

CAN WORKING WITH AN EXECUTIVE COACH IMPROVE MULTISOURCE FEEDBACK RATINGS OVER TIME? A QUASI-EXPERIMENTAL FIELD STUDY

JAMES W. SMITHER
Management Department
La Salle University

MANUEL LONDON
Harriman School of Management and Policy
State University of New York–Stony Brook

RAYMOND FLAUTI, YVETTE VARGAS, IVY KUCINE
JP Morgan Chase

This study examined the effects of executive coaching on multisource feedback over time. Participants were 1,361 senior managers who received multisource feedback; 404 of these senior managers worked with an executive coach (EC) to review their feedback and set goals. One year later, 1,202 senior managers (88% of the original sample) received multisource feedback from another survey. Managers who worked with an EC were more likely than other managers to set specific (rather than vague) goals ($d = .16$) and to solicit ideas for improvement from their supervisors ($d = .36$). Managers who worked with an EC improved more than other managers in terms of direct report and supervisor ratings, however, the effect size ($d = .17$) was small.

This study examines two interventions that have become increasingly widespread in leadership development—multisource feedback and executive coaching. Although multisource feedback had its origins many years earlier (Klimoski & London, 1974), its popularity has soared in recent years (Hedge, Borman, & Birkeland, 2001) and there have been an increasing number of books (e.g., Bracken, Timmreck, & Church, 2001; Tornow & London, 1998; Waldman & Atwater, 1998) and articles that address the process. The use of external executive coaches is also an increasingly popular trend in corporations (Graddick & Lane, 1998; Smither & Reilly, 2001). Despite the growing popularity of these interventions in applied settings, they have received relatively little attention in empirical research studies. Only a modest number of longitudinal

The authors thank Dick Reilly and three anonymous reviewers for their helpful comments concerning an earlier draft of this paper.

Correspondence and requests for reprints should be addressed to James W. Smither, Management Department, La Salle University, Philadelphia, PA 19141; Smither@lasalle.edu.

studies have examined the impact of multisource or upward feedback on performance, and longitudinal research on the impact of executive coaching is scarce.

This field study extends the current literature on multisource feedback and executive coaching in several ways. The data represent the first test of whether working with an external executive coach can improve multisource feedback ratings over time. We also examine whether working with an executive coach affects the goals that feedback recipients set and the likelihood that feedback recipients will discuss their feedback with raters to solicit ideas for improvement. In addition, we examine whether feedback recipients' reactions to executive coaching are related to improvement in multisource feedback ratings over time. These data are important from a practical perspective because, as Hall, Otazo, and Hollenbeck (1999), Hollenbeck and McCall (1999), and Harris (1999) have noted, organizations that are implementing coaching on a large scale are becoming concerned about cost. If the impact of executive coaching cannot be demonstrated, it could be viewed by some organizations as too expensive and might be eliminated when budgets get tight. These data also contribute to the emerging literature on factors that might enhance the impact of multisource feedback in organizations.

Dalessio (1998) defines multisource feedback as evaluations gathered about a target participant from two or more rating sources, which can include: self, supervisor, peers, direct reports, internal customers, external customers, and vendors/suppliers. Executive coaching can take a number of different forms. Some executives use coaching to learn specific skills, others to improve performance on the job or to prepare for advancement in business or professional life, and still others to support broader purposes, such as an executive's agenda for major organizational change (Witherspoon & White, 1996). Thus, executive coaching may involve a small number of meetings (as in the current study) where the coach and a senior-level manager are focused on a relatively narrow goal, or it may involve a longer relationship with multiple and complex goals. Hall et al. (1999) state that executive coaching is "a practical, goal-focused form of personal one-to-one learning for busy executives. It may be used to improve performance, to improve or develop executive behaviors, to work through organizational issues, to enhance a career, or to prevent derailment." (p. 40)

The Impact of Multisource and Upward Feedback on Behavior Change

Although there has been considerable research on immediate reactions to multisource feedback as well as to the level of agreement within- and between-rater sources (Mount & Scullen, 2001; Murphy, Cleveland,

& Mohler, 2001), less attention has been directed to the impact of multi-source feedback over time. Smither, London, Flautt, Vargas, and Kucine (2002) summarized 13 longitudinal studies that examined whether performance (as measured by subsequent feedback scores) improved following multisource or upward feedback. They 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 or upward feedback. Of the three studies that included a control group (Atwater, Waldman, Atwater, & Cartier, 2000; Hegarty, 1974; Heslin & Latham, 2001), two found that managers who received feedback improved more than managers who did not receive feedback. Smither et al. (2002) concluded that managers often (but not always) improve their performance (at least as reflected by subsequent feedback) after receiving multisource or upward feedback, and improvement is greatest among managers who initially receive the most negative feedback or who initially overrate themselves.

The Emerging Literature on Executive Coaching

Many recent articles approach the topic of executive coaching from the perspective of counseling or clinical psychology (e.g., multimodal therapy, psychodynamic theory or psychotherapy, eye movement desensitization and reprocessing), and rely on case studies or vignettes as illustrations or sources of evidence (Foster & Lendl, 1996; Kilburg, 1996, 1997; Laske, 1999; Levinson, 1996; Richard, 1999; Rotenberg, 2000; Sperry, 1993). Hollenbeck and McCall (1999) described a follow-up study of a coaching-based executive development process where participants rated the process as very valuable (3.95 out of 5.00) and reported that they changed (4.07 out of 5.00; Edelman & Armstrong, 1993). Hall et al. (1999) asked executives to rate the overall effectiveness of their experiences with coaching and concluded that "the positive image of coaching that is presented in the business media is supported by the experiences of the people we interviewed."

