

CAN CHARISMA BE TAUGHT? TESTS OF TWO INTERVENTIONS

JOHN ANTONAKIS
University of Lausanne
Faculty of Business and Economics
Internef #618
Lausanne 1015
Switzerland

MARIKA FENLEY
University of Lausanne
Faculty of Business and Economics
Lausanne 1015
Switzerland

SUE LIECHTI
University of Lausanne
Institute of Psychology
Lausanne 1015
Switzerland

In press:

Academy of Management Learning & Education

Address correspondence to: john.antonakis@unil.ch, Tel: +41 21 692-3438

We are grateful to the special issue editors and the reviewers for their helpful feedback; our Action Editor, Sim Sitkim, was particularly constructive in shepherding our manuscript. We also thank Marius Brulhart, Fabrizio Butera, Anne d'Arcy, Saskia Faulk, Deanne den Hartog, Klaus Jonas, Alexis Kunz, Rafael Lalive, and Boas Shamir for their helpful suggestions on earlier versions of this work; any errors or omissions are our responsibility. Finally, we thank the participant depicted in Figure 1A for giving us permission to publish pictures of him.

CAN CHARISMA BE TAUGHT? TESTS OF TWO INTERVENTIONS

ABSTRACT

We tested whether we could teach individuals to behave more charismatically, and whether changes in charisma affected leader outcomes. In Study 1, a mixed-design field experiment, we randomly assigned 34 middle-level managers to a control or an experimental group. Three months later, we reassessed the managers using their co-worker ratings (Time 1 raters = 343; Time 2 raters = 321). In Study 2, a within-subjects laboratory experiment, we videotaped 41 MBA participants giving a speech. We then taught them how to behave more charismatically and they redelivered the speech six weeks later; independent assessors ($n = 135$) rated the speeches. Results from the studies indicated that the training had significant effects on ratings of leader charisma (mean $d = .62$) and that charisma had significant effects on ratings of leader prototypicality and emergence.

Keywords: charisma, experiment, leadership development, rhetoric, quasi-experimentation, field study, laboratory experiment, evidence-based practice.

Can leadership, and in particular charisma be taught? If answered in the affirmative, this question has important implications for training and practice. The management and economics literatures have established that leadership matters a great deal, whether at the country, organization, or team level of analysis (Chen, Kirkman, & Kanfer, 2007; Flynn & Staw, 2004; House, Spangler, & Woycke, 1991; Jones & Olken, 2005; Judge & Piccolo, 2004). More importantly, when measuring specific components of leadership, charismatic leadership demonstrates strong effects on leader outcomes, as meta-analyses have repeatedly established (DeGroot, Kiker, & Cross, 2001; Dum Dum, Lowe, & Avolio, 2002; Gasper, 1992; Judge & Piccolo, 2004; Lowe, Kroeck, & Sivasubramaniam, 1996). Knowing whether charisma can be learned is therefore an important practical question.

Casting some doubt on whether leadership can be taught is evidence that heritable traits including intelligence and personality (Bouchard & Loehlin, 2001; Bouchard & McGue, 2003) predict leadership (Judge, Bono, Ilies, & Gerhardt, 2002; Judge, Colbert, & Ilies, 2004; Lord, De Vader, & Alliger, 1986). Such findings feed into popular notions that “leaders are born.” Teaching adults to be better leaders should thus be quite a feat. A recent meta-analysis, though, indicates that leadership is learnable (Avolio, Reichard, Hannah, Walumbwa, & Chan, 2009); however, the specific mechanisms by which charisma can be trained are still largely unknown.

Our objective, therefore, was to add to the literature on leader development by testing whether a theoretically-designed intervention can make individuals appear more charismatic to independent observers. In a field and laboratory setting, we investigated whether we could manipulate participant charisma and whether charisma would have a significant effect on leader outcomes (i.e., observer perceptions); as we discuss below, it is only via observer attributions that we can study the impact of charisma.

Apart from our substantive findings, which have important implications for theory and training, our paper has some methodological novelties: In Study 1, we used two-stage least squares, a mainstay of economics, to estimate the causal effect of trained charisma on leader outcomes. In Study 2, which was conducted in the context of a MBA program wherein we were unable to have a control group, we used a nonequivalent dependent variable that we did not manipulate-- communication skills--as a control variable to remove learning effects inherent to within-subject designs. Also, we used a naturalistic design by manipulating charisma in working adults (i.e., we did not use confederates).

Our paper is organized as follows. First, we discuss how charisma has been conceptualized and defined; we differentiate sociological from psychological perspectives. Next, we discuss observable markers of charisma. We then present our hypotheses, as well as an overview of the two studies we conducted. We report detailed results and consider their theoretical and practical implications.

Charismatic Leadership

Before providing a specific definition of charisma that will guide our research, we first distinguish charisma from a related construct, transformational leadership. Bass (1985) suggested that charisma is a subcomponent of transformational leadership; however, some scholars state that these constructs are isomorphic or have similar effects (cf. Judge & Piccolo, 2004).

Charisma and transformational leadership are related but theoretically distinct (Antonakis, 2011; see Yukl, 1999). Transformational leadership is much broader and includes means of influence predicated on the leader having a developmental and empowering focus (e.g., individualized consideration) and on using “rational” influencing means (e.g., intellectual stimulation).

However, key in neo-charismatic perspectives is that charismatic leadership uses symbolic

influence and stems from certain leader actions and attributions that followers make of leaders, which produces the alchemy known as charisma (Conger & Kanungo, 1998; House, 1999; Shamir, 1999). Charisma's consequences are only evident in the perceptions of followers who "validate" the leader's charisma. Only once followers have accepted the leader as a symbol of their moral unity can the leader have charisma (Keyes, 2002).

The concept of charisma was first proposed by Weber (1947, 1968). He referred to charisma as a gift "of the body and spirit not accessible to everybody" (Weber, 1968: 19). House (1977: 189), who was the first to present an integrated psychological theory of charisma, suggested it referred to "leaders who by force of their personal abilities are capable of having profound and extraordinary effects on followers." He also noted that the foundation of charisma is an emotional interaction that leaders have with followers and that the "gift" was due, in part to the leader's "personal characteristics, and the behavior the leader employs" (House, 1977: 193). He stated further that, "Because of other 'gifts' attributed to the leader, such as extraordinary competence, the followers believe that the leader will bring about social change and will thus deliver them from their plight" (House, 1977: 204). Followers of charismatic leaders show devotion and loyalty toward the cause that the leader represents (Bass, 1985) and willingly place their destiny in their leader's hands (Weber, 1968). Such leaders reduce follower uncertainty or feelings of threat (Hogg, 2001); simply put, charismatic leaders are highly influential leaders.

Sociological accounts of charisma were usually associated with high-level political leadership and crisis (cf. Downton, 1973; Weber, 1947); however, modern notions of charisma have departed from the original conceptualization, and do not see charisma as a rare and unusual quality (cf. Beyer, 1999). There is debate as to whether charisma can be studied in organizational settings: Neo-charismatic theorists take more of a psychological and organizational tack with

charisma and do not agree with some assumptions made by sociologists. Neo-charismatic theorists view charisma in a more “tame” way, suggesting it can be studied in a variety of organizational milieus (Antonakis & House, 2002; House, 1999; Shamir, 1999) and care more about its “close” rather than “distant” manifestation (cf. Antonakis & Atwater, 2002; Shamir, 1995). Even Shils (1965: 202), who wrote from a sociological perspective, noted that Weber “did not consider the more widely dispersed, unintensified operation of the charismatic element in corporate bodies governed by the rational-legal type of authority.” Furthermore, although crisis can facilitate the emergence of charismatic leadership it is not necessary for its emergence (Conger & Kanungo, 1998; Etzioni, 1961; Shamir & Howell, 1999; Shils, 1965).

Following House (1999), we focus on “organizational charisma,” which depends on leaders and followers sharing ideological values and is predicated on the leader’s use of symbolic power; this type of power is emotions and ideology based and does not use reward, coercive, or expert influence indicative of leadership styles like transactional or task-focused leadership (cf. Antonakis & House, 2002; Etzioni, 1964; French & Raven, 1968). Antonakis and House (2002: 8), who reviewed and integrated theories of charisma noted that charisma is observable in organizations and concerns “leaders who use symbolic means to motivate followers . . . and in whom followers can express their ideals. Charismatic leaders are viewed as strong and confident based on attributions that followers make of these leaders. Followers respect and trust these leaders [who display] moral conviction and are idealized and highly respected by followers. This perspective contrasts with the ‘take-it-or-leave-it’ radical perspective of revolutionary leaders.”

As concerns definitions of charisma, scholars usually describe it by its antecedents and outcomes or by identifying exemplars (e.g., see Bryman, 1993; Conger & Kanungo, 1988; House, 1977; Shamir, House, & Arthur, 1993; Yukl, 2006). As argued by MacKenzie (2003),

using only these elements in a definition is not helpful because they do not define the nature or underlying themes of the phenomenon and turn the empirical test into a tautology. Following MacKenzie's recommendations, we therefore propose a general definition of charisma having a unifying theoretical denominator, *symbolic* leader power (cf. Etzioni, 1961). We thus define charisma as symbolic leader influence rooted in emotional and ideological foundations; charisma's effects are evident on observer attributions of the leader and its antecedents stem from nonverbal and verbal influencing tactics that reify the leader's vision. Of course, our conceptualization of charisma is a neo-charismatic one and other conceptualizations, (e.g., sociological) are equally valid. We focus on a neo-charismatic conceptualization because it can be observed in organizations and can, theoretically, be manipulated. Using our definition as a guiding framework, we sought to manipulate charisma's antecedents or what we call "markers" and observe their effects on follower attributions of the leader's charisma and other outcomes (e.g., affect for the leader, trust in the leader).

The Making of Charisma: Charismatic Leadership Tactics (CLTs)

Neo-charismatic scholars have suggested that leaders are attributed charisma because they can communicate in vivid and emotional ways that federate collective action around a vision (Den Hartog & Verburg, 1997; Shamir et al., 1993). The influencing tactics used by charismatic leaders depend not only on the content (i.e., verbal) of what they say but also on the delivery mode (i.e., nonverbal, Awamleh & Gardner, 1999). These Charismatic Leadership Tactics (CLTs) are very potent devices that affect followers' emotions and information processing (cf. Aristotle, 1954, and the "artistic" influencing means he identified). We identified nine core verbal and three core nonverbal strategies used by charismatic leaders, to which neo-charismatic scholars often refer (Awamleh & Gardner, 1999; Den Hartog & Verburg, 1997;

Shamir, Arthur, & House, 1994). Note that the nine verbal charismatic strategies, as measured in the nomination speeches of the democratic and republican contenders for the U.S. presidency between 1916-2008, significantly predict the outcomes of the U.S. presidential elections (Jacquart & Antonakis, 2010) beyond the effects of a well-established macroeconomic voting model (i.e., the model of Fair, 2009); also, these verbal charisma measures correlate strongly with other independent measures of charisma.

The verbal (i.e., rhetorical) tactics charismatic leaders use include *metaphors*, which are very effective persuasion devices that affect information processing and framing by simplifying the message, stirring emotions, invoking symbolic meanings and aiding recall (Charteris-Black, 2005; Emrich, Brower, Feldman, & Garland, 2001; Mio, 1997; Mio, Riggio, Levin, & Reese, 2005). Charismatic leaders frequently use *stories and anecdotes* (Frese, Beimel, & Schoenborn, 2003; Towler, 2003), which make the message understandable and easy to remember (Bower, 1976), and induce identification with the protagonists (Altenbernd & Lewis, 1980). Charismatic leaders demonstrate *moral conviction* (Conger & Kanungo, 1998; House, 1977) and share the *sentiments of the collective* (Shamir et al., 1994; Shamir et al., 1993); such an orientation aids in identification to the extent that the morals and sentiments overlap with those of followers, and the leader is seen as a representative of the group (cf. Hogg, 2001; Kark & Shamir, 2002). Furthermore, these leaders *set high expectations* for themselves and their followers and *communicate confidence* that these goals can be met (House, 1977); theoretically, these strategies are catalysts of motivation (Eden, 1988; Eden et al., 2000; Locke & Latham, 2002) and increase self-efficacy belief (Bandura, 1977; Shamir et al., 1993). Finally, charismatic leaders use specific rhetorical devices including *contrasts* (frame and focus the message), *lists* (given the impression of completeness), as well as *rhetorical questions* (create anticipation and puzzles,

see Atkinson, 2004; Den Hartog & Verburg, 1997; Willner, 1984).

On a nonverbal level, charismatic leaders are masters at conveying their emotional states, whether positive or negative, to demonstrate passion and obtain support for what is being said (Bono & Ilies, 2006; Frese et al., 2003; Wasielewski, 1985). They use *body gestures*, as well as *facial expressions* (Frese et al., 2003; Towler, 2003; Wasielewski, 1985) and an *animated voice tone* (Frese et al., 2003; Towler, 2003). Finally, both the verbal and nonverbal CLTs make the message of the leader more memorable (cf. Bower, 1976; Charteris-Black, 2005; Emrich et al., 2001; Mio, 1997; Mio et al., 2005; Wasielewski, 1985).

