

DOES PERFORMANCE IMPROVE FOLLOWING MULTISOURCE FEEDBACK? A THEORETICAL MODEL, META-ANALYSIS, AND REVIEW OF EMPIRICAL FINDINGS

JAMES W. SMITHER
La Salle University

MANUEL LONDON
State University of New York at Stony Brook

RICHARD R. REILLY
Stevens Institute of Technology

We review evidence showing that multisource feedback ratings are related to other measures of leadership effectiveness and that different rater sources conceptualize performance in a similar manner. We then describe a meta-analysis of 24 longitudinal studies showing that improvement in direct report, peer, and supervisor ratings over time is generally small. We present a theoretical framework and review empirical evidence suggesting performance improvement should be more likely for some feedback recipients than others. Specifically, improvement is most likely to occur when feedback indicates that change is necessary, recipients have a positive feedback orientation, perceive a need to change their behavior, react positively to the feedback, believe change is feasible, set appropriate goals to regulate their behavior, and take actions that lead to skill and performance improvement.

It has been nearly 10 years since London and Smither (1995) evaluated the state of multisource feedback practice and offered theory-based propositions for understanding how people process and use the feedback. This article assesses progress in the field, especially focusing on the extent to which feedback recipients improve their performance after receiving multisource feedback. We argue that practitioners should not expect large, widespread performance improvement after employees receive multisource feedback. Instead, we present a theoretical model that suggests some feedback recipients should be more likely to improve than others. First, we review empirical evidence concerning the validity of multisource feedback. This is important because it would make little sense to focus

The authors thank Murray Barrick (Associate Editor), three anonymous reviewers, and Pete Dominick for their helpful comments concerning an earlier draft of this paper.

Correspondence and requests for reprints should be addressed to James W. Smither, La Salle University, 12 Cascades Terrace, Branchburg, NJ 08876; amysean@aol.com.

our attention on multisource feedback unless there is convincing evidence that multisource ratings are related to other indices of performance and effectiveness. Second, we present a meta-analysis of longitudinal studies of multisource feedback in order to examine the amount of improvement (i.e., in multisource feedback ratings) that occurs following receipt of multisource feedback. Third, we present a preliminary theoretical model and review evidence concerning factors that might enhance the likelihood of performance improvement following receipt of multisource feedback.

Validity of Multisource Feedback

Two lines of evidence speak to the validity of multisource ratings. A number of studies have shown that multisource ratings are related to a variety of performance measures and hence provide evidence concerning the concurrent validity of multisource ratings. Other studies have examined whether different rater sources (e.g., peers, direct reports) conceptualize performance in a similar manner.

Concurrent Validity of Multisource Ratings

Evidence for the validity of multisource ratings is broadly considered here and includes research comparing multisource ratings with assessment center performance, annual appraisals, objective performance data, and the satisfaction and turnover intentions of subordinates.

Three studies (Atkins & Wood, 2002; Helland, Hoffman, & Smith, 2003; Warech, Smither, Reilly, Millsap, & Reilly, 1998) have shown a positive relationship between multisource ratings and assessment center performance. In a study of upward feedback, Smither and Walker (2004) found that ratees who received more favorable narrative comments from their direct reports also received a more favorable annual review from their supervisors. Ostroff, Atwater, and Feinberg (2004) reported that subordinate and peer ratings were positively related to appraisals of overall performance made by the feedback recipients' supervisors. Subordinate and peer ratings also displayed a small (albeit statistically significant) positive relationship with feedback recipients' compensation and organizational level. Conway, Lombardo, and Sanders (2001), conducting a meta-analysis of correlations, found that both direct report and peer ratings accounted for significant variation in objective measures (e.g., production, profit) over and above other sources. Erickson and Allen (2003) found that multisource feedback ratings were positively related to retail store outcomes, such as revenue, gross margin, and sales of accessories and service contracts in a sample of 573 store managers. Smither and Walker (2001) found that upward feedback ratings of bank branch managers correlated

significantly with branch measures of customer loyalty (e.g., a customer's stated intentions to remain a customer). Church (2000) found that managers who received more favorable multisource feedback had lower turnover and higher service quality in their workgroups. Finally, Atwater, Brett, and Ryan (2004) found that increases in subordinate ratings over a 1-year period were related to increases in subordinate satisfaction and engagement with work and decreases in subordinates' turnover intentions.

Do Different Rater Sources Conceptualize Performance in a Similar Manner?

Cheung (1999) has noted the potential problems that could arise from conceptual disagreement that might occur between rater sources (i.e., concerning the number of factors measured by multisource survey items or the items associated with each factor). Several recent studies have examined this question. Mount, Judge, Scullen, Sytsma, and Hezlett (1998) found that method effects were more strongly associated with individual raters than with the rater's role, with the possible exception that supervisor ratings may constitute a separate method factor because they are more likely to share a common frame of reference based on their experience and training. Their study also showed trait effects (human relations, technical, and administrative performance) suggesting that raters have similar views on the distinctiveness of traits. Fecteau and Craig (2001) and Maurer, Raju, and Collins (1998), using confirmatory factor analysis, examined covariance matrices across groups and found that the number of factors and the item loadings on those factors were invariant across rater sources. That is, raters in different roles shared a common conceptualization of the performance dimensions. Scullen, Mount, and Judge (2003) found that raters from all perspectives attended to a similar set of core performance factors and that subordinate ratings were more highly correlated across dimensions than ratings from other perspectives. They speculated that subordinates might be less experienced in providing ratings compared to supervisors, and their ratings may become more differentiated as they gain more experience. Birkeland, Borman, and Brannick (2003) studied the personal work constructs used by subordinates, incumbents, and supervisors to distinguish effective from ineffective performers for a targeted job. Although the constructs were more homogeneous from raters within the same organizational level than constructs from raters across different organizational levels, these source effects were small. Birkeland et al. (2003) concluded that conceptualizations of job performance do not vary substantially by level. In sum, recent research has shown that raters in different roles share a common conceptualization of managerial performance dimensions.

A Meta-Analysis of Longitudinal Studies of Multisource Feedback

In a recent narrative review, Seifert, Yukl, and McDonald (2003) concluded that there is little evidence that such feedback consistently results in behavior change or performance improvement. A narrative review by Smither, London, Flautt, Vargas, and Kucine (2002) of 13 longitudinal studies noted that, despite considerable variability in the magnitude of effect sizes across studies, 11 of the 13 studies found evidence of improvement over time for people receiving multisource feedback. In an effort to more accurately evaluate the impact of multisource feedback, we conducted a meta-analysis of longitudinal studies of multisource feedback to determine whether multisource feedback ratings improve over time.

Method

Literature search. We searched the PsycINFO database using terms such as *multisource feedback*, *upward feedback*, *multirater feedback*, and *360-degree feedback*. We searched for articles in journals and dissertations. We wrote to the authors of any studies identified through this search and asked if they were aware of any additional studies relevant to this meta-analysis. In addition, we reviewed the references in each of these articles to identify additional studies that might be relevant.

Inclusion criteria. We sought studies where multisource feedback was collected for the same focal individuals (ratees) on more than one occasion. We included studies where ratings were provided by direct reports, peers, supervisors, and/or self. To be included in the meta-analysis, a study had to contain sufficient information to calculate an effect size (i.e., either means and standard deviations, or values of t , F , or p). Two-thirds of these studies used customized multisource feedback instruments developed for the organization whereas the remaining studies used commercially available (e.g., from consulting firms) feedback instruments. The content of these instruments covered several domains such as leadership behaviors, managerial skills and practices, influence tactics, and coaching and appraising employees.

Three longitudinal studies (Conway, 1999; Hazucha, Hezlett, & Schneider, 1993; Luthans & Peterson, 2003) were not included in the meta-analysis because they combined peer, subordinate, and supervisor feedback into a single category called "other"; hence, we were unable to assess the effect size for specific rater sources. We also did not include data from Smither, London, Flautt, Vargas, and Kucine (2003) because the data they presented were for a subset of ratees that had been included in Smither, London, Flautt, Vargas, and Kucine (2002). We identified only two studies that included more than two waves of multisource ratings

(Reilly, Smither, & Vasilopoulos, 1996; Walker & Smither, 1999). We included only the first two waves of ratings from Walker and Smither. We did not use data from Reilly et al. (1996) because the first two waves of data in their study were identical to data presented in Smither et al. (1995), which was included in our analyses.

We did not include studies (or conditions within studies) where multisource feedback was combined with training (e.g., Bonnett, 1995; the feedback plus 24 hours of training condition in Rosti & Shipper, 1998; and the feedback plus 7-hour workshop condition in Seifert, Yukl, & McDonald, 2003) because it was not possible to determine whether any subsequent change in behavior (reflected in subsequent multisource ratings) was due to the feedback, the training, or a combination of the two. Finally, we did not include studies using college students conducted in a classroom setting (e.g., Dominick, Reilly, & Byrne, 2004).

Nonindependence. In most multisource feedback programs, ratings are collected on many items that are sometimes intended to assess multiple dimensions of the ratee's performance (e.g., communication, decisiveness, risk taking). However, 15 of the 24 studies that met our inclusion criteria calculated only a single score (i.e., the average rating across all items) for each ratee at Time 1 and at Time 2 for each rater source. Three other studies presented sufficient information only to calculate an overall effect size (across all items) although they presented means (but not *SDs*) for multiple dimensions. Only 6 of the 24 studies presented means and *SDs* for more than one dimension. These data points are not independent. In such instances, we averaged across performance dimensions to calculate an average effect size for each rater source in the study. As a result, our meta-analysis uses only one effect size estimate from each study for any rater source.

