Consulting Psychology Journal: Practice and Research
2014, Vol. 66, No. 2, 118–137

1065-9293/14/$12.00 http://dx.doi.org/10.1037/cpb0000005

THE EFFECTIVENESS OF STRENGTH-BASED EXECUTIVE COACHING IN ENHANCING FULL RANGE LEADERSHIP DEVELOPMENT: A CONTROLLED STUDY

Doug MacKie
CSA Consulting, Brisbane, Australia

This study attempts to investigate the effectiveness of a strength-based coaching methodology in enhancing elements of the full range leadership model, especially transformational leadership. Transformational leadership is the process whereby leaders engage and influence their followers toward attaining a shared vision through their capacity to inspire, innovate and personalize their attention. A between-subjects nonequivalent control group design was used to explore the impact of strength-based coaching on transformational and transactional leadership behaviors measured in a 360-degree feedback process. Thirty-seven executives and senior managers from a large not-for-profit organization were nonrandomly assigned to either a coaching or waitlist cohort. The coaching cohort received six sessions of leadership coaching involving feedback on leadership and strengths, goal setting, and strengths development. The coaching protocol was manualized to ensure some methodological consistency between the 11 executive coaches providing the intervention. This involved providing a written manual to each coach and coachee that outlined the required coaching process for each session. After six sessions of coaching over 3 months, cohorts then switched roles. The results showed that participants experienced highly statistically significant increases in their transformational leadership behavior after coaching and this difference was perceived at all levels within the organization but not by the participants themselves. Adherence to the strength-based protocol was also a significant predictor of ultimate degree of change in transformational leadership behavior. The results suggest that strength-based coaching may be effective in the development of transformational leaders.

Keywords: strengths leadership, development leadership, coaching, executive coaching, transformational leadership

The challenges for contemporary leadership in organizations are profound, dynamic, and complex (Youssef & Luthans, 2012). This places increasing demands and expectations on leadership development methodologies to cultivate leaders whose capacity matches these challenges. Executive

Dr. Doug MacKie, CSA Consulting, Brisbane, Australia.
I thank Dr J. Maltby, Dr. S. Harrington, Dr. R. Randall, and Prof D. Clutterbuck for their comments and suggestions on this research.
Correspondence concerning this article should be addressed to Dr. Doug MacKie, L7, 320 Adelaide Street, Brisbane, QLD 4000, Australia. E-mail: doug@csaconsulting.biz

coaching is now one of the dominant methodologies for developing leaders and yet there remains significant debate about what the effective components are, what outcomes can be achieved and what are the qualities of an effective coachee (Grant, Passmore, Cavanagh, & Parker, 2010; Passmore & Fillery-Travis, 2011). This gap in the evidence base is in part a function of the fact that conducting coaching research in organizations presents at least four significant challenges. First, there is no established universal coaching methodology so establishing any kind of consistency in the delivery of executive coaching is problematic and may attenuate the very idiographic focus of coaching that may significantly add to its effectiveness (De Haan & Duckworth, 2013; Passmore & Fillery-Travis, 2011). Second, the range of potential outcomes is vast, making cross-study comparisons nearly impossible. This has led to a focus on process (e.g., goal attainment) rather than content (e.g., leadership skills) and a tendency to focus on self-reported outcomes rather than assessing the broader impact on the organization (Spence, 2007). Third, coachees engage in coaching with a variety of abilities, motivation, and capacity to change. Establishing the effectiveness of executive coaching means establishing what coachee variables predict better outcomes and where coaching resources should be targeted (Best, 2010; Stewart, Palmer, Wilkin, & Kerrin, 2008). Finally, organizations are increasingly dynamic and complex places (Luthans, Luthans, Hodgetts, & Luthans, 2001). Other interventions and initiatives can be going on in parallel, making it difficult to attribute change purely to coaching. Only controlled interventions utilizing objective criteria can disentangle the myriad of contaminating factors and they are difficult and demanding to perform in organizations (Grant et al., 2010).

This article will provide an overview of the existing research on the effectiveness of workplace coaching with a particular emphasis on leadership as a core outcome criterion. It will also position a strength-based methodology as one potential solution to the challenge of methodological heterogeneity and transformational leadership as a generic and desirable coaching outcome.

Evidence for the Effectiveness of Workplace Coaching

The evidence for the effectiveness of executive coaching in the workplace is surprisingly limited. A number of reviews of the effectiveness of executive coaching have been conducted (De Meuse, Dai, & Lee, 2009; Grant et al., 2010; MacKie, 2007; Theeboom, Beersma, & van Vianen, 2014). Grant et al. (2010) found 39 within-subject and 16 between-subjects coaching outcome studies. Of those studies with within-subject designs, there was some evidence that 360-degree feedback and coaching was correlated with enhanced workplace performance (Smither, London, Flautt, Vargas, & Kucine, 2003). In the between-subjects studies reviewed, only 11 of those were randomized. Of those 11, only two were conducted in the workplace. Deviney (1994) used a randomized controlled design with 45 line supervisors and found no difference in their feedback skills following coaching from their managers. Duijts, Kant, van den Brandt, and Swaen (2008) found participants increase on subjective ratings of well being and reduced symptoms of burnout after receiving between 7 and 9 hr of preventative coaching. However, no significant difference was reported on the primary objective measure of sickness absence was found. Grant, Curtayne, & Burton (2009) in a study of 41 executives in a public health agency used a randomized controlled design to assess the effectiveness of 360-degree feedback and four sessions of coaching on enhancing goal attainment, resilience, and well-being. Utilizing a cognitive–behavioral solution focused methodology (where the focus is on thoughts and behaviors that lead to a positive and pragmatic solutions to an issue or challenge), the results showed a subjective perception of an increase in goal attainment, resilience, and well-being and a decrease in stress and anxiety. No objective performance data was gathered and the 360-degree feedback was not repeated after the coaching making the impact on leadership uncertain.

Theeboom et al. (2014) conducted one of the first meta-analytic studies into coaching effectiveness. They screened 107 articles and excluded all cross-sectional studies, those not conducted by a professional external coach and those without any quantitative analysis. Eighteen studies passed the initial selection criteria but only four of those were conducted in the workplace, used a between subjects methodology and collected data other than self-report (Bozer & Sarros, 2012; Cerni, Curtis,

& Colmar, 2010; Grant, Green, & Rynsaardt, 2010; Smither et al., 2003). Average effect sizes for these four studies ranged from 0.08 to 0.36 that would be considered small to medium effect sizes (Cohen, 1988). Smither et al. (2003) examined whether coaching could improve the effect of 360-degree feedback in enhancing performance. Of 1,202 senior managers who received 360-degree feedback, 404 were selected for subsequent coaching. Those who participated in coaching were reported to set more specific goals, solicit more ideas, and improve more in terms of others' ratings. However, despite some positive changes in goal setting and performance (effect size, $d = 0.17$) as measured by supervisor ratings in a repeat 360 feedback process, only 30% of the original participants participated in the coaching and the selection criteria for their inclusion was mixed (some were required to participate by their managers) making the results prone to selection effects and difficult to generalize.

Grant et al. (2010) conducted a further randomized controlled trial of executive coaching in the educational setting. The study again utilized a cognitive–behavioral solution focused approach and randomly assigned 44 teachers to an experimental group receiving 10 coaching sessions or a control group. The results showed a significant increase in goal attainment and well-being and a reduction in stress. The leadership styles inventory 360 also found significant improvements in constructive leadership styles on the self-reports of the coachees over time. Constructive leadership styles include an achievement focus, promoting development in self and others and engaging others in a cooperative and affiliative manner. However, there was no significant difference in constructive leadership when the ratings of their managers and peers were analyzed. This absence of significant change was attributed to rater inconsistency over time. Consequently, data demonstrating the impact of executive coaching on leadership behavior beyond the realms of self-report remains limited.

Bozer et al. (2012) compared a sample of 72 executives on a range of self-ratings of performance and satisfaction as well as ratings of job performance by their supervisor. The results showed that only career satisfaction improved more in the coaching group that in the control group, providing partial support for the effectiveness of coaching interventions. Scores on supervisory-rated job performance were significantly higher after coaching. However, it is of note that this study did not use a 360 methodology and relied instead on one supervisory rating for comparison.

In conclusion, the limited research on workplace coaching has not provided convincing evidence of which coaching methodologies are effective nor converged on generic and relevant evaluation criteria. Data demonstrating the impact of executive coaching beyond the realms of self-report remains sporadic and there remains limited evidence for a significant impact on core outcome criteria like leadership effectiveness. Consequently, clarifying core outcome criteria in coaching and investigating the impact of a specified methodology on external raters would appear to be a crucial step forward.