Research suggests that many of these tactics can be manipulated and taught. For instance Howell and Frost (1989) conducted a laboratory experiment where they manipulated some markers of charisma (e.g., communicating high expectations and confidence, using nonverbal influencing means); however, the manipulation was demonstrated by a confederate. Turning to field interventions, there is evidence that transformational leadership, which includes some aspects of charisma, is learnable and that it can have effects on real-world outcomes (Barling, Weber, & Kelloway, 1996; Dvir, Eden, Avolio, & Shamir, 2002; Morhart, Herzog, & Tomczak, 2009). However, these studies did not specifically focus on charisma and trained mostly transformational leadership.

More relevant is a recent study by Frese, Beimeel, and Schoenborn (2003), who taught aspects of the CLTs to practicing managers. Similarly Towler (2003) showed that vision and charisma behavior can be improved in students and that participants exposed to the charismatic speeches had better outcomes than did those exposed to speeches from the control group. The Frese et al. and Towler studies had in common the fact that they provided participants with very specific training and feedback on charismatic content and use of effective verbal and nonverbal

strategies indicative of charismatic leadership. Our study goes beyond these studies in two ways: First, Frese et al. did not use a control group (i.e., they used a within-subjects design only). Importantly, they found that the intervention had a significant effect on a nonequivalent dependent variable (one that they did not intend to change); thus, the causal effects they report may be confounded because they did not use any corrective procedure to partial out learning effects. Second, although the Towler study had a control group, she used only students who were quite young: mean age = 28.95, SD = 6.91 for Study 1; Mean age = 19.31, SD = 2.02 for Study 2 (these descriptive data were not published by Towler, but provided to us by Towler). Thus, her results are less generalizable to working populations.

Using two samples of mature working adults, we based our intervention method using a similar approach to the aforementioned studies. We used different assessment methods to gauge charisma (i.e., attributed charisma and the markers of charisma), included stronger controls in our regression models, and conducted a field as well as a laboratory experiment. We expected that an intervention with substantial feedback would engender higher levels of charisma, which would, in turn, predict various leader outcomes (discussed in the next section). Specifically, we hypothesized the following:

H1a: an intervention group having received charismatic leadership training will score significantly higher on ratings of attributed charisma than will a control group (Study 1).

H1b: controlling for learning effects, post-training ratings of charisma markers will show a significant improvement as compared to pre-training ratings of charisma (Study 2).

Leadership prototypicality and emergence

Although we wished to determine whether participants could learn to behave more charismatically, we were also interested in observing whether individuals receiving training were

rated higher in leader prototypicality and would more likely emerge as leaders. As discussed before, charisma can only be “validated” in perceptions of followers. Thus, given that charisma is theorized to have very potent effects on observers, we also examined whether an increase in charisma affected leader outcomes in the following process model: training → charisma → outcomes. As mentioned, charismatic leader behavior predicts subjective or objective leader outcomes; therefore, if individuals have developed archetypes based on the leaders they have observed in practice, charismatic leaders should be representative of prototypical leaders (because charismatic leaders are effective). Indeed, prototypes of leadership are universally endorsed to be strongly associated with neo-charismatic forms of leadership (Bass & Avolio, 1997; Brodbeck et al., 2000; Den Hartog, House, Hanges, & Ruiz-Quintanilla, 1999).

Leader categorization theory suggests too that observers draw on their implicit contextual prototypes regarding leaders and then compare the target individual to that prototype (Lord, Brown, Harvey, & Hall, 2001; Lord, Foti, & De Vader, 1984). The extent to which the actual leader schema overlaps with the prototype will determine the extent to which the target will be accorded leader status. Because charismatic leadership is considered to be indicative of effective leadership, we expected that individuals who behave charismatically will be seen as more prototypic (i.e., closer to the “stereotype” of what observers consider to be a good leader) and more likely to emerge as leaders (Hogg, 2001). Thus, the consequences of charisma would be perceived by independent observers, who would rate charismatic leaders as being more leader-like (note, to our knowledge, an estimate of the relation between ratings of charisma and prototypicality of actual leaders has never been reported). Specifically we hypothesized that:

H2a: charisma will predict leader prototypicality (Study 1 & 2)

H2b: charisma will predict leader emergence (Study 2)

Finally, our review of the literature suggests that charismatic leaders would (a) create affect-laden relationships with followers, (b) induce trust, (c) be seen as very competent, and (d) be easily able to influence followers. Thus, we hypothesized:

H3: charisma will predict outcomes associated with it including affect for the leader, trust in the leader, and attributions of competence and influence (Study 1 & 2).

Overview of the Experiments

We examined, in two studies, whether charisma could be taught. In Study 1, a between-subject repeated measures field experiment, we randomly assigned managers to an experimental or a control group. We controlled for preexisting differences in charisma and examined the impact of the training three months later. In Study 2, a within-subjects laboratory experiment, we specifically measured changes in the CLTs of MBA students to determine whether charisma predicted leader outcomes.

Both studies shed unique light on the process by which individuals are accorded leader status in contexts where external validity (Study 1) and internal validity (Study 2) were assured. The field study was a rigorous test of our hypotheses given that significant findings would indicate that leaders in the experimental group had to change their behaviors and maintain these changes in interactions with multiple workplace observers who already had categorized the target across multiple occasions and situations. However, the field study might have potential rival explanations regarding the presumed training effect. Thus, in the laboratory study we precisely measured the CLTs exhibited by the trained participants to determine whether they would be seen as more leader-like by independent observers as a function of these tactics.

Study 1

Design

This study was a mixed design field experiment with a control group with a pre- and post-test of charisma (that can be analyzed using ANCOVA or repeated-measures ANOVA). We obtained 360-degree ratings (i.e., ratings from subordinates, peers, and bosses) on the perceived charisma of the managers one month before the intervention (Time 1). We then randomly assigned managers to an experimental or a control group. We provided the intervention to the experimental group and re-measured charisma and leader outcomes in both groups three months after the workshop (Time 2). Thereafter, we gave the workshop to the control group. This three-month delay had three purposes: (a) we provided managers time to implement what they learned (b) we evaluated stability of the intervention over time, and (c) we reduced the effects of common-method variance.

We measured Time 1 Charisma, which served as a robust covariate in an ANCOVA/regression-type design to control for pre-existing differences in leadership ability when predicting Time 2 Charisma; more importantly, it increased statistical power and precision (Maxwell, Cole, Arvey, & Salas, 1991). We included the pretest because the company that provided us the participants limited the amount of participants we could enroll in the study. We thus designed the study appropriately for robust inference and to avoid making a Type II error. We estimated that Time 1 Charisma would correlate very strongly with Time 2 Charisma, at about .80 or greater (see Bass & Avolio, 1997); along with control covariates, we estimated that the full regression model would predict about 75% of the variance in the regression model. We also assumed that the intervention dummy would predict about 7% of the variance (i.e., $r = .26$) in Time 2 Charisma (see results of Dvir et al., 2002 for estimates). Using STATA's "powerreg" module (Ender & Chen, 2008), we conducted a power analysis: A sample size of about 30 would be sufficient to detect the effect (i.e., power of .80). Not having a pretest would have required a

sample size of at least 100 to detect a significant finding with that particular effect size (Cohen, 1988).

We informed raters that the Time 2 ratings were to test for the stability of leadership patterns of their manager (and thus not draw attention to them that we were conducting an experiment). Also prior to the workshop and again during the workshop, we asked managers in the experimental group not to inform any of their colleagues that they had received leadership training so as to avoid raters having any expectations when re-rating the participants. In this way, we ensured that ratings for the control and experimental group leaders were conducted under the same expectancy conditions. Also, we expected minimal contamination effects between the experimental and control group given that the size and the operations of the company are such that the participants were distributed across Switzerland (the location of the experiment).

Participants

Participants who qualified for inclusion were 34 middle managers of a large Swiss company (circa 20,000 employees) with operations across Switzerland; we provided the training to this company gratis for research purposes. Management supported this training program and raters were informed beforehand that the target managers would be participating in an assessment provided by a university. Leaders were self-selected and participated for developmental reasons¹. The working language of the company is mostly German, though French is used too (most managers had German as their first language). All managers spoke English, the language in which the training intervention was conducted. The mean age of the

¹ Because the managers were self-selected, it may be possible that they were not representative of managers in general. We thus compared the Time 1 Charisma--the pretest variable used in this experiment--of this group (n = 34) to 10 groups of male managers (n = 197) on whom we had data on the same Charisma variable. These managers, who were from various Swiss and European companies, were not self-selected and attended company-sponsored training programs. Overall, the 11 groups did not differ significantly ($F_{10,220} = 1.67, p > .05$). None of the post-hoc Scheffé or SNK tests were significant.

managers was 42.44 years ($SD = 5.75$). Of the 34 participants, three were women. Most of the raters were German speaking and provided ratings in German; the rest responded in French.

We took the usual precautions regarding the translations of the questionnaire (i.e., forward and backward translation by independent translators). To ensure that responses were unaffected by social desirability, raters participated anonymously (Antonioni, 1994). Surveys were completed on-line via a secure university server. To ensure we had a large and representative sample of raters, we asked leaders in both the experimental and control group to provide the e-mail details of colleagues with whom they worked directly and frequently; we instructed them to select, at a minimum, six subordinates, four to six peers, and their boss/es. The HR office gave us the contact particulars of the leaders and also verified that the number of raters provided by the leaders was sufficient and representative. Including as many raters as possible ensured that leaders did not only select targeted raters who would give them higher ratings; this point is important, because at Time 2 the experimental group may have selected specific raters with whom they had good relations to obtain higher ratings and thus “please” the experimenters. Overall, response rates were over 70%: For the first wave of data gathering, we contacted 463 raters and obtained 343 responses (i.e., 10.08 raters per leader). For the second wave, raters contacted (444) and ratings received (321) were about the same (i.e., 9.44 raters per leader). The difference in response distributions from Time 1 to Time 2 was not significant, $\chi^2(1) = .06, p > .10$.

Because raters participated anonymously (i.e., only their leader group was identifiable), we could not model rater nestings across time nor include a dummy variable to indicate rater position. We therefore aggregated ratings to the leader level of analysis, using the appropriate statistical justifications. Furthermore, some of the raters who responded at Time 1 and Time 2

were not all the same, which means that potential common source variance problems were further reduced. That is, the fact that different raters rated the leaders in the two assessment periods made for a more objective assessment of Time 2 leadership; raters responding the first time in Time 2 do not have a “consistency motif” to defend as compared to raters responding both in Time 1 and Time 2 (cf. Podsakoff, MacKenzie, Lee, & Podsakoff, 2003).

Also, having different sets of raters assumes that a leader will be evaluated in a reliable way as a function of the leader’s demonstrated behavior (irrespective of which raters are included). This is a safe assumption to make provided that the number of raters is sufficiently large (Mount & Scullen, 2001; Scullen, Mount, & Goff, 2000). In fact, pooling raters with different perspectives reduces idiosyncratic rater bias and measurement error (Scullen et al., 2000). Using a minimum of six raters from different organizational perspectives (i.e., boss, peer, and subordinate) captures about 68% of the true variance in ratings. However, an individual boss or peer rating only captures about 31% true variance; the proportion of true variance for one subordinate is only 17% (Mount & Scullen, 2001). Thus, although raters have different perspectives and different information on the leader, averaging the data of *a large number* of raters from different rater perspectives considerably reduces rater and measurement bias and makes for a more objective measurement of leader behavior. Because we had about 10 raters per leader in both measurement periods, we can be sure that the ratings provided in the two time periods are accurate representations of the leaders’ behaviors.

Measures

Charismatic leadership: We used only the attributed charismatic leadership scale (i.e., attributed idealized influence) of the Multifactor Leadership Questionnaire, Form 5X (Bass & Avolio, 1995); this questionnaire is the best-validated neo-charismatic leadership instrument and

the scale we used has the strongest predictive effects (Aditya, 2004; Antonakis, Avolio, & Sivasubramaniam, 2003; Lowe et al., 1996). We asked raters to estimate how frequently their leader demonstrated the items described. At Time 1, the scale ($\alpha = .70$) was calibrated from 0 (not at all) to 4 (frequently if not always).

At post intervention (Time 2), we re-gathered data on the same charisma scale ($\alpha = .88$), which was reworded to accommodate different anchor points: 0 (strongly disagree) to 8 (strongly agree). We did this to reduce recency and recall effects (i.e., common-source effects). For the purpose of estimating a repeated-measures ANOVA analysis, we converted the scale on the same metric as the Time 1 scale of Charisma by dividing the Time 2 scores by 2. Note that as demonstrated by Preston and Colman (2000), increasing scale points from a 5-point scale to a 9-point scale does not perturb reliability or criterion validity (see also Aiken, 1987; Lissitz & Green, 1975). We also show later that the change in scale did not induce any bias in the results by comparing results from the rescaled data to those where both Time 1 and Time 2 Charisma were standardized or equipercentile equated. We also compared the repeated-measures ANOVA results to those of an ANCOVA analysis, and the ANCOVA predicted means were almost precisely the same as the repeated-measures ANOVA means. Also, using an ANCOVA model does not require that the pretest be measured on the same metric as a posttest; any change in the mean of Time 2 Charisma that may be due to change in response scale is irrelevant when predicting Time 2 Charisma because we have a control group; thus, both groups would benefit equally and any systematic error across groups would be reflected in a change of the intercept in both groups but not the slope of the relationship. Thus, the experimental effect can be consistently estimated, particularly when using the ANCOVA model.

