Managerial Self-Awareness in High-Performing Individuals in Organizations

Allan H. Church
W. Warner Burke Associates

Individual differences in the construct of managerial self-awareness (MSA)—operationalized as congruence between self and direct reports' behavioral ratings—were examined in 134 high-performing (HP) and 470 average-performing (AV) managers obtained from 4 independent datasets. Results based on several different approaches to measuring ratings agreement indicated that HPs were significantly more managerially self-aware compared with AVs. This relationship was consistent regardless of data source, organization, or method of assessing managerial performance. No overall relationships were found between congruence and level of item importance, gender, management level, age, or tenure. When compared with other measures for assessing self-focus, the construct of self-monitoring was found to be convergent with managerial self-awareness, whereas the construct of self-consciousness appeared to reflect primarily rating leniency effects. The article concludes with a comparison of the measurement approaches used, limitations, and suggestions for further study.

With the rising popularity of multirater or 360 degree feedback systems in organizations (e.g., Church, 1995; Hazucha, Hezlett, & Schneider, 1993; London & Beatty, 1993) designed explicitly to enhance self-knowledge and improve managerial behavior, a number of researchers have begun exploring the psychological concept of self-awareness from an applied ratings-based perspective (e.g., Atwater & Yammarino, 1992; Van Velsor, Taylor, & Leslie, 1993; Wohlers & London, 1989). From this perspective, the greater a given individual's level of managerial self-awareness (MSA) (i.e., the ability to reflect on and accurately assess one's own behaviors and skills as they are manifested in workplace interactions), the greater the expected level of congruence between his or her self-assessments and those provided by direct reports or by peers. Conceptualized as both a personality trait and a skill, the construct of MSA is likely to interact with the self-perception process prior to assessment (Campbell & Lee, 1988; Figurski, 1987; Yammarino & Atwater, 1993), which in turn moderates the "self-other" rating comparison. By applying the notion of MSA to performance ratings research, this perspective has provided some new and interesting insights into the general lack of agreement observed in ratings from multiple sources (e.g., Furnham & Stringfield, 1994; Harris & Schaubroeck, 1988; Landy & Farr, 1980).

Although MSA is intrinsically interesting, its significance to researchers and practitioners is limited unless differences in self-awareness, as measured by congruence in ratings, can be linked to other key variables such as individual performance. Unfortunately, although much applied research has explored managers' perceptions of their own behaviors compared with those of their subordinates, peers, and supervisors (e.g., Arnold & Davey, 1992; Borman, White, & Dorsey, 1995; Riggio & Cole, 1992; Wohlers, Hall, & London, 1993), the number of studies specifically linking levels of self-other ratings congruence to managerial effectiveness or performance are relatively few. Moreover, the generalizability of the findings that do exist (e.g., Atwater & Yammarino, 1992; Van Velsor et al., 1993), although important in establishing a relationship between performance and MSA, has been somewhat limited because of small sample sizes and the unique nature of the observed organizations (i.e., naval officers and hospital administrators). Thus, even though practitioners in the field of organizational change and development have been operating according to the principle of enhancing self-awareness through data-based methods for years (e.g., Burke, 1982; Nadler, 1977), and numerous leadership, executive, and management development efforts have been initiated on the steadfast belief that enhanced self-awareness leads to increased performance (e.g., Bernardin, 1986; Burke, Richley, & DeAngelis, 1985; Ship-
per & Neck, 1990), researchers are just beginning to build evidence to support this contention. Such proof is needed, however, to (a) justify the staffing and production costs associated with multirater feedback efforts and (b) advance understanding of the rating process in general.

The purpose of this article is to replicate and extend the results of previous studies (e.g., Atwater & Yammarino, 1992; Van Velsor et al., 1993) regarding the construct of MSA and its link to performance vis-à-vis self–other ratings congruence. This will be achieved through the use of (a) several different organizational samples that are more representative of general management populations and (b) a variety of different measures of managerial performance. In addition, given the need to validate the use of self–other ratings congruence as a means of assessing MSA, the relationship among this operationalization and other more established measures of self-directed attention—for example, self-monitoring (Snyder, 1974) and self-consciousness (Fenigstein, Scheier, & Buss, 1975)—is also examined. Finally, the impact of various moderating variables (e.g., gender, management level, age, and tenure) is explored as well.

More specifically, several research hypotheses were tested. As noted above, many practitioners and researchers (e.g., Ashford & Tsui, 1991; McCall, Lombardo, & Morrison, 1988; Sinetar, 1988) have argued that self-awareness is an important ability and skill set for achieving managerial excellence. Thus, better managers should be more adept at assessing the level of their own behaviors and the impact those behaviors have on others. Moreover, high performers and high potentials in organizations should also exhibit greater levels of MSA because they presumably have already been identified as possessing promising managerial characteristics such as strong interpersonal, leadership, and communication skills (Derr, Jones, & Toomey, 1988; Garfield, 1986) and are often given significantly greater developmental opportunities that serve to enhance their abilities.