Olivero, Bane, and Kopelman (1997) examined the effects of executive coaching in a public sector agency where managers participated in a 3-day management development program and then worked with an internal executive coach for 8 weeks. The authors found that both the management development program and coaching increased productivity with executive coaching resulting in a significantly greater gain compared to the management development program alone. McGovern et al. (2001) examined the impact of executive coaching on 100 executives from 56 organizations. Coaching programs generally ranged from 6 to 12 months in duration. Based on interviews, they found that 86% of par-

ticipants and 74% of stakeholders (immediate supervisors or HR representatives) indicated they were very satisfied or extremely satisfied with the coaching process. Participants estimated that the return on coaching was nearly 5.7 times the investment in coaching. However, these results relied on executives' estimates of impact, as contrasted with input from other stakeholders.

In sum, preliminary evidence indicates that managers generally have a favorable reaction to executive coaching, but very little attention has been directed to examining the impact of executive coaching on behavior change and performance improvement. It is especially noteworthy that none of the studies cited above compared managers who received executive coaching with managers who did not receive executive coaching.

How Executive Coaches Might Enhance the Impact of Multisource Feedback

There are several ways that executive coaches may enhance the impact of multisource feedback on behavior change. For example, a coach can help feedback recipients acquire new skills (e.g., offering useful suggestions about how feedback recipients can better manage subordinates with performance problems). In addition, discussing the feedback with an executive coach may increase the feedback recipient's sense of accountability to use the feedback to guide behavior change (rather than ignoring the feedback; London, Smither, & Adsit, 1997). Moreover, executive coaches can help feedback recipients navigate through the stages of change (Dalton & Hollenbeck, 2001; Prochaska, Norcross, & DiClemente, 1995).

In the current study, we also examine two specific ways that executive coaches could affect the impact of multisource feedback. First, executive coaches can help feedback recipients set appropriate goals based on the feedback. 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. To the extent that executive coaches encourage and help feedback recipients to set specific goals, they should enhance the impact of the feedback on behavior change. Second, executive coaches can encourage feedback recipients to share their feedback with and solicit suggestions for improvement from raters. Walker and Smither (1999) found that feedback recipients who met with direct reports to discuss their upward feedback improved more than other feedback recipients, and feedback recipients 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. In the current study, we examine whether executive coaching

affects the extent to which feedback recipients set specific goals and discuss their feedback with raters. We also examine whether goal specificity and discussing feedback with raters partially mediate the relationship between executive coaching and performance improvement.

Summary and Hypotheses

Figure 1 presents a summary of the conceptual model tested here. First, we hypothesize that feedback recipients who work with an executive coach will be more likely than other feedback recipients to set specific (rather than vague or general) goals. Second, we hypothesize that feedback recipients who work with an executive coach will be more likely than other feedback recipients to share their feedback with raters and solicit suggestions for improvement. Third, we hypothesize that executive coaching will be positively related to improvement in multisource ratings and, consistent with the feedback and goal setting literature (Locke & Latham, 1990; London & Smither, 1995; Walker & Smither, 1999), goal specificity and sharing feedback/soliciting suggestions from raters will partially mediate this relationship. Fourth, we hypothesize that the reactions of feedback recipients to the executive coach and the coaching process will be positively related to improvements in multisource ratings.

Method

Overview

The current study used a quasi-experimental pre-post control group (executive coaching vs. no executive coaching) design. The key dependent variable was improvement in multisource ratings over a 1-year period. We also examined goal specificity and sharing feedback/soliciting suggestions for improvement from raters. The subjects were 1,361 senior managers in a large, global corporation who received multisource feedback in the autumn of 1999 (as part of a broader company-wide multisource feedback program). The senior manager's supervisor also received a copy of the feedback report and was allowed to use the results as input to the formal appraisal process. (Raters were aware of this feature in the feedback process.) After receiving their feedback, 404 (29.7%) of these senior managers worked with an external executive coach (EC). After working with the executive coach, 286 (70.8%) of these 404 senior managers responded to a brief online survey that gathered their reactions to the executive coach and the coaching process.

In July 2000, the company administered a brief survey where raters evaluated the extent of progress made toward individual goals set by each manager based on the initial multisource feedback. At the same

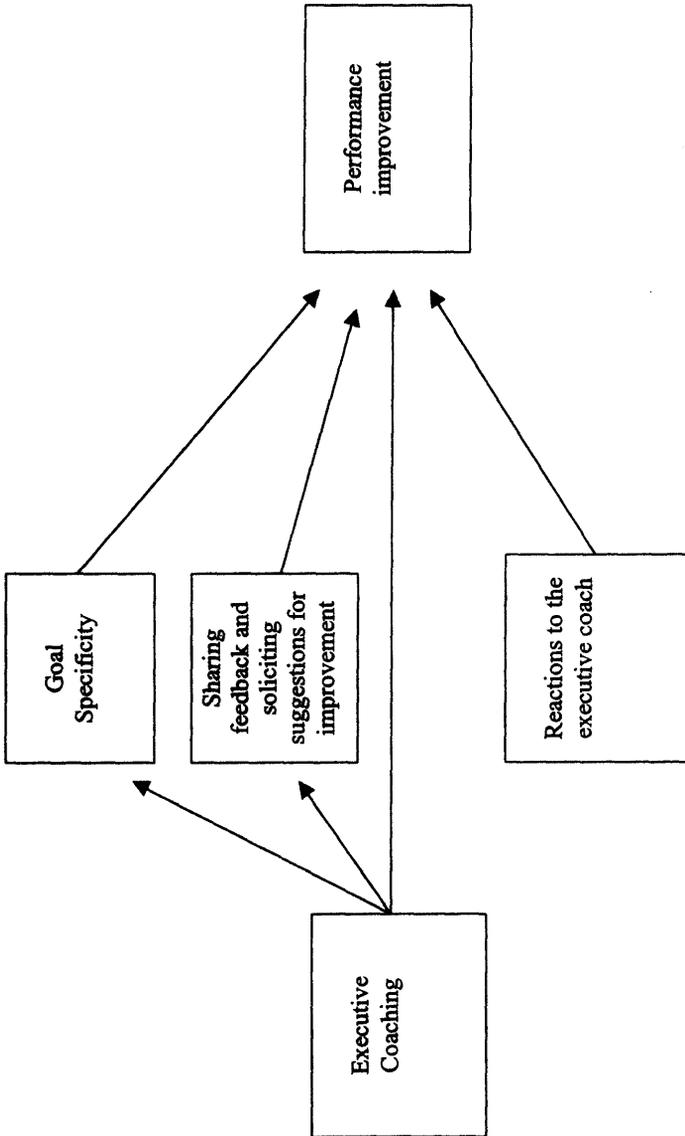


Figure 1: Summary of Conceptual Model of the Relationship Between Coaching and Performance Improvement

