Journal of Applied Psychology

Does What Others Can(not) See Really Matter? The Relationship Between Leadership Arena–Reputation–Identity (LARI) Model and Leader Effectiveness

Andrew C. Loignon, John W. Fleenor, Stephen Jeong, and David J. Woehr

Online First Publication, September 26, 2024. https://dx.doi.org/10.1037/apl0001238

AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Applied Psychology

ISSN: 0021-9010

https://doi.org/10.1037/apl0001238

Does What Others Can(not) See Really Matter? The Relationship Between Leadership Arena–Reputation–Identity (LARI) Model and Leader Effectiveness

Andrew C. Loignon[1], John W. Fleenor[1], Stephen Jeong[1], and David J. Woehr[2]
[1] Center for Creative Leadership, Greensboro, North Carolina, United States
[2] Department of Management, University of North Carolina at Charlotte

Leadership scholars recognize that there is value in capturing how leaders view themselves and how they are viewed by others. Recently, the leadership Arena–Reputation–Identity (LARI) model has been advanced as a means of more precisely capturing the shared and unique perspectives that underlie multisource ratings of leadership. Despite its strengths, several critical questions pertaining to this model remain unanswered: (1) Does the wealth of information provided by the LARI model have any bearing on the effectiveness of a leader? (2) Does the amount of variance explained by a particular source within the LARI model depend on the observability of the leadership dimension being rated? (3) Does the LARI model generalize to the upper echelon of the firms (i.e., senior executives) while also accommodating additional source effects (i.e., board members)? Drawing on multisource ratings of 491 senior executives' leadership competencies, as well as a team-based assessment of their effectiveness, we first conceptually and empirically extend this Model 1 that can accommodate predictive relationships, that is, LARI (S-1) model, and then find that the LARI (S-1) model functions well as a means of conceptualizing multisource ratings of leadership (even in a distinct context and additional sources of ratings). We also find that the LARI (S-1) model captures a significant, and at times, substantial portion of variability in leader effectiveness. Our results also suggest that the extent to which a particular source of leadership ratings predicts a leader's effectiveness is based, in part, on the observability of the leadership dimension being assessed. Implications and future directions are discussed.

Keywords: multisource leadership ratings, 360-degree feedback, self–other agreement, leader effectiveness

In the field of leadership, it has long been acknowledged there is value in capturing how leaders view themselves and how they are viewed by others (Fleenor et al., 2010; Lee & Carpenter, 2018). Armed with these distinct perspectives, one can determine where a leader and other stakeholders come to consensus about the leader's capabilities, as well as potential "blind spots" or areas of disagreement. These two perspectives are useful in facilitating leader development by raising a leader's self-awareness using multisource ratings of leadership (Atwater & Yammarino, 1992; Fleenor et al., 2010). Often referred to as 360s, these ratings are a mainstay in leadership development programs around the world (Fleenor et al., 2020; Slater & Coyle, 2014).

Recently, leadership scholars advanced the leadership Arena–Reputation–Identity (LARI) model as a means of more precisely capturing the shared and unique perspectives that underlie multisource ratings of leadership (Vergauwe et al., 2022). This model, which draws on both the Johari window (Luft & Ingram, 1955) and research on multisource ratings of personality (B. S. Connelly et al., 2022; McAbee & Connelly, 2016), maps information provided by 360 assessments onto a 2 × 2 grid depending on whether it is known or unknown to self or others (see top of Figure 1). In the LARI model, *Arena* reflects information shared between the leader and others (i.e., known to the self and others). *Reputation* corresponds to aspects of leaders that others see but are unknown to themselves. *Identity* reflects information known to the leader, but unknown to others. Last, there are aspects that are unknown to both the leader and others (i.e., "unknown").

Vergauwe et al. (2022) demonstrated that the LARI model and this framework can be translated into a bifactor model when analyzing multisource ratings of leadership (see the bottom left of Figure 1). Across two samples, the authors found the LARI model fit the data well and better than several alternative models. Furthermore, they found that shared perspectives (i.e., Arena + Reputation) accounted for only half of the variance in these ratings, while the remaining variability could be attributed to specific sources (e.g., peers, direct reports).

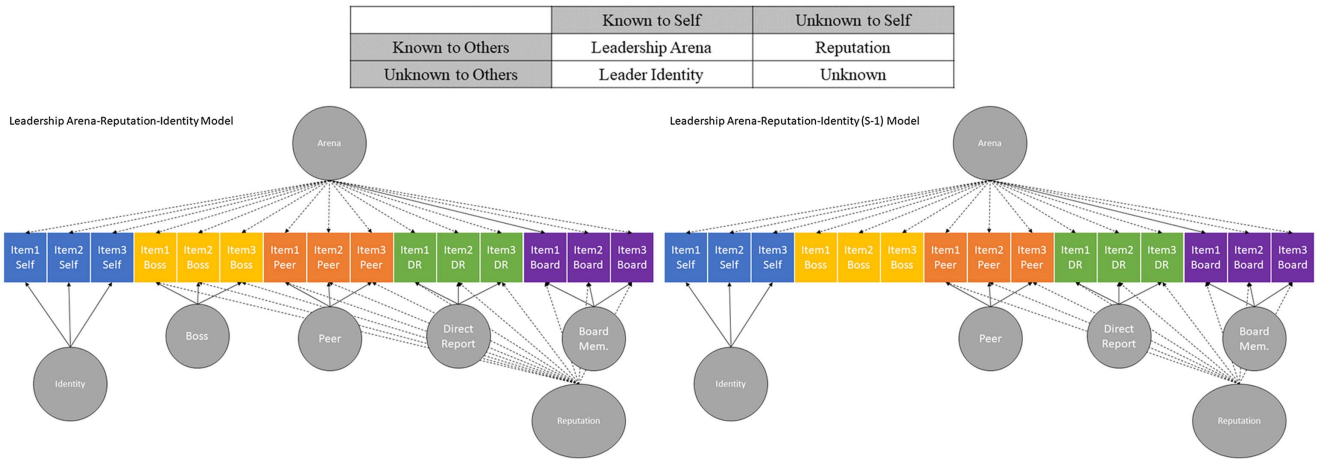
Talya N. Bauer served as action editor.
Andrew C. Loignon https://orcid.org/0000-0002-0073-3963
John W. Fleenor https://orcid.org/0000-0003-3384-7956
An earlier version of this article was presented at the Annual Meeting of the Society for Industrial and Organizational Psychology in 2023 in Boston, Massachusetts.

Andrew C. Loignon played a lead role in conceptualization, data curation, formal analysis, methodology, writing–original draft, and writing–review and editing. John W. Fleenor played an equal role in conceptualization, methodology, and writing–original draft. Stephen Jeong played an equal role in conceptualization, methodology, and writing–original draft. David J. Woehr played an equal role in conceptualization, investigation, writing–original draft, and writing–review and editing.

Correspondence concerning this article should be addressed to Andrew C. Loignon, Center for Creative Leadership, Greensboro, NC 27410, United States. Email: loignona@ccl.org

Figure 1

The Johari Window (Top) and the Leadership Arena–Reputation–Identity (LARI) Model and LARI (S-1) Model (Bottom)



Note. LARI (S-1) = LARI model extension; DR = direct report; Board Mem. = board member. See the online article for the color version of this figure.

Although the LARI model holds promise in disentangling the shared and unique perspectives that underlie multisource ratings of leadership, several questions remain to be answered and are the focus of this study. First, it remains unclear whether the information provided by the LARI model has any bearing on the effectiveness of a leader. Do the shared and unique perspectives relate to criteria that are important to leaders and other stakeholders? Answers to such questions may provide explicit tests of competing theories of multisource ratings. The LARI model allows us to examine whether leader effectiveness is impacted more by a leader's general standing on a particular dimension (i.e., Arena), as presumed in self–other agreement research (Fleenor et al., 2010), or by adaptive responses of leaders to the demands of specific stakeholders (e.g., superiors vs. direct reports). Additional perspectives, like stereotype-based or reputational models of leadership, emphasize the role of a leader's reputation, which corresponds to information known solely by others (McAbee & Connelly, 2016). Importantly, to address these questions, prevailing recommendations would require researchers to modify the existing LARI model such that a specific source (e.g., a leader's superior) serves a reference domain (bottom right of Figure 1; Eid et al., 2017; Zhang et al., 2021). Thus, by extending the LARI model so that it can incorporate criteria, we seek to understand how leaders shift in response to specific stakeholders or general situations and how important these rater-specific versus general manifestations are in predicting leader effectiveness.

Second, Vergauwe et al. (2022) found that the relative amount of variance captured by different LARI factors depends on the specific dimension being rated. Returning to the theoretical roots of the LARI model, we extend these findings by drawing on personality research that found the observability of traits has implications for multisource ratings (John & Robins, 1993; Vazire, 2010; Tett et al., 2021). We examine whether the amount of variance explained by a particular source depends on the level of observability of the dimension being rated. By considering the moderating effects of the content being assessed, we test whether the components of the LARI model are conceptually distinct and driven by separate elements (Vergauwe et al., 2022).

Third, much of the existing research on multisource ratings of leadership is drawn from entry- and midlevel leaders (Fleenor et al., 2010). The initial application of the LARI model was no different with 77%–93% of both samples consisting of leaders from these levels (Vergauwe et al., 2022). By situating our study in the upper echelons of the firm, we acknowledge that top leaders are not only evaluated by the typical raters in multisource leadership studies (e.g., superiors, peers, direct reports), but they are also evaluated by powerful external stakeholders (i.e., board members). We replicate prior tests of the LARI model in this context and address several issues including: Does the LARI model generalize to distinct contexts while accommodating additional source effects (i.e., board members)? Answers to these questions are necessary to extend and strengthen evidence supporting the LARI model (Kohler & Cortina, 2021).

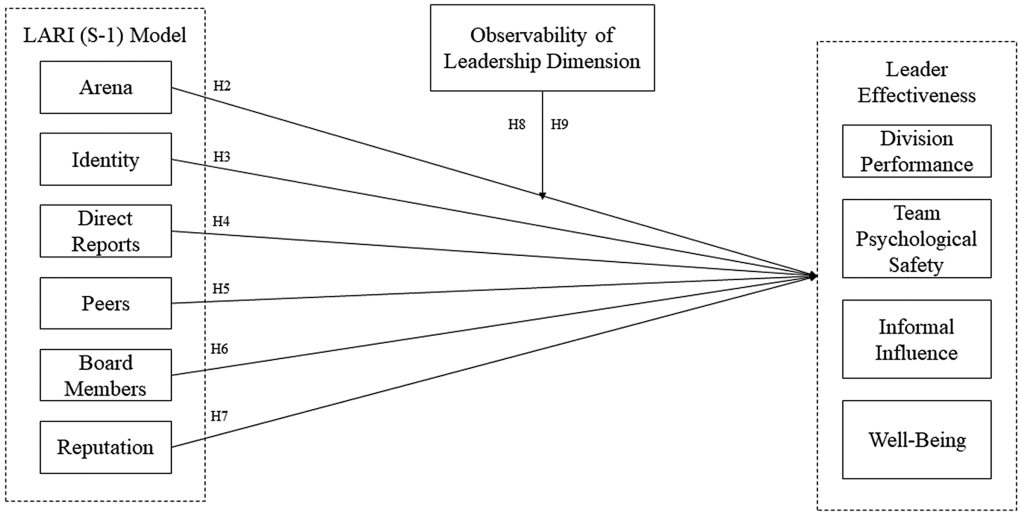
Hypotheses Development

In this section, we discuss the distinction between the LARI model, which was recently advanced for multisource leadership ratings, and our extension (LARI [S-1]) that is intended to incorporate criterion measures (Eid et al., 2017; Zhang et al., 2021). We develop hypotheses regarding the predictive validity of each component of the LARI model. Specifically, we consider the type of relationship hypothesized for Arena, Reputation, Identity, and source-specific factors in the model and leader effectiveness. We also consider the extent to which these relationships depend on the observability of the dimension of leadership being evaluated. Figure 2 features a conceptual model of our hypotheses.