General prototypicality: Given that we wished to determine whether post-intervention

leader charisma predicted prototypicality, we also gathered data using Cronshaw and Lord's (1987) prototypicality measure. We adapted the measure to suit the experimental condition. We asked raters to estimate the extent to which they agreed with the items using a scale from 0 (strongly disagree) to 8 (strongly agree). We used the following three items from Cronshaw and Lord: "The person I am rating frequently demonstrates leader behavior," "The person I am rating acts like a typical leader," and "The person I am rating fits my image of a leader" ($\alpha = .92$).

Leader Outcomes: Apart from charisma, we included four single item dependent measures, which are theoretically indicative of attributions and outcomes of charismatic leadership: (a) affect for the leader: "I like this person as a leader" (cf. Antonakis & House, 2002; Bass, 1985; Conger & Kanungo, 1998; House, 1977); (b) trust in the leader: "The person I am rating is easily trusted" (cf. Antonakis & House, 2002; Conger & Kanungo, 1998; House, 1977; Shamir et al., 1993); (c) leader competence: "The person I am rating is competent as a leader" (cf. Conger & Kanungo, 1998; House, 1977); and (d) leader influencing ability: "The person that I am rating is able to easily influence others" (cf. House, 1977; Shamir et al., 1993). Of course, single-measure dependent variables of this nature are individually limited in scope; however, together these items capture important outcomes of charismatic leadership in multivariate space. Also, single-item measures are not necessarily less reliable than multi-item measures (Bernard, Walsh, & Mills, 2005); more importantly, there is no reason to be concerned about measurement errors in dependent variables because this error is absorbed in the disturbance and is orthogonal to the regressors (Antonakis, Bendahan, Jacquart, & Lalive, 2010; Ree & Carretta, 2006).

Intervention

The intervention was conducted by the first author of this study and took place at a

training facility provided by the company. In addition to the 360-degree ratings, we gathered pre-workshop self-rating data on leadership and personality, which we used for feedback purposes (Atwater, Roush, & Fischthal, 1995). Because it is likely that completing the leader self-ratings and personality test promoted leader self-reflection between the two data-gathering periods, which may have consequently affected leader behaviors, we also asked the control group to provide self-ratings at Time 1. In this way, we could statistically remove any “placebo” or learning effects due to the self-ratings.

The first phase of the intervention lasted five hours. We used an action training approach (cf. Frese et al., 2003). We provided participants with general principles of what constitutes effective and charismatic leadership, we allowed participants to learn by doing, and gave participants the ambitious goal of demonstrating charisma, which they had not yet mastered so as to create a positive tension and hence motivation for them to improve. We also gave participants learning-oriented feedback regarding their charisma.

We focused on explaining to managers the importance of charismatic leadership and highlighted how charisma could be engendered by displaying the CLTs. We also made extensive use of film scenes (e.g., from “True Blue,” “Reversal of Fortune,” “Dead Poets Society,” “Any Given Sunday”) demonstrating charisma so participants could see the theory behind the CLTs in practice. During the workshop participants developed a short speech (in dyads). We asked participants to create a scenario (hypothetical or one based on a real situation) where they were addressing their followers about a particular issue. Each dyad nominated one person to present the speech, and the intervener, together with the other participants, provided the presenters with feedback on their use of Charismatic Leader Tactics.

We also gave participants a leadership feedback report (based on the Time 1 360-degree

measures), which they read after the workshop. This feedback report included data on self and other perceptions of their leadership style. To maximize the impact of this feedback, we gave the leaders the feedback in ways that were not negative, personal or threatening to the self-esteem (Kluger & DeNisi, 1996). A few days after the workshop, participants made a telephone appointment with the intervener to discuss their leadership profile and to present a leadership development plan (a couple of pages long) showing how they would improve their leadership and in particular their charisma. Plans were discussed privately for about an hour with each participant. The purpose of writing the plans and having the coaching sessions was to help the leaders formulate explicit developmental goals and to provide concrete example of how the goals could be enacted (following the principles of control and goal-setting theory, Carver & Scheier, 1990; Locke & Latham, 2002). We encouraged participants to plan these goals into their diaries and to practice the charismatic tactics as often as possible (i.e., a few times per week).

We also gave the participants the audio and transcripts of two speeches, which are indicative of exceptional charismatic content: Martin Luther King's "dream" speech and Jesse Jackson's 1998 speech to the Democratic National Convention. Apart from the first speech which is universally acclaimed to demonstrate excellent rhetoric, the latter speech is also widely touted to be a very effective and charismatic speech (Shamir et al., 1994).

Data aggregation

Because 360-degree ratings at Time 1 and Time 2 were anonymous we averaged other ratings at the leader level of analysis. The aggregation was justified for all scales, based on the ICC₁ (Bliese, 2000) and r_{wg} (Lindell & Brandt, 1999, 2000; Lindell, Brandt, & Whitney, 1999). For the r_{wg} , we assumed a maximum variance null distribution with a Spearman-Brown correction (given the large amount of raters we had per leader, which suggested increased

variation in ratings). The aggregation statistics showed the following: Time 1 Charisma measure, ($ICC_1 = .15$, $F_{33,309} = 2.51$, $p < .01$; $r_{wg} = .92$); Time 2 Charisma measure ($ICC_1 = .14$, $F_{33,287} = 2.06$, $p < .001$; $r_{wg} = .89$); Time 2 General Prototypicality measure ($ICC_1 = .19$, $F_{33,287} = 2.51$, $p < .001$; $r_{wg} = .89$); Time 2 Affect measure ($ICC_1 = .09$, $F_{33,287} = 1.88$, $p < .01$; $r_{wg} = .86$); Time 2 Trust measure ($ICC_1 = .07$, $F_{33,287} = 1.70$, $p < .05$; $r_{wg} = .89$); Time 2 Competence measure ($ICC_1 = .17$, $F_{33,287} = 2.91$, $p < .001$; $r_{wg} = .83$); and the Time 2 Influence measure ($ICC_1 = .22$, $F_{33,287} = 3.60$, $p < .001$; $r_{wg} = .81$).

Results

Manipulation check: At Time 1, the means of Charisma in the experimental ($M = 2.46$, $SD = .25$) and the control groups ($M = 2.49$, $SD = .42$) were statistically equal ($F_{1,32} = .05$, $p > .10$). We estimated a repeated measures mixed-design model including the intervention (dummy coded 1 = experimental group, else 0) and Time 1 and Time 2 Charisma as the repeated measures. We controlled for leader age, sex (dummy coded 1 = female, else 0), and several contextual effects (Liden & Antonakis, 2009) including rater language of response (dummy coded 1 = German if the majority of raters responded in German, else 0), leader first language (dummy coded 1 = German, else 0), and Time 1 and Time 2 rater response percentage. We controlled for the latter in case response rates are correlated with ratings and given the possibility that leaders in the experimental group “motivated” their raters to respond by virtue of the fact that they were trained to demonstrate better leadership. Moreover, we controlled for ethnic group composition of the leader’s group as well as the leader’s first language because rater ethnicity is related to prototypical expectations and behaviors as a function of culture. Note, in Switzerland, one’s first language is indicative of one’s culture because language divides follow cultural divides; also, the Swiss French and German cultures are different in many fundamental ways,

which can affect leadership styles (House, Hanges, Javidan, Dorfman, & Gupta, 2004).

At Time 2, the mean of the experimental group ($M = 2.99, SD = .25$) and the control group ($M = 2.88, SD = .39$) were different (i.e., the within-between interaction was significant: $F_{1,26} = 7.29, \eta^2 = .219, p < .05$)². The estimated Time 1 (and Time 2) marginal means were 2.49 (and 2.89) for the control group, and 2.45 (and 3.00) for the experimental group; estimates were similar to the observed means, suggesting that the randomization worked well. This result provides support for *H1a*.

Effect of Charisma on Leader Outcomes: We report the correlation matrix of the measures in Table 1. As expected, the correlation between Time 1 and 2 Charisma was very high ($r = .87$). Interestingly, the bivariate correlation ($r = .16$) between the intervention dummy and Time 2 Charisma was not significant (given the low power without statistical control for the regression error, i.e., the power without the pretest covariate at this sample size is only .36).

 Insert Table 1 here

The partial regression coefficient of the intervention on Time 2 charisma becomes significant when estimating the following regression (ANCOVA) model, for leader i in group j :

$$Y_i = \beta_0 + \beta_1 \text{Treat}_j + \beta_2 \text{Ch_old}_i + \beta_3 \text{T1_raters}_i + \beta_4 \text{Female}_i + \beta_5 \text{Age} + \beta_6 \text{Lan_rat}_i + \beta_7 \text{Lan_lead}_i + u_i \quad (1)$$

Where, Y = Time 2 Charisma, Treat = Dummy variable for experimental treatment (coded 1, else 0), Ch_old = Time 1 Charisma (pretest), T1_raters = raters response percentage, Female = Dummy variable indicating female leader sex (coded 1, else 0), Age = Leader age,

² As a check, we re-estimated the repeated-measures mixed-design model using the standardized Time 1 and Time 2 Charisma measures (instead of using the rescaled Time 2 Charisma measure). Results were almost identical with respect to the within-between interaction as indicated by the F -statistic and the eta-square: $F_{1,26} = 7.33, \eta^2 = .22, p < .05$, suggesting that the change of scale for the Time 2 measure did not affect the substantive results. We obtained very similar results using the original Time 1 measure and the equipercntile equated Time 2 measure.

Lan_rat = majority German rater responses (coded 1, else 0), Lan_lead = first language of leader German (coded 1, else 0).

Refer to Table 2 (Model 1) for estimates. When controlling for Time 1 Charisma the intervention dummy was significant (standardized $\beta = .23$).³ This regression model is the alternative manipulation check that is unaffected by the differences in ratings scales of Time 1 and Time 2 Charisma.⁴ This result provides a textbook example showing how the pretest-posttest control group design, which includes covariates, improves power and measurement precision by reducing experimental error (Keppel & Wickens, 2004).

Given the limitation of measuring only Time 2 Charisma, we also estimated the direct effect of the treatment on the other Time 2 measures (i.e., prototypicality, affect, trust, competence, and influence) using a multivariate dependent variable specification of in Eq. 1. As indicated in Table 2 (Models 2-6), the intervention had a significant effect on three of the five measures. This result suggests that the treatment affected a broad range of dependent variables that are theoretical outcomes of charisma training.

Next, we estimated the effect of Time 2 Charisma on the Time 2 dependent variables (prototypicality, affect, trust, competence, influence). Because these variables were gathered concurrently from the same source, we did not assume that their disturbances were independent; also, we did not assume that Time 2 Charisma was exogenous. We thus estimated the following systems of equations using the two-stage least squares (2SLS) estimator, for leader i in group j , in the first stage:

³When using the original metric of the Time 2 measure, we obtain exactly the same standardized effect for the treatment, which shows that the scale change in Time 2 had no effect on the treatment. Using equipercentile equating gave a beta of .21.

⁴Controlling for all other factors in the regression model, the estimated marginal means of the dependent variable Time 2 Charisma were 2.87 for the control and 3.01 for the experimental group. These means are almost precisely the same as the repeated measures estimates of 2.89 and 3.00 respectively. Therefore, the change of scale did not have any impact on the mean of Time 2 Charisma.

$$Ch_new_i = \beta_0 + \beta_1 Treat_j + \beta_2 Ch_old_i + \beta_3 T1_raters_i + \beta_4 Female_i + \beta_5 Age + \beta_6 Lan_rat_i + \beta_7 Lan_lead_i + u_i \quad (2)$$

$$Y_{(prot)_i} = \gamma_0 + \gamma_1 \hat{Ch_new}_i + \gamma_2 T2_raters_i + \gamma_3 Lan_rat_i + \gamma_4 Lan_lead_i + \varepsilon_i \quad (3)$$

$$Y_{(aff)_i} = \delta_0 + \delta_1 \hat{Ch_new}_i + \delta_2 T2_raters_i + \delta_3 Lan_rat_i + \delta_4 Lan_lead_i + v_i \quad (4)$$

$$Y_{(trust)_i} = \eta_0 + \eta_1 \hat{Ch_new}_i + \eta_2 T2_raters_i + \eta_3 Lan_rat_i + \eta_4 Lan_lead_i + q_i \quad (5)$$

$$Y_{(comp)_i} = \lambda_0 + \lambda_1 \hat{Ch_new}_i + \lambda_2 T2_raters_i + \lambda_3 Lan_rat_i + \beta_4 Lan_lead_i + w_i \quad (6)$$

$$Y_{(infl)_i} = \mu_0 + \mu_1 \hat{Ch_new}_i + \mu_2 T2_raters_i + \mu_3 Lan_rat_i + \mu_4 Lan_lead_i + o_i \quad (7)$$

Where, Y_{prot} = general prototypicality, Y_{aff} = affect for the leader, Y_{trust} = trust in the leader, Y_{comp} = refers to leader competence, Y_{infl} = refers to leader influence, and $\hat{Ch_new}$ = the predicted value of Time 2 Charisma from Equation 2.

