$). Results from additional hierarchical regression analyses indicated that self-ratings incrementally contributed to predictions of in-basket performance ($\Delta R^2 = .04, p < .05$) above that of manager ratings on task management skills ($\Delta R^2 = .33, p < .01$). In-agreement/poor raters and underestimators had lower overall in-basket scores than in-agreement/good raters or overestimators. Implications of these findings in terms of previous and future research are discussed.

Extensive research on assessment centers has shown that overall assessor ratings are valid predictors of job performance and career success (Gaugler, Rosenthal, Thornton, & Bentson, 1987; Russell, 1987; Schneider & Schmitt, 1992; Shore, Shore, & Thornton, 1992; Shore, Thornton, & Shore, 1990; Thornton & Cleveland, 1990; Thornton & Byham, 1982). However, researchers have continued to raise the question of how and why assessment centers have shown predictive validity despite a growing literature supporting their overall effectiveness (e.g., Klimoski & Brickner, 1987; Sackett & Dreher, 1982). Assessor ratings are typically based on many types of data, including paper-and-pencil instruments, interactive group exercises, role plays, games/simulations,

Riggio, R.E., & Mayes, B.T. (Eds.). (1997). Assessment centers: Research and applications [Special issue]. *Journal of Social Behavior and Personality*, Vol. 12, No. 5, 145-166.
©1997 Select Press, Corte Madera, CA, 415/924-1612.

in-basket exercises, structured interviews, peer and supervisory ratings and self-evaluations (360-degree feedback).

Although many of these types of data have been evaluated for their relative contribution to assessment center validity (e.g., Brannick, Michaels, & Baker, 1989; Reilly & Chao, 1982; Schippmann, Prien, & Katz, 1990), in-basket exercises, self-ratings, and 360-degree feedback (peer and supervisory ratings and self-evaluations) have received less attention in the research literature. In light of the growing popularity and use of multi-rater 360-degree feedback instruments (cf., Antonioni, 1996; London & Beatty, 1993; London, Wohlers, & Gallagher, 1990; Hoffman, 1995; O'Reilly, 1994; Tornow, 1993), it is clear that this is an area in which practice is well ahead of theory and empirical research.

Despite the paucity of research on self-evaluations within the assessment center, there has been extensive research on self-ratings in other settings (Harris & Schaubroeck, 1988; Mabe & West, 1982; Nowack, 1992; Thornton, 1980). For example, Harris and Schaubroeck's (1988) meta-analysis found that peer and supervisor ratings were relatively highly correlated, but that only moderate correlations existed between self-supervisor and self-peer ratings. The results of their meta-analysis indicated a relatively high correlation between peer-supervisor ratings ($r = .65$), but only a modest correlation between self-peer ($r = .36$) and self-supervisor ratings ($r = .35$). Furthermore, while rating format (global versus dimensional) and rating scale (behavioral vs. trait) did not moderate the results, job type (managerial vs. blue collar) did appear to moderate self-peer and self-supervisor ratings. Specifically, self-supervisor and self-peer correlations were much lower for managerial employees than for blue-collar types. However, these effects were not seen for peer-supervisor correlations.

In their review of over 50 studies, Shrauger and Schoeneman (1979) examined evidence relating self-perceptions to evaluations by significant others. In general, they found little evidence of congruence between self-perceptions and evaluations by others, nor did they find consistent evidence that self-evaluations are strongly influenced by other's feedback. Nowack (1992) investigated differences between the self-ratings of 335 managers to others (supervisors, subordinates, peers) using an instrument assessing 20 specific management skills areas derived from job analysis procedures. The results indicated that the amount of agreement between managers and the others describing them was moderately low with correlation coefficients ranging from .12 to .30 ($p < .05$). Paired *t* tests revealed that managers consistently reported practicing specific skills more frequently than others who rated them across the majority of skill

areas. These results are consistent with the findings of Thornton (1980) who found that self-assessments tend to be lenient.

In general, the predictive validity for self-evaluations in assessment centers is inconclusive and mixed (Bray & Grant, 1966; Shore et al., 1992; Tziner, 1984;). Bray and Grant (1966) reported that overall assessor ratings and peer evaluations correlated higher than self-assessor ratings or self-peer ratings. These findings are consistent with the meta-analytic findings of Harris and Schaubroeck (1988) but specific to assessment center performance.

In their assessment center study of 394 employees in a large petroleum company, Shore et al. (1992) found evidence that the construct validity was stronger for peer- than for self-evaluations, and for more easily observable dimensions than for dimensions requiring greater judgment. Self-evaluations were found to be significantly associated with overall assessor ratings for management potential, but only peer ratings significantly contributed to predictions of job advancement.

Reluctance to use self-ratings in assessment centers seems largely due to concerns that many individuals possess an overly positive, albeit unrealistic, view of themselves and tend to present themselves in socially desirable or self-deceptive ways (e.g., Crowne & Marlowe, 1964; DeNisi & Shaw, 1977; Edwards, 1957; Kagan, 1988; Schwartz, 1990; Taylor & Brown, 1988). Furthermore, research on self- and other ratings has generally found that the distribution of self-ratings is negatively skewed (Mabe & West, 1982).

The relationship between rating congruence and managerial effectiveness is a cornerstone of multi-rater 360-degree feedback. Most developmental assessment centers and leadership programs using 360-degree feedback attempt to identify specific areas where incongruence between self-ratings and other ratings exist and attempt to reduce these differences through developmental interventions, such as training. Several popular methods have emerged to determine congruence (or lack thereof) between self-ratings and the ratings of others.

Atwater and Yammarino (1992, in press) introduced the idea of agreement categories to classify individuals whose self-evaluations and self-ratings agreed in varying degrees with ratings of relevant others. In particular, Atwater and Yammarino (1992, in press) distinguish between overestimators (those who rate themselves higher than others do), underestimators (those who rate themselves lower than others do), in-agreement/good raters (those who rate themselves favorably and similar to other's ratings) and in-agreement/poor raters (those who rate themselves unfavorably and similar to other's ratings).