time, raters indicated whether the manager had shared his or her feedback and solicited suggestions for improvement from the rater. To participate in the survey, each manager was first required to create one to three development goals based on the initial multisource feedback. Of the 1,361 senior managers who received multisource feedback in the autumn of 1999, 1,229 (90.3%) participated in the survey. Of the 404 senior managers who worked with an executive coach, 400 (99%) participated in the survey. Of the 957 senior managers who did not work with an executive coach, 829 (86.6%) participated in the survey.

In the autumn of 2000, there was another administration of the company-wide multisource feedback program. Of the 1,361 senior managers who received multisource feedback in the autumn of 1999, 1,202 (88.3%) received feedback from the autumn 2000 multisource feedback survey. Of those who worked with an executive coach, 94.6% received feedback from the autumn 2000 multisource feedback survey; of those who did not work with an executive coach, 85.7% received feedback from the autumn 2000 multisource feedback survey ($\chi^2 = 21.66$, $df = 1$, $p < .01$).

The Executive Coaching Program

In the autumn of 1999, the company allowed the leader in each line of business to decide whether to require all senior managers in the line of business to use an EC. Some of these leaders required all senior managers to work with an EC, some declined to participate, and some made the process optional (that is, individual senior managers could decide whether they wanted to work with an EC). In most instances, if the leader of the line of business worked with an EC, their subordinate senior managers also worked with ECs. Of the 286 senior managers who worked with an EC and completed the online survey, 7.7% indicated that they were hesitant to work with a coach but it was a requirement in their business unit, 79.6% indicated that they welcomed the opportunity to work with a coach and it was a requirement in their business unit, and 12.7% indicated that they requested a coach but it was not a requirement in their business unit.

Those lines of business that elected to participate in the EC process agreed to pay for 5 hours (and in a few instances 7 hours) of EC time for all senior managers. For most ECs, the 5 hours were allocated to preparation time (reviewing the senior manager's multisource feedback report and other background material) and two or three in-person meetings with the senior manager.

Identifying, screening, and matching ECs. Several years earlier, the company identified a small number of ECs who had previously worked

with leaders of the business. These ECs provided names of other potential ECs and still other potential ECs were identified via word of mouth. Potential ECs were asked to describe their philosophy and approach to coaching and specific examples where they had worked as an EC. In addition, the corporate human resource department checked references provided by each EC. The majority of ECs had a master's degree with a smaller percentage having a doctorate. Most had backgrounds in organization development, psychology, human resources, or relevant industry experience.

Orientation and instructions for ECs. The company sponsored a half-day orientation session for all ECs. The company also subsequently provided forums where ECs could talk to each other in order to leverage the experience, knowledge, and successes acquired by individual ECs. ECs were told that their goal was to help the senior manager interpret his or her feedback results, link the feedback to the senior manager's business plan or situation, offer suggestions to help the senior manager create a development plan, help the senior manager identify ways to share feedback with his or her raters and solicit ideas for improvement, and encourage the senior manager to coach and give feedback to others.

Matching senior managers to ECs. ECs were invited to a meeting with the senior HR and development leaders for a specific business unit. Based on these meetings and a review of ECs' biographical sketches, ECs were matched to senior managers by the senior HR and development leaders in the line of business. The number of senior managers with whom each EC worked ranged from 1 to 36 ($M = 5.47$, $SD = 6.21$); the distribution was positively skewed (skewness index = 2.51) with a median of 3.0.

Interim Survey

This survey was administered in July 2000. To create the survey, each feedback recipient (ratee) was first asked to list up to three development goals. The mean number of goals set per ratee was 2.59 ($SD = .59$). Raters then evaluated the short-term progress of ratees toward these goals. As part of the survey, raters responded to two other items: "To what extent did this person share their feedback with you?" and "To what extent did this person solicit suggestions for how to improve?" Each of these items was rated on a 5-point scale ranging from 1 = *to a great extent* to 5 = *not at all*. For our analyses, we reversed this scale so that 5 = *to a great extent* and 1 = *not at all*. Across all raters, the correlation between sharing feedback and asking for suggestions for how to improve was .87 ($p < .01$). We therefore created a 2-item scale based on the mean of these two items ($\alpha = .93$).

Autumn 1999 and Autumn 2000 Multisource Feedback Surveys

All ratings were completed on a 5-point rating scale where 1 = *outstanding* and 5 = *unsatisfactory*. For purposes of all analyses, this rating scale was reversed (so that 5 = *outstanding* and 1 = *unsatisfactory*). In the autumn of 1999, the number of items on the survey varied depending on the type of rater (see Appendix). In the autumn of 2000, all rater types completed ratings on 20 items and also provided an overall rating (see Appendix). Because we were interested in comparing autumn 1999 ratings to ratings collected in the autumn of 2000, we focused on the subset of items for each rater group from the autumn 1999 survey that were also included on the autumn 2000 survey. For example, there were eight identical items that peers rated on both surveys. The alpha coefficients for peer ratings on these eight items in 1999 and 2000 were .89 and .89, respectively. The average of peer ratings on these eight items in 1999 was highly correlated with the average of peer ratings on the other survey items that were rated by peers in 1999 ($r = .91$). The average of peer ratings on these eight items in 2000 was highly correlated with the average of peer ratings on the other survey items that were rated by peers in 2000 ($r = .87$). Based on the above, we used these eight items when comparing peer ratings in 1999 and 2000. We conducted parallel analyses to those described above for self, supervisor, and direct report ratings (see summary in Table 1). For each rater source, the common item set (i.e., items rated on both surveys) was highly correlated with other items on the surveys and internally consistent (see Table 1). All subsequent analyses are therefore based only on those items that each rater source rated in both the autumn 1999 and autumn 2000 surveys. Our dataset therefore contained the mean ratings from peers, supervisors, and direct reports in autumn 1999 and autumn 2000 where each of these mean ratings was based on the subset of items described above.