The 2SLS estimator uses exogenous variation (e.g., random assignment to treatment, and in our case Time 1 Charisma as a lagged control) and the predicted value of Time 2 Charisma in the second stage (i.e., $\hat{Ch_new}$) to ensure consistent (i.e., unbiased) estimation of γ , δ , η , λ , and μ ; that is, because the treatment dummy does not correlate with the disturbances in Eq. 3 to Eq. 7 by experimental design, the estimates of $\hat{Ch_new}$ in Eq. 3 to Eq. 7 are purged from simultaneity, common methods bias, and measurement error, and that their causal direction is “locked-in,” provided that endogenous disturbances are correlated (Antonakis et al., 2010; Foster & McLanahan, 1996; Gennetian, Magnuson, & Morris, 2008; Shaver, 2005; Wooldridge, 2002). Also, 2SLS has several advantages: It makes no distributional assumptions on the independent

variables (and thus does not require multivariate normal data); it always converges, it has good small-sample properties, and as a limited-information estimator, it does not spread any potential misspecification bias across equations (Baltagi, 2002; Bollen, Kirby, Curran, Paxton, & Chen, 2007; Kennedy, 2003). Refer to Table 2 for estimates (Models 7-11).

Insert Table 2 here

The results indicated the effect of Time 2 Charisma strongly predicted Prototypicality (standardized $\beta = .84$) supporting *H2*. Time 2 Charisma also significantly predicted the other leader outcomes supporting *H3*. Given that the small sample size of our study, we also estimated the models both with and without the control variables and results did not change. We also estimated the model using bootstrapped ($k = 1000$ replications) and jackknifed standard errors; all coefficients remained significant.

Study 2

In this within-subject's study, we isolated the effects of the CLTs on outcomes by coding the content of pre- and post-training speeches and using the CLTs as predictors of observer ratings. This study complements the first one wherein we could not ascertain whether changes in charisma were due to the CLTs or a competing explanation (e.g., the fact that those leaders who were in the experimental group felt more efficacious and confident, and consequently appeared more leader-like and charismatic).

Design

One week before the training, we gave participants the following instructions: "Imagine that you are a manager of a division of a multinational company. You are faced with a serious

problem that requires you to take drastic and unprecedented action. This action will force you to make some very difficult decisions, which might not go down well with your staff (i.e., the decisions will be difficult to “sell”). Prepare a four minute speech, where you detail the rationale and major points of your plan to your staff.”

We filmed the speeches in standardized conditions. We then provided participants with training in charismatic leadership and asked them to redeliver the same speech six weeks later. We did not allow participants to change the major content of their second speech (i.e., the context and the major decisions remained the same). We videotaped both speeches under the same conditions and each participant wore the same attire in the first and second filming session. The advantage of using this type of within-subjects design is to determine whether variation in charisma predicted subjective ratings of leader prototypicality and other outcomes beyond participant constant (i.e., fixed) effects. These fixed effects, whether observed or not, can be an important source of variance because raters are highly biased by stereotypes when rating leaders, including simple factors like facial appearances (Antonakis & Dalgas, 2009; Todorov, Mandisodza, Goren, & Hall, 2005; Willis & Todorov, 2006). Although we cannot directly measure all the fixed effects, we can capture the effects using leader ($k - 1$) dummy variables.

Participants

Participant leaders were 41 students registered in an English-speaking MBA program at a Swiss public university, and were fluent in English. We gathered this data over a three-year period from three cohorts (the first was a full-time MBA program; the rest were from an EMBA program); note the effect of year/cohort was captured by the leader fixed effects. Four participants were female. The mean age of the participants was 34.74 years ($SD = 5.89$) and they had 9.11 years of work experience ($SD = 4.95$).

The participants were enrolled in a Leadership course, focused mostly on charismatic leadership, and delivered the speeches in English as part of their coursework. The speeches were evaluated by 135 students recruited from several master-level organizational behavior classes (all taught in English). These raters participated in the study for course credit. More males ($n = 82$) than female raters ($n = 53$) participated: $\chi^2(1) = 6.23, p < .05$. The mean age of raters was 27.06 years ($SD = 6.18$). We randomly distributed the raters into 24 groups. We then randomly assigned the 82 speeches to the groups of raters in batches of between 3 to 4 speeches with the constraint that no Time 1 and Time 2 speech of the same leader was viewed in the same group. Raters did not know the purpose of the experiment or that we used two sets of speeches.

Measures

We first provide an overview of the dependent measures used. The first were absolute measures, which captured how the same person was rated in Times 1 and 2 on the leader outcomes (affect, trust, competence, and influence) by raters. The second was a relative measure of leader emergence, to test if leaders giving Time 2 speeches were rated higher, on average, than leaders giving Time 1 speeches. This second measure, which required raters to rank-order the leaders in each group, makes for a particularly strong test because Time 1 speeches served as controls (counterfactuals) for Time 2 speeches.

We also measured the charismatic leader tactics and communication skills of the leaders; these served as manipulation checks, which were coded by four trained coders blinded to the purpose of the experiment. The first check was to ensure that charisma training had its intended effect. The second check was on a nonequivalent dependent variable (Shadish, Cook, & Campbell, 2002)--communication skills--which we did not teach. We measured this variable to control for any possible learning effects that may have occurred given that we did not have a

control group (see Antonakis et al., 2010).

Absolute measure (prototypicality): We used the same measure of leader prototypicality as in Study 1. The α reliability coefficient for this measure was high both at the rater (.92) and the leader level (.97). We used the same leader outcomes measures as in Study 1: Affect for the leader, leader trustworthiness, leader competence, and leader influencing ability. Note that after raters viewed a speech, they immediately rated it on the absolute measure and we then immediately collected their ratings.

Relative measure (emergence): We asked raters to rank the leaders they saw from best (1) to worst (4) (for ease of interpretation, we reversed the ordering for the analysis). If the intervention had an effect, Time 2 leaders should, on average, be rated higher than Time 1 leaders across all groups (all else being equal) irrespective of who the Time 2 leaders were pitted against. Only after raters had viewed all speeches and submitted all ratings on leader outcomes did they rank-order the leaders.

Charisma markers (manipulation check): Two trained coders coded for the presence or absence of the 12 CLTs, using a binary measure (0, 1). We used a binary measure because either a CLT was appropriately demonstrated or it was not, regardless of the frequency of demonstration of the CLT. Using a metaphor that is inappropriate will not induce charisma; for example, when role playing in class one student noted: “Our company has been cruising along with no cares in the world. We have been flying high like an eagle over the mountains. Now, someone has shot off our wing, but we will continue to fly on one wing!” Another example concerns a student whose attempt to demonstrate hand gestures made him continuously and almost without interruption point with his index finger downwards towards his notes. This overuse of a CLT is also inappropriate and will not engender charisma. A more profound

example providing a justification for using a binary measure is the case of Phil Davison who sought the Republican party nomination for Stark County Treasurer in September 2010. Although he used many CLTs, he overused them, particularly emotional displays. Descriptions of Davison's speech included terms such as "crazed," "berserk" and "bizarre" (ABC News, 2010; CBS, 2010; Daily Mail, 2010), "the likes of which few politically involved citizens have ever seen" (Huffington Post, 2010). Such was the surrealistic nature of the speech that it went viral; at the time of writing this paper the speech has been watched about 2.5 million times on *YouTube* (in fact, we currently use this speech as an example of how to not use the CLTs)!

We thus trained the coders to only code for CLTs that were used appropriately, which entails a judgment call. It was important, therefore, to ensure that the raters made the call in a consistent way. After several rounds of training on leaders filmed in similar conditions we tested the coders' inter-rater reliability on six new leaders (72 item observations) who were not included in the experiment. The coders had very high agreement: 85.03%. The Kappa agreement statistic was .67 ($SE = .08$), which can be characterized as substantial (Landis & Koch, 1977). Once the coders' reliability was established, they then independently coded the speeches: We assigned one coder for the Time 1 speeches and the other coder for the Time 2 speeches. Using these 12 items, we created a composite index, indicating the percentage of the items the leader demonstrated, on a scale from zero (0%) to one (100%). As stated by Podsakoff, MacKenzie, Podsakoff, and Lee (2003), a composite index is suitable for our measure because we expected the items to "form" the measure of charisma, we did not expect the items to covary, nor were the items interchangeable (i.e., some leaders may focus on using some types of strategies to appear charismatic, whereas other leaders might use fundamentally different strategies). Nonetheless the 12 items were correlated, given that they were concurrently trained across time. For information

purposes we can report that the α reliability coefficient was .63.

Communication skills (manipulation check): Two other coders coded for the extent to which leaders used appropriate communication skills. We used a scale from zero (strongly disagree) to eight (strongly agree) on the following nine dimensions, which we gleaned from previous research: Speech structure (as communicated in the introduction), framework (beginning and end well-connected), logic of the speech, simple and easy-to-understand language, nonlexical utterances, clear pronunciation, correct English, nervousness, and tempo of speech (see Frese et al., 2003; Magin & Helmore, 2001; Towler, 2003). After several rounds of training, we tested the coders on four leaders (36 item observations) not included in the experiment. The concordance correlation coefficient (Lin, 1989) was .62 (SE = .10); regressing one coder's score on the others (and controlling for leader fixed effects) indicated that the standardized beta of the coders was .72. Given their high level of reliability we then obtained ratings from one coder on the Time 1 speeches, and ratings from the second coder for the Time 2 speeches (overall, the α reliability for the scale was .71).

Intervention

The first author of the study gave the intervention over a seven-week period (20 hours). Again, the approach we used is similar to that of Study 1, except that the intervention was longer; thus, we spent more time showing film scenes (of charismatic and non-charismatic leaders), as well as in theoretical explanations, role playing, and classroom discussions. We attempted to maximize knowledge transfer by getting participants to focus on what new behaviors they would need to demonstrate to look more charismatic, while providing participants with extensive feedback on their charisma.

The other major difference between this intervention and the first one was that we filmed

the participants' speeches. After the first speech was delivered, we gave participants a video copy of their speech and specific feedback on their demonstration of the tactics. We asked them to study their speech and focus on improving it. We also encouraged students to give feedback to each other on the CLTs. Finally, participants submitted a more extensive leadership development plan (10-12 pages, double-spaced).

Results

For descriptive statistics refer to Table 3.

[Insert Table 3 here]

Manipulation checks: To determine whether the manipulation worked, we predicted the mean Time 1 and Time 2 scores of the Charisma (CLTs) and Communications (Comm) skills measures using the following simultaneous equations (with correlated disturbances) with cluster-robust variance estimation at the leader level for leader i , coder c (for charisma), coder k (for communication), in period j (pre- and post-intervention), with equation disturbances e and u :

$$\text{CLTs}_{icj} = \beta_{00} + \beta_{01}\text{Pre_post}_{ij} + \beta_{02}\text{time}_{ij} + \sum_{i=2}^{41} \alpha_i \mathbf{L}_i + e_{ijc} \quad (8)$$

$$\text{Comm}_{ikj} = \beta_{01} + \beta_{11}\text{Pre_post}_{ij} + \beta_{12}\text{time}_{ij} + \sum_{i=2}^{41} \gamma_i \mathbf{L}_i + u_{ijk} \quad (9)$$

Pre_post refers to the dummy variable indicating pre-training (coded 0) or post-training (coded 1). Time refers to length of speech. L refers to 40 dummy variables ($k - 1$ leaders) to capture the fixed effects common to a specific leader. As mentioned, this specification ensures that all effects, whether observed or not, and common to each leader (e.g., facial appearance, leader style, dress style, sex, etc.), are explicitly modeled. Thus, estimates for person-varying covariates are consistently estimated. We controlled for length of the speech because mean Time 1 and Time 2 speeches across participants were different: 3.45 minutes and 4.06 minutes

respectively ($\text{Pre_post } \beta = .60, SE = .14, z = 4.32, p < .001$). We included length of speech time in subsequent analyses to eliminate potential duration effects correlated with independent or dependent variables. The mean percentage of charismatic tactics (from 0 to 1) we observed at Time 1 was .24 ($SD = .13$), whereas for Time 2 it was .48 ($SD = .16$). The coefficient for Pre_post from Eq. 8 was .21 ($SE = .03, z = 7.00, p < .001$); the standardized coefficient was .57. For an example of the improvement of nonverbal ability, refer to Figure 1 (Panel A).