Description of variables. The literature search and inclusion criteria described above yielded 24 studies for our meta-analyses. We conducted separate meta-analyses for each rater source (direct reports, peers, supervisors, and self).

Three study designs were represented in our data: (a) single-group pretest-posttest design (repeated measures or RM)—17 studies, (b) independent groups posttest design (IG)—four studies, and (c) independent groups pretest-posttest design (IGPP)—three studies. In RM studies, all participants were rated and received feedback at Time 1 and Time 2. RM studies examined the change in ratings from Time 1 to Time 2. In IG studies, one group was rated and received feedback at Time 1 and Time 2, whereas a comparison or control group was rated and received feedback only at Time 2. IG studies compare the Time 2 ratings of the two groups. In IGPP studies, one group was rated and received feedback at Time 1 and Time 2, and a comparison or control group was rated but did not receive

feedback at Time 1 and was rated and received feedback at Time 2. IGPP studies compared the change in ratings from Time 1 to Time 2 for the two groups. For the RM studies, 7,355 ratees were rated by direct reports, 5,331 were rated by peers, 5,358 were rated by supervisors, and 3,684 ratees completed self-ratings. All data for IG and IGPP designs were based on direct report ratings with 192 ratees in the IG studies and 158 ratees in the IGPP studies.

Where sufficient information was provided, we coded several additional characteristics of the studies. We noted whether the study involved only upward feedback (i.e., only from direct reports) versus feedback from multiples sources (i.e., direct reports, peers, supervisor). We noted whether the feedback was used solely for developmental purposes (where only the ratee received the feedback report) versus studies where the ratee's supervisor also received the feedback report and multisource ratings could affect annual performance reviews, promotion decisions, or identifying high-potential leaders. We noted whether a facilitator was used to help feedback recipients interpret the feedback report and to offer guidelines for using the feedback. In two studies (Bailey & Fletcher, 2002; Smither et al., 2002), a facilitator was provided for some feedback recipients but not for others; we therefore excluded these studies when examining facilitation as a possible moderator of effect size. We also noted the length of time between the two administrations of the feedback program.

Calculating the effect size statistic and analyses. We used the meta-analysis procedures and formulas presented by Morris and DeShon (2002). Morris and DeShon present formulas for meta-analyzing data from each of the three study designs mentioned above. We used their formulas to calculate the effect size for each rater source in each study (e.g., for RM studies, the improvement from Time 1 to Time 2 feedback; for IG studies, the difference between Time 2 ratings for ratees who had previously received feedback and ratees who had not). For RM studies, the effect size was calculated as:

$$d = (M_{\text{Time 2}} - M_{\text{Time 1}}) / SD_{\text{Change Scores}}$$

Morris and DeShon refer to this as a change-score metric of effect size. We calculated the standard deviation of change scores as

$$SD_{\text{Change Scores}} = [SD_{\text{Time 1}}^2 + SD_{\text{Time 2}}^2 - (2)(r_{\text{Time 1, Time 2}})(SD_{\text{Time 1}})(SD_{\text{Time 2}})]^{1/2}$$

For studies that did not report $r_{\text{Time 1, Time 2}}$, we used the mean $r_{\text{Time 1, Time 2}}$ across studies for the same rater source and study design.

Morris and DeShon note that effect sizes from different designs cannot be directly compared without transforming all effect sizes to a common metric and that the choice of metric should be guided by the research

question. They state that the change-score metric is most appropriate when the focus of the research is on individual change (as it is here). We therefore used formulas presented by Morris and DeShon to convert all effect sizes to a common, change-score metric (i.e., effect sizes for IG and IGPP studies were initially calculated in the original metric and then converted to a change-score metric). We also used Morris and DeShon's formulas to compute the sampling variance (in the change score metric) associated with each effect size.

Following the recommendation of Morris and DeShon, the mean effect size was weighted by the reciprocal of the sampling variance. Because sampling variance is generally a decreasing function of sample size, this procedure results in studies with large sample sizes being weighted more heavily than those with small sample sizes. We evaluated the homogeneity of effect sizes using Hedge's (1982) Q , which is tested against a chi-square distribution with $k - 1$ degrees of freedom. We tested for differences between levels of a moderator using the between-groups statistic, Q_B , with $df = J - 1$, where J is the number of levels of the moderator.

Corrections for unreliability of performance ratings. Hunter and Schmidt (1990) note that the results from "bare bones" meta-analyses (i.e., which correct only for sampling error) provide biased (low) estimates of true effect sizes. They therefore argue that meta-analyses should correct for error of measurement (unreliability) whenever some information concerning the reliability of the measures is available. When information concerning reliability is available for each study in a meta-analysis, then each effect size can be corrected for unreliability and the corrected effect sizes can then be analyzed by meta-analysis to eliminate sampling error. However, in our set of studies (and many other meta-analysis studies), information about the reliability of ratings was available only for some studies. We therefore followed Hunter and Schmidt's (1990, p. 315) procedures to correct the mean and variance of the effect sizes by using the mean attenuation factor (i.e., the square root of the reliability coefficient) and the variance of the attenuation factors from studies that did provide information about the reliability of performance ratings.

The internal consistency (coefficient alpha) of performance ratings overestimates the reliability of performance ratings because it treats variance idiosyncratic to raters (e.g., halo) as true score (i.e., reliable) variance (Schmidt & Hunter, 1996). Instead, we followed the recommendation of Schmidt, Viswesvaran, and Ones (2000) and used measures of interrater reliability (e.g., intraclass coefficients) within each rater source (except self-ratings) as estimates of reliability.

We calculated the mean and variance of the attenuation factors (i.e., the square root of intraclass coefficients of interrater reliability) that were available concerning direct report ratings (mean reliability = .63, mean

attenuation factor = .79) and peer ratings (mean reliability = .55, mean attenuation factor = .74) in our studies. None of the studies in our meta-analysis provided data on interrater reliability of supervisor ratings. Previously, Rothstein (1990) examined the interrater reliability of supervisory ratings of 9,975 employees from 79 different organizations and reported that the mean reliability was .48. We therefore used .48 as an estimate of interrater reliability (and therefore .69 as the attenuation factor) of supervisor ratings. Because Rothstein did not provide the variance of supervisor reliabilities, we used the average of the variance of the direct report and peer attenuation factors in our studies as an estimate of the variance of attenuation factors for supervisor ratings. Finally, we used the mean correlation across studies between Time 1 and Time 2 ratings (.61) as an estimate of the reliability of self-ratings.

Results

Overall results from the meta-analyses are presented in Table 1. We present the unweighted (and uncorrected) effect sizes, the weighted (and uncorrected) effect sizes, and the weighted effect sizes corrected for measurement error. For direct report feedback, 19 of the 21 effect sizes were positive but the corrected mean effect size was only .15. For peer feedback, 6 of the 7 effect sizes were positive but the corrected mean effect size was only .05. For supervisor feedback, 8 of the 10 effect sizes were positive but the corrected mean effect size was only .15. For self-ratings, 6 of the 11 effect sizes were positive; the corrected mean effect size was $-.04$. For each rater source except supervisors, the 95% confidence interval included zero. Hedge's Q statistic indicated that the effect sizes were not homogeneous.

The study with the largest sample size (Smither, London, Flautt, Vargas, & Kucine, 2002) reported relatively small effect sizes ($d = .09$, $.03$, and $.09$ and $N = 2,763$, $4,319$, and $4,092$ for direct report, peer, and supervisor feedback, respectively). Only one other direct report effect size was based on more than 1,000 ratees (Johnson & Ferstl, 1999) and all other peer and supervisor effect sizes were based on less than 600 ratees. We therefore examined the mean unweighted (and uncorrected) effect sizes and compared them to the mean weighted (and uncorrected) effect sizes. The mean *unweighted* (and uncorrected) effect sizes were $.24$, $.12$, $.14$, and $.00$ for direct report, peer, supervisor, and self-ratings, respectively. The mean *weighted* (and uncorrected) effect sizes were $.12$, $.04$, $.10$, and $-.03$ for direct report, peer, supervisor, and self-ratings, respectively. Although the unweighted effect sizes are somewhat larger than the weighted effect sizes, they remain small in magnitude. According to Cohen (1988), small, medium, and large values of d are about $.20$, $.50$, and $.80$, respectively.

TABLE 1
Results of Meta-Analyses

Source of ratings	Number of data points (<i>k</i>)	<i>N</i>	Unweighted mean <i>d</i>	Weighted mean <i>d</i>	Corrected mean <i>d</i> ^a	Corrected <i>SD</i> ^b	Variance due to sampling error (%)	95% Confidence interval		<i>Q</i>
								Lower	Upper	
Direct reports: All studies	21	7705	.24	.12	.15	.15	17	-.14	.44	125.61**
Peers: All studies	7	5331	.12	.04	.05	.15	9	-.24	.34	74.14**
Supervisors: All studies	10	5358	.14	.10	.15	.07	47	.01	.29	21.13*
Self: All studies	11	3684	.00	-.03	-.04	.03	12	-.10	.02	93.83**

p* < .05 *p* < .01.

^aWeighted mean *d* corrected for measurement error (unreliability).

^bCorrected for sampling error and measurement error (unreliability).

That is, the results and conclusions remain largely unchanged when the Smither et al. (2002) data are given the same weight as data from the other studies.

Next, we examined three moderator variables related to study methodology. Results of the moderator analyses are presented in Table 2. First, we examined whether the study design was related to effect sizes of feedback from direct reports. Most studies of direct report feedback were repeated measures studies ($k = 14$, $N = 7,355$) without a control or comparison group. A small number of studies ($k = 4$, $N = 192$) used an independent group design (with a control or comparison group that did not receive feedback at Time 1) and three studies ($N = 158$) used an independent group pre-post design. Corrected effect sizes for the three designs were .15, .53, and .21, respectively. Hedges Q_B statistic indicates that differences in effect sizes between levels of this moderator variable were not significant ($Q_B = 3.84$, $df = 2$, $p > .10$).