The Coaching Effectiveness Survey

As noted above, 286 of the 404 senior managers who worked with an EC responded to a brief online survey that gathered their reactions to the EC and the coaching process. Respondents were asked "How many face-to-face/phone conversations did you have with your external coach?" Responses were coded 1, 2, or 3 or more. Over half (55.0%) of the senior managers reported having three or more conversations with the coach, 29.4% had two conversations with the coach, and 15.6% reported having only one conversation with the coach. Respondents were also asked, "For the next 360-degree feedback process, do you want to work with a coach again?" We coded responses 1 = *no* and 2 = *yes*. Ratees

TABLE 1
Summary of Items Rated by Each Rating Source in Autumn 1999 and Autumn 2000

Rater source	Number of items rated on fall 1999 survey	Number of items rated on fall 2000 survey	Number of common items rated in both fall 1999 and fall 2000	Correlation of common items with other items rated in fall 1999 ¹	Correlation of common items with other items rated in fall 2000 ²	Alpha coefficient of common items in fall 1999	Alpha coefficient of common items in fall 2000
Supervisor	29	21	11	.91	.88	.86	.86
Direct reports	28	21	9	.93	.90	.91	.90
Peers	22	21	8	.91	.87	.89	.89

¹ This column contains correlations between (a) items in the common item set (those items rated in both fall 1999 and fall 2000) and (b) those items rated in fall 1999 that were not part of the common item set.

² This column contains correlations between (a) items in the common item set (those items rated in both fall 1999 and fall 2000) and (b) those items rated in fall 2000 that were not part of the common item set.

who answered "yes" to the this question were asked, "Would you like to work with someone else or the same coach?" We coded 1 = *someone else* and 2 = *the same coach*.

Respondents were asked to rate the effectiveness of their coach on six items: Helping you interpret your feedback results by asking questions to uncover reasons for the feedback; helping you link your feedback to your business plan/situation; offering you useful suggestions, advice, or insights to set goals for development; helping you identify ways to share feedback with your raters and to solicit ideas for improvement; encouraging you to coach and give feedback to others; and, overall, contributing to your job performance and career development. Each of these items was rated on a 5-point scale where 1 = *very effective* and 5 = *very ineffective*. For our analyses, we reversed this scale (i.e., higher numbers indicated more effective coaching). We conducted a factor analysis on these six items using principal axis extraction. The first eigenvalue was 4.45 and no other eigenvalue exceeded 1.0, thereby suggesting that a single factor explained most of the common variance among the six items. We created a coaching effectiveness scale based on the mean of these six items ($\alpha = .93$) and used this as an overall measure of the coach's effectiveness (as perceived by the senior manager who received the coaching).

Results

Table 2 presents the correlations among the major variables. Ratings from different sources were modestly correlated on the autumn 1999 survey and on the autumn 2000 survey. Same-source ratings of autumn 1999 with autumn 2000 ratings were also moderately correlated. Senior managers who worked with an executive coach (EC) did not differ from senior managers who did not work with an EC in terms of their self-ratings, supervisor ratings, direct report ratings, or peer ratings (all $p > .50$), thereby indicating there were no preexisting differences between the two groups on the target behaviors.

On the coaching effectiveness survey, 86.3% of the senior managers indicated that they wanted to work with a coach again. Of these, 78.5% wanted to work with the same coach. The mean of the coaching effectiveness scale was 4.04 ($SD = .76$). Taken together, these findings indicate that senior managers generally had favorable reactions to their executive coach and the coaching process.

Hypothesis One

Using a 1 = *very general* to 4 = *very specific* rating scale, the first two authors independently rated each of the 2,927 developmental goals that

TABLE 2
Correlations Among Major Variables

Mean	SD	A	B	C	D	E	F	G	H	I	J	K	L	M	N
A. Self-ratings 1999	4.31	.44	-												
B. Direct report ratings 1999	4.28	.42	.30**	-											
C. Supervisor ratings 1999	4.29	.44	.20**	.26**	-										
D. Peer ratings 1999	4.24	.35	.30**	.39**	.37**	-									
E. Number of coaching conversations	2.39	.74	.05	-.02	.04	-.06	-								
F. Coaching effectiveness	4.04	.76	.06	.07	.04	-.01	.39**	-							
G. Goal specificity	2.34	.72	.00	-.03	.01	.00	-.10	-.06	-						
H. Shared feedback/solicited suggestions from direct reports	2.85	.88	-.02	.11**	-.05	-.04	-.04	-.07	.07*	-					
I. Shared feedback/solicited suggestions from supervisor	3.54	1.02	-.02	-.06	.05	-.01	.01	-.01	.01	.13**	-				
J. Shared feedback/solicited suggestions from peers	2.10	.69	-.03	-.06	-.02	.05	.09	.13*	-.03	.23**	.11**	-			
K. Self-ratings 2000	4.35	.41	.68**	.29*	.26**	.33**	.06	.12	.01	-.04	-.03	-.06	-		
L. Direct report ratings 2000	4.31	.39	.31**	.59*	.21**	.38**	-.06	.08	-.05	.15**	.02	-.04	.34**	-	
M. Supervisor ratings 2000	4.28	.42	.16**	.27**	.51**	.34**	.11	.04	.00	-.04	.11**	-.05	.25**	.27**	-
N. Peer ratings 2000	4.27	.34	.28**	.33**	.32**	.62**	-.09	.06	-.02	-.02	.00	.07*	.33**	.41**	.37**

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

senior managers wrote as part of their participation in the interim survey described above. At the time we completed our ratings, we did not know whether the senior manager had worked with an EC. For each goal, we then calculated the mean of our ratings (interrater reliability was .73). The mean rating of goal specificity was 2.34 ($SD = .84$, $n = 2,927$). For each senior manager, we then calculated the average goal specificity across goals. Consistent with Hypothesis 1, senior managers who worked with a coach ($M = 2.42$, $SD = .78$) set more specific goals ($t = 2.60$, $df = 1,124$, $p < .01$) than senior managers who did not work with a coach (2.30, $SD = .69$), however the effect size ($d = .16$, where d equals the mean difference divided by the pooled standard deviation, Cohen, 1988) was small.

