

Running Head: EFFECTIVENESS OF MULTI-SOURCE FEEDBACK

A Field Study on the Impact of Multi-source Feedback
on Leadership Competency Development

Clarke Peterson
Atlanta Leadership Consulting, LLC

Guangrong Dai Kenneth P. De Meuse
Lominger International—A Korn/Ferry Company

Abstract

This longitudinal study examined the effectiveness of a developmental multi-source feedback program implemented in a large financial company in the U.S.A. It was found that managers improved more on the leadership competencies they selected for development than those they did not select and the overall composite score. The improvement was more than explained for by the regression to mean, and was sustained across multiple administrations of the multi-source feedback. It was further found that the improvement was correlated to the development difficulty of the leadership competency in that managers improved more on easier competencies than on harder competencies.

Multi-source feedback interventions have become increasingly widespread in many organizations around the world. It has been reported that nearly 90% of Fortune 1000 firms use some form of multi-source assessment (Atwater & Waldman, 1998). The popularity of such a human resource practice has stimulated much research enthusiasm in the academic field. According to Morgeson, Mumford, and Campion (2005), more than 100 multi-source feedback related articles have been published since 1990. The practice of providing multi-source feedback is generally viewed favorably, with researchers espousing the potential benefits for performance improvement at both the individual and organizational level (Bailey & Fletcher, 2002). Yet empirical research on the effectiveness of multi-source feedback is not conclusive. The link between providing feedback and performance improvement is not uniform (Kluger & DeNisi, 1996). A recent meta-analysis with the largest sample size in this field so far reported an overall positive but very small effect size of multi-source feedback effectiveness (Smither, London, & Reilly, 2005). Based on this research, these authors argued that practitioners should not expect large, widespread performance improvement after employees receive multi-source feedback.

From a review of the literature it appears that in virtually all cases studies investigated the performance improvement by comparing the average or composite performance rating at Time 1 (before feedback) to the one at Time 2 (after feedback) (e.g., Johnson & Ferstl, 1999; Reilly, Smither, & Vasilopoulous., 1996; Smither, London, Flautt, Vargas, & Kucine, 2003). Smither et al. (2005) suggested that this comparison may be too crude a measure of behavioral change. Many organizations used managerial performance multi-source feedback for leadership development. Such developmental

programs generally recommend that feedback recipients should focus on no more than two or three leadership areas for improvement (Antonioni, 1996). Indeed, in many feedback programs, managers select a small number of improvement goals. For example, many organizations provide managers multi-source feedback on leadership competencies. Managers usually will select and work on a few bottom-ranked (but important) competencies for development. These managers might make substantial improvement on these selected leadership competencies, but this would have little effect on the overall leadership as indicated by the composite score. Therefore, the amount of improvement observed in the literature may not reflect the true success of the managerial performance multi-source feedback. For this reason, the primary purpose of the current study is to investigate the improvement on the leadership competencies feedback recipients specifically selected for development following multi-source feedback. We predict that feedback recipients will make more improvement on the selected leadership competencies than on other competencies or the overall leadership score.

Another critical question regarding multi-source feedback is whether feedback recipients make or sustain the improvement over a long period of time. The cost of multi-source feedback is usually expensive. If feedback recipients return to their previous level of competency, the substantial financial investments made in such programs will be of little value (Reilly et al., 1996). Walker and Smither (1999) followed 252 managers over multiple administrations of an upward feedback program. They found no improvement between the first and second feedback, but they observed improvement between the second and the third and again between the third and fourth administration of the upward feedback. Time therefore is an important consideration to document the multi-source

feedback effectiveness. In managerial performance feedback programs, feedback recipients typically select and develop their bottom-ranked competencies. One might argue that the improvement we observe after the initial administrations of the feedback program is due to the regression to mean. Hence, the second purpose of the current study is to examine whether the initial gains following the first administration of the multi-source feedback were sustained over time and across multiple administrations of the feedback. If the improvement is sustained across multiple administrations of the feedback, we will have confidence to say that the improvement is real and not because of the regression to mean.

