SINCE WHEN IS NO NEWS GOOD NEWS?
The Relationship Between Performance
And Response Rates in Multirater Feedback

ALLAN H. CHURCH
W. Warner Burke Associates, Inc.

STEVEN G. ROGELBERG
Bowling Green State University

JANINE WACLAWSKI
W. Warner Burke Associates, Inc.

Despite the popularity of multirater feedback in practice and research, few studies have examined the issue of response rates in these efforts. This study explored the relationship between performance, vis-à-vis a measure of service quality, and feedback response rates in a large-scale developmental multirater feedback initiative using data from 538 senior service providers, 4,446 coworkers and supervisors, and 1,617 clients. The number of rater responses that the focal individual received was largely unrelated to his or her performance level as rated by his or her clients. More specifically, less than 2% of the variance in response rates was explained by the focal individual's performance. Data representativeness and feedback acceptance implications are discussed.

As Huron (1993) predicted, data-driven measures have indeed become the vital signs of organizational and individual effectiveness. These days, it seems that corporations are interested in measuring everything and everyone on some aspect of their organizational existence. Whether it is personality assessments for determining which individuals will make the most effective leaders (e.g., Hogan, Curphy, & Hogan, 1994; McCaulley, 1990), surveys for large-scale organization development and change efforts (e.g., Church, Margiloff, & Coruzzi, 1995; Kraut, 1996; Nadler, 1977; Waclawski, 1996), collecting individual feedback for initiating specific targeted behavioral change (e.g., Burke, Richley, & DeAngelis, 1985; Church, Javitch, & Burke, 1995; Dalton, 1996; London & Beatty, 1993), or establishing measures of training effectiveness over

The research described here was first presented in Dallas, Texas at the 13th Annual Conference of the Society for Industrial and Organizational Psychology. The authors thank Jim Smither and two anonymous reviewers for their helpful comments and suggestions.

Correspondence and requests for reprints should be addressed to Allan H. Church, Principal, W. Warner Burke Associates Inc., 201 Wolfs Lane, Pelham, NY 10803; Allanhc@aol.com.
time (e.g., Holton, Bates, Seyler, & Carvalho, 1997), part of contemporary organizational experience is to be involved in some form of applied measurement related initiative. Given these trends, it should come as no surprise that the process of providing leaders and managers with individualized feedback from coworkers, supervisors and clients on their behavioral tendencies has become a popular organizational consulting tool (Bracken, 1994; Church, 1995; Hazucha, Hezlett, & Schneider 1993; London & Smither, 1995; Tornow, 1993; Waldman, Atwater, & Antonioni, 1998). Although the terminology varies from application to application (including such terms as multirater, multisource, full-circle, and 360-degree), the fundamental assumptions about the purpose and efficacy of providing feedback are generally consistent across interventions.

Simply stated, individualized feedback, regardless of the behavioral content on which it is based or the purpose to which it is directed (i.e., development or appraisal), is intended to provide a vehicle for enhancing self-awareness with respect to one’s strengths and areas for improvement. This, in turn, when combined with a host of possible interventions, is expected to result in positive behavior change and ultimately improved individual performance (Antonioni, 1996; Church & Bracken, 1997; Church & Waclawski, 1998a; London, 1997; London & Smither, 1995; Tornow, 1993; Yammarino & Atwater, 1997). Although recent reviews of the feedback-based literature have raised significant questions regarding the extent to which positive gains are actually realized (e.g., Kluger & DeNisi, 1996), our collective professional and personal belief in the fundamental “goodness” of feedback as a tool for change continues to drive an entire industry of developmental and assessment related interventions worth billions of consulting dollars.