Leadership Effectiveness as a Core Outcome of Coaching

One way to address the challenge of the multiplicity of potential outcomes in executive coaching is to identify the core criterion for change. Leadership coaching provides this focus with an explicit agenda to positively increase knowledge and enhance effective behavior in the leadership domain (Elliott, 2011). In addition, comparisons across different coaching evaluation studies require some consistency in the measurement and reporting of outcome domains to insure their validity (Grant et al., 2010; MacKie, 2007; Theeboom et al., 2014). Executive coaching that explicitly targets leadership development must by necessity, use reliable and valid measures of leadership behavior that gather data from a wide range of stakeholders to assess the impact of the coaching intervention. This approach to outcomes can circumvent the relatively context specific elements of coaching evaluation and assess the reliability and validity of these measures across different coaching contexts (Bowles, Cunningham, Gabriel, & Picano, 2007). Leadership has been conceptualized and measured in numerous ways but one of the most consistent, researched, and comprehensive models has been the multifactor leadership theory or full range leadership model (FRLM; Bass & Avolio, 1993). This model encompasses both the transformational elements of leadership (that is building trust, acting with principle and integrity, inspiring others, innovating, and developing others), transactional

elements (that includes both constructive elements, e.g., contingent reward and corrective elements and management by exception) and avoidant or laissez-faire leadership behaviors. Transformational leadership was hypothesized to add to the benefits of transactional leadership through the *augmentation effect* where the transformational engagement of followers encouraged their enhanced performance through increased discretionary effort (Bass, 1999; Bass, Avolio, Jung, & Berson, 2003). Thus, the more transactional elements of leadership like goal setting can be enhanced with the addition of transformational elements of leadership where followers are inspired to give more of their time and effort by the vision and charisma of the leader.

The FRLM has begun to be utilized in coaching research as an outcome variable. Cerni et al. (2010) utilized the Multifactor Leadership Questionnaire (MLQ; Bass & Avolio, 1997) in a controlled investigation of the impact of executive coaching on 14 senior school principles. The MLQ offers both a comprehensive theory of leadership ranging from the passive and ineffective to the active and effective elements and a reliable and valid psychometric assessment of the construct (Antonakis, Avolio, & Sivasubramaniam, 2003). The nine elements give rise to the FRLM (Avolio & Bass, 1991) and offer coaching research a robust and targeted broad based leadership construct on which to measure its efficacy.

Cerni et al. (2010) found significant differences in the ratings of transformational leadership in the principals after coaching when rated by the school staff. These changes were apparent at the broad composite transformational leadership level and well as two of the transformational leadership subscales, Idealized influence and Individualized consideration. This suggests both that coaching can enhance elements of transformational leadership and that the MLQ is sensitive to changes in leadership behavior.

Strength-Based Coaching as a Coherent Methodology

The issue of methodological coherence also needs to be addressed to control some of the multiplicity of variables that can occur under a typical coaching methodology. There is increasing evidence that effective leadership coaching requires a core set of common principles at its foundation (Grant, Green & Rynsaardt, 2010; McKenna & Davis, 2009). These include a collaborative working alliance between coach and coachee, the integration of activities to raise self-awareness of the coachee, some clearly defined goals and specific actions to achieve them. However, there remains an ongoing debate about both the relative contributions these core components make to a successful outcome in coaching and the degree to which differing theory and techniques influence successful coaching interventions (De Haan & Duckworth, 2013; MacKie, 2007). The strength-based approach in positive psychology offers both a coherent theoretical framework, empirical validation and a well developed range of reliable and valid psychometric assessment tools that could bring some methodological consistency to the delivery of executive coaching (Kauffman, 2006; Seligman, 2007).

Positive psychology with its emphasis on building on strengths and enhancing confidence and positive emotion is increasingly being applied in an executive coaching context (Biswas-Diener & Dean, 2007). There is growing evidence of the effectiveness of positive psychology interventions in clinical populations (Seligman, Steen, Park, & Petersen, 2005) but to date, the majority of the organizational research has focused on the enhancement of well-being criteria such as mental health and engagement rather than performance criteria like the development of transformational leadership behaviors (Linley, Harrington, & Garcea, 2010; Wood, Linley, Maltby, Kashdan, & Hurling, 2011). Focusing on a strength-based approach to leadership coaching provides the opportunity to test the performance impact of a specific element of the positive psychology paradigm.

A strength-based methodology, however, requires more than just the identification and leverage of strengths in the coachee. Part of the challenge of advocating a strength-based approach is to define exactly what that entails. The construct of what entails a strength has begun to be categorized and has been assessed through a variety of standardized inventories like the Realise2 (Linley & Stoker, 2012), the Values in Action (Peterson, Stephens, Park, Lee, & Seligman, 2010), and the StrengthFinder (Rath & Conchie, 2008). However, only the Realise2 inventory has taken a broader view of

assessment by identifying realized strengths that are known and used, unrealized strengths that are known but underused, learned behaviors where performance has been acquired but is not energizing and weaknesses where both competence and energy are low. How identified strengths are subsequently developed requires some consistency that a similar process is applied across different coaching engagements. Manualization provides a potential solution to the challenge of methodological inconsistency by operationalizing the coaching intervention and provides an objective index of adherence to the coaching protocol. Manualization is the process by which coaches follow a defined written protocol and track coachee behavior within this framework. It also provides a structure for each coaching session that includes reflection on the process, goal attainment, explicit actions, and self-ratings on progress to date that is consistent across different coaches. Manualization also offers the opportunity to be specific and consistent about what is meant by strengths development (Biswas-Diener, Kashdan, & Minhas, 2011) by requiring the coachee to rate themselves on the following four criteria. Strengths awareness for example, undertaking a strength diagnostic and consciously applying strengths to performance issues (Linley, Nielsen, Gillett, & Biswas-Diener, 2010), optimal titration or managing the potential overuse of strengths for example, when confidence becomes arrogance (Kaiser & Kaplan, 2009), pairing strengths with other complimentary competencies for example, technical expertise with building relationships (Zenger, Folkman, & Edinger, 2010) and aligning strengths with the broader business goals and with intrinsic interests for example, linking problem solving with an organizational need for innovation (Govindji & Linley, 2007; Linley, Woolston, & Biswas-Diener, 2009). These four elements form the core of the strength-based leadership coaching protocol.

Rationale and Aims

The limited number of controlled trials in coaching that have been performed to date use a variety of methodologies, draw on differing theoretical orientations, and rely largely on self-report outcome data, making conclusions about effectiveness difficult to generalize (Grant et al., 2010; MacKie, 2007; Theeboom et al., 2014). This study aims to address these three core issues in coaching research by investigating some of the specific active components of executive coaching using a standardized strength-based coaching methodology. Standardization was achieved by way of a manualized strengths-based coaching intervention that explicitly aimed to identify and develop participant's strengths in a leadership development context. The process of standardization is also enhanced by adopting a between subjects design and recruiting subjects from the same organization. Second this study aims to examine the effects of executive coaching on a specific outcome criterion, namely the FRLM that includes transformational, transactional and laissez-faire elements of leadership. This leadership outcome provides 360-degree feedback on changes in leadership behavior throughout the organization and moves the assessment of coaching outcomes beyond the reliance on self-report measures (Ely et al., 2010). By focusing on a specific strength-based methodology, using a reliable and valid measure of transactional and transformational leadership as the dependent variable (DV) and assessing outcomes by way of a 360-degree feedback methodology this study aims to bring some clarity and specificity to the question of coaching efficacy in the workplace.

Hypotheses

The following specific research hypotheses will be addressed in an attempt to address the aims of the study.

1. Leaders who experience strength-based leadership coaching first (Cohort 1) will show a significantly greater increase in transformational leadership behavior than those on the waiting list first (Cohort 2) at Time 2. Equally, Cohort 2 will show a greater increase in transformational leadership behaviors at Time 3. This will address the question of the relative efficacy of workplace coaching over a control group.

2. Both Cohorts 1 and 2 will show a significant increase in transformational leadership behaviors (as reported by all raters) after strength-based leadership coaching at Time 3 as compared to their precoaching leadership scores at Time 1. This will address the issue of the overall effectiveness of workplace coaching over time as perceived beyond the level of self-report.
3. Participants who adhere to the strength-based coaching methodology will show a greater increase in transformational leadership behavior than those who do not adhere. This will test the specific effectiveness of a strength-based methodology in developing transformational leadership behaviors.

Method

Participants

In total, 37 senior managers (17 male, 20 female) were recruited from the same organization in the not-for-profit sector. They were all senior managers and leaders in the Australian arm of a multinational not-for-profit organization. The average age was 45 years (range 31–62 years). This represented all available senior mangers from the top two levels in the organization and included the executive director and the leadership team. In total, 41 individuals were invited to participate but four declined because of overseas postings and maternity leave. Having managerial responsibility for a number of direct reports was a prerequisite of participating in the study. The participants were then divided into two groups—the coaching first group (Cohort 1) and the waitlist first group (Cohort 2). The process of group allocation was not random as it depended on the availability of the participants and the preferences of the organization. All participants gave their written informed consent to participate in the study.

