American Psychological Association

Journal of Applied Psychology

0021-9010/18/$12.00

2018, Vol. 103, No. 7, 738–752
http://dx.doi.org/10.1037/apl0000302

An Exploration of the Role of Subordinate Affect in Leader Evaluations

Mark J. Martinko
Florida A & M University

Jeremy D. Mackey
Auburn University

Sherry E. Moss
Wake Forest University

Paul Harvey
University of New Hampshire

Charn P. McAllister
Northeastern University

Jeremy R. Brees
Northern Arizona University

Leadership research has been encumbered by a proliferation of constructs and measures, despite little evidence that each is sufficiently conceptually and operationally distinct from the others. We draw from research on subordinates' implicit theories of leader behavior, behaviorally anchored rating scales, and decision making to argue that leader affect (i.e., the degree to which subordinates have positive and negative feelings about their supervisors) underlies the common variance shared by many leadership measures. To explore this possibility, we developed and validated measures of positive and negative leader affect (i.e., the Leader Affect Questionnaires; LAQs). We conducted 10 studies to develop the five-item positive and negative LAQs and to examine their convergent, discriminant, predictive, and criterion-related validity. We conclude that a) the LAQs provide highly reliable and valid tools for assessing subordinates' evaluations of their leaders; b) there is significant overlap between existing leadership measures, and a large proportion of this overlap is a function of the affect captured by the LAQs; c) when the LAQs are used as control variables, in most cases, they reduce the strength of relationships between leadership measures and other variables; d) the LAQs account for significant variance in outcomes beyond that explained by other leadership measures; and e) there is a considerable amount of unexplained variance between leadership measures that the LAQs do not capture. Research suggestions are provided and the implications of our results are discussed.

Keywords: affect, leadership, implicit leadership theory, behaviorally anchored rating scales, decision making

Supplemental materials: http://dx.doi.org/10.1037/apl0000302.supp

In this article, we describe the development and validation of leader affect scales based on the premise that subordinate evaluations of leaders are largely a function of affect, which we conceptualize as the degree to which subordinates have positive and negative feelings about their supervisors. We argue that such scales are necessary to identify the effects of affect on subordinates' ratings of leaders. We also propose that affect is a primary driver of subordinate evaluations of their leaders and should not be ignored or viewed only as a nuisance factor or control variable.

The foundation for our research relies on three bodies of research. First, the work of Lord and his colleagues (e.g., Lord, Binning, Rush, & Thomas, 1978; Lord, Foti, & De Vader, 1984; Lord & Shondrick, 2011; Rush, Thomas, & Lord, 1977; Shondrick, Dinh, & Lord, 2010) demonstrates that evaluations of lead-

This article was published Online First March 26, 2018.

Mark J. Martinko, School of Business and Industry, Florida A & M University; Jeremy D. Mackey, Raymond J. Harbert College of Business, Auburn University; Sherry E. Moss, School of Business, Wake Forest University; Paul Harvey, Peter T. Paul College of Business and Economics, University of New Hampshire; Charn P. McAllister, D'Amore-McKim School of Business, Northeastern University; Jeremy R. Brees, W. A. Franke College of Business, Northern Arizona University.

We thank Neal Ashkanasy (University of Queensland) and Robert Lord (Durham University) for their helpful feedback regarding a draft of this article. Thanks also go to the National Institute for Occupational Safety and Health (NIOSH) Pilot Research Project through the Sunshine Education and Research Center at the University of South Florida for funding a research grant for Charn McAllister that provided the data that were examined in several of the studies. An earlier version of this paper Martinko, M., Mackey, J., McAllister, C., Harvey, P., Moss, S., &. Brees, J. "An exploration of the role of subordinate affect in leader evaluations" was presented at the annual meeting of the Academy of Management, in Atlanta, August, 2017.

Correspondence concerning this article should be addressed to Mark J. Martinko, School of Business and Industry, Florida A & M University, 500 Gamble Street, Tallahassee, FL 32307-5200. E-mail: mark.martinko@famu.edu

ers are typically a function of categorization processes that are dependent upon subordinates' implicit theories of leader behavior. According to this perspective, leaders who match subordinates' implicit templates of good leaders are evaluated more positively than those who do not. Lord and his colleagues contend that this template matching process is not the result of deliberate cognitive processes, but is more appropriately described as an instantaneous recognition process. We argue that this instantaneous recognition process is cued by the affect that subordinates feel toward their supervisors.

A second influence on our premise is the work on developing and validating performance appraisals, particularly the work on behaviorally anchored rating scales (BARS; Smith & Kendall, 1963). The initial proposition of this work was that requiring supervisors to evaluate specific performance-related behaviors would lead to more valid performance ratings than traditional appraisal formats that are based on general dimensions such as productivity, quality, and dependability. However, researchers found that ratings of performance were based more on a general factor relating to performance, rather than the specific criteria on the appraisal forms (e.g., Viswesvaran, Schmidt, & Ones, 2005). Similarly, we argue that impressions of supervisor effectiveness are largely guided by a general factor relating to subordinate affect.

A third influence on our basic thesis is research on the role of affect in decision making (e.g., Damasio, 1994; Lerner, Li, Valdesolo, & Kassam, 2015). In general, this research demonstrates that affect is the primary stimulus driving most important decisions. We also draw from affective events theory (Weiss & Cropanzano, 1996), which contends that affective reactions to work related events (e.g., positive or negative feedback by a supervisor) are primary determinants of subordinates' moods and emotions, which, in turn, affect both cognition and behavior. Thus, we argue that the affect subordinates feel about their supervisors drives the decision-making processes subordinates use to evaluate their leaders.

We draw from these three bodies of research to support our argument that the majority of variance in subordinates' leadership ratings is attributable to affect regarding their leaders. To test this premise, we conducted a series of 10 studies in order to a) develop and validate reliable measures of positive and negative leader affect (i.e., the Leader Affect Questionnaires; LAQs); b) show that there is significant overlap between leadership measures related to the affect captured by the LAQs; c) demonstrate that the LAQs can reduce the strength of the relationships between leadership measures and other variables when they are used as control variables; and d) show that the LAQs account for significant variance in outcomes beyond that explained by other leadership measures. Ultimately, our findings demonstrate that affect may be the common factor explaining why many leadership measures are so highly intercorrelated.

Conceptual Foundations

Subordinate Evaluations of Leadership

A series of empirical and theoretical works by Lord and his colleagues (e.g., Cronshaw & Lord, 1987; Lord et al., 1984; Lord, Brown, Harvey, & Hall, 2001; Lord & Shondrick, 2011; Rush et al., 1977; Shondrick et al., 2010), as well as others (e.g., Eden & Leviatan, 1975; Hunter, Bedell-Avers, & Mumford, 2007) provides convincing evidence that the accuracy of subordinates' evaluations of leaders is questionable and subject to a wide range of influences. Much of the research and discussion of subordinate biases has centered on the construct of implicit leadership theories (ILTs). According to this perspective, followers' evaluations of leaders are typically a result of relatively automatic and reactive categorization processes instead of systematic rational processes such as those described by attribution theory (e.g., Weiner, 1986). Early research by Eden and Leviatan (1975) demonstrated that the factor structure of ratings of both real and hypothetical leaders were almost identical. Constructive replications by Rush et al. (1977) as well as by Weiss and Adler (1981) supported the notion that ratings of leader behaviors were strongly influenced by raters' ILTs. A series of studies followed that demonstrated support for the notion that evaluations of leaders are influenced by a variety of factors, including ILTs (Cronshaw & Lord, 1987), performance information (Cronshaw & Lord, 1987; Lord et al., 1978), and context (Lord et al., 2001).

A recent article by Hansbrough, Lord, and Schyns (2015) attempted to integrate the research and theory regarding the factors influencing subordinates' evaluations of leaders. The authors discussed a variety of factors that have been found to influence, or are expected to influence, subordinates' ratings of leaders. These factors included followers' individual differences (e.g., personality, positive and negative affectivity, attribution styles, and motives), contextual factors (e.g., leader individual differences such as height, facial appearance, and gender, as well as culture and research methods), and mediating factors (e.g., stereotype activation, perceived similarity, liking, and mood). Hansbrough et al. concluded that all of these factors interact and contribute to inaccuracy in subordinate ratings. They noted that many of the factors are endogenous, which makes it almost impossible to disaggregate their specific effects on leader evaluations. The authors did not claim precedence for any of these factors in the prediction of subordinates' ratings of leaders. In contrast, we propose that affect toward their supervisors is one of the most important factors influencing how subordinates rate their leaders.

Hansbrough et al. (2015) discussed liking as one of several factors influencing leadership evaluations, and reviewed two studies that suggest liking (i.e., affect) has a major influence on leader evaluations. The first study (i.e., Brown & Keeping, 2005) explored how liking impacted ratings of transformational leadership. Similar to our conceptualization of affect, Brown and Keeping defined liking as the amount of interpersonal affect that subordinates have toward their supervisors. Their study found that liking was significantly related to all five dimensions of transformational leadership (i.e., idealized influence attributes, idealized influence behaviors, inspirational motivation, intellectual stimulation, and individualized consideration), and accounted for so much variance that, when they controlled for liking, many of the relationships between transformational leadership and the outcome variables failed to achieve significance.

The second study (i.e., Naidoo, Kohari, Lord, & DuBois, 2010) found that liking accounted for more than 60% of the variance in subordinates' LMX ratings, as well as a majority of the variance in trust and charisma. These results replicated a study by Engle and Lord (1997), which also found that more than 60% of the variance in LMX ratings was accounted for by subordinates' liking. Earlier

studies (Lewter & Lord, 1992; Wayne & Ferris, 1990) provide further support for the notion that a large proportion of the variance in leader ratings is related to subordinate affect. Although this line of research appears promising, we were unable to find any additional work relating to the role of affect in subordinates' leadership ratings. Perhaps the choice of researchers not to explore the role of affect further is because their intent was to develop reliable estimators for the effects of leader behavior rather than explore the process underlying these estimates. For example, Hansbrough et al. (2015, p. 226) commented "because liking is not exogenous, but rather part of the process that causes leadership, controlling for liking will still not produce consistent estimators for effects of leader behavior".

