An Empirical Study Comparing the Effect of Feedback, Training and Executive Coaching on Leadership Behavior Change

Nona Saling
North Carolina State University

Despite the phenomenal growth of executive coaching in recent years, there has been little empirical research on its effectiveness. Executive coaching is typically delivered with 360 degree feedback and training. This study tests whether there are significant differences in leadership behavior change for participants who: received feedback alone; received feedback and attended a leadership training program; and received feedback, attended training and received executive coaching. Preliminary results, conclusions and recommendations are discussed.

Keywords: Leadership; Researching Human Resource Development; Training and Development

Coaching has been one of the fastest growing new trends in leadership development over the past 10 years. Most issues of professional publications in human resources and training and development, like Training (Bolch, 2001), Training & Development Journal (Nowack and Wimer, 1997; Thach and Heinselman, 1999; Witherspoon, White and Hutchinson, 1996), HR Reporter (Peterson and Hicks, 1998), and Human Resource Executive (Patton, 2001, May and Apr; Jossi, 2000) include articles on coaching—defining it, exhorting people to measure its effectiveness, providing guidelines for selecting coaches or relating success stories. Even the popular press, like Newsweek (Hamilton, 1996) and Fortune (Morris, 2000), has reported the phenomenon.

In May 2001, Training magazine reported that there are some 10,000 full and part-time coaches worldwide, according to the International Coach Federation (ICF) (Bolch, 2001). In 1996, Newsweek (Hamilton) reported that the number of coaches in the United States was 1,000. Over a five-year period this sector has grown by a factor of 10. This growth trend is not limited to the boardroom. Training (Bolch, 2001) reported that junior executives and newly promoted managers and those judged to have high potential to succeed are often provided coaches. Although it was more common in the past for coaches to work with managers who were in trouble or likely to “derail,” coaching is now more commonly growth-oriented, helping employees develop skills and competencies that will help them succeed and advance (Bolch, 2001).

Definition and Purposes of Executive Coaching

The roots of coaching are deep and the types of coaching included in the field are many. The sports analogy and tradition are pervasive (Kilburg, 2000). For many years managers have been expected to coach as a vital part of their jobs (Gilley & Boughton, 1996; Peterson & Hicks, 1996). Career or outplacement coaching has been an established process provided for many employees terminated by companies during business lay-offs and downsizing. The term executive coaching has been widely accepted as a way to differentiate a particular type of coaching. There are many definitions of executive coaching (Douglas & Morley, 2000; Hudson, 1999; Smither & Walker, 1995; Witherspoon & White, 1998). One of the most comprehensive was developed in 1996 by Kilburg:

…helping relationship formed between a client who has managerial authority and responsibility in an organization and a consultant who uses a wide variety of behavioral techniques and methods to help the client achieve a mutually identified set of goals to improve his or her professional performance and personal satisfaction and, consequently, to improve the effectiveness of the client’s organization within a formally defined coaching agreement. (Kilburg, 1996)

It should be noted that executive is frequently taken very broadly and could easily mean manager or employee. There are four aspects of this definition that are commonly accepted within the practice of executive coaching (1) coaching is a one-on-one relationship between a coach and a manager or leader, (2) the coach uses a wide variety of behavioral methods in working with the client or coachee, (3) there is a mutually agreed set of goals aimed at improving performance and satisfaction, and (4) this is bounded by a formal coaching agreement (Thach, 2002).

In Four Essential Ways that Coaching Can Help Executives, Witherspoon and White (Witherspoon & White, 1998) describe four distinct purposes of executive coaching: improvement of managerial skills, correcting serious
long-term development, often for a future leadership role in the organization (Witherspoon & White, 1998).

The Coaching Process

Although there are a number of different goals, strategies and styles used in executive coaching, the process usually has four parts: 1) goal-setting, 2) assessment, 3) awareness and action planning, and 4) implementation and monitoring (Douglas & Morley, 2000).