Although multirater feedback applications are indeed a popular research and conference topic, one area that has been neglected is the issue of participation—that is, feedback-provider response rates in these types of efforts (Westerman & Rosse, 1997). This is somewhat surprising considering the amount of effort that has been directed at understanding mail survey response rates in general (Aiken, 1988; Church, 1993; Dillman, 1978; Rogelberg & Luong, 1998), and employee attitude survey response rates in particular (Church & Waclawski, 1998b; Macey, 1996; Rogelberg, Luong, Sederburg, & Cristol, in press; Viswesvaran, Barrick, & Ones, 1993). Given their widespread usage, why haven’t response rates in these individualized “surveys for one” (Waclawski, 1998) been explored? Certainly 100% participation rates are not achieved. In fact, anecdotal reports estimate response rates ranging from 40% to 90% (Sederburg & Rogelberg, 1998)—and even lower (e.g., 29% to 46%) in university settings (Westerman & Rosse, 1997).
Another reason for exploring this response rate issue is related to practitioner effectiveness. As a practitioner or coach charged with helping focal individuals interpret and use their data to create behavioral changes, some knowledge of response rate relationships may increase professional credibility, and perhaps even the perceived validity of the results collected, in the eyes of the feedback recipient, in that ratee concerns of data representativeness can be addressed. Thus, we need to have a better understanding of whether certain variables or personal characteristics (e.g., gender, age, prior performance) impact the number of responses received. The central question, then, is do focal individuals (e.g., executives, managers, service providers, staff professionals, etc.) who receive more or less feedback differ in significant ways from each other?

One variable that might help explain response and nonresponse tendencies is individual performance. Although performance at the managerial level and the quality of service provided have been extensively examined as potential correlates of ratings levels and congruence among multiple sources (e.g., Atwater & Yammarino, 1992; Church, 1997a; 1997b; Furnham & Stringfield, 1994, 1998; Van Velsor, Taylor, & Leslie, 1993; Wohlers & London, 1989), the relationship between individual performance and rater response rates remains to date unexplored. The central purpose of this study is to answer the question, “To what extent does performance by the focal individual (the person receiving feedback) relate to rater response rates?” Given the considerable emphasis placed on client ratings of performance in professional service firms and consulting organizations (e.g., Church et al., 1995; Maister, 1993), the service setting represents an ideal context for exploring this research question vis-à-vis individual service quality ratings.

As is often the case in applied research, the expected relationship between these two variables can be hypothesized as occurring in either direction. For example, one could argue that coworkers of higher-performing service providers would be more likely to participate by completing a rating form in order to further help increase the focal individual’s performance (i.e., they may assume that the focal manager is committed to delivering high quality client service), or perhaps due to greater levels of positive affect and friendship. From this perspective, lower performers might be expected to receive fewer responses due to a mindset such as “this individual is too poor for my feedback to make a significant impact.”

Conversely, it could also be defensibly argued that coworkers of lower-performing service providers would be more likely to feel the need to respond to a feedback effort for a variety of reasons, including interest in helping to improve the focal individual’s behavior (and thereby en-
hance his or her own workplace experience and/or performance) or the simple chance to provide criticism of that individual. From this perspective, higher performers might be expected to receive fewer responses due to a mindset such as “this individual is very good and does not need my feedback.”

Of course, there is also the distinct possibility that no relationship exists between service performance and responses received. Given the widespread usage of survey and multirater feedback methodologies to assess all sorts of organizational issues and initiatives, one could argue that perceptions of survey overload (Rogelberg, 1998) and ratings utility may have more to do with response rates than would the performance of the person actually being rated. Similarly, trust in the process itself or perceived intentions and commitment on the part of senior management may also be important factors.

**Objectives of This Study**

Given the significant problems associated with nonresponse in any type of diagnostic data-collection effort (including multirater feedback), this study was conducted to explore the relationship between service performance and the number of responses received. As is common in ratings research, we were also interested in examining any possible interactions and/or main effects on response rates with respect to individual-level variables such as gender, region, and age. Findings from this study should prove useful for raters and ratees, practitioners delivering feedback in various settings, and those individuals in administrative positions who might be designing or implementing such systems in organizations, because by examining rater response rates insights into data representativeness are revealed. Consequently, questions, concerns, and suspicions about response rates can be addressed to some extent.