While Hansbrough et al. (2015) considered liking as one of many factors biasing subordinate ratings of their leaders, we propose that subordinates' affective reactions to their supervisors are the primary mechanism by which their ILTs are triggered and encoded. More specifically, the work of Srull and Wyer (1989) and Zajonc (1980) indicates that affective reactions are instantaneous and initiate cognitive processes, such as the recall of schemata. Thus, we propose that the positive and negative affective experiences that subordinates recall when responding to leadership questionnaires trigger schemata they have built regarding their supervisors. If the triggered affective reaction is positive, responses to questions that relate to the subordinates' ILTs are positive. If the triggered affective reaction is negative, it invokes a negative comparison to the ILT, resulting in negative responses to the specific items in questionnaires, regardless of their content. Although this process can be viewed as a halo effect, we view the affective reaction as an antecedent that triggers the ILTs that are specific to each subordinate and the schemata that are specific to each subordinate's supervisor. As a result, we expect that the majority of the variance in leadership questionnaires is related to the amount of positive and negative affect subordinates experience when asked to evaluate their supervisor.

Behaviorally Anchored Rating Scales (BARS)

A second body of literature which informed our thinking is the research on BARS (Smith & Kendall, 1963). BARS were developed to guide evaluators to use specific job-oriented behaviors to overcome the biases associated with graphic rating scales, which are numerically, rather than behaviorally, anchored. By focusing on behaviors, it was theorized that biases and the high intercorrelations between scales could be overcome, resulting in more valid evaluations.

Research failed to establish the superiority of BARS. An early review by Schwab, Heneman and DeCotis (1975) concluded that there was little difference in the reliability and validity of BARS versus other rating formats. Other reviews (Atkin & Conlon, 1978; Kingstrom & Bass, 1981; Landy & Farr, 1980; Murphy & Constans, 1987; Nathan & Alexander, 1985) reached similar conclusions. A comprehensive review by Viswesvaran et al. (2005) proposed that the failure of BARS to demonstrate the independence of appraisal dimensions reflects the existence of a general factor common to all of the performance dimensions. Viswesvaran et al. speculated that this common factor was related to citizenship behavior but acknowledged that the exact nature of the factor was unclear.

Nathan and Alexander (1985) contended that performance ratings result from impression-based cognitive structures, as opposed to observations of specific behaviors. According to this perspective, evaluators have implicit theories about performance. When evaluating an individual, raters rely upon their general impression (e.g., good or bad), which is then matched to their implicit theory of performance. This position is supported by several studies (e.g., Binning, Zaba, & Whattam, 1986; McElroy & Downey, 1982; Phillips & Lord, 1982). We expect that a similar evaluative process affects subordinates' judgments of their supervisors and that affective reactions to supervisors are the primary stimulus which initiates this process. Although it may be coincidental, a general factor accounted for about 60% of the variance in subordinate performance appraisals (Viswesvaran et al., 2005) and about 60% of the variance in ratings found in the two studies on liking described earlier (i.e., Brown & Keeping, 2005; Naidoo et al., 2010). Thus, we propose that the general factor that Viswesvaran et al. found is related to liking and that a large proportion of the variance in subordinates' ratings of their supervisors is attributable to subordinate affect (perhaps as much as 60%).

Research on Decision Making and Emotion

We view subordinates' responses to questionnaires as a series of decisions. Research on decision-making often questions the rationality of the decision process. Decision makers frequently use heuristics (i.e., cognitive short-cuts; Simon, 1967) to simplify the process. Fiske and Taylor (1991) argued that in many situations, decision makers are motivated to conserve effort and cognitive resources. They described decision makers as cognitive misers who resort to heuristics when resources are strained or the task is not viewed as important.

Relatedly, Baumeister and Bushman (2008) drew from theorizing in social psychology (e.g., Chaiken & Trope, 1999; Srull & Wyer, 1988) to describe a dual process model of cognitive processing. One model is automatic, quick, and efficient, relying on short-cuts such as heuristics and implicit theories, whereas the other model is more conscious and thoughtful, such as the processes described by attribution (Weiner, 1985) and stress theories (Lazarus & Folkman, 1984). The second and more elaborate process is activated when the situation is viewed as surprising, important, or stressful. The more efficient and automatic processes are activated when the outcome is not viewed as particularly important or stressful. We expect that subordinates do not generally view situations in which they are asked to rate their leaders as particularly stressful or important. As a result, we believe that respondents are motivated to conserve resources, which is consistent with Fiske and Taylor's (1991) cognitive miser analogy and the types of heuristics (e.g., implicit theories) described by the automatic part of the dual process model.

Research also indicates that emotions play a central role in decision making. The neuropsychological work by Damasio (1994) and Bechara, Damasio, Damasio, and Lee (1999) demonstrated that profound deficits in decision making capacities exist for individuals who have experienced neurological impairments in the emotional centers of the brain. In a recent review of the relationships between motivation and decision making, Lerner et al. (2015, p. 801) concluded that "many psychological scientists

now assume that emotions are, for better or worse, the dominant driver of most meaningful decisions in life".

In their review, Lerner at al. (2015) emphasized that emotions typically precede cognitions, and emotions are often viewed as initiating cognitive processes. The models of stress by Lazarus and Folkman (1984) and the attribution process by Weiner (1986) both begin with a primary appraisal of the situation that results in an affective reaction as to whether the situation is "good" or "bad". We expect that when subordinates are asked to evaluate their supervisors via questionnaires, they also have immediate affective reactions depending on the scripts and schemata they have stored regarding their supervisor. Further, we expect that the scripts and schemata are then matched to their ILTs. Also, we expect that for many subordinates, initial affective reactions to their supervisors drive their leadership ratings. We also recognize that because of dual processing and individual differences, some subordinates may engage in the conscious processing described by the dual process model. Thus, although a major proportion of the variance in subordinates' ratings of their supervisors may be attributable to affect, other factors are also important.

In addition to the perspective we gained from the neuropsychological work, recent theory and research concerning affective events theory (Weiss & Cropanzano, 1996) also informed our thinking. According to affective events theory, individuals experience a continuous series of affective events throughout their workday that influences their moods and emotions which, in turn, influence their work attitudes and behaviors. Given that supervisors are a major factor in subordinates' work environments, we expect that the affective events associated with their supervisors have a considerable influence on their evaluations of their leaders. We reason that as subordinates experience affective events involving their supervisors, the events are continuously encoded into affective schemata. When asked to evaluate their supervisors, we expect that these affective schemata are activated and matched with subordinates' ILTs.

Summary: Theoretical Foundations

Combining what we have learned from the three perspectives described above, we believe that when subordinates are asked to rate their supervisors via questionnaires, the process begins with their affective reactions to their supervisors, both positive and negative. Subordinates then match their affective supervisor schemata to their ILT so the degree to which supervisors are rated positively or negatively corresponds to the degree that subordinates' supervisory schemata match their ILTs. As a result, we expect that subordinate affect accounts for a major proportion of the variance in subordinates' ratings of their leaders. In order to explore the contribution of subordinate affect to leadership ratings, we developed and validated the positive and negative LAQs, which we describe below. After developing these measures, we conducted constructive replications of three leadership studies and included the LAQs to explore the role that affect plays in subordinates' leadership ratings.

Method: Development and Validation of the LAQs[1]

We followed the recommendations of Hinkin (1995, 1998) and MacKenzie, Podsakoff, and Podsakoff (2011) to use a multistep procedure to develop and validate the positive and negative LAQs. We developed the positive LAQ measure in Studies 1–3 using content validity assessments, exploratory factor analyses (i.e., EFAs), and confirmatory factor analyses (i.e., CFAs). Then, we developed the negative LAQ measure in Studies 4–6 using the same measure development procedures we used for the positive LAQ. In Study 7, we assessed the convergent and discriminant validity of both measures. In Studies 8–10, we demonstrated criterion-related and predictive validity by examining partial correlations and controlling for affect with the LAQs while replicating prior leadership studies in the areas of abusive supervision, transformational leadership, transactional leadership, ethical leadership, and LMX.

Study 1: Item Development and Content Validity Assessment of the Positive LAQ

Item Development

We used Hinkin's (1998) deductive scale development approach to generate items for the positive LAQ. Specifically, we used our definition of positive leader affect (i.e., the degree to which subordinates have positive feelings about their supervisors) to create potential items for the positive LAQ. We drew inspiration from prior theories and measures of affect (i.e., Watson, Clark, & Tellegen, 1988; Weiner, 1985; Weiss & Cropanzano, 1996), social exchange (i.e., Bernerth, Armenakis, Feild, Giles, & Walker, 2007; Colquitt, Baer, Long, & Halvorsen-Ganepola, 2014), and LMX (i.e., Graen & Uhl-Bien, 1995; Liden & Maslyn, 1998) when creating the initial pool of items for the positive LAQ. Ultimately, we developed an initial list of 78 potential items that included content relevant to our conceptualization of leader affect (Podsakoff, MacKenzie, & Podsakoff, 2016). We followed Hardy and Ford's (2014) recommendations to assess the initial list of 78 items for content validity.

Content Validity

We selected five faculty members and six doctoral students from universities in the United States as respondents for a Q-sort activity. Each respondent was sent an electronic survey that included all 78 of the items initially developed. Respondents were asked to indicate the extent to which they agreed or disagreed that each of the items corresponded with our definition of positive leader affect according to an agreement scale (1 = "Strongly disagree" to 7 = "Strongly agree"). The definition of leader affect was presented before each set of 10 items to keep the definition salient for respondents. The mean score for each group (i.e., faculty member, doctoral student) was weighted equally in order to create a composite score. We retained the 10 items for further analysis that had a weighted mean of at least 6.00.

[1] The Florida State University Institutional Review Board approved this study (protocol number: 2014.12636, title of study: "Stress MTurk Spring, 2014"; protocol number: 2017.20247, title of study: "Leader Affect Replications).