In the initial meeting, goals are set and the coach forms a contract with the client. The assessment phase frequently uses 360-degree feedback, interviews and personality measurements to gather information about the client’s strengths and weaknesses. In the next phase, the coach provides the assessment information to the client in order to build awareness of how they are perceived by others and assists the client in developing a personal behavior change plan. Over a period of six to 12 months, the client implements this action plan with interim discussions and monitoring with the coach (Douglas & Morley, 2000).

Problem Statement

Despite the phenomenal growth of executive coaching, a survey commissioned by Personnel Decisions International reported that only 10 per cent of companies measure the effectiveness of executive coaching (Peterson & Hicks, 1998). There is extensive anecdotal, case study and other qualitative research on executive coaching (Dotlich & Cairo, 1999; M. Goldsmith, Lyons, & Freas, 2000; Hudson, 1999; Kilburg, 2000); however, these studies describe the process of executive coaching, not its effects.

In addition, there are also a number of return-on-investment (ROI) studies (Patton, 2001; Peterson & Hicks, 1998) based on Phillips’ (Phillips, 2003) ROI model of evaluation, which report returns on investment as high as 700 per cent. Unfortunately, these studies are based on unvalidated self-report performance data from executive coaching clients and cannot be considered as valid empirical evidence of the effectiveness of executive coaching.

Limited Empirical Evidence of Effectiveness

There are surprisingly few (Douglas & Morley, 2000; Kilburg, 2000) research studies that document the effectiveness of executive coaching as measured by changed leadership behavior using accepted empirical methods. Of the few studies done, results have been conflicting. Only seven of the nine studies reported have measured effectiveness by examining changed leadership behavior; the other two used other measures (productivity and measures of cognitive learning) to assess the effectiveness of the coaching intervention (Dore, 2001; Olivero, Bane, & Kopelman, 1997). Three of these studies did not subject their results to standard tests of statistical inference (Marshall Goldsmith, 2004; Marshall Goldsmith & Morgan, 2004; Thach, 2002).

In only one of the four remaining empirical studies that examined the effect of executive coaching on leadership behavior (Smither, London, Flutt, Vargas, & Kuchine, 2003) has executive coaching been shown to make a consistently significant difference in changing leadership behavior. Of the other three studies (Luthans & Peterson, 2003; McCauley & Hughes-James, 1994; Young & Dixon, 1996) results were mixed, with significant differences demonstrated on only two-thirds or fewer outcome measures.

In the largest study with over 1,200 participants, the researchers found that coaching had a significant impact on leadership behavior change, but the author himself noted that the size of the effect found was small, given the large number of participants (Smither et al., 2003). The practice of executive coaching continues to gain momentum, despite such scant empirical evidence of its effect on leadership behavior change. Therefore, it is essential to continue to empirically examine whether executive coaching positively impacts leadership behavior change.

Isolating the Effects of Executive Coaching

Additionally, isolating the effect of executive coaching from other interventions, such as 360-degree feedback and training, which quite frequently accompany it, is difficult (Chappelow, 2004). The executive coaching process includes assessment, which in almost all cases, is carried out by administering 360-degree feedback on the leader’s behavior from their superior(s), peers and subordinates (Douglas & Morley, 2000; Peltier, 2001; Smither & Reilly, 2001). Feedback alone has been reported to be an effective motivator of behavior change (Atwater, Roush, & Fischthal, 1995; Johnson & Ferstl, 1999; Reilly, Smither, & Vasilopoulos, 1996; Smither & Walker, 1995; Walker & Smither, 1999).
None of the cited studies has tested whether feedback alone leads to significantly more change in leadership behavior than feedback and coaching does. There are a large number of validated, 360-degree feedback instruments available (Fleenor & Prince, 1997; Leslie & Fleenor, 1998; Tornow, London, & Associates, 1998), so feedback is an increasingly accessible and inexpensive leadership development intervention. It is important to compare the effect of feedback alone and the effect of a full executive coaching intervention on change in leadership behavior in a controlled study.