**Method**

**Participants**

The data used in the present study were collected over a 3-year period from 538 senior service providers (consultants who were the focal individuals in the rating process), 2,076 of their direct reports, 1,873 peers, 497 supervisors, and 1,617 clients. The focal individuals were executives in a global consulting and professional services organization. Members of this organization typically provide their clients with a variety of specialty consulting services across a wide range of industries. Based on clients’ demographics in the present sample, the industries rep-
resented included manufacturing (30.4%), financial services (15.5%),
energy and telecommunications (12.8%), consumer products (11.3%),
commercial services (10.0%), and healthcare (6.9%). With respect to
the global representation of service providers in this sample, 50.4%
were from the United States or Latin America (with the majority com-
ing from North America), 21.2% represented Asia/Pacific, and 28.4%
were from the area known as EMEA, which was comprised of Europe,
Middle East, India and Africa.¹ The mean age of participants was 42.7
($SD = 5.26$). Consultants were primarily male—474 (91.2%)—with only
46 (8.8%) females. Eighteen individuals chose not to provide gender in-
formation on the response form.

All the senior service providers included in the present study were
participants in a 1-week residential feedback-based development pro-
gram designed to enhance their consulting and client management and
delivery skills. These programs were held in a variety of offsite manage-
ment training facilities throughout the world. Prior to their attendance,
behavioral and performance ratings were obtained from these individu-
als’ coworkers (including direct reports, peers, and supervisors), as well
as from their clients for use during the developmental program. The
focal individuals were instructed to personally select and distribute all
the feedback questionnaires themselves to a mixed group of five of their
direct reports, five of their peers, four of their clients, and their immedi-
ate supervisor. This commonly used form of distribution was intended to
increase rater motivation (through personal contact) and feedback ac-
cceptance on the part of the focal individual. Any and all unused forms
were to be returned to the external consulting firm.

Given that these individuals represent the most senior partnership
ranks of their firm with over 70,000 employees, the probability of there
being a limited number of direct reports, peers, or clients from which to
choose potential raters was extremely low. Representatives from the in-
ternal corporate training function as well as an external consulting firm
were in frequent contact with focal individuals to encourage complete
sets of responses for each program participant. Because internal funding
for program attendance required some level of complete responses, the
focal managers were sufficiently motivated to follow up with their con-
stituents. Moreover, because they personally knew exactly to whom the
forms were distributed, it was a relatively simple task to remind people to
complete and return the forms. During the tracking procedure, no situa-
tions were reported among this group of focal individuals where they did
not have enough direct reports or peers to whom they could distribute

¹These three categories represented the business unit groups for this organization and
could not be reconfigured.
five forms (although we have experienced this situation with other types of samples, such as individuals in nonmanagerial roles, those in lower management positions, or those in smaller organizations). Finally, no sets of blank forms were returned as having not been distributed.

The data were delivered to participants during the program in the form of individualized reports. During the week, in team and one-on-one sessions, professionally trained staff provided individual coaching and counseling to help participants understand, and make effective use of the feedback. As with many professional service firms, these senior service providers had little to no experience with internal assessments of their behavior or performance (e.g., up to this time the method of performance appraisal for these individuals consisted primarily of financial performance indicators—e.g., billable hours with clients). Thus, their involvement in the present rating process, despite its developmental orientation, represented a new experience for them.

All individuals involved in the process were informed of the solely developmental nature of the feedback and were instructed to be as candid in their ratings as possible. They were also informed that no one in the firm other than the focal individual and the external executive coach would receive a copy of the feedback results. Moreover, additional efforts to protect confidentiality, ensure candor, and enhance rater motivation were implemented, including frequent communication and follow-up efforts, guarantees of protection by not reporting data when fewer than two direct reports, peers or clients responded, and the use of an external and independent consulting firm to collect the assessment data (e.g., Church, 1995; Harris, 1994; Ostroff, 1993).

Measures

Number of returns. The primary dependent measure of interest—the number of completed feedback returns—was computed for each focal individual. This was done at a summary level (i.e., total number of questionnaires returned from all coworkers) and for each of the rater groups (i.e., direct reports, peers, and supervisor). The total number of coworker responses received per service provider was 8.25 (SD = 1.90) and ranged, as would be expected by the distribution limits, from 0 to 11. There were 4,446 multirater feedback returns overall. The average number of responses for each rater group was 3.86 for direct reports (SD = 1.04), 3.47 for peers (SD = 1.11), and 0.92 for supervisors (SD = .27). Given the distribution limits and consistent with the formal instructions provided, repeated follow-up contact, and anecdotal evidence that focal individuals did in fact distribute all their questionnaires, these averages represent general response rates of 77.2%, 69.6%,
and 92.4% for each rater group, respectively. Although not included as part of the dependent variable, in order to prevent contamination with the predictor measure, return rates were also calculated for clients. In total, 1,617 client forms were returned, which represented an average of 3.00 (SD = .86) per focal manager and a response rate of 75.1%.