The LARI (S-1) Model as an Extension of Earlier Models of Multisource Ratings

Prior studies of multisource ratings of leadership presume that each rating source provides idiosyncratic information about a leader (i.e., the "discrepancy hypothesis"; Borman, 1997; Fleenor et al., 2010; LeBreton et al., 2003; Lee & Carpenter, 2018). Such differences arise

Figure 2
Conceptual Model for Hypotheses



Note. LARI (S-1) = leadership Arena–Reputation–Identity model extension.

because the leader may not act consistently with each rater group, or raters may have different opportunities to observe leaders (Borman, 1997). From this perspective, rating sources are not interchangeable; therefore, the variance unique to each source is not error but provides valid information. This assumption is captured directly in the LARI model via a common Reputation factor that is distinct from each of the unique sources (Vergauwe et al., 2022). The Reputation factor represents the characteristics of a leader that sources agree on while preserving the perspective of each individual source (see bottom left of Figure 1).

A natural extension of the LARI model is to consider whether the distinct sources provide predictive information. Unfortunately, the initial formulation of the LARI model may not lend itself to answering these questions (Vergauwe et al., 2022). Because the LARI model represents a bifactor model, there is a distinct risk that simply incorporating criteria could lead to anomalous results (e.g., irregular factor loadings, indeterminate structural pathways; Eid et al., 2018; Zhang et al., 2021). Many of these issues can be traced to tensions between the general factor (Arena) and the specific sources of leadership ratings. Conceptually, scholars may be tempted to view the Arena factor as a shared rating of the leader's competencies (Eid et al., 2017). However, because the LARI model does not randomly select sources from a broader universe of raters, the specific factors are not interchangeable. Instead, each rating source is distinct and should be viewed as fixed factors.

To overcome these limitations, we developed the LARI (S-1) model as an alternative specification, which would limit the likelihood of these anomalous results (see the bottom right of Figure 1).[1] The LARI (S-1) model takes a specific source of leadership ratings and treats them as the reference indicator for the Arena factor (Eid et al., 2017). We chose a leader's superior as the reference source because these individuals occupy formal positions of responsibility and control over an individual and are, in many ways, seen as a critical point of comparison. This approach also allowed us to preserve the three quadrants of the Johari window (top half of Figure 1). Within the LARI (S-1), the Arena factor reflects, primarily, the extent to which the sources of leadership ratings agree with one's superior, and specific sources capture deviations of these sources from the values expected based on the Arena variable. Thus, a positive score for a specific source (e.g., Identity or Board Member factors) would indicate that a source provides higher ratings than one would expect based on how the superior and others generally view the focal individual.

Our decision to develop the LARI (S-1), as opposed to employing a traditional bifactor model, is largely a reflection of the advantages it affords in the context of multisource leadership ratings for executives. Namely, S-1 bifactor models appear to be especially useful when confronted with relatively smaller sample sizes, having fewer items, and not having interchangeable source factors (i.e., distinct types of raters; Eid et al., 2017; Zhang et al., 2021). Because we derived the LARI (S-1) model from the LARI model, which was designed specifically for multisource ratings of leadership, we hypothesize that:

> *Hypothesis 1:* The LARI (S-1) model adequately fits multisource ratings of leadership.

Arena

As noted previously, in the LARI (S-1) model, Arena is a general factor that is defined primarily by the evaluation of one's superior and captures variance common across both self- and observer reports (McAbee & Connelly, 2016; Vergauwe et al., 2022). Because the Arena factor captures shared variance in a given dimension of

[1] We decided to develop the LARI (S-1), as opposed to employing a traditional bifactor model. This is largely a reflection of the advantages it affords in the context of multisource leadership ratings for executives (e.g., sample size, number of items, having items that meet the proportionality constraint; Eid et al., 2017; Zhang et al., 2021). Other contexts, measures, and research questions may be better suited for traditional bifactor models.

leadership, it corresponds directly to research on agreement in multisource ratings of leadership (Fleenor et al., 2010; Lee & Carpenter, 2018).

Differing theories led us to hypothesize that higher scores on Arena factors are predictive of effective leadership (Atwater et al., 1998; Bass & Yammarino, 1991). For example, a meta-analysis involving 128 independent samples found an overall correlation of .22 ($\rho = .34$) between self–supervisor rating agreement and job performance (Heidemeier & Moser, 2009). In personality research, agreement between the self and observers reflects a consistency between internal (i.e., affect, cognition, and desire) and external (i.e., behavior) manifestations of an underlying trait (B. S. Connelly et al., 2022). Likewise, in leadership research, higher levels of self–other agreement among leaders and specific sources are associated with leader effectiveness (Fleenor et al., 2010). Therefore, we hypothesize:

Hypothesis 2: Higher standing on the Arena factor is positively associated with leader effectiveness.

Identity

Within the LARI (S-1) model, like its predecessor, the Identity factor reflects variance in a leader's self-evaluation not shared with observers. The unique variance captured in Identity comprises both errors in self-perception (e.g., erroneously believing one is inspirational) and information about leaders that is not available to or shared with raters completing the 360 assessment (e.g., not openly expressing thoughts on one's true motivation; McAbee & Connelly, 2016).

Prior studies of multisource leadership consider the Identity factor as a form of bias or inaccurate self-perception (Fleenor et al., 2010; Lee & Carpenter, 2018). These biases may arise from the lack of opportunities to receive feedback from others that limit the leader's ability to (dis)confirm their perceptions; a lack of motivation to use others' perceptions as a relevant source of feedback; or thoughts, perceptions, or attitudes that are kept private (McAbee & Connelly, 2016; Vergauwe et al., 2022). Given these considerations, we hypothesize that:

Hypothesis 3: Higher standing on the Identity factor is negatively associated with leader effectiveness.

Unique Perspectives in Multisource Ratings of Leadership

Like its predecessor, the LARI (S-1) model presumes that the unique perspectives of the different raters afford relevant information. Theoretically, others' ratings will exhibit criterion-related validity because such evaluations are informed, at least in part, by their awareness of leaders' prior performance, which itself is a predictor of future performance (Oh et al., 2011). Furthermore, multiple theoretical frameworks (e.g., signaling theory, socio-analytic theory) presume that raters are motivated to form accurate perceptions of a leader to inform their own behaviors and responses (e.g., Banks et al., 2021; B. L. Connelly et al., 2011; McAbee & Connelly, 2016; Vergauwe et al., 2022). Such accuracy should, in turn, contribute to the predictive validity of different raters' evaluations.

Relatedly, Lance et al. (2006) compared multisource ratings from two perspectives: a normative accuracy model and an ecological perspective. The normative accuracy model treats source differences as bias. The ecological perspective suggests that different sources have valid but differing views on the performance of a leader. Specifically, the ecological approach postulates that individuals (e.g., peers, supervisors, subordinates) in different relationships with the leader will interact in different ways and have different expectations and attentional foci when observing and processing information about the leader. Lance et al. (2006) presented strong empirical support for the ecological perspective relative to the normative accuracy model. Their results indicate that different rating sources provide valuable and complementary information about the leader. Given these findings, we also form hypotheses for each source in the LARI (S-1) model.[2]

Direct Reports

Direct reports, who have many opportunities to witness their leader in action, are commonly incorporated into 360 assessments. Research presumes that direct reports, compared to other sources, have a unique perspective on various dimensions of leadership (e.g., leveraging differences, interpersonal savvy, leading change; Fleenor et al., 2010). To the extent that these unique perspectives are favorable, we hypothesize:

Hypothesis 4: Higher standing on the Direct Report factor is positively associated with leader effectiveness.

Peers

Peers, relative to other sources, are unlikely to be the direct target of leader actions and influence. Because a leader's peers are leaders themselves, they are not likely to depend on the focal leader for resources or experience power differences. On the other hand, peers are privy to leader behaviors that others rarely see (Braddy et al., 2014), such as dimensions capturing a leader's judgment, ability to lead change, or ability to develop and empower others. As such, we hypothesize:

Hypothesis 5: Higher standing on the Peer factor is positively associated with leader effectiveness.

Board Members

For senior leaders, board members represent a unique and important stakeholder. As a source of corporate governance, board members influence the strategic decisions made in organizations and the plans that are ultimately implemented (e.g., Beekun et al., 1998; Hill, 1990). Although board members are removed from the day-to-day operations of an organization, we hypothesize that more positive evaluations of a senior manager's leadership from this source are indicative of more effective leadership. Specifically, we hypothesize:

Hypothesis 6: Higher standing on the Board Member factor is positively associated with leader effectiveness.

[2] As noted previously, because one's superior is used as the reference source for the Arena factor and is not modeled separately in the LARI (S-1) model, we do form hypotheses for this source of leadership ratings.

Reputation

The last component of the LARI model is the Reputation factor. With the LARI model, Reputation is defined as shared residual variance that is unique to the rater sources that arises both from errors in raters' perceptions (e.g., stereotypes) and from information relevant to the leadership dimensions unavailable to the self (e.g., subtle facial or nonverbal responses that arise during meetings; McAbee & Connelly, 2016; Vergauwe et al., 2022).[3]

Reputation directly corresponds with blind spots or to aspects of one's leadership that are collectively known to others but unknown to the self (Fleenor et al., 2010). There are several ways that a leader's reputation can emerge including (a) information that a leader is simply unaware of, (b) information that is intentionally withheld by the leader but is still picked up by other raters, or (c) systematic bias that is shared across all observer groups (e.g., physical appearance stereotypes or a leniency bias; Vergauwe et al., 2022). Because Reputation corresponds to a breakdown in the shared reality between leaders and other stakeholders (Fleenor et al., 2010; Gooty & Yammarino, 2011), we hypothesize:

> *Hypothesis 7:* Higher standing on the Reputation factor is negatively associated with leader effectiveness.

Moderating Effects of the Observability of Leadership Dimensions

Research has yet to explicitly consider the nature of *what* is being evaluated when applying the LARI model to multisource leadership ratings. This is critical given that multisource personality research, from which the original LARI model was derived (Vergauwe et al., 2022), emphasizes the importance of the observability of the traits that raters are evaluating. For example, the self–other knowledge accuracy (SOKA) model proposes that a leader's self-evaluation and other's evaluations are likely to diverge depending on the information available as well as the salience of that information (Vazire, 2010). As an example, with more observable traits, the information that is available to the leader, relative to other sources, is likely to be limited. Similarly, notions of observability underlie trait activation theory. Trait activation theory proposes that situations differ in the extent to which they provide cues for trait-relevant behavior or demand responses that not all individuals are equally capable of displaying (Tett et al., 2021).