[Insert Figure 1 here]

The above result suggests that charisma can be learned; however, this coefficient may be biased by learning effects. Thus, we first analyzed whether the treatment had an effect on the non-equivalent dependent variable. The mean of communication skills was 6.60 at Time 1, and 7.13 at Time 2. The coefficient for Pre_post from Eq. 9 was .69 ($SE = .11, z = 6.46, p < .001$). Thus, the nonequivalent dependent variable, which we did not attempt to manipulate, changed too, suggesting that there were significant learning effects between the two time periods. Given that our communications skills measure was rather broad, it captured a large class of learning effects that might have occurred. Thus, we included communications skills as a control variable in all specifications to remove these learning effects from the charisma measure. Indeed, including communication skills in Equation 7 reduced the standardized coefficient of Pre_post on the CLTs from .57 to .35, the latter being a lower-bound experimental effect, to the extent that charisma might have induced an increase in communication skills too (which is likely). This result indicates that the tactics can be learned providing strong support for *H1b*. Now we turn to whether these tactics had an effect on leader outcomes.

Effect on leader prototypicality, outcomes, and emergence: We then estimated five models simultaneously, using mixed-process maximum likelihood estimation for leader i , in

period j , for coder c coding CLTs and coder k coding communications skills (Comm), and rater r , with equation disturbances v , w , q , s , g , u , and correlating cross-equation disturbances (Roodman, 2008):

$$Y_{(Prot)ijckr} = \beta_0 + \beta_1 CLTs_{ijc} + \beta_2 Comm_{ijk} + \beta_3 time_{ij} + \beta_4 Rater_male_r + \beta_5 Rater_age_r + \sum_{i=2}^{41} B_i L_i + v_{ijckr} \quad (10)$$

$$Y_{(aff)ijckr} = \gamma_0 + \gamma_1 CLTs_{ijc} + \gamma_2 Comm_{ijk} + \gamma_3 time_{ij} + \gamma_4 Rater_male_r + \gamma_5 Rater_age_r + \sum_{i=2}^{41} \Gamma_i L_i + w_{ijckr} \quad (11)$$

$$Y_{(trust)ijckr} = \delta_0 + \delta_1 CLTs_{ijc} + \delta_2 Comm_{ijk} + \delta_3 time_{ij} + \delta_4 Rater_male_r + \delta_5 Rater_age_r + \sum_{i=2}^{41} \Delta_i L_i + q_{ijckr} \quad (12)$$

$$Y_{(comp)ijckr} = \eta_0 + \eta_1 CLTs_{ijc} + \eta_2 Comm_{ijk} + \eta_3 time_{ij} + \eta_4 Rater_male_r + \eta_5 Rater_age_r + \sum_{i=2}^{41} H_i L_i + s_{ijckr} \quad (13)$$

$$Y_{(infl)ijckr} = \lambda_0 + \lambda_1 CLTs_{ijc} + \lambda_2 Comm_{ijk} + \lambda_3 time_{ij} + \lambda_4 Rater_male_r + \lambda_5 Rater_age_r + \sum_{i=2}^{41} \Lambda_i L_i + g_{ijckr} \quad (14)$$

$$Y_{(rank)ijckr} = \mu_0 + \mu_1 CLTs_{ijc} + \mu_2 Comm_{ijk} + \mu_3 time_{ij} + \mu_4 Rater_male_r + \mu_5 Rater_age_r + \sum_{i=2}^{41} M_i L_i + u_{ijckr} \quad (15)$$

Where, $Y_{(prot)}$ = general prototypicality, $Y_{(aff)}$ = affect for the leader, $Y_{(trust)}$ = trust in the leader, $Y_{(comp)}$ = refers to leader competence, $Y_{(infl)}$ = refers to leader influence; these constituted the absolute measures. $Y_{(rank)}$ is the relative rating (from 1 to 4, estimated with an ordered probit). We also controlled for time, rater male (= 1, else 0) and rater age. We estimated cluster-robust

standard errors on the rater level (given that each rater rated four leaders). Refer to Table 4 for results.

Insert Table 4 about here

Results indicate that the CLTs had a significant effect on all the dependent variables, Wald $\chi^2(6) = 46.67, p < .001$. The standardized betas for the respective dependent variables (in parentheses) were: prototypicality (.49), affect for the leader (.33), trust in the leader (.25), leader competence (.45), leader influence, (.37), leader rank (.75). These results provide support for *H2a*, *H2b*, and *H3*. As another indication of the experimental effect we calculated the predicted probability of obtaining a particular ranking on the emergence measure. Using the CLTs to a high degree (+2SD's from the overall mean of the measures) was more strongly associated with obtaining a rank of 4 (best ranking) than a rank of 1 (worst), as compared to using the CLTs to a low degree (-2SD's from the mean), which further supports *H2a*. Refer to Figure 1 (Panel B).

Discussion

The results of the studies suggest that charisma can be taught. We obtained these results in a field and laboratory setting, using a similar intervention with different designs and measures. The change in charisma we induced increased leader prototypicality and participant charisma predicted leader outcomes as well as leader emergence. These findings occurred in a European setting, thus enhancing the generalizability of other experimental studies on charisma conducted mostly in North America.

The results have important industrial and educational implications because they demonstrate that charisma is learnable. The effects were significant, as are their practical consequences. As a measure effect for charisma, the treatment yielded an *r* of .23 in Study 1 and

.35 in Study 2. These effect sizes can be qualified as being close to medium ($d = .47$) for Study 1 and medium ($d = .75$) for Study 2; the mean sample-weighted effect across the two studies being $d = .62$ (Cohen, 1988). We probably obtained stronger effects for charisma in Study 2 because of the greater amount of time we invested in the intervention. Our results compare favorably with a recent meta-analysis, which for broad classes of leadership models found a medium effect size too ($d = .60$) (Avolio et al., 2009). Furthermore, our measure, in terms of success rate (weighted by sample size at the leader level, i.e., $r = .29$), suggests that similar interventions will produce a success rate of about 65% in the experimental group; that is, 65% of individuals in the experimental group will have above median performance, whereas only 35% of the individuals in the control group will have above median performance (Rosenthal & Rubin, 1982).

Our results also provide support for the theorists' propositions concerning the CLTs. These CLTs predicted prototypicality and various leader outcome ratings. These indicators could be used in future research to develop more comprehensive interventions and to test the effects of the tactics on leader outcomes. Our paper makes methodological contributions too because it is, to our knowledge, one of the first in the leadership discipline to use a 2SLS regression design to rule-out confounds and to correct estimates that may be biased because of endogeneity, simultaneity, common-methods or measurement error. We also show, for the first time, how learning effects can be removed from intervention effects by using a non-equivalent dependent variable as a control variable (when using a control group is infeasible, cf. Antonakis et al., 2010).

Theoretical Implications

Charisma is a component of leadership that has created much definitional confusion (Yukl, 1999). Also, what we know about charisma from standardized questionnaires measures is

wholly inadequate. An inherent limitation of measures like the venerable Multifactor Leadership Questionnaire is that the items constituting the charisma scales, particularly attributed charisma, are quite vague and their processual antecedents are poorly understood (Yukl, 1999). For instance, what makes a leader seem powerful and confident? Given the haziness surrounding some of these questionnaire measures, it is no wonder that some have suggested that charisma is an “illusionary . . . U.F.O. phenomenon,” a “black hole,” and a “social delusion” that is unmeasurable (Gemmill & Oakley, 1992: 119).

In addition to not being prospectively defined, charisma measures depend on certain antecedents; our study shows that leaders appear charismatic because they use a wide array of verbal and nonverbal CLTs. These CLTs are not easily captured in behavioral questionnaires because the CLTs probably would go unnoticed to untrained observers. As our study suggests, one way to sharpen the conceptualization of a construct is to identify markers of the construct that can be manipulated and measured (ideally in a more objective way than using questionnaires).

Apart from shedding light on the charisma phenomenon our study’s main contribution is that leadership is learnable; research should focus on uncovering which learning processes are most relevant and how they can be managed and accelerated. For example, clinical psychology has made many inroads with cognitive-behavioral theory, whose interventions can produce strong effects (Butler, Chapman, Forman, & Beck, 2006); however, such methods have not been adequately leveraged in leadership studies.

Practical Implications

Turning to practice, because the education and training industries are high-stakes games, it is important that evidence-based approaches for leader development should be brought to the

fore. Unfortunately, there is a well-known rift between science and practice in general (Rynes, Colbert, & Brown, 2002) including leadership practice in particular (Zaccaro & Horn, 2003). Worse, evidence-based principles are not being used to the extent that they should be in management education programs such as MBAs (Charlier, Brown, & Rynes, in press).

At this time, there is not much theoretical or empirical work regarding how to develop charisma. As we demonstrated with our approach, leaders must be provided with extensive feedback on their styles in a participant-centered approach; more importantly, leaders need focused and explicit development goals, and to be placed in a “maximum-performance situation,” with knowledge that they will be re-measured (cf. Reilly, Smither, & Vasilopoulos, 1996). We believe that this type of structured approach is required given leaders’ tendencies to overestimate their leadership abilities (Atwater & Yammarino, 1992; Podsakoff & Organ, 1986); that is, most leaders think that they are more effective than they actually are. Without being made cognizant that they need to improve their leadership and that they will be reevaluated, leaders simply might not be motivated to “self-improve” (London, 2002; Reilly et al., 1996).

Also, it is evident that training efforts cannot be done in a cursory manner; there are no quick fixes to learning leadership. For instance, in Study 1--including the workshop, the coaching session, completing self-ratings, reading their feedback reports, take-home readings, and time spent developing their leadership plans--we estimate that participants invested about 16 hours directly in the intervention as well as about one to two hours per week (for about 12 weeks) to plan and practice their newly acquired knowledge (thus about 30 hours at least). For Study 2, the MBA students spent considerably more time (about 80-90 hours in total) on their leadership development. Furthermore, Study 2 included healthy doses of the theory as well as many classroom discussions, where experiences were shared and students gave feedback to each

other. We also believe that the key success factors of our interventions were (a) the use of video cases to demonstrate effective leadership and (b) role-playing; in the case of the MBA group, (c) the filming of participants and the feedback they received was instrumental.

Based on participants' evaluation of the intervention, they appreciated the training approach used. The MBA participants particularly liked the filming sessions and feedback received. Interestingly--and this we report anecdotally--some participants initially felt uncomfortable using the CLTs thinking that they would seem inauthentic, which is reminiscent of the "illusion of transparency" phenomenon (Gilovich, Savitsky, & Medvec, 1998); that is, individuals incorrectly believe that their knowledge or feelings "leak out" to others. However, after the second filming session, they felt more comfortable about using them. Also, in debriefing sessions with the raters, none of them suspected that the video-taped participants were putting on an act or that charisma was manipulated.

Next, the training approach used seems to be useful for mature participants. Both the manager group and MBA groups had extensive work experience with mean ages of 42 and 35 years respectively. We do not know whether participants who are substantially younger will be able to also improve their leadership in the same way, though we are currently trying out interventions with younger master of management students.

Finally, a division president at a Fortune 50 company once noted his disillusionment with leadership training: "We spend \$120 million a year on this stuff [i.e., leadership development], and if it all went away tomorrow, it wouldn't matter one bit" (Ready & Conger, 2003: 86). To avoid such outcomes, companies should consider working with academia or with consultants who have academic training so as to assess the efficacy of interventions. As noted by Kluger and DeNisi "those who have a financial stake in the assumption that [feedback interventions] always

improves performance would have very little interest in carefully testing this assumption” (1996: 277). We would also recommend that companies only use validated interventions or consider conducting an experiment to test the intervention; companies should even consider paying consultants only if significant results can be demonstrated. Also, business-school educators should ensure that they use evidence-based practices (cf. Charlier et al., in press).

Limitations and Future Research

Both studies shed interesting findings on leadership development; however, there are limitations that temper our conclusions. First, it is not clear whether we have identified the best markers of charisma; future research should gather more complete data on these and other markers to see which ones better predict charismatic outcomes. Also, provided the sample size is large enough (in the context of a latent class analysis), it would be interesting to know which combination of markers work best in predicting outcomes (cf. Muthén & Shedden, 1999).

As for our dependent variables, although we obtained similar results in the field and laboratory experiments (with respect to charisma’s effect on prototypicality and outcomes), the laboratory experiment gauged leader emergence and not actual leader behavior *in situ*. That is, emerging as a leader does not necessarily mean that the leader is effective; thus, it is not clear from Study 2 whether participants just behaved leader-like or whether this change could transfer to industrial settings. Study 1, however, suggests that the effects of the intervention were evident in the work place and did demonstrate some temporal stability. Even then, however, the effect of the intervention was not measured objectively. Future research should gather objective outcomes measures (e.g., financial-based) that can examine whether leaders’ work groups actually become more effective because of a charisma manipulation (e.g., as in the study of Barling et al., 1996).

Although the Study 2 design allowed us to make relatively strong causal inferences it

would be desirable to replicate these results with a design that includes a control group. Study 1 showed that the control group increased significantly in charisma, which was probably due to a testing effect (Wilson & Putnam, 1982)--a widely observed phenomenon in a variety of domains. Theoretically, this testing effect will occur in conditions where subjects know that they will be remeasured and particularly when they can influence their future test scores. We presume that the subjects in the control group of Study 1 actually changed their leadership behavior (whether consciously or not) after Time 1 so as to receive better leader ratings in Time 2. Without a control group, the effects of the treatment in Study 1 would have been overstated; although we believe we removed the variance due to a testing effect in Study 2 (by controlling for the nonequivalent dependent variable), a cleaner design would include a control group.