Although these ratios can be treated as response rates, it is important to note that technically they do not reflect the total ratio of direct reports, peers or clients capable of responding for a given individual (e.g., someone receiving five completed responses would have a return rate of 100% even though he or she may well have had 40 individuals in his or her span of control).

**Performance.** Service provider performance was measured using 9 outcome ratings regarding the quality of services provided. These ratings were provided by actual clients of the focal individuals. The organizational level of the primary clients providing ratings of these service providers was quite high and distributed as follows: 30.2% CEOs, 61.9% CFOs and 7.9% corporate tax directors. Clients completed a custom-designed instrument, which was developed from a combination of interview data, prior feedback-based consulting work with similar organizations, archival research, and literature in the field regarding effective service performance (e.g., Maister, 1993; Schneider, 1980). Clients were asked to rate the extent to which the senior service provider achieved each specific outcome, using a scale from 1 to 5 (where 1 = *to a very small extent* and 5 = *to a very great extent*). These outcomes assessed several different aspects of service quality including reliability, innovativeness, and professionalism (e.g., "meets commitments," "demonstrates strong professional skills," "shows creativity in providing client solutions," and "provides high quality service"). The interitem reliability for the nine outcome measures as rated by clients in this sample was .89.

It is important to note that although each rater group (e.g., direct reports, peers, supervisors, and clients) offered a unique perspective on the behavior of the focal individuals, client ratings were chosen as the service performance criterion measure in this study. The individuals participating in the development program were external consultants, therefore clients represent a very important (perhaps the most important) source for evaluating performance. Moreover, for this group, client ratings of effectiveness are directly related to customer satisfaction, billable hours, and repeat business, which ultimately lead to increased sales and profits (Bitner, Booms, & Mohr, 1994; Davidow & Uttal, 1990; Maister, 1993). Further, because no internal formal appraisal measures existed, and the supervisors of these consultants were, in actuality, one level higher partners themselves in a partnership-driven structure, client service satisfac-
tion and therefore client ratings represented the best and most meaningful measure of performance for this population.

The remaining individual level variables—that is, gender, region, and age—were obtained from the focal individuals themselves as part of a self-report questionnaire.

Results

Table 1 provides an overview of the descriptive statistics and correlations among the variables. As is typical in ratings research (e.g., Atwater, Ostroff, Yammarino, & Fleenor, 1998; Church, 1997b; Furnham & Stringfield, 1998; Harris & Schaubroeck, 1988; London & Wohlers, 1991), correlations among performance ratings from each of the five observer perspectives were significant and positive but only moderate in strength. The strongest set of relationships were among the internal raters—for example, direct reports with peers ($r = .40, p < .01$), direct reports with supervisors ($r = .39, p < .01$), and peers with supervisors ($r = .37, p < .01$). Correlations with client ratings of service performance were slightly lower, but still positive and significant. The strongest relationship here was with direct report ratings ($r = .25, p < .01$). Self-other correlations for this sample were generally low.

There were small but significant negative relationships between client service performance ratings and the total number of responses received from total internal raters ($r = -.13, p < .01$), returns from direct reports only ($r = -.10, p < .05$), and returns from peers only ($r = -.12, p < .01$). This finding demonstrates that the more effective the service provider was, the fewer returns he or she received. Conversely stated, less effective service providers received a greater number of returns than more effective service providers.

Relatively high positive correlations were present among the number of returns received by each observer group. Particularly strong for direct reports and peers ($r = .43, p < .01$), it was apparent that the more one group responded, the more others did as well. Finally, although no effects were present for gender, there was a small effect for age such that older service providers were less likely to receive feedback from their supervisor ($r = -.09, p < .05$).