Similar ideas have been raised in other multisource assessments like assessment centers. Because interactions with different rating sources unfold within distinct situations (e.g., formal authority, hierarchy, typical task demands), leaders are expected to exhibit different behaviors relevant to different leadership dimensions. In fact, the lack of convergence in the dimensions being assessed in assessment centers has been attributed to discrepancies in the extent that a particular exercise afforded raters the opportunity to observe relevant behaviors (Fleenor, 1996; Haaland & Christiansen, 2002; Lievens et al., 2015; Schollaert & Lievens, 2012).

The importance of observability has been alluded to in leadership research. For example, when considering multisource ratings of personality, McKee et al. (2018) discussed how aspects of a leader (e.g., behaviors, performance) are more or less observable depending on who is rating those aspects. They highlighted how observability depends on not only what is being evaluated (e.g., leadership dimensions) but in what context the evaluations unfold (e.g., senior leaders; McKee et al., 2018).

Vergauwe et al. (2022) found evidence that the observability of the leadership dimension being evaluated may influence the results of the LARI model. Consistent with the SOKA model (Vazire, 2010), the authors proposed that with more observable dimensions, the shared perceptions of leadership (i.e., Arena + Reputation) would account for relatively more variance, while unique perceptions (e.g., Identity) would account for less.

Observability also has implications for the predictive validity of the components of the LARI (S-1) model. Again, drawing on trait activation theory (Haaland & Christiansen, 2002; Tett et al., 2021) and the SOKA model (Vazire, 2010), any given source of leadership ratings may have limited opportunity to observe specific relevant information, or this information may be less salient to them (Braddy et al., 2014). Thus, different raters may interact with the leader for relatively brief periods of time or have their interactions in an environment that elicits contextually bound manifestations of leadership qualities (McAbee & Connelly, 2016; Oh et al., 2011). Take, for example, a senior leader and their board members. Typically, such interactions span organizational boundaries, focus primarily on strategic-level decisions, and occur relatively infrequently. Thus, as a leader interacts with board members, we would expect the situation to require more behaviors related to task-focused dimensions (e.g., a leader's understanding of the firm's business) rather than elicit behaviors pertaining interpersonal-focused dimensions (e.g., a leader's ability to cultivate relationships with colleagues in different departments). Ultimately, the extent to which a leader has the opportunity to exhibit behaviors with certain raters, thus making the corresponding leadership dimension more or less observable for those raters, has implications for the predictive validity for that source's ratings (Tett et al., 2021). In fact, prior research on multisource personality ratings has found that personality dimensions are relatively poor predictors when evaluated under conditions that afford raters with limited observability (Oh et al., 2011; Tett et al., 2021).

In this study, we test and extend these propositions. We expect that for more observable dimensions of leadership, Arena and Reputation will not only capture more variance in ratings of leadership but will emerge as stronger predictors of leader effectiveness. Much like a meta-analysis, we are seeking to explain the heterogeneity in effect sizes (e.g., factor loadings, β coefficients) based on our understanding of the observability of the leadership dimension for each source in the LARI (S-1) model. Specifically, we hypothesize:

> *Hypothesis 8:* The greater the observability of a leadership dimension, the more explained common variance in multi-source leadership ratings is attributed to general/shared factors (i.e., Arena and Reputation), and the less explained common

[3] There is a tradition of examining reputations in the organizational sciences (e.g., Ferris et al., 2003). Although this work spans multiple disciplines and levels of analysis, it shares many similarities with the concept of reputation examined here. Namely, reputations, in general, are thought to be perception-based, based on the evaluations of multiple observers, and vary along multiple dimensions. The primary difference, though, may be that past work on reputation has typically emphasized a target's broader reputation rather than simultaneously considering the specific sources whose ratings combine to form that reputation (i.e., emphasizing the sum over its parts).

variance is attributed to specific source factors (i.e., Direct Report, Peer, Board Member, Identity).

Hypothesis 9: The greater the observability of a leadership dimension, the stronger the predictive relationship between the general/shared factors (i.e., Arena and Reputation) and leader effectiveness compared to specific source factor (i.e., Direct Report, Peer, Board Member, Identity).

Method

Participants and Procedure

Our data consist of multisource ratings of 491 senior leaders attending a 5-day leadership development program for *C*-level executives. The program provides insights into their leadership effectiveness, suggestions for enhancing their influence with stakeholders (e.g., boards, shareholders), and ways to improve their well-being. Leaders completed the program between August 2019 and July 2023.[4]

The senior leaders in this study represent a diverse set of business sectors and leadership experiences. The leaders led organizations operating in 32 different industries (e.g., government or public sector, $n = 52$; manufacturing, $n = 38$; aerospace and defense, $n = 37$). They led small organizations (fewer than 100 employees) to much larger firms (more than 10,000 employees). The average senior leadership team consisted of seven members ($M = 7.36$, $SD = 2.49$). Team leaders, were, on average 51 years old ($SD = 6.85$ years) and were born in 32 different countries. Most had worked for their current organization for 13 years ($SD = 10.90$), identified as male (73%), held graduate degrees (68%), and self-identified as White (82%).

Measures

Multisource Ratings of Leadership

Prior to the program, leaders and their colleagues completed a multisource leadership assessment (Leslie et al., 2015) that measures dimensions relevant for senior-level leaders. To ensure that this measure exhibited sufficient unidimensional and multidimensional validity, we conducted an exploratory factor analysis of all possible items and retained only those dimensions and items with factor loadings greater than .60 (see Appendix A for a complete description of these analyses). This yielded a three-factor solution: (a) Forging Synergy—seeks common ground in an effort to resolve conflicts; (b) Results Orientation—clearly conveys objectives, deadlines, and expectations; and (c) Business Understanding—understands the strengths and weaknesses of major competitors. The three scales were measured using three items rated on a 5-point scale ranging from 1 (*deficient*) to 5 (*exceptional*). Aside from the leader's self-ratings, an average of nine raters ($SD = 3.81$) rated each leader. A total of 4,661 raters, including 2,118 direct reports, 1,701 peers, 474 superiors, and 368 board members provided ratings. Across dimensions and sources, these scales exhibited adequate internal consistency (avg. $\alpha = .83$, min. = .75, max = .87), interrater agreement (grand mean $r_{wg(j)} = .85$, min. = .83, max = .90), and interrater reliability, grand mean intraclass correlation coefficients, ICC, $(2, k) = .72$, min. = .65, max. = .80 (see Appendices B and C for more details, including descriptive statistics and correlations). To avoid potential aggregation and missing data biases, we proceeded with the median number of raters for each rater type (one superior, one board member, four direct reports, and three peers), which served as input for our analyses and used full-information maximum likelihood estimation with our models (Newman, 2014).

Observability of Leadership Dimensions

Members of the authorship team, with an average of 14 years of experience with multisource assessments, independently evaluated the level of observability for each leadership dimension as evaluated by a particular rating source. Each member considered the extent that a given source (i.e., the leaders themselves, direct reports, peers, superiors, or board members) would be able to observe certain behaviors or have more information available to them. Ratings were made using a 5-point scale ranging from 1 (*very difficult to observe*) to 5 (*very easy to observe*). The average interrater reliability, ICC$(2, k)$ across all sources was .68 (.60–.74; Shrout & Fleiss, 1979; see Appendix D for descriptives across dimensions and sources).

Division Performance

During the program, leaders and their team members completed a team-based assessment (Loignon & Wormington, 2022). This assessment featured several different measures related to a leader's effectiveness (Ulrich et al., 1999). First, the leader and their teammates evaluated the performance of their division using a five-item scale (e.g., "Overall, the division/function is effective") with response options ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). The coefficient α was .93, the average $r_{wg(j)}$ while assuming a rectangular distribution was .85 ($SD = .12$), and the ICC(1) was .22. Because a primary objective for senior leaders is the performance of their business areas, this is a key component of leader effectiveness.

Psychological Safety

The leader and their team members rated the psychological safety in the group using Edmondson's (1999) seven-item scale (e.g., "It is safe to take a risk in this team."). Responses were provided using a 5-point scale (1 = *strongly disagree* to 5 = *strongly agree*). The coefficient α was .86, the average $r_{wg(j)}$ while assuming a rectangular distribution was .84 ($SD = .12$), and the ICC(1) was .20. Psychological safety is a consistent predictor of team effectiveness (Frazier et al., 2017), and it is recognized as a key consideration for leaders in their teams (Edmondson, 1999).

Informal Leadership

Senior leaders and their team members completed round-robin ratings of informal leadership. Each person indicated who else on the team leads them (1 = *this person leads me*, 0 = *this person does not lead me*). For each senior leader, we calculated an in-degree or target effect using the social relations model (Kenny, 1994). Thus, higher scores indicate the extent that a senior leader is recognized as a source of informal influence in their teams ($r_{wg} = .78$, $SD = .36$). Given that leadership is increasingly recognized as something

[4] These data are part of a broader data collection effort and have been used in prior research (Loignon & Wormington, 2022).

granted by others (DeRue, 2011), this measure captures the level of informal leadership afforded to a senior leader in their team.

Well-Being

We included leaders' self-reported levels of thriving as a measure of their well-being (Spreitzer et al., 2005). Well-being is increasingly seen as a key criterion within organizations (Tay et al., 2023). Leaders rated four items (e.g., "I have energy and spirit.") using a 5-point scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). The coefficient α was .91.

Transparency and Openness

We tested our hypotheses using existing data from a U.S.-based global leadership development provider. We describe all measures and our sample, adhering to the *Journal of Applied Psychology* methodological checklist. Data were analyzed using R, Version 4.0.0 (R Development Core Team, 2010) as well as the psych (Version 2.3.9; Revelle, 2023) and lavaan (Version 0.6-17; Rosseel, 2012) packages. In an Open Science Forum repository, we provide our preregistered hypotheses and analysis plan[5] and additional online material allowing others to replicate our analyses, which is accessible at https://osf.io/f7n2r/?view_only=4914a29e071946b380161b4d6b082c7d. Because of their proprietary nature, the data used in this study are not available. Also, as members of a nonprofit organization, we did not have access to an institutional review board. Instead, we used data that is collected as part of ongoing business practices. Nevertheless, we endeavored to ensure that our article conforms to prevailing ethical standards (e.g., using archival data, ensuring that participants sign an informed consent form acknowledging that their data can be used for research purposes).

Results

Testing the LARI (S-1) Model

We first tested whether the LARI (S-1) model was a reasonable representation of the multisource leadership ratings (Hypothesis 1, Table 1). Factor loadings for these models are reported in Appendix E. For each dimension, the LARI (S-1) model, which included predictive pathways to the measures of leader effectiveness, was either the best fitting (i.e., provided statistically significant improvement in fit) or exhibited adequate overall levels of fit (minimum comparative fit index = .96, minimum Tucker–Lewis index = .96, maximum root-mean-square error of approximation = .03). Other models that exhibited better fit (e.g., LARI) have a tendency to overfit the data and can yield anomalous results (e.g., negative factor loadings; Eid et al., 2017; Zhang et al., 2021). These findings replicate those reported in previous research (Vergauwe et al., 2022) and also extend this study by showing that a LARI (S-1) model that includes criterion measures adequately represents the data while affording an opportunity to consider the predictive validity of the different sources of leadership ratings.

