The Effects of Peer Feedback on Team Member Behavior

PETER G. DOMINICK
Assessment Alternatives, Inc.

RICHARD R. REILLY
Stevens Institute of Technology

JACK W. McGOURTY
Assessment Alternatives, Inc.

We examined the effects of peer feedback on subsequent behavior using a four-dimensional model of team behavior. Participants \((N = 75)\) were randomly assigned to teams, and teams were randomly assigned to one of three experimental conditions: feedback, exposure, or control. In the feedback condition, participants rated themselves and each other using a 24-item behavioral observation scale after completing the first of two decision-making tasks. Before performing the second task, they received individualized feedback reports summarizing their self- and peer ratings. Those assigned to the exposure condition completed the behavioral observation scale after the first task but did not receive feedback. The second task was videotaped and rated by experts blind to experimental condition. Results showed significantly higher ratings for participants in the feedback and exposure conditions. The findings extend previous research on multisource feedback by isolating exposure to key behaviors as an important variable in behavioral improvement.

Numerous authors have stressed the pivotal role that autonomous or semiautonomous teams play in the success and effectiveness of modern firms (e.g. Katzenbach & Smith, 1993; Peters, 1988; Reich, 1987). Increasingly, teams have become integral parts of organizations’ structures. For example, by 1990, 47% of Fortune 1,000 companies reported that they used work teams compared with 28% just 3 years earlier in 1987 (Lawler, Mohrman, & Ledford, 1992).

Regardless of the setting, the organization of workers into teams implies an increasing emphasis on self-management both by individuals and the team as a unit (Hackman, 1987). Moreover, successful team outcomes

The authors would like to thank James Smither for his helpful comments on an earlier draft of this manuscript. They would also like to thank Veona Martin for her assistance in conducting this study.

Group & Organization Management, Vol. 22 No. 4, December 1997 508-520
© 1997 Sage Publications, Inc.

508
depend heavily on effective interaction among team members. Virtually all models of team and work group effectiveness acknowledge that the interpersonal process is a crucial antecedent of team performance (Gladstein, 1984; Hackman, 1983; Nieva, Fleishman, & Reich, 1978; Sundstrom, DeMeuse, & Futrell, 1990). Thus, the effective transition to a team-based organization can be facilitated by organizational interventions designed to promote and reinforce the behaviors necessary for effective teamwork.

One such intervention is a behavioral feedback program. For example, studies of upward feedback have shown significant changes in behavior after the implementation of behaviorally based feedback programs (e.g., Atwater, Roush, & Fischthal, 1995; Reilly, Smither, & Vasilopoulos, 1996; Smither et al. 1995). On the other hand, there have been virtually no empirical investigations of peer feedback in team settings. The primary purpose of the present study was to examine the effects of behaviorally based peer feedback on subsequent behavior in a team-based task. Because earlier quasi-experimental research on upward feedback (Reilly et al., 1996; Smither et al., 1995) suggested that mere exposure to the desired behaviors was a critical factor, we also examined whether exposure alone would lead to behavior change.

PEER EVALUATION AND FEEDBACK

The self-managing context of many teams suggests that team members themselves can play an important role in enhancing and sustaining team effectiveness by providing feedback to each other. Mechanisms that assist team members in such assessment and feedback can be crucial from both an evaluative and developmental perspective (Murphy & Cleveland, 1991; Saavedra & Kwon, 1994). Ideally, behavioral evaluation and feedback from other team members should be an important developmental tool fostering more positive team behaviors from members. In fact, several organizations have reported successfully using peer feedback in team settings as the basis for both development and evaluation (e.g. Ramsay & Letho, 1994; Zigon, 1994).

Several studies have examined the effects of feedback on subsequent behavior. Bernardin, Hagan, and Kane (1995) found improvement in subordinate and peer ratings but no changes in supervisor or customer ratings after managers received 360-degree feedback. Hazucha, Gentile, and Schneider (1993) reported skill increases 2 years after receiving 360-degree feedback, but the absence of a control and subject loss makes it difficult to determine the cause of the improvement. Hegarty (1974) found that upward feedback leads to subordinates perceiving positive changes in the boss’s subsequent
behavior. Atwater et al. (1995) found that follower ratings of student leaders improved after feedback was given to leaders and that leaders receiving negative feedback (defined as self-ratings that were considerably higher than follower ratings) improved the most. Smither et al. (1995) found a significant but small improvement in subordinate ratings across a sample of 238 managers receiving feedback. Consistent with Atwater et al. (1995), Smither et al. (1995) and Reilly et al. (1996) found that improvement was greatest for managers who initially received the most negative feedback (i.e., lowest ratings).