Multiple regression was employed to look for interactions and main effects for each of the predictors (service performance, gender, age, and region) with total number of responses received as the dependent measure. Of the 538 individuals in the dataset, only 506 had complete demographic responses and were included in this set of analyses. Although the overall model was significant $F(14,491) = 1.92, p < .05$ with an $R^2 = .05$, none of the higher order interaction effects (e.g., age by performance
<table>
<thead>
<tr>
<th>Variable</th>
<th>M</th>
<th>SD</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Client performance ratings</td>
<td>3.90</td>
<td>0.43</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Self ratings</td>
<td>3.86</td>
<td>0.40</td>
<td>.16**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Direct report ratings</td>
<td>3.85</td>
<td>0.42</td>
<td>.25**</td>
<td>.14*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Peer ratings</td>
<td>3.87</td>
<td>0.37</td>
<td>.16**</td>
<td>.14*</td>
<td>.40**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Supervisor ratings</td>
<td>3.86</td>
<td>0.53</td>
<td>.19**</td>
<td>.22**</td>
<td>.39**</td>
<td>.37**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Number of clients responding</td>
<td>3.00</td>
<td>0.86</td>
<td>-.04</td>
<td>.00</td>
<td>-.01</td>
<td>-.04</td>
<td>-.02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Number of direct reports responding</td>
<td>3.86</td>
<td>1.04</td>
<td>-.10*</td>
<td>.09*</td>
<td>-.03</td>
<td>.00</td>
<td>.02</td>
<td>.31**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Number of peers responding</td>
<td>3.47</td>
<td>1.11</td>
<td>-.12**</td>
<td>.01</td>
<td>-.01</td>
<td>-.01</td>
<td>.01</td>
<td>.31**</td>
<td>.43**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Number of supervisors responding</td>
<td>0.92</td>
<td>0.27</td>
<td>-.01</td>
<td>.02</td>
<td>.03</td>
<td>-.03</td>
<td>NA</td>
<td>.16**</td>
<td>.24**</td>
<td>.19**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Total number of internal responses</td>
<td>8.25</td>
<td>1.90</td>
<td>-.13**</td>
<td>.06</td>
<td>-.02</td>
<td>-.01</td>
<td>.02</td>
<td>.37**</td>
<td>.83**</td>
<td>.84**</td>
<td>.39**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. Participant gender</td>
<td>1.09</td>
<td>0.28</td>
<td>-.05</td>
<td>.00</td>
<td>.04</td>
<td>-.08</td>
<td>-.04</td>
<td>.05</td>
<td>.03</td>
<td>.04</td>
<td>.03</td>
<td>.05</td>
<td></td>
</tr>
<tr>
<td>12. Participant age</td>
<td>42.69</td>
<td>5.26</td>
<td>-.04</td>
<td>-.01</td>
<td>-.05</td>
<td>-.07</td>
<td>-.14*</td>
<td>-.05</td>
<td>-.02</td>
<td>-.09*</td>
<td>-.11*</td>
<td>-.05</td>
<td>-.16**</td>
</tr>
</tbody>
</table>

\*Gender was coded 1 for male, 2 for female.

\*p < .05  **p < .01
interaction) were significant. Next, a more simple, main-effects-only regression model was examined. The results of this model indicated a relationship between service performance and total responses received $F(5,500) = 3.40, p < .01$ with a $R^2 = .03$. More specifically, the number of feedback responses received by the focal individual was significantly predicted by his or her service performance ($B = -.13, p < .01$). Although neither gender nor age contributed significantly to the regression model as main effects, there was evidence of a main effect for one of the region contrast codes on number of returns ($B = .11, p < .05$). More specifically, focal managers in the United States/South America region received significantly more responses than individuals in the EMEA region (no mean differences in client performance ratings were found by region; $p < .05$).

Next, three additional regression models were examined. Each regression model used the same set of predictors specified above; however, the number of direct report returns, peer returns, and supervisor returns served as the dependent measures respectively. In general, the results of these were relatively consistent with the results for the overall returns model. Service performance proved to have a significant Beta with respect to number of direct report ratings received ($B = -.11, p < .05$) as well as number of peer ratings received ($B = -.10, p < .05$). Service performance was not a significant predictor, however, of whether or not supervisors chose to respond. Although gender did not contribute significantly to any of the models, service provider age did emerge as a significant predictor for supervisor responses ($B = -.11, p < .05$), whereas the region effect noted above was evident only in the equation involving the number of peer ratings returned ($B = .14, p < .01$).