Explained Common Variance in Multisource Leadership Ratings Across LARI Factors

We considered the extent to which the proportion of explained common variance (ECV) in multisource leadership ratings varied across the factors in the LARI (S-1) model (Reise, 2012). We squared then summed the loadings from each factor and then divided this total by the sum of the squared factor loadings for all factors (McAbee & Connelly, 2016; Vergauwe et al., 2022; see Figure 3). Across the leadership dimensions, the factors representing shared perspectives explained relatively large proportions of variance (Arena: $M = 40\%$, $SD = 5\%$; Reputation: $M = 20\%$, $SD = 6\%$). However, several of the specific sources also explained substantial portions of variance (Self: $M = 11\%$, $SD = 3\%$; Board Members: $M = 13\%$; $SD = 3\%$). These results, estimated using the LARI (S-1) model, correspond with previous studies indicating that no single factor accounts for all, let alone the majority, of the variance in the multisource leadership ratings (Vergauwe et al., 2022).

We then considered whether the overall pattern of ECV in leadership ratings varied according to the observability of the dimensions (Hypothesis 8). Specifically, we regressed each item's level of ECV on the interaction between its respective source within the LARI (S-1) model and the level of observability for these sources (Table 2). Several of the interactions were significant and contributed to a significant change in R^2 beyond a main effects model. To better understand these effects, we plotted the predicted levels of ECV (see Figure 4). In general, we observe positive moderating effects of observability on ECV for several of the source factors (i.e., Identity, Board Members, and Direct Reports), while the Arena and Reputation factors largely exhibited null relationships. These findings fail to support Hypothesis 8. Instead of observability increasing the proportion of ECV in just the general factors (i.e., Arena and Reputation), more observable leadership dimensions are associated with greater proportions of variance explained, and this effect is particularly pronounced for unique sources of ratings. We revisit these findings in the Discussion section.

Predictive Validity of LARI Factors

To examine the predictive validity of the factors within the LARI (S-1) model (Hypotheses 2–7), we considered the path estimates between each LARI factor and the measures of leader effectiveness (Table 3). Across the four criteria of leader effectiveness, the multisource leadership ratings were stronger predictors of division performance (average $R^2 = .31$), psychological safety (average $R^2 = .25$) than well-being (average $R^2 = .11$) than informal leadership (average $R^2 = .12$). We first consider the shared perspectives in the LARI (S-1) model. When significant, the Arena factor was a positive predictor of division performance, psychological safety, and well-being, providing support for Hypothesis 2. Reputation was a significant predictor of division performance and psychological safety and tended to exhibit positive relationships with the criteria. This fails to support Hypothesis 7. Among the unique perspectives in the LARI (S-1) model, the Identity factor was a positive predictor of well-being and a negative predictor of informal leadership, which provides mixed support for Hypothesis 3. Direct reports' ratings were positive and consistent predictors of division performance,

[5] We preregistered hypotheses regarding the replication of the LARI model (i.e., its fit compared to other models and the proportion of variance explained across leadership dimensions). However, for the sake of parsimony, those have been relegated to the online repository. Our first hypothesis, which considers the fit of the LARI (S-1) model, was not preregistered as this emerged during the review process.

Table 1
Model Fit of the LARI (S-1) Models Compared to Alternative Models

Model	χ^2	*DF*	*p*	CFI	TLI	RMSEA	RMSEA—95% CI	
							LL	*UL*
1. Results orientation								
LARI (S-1) model w/criteria	615.08	471	.00	.97	.97	.03	.02	.03
LARI (S-1) model	495.67	375	.00	.97	.97	.03	.02	.03
LARI model	399.70	365	.10	.99	.99	.01	.00	.02
TRI model	449.54	378	.01	.98	.98	.02	.01	.03
LARI model w/o G-Reputation	582.83	385	.00	.96	.95	.03	.03	.04
LARI model w/o S-Reputation	901.03	397	.00	.89	.88	.05	.05	.06
Higher order model	598.04	388	.00	.95	.95	.03	.03	.04
Correlated factors model	512.44	383	.00	.97	.97	.03	.02	.03
2. Business understanding								
LARI (S-1) model w/criteria	694.76	471	.00	.96	.96	.03	.03	.04
LARI (S-1) model	438.92	365	.01	.99	.98	.02	.01	.03
LARI model	462.59	378	.00	.99	.98	.02	.01	.03
TRI model	638.12	385	.00	.96	.95	.04	.03	.04
LARI model w/o G-Reputation	865.47	397	.00	.92	.91	.05	.05	.05
LARI model w/o S-Reputation	631.61	387	.00	.96	.95	.04	.03	.04
Higher order model	503.63	383	.00	.98	.98	.03	.02	.03
Correlated factors model	438.92	365	.01	.99	.98	.02	.01	.03
3. Forging synergy								
LARI (S-1) model w/criteria	570.35	471	.00	.99	.98	.02	.01	.03
LARI (S-1) model	426.25	365	.02	.99	.99	.02	.01	.03
LARI model	448.15	378	.01	.99	.99	.02	.01	.03
TRI model	536.28	385	.00	.98	.97	.03	.02	.03
LARI model w/o G-Reputation	840.16	397	.00	.93	.92	.05	.04	.05
LARI model w/o S-Reputation	551.77	387	.00	.97	.97	.03	.02	.04
Higher order model	438.94	367	.01	.99	.99	.02	.01	.03
Correlated factors model	426.25	365	.02	.99	.99	.02	.01	.03

Note. $n = 491$. All models were estimated using full maximum likelihood estimation to account for missing data (Newman, 2014). LARI model reflects a bifactor model with two general factors (i.e., Arena, Reputation) and specific factors, which are forced to be uncorrelated. LARI (S-1) model is an extension of the LARI model that excludes the specific superior factor and defines the Arena factor with the superior's ratings. LARI (S-1) model with criteria references extends this model to include predictive pathways between the LARI factors and measures of leader effectiveness (i.e., division performance, psychological safety, informal influence, and well-being). We also tested several other alternative models: (a) a TRI or Trait–Reputation–Identity model, which models the Reputation factor as a higher order factor; (b) a LARI bifactor model without a general Reputation factor, which tests whether there is sufficient communality shared by external observers; (c) a LARI bifactor model without source-specific Reputation factors, which tests whether there is sufficient shared variance among specific groups of raters; (d) a higher order factor model, which tests whether the source-specific factors can be more parsimoniously captured with mere higher order factors; and (e) a correlated factors model, which tests whether a specific higher order factorial structure is necessary in the multisource leadership ratings. All models treat multiple raters within a particular source as interchangeable by including equality constraints across the raters within this group. Also, small negative factor loadings were constrained to zero for all models (B. S. Connelly et al., 2022). LARI = leadership Arena–Reputation–Identity; *DF* = degrees of freedom; CFI = comparative fit index; TLI = Tucker–Lewis index; RMSEA = root-mean-square error of approximation; CI = confidence interval; *LL* = lower limit; *UL* = upper limit; G = general; S = specific.

psychological safety, and informal leadership, which supports Hypothesis 4. Interestingly, Board Members exhibited a significant negative relationship with well-being, which fails to support Hypothesis 6, while Peers were a significant predictor of informal influence, supporting Hypothesis 5.

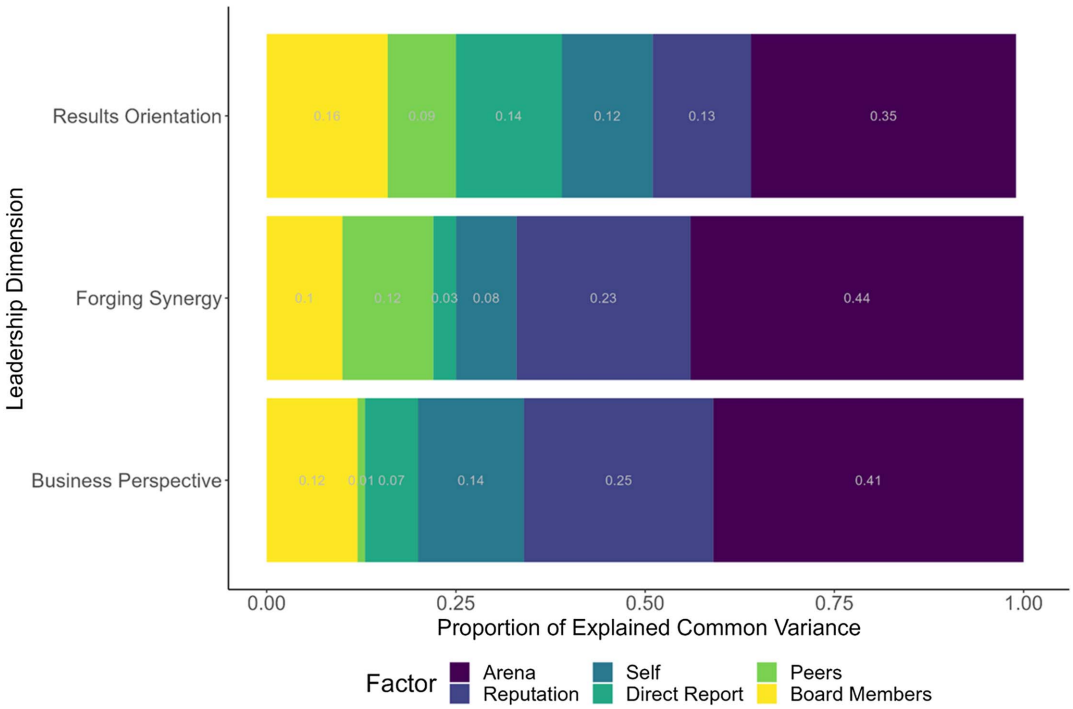
Relative Importance of LARI Factors

To test Hypothesis 9, which considered the moderating effects of the observability of a leadership dimension on predictive validity, we examined the proportion of variance explained by each LARI factor across the measures of leadership effectiveness. Because the LARI (S-1) model constrains the factors to be independent, the squared standardized coefficients sum to the overall model's R^2. Thus, we examined the squared standardized coefficients to determine the relative contribution of predictors which is akin to a dominance analysis (Braun et al., 2019). We regressed the squared standardized coefficients (i.e., each factor's relative predictive contribution) onto a series of dummy variables corresponding to specific LARI factors, the level of observability of the leadership dimensions, and the interaction between these sets of variables (see Model 2 in Table 2).

Our findings indicated that, even while controlling for other design features (e.g., the specific criteria of leader effectiveness and leadership dimension under consideration), the interactions between the level of observability and the LARI-specific factors provided an improvement in R^2 ($\Delta R^2 = .10$, $p < .05$). To better understand the nature of this moderating effect, we plotted the contribution of the different LARI factors across levels of observability (±1 *SD*) for the leadership dimensions (Figure 5). This plot indicates that, as the level of observability increased, the relative contribution of the Reputation and Direct Report factors decreased, while the effect of Peers increased. These results fail to support Hypothesis 9 in that greater levels of observability for a leadership dimension yielded

Figure 3

Proportion of Explained Common Variance in Multisource Leadership Ratings Across LARI Factors and Leadership Dimensions



Note. $n = 491$ senior leaders. LARI = leadership Arena–Reputation–Identity. See the online article for the color version of this figure.

stronger predictive relationships for a specific source factor (i.e., Peers) but not shared factors (i.e., Arena and Reputation).