Study 2: Initial Factor Structure Assessment of the Positive LAQ via EFAs

Sample and Procedures

We recruited 300 respondents from Amazon's Mechanical Turk (MTurk; Buhrmester, Kwang, & Gosling, 2011) platform to examine the factor structure of the 10 items that were retained from Study 1. Respondents were paid 50 cents each to complete a survey that included demographic items and the 10 positive LAQ items from Study 1. Respondents were instructed to think about their immediate supervisors and indicate the extent to which they agreed or disagreed with the 10 positive LAQ items by selecting the appropriate choice along a seven-point agreement scale (1 = "Strongly disagree" to 7 = "Strongly agree").

The final sample consisted of 272 respondents who provided usable data and met this study's inclusion criteria (i.e., 90.67% usable data), which required that respondents were adults at least 18 years of age who worked with direct supervisors at least 20 hr per week in U.S. organizations other than MTurk. Also, we included an instructed item in the survey to ensure that respondents were dedicating appropriate levels of attention throughout the completion of the survey (DeSimone, Harms, & Desimone, 2015; Meade & Craig, 2012), and screened respondents' MTurk identification codes (IDs) to ensure that no respondent was included in the study more than once. The 272 respondents in the final sample averaged 31.99 years of age (SD = 8.79), 5.11 years of tenure with their organizations (SD = 4.97), 3.20 years of tenure with their supervisors (SD = 3.15), and 37.43 hr of work per week at organizations other than MTurk (SD = 9.09). The majority of respondents (63.97%) were male. Respondents represented various levels of education (i.e., 45.59% held bachelor's degrees, 21.32% completed some college), job functions (e.g., 14.71% sales, 12.50% administration), levels in U.S. organizations (e.g., 48.16% staff/associate level, 22.06% middle management), and industries (e.g., 17.65% educational, health, and social services, 15.07% retail trade).

EFA Results

We conducted EFAs (see Preacher & MacCallum, 2003) with SPSS 23.0 using principal axis factor extraction and varimax (i.e., orthogonal) rotation, per Hinkin's (1998) recommendations. We required that items demonstrate a loading of .40 or higher on only one factor in order to be retained. We also required Eigenvalues greater than 1.00 and used visual examination of the scree plot to retain factors. The results of the EFAs are presented in Appendix A and demonstrate that all 10 items loaded strongly on the first and only identified factor, which accounted for about 81.71% of the explained variance.

A purpose of this study was to develop a brief measure, so we retained the five highest loading items. Then, we examined the factor structure of the five-items to determine whether the construct validity suffered by shortening the measure. The overlapping content in the retained and deleted items, the loading of all of the items on a single factor with reflective indicators, the high interitem correlations, and the nearly perfect correlation between the five-item and 10-item measures of positive LAQ (i.e., r = .99, p < .01) justified the criteria we applied for retaining items. The five items that comprised the final positive LAQ strongly loaded on a single factor in the EFA, which accounted for 86.09% of the explained variance. Table 1 provides a full list of the positive LAQ items. Next, we conducted CFAs using an organizational sample of data so we could further examine the factor structure of the positive LAQ.

Table 1
Summary of CFA Results and Descriptive Statistics for the Positive Leader Affect Questionnaire (Positive LAQ) Across Studies

Items and statistics	Study 3 Initial factor structure (n = 342)	Study 7 Convergent and discriminant validity (n = 296)	Study 8 Abusive supervision replication (n = 806)	Study 9 MLQ Replication (n = 507)	Study 10 LMX and ethical leadership replication (n = 213)
Confirmatory Factor Analysis (CFA) Results					
I feel positively about my supervisor.	.94	.94	.96	.93	.91
I like my supervisor.	.92	.94	.96	.95	.91
I like to work with my supervisor.	.97	.93	.97	.95	.95
I value the relationship I have with my supervisor.	.92	.96	.95	.92	.93
I have been happy with my supervisor.	.97	.96	.97	.95	.95
Chi-Square (χ^2; df = 5)	5.71	23.50	29.56	22.89	27.71
p-value for Chi-Square	.34	<.01	<.01	<.01	<.01
Comparative Fit Index (CFI)	1.00[a]	.99	1.00[a]	1.00[a]	.98
Tucker-Lewis Index (TLI)	1.00[a]	.98	.99	.99	.97
Root Mean Square Error of Approximation (RMSEA)	.02	.11	.08	.08	.15
Standardized Root Mean Square Residual (SRMR)	.00[a]	.01	.00[a]	.01	.01
Descriptive Statistics					
Mean ($\bar{x}$)	5.84	5.44	5.17	5.38	5.26
Standard Deviation (SD)	1.41	1.60	1.70	1.49	1.53
Cronbach's Alpha (α)	.98	.98	.98	.97	.97
Composite Reliability	.98	.98	.98	.97	.97
Average Variance Extracted (AVE)	.89	.90	.92	.88	.87

Note. We report the standardized item loadings for the confirmatory factory analysis (CFA) results.
[a] None of the values were exactly .000 or 1.000, but some results rounded down to .00 or up to 1.00.

Study 3: Initial Factor Structure Assessment of the Positive LAQ via CFAs

Sample and Procedures

We recruited participants from the North American headquarters of a large truck manufacturer to confirm the factor structure of the positive LAQ. Employees were encouraged by their supervisors to participate in the data collection effort. Ultimately, 342 respondents provided complete and usable data. The respondents averaged 45.84 years of age (SD = 11.48) and 80.41% were male. Respondents reported various levels of organizational tenure (e.g., 28.65% reported 1–5 years, 23.98% reported 6–10 years) and education (e.g., 45.32% held bachelor's degrees, 24.85% held master's degrees).

We conducted CFAs in AMOS 23.0 (Arbuckle, 2013) to confirm the factor structure of the positive LAQ. According to Hu and Bentler's (1999) standards, the results demonstrated acceptable fit to the data ($\chi^2[5] = 5.71$, Comparative Fit Index [CFI] = 1.00, Tucker-Lewis Index [TLI] = 1.00, root mean square error of approximation [RMSEA] = .02, standardized root mean residual [SRMR] = .00). Further, the five-item positive LAQ measure demonstrated acceptable levels of convergent validity because each item loaded significantly on the positive LAQ construct, had standardized item loadings greater than .50, and the average variance extracted (AVE) value was above .50 (i.e., AVE = .89; Fornell & Larcker, 1981). The five items also demonstrated acceptable levels of internal consistency, as evidenced by the estimates of Cronbach's alpha (α = .98) and composite reliability (.98) reported in Table 1. Table 1 provides CFA results, fit statistics, and descriptive statistics for the five-item positive LAQ across studies. The results of this initial factor structure assessment supported the use of the five-item positive LAQ.

Study 4: Item Development and Content Validity Assessment of the Negative LAQ

At the suggestion of the action editor and reviewers, we also developed a negative LAQ measure. The negative LAQ was an important addition because positive and negative affect can have independent effects on organizational phenomena that contribute incremental predictive and explanatory power (Tellegen, Watson, & Clark, 1999). We followed the procedures outlined in Study 1 to develop items that corresponded with our conceptualization of negative leader affect (i.e., the degree to which subordinates have negative feelings about their supervisors). We created an initial pool of 52 items. Then, we selected five faculty members and five doctoral-level students from universities in the United States as respondents for a Q-sort activity that was structured and evaluated identically to the Q-sort activity described in Study 1. Ultimately, we retained 15 items that had a weighted mean of at least 6.00 on our seven-point agreement scale.

Study 5: Initial Factor Structure Assessment of the Negative LAQ via EFAs

We compensated 300 respondents from MTurk 50 cents each to complete a survey that included demographic items and the 15 negative LAQ items identified in Study 4. We used the same data collection procedures and inclusion criteria described in Study 2. All of the respondents met the inclusion criteria. The 300 respondents in the final sample averaged 33.61 years of age (SD = 9.80), 5.59 years of tenure with their organizations (SD = 5.35), 3.57 years of tenure with their supervisors (SD = 3.39), and 39.17 hr of work per week at organizations other than MTurk (SD = 7.06). The majority of respondents (61.00%) were male. Respondents represented various levels of education (i.e., 43.33% held bachelor's degrees, 23.00% completed some college), job functions (e.g., 15.00% administration, 14.67% sales), levels in U.S. organizations (e.g., 43.33% staff/associate level, 29.67% middle management), and industries (e.g., 20.67% educational, health, and social services, 13.00% information).

EFA Results

We followed the same EFA procedures described in Study 2. The results are presented in Appendix B. Initially, we chose the five highest loading items for the negative LAQ, but one of the items was problematic (i.e., "I do not like my supervisor") because low liking or the absence of liking does not necessarily equate to disliking (e.g., Daniels & Berkowitz, 1963). Thus, we did not include the problematic item in the final measure. Ultimately, the five items that comprise the final negative LAQ correlated highly with the full 15-item measure, $r = .99$, $p < .01$ and strongly loaded on a single factor in the EFA that accounted for 88.11% of the explained variance. Next, we collected a sample from working professional MBA students to examine the factor structure of the negative LAQ.

Study 6: Initial Factor Structure Assessment of the Negative LAQ via CFAs

We recruited working professional MBA students who attended a private university in the Southeastern United States to complete a survey that included the negative LAQ and demographic items. Ultimately, 243 respondents were at least 18 years old, worked at least 20 hr per week, reported to a direct supervisor, and provided complete and usable data. The respondents averaged 33.58 years of age (SD = 6.67), 4.34 years of tenure with their organizations (SD = 3.58), 2.08 years of tenure with their supervisors (SD = 2.15), and 46.94 hr of work per week at their organizations (SD = 7.10). The majority of respondents (60.91%) were male. Respondents represented various levels of education (i.e., 48.56% held master's degrees or were about to be conferred with a master's degree, 48.15% held bachelor's degrees), job functions (e.g., 20.58% management, 16.87% accounting/finance), levels in U.S. organizations (e.g., 40.74% middle management, 30.45% staff/associate level), and industries (e.g., 22.63% manufacturing, 19.75% finance, insurance, real estate, and rental and leasing).