Their model proposes that over-raters and in-agreement/poor raters have negative individual and organizational outcomes, whereas in-agreement/good raters would have positive outcomes. Underestimators are expected to have mixed outcomes in that as modest self-evaluators, they are open to self-improvement but might have low aspirations and self-efficacy (cf. Ashford, 1989; Bandura, 1992).

With respect to congruence between self- and other ratings, work done by Bass and Yammarino (1991) and Smircich and Chesser (1981) would suggest that greater agreement (either poor or good) is preferable to disagreement because it indicates some level of self-insight and understanding. Unfortunately, inflated self-evaluations are been associated with career failure and derailment (McCall & Lombardo, 1983; Van Velsor, Taylor, & Leslie, 1993), lower promotability (Bass & Yammarino, 1991; McCauley & Lombardo, 1990) and low job performance (Yammarino & Atwater, 1993). In general, research on leadership also supports the positive outcomes of accurate self-evaluation. Charismatic leadership was found to be positively associated with accurate self-insight (Atwater & Yammarino, in press) and leaders with more accurate self-evaluations have been found to be more effective and successful than those with poor self-evaluations (cf. Atwater & Yammarino, 1992; Bass & Yammarino, 1991).

Individuals with inflated or deflated ratings (underestimators and overestimators) will typically misdiagnose their strengths and development areas. These inaccurate self-evaluations can adversely affect job decisions and actions such as not pursuing developmental opportunities or seeking feedback from others (Atwater & Yammarino, in press). As such, overestimators perceive that they possess strengths that other do not see or deny that they have weaknesses that are perceived by others (Bass & Yammarino, 1991). As a result, these individuals are less motivated to seek self-improvement or seek additional input and feedback from others because they believe their performance is adequate (Fedor, Rensvold, & Adams, 1992). However, two recent studies suggest that in response to feedback from others, overestimators showed the greatest improvements relative to underestimators or in-agreement/poor raters (Atwater, Roush, & Fischthal, 1995; Smither, London, Vasilopoulos, Reilly, Millsap, & Salvemini, 1995).

The present study attempted to explore congruence between self- and other ratings and measures of assessment center performance. Specifically, the present study examined the interrelationships between self-ratings and manager ratings (measured through use of a 360-degree management practices skills inventory providing for an overall score and separate factor scores for communication, interpersonal and task manage-

ment skills) with overall assessment center performance using both objective (in-basket simulation performance) and subjective measures (overall assessor ratings of assessment center performance on behavioral and interactive exercises). Based upon existing literature on the relationship between self-other ratings and individual/organizational performance, three specific hypotheses are proposed:

- H1. Self-ratings and manager ratings will significantly contribute to predictions of objective (in-basket scores) and subjective (OAR) assessment center performance.
- H2. In-agreement/good raters will be better performers than overestimators, underestimators and in-agreement/poor raters on objective and subjective measures of assessment center performance.
- H3. Overestimators will be better performers than underestimators on objective and subjective measures of assessment center performance.

METHOD

Participants

Participants included 144 supervisors working in a production environment at three separate locations for a large communications company in the Los Angeles area. The program participants were primarily males (95.8%) with a mean age of 45.36 (SD = 8.31) and average tenure of approximately 15.5 years with the company. Participants were selected to attend one of twelve scheduled one-day assessment centers as part of a comprehensive training and development program targeted to enhance overall supervisory knowledge and skills. Participants were given oral and written feedback approximately 2–3 weeks after the assessment center program was completed. This feedback meeting resulted in the initiation of a written professional development plan for each assessment center participant. Participants were also scheduled to attend a three-day supervisory skills training program later in the same year.

Assessment Center Procedures

The developmental assessment center targeted 14 dimensions based upon a job analysis of the supervisory position (Oral Communication, Presentation, Written Communication, Listening, Initiative, Delegation, Administrative Control, Leadership/Influence, Conflict Management, Interpersonal Sensitivity, Decisiveness, Judgment, Problem Analysis, and Coaching). Written and behavioral group exercises were designed to assess these 14 dimensions. Behavioral group exercises included two

leaderless group discussions (a planning consensus exercise and manufacturing problem-solving exercise based on actual employee satisfaction opinion survey data from the organization) and a 30-minute performance management coaching role-play with one of the assessors. The written exercises consisted of 90-minute manufacturing in-basket simulation, career interest measure (Career Profile Inventory), personality inventory (FIRO-B), and self-ratings on each of the 14 assessment center dimensions using a 1 to 5 rating scale of competence.

During each one-day assessment center session, twelve supervisors participated in these written and behavioral group exercises and completed several paper-and-pencil assessments. The participants were observed by three external assessors who had received a full-day of training in the assessment center method, observational ratings, and specific group exercises used in the center. Assessors observed participant behaviors during the exercises and prepared dimensional ratings and observational comments after each group exercise was completed. At the conclusion of each group exercise, the participants ranked one another on overall effectiveness in each group exercise. Participants were not specifically told how the peer- and self-evaluations would be used, but there was an understanding that all information generated in the assessment center might be used for developmental planning efforts.

Each participant was provided a one-hour feedback session conducted by the administrator and a written developmental action plan approximately two-weeks following the assessment center. During the feedback meeting, results from the individual, group, and paper-and-pencil exercises were summarized. Each participant was a member of a group of six that rotated through the group exercises. Two assessors rotated their observations and evaluations of the six group members during the two large leaderless group exercises. The first large group involved a consensus seeking task involving the steps involved in the project planning process. Participants were first asked to individually rank the project planning steps and then reach consensus as a large group about these rankings. In the second large group exercise, participants were given actual data extracted from a recent company-wide employee satisfaction survey. This data compared employee responses to specific questions between the production department and the rest of the company. The group was given the task to interpret these findings and suggest specific interventions or recommendations for improving the effectiveness of the production department. The last exercise involved a performance coaching role-play. The participant was asked to play the role of a production supervisor who had to conduct a discipline coaching meeting with a poor performing employee, played by one of the assessors. The

participant was given a one-page summary describing recent performance issues and problems to be addressed during the coaching meeting. At the conclusion of the assessment center, assessors were joined by the administrator to evaluate participants.