Finally, we will examine the moderating effect of the “developmental difficulty” of the leadership competency on the multi-source feedback effectiveness. Smither et al. (2005) articulated a theoretical model for multi-source feedback and performance improvement. In this model, there are eight factors that play a role in determining the extent of behavioral change and performance improvement following multi-source feedback. These eight factors fall in or relate to three categories: (a) characteristics of the feedback such as self-other agreement (e.g., Atwater & Yammarino, 1992; Fletcher & Baldry, 2000; Johnson & Ferstl, 1999) and favorability of the feedback (e.g., Atwater & Brett, 2005), (b) characteristics of feedback recipient such as self-efficacy (e.g., Bailey & Austin, 2006) and goal orientation (e.g., Heslin & Latham, 2004), and (c) characteristics of the rater source such as credibility of raters (e.g., Albright & Levy, 1995). We feel another important factor is missing from the model—the nature of the performance dimension. The nature of the performance dimension is likely to influence the feedback effectiveness directly or indirectly. First of all, not all performance dimensions are equal

on cognitive complexity, experience requirement, and emotional involvement. Some performance dimensions are harder than others to improve. The nature of the performance dimension will impact the feedback effectiveness directly in that feedback recipients likely will make more improvement on easier performance dimensions than harder ones. The nature of performance dimension also may interact with other factors to influence feedback effectiveness indirectly. Bailey and Austin (2006) examined how the perceived importance of the feedback behaviors to the individual's current role affects performance improvement based on the assumption that the perceived importance is related to the feedback salience (London & Smither 1995). Another example of the indirect effects of the nature of the performance dimension on feedback effectiveness is its relationship with self-other agreement. Self-other agreement was found to be influenced by performance dimension rating difficulty (Wohlers & London, 1989). The credibility of rater source probably also is determined by the performance dimension in that rating on a specific dimension (e.g., developing direct reports) from a particular rater source (e.g., direct reports) is likely to be perceived more credible than from other rater sources (e.g., boss and peers).

In the current study, the factor related to the nature of the performance dimension is the developmental difficulty of the leadership competencies. We predict that feedback recipients will make more improvement on easier leadership competencies than on harder ones.

Method

Participants and Organizational Context

Participants consisted of 78 managers from a large financial service company in the United States. The company started a Leadership Development Program in 1997 in an effort to improve the quality of their current management team and to enhance the bench strength. The central purpose of the program has been to provide 360-degree feedback to managers and to build development plans that focus on competencies in need of development for either the current role or for some future role.

All participants had at least one development cycle. A development cycle begins with one multi-source feedback (pre-assessment) and ends with another multi-source feedback (post-assessment). The time interval was usually between one and two years. In a few instances, the time interval was up to three years. For each administration of the multi-source feedback, raters provided both skill and importance ratings. Between the two feedbacks, participants worked with an executive coach to select specific leadership competencies and develop an action plan. Typically, the competencies selected were from the bottom-ranked on skill rating but top-ranked on importance rating. The action plans generally run 10-15 pages and specified many potential resources that would help the manager's development. Additionally, each manager was provided a book containing a great many practical suggestions on how to develop specific leadership competencies.

The company tried to hold managers accountable to do the work in their development plans. For example, managers were asked to supply a progress report regularly. All reports were read and scored. Substantive responses were written and provided to each manager. Most managers scored "B", meaning that they were

approaching their plan well and making good progress. Those scoring “C” were given some suggestions, and it was made clear that they were not making sufficient progress. For those managers scoring “A”, they had done something impressive, and the group received a congratulatory e-mail with a copy going to their bosses up to and including the CEO. Additionally, managers had 20% of their bonus at stake, dependent on their completion of their plan within specified time period.

Fifty five of the 78 managers (70.5%) had two development cycles. As a result, these managers experienced three administrations of the multi-source feedback. Sixteen (20.5%) had three development cycles. These 16 managers experienced four administrations of the multi-source feedback. For managers who had more than one development cycles, they mostly selected and worked on different leadership competencies for different development cycles. Less than 7% of the times the competencies selected were the same over multiple development cycles. For each administration of the multi-source feedback, each of the managers had been rated by usually eight people—one immediate manger, three peers, three direct reports, and self. In order to maintain respondent confidentiality, no information on age, ethnicity, or gender was reported.

Multi-source feedback instrument

The multi-source feedback instrument consisted of 67 items designed to assess various leadership competencies. The theoretical foundation and research on the development of this instrument are summarized in Lombardo and Eichinger (2003, 2004). Each of the managerial leadership competencies was rated on a 5-point scale to assess how skilled the ratee was. Responses ranged from “a towering strength” (5),

“talented” (4), “skilled/ok” (3), “a weakness” (2), to “a serious issue” (1). According to Lombardo and Eichinger (2003), test-retest reliability was .75 for self-rating and .82 for others’ rating (i.e., peers, boss, and direct reports). In addition, each of the leadership competencies was rated on a 5-point scale to assess how important it was for the ratee’s job. The scale is “mission critical” (5), “very important” (4), “nice to have” (3), “less important” (2), and “least important” (1).