Research Design

The study utilized a nonequivalent control group design with two cohorts: a coaching first group (Cohort 1) and a waitlist first group (Cohort 2). While Cohort 1 was engaged in the coaching, Cohort 2 acted as the control group. Cohorts then switch roles at the midpoint (Time 2). However, because Cohort 1 had had the coaching intervention at this stage, it was not able to act as an independent control group for Cohort 2. Each participant received six sessions (9 hr) of strength-based leadership coaching.

Coaches. In total, 11 coaches provided their services probono for the research. They were highly experienced practitioners who were mainly recruited from the local executive education department of a prestigious business school and had been preselected for both psychological mindedness and business acumen. All coaches were self-employed practitioners who earned a significant part of their income from providing executive coaching services to corporate entities. On average they had 12 years of experience providing executive coaching in organizations and had been working in organizations for an average of 28 years. The majority (70%) were qualified at Masters level or above and were registered practicing psychologists. Each coach was trained in the author's strength-based methodology by way of a half-day training program. This process described the underlying rationale for strength-based approaches to leadership and provided a structured strength-based coaching manual for the coach to follow. The induction particularly focused on the identification of strengths through interview data, 360-degree feedback and the Realise2 inventory. The Realise2 four-quadrant model was also used to give the coaches a format for setting strength-based development goals. The induction also provided the format for strengths development through the sessional rating of strengths awareness, alignment, pairing, and utilization. Each coach provided leadership coaching to between one to two participants per cohort.

Procedure

Strength-based protocol. Each coachee received six 90-min coaching sessions that followed a format articulated in their coaching manual. Initially, coaches began with a strength-based

interview followed by feedback for the coachee on their MLQ 360 report and Realise2 inventory (Linley, Willars, & Biswas-Diener, 2010). The strength-based interview focused on their peak experiences and what energized them about their work. The Realise2 questionnaire provided feedback on what energized them, where they felt competent and where they had the opportunity to apply their strengths. This led to structured feedback on their realized strengths (those that were known and utilized), unrealized strengths (those that were known but underutilized), learned behaviors (those that were competent but not energizing), and weaknesses (where both competence and energy were low). The MLQ 360 provided qualitative and quantitative multirater feedback on their scores on the FRLM that included transformational, transactional, and laissez-faire leadership styles. Coachees were then required to select three goals they would like to focus on during the coaching; a realized strength, an unrealized strength, and a learned behavior or weakness. Coachees then tracked their progress on these goals for the remaining five sessions and committed to actions designed to help their goal attainment. Coachees also tracked their progress on a sessional basis by reflecting on and rating their strength awareness, alignment, pairing, and utilization in their coaching manual.

Measures

Each participant received.

Realise2 strengths inventory. This is an online strengths assessment and development tool that assesses 60 different attributes or strengths in the individual (e.g., curiosity, authenticity, and action). Participants respond on a 7-point Likert scale for each attribute across three dimensions of energy, performance, and use. The responses are then classified into realized strengths, unrealized strengths, learned behaviors, and weaknesses (Linley et al., 2010). The mean Cronbach's α across all 60 attribute item groupings was 0.82 (Linley & Stoker, 2012). Criterion validities for the individual strengths can be found in the manual (Linley & Stoker, 2012).

The MLQ. The MLQ (Bass & Avolio, 1997) is a 49-item questionnaire that measures nine elements of the FRLM namely idealized influence attributes (e.g., Display a sense of power and confidence), idealized influence behavior (e.g., Talk about my most important values and beliefs), inspirational motivation (e.g., Articulate a compelling vision of the future), intellectual stimulation (e.g., Seek different perspectives when solving problems), individualized consideration (e.g., Help others to develop their strengths), contingent reward (e.g., Provide others with assistance in exchange for their efforts), management by exception active (e.g., Keep track of all mistakes), management by exception passive (e.g., Fail to interfere until things become serious), and laissez-faire (e.g., Avoid making decisions). The inventory also has three measures of leadership outcomes; extra effort (e.g., Heighten others' desire to succeed), effectiveness (e.g., Lead a group that is effective), and satisfaction (e.g., Work with others in a satisfactory way) (Bass & Avolio, 1997). It measures all items on a 5-point Likert scale from *not at all* to *frequently if not always*. Cronbach's α for the main transformational leadership factor has been reported as 0.85 (Antonakis et al., 2003) and criterion validities vary for satisfaction (0.71), effectiveness (0.64), and performance (0.27; Judge & Piccolo, 2004).

Each coach received.

Adherence to strengths protocol scale. A 14-item checklist constructed by the author that surveyed the coachee's adherence to the strength-based elements of the protocol including strengths awareness, alignment, pairing, and overuse (e.g., "To what extent did your coachee engage in actions that developed their strengths?"). The scale identified the strength-based elements of the manual and asked the coach to rate adherence to these elements during the coaching on a 5-point Likert scale from *not at all* to *almost always*. In total, protocol adherence score was then calculated. Scores ranged from 14 (*no adherence*) to 70 (*total adherence*). Cronbach's α for this scale was 0.94.

Each coach manual was scored using.

Manual adherence checklist. A 43-item checklist completed by the author that scored the coach's manual according to how much of the strength-based criteria had been completed in the manual. This included for each session, the completion of the identification of strengths, the strengths development plan, tracking of strengths awareness, alignment, pairing, and utilization.

Coach's adherence was measured on a 3-point Likert scale from *not at all*, *somewhat*, to *almost always*, and a total manual adherence score was then calculated.

Data Analysis

The data was analyzed using a variety of parametric statistics. Given that the control group only functioned as such between Times 1 and 2, the goal of the analysis was to conserve as much data as possible at Time 2 to maximize the validity of the findings. Consequently, an analysis of covariance (ANCOVA) was calculated to compare the rate of change on transformational leadership at Time 2 between the two cohorts. The ANCOVA controlled for the fact that the samples were nonrandomized and had nonsignificantly different aggregate scores on transformational leadership at Time 1. In addition, mixed ANOVAs (Tabachnick & Fidell, 2007) were calculated separately at Time 2 and Time 3 to avoid losing the data from those individuals who had dropped out between Time 2 and 3. This allowed analysis of both within-subject changes in the DV over time and between group differences in terms of rate of change on the DV. The five transformational leadership subscales as measured by all raters, were aggregated into one combined transformational leadership score (Bass & Avolio, 1997) to provide an overall index of change. In addition, the different levels of raters (self, boss, direct report, and peer) were combined to maximize the number of observations. The investigation was conducted at the leader level of analysis. In addition, while the data presented herein represents an aggregation of all data levels (self, manager, peer, and direct report) analysis by level showed that all rater groups reported a significant positive change in transformational leadership scores over time. The only exception was self-ratings in Cohort 1 that were significantly higher than all other rater groups at Time 1 and subsequently dropped over time.[1] This is consistent with a general trend in multisource feedback that participant self-ratings become more accurate (i.e., more aligned with the ratings of others) over time as their awareness of the ratings of others increases (Atwater, Brett, & Charles, 2007; Atwater, Waldman, Atwater, & Cartier, 2000).

Results

Descriptive Statistics MLQ Leadership Scores at Time 1

Although 37 participants began the program at Time 1, 6 dropped out (3 from each cohort) before the coaching began and were not included in the analysis. An independent sample *t* test conducted on the data showed that there was no significant difference in the transformational leadership scores at Time 1 for the two cohorts. The only exception was the subscale Idealized Influence Attribute, $t(29) = 2.24, p = .033$. Thus, leaders in Cohort 2 were perceived as more respected, confident, and altruistic that their peers in Cohort 1.

In addition to checking for initial differences in transformational leadership behaviors between the two cohorts, it was also important to check for differences in the transactional, passive, and avoidant styles of leadership. An independent sample *t* test showed that there was no significant difference in the transactional, passive, and leadership outcome scores at Time 1 for the two cohorts. The only exception was effectiveness, $t(29) = 2.07, p = .047$. Thus, leaders in the waitlist first group (Cohort 2) were seen as generally more effective in terms of their leadership effectiveness at the start of the study. Given that Cohort 2 was rated higher at Time 1 in both idealized influence and effectiveness, it may be that the organization had recognized their higher capability by emphasizing the need for others to receive their coaching first in Cohort 1. However, there were only significant differences between cohorts in 2 out of 12 possible criteria at Time 1.

[1] Space precludes a more detailed analysis of the results by rater level. Please contact the author if you would like further information.