MEASURES

Overall assessor ratings. Overall assessor ratings (OAR) for each dimension were obtained through consensus ratings of all three assessors based on performance across the three interactive behavioral exercises used in the assessment center for each participant (two leaderless group exercises and a coaching role-play). No paper-and-pencil results were made available to the assessors before they reached consensus and assigned the OAR. It is important to emphasize that the assessors did not utilize in-basket scores, career interest, or personality inventory results as they reached their overall assessment ratings (OAR) in this particular developmental assessment center. Therefore, the OAR was solely based upon performance during the two leaderless group and coaching exercises which emphasized largely interpersonal, rather than, performance style behaviors (presentation, oral communication, listening, initiative, leadership/influence, conflict management, interpersonal sensitivity, coaching, and problem analysis). In this study, the OAR reflected the sum of the nine dimensional ratings based upon a 1 to 5 scale (1 = *very low level observed* to 5 = *very high level observed*) observed in the three behavioral exercises derived by consensus by the three assessors ($M = 22.9$; $SD = 6.48$; $\alpha = .82$).

Manager ratings. In this study, the 360-degree inventory, the Management Practices Questionnaire (MPQ), was used to assess supervisory and management skills (Nowack, 1992) from both the participant and his/her manager. The MPQ was designed to provide participants with specific feedback on twenty specific job skills and behaviors. The instrument is completed by the assessment center participant and his/her current manager. Each question is behaviorally focused and respondents rate the frequency or practice of each behavioral skill using a 7-point scale where 1 = *to an extremely small extent* and 7 = *to an extremely large extent*.

The MPQ is composed of 100-items measuring 20 specific management practices areas based upon job analysis and organized into four general areas: (a) *Communication Skills*—Listening, Two-Way Feedback, Oral Communication, Written Communication, Oral Presentation; (b) *Task Management Skills*—Vision & Goal Setting, Performance Evaluation, Planning/Organizing, Delegation, Follow-Up/Administrative Control, Performance Management; (c) *Interpersonal Skills*—Team Building/Team Development, Recognizing & Rewarding Performance, Inter-

personal Sensitivity, Conflict Management, Coaching/Developing Others, Leadership/Influence, Employee Involvement; and (d) *Decision Making Skills*—Strategic Problem-Analysis and Decisiveness/Judgment.

The internal consistency reliability of the MPQ scales are fairly high, ranging from .71 to .91 as measured by Cronbach's alpha. The MPQ has shown acceptable test retest reliability ranging from .45 to .82 over a three-month period in previous research (Nowack, 1989). The average test-retest reliability across all 20 scales was .65. Item-total analyses indicate that items within each scale are highly intercorrelated with each other, however, items across certain scales also correlate modestly with each other (data not shown).

A principal components factor analysis across the entire sample, with varimax rotation, yielded a three factor solution explaining a total of 74.2% of the common variance. These three factors included: (a) Interpersonal skills, (b) Task Management skills, and (c) Communication skills which correspond to the general managerial skill areas identified in the original job analysis. However, the first factor (Interpersonal skills) was clearly the largest factor accounting for 61.4% of the common variance. In this study, four separate manager ratings were calculated based on an overall and factor related scores: (a) Overall manager rating based on a cumulative scale score across the 100-items composing the MPQ, (b) Communication skills rating score, (c) Interpersonal skills rating score, and (d) Task management skills rating score.

Self-ratings. Self-ratings on the 14 assessment center dimensions were obtained for each participant. Each dimension was behaviorally defined and participants were asked to rate self-perceived competence on a 1 (*low level of competence*) to 5 (*high level of competence*) scale ($M = 51.50$; $SD = 6.79$; $\alpha = .72$). The self-rating also included a section for a brief narrative summary of strengths, development areas, and career objectives. This information was used in the feedback meeting after the assessment center to assist with developmental planning efforts.

In-basket management simulation. An in-basket exercise was used to simulate the task management, administrative, and supervisory problems of a typical manager in a production and manufacturing setting. With the aid of background information on a fictitious organization, the participant assumes the role of a new department manager and is asked to respond to 23 specific letters, memos, reports, requests, and problems that have accumulated on a predecessor's desk in a 90-minute period. The participant must make decisions, take actions, delegate responsibility, write letters, initiate meetings, assign work, plan, organize, and schedule activities based on the material in the in-basket exercise.

The production in-basket simulation measures eight separate scale scores (initiative, interpersonal sensitivity, planning/organizing, delegation, administrative follow-up, problem analysis, decisiveness, judgment) and an overall score. This simulation is scored using an objective scoring key providing points for desired behaviors expressed, decisions made, or actions taken. The in-basket simulation has demonstrated high interrater reliability (.93) and criterion-related validity with supervisory ratings (average $r = .27$) in a previous validation study (Nowack, 1989). Each in-basket scale has different scale ranges, means, and standard deviations based on the objective scoring key. To standardize the eight in-basket scale scores, separate z scores were first calculated. In this study, overall in-basket performance was measured by summing the z scores across each of the eight scales to provide for a composite score.

ANALYSES AND RESULTS

Pearson correlation coefficients were calculated to explore the relationships between self-ratings, supervisor ratings, and assessment center performance (Table 1). Participant self-ratings were significantly associated with in-basket performance ($r = .27$, $p < .01$) and overall assessor ratings ($r = .26$, $p < .01$), but not overall manager ratings ($r = .10$, $p > .05$). Overall manager ratings were significantly associated with OAR ($r = .25$, $p < .02$) and overall in-basket performance ($r = .28$, $p < .01$). These results provide limited evidence of criterion-related validity of the in-basket simulation with manager and OAR ratings despite recent questions and criticisms about the validity of in-basket scores (cf. Brannick et al., 1989; Schippmann et al., 1990). Greater congruence was observed between assessor-supervisor ratings than between self-supervisor ratings. These findings are consistent with previous research (cf. Harris & Schaubroeck, 1988; Mabe & West, 1982) comparing self and other ratings.