Development difficulty

The instrument publisher developed an index of “developmental difficulty” to indicate how hard it is to develop each of the 67 leadership competencies (Lombardo & Eichinger, 1995). This index was made up of three components: the underlying nature of the competency, the actual importance, and the perceived importance. The underlying nature of the competency was assessed on six factors: complexity of the skills, experience requirement, beliefs (e.g., attitudes, values, and opinions), cognitive complexity, emotion involvement, and human make-up. The actual importance was determined by the relationship between each of the 67 leadership competencies and job performance or promotion. The assumption is that the lower the actual importance, the more difficult the skill will be to develop as it rarely relates to performance or promotion and there would not be many support systems in place to help people. The scale was inverted so that more important competencies are easier for development on the job. The perceived importance was based on normative data. The instrument publisher collected importance rating from thousands of managers. The theory of the case is that the lower the perceived importance of the competency, the lower the motivation to develop the skill, and the less organizational support would be available and therefore the harder to develop the skill.

Again, the perceived importance scale was inverted to indicate the development difficulty. An overall difficulty index was computed by adding the above three components. The 67 competencies were then grouped into five categories based upon the distribution of the overall difficulty index: easiest (1), easier (2), moderate (3), harder (4), and hardest (5).

Analysis

We compared the pre-post averages of “all others” ratings (boss, peers, and direct reports) to indicate the improvement in each development cycle. Self-rating is least accurate compared with the ratings from others (Dunning, Heath, & Suls, 2004; Eichinger & Lombardo, 2004) and therefore not included in the analysis. Raters from each rater source (peers and direct reports) were first averaged. The score on a competency is the average of the rater source averages. The improvement score (the difference between post and pre) was correlated with the coding of the development difficulty to examine whether the improvement was influenced by the developmental difficulty.

Results

As mentioned previously, all the 78 managers had at least one development cycle. More than half (55) had two development cycles, whereas 16 had three. The results are presented below in three sections pertain to the three development cycles.

First Development Cycle

For each manager, three leadership competency scores were computed: the average score on the competencies selected for development, the average score on the

competencies not selected for development, and the average score on all the 67 competencies. The means and standard deviations of the three types of average scores in the first development cycle are presented in the Table 1.

As can be seen, all the three types of averages are statistically different between pre and the immediate post administrations. Managers improved on the competencies they did not select for development. However, managers improved much more on competencies selected ($d=.94$) than on competencies not selected ($d=.38$) and all competency composite ($d=.42$).

Since the average score of the competencies selected for development ($M=3.25$) before action plan was statistically lower than the average score of the competencies not selected for development ($M=3.71$) ($p<.05$), it is necessary to examine if the improvement on the competencies selected for development is due to the regression to mean. Regression to the mean happens when the differences on the means between the bottom-ranked competencies and the overall mean (the mean on all 67 competencies) on the pre-assessment decreases on the post-assessment. In other words, the mean differences between the bottom-ranked competencies and overall mean on the pre-assessment will be larger than the mean differences between the same competencies and overall mean on the post-assessment. See the figure for the illustration of the regression to mean.

If there is truly improvement, the observed difference between the mean of the selected competencies and the overall mean on post-assessment must be smaller than the expected difference due to regression. To test this, we computed the correlation between the pre- and the immediate post-assessment for each of the participants. The expected

difference between the mean of the competencies selected for development and the overall mean on the post-assessment was then computed for each of the participants using the following formula: Expected d at Time 2 = Observed d at Time 1 * r

Where r is the correlation between the pre- and post assessment, observed d at time 1 is the difference between overall mean and the mean of the competencies selected for development at Time 1.

Paired-sample T test indicated that the observed difference at Time 2 ($M=.19$, $s=.22$) was statistically smaller than the expected difference at Time 2 ($M=.25$, $s=.15$) ($p<.05$), suggesting that the observed improvement on the competencies selected for development was more than explained by the regression to mean.

Fifty five of the 78 managers had three administrations of the multi-source feedback. Sixteen of them had four administrations. Table 2 shows the means and standard deviations of the ratings on the competencies selected for development across multiple administrations of the multi-source feedback for the two groups of managers.

One-way ANOVA indicated a significant main effect of time for the group with three administrations ($F(2, 108)=66.17$, $p<.001$) and the group with four administrations ($F(3, 45)=18.31$, $p<.001$). The results of post hoc paired t tests show that for participants with three administrations, the post 1 mean rating on the selected competencies was significantly higher than the pre mean rating ($p<.001$), and the post 2 mean rating was significantly higher than the post 1 mean rating ($p<.001$). For participants with four administrations, the post 1 mean rating on the selected competencies was significantly higher than pre mean rating ($p<.001$), and the post 2 mean rating was significantly higher than the post 1 mean rating ($p<.001$). But the post 3 mean rating and the post 2 mean

rating were not statistically different ($p=.10$). Therefore, the improvement that occurred in the first development cycle continued at the third administration of the multi-source feedback, but maintained at the same level at the fourth administration of the feedback. Figure 2 depicts the mean ratings on the competencies selected for development in the first development cycle across multiple administrations of the multi-source feedback.