Robustness Checks

Along with these focal analyses, we conducted a series of robustness checks described in Appendix F. They include: (a) replicating the relative contributions of the LARI model factors using a random-effects model, (b) ensuring that the moderating effects of the observability of leadership dimensions are distinct from their reliability (Lance et al., 2010; Oh et al., 2011), and (c) conducting a small-scale simulation study to ensure the statistical validity of our models. The findings of these additional analyses replicate and are fully consistent with the results reported here.

Discussion

Multisource ratings of leadership can open windows into powerful and revealing insights (Fleenor et al., 2010; Lee & Carpenter, 2018). Until the application of the LARI model (Vergauwe et al., 2022), the contribution of the different perspectives reflected in multisource ratings has remained obscured. Drawing on the assessment of over 400 senior leaders, this study advances findings that replicate and extend the initial applications of the LARI model. As with prior research (Vergauwe et al., 2022), we found support for the LARI model as a useful way of conceptualizing multisource ratings of leadership.

This study extends understanding of the LARI model and multisource assessments of leadership in several ways. First, we demonstrated that the LARI model generalizes to the upper echelon of organizations. Even when considering alternative leadership dimensions and incorporating additional context-specific sources (i.e., board members), the LARI model functions quite well. Our results suggest the different sources collected in multisource studies afford distinct perspectives that are not interchangeable and afford unique information (Borman, 1997; Fleenor et al., 2010; Lee & Carpenter, 2018).

We also answer recent calls to consider the predictive validity of the LARI model. By developing and testing the LARI (S-1) model, our findings suggest that across a range of leadership effectiveness criteria, the distinct sources reflected in multisource leadership ratings capture a significant, and at times, substantial portion of variability. Drawing on prior models of multisource ratings in both the personality and assessment center literature (Schollaert & Lievens, 2012; Vazire, 2010; Tett et al., 2021), we found that the level of observability of the leadership dimensions affects how the components of this model function. The variability in leadership ratings captured by specific factors in the LARI (S-1) model and the ability of these factors to predict leader effectiveness depend on the observability of the leadership dimension being assessed.

Theoretical Contributions

Our findings provide three contributions. First, we replicate and extend emerging models of multisource ratings of leadership.

Table 2

Moderating Effects of Observability on Explained Common Variance for LARI (S-1) Factors and Predictive Relationships With Leader Effectiveness

Predictor	Model 1 Dependent variable: ECV			Model 2 Dependent variable: β weights[2]		
	B	*SE*	*p*	*B*	*SE*	*p*
Intercept	.04	.03	.19	.64	.22	.01
Observability of leadership dimension	−.01	.01	.30	−.14	.06	.02
Observability × Arena	.00	.01	.98	.14	.07	.07
Observability × Identity	.03	.01	.04	.14	.06	.04
Observability × Direct Reports	.02	.01	.12	.08	.07	.26
Observability × Peers	−.01	.01	.55	.25	.07	<.001
Observability × Board Members	.02	.01	.08	.15	.06	.02
LARI factors (reference—Reputation)						
Arena	.01	.03	.86	−.59	.26	.03
Identity	−.07	.05	.14	−.61	.23	.01
Direct Reports	−.07	.04	.10	−.20	.26	.45
Peers	.03	.04	.54	−1.03	.26	<.001
Board Members	−.04	.04	.34	−.64	.22	<.001
Leadership dimension (reference—Business understanding)						
Strategic planning	.00	.00	.29	−.01	.01	.62
Leading change	.00	.00	.26	−.01	.01	.44
Criterion (reference—Division performance)						
Informal leadership				−.11	.03	.00
Psychological safety				−.02	.03	.47
Well-being				−.11	.03	<.001
Factor × Criterion						
Arena × Informal Leadership				.09	.04	.03
Identity × Informal Leadership				.12	.04	.01
Direct Reports × Informal Leadership				−.03	.04	.41
Peers × Informal Leadership				.17	.04	<.001
Board Members × Informal Leadership				.11	.04	.01
Arena × Psychological Safety				.03	.04	.44
Identity × Psychological Safety				.02	.04	.62
Direct Reports × Psychological Safety				−.03	.04	.49
Peers × Psychological Safety				.03	.04	.48
Board Members × Psychological Safety				.02	.04	.64
Arena × Well-Being				.11	.04	.01
Identity × Well-Being				.15	.04	<.001
Direct Reports × Well-Being				−.05	.04	.26
Peers × Well-Being				.12	.04	.01
Board Members × Well-Being				.15	.04	<.001
F		12.34**			4.40**	
R^2		.41			.77	
ΔF		5.52*			3.56*	
ΔR^2		.04			.10	
n		243			72	

Note. Change statistics (Δ*F* and Δ*R*) compare a model that excludes the interaction terms for observability and LARI factors to those presented here. LARI = leadership Arena–Reputation–Identity; ECV = explained common variance; *SE* = standard error; *n* = number of parameters estimated across the three LARI models.

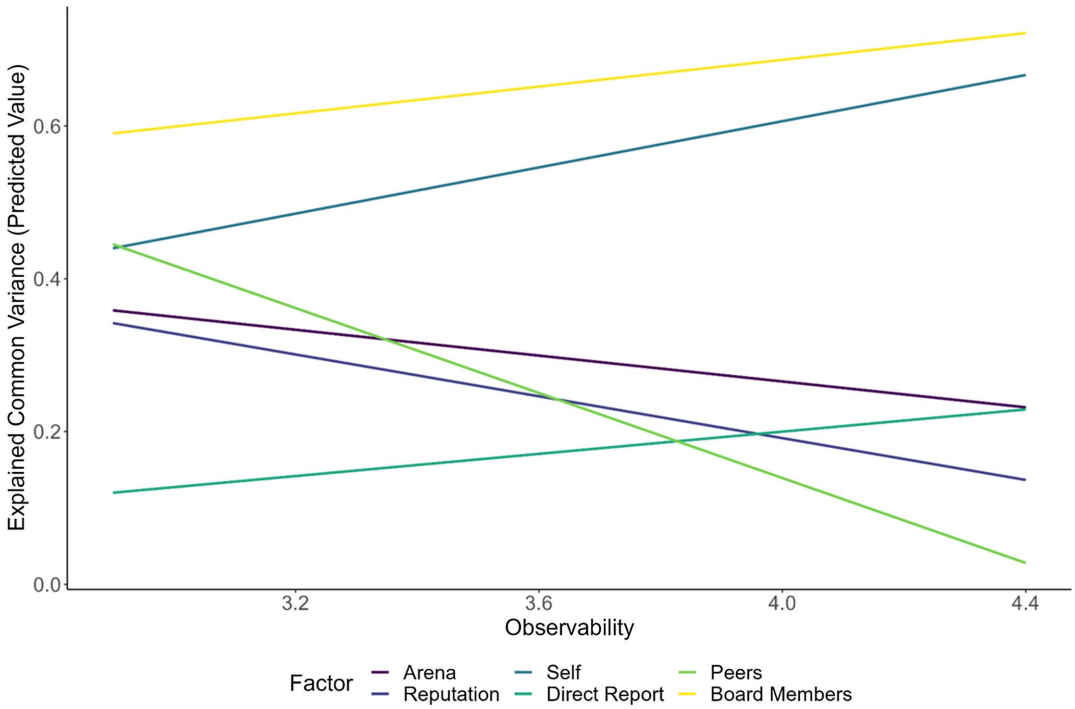
* $p < .05$. ** $p < .01$.

In particular, our findings further support the original LARI model's assumption that each source provides unique criterion-relevant information (Vergauwe et al., 2022). Because leaders work with theoretically distinct stakeholders (e.g., vertical reporting relationships vs. horizontal collaborations) across multiple contexts, they are expected to engage differently with these groups. Thus, each type of rater is afforded unique opportunities to observe the leader (Borman, 1997; Fleenor et al., 2010; Hoffman & Woehr, 2009). We also extend this model to a form, the LARI (S-1), that can examine whether the unique information is informative and criterion-relevant (Eid et al., 2017; Zhang et al., 2021). Importantly, this model treats a specific source (i.e., one's superior) as a reference point for the shared perceptions (i.e., Arena) and thus considers the extent to which remaining sources (e.g., Identity, Board Members) evaluate a leader higher or lower relative to this benchmark. Perhaps the most striking specific source in the present study, especially given our current context, is the role that board members' ratings play. This source accounts for 10%–16% of the variability in multisource leadership ratings and appears to have distinct predictive effects as their ratings diverge from other sources. Thus, taken as a whole, we find that the unique perspectives captured in multisource leadership ratings provide criterion-relevant information.

Second, we show that the level of observability of a leadership dimension is a critical consideration when theorizing about multisource

Figure 4

Moderating Effects of Observability on Factor Loadings in Multisource Leadership Ratings by Leadership Arena–Reputation–Identity (LARI) Factors



Note. $n = 243$ parameter estimates from LARI (S-1) models. Predicted values were estimated using Model 1 in Table 2 and assuming observability scores that ranged from ±1 *SD*. LARI (S-1) = LARI model extension. See the online article for the color version of this figure.

leadership ratings. By drawing on earlier models of interpersonal perception (e.g., Johari window, SOKA model; McAbee & Connelly, 2016; Vazire, 2010; Vergauwe et al., 2022), we theorized that self-evaluation and other's evaluations diverge depending on the information available. Contrary to what we hypothesized, however, observability has a limited effect on the predictive utility of the Arena and a negative relationship with the Reputation factor. Thus, for more observable dimensions of leadership, blind spots (i.e., what is unknown to the self and known to others) contribute less criterion-relevant information (Fleenor et al., 2010).

Third, our findings introduce a caveat for the discrepancy hypothesis, which presumes each rating source provides specific information about a leader (Borman, 1997). Across four distinct criteria of leadership effectiveness, much of the *unique* predictive validity is attributed to three sources (i.e., Arena, Reputation, and Identity). The remaining sources (i.e., Board Members, Direct Reports, and Peers), provide less unique information. However, their contributions are often dependent on the observability of the leadership dimensions being evaluated. This suggests that some of the effects of specific sources, as described in the discrepancy hypothesis, may emerge only under certain conditions (Tett et al., 2021). For instance, these findings may also suggest that observability, in fact, represents the salience of the leadership dimensions for specific sources making them more important for leader effectiveness.

Practical Implications, Future Research, and Limitations

Our findings inform the application of 360 leadership ratings. When reviewing multisource feedback, leaders often question what it means when sources' ratings diverge (Bracken et al., 2001). Because the LARI model captures both shared and unique factors, resulting in differential prediction, it may be appropriate to employ multiple interpretations and avoid aggregating ratings when interpreting multisource feedback. That is, it is likely that some leaders will need to prioritize specific stakeholders (e.g., Direct Reports vs. Board Members) while others should attend to more general trends (e.g., Arena). Put differently, when employing the Johari window as a model for interpreting multisource ratings (Luft & Ingram, 1955), it would behoove leaders to peer through each pane separately rather than gazing through the entire window.

As with any research, this study has limitations. Although there is a time lag between the multisource ratings of leadership and the criteria use distinct referents, there is likely some overlap among the raters who completed these assessments. This situation is not unlike other multirater studies that either used self-reported criteria or exhibited overlap among the raters completing assessments (Braddy et al., 2014; B. S. Connelly et al., 2022). Future research that employs longitudinal designs and objective criteria would complement our work.