To test the hypothesis that self-ratings and manager ratings would significantly contribute towards predicting objective (in-basket simulation) and subjective (overall assessor ratings) measures of assessment center performance, a series of hierarchical regression analyses were performed. The results of these analyses are summarized in Tables 2 and 3. In the first hierarchical regression analysis (Table 2), the independent variables included self-ratings and overall manager ratings entered as a block in the first step followed by the interaction term (Self-ratings \times Overall manager ratings).

Additional hierarchical regression analyses were performed in the same manner but using the three factor derived manager rating scales as separate independent variables entered with self-ratings in the first step

TABLE 1 Means, Standard Deviations and Intercorrelations of Assessment Center Scales (N = 144)

Scale	M	SD	Self-Ratings					In-Basket Score
			Overall Assessor Rating (OAR)	Overall Manager Rating	Manager Rating Interpersonal	Manager Rating Performance	Manager Rating Communication	
Self-Rating	51.50	6.79	.26**	.10	.03	.16	.09	.27**
Overall Assessor Rating (OAR)	22.99	6.48		.25**	.25**	.17	.24**	.38**
Overall Manager Rating ^a	95.45	14.57			.92**	.89**	.79**	.28**
Manager Interpersonal Rating ^a	47.38	7.72				.69**	.60**	.19*
Manager Performance Rating ^a	33.74	5.87					.67**	.26**
Manager Communication Rating ^a	14.40	2.83						.38**
In-Basket Score ^b	0.03	5.46						—

^aOverall manager ratings were derived by a cumulative score across all 20 scales of the Management Practices Questionnaire (MPQ). Separate manager ratings of interpersonal, performance, and communication skills were calculated from subscales of the MPQ based on previous factor analyses.

^bAn overall In-Basket score was derived by summing z-scores of the eight scales (initiative, interpersonal sensitivity, delegation, follow-up, problem analysis, judgment, decision making)

* $p < .05$; ** $p < .01$.

TABLE 2 Summary of Hierarchical Regression Analysis for Variables Predicting Assessment Center Performance with Overall Manager Ratings as the Independent Variable (N = 144)

Variable	MultiR	R ²	F(Eqn)	ΔR^2	Fch	SE	β
In-Basket Performance							
Overall Manager Ratings	.51	.26	39.94**	.26	39.94**	.61	.49**
Self-Ratings	.57	.33	26.8**	.06	10.34**	.06	.25**
Overall Manager \times Self-Ratings	.57	.00	.55	.00	.00	.00	.00
Total	.57	.33					
Overall Assessor Ratings (OAR)							
Self-Ratings	.26	.07	8.49**	.07	8.49**	.08	.24**
Overall Manager Ratings	.35	.12	7.46*	.05	6.10*	.82	.22**
Overall Manager \times Self-Ratings	.35	.00	.01	.00	.00	.00	.00
Total	.35	.12					

* $p < .05$; ** $p < .01$.

TABLE 3 Summary of Hierarchical Regression Analysis for Variables Predicting Assessment Center Performance with Factor Scale Manager Ratings as Independent Variables (N = 144)

Variable	MultiR	R ²	F(Eqn)	ΔR ²	F _{ch}	SE	β
In-Basket Performance							
Manager Performance Ratings	.57	.33	53.56**	.33	53.56**	.09	.38*
Self-Ratings	.61	.37	32.26**	.04	7.69*	.06	.22*
Manager Communication Ratings	.63	.40	24.22*	.03	5.50*	.19	.23*
Total	.63	.40					
Overall Assessor Ratings (OAR)							
Self-Ratings	.26	.07	8.21**	.07	8.21**	1.15	.26*
Manager Interpersonal Ratings	.36	.13	8.19**	.06	7.67**	.08	.24*
Total	.36	.13					

Note: No interaction items significantly contributed to in-basket performance or assessment center performance when entered as a block in step 2.
*p < .05; **p < .01.

(Table 3). Interaction terms were again entered as a block in the second step to test the incremental contribution to the prediction of assessment center performance above that of self-ratings and factor-derived manager ratings.

Self-ratings added incrementally to predictions of in-basket performance above that of overall manager ratings ($\Delta R^2 = .06$, $p < .01$). Self-ratings and overall manager ratings also significantly contributed towards prediction of overall assessment center performance, jointly accounting for 12% of the variance in this dependent variable (Table 2).

Table 3 summarizes additional hierarchical regression analyses with self-ratings, manager ratings of communication skills, manager ratings of interpersonal skills, and manager ratings of task management skills as independent variables entered as a block in the first step, and in-basket performance and overall assessment center performance (OAR) as the dependent variables. Self-ratings and manager ratings of communication skills incrementally contributed towards prediction of in-basket performance above that of manager ratings of task management skills ($\Delta R^2 = .03$, $p < .05$ and $\Delta R^2 = .03$, $p < .05$, respectively). Manager ratings of interpersonal skills, but not communication skills or task management skills, significantly contributed towards prediction of overall assessment center performance ($\Delta R^2 = .06$, $p < .01$). These results provide support for the hypothesis that both self-ratings and manager ratings are significant predictors of both subjective and objective measures of assessment center performance.

Agreement grouping (in-agreement/good raters, in-agreement/poor raters, underestimators and overestimators) and assessment center performance served as the independent and dependent variables, respectively in an analysis of variance. The four-group categorization technique, introduced by Atwater and Yammarino (in press), uses a two-step process to categorize individuals in one of four categories. The first step consists in categorizing individuals whose self-manager difference (based on z scores of self-ratings and manager ratings) scores lie within one-half standard deviation of the mean difference; those who meet this criterion are in-agreement raters.