Although small, the sample size of Study 1 was sufficiently large for the purpose of our experiment (i.e., we detected significant effects). Nonetheless, we hope that companies will become more committed to participating in large-scale experimental studies as more data showing the beneficial effects of interventions becomes published. Finally, to ensure that raters participated in a completely anonymous way, we did not record rater hierarchical relationships to leaders in Study 1; although this information loss was probably offset by more valid ratings, future research should control for these fixed-effects to see whether changes in leader behavior are more evident in some rater groups.

Conclusion

The results of our field and laboratory study indicated that charisma can be taught and that this effect had an impact on leader outcomes. We trust that the intervention approach we have developed will prove useful for educators and consultants, and this in an area where practice is often not evidence-based. Lamenting on the science-practice divide, Zaccaro and

Horn sounded a strong warning about practitioners' "undue reliance on popular ideas and fads without sufficient consideration given to the validity of these ideas" (2003: 779). That the practitioner literature is replete with hyperbolic yet untested claims regarding leader development is problematic both economically and ethically. Companies and business schools too (Charlier et al., in press) may be using methods that might not work, or worse, methods that are detrimental to leader outcomes.

Our results also showed that a substantial investment must be made and this to produce medium effects: However, the popular literature and consulting companies are a lot more sanguine about leadership development. Indeed, there are a plethora of leader development programs and cookbooks making lofty claims about the ease with which leadership can be taught; a simple internet search will demonstrate just how much optimism abounds. If these programs are much like evaluated feedback interventions, many if not most programs do not work: About 38% of rated interventions decrease future performance and 14% have no effect (Kluger & DeNisi, 1996).

To conclude, we hope that our results as well as others that show that leadership interventions work will not be used as a pretext to conduct any sort of intervention. Leadership, including charisma, can be developed and given the importance of leadership for organizational performance it is imperative to better understand the leadership development process. By understanding leader development we will also better understand leadership per se, as nicely summed up by an oft-quoted saying attributed to Kurt Lewin: *The best way to understand something is to try to change it.*

TABLE 1
Correlation Matrix of Key Variables (Study 1)

	Mean	St. dev.	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Prototypicality	5.68	.75													
2. Affect	6.18	.80	.82												
3. Trust	5.92	.73	.57	.73											
4. Competence	6.19	.82	.83	.85	.72										
5. Influence	5.35	.82	.71	.55	.46	.69									
6. Time 2 Charisma	2.94	.33	.78	.87	.80	.85	.57								
7. Experimental Treatment (=1; else = 0)	.50	.51	.18	.11	.06	.14	.05	.16							
8. Time 1 Charisma	2.47	.34	.70	.79	.74	.77	.59	.87	-.04						
9. Leader female (=1; else = 0)	.09	.29	-.04	.09	.14	-.06	-.18	.10	.31	.11					
10. Leader age	42.44	5.75	-.17	-.14	-.01	-.04	-.13	-.13	.05	-.14	-.04				
11. Time 1 Rater response %	74.19	12.52	.14	.22	.14	.15	.10	-.03	-.09	-.14	-.26	.25			
12. Time 2 Rater response %	74.02	20.67	.01	-.08	-.22	-.12	-.13	-.25	.12	-.27	-.19	.20	.32		
13. Majority German Raters (=1; else = 0)	.94	.24	.14	.14	.01	.17	.16	.06	.00	.00	-.36	.11	.05	.11	
14. Leader Language German (=1; else = 0)	.74	.45	-.22	-.01	-.05	-.08	-.29	-.09	-.07	-.23	-.05	.05	.03	-.06	.42

Note: n = 34 leaders (rated by n = 343 at Time1 and n = 321 at Time 2 raters); Correlations greater than |.50| are significant at $p < .001$; correlations greater than |.41| are significant at $p < .01$; correlations greater than |.33| are significant at $p < .05$; correlations greater than |.28| are significant at $p < .10$. Rater response percentages are mean response percentages on the leader level (on the group level the Time 1 and Time 2 response % was 74.08% and 73.42% respectively)

TABLE 2
OLS and 2SLS Regression Estimates (Study 1)

VARIABLES: DV in rows ; IVs in columns	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	T2 Charisma	Leader Prototyp	Leader Affect	Leader Trust	Leader Competen.	Leader Influen.	Leader Prototyp	Leader Affect	Leader Trust	Leader Competen.	Leader Influen.
Exp. Treatment	.15** (2.99)	.38* (2.39)	.27* (2.14)	.14 (.87)	.40** (2.66)	.21 (.99)					
T1 Charisma	.88** (12.27)	1.57** (6.71)	2.01** (10.90)	1.73** (7.45)	2.04** (9.30)	1.35** (4.34)					
Age	-.00 (-.62)	-.02 (-1.59)	-.02 (-1.82)	.00 (.24)	-.00 (-.13)	-.02 (-.92)					
Female	-.05 (-.48)	-.20 (-.64)	.24 (.96)	.23 (.72)	-.38 (-1.28)	-.51 (-1.23)					
T1 Rater resp.	.00 (1.50)	.02** (2.65)	.03** (5.06)	.02* (2.55)	.02** (2.77)	.01 (1.28)					
Maj. Germ. raters	-.01 (-.10)	.53 (1.37)	.33 (1.09)	-.10 (-.27)	.24 (.66)	.70 (1.36)	.48 (1.41)	.08 (.25)	-.19 (-.54)	.38 (1.11)	.95 (1.92)
Leader German	.10 (1.69)	-.19 (-.99)	.28 (1.82)	.25 (1.28)	.17 (.93)	-.44 (-1.72)	-.33 (-1.83)	.14 (.87)	.09 (.45)	-.06 (-.34)	-.65* (-2.45)
T2 Rater resp.							.01 (1.83)	.01 (1.93)	-.00 (-.02)	.00 (1.06)	-.00 (-.30)
T2 Charisma							1.94** (7.56)	2.44** (10.48)	1.92** (7.10)	2.33** (9.00)	1.46** (3.89)
Constant	.51 (1.70)	.93 (.96)	-.55 (-.71)	.11 (.12)	-.53 (-.58)	1.53 (1.18)	-.74 (-.84)	-1.65* (-2.05)	.38 (.41)	-1.27 (-1.42)	.76 (.58)
R-squared	.83	.67	.82	.65	.75	.50	.69	.78	.63	.73	.44

Note: $n = 34$; t-statistics in parentheses; ** $p < 0.01$, * $p < 0.05$; Estimates are unstandardized. Note: Models 1-6 are OLS models estimating the effect of the treatment on the dependent variables in a multivariate regression; Models 7-11 are the 2SLS estimates of the effect of T2 Charisma on the rest of the dependent variables; the chi-square test (Sargan, 1958) for overidentifying constraints indicated that the model was tenable: $\chi^2(21) = 32.44$, $p = .32$.

TABLE 3
Correlation Matrix of Key Variables (Study 2)

	Mean ¹	St. dev. ¹	1	2	3	4	5	6	7	8	9	10	Mean ²	St. dev. ₂
1. Prototypicality	3.55	2.20		.80	.76	.92	.93	.67	.32	.09	.27	.14	3.56	1.52
2. Affect	3.97	2.00	.67		.78	.78	.84	.68	.28	.07	.23	.08	3.97	1.18
3. Trust	4.10	2.09	.69	.74		.77	.77	.62	.12	-.01	.08	-.11	4.10	1.16
4. Competence	3.98	2.23	.86	.68	.69		.85	.69	.31	.07	.25	.12	3.98	1.40
5. Influence	3.64	2.29	.84	.75	.70	.78		.70	.36	.14	.32	.20	3.64	1.52
6. Rank	2.69	1.07	.59	.50	.48	.56	.59		.30	.10	.26	.05	2.67	.77
7. Charisma (verbal)	.28	.18	.24	.17	.08	.20	.26	.23		.46	.92	.58	.28	.18
8. Charisma (non-verbal)	.60	.32	.10	.07	.02	.08	.14	.11	.47		.77	.31	.61	.32
9. Charisma (overall)	.36	.19	.22	.15	.06	.18	.24	.21	.92	.77		.55	.36	.19
10. Communication skills	6.80	.67	.10	.05	-.05	.09	.14	.05	.59	.31	.56		6.81	.66

¹At the rater level (with correlations below the diagonal); ²at the leader level (with correlations above the diagonal). $n = 458$ observations; $n = 135$ raters evaluating $n = 41$ leaders. Note, for descriptive purposes, we include a verbal charisma composite score as well as a nonverbal composite score. The overall measure is the one used in the predictive model.

TABLE 4
Mixed-Process Estimates (Study 2)

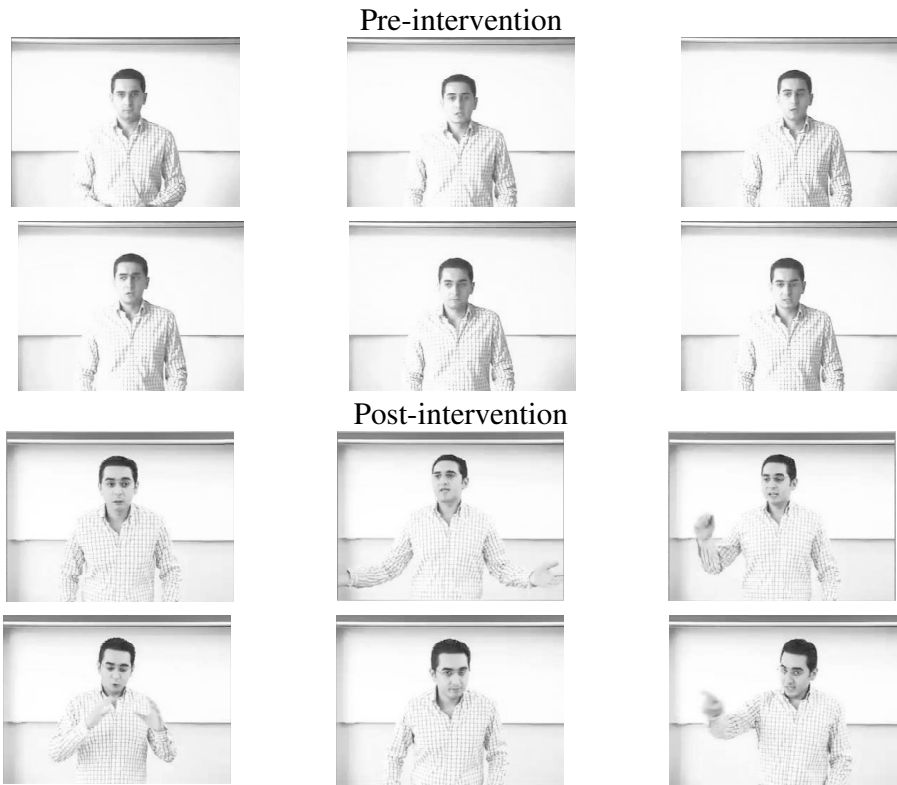
VARIABLES: Dependent variables in columns; Independent variables in rows	(1)	(2)	(3)	(4)	(5)	(6)
	Leader Prototyp.	Leader Affect	Leader Trust	Leader Competen.	Leader Influen.	Leader Rank
Charismatic leadership tactics	5.78** (5.27)	3.53** (3.29)	2.80* (2.29)	5.31** (4.06)	4.56** (3.55)	4.01** (5.59)
Communication Skills	-.43 (-1.30)	-.25 (-.88)	-.33 (-.97)	-.29 (-.81)	-.12 (-.33)	-.51** (-2.64)
Time	-.13 (-.67)	-.01 (-.04)	-.30 (-1.49)	-.37 (-1.75)	.01 (.06)	-.04 (-.30)
Rater male	-.27 (-1.46)	-.42* (-2.36)	-.59** (-2.91)	-.47* (-2.30)	-.35 (-1.73)	.08 (1.28)
Rater age	.03 (1.88)	-.01 (-.40)	.03 (1.50)	.03 (1.75)	.01 (.74)	.00 (.09)
Leader fixed-effects	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.
Constant	3.69 (1.74)	4.20* (2.31)	5.81** (2.88)	3.84 (1.72)	2.64 (1.18)	
Cut 1						-2.94
Cut 2						-1.98
Cut 3						-.99
F(45,134)	7.71**	4.03**	4.20**	6.03**	4.67**	
Wald $\chi^2(45)$						402.57**
R-square	.30	.22	.22	.25	.26	
Pseudo-R						.16

$n = 458$ observations; $n = 135$ raters evaluating $n = 41$ leaders (average raters per leader $n = 11.2$). Cluster robust z-statistics in parentheses. ** $p < .01$, * $p < .05$

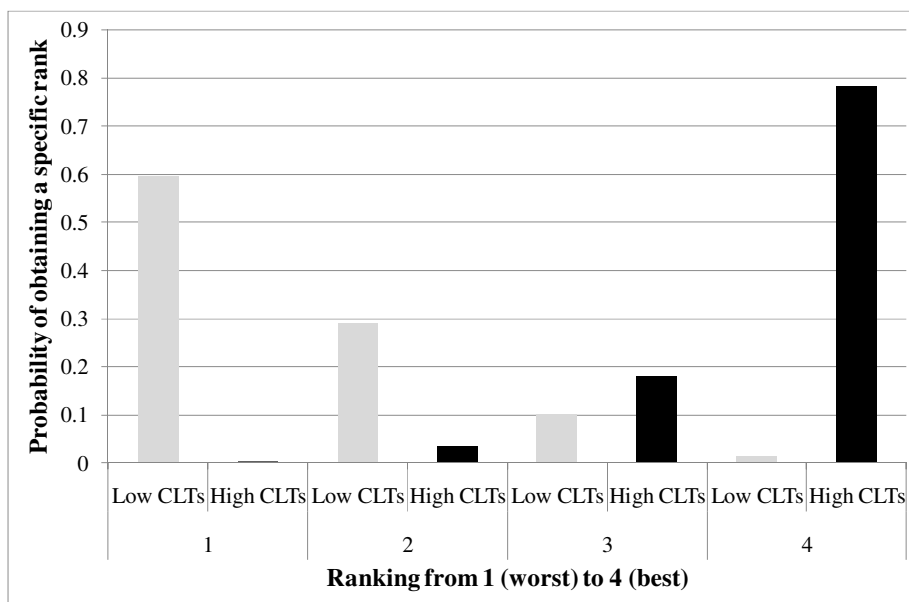
Note, Models 1-5 estimated with OLS regression; Model 6 estimated with ordered probit regression. Leader Rank refers to the emergence measure and is coded 4 (best) to 1 (worst). F -statistics, Wald χ^2 and R -squares of equation from independent OLS and ordered-probit models (and leader fixed-effects were significant). Estimates are unstandardized.