Second Development Cycle

The means and standard deviations for the immediate pre and post assessment in the second development cycle are presented in the Table 3. As can be seen in the table, all the three types of averages are statistically different between the immediate pre- and the immediate post-assessment. But managers improved much more on competencies selected ($d=.97$) for development than on competencies not selected ($d=.38$) and all competency composite ($d=.38$). Again, we tested whether the improvement was explained for by the regression to mean. The correlation between the immediate pre- and post- assessment was calculated for each of the 55 participants. The expected difference between the mean rating of the competencies selected for development and the overall mean rating on the 67 competencies was calculated using the formula presented above. Paired-sample T test indicated that the observed difference ($M=.20$, $s=.20$) was statistically smaller than the expected difference ($M=.24$, $s=.12$) ($p<.05$), suggesting that the improvement occurred in the second development cycle was more than the regression to mean.

Sixteen of the 55 participants had another development cycle. Table 4 shows the means and standard deviations of the ratings on the competencies selected for development in the second development cycle across multiple administrations of the

multi-source feedback. One-way ANOVA indicated a significant main effect of time for the group with three administrations ($F(2, 108)=25.07, p<.001$) and the group with four administrations ($F(3, 45)=15.14, p<.001$). The results of post hoc paired-sample T tests show that for participants with three administrations, the pre 2 mean rating on the selected competencies was not statistically different with the pre 1 mean rating ($p=.15$), but the post 1 mean rating was significantly higher than the pre 2 mean rating ($p<.001$). For participants with four administrations, the pre 2 mean rating on the selected competencies was not statistically different with the pre 1 mean rating ($p>.05$). The post 1 mean rating was significantly higher than the pre 2 mean rating ($p<.001$), but the post 2 mean rating was not statistically different with the post 1 mean rating ($p=.63$). The longitudinal analysis indicated again that the improvement after intervention sustained. Figure 3 depicts the mean ratings on the competencies selected for development in the second development cycle across multiple administrations of the multi-source feedback.

Third Development Cycle

The means and standard deviations for the immediate pre and post assessment in the third development cycle are presented in Table 5. The mean rating on the competencies not selected for development and the overall average (on 67 competencies) improved in the third development cycle, but were not statistically significant. Participants improved significantly on the competencies selected for development. The same method was used to test the regression to mean. The observed difference between the mean rating on the competencies selected for development ($M=.14, s=.22$) was smaller than the expected difference ($M=.19, s=.11$). But the difference was not statistically significant ($p>.05$), because of the small sample size ($N=16$).

These participants had four administrations of the multi-source feedback, including three pre assessments and one post assessment. Table 6 presents the means and standard deviations on the competencies selected for development in the third development cycle across multiple administrations of the multi-source feedback. One-way ANOVA indicated a significant main effect of time ($F(3, 45)=7.69, p<.001$). On the other hand, post hoc paired-sample T tests show that the three pre assessments are not statistically different with each other ($p>.05$). The post assessment is statistically higher than the pre3 assessment ($p<.001$).

Development Difficulty

The average improvement on each of the competencies selected by the participants for development in the three development cycles was computed. The improvement was negatively correlated with the leadership competency development difficulty coding composite index ($r=-.27, p<.05$) and development difficulty category ($r=-.31, p<.05$), suggesting that participants made less improvement on harder leadership competencies.

Summary and Conclusions

This study sought to overcome the methodological weakness of the previous research on the multi-source feedback effectiveness by examining the improvement on the leadership competencies participants selected for development. The results consistently demonstrated that participants improved more on competencies selected for development than on competencies not selected. This improvement was more than could be explained by the regression to mean, and sustained over multiple administrations of

the multi-source feedback. Finally, it was found that the improvement in competencies was related to the developmental difficulty. Participants improved more on easier competencies than on harder competencies. To our knowledge, this is the first study to investigate how the nature of the leadership competency influences the multi-source feedback effectiveness.