Second, effect sizes for direct report feedback were also not related to whether ratees received feedback only from direct reports versus receiving feedback from direct reports and other sources such as peers and supervisors ($Q_B = .03$, $df = 1$).

Third, we examined whether time between administrations of the feedback program was related to effect size. We sorted studies into two categories based on whether the two administrations of the feedback program were separated by 12 or more months versus less than 12 months. We could not examine this moderator variable for supervisor and peer feedback because the time between the two administrations of the feedback program was 12 months or longer in all but one of these studies. As can be seen in Table 2, for direct report and self-ratings, effect sizes were significantly larger when the two administrations of the feedback program were separated by less than 12 months. This finding might be counterintuitive because one might have expected that a longer time span between administrations of multisource feedback would have provided feedback recipients with more time and opportunity to develop their skills and make performance improvements.

Discussion

Nearly all of the effect sizes for direct report, peer, and supervisor feedback were positive. However, the magnitude of improvement was very small. It is also noteworthy that in most instances there remained a large percentage of variance in effect sizes that was not explained by sampling error, even after accounting for the effects of moderator variables. This suggests that other factors might affect the extent of behavior change associated with multisource feedback.

TABLE 2
Results of Moderator Analyses Related to Research Method

	Number of data points (<i>k</i>)	<i>N</i>	Unweighted mean <i>d</i>	Weighted mean <i>d</i>	Corrected mean <i>d</i> ^a	Corrected <i>SD</i> ^b	Variance due to sampling error (%)	95% Confidence interval		<i>Q</i>
								Lower	Upper	
Pre-post studies	14	7,355	.24	.12	.15	.15	12	-.14	.44	119.04**
Independent group studies	4	192	.37	.42	.53	.00	100	.53	.53	2.41
Independent group, pre-post studies	3	158	.06	.17	.21	.00	100	.21	.21	.32
Direct reports: Study design ($Q_B = 3.84$)										
Upward feedback only	11	3,755	.24	.12	.15	.10	37	-.05	.35	29.86**
360° Feedback	10	3,950	.24	.12	.15	.19	10	-.22	.52	95.72**
Direct reports: Received only upward feedback versus received 360° feedback ($Q_B = 0.03$)										
Direct reports: Time ($Q_B = 17.53^{**}$)										
Less than 12 months	9	1,615	.29	.22	.28	.00	100	.28	.28	4.66
12 or more months	11	6,076	.24	.10	.13	.16	11	-.18	.44	103.21**
Self: Time ($Q_B = 40.77^{**}$)										
Less than 12 months	5	1,302	.08	.11	.14	.13	27	-.11	.39	18.36**
12 or more months	6	2,382	-.07	-.11	-.14	.14	17	-.41	.13	34.70**

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

^aWeighted mean *d* corrected for measurement error (unreliability).

^bCorrected for sampling error and measurement error (unreliability).

Why does the improvement in multisource feedback appear to be small? We consider several possible reasons. First, the small effect sizes might be related to the lack of discriminant validity in the items and dimensions contained in most multisource feedback instruments. It appears that raters often do not discriminate between the *a priori* dimensions very well. Although instruments often list many *a priori* dimensions, factor analyses of ratings generally find only a small number (e.g., one to three) of factors. Even when there is a more differentiated factor structure (e.g., Scullen, Mount, & Judge, 2003), the factors themselves are highly intercorrelated. The poor discriminant validity of dimensions and items could make it difficult for feedback recipients to easily and unambiguously identify areas of strength and weakness from the feedback report.

Second, this meta-analysis examined only two administrations of multisource feedback. We know of only two studies that have examined three or more administrations of multisource feedback. Reilly et al. (1996) followed 92 managers during four administrations of an upward feedback program over $2\frac{1}{2}$ years. They found that most of the performance improvement occurred between the first and second administrations (over 6 months) of the program but that these gains were sustained 2 years later. In contrast, Walker and Smither (1999) followed 252 managers over five annual administrations of an upward feedback program. They found no improvement between the first and second year, but they did find improvement between the second and third year and again between the third and fourth year. This suggests that multiple administrations of a feedback program might sometimes be required before significant improvement is observed.

Third, looking at the average rating across items at Time 1 versus Time 2 (as was done by the majority of studies in this meta-analysis) may be too crude a measure of behavior change. In many feedback programs, recipients are encouraged to select a small number (e.g., two or three) of improvement goals. A feedback recipient might subsequently make meaningful behavior changes related to two or three items but this would have little effect on the average rating across items. From this perspective, even small changes (effect sizes) over time might represent meaningful improvement in response to multisource feedback. For this reason, Smither and Walker (2001) have suggested alternative approaches to measuring the impact of multisource feedback such as retrospective degree of change ratings, where coworkers rate the extent to which the feedback recipient's performance has improved (or declined) with respect to each of the improvement goals the feedback recipient set after receiving feedback at Time 1. For example, Smither et al. (2002) asked each feedback recipient to set improvement goals. Multisource ratings of improvement (where 1 = *less effective*, 3 = *no change*, and 5 = *more effective*) collected

8 months later indicated that raters thought feedback recipients were making progress toward their individual goals (from direct reports, $M = 4.09$, $SD = 0.51$; from supervisors, $M = 3.99$, $SD = 0.50$; from peers, $M = 3.98$, $SD = 0.41$). Ideally, it would also be desirable to gather retrospective degree of change ratings on control items (i.e., behaviors and skills that were not improvement goals for the feedback recipient). A comparison between retrospective degree of change ratings on the feedback recipient's goals versus control items would be informative. That is, change ratings on the feedback recipient's goals should indicate improvement, whereas change ratings on control items should not (see Peterson, 1993, for an example).

Four of the 24 studies in this meta-analysis examined whether feedback recipients improved more on a subset of dimensions where improvement would be especially likely. Nemeroff and Cosentino (1979) found that feedback recipients improved on the items for which they set goals but not on other items. However, Avery (2000) found that feedback recipients did not improve more on the dimensions where they were rated lowest (i.e., where improvement was arguably most needed). Quadracci (1995) found that feedback recipients did not improve more on a subset of dimensions designated by the organization as being critical. And Hezlett and Ronnkvist (1996) found that managers did not improve more on skills they targeted for development than other skills.

Fourth, research in social cognition and performance appraisal (Buda, Reilly, & Smither, 1991; Smither, Reilly, & Buda, 1988) has shown that once a person (e.g., a ratee) has been categorized by others (e.g., raters), raters will subsequently be likely to attend to and recall information about the ratee in a way that is consistent with their initial categorization of the ratee. This suggests that raters might not notice or recall small or modest improvements in ratee behavior. This may be another factor that explains why we do not see substantial changes in multisource ratings over time.

Each of the issues described above might have contributed to the small effect sizes in this meta-analysis. However, we think that the most likely explanation for the small effect sizes is that some feedback recipients are more likely to change than others. We therefore turn our attention to variables that might influence the extent of performance improvement following multisource feedback.

Theoretical Framework: Multisource Feedback and Performance Improvement

London and Smither (1995) argued that these positive outcomes would be greater for feedback recipients with high self-efficacy and lower for feedback recipients who focus on managing others' impressions of them rather than actually changing their behavior. In this article, we offer a more

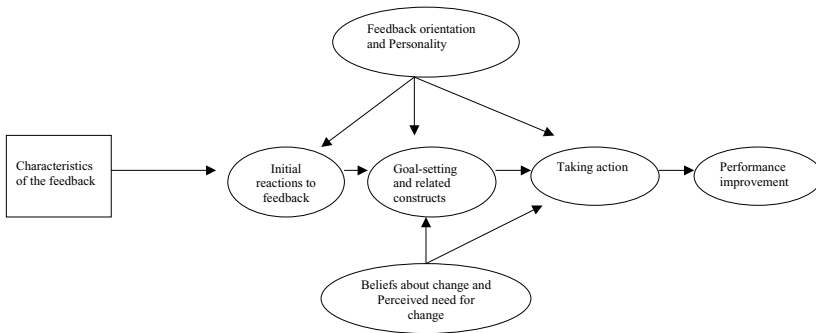


Figure 1: Theoretical Model for Understanding Performance Improvement Following Multisource Feedback.

developed but still preliminary theoretical model (see Figure 1). Empirical research since London and Smither (1995) has not been organized around any single conceptual model. Although each study has examined interesting constructs or variables, results from these studies appear at first glance to be unconnected. Thus, the first purpose of our model is to provide a way of organizing recent research so that converging evidence and emerging conclusions can be more easily detected. Second, we use the model to suggest directions for future research.

We believe that the accumulated evidence (described below) points to eight broad factors that play a role in determining the extent of behavior change and performance improvement following multisource feedback: characteristics of the feedback, initial reactions to feedback, personality, feedback orientation, perceived need for change, beliefs about change, goal setting, and taking action. The model begins with the characteristics of the feedback itself. These characteristics influence recipients' initial reactions, which in turn affect goal setting. Goal setting influences taking actions, which in turn determine performance improvement. Personality and feedback orientation influence reactions to feedback, goal setting, and taking action. Beliefs about change influence initial reactions and goal setting, and perceived need for change influences goal setting and taking action.