There will be a significant relationship between managerial performance and congruence in self–other behavioral ratings. High-performing and/or high-potential individuals will be more managerially self-aware and will, therefore, achieve higher congruence between self-based and other behaviorally based ratings compared with average-performing individuals. (Hypothesis 1)

The contribution of individual cognitions to the processing of information, decision making, and memory functioning (e.g., Bandura & Jourden, 1991; Hendrick, 1990; Lewicki, 1983) must also be considered as a potential moderator in any ratings-based self-awareness perspective. If a given behavior is not considered meaningful; that is, its standard of comparison is considerably lower relative to other behaviors (Carver & Scheier, 1981; Klein, 1989), then the individual is far less likely to monitor, assess, and cognitively record his or her actions in that given area (DeNisi & Williams, 1988; Zalesny & House, 1992). Conversely, those behaviors perceived as important to an individual should receive greater attention in the observation process, yielding more accurate self-ratings (i.e., in agreement with others' observations).

There will be a significant relationship between level of perceived importance of the behaviors being rated and overall congruence in self–other ratings received. The more an individual manager identifies a behavior as important to his or her managerial success, the more perceptual attention will be directed to its practice and, therefore, the greater the probability that MSA (i.e., overall congruence in self-ratings vs. others' ratings) will be enhanced. (Hypothesis 2)

Individual differences in convergent theories of personality must also be considered in any attempt to explain differences in the level of agreement in performance ratings. In particular, both the self-monitoring (Snyder, 1974, 1987) and the self-consciousness constructs (e.g., Buss, 1980; Fenigstein, 1987; Fenigstein et al., 1975) appear to have direct relevance to the level of self–other agreement obtained if this measure in fact reflects an appropriate operationalization of MSA. In summary, self-monitoring theory posits that individuals differ on the extent to which they actively monitor, regulate, and/or control their own expressive behaviors in the presence of others. Low self-monitors tend to behave according to their own inner states rather than external socially driven cues, and high self-monitors behave in the opposite; that is, they are driven by interpersonal situations rather than internal convictions (Briggs & Cheek, 1988; Snyder, 1987). Thus, high self-monitors are more likely to attend to and assess their own behaviors in the workplace.

There will be a significant relationship between self-monitoring and MSA (agreement in ratings) such that high self-monitors will demonstrate greater congruence. (Hypothesis 3)

Similarly, Fenigstein et al.'s (1975) notion of self-consciousness conceptualizes individuals as falling along a continuum with regard to level of self-focus. They described two distinct components of self-focus: public and private self-consciousness. Public self-consciousness (SC) occurs when one feels under the scrutiny of a "social other" in regard to personal appearance, manners, presentation style, or physical quirks. People high in public SC are generally more sensitive to the opinions and evaluations of others based on their physical appearance in social situations (Bennett & Buchanan-Barrow, 1993; Buss, 1980; Schlenker & Weigold, 1990). In comparison, private self-consciousness reflects an orientation toward the inner world of ideas, concepts, motives, emotions, memories, and other types of self-reflection. This focus
on the purely internal aspects of self results in both an intensification of the affective charge, deepening depression, heightening elation, and intensifying pain, and in a greater clarification of the phenomena to the self, thus allowing more precise knowledge about the given experience (Buss, 1980). Individuals high in private SC are better able to assess their own behaviors in such areas as aggression, humorous stimuli, speed of processing information, and compliance; they also have quicker access to and more highly developed self-schemata (e.g., Brière & Vallerand, 1990; Gibbons, 1983; Nasby, 1989).

There will be a significant relationship between private SC and MSA (agreement in ratings) such that people high in private SC will achieve greater congruence. (Hypothesis 4)

Method

Participants

Data were collected from 134 high-performing (HPs) and 470 average-performing (AVs) middle- and senior-level managers and their direct reports from three different organizations and industries: technological, pharmaceutical, and airline services. Although three organizations are represented, four independent datasets were used for analytic purposes (one senior-level sample and one middle management-level sample from the same technological firm). All individuals were involved in some form of leadership, management, or executive training seminar prior to which various measures of managerial behavior were collected from their direct reports. Responses were collected with an anonymous distribution process. This process ensured as much as possible that the ratings obtained from others reflected an accurate assessment of that manager’s behavior in the workplace. Moreover, because ratings were completed independently and were not divulged until the conduct of the program, it is extremely unlikely that direct reports’ scores were influenced by the manager’s self-assessments or vice versa. These data were subsequently delivered to each participant during the course of the program in the form of individual feedback reports for development purposes. The selection process for inclusion in the development programs differed somewhat among the four samples. In the airline service and pharmaceutical organizations, the entire senior management population was included in the feedback process. In the technological organization, both the senior-executive and middle-manager samples were a combination of (a) selection by superiors for attendance and (b) interest in development and training.