Participants also improved on competencies not selected for development and on the overall average. The improvement was statistically significant, suggesting that multi-source feedback alone is likely to result in improvement, but in a relatively small effect size. In the current study, participants improved about .4 standard deviations on the composite leadership competency score. The effect size is higher than the ones observed in Smither et al. (2005), but similar to the one reported by Kluger & DeNisi (1996). Researchers have proposed theories on why multi-source feedback is likely to result in improvement on performance, such as self-consistency theory (Korman, 1970 and 1976; Johnson & Ferstl, 1999), control theory (Carver & Scheier, 1981), and self-awareness theory (Atwater & Yammarino, 1992; Wegner & Vallacher, 1980). However, researchers have debate on how to measure performance improvement. Many contrasted the post-feedback performance composite score with the pre-feedback composite score (as discussed previously). Some argued against the direct measure of the performance improvement (Buda, Reilly, & Smither, 1991; Smither, Reilly, & Buda, 1988). For example, Smither and Walker (2001) suggested alternative approaches to measuring the impact of multi-source feedback such as retrospective degree of change ratings, where coworkers rate the extent to which the feedback recipient's performance has improved with respect to each of the improvement goals the feedback recipient set after receiving

feedback. Built upon Smither et al.'s (2005) observation, the current study examined the amount of improvement on the leadership competencies participants selected for development after multi-source feedback. In all three development cycles, participants improved more on selected competencies than on competencies not selected for development and composite competency score. The current study suggested that the small effect size observed in previous research partly was due to the method used to measure the performance improvement.

Two of the findings from the current study are noteworthy. First, the improvement participants made on the competencies they selected for development was more than the regression to mean. Second, the improvement was sustained across multiple administrations of the multi-source feedback. These findings are consistent to Reilly et al. (1996). In their study, most of the performance improvement occurred between the first and second administrations of the feedback program, and the improvement was not attributed solely to regression to the mean and was sustained over a fairly long period of time. The uniqueness of the current study was that we tested the regression to mean and the longitudinal effects of the multi-source feedback through multiple development cycles. In addition, we focused on the competencies participants selected for development rather than an overall index of leadership competency. The consistent findings in the current study suggest that overall the managerial leadership competency multi-source feedback program achieved its objective to help managers develop their leadership competencies.

Probably the most intriguing finding of the present study has to do with the relationship between the multi-source feedback effectiveness and leadership competency

development difficulty. Studies investigating the performance improvement within the context of multi-source feedback were not conclusive (Bailey & Fletcher, 2002). The developmental outcomes following a multi-source feedback process likely depend on a complex interplay of factors (Craig & Hannum, 2006). In addition to the factors articulated in Smither et al. (2005), we contend that the nature of the competencies would be an important factor influencing the developmental outcomes following the multi-source feedback. Bailey and Austin (2006) tested the indirect effects of the perceived importance of performance dimension. In their study, perceived importance of feedback did not significantly moderate the association between feedback favorability and amount of developmental activity or changes in self-assessments. The current study found that the development difficulty was statistically correlated to the observed improvement on the leadership competencies participants selected for development. This finding has theoretical as well as practical implications. Theoretically, it expands our knowledge base regarding the situations in which feedback recipients are likely to make measurable performance improvement. It has practical implications for organizations that administer multi-source feedback program. For example, it can help coaches and feedback recipients decide the number of development goals. A general rule of thumb was that feedback recipients should focus on no more than two or three areas for improvement (Antonioni, 1996). The findings of the current study imply that the number probably depends on the developmental difficulty of the leadership competencies selected for development. Feedback recipients may decide to work on more or fewer competencies depends on how difficult the competencies selected are. Another practical implication is that organizations may provide more resources to support feedback recipients to develop hard

competencies. The awareness of the nature of the leadership competency can help organizations design action plans and provide supporting resources that maximize the multi-source feedback effectiveness.

A limitation in this study was that the effects of the multi-source feedback might be confounded by other factors. As described in the method section, the company provided coaching and implemented other interventions to hold the managers accountable. Multi-source feedback combined with coaching may have led to the improved effectiveness of the intervention (Smither, London, Flautt, Vargas, & Kucine, 2003). It is therefore hard to tell if the positive effects observed in this study were due to the multi-source feedback or other practices. Nevertheless, the findings in the current study are encouraging to companies using multi-source feedback. Organizations rarely use multi-source feedback alone. Other organizational practices, particularly executive coaching, are often implemented together with multi-source feedback. The findings from the current study largely support the use of multi-source feedback. HR professionals or other administrators of multi-source feedback programs should recognize the importance of monitoring the development difficulty of the leadership competencies selected for improvement.