The Impact of Rater Consistency on Transformational Leadership Scores

The core aim of the study was to investigate the impact of strength-based leadership coaching on transformational leadership. However, to measure this, it was first necessary to investigate the impact of rater consistency on the results. Rater consistency is an issue in 360 measurement processes as, if original raters leave over time and new raters are added, this can potentially compromise the validity of the study. Table 1 shows the rater consistency over time as well as the number of raters and number of missing data at each time period in this study. Each participant had an average of 9.86, 9.70, and 9.62 rater responses per participant at Time 1, Time 2 and Time 3, respectively. Calculating a ratio for new and original raters at Time 2 and Time 3 gives an index of rater consistency of 92.5% and 88.8%, respectively. A paired sample *t* test of transformational leadership scores with original against new raters at both Times 2 and 3 showed no significant difference in mean total transformational leadership scores. Consequently, the full compliment of raters (original plus new raters) were used for the subsequent analysis.

Hypothesis 1: The Impact of Strength-Based Leadership Coaching on Transformational Leadership

The first analysis was designed to assess whether there was a significant difference in transformational leadership scores after Cohort 1 completed coaching and before Cohort 2 commenced. This is the only stage where the waitlist first cohort (Cohort 2) could act as a real control group for the coaching first cohort (Cohort 1). To assess changes in total transformational leadership ratings, a composite mean was calculated for the five elements of transformational leadership (MLQ5I) using the recommended minimum number of responses by the test author (a minimum of three-quarters responses for each subscale is recommended to calculate a composite mean; Bass & Avolio, 1997).

As this was the only time in which a valid control group comparison could be made, it was important to fully investigate the relative difference in the rate of change between the two cohorts at Time 2. Given that the sample was nonrandomized and there were some nonsignificant differences between the two cohort's MLQ5I scores at Time 1, an ANCOVA was calculated that used the scores on MLQ5I Time1 as the covariate. There were no violations of the assumption of normality, linearity, and homogeneity of variance. After adjusting for preintervention scores at Time 1, there was a significant difference between the MLQ5I scores between the two groups at Time 2, $F(1, 28) = 5.218$, $p = .030$, partial $\eta^2 = 0.157$. This confirmed that the rate of change for Cohort 1 was significantly greater than for Cohort 2: Cohort 1 Wilks' $\lambda = 0.52$, $F(1, 29) = 27.05$, $p < .001$, partial $\eta^2 = 0.483$; Cohort 2 Wilks' $\lambda = 0.86$, $F(1, 29) = 4.81$, $p < .037$, partial $\eta^2 = 0.142$. These within- and between-group effects would be considered large effect sizes (Cohen, 1988).

Given that the rate of change of transformational leadership behavior between the two groups at Time 2 did differ significantly, the size of that difference was further explored using a mixed ANOVA. Table 2 shows that all the elements of transformational leadership improved significantly for Cohort 1 after they had received strength-based leadership coaching although

Table 1
MLQ 360 Rater Information for Both Cohorts Across the Three Time Periods

Time period Number of participants	Time 1 ($N = 37$)	Time 2 ($N = 31$)	Time 3 ($N = 24$)
Total number of raters	395	345	265
Missing data	30	44	34
New raters	0	26	41
Rater consistency	100%	92.5%	88.8%

Table 2
Average and Individual Transformational Leadership Scores for Cohort 1 and 2 Across Time 1 and 2

	Time 1		Time 2					
MLQ scores	Mean	*SD*	Mean	*SD*	*F*	*df*	*p*	*Partial* η^2
Cohort 1 coaching first ($N = 14$)								
MLQ5I	2.63	0.30	2.85	0.35	27.054	1,29	.000	.483
IIA	2.70	0.29	2.95	0.41	18.855	1,29	.000	.394
IIB	2.68	0.39	2.88	0.48	15.765	1,29	.000	.352
IM	2.63	0.31	2.81	0.29	6.423	1,29	.017	.181
IS	2.60	0.28	2.81	0.29	16.703	1,29	.000	.365
IC	2.58	0.43	2.80	0.41	13.847	1,29	.001	.323
Cohort 2 waitlist first ($N = 17$)								
MLQ5I	2.77	0.34	2.85	0.37	4.807	1,29	.037	.142
IA	2.97	0.37	3.05	0.38	2.457	1,29	.039	.139
IIB	2.69	0.41	2.88	0.44	17.393	1,29	.000	.375
IM	2.72	0.44	2.79	0.45	1.420	1,29	.243	.047
IS	2.75	0.35	2.82	0.35	2.456	1,29	.128	.078
IC	2.71	0.39	2.80	0.41	2.641	1,29	.115	.083

Note. Repeated measures mixed ANOVA within group comparison of means between waitlist first group and coaching first group at Times 1 and 2. MLQ5I = average of the five transformational leadership scores; IIA = idealized influence attributes; IIB = idealized influence behavior; IM = inspirational motivation; IS = intellectual stimulation; IC = individualized consideration.

inspirational motivation showed a lower effect size than the other four elements of transformational leadership. The change in the MLQ5I score, which is a composite of the five elements of transformational leadership, significantly increased over time, $t(14) = 4.88$, $p < .001$, two-tailed. An interesting find was that in the waitlist first cohort (Cohort 2), two of the subscales and the composite MLQ5I mean score also reported significant changes in transformational leadership. However, the change was less significant for MLQ5I score than the coaching first group, $t(17) = 2.35$, $p < .032$, two-tailed.

At Time 2, the waitlist first group (Cohort 2) began their coaching whereas Cohort 1 continued to be monitored to see if their gains were maintained or increased over time. Table 3 shows that both groups changed significantly on the mean transformational leadership scores at Time 3 so Hypothesis 1 was only partially supported. Thus, Cohort 1 continued to *enhance* their transformational leadership ratings after their coaching stopped at Time 2 and Cohort 2 showed a significant difference in transformational leadership after their coaching completed at Time 3. The partial η^2 effect sizes for the mean composite transformational leadership scores were very similar (C1 = 0.329, C2 = 0.343) and a mixed ANOVA revealed there was no significant difference in the overall mean subscale score change between the two groups, $F(1, 22) = 0.105$, $p = .749$, partial $\eta^2 = 0.005$, from Time 2 to Time 3.

The Impact of Strength-Based Leadership Coaching on Transactional, Avoidant, and Laissez-Faire Leadership and Leadership Outcomes

As a further test of the impact of strength-based leadership coaching, the impact on the transactional and management by exception elements of the FRLM was also examined. Transactional leadership is still an effective form of leadership that involves setting goals and expectations so this would be predicted to also increase after the leadership coaching. The corrective, passive, and avoidant scales by contrast, would be predicted to decrease as a

Table 3
Individual and Mean Transformational Leadership Scores for Cohort 1 and 2 Across Time 2 and 3

	Time 2		Time 3					
MLQ scores	Mean	*SD*	Mean	*SD*	*F*	*df*	*p*	*Partial* η^2
Cohort 1 coaching first ($N = 10$)								
MLQ5I	2.75	0.30	2.96	0.26	10.808	1,22	.003	.329
IIA	2.88	0.40	3.07	0.30	7.392	1,22	.013	.252
IIB	2.70	0.39	2.97	0.32	14.931	1,22	.001	.404
IM	2.67	0.41	2.94	0.36	9.942	1,22	.005	.311
IS	2.77	0.25	2.94	0.20	5.831	1,22	.025	.210
IC	2.72	0.42	2.91	0.28	4.267	1,22	.051	.162
Cohort 2 waitlist first ($N = 14$)								
MLQ5I	2.94	0.28	3.12	0.24	11.476	1,22	.003	.343
IA	3.16	0.30	3.31	0.28	6.477	1,22	.018	.227
IIB	2.95	0.39	3.08	0.28	4.813	1,22	.039	.180
IM	2.86	0.41	3.07	0.33	8.676	1,22	.007	.283
IS	2.92	0.30	3.08	0.26	7.276	1,22	.013	.249
IC	2.89	0.31	3.08	0.25	6.373	1,22	.019	.225

Note. Repeated measure mixed ANOVA within group comparison of means between waitlist first group and coaching first group at Times 2 and 3. MLQ5I = combined average transformational leadership score; IIA = idealized influence attributes; IIB = idealized influence behavior; IM = inspirational motivation; IS = intellectual stimulation; IC = individualized consideration.

function of the leadership coaching. Finally, the leadership outcomes would be expected to increase as more effective leadership is correlated with others providing greater effort and reporting greater satisfaction and perceived effectiveness (Judge & Piccolo, 2004). Table 4 again shows that these changes for Cohort 1 are all in the expected direction and all the changes are significant apart from management by exception active (MBEA). It is also important to note that the outcomes of leadership are all significantly improved, especially in terms of leadership effectiveness. By contrast the waitlist has only two significant reductions in management by exception active and passive (MBEA and MBEP). There were no significant increases in leadership outcomes in the waitlist first cohort.