In total, high performers were rated by 973 direct reports and average performers by 3,398 direct reports. The mean age of the manager sample among the 74.1% that provided this information was 49.4 years (SD = 9.44). The average tenure with the organization, based on responses from 76.8% of respondents, was 16.2 years (SD = 9.14). There were no significant differences between the two groups in mean age or overall tenure. In terms of gender, the sample was heavily biased, with 87.7% men and 12.3% women. Most women were from the middle-manager sample in the technological organization (28.6% of Sample 1). A significant chi-square indicated that there was a greater proportion of women among the HPs than in the AVs; 20% vs. 10%; χ²(1, N = 514) = 8.3, p < .01. With respect to management level, 55.6% of the sample were middle managers, and 44.4% were senior executives. Each of the three organizations had adequate representation at both levels. For the technological organization, Sample 1 was comprised entirely of middle managers, whereas Sample 2 included only senior managers. Sample 3 and Sample 4 were also split between middle and senior levels at 49.4% versus 50.6% and 64.9% versus 35.1%, respectively. There were no significant demographic relationships between level and performance group.

Procedure

Behavior ratings and demographic information were obtained from a database containing the accumulated scores from each company’s training program. Because individuals were assessed on a variety of different management behaviors, some of which were similar in terms of content, individual item—content level comparisons were not investigated, as the specific wording was such that direct content comparisons between samples would have been problematic. Instead, analyses were based on the overall level of agreement in self-reports’ versus direct reports’ average ratings across each organization’s instrument.

Individuals in each of the four datasets were assigned to either a high-performing or an average-performing classification on the basis of information obtained from their respective organizations. Table 1 provides an overview of the selection criteria used for each sample and the resulting number of managers selected in the HP and AV groups. Methods of performance assessment ranged from external assessment that was based on case histories of each manager to a more informal selection by the highest level managers, a process that despite its subjectivity is a quite common method of selecting future leaders in organizations (e.g., Derr et al., 1988; Garfield, 1986). Although quantitative measures of performance are preferable for analytic purposes, it was not possible to obtain this level of detail. It should also be noted that at the time of the self–other data collection process, the performance group designation was neither known nor communicated to individual participants. Thus, the level of performance and/or the designation of being a high performer would not have had a direct impact on the actual ratings obtained from managers or their direct reports.

Measures

The degree of MSA for an individual was operationalized as the collective level of agreement or congruence between self-ratings and averaged others’ ratings across a number of behaviorally based indices. Because there is considerable diversity of opinion about how such agreement or congruence should be measured (e.g., Berger-Gross, 1982; Edwards, 1994; Johns, 1981; Tisak & Smith, 1994; Zimmerman, 1994), several different methods were generated and compared for analysis purposes. The primary method used for the results in this study consisted of the following.

1 Given the increasing popularity of research on behavioral ratings, a number of alternative methods of assessing congruence in ratings are available. Although the measures used in the present research represent more traditional methods for comparison,
Difference score approach. This measure of MSA used a standard index of similarity that was based on the absolute average profile difference in the ratings obtained. Referred to in the literature as $d$ (Edwards, 1994; Nunnally, 1978), this index is computed with the square root of the sum of squared (i.e., absolute) differences between self-report and average direct report score for each self–other item comparison divided by the total number of items for that sample. This measure was chosen because of its superior sensitivity to differences in profile level, dispersion, and shape (Nunnally, 1978; Tisak & Smith, 1994; Zimmerman, 1994). For comparison purposes, a second difference score measure was also generated. It was based on the relative average item difference between self-reports and direct reports across each of the behaviorally based questions. Computed with the actual arithmetic average of the differences obtained, this measure reflects the degree or direction, as well as the magnitude of the differences obtained (Wohlers & London, 1989). Difference scores of these types have been one of the primary means by which ratings congruence has been assessed in the literature (e.g., Ashford & Tui, 1991; Wexley & Pulakos, 1983; Wohlers et al., 1993; Zalesny & Highhouse, 1992).\(^2\) Because the formula for both of these measures utilized an average of uniquely matched self–direct report comparisons across a series of item ratings for each individual, these data had the potential to yield more sensitive results than would analyses that were based on more global comparison of the difference between total self and total direct reports’ scores. Other types of summary driven approaches (e.g., those used by Atwater & Yammarino, 1992, and Van Velsor et al., 1993) are ultimately measuring the general rating tendency of an individual—that is, as an underrater or an overrater—and not the cumulative degree of congruence across a series of specific comparisons.