References

- Albright, M. D., & Levy, P. E. (1995). The effects of source and performance rating discrepancy on reactions to multiple raters. *Journal of Applied Social Psychology, 25*, 577-600.
- Antonioni, D. (1994). The effects of feedback accountability on upward appraisal ratings. *Personnel Psychology, 47*, 349-356.
- Atwater, L. E., & Bertt, J. F. (2005). Antecedents and consequences of reactions to developmental 360-degree feedback. *Journal of Vocational Behavior, 66*, 532-548.
- Atwater, L. E., & Waldman, D. (1998). Accountability in 360 degree feedback. *HR Magazine, 43*, 96-104.
- Atwater, L. E., & Yammarino, F. (1992). Does self-other agreement on leadership perceptions moderate the validity of leadership and performance predictions. *Personnel Psychology, 45*, 141-164.
- Bailey, C., & Austin, M. (2006). 360-degree feedback and developmental outcomes: The role of feedback characteristics, self-efficacy and importance of feedback dimensions to focal managers' current role. *International Journal of Selection and Assessment, 14*, 51-66.
- 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.

- Buda, R., Reilly, R. R., Smither, J. W. (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 Science*, 6, 89-99.
- Carver, C. & Scheier, M. (1981). *Attention and self-regulation: A control theory approach to human behavior*. New York: Springer.
- Craig, S. B., & Hannum, K. (2006). Research update: 360-degree performance assessment. *Consulting Psychology Journal: Practice and Research*, 58, 117-124.
- Dunning, D., Heath, C., & Suls, J. M. (2004). Flawed self-assessment: Implications for health, education, and the workplace. *Psychological Science in the Public Interest*, 5, 69-106.
- Eichinger, R. W., & Lombardo, M. M. (2004). Patterns of rater accuracy in 360 degree feedback. *Human Resource Planning*, 27, 23-25.
- 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.
- Heslin, P. A., & Latham, G. P. (2004). The effect of upward feedback on managerial behavior. *Applied Psychology: An International Review*, 53, 23-37.
- Johnson, J. W., & Ferstl, K. L. (1999). The effects of interrater and self-other agreement on performance improvement following upward feedback. *Personnel Psychology*, 52, 271-303.

- Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119*, 254-284.
- Korman, A. K. (1970). Toward a hypothesis of work behavior. *Journal of Applied Psychology, 54*, 31-41.
- Korman, A. K. (1976). Hypothesis of work behavior revisited and an extension. *Academy of Management Review, 1*, 50-63.
- Lombardo, M. M., & Eichinger, R. W. (1995). The career architect user's manual. Minneapolis, MN: Lominger Limited, Inc.
- Lombardo, M. M., & Eichinger, R. W. (2003). *Leadership architect norms and validity report*. Minneapolis, MN: Lominger Limited, Inc.
- Lombardo, M. M., & Eichinger, R. W. (2004). *The leadership machine*. Minneapolis MN: Lominger Limited, Inc.
- Lombardo, M. M., & Eichinger, R. W. (1995) The CAREER ARCHITECT user's manual. Minneapolis, MN: Lominger Limited, Inc.
- London, M., & Smither, J. W. (1995). Can multi-source feedback change perceptions of goal accomplishment, self-evaluations, and performance-related outcomes? Theory based applications and directions for research. *Personnel Psychology, 48*, 803-839.
- Morgeson, F. P., Mumford, T. V., & Campion, M. A. (2005). Coming full circle: Using research to address 27 questions about 360-degree feedback programs. *Consulting Psychology Journal: Practice and Research, 57*, 196-209.

- Reilly, R. R., Smither, J. W. & Vasilopoulos, N. L. (1996). A longitudinal study of upward feedback. *Personnel Psychology*, 49, 599-612.
- Smither, J. W., London, M., Flautt, R., Vargas, Y., & Kucine, I. (2003). Can working with an executive coach improve multisource feedback ratings over time? A quasi-experimental field study. *Personnel Psychology*, 56, 23-44.
- Smither, J. W., London, M., & Reilly, R. R. (2005). Does performance improve following multisource feedback? A theoretical model, meta-analysis, and a review of empirical findings. *Personnel Psychology*, 58, 33-66.
- Smither, J. W., Reilly, R. R., & Buda, R. (1988). The effect of prior performance information on rating of present performance: Contrast versus assimilation revisited. *Journal of Applied Psychology*, 73, 487-496.
- Smither, J. W., & Walker, A. G. (2001). Measuring the impact of multisource feedback. In Bracken D., Timmreck C., & Church, A. (Eds.), *Handbook of multisource feedback*. San Francisco: Jossey-Bass.
- Walker, A. G., & Smither, J. W. (1999). A five-year study of upward feedback: What managers do with their results matters. *Personnel Psychology*, 52, 393-423.
- Wegner, D., & Vallacher, R. (1980). *The self in social psychology*. New York: Oxford University Press.
- Wohlers, R. J., & London, M. (1989). Ratings of managerial characteristics: Evaluation difficulty, co-worker agreement, and self-awareness. *Personnel Psychology*, 42, 235-261.