Hypothesis Two

An omnibus test (multivariate analysis of variance, MANOVA) indicated that senior managers who worked with a coach differed from senior managers who did not work with a coach on the set of three dependent variables: sharing feedback and soliciting ideas for improvement from their supervisor, direct reports, and peers (Wilks $F = 10.82$; $df = 3, 926$; $p < .01$). The results of follow-up t -tests indicated that senior managers who worked with a coach ($M = 3.78$, $SD = 0.95$) were more likely than other senior managers ($M = 3.42$, $SD = 1.03$, $t = 5.60$, $df = 1,074$, $p < .01$, $d = .36$) to share their feedback and solicit ideas for improvement from their supervisors. The effect size was moderate (according to Cohen, 1988, small, medium, and large values of d are about .20, .50, and .80 respectively). However, senior managers who worked with a coach ($M = 2.91$, $SD = 0.84$) were not more likely than other senior managers ($M = 2.82$, $SD = 0.89$) to share their feedback and solicit ideas for improvement from direct reports ($t = 1.63$, $df = 1,081$, $p = .10$, $d = .10$) or peers ($M = 2.08$, $SD = 0.68$; $M = 2.10$, $SD = 0.70$; $t = 0.53$, $df = 1,188$, $p = .60$, $d = .03$).

Finally, coaching effectiveness was positively related to whether senior managers shared their feedback and solicited ideas for improvement from peers ($r = .13$, $p < .05$, see Table 2).

Hypothesis Three

Hypothesis 3 stated executive coaching will be positively related to improvement in multisource ratings, and goal specificity and sharing feedback/soliciting suggestions from raters will partially mediate this relationship. In each analysis described below, improvement in ratings is assessed by treating autumn 2000 ratings from a specific source as the

TABLE 3
Mean Ratings from Each Source for Senior Managers Who Were Coached Versus Not Coached

	Not coached		Coached	
	1999	2000	1999	2000
Direct reports	4.29 (.43)	4.30 (.41)	4.29 (.35)	4.34 (.33)
Supervisors	4.31 (.44)	4.26 (.43)	4.30 (.37)	4.31 (.37)
Peers	4.26 (.35)	4.26 (.35)	4.25 (.32)	4.26 (.32)

Note: Standard deviations are in parentheses.

criterion and autumn 1999 ratings from the same source as a covariate (Cronbach & Furby, 1970). To show mediation, four relationships must be demonstrated (Baron & Kenny, 1986). First, the relationship between the predictor (executive coaching) and criterion (improvement in multisource ratings) must be significant. Second, the relationship between the predictor and mediator variables (goal specificity and sharing feedback/soliciting suggestions for improvement) must be significant. Third, the relationship between the mediator variables and the criterion must be significant. Fourth, a regression equation is examined in which the predictor and mediator variables are used to predict the criterion. If the regression weight of the predictor decreases or becomes non-significant and the mediator remains significant, this serves as evidence of mediation.

Mediation Test 1. The mean autumn 1999 and autumn 2000 ratings for senior managers who worked with an executive coach and those who did not work with an executive coach are provided in Table 3. To determine whether executive coaching was related to improvement in ratings, we conducted hierarchical regression analyses with autumn 2000 ratings from a specific source as the criterion, autumn 1999 ratings from the same source as the covariate, and working with an EC (1 = no, 2 = yes) as the predictor. For direct report ratings, working with a coach was positively related to improvement from autumn 1999 to autumn 2000 (see Table 4). The interaction term (i.e., the cross product of coaching and autumn 1999 ratings) was not significant. To illustrate this finding we calculated an adjusted autumn 2000 direct report score (actual autumn 2000 direct report rating minus predicted autumn 2000 direct report rating based on the autumn 1999 direct report rating) and conducted a *t*-test that compared the adjusted autumn 2000 direct report scores of senior managers who were coached ($M = .03$, $SD = .27$) versus senior managers who were not coached ($M = -.02$, $SD = .33$, $t = 2.15$, $p < .05$). The effect size ($d = .17$) associated with executive coaching was positive and statistically significant but small (Cohen, 1988).

TABLE 4
Results of Hierarchical Regression Analyses

Dependent variable: Direct report ratings 2000						
Step	Variable entered	R^2	Increase in R^2	Standardized β	t	P
1	Direct report ratings 1999	.353	-	.59	23.35	.00
2	Coached (1 = no, 2 = yes)	.356	.003	.05	2.15	.03
3	Interaction	.356	.000	-.21	-.69	.49
Dependent variable: Supervisor ratings 2000						
Step	Variable entered	R^2	Increase in R^2	Standardized β	t	P
1	Supervisor ratings 1999	.264	-	.51	19.02	.00
2	Coached (1 = no, 2 = yes)	.268	.004	.07	2.48	.01
3	Interaction	.268	.000	.01	.04	.97
Dependent variable: Peer ratings 2000						
Step	Variable entered	R^2	Increase in R^2	Standardized β	t	P
1	Peer ratings 1999	.381	-	.62	26.70	.00
2	Coached (1 = no, 2 = yes)	.382	.001	.02	.64	.52
3	Interaction	.385	.003	-.75	-2.64	.01
Dependent variable: Supervisor ratings 2000						
Step	Variable entered	R^2	Increase in R^2	Standardized β	t	P
1	Supervisor ratings 1999	.288	-	.54	15.75	.00
2	Shared feedback/Solicited suggestions from supervisor	.295	.007	.09	3.11	.00
3	Interaction	.298	.003	.56	1.87	.06

For supervisor ratings, working with a coach was positively related to improvement from autumn 1999 to autumn 2000. The interaction term was not significant. A t -test that compared the adjusted autumn 2000 supervisor scores of senior managers who were coached ($M = .04$, $SD = .32$) versus senior managers who were not coached ($M = -.02$, $SD = .37$, $t = 2.48$, $p < .05$). The effect size ($d = .17$) associated with executive coaching was positive and statistically significant but small.