Between-manager correlations. Another commonly reported measure of self–other congruence in the literature is the Pearson correlation between the average self-score and the corresponding average others’ ratings (e.g., Furnham & Stringfield, 1994; Nowack, 1992; Riggio & Cole, 1992). Although this value has been found to be notoriously low in various quantitative reviews (e.g., Harris & Schaubroeck, 1988; R. L. Heneman, 1986; Mabe & West, 1982) because it reflects overall standing in a given sample on the total instrument and not agreement in individual ratings, it is a useful measure to compute for comparison with previous research.

The behavioral items from each of the organization-specific instruments ranging from 21 for the middle managers from the technological firm to 49 for the service executives,\(^3\) were gener-

---

\(^2\) Although some researchers (e.g., Edwards, 1994) have argued that polynomial regression models are the preferred alternative to any type of difference score analysis, this approach was not used because it did not appropriately fit the categorical nature of the performance data collected. Moreover, this type of analysis assumes that the predictor is actually intended to account for the variance in a grouping variable (e.g., performance). Such a causal relationship was neither hypothesized in the present study nor testable, given the type and content of the data collected.

\(^3\) The impact of differences in the specific content and number of items across the four samples on the various measures of congruence in ratings was examined indirectly by including data source in each of the analyses. None of the results identified data source as a significant moderating variable.

---

**Table 1**

**Composition of Datasets**

<table>
<thead>
<tr>
<th>Industry, program source, and sample description</th>
<th>Performance group selection criteria</th>
<th>No. high performers</th>
<th>No. average performers</th>
<th>No. items</th>
<th>Self-ratings $\alpha$</th>
<th>Direct report ratings $\alpha$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Technology</td>
<td>Selection into a very competitive high potential senior executive service candidate development program based on work history, performance and achievements, and supervisor evaluations. Total pool of applicants over 1,000.</td>
<td>56</td>
<td>131</td>
<td>21</td>
<td>.94</td>
<td>.90</td>
</tr>
<tr>
<td>Technology</td>
<td>Recipient of any one of the following: meritorious, distinguished, or other position awards.</td>
<td>49</td>
<td>242</td>
<td>40</td>
<td>.90</td>
<td>.97</td>
</tr>
<tr>
<td>Pharmaceutical</td>
<td>Identified by senior most members of the organization as being high performers or high potential for future leadership role.</td>
<td>15</td>
<td>74</td>
<td>37</td>
<td>.89</td>
<td>.96</td>
</tr>
<tr>
<td>Airline service</td>
<td>Independent study by external consulting firm of executive performance and future potential based on combination of personality assessment, performance history, skills, and one-on-one interviews.</td>
<td>14</td>
<td>23</td>
<td>49</td>
<td>.97</td>
<td>.98</td>
</tr>
</tbody>
</table>

---

other approaches such as the three-group classification into underraters, accurate-raters, and overraters (e.g., Atwater & Yammarino, 1992; Van Velsor et al., 1993), the T index and the six category scheme (e.g., Brutts, Fleenor & Taylor, 1996), and $r_{sub(WD)}$ (e.g., James, Demaree, & Wolf, 1993) are also available.
ated on the basis of a series of individual interviews and focus groups with cross-sections of managers followed by piloting and subsequent refinement through statistical and item-response techniques. For each item, managers were asked to rate themselves on a scale of 1 to 5 (where 1 = a very small extent and 5 = a very great extent) as to the extent to which they engage in a given behavior (see the Appendix for sample items). These items were also independently rated by direct reports from individuals who worked closely with that individual. All items reflected behaviors exhibited by managers in the work setting. Because research has demonstrated that averaged ratings are more reliable and therefore better indicators of the behavior being rated than any single assessment (e.g., Furnham & Stringfield, 1994; London & Wohlers, 1991), the direct report scores were averaged by item for each manager.

In the technology organization (Samples 1 and 2), additional information regarding the importance of each of the behaviors being assessed was also collected. Individuals were also asked to assess the degree to which the practice of each behavior was important for the success of the manager being rated on a 1-to-5 scale where 1 = not very important and 5 = of critical importance. An average rating of this type represents the overall degree of perceived importance for a given behavior or set of behaviors.

Internal consistency coefficients and response analysis tables were computed for both self-reports and direct reports' average ratings for all items across each of the four datasets to test for problematic items. No items were identified for removal. Overall, the values of these alpha coefficients ranged from .89 to .97 for self-reports and .90 to .98 for averaged direct report ratings, indicating a high degree of internal consistency inherent in the behavioral measures for each organization.