Those whose self-manager difference scores lie above one-half standard deviation are categorized as overestimators and those whose difference scores lie below are underestimators. The one-half standard deviation is an arbitrary margin set to create a broader band for inclusion in the in-agreement group. The second step of this process involves dichotomizing the in-agreement group based on whether the manager's ratings are above (good) or below (poor) the average rating of managers in the sample. If the average manager rating is above average, the

TABLE 4 Mean Assessment Center Performance by Agreement Groups Based on Standardized Self-Ratings and Overall Manager Ratings

Agreement Group	Mean	SD	N
In-Basket Performance¹			
Underestimators	-2.77	4.87	26
In-Agreement/Poor	-1.67	5.13	43
In-Agreement/Good	.51	5.08	28
Overestimators ²	1.61	5.84	27
Overall Assessor Ratings (OAR)³			
Underestimators	18.69	6.60	26
In-Agreement/Poor	22.51	6.67	43
In-Agreement/Good ⁴	24.65	6.73	29
Overestimators ⁴	24.29	4.52	27

¹ Significant one-way analysis of variance of in-basket performance by agreement group [$F(3,113) = 3.50, p = .018$].

² Significantly higher performance on the in-basket by overestimators compared to in-agreement/poor or overestimator groups based on Tukey's HSD test, $p < .01$.

³ Significantly one-way analysis of variance of overall assessment center performance ratings by agreement group [$F(3,114) = 3.69, p = .014$].

⁴ Significantly higher overall assessor ratings by in-agreement/good and overestimator groups compared to the underestimator group based on Tukey's HSD test, $p < .01$.

individual is categorized as an in-agreement/good rater. If the average manager rating is below average, the individual is categorized as an in-agreement/poor rater.

A univariate F value was computed and post hoc comparisons were made using Tukey's HSD method for both subjective and objective assessment center performance measures using the overall manager rating score (Table 4). The results of the ANOVA indicated that the effects of agreement group on overall assessment center performance (OAR) and in-basket performance were significant [$F(3, 114) = 3.69, p = .014$ and $F(3, 113) = 3.51, p = .018$, respectively]. Differences in effectiveness ratings among the agreement groups for both measures of assessment center performance are shown in Table 4. With respect to overall assessment center performance (OAR), assessors rated in-agreement/good rat-

ers and overestimators significantly higher than underestimators. There were no significant differences, however, between in-agreement/good raters and in-agreement/poor raters. With respect to in-basket performance, overestimators were observed to have higher scores compared to in-agreement/poor or overestimator groups based on Tukey's HSD test, $p < .01$. There were no significant differences among the other agreement groups for in-basket performance (Table 4). Separate post hoc analyses were also performed using the separate factor-derived manager ratings scores to create the agreement groupings (manager ratings of communication, interpersonal, and task management skills). In each of the subsequent ANOVAs, no significant differences were observed between in-agreement/poor raters, in-agreement/good raters, overestimators and underestimators.

DISCUSSION

This study investigated the relationship between self-ratings and manager ratings on assessment center performance with 144 production supervisors. Results of the hierarchical multiple regression analyses provided support for the hypothesis that self-ratings and manager ratings significantly contribute independently towards prediction of both objective and subjective measures of assessment center performance (Tables 2 and 3).

Results from the ANOVA tests and post hoc comparisons provided support for the hypothesis that in-agreement/good raters performed significantly better on an in-basket simulation and were rated higher in assessment center performance than in-agreement/poor raters (Table 4). However, the results of these analyses did not support differences between underestimator and overestimator groups on either the objective (in-basket scores) or subjective measures of assessment center performance (OAR).

Manager ratings of task management skills and communication skills were predictive of in-basket scores but not overall assessment center performance (Table 3). Additionally, manager interpersonal skill ratings, but not communication or task management skills, significantly contributed towards overall assessment center performance ($\Delta R^2 = .06, p < .01$). These results provide limited construct validity for the measures of assessment center performance used in this study in that the in-basket simulation emphasizes largely task and project management skills, whereas, the overall assessor ratings were based solely on performance on three separate behavioral group and interactive exercises emphasizing largely interpersonal dimensions.

The results of this study support the recent findings of Shore et al. (1992), Bass and Yammarino (1991), McCall and Lombardo (1983), and McCauley and Lombardo (1990) and providing limited construct validity for the use of self-evaluations in assessment centers. These results would appear to be promising in light of previous research suggesting that accurate self-insight and awareness is associated with diverse measures of individual and organizational success and effectiveness outcomes (Atwater & Yammarino, in press).

For example, McCall and Lombardo (1983) found that inflated self-ratings, characteristic of overestimators, were significantly associated with career derailment and failure in managers and executives. Derailment might have resulted due to self-aspirations and self-expectations far exceeding others' perceptions of skills and abilities. With respect to career advancement, McCauley and Lombardo (1990) found that self-rating accuracy was positively associated with supervisors' assessments of promotability; that is, the more accurate the self-rating, the more likely an individual was promoted. In their study of naval officers, Bass and Yammarino (1991) found that those with more accurate self-ratings compared to those of their supervisors attained higher ranks and were rated as more promotable by their senior officers. Additional support for the value of self-insight for managerial effectiveness comes from the work of Smither, Wohlers, and London (1995) who found that managers who rated themselves low and received low subordinate ratings (in-agreement/poor estimators) did not improve their performance over time, while those with low subordinate and moderate or high self-ratings did improve.

In this study, self-rating scores across the 14 assessment center dimensions were found to be significantly, albeit modestly, correlated with OAR ($r = .26, p < .01$) and specific in-basket dimensions (Table 1). Although self-ratings may be influenced by the purpose by which they are used (e.g., Farh & Werbel, 1986; Nowack, 1992; Tetlock & Boettger, 1989), future research should continue to explore the use and validity of self-evaluations in assessment centers. Overall, these results suggest that greater theoretical, practical, and empirical exploration of the contribution of self-assessment to the validity of assessment center is warranted.