The results of the transactional, passive and avoidant subscales of the MLQ were largely consistent with the hypothesis that While transformational leadership would increase as a result of coaching for Cohort 1, the less functional elements of the FRLM would display an inverse effect. Of note is that the contingent reward (CR) element of the model increased significantly in the coaching first group but there was almost no change in the waitlist group. CR is about setting goals and expectations and is positively impacted by the leadership coaching intervention. The one anomaly is that there was almost no change in the monitoring mistakes (MBEA) element in the coaching group whereas this reduced significantly in the waitlist group. Another substantive and supportive finding is that all the outcomes of leadership namely extra effort, effectiveness, and satisfaction significantly improved in the coaching group at Time 2 whereas there was no change at all in the waitlist group. This demonstrates that Cohort 1 participants were rated as more effective and more satisfying leaders and raters were willing to give greater discretionary effort to them after they had received the strength-based leadership coaching.

Given that at Time 2, the waitlist first group (Cohort 2) began their coaching, it was expected that they would show the greater reduction in their passive and avoidant leadership

Table 4
Transactional, Passive, and Avoidant Leadership Styles and Leadership Outcomes for Cohort 1 and 2 Across Time 1 and 2

	Time 1		Time 2					
MLQ scores	Mean	*SD*	Mean	*SD*	*F*	*df*	*p*	*Partial* η^2
Cohort 1 coaching first ($N = 14$)								
CR	2.66	0.38	2.89	0.35	9.122	1,29	.005	.239
MBEA	1.50	0.29	1.43	0.37	0.876	1,29	.357	.029
MBEP	1.11	0.48	0.89	0.48	16.166	1,29	.000	.358
LF	0.85	0.43	0.62	0.35	11.954	1,29	.002	.292
Effectiveness	2.73	0.33	3.01	0.39	17.646	1,29	.000	.378
Satisfaction	2.92	0.32	3.15	0.38	10.551	1,29	.003	.267
Extra effort	2.41	0.34	2.70	0.48	11.415	1,29	.002	.282
Cohort 2 waitlist first ($N = 17$)								
CR	2.78	0.32	2.85	0.42	1.326	1,29	.259	.044
MBEA	1.58	0.38	1.39	0.41	8.708	1,29	.006	.231
MBEP	0.93	0.36	0.79	0.36	7.798	1,29	.009	.212
LF	0.63	0.29	0.57	0.30	0.740	1,29	.397	.025
Effectiveness	2.98	0.35	3.04	0.40	0.894	1,29	.352	.030
Satisfaction	3.10	0.43	3.11	0.40	0.041	1,29	.906	.000
Extra effort	2.61	0.45	2.66	0.35	0.474	1,29	.497	.016

Note. Repeated measures mixed ANOVA within group comparison of means between waitlist first group and coaching first group at Times 1 and 2. CR = contingent reinforcement; MBEA = management by exception active; MBEP = management by exception passive; LF = laissez-faire.

behavior and the greater increase in their leadership outcomes over this time period. However, possibly because of the drop out of three members of this cohort, the mean score increased at Time 2 and this reduced the significance of the size of the increase in CR (rewards achievement) at Time 3. Table 5 illustrates these results. Surprisingly there was also no significant reduction in monitoring mistakes (MBEA) or fighting fires (MBEP) in Cohort 2. There was, however, a significant decrease in avoiding involvement (laissez-faire) and all the three leadership outcomes (effectiveness, satisfaction, and extra effort) increased significantly in Cohort 2 after they had received the coaching intervention. Overall, Cohort 2 was a higher scoring group before the leadership coaching so they may have been approaching a ceiling in terms of some of the scores.

Hypothesis 2: Transformational Leadership Scores and Leadership Outcomes Over Time

The longitudinal MLQ scores were also analyzed to track the trends in both Cohorts over time. Cohort 1 responded positively and significantly immediately after the coaching but also kept increasing at 3 month follow up. Cohort 2 showed a smaller but still significant increase in transformational leadership scores while on the waitlist for their leadership coaching.

A three way repeated measures mixed ANOVA showed that both groups increased their scores on transformational leadership at a very similar rate between Times 1 and 3: Cohort 1, $F(2, 21) = 10.51, p = .001$, partial $\eta^2 = 0.50$; Cohort 2, $F(2, 21) = 8.92, p = .002$, partial $\eta^2 = 0.46$. These would be considered very significant effect sizes in terms of increasing transformational leadership ratings over time (Cohen, 1988). There was no significant difference in the rate of change between the two cohorts from Time 1 to Time 3 after both groups have received the coaching intervention.

Table 5
Transactional, Passive, and Avoidant Leadership Styles and Leadership Outcomes for Cohort 1 and 2 Across Time 2 and 3

	Time 2		Time 3					
MLQ scores	Mean	*SD*	Mean	*SD*	*F*	*df*	*p*	*Partial* η^2
Cohort 1 coaching first ($N = 10$)								
CR	2.86	0.35	3.02	0.39	2.947	1,22	.100	.118
MBEA	1.49	0.38	1.45	0.44	0.076	1,22	.785	.003
MBEP	0.80	0.38	0.68	0.27	3.701	1,22	.067	.144
LF	0.57	0.36	0.55	0.25	0.075	1,22	.787	.003
Effectiveness	2.99	0.44	3.16	0.22	2.787	1,22	.109	.112
Satisfaction	3.11	0.38	3.18	0.24	0.455	1,22	.507	.020
Extra effort	2.54	0.37	2.78	0.27	7.387	1,22	.013	.251
Cohort 2 waitlist first ($N = 14$)								
CR	2.96	0.31	3.03	0.24	0.913	1,22	.350	.040
MBEA	1.38	0.36	1.38	0.53	0.004	1,22	.949	.000
MBEP	0.73	0.32	0.65	0.27	1.889	1,22	.183	.079
LF	0.51	0.27	0.33	0.11	9.167	1,22	.006	.294
Effectiveness	3.14	0.36	3.39	0.31	9.490	1,22	.005	.301
Satisfaction	3.23	0.29	3.42	0.27	4.961	1,22	.036	.184
Extra effort	2.77	0.27	3.04	0.33	14.283	1,22	.001	.394

Note. Repeated measures mixed ANOVA within group comparison of means between waitlist first group and coaching first group at Times 2 and 3. CR = contingent reinforcement; MBEA = management by exception active; MBEP = management by exception passive; LF = laissez-faire.

The changes in MLQ outcome scores are represented in Figures 1 and 2. Figure 1 shows that for Cohort 1, the most significant changes occur immediately after the coaching has occurred at Time 2. Coachees are seen as significantly more effective, more satisfying and more inspiring in terms of influencing followers to donate their extra discretionary effort. There were no significant changes in the waitlist cohort on leadership outcomes at Time 2. The same trend is

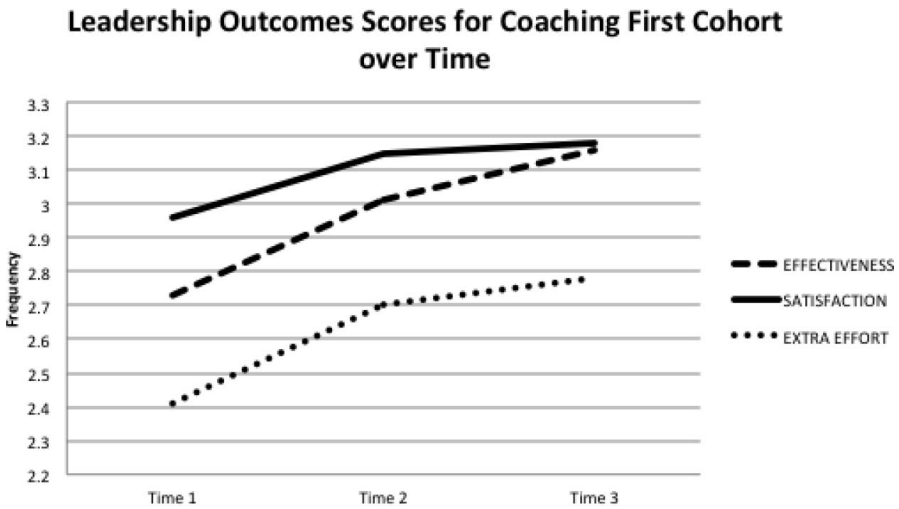

Figure 1. Longitudinal changes in MLQ leadership outcomes in coaching first cohort over time (scores: 0 = *not at all* to 4 = *frequently if not always*).