Additional personality measures reflecting managers' level of self-focus were assessed with the 18-item true-false revised version of the self-monitoring questionnaire (Snyder, 1987) and the original 23-item 0 to 4 Likert-type scale version of the SC scale (Fenigstein et al., 1975). Both instruments have been widely used for research purposes and have been tested on multiple samples. Unfortunately, these additional data were available only for the high-performing middle managers in the technology organization (n = 53) assessed at the same time as the other measures. Thus, the resulting analyses with these variables were limited to this specific subsample. The internal reliability of the self-monitoring scale for these managers was .79. The alpha coefficients for the four self-consciousness measures were computed at .86 for the total score, .77 for the 10-item private scale, .79 for the 7-item public scale, and .78 for the 6-item social anxiety scale.

Analyses

Given the need to accommodate multiple measures of the level of agreement in self-direct report ratings, it was necessary to analyze the data with several different statistical approaches. First, because of the categorical nature of the performance data as well as the fact that a predictive relationship between performance and ratings congruence was neither hypothesized nor testable given the present data, analysis of variance (ANOVA) models were selected to test the relationships among the categorical variables (i.e., performance group, gender, level of management, and data source) and the two continuous criterion measures; that is, absolute average profile difference and relative average item difference score. Additional ANOVAs were also conducted on the total average self-reports and direct reports' scores by the same grouping variables to look for general trends in ratings. Simple correlation analyses and Fisher's z scores were used to test associations among importance scores, between manager correlations, and other individual variables (e.g., age, tenure or personality measures).

Results

Table 2 provides descriptive statistics and correlation coefficients for all the continuous variables included in the analyses. As reported in previous research (e.g., Arnold & Davey, 1992; Mabe & West, 1982; Nowack, 1992; Van Velsor et al., 1993), there is a general tendency on the part of managers to rate themselves somewhat higher across all behaviors than did their direct reports, t(603) = 9.13, p < .001. This inflation was also evident among importance ratings between self-reports and direct reports, t(477) = 12.49, p < .001. The high correlations between total behavior and importance scores for both self-reports (r = .63) and direct reports (r = .63) indicated some degree of common method variance (Mitchell, 1985; Spector, 1987; Williams, Cote, & Buckley, 1989) between the two sets of ratings as well.

In terms of the overall level of managerial self-awareness vis-à-vis congruence between self-reports and direct report ratings, the absolute average profile difference for all managers was 0.82 (SD = 0.32) and a range from 0.32 to 2.13. The relative average item difference score across these same items was 0.17 (SD = 0.46) and ranged from -.12 to 1.86. The correlation between these two difference score measures was r = .48 (p < .001). Both approaches reflect the average degree of difference observed across a series of individual item rating comparisons. The average between manager correlation, reflecting similarities in relative standing between self-reports and direct reports as compared with the total sample, was r = .25, p < .001. Although relatively weak, this correlation is similar to those reported in previous research and meta-analytic comparisons (e.g., Furnham & Stringfield, 1994; Harris & Schaubroeck, 1988; Wohlers & London, 1989).

In terms of demographic variables, results of a 2 X 2 X 2 X 4 factorial ANOVA model that included management level (middle vs. senior manager status), gender, performance group, and data source as the grouping variables and profile difference as the criterion yielded no significant effects for any variable except performance group. Management level, gender, and data source did not significantly contribute to the variance explained in the analysis as either single independent variables or in conjunction with others. Similar results were found by using
the average item difference score as the criterion. Although there were no significant relationships between age and any of the congruence measures, there was a small significant positive correlation between age and total average behavioral ratings for both self-reports and direct reports \( (r = .21, p < .001; r = .15, p < .001) \). Although the relationship between age and inflation of self-ratings appears to be a consistent phenomenon, prior research has suggested that older managers tend to be rated more negatively (e.g., Lawrence, 1988) or not differentially at all by others (London & Wohlers, 1991). Organizational tenure was significantly correlated only with age \( (r = .32, p < .001) \).

Thereupon, analyses testing each of the four research hypotheses were conducted. Support for Hypothesis 1 was provided by several different analyses. First, a \( 2 \times 4 \) factorial ANOVA model was conducted with the absolute profile difference measure as the criterion and performance group and data source as the grouping variables. Overall, those individuals identified as HPs yielded significantly lower profile difference scores, indicating greater absolute item level agreement than did NVs; \( M = .74 \) versus \( M = .85 \), \( F(1, 596) = 10.70, p < .001 \). Neither data source nor the interaction between performance group and data source was significant. The effect size \( (partial \ r^2) \) for performance group was .033, somewhere between a small and medium result (Cohen, 1977; Stevens, 1992). The same \( 2 \times 4 \) model with average item difference as the criterion instead yielded exactly the same result, \( F(1, 596) = 9.73, p < .01 \), with HPs \( (M = .07) \) demonstrating significantly greater average agreement in ratings compared with NVs \( (M = .20) \). This significant effect was also present for the between-manager correlations \( (r = .38 \text{ for HPs vs. } r = .20 \text{ for AVs, Fisher's } z = 2.03, p < .05) \).