Despite limited support for the utility of accurate self-awareness (agreement with manager ratings) in this study, several limitations should be pointed out. First, this study was cross-sectional in design, utilized a relatively small sample size (144) that was exclusively male and, as such, may not be highly generalizable to other nonmanufacturing and production environments.

Second, self-ratings collected in this study were used largely for developmental purposes. In general, self-ratings and other ratings are more likely to be inflated when they are used for evaluation, rather than, developmental purposes (Farh & Werbel, 1986; Mabe & West, 1982) and perhaps less accurate overall when there is no pressure to justify one's ratings to others (Tetlock, 1983). As such, the results of this study must be considered in light of the developmental emphasis of this assessment center without any accountability to provide accurate ratings.

Another limitation might be the approach used for assessing self-other agreement in this study. Self-other agreement was defined based upon the agreement category model first introduced by Atwater and Yammarino (1992) to specify a classification of individuals whose self-ratings agreed in varying degrees with ratings of others. Other approaches include calculations of difference scores, "Within and Between Analysis" and polynomial regression analysis (Atwater & Yammarino, in press). Atwater and Yammarino (in press) point out, "all three approaches offer distinct advantages over the use of difference scores and could be used alone or in combination for assessing self-other agreement."

Recently, Brutus, Fleenor, and Taylor (1996) building on the four-group categorization technique introduced by Atwater and Yammarino (in press) introduced further distinctions within each performance level—overestimators/good and underestimators/poor categories. They expanded the four-group categorization to six groups conceptualizing two additional groups that are based upon the degree of difference between self and other ratings for both good performance and poor performance. If the standardized self-rating of a good performance is more than one-half standard deviation above the standardized other rating, the individual is categorized as an overestimator/good rater. The same procedure is followed for the poor performers resulting in overestimator/poor, underestimator/poor and in-agreement/poor agreement groups.

In their study exploring rating congruence of 2,292 managers with a measure of leader effectiveness, Brutus et al. (1996) concluded that the addition of two groups to the Atwater and Yammarino (in press) four-group agreement categorization eased the interpretation of the results indicating that overestimator/good, in-agreement/good and underestimator/good raters were found to be significantly more effective than overestimator/poor, in-agreement/poor and underestimator/poor groups. Clearly, further research is needed to explore the methodological strengths and weaknesses of various agreement categorization methods in the study of self-other ratings and diverse measures of job performance and organizational success.

Although not a direct focus of the current study, future research should also investigate the extent to which self-evaluations and subsequent behavior change in program participants after receiving multi-source feedback. In general, the few empirical results to date have yielded mixed findings despite general findings that support the argument that employees tend to maintain a favorable view of themselves that leads to higher self-evaluations relative to ratings provided by others (cf. Campbell & Lee, 1988; Hazucha, Hezlett, & Schneider, 1993; Klimoski & Brice, 1991; Yu & Murphy, 1993).

For example, Yammarino and Atwater (1993) reported that feedback from subordinates altered self-evaluations of leadership but London and Smither (1995) did not find evidence of changes in self-ratings in a 6-month follow-up study. Atwater et al. (1995) reported that self-ratings for leaders who rated themselves high initially and received negative feedback declined whereas self-ratings for leaders receiving positive feedback who initially rated themselves low increased. As London and Smither (1995, p. 817) point out, "the organization wants the discrepancy to be reduced through improved performance by the employee—not through the employee lowering his or her self-evaluation, and not through the employee merely using impression management techniques to increase the level of feedback from others." Additional research in this area would help to clarify for whom developmental feedback from others leads to behavioral change, how changes in self-awareness lead to modifications in actual management practices, and whether these changes result in increased managerial effectiveness and career success.

Future research should attempt to replicate this study utilizing prospective designs and the full-range of feedback sources (e.g., direct reports, peers, team members, and customers). Other organizational measures of job performance and success should also be included in future empirical studies including career promotion and advancement, derailment, promotion, and salary increases. Finally, future research should attempt to investigate differences in self-other rating agreement with respect to gender, tenure, and relevant personality variables that might moderate these relationships (e.g., repressive coping style).

For example, it might be hypothesized that in-agreement/good raters who have a repressive coping style (characterized as those high in self-deception and low in negative affectivity) might be as vulnerable to derail in their managerial roles as other high risk groups (e.g., overestimators). Although it would appear that such individuals possess accurate self-insight based upon the self-other agreement model proposed by Atwater and Yammarino (1992), such individuals might be prone to creating favorable self-impressions and therefore, blind to their weaknesses. These

individuals might be expected to be least likely to solicit feedback from others, accept critical information about their behavior from others, and ultimately modify his/her management practices.

In summary, this study provides support for the construct validity of self-evaluation in assessment centers as well as demonstrating that accurate self-insight is associated with both objective (in-basket simulation) and subjective (overall assessor ratings) measures of performance. Additional research is needed to clarify the nature of self-other rating.