**Discussion**

The objective of this study was to determine whether a relationship exists between the number of responses received in a multi-rater assessment process and the performance of the focal individual. Although the results indicated a small statistically significant relationship between performance and response rates, practically speaking, the number of responses that the focal individual received was largely unrelated to his or her performance level. More specifically, less than 2% of the variance in response rates was explained by the focal individual's performance as indicated by his or her clients.

A few explanations exist that may account for the lack of a meaningful relationship between focal individual performance and response rates. One possible explanation is the presence of an overarching fundamental belief in feedback. That is, raters may have felt, given the de-
velopmental purpose of the program, that all ratees, regardless of their present performance levels, deserved and should receive developmental feedback. In addition, or alternatively, variables other than performance may be more salient than performance in understanding response rates. Studies of response rates in other survey research (e.g., Church & Waclawski, 1998b; Kraut, 1996; Rogelberg & Luong, 1998) suggest that accessibility, interest, and general attitudes toward compliance may also play a strong role in affecting response rates. Other factors might include the purpose of the feedback effort itself (development vs. pay), linkages with existing strategic initiatives in the larger organization, trust in senior management, prior experience with multirater systems, cultural and political contexts, rater anonymity, and rater incentives (Antonioni, 1996; Bracken, 1994; Church 1995; Dalessio, 1998; Funderberg & Levy, 1997; Harris, 1994; Judge & Ferris, 1993; Landy & Farr, 1980; Maurer & Tarulli, 1996; Ostroff, 1993). Westerman and Rosse (1997), for example, found that rater knowledge, time required for completion, perceived relevance of the questions asked, and the developmental intent of the process all contributed significantly to the decision to participate in a multirater feedback project in a university setting.

Although future research is needed to study these additional personal, situational, and contextual variables in order to enhance our understanding of the conditions that affect response rates in multirater feedback applications, our findings are potentially useful to a variety of individuals with roles in the feedback process (e.g., feedback recipient, practitioner or coach delivering the results, and system administrator). Namely, feedback recipients can have increased confidence that the number of responses they received is not likely to have been strongly affected by their performance level. This may offer some comfort to focal individuals whose coworker response rates are less complete than they might have hoped for. In addition, a feedback recipient who receives a relatively small number of responses that yield lower ratings cannot so easily argue that his or her ratings would have been higher (i.e., more favorable) if only more people had responded. Overall, if our findings can be replicated in other settings, it would suggest that response rates do not create troubling biases (as is frequently assumed in other data collection efforts) for focal individuals in the average multirater feedback process.

Other Relationships

A small age effect on response rates was found. More specifically, regardless of the level of performance, older consultants tended to receive fewer responses from their peers and were less likely to receive a
supervisor rating. Although this effect was quite small and may not be practically meaningful, the most compelling explanation for such a finding is that there may be less of a perceived developmental opportunity for older consultants. Although entirely speculative, such disinterest in providing feedback could be due to one of several factors: (a) Older consultants are presumably more experienced and therefore need less feedback; (b) older consultants may be perceived to be less likely to change based on feedback (e.g., you can't teach an old dog new tricks); (c) older consultants are closer to retirement, therefore, it is better to invest more feedback on service providers who are "on their way up" as opposed to those who are "on their way out"; or (d) supervisors simply feel less comfortable giving older consultants feedback. In any event, this trend is in need of further study.

The findings with respect to region were interesting, but not particularly surprising. In general and regardless of service performance levels, focal individuals in the EMEIA region received significantly fewer responses from their peers. Again, although the observed effect was quite small, we speculate that feedback may be less universally culturally appropriate or accepted among people in these areas. In cultures where there is a strong sense of hierarchy and power-distance (Hofstede, 1980), feedback from others, especially direct report and peers, may be frowned upon and even actively resisted. Alternatively, the issue may simply be instrumentation—that is, individuals in these regions may have less exposure to and familiarity with formal rating forms. We have often heard from participants in coaching sessions that individuals in the EMEIA regions are less comfortable with formal rating forms in general. Regardless of the reason, these trends suggest that feedback interventions involving individuals from EMEIA and similar regions should be carefully designed to consider possible lower response rate tendencies. For example, greater distribution of forms may be needed to yield similar amounts of completed responses.