Next, we conducted CFAs using the same procedures from Study 3. The results demonstrated acceptable fit to the data ($\chi^2[5] = 21.37$, CFI = .99, TLI = .98, RMSEA = .12, SRMR = .02). The five-item negative LAQ measure demonstrated acceptable levels of convergent validity because each item loaded significantly on the negative LAQ construct, had standardized item loadings greater than .50, and the average variance extracted (AVE = .82) value was above .50 (Fornell & Larcker, 1981). The five items also demonstrated acceptable levels of internal consis-

tency, as evidenced by the estimates of Cronbach's alpha (α = .96) and composite reliability (.96) reported in Table 2. Table 2 also provides CFA results, fit statistics, and descriptive statistics for the negative LAQ across studies. Overall, the results of the factor structure assessment in Study 6 support the use of the five-item negative LAQ. Thus, we proceeded to test the convergent and discriminant validity of the LAQs in Study 7.

Study 7: Convergent and Discriminant Validity Study

Sample and Procedures

One dollar each to complete an electronic survey. The 296 respondents in the final sample (i.e., 98.67% usable data) averaged 34.60 years of age (*SD* = 9.94), 5.68 years of tenure with their organizations (*SD* = 4.87), 4.18 years of tenure with their supervisors (*SD* = 3.79), and 39.70 hr of work per week at organizations other than MTurk (*SD* = 6.72). A slight majority (53.72%) of the respondents were male. Respondents represented various levels of education (e.g., 41.89% held bachelor's degrees, 22.30% completed some college), job functions (e.g., 15.20% sales, 12.84% administration), levels in U.S. organizations (e.g., 42.23% staff/associate level, 27.70% middle management), and industries (e.g., 16.55% educational, health, and social services, 13.18% finance, insurance, real estate, and rental and leasing).

Measures

We used a seven-point agreement response format (1 = "Strongly disagree" to 7 = "Strongly agree") for all measures, unless otherwise noted. Higher scores for each measure indicated greater perceptions of each variable.

The LAQs. The five items developed in the previous studies were used to measure the positive LAQ (α = .98) and negative LAQ (α = .96).

Abusive supervision. We used Tepper's (2000) 15-item measure of abusive supervision (α = .96). Responses to the statement "My boss . . ." were recorded on a 5-point frequency scale (1 = "I cannot remember him/her using this behavior with me," 5 = "He/she uses this behavior very often with me"). "Puts me down in front of others" and "Tells me I'm incompetent" were sample scale items.

Authentic leadership. We used Walumbwa, Avolio, Gardner, Wernsing, and Peterson's (2008) 16-item measure of authentic leadership (α = .97). Responses to the statement "My leader:" were recorded on a 5-point frequency scale (1 = "Not at all," 5 = "Frequently, if not always"). "Says exactly what he or she means" and "Encourages everyone to speak their mind" were sample scale items.

LMX. We used Liden and Maslyn's (1998) LMX-12 measure of LMX (α = .96). "My supervisor would come to my defense if I were 'attacked' by others" and "I do work for my supervisor that goes beyond what is specified in my job description" were two of the scale items. In addition to examining the LMX-12 measure in its entirety, we examined the three-item dimension labeled affect (α = .94). We also used Graen and Uhl-Bien's (1995) LMX-7 measure of LMX (α = .95) in order to provide a holistic view of the overlap between the LAQs and widely used measures of LMX. "My leader understands my job problems and needs" and "I would characterize my working relationship with my leader as effective" were two scale items.

Table 2
Summary of CFA Results and Descriptive Statistics for the Negative Leader Affect Questionnaire (Negative LAQ) Across Studies

Items and statistics	Study 6 Initial factor structure (*n* = 243)	Study 7 Convergent and discriminant validity (*n* = 296)	Study 8 Abusive supervision replication (*n* = 806)	Study 9 MLQ Replication (*n* = 507)	Study 10 LMX and ethical leadership replication (*n* = 213)
CFA Results					
I dislike my supervisor.	.89	.89	.92	.92	.90
I have disdain for my supervisor.	.94	.90	.95	.96	.98
I get irritated when I'm around my supervisor.	.88	.91	.91	.92	.90
I resent my supervisor.	.88	.89	.93	.91	.96
I get upset when I think about my supervisor.	.93	.94	.93	.96	.96
Chi-Square (χ^2; *df* = 5)	21.37	17.60	85.72	22.88	30.66
p-value for Chi-Square	<.01	<.01	<.01	<.01	<.01
Comparative Fit Index (CFI)	.99	.99	.98	1.00[a]	.98
Tucker-Lewis Index (TLI)	.98	.98	.97	.99	.97
Root Mean Square Error of Approximation (RMSEA)	.12	.09	.14	.08	.16
Standardized Root Mean Square Residual (SRMR)	.02	.01	.01	.01	.01
Descriptive Statistics					
Mean ($\bar{x}$)	2.19	2.21	2.24	2.18	2.28
Standard Deviation (*SD*)	1.49	1.39	1.45	1.54	1.62
Cronbach's Alpha (α)	.96	.96	.97	.97	.97
Composite Reliability	.96	.96	.97	.97	.97
Average Variance Extracted (AVE)	.82	.82	.85	.87	.88

Note. We report the standardized item loadings for the confirmatory factory analysis (CFA) results.
[a] None of the values were exactly .000 or 1.000, but some results rounded down to .00 or up to 1.00.

Social desirability. We used Fischer and Fick's (1993) Revised Form X1 to measure social desirability ($\alpha = .76$). Four items were negatively worded, so we reverse-scored these items. "I'm always willing to admit it when I make a mistake" and "I have never deliberately said something that hurt someone's feelings" were sample scale items.

Positive affectivity. We used Watson et al.'s (1988) 10-item measure of positive affectivity ($\alpha = .93$). Respondents were asked to report how they generally felt using a 5-point frequency scale (1 = "Very slightly or not at all," 5 = "Extremely"). "Excited" and "Inspired" were sample scale items.

Negative affectivity. We used Watson et al.'s (1988) 10-item measure of negative affectivity ($\alpha = .95$). Respondents were asked to report how they generally felt using a 5-point frequency scale (1 = "Very slightly or not at all," 5 = "Extremely"). "Distressed" and "Upset" were sample scale items.

Measure Assessment

The positive LAQ ($\chi^2[5] = 23.50$, CFI = .99, TLI = .98, RMSEA = .11, SRMR = .01) and negative LAQ ($\chi^2[5] = 17.60$, CFI = .99, TLI = .98, RMSEA = .09, SRMR = .01) demonstrated acceptable fit, as reported in Tables 1 and 2, respectively. Additionally, the LAQs demonstrated acceptable levels of convergent validity because all items significantly and strongly loaded on the LAQ constructs (positive LAQ: AVE = .90; negative LAQ: AVE = .82). Further, high levels of internal consistency were present for both LAQs (positive LAQ: $\alpha = .98$; composite reliability = .98; negative LAQ: $\alpha = .96$; composite reliability = .96).

Table 3 provides the results of two-factor and one-factor CFAs across the studies that included both LAQs (i.e., Studies 7–10). As indicated, a two-factor CFA that included both LAQs demonstrated acceptable fit in Study 7 ($\chi^2[34] = 134.91$, CFI = .98, TLI = .97, RMSEA = .10, SRMR = .03), as well as in Studies 8–10. Although the correlations between the latent factors were high in each study (range: −.77 to −.86), the results of the one-factor solutions demonstrated consistently unacceptable fit. Thus, we proceeded with the two-factor solution.

We assessed discriminant validity by examining the zero-order correlations between the LAQs and the other study variables. The LAQs were strongly correlated, $r = -.80$, $p < .01$, but not so strongly correlated that they did not demonstrate independent relationships with the other study variables. Not surprisingly, the results presented in Table 4 demonstrated that the LAQs were highly correlated with abusive supervision, authentic leadership, LMX-12, the affect dimension of LMX-12, and LMX-7 ($|.64| < r < |.92|$). In addition, all of the correlations between abusive supervision, authentic leadership, LMX-12, the affect dimension of LMX-12, and LMX-7 were high ($|.55| < r < |.93|$), which lends support to the argument that many leadership constructs likely are conceptually and empirically confounded, or at least share some common source of variance (Martinko, Harvey, & Mackey, 2014).

Next, we examined partial correlations between study variables that controlled for the effects of the LAQs. As shown above the diagonal in Table 4, the partial correlations demonstrated that the LAQs generally had strong effects on the obtained correlations among leadership variables. We used 95% confidence intervals (CIs) with bias-corrected and accelerated (BCa) bootstrapping ($n = 5{,}000$) to determine whether there were significant differences between the zero-order correlations and partial correlations. The differences between all of the zero-order correlations and partial correlations for the abusive supervision, authentic leadership, LMX-12, the affect dimension of LMX-12, and LMX-7 were significant because the 95% CIs did not overlap. For example, the zero-order correlations between abusive supervision and the other forms of leadership (authentic leadership, LMX-12, LMX-7) were all strong and significant ($-.62 < r < -.55$), but none of the partial correlations were significant ($-.08 < r < .05$). Additionally, the zero-order correlation between abusive supervision and LMX-12 (i.e., $r = -.61$, 95% CI [−.70, −.51]) significantly differed from the

Table 3
Summary of CFA Results for the Positive and Negative Leader Affect Questionnaires Across Studies

Statistics	Study 7 Convergent and discriminant validity ($n = 296$)	Study 8 Abusive supervision replication ($n = 806$)	Study 9 MLQ Replication ($n = 507$)	Study 10 LMX and ethical leadership replication ($n = 213$)
Two-Factor Solution				
Chi-Square (χ^2; $df = 34$)	134.91	269.41	160.89	120.03
p-value for Chi-Square	< .01	< .01	< .01	< .01
Comparative Fit Index (CFI)	.98	.98	.98	.97
Tucker-Lewis Index (TLI)	.97	.98	.98	.97
Root Mean Square Error of Approximation (RMSEA)	.10	.09	.09	.11
Standardized Root Mean Square Residual (SRMR)	.03	.02	.02	.02
Correlation between Factors	−.83	−.86	−.84	−.77
One-Factor Solution				
Chi-Square (χ^2; $df = 35$)	723.96	2016.66	1412.07	889.15
p-value for Chi-Square	< .01	< .01	< .01	< .01
Comparative Fit Index (CFI)	.84	.85	.82	.74
Tucker-Lewis Index (TLI)	.79	.81	.77	.66
Root Mean Square Error of Approximation (RMSEA)	.26	.27	.28	.34
Standardized Root Mean Square Residual (SRMR)	.08	.07	.07	.12

Note. We report the standardized item loadings for the confirmatory factory analysis (CFA) results.