For peer ratings, the main effect of working with a coach was not significant. Although not hypothesized, the interaction term was significant however the effect size was small. To explore this interaction, we split senior managers into three groups based on their autumn 1999 peer ratings (bottom third, middle third, top third). For senior managers whose autumn 1999 peer ratings were in the bottom third, working with a coach was positively related to autumn 2000 ratings ($r = .11$, $p < .05$). For senior managers whose autumn 1999 peer ratings were in the middle third,

working with a coach was unrelated to autumn 2000 ratings ($r = .03$, $p = .57$). For senior managers whose autumn 1999 peer ratings were in the top third, working with a coach was negatively related to autumn 2000 ratings ($r = -.10$, $p < .05$).

In sum, the relationship between executive coaching and improvement in direct report and supervisor ratings was positive but small.

Mediation Test 2. As described above, executive coaching had a significant, albeit modest, relationship to goal specificity and sharing feedback/soliciting suggestions from supervisors (but not from direct reports and peers).

Mediation Test 3. We conducted hierarchical regression analyses with autumn 2000 ratings from a specific source as the criterion, autumn 1999 ratings from the same source as the covariate, and goal specificity as the predictor. Results showed that goal specificity was unrelated to improvement in direct report, supervisor, or peer ratings, thereby indicating that goal specificity does not mediate the relationship between executive coaching and improvement in ratings.

We also conducted hierarchical regression analyses with autumn 2000 ratings from a specific source as the criterion, autumn 1999 ratings from the same source as the covariate, and sharing feedback/soliciting suggestions from supervisors as the predictor. (Note that we did not conduct similar analyses for sharing feedback/soliciting suggestions from direct reports and peers because we determined at Step 2 of the mediation analysis that these variables cannot mediate the relationship between executive coaching and improvement in ratings.) Results (see Table 4) indicated that sharing feedback/soliciting suggestions from the supervisor was significantly and positively related to improvement in supervisor ratings, however, the variance explained was small.

Mediation Test 4. Using autumn 2000 ratings from the supervisor as the criterion and autumn 1999 ratings from supervisors as a covariate, the regression weight for coaching remained nearly unchanged when sharing feedback/soliciting suggestions was added to the regression equation (standardized beta = .067, $t = 2.48$, $p = .01$ vs. standardized beta = .061, $t = 2.17$, $p = .03$, respectively), thereby indicating that sharing feedback/soliciting suggestions from supervisors does not mediate the small relationship we observed between executive coaching and improvement in supervisor ratings.

Hypothesis Four

We examined this hypothesis by looking at whether improvement in ratings from autumn 1999 to autumn 2000 was related to the number of conversations senior managers had with the coach or to the senior

manager's ratings of the coach's effectiveness. The effectiveness of the coach as rated by the senior manager (entered on the second step of each analysis) was unrelated to improvement from autumn 1999 to autumn 2000 as rated by direct reports, supervisors, or peers. These results fail to support Hypothesis 4.

Discussion

There may be no simple answer to the broad question of whether executive coaching enhances performance. Instead, we believe that research should examine how the impact of executive coaching is shaped by a variety of factors such as its purpose (e.g., to facilitate use of multi-source feedback, to overcome behavior problems that limit career advancement, to support a transition to a new leadership assignment), length (several meetings over several weeks vs. many meetings over many months), organizational context, and individual differences among coaches and among those being coached. Ultimately, the goal is to identify situations where the use of executive coaching is likely to have a positive impact versus situations where executive coaching is likely to be less effective.

The current results are important because they provide the first data on the impact of working with an executive coach as judged by other sources (supervisors, direct reports, and peers). We found that managers who worked with an EC were more likely than other managers to set specific (rather than vague) goals ($d = .16$), solicit ideas for improvement from their supervisors ($d = .36$), and receive improved ratings ($d = .17$) from direct reports and supervisors. In addition, the small relationship between executive coaching and improvement in ratings was not mediated by goal specificity and sharing feedback/soliciting suggestions for improvement.

Practical Implications, Limitations, and Directions for Future Research

Practitioners may wonder whether the small, albeit positive, effect sizes observed in the current study are sufficient to justify the investment in executive coaching. Because the standard deviation of job performance in dollars (SD_y) is likely to be large for senior managers, even small improvements in performance may be associated with meaningful economic benefits. In the end, we think that judgments about the practical (e.g., economic) value of executive coaching must await further research. For example, the generalizability of the results described here might be limited because multisource feedback was shared with the feedback recipient's supervisor, who could use the information to influence

compensation, promotion, and so on. This may have increased recipients' sense of accountability to respond to the feedback even in the absence of executive coaching (London et al., 1997). When accountability to act on the feedback is already high, the incremental impact of working with an executive coach may be limited. Future research should examine the impact of executive coaches in an environment where accountability is relatively low (e.g., where multisource feedback is confidential).

Because research on executive coaching is in its infancy, there are many other areas where additional research would be desirable. For example, future research should examine the impact of longer term executive coaching relationships. Another promising avenue for future research would be to focus on individual differences in readiness for change (Prochaska, DiClemente, & Norcross, 1992; Prochaska et al., 1995) or receptivity to coaching and feedback (London & Smither, 2002).