We were guided by two previous studies (Reilly et al., 1996; Smither et al., 1995) suggesting that feedback itself may not be the critical variable in producing change. In these studies, improvement for individuals who did not receive feedback but were exposed to the feedback instrument (by completing self-ratings and ratings for their bosses) improved as much as those who actually got feedback reports. Along similar lines, Smither, Wohlers, and London (1992) found that the expressed intentions of team leaders to change their behavior were the same regardless of whether the leaders received individualized, upward feedback or only normative feedback (average of team leaders’ ratings). Because of the quasi-experimental nature of these studies, however, it was difficult to determine whether exposure or feedback was the critical variable in producing change. From the perspectives of control theory (Carver & Scheier, 1981), receipt of feedback should be critical. Control theory suggests that specific feedback is the basis for identifying goal-feedback discrepancies, which in turn direct one’s attention toward change and improvement.

On the other hand, there are several reasons to believe that behavioral improvements should occur simply through exposure to and completion of the feedback instrument. First, a feedback instrument consisting of behavioral items is, by nature, prescriptive (Van Velsor & Leslie, 1991) and provides people with a learning opportunity. By reviewing the items, individuals are exposed to examples of effective behavior. Second, Locke and Latham (1990) have argued that the act of introducing a formal feedback system into a work group may sometimes be sufficient to cause spontaneous goal setting. That is, there is an unmistakable message that performance should be improved in those areas that are being measured. Third, the process of completing self- and peer ratings provides individuals with an opportunity to reflect on their own behavior and to establish normative standards and personal improvement objectives.

The present study provided an opportunity to examine two related questions. First, we were interested in whether behavior in team settings could be
changed by the administration of a feedback intervention. Second, we wanted to determine whether feedback, or exposure to the behaviors, was the critical factor in producing change. Specifically, we made the following predictions:

_Hypothesis 1:_ Participants who gave and received behavioral peer feedback will demonstrate effective team behavior more frequently than participants who did not give or receive feedback.

_Hypothesis 2:_ Participants who gave, but did not receive, feedback will demonstrate effective team behavior more frequently than participants who did not give or receive feedback.

_Hypothesis 3:_ Participants who gave and received feedback will not differ from participants who gave feedback only.

**METHOD**

**SAMPLE AND PROCEDURE**

Participants were 75 graduate and undergraduate students in organizational behavior and group dynamics classes. Undergraduates were senior and junior engineering students. Graduate students were enrolled in master’s courses in management and had at least 1 year of work experience. All subjects participated in the study as part of a class assignment. The sample included 54 males and 23 females, with 51 White and 24 minority participants. Participants were randomly organized into teams of four or five. The teams were randomly assigned to either a feedback, exposure, or control condition and performed two counterbalanced group decision-making tasks—one at each of two sessions in successive weeks. Chi-square tests showed no significant differences in the assignment of subjects to condition by gender ($\chi^2 = 2.40, df = 2$) or by race $\chi^2 = 4.03, df = 2$.

**FEEDBACK CONDITION**

Participants in the feedback condition completed ratings on themselves and their teammates following completion of the first task. The feedback form was based on a behavioral model of team performance that specified four dimensions of team performance: collaboration, communication, decision making, and self-management (McGourty, DeMeuse, & Dominick, 1994). Before beginning a planning session for the second task, participants in the feedback condition received their reports and were given time to
review the contents. Individualized feedback reports included average dimension scores on each participant's self- and peer ratings and a listing of self-ratings and average peer ratings for each of the 24 items.

EXPOSURE CONDITION

Participants in the exposure condition also completed ratings on themselves and their teammates following completion of the first task. Subjects in the exposure condition were told that they would receive feedback at some point in the future but were not actually given feedback until the completion of the second task.

CONTROL CONDITION

Participants in the control condition completed a placebo instrument on the task content but did not complete self- or peer ratings and received no feedback.