One implication of this model is that multisource feedback should not be expected to lead to substantial performance improvement for all feedback recipients. Another implication of our model is that the effects of some factors on performance improvement are likely to be indirect rather than direct. Unfortunately, research has not yet explored many of the causal links proposed in this model. As a result, we sometimes describe research showing the relationship between a factor in our model and performance

improvement even though we propose that the factor has an indirect rather than a direct effect on performance improvement. Future research is needed to test many of the indirect effects proposed in the model.

Characteristics of Feedback

Several characteristics of multisource feedback are likely to influence initial reactions and subsequent behavior. The most obvious characteristic is whether the feedback is positive or negative. Recipients may interpret feedback as negative when ratings from other sources are lower than self-ratings.

Several studies have examined correlates of self–other rating differences (Church, 1997; Fletcher & Baldry, 2000; McEvoy, 1990; Pollack & Pollack, 1996; Smith & Helland, 2002). Unfortunately, studies have generally used a single index of agreement (e.g., a difference score) rather than multivariate procedures that would enhance our understanding concerning the underlying reason for the extent of disagreement. For example, if it were discovered that extraverted employees are more likely to be overraters, this could be due to higher self-ratings, lower ratings from others, or a combination of the two. A recent exception is Ostroff, Atwater, and Feinberg's (2004) study of multisource feedback received by over 4,000 managers across 650 organizations. They used multivariate regression analyses and found that differences between self and other ratings were related to gender (not because of differences in self-ratings but because men were rated less favorably than women by others), race (because non-Whites rated themselves higher than Whites), age (because older managers rated themselves higher but were rated lower by others), experience (because more experienced managers rated themselves more favorably than less experienced managers), and education (managers with less education were rated lower by others but did not differ from more educated managers in their self-ratings).

Nilsen and Campbell (1993) showed that in the absence of feedback, self–other rating differences were stable over time. However, Yammarino and Atwater (1993) found that feedback from subordinates altered self-evaluations of leadership. Atwater, Rouch, and Fischthal (1995) found that, after receiving upward feedback, leaders' subsequent self-ratings more closely matched those of their subordinates. Overestimators lowered their self-ratings and underestimators raised their self-ratings. However, Smither et al. (1995) did not find evidence of changes in self-evaluations following receipt of upward feedback.

Finally, Atwater and Brett (in press) found that, when leaders received unfavorable feedback, those who agreed with others about the feedback (i.e., who rated themselves low) were less motivated than those who

received low ratings and overrated themselves. For leaders who received favorable feedback, agreement between self and others did not influence their motivation.

In sum, research indicates that self–other rating differences are related to individual characteristics, stable over time in the absence of feedback, reduced after receiving multisource feedback, and related to motivation to use the feedback.

Initial Reactions to Feedback

Initial reactions to feedback, especially affective reactions, can be expected to influence whether feedback recipients use the feedback to set goals and make performance improvements. For example, unfavorable feedback will not lead to performance improvement when recipients reject the feedback (Ilgen, Fisher, & Taylor, 1979). Kluger and DeNisi (1996) note that extremely negative feedback can lead recipients to abandon their goals (e.g., to abandon the goal of performing effectively in the view of others). Their meta-analysis found that discouraging feedback and feedback that threatens the recipient's self-esteem decreased the effectiveness of feedback interventions. In sum, if initial reactions to feedback are extremely negative or lead recipients to reject the feedback, subsequent performance improvement is unlikely.

In the context of multisource feedback, Brett and Atwater (2001) found that leaders who received negative feedback reacted with anger and felt discouraged; however, those who received positive feedback did not react with positive affect. Reactions immediately following feedback were positively related to feelings that the feedback was useful; however, this reaction was short-lived. Three weeks later, the sign of the feedback was not related to perceptions of usefulness. How recipients react to negative or positive feedback can also depend on the source of the feedback. For example, ratees react more positively when the feedback is perceived to come from a credible source, such as peers (Albright & Levy, 1995).

Recently, Atwater and Brett (in press) found that leaders who expressed more motivation and had more positive emotions immediately after receiving multisource feedback subsequently improved in terms of direct reports' ratings (1 year later) but those who expressed negative emotions showed a decline in direct reports' ratings. Leaders who were more motivated also improved in terms of peer ratings; however, leaders' reactions were unrelated to changes in supervisors' ratings. These findings are important because they demonstrate that immediate reactions to feedback are not merely transitory mood states without relevance to subsequent behavior.

Feedback Orientation

London and Smither (2002) described feedback orientation as an individual-difference construct related to an employee's predisposition to seek and use feedback. They proposed that people who are high in feedback orientation are not afraid of being evaluated, like and seek feedback, process it carefully, care about how others view them, believe that feedback offers insights that may help them become more effective, and feel accountable to use feedback. Rutkowski, Steelman, and Griffith (2004) recently found that feedback orientation is positively correlated with feedback acceptance. A related individual difference variable is propensity for continuous learning. Continuous learners anticipate changes in job requirements, request and use feedback, set development goals, participate in learning activities, practice new behaviors, apply learning on the job, and improve their performance (Vicere & Fulmer, 1998). Empirical research on individual differences in feedback orientation and continuous learning can help us better understand the characteristics of employees who are most likely to benefit from multisource feedback.

Personality

A number of personality variables are logically related to how one might react to feedback or use feedback to set goals and take action. For example, Smither, London, and Richmond (in press) found that leaders' emotional stability was positively related to a psychologist's ratings (completed immediately after the leader received multisource feedback) of the leaders' motivation to use the results from multisource feedback. Smither et al. (in press) also found that leaders' extraversion was positively related to requesting additional feedback and conscientiousness was positively related to subsequently participating in developmental activities and believing that their performance had improved 6 months after receiving multisource feedback. Leaders who were high in conscientiousness, particularly the responsibility component, indicated that they felt obligated to use the feedback. Maurer and Palmer (1999) found that intentions to participate in development following feedback from peers and subordinates were related to recipients' sense of having control over their own improvement. Dominick, Reilly, and Byrne (2004) argued that conscientiousness (being dutiful, reliable, organized, responsible, and achievement oriented) should be related to setting and attaining goals after receiving peer feedback. They found that conscientiousness was positively related to performance improvement (i.e., enhanced effectiveness as a team member) after receiving peer feedback from classmates (where ratings were collected from different peers over a two-semester period). In the same study

Dominick et al. found that openness to experience was also positively related to performance improvement. Atwater and Brett (in press) found that dispositional distrust was negatively related to feedback recipients' attitudes toward using feedback. In addition, leaders low in emotional stability reported more negative emotions (describing themselves as angry, frustrated, unhappy, discouraged, and disappointed) after receiving feedback, although they did not receive less favorable feedback than other leaders. Research has shown that high self-monitors get better supervisor ratings but also tend to inflate their self-ratings (Day, Schleicher, Unckless, & Hiller, 2002; Miller & Cardy, 2000). Inflated self-ratings are likely to produce a gap for self-monitors, especially relative to peer and subordinate ratings. Because a characteristic of high self-monitors is adaptation to situational requirements (Gangestad & Snyder, 2000; Miller & Cardy, 2000), it seems reasonable to expect that self-monitoring should be related to performance improvement following feedback.

Beliefs About Change

Even if feedback recipients accept feedback and believe they need to change their behavior, they might not exert effort to change because they do not believe change is possible or that change will result in some positive outcome. Examples of variables related to beliefs about change include self-efficacy, implicit personality theory (i.e., beliefs about the malleability of personal attributes such as ability or personality), and organizational cynicism. An entity implicit theory refers to the belief that personal attributes are largely fixed whereas an incremental implicit theory refers to the belief that personal attributes are malleable (e.g., Dweck & Leggett, 1988). Feedback recipients with low self-efficacy and entity implicit theories are likely to believe that change is futile and therefore can be expected to exert little effort to change their behavior after receiving multisource feedback.

Several studies have examined individual-difference variables that are related to feedback recipients' beliefs about change. Atwater and Brett (in press) found that feedback recipients with higher self-efficacy were more likely than others to engage in follow-up activities. Heslin and Latham (2004), in a study of Australian managers in a professional services firm who received upward feedback, found that those with high self-efficacy subsequently improved more than other managers.

We are unaware of any research that has examined the role of implicit personality theory in reactions to or use of multisource feedback. However, Heslin, Latham, and VandeWalle (in press) found that managers who held incremental beliefs were more likely than managers who held entity beliefs to recognize both improvements and declines in employee performance.

In a separate study, these authors used a 90-minute workshop based on self-persuasion techniques to help participants who initially held entity implicit theory beliefs to acquire incremental implicit theory beliefs and to sustain those beliefs over a 6-week period. This change led to greater acknowledgement of improvement in employee performance than was exhibited by entity theorists in a placebo control group. In a related study, Heslin, Latham, and VandeWalle (2004) found that inducing incremental implicit theory beliefs was positively related to willingness to coach a poor-performing employee and to the quality and quantity of performance improvement suggestions offered. These studies are promising because they suggest that incremental implicit theory beliefs can be developed and that they are related to work-related behavior. Perhaps, feedback recipients can be induced to hold incremental implicit theory beliefs and thereby enhance the likelihood that they will subsequently change their behavior and improve their performance.

Organizational cynicism may lead some employees to question the utility of changing their behavior. For example, cynics do not believe that good performance will result in earning more money (Wanous, Reichers, & Austin, 2000). It is reasonable to expect that cynics would question the value of setting goals to improve their performance in response to feedback. Atwater, Waldman, Atwater, and Cartier (2000), in a study of police supervisors who received upward feedback, found that supervisors who were low in organizational cynicism subsequently improved their performance more than other supervisors.

In sum, emerging evidence indicates that beliefs about change are related to whether performance improves after receiving multisource feedback.