In the current study, we know little about the nature or content of the executive coaching conversations. Future research could examine whether the effects of coaching are related to the coach's style. Moreover, we do not know how the background of executive coaches (e.g., counseling psychology, organization development) might affect the coaching process or the outcomes of coaching. Executive coaches with specific backgrounds (e.g., counseling psychology) may be well suited for some situations (e.g., helping a senior manager overcome an aggressive or demeaning interpersonal style) but not for others (e.g., helping a senior manager integrate organizational cultures during a merger or acquisition). In addition, the match or "chemistry" between the executive coach and the senior manager may be important. It would also be desirable for future research to examine the effects of executive coaching where participants are randomly assigned to conditions.

A final issue that will be important in future research is determining the appropriate criteria for assessing the impact of executive coaching. For example, in the current study, we examined changes in multisource ratings over time. In other settings (e.g., where coaching is focused on working with career transitions or on learning specific skills) different criteria will be required. It is also possible that multisource feedback ratings (collected before and after coaching) are simply too broad a measurement tool to detect the impact of executive coaching. Future research could examine more individualized criteria of behavior change (e.g., long-term measures of progress toward specific goals set by each feedback recipient). It would also be desirable to use retrospective pretests (Smither & Walker, 2001) to evaluate behavior change.

The process and outcomes of executive coaching are likely to be influenced by many variables not measured here. The popularity of executive coaching, coupled with the number of potentially relevant ex-

planatory variables, should make this an important area for future research. Such research can enhance our theoretical understanding of executive development, feedback processes, and behavior change while at the same time pointing to situations in which executive coaching can and cannot provide a valuable return on investment.

REFERENCES

- 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.
- Baron RM, Kenny DA. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51, 1173–1182.
- Bracken D, Timmreck C, Church A (Eds.). (2001). *Handbook of multisource feedback*. San Francisco: Jossey-Bass.
- Cohen J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.
- Cronbach LJ, Furby L. (1970). How should we measure “change”—or should we? *Psychological Bulletin*, 74, 68–80.
- Dallessio AT. (1998). Using multisource feedback for employee development and personnel decisions. In Smither JW (Ed.), *Performance appraisal: State-of-the-art in practice*. San Francisco: Jossey-Bass.
- Dalton MA, Hollenbeck GP. (2001). After feedback: How to facilitate change in behavior. In Bracken DW, Timmreck CW, Church AH (Eds.), *The handbook of multisource feedback* (pp. 352–367). San Francisco: Jossey-Bass.
- Edelstein BC, Armstrong DJ. (1993). A model for executive development. *Human Resource Planning*, 16(4), 51–64.
- Foster S, Lendl J. (1996). Eye movement desensitization and reprocessing: Four case studies of a new tool for executive coaching and restoring employee performance after setbacks. *Consulting Psychology Journal: Practice and Research*, 48, 155–161.
- Graddick MM, Lane P. (1998). Evaluating executive performance. In Smither JW (Ed.), *Performance appraisal: State-of-the-art in practice*. San Francisco: Jossey-Bass.
- Hall DT, Otazo KL, Hollenbeck GP. (1999). Behind closed doors: What really happens in executive coaching. *Organizational Dynamics*, 39–52.
- Harris M. (1999). Practice network: Look, it's an I-O psychologist... no, it's a trainer... no, it's an executive coach! *The Industrial-Organizational Psychologist*, 36(3), 38–42.
- Hedge JW, Borman WC, Birkeland SA. (2001). History and development of multisource feedback as a methodology. In Bracken D, Timmreck C, Church A (Eds.), *Handbook of multisource feedback*. San Francisco: Jossey-Bass.
- Hegarty WH. (1974). Using subordinate ratings to elicit behavioral changes in managers. *Journal of Applied Psychology*, 59, 764–766.
- Heslin PA, Latham GP. (2001, April). *The effect of upward feedback on managerial behavior*. Paper presented at the 16th Annual Conference of the Society for Industrial and Organizational Psychology, San Diego, CA.
- Hollenbeck GP, McCall MW. (1999). Leadership development: Contemporary practices. In Kraut AI, Korman AK (Eds.), *Evolving practices in human resource management*. San Francisco: Jossey-Bass.
- Kilburg RR. (1996). Toward a conceptual understanding and definition of executive coaching. *Consulting Psychology Journal: Practice and Research*, 48, 134–144.

- Kilburg RR. (1997). Coaching and executive character: Core problems and basic approaches. *Consulting Psychology Journal: Practice and Research*, 49, 281–299.
- Klimoski RJ, London M. (1974). Role of the rater in performance appraisal. *Journal of Applied Psychology*, 59, 445–451.
- Laske OE. (1999) An integrated model of developmental coaching. *Consulting Psychology Journal: Practice and Research*, 51, 139–159.
- Levinson H. (1996). Executive coaching. *Consulting Psychology Journal: Practice and Research*, 48, 115–123.
- Locke EA, Latham GP. (1990). *A theory of goal setting and task performance*. Englewood Cliffs, NJ: Prentice Hall.
- London M, Smither JW. (2002). Feedback orientation, feedback culture, and the longitudinal performance management process. *Human Resource Management Review*, 12(1), 81–101.
- London M, Smither JW. (1995). Can multisource feedback change perceptions of goal accomplishment, self-evaluations, and performance-related outcomes? Theory-based applications and directions for research. *PERSONNEL PSYCHOLOGY*, 48, 803–839.
- London M, Smither JW, Adsit DJ. (1997). Accountability: The Achilles heel of multisource feedback. *Group and Organization Management*, 22, 162–184.
- McGovern J, Lindemann M, Vergara M, Murphy S, Barker L, Warrenfeltz R. (2001). Maximizing the impact of executive coaching: Behavioral change, organizational outcomes, and return on investment. *The Manchester Review*, 6(1), 1–9.
- Mount MK, Scullen SE. (2001). Multisource feedback ratings: What do they really measure? In London M (Ed.), *How people evaluate others in organizations* (pp. 155–176). Mahwah, NJ: Erlbaum.
- Murphy KR, Cleveland JN, Mohler CJ. (2001). Reliability, validity, and meaningfulness of multisource ratings. In Bracken DW, Timmreck CW, Church AH (Eds.), *The handbook of multisource feedback: The comprehensive resource for designing and implementing MSF processes* (pp. 130–148). San Francisco: Jossey-Bass.
- Olivero G, Bane KD, Kopelman RE. (1997). Executive coaching as a transfer of training tool: Effects on productivity in a public agency. *Public Personnel Management*, 26, 461–469.
- Prochaska JO, DiClemente CC, Norcross JC. (1992). In search of how people change: Applications to addictive behaviors. *American Psychologist*, 47, 1102–1114.
- Prochaska JO, Norcross JC, DiClemente CC. (1995). *Changing for good*. New York: Avon Books.
- Richard JT. (1999). Multimodal therapy: A useful model for the executive coach. *Consulting Psychology Journal: Practice and Research*, 51, 24–30.
- Rotenberg CT. (2000). Psychodynamic psychotherapy and executive coaching—overlapping paradigms. *Journal of the American Academy of Psychoanalysis*, 28, 653–663.
- Smither JW, London M, Flautt R, Vargas Y, Kucine I. (2002, April). *Does discussing multisource feedback with raters enhance performance improvement?* Paper presented at the 17th Annual Conference of the Society for Industrial and Organizational Psychology, Toronto, ON.
- Smither JW, Reilly SP. (2001). Coaching in organizations: A social psychological perspective. In London M (Ed.), *How people evaluate others in organizations*. Mahwah, NJ: Erlbaum.
- 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.
- Sperry L. (1993). Working with executives: Consulting, counseling, and coaching. Individual Psychology: *Journal of Adlerian Theory, Research and Practice*, 49, 257–266.