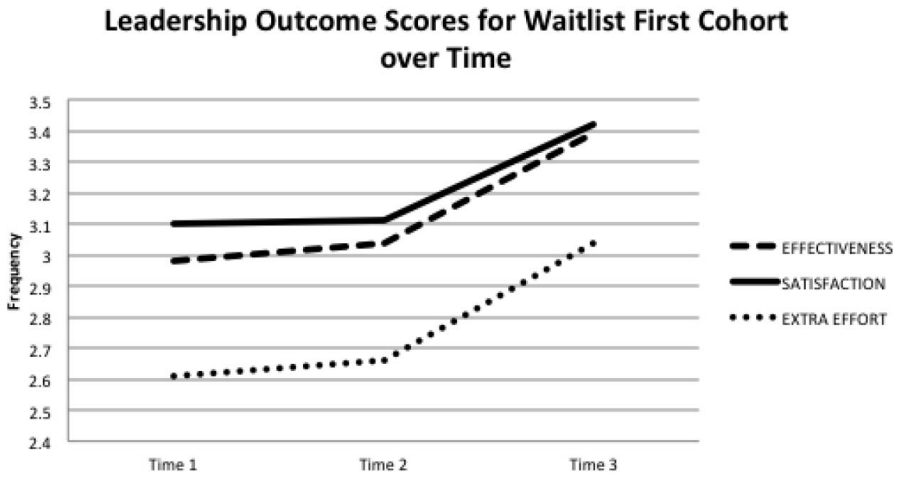

Figure 2. Longitudinal changes in MLQ leadership outcomes in waitlist first cohort over time (scores: 0 = *not at all* to 4 = *frequently if not always*).

found in Figure 2 for the waitlist first group at Time 3 immediately after their coaching. A mixed repeated measures ANOVA was calculated to assess differences in the rate of change on leadership outcomes between the two cohorts at Time 2. The results showed there was a significant difference in the leadership outcome scores between the two groups at Time 2 in rating of effectiveness, $F(1, 28) = 6.128$, $p = .019$, partial $\eta^2 = 0.174$, and satisfaction, $F(1, 28) = 5.405$, $p = .027$, partial $\eta^2 = 0.157$, but the differences extra effort were just short of being significant, $F(1, 28) = 4.159$, $p = .051$, partial $\eta^2 = 0.125$.

Hypothesis 3: Leadership Change as a Function of Adherence to Strengths Protocol

Hypothesis 3 predicted that those who adhere more closely to the strength-based methodology would show greater change in their transformational leadership scores. This hypothesis attempts to test the active ingredients of leadership coaching and to establish that the focus on strengths is a mediator or moderator of positive leadership change. To test this hypothesis, two indices of adherence were calculated. First, each coach was asked to complete a short questionnaire indicating how much they and the coachee had adhered to the strength-based approach during the coaching process. Second, the participant manuals were scored according to how much of the strength-based criteria had been completed. Thus, the perspective of both coach and coachee were used to measure protocol and manual adherence, respectively. Analysis was by way of linear regression analysis. Table 6 shows the results.

The data in Table 6 indicate that both manual adherence and coach adherence to the strength-based protocol are significant predictors of final transformational leadership scores at Time 3. This suggests that adherence to the strengths based component of the leadership

Table 6
The Unstandardized and Standardized Regression Coefficients for Manual Adherence and Coach Adherence to the Strengths Protocol as Predictors of Final MLQ5IT3 Score

Variable	*B*	*SE* B	β	*p*
Manual adherence	.004	.002	.380	.024
Coach adherence	.017	.006	.408	.016

coaching played a significant role in increasing the transformational leadership scores of participants.

Discussion

In this study, the effects of strength-based leadership coaching on transformational leadership were explored using a nonequivalent between-subjects controlled design. It is a major strength of the study that it received almost total participation from the top two layers of management within the participating organization. In the first instance, a controlled nonequivalent design was utilized to see if leaders in the coaching first cohort (Cohort 1) who received strength-based executive coaching first, received higher ratings in transformational leadership than leaders in Cohort 2 who acted as their waitlist control group. At the midway point, the two groups swapped roles and the second cohort (waitlist first) received their leadership coaching while the first group was monitored to see if their gains were maintained at follow up. The results clearly demonstrated a significant increase in the other-rater feedback on transformational leadership behaviors after strength-based executive coaching with an effect size more than three times greater in the intervention group versus the control group at the midpoint of the study.

It is important to note that even being on the waitlist (Cohort 2) and anticipating their coaching was enough to produce a significant increase in transformational leadership ratings by all others. This perceived change in leadership behavior may have been mediated by participant expectations regarding the forthcoming coaching program and participating (but not receiving the results of) a multisource feedback process. Although this increase was much smaller than the experimental effect, it does suggest that completing a 360-feedback process even without the debrief and anticipating future executive coaching is an intervention in itself (albeit of significantly less potency than a structured strength-based coaching intervention). This increase was, however, found only in the other ratings of transformational leadership subscales at Time 2. The self-ratings of transformational leadership and leadership outcomes did not change for Cohort 2 while on the waitlist. This finding is consistent with Grant et al. (2009) who found no change on a range of self-report measures from goal-attainment, stress, and resilience in participants on a waitlist before receiving cognitive–behavioral coaching.

Subsequently, both groups significantly increased on transformational leadership between Times 2 and 3 and demonstrated very similar effect sizes. This may be a function of Cohort 2 approaching a ceiling in the scoring as they began the study with higher scores and still managed significant differences. It may also be a function of Cohort 1 maintaining and enhancing the gains they made after coaching through the continued application of skills and techniques acquired through the coaching process. This is not a surprising finding given that coaching aims to increase an executive's effectiveness long after the coaching has concluded. This post-intervention increase in leadership effectiveness also suggests skills acquired in the coaching process has been effectively transferred and applied to the workplace.

It is also apparent that other elements of the FRLM changed significantly as a function of the coaching process. In Cohort 1, participants significantly increased their transactional leadership after their coaching intervention and significantly reduced their corrective (management by exception) and laissez-faire styles of leadership. Cohort 1 also significantly increased their leadership outcomes indicating that others perceived them to be more effective, reported greater satisfaction with their leadership and were willing to provide extra discretionary effort as a consequence. Cohort 2 showed a somewhat different pattern after receiving their coaching intervention. There was no significant increase in their levels of transactional leadership but they did significantly reduce their levels of laissez-faire leadership style and demonstrated significant increases on the three outcomes of leadership, effectiveness, satisfaction, and extra effort. Thus, it is apparent that leadership coaching reduces the dysfunctional elements of leadership as well as enhances the more functional elements by increasing both transactional and transformational leadership behaviors.

Despite some turnover of raters during the study, there was no significant difference on the ratings of transformational leadership from original and new raters. Rater turnover is a rarely

reported issue in the literature and yet it is often assumed that rater consistency is key to successfully and reliably mapping changes in leadership behavior over time (Grant et al., 2010). The results in this study suggest that leader behavior is consistent, independent of individual raters, and that moderate turnover in raters over time does not adversely impact the reliability of the mean ratings. This represents important confirmation of the validity of a multirater methodology in assessing leadership outcomes in executive coaching.

Finally, the importance of the strength-based methodology in mediating increases in transformational leadership was assessed. Both perspectives of methodological adherence, that is manual and coach adherence, showed a significantly positive prediction of transformational leadership scores at Time 3. This confirms that adherence to a strength-based protocol predicts enhanced leadership performance but it cannot tell us whether a strength-based approach is superior to other structured methodologies or indeed which specific elements of the strength-based protocol (strengths identification, goal setting, strengths development, etc.) were most effective in increasing transformational leadership behaviors. However, contrasting the effect sizes obtained in this study with those reported by Theeboom et al. (2014), these results would be at the very top end of effects reported in their meta-analysis. This suggests that a structured strength-based methodology is highly effective at enhancing transformational leadership behavior and moreover, these changes are perceived at al levels in the organization. The manualization process also gives us for the first time, clarity into exactly what occurs in executive coaching and once alternative approaches are manualized, future research will provide data on the comparative efficacy of different coaching methodologies.

This results sits at the nexus of calls for more randomized controlled trials to conclusively demonstrate the effectiveness of coaching (Passmore & Fillery-Travis, 2011) versus suggestions that effectiveness has been sufficiently demonstrated and it is now time to compare conditions and methodologies to find the optimum blend of critical components of effective coaching (De Haan & Duckworth, 2013). However, the comparative methodologies approach currently tends to focus on coach and coachee variables rather than theoretical distinctions because of the adoption of the common factors approach. This asserts that there are common factors like the quality of the coaching relationship between coach and coachee that are much more predictive of outcomes that theoretical orientation. This contentious concept originally derived from psychotherapy outcome studies, has been adopted somewhat uncritically into the coaching research literature and has significant implications for the direction of coaching research especially in the delineation and testing of different coaching methodologies (MacKie, 2007).

Practical Implications

This study has a number of practical implications for both researchers and practitioners in the leadership coaching arena. First, it supports the notion that leadership can be significantly enhanced in a relatively short period of time through a structured multirater feedback and executive coaching process (Kelloway & Barling, 2000; Walumba & Wernsing, 2013). Second it supports the utility of the FRLM as a sensitive and discriminating mechanism for tracking changes in perceived leadership effectiveness over time. This gives executive coaching a reliable and valid outcome criterion that can be used both to demonstrate increasing leadership expertise and connect with performance criterion that may be more challenging to measure and attribute to individual performance. This utilization of multirater feedback that is grounded in leadership theory is an important confirmation of prior research that has utilized transformational leadership as a DV (Cerni et al., 2010) and addresses the overreliance of self-reports measures in assessing leadership coaching outcomes (Dunning, Johnson, Ehrlinger, & Kruger, 2003; Grant et al., 2010).