Perceived Need for Change

Perceiving a need to change should enhance the likelihood of setting performance improvement goals and taking action. London and Smither (1995) argued that discrepancies between self-ratings and feedback from others can lead feedback recipients to perceive a need to change their behavior. Merely receiving unfavorable ratings or ratings that are below average (e.g., below norms) might also increase feedback recipients' perceived need to improve their performance.

If receiving unfavorable feedback or feedback that differs from one's self-evaluations leads feedback recipients to perceive a need to change and, hence, to set goals, then we would expect these feedback recipients to subsequently improve their performance more than others. Two studies (Atwater, Rouch, & Fischthal, 1995; Johnson & Ferstl, 1999) have found that feedback recipients who initially overrated themselves subsequently

improved more than others. Two studies (Smither et al., 1995; Walker & Smither, 1999) found that feedback recipients who initially received low ratings subsequently improved more than others. In each of these studies, the improvement of feedback recipients who initially overrated themselves or who initially received unfavorable feedback was greater than what would be expected on the basis of statistical regression to the mean. Taken together, these findings show that feedback recipients who receive unfavorable feedback or who initially overrate themselves tend to improve more than others and thereby provide preliminary support for the idea that perceived need to change plays a role in determining whether feedback recipients subsequently improve their performance.

But it is unrealistic to expect that everyone who receives multisource feedback will perceive a need to change. Some, perhaps many, may be satisfied with their feedback either because it is generally favorable (e.g., above the midpoint of the rating scale) or because their unfavorable feedback was consistent with a low self-evaluation (e.g., Johnson & Ferstl, 1999). Individuals who receive feedback that is favorable (either in an absolute sense or relative to item norms) or that is reasonably consistent with their self-ratings may see no need to change.

Finally, some feedback recipients may perceive a need to change but might not set goals because they believe that change is not feasible (see discussion concerning *Beliefs About Change*). In addition, as noted below, people may sometimes set performance improvement goals even after receiving favorable feedback (i.e., they want to improve even though the feedback does not point to a need to change).

Goal Setting and Related Constructs

Locke and Latham (1990) have shown that feedback alone is not the cause of behavior change; instead it is the goals that people set in response to feedback. Brutus, London, and Martineau (1999) reported that ratees who received negative feedback were more likely to set goals, although this may have been influenced by participation in a leadership development program following feedback. Smither et al. (in press) found that leaders who received unfavorable feedback had set more improvement goals than other leaders 6 months later. Atwater et al. (2000) found that a measure of reactions to feedback that included using the feedback to set improvement goals was positively related to improvement after receiving upward feedback.

Next, we describe goal-related constructs that are likely to affect whether feedback recipients set performance improvement goals and the nature of any goals that are set.

Regulatory focus and goal orientation. Higgins (1998; Brockner & Higgins, 2001) has distinguished between a promotion regulatory focus

that orients the person toward attaining positive outcomes versus a prevention regulatory focus that orients the person toward minimizing negative outcomes. Promotion focus concerns striving for ideals (presumably set by the self) whereas prevention focus concerns being motivated by oughts (often expectations established by others). Higgins has argued that regulatory focus, in part, develops as a consequence of nurturance-oriented parenting (which instills a promotion focus in children) or security-oriented parenting (which instills a prevention focus). In this sense, regulatory focus can be thought of as a dispositional variable. However, Higgins and others also acknowledge that regulatory focus can be situationally induced (e.g., Thorsteinson & Highhouse, 2003; Van-Dijk & Kluger, 2004). Moreover, Higgins (1999) points out that the effects of regulatory focus are comparable regardless of whether it varies as a function of persons (i.e., a dispositional variable) or situations. Indeed, research has shown that situational features can make one or the other regulatory focus more accessible (at least temporarily) and thereby influence the goals that people set and their persistence and achievement (Roney, Higgins, & Shah, 1995).

Goal orientation is a construct very similar to regulatory focus. VandeWall, Cron, and Slocum (2001), building on the work of Dweck, described a three-factor model of goal orientation. A learning goal orientation (similar to a promotion regulatory focus) refers to a focus on becoming more competent by acquiring new skills, mastering new situations, and learning from experience. A proving goal orientation refers to a focus on demonstrating competence and gaining favorable judgments from others. An avoiding goal orientation (similar to a prevention regulatory focus) refers to a focus on avoiding negative outcomes and negative judgments from others. VandeWall et al. (2001) found that learning goal orientation was positively related to effort, self-efficacy, goal setting, and performance on exams and avoiding goal orientation was negatively related to self-efficacy, goal setting, and performance. One recent study (Heslin & Latham, 2004) found that a learning goal orientation was positively related to improvement in upward feedback.

In the context of multisource feedback, we suggest that using the feedback solely to guide the recipient's development (where only the ratee receives the feedback) will create a promotion focus (or learning goal orientation) and thereby enhance subsequent performance improvement. In contrast, we suggest that using multisource feedback for administrative purposes (where the ratee's supervisor receives a copy of the feedback report and can use the information to influence decisions related to the ratee's formal performance appraisal, pay, or promotion opportunities) will create a prevention focus (or avoiding goal orientation).

Research has shown that employees believe that appraisals used for development are more likely to produce positive outcomes than appraisals used for administrative purposes (Bettenhausen & Fedor, 1997). Greguras,

Robie, Schleicher, and Goff (2003) found that ratings collected for administrative purposes were more likely to be affected by rater biases than were ratings collected for developmental purposes. For example, subordinate ratings showed higher interrater reliability for developmental than for administrative purposes. Brutus and Petosa (2002) found that when multisource ratings were to be used for administrative purposes and ratees were allowed to select raters, ratees selected raters whom they liked, increasing the chances that the ratings would be inflated. However, when the purpose was developmental, ratees selected raters who were more likely to provide accurate evaluations. In sum, it appears that multisource ratings (like performance appraisal ratings in general) are less likely to be subject to biases and other rating errors when the ratings are collected for developmental rather than administrative purposes.

We examined whether effect sizes in our meta-analysis were related to the purpose of the feedback (i.e., whether the feedback was used solely for developmental purposes vs. also being used for administrative purposes). The Q_B statistics in Table 3 show that purpose was significantly related to effect size. The moderator analyses indicated that effect sizes were larger when feedback was used only for developmental purposes. This was true regardless of whether one looks at the unweighted or weighted effect sizes, thereby indicating that this finding was not solely the result of the large sample sizes and small effect sizes in Smither et al. (2002) where the feedback was used for administrative purposes. Across rater sources (excluding self-ratings), the average effect size in the developmental purpose studies was .25 (vs. .08 in the administrative purpose studies—see Table 3 for details). It is important that these findings be interpreted with caution because very few studies used the feedback for administrative purposes. Nevertheless, these results reveal that when multisource feedback is used solely for developmental purposes, it produces more positive gains in performance than when it is used for administrative purposes. This suggests the important role of a promotion focus (or learning goal orientation) in this process.

Discrepancy reduction versus discrepancy production. Social Cognitive Theory (Bandura, 1991) describes two mechanisms that people use to regulate their performance via personal goals. *Discrepancy reduction* occurs when individuals monitor and work to reduce any discrepancy between their goals and performance. *Discrepancy production* occurs when individuals have met or exceeded their goals and then set a more challenging goal. That is, individuals create a discrepancy between their current performance and their goals to motivate themselves to attain a higher level of performance (e.g., Donovan & Williams, 2003; Phillips, Hollenbeck, & Ilgen, 1996; Williams, Donovan, & Dodge, 2000). This process enables individuals to attain satisfaction by progressively improving their

TABLE 3
Results of Moderator Analyses Related to Theoretical Framework

	Number of data points (<i>k</i>)	<i>N</i>	Unweighted mean <i>d</i>	Weighted mean <i>d</i>	Corrected mean <i>d</i> ^a	Corrected <i>SD</i> ^b	Variance due to sampling error (%)	95% Confidence interval		<i>Q</i>
								Lower	Upper	
Developmental only Administrative	15	2,504	.28	.22	.28	.21	18	-.13	.69	81.84**
	4	5,015	.02	.07	.09	.01	94	.07	.11	4.25
Developmental only Administrative	5	903	.28	.17	.23	.19	22	-.14	.60	23.18**
	2	4,428	-.25	.02	.03	.11	6	-.19	.25	32.92**
Developmental only Administrative	8	1,145	.17	.16	.23	.07	70	.09	.37	11.42
	2	4,213	-.01	.08	.12	.03	43	.06	.18	4.66**
Developmental only Administrative	9	1,735	.01	.09	.12	.19	19	-.25	.49	47.51**
	2	1,949	-.05	-.13	-.17	.04	57	-.25	-.09	3.50
Yes No	9	912	.21	.12	.15	.20	33	-.24	.54	27.44**
	11	4,030	.29	.14	.18	.19	12	-.19	.55	94.16**
Yes No	4	449	.01	.04	.05	.45	8	-.83	.93	52.14**
	2	563	.40	.14	.19	.22	12	-.24	.62	16.10**
Yes No	4	459	.17	.15	.22	.20	31	-.17	.61	12.97**
	5	807	.12	.12	.17	.04	84	.09	.25	5.96
Yes No	3	401	.01	.02	.03	.08	61	-.13	.19	4.89
	7	3,205	.01	-.03	-.04	.20	8	-.43	.35	87.21**

p* < .05 *p* < .01.

^aWeighted mean *d* corrected for measurement error (unreliability).

^bCorrected for sampling error and measurement error (unreliability).

task performance. According to Bandura (1986), high self-efficacy is an important determinant of discrepancy production. The concept of discrepancy production is important because it suggests that feedback recipients might set goals to improve their performance even after receiving favorable multisource feedback. Unfortunately, we know little about the conditions under which discrepancy production is likely to occur in the context of multisource feedback.