TASKS

All teams completed the same two group decision-making tasks, Black Bear (B. Glaser, 1993) and Outback (R. Glaser & Glaser, 1993). Both are tasks requiring team members to come to consensus regarding selection of a strategy alternative and the prioritized usefulness of several resources. Completion of these tasks required significant interaction and discussion among participants. Team members were required to work interdependently as a self-managed unit. As in many actual workplace teams, members had to establish procedures for working together, jointly diagnose problems and alternatives, and collaborate to develop solutions (Van de Ven, Delbecq, & Koenig, 1976).

DEPENDENT VARIABLES

Unlike other studies of the effects of feedback (e.g., Reilly et al., 1996; Smither et al., 1995), the dependent variable in the present study was not based on participants' perceptions of their peers' behavior. Instead, we used objective ratings made by experts blind to the experimental condition. Experienced assessors were given standard instructions on observing and rating the subjects on the same 24-item rating form completed by participants in the feedback and exposure conditions. All raters had graduate degrees in
industrial/organizational psychology and had assessment experience. Our raters met the definition of experts as used in previous studies (e.g., Murphy & Balzer, 1986; Smither, Barry, & Reilly, 1989). Videotapes of each team’s second task were randomly distributed to the raters, who were blind to condition. Of the 75 participants, 49 were rated by one assessor and 26 were independently rated by two assessors. Where two ratings were available, the average of the two ratings was used. Average ratings for the items measuring collaboration, communication, decision making, and self-management were computed. In addition, an overall team behavior score was derived for each participant by calculating the average rating across the 24 behaviors.

RESULTS

Intercorrelations, internal consistency reliabilities (alpha), means, and standard deviations for all variables are shown in Table 1. Interrater reliabilities, based on the subsample of 26 participants with two ratings, are also shown in Table 1. Correlations between dimensions ranged from .72 (communication vs. self-management) to .89 (communication vs. decision making) and all were significant ($p < .01$). Internal consistency reliabilities were also high, ranging from .89 for communication to .97 for the overall score. Interrater reliabilities for the dimension scores ranged from .37 for communication to .72 for self-management. The interrater reliability for the overall score was .68 for two raters and .52 for one rater.

Although it had originally been planned to use assessor ratings of the first task as a covariate, malfunctioning video equipment made several of the tapes unusable. This made it necessary to use a nested (teams within condition) ANOVA to compare the behavioral performance of subjects in each of the three conditions. The overall score across the 24 items was used to perform an initial omnibus test. Separate ANOVAs were then done for each of the four team behavior dimensions. Post hoc Scheffe tests were used to compare pairs of groups.

Both hypotheses were supported. The ANOVA for overall team behavior indicated a significant difference between conditions ($F = 47.87, df = 2, 58; p < .01$). Similar results were obtained for ratings on collaboration ($F = 46.66, p < .01$), communication ($F = 59.94, p < .01$), decision making ($F = 40.53, p < .01$), and self-management ($F = 23.39, p < .01$). A post hoc comparison of means on overall team behavior indicated a significant difference between the feedback and control groups ($p < .05$) and between the exposure and control groups ($p < .05$), but no significant difference between the exposure and
TABLE 1
Correlations, Alpha Coefficients, Means, Standard Deviations, and Interrater Reliabilities for Measures of Team Behavior

<table>
<thead>
<tr>
<th>Variable</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collaboration</td>
<td>.91</td>
<td>.83</td>
<td>.79</td>
<td>.76</td>
<td>.91</td>
</tr>
<tr>
<td>Communication</td>
<td>.89</td>
<td>.89</td>
<td>.79</td>
<td>.72</td>
<td>.93</td>
</tr>
<tr>
<td>Decision making</td>
<td></td>
<td>.93</td>
<td>.78</td>
<td>.94</td>
<td></td>
</tr>
<tr>
<td>Self-management</td>
<td></td>
<td></td>
<td>(.92)</td>
<td>.89</td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td></td>
<td></td>
<td></td>
<td>(.97)</td>
<td></td>
</tr>
<tr>
<td>Means</td>
<td>2.92</td>
<td>2.92</td>
<td>2.66</td>
<td>2.33</td>
<td>2.71</td>
</tr>
<tr>
<td>SD</td>
<td>.86</td>
<td>.87</td>
<td>1.00</td>
<td>1.01</td>
<td>.86</td>
</tr>
<tr>
<td>Interrater reliability</td>
<td>.62</td>
<td>.37</td>
<td>.56</td>
<td>.72</td>
<td>.68</td>
</tr>
</tbody>
</table>

NOTE: Alpha coefficients are shown in parentheses.

feedback groups. A final post hoc t test was performed to test the difference between the pooled exposure/feedback groups and the control group using teams as the unit of analysis. A significant difference (p < .05) was found in favor of the exposure/feedback groups (t = 1.86, df = 15). Table 2 shows the means and standard deviations for all variables by condition.