Table 4
Descriptive Statistics, Zero-Order Correlations, and Partial Correlations for Study 7: Convergent and Discriminant Validity

Variable	1	2	3	4	5	6	7	8	9	10
1. Abusive Supervision	—	**−.04**	**−.03**	**.05**	**−.08**	−.03	**.17****	.33**	—	—
2. Authentic Leadership	−.55**	—	**.50****	**.29****	**.42****	.10	.15*	**.05**	—	—
3. LMX-12	−.61**	.85**	—	**.57****	**.58****	.05	.22**	.02	—	—
4. LMX-12 Affect	−.62**	.81**	.93**	—	**.16****	.01	.08	**.07**	—	—
5. LMX-7	−.61**	.83**	.91**	.84**	—	−.02	.14*	.01	—	—
6. Social Desirability	−.07	.07	.03	.02	.00	—	.12*	−.14*	—	—
7. Positive Affectivity	−.06	.31**	.34**	.28**	.31**	.11	—	−.05	—	—
8. Negative Affectivity	.56**	−.22**	−.27**	−.27**	−.26**	−.16**	−.09	—	—	—
9. Positive LAQ	−.67**	.80**	.91**	.92**	.88**	.01	.28**	−.30**	—	—
10. Negative LAQ	.82**	−.64**	−.73**	−.76**	−.71**	−.07	−.18**	.48**	−.80**	—
Mean ($\bar{x}$)	1.57	3.63	5.17	5.04	5.21	4.20	3.29	1.56	5.44	2.21
Standard Deviation (*SD*)	.81	.98	1.34	1.66	1.39	1.18	.93	.79	1.60	1.39
Cronbach's Alpha (α)	.96	.97	.96	.94	.95	.76	.93	.95	.98	.96

Note. Zero-order correlations are reported below the diagonal (n = 296). Partial correlations while controlling for positive LAQ and negative LAQ are reported above the diagonal. All statistical significance tests were based on two-tailed tests ($\alpha = .05$). Bold partial correlations significantly differed from zero-order correlations (i.e., the 95% confidence intervals for the zero-order and partial correlations did not overlap). LMX = leader-member exchange; LAQ = leader affect questionnaire.
* $p < .05$. ** $p < .01$.

partial correlation for this relationship (i.e., $r = -.03$, 95% CI [−.17, .11]), and the correlations between authentic leadership and the LMX measures all decreased significantly (i.e., zero-order correlations: $.81 < r < .85$; partial correlations: $.29 < r < .50$).

Finally, we conducted regression analyses (Cohen, Cohen, West, & Aiken, 2003) to determine the amount of variance in the various leadership constructs that was explained by the LAQs. Notably, the LAQs explained the majority of the variance in abusive supervision ($R^2 = .67$), authentic leadership ($R^2 = .64$), LMX-12 ($R^2 = .82$), the affect dimension of LMX-12 ($R^2 = .85$), and LMX-7 ($R^2 = .78$).

The results of Study 7 demonstrated that the LAQs had the expected levels of convergent and discriminant validity, and showed that the LAQs were only moderately correlated with social desirability, positive affectivity, and negative affectivity. The results also established that leader affect is highly associated with subordinates' favorable (i.e., authentic leadership, LMX) and unfavorable (i.e., abusive supervision) perceptions of leaders. Thus, we proceeded to examine the criterion-related and predictive validity of the LAQs in Studies 8–10 by replicating prior leadership studies while controlling for the LAQs.

Study 8: Criterion-Related and Predictive Validity

In Study 8, we conducted a constructive replication of Tepper, Carr, Breaux, Geider, Hu, and Hua's (2009) abusive supervision study

Sample and Procedures

We replicated the first study from Tepper et al.'s (2009) examination of the interactive effects of abusive supervision and intention to quit on organization- and supervisor-directed deviance. We compensated 825 respondents from MTurk 50 cents each to complete a survey that included the LAQs and all of the measures reported in Tepper et al.'s first study. The final sample consisted of 806 respondents who provided usable data and met the inclusion criteria and data screening requirements described in Study 2 (i.e., 97.70% usable data).

Similar to Tepper et al.'s (2009) sample, our sample consisted of employees across job functions (e.g., 14.64% administration, 13.77% sales), levels in organizations (e.g., 46.15% staff/associate level, 27.05% middle management), and across industries (e.g., 22.21% educational, health, and social services, 13.28% finance, insurance, real estate, and rental and leasing). Respondents' genders (Tepper et al.: 47% male; current study: 44.29% male), ages (Tepper et al.: $\bar{x} = 2.24$; current study: $\bar{x} = 2.53$; age was coded such that 18–25 = 1, 26–35 = 2, 36–45 = 3, 46–55 = 4, 56–65 = 5, and over 65 = 6), and tenure with supervisor (Tepper et al.: $\bar{x} = 4.12$ years; current study: $\bar{x} = 3.76$ years) were similar across studies. Further, the results in Appendix C demonstrated that the means ($\bar{x}$), standard deviations (*SD*), internal consistency (α), and zero-order correlation estimates for all variables generally were similar across studies. Additionally, we report partial correlations for study variables in Appendix D.

Analyses and Results

We replicated Tepper et al.'s (2009) use of hierarchical moderated multiple regression (Cohen et al., 2003) to test the hypothesized model in SPSS 23.0. The results from Tepper et al.'s Study 1, as well as our replication study with and without the LAQs as control variables, are reported in Appendix E. The results in Appendix E demonstrated that we were able to replicate most of Tepper et al.'s Study 1 findings, except for the interaction effect of abusive supervision and intention to quit on organization- and supervisor-directed deviance. The results also demonstrated that the positive LAQ did not significantly predict organization- ($B = .04$, *ns*) or supervisor-directed deviance ($B = .00$, *ns*) in the final step of the regression equation. Further, the results demonstrated that the negative LAQ significantly predicted supervisor-directed deviance ($B = .11$, $p < .01$), but not organization-directed deviance ($B = .01$, *ns*).

Partial correlations. The LAQs demonstrated moderate-to-high (i.e., $|.32| < r < |.70|$) zero-order correlations with the nondemographic variables in our study. Also, the positive LAQ, $r = -.56$, $p < .01$ and negative LAQ, $r = .70$, $p < .01$ were both strongly associated with abusive supervision. We examined partial correlations between the focal variables while controlling for the LAQs (see Appendix D). Notably, several of the zero-order correlations and partial correlations significantly differed because the 95% CIs did not overlap. For example, the zero-order correlation between abusive supervision and supervisor-directed deviance (i.e., $r = .59$, 95% CI [.50, .66]) significantly differed from the partial correlation for this relationship (i.e., $r = .33$, 95% CI [.21, .44]). Additional supplementary analyses are available in Appendixes F-I.

Study 9: Criterion-Related and Predictive Validity.

In Study 9, we conducted a constructive replication of Epitropaki and Martin's (2005) transformational and transactional leadership study

Sample and Procedures

We replicated Epitropaki and Martin's (2005) examination of the interactive effects of leadership perceptions (i.e., transformational leadership and transactional leadership) and individual difference variables (i.e., separateness-connectedness self-schema, positive affectivity, and negative affectivity) on organizational identification. Epitropaki and Martin collected cross-sectional data from 502 employees who worked at banks in Greece. We compensated 525 respondents from MTurk 50 cents each to complete a survey that included the LAQs and all of the measures reported in Epitropaki and Martin's study. We received complete and usable data from 507 respondents who met the inclusion criteria described in the previous studies (i.e., 96.57% usable data).

Similar to Epitropaki and Martin's (2005) sample, ours was comprised of relatively equal numbers of males and females (Epitropaki and Martin: 50.4% males; current study: 46.94% males). Respondents' ages (Epitropaki and Martin: $\bar{x} = 34.76$ years; current study: $\bar{x} = 35.00$ males) and organizational tenure (Epitropaki and Martin: $\bar{x} = 6.02$ years; current study: $\bar{x} = 5.41$ years) were also similar, as were respondents' roles in managerial/supervisory jobs (Epitropaki and Martin: 25.4%; current study: 29.19%) and clerical/administrative jobs (Epitropaki and Martin: 61.2%; current study: 61.93%). The results in Appendix J demonstrate that the means ($\bar{x}$), standard deviations (*SD*), internal consistency (α), and zero-order correlation estimates for all variables were similar across studies as well.

Analyses and Results

We conducted moderated hierarchical regression analyses (Cohen et al., 2003) in SPSS 23.0 to replicate the results obtained in Epitropaki and Martin's (2005) study. The results from Epitropaki and Martin's study, as well as our replication study with and without the LAQs as control variables, are reported in Appendix K. The results in Appendix K demonstrated that we were able to replicate most of Epitropaki and Martin's findings, except for the main effect of transactional leadership on organizational identification. Further, our results demonstrated that the positive LAQ was a significant predictor of organizational identification in the final step of the regression equation ($B = .16$, $p < .01$), whereas the negative LAQ was not ($B = .02$, *ns*).

Partial correlations. The LAQs were both strongly associated with transformational leadership (positive LAQ: $r = .77$, $p < .01$; negative LAQ: $r = -.65$, $p < .01$), but were weakly associated with transactional leadership (positive LAQ: $r = .09$, $p < .05$; negative LAQ: $r = .08$, *ns*). Also, we examined partial correlations between the focal variables while controlling for the LAQs (see Appendix J). Importantly, the zero-order correlation between transformational leadership and organizational identification ($r = .68$, 95% CI [.62, .74]) significantly differed from the partial correlation for this relationship ($r = .40$, 95% CI [.30, .49]). Additional supplementary analyses are shown in Appendixes L-N.

The findings reported above are not surprising because the dimensions of transformational leadership (i.e., idealized influence attributes, idealized influence behaviors, inspirational motivation, intellectual stimulation, and individual consideration) likely are more susceptible to influence from subordinates' affective evaluations than the dimensions of transactional leadership (i.e., contingent reward, active management-by-exception, and passive management-by-exception). Thus, the results make conceptual sense given that transformational leadership assesses followers' perceptions of leadership behaviors meant to develop their followers, whereas transactional leadership assesses followers' perceptions of supervisory interactions with specific exchange purposes (Bass & Avolio, 1995).