Our final and perhaps somewhat unanticipated finding was that response rates among the different internal rating groups (e.g., direct reports, peers, and supervisors) were significantly correlated with each other. At least three possible explanations come to mind for this effect. First, certain personal attributes of the focal individual (i.e., ones not examined here such as introversion–extraversion) might have led internal raters to be less likely to respond in general. Second, certain behaviors (or messages) on the part of the focal individual when distributing and/or following up with his or her raters may have lead to consistency in response rates. Finally, other ratee characteristics such as time in current position, cultural differences (such as those mentioned above), organi-
zational location, and so forth, might explain this phenomenon as well. A precise explanation, however, awaits additional research.

**Limitations**

One limitation of the present study was the nature of the sample. Aside from the low proportion of women, the data were collected from the senior population of a single global professional services firm. Moreover, the majority of these individuals originated primarily from the Americas, and North America in particular. Thus, the generalizability of these findings may be limited to this type of population and occupational setting. This limitation, however, is also arguably a strength in that it somewhat controls for a number of extraneous contextual variables that could also impact response patterns if different organizations were examined simultaneously (e.g., prior experience with multirater assessment instruments, existing performance appraisal process, recent mergers, acquisitions, or other large-scale structural or personnel changes, as well as norms regarding giving and receiving feedback). Future research needs to be conducted across levels of management, service lines, industries, various regions of the world, and in more traditional organizational settings before broader conclusions regarding the relationships observed here can be drawn.

Another limitation was our inability to explore the effect of different distribution methods (e.g., self-selected vs. externally selected) on feedback participation rates. All focal individuals in the present sample selected and distributed materials to their direct reports, peers and clients. Although this type of distribution method is common among practitioners (e.g., Dalessio, 1998; Sederburg & Rogelberg, 1998), it is impossible to know whether decisions in the self-selection process (e.g., in criteria for inclusion, the way in which observers were asked to participate, attempts to follow-up by focal managers, messages communicated regarding the importance of the feedback effort, etc.) affected response rates.

Moreover, although unlikely, the possibility exists that some participants may have selected fewer individuals to provide them with feedback. Such a situation would raise doubt as to the appropriateness of response rates being operationalized as the number of responses received. There are several factors that argue against this possibility. First, the span of control and level of management at which these focal individuals operated in the organization resulted in an abundance of direct reports and peers from which to choose five for rating purposes. Second, considerable vigilance was maintained in conducting follow-up efforts by internal and external administrative staff wherein the importance of
having complete data from each perspective for developmental coaching purposes was emphasized. Third, the fact that each department or business unit was required to pay a substantial sum to the internal training function (e.g., hotel, travel, materials, coaching staff for each senior partner) created a climate where attending one of these programs and collecting the appropriate feedback data were taken relatively seriously internally.

It should also be noted that the complete range of feedback responses was restricted by the number of forms distributed (i.e., participants could only distribute a maximum of five forms to direct reports, five to their peers, and four to their clients). A better method from a research perspective would have been to provide participants with an unlimited number of feedback forms to distribute (e.g., to all of their direct reports). Assuming that the total number of available raters in each observer pool were calculable for each focal individual such a process may have allowed for greater numbers variability in response rates (i.e., percentage of distributed forms that are subsequently returned) and therefore the possibility of stronger effect sizes. Unfortunately, this type of approach is often impractical in practice from both an administrative and follow-up perspective (Dalessio, 1998).

Finally, client response rates to the feedback effort were subject to some degree of nonresponse as well (24.9%). Client performance assessments of the focal individuals did not vary significantly with their own return rates. Client response rates were, however, correlated with direct report, peer and supervisor return rates, which may suggest the presence of other potential predictor variables (as described above) with respect to response rates.

Summary

In summary, our findings based on a sample of over 500 senior service providers suggest that the number of responses that the focal individual received was largely unrelated to his or her performance level. Although this relationship (or lack thereof) needs to be examined in other settings before it can be considered stable, our findings should be useful to practitioners who often have to deal with and communicate to ratees the meaning of obtained response rates and missing data in practice.

REFERENCES


Rogelberg SG. (1998, April). Surveys and more surveys: Addressing and dealing with over-surveying. Rogelberg SG (Chair). Symposium conducted at the Annual Conference of the Society for Industrial and Organizational Psychology, Dallas, TX.