Finally, although the LARI (S-1) shows promise as a means of advancing our understanding of multisource leadership ratings, it

Table 3

Summary of Path Estimates for Predicting Leader Effectiveness With LARI (S-1) Factors

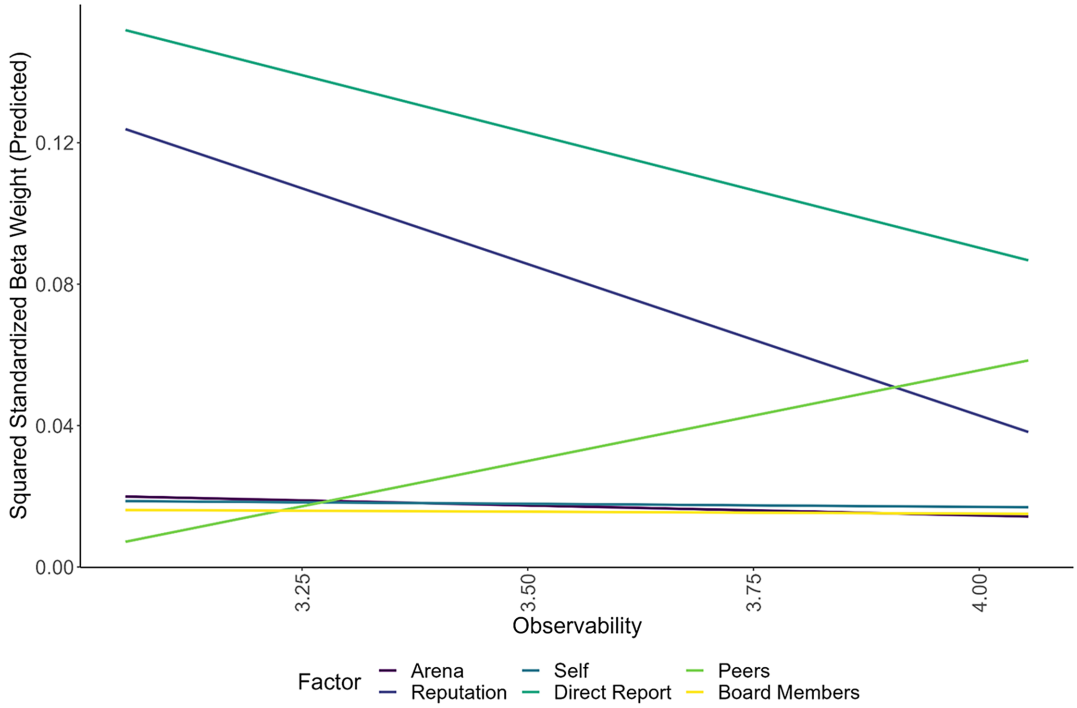
Outcome	Competency	Arena			Reputation			Identity			Board Member			Direct Report			Peer			
		β	*SE*	*p*	β	*SE*	*p*	β	*SE*	*p*	β	*SE*	*p*	β	*SE*	*p*	β	*SE*	*p*	R^2
Division performance	Business understanding	0.09	0.02	.07	**0.22**	**0.04**	**.03**	0.00	0.02	.94	−0.04	0.04	.69	**0.51**	**0.04**	**.00**	0.02	0.06	.88	0.32
	Forging synergy	0.06	0.02	.23	**0.42**	**0.04**	**.00**	−0.03	0.02	.51	−0.10	0.03	.28	**0.37**	**0.05**	**.02**	−0.14	0.03	.10	0.35
	Results orientation	**0.20**	**0.02**	**.00**	**0.34**	**0.03**	**.00**	0.03	0.02	.56	0.00	0.03	.98	**0.31**	**0.03**	**.00**	−0.01	0.03	.92	0.26
Psychological safety	Business understanding	**0.15**	**0.02**	**.00**	0.16	0.04	.11	0.02	0.02	.66	−0.05	0.04	.62	**0.37**	**0.04**	**.00**	−0.18	0.06	.31	0.22
	Forging synergy	**0.15**	**0.02**	**.00**	**0.41**	**0.04**	**.00**	−0.02	0.02	.73	0.00	0.03	.98	**0.39**	**0.06**	**.01**	−0.09	0.03	.31	0.35
	Results orientation	**0.20**	**0.02**	**.00**	**0.30**	**0.03**	**.00**	0.00	0.02	.96	0.06	0.03	.45	**0.24**	**0.03**	**.00**	−0.05	0.03	.56	0.20
Informal leadership	Business understanding	−0.06	0.01	.23	−0.06	0.01	.46	**−0.12**	**0.01**	**.03**	0.05	0.01	.53	**0.20**	**0.01**	**.03**	**0.46**	**0.04**	**.05**	0.28
	Forging synergy	−0.02	0.01	.68	0.14	0.01	.06	**−0.11**	**0.01**	**.05**	−0.02	0.01	.81	−0.03	0.02	.80	−0.05	0.01	.54	0.03
	Results orientation	−0.08	0.01	.13	0.06	0.01	.38	−0.05	0.01	.33	0.07	0.01	.34	**0.17**	**0.01**	**.02**	0.01	0.01	.90	0.05
Well-being	Business understanding	0.08	0.03	.15	0.06	0.05	.50	0.12	0.03	.03	**−0.28**	**0.06**	**.01**	−0.05	0.06	.58	0.15	0.11	.42	0.13
	Forging synergy	**0.16**	**0.03**	**.00**	−0.05	0.04	.51	**0.23**	**0.03**	**.00**	−0.13	0.06	.20	0.06	0.06	.53	0.00	0.05	.95	0.10
	Results orientation	0.07	0.03	.19	0.03	0.04	.70	**0.23**	**0.03**	**.00**	−0.15	0.06	.09	−0.07	0.04	.35	0.06	0.05	.43	0.09

Note. $n = 491$. Figures in bold are significant at p value of .05. Estimates represent predictive pathways between the LARI factors and measures of leaders' effectiveness. Coefficients are fully standardized. LARI (S-1) = leadership Arena–Reputation–Identity extension model; *SE* = standard error.

makes several assumptions that should be noted. Like all bifactor models, the LARI (S-1) model orthogonalizes the specific source factors (e.g., Identity, Board Members) from the general factors (e.g., Arena). Also, the LARI (S-1) model assumes there is a causal effect for the higher order factor (i.e., the joint influence of the shared factor and specific factors on item-level ratings; see Lang et al., 2010; Mulaik & Quartetti, 1997; Rindskopf & Rose, 1988). These assumptions are consistent with several frameworks within the literature (e.g., the Johari window) and receive initial support in this study but should be considered and evaluated as the literature continues to evolve.

Figure 5

Moderating Effects of Observability on Squared Standardized Regression Coefficients for Leader Effectiveness



Note. $n = 72$ parameter estimates from LARI (S-1) models. Predicted values were derived using Model 2 in Table 2 and assuming observability scores that ranged from ±1 *SD*. LARI (S-1) = leadership Arena–Reputation–Identity extension model. See the online article for the color version of this figure.

Conclusion

By replicating and extending the LARI model (Vergauwe et al., 2022), we find that this model generalizes to ratings of senior executives, specific sources in these ratings are not interchangeable, and each source affords unique predictive information. Our findings also point to a need to better integrate prevailing conceptual models of interpersonal perception and context-specific perspectives, which place an onus on the observability of the leadership dimension being rated, that is, the specific dimension of leadership being evaluated affects the predictive utility of the LARI model and, thus, merits further inquiry.

References

Atwater, L., Ostroff, C., Yammarino, F. J., & Fleenor, J. (1998). Self-other agreement: Does it really matter? *Personnel Psychology*, *51*(3), 577–598. https://doi.org/10.1111/j.1744-6570.1998.tb00252.x

Atwater, L., & Yammarino, F. J. (1992). Does self-other agreement on leadership perceptions moderate the validity of leadership and performance predictions? *Personnel Psychology*, *45*(1), 141–164. https://doi.org/10.1111/j.1744-6570.1992.tb00848.x

Banks, G. C., Fischer, T., Gooty, J., & Stock, G. (2021). Ethical leadership: Mapping the terrain for concept cleanup and a future research agenda. *The Leadership Quarterly*, *32*(2), Article 101471. https://doi.org/10.1016/j.leaqua.2020.101471

Bass, B. M., & Yammarino, F. J. (1991). Congruence of self and others' leadership ratings of naval officers for understanding successful performance. *Applied Psychology*, *40*(4), 437–454. https://doi.org/10.1111/j.1464-0597.1991.tb01002.x

Beekun, R. I., Stedham, Y., & Young, G. J. (1998). Board characteristics, managerial controls, and corporate strategy: A study of U.S. hospitals. *Journal of Management*, *24*(1), 3–19. https://doi.org/10.1177/014920639802400104

Borman, W. C. (1997). 360 degree ratings: An analysis of assumptions and a research agenda for evaluating their validity. *Human Resource Management Review*, *7*(3), 299–315. https://doi.org/10.1016/S1053-4822(97)90010-3

Bracken, D. W., Timmreck, C. W., & Church, A. H. (Eds.). (2001). *The handbook of multisource feedback*. Wiley.

Braddy, P. W., Gooty, J., Fleenor, J., & Yammarino, F. J. (2014). Leader behaviors and career derailment potential: A multi-analytic method examination of rating source and self–other agreement. *The Leadership Quarterly*, *25*(2), 373–390. https://doi.org/10.1016/j.leaqua.2013.10.001

Braun, M. T., Converse, P. D., & Oswald, F. L. (2019). The accuracy of dominance analysis as a metric to assess relative importance: The joint impact of sampling error variance and measurement unreliability. *Journal of Applied Psychology*, *104*(4), 593–602. https://doi.org/10.1037/apl0000361

Collins, L. M., Schafer, J. L., & Kam, C. M. (2001). A comparison of inclusive and restrictive strategies in modern missing data procedures. *Psychological Methods*, *6*(4), 330–351. https://doi.org/10.1037/1082-989X.6.4.330

Connelly, B. L., Certo, S. T., Ireland, R. D., & Reutzel, C. R. (2011). Signaling theory: A review and assessment. *Journal of Management*, *37*(1), 39–67. https://doi.org/10.1177/0149206310388419

Connelly, B. S., McAbee, S. T., Oh, I.-S., Jung, Y., & Jung, C.-W. (2022). A multirater perspective on personality and performance: An empirical examination of the trait-reputation-identity model. *Journal of Applied Psychology*, *107*(8), 1352–1368. https://doi.org/10.1037/apl0000732

DeRue, D. S. (2011). Adaptive leadership theory: Leading and following as a complex adaptive process. *Research in Organizational Behavior*, *31*, 125–150. https://doi.org/10.1016/j.riob.2011.09.007

Edmondson, A. (1999). Psychological safety and learning behavior in work teams. *Administrative Science Quarterly*, *44*(2), 350–383. https://doi.org/10.2307/2666999

Eid, M., Geiser, C., Koch, T., & Heene, M. (2017). Anomalous results in G-factor models: Explanations and alternatives. *Psychological Methods*, *22*(3), 541–562. https://doi.org/10.1037/met0000083

Eid, M., Krumm, S., Koch, T., & Schulze, J. (2018). Bifactor models for predicting criteria by general and specific factors: Problems of nonidentifiability and alternative solutions. *Journal of Intelligence*, *6*(3), Article 42. https://doi.org/10.3390/jintelligence6030042

Ferris, G. R., Blass, F. R., Douglas, C., Kolodinsky, R. W., & Treadway, D. C. (2003). Personal reputation in organizations. In J. Greenburg (Ed.), *Organizational behavior: A management challenge* (pp. 211–246). Lawrence Erlbaum.