Study 10: Criterion-Related and Predictive Validity.

In Study 10, we conducted a constructive replication of Mahsud, Yukl, and Prussia's (2010) ethical leadership and LMX study

Sample and Procedures

We replicated Mahsud et al.'s (2010) study of the impact of leader empathy, ethical leadership, and relation-oriented behavior on LMX. Mahsud et al. collected responses from 218 business students who were employed and reported to a supervisor. We compensated 225 respondents from MTurk 50 cents each to complete a survey that included the LAQs and all of the measures reported in Mahsud et al.'s study. The final sample consisted of 213 individuals who met the inclusion criteria described in the previous studies (i.e., 94.67% usable data).

Respondents' average age was 34.16 years (Mahsud et al., 2010: $\bar{x} = 38$ years) and their average organizational tenure was 6.18 years (Mahsud et al.: average reported as two to three years). All respondents in both samples had direct managers to whom they reported, and many were also managers themselves (Mahsud et al.: 40%; current study: 35.68%). The sample populations differed regarding the gender of respondents (Mahsud et al.: 43% male; current study: 56.81% male) and the gender of respondents' supervisors (Mahsud et al.: 36% male; current study: 61.03% male). The results in Appendix O demonstrated that the means ($\bar{x}$), standard deviations (*SD*), internal consistency (α), and zero-order correlation (*r*) estimates generally were similar across the two studies.

Analyses and Results

We used SPSS 23.0 to conduct hierarchical multiple regression analyses (Cohen et al., 2003) in order to replicate Mahsud et al.'s (2010) results. The results from Mahsud et al.'s study, as well as our replication study with and without the LAQs as control variables, are reported in Appendix P. Overall, we were able to replicate most of Mahsud et al.'s findings. Our results also demonstrated that the positive LAQ was a significant predictor of LMX in the final step of the regression equation ($B = .15$, $p < .01$), whereas the negative LAQ was not ($B = -.01$, *ns*).

Partial correlations. An examination of the zero-order correlations (see Appendix O) indicated that the positive LAQ and the negative LAQ were both strongly associated with ethical leadership (positive LAQ: $r = .81$, $p < .01$; negative LAQ: $r = -.63$, $p < .01$) and LMX (positive LAQ: $r = .85$, $p < .01$; negative LAQ: $r = -.65$, $p < .01$). Importantly, the zero-order correlations were significantly different from the partial correlations for all of the relationships examined. For example, the zero-order correlation between ethical leadership and LMX ($r = .78$, 95% CI [.71, .84]) significantly differed from the partial correlation for this relationship ($r = .31$, 95% CI [.17, .43]). Additional supplementary analyses are available in Appendix Q.

Discussion

Our findings support the conclusions that a) the LAQs provide highly reliable and valid measures for assessing subordinates' evaluations of their leaders; b) there is significant overlap between leadership measures, and much of this overlap appears to be a function of the affect captured by the LAQs; c) when the LAQs are used as control variables, in most cases, it reduces the strength of the relationships between leadership measures and outcome variables and, in some instances, renders these relationships nonsignificant; d) the LAQs account for small but significant increases in the variance of outcomes beyond that explained by other leadership measures; e) there is still a considerable amount of unexplained variance between leadership measures that the LAQs do not capture; and f) there is a considerable amount of variance in outcome measures that is not explained by subordinates' ratings of their leaders.

Although our results are clear, the implications are less obvious and raise a number of intriguing issues. First and foremost is the role of affect in subordinates' evaluations of their leaders. As discussed earlier, emotions often precede cognitions and initiate cognitive processes (Lerner at al., 2015). Because our research designs were cross-sectional, we were unable to directly test where and when affect enters into subordinates' rating processes. Nonetheless, the strength of our results is consistent with our theoretical framework, which suggests that affect cues an overall judgment based on subordinates' cognitive schemata regarding their supervisors and their ILTs. For the majority of respondents, this results in ratings of their leaders that are consistent with their ILTs as opposed to assessing the specific leadership behaviors. Research is needed to more clearly understand, identify, and verify when and how affect enters into subordinates' evaluations of their leaders. Longitudinal and qualitative studies would be helpful.

It would also be informative to test the effects of affect in situations perceived to be more versus less important. This could be done in laboratory situations by manipulating the purpose and effects of the rating processes. If the results are consistent with our theorizing, affect would have more impact in routine and relatively unimportant evaluations versus evaluations with serious consequences where we would expect more deliberate cognitive processing. It would also be interesting to contrast subordinates' evaluations in new versus mature leader-member dyads. Affective events and the development of cognitive schemata occur over time, so we expect the affective reactions measured by the LAQs to influence evaluations more in mature than less mature subordinate-supervisor dyads. Finally, it might prove worthwhile to employ neurophysiological methodologies to map brain activity. Identifying which parts of the brain are activated at the beginning and during the completion of the questionnaires would be informative and demonstrate whether and to what extent emotional centers of the brain are activated.

A second major issue concerns the high correlations between and among leadership measures that suggests that all of these measures share common variance. Given the high face validity and rigorous process we used to develop the LAQs, it is reasonable to attribute this overlap to affect. However, we recognize the possibility that the common factor underlying the integrity of the LAQs and other leadership measures could be something other than affect. Additional research examining the factor structure of leadership questionnaires is needed. In particular, an EFA that simultaneously evaluates the factor structures for multiple questionnaires would be informative. Some minor rewording of scales and items might be necessary to conduct these analyses, but it would be worth the effort to assess the factor structure and content validity of existing leadership measures. Likewise, it may be possible to conduct a CFA similar to the Harman one-factor test for common method variance (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003), but assess affect as the latent factor accounting for the shared variance between measures (see Martinko et al., 2014).

Our findings regarding the considerable overlap among leadership measures also raise serious questions about the differences between leadership theories and measures. Although the theoretical foundations for authentic leadership, LMX, ethical leadership, and transformational leadership appear to be very different, our results demonstrate that a large percentage of the variance measured by the scales that represent these leadership constructs is the same and appears to be related to affect. Thus, future research would benefit from using the LAQs as control variables. Assuming that the objective of research designs is to explain variance in the criterion variables, using the LAQs as control variables will enable researchers to isolate and more clearly identify and explain the remaining variance in traditional leadership questionnaires. Thus, a major focus of future research should be on explaining the proportion of variance between leadership measures and outcome variables that is not explained by the LAQs. It is possible that the majority of this variance may be error variance, but it is also possible that other constructs such as perceived competence, intelligence, and motivation to lead may explain additional variance and offer new insights into the factors that subordinates consider when evaluating their leaders. Thus, we may be able to develop a more complete understanding of subordinates' rating processes and the distinctions between various theories and their measures.

Another major issue is the use and utility of the LAQs. Our findings demonstrate that the LAQs are highly reliable and that the predictive validity of the LAQs is comparable to other leadership

measures. Thus, a case can be made that the affect measured by the LAQs deserves attention as a leadership construct. Unlike Hansbrough et al. (2015, p. 226), who concluded that "we anticipate that follower liking will be associated with inflated ratings of leader behavior", the evidence we developed supports the position that affect is an integral part of the leader rating process and deserves consideration in its own right. Although we suggest that the LAQs can be used as important control variables, much like our use of the LAQs as control variables in Studies 8–10, we also argue that the LAQs can be important focal constructs to include in future research. As our theoretical perspective indicates, we view affect as an important part of the evaluative process that cues the initial reaction, which, in turn, cues the recall of causal schemata and the subordinates' ILTs that shape their evaluative responses. Thus, we view affect as a major factor that both initiates and continues to influence the evaluative process.

The term "bias" implies that affect is a factor that, if removed, would result in "objective" evaluations. We view affect as an important endogenous factor that is part of the evaluative process, which is influenced by multiple factors. Thus, instead of viewing affect only as a nuisance variable that needs to be controlled, our results suggest that more research needs to be focused on affect as a substantive variable and its relationship with leader-member relations, subordinate evaluations of leadership behaviors, and as a predictor of individual and organizational outcomes. Moreover, we encourage additional research to explore differences in the development and effects of positive and negative leader affect.

Another issue suggested by the above discussion is whether or not the LAQs are viable substitutes and/or proxies for the other leadership questionnaires. As noted earlier, the LAQ ratings are highly correlated with transformational leadership, authentic leadership, ethical leadership, abusive supervision, and LMX. Also, the predictive validity of the LAQs is comparable to that of the other leadership measures. Due to the high correlations between the LAQs and the other leadership scales, it is likely that subordinates who express high positive and low negative affect toward their leaders will evaluate their leader-member relations positively and also rate their leaders as highly authentic, transformational, and ethical, as well as less abusive. It is also likely that the outcomes of the LAQ ratings will be similar to those of the other leadership measures. Thus, the case can be made that the LAQs can serve as a viable substitute for the other measures. Moreover, from a practical standpoint, managers interested in how their employees view their supervisors could administer the five-item LAQs, which are highly correlated with all of the other leadership measures, instead of multiple and lengthier assessments of leadership.

Our findings may also have implications beyond helping to explain how subordinates evaluate their leaders. More specifically, we believe it would be worthwhile to investigate the role of affect in other important evaluation processes in organizations, such as performance evaluations and hiring decisions, but also decisions made outside of organizational contexts such as personal investments and voting. As discussed earlier, Viswesvaran et al. (2005) identified a single factor that accounted for almost 60% of the variance in performance evaluations. It may be more than coincidental that the 60% of the variance due to the common factor in BARS ratings is similar to the magnitude of the variance we typically found for affect in subordinates' leadership evaluations. The broader literature on decision making (e.g., Lerner et al., 2015) supports the notion that affect plays a critical role in many of the important decisions both within and outside of organizational contexts. Thus, more research into the role affect plays in decision processes within and outside of organizational contexts appears warranted.