REFERENCES

- Antonioni, D. (1996, April). *Designing an effective 360-degree appraisal feedback process*. Paper presented at the Annual Conference of the Society of Industrial and Organizational Psychology, San Diego, CA.
- Ashford, S. (1989). Self assessments in organizations: A literature review and integrative model. In L.L. Cummings & B.M. Staw (Eds.), *Research in organizational behavior* (Vol. 11, pp. 133-174), Greenwich, CT: JAI Press.
- Atwater, L., & Yammarino, F. (1992). Does self-other agreement on leadership perceptions moderate the validity of leadership and performance predictions? *Personnel Psychology, 45*, 141-164.
- Atwater, L., Roush, P., & Fischthal, A. (1995). The influence of upward feedback on self- and follower ratings of leadership. *Personnel Psychology, 48*, 35-59.
- Atwater, L., & Yammarino, F. (in press). Self-other rating agreement: A review and model. *Research in Personnel and Human Resource Management*.
- Bandura, A. (1982). Self-efficacy mechanisms in human agency. *American Psychologist, 37*, 122-147.
- Bass, B.M., & Yammarino, F. (1991). Congruence of self and other's leadership ratings of naval officers for understanding successful performance. *Applied Psychology: An International Review, 40*, 437-454.
- Brannick, M.T., Michaels, C.E., & Baker, D.P. (1989). Construct validity of in-basket scores. *Journal of Applied Psychology, 74*, 957-963.
- Bray, D.W., & Grant, D.L. (1966). The assessment center in the measurement of potential for business management. *Psychological Monographs, 80*(17, Whole No. 625), 1-27.
- Brutus, S., Fleenor, J.W., & Taylor, S. (1996, April). Paper presented at the Eleventh Annual Conference for the Society for Industrial and Organizational Psychology, San Diego, CA.
- Campbell, D., & Lee, C. (1988). Self-appraisal in performance evaluation: Development versus evaluation. *Academy of Management Review, 13*, 302-314.
- Crowne, D.P., & Marlowe, D. (1964). *The approval motive: Studies in evaluative dependence*. New York: Wiley.
- DeNisi, A., & Shaw, J. (1977). Investigation of the uses of self-reports of abilities. *Journal of Applied Psychology, 62*, 641-644.
- Edwards, A.L. (1957). *The social desirability variable in personality assessment and research*. New York: Dryden.
- Farh, J., & Werbel, J. (1986). The effects of purpose of the appraisal and expectation of validation on self-appraisal leniency. *Journal of Applied Psychology, 62*, 641-644.
- Fedor, D., Rensvold, R., & Adams, S. (1992). An investigation of factors expected to affect feedback seeking: A longitudinal field study. *Personnel Psychology, 45*, 779-805.
- Gaugler, B.B., Rosenthal, D.B., Thornton, G.C., III, & Bentson, C. (1987). Meta-analysis of assessment center validity. *Journal of Applied Psychology, 72*, 493-511.
- Harris, M.M., & Schaubroeck, J. (1988). A meta-analysis of self-supervisor, self-peer, and peer-supervisor ratings. *Personnel Psychology, 41*, 43-62.
- Hazucha, J., Hezlett, S., & Schneider, R. (1993). The impact of 360-degree feedback on management skills development. *Human Resource Management, 32*, 325-351.
- Hoffman, R. (1995, April). Ten reasons you should be using 360-degree feedback. *HR magazine, 40*, 82-85.
- Kagan, J. (1988). The meanings of personality predicates. *American Psychologist, 43*, 614-620.
- Klimoski, R.J., & Brice, T. (1991, April). *The nature and impact of self-appraisals in a manufacturing environment*. Paper presented at the Sixth Annual Conference of the Society for Industrial and Organizational Psychology, St. Louis, MO.
- Klimoski, R.J., & Brickner, M. (1987). Why do assessment centers work? The puzzle of assessment center validity. *Personnel Psychology, 40*, 243-260.
- London, M., & Beatty, R. (1993). 360-degree feedback as competitive advantage. *Human Resource Management, 32*, 353-373.
- London, M., Wohlers, A., & Gallagher, P. (1990). 360-degree feedback surveys: A source of feedback to guide management development. *Journal of Management Development, 9*, 17-31.
- London, M., & Smither, J.W. (1995). Can multi-source feedback change perceptions of goal accomplishment, self-evaluations, and performance related outcomes? Theory-based applications and directions for research. *Personnel Psychology, 48*, 803-839.
- Mabe, P.A., & West, S.G. (1982). Validity of self-evaluation of ability: A review and meta-analysis. *Journal of Applied Psychology, 67*, 280-296.
- McCall, M., & Lombardo, M. (1983). *Off the track: Why and how successful executives get derailed*. (Technical Report No. 21). Greensboro, NC: Center for Creative Leadership.
- McCauley, C., & Lombardo, M. (1990). Benchmarks: An instrument for diagnosing managerial strengths and weaknesses. In K.E. Clark & M.B. Clark (Eds.), *Measures of leadership* (pp. 535-545). West Orange, NJ: Leadership Library of America.
- Nowack, K.M. (1992). Self-assessment and rater-assessment as a dimension of management development. *Human Resources Development Quarterly, 3*, 141-155.
- Nowack, K.M. (1989). *OPD in-basket software and administration manual*. Los Angeles: Organizational Performance Dimensions.
- O'Reilly, B. (1994, October 17). 360-degree feedback can change your life. *Fortune, 129*, 93-96.
- Reilly, R.R., & Chao, G. (1982). Validity and fairness of some alternative employee selection procedures. *Personnel Psychology, 35*, 1-62.
- Russell, C.J. (1987). Person characteristics versus role congruency explanations for assessment center findings. *Academy of Management Journal, 30*, 817-826.
- Sackett, P.R., & Dreher, G.F. (1982). Constructs and assessment center dimensions: Some troubling empirical findings. *Journal of Applied Psychology, 67*, 401-410.
- Schippmann, J.S., Prien, E.P., & Katz, J. (1990). Reliability and validity of in-basket performance measures. *Personnel Psychology, 43*, 837-859.
- Schneider, J.R., & Schmitt, N. (1992). An exercise approach to understanding assessment center dimension and exercise constructs. *Journal of Applied Psychology, 1*, 32-41.
- Shrauger, J.S., & Schoeneman, T.J. (1979). Symbolic interactionist view of self-concept: Through the looking glass darkly. *Psychological Bulletin, 86*, 549-572.

Schwartz, G. (1990). Psychobiology of repression and health. In J.L. Singer (Ed.), *Repression and dissociation: Implications for personality theory* (pp. 405-434). Chicago: University of Chicago Press.

Shore, T.H., Thornton, G.C., III, & Shore, L.M. (1990). Construct validity of two categories of assessment center dimension ratings. *Personnel Psychology*, *43*, 101-116.