In terms of gathering objective outcome data to demonstrate coaching effectiveness, this study supports the use of the multirater methodology in providing feedback on behavioral change from a variety of levels with the organization. This study has also shown that minor rater turnover did not influence the results in any significant way suggesting that behaviors were consistently observed, independent of raters. This is an important finding as many research approaches do not use a multirater methodology and rely of self-report instead, partly because of the challenge of maintaining rater consistency over time (Grant et al., 2010).

Finally, the results do support the notion that a structured methodology is beneficial to the leadership coaching process. There was a strong predictive effect of methodological adherence to final changes in leadership behavior. As this methodology was primarily focused around the identification and development of strengths and talents, it is reasonable to assume in this case that the strength-based methodological adherence was a significant moderator of the improvement in transformational leadership. However, until this is contrasted with another equally coherent and structured methodology in a between subjects design, the relative efficacy of the strengths based component will be unknown.

Limitations of the Study

The study used a nonequivalent between-subjects design that utilized a control group to assess the impact of a leadership coaching intervention on transformational leadership behaviors. It was not possible to randomly assign subjects to each cohort as the availability of participants as the logistical needs of the organization took precedence. Despite this nonrandomization, however, there was no significant difference between the two cohorts at Time 1 in the aggregated transformational leadership scores (MLQ5I) suggesting the allocation of participants did not unduly influence the study. However, the between subjects design only allowed the first cohort to be fully controlled as at Time 2 when the two cohorts crossed over, Cohort 1 had already had the intervention and could no longer function as an independent control group. Ideally, a third group would have been utilized as a control group throughout the study and only offered the intervention as the end of the research process. However, the number of eligible participants and the requirements of the organization to urgently offer leadership development to their senior managers prevented this.

Second, the utilization of only one methodology prevented conclusions being drawn about the differential effectiveness of two contrasting interventions. Ideally, the strength-based methodology would have been contrasted with another equally coherent methodology to delineate the relative efficacy of the specific factors in the intervention. Solution focused approaches (Spence et al., 2008) or interventions based on acceptance and commitment theory (Bond, Hayes, & Barnes-Holmes, 2006) could offer such comparisons. This would also allow further delineation of the relative importance of common versus specific factors in leadership coaching outcomes (De Haan & Duckworth, 2013).

Third the sample size was relatively small and the drop out rate of 15% could potential limit the generalizability of the study. The sample was also from the not-for-profit sector although given this size of the organization, the organizational structure (with traditional functions and reporting line), and the challenges they faced, this is unlikely to affect generalization to other sectors. A larger sample size would help address some of these issues and allow the further investigation of coach and coachee variables like developmental readiness, that may well be mediating and moderating coaching outcomes (Hannah & Avolio, 2010).

Finally, using transformational leadership as the DV is only one possible measure of leadership effectiveness. However, it is the most researched theory over the last 30 years and has established significant correlations between increases in transformational leadership and objective performance outcomes including financial performance, job satisfaction, follower satisfaction, and organizational commitment (Avolio, 2011). While the emphasis on "heroic" leadership within the transformational leadership model has been criticized (Alimo-Metcalfe et al., 2008), it still remains one of the most validated and researched leadership assessment tools available (Avolio, 2011).

Further Research

This study embraced a positivist, nomothetic, and inductive approach to the investigation of the effectiveness of leadership coaching as a means to enhance transformational leadership behavior. There are significant assumptions in such an approach that need to be made explicit. A core assumption is that there is sufficient communality among participants that warrants a common methodology and assessment process. Such assumptions promote the use of cross-research comparisons but may ignore or omit the more idiographic elements of individual coachee characteristics. Coaching research is proceeding down both idiographic and nomothetic approaches and clearly both

are required to fully comprehend the process of effective coaching. The exact point of transition from one approach to the other and how the two interact is the cause of some healthy tension within the profession (De Haan & Duckworth, 2013).

In addition this research highlights the need for more comparative coaching research where different theoretical methodologies are compared in a between-subjects design. Contrasting different theory-based methodologies would provide comparative data on the relative effectiveness of coaching models primarily focused around strengths, solutions, or other key potential mediators of leadership effectiveness. Finally, the concepts of change and developmental readiness require further investigation to further explore coachee variables in maximizing outcomes from the coaching process. The underlying constructs require to be operationalized into reliable and valid psychometrics so that their predictive utility can be formally tested in a leadership coaching context.

Conclusion

This research is one of the first controlled studies using a between subjects design to show significant changes in transformational leadership behavior following workplace executive coaching, that are perceived beyond the level of self-report. It confirms the trainability of transformational leadership and emphasizes the efficacy of individual executive coaching as an effective leadership development methodology. This research sits in the context of promoting a more balanced and constructive perspective on positive leadership development that supports the identification and development of strengths as a core element in developing leadership capability. Second, it supports the notion that methodology matters and that a structured and systematic approach to the provision of executive coaching significantly predicts the enhancement of leadership behaviors. Further research is required to help identify both the key individual variables that predict positive outcomes in leadership development and identify the relative importance of common versus specific factors in leadership coaching methodology. This will help refine and enhance what are already effectiveness mechanisms for developing leaders capable of navigating the complexity of contemporary organizations.

References

Alimo-Metcalfe, B., Alban-Metcalfe, J., Bradley, M., Mariathasan, J., & Samele, C. (2008). The impact of engaging leadership on performance, attitudes to work and wellbeing at work: A longitudinal study. *Journal of Health Organization and Management, 22,* 586–598. doi:10.1108/14777260810916560

Antonakis, J., Avolio, B. J., & Sivasubramaniam, N. (2003). Context and leadership: An examination of the nine-factor full-range leadership theory using the Multifactor Leadership Questionnaire. *The Leadership Quarterly, 14,* 261–295. doi:10.1016/S1048-9843(03)00030-4

Atwater, L. E., Brett, J. F., & Charles, A. C. (2007). Multisource feedback: Lessons learned and implications for practice. *Human Resource Management, 46,* 285–307. doi:10.1002/hrm.20161

Atwater, L. E., Waldman, D. A., Atwater, D., & Cartier, P. (2000). An upward feedback field experiment: Supervisors' cynicism, reactions, and commitment to subordinates. *Personnel Psychology, 53,* 275–297. doi:10.1111/j.1744-6570.2000.tb00202.x

Avolio, B. J. (2011). *Full range leadership development*. Thousand Oaks, CA: SAGE Publications, Incorporated. doi:10.4135/9781483349107

Avolio, B. J., & Bass, B. M. (1991). *The full range leadership development programs: Basic and advanced manuals*. Binghamton, NY: Bass, Avolio & Associates.

Avolio, B. J., & Gardner, W. L. (2005). Authentic leadership development: Getting to the root of positive forms of leadership. *The Leadership Quarterly, 16,* 315–338. doi:10.1016/j.leaqua.2005.03.001

Bass, B. M. (1999). Two decades of research and development in transformational leadership. *European Journal of Work and Organizational Psychology, 8,* 9–32. doi:10.1080/135943299398410

Bass, B. M., & Avolio, B. J. (1993). Transformational leadership: A response to critiques. In M. M. Chemers & R. Ayman (Eds.), *Leadership theory and research: Perspectives and directions* (pp. 81–107). Sydney, Australia: Academic Press Inc.

Bass, B. M., & Avolio, B. J. (1997). *Full range leadership development: Manual for the Multifactor Leadership Questionnaire*. Palo Alto, CA: Mind Garden Inc.

Bass, B. M., Avolio, B. J., Jung, D. I., & Berson, Y. (2003). Predicting unit performance by assessing transformational and transactional leadership. *Journal of Applied Psychology, 88,* 207–218. doi:10.1037/0021-9010.88.2.207

Best, K. C. (2010). Assessing leadership readiness using developmental personality style: A tool for leadership coaching. *International Journal of Evidence Based Coaching and Mentoring, 8,* 22–38.

Biswas–Diener, R., & Dean B (2007). *Positive psychology coaching: Putting the science of happiness to work for your clients*. New York, NY: Wiley.

Biswas-Diener, R., Kashdan, T. B., & Minhas, G. (2011). A dynamic approach to psychological strength development and intervention. *Journal of Positive Psychology, 6,* 106–118. doi:10.1080/17439760.2010.545429

Bond, F. W., Hayes, S. C., & Barnes-Holmes, D. (2006). Psychological flexibility, ACT, and organizational behavior. *Journal of Organizational Behavior Management, 26,* 25–54. doi:10.1300/J075v26n01_02

Bowles, S., Cunningham, C. J., Gabriel, M., & Picano, J. (2007). Coaching leaders in middle and executive management: Goals, performance, buy-in. *Leadership & Organization Development Journal, 28,* 388–408. doi:10.1108/01437730710761715

Bozer, G., & Sarros, J. C. (2012). Examining the effectiveness of executive coaching on coachees' performance in the Israeli context. *International Journal of Evidence Based Coaching and Mentoring, 10,* 14–32.