Some thought should also be given to the nature and purpose of subordinates' leadership evaluations. Because there appears to be limited discriminant validity among the various forms of leadership questionnaires, inferences regarding the effects of leadership styles and behaviors on outcomes appear uncertain. The links between the leadership styles, behaviors, and organizational outcomes inferred by typical leadership questionnaire studies are also tentative because the majority of the studies linking subordinate leader evaluations with organizational outcomes are cross-sectional, which precludes inferences of causation. In addition, the relationships between subordinates' leadership evaluations and organizational outcomes warrant additional attention. If, as suggested above, affect also plays a role in the evaluations of outcome variables, there is the possibility that affect is the underlying factor in the correlations between subordinates' evaluations of their leaders and subordinates' ratings of outcomes such as satisfaction and stress.

Studies assessing subordinates' affect in relation to both their evaluations of leadership and other outcome variables may help deconstruct and explain the relationship between subordinates' ratings of their leaders and organizational outcomes. In addition, it must be recognized that because the majority of the variance in the outcome variables is typically attributable to factors other than subordinate evaluations, there may be factors that are much more important in evaluating leader effectiveness than subordinates' evaluations. Additionally, it will be important for future research to determine the role of leader affect as a boundary condition that can attenuate or enhance the effects of leadership on followers' outcomes. Thus, we encourage future research to examine the interactive effects of various forms of leadership and the LAQs on employees' outcomes.

A potential limitation of our study is the inconsistent pattern of RMSEA values we obtained for some of the positive LAQ, negative LAQ, and two-factor CFAs across studies. Prior research emphasizes the importance of examining fit indices in tandem rather than isolation (Hu & Bentler, 1999), and has documented the difficulty of obtaining acceptable RMSEA values (Kenny, Kaniskan, & McCoach, 2015). In particular, Hu and Bentler reported that the RMSEA statistic tends to overreject models for analyses run with small sample sizes (i.e., $n \leq 250$), whereas the findings for models with larger samples (e.g., $n = 500$, $n = 1{,}000$) tend to be more stable. Kenny et al. built on Hu and Bentler's work by running Monte Carlo simulations, which demonstrated that the low power associated with RMSEA computations for models with low degrees of freedom and/or small sample sizes was problematic. Kenny et al. concluded by stating "that it is advisable for researchers to completely avoid computing the RMSEA when model *df* are small" (p. 503). Thus, the extent to which the inconsistent RMSEA values in our studies indicate poor fit is not clear because many of our sample sizes were below or around the thresholds Hu and Bentler identified (e.g., Study 3: $n = 342$; Study 6: $n = 243$; Study 7: $n = 296$), and our CFA models had low degrees of freedom (range: 5 to 35).

Another potential limitation of our study is that our scale is similar to the scales used by Engle and Lord (1997) and Turban, Jones, and Rozelle (1990) to measure liking. However, our scale extends the contributions of these earlier scales by extensively developing and validating the LAQs. Thus, we encourage future research to extend our findings using the LAQs as control and substantive variables in studies that examine leadership.

Conclusion

The objective of this study was to explore the premise that a large proportion of the variance in subordinates' responses to leadership questionnaires can be attributed to the affect that subordinates feel about their supervisors. Thus, we developed and validated the positive and negative LAQs. Our research demonstrated that the LAQs are highly reliable and have both high construct and predictive validity. When we compared the LAQs to traditional leadership measures, we found significant overlap between all of the measures. When the LAQ measures were used as control variables to examine partial correlations, we found that much of the overlap between leadership measures appears to be attributable to the affect that subordinates feel toward their leaders. We also found that when the LAQ measures were used as control variables, they reduced the correlations between leadership measures and outcome measures, and sometimes accounted for variance in outcomes beyond that attributable to other leadership questionnaires.

Our recommendations for research and theory building include: 1) more fully exploring the overlap among leadership measures using the LAQs as control variables; 2) developing studies to explore where and when affect enters into subordinates' evaluations processes; 3) exploring how affect between leaders and subordinates develops over time; 4) further exploration of the predictive validity of the LAQ measures; 5) investigations of the variance in leadership measures not explained by the LAQs; 6) a reconsideration of the utility of subordinates' evaluations of leadership; 7) a more thorough examination of the similarities and differences between and among the various theories of leadership; and 8) exploring the role of affect in other evaluative situations such as selection, job assignment, voting, and investment decisions.

The major contribution of our research is the insight we offer into the foundational assumptions underlying traditional methods of evaluating leadership via subordinate evaluations. Our research suggests that affect is much more than a nuisance variable and plays a critical role in cuing subordinates' evaluations of their leaders, as well as in understanding the relationships between and among the various measures and theories of leadership. Although we have offered numerous suggestions for research, it is ultimately the responsibility of the community of leadership researchers to choose and explore the paths that most clearly illuminate and explicate the process and meaning of subordinates' leadership evaluations. We hope that our findings and the discussion of those findings are constructive and make meaningful contributions that facilitate a more informed understanding of the role that affect plays in subordinates' evaluations of their leaders.

References

Aguinis, H., & Gottfredson, R. K. (2010). Best-practice recommendations for estimating interaction effects using moderated multiple regression. *Journal of Organizational Behavior, 31,* 776–786. http://dx.doi.org/10.1002/job.686

Arbuckle, J. L. (2013). *IBM SPSS AMOS 22 user's guide*. Chicago, IL: SPSS, Inc.

Atkin, R. S., & Conlon, E. J. (1978). Behaviorally anchored rating scales: Some theoretical issues. *Academy of Management Review, 3,* 119–128.

Bass, B. M., & Avolio, B. J. (1995). *MLQ multifactor leadership questionnaire*. Redwood City, CA: Mind Garden.

Baumeister, R. F., & Bushman, B. J. (2008). *Social psychology and human nature*. Belmont, CA: Thomson Wadsworth.

Bechara, A., Damasio, H., Damasio, A. R., & Lee, G. P. (1999). Different contributions of the human amygdala and ventromedial prefrontal cortex to decision-making. *The Journal of Neuroscience, 19,* 5473–5481.

Bernerth, J. B., Armenakis, A. A., Feild, H. S., Giles, W. F., & Walker, H. J. (2007). Leader-member social exchange (LMSX): Development and validation of a scale. *Journal of Organizational Behavior, 28,* 979–1003. http://dx.doi.org/10.1002/job.443

Binning, J. F., Zaba, A. J., & Whattam, J. C. (1986). Explaining the biasing effects of performance cues in terms of cognitive categorization. *Academy of Management Journal, 29,* 521–535. http://dx.doi.org/10.2307/256222

Brown, D. J., & Keeping, L. M. (2005). Elaborating the construct of transformational leadership: The role of affect. *The Leadership Quarterly, 16,* 245–272. http://dx.doi.org/10.1016/j.leaqua.2005.01.003

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science, 6,* 3–5. http://dx.doi.org/10.1177/1745691610393980

Chaiken, S., & Trope, Y. (1999). *Dual-process theories in social psychology*. New York, NY: Guilford Press.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Hillsdale, NJ: Erlbaum.

Colquitt, J. A., Baer, M. D., Long, D. M., & Halvorsen-Ganepola, M. D. K. (2014). Scale indicators of social exchange relationships: A comparison of relative content validity. *Journal of Applied Psychology, 99,* 599–618. http://dx.doi.org/10.1037/a0036374

Cronshaw, S. T., & Lord, R. G. (1987). Effects of categorization, attribution, and encoding processes on leadership perceptions. *Journal of Applied Psychology, 72,* 97–106. http://dx.doi.org/10.1037/0021-9010.72.1.97

Damasio, A. R. (1994). *Descartes' error: Emotion, reason, and the human brain*. New York, NY: Putnam.

Daniels, L. R., & Berkowitz, L. (1963). Liking and response to dependency relationships. *Human Relations, 16,* 141–148. http://dx.doi.org/10.1177/001872676301600203

DeSimone, J. A., Harms, P. D., & Desimone, A. J. (2015). Best practice recommendations for data screening. *Journal of Organizational Behavior, 36,* 171–181. http://dx.doi.org/10.1002/job.1962

Eden, D., & Leviatan, U. (1975). Implicit leadership theory as a determinant of the factor structure underlying supervisory behavior scales. *Journal of Applied Psychology, 60,* 736–741. http://dx.doi.org/10.1037/0021-9010.60.6.736

Engle, E. M., & Lord, R. G. (1997). Implicit theories, self-schemas, and leader-member exchange. *Academy of Management Journal, 40,* 988–1010. http://dx.doi.org/10.2307/256956

Epitropaki, O., & Martin, R. (2005). The moderating role of individual differences in the relation between transformational/transactional leadership perceptions and organizational identification. *The Leadership Quarterly, 16,* 569–589. http://dx.doi.org/10.1016/j.leaqua.2005.06.005

Fischer, D. G., & Fick, C. (1993). Measuring social desirability: Short forms of the Marlowe-Crowne social desirability scale. *Educational and Psychological Measurement, 53,* 417–424. http://dx.doi.org/10.1177/0013164493053002011

Fiske, S. T., & Taylor, S. E. (1991). *Social cognition* (2nd ed.). New York, NY: McGraw-Hill Book Company.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18,* 39–50. http://dx.doi.org/10.2307/3151312

Graen, G. B., & Uhl-Bien, M. (1995). Relationship-based approach to leadership: Development of leader-member exchange (LMX) theory of leadership over 25 years: Applying a multi-level multi-domain perspective. *The Leadership Quarterly, 6,* 219–247. http://dx.doi.org/10.1016/1048-9843(95)90036-5

Hansbrough, T. K., Lord, R. G., & Schyns, B. (2015). Reconsidering the accuracy of follower leadership ratings. *The Leadership Quarterly, 26,* 220–237. http://dx.doi.org/10.1016/j.leaqua.2014.11.006

Hardy, B., & Ford, L. R. (2014). It's not me, it's you: Miscomprehension in surveys. *Organizational Research Methods, 17,* 138–162. http://dx.doi.org/10.1177/1094428113520185

Hinkin, T. R. (1995). A review of scale development practices in the study of organizations. *Journal of Management, 21,* 967–988. http://dx.doi.org/10.1177/014920639502100509

Hinkin, T. R. (1998). A brief tutorial on the development of measures for use in survey questionnaires. *Organizational Research Methods, 1,* 104–121. http://dx.doi.org/10.1177/109442819800100106

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6,* 1–55. http://dx.doi.org/10.1080/10705519909540118

Hunter, S. T., Bedell-Avers, K. E., & Mumford, M. D. (2007). The typical leadership study: Assumptions, implications, and potential remedies. *The Leadership Quarterly, 18,* 435–446. http://dx.doi.org/10.1016/j.leaqua.2007.07.001

Johnson, J. W. (2000). A heuristic method for estimating the relative weight of predictor variables in multiple regression. *Multivariate Behavioral Research, 35,* 1–19. http://dx.doi.org/10.1207/S15327906MBR3501_1

Kenny, D. A., Kaniskan, B., & McCoach, D. B. (2015). The performance of RMSEA in models with small degrees of freedom. *Sociological Methods & Research, 44,* 486–507. http://dx.doi.org/10.1177/0049124114543236

Kingstrom, P. O., & Bass, A. R. (1981). A critical analysis of studies comparing behaviorally anchored rating scales (BARS) and other rating formats. *Personnel Psychology, 34,* 263–289. http://dx.doi.org/10.1111/j.1744-6570.1981.tb00942.x

Landy, F. J., & Farr, J. L. (1980). Performance rating. *Psychological Bulletin, 87,* 72–107. http://dx.doi.org/10.1037/0033-2909.87.1.72

Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal, and coping*. New York, NY: Springer.