Shore, T.H., Shore, L.M., & Thornton, G.C., III (1992). Construct validity of self- and peer evaluations of performance dimensions in an assessment center. *Journal of Applied Psychology*, *77*, 42-54.

Smither, J.W., Wohlers, A., & London, M. (1995). A field study of reactions to normative versus individualized upward feedback. *Group and Organization Management*, *20*, 61-89.

Smither, J.W., London, M., Vasilopoulos, N., Reilly, R.R., Millsap, R., & Salvemini, N. (1995). An examination of the effects of an upward feedback program over time. *Personnel Psychology*, *48*, 1-33.

Smircich, L., & Chesser, R. (1981). Superior's and subordinates' perceptions of performance: Beyond disagreement. *Academy of Management Journal*, *24*, 198-205.

Taylor, S.E., & Brown, J.D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin*, *103*, 193-210.

Tetlock, P. (1983). Accountability and the perseverance of first impressions. *Social Psychology Quarterly*, *46*, 74-83.

Tetlock, P., & Boettger, R. (1989). Accountability: A social magnifier of the dilution effect. *Journal of Personality and Social Psychology*, *57*, 388-397.

Tomow, W. (1993). Perceptions or reality: Is multi-perspective measurement a means or an end? *Human Resource Management*, *32*, 221-230.

Thornton, G.C., III (1980). Psychometric properties of self-appraisals of job performance. *Personnel Psychology*, *33*, 263-271.

Thornton, G.C., III, & Byham, W.C. (1982). *Assessment centers and managerial performance*. New York: Academic Press.

Thornton, G.C., III, & Cleveland, J.N. (1990). Developing managerial talent through simulation. *American Psychologist*, *45*, 190-199.

Tziner, A. (1984). Prediction of peer rating in a military assessment center: A longitudinal follow-up. *Canadian Journal of Administrative Sciences*, *1*, 146-160.

Van Velsor, E., Taylor, S., & Leslie, J. (1993). An examination of the relationships among self-perception accuracy, self-awareness, gender and leader effectiveness. *Human Resource Management*, *32*, 249-263.

Yammarino, F., & Atwater, L. (1993). Understanding self-perception accuracy: Implications for human resources management. *Human Resource Management*, *32*, 231-247.

Yu, J., & Murphy, K.R. (1993). Modesty bias in self-ratings of performance: A test of cultural relativity hypothesis. *Personnel Psychology*, *46*, 357-363.

Journal of Social Behavior and Personality

How to Submit Papers

The submission of good papers is very important to us as a new Journal. We promise to mail a decision on your paper within 5 weeks of receipt by us. We promise to publish your paper within an average of 4 months of receipt of a final accepted version. Our submission and publication requirements are rather unusual, so please inquire about details *before* submitting to the editor, Dr. Rick Crandall, Journal of Social Behavior and Personality, P.O. Box 37, Corte Madera, CA 94976-0037, 415/924-1612.

Because of our ongoing research on the review process, new reviewers and articles on the publication process are always wanted. Consulting editors are selected on the amount and quality of editorial work done.

Subscriptions

All subscriptions are on a calendar year basis only. If you order late in the year, the earlier issues will be sent to you. Institutional rates, \$70. Individual rates, paid by personal check or credit card only, \$30. Add \$8 for subscriptions outside the United States. A subscription includes four regular issues *plus* any special issue(s) published during the year.

Order Form

<input type="checkbox"/> Yes! Start/renew my subscription for 1997.	Amount
<input type="checkbox"/> Four regular issues and one special issue (\$30 individuals; \$70 institutions)	\$
Please send me the following back issues: (Special issues are included)	
<input type="checkbox"/> 1996 (4 regular issues, 1 special issue) (\$30 individuals; \$65 institutions)	
<input type="checkbox"/> 1995 (4 regular issues, 2 special issues) (\$30 individuals; \$65 institutions)	
1986-1994 also available at the same price.	
Special issues are included in yearly subscriptions. They may also be purchased individually.	
<input type="checkbox"/> Assessment Centers: Research and Applications (1997, 338 pp., \$30)	
<input type="checkbox"/> Handbook of Gender Research (1996, 322 pp., \$30)	
<input type="checkbox"/> Gender in the Workplace (1995, 308 pp., \$30)	
<input type="checkbox"/> Psychosocial Perspectives on Disability (1994, 424 pp., \$30)	
<input type="checkbox"/> Replication Research in the Social Sciences (1993, 196 pp., \$20)	
<input type="checkbox"/> Handbook of Post-Disaster Interventions (1993, 488 pp., \$30)	
<input type="checkbox"/> Handbook of Self-Actualization (1991, 362 pp., \$20)	
<input type="checkbox"/> Type A Behavior (1990, 460 pp., \$20)	
<input type="checkbox"/> Communication, Cognition, and Anxiety (1990, 225 pp., \$20)	
<input type="checkbox"/> Handbook of Replication Research (1990, 530 pp., \$20)	
<input type="checkbox"/> Mood and Memory (1989, 190 pp., \$15)	
<input type="checkbox"/> The State of Social Psychology (1988, 145 pp., \$15)	
<input type="checkbox"/> Work & Family: Theory, Research and Applications (1988, 440 pp., \$20)	
<input type="checkbox"/> Loneliness: Theory, Applications, and Research (1987, 302 pp., \$20)	
<i>Shipping: Included for all U.S. orders. For shipments outside the U.S., add \$8 for each yearly subscription, add \$3 for each special issue ordered individually.</i>	
TOTAL (All payments must be in U.S. dollars, checks drawn on a U.S. bank. Visa / Master Card and AMEX orders accepted—include card number, signature, and expiration date.)	\$
<i>Thank you for your order.</i>	

Name _____

Address _____

City, State, Zip _____

Mail to: Select Press • P.O. Box 37 • Corte Madera, CA 94976-0037 • 415/924-1612