Furthermore, most often executive coaching engagements include not only these two interventions (360-degree feedback and executive coaching) but also a third intervention, attendance at a leadership development training program, which is designed to help leaders recognize effective and ineffective leadership behaviors, become positively motivated to change their leadership behavior, and set goals for changing leadership behavior. Many of these factors are designed into these training programs precisely because they theoretically or empirically lead to behavior change (Conger, 1992; Guthrie & King, 2004). Studies previously reported in the literature have not compared the effect of feedback and training with that of a full executive coaching engagement on leadership behavior change.

It is important to understand whether there are empirically validated differences between the effect of feedback and training and the effect of feedback, training and executive coaching on leadership behavior change. The following theoretical model explains why there are good reasons to expect that the effect of feedback alone, feedback and training, and feedback, training and coaching would differ.

**Conceptual Framework for the Study**

The Center for Creative Leadership has created a model that identifies three factors that make experiences developmental: assessment, challenge, and support (Van Velsor & McCauley, 2004). This is shown in Figure 2.

**Figure 1. The Center for Creative Leadership leadership development model**


Assessment can be both informal and formal. Informal assessment can include asking colleagues for feedback, observing how others react or feedback received from one’s boss. Formal assessments include: performance appraisals, 360-degree feedback, employee satisfaction surveys, or psychological tests. Assessment data show a gap
between the current state of a leader’s skills and another measure, which could be the perceptions of others or normative data. The gap between the current state and some desired future state provides motivation to change (Van Velsor & McCauley, 2004).

Challenge is present in situations that call for new skills and behaviors that have not been used before, commonly referred to as “moving out of your comfort zone.” There are five different types of challenge: new situations, such as newly promoted managers; difficult goals; conflict situations, such as dealing with conflict with another person or group or reconciling work and family responsibilities; losses, failures and disappointments, such as losing a job or being overlooked for a promotion; and having a wide variety of experiences (Van Velsor & McCauley, 2004).

Support is provided by other people: co-workers, peers, a boss, mentor or coach, sometimes family members. Support helps leaders maintain a positive view of themselves during the trial and error process of trying new behaviors. Support helps build self-efficacy, the belief that one can be successful in accomplishing a goal (Bandura, 1986). It helps leaders stay on course and supporters may provide learning resources through their own expertise or access to other people or assignments (Van Velsor & McCauley, 2004).

The authors state that any developmental experience, a training program, special project assignment, or executive coaching, will be more developmental the more multiple sources of assessment, challenge and support are built in to the experience (Van Velsor & McCauley, 2004). The level of assessment, challenge and support provided in a feedback only intervention can be characterized as low, compared to the others.

Similarly, the degree of the three dimensions in a feedback and training situation can be characterized as moderate, relative to the other two. The level of assessment and support in a feedback, training and coaching application are clearly much higher than the other two. The level of challenge in this latter intervention may range from moderate to high, because there is an increasing trend for more use of executive coaching for development of executives and managers at all levels of the organization (Bolch, 2001; Witherspoon, 2000). These relationships are shown in Figure 2.

Figure 2. Theoretical construct of the study

![Theoretical Construct of the Study](image)

Purpose of the Study

The purpose of this study was to examine the differences in leadership behavior change due to these three different interventions: 360-degree feedback alone; 360-degree feedback and training; and 360-degree feedback, training, and executive coaching. This question tested a basic assumption of the Center for Creative Leadership leadership development model, which is that adding more assessment, challenge and support methods to a developmental experience will increase learning, which will result in enhanced levels of leadership behavior change. Thus, a feedback only intervention has a relatively low amount of assessment, challenge and support methods and will result in a lower level of behavior change. A feedback and training intervention has moderate levels of all three elements, which will result in a moderate level of behavior change. In an executive coaching engagement, additional
assessment, challenge and support processes are added, resulting in a high level of assessment and support and a moderate to high level of challenge, which will lead to the highest level of behavior change.