- Tornow WW, London M (Eds.). (1998). *Maximizing the value of 360-degree feedback: A process for successful individual and organizational development*. San Francisco: Jossey-Bass.
- Waldman DA, Atwater LE. (1998). *The power of 360-degree feedback: How to leverage performance evaluations for top productivity*. Houston, TX: Gulf.
- Walker AG, Smither JW. (1999). A five-year study of upward feedback: What managers do with their results matters. *PERSONNEL PSYCHOLOGY*, 52, 393–423.
- Witherspoon R, White RP. (1996). Executive coaching: A continuum of roles. *Consulting Psychology Journal: Practice & Research*, 48, 124–133.

APPENDIX

Autumn 1999 Items

- Partnership/Teamwork (M, P, D)
- Client skills/Customer focus (M, P, D)
- Technical skills (M, P, D)
- Responsiveness/Dependability (M, P, D)
- Judgment/Decision making (M, P, D)
- Management (M, P, D)
- Leadership (M, P, D)
- Quickly assesses the “big picture” in complex situations and identifies what is critical. (M)
- Recognizes patterns and connections in information from different sources and their business implications. (M)
- Creates simple, compelling messages and few key priorities that guide and focus the efforts of others. (D)
- Simplifies complex projects or situations by focusing on key issues, activities and goals. (M, D)
- Delegates detailed oversight and responsibility to those with the necessary skills and information. (D)
- Provides clear goals, written performance appraisals and follow-up discussions annually. (D)
- Provides coaching and feedback to improve performance. (P, D)
- Negotiates realistic resources to achieve results. (M)
- Respectfully confronts problematic behavior. (M, P, D)
- Encourages and is open to feedback and coaching from others. (M, P, D)
- Stands firm in the face of opposition/disagreement from influential others when appropriate. (M, P, D)
- Makes tough choices and decisions in a timely fashion. (M, D)
- Invests time and resources to enhance the effectiveness of management team. (M, D)
- Responds to others’ needs to balance personal and work demands. (M, D)
- Seeks out and listens to customers’ and colleagues’ views to establish their concerns. (M, P, D)
- Involves those who are directly affected by decisions in the decision-making process. (D)
- Gains cooperation by explicitly addressing others’ interests and concerns. (P, D)
- Accurately assesses the impact of own behavior and decisions on others. (M, P, D)
- Accurately identifies own strengths and weaknesses and works to overcome weaknesses. (M, P, D)
- Treats people respectfully regardless of personal views, disagreements, or level. (M, P, D)
- Takes calculated risks needed to achieve results. (M)

APPENDIX (continued)

Autumn 1999 Items (continued)

- Demonstrates technical expertise to resolve business issues. (M, P)
 - Maintains composure and positive attitude during stressful situations. (M, P)
 - Proactively seeks new experiences and knowledge. (M)
 - Quickly adjusts in response to changing situations. (M, P, D)
 - Looks for ways to do things better, faster and more cost efficient. (M, P, D)
 - Shows by his or her actions a strong commitment to diversity. (M, P, D)
 - Overall rating (M, P, D)
- (M = Rated by supervisors, P = Rated by peers, D = Rated by direct reports)

Autumn 2000 Items

- Demonstrates technical expertise to resolve business issues.
 - Gains cooperation by explicitly addressing others' interests and concerns.
 - Seeks out and listens to customers' and peers' views to establish their concerns.
 - Sets a strong personal example by consistently and visibly demonstrating the Chase Values.
 - Provides clear goals, written performance appraisals, coaching and feedback.
 - Takes calculated risks needed to achieve results.
 - Maintains composure and positive attitude during stressful situations.
 - Takes a view on issues and constructively stands behind them.
 - Makes tough choices and decisions in a timely fashion.
 - Builds strong partnerships with colleagues and customers.
 - Shows by his or her actions a strong commitment to diversity.
 - Looks for ways to do things better, faster, and more cost efficiently.
 - Simplifies complex projects or situations by focusing on key issues, activities and goals.
 - Treats people respectfully regardless of personal views, disagreements, or level.
 - Responds to others' needs to balance personal and work demands.
 - Demonstrates passion and commitment in order to continually drive change.
 - Uses vision and creativity to develop innovative approaches to business challenges.
 - Takes the lead in understanding and applying internet thinking.
 - Consistently takes initiative to make things happen.
 - Client driven in all interactions with customers and colleagues.
 - Please provide a rating based on your general impression of XXX: Overall rating (1-5)
-

Copyright of Personnel Psychology is the property of Blackwell Publishing Limited and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.