Hypothesis 2 was tested with a correlational approach; that is, the Pearson correlation between managers' self-ratings of the degree of behavioral importance and their level of agreement in ratings. No significant correlations existed overall between importance ratings by self and level of profile difference (see Table 2). Interestingly, the relative average item difference between self-reports and direct reports' ratings was actually positively correlated with total ratings of item importance by self \( (r = .37, p < .001) \), suggesting instead that the more important the behaviors the greater the level of difference observed.

Finally, correlation analyses and one-way ANOVAs were used to test the remaining two hypotheses. Although the relationship predicted in Hypothesis 3 between greater self-monitoring tendencies and agreement in ratings was supported through a significant correlation with the absolute average profile difference \( (r = -.31, p < .05) \), the relationship with high private self-consciousness specified
in Hypothesis 4 was not significant (see Table 2). The only significant relationship between any of the self-consciousness scores and the other variables was a negative Pearson value between average item difference and the total self-consciousness scale ($r = -0.27, p < .05$). Not surprisingly, the two personality constructs (e.g., self-monitoring and self-consciousness) were not significantly correlated with one another. Because these additional personality measures were only available for one of the HP groups, a comparison by performance group was not possible.

Given the stated concern with the use of difference scores in statistical analyses, additional analyses were conducted on these same data with the categorical approach to operationalizing self-awareness following Atwater and Yammarino’s (1992) and Van Velsor et al.’s (1993) research. As expected, similar results were obtained with this approach as well; that is, there was a significantly greater proportion of HPs among accurate raters or self-aware individuals as compared with either under- or overraters, $X^2(2, N = 60) = 20.24, p < .001$. When mean scores and correlations for the other measures of MSA were computed by agreement group, the results indicated significant convergence among the different methods. Table 3 provides the details of these additional analyses.

Discussion

The purpose of this study was to replicate and extend existing research on the nature and possible moderators of managerial self-awareness, operationalized as agreement between self and others’ behavioral ratings, and its link to managerial performance. The results obtained met these objectives. With data collected from four independent samples representing three different types of organizations and regardless of the nature of the performance criteria, the results consistently indicated that high-performing managers were able to assess more accurately their own behaviors in the workplace, yielding greater congruence in self-reports versus direct reports’ ratings compared with average performers. This relationship was supported with both item level difference score measures and the comparison of between manager correlations and the agreement categorical approach as well. Thus, it is clear from this research that high-performing managers have a greater level of MSA or, at the very least, are better at assessing and/or rating their own behaviors in the workplace. Although it is impossible given the present data to assert that enhanced degrees of MSA resulted in their greater performance, the pattern of enhanced congruence among the HPs, regardless of the data source, organization, criteria for performance group assignment, and number or content of the items used in the comparison, is consistent with and lends additional support for the conceptualization of MSA. Moreover, these results serve to replicate Atwater and Yammarino’s (1992) findings between accurate estimators and performance ratings rather than Van Velsor et al.’s (1993) research. Given the manner in which performance was operationalized in the latter study, however (i.e., as another set of similarly rated items), this is not surprising. The general tendencies in self-reports’ and direct reports’ ratings in the present research, however, did replicate trends evident in both these and other earlier studies.

Though it can be argued that this degree of improvement in ratings congruence by performance group reflects some other cognitive process (e.g., other-directedness) rather than managerial self-awareness, few other theoretical orientations explain such differences in performance as well as does the conceptualization of MSA. For example, although it is possible that HPs might merely be more adept at estimating how their direct reports would rate them, given the number of direct reports providing ratings and that the comparison scores were computed with averages, it is unlikely that managers would be “second guessing” an average direct report rating on each item. For most individuals, it is probably far easier to assess the relative degree to which they exhibit a certain behavior than to calculate some arbitrary average. Moreover, even if such a self-referenced process were occurring, it would have to be moderated at least to some degree by the manager’s own perceptions of how he or she behaves in the workplace (which reflects the MSA concept as well). Finally, because both performance groups had sizable portions of overraters and underraters, it is unlikely that simple measurement issues such as modesty in self-ratings or inflated direct reports’ ratings for HPs in general are appropriate alternative explanations for the results.