Research Question

The research question examined in this study was: are there significant differences in the mean change scores in leadership behavior for each of the three following groups:

(A) leaders who received 360-degree feedback only on their leadership behavior from their superiors, peers and subordinates;
(B) leaders who received 360-degree feedback on their leadership behavior from superiors, peers and subordinates; and also attended a five-day leadership training program, designed to promote leadership behavior change;
(C) leaders who: a) received 360-degree feedback on their leadership behavior from superiors, peers and subordinates; b) attended a five-day leadership training program, designed to promote leadership behavior change; and c) received executive coaching, which consisted of ten to 20 hours of one-on-one coach interaction with the client over a six to 12 month period.

Methodology

A quasi-experimental, pre-test post-test design was used for this study to investigate the research question. The three groups in the study were F, FT and FTC: feedback only (F); feedback and training (FT); and feedback, training and coaching (FTC). The dependent variable was change in leadership behavior.

The research design was as follows. In the group that received feedback only (F), participants distributed surveys of their leadership behavior to a group of eight to 15 people, composed of their superior(s), peers and subordinates. They received the aggregated results of the survey in a written, one-page feedback report. For the post-test measure, the leadership behavior surveys were redistributed to the same or a similar group of colleagues six months after the administration of the pre-test. Post-test results were also distributed to participants as a courtesy; however, this was outside the parameters of this study.

In the group receiving feedback and training, participants distributed surveys of their leadership behavior to a group of eight to 15 people, composed of their superior(s), peers and subordinates. They attended a five-day leadership development training program, in which they received the aggregated results of the survey in a written, one-page feedback report. For the post-test measure, the leadership behavior surveys were redistributed to the same or a similar group of colleagues from six to 24 months after the administration of the pre-test. Post-test results were also distributed to participants as a courtesy; however, this was outside the parameters of this study.

In the group that received feedback, training and coaching (FTC), participants distributed surveys of their leadership behavior to a group of eight to 15 people, composed of their superior(s), peers and subordinates as the pre-test measure. They attended a five-day leadership development training program, in which they received the aggregated results of the survey in a written, one-page feedback report, and they also received ten to 20 hours of individualized executive coaching in the six to 12 months following training. For the post-test measure, the leadership behavior surveys were redistributed to the same or a similar group of colleagues from six to 24 months after the administration of the pre-test. Post-test results were also distributed to participants as a courtesy; however, this was outside the parameters of this study.

Population and Sample

The population was 2,837 clients of a medium sized management consulting firm in the Southeast specializing in leadership development and executive coaching, from 2001 to 2004. Clients are typically middle to senior managers of medium to large for-profit businesses ranging in age from 32-57.

The sample of 100 participants represents 4.2 per cent of the population. The participants in this study were selected in the following ways. The 40 participants in the feedback only (F) group were randomly selected from a convenience sample of the firm’s clients who were offered the opportunity to receive feedback, but no training or executive coaching, over a six-month period from June 2004 through May 2005.

The 40 participants in the feedback and training (FT) group were randomly selected from over 1,000 clients of the same client population who received feedback and training from January 2002 through June 2004. The 20 participants in the executive coaching group (FTC) represent all executive coaching clients who agreed to participate in the study from January 2001 through June 2005. All of these clients received feedback, training and coaching.
In addition, age and sex of participants were controlled statistically. Consultants at the firm reported that executive coaching clients (FTC) tended to be older and more frequently males than clients who receive feedback and training (FT). However, these data were not routinely collected as a part of their normal operating procedures. Age, sex and other demographic data was collected in the post-test administration. The size of the quasi-experimental groups was determined by the population of executive coaching clients treated during the time frame of the study.

The consulting firm’s proprietary instrument was used to collect pre-test and post-test leadership behavior data on participants. This instrument is a 28-item Likert-type scale composed of nine sub-scales, shown in Table 1. Each question was measured on a range from zero (strongly disagree) to 10 (strongly agree). Each sub-scale measured commonly accepted leadership constructs, including the following sample items shown in Table 1. Individual items were descriptions of behaviors typical of each construct.

The Leadership Change Score was operationally defined as the difference between the average scores of selected items from the Leadership Impact Assessment instrument in pre and post tests. Change in leadership behavior was the dependent variable in this study and it was calculated by subtracting the average pre-test score on selected items from the average post-test score on selected items.