Fleenor, J. W. (1996). Constructs and developmental assessment centers: Further troubling empirical findings. *Journal of Business and Psychology*, *10*(3), 319–335. https://doi.org/10.1007/BF02249606

Fleenor, J. W., Smither, J. W., Atwater, L., Braddy, P. W., & Sturm, R. E. (2010). Self–other rating agreement in leadership: A review. *The Leadership Quarterly*, *21*(6), 1005–1034. https://doi.org/10.1016/j.leaqua.2010.10.006

Fleenor, J. W., Taylor, S., & Chappelow, C. (2020). *Leveraging the impact of 360-degree feedback* (2nd ed.). Berrett-Koehler.

Frazier, M. L., Fainschmidt, S., Klinger, R. L., Pezeshkan, A., & Vracheva, V. (2017). Psychological safety: A meta-analytic review and extension. *Personnel Psychology*, *70*(1), 113–165. https://doi.org/10.1111/peps.12183

Gooty, J., & Yammarino, F. J. (2011). Dyads in organizational research: Conceptual issues and multilevel analyses. *Organizational Research Methods*, *14*(3), 456–483. https://doi.org/10.1177/1094428109358271

Haaland, S., & Christiansen, N. D. (2002). Implications of trait-activation theory for evaluating the construct validity of assessment center ratings. *Personnel Psychology*, *55*(1), 137–163. https://doi.org/10.1111/j.1744-6570.2002.tb00106.x

Heggestad, E. D., Scheaf, D. J., Banks, G. C., Monroe Hausfeld, M., Tonidandel, S., & Williams, E. B. (2019). Scale adaptation in organizational science research: A review and best-practice recommendations. *Journal of Management*, *45*(6), 2596–2627. https://doi.org/10.1177/0149206319850280

Heidemeier, H., & Moser, K. (2009). Self–other agreement in job performance ratings: A meta-analytic test of a process model. *Journal of Applied Psychology*, *94*(2), 353–370. https://doi.org/10.1037/0021-9010.94.2.353

Hill, C. W. L. (1990). Cooperation, opportunism, and the invisible hand: Implications for transaction cost theory. *Academy of Management Review*, *15*(3), 500–513. https://doi.org/10.2307/258020

Hoffman, B. J., & Woehr, D. J. (2009). Disentangling the meaning of multisource performance rating source and dimension factors. *Personnel Psychology*, *62*(4), 735–765. https://doi.org/10.1111/j.1744-6570.2009.01156.x

Hoogland, J. J., & Boomsma, A. (1998). Robustness studies in covariance structure modeling. *Sociological Methods & Research*, *26*(3), 329–367. https://doi.org/10.1177/0049124198026003003

Human, L. J., & Biesanz, J. C. (2011). Target adjustment and self–other agreement: Utilizing trait observability to disentangle judgeability and self-knowledge. *Journal of Personality and Social Psychology*, *101*(1), 202–216. https://doi.org/10.1037/a0023782

John, O. P., & Robins, R. W. (1993). Determinants of interjudge agreement on personality traits: The big five domains, observability, evaluativeness, and the unique perspective of the self. *Journal of Personality*, *61*(4), 521–551. https://doi.org/10.1111/j.1467-6494.1993.tb00781.x

Kenny, D. A. (1994). *Interpersonal perception: A social relations analysis*. Guilford Press.

Kenny, D. A. (2012). *Multiple latent variable models: Confirmatory factor analysis*. http://davidakenny.net/cm/mfactor.htm

Kohler, T., & Cortina, J. M. (2021). Play it again, Sam! An analysis of constructive replication in the organizational sciences. *Journal of Management*, *47*(2), 488–518. https://doi.org/10.1177/0149206319843985

Lance, C. E., Baranik, L. E., Lau, A. R., & Scharlau, E. A. (2010). If it ain't trait it must be method: (Mis) application of the multitrait–multimethod design in organizational research. In C. E. Lance & R. J. Vandenberg (Eds.), *Statistical and methodological myths and urban legends* (pp. 357–380). Routledge. https://doi.org/10.4324/9780203867266-23

Lance, C. E., Baxter, D., & Mahan, R. P. (2006). Multi-source performance measurement: A reconceptualization. In W. Bennett, C. E. Lance, & D. J. Woehr (Eds.), *Performance measurement: Current perspectives and future challenges* (pp. 49–76). Erlbaum.

Lang, J. W., Kersting, M., Hülsheger, U. R., & Lang, J. (2010). General mental ability, narrower cognitive abilities, and job performance: The perspective of the nested-factors model of cognitive abilities. *Personnel psychology*, *63*(3), 595–640. https://doi.org/10.1111/j.1744-6570.2010.01182.x

LeBreton, J. M., Burgess, J. R. D., Kaiser, R. B., Atchley, E. K., & James, L. R. (2003). The restriction of variance hypothesis and interrater reliability and agreement: Are ratings from multiple sources really dissimilar? *Organizational Research Methods*, *6*(1), 80–128. https://doi.org/10.1177/1094428102239427

LeBreton, J. M., & Senter, J. L. (2008). Answers to 20 questions about interrater reliability and interrater agreement. *Organizational Research Methods*, *11*(4), 815–852. https://doi.org/10.1177/1094428106296642

Lee, A., & Carpenter, N. C. (2018). Seeing eye to eye: A meta-analysis of self–other agreement of leadership. *The Leadership Quarterly*, *29*(2), 253–275. https://doi.org/10.1016/j.leaqua.2017.06.002

Leslie, J., Braddy, P. W., & Chappelow, C. (2015). *Benchmarks for executives technical manual*. Center for Creative Leadership.

Lievens, F., Schollaert, E., & Keen, G. (2015). The interplay of elicitation and evaluation of trait-expressive behavior: Evidence in assessment center exercises. *Journal of Applied Psychology*, *100*(4), 1169–1188. https://doi.org/10.1037/apl0000004

Loignon, A., & Wormington, S. (2022). *Psychologically safe for some, but not all?: The downsides of assuming shared psychological safety among senior leadership teams*. Center for Creative Leadership. https://doi.org/10.35613/ccl.2022.2048

Luft, J., & Ingram, H. (1955). *The Johari window: A graphic model of interpersonal awareness*. Proceedings of the western training laboratory in group development. University of California.

McAbee, S. T., & Connelly, B. S. (2016). A multi-rater framework for studying personality: The trait-reputation-identity model. *Psychological Review*, *123*(5), 569–591. https://doi.org/10.1037/rev0000035

McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods*, *1*(1), 30–46. https://doi.org/10.1037/1082-989X.1.1.30

McKee, R. A., Lee, Y.-T., Atwater, L., & Antonakis, J. (2018). Effects of personality and gender on self–other agreement in ratings of leadership. *Journal of Occupational and Organizational Psychology*, *91*(2), 285–315. https://doi.org/10.1111/joop.12209

Mulaik, S. A., & Quartetti, D. A. (1997). First order or higher order general factor? *Structural Equation Modeling: A Multidisciplinary Journal*, *4*(3), 193–211. https://doi.org/10.1080/10705519709540071

Newman, D. A. (2014). Missing data: Five practical guidelines. *Organizational Research Methods*, *17*(4), 372–411. https://doi.org/10.1177/1094428114548590

Oh, I.-S., Wang, G., & Mount, M. K. (2011). Validity of observer ratings of the five-factor model of personality traits: A meta-analysis. *Journal of Applied Psychology*, *96*(4), 762–773. https://doi.org/10.1037/a0021832

Pornprasertmanit, S., Miller, P., Schoemann, A., & Jorgensen, T. D. (2021). *simsem: SIMulated structural equation modeling R package*. https://CRAN.R-project.org/package=simsem

R Development Core Team. (2010). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research*, *47*(5), 667–696. https://doi.org/10.1080/00273171.2012.715555

Revelle, W. (2023). *psych: Procedures for psychological, psychometric, and personality research* (R package Version 2.3.9) Northwestern University. https://CRAN.R-project.org/package=psych

Rindskopf, D., & Rose, T. (1988). Some theory and applications of confirmatory second-order factor analysis. *Multivariate Behavioral Research*, *23*(1), 51–67. https://doi.org/10.1207/s15327906mbr2301_3

Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016). Applying bifactor statistical indices in the evaluation of psychological measures. *Journal of Personality Assessment*, *98*(3), 223–237. https://doi.org/10.1080/00223891.2015.1089249

Rönkkö, M., & Cho, E. (2022). An updated guideline for assessing discriminant validity. *Organizational Research Methods*, *25*(1), 6–14. https://doi.org/10.1177/1094428120968614

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

Schollaert, E., & Lievens, F. (2012). Building situational stimuli in assessment center exercises: Do specific exercise instructions and role-player prompts increase the observability of behavior? *Human Performance*, *25*(3), 255–271. https://doi.org/10.1080/08959285.2012.683907

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, *86*(2), 420–428. https://doi.org/10.1037/0033-2909.86.2.420

Slater, R., & Coyle, A. (2014). The governing of the self/the self-governing self: Multi-rater/source feedback and practices 1940–2011. *Theory & Psychology*, *24*(2), 233–255. https://doi.org/10.1177/0959354313520087

Spreitzer, G., Sutcliffe, K. M., Dutton, J. E., Sonenshein, S., & Grant, A. M. (2005). A socially embedded model of thriving at work. *Organization Science*, *16*(5), 537–549. https://doi.org/10.1287/orsc.1050.0153

Tay, L., Batz-Barbarich, C., Yang, L. Q., & Wiese, C. W. (2023). Well-being: The ultimate criterion for organizational sciences. *Journal of Business and Psychology*, *38*(6), 1141–1157. https://doi.org/10.1007/s10869-023-09908-5

Tett, R. P., Toich, M. J., & Ozkum, S. B. (2021). Trait activation theory: A review of the literature and applications to five lines of personality dynamics research. *Annual Review of Organizational Psychology and Organizational Behavior*, *8*(1), 199–233. https://doi.org/10.1146/annurev-orgpsych-012420-062228

Ulrich, D., Zenger, J., & Smallwood, N. (1999). *Results-based leadership*. Harvard Business School Press.

Vazire, S. (2010). Who knows what about a person? The self–other knowledge asymmetry (SOKA) model. *Journal of Personality and Social Psychology*, *98*(2), 281–300. https://doi.org/10.1037/a0017908

Vergauwe, J., Hofmans, J., & Wille, B. (2022). The leadership arena-reputation-identity (LARI) model: Distinguishing shared and unique perspectives in multisource leadership ratings. *Journal of Applied Psychology*, *107*(12), 2243–2268. https://doi.org/10.1037/apl0001012

Zhang, B., Sun, T., Cao, M., & Drasgow, F. (2021). Using bifactor models to examine the predictive validity of hierarchical constructs: Pros, cons, and solutions. *Organizational Research Methods*, *24*(3), 530–571. https://doi.org/10.1177/1094428120915522

(*Appendices follow*)

Appendix A

Summary of Psychometric Properties of Multisource Ratings of Leadership

Exploratory Factor Analysis

For this study, we used an adapted version of the original multisource leadership assessment (Leslie et al., 2015). Thus, we reassessed the psychometric properties of this adapted version (Heggestad et al., 2019). First, we conducted a parallel analysis using all available items within the assessment, which indicated that 12 factors exhibited eigenvalues that exceeded the amount obtained from randomly generated data (Revelle, 2023). We then submitted the entire set of items to an exploratory factor analysis using an orthogonal rotation and specified a 12-factor solution. After inspecting the factor loadings, we determined that three dimensions had sufficiently high loadings on the intended factors (>.60) and minimal cross loadings (<.25). The table below summarizes these loadings.