FIGURE 1

Panel A: Example of representative improvement in nonverbal behavior



Panel B: Probability of being ranked worst to best as a function of the Charismatic Leadership Tactics (CLTs)



REFERENCES

- ABC News. 2010. Local Candidate Gets Red in the Face *Available:* <http://abcnews.go.com/Politics/video/politician-phil-davison-freaks-out-during-speech-11600742> [accessed: 05.21.2011].
- Aditya, R. N. 2004. Leadership. In M. Hersen (Ed.), *Comprehensive handbook of psychological assessment*, Vol. 4: 216-239. Hoboken, N.J.: John Wiley & Sons.
- Aiken, L. R. 1987. Formulas for equating ratings on different scales. *Educational and Psychological Measurement*, 47: 51-54.
- Altenbernd, L., & Lewis, L. L. 1980. *Introduction to literature, stories* (3d ed.). New York: Macmillan.
- Antonakis, J. 2011. Transformational and Charismatic Leadership. In D. V. Day, & J. Antonakis (Eds.), *The nature of leadership*, 2nd ed.: 256-288. Thousand Oaks: Sage Publications.
- Antonakis, J., & Atwater, L. 2002. Leader distance: A review and a proposed theory. *The Leadership Quarterly*, 13: 673-704.
- 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(3): 261-295.
- Antonakis, J., Bendahan, S., Jacquart, P., & Lalive, R. 2010. On making causal claims: A review and recommendations. *The Leadership Quarterly*, 21(6): 1086-1120.
- Antonakis, J., & Dalgas, O. 2009. Predicting Elections: Child's Play! *Science*, 323(5918): 1183.
- Antonakis, J., & House, R. J. 2002. An analysis of the full-range leadership theory: The way forward. In B. J. Avolio, & F. J. Yammarino (Eds.), *Transformational and Charismatic Leadership: The Road Ahead*: 3-34. Amsterdam: JAI Press.
- Antonioni, D. 1994. The Effects of Feedback Accountability on Upward Appraisal Ratings. *Personnel Psychology*, 47: 349-356.
- Aristotle. 1954. *Rhetoric* (W. R. Roberts, I. Bywater, & F. Solmsen, Trans.) (1st Modern Library ed.). New York,: Modern Library.
- Atkinson, J. M. 2004. *Lend me your ears: All you need to know about making speeches and presentations*. London: Vermilion.
- Atwater, L., Roush, P., & Fischthal, A. 1995. The Influence of Upward Feedback on Self- and Follower Ratings and Leadership. *Personnel Psychology*, 48: 35-59.
- Atwater, L. E., & Yammarino, F. J. 1992. Does self-other agreement on leadership perceptions moderate the validity of leadership and performance predictions? *Personnel Psychology*, 45(1): 141-164.
- Avolio, B. J., Reichard, R. J., Hannah, S. T., Walumbwa, F. O., & Chan, A. 2009. A meta-analytic review of leadership impact research: Experimental and quasi-experimental studies. *The Leadership Quarterly*, 20(5): 764-784.
- Awamleh, R., & Gardner, W. L. 1999. Perceptions of leader charisma and effectiveness: The effects of vision content, delivery, and organizational performance. *The Leadership Quarterly*, 10(3): 345-373.
- Baltagi, B. H. 2002. *Econometrics*. New York: Springer.
- Bandura, A. 1977. *Social Learning Theory*. Englewood Cliffs, NJ: Prentice-Hall.

- Barling, J., Weber, T., & Kelloway, E. K. 1996. Effects of transformational leadership training on attitudinal and financial outcomes: A field experiment. *Journal of Applied Psychology*, 81(6): 827-832.
- Bass, B. M. 1985. *Leadership and performance beyond expectations*. New York: The Free Press.
- Bass, B. M., & Avolio, B. J. 1995. *MLQ Multifactor leadership questionnaire for research: Permission set*. Redwood City, CA: Mindgarden.
- Bass, B. M., & Avolio, B. J. 1997. *Full range leadership development: Manual for the multifactor leadership questionnaire*. Palo Alto, CA: Mindgarden.
- Bernard, L. C., Walsh, R. P., & Mills, M. 2005. Ask once, may tell: Comparative validity of single and multiple item measurement of the big-five personality factors. *Counseling and Clinical Psychology Journal*, 2: 40-57.
- Beyer, J. M. 1999. Taming and promoting charisma to change organizations. *The Leadership Quarterly*, 10(2): 307-330.
- Bliese, P. D. 2000. Within-group agreement, non-independence, and reliability: Implications for aggregation and analysis. In S. W. J. Kozlowski, & K. J. Klein (Eds.), *Multilevel theory, research, and methods in organizations*: 349-381. San Francisco, CA: Jossey-Bass.
- Bollen, K. A., Kirby, J. B., Curran, P. J., Paxton, P. M., & Chen, F. N. 2007. Latent variable models under misspecification - Two-stage least squares (2SLS) and maximum likelihood (ML) estimators. *Sociological Methods & Research*, 36(1): 48-86.
- Bono, J. E., & Ilies, R. 2006. Charisma, positive emotions and mood contagion. *The Leadership Quarterly*, 17: 317-334.
- Bouchard, T. J., & Loehlin, J. C. 2001. Genes, evolution, and personality. *Behavior Genetics*, 31(3): 243-273.
- Bouchard, T. J., & McGue, M. 2003. Genetic and environmental influences on human psychological differences. *Journal of Neurobiology*, 54(1): 4-45.
- Bower, G. H. 1976. Experiments on story understanding and recall. *Quarterly Journal of Experimental Psychology*, 28: 511-534.
- Brodbeck, F. C., Frese, M., Akerblom, S., Audia, G., Bakacsi, G., Bendova, H., Bodega, D., Bodur, M., Booth, S., Brenk, K., Castel, P., Den Hartog, D., & Donnelly-Cox, G. 2000. Cultural variation of Leadership Prototypes Across 22 European Countries. *Journal of Occupational and Organizational Psychology*, 73: 1-73.
- Bryman, A. 1993. Charismatic leadership in business organizations: Some neglected issues. *The Leadership Quarterly*, 4(3-4): 289-304.
- Butler, A. C., Chapman, J. E., Forman, E. M., & Beck, A. T. 2006. The empirical status of cognitive-behavioral therapy: A review of meta-analyses. *Clinical Psychology Review*, 26(1): 17-31.
- Carver, C. S., & Scheier, M. F. 1990. Origins and Functions of Positive and Negative Affect: A Control-Process View. *Psychological Review*, 97(1): 19-35.
- CBS. 2010. Phil Davison, Local GOP Candidate in Ohio, Gives Screeching Stump Speech. Available: http://www.cbsnews.com/8301-503544_162-20015989-503544.html [accessed: 05.21.2011].
- Charlier, S. D., Brown, K. G., & Rynes, S. L. in press. Teaching evidence-based management in MBA programs: What evidence is there? *Academy of Management Learning & Education*.

- Charteris-Black, J. 2005. *Politicians and rhetoric: The persuasive power of metaphor*. Basingstoke, UK: Palgrave-MacMillan.
- Chen, G., Kirkman, B. L., & Kanfer, R. 2007. A multilevel study of leadership, empowerment, and performance in teams. *Journal of Applied Psychology*, 92: 331–346.
- Cohen, J. 1988. *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Conger, J. A., & Kanungo, R. N. (Eds.). 1988. *Charismatic leadership: The elusive factor in organizational effectiveness*. San Francisco: Jossey-Bass Publishers.
- Conger, J. A., & Kanungo, R. N. 1998. *Charismatic leadership in organizations*. Thousand Oaks, CA: Sage Publications.
- Cronshaw, S. F., & Lord, R. G. 1987. Effects of categorization, attribution, and encoding processes on leadership perceptions. *Journal of Applied Psychology*, 72(1): 97-106.
- Daily Mail. 2010. 'I have a masters degree in communication': Republican candidate's bizarre speech becomes internet hit. *Available:* <http://www.dailymail.co.uk/news/worldnews/article-1310777/Phil-Davison-speech-YouTube-hit-I-degree-communications.html> [accessed: 05.21.2011].
- DeGroot, T., Kiker, D. S., & Cross, T. C. 2001. A Meta-Analysis to Review Organizational Outcomes Related Charismatic Leadership. *Canadian Journal of Administrative Sciences*, 17(4): 356-371.
- Den Hartog, D. N., House, R. J., Hanges, P. J., & Ruiz-Quintanilla, S. A. 1999. Culture specific and cross-culturally generalizable implicit leadership theories: Are attributes of charismatic/transformational leadership universally endorsed? *The Leadership Quarterly*, 10(2): 219-256.
- Den Hartog, D. N., & Verbarg, R. M. 1997. Charisma and rhetoric: Communicative techniques of international business leaders. *The Leadership Quarterly*, 8(4): 355-391.
- Downton, J. V. 1973. *Rebel leadership: Commitment and charisma in the revolutionary process*. New York: The Free Press.
- Dumdum, U. R., Lowe, K. B., & Avolio, B. J. 2002. A meta-analysis of transformational and transactional leadership correlates of effectiveness and satisfaction: An update and extension. In B. J. Avolio, & F. J. Yammarino (Eds.), *Transformational and charismatic leadership: The road ahead*: 35-66. Amsterdam: JAI.
- Dvir, T., Eden, D., Avolio, B. J., & Shamir, B. 2002. Impact of transformational leadership on follower development and performance: A field experiment. *Academy of Management Journal*, 45(4): 735-744.
- Eden, D. 1988. Pygmalion, Goal Setting, and Expectancy: Compatible Ways to Boost Productivity. *Academy of Management Review*, 13(4): 639-652.
- Eden, D., Geller, D., Gewirtz, D., Gordon-Terner, R., Inbar, I., Liberman, M., Pass, Y., Salomon-Segev, I., & Shalit, M. 2000. Implanting Pygmalion Leadership Style through Workshop Training: Seven Field Experiments. *The Leadership Quarterly*, 11(2): 171-210.
- Emrich, C. G., Brower, H. H., Feldman, J. M., & Garland, H. 2001. Images in words: Presidential rhetoric, charisma, and greatness. *Administrative Science Quarterly*, 46: 527–557.