In sum, we suggest that feedback recipients who have a promotion (rather than prevention) self-regulatory focus (or a learning orientation) and engage in discrepancy production should be especially likely to set performance improvement goals after receiving multisource feedback.

Unfortunately, it is impossible to know the extent of goal setting that followed receipt of multisource feedback in most studies. Most studies we reviewed for this meta-analysis indicated that feedback recipients were encouraged to set goals, but there were no measures of the extent to which feedback recipients actually set goals or of the quality or commitment to any goals that might have been set. One exception was Smither et al. (2002) where nearly all feedback recipients set one to three developmental goals. Multisource feedback on a follow-up survey indicated that feedback recipients were generally making progress toward the specific developmental goals they had set. It would be valuable for future research to collect data concerning the extent to which feedback recipients set goals after receiving multisource feedback.

In practice, multisource feedback recipients often meet (usually in groups) with a facilitator who, among other things, encourages them to use the feedback to set developmental goals. We compared studies that explicitly indicated a facilitator met with feedback recipients to help them interpret the feedback report and encourage them to set goals with studies that did not mention the availability of a facilitator. The Q_B statistics in Table 3 show that facilitation was not significantly related to effect size.

Taking Action

Ultimately, performance improvement is likely only for feedback recipients who take appropriate action (e.g., working with a coach, discussing one's feedback with others, participating in development activities) in response to their feedback. Sometimes, organizations provide support to help feedback recipients take appropriate action. Examples include supportive supervisors or executive coaches who emphasize the importance of acting on the feedback and help feedback recipients establish and follow through on appropriate goals. Feedback recipients can talk to others about the feedback, thereby helping to clarify the feedback and identify useful actions to enhance their effectiveness. Finally,

feedback recipients can engage in a variety of informal or formal developmental activities to address skill deficiencies (e.g., discrepancy reduction), acquire new skills, or enhance existing skills (e.g., discrepancy production).

Two studies suggest that working with an executive coach can be helpful for feedback recipients. In a study where 20 managers received multi-rater feedback *and* executive coaching, Luthans and Peterson (2003) found that self-ratings of managers remained the same over time (3 months) but coworker ratings increased. In addition, for both the managers and their employees, work attitudes (job satisfaction, organizational commitment, and turnover intentions) improved following the feedback-coaching intervention. Using a quasi-experimental design, Smither et al. (2003) found that managers who worked with an executive coach were more likely than other managers to set specific (rather than vague) goals, to solicit ideas for improvement from their supervisors, and to improve in terms of subsequent direct report and supervisor ratings. However, the differences between managers who worked with a coach and those who did not were small in magnitude (albeit statistically significant).

In a 5-year study of upward feedback, Walker and Smither (1999) found that (a) managers who met with direct reports to discuss their upward feedback improved more than other managers, and (b) managers improved more in years when they discussed the previous year's feedback with direct reports than in years when they did not discuss the previous year's feedback with direct reports. Smither et al. (2004) also found that sharing multisource feedback and asking for suggestions from raters was positively related to improvement over time.

Maurer and Palmer (1999) found that managers' intentions to participate in development following feedback from peers and subordinates were positively related to their perceptions of social pressures for improvement and rewards or benefits from improvement. Finally, Hazucha et al. (1993) found that managers who participated in training programs and other development activities (e.g., receiving coaching and feedback, reviewing progress quarterly) after receiving multisource feedback were more likely to improve than other managers.

In summary, these studies provide evidence that taking action is positively related to performance improvement after receiving multisource feedback.

Directions for Future Research

We have noted above several promising directions for future research that are related to our theoretical model. We believe that several additional

individual-level and organizational-level variables might facilitate or detract from behavior change and performance improvement.

Another set of individual-level variables includes attitudes such as job satisfaction and organizational commitment (Meyer & Allen, 1997). For example, employees who are satisfied with their jobs and who have high affective commitment to the organization (Allen & Meyer, 1996) should be more likely to use multisource feedback to set goals to improve their performance and contributions to the organization.

Organizational-level variables might also influence the impact of multisource feedback. For example, a strong feedback culture (London & Smither, 2002) may increase the likelihood that feedback recipients will respond positively to feedback and use it to guide behavior change. London and Smither (2002) described a feedback-oriented culture as one that implements multisource feedback for development, trains supervisors to be better coaches, and strengthens the link between performance improvement and valued outcomes, such as pay and promotion. A positive development in this area is found in recent work by Steelman, Levy, and Snell (2002; see also Steelman & Levy, 2001) who developed a Feedback Environment Scale (FES) that measures the extent to which the organization supports the use of feedback, including the formal presentation of the annual performance appraisal and informal daily interactions during which performance is discussed.

Research is also needed to understand how organizational interventions such as training can enhance the impact of multisource feedback. For example, Seifert, Yukl, and McDonald (2003) compared the joint effects of a multisource feedback measure of influence tactics and a 7-hour workshop to (a) a control condition that received no feedback and did not attend the workshop, and (b) a comparison group of managers who received a feedback report but did not participate in the workshop. There were seven middle managers in each of the three conditions. The workshop included a description of effective influence tactics, a video that illustrated the use of these tactics, and an exercise in which feedback recipients considered ways to use the influence tactics in situations they were likely to encounter at work. Participants in the workshop also reviewed their feedback reports and created action plans to improve their influence behavior. Influence tactics with subordinates improved for managers in the experimental condition (feedback plus workshop) but did not change for managers in the control or comparison groups. Seifert et al. (2003) concluded that the effectiveness of multisource feedback might be partially dependent on facilitating conditions during the feedback process, such as training to explain and demonstrate the desired behaviors. Unfortunately, other studies that have included a training intervention after multisource feedback have not compared the combined effects of

multisource feedback and training to the effect of multisource feedback alone (e.g., Rosti & Shipper, 1998). Seifert et al. (2003) are an exception but the very small sample size limits confidence in the generalizability of their findings. In sum, research is needed to understand whether organizational investments in training linked to multisource feedback increase the impact of the feedback beyond what would have occurred with only feedback or only training.

London, Smither, and Adsit (1997) described organizational actions that might enhance feedback recipients' sense of accountability to use their feedback. Examples include providing adequate resources to support employee development, linking training and development opportunities to the competencies assessed in multisource ratings, providing feedback recipients with facilitation and coaching to help them interpret and use the feedback, and rewarding behavior change and performance improvement following multisource feedback. Rutkowski, Steelman, and Griffith (2004) found that police officers who had positive perceptions concerning the feedback environment (as measured via the Feedback Environment Scale) in their organization felt more accountable to use feedback to improve their performance. We know of only two empirical studies that have examined whether accountability is related to behavior change following multisource feedback. Leonard and Williams (2001) found that ratees' sense of accountability (measured before feedback was received) was a better predictor of developmental behaviors taken in response to 360° feedback (measured 2 months after feedback was received) than other variables including self-efficacy, need for achievement, supervisor support, and situational constraints. And Atwater and Brett (in press) found that leaders who thought it was important to use the feedback given the time their raters invested in the process (a component of accountability) improved in terms of peer (but not direct report or supervisor) feedback over a 1-year period.

In addition to being affected by organizational-level variables, it is also possible that multisource feedback programs may affect organizational-level variables such as organizational culture. As one reviewer noted, relevant research and thinking might be found in literatures heretofore rarely considered by multisource feedback researchers (e.g., the strategic human resource management literature). Research that simultaneously examines both individual-level factors and organizational-level factors would help illuminate our understanding of how organizational factors enhance or limit the impact of multisource feedback. Multilevel analytic techniques such as hierarchical linear modeling (Bryk & Raudenbush, 2001) are well suited to answering such questions. The interaction between organizational-level variables and individual-level variables is especially interesting. For example, organizational factors (such as a strong feedback culture or the

availability of training) might enhance the impact of multisource feedback for some employees (e.g., those who are high in openness to experience) but not others (those who are low in conscientiousness).

Conclusion

The results of this meta-analysis and evidence related to the theoretical model indicate that it is unrealistic for practitioners to expect large across-the-board performance improvement after people receive multisource feedback. Instead, it appears that some feedback recipients will be more likely to improve than others. We therefore think it is time for researchers and practitioners to ask “Under what conditions and for whom is multisource feedback likely to be beneficial?” (rather than asking “Does multisource feedback work?”). Although emerging evidence is supportive of this preliminary theoretical model, a great deal of additional research is needed before unambiguous guidance can be offered to practitioners about the precise conditions where multisource feedback is likely to be most beneficial.

In the interim, we think that our preliminary theoretical model offers some promising suggestions for practitioners. For example, they can send messages that emphasize a learning orientation and a promotion focus (e.g., on attaining positive outcomes and acquiring new skills rather than on how feedback recipients are performing relative to each other). They can encourage recipients who receive favorable feedback to set even more challenging goals (discrepancy production) that build on their strengths (e.g., Buckingham & Clifton, 2001). They can take steps to enhance the self-efficacy of feedback recipients (e.g., Noe, 2005, p. 89) and they can use self-persuasion techniques (see Heslin, Latham, & VandeWalle, in press) to help feedback recipients develop an incremental implicit theory where they see personal change and performance improvement as not only possible but also as probable. And they can encourage feedback recipients to talk to others to clarify their feedback, to set specific goals, and to participate in formal and informal development activities. At the same time, practitioners and managers will need to remember that, like any other training and development investment, not all participants will benefit equally.

REFERENCES

- References marked with an asterisk indicate that the studies were included in the meta-analysis.
- Albright MD, Levy PE. (1995). The effects of source credibility and performance rating discrepancy on reactions to multiple raters. *Journal of Applied Social Psychology*, 25, 577–600.