With respect to the appropriateness of using any form of congruence in self–other ratings to measure the construct self-awareness, the significant correlation between managers’ self-scores on the self-monitoring questionnaire (Snyder, 1974, 1987) and the absolute profile difference index support this method of operationalization. Individuals high in the personality trait of self-monitoring tended to have more accurate self-assessments, regardless

---

4 Using this approach, managers were categorized according to the following parameters: Accurate raters were those managers whose average total self scores across all items fell within one-half standard deviation of the mean overall direct report rating for the entire sample plus or minus the actual rating given by their direct reports; underraters were those managers whose self scores were lower than their direct reports’ ratings, plus $0.5 SD$ of the total others’ ratings; overraters were those individuals with total self scores $0.5 SD$ above their direct reports’ overall ratings.
Table 3

<table>
<thead>
<tr>
<th>Group</th>
<th>Absolute average profile difference</th>
<th>Relative item average difference</th>
<th>Self-monitoring</th>
<th>Self-consciousness</th>
<th>Between-manager correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Underestimators</td>
<td>133</td>
<td>22.0</td>
<td>78.8</td>
<td>3.62</td>
<td>0.03</td>
</tr>
<tr>
<td>Accurate rats</td>
<td>200</td>
<td>33.1</td>
<td>67.0</td>
<td>3.91</td>
<td>0.27</td>
</tr>
<tr>
<td>Overestimators</td>
<td>271</td>
<td>44.9</td>
<td>15.1</td>
<td>3.56</td>
<td>0.28</td>
</tr>
<tr>
<td>Total</td>
<td>604</td>
<td>100</td>
<td>77.8</td>
<td>3.95</td>
<td>0.38</td>
</tr>
</tbody>
</table>

Note. HPs = high performers; ARs = average performers. Between-manager correlations were self-report ratings correlated with direct report ratings. All behavior ratings and difference of direction, on an absolute average item level, supporting Hypothesis 3. Although low self-monitors in this study tended to behave in consistent ways across situations, it appears that they are likely to be less focused on these behaviors. On the other hand, people high in self-monitoring tend to focus on the manner in which they express themselves, including their behaviors, in social situations. Thus, the notion of self-monitoring provides some convergence on the construct of MSA vis-à-vis difference score ratings comparisons.

Interestingly enough, however, none of the self-consciousness scores, which represent the second alternative approach to measuring self-awareness, yielded significant correlations with the absolute profile difference measure. The only significant correlation among these scale scores was a negative relationship between the relative average item difference measure and for the total SC score. Given the complexities of interpreting a correlation coefficient that was based on a bipolar measure, a subsequent examination of the scatterplot of these two variables indicated that managers high in total self-consciousness tended to have average difference scores below zero, whereas those low in the total SC score tended to rate themselves higher than did their direct reports. In general, the linear trend in the means and the scatterplot suggest that the SC measure is more reflective of a modesty or evaluative bias than a representation of MSA. Given the conceptual underpinnings of the self-consciousness construct (e.g., Buss, 1980; Fenigstein et al., 1975), this makes sense. Individuals high in overall self-consciousness are focused on the self in terms of internal states, public appearances, and social anxiety. It seems likely that such a degree of self-focus can result in a tendency to underrate oneself, particularly given the evaluative component of the self-consciousness framework. Thus, the self-monitoring construct seems to reflect more of a self-assessment perspective, and the self-consciousness approach appears to have more of a ratings leniency or self-evaluative focus. This might also explain why these two personality measures were not significantly correlated in the present sample. Furthermore, these results provide additional insight into the present debate in the literature regarding which of the two conceptualizations represents a truer assessment of self-focus (e.g., Briggs & Cheek, 1988; Fenigstein, 1987; Miller & Thayer, 1989; Snyder & Gangestad, 1986; Wicklund & Gollwitzer, 1987).

Another variable hypothesized to be related to MSA as measured by congruence concerned the extent to which a manager perceived a specific set of practices to be critical to his or her successful performance in the organization. In general, the results of the simple correlational analyses failed to support Hypothesis 2, that is, the relationship between importance ratings and the absolute profile difference measure was nonsignificant. In fact, the
significant correlation between the relative average item difference score and total ratings of item importance by self suggested the opposite relationship to the one originally proposed. That is, the more important the behaviors being rated, the greater the average level of difference observed between self versus direct reports' ratings. While such a conclusion might be warranted from these results, it is also possible that scale limitations inherent in the use of importance ratings (e.g., ceiling effects or overrating tendencies) may have contributed to these unanticipated findings, particularly given the importance mean value ($M = 4.17$) observed for self-ratings. Additionally, because the correlational approach used to test this hypothesis was conducted at the summary score rather than individual item level of analysis, it is also possible that more complex relationships may have been present but less easily detectable. In any case, more research is needed to clarify the exact nature of the relationship between importance ratings and MSA.

Finally, a few comments are warranted regarding the comparative utility of each of the different approaches to measuring self- and others' ratings agreement. Of the two difference score measures, both of which reflect item level comparisons, the absolute profile difference yielded more significant relationships with the grouping variables than did the relative average item difference score. Although the average item difference provided additional information regarding the direction of the difference obtained, and is generally more intuitively interpretable from a ratings perspective, significant detail is lost due to the averaging process; for example, a series of large overratings and large underratings for the same individual would net out to an accurate average score indicating little difference. The absolute profile difference measure is not susceptible to this issue since all differences are squared before being averaged. In part, this may explain the greater utility of profile difference in the present study. Thus, in general, it is recommended that both the profile difference measure and the item difference measures be used for future item level difference score computations in order to adequately account for the item-by-item ratings profile, the directionality of the ratings obtained, and general level effects among different sets of raters.