Cerni, T., Curtis, G. J., & Colmar, S. H. (2010). Executive coaching can enhance transformational leadership. *International Coaching Psychology Review, 5,* 81–84.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. London: Routledge.

De Haan, E., & Duckworth, A. (2013). Signalling a new trend in executive coaching outcome research. *International Coaching Psychology Review, 8,* 6–19.

De Meuse, K. P., Dai, G., & Lee, R. J. (2009). Evaluating the effectiveness of executive coaching: Beyond ROI? *Coaching: An international Journal of Theory, Research and Practice, 2,* 117–134.

Deviney, D. E. (1994). The effects of coaching using multiple rater feedback to change supervisor behavior. *Dissertation Abstracts International Section A, 55,* 114.

Duijts, S. F., Kant, I., van den Brandt, P. A., & Swaen, G. M. (2008). Effectiveness of a preventive coaching intervention for employees at risk for sickness absence due to psychosocial health complaints: Results of a randomized controlled trial. *Journal of Occupational and Environmental Medicine, 50,* 765–776. doi:10.1097/JOM.0b013e3181651584

Dunning, D., Johnson, K., Ehrlinger, J., & Kruger, J. (2003). Why people fail to recognize their own incompetence. *Current Directions in Psychological Science, 12,* 83–87. doi:10.1111/1467-8721.01235

Elliott, R. (2011). Utilizing evidence-based leadership theories in coaching for leadership development: Towards a comprehensive integrating conceptual framework. *International Coaching Psychology Review, 6,* 46–70.

Ely, K., Boyce, L. A., Nelson, J. K., Zaccaro, S. J., Hernez-Broome, G., & Whyman, W. (2010). Evaluating leadership coaching: A review and integrated framework. *The Leadership Quarterly, 21,* 585–599. doi:10.1016/j.leaqua.2010.06.003

Govindji, R., & Linley, P. A. (2007). Strengths use, concordance and well-being: Implications for strengths coaching and coaching psychologists. *International Coaching Psychology Review, 2,* 143–153.

Grant, A. M., Curtayne, L., & Burton, G. (2009). Executive coaching enhances goal attainment, resilience and workplace well-being: A randomized controlled study. *The Journal of Positive Psychology, 4,* 396–407. doi:10.1080/17439760902992456

Grant, A. M., Green, S. L., & Rynsaardt, J. (2010). Developmental coaching for high school teachers: Executive coaching goes to school. *Consulting Psychology Journal: Practice and Research, 62,* 151–168. doi:10.1037/a0019212

Grant, A. M., Passmore, J., Cavanagh, M. J., & Parker, H. M. (2010). The state of play in coaching today: A comprehensive review of the field. *International Review of Industrial and Organizational Psychology, 25,* 125–167.

Hannah, S. T., & Avolio, B. J. (2010). Ready or not: How do we accelerate the developmental readiness of leaders? *Journal of Organizational Behavior, 31,* 1181–1187. doi:10.1002/job.675

Judge, T. A., & Piccolo, R. F. (2004). Transformational and transactional leadership: A meta-analytic test of their relative validity. *Journal of Applied Psychology, 89,* 755–768. doi:10.1037/0021-9010.89.5.755

Kaiser, R. B., & Kaplan, R. E. (2009). When strengths run amok. In R. B. Kaiser (Ed.), *The perils of accentuating the positive* (pp. 57–76). Tulsa, OK: Hogan Press.

Kauffman, C. (2006). Positive psychology. The science at the heart of coaching. In D. Stober & A. M. Grant (Eds.), *Evidence based coaching handbook* (pp. 193–225). New York, NY: Wiley.

Kelloway, E. K., & Barling, J. (2000). What we have learned about developing transformational leaders. *Leadership & Organization Development Journal, 21,* 355–362. doi:10.1108/01437730010377908

Linley, A., & Stoker, H. (2012). Technical manual and statistical properties for Realise2. Coventry: Centre of Applied Positive Psychology.

Linley, P. A., Harrington, S., & Garcea N. (2010). *Oxford handbook of positive psychology and work*. Oxford: Oxford University Press.

Linley, P. A., Nielsen, K. M., Gillett, R., & Biswas-Diener, R. (2010). Using signature strengths in pursuit of goals: Effects on goal progress, need satisfaction, and well-being, and implications for coaching psychologists. *International Coaching Psychology Review, 5,* 6–15.

Linley, P. A., Willars, J., & Biswas-Diener, R. (2010). *The strengths book*. Coventry, United Kingdom: CAPP Press.

Linley, P. A., Woolston, L., & Biswas-Diener, R. (2009). Strengths coaching with leaders. *International Coaching Psychology Review, 4,* 37–48.

Luthans, F., & Avolio, B. J. (2003). Authentic leadership: A positive developmental approach. In K. S. Cameron, J. E. Dutton, & R. E. Quinn (Eds.), *Positive organizational scholarship* (pp. 241–261). San Francisco, CA: Barrett-Koehler.

Luthans, F., Luthans, K. W., Hodgetts, R. M., & Luthans, B. C. (2001). Positive approach to leadership (PAL): Implications for today's organizations. *Journal of Leadership Studies, 8,* 3–20. doi:10.1177/107179190100800201

MacKie, D. (2007). Evaluating the effectiveness of executive coaching: Where are we now and where do we need to be? *Australian Psychologist, 42,* 310–318. doi:10.1080/00050060701648217

McKenna, D., & Davis S. L. (2009). Hidden in plain sight: The active ingredients of executive coaching. *Industrial and Organizational Psychology, 2,* 244–260. doi:10.1111/j.1754-9434.2009.01143.x

Passmore, J., & Fillery-Travis, A. (2011). A critical review of executive coaching research: A decade of progress and what's to come. *Coaching: An International Journal of Theory, Research and Practice, 4,* 70–88.

Peterson, C., Stephens, J. P., Park, N., Lee, F., & Seligman, M. E. P. (2010). Strengths of character and work. In P. A. Linley, S. Harrington, & N. Garcea. (Eds.), *Oxford handbook of positive psychology and work* (pp. 221–231). Oxford: Oxford University Press.

Rath, T., & Conchie, B. (2008). *Strengths based leadership*. New York, NY: Gallup Press.

Seligman, M. E. P. (2007). Coaching and positive psychology. *Australian Psychologist, 42,* 266–267. doi:10.1080/00050060701648233

Seligman, M. E. P., Steen, T. A., Park, N., & Petersen, C. (2005). Positive psychology progress: Empirical validation of interventions. *American Psychologist, 60,* 410–421. doi:10.1037/0003-066X.60.5.410

Smither, J. W., London, M., Flautt, R., Vargas, Y., & Kucine, I. (2003). Can working with an executive coach improve multi-source feedback ratings over time? A quasi-experimental field study. *Personnel Psychology, 56,* 23–44. doi:10.1111/j.1744-6570.2003.tb00142.x

Spence, G. B. (2007). GAS powered coaching: Goal Attainment Scaling and its use in coaching research and practice. *International Coaching Psychology Review, 2,* 155–167.

Spence, G. B., Cavanagh, M. J., & Grant, A. M. (2008). The integration of mindfulness training and health coaching: An exploratory study. *Coaching: An International Journal of Theory, Research and Practice, 1,* 145–163.

Stewart, L. J., Palmer, S., Wilkin, H., & Kerrin, M. (2008). The influence of character: Does personality impact coaching success. *International Journal of Evidence Based Coaching and Mentoring, 6,* 32–42.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics*. Toronto, Ontario: Pearson Education.

Theeboom, T., Beersma, B., & van Vianen, A. E. (2014). Does coaching work? A meta-analysis on the effects of coaching on individual level outcomes in an organizational context. *The Journal of Positive Psychology, 9,* 1–18. doi:10.1080/17439760.2013.837499

Walumba, F. O., & Wernsing, T. (2013). From transactional and transformational to authentic leadership. In M. G. Rumsey (Ed.), *The Oxford handbook of leadership* (pp. 392–400). Oxford: Oxford Press.

Wood, A. M., Linley, P. A., Maltby, J., Kashdan, T. B., & Hurling, R. (2011). Using personal and psychological strengths leads to increase in well-being over time: A longitudinal study and the development of the Strengths Use Questionnaire. *Personality & Individual Differences, 50,* 15–19. doi:10.1016/j.paid.2010.08.004

Youssef, C. M., & Luthans, F. (2012). Positive global leadership. *Journal of World Business, 47,* 539–547. doi:10.1016/j.jwb.2012.01.007

Zenger, J. H., Folkman, J. R., & Edinger, S. K. (2011). Making yourself indispensable. *Harvard Business Review, 89,* 84–90.

Received January 16, 2014
Latest revision received April 21, 2014
Accepted April 22, 2014 ■