- Allen NJ, Meyer JP. (1996). Affective, continuance, and normative commitment to the organization: An examination of construct validity. *Journal of Vocational Behavior*, 49, 252–276.
- *Antonioni D. (1995). Problems associated with implementation of an effective upward appraisal process: An experimental field study. *Human Resource Development Quarterly*, 6, 157–171.
- Atkins PWB, Wood RE. (2002). Self-versus others' ratings as predictors of assessment center ratings: Validation evidence for 360-degree feedback programs. *PERSONNEL PSYCHOLOGY*, 55, 871–904.
- *Atwater LE, Brett J. (2003). Unpublished data.
- Atwater LE, Brett JF. (in press). Antecedents and consequences of reactions to 360-degree feedback. *Journal of Vocational Behavior*.
- Atwater LE, Brett JF, Ryan JM. (2004, April). *360 degree feedback to managers: Does it result in changes in employee attitudes?* Paper presented in a symposium titled "Understanding the Effectiveness of 360-Degree Feedback Programs" (Chair: Joan Brett) at the 19th Annual Conference of the Society for Industrial and Organizational Psychology, Chicago.
- *Atwater LE, Rouch P, Fischthal A. (1995). The influence of upward feedback on self- and follower raters of leadership. *PERSONNEL PSYCHOLOGY*, 48, 34–60.
- *Atwater LE, Waldman DA, Atwater D, Cartier P. (2000). An upward feedback field experiment: Supervisors' cynicism, reactions, and commitment to subordinates. *PERSONNEL PSYCHOLOGY*, 53, 275–297.
- *Avery K. (2000). *The effects of 360 degree feedback over time*. Unpublished doctoral dissertation. Chicago School of Professional Psychology.
- *Bailey C. (2003). *360-degree feedback and developmental outcomes: The role of feedback characteristics, self-efficacy, and importance of feedback dimensions to focal managers' current role*. Manuscript submitted for publication.
- *Bailey C, Fletcher C. (2002). The impact of multiple source feedback on management development: Findings from a longitudinal study. *Journal of Organizational Behavior*, 23, 853–867.
- Bandura A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.
- Bandura A. (1991). Social cognitive theory of self-regulation. *Organizational Behavior and Human Decision Processes*, 50, 248–287.
- *Bernardin HJ, Hagan C, Ross S, Kane JS. (1995, May). *The effects of a 360-degree appraisal system on managerial performance*. Paper presented in a symposium titled "Upward feedback: The ups and downs of it" (Chair: W.W. Tornow) at the 10th Annual Conference of the Society for Industrial and Organizational Psychology, Orlando, FL.
- Bettenhausen KL, Fedor DB. (1997). Peer and upward appraisals: A comparison of their benefits and problems. *Group & Organization Management*, 22, 236–263.
- Birkeland SA, Borman WC, Brannick MT. (2003). *Using personal construct theory to investigate 360-degree source effects*. Presented at the 18th Annual Conference of the Society for Industrial and Organizational Psychology, Orlando, FL.
- Bonnett KL. (1995). *Transfer of training: The effect of accountability and the role of upward feedback*. Unpublished doctoral dissertation. California School of Professional Psychology.
- Brett JF, Atwater LA. (2001). 360-degree feedback: Accuracy, reactions and perceptions of usefulness. *Journal of Applied Psychology*, 86, 930–942.
- Brockner J, Higgins ET. (2001). Regulatory focus theory: Implications for the study of emotions at work. *Organizational Behavior and Human Performance*, 86, 35–66.

- *Brooks JH. (2000). *An examination of ratings congruence and organizational level on longitudinal multi-source feedback ratings*. Unpublished doctoral dissertation. De Paul University.
- Brutus S, Petosa S. (2002). *Rater selection in multi-source assessment: Evidence for the use of different strategies*. Molson School of Business, Concordia University, Montreal, Canada.
- Brutus S, London M, Martineau J. (1999). The impact of 360-degree feedback on planning for career development. *Journal of Management Development*, 18, 676–693.
- Bryk AS, Raudenbush SW. (2001). *Hierarchical linear models: Applications and data analysis methods*. Newbury Park, CA: Sage.
- Buckingham M, Clifton DO. (2001). *Now, discover your strengths*. New York: Free Press.
- Buda R, Reilly RR, Smither JW. (1991). The influence of indirect knowledge of prior performance on evaluations of present performance: The generalizability of assimilation effects. *Journal of Psychology and the Behavioral Sciences*, 6, 89–99.
- Cheung GW. (1999). Multifaceted conceptions of self-other ratings disagreement. *PERSONNEL PSYCHOLOGY*, 52, 1–36.
- Church AH. (1997). Do you see what I see? An exploration of congruence in ratings from multiple perspectives. *Journal of Applied Social Psychology*, 27, 983–1020.
- Church AH. (2000). Do higher performing managers actually receive better ratings? A validation of multirater assessment methodology. *Consulting Psychology Journal: Practice and Research*, 52, 99–116.
- Cohen J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.
- Conway RL. (1999). *The impact of coaching mid-level managers utilizing multi-rater feedback*. Doctoral dissertation. University of La Verne, La Verne, CA.
- Conway JM, Lombardo K, Sanders KC. (2001). A meta-analysis of incremental validity and nomological networks for subordinate and peer ratings. *Human Performance*, 14, 267–303.
- Day DV, Schleicher DJ, Unckless AL, Hiller NJ. (2002). Self-monitoring personality at work: A meta-analytic investigation of construct validity. *Journal of Applied Psychology*, 87, 390–341.
- Dominick PG, Reilly RR, Byrne JC. (2004, April). *Individual differences and peer feedback: Personality's impact on behavior change*. Paper presented at the 19th Annual Conference of the Society for Industrial and Organizational Psychology. Chicago, Illinois.
- Donovan JJ, Williams KJ. (2003). Missing the mark: Effects of time and causal attributions on goal revision in response to goal-performance discrepancies. *Journal of Applied Psychology*, 88, 379–390.
- Dweck CS, Leggett ELA. (1988). A social-cognitive approach to motivation and personality. *Psychological Review*, 95, 256–273.
- Erickson A, Allen T. (2003). *Linking 360° feedback to business outcome measures*. Presented at the 18th Annual Conference of the Society for Industrial and Organizational Psychology, Orlando, FL.
- Facteau JD, Craig SB. (2001). Are performance appraisal ratings from different rating sources comparable? *Journal of Applied Psychology*, 86, 215–227.
- Fletcher C, Baldry C. (2000). A study of individual differences and self-awareness in the context of multi-source feedback. *Journal of Occupational and Organizational Psychology*, 73, 303–319.
- Gangestad SW, Snyder M. (2000). Self-monitoring: Appraisal and reappraisal. *Psychological Bulletin*, 126, 530–555.

- Greguras GJ, Robie C, Schleicher DJ, Goff M III. (2003). A field study of the effects of rating purpose on the quality of multisource ratings. *PERSONNEL PSYCHOLOGY*, 56, 1–21.
- Hazucha JF, Hezlett SA, Schneider RJ. (1993). The impact of multisource feedback on management skills development. *Human Resource Management*, 32, 325–351.
- Hedges LV. (1982). Estimation of effect size from a series of independent experiments. *Psychological Bulletin*, 92, 490–499.
- *Hegarty WH. (1974). Using subordinate ratings to elicit behavioral changes in managers. *Journal of Applied Psychology*, 59, 764–766.
- Helland KR, Hoffman BC, Smith EM. (2003, April). *What does it all mean? Differences in multisource ratings*. Presented at the 18th Annual Conference of the Society for Industrial and Organizational Psychology, Orlando, FL.
- *Heslin PA, Latham GP. (2004). The effect of upward feedback on managerial behavior. *Applied Psychology: An International Review*, 53, 23–37.
- Heslin PA, Latham GP, VandeWalle D. (2004, April). *Implicit-person theory effects on employee coaching*. Paper presented at the 19th annual conference of the Society for Industrial and Organizational Psychology, Chicago.
- Heslin PA, Latham GP, VandeWalle D. (in press). The effect of implicit person theory on performance appraisals. *Journal of Applied Psychology*.
- *Hezlett SA, Ronnkvist AM. (1996). *The effects of multi-rater feedback on managers' skill development: Factors influencing behavior change*. Paper presented at the 11th Annual Conference of Society for Industrial and Organizational Psychology. San Diego, CA.
- Higgins ET. (1998). From expectancies to worldviews: Regulatory focus in socialization and cognition. In Darley JM, Cooper J (Eds.), *Attribution and social interaction: The legacy of Edward E. Jones* (pp. 243–269). Washington, DC: American Psychological Association.
- Higgins ET. (1999). Persons or situations: Unique explanatory principles or variability in general principles? In Cervone D, Shoda Y (Eds.), *The coherence of personality: Social-cognitive bases of consistency, variability, and organization* (pp. 61–93). New York: Guilford.
- *Hilario FL. (1998). *Effects of 360-degree feedback on managerial effectiveness ratings: A longitudinal field study*. Unpublished doctoral dissertation. New York University.
- Hunter JE, Schmidt FL. (1990). *Methods of meta-analysis: Correcting error and bias in research findings*. Newbury Park, CA: Sage.
- Ilgén DR, Fisher CD, Taylor MS. (1979). Consequences of individual feedback on behavior in organizations. *Journal of Applied Psychology*, 64, 349–371.
- *Ivanitskaya LV. (1999). *Cognitive and motivational influences on self-other congruence in multisource feedback*. Unpublished doctoral dissertation. Central Michigan University.
- *Johnson JW, Ferstl KL. (1999). The effects of interrater and self-other agreement on performance improvement following upward feedback. *PERSONNEL PSYCHOLOGY*, 52, 271–303.
- *Johnson KM, Johnson JW. (2001, April). *The influence of self-other agreement on performance improvement following feedback from multiple sources*. Paper presented at a symposium titled “If I could put time in a model: Understanding constructs longitudinally” (Chair: F.L. Oswald) at the 16th Annual Conference of the Society for Industrial and Organizational Psychology, San Diego, CA.
- Kluger AN, DeNisi A. (1996). The effects of feedback intervention on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119, 254–284.