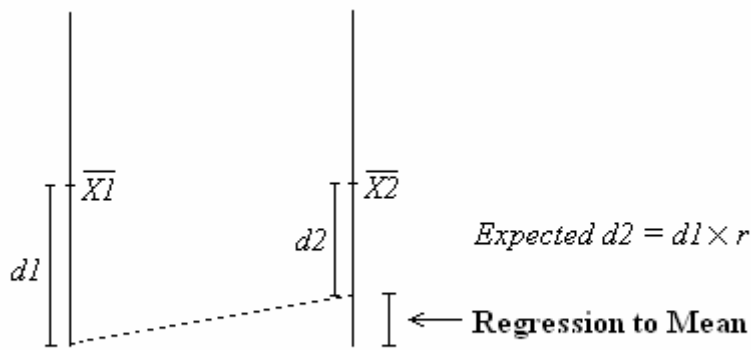


Figure 1. Regression to Mean

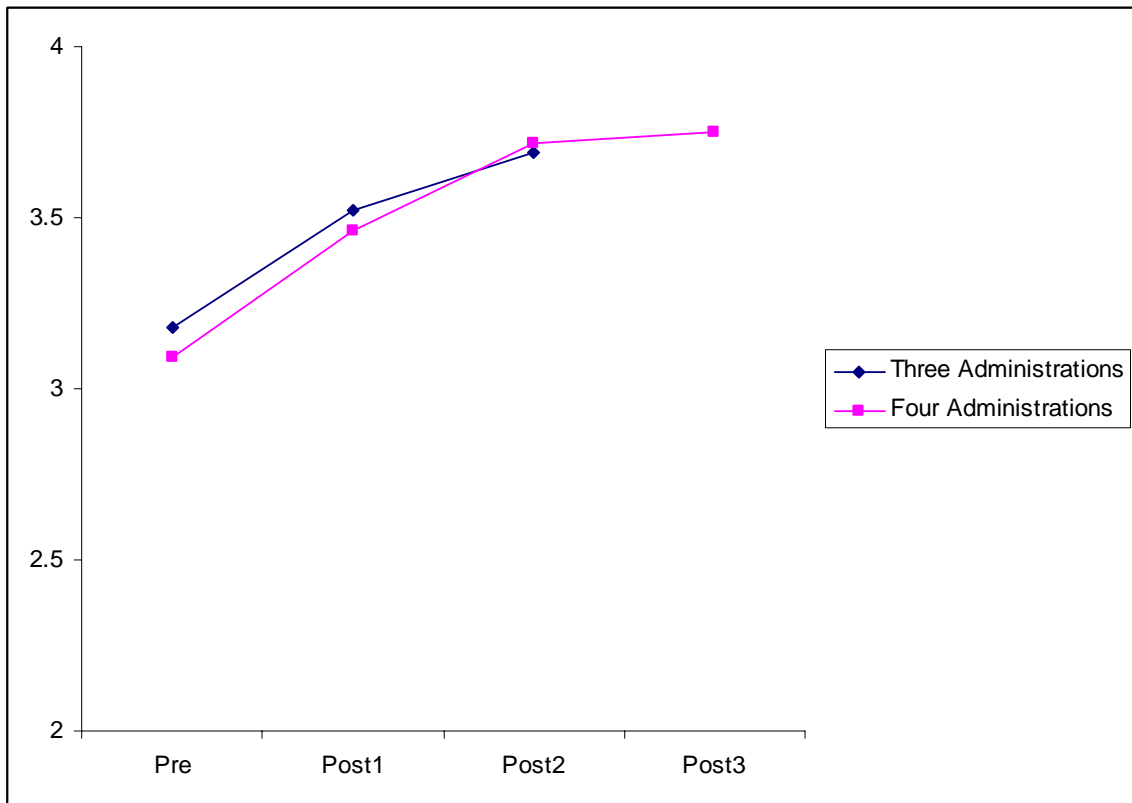


Figure 2. The mean ratings on the competencies selected in the first development cycle across multiple administrations of the multi-source feedback