DISCUSSION

These results suggest that peer feedback can be a useful approach for helping team members to improve their interpersonal effectiveness. The results also suggest that it is not so much the feedback itself that drives change but exposure to and completion of the feedback instrument. These results are consistent with those reported by Smither et al. (1995) and Reilly et al. (1996), who found that exposure to a feedback instrument without feedback resulted in as much improvement as feedback itself. Unlike the present study, however, their results did not include a sample not exposed to the instrument. Thus, the present study more clearly identifies exposure as the critical mechanism in creating behavior change.

A limitation of most previous studies of feedback was the use of study participants to provide measurement of behavioral change (e.g., Atwater et al., 1995; Reilly et al., 1996; Smither et al., 1995). In contrast, the present study measured behavioral change using external raters who were blind to condition. Our findings strengthen previous research by showing that it is actual behavior that changes and not merely a change in the perceptions of participants in feedback programs.
TABLE 2
Means and Standard Deviations for Team Behavior Variables Within Experimental Condition

<table>
<thead>
<tr>
<th>Variable</th>
<th>Control (N = 23)</th>
<th>Feedback (N = 30)</th>
<th>Exposure (N = 22)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
</tr>
<tr>
<td>Collaboration</td>
<td>2.37</td>
<td>0.75</td>
<td>3.20</td>
</tr>
<tr>
<td>Communication</td>
<td>2.35</td>
<td>0.77</td>
<td>3.10</td>
</tr>
<tr>
<td>Decision making</td>
<td>2.13</td>
<td>0.82</td>
<td>2.81</td>
</tr>
<tr>
<td>Self-management</td>
<td>1.96</td>
<td>0.75</td>
<td>2.63</td>
</tr>
<tr>
<td>Overall</td>
<td>2.21</td>
<td>0.75</td>
<td>2.94</td>
</tr>
</tbody>
</table>

From a theoretical perspective, there are several possible explanations for our findings. First, the exposure to the instrument gave team members a framework for understanding their own behavior in a team context and focused their attention on the key behaviors relevant to effective performance. Second, introducing the feedback instrument communicated to participants that certain behaviors were important and valued. In turn, this allowed team members to set informal goals to improve on the second task (Locke & Latham, 1990). For example, understanding that effective decision making includes anticipating problems and developing contingency plans (see the appendix) allowed participants to increase their efforts to develop alternative solutions to problems in the subsequent task. Finally, results might also be explained from a control theory perspective (Carver & Scheier, 1981; Lord & Hanges, 1987). Participants who perceived a discrepancy between their behavior on the first task and the standard behavior presented might be most motivated to change their behavior (Atwater et al., 1995; Reilly et al., 1996; Smither et al., 1995).

There are two ways practitioners might be able to interpret the results of this study. One is to conclude that feedback itself may not be worthwhile and that simply introducing key behaviors and stressing their importance is enough to bring about change. Recent research by Reilly et al. (1996) suggests that this assumption would be unwise for two reasons: first, the knowledge that behavior is being measured and that feedback will occur at some later point may be equal in importance to exposure, and second, the repeated administration of a feedback instrument appears to result in sustaining initial change over a fairly long (e.g., 2½ years) period of time. The knowledge that feedback will not be provided may diminish team members’ motivation to set goals or regulate their behavior over time. In this regard, it is noteworthy that
participants in the exposure condition expected feedback after completion of the second task.

It should be noted that change as a result of feedback could probably be strengthened by supporting feedback initiatives with broader and more concerted change efforts. For example, we might expect even greater changes in behavior when information obtained through peer feedback instruments is combined with additional interventions such as the establishment of specific performance goals or by linking improvements to pay and other rewards.