- Leonard E, Williams JR. (2001). *An empirical examination of accountability perceptions within a multi-source feedback system*. Presented at the 16th Annual Conference of the Society for Industrial and Organizational Psychology, San Diego, CA.
- Locke EA, Latham GP. (1990). *A theory of goal setting and task performance*. Englewood Cliffs, NJ: Prentice Hall.
- London M, Smither JW. (1995). Can multi-source feedback change self-awareness and behavior? Theoretical applications and directions for research. *PERSONNEL PSYCHOLOGY*, *48*, 803–840.
- London M, Smither JW. (2002). Feedback orientation, feedback culture, and the longitudinal performance management process. *Human Resource Management Review*, *12*, 81–100.
- London M, Smither JW, Adsit DJ. (1997). Accountability: The Achilles heel of multisource feedback. *Group and Organization Management*, *22*, 162–184.
- Luthans F, Peterson SJ. (2003). 360-degree feedback with systematic coaching: Empirical analysis suggests a winning combination. *Human Resource Management*, *42*, 243–256.
- Maurer TJ, Palmer JK. (1999). Management development intentions following feedback: Role of perceived outcomes, social pressures, and control. *Journal of Management Development*, *18*, 733–751.
- Maurer TJ, Raju NS, Collins WC. (1998). Peer and subordinate performance appraisal measurement equivalence. *Journal of Applied Psychology*, *83*, 693–702.
- McEvoy GM. (1990). Public sector managers' reactions to appraisals by subordinates. *Public Personnel Management*, *19*, 201–212.
- Meyer JP, Allen NJ. (1997). *Commitment in the workplace*. Thousand Oaks, CA: Sage.
- Miller JS, Cardy RL. (2000). Self-monitoring and performance appraisal: Rating outcomes in project teams. *Journal of Organizational Behavior*, *21*, 609–626.
- Morris SB, DeShon RP. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Journal of Applied Psychology*, *7*, 105–125.
- Mount MK, Judge TA, Scullen SE, Sytsma MR, Hezlett SA. (1998). Trait, rater, and level effects in 360-degree performance ratings. *PERSONNEL PSYCHOLOGY*, *51*, 557–569.
- *Nemeroff WF, Cosentino J. (1979). Utilizing feedback and goal setting to increase performance appraisal interview skills of managers. *Academy of Management Journal*, *22*, 566–576.
- Nilsen D, Campbell D. (1993). Self-observer rating discrepancies: Once an overrater, always an overrater? *Human Resource Management*, *32*, 265–281.
- Noe RA. (2005). *Employee training and development*. Boston: McGraw Hill Irwin.
- Ostroff C, Atwater LE, Feinberg BJ. (2004). Understanding self–other agreement: A look at rater characteristics, context, and outcomes. *PERSONNEL PSYCHOLOGY*, *57*, 333–375.
- Peterson DB. (1993). *Measuring change: A psychometric approach to evaluating individual coaching outcomes*. Presented at the 8th Annual Conference of the Society for Industrial and Organizational Psychology, San Francisco, CA.
- Phillips JM, Hollenbeck JR, Ilgen DR. (1996). The prevalence and prediction of positive discrepancy creation: Examining a discrepancy between two self-regulation theories. *Journal of Applied Psychology*, *81*, 498–511.
- Pollack DM, Pollack LJ. (1996). Using 360° feedback in performance appraisal. *Public Personnel Management*, *25*, 507–528.
- *Quadracci RH. (1995). *The effects of 360 degree feedback, method of feedback and criticality of skill sets on changes in self-perception*. Unpublished doctoral dissertation. California School of Professional Psychology—Los Angeles.

- Reilly RR, Smither JW, Vasilopoulos NL. (1996). A longitudinal study of upward feedback. *PERSONNEL PSYCHOLOGY*, 49, 599–612.
- Roney CJR, Higgins ET, Shah J. (1995). Goals and framing: How outcome focus influences motivation and emotion. *Personality and Social Psychology Bulletin*, 21, 1151–1160.
- *Rosti RT, Shipper F. (1998). A study of the impact of training in a management development program based on 360-degree feedback. *Journal of Managerial Psychology*, 13, 77–89.
- Rothstein HR. (1990). Interrater reliability of job performance ratings: Growth to asymptote level with increasing opportunity to observe. *Journal of Applied Psychology*, 75, 322–327.
- Rutkowski KA, Steelman LA, Griffith RL. (2004, April). Paper presented at the 19th Annual Conference of the Society of Industrial and Organizational Psychology, Chicago.
- Schmidt FL, Hunter JE. (1996). Measurement error in psychological research: Lessons from 26 research scenarios. *Psychological Methods*, 1, 199–223.
- Schmidt FL, Viswesvaran C, Ones DS. (2000). Reliability is not validity and validity is not reliability. *PERSONNEL PSYCHOLOGY*, 53, 901–912.
- Scullen SE, Mount MK, Judge TA. (2003). Evidence of the construct validity of developmental ratings of managerial performance. *Journal of Applied Psychology*, 88, 50–66.
- *Seifert CF, Yukl G, McDonald RA. (2003). Effects of multisource feedback and a feedback facilitator on the influence behavior of managers toward subordinates. *Journal of Applied Psychology*, 88, 561–569.
- Smith EM, Helland KR. (2002). *Self-other rating agreement in multi-source feedback: Antecedents and correlates*. Paper presented at the 17th Annual Conference of the Society for Industrial and Organizational Psychology, Toronto, Canada.
- *Smither JW, London M, Flautt R, Vargas Y, Kucine I. (2002, April). *Does discussing multi-source feedback with raters enhance performance improvement?* Paper presented at the 17th Annual Conference of the Society for Industrial and Organizational Psychology, Toronto.
- Smither JW, London M, Flautt R, Vargas Y, Kucine I. (2003). Can executive coaches enhance the impact of multisource feedback on behavior change? A quasi-experimental field study. *PERSONNEL PSYCHOLOGY*, 56, 23–44.
- Smither JW, London M, Reilly RR, Flautt R, Vargas Y, Kucine I. (2004). Discussing multi-source feedback with raters and performance improvement. *Journal of Management Development*, 23.
- Smither JW, London M, Richmond KR (in press). Relationships between leaders' personality and reactions to, and use of, multisource feedback: A longitudinal study. *Group and Organization Management*.
- *Smither JW, London M, Vasilopoulos NL, Reilly RR, Millsap RE, Salvemini N. (1995). An examination of the effects of an upward feedback program over time. *PERSONNEL PSYCHOLOGY*, 48, 1–34.
- Smither JW, Reilly RR, Buda R. (1988). The effect of prior performance information on ratings of present performance: Contrast versus assimilation revisited. *Journal of Applied Psychology*, 73, 487–496.
- Smither JW, Walker AG. (2001). Measuring the impact of multisource feedback. In Bracken D, Timmreck C, Church A (Eds.), *Handbook of multisource feedback*. San Francisco: Jossey-Bass.
- Smither JW, Walker AG. (2004). Are the characteristics of narrative comments related to improvement in multirater feedback ratings over time? *Journal of Applied Psychology*.

- Steelman LA, Levy PE. (2001). *The feedback environment and its potential role in 360-degree feedback*. Paper presented at the 16th Annual Conference of the Society for Industrial and Organizational Psychology, San Diego, CA.
- Steelman LA, Levy PE, Snell AF. (2002). *The feedback environment scale (FES): Construct definition, measurement and validation*. The University of Akron. Unpublished manuscript.
- Thorsteinson TJ, Highhouse S. (2003). Effects of goal framing in job advertisements on organizational attractiveness. *Journal of Applied Social Psychology, 33*, 2393–2412.
- VandeWall D, Cron W, Slocum J. (2001). The role of goal orientation following performance feedback. *Journal of Applied Psychology, 86*, 629–640.
- Van-Dijk D, Kluger AN. (2004). Feedback sign effect on motivation: Is it moderated by regulatory focus? *Applied Psychology: An International Review, 53*, 113–135.
- Vicere AA, Fulmer RM. (1998). *Leadership by design: How benchmark companies sustain success through investment in continuous learning*. Boston: Harvard Business Review.
- *Waldman DA, Atwater LE. (2001). Attitudinal and behavioral outcomes of an upward feedback process. *Group and Organization Management, 26*, 189–205.
- *Walker AG, Smither JW. (1999). A five-year study of upward feedback: What managers do with their results matters. *PERSONNEL PSYCHOLOGY, 52*, 393–423.
- Wanous JP, Reichers AE, Austin JT. (2000). Cynicism about organizational change: Measurement, antecedents and correlates. *Group & Organization Management, 25*, 132–153.
- Warech MA, Smither JW, Reilly RR, Millsap RE, Reilly SP. (1998). Self-monitoring and 360-degree ratings. *Leadership Quarterly, 9*, 449–473.
- Williams KJ, Donovan JJ, Dodge TL. (2000). Self-regulation of performance: Goal establishment and goal revision processes in athletes. *Human Performance, 13*, 159–180.
- Yammarino FJ, Atwater LE. (1993). Understanding self-perception accuracy: Implications for human resources management. *Human Resource Management, 32*, 231–247.