- Ender, P. B., & Chen, X. 2008. Powerreg STATA ADO file. *UCLA: Academic Technology Services, Statistical Consulting Group*.
- Etzioni, A. 1961. *A comparative analysis of complex organizations*. New York: The Free Press.
- Etzioni, A. 1964. *Modern organizations*. Englewood Cliffs, New Jersey: Prentice-Hall.
- Fair, R. C. 2009. Presidential and Congressional Vote-Share Equations. *American Journal of Political Science*, 53(1): 55-72.
- Flynn, F. J., & Staw, B. M. 2004. Lend me your wallets: The effect of charismatic leadership on external support for an organization. *Strategic Management Journal*, 25: 309-330.
- Foster, E. M., & McLanahan, S. 1996. An Illustration of the Use of Instrumental Variables: Do neighborhood conditions affect a young person's change of finishing high school? *Psychological Methods*, 1(3): 249-260.
- French, J. R. P., & Raven, B. H. 1968. The bases of social power. In D. Cartwright, & A. F. Zander (Eds.), *Group dynamics: Research and theory*, 3rd ed.: 259-269. New York: Harper & Row.
- Frese, M., Beimeel, S., & Schoenborn, S. 2003. Action training for charismatic leadership: Two evaluations of studies of a commercial training module on inspirational communication of a vision. *Personnel Psychology*, 56: 671-697.
- Gasper, J. M. 1992. *Transformational leadership: An integrative review of the literature*. Western Michigan University, Michigan.
- Gemmill, G., & Oakley, J. 1992. Leadership: An Alienating Social Myth? *Human Relations*, 45(2): 113-129.
- Gennetian, L. A., Magnuson, K., & Morris, P. A. 2008. From statistical associations to causation: What developmentalists can learn from instrumental variables techniques coupled with experimental data. *Developmental Psychology*, 44(2): 381-394.
- Gilovich, T., Savitsky, K., & Medvec, V. H. 1998. The illusion of transparency: Biased assessments of others' ability to read one's emotional states. *Journal of Personality and Social Psychology*, 75(2): 332-346.
- Hogg, M. A. 2001. A social Identity theory of Leadership. *Personality and Social Psychology Review*, 5(3): 184-200.
- House, R. J. 1977. A 1976 Theory of Charismatic Leadership. In J. G. Hunt, & L. L. Larson (Eds.), *The Cutting Edge*. Carbondale: Southern Illinois: University Press.
- House, R. J. 1999. Weber and the Neo-charismatic Leadership Paradigm: A Response to Beyer. *The Leadership Quarterly*, 10(4): 563-574.
- House, R. J., Hanges, P. J., Javidan, M., Dorfman, P. W., & Gupta, V. 2004. *Culture, leadership, and organizations: The GLOBE study of 62 societies*. Thousand Oaks, Calif.: Sage Publications.
- House, R. J., Spangler, W. D., & Woycke, J. 1991. Personality and charisma and the U.S. presidency: A psychological theory of leader effectiveness. *Administrative Science Quarterly*, 36: 364-396.
- Howell, J. M., & Frost, P. J. 1989. A laboratory study of charismatic leadership. *Organizational Behavior and Human Decision Processes*, 43(2): 243-269.
- Huffington Post. 2010. Phil Davison, GOP Candidate, FREAKS OUT Trying To Sell Candidacy (VIDEO). Available: http://www.huffingtonpost.com/martin-olson/phil-davison-gop-speech_b_710642.html [accessed: 05.21.2011].

- Jacquart, P., & Antonakis, J. 2010. "It's the economy stupid," but charisma matters too: A dual-process model of presidential election outcomes, *Academy of Management, Organizational Behavior Division*. Montréal, Canada.
- Jones, B. F., & Olken, B. A. 2005. Do leaders matter? National leadership and growth since World War II. *The Quarterly Journal of Economics*: 835-864.
- Judge, T. A., Bono, J. E., Ilies, R., & Gerhardt, M. W. 2002. Personality and leadership: A qualitative and quantitative review. *Journal of Applied Psychology*, 87(4): 765-780.
- Judge, T. A., Colbert, A. E., & Ilies, R. 2004. Intelligence and leadership: A quantitative review and test of theoretical propositions. *Journal of Applied Psychology*, 89(3): 542-552.
- Judge, T. A., & Piccolo, R. F. 2004. Transformational and transactional leadership: A meta-analytic test of their relative validity. *Journal of Applied Psychology*, 89(5): 755-768.
- Kark, R., & Shamir, B. 2002. The dual effect of transformational leadership: Priming relational and collective selves and further effects on followers. In B. J. Avolio, & F. J. Yammarino (Eds.), *Transformational and charismatic leadership: The road ahead*: 67–91. Amsterdam: JAI Press.
- Kennedy, P. 2003. *A guide to econometrics* (5th ed.). Cambridge, MA: MIT Press.
- Keppel, G., & Wickens, T. D. 2004. *Design and Analysis: A researcher's handbook*. Upper Saddle River, NJ: Pearson.
- Keyes, C. F. 2002. Weber and Anthropology. *Annual Review of Anthropology*, 31: 233-255.
- Kluger, A. N., & DeNisi, A. 1996. The Effects of Feedback Interventions on Performance: A Historical Review, a Meta-Analysis, and a Preliminary Feedback Intervention Theory. *Psychological Bulletin*, 119(2): 254-284.
- Landis, J. R., & Koch, G. G. 1977. The measurement of observer agreement for categorical data. *Biometrics*, 33(1): 159-174.
- Liden, R. C., & Antonakis, J. 2009. Considering context in psychological leadership research. *Human Relations*, 62(11): 1587-1605.
- Lin, L. I. 1989. A concordance correlation-coefficient to evaluate reproducibility. *Biometrics*, 45(1): 255-268.
- Lindell, M. K., & Brandt, C. J. 1999. Assessing interrater agreement on the job relevance of a test: a comparison of the CVI, T, rwg(J), and r*WG(J) indexes. *Journal of Applied Psychology*, 84(4): 640-647.
- Lindell, M. K., & Brandt, C. J. 2000. Climate quality and climate consensus as mediators of the relationship between organizational antecedents and outcomes. *Journal of Applied Psychology*, 85(3): 331-348.
- Lindell, M. K., Brandt, C. J., & Whitney, D. J. 1999. A revised index of interrater agreement for multi-item ratings of a single target. *Applied Psychological Measurement*, 23(2): 127-135.
- Lissitz, R. W., & Green, S. B. 1975. Effect of the number of scale points on reliability: A Monte Carlo approach. *Journal of Applied Psychology*, 60: 10–13.
- Locke, E. A., & Latham, G. P. 2002. Building Practical Useful Theory of Goal Setting and Task Motivation. *American Psychologist*, 57(9): 705-717.
- London, M. 2002. *Leadership development: Paths to self-insight and professional growth*. Mahwah, N.J.: Lawrence Erlbaum Associates.

- Lord, R. G., Brown, D. J., Harvey, J. L., & Hall, R. J. 2001. Contextual Constraints on Prototype Generation and their Multilevel Consequences for Leadership Perceptions. *The Leadership Quarterly*, 12: 311-338.
- Lord, R. G., De Vader, C. L., & Alliger, G. M. 1986. A Meta-analysis of the relation between personality traits and leadership perceptions: An application of validity generalization procedures. *Journal of Applied Psychology*, 71(3): 402-410.
- Lord, R. G., Foti, R. J., & De Vader, C. L. 1984. A Test of Leadership Categorization Theory: Internal Structure, Information Processing, and Leadership Perceptions. *Organizational Behavior and Human Performance*, 34: 343-378.
- Lowe, K. B., Kroeck, K. G., & Sivasubramaniam, N. 1996. Effectiveness correlates of transformational and transactional leadership: A meta-analytic review of the MLQ literature. *The Leadership Quarterly*, 7(3): 385-425.
- MacKenzie, S. B. 2003. The dangers of poor construct conceptualization. *Journal of the Academy of Marketing Science*, 31: 323-326.
- Magin, D., & Helmore, P. 2001. Peer and teacher assessments of oral presentation skills: how reliable are they? *Studies in Higher Education*, 26(3): 287-298.
- Maxwell, S. E., Cole, D. A., Arvey, R. D., & Salas, E. 1991. A comparison of methods for increasing power in randomized between-subjects designs. *Psychological Bulletin*, 110(2): 328-337.
- Mio, J. S. 1997. Metaphor and politics. *Metaphor and symbol*, 12: 113-133.
- Mio, J. S., Riggio, R. E., Levin, S., & Reese, R. 2005. Presidential leadership and charisma: The effects of metaphor. *The Leadership Quarterly*, 16(2): 287-294.
- Morhart, F. M., Herzog, W., & Tomczak, T. 2009. Brand-Specific Leadership: Turning Employees into Brand Champions. *Journal of Marketing*, 73(5): 122-142.
- Mount, M. K., & Scullen, S. E. 2001. Multisource feedback ratings: What do they really measure? In M. London (Ed.), *How people evaluate others in organizations*: 155-176. Mahwah, NJ: Lawrence Erlbaum.
- Muthén, B. O., & Shedden, K. 1999. Finite mixture modeling with mixture outcomes using the EM algorithm. *Biometrics*, 55(2): 463-469.
- Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. 2003. Common Method Biases in Behavioral Research: A Critical Review of the Literature and Recommended Remedies. *Journal of Applied Psychology*, 89(5): 879-903.
- Podsakoff, P. M., MacKenzie, S. B., Podsakoff, N. P., & Lee, J. Y. 2003. The mismeasure of man(agement) and its implications for leadership research. *The Leadership Quarterly*, 14(6): 615-656.
- Podsakoff, P. M., & Organ, D. W. 1986. Self-reports in organizational research: Problems and prospects. *Journal of Management*, 12(4): 531-544.
- Preston, C. C., & Colman, A. M. 2000. Optimal number of response categories in rating scales: Reliability, validity, discriminating power, and respondent preferences. *Acta Psychologica*, 104: 1-15.
- Ready, D. A., & Conger, J. A. 2003. Why Leadership-Development Efforts Fail. *Sloan Management Review*, 44: 83-88.
- Ree, M. J., & Carretta, T. R. 2006. The role of measurement error in familiar statistics. *Organizational Research Methods*, 9(1): 99-112.
- Reilly, R. R., Smither, J. W., & Vasilopoulos, N. L. 1996. A longitudinal study of upward feedback. *Personnel Psychology*, 49(3): 599-612.

- Roodman, D. M. 2008. Cmp: Stata module to implement conditional (recursive) mixed process estimator. <http://ideas.repec.org/c/boc/bocode/s456882.html>.
- Rosenthal, R., & Rubin, D. B. 1982. A simple, general purpose display of magnitude of experimental effect. *Journal of Educational Psychology*, 74(2): 166-169.
- Rynes, S. L., Colbert, A. E., & Brown, K. G. 2002. HR professionals' beliefs about effective human resource practices: Correspondence between research and practice. *Human Resource Management*, 41(2): 149-174.
- Sargan, J. D. 1958. The Estimation of Economic Relationships Using Instrumental Variables. *Econometrica*, 26: 393-415.
- Scullen, S. E., Mount, M. K., & Goff, M. 2000. Understanding the latent structure of job performance ratings. *Journal of Applied Psychology*, 85(6): 956-970.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. 2002. *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin.
- Shamir. 1999. Taming charisma for better understanding and greater usefulness: a response to Beyer. *The Leadership Quarterly*, 10(4).
- Shamir, B. 1995. Social Distance and Charisma - Theoretical Notes and an Exploratory-Study. *The Leadership Quarterly*, 6(1): 19-47.
- Shamir, B., Arthur, M. B., & House, R. J. 1994. The rhetoric of charismatic leadership: A theoretical extension, a case study, and implications for research. *The Leadership Quarterly*, 5(1): 25-42.
- Shamir, B., House, R. J., & Arthur, M. B. 1993. The motivational effects of charismatic leadership: A self-concept based theory. *Organization Science*, 4(4): 577-594.
- Shamir, B., & Howell, J. M. 1999. Organizational and contextual influences on the emergence and effectiveness of charismatic leadership. *The Leadership Quarterly*, 10(2).
- Shaver, J. M. 2005. Testing for mediating variables in management research: Concerns, implications, and alternative strategies. *Journal of Management*, 31(3): 330-353.
- Shils, E. 1965. Charisma, Order, and Status. *American Sociological Review*, 30(2): 199-213.
- Todorov, A., Mandisodza, A. N., Goren, A., & Hall, C. C. 2005. Inferences of competence from faces predict election outcomes. *Science*, 308(5728): 1623-1626.
- Towler, A. J. 2003. Effects of charismatic influence training on attitudes, behavior, and performance. *Personnel psychology*, 56(2): 363 - 381.
- Wasielewski. 1985. The emotional basis of charisma. *Symbolic interaction*, 8(2): 207-222.
- Weber, M. 1947. *The theory of social and economic organization* (T. Parsons, Trans.). New York: The Free Press.
- Weber, M. 1968. *Max Weber on charisma and institutional building* (Ed. S. N. Eisenstadt). Chicago: The University of Chicago Press.
- Willis, J., & Todorov, A. 2006. First impressions: Making up your mind after a 100-ms exposure to a face. *Psychological Science*, 17(7): 592-598.
- Willner, A. R. 1984. *The spellbinders: Charismatic political leadership*. New Haven: CT.: Yale University Press.
- Wilson, V. L., & Putnam, R. R. 1982. A meta-analysis of pretest sensitization effects in experimental design. *American Educational Research Journal*, 19: 249-258.
- Wooldridge, J. M. 2002. *Econometric analysis of cross section and panel data*. Cambridge, MA: MIT Press.

- Yukl, G. A. 1999. An evaluation of conceptual weaknesses in transformational and charismatic leadership theories. *The Leadership Quarterly*, 10(2): 285-305.
- Yukl, G. A. 2006. *Leadership in organizations* (6th ed.). Upper Saddle River, NJ: Pearson/Prentice Hall.
- Zaccaro, S. J., & Horn, Z. N. J. 2003. Leadership theory and practice: Fostering an effective symbiosis. *The Leadership Quarterly*, 14(6): 769-806.