Table 1 Number of Questions per Sub-scale and Sample Questions

<table>
<thead>
<tr>
<th>Sub-scale Title</th>
<th>Number</th>
<th>Sample Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>Communication</td>
<td>2</td>
<td>Demonstrate effective listening and questioning skills</td>
</tr>
<tr>
<td>Interpersonal Skills</td>
<td>4</td>
<td>Earns trust and loyalty.</td>
</tr>
<tr>
<td>Teamwork</td>
<td>5</td>
<td>Places the team before personal interest.</td>
</tr>
<tr>
<td>Initiative</td>
<td>3</td>
<td>Seeks new challenges and responsibilities.</td>
</tr>
<tr>
<td>Creativity</td>
<td>1</td>
<td>Develops innovative ideas and approaches that work.</td>
</tr>
<tr>
<td>Adaptability</td>
<td>3</td>
<td>Flexible in dealing with different viewpoints and styles.</td>
</tr>
<tr>
<td>Judgment</td>
<td>5</td>
<td>Considers alternatives and consequences before making a judgment.</td>
</tr>
<tr>
<td>Leadership</td>
<td>3</td>
<td>Provides purpose, values, vision.</td>
</tr>
<tr>
<td>Maturity</td>
<td>2</td>
<td>Effectively deals with stress and frustration.</td>
</tr>
</tbody>
</table>

Reliability and Validity

Validity and reliability analyses had not been done on this instrument prior to this study. Approximately half of the previous research studies cited have used instruments which had prior validation studies. The others used unvalidated instruments. Recent research on the reliability of multisource instruments suggests that most of the variance in ratings can be explained by rater idiosyncrasy (Craig & Kaiser, 2003; Yammarino, 2003). This casts doubt on the usefulness of reliability based on previous research studies. Recent research has also shown that the actual factors identified in multisource leadership behavior data rarely match the factors identified in earlier data collections using the same instrument (Craig & Kaiser, 2003). For this reason, a factor analysis was performed on the pre-test data and only items that demonstrated high primary factor loading and low cross-factor loading were selected for data analysis. The procedure described in Craig and Kaiser’s (2003) study of the application of item response theory to multisource ratings was used. This procedure addresses many of the statistical concerns surrounding the reliability of group mean difference scores.

Data Analysis

The research question examined in this study was: were there significant differences in the mean change scores in leadership behavior for each of the three following groups:

(F) leaders who received 360-degree feedback only on their leadership behavior from their superiors, peers and subordinates;

(FT) leaders who received 360-degree feedback on their leadership behavior from superiors, peers and subordinates; and also attended a five-day leadership training program, designed to promote leadership behavior change;
(FTC) leaders who: a) received 360-degree feedback on their leadership behavior from superiors, peers and subordinates; b) attended a five-day leadership training program, designed to promote leadership behavior change; and c) received executive coaching, which consisted of ten to 20 hours of one-on-one coach interaction with the client over a six to 12 month period.

MANOVA (Multivariate Analysis of Variance) was used to test the mean differences in leadership behavior change scores and control for age and sex for the three quasi-experimental groups (Agresti & Finlay, 1997). In addition, the time between administration of the pre-and post-test surveys in the FT and FTC varied from six months to two years. This was also controlled in the multivariate analysis of variance.

Limitations and Delimitations
Conclusions must be limited to clients of medium-sized consulting forms, the population for the study. Because the treatment conditions reflect widespread leadership development practices (Chappelow, 2004), which link the delivery of feedback with training and/or training and coaching, the effects of training and coaching cannot be isolated. Differences in age, sex and time between the administration of the pre-test and post-test were controlled statistically; however, the comparability of the three quasi-experimental groups may be limited. Participants in the FTC group are more likely to be at risk of derailment (Day, 2000; Witherspoon, 2000).

Results, Findings, Conclusions and Recommendations
Preliminary results, findings, conclusions and recommendations will be available in February. Contributions to HRD will be reported with preliminary results, findings, conclusions and recommendations.

References