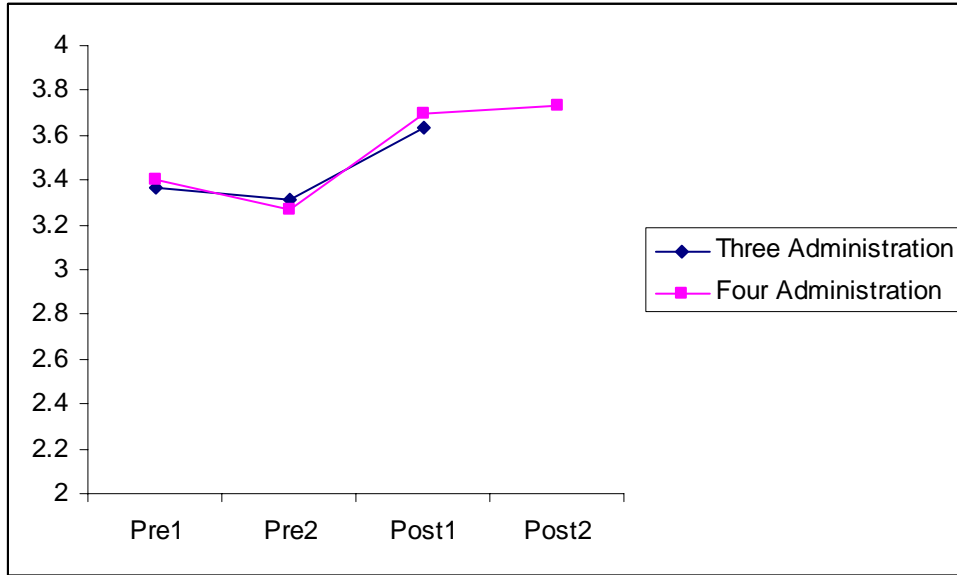


Figure 3. The mean ratings on the competencies selected in the second development cycle across multiple administrations of the multi-source feedback

Table 1.

Pre and the immediate post administration of the feedback in development cycle one

	Pre		Post1		<i>d</i>	<i>p</i>
	<i>Mean</i>	<i>Std.</i>	<i>Mean</i>	<i>Std.</i>		
Competencies Selected	3.25	.35	3.59	.36	.94	<i>p</i> <.001
All 67 Competencies	3.69	.22	3.79	.25	.42	<i>p</i> <.001
Competencies Not Selected	3.71	.22	3.80	.25	.38	<i>p</i> <.001

N=78

Table 2.

Means and standard deviations of the ratings on the competencies selected in the first development cycle across multiple administrations of the multi-source feedback

	Pre	Post1	Post2	Post3
Three administrations (<i>N</i> =55)	3.18 (<i>s</i> =.34)	3.52 (<i>s</i> =.34)	3.69 (<i>s</i> =.38)	
Four administrations (<i>N</i> =16)	3.08 (<i>s</i> =.35)	3.46 (<i>s</i> =.31)	3.72 (<i>s</i> =.38)	3.75 (<i>s</i> =.29)

Table 3.

Means and standard deviations for the immediate pre and post assessment in the second development cycle

	Pre		Post1		<i>d</i>	<i>p</i>
	<i>Mean</i>	<i>Std.</i>	<i>Mean</i>	<i>Std.</i>		
Competencies Selected	3.31	.30	3.63	.34	.97	<i>p</i> <.001
All 67 Competencies	3.74	.23	3.83	.25	.38	<i>p</i> <.01
Competencies Not Selected	3.75	.23	3.84	.25	.38	<i>p</i> <.01

N=55

Table 4.

Means and standard deviations of the rating on the competencies selected in the second development cycle across multiple administrations of the multi-source feedback

	Pre1	Pre2	Post1	Post2
Three Administrations (N=55)	3.37 (s=.29)	3.31 (s=.30)	3.63 (s=.34)	
Four Administrations (N=16)	3.40 (s=.27)	3.27 (s=.27)	3.70 (s=.17)	3.73 (s=.22)

Table 5.

Means and standard deviations for the immediate pre and post assessment in the third development cycle

	Pre		Post1		<i>d</i>	<i>p</i>
	<i>Mean</i>	<i>Std.</i>	<i>Mean</i>	<i>Std.</i>		
Competencies Selected	3.49	.32	3.77	.26	.93	<i>p</i> <.001
All 67 Competencies	3.83	.22	3.91	.17	.40	<i>p</i> =.15
Competencies Not Selected	3.85	.22	3.94	.17	.45	<i>p</i> =.13

N=16

Table 6.

Means and standard deviations on the competencies selected in the third development cycle across multiple administrations of the multi-source feedback

Pre1	Pre2	Pre3	Post1
3.40 (s=.31)	3.45 (s=.30)	3.48 (s=.32)	3.77 (s=.26)

N=16