Before drawing more definitive conclusions however, a number of limitations to this study should be noted. First, it might be argued that the teams we studied were, in fact, not teams but groups. Although previous literature is not consistent on the definition of teams, Salas, Dickinson, Converse, and Tannenbaum (1992) define teams as including the following characteristics: (a) two or more people who interact dynamically and interdependently toward a common and valued goal, objective, or mission and (b) a limited life span of membership. These same characteristics have been ascribed to groups (e.g., Johnson & Johnson, 1994, pp. 12-13). In any event, our teams/groups possessed these characteristics. Although our teams met only twice for approximately 1 hour each time, the interactive tasks they performed were designed specifically to teach people about teamwork and what some researchers (e.g., Gaddy & Wachtel, 1992) refer to as generic team skills (behaviors that are beneficial to team members regardless of the work setting). Nonetheless, an increase in the number and length of team meetings would have provided increased opportunities to observe behavior prior to the intervention. Additional meetings would also have allowed more opportunities for subjects to modify behavior in response to feedback.

A second limitation has to do with the simulated environment of the study. Although our results were consistent with findings in upward feedback research (e.g., Reilly et al., 1996), field replication of these results would strengthen our conclusions. Organizational factors such as feedback from other sources and reward structures could confound and weaken the effects of feedback/exposure found in the present study.

A final limitation has to do with the effects of exposure versus the effects of anticipation of feedback. It is possible that participants in the exposure condition were motivated to change behavior not only by exposure to the dimensions and items but also by the knowledge that feedback would be
received at some point in the future. Future research should control for this possibility.

The reliability of the observers on the overall score was .68 for two raters and .52 for one rater, which is comparable to field ratings of performance by supervisors. Rothstein (1990), for example, found the highest reliabilities for supervisor ratings of employees in field settings to be .55. Nonetheless, efforts should be made to improve the interrater reliability of the observed behaviors. This could be done by obtaining more ratings, providing more training to raters, and requiring raters to review the tapes at least twice. On the other hand, the relatively low interrater reliabilities make the obtained effect sizes even more compelling.

Future investigations should explore the extent to which self-evaluations moderate behavioral changes due to peer feedback on team behavior. Attempts to replicate these findings in field settings should be undertaken as well. The impact of peer feedback in combination with other interventions such as rewards and developmental training could also be explored. Finally, attempts should be made to identify relationships between behavioral changes due to peer feedback and overall team task performance.

APPENDIX
Team Behavior Checklist

Collaboration
- Acknowledged conflict and worked to resolve issues among team members.
- Negotiated solutions and compromises that were acceptable to all team members.
- Was cooperative rather than competitive.
- Respected and capitalized on team members’ diverse knowledge, skills, and abilities.
- Encouraged and accepted points of view different from his or her own.
- Understood the team’s strategies and goals.

Communication
- Demonstrated sensitivity to other team members’ feelings and interests.
- Gave specific and constructive feedback to others.
- Listened attentively to others without interrupting.
- Restated what had been said to show understanding.
- Effectively used facts to get points across to other team members.
- Articulated ideas clearly and concisely.

(continued)
APPENDIX Continued

Decision making
- Anticipated problems and developed contingency plans.
- Analyzed problems from different points of view.
- Helped the team generate alternative solutions.
- Discouraged team members from rushing to conclusions.
- Made decisions based on factual information rather than gut-feel or intuition.
- Solicited input from all team members.

Self-management
- Actively monitored progress to ensure completion according to plan.
- Determined action steps necessary to complete projects.
- Kept the team focused on its task.
- Shared knowledge and expertise with other team members.
- Provided clear direction and defined priorities.
- Helped team members clarify roles and responsibilities.

NOTE: Rating scale: 1 (never) to 5 (always); N = does not apply.

REFERENCES


Peter G. Dominick is a senior associate with Assessment Alternatives, Inc. and is completing his Ph.D. in applied psychology at Stevens Institute of Technology. He is also an adjunct faculty member at Stevens where he has taught courses in group dynamics and human resources. His research interests include team behavior and effectiveness, organizational change, and innovation.

Richard R. Reilly is a professor of management at Stevens Institute of Technology. He received his Ph.D. from the University of Tennessee in 1969 and has been a research psychologist for Educational Testing Service and AT&T. He has published widely in areas such as employee assessment, appraisal, and the effects of feedback programs. His current research interests include team performance and team reward systems.

Jack W. McGourty is a principal of Assessment Alternatives, Inc. A former president and chief operations officer of a national retail organization, he consults with senior managers on such areas as organizational development and team effectiveness. He received his Ph.D. from Stevens Institute of Technology where he is presently conducting research with technology-based organizations on idea generation, innovation, and teamwork.