**Limitations of the Research**

There are several limitations to this research that need to be addressed. First, although the reliabilities of the items used to generate the comparison measures were high, the content among the different sets (e.g., representing different dimensions of behavior) was not explored. Another limitation concerns the composition and comparative sample sizes of the individuals from whom the data were collected. It is likely that there may have been insufficient power to detect significant differences among certain groups (e.g., between female and male HPs). Additionally, other potential moderators, such as education, training, ethnicity, span of control, frequency of contact with people providing ratings, remain unexamined. The generalizability of the findings obtained is also a concern. Although managers' data from three different organizations were involved in the analyses, all of these individuals were at the very least adequate enough performers to be included in some form of developmental training program. How these average individuals actually compare with those who perform poorly is unknown. Moreover, given the imperfect nature of the performance criteria, it is possible that other factors and concerns were involved in the HP identification process. Because HPs did yield significantly higher congruence in ratings regardless of the selection criteria, however, the results should have some degree of external validity. Finally, although the differences obtained were consistently significant across a variety of analyses, it is important to remember that the present research is descriptive rather than causal. Thus, how the HPs actually became more self-aware (e.g., an inherent ability or a skill developed through social interactions) or whether the presence of MSA actually resulted in or contributed to their achieving a high-performance designation are unanswerable questions, given the present research design. Although it is plausible that MSA played at least some part in these individuals' successful performance, this relationship cannot be stated with absolute confidence.

**Suggestions for Future Research**

Future research is needed to test the assumption that higher levels of MSA eventually lead to increased performance in the workplace. The presence of a causal linkage can be examined through the use of a longitudinal design, whereby MSA is assessed initially and then tracked for several years and across various interventions (e.g., multitrait feedback) to determine changes in MSA and career success. Such research could also be used to identify the degree to which MSA is an inherent trait versus a skill set, as well as to help clarify conflicting results regarding the stability of self-other discrepancies over time (e.g., London & Wohlers, 1991; Nilsen & Campbell, 1993). Future studies could also explore the extent to which managers differ in assessing themselves in specific content areas (e.g., task vs. people behaviors, work group climate, leadership skills, or customer service), and with respect to different constituents in the workplace (e.g., peers, subordinates, supervisors, and customers). Related to this issue is the need to link MSA with other types of self-awareness and situations in which it might be present such as whether self-awareness in the workplace translates
to self-awareness at home with family and friends. These types of studies contribute to establishing a theoretical foundation for MSA, as would research exploring self-awareness in relation to various other perspectives such as information processing or other-directedness. Finally, future work could also be directed at determining alternative and more direct methods for assessing MSA beyond the ratings congruence paradigm.

References


(Appendix follows on next page)
Sample Behavioral Items Used for Rating Purposes

1. You pay close attention and seek to understand direct reports when interacting.
2. You delegate authority to enable direct reports to make decisions and take action in a timely manner.
3. You provide clarity about individuals' roles and responsibilities.
4. You emphasize a team approach in accomplishing work.
5. You give direct reports the opportunity to learn things and try different activities on the job.
6. You coach and train people in new assignments.
7. You give your direct reports an opportunity to express openly their disagreements or to voice objections to your proposed actions/decisions.
8. You help direct reports to attain rewards and recognition.
9. You try to make the best use of your direct reports' skills and abilities when making assignments.
10. You establish trust and mutual respect in relating with your direct reports.
11. You solicit appropriate information from your direct reports—facts, opinions and concerns about their work.
12. You encourage your direct reports to submit new ideas and suggestions for improvement.
13. You give your direct reports full credit for their ideas.
14. You explain to your direct reports what is required to achieve work objectives.
15. You work with direct reports to establish mutually acceptable performance objectives and requirements.
16. You take time to provide direct reports with informal feedback on their performance.

Received May 10, 1996
Revision received September 25, 1996
Accepted September 26, 1996

New Editor Appointed for *Contemporary Psychology*: 1999-2004

The Publications and Communications Board of the American Psychological Association announces the appointment of Robert J. Sternberg (Yale University) as editor of *Contemporary Psychology*, for a 6-year term beginning in 1999. The current editor, John H. Harvey (University of Iowa), will continue as editor through 1998.

All reviews are written by invitation only, and neither the current editor nor the incoming editor receives books directly from publishers for consideration. Publishers should continue to send two copies of books for consideration, along with any notices of publication, to PsycINFO Services Department, APA, Attn: *Contemporary Psychology* Processing, P.O. Box 91700, Washington, DC 20090-1700 or (for UPS shipments) 750 First Street, NE, Washington, DC 20002-4242.