Lerner, J. S., Li, Y., Valdesolo, P., & Kassam, K. S. (2015). Emotion and decision making. *Annual Review of Psychology, 66,* 799–823. http://dx.doi.org/10.1146/annurev-psych-010213-115043

Lewter, J., & Lord, R. G. (1992, August). *Affect, self-schemas, and transformational leadership*. Paper presented at the Academy of Management Convention, Las Vegas, NV.

Liden, R. C., & Maslyn, J. M. (1998). Multidimensionality of leader-member exchange: An empirical assessment through scale development. *Journal of Management, 24,* 43–72.

Lord, R. G., Binning, J. F., Rush, M. C., & Thomas, J. C. (1978). The effect of performance cues and leader behavior on questionnaire ratings of leadership behavior. *Organizational Behavior & Human Performance, 21,* 27–39. http://dx.doi.org/10.1016/0030-5073(78)90036-3

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. http://dx.doi.org/10.1016/S1048-9843(01)00081-9

Lord, R. G., Foti, R. J., & DeVader, C. L. (1984). A test of leadership categorization theory: Internal structure, information processing, and leadership perceptions. *Organizational Behavior & Human Performance, 34,* 343–378. http://dx.doi.org/10.1016/0030-5073(84)90043-6

Lord, R. G., & Shondrick, S. J. (2011). Leadership and knowledge: Symbolic, connectionist, and embodied perspectives. *The Leadership Quarterly, 22,* 207–222. http://dx.doi.org/10.1016/j.leaqua.2010.12.016

MacKenzie, S. B., Podsakoff, P. M., & Podsakoff, N. P. (2011). Construct measurement and validation procedures in MIS and behavioral research: Integrating new and existing techniques. *Management Information Systems Quarterly, 35,* 293–334. http://dx.doi.org/10.2307/23044045

Mahsud, R., Yukl, G., & Prussia, G. (2010). Leader empathy, ethical leadership and relations-oriented behaviors as antecedents of leader-member exchange quality. *Journal of Managerial Psychology, 25,* 561–577. http://dx.doi.org/10.1108/02683941011056932

Martinko, M. J., Harvey, P., & Mackey, J. D. (2014). Conceptual and empirical confounds in the organizational sciences: An explication and discussion. *Journal of Organizational Behavior, 35,* 1052–1063. http://dx.doi.org/10.1002/job.1961

McElroy, J. C., & Downey, H. K. (1982). Observation in organizational research: Panacea to the performance-attribution effect? *Academy of Management Journal, 25,* 822–835. http://dx.doi.org/10.2307/256101

Meade, A. W., & Craig, S. B. (2012). Identifying careless responses in survey data. *Psychological Methods, 17,* 437–455. http://dx.doi.org/10.1037/a0028085

Murphy, K. R., & Constans, J. I. (1987). Behavioral anchors as a source of bias in rating. *Journal of Applied Psychology, 72,* 573–577. http://dx.doi.org/10.1037/0021-9010.72.4.573

Naidoo, L. J., Kohari, N. E., Lord, R. G., & DuBois, D. A. (2010). "Seeing" is retrieving: Recovering emotional content in leadership ratings through visualization. *The Leadership Quarterly, 21,* 886–900. http://dx.doi.org/10.1016/j.leaqua.2010.07.014

Nathan, B. R., & Alexander, R. A. (1985). The role of inferential accuracy in performance rating. *The Academy of Management Review, 10,* 109–115.

Phillips, J. S., & Lord, R. G. (1982). Schematic information processing and perceptions of leadership in problem-solving groups. *Journal of Applied Psychology, 67,* 486–492. http://dx.doi.org/10.1037/0021-9010.67.4.486

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, 88,* 879–903. http://dx.doi.org/10.1037/0021-9010.88.5.879

Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P. (2016). Recommendations for creating better concept definitions in the organizational, behavioral, and social sciences. *Organizational Research Methods, 19,* 159–203. http://dx.doi.org/10.1177/1094428115624965

Preacher, K. J., & MacCallum, R. C. (2003). Repairing Tom Swift's electric factor analysis machine. *Understanding Statistics, 2,* 13–43. http://dx.doi.org/10.1207/S15328031US0201_02

Rush, M. C., Thomas, J. C., & Lord, R. G. (1977). Implicit leadership theory: A potential threat to the internal validity of leader behavior questionnaires. *Organizational Behavior & Human Performance, 20,* 93–110. http://dx.doi.org/10.1016/0030-5073(77)90046-0

Schwab, D. P., Heneman, H. G., III, & DeCotis, T. A. (1975). Behaviorally anchored rating scales: A review of the literature. *Personnel Psychology, 28,* 549–562. http://dx.doi.org/10.1111/j.1744-6570.1975.tb01392.x

Shondrick, S. J., Dinh, J. E., & Lord, R. G. (2010). Developments in implicit leadership theory and cognitive science: Applications to improving measurement and understanding alternatives to hierarchical

leadership. *The Leadership Quarterly, 21,* 959–978. http://dx.doi.org/10.1016/j.leaqua.2010.10.004

Simon, H. A. (1967). Motivational and emotional controls of cognition. *Psychological Review, 74,* 29–39. http://dx.doi.org/10.1037/h0024127

Smith, P. C., & Kendall, L. M. (1963). Retranslation of expectations: An approach to the construction of unambiguous anchors for rating scales. *Journal of Applied Psychology, 47,* 149–155. http://dx.doi.org/10.1037/h0047060

Srull, T. K., & Wyer, R. S. (1988). *Advances in social cognition: A dual process model of impression formation* (Vol. 1). New York, NY: Erlbaum, Inc.

Srull, T. K., & Wyer, R. S., Jr. (1989). Person memory and judgment. *Psychological Review, 96,* 58–83. http://dx.doi.org/10.1037/0033-295X.96.1.58

Tellegen, A., Watson, D., & Clark, L. A. (1999). On the dimensional and hierarchical structure of affect. *Psychological Science, 10,* 297–303. http://dx.doi.org/10.1111/1467-9280.00157

Tepper, B. J. (2000). Consequences of abusive supervision. *Academy of Management Journal, 43,* 178–190. http://dx.doi.org/10.2307/1556375

Tepper, B. J., Carr, J. C., Breaux, D. M., Geider, S., Hu, C., & Hua, W. (2009). Abusive supervision, intentions to quit, and employees' workplace deviance: A power/dependence analysis. *Organizational Behavior and Human Decision Processes, 109,* 156–167. http://dx.doi.org/10.1016/j.obhdp.2009.03.004

Tonidandel, S., & LeBreton, J. M. (2015). RWA Web: A free, comprehensive web-based, and user-friendly tool for relative weight analyses. *Journal of Business and Psychology, 30,* 207–216. http://dx.doi.org/10.1007/s10869-014-9351-z

Turban, D. B., Jones, A. P., & Rozelle, R. M. (1990). Influences of supervisor liking of a subordinate and the reward context on the treatment and evaluation of that subordinate. *Motivation and Emotion, 14,* 215–233. http://dx.doi.org/10.1007/BF00995570

Viswesvaran, C., Schmidt, F. L., & Ones, D. S. (2005). Is there a general factor in ratings of job performance? A meta-analytic framework for disentangling substantive and error influences. *Journal of Applied Psychology, 90,* 108–131. http://dx.doi.org/10.1037/0021-9010.90.1.108

Walumbwa, F. O., Avolio, B. J., Gardner, W. L., Wernsing, T. S., & Peterson, S. J. (2008). Authentic leadership: Development and validation of a theory-based measure. *Journal of Management, 34,* 89–126. http://dx.doi.org/10.1177/0149206307308913

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54,* 1063–1070. http://dx.doi.org/10.1037/0022-3514.54.6.1063

Wayne, S. J., & Ferris, G. R. (1990). Influence tactics, affect, and exchange quality in supervisor-subordinate interactions: A laboratory experiment and field study. *Journal of Applied Psychology, 75,* 487–499. http://dx.doi.org/10.1037/0021-9010.75.5.487

Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review, 92,* 548–573. http://dx.doi.org/10.1037/0033-295X.92.4.548

Weiner, B. (1986). *An attributional theory of motivation and emotion*. New York, NY: Springer. http://dx.doi.org/10.1007/978-1-4612-4948-1

Weiss, H., & Adler, S. (1981). Cognitive complexity and the structure of implicit leadership theories. *Journal of Applied Psychology, 66,* 69–78. http://dx.doi.org/10.1037/0021-9010.66.1.69

Weiss, H. M., & Cropanzano, R. (1996). Affective events theory: A theoretical discussion of the structure, causes and consequences of affective experiences at work. In B. M. Staw & L. L. Cummings (Eds.), *Research in organizational behavior* (Vol. 18, pp. 1–74). Greenwich, CT: JAI Press Inc.

Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist, 35,* 151–175. http://dx.doi.org/10.1037/0003-066X.35.2.151

Received October 10, 2016
Revision received January 18, 2018
Accepted January 23, 2018 ■