Confirmatory Factor Analyses

Next, within each rating source, we fit a one-factor confirmatory factor analysis model to each of the three dimensions. The table below reports the standardized factor loadings and model fit indices chi-square, *df*, comparative fit index, Tucker–Lewis index, root-mean-square error of approximation, standardized root-mean-square residual). These analyses provide evidence of the degree of unidimensionality of each subscale.

Table A1
Factor Loadings From Exploratory Factor Analysis

Item label	Item stem	Factor loading		
		Factor 1	Factor 2	Factor 3
RO.1	Assigns clear accountability for important objectives	0.36	0.61	0.24
RO.2	Clearly conveys objectives, deadlines, and expectations	0.32	0.64	0.22
RO.3	Acts with a sense of urgency	0.18	0.60	0.26
BU.1	Understands the strengths and weaknesses of major competitors	0.19	0.23	0.69
BU.2	Has a firm grasp of external conditions affecting the organization	0.25	0.29	0.67
BU.3	Stays informed about the strategic moves of major competitors	0.16	0.18	0.78
FS.1	Seeks common ground in an effort to resolve conflicts	0.71	0.22	0.16
FS.2	Works harmoniously with key stakeholders	0.76	0.16	0.18
FS.3	Maintains smooth, effective working relationships	0.76	0.20	0.19

Note. n = 5,150 ratings for 491 leaders. RO = results orientation; BU = business understanding; FS = forging synergy.

(*Appendices continue*)

Table A2
Factor Loadings From Unidimensional Confirmatory Factor Analysis

Item number	Source	Standardized factor loading		
		Results Orientation	Business Understanding	Forging Synergy
1	Board member	.78	.73	.71
2	Board member	.87	.69	.84
3	Board member	.78	.88	.87
1	Superior	.74	.83	.71
2	Superior	.89	.75	.87
3	Superior	.68	.83	.83
1	Direct report (1)	.38	.51	.42
2	Direct report (1)	.47	.49	.52
3	Direct report (1)	.45	.53	.52
1	Direct report (2)	.39	.50	.41
2	Direct report (2)	.45	.50	.53
3	Direct report (2)	.48	.53	.49
1	Direct report (3)	.38	.48	.42
2	Direct report (3)	.44	.49	.52
3	Direct report (3)	.48	.48	.50
1	Direct report (4)	.40	.50	.42
2	Direct report (4)	.45	.48	.55
3	Direct report (4)	.48	.50	.54
1	Peer (1)	.53	.56	.57
2	Peer (2)	.50	.45	.58
3	Peer (3)	.48	.61	.61
1	Peer (1)	.53	.55	.56
2	Peer (2)	.51	.44	.60
3	Peer (3)	.49	.57	.59
1	Peer (1)	.54	.53	.58
2	Peer (2)	.51	.43	.57
3	Peer (3)	.51	.56	.59
1	Self	.70	.81	.62
2	Self	.86	.62	.80
3	Self	.58	.86	.76

Note. $n = 491$ leaders. All factor loadings are significant at the $p < .001$. Confirmatory factor analysis models were run separately for each dimension and source. For board member, superior, and self-ratings, the models were saturated. For peers and direct reports, the models exhibited adequate fit: maximum $\chi^2 = 71.79$, minimum CFI = .98, minimum TLI = .99, maximum RMSEA = .04, maximum SRMR = .05. These models included correlated residuals among items from the same rater. Within each rating source, we then fit a three-correlated-factors model to better understand the structural validity of the scale. Each of these models fit the data well (minimum CFI = .97, minimum TLI = .95, maximum RMSEA = .07, maximum SRMR = .06). All the factor loadings were positive and statistically significant. We also found that the largest factor correlation was .73 (95% confidence interval [.63, .82]), which is less than traditional cutoff scores (e.g., Kenny, 2012) and significantly different from 1.00 (Rönkkö & Cho, 2022). This suggests that there is sufficient discriminant validity among the leadership dimensions with this adapted version of the measure. CFI = comparative fit index; TLI = Tucker–Lewis index; RMSEA = root-mean-square error of approximation; SRMR = standardized root-mean-square residual.

(*Appendices continue*)

Table A3

Factor Loadings and Factor Correlations From Multidimensional Confirmatory Factor Analysis

Item number	Source	Standardized factor loading: Results Orientation	Standardized factor loading: Business Understanding	Standardized factor loading: Forging Synergy		Factor correlations	
1	Board member	.86	.84	.83		RO	BU
2	Board member	.78	.73	.87	BU	0.42 (0.26, 0.58)	
3	Board member	.78	.73	.73	FS	0.66 (0.55, 0.77)	0.31 (0.14, 0.47)
1	Superior	.88	.81	.85		RO	BU
2	Superior	.75	.81	.84	BU	0.60 (0.52, 0.68)	
3	Superior	.68	.78	.73	FS	0.51 (0.42, 0.60)	0.33 (0.23, 0.44)
1	Direct report (1)	.42	.48	.49		RO	BU
2	Direct report (1)	.36	.47	.55	BU	0.70 (0.60, 0.79)	
3	Direct report (1)	.45	.46	.37	FS	0.48 (0.35, 0.61)	0.54 (0.40, 0.68)
1	Direct report (2)	.40	.58	.56			
2	Direct report (2)	.32	.61	.51			
3	Direct report (2)	.47	.57	.49			
1	Direct report (3)	.40	.43	.48			
2	Direct report (3)	.30	.43	.49			
3	Direct report (3)	.40	.39	.35			
1	Direct report (4)	.43	.45	.52			
2	Direct report (4)	.43	.46	.31			
3	Direct report (4)	.46	.48	.25			
1	Peer (1)	.53	.57	.58		RO	BU
2	Peer (2)	.53	.51	.62	BU	0.73 (0.63, 0.82)	
3	Peer (3)	.47	.49	.49	FS	0.65 (0.56, 0.75)	0.53 (0.42, 0.65)
1	Peer (1)	.50	.48	.55			
2	Peer (2)	.51	.49	.52			
3	Peer (3)	.45	.41	.54			
1	Peer (1)	.44	.66	.62			
2	Peer (2)	.56	.65	.60			
3	Peer (3)	.42	.47	.65			
1	Self	.79	.85	.78		RO	BU
2	Self	.75	.82	.78	BU	0.38 (0.28, 0.48)	
3	Self	.61	.64	.62	FS	0.47 (0.37, 0.57)	0.40 (0.30, 0.50)

Note. $n = 491$ leaders. All factor loadings are significant at the $p < .001$. Confirmatory factor analysis models were run separately for each source. Models featuring peers' and direct reports' ratings included correlated residuals among items from the same rater. Numbers in parentheses represent 95% confidence intervals for factor correlations. RO = results orientation; BU = business understanding; FS = forging synergy.

(*Appendices continue*)

Appendix B

Item-Level Descriptive Statistics and Correlations for Multisource Leadership Ratings

Variable	*M*	*SD*	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1. RO.1.B.1	3.85	.85	—																	
2. RO.1.BM.1	3.87	.84	.24	—																
3. RO.1.DR.1	3.77	.94	.23	.05	—															
4. RO.1.DR.2	3.84	.92	.12	−.05	.17	—														
5. RO.1.DR.3	3.8	.92	.16	.05	.16	.21	—													
6. RO.1.DR.4	3.91	.89	.20	.09	.16	.22	.22	—												
7. RO.1.P.1	3.84	.83	.25	.18	.15	.21	.07	.21	—											
8. RO.1.P.2	3.81	.85	.26	.12	.18	.16	.15	.22	.29	—										
9. RO.1.P.3	3.71	.85	.28	.21	.19	.20	.16	.26	.32	.36	—									
10. RO.1.S.1	3.54	.77	.13	.08	.08	.16	.21	.16	.08	.13	.16	—								
11. RO.2.B.1	3.95	.85	.66	.26	.19	.12	.13	.15	.25	.24	.24	.05	—							
12. RO.2.BM.1	4.02	.87	.25	.67	.11	.03	.02	.16	.32	.24	.35	.14	.29	—						
13. RO.2.DR.1	3.85	.92	.16	.18	.62	.19	.14	.22	.18	.12	.18	.08	.15	.08	—					
14. RO.2.DR.2	3.86	.96	.12	−.06	.14	.65	.24	.22	.20	.18	.20	.17	.13	.06	.19	—				
15. RO.2.DR.3	3.87	.98	.14	.04	.16	.22	.7	.27	.10	.20	.15	.15	.17	.07	.15	.31	—			
16. RO.2.DR.4	3.89	.96	.28	.08	.20	.21	.16	.67	.22	.19	.21	.19	.19	.10	.24	.24	.19	—		
17. RO.2.P.1	3.92	.87	.25	.21	.14	.21	.09	.23	.71	.27	.32	.08	.29	.35	.15	.19	.15	.22	—	
18. RO.2.P.2	3.9	.87	.27	.14	.19	.18	.18	.20	.29	.65	.29	.13	.26	.18	.14	.22	.20	.27	.28	—
19. RO.2.P.3	3.86	.86	.19	.23	.18	.19	.13	.14	.24	.29	.70	.18	.24	.27	.13	.18	.11	.17	.29	.25
20. RO.2.S.1	3.61	.82	.13	.21	.13	.16	.24	.08	.09	.14	.20	.60	.09	.17	.13	.19	.20	.18	.12	.14
21. RO.3.B.1	4.12	.84	.50	.08	.14	.11	.09	.12	.19	.22	.11	.01	.61	.13	.08	.11	.17	.07	.19	.21
22. RO.3.BM.1	4.12	.82	.09	.59	.05	.03	.04	.14	.17	.14	.19	.09	.16	.67	.01	.08	.01	.00	.17	.10
23. RO.3.DR.1	4.11	.85	.12	.06	.47	.09	.19	.16	.14	.07	.10	.09	.15	.06	.51	.11	.22	.20	.09	.14
24. RO.3.DR.2	4.19	.82	.14	.01	.19	.48	.19	.21	.24	.20	.11	.18	.17	.06	.22	.51	.18	.23	.20	.21
25. RO.3.DR.3	4.15	.81	.17	.17	.18	.14	.53	.13	.08	.17	.14	.20	.17	.10	.19	.26	.56	.15	.07	.17
26. RO.3.DR.4	4.21	.81	.25	.02	.10	.24	.12	.48	.08	.17	.25	.13	.20	.06	.18	.23	.18	.49	.10	.25
27. RO.3.P.1	3.99	.89	.21	.09	.09	.18	.02	.14	.62	.25	.22	−.01	.28	.13	.12	.17	.06	.14	.60	.29
28. RO.3.P.2	4.04	.88	.21	.07	.15	.12	.09	.19	.20	.61	.30	.12	.19	.24	.08	.13	.14	.20	.19	.60
29. RO.3.P.3	4.03	.84	.24	.28	.14	.20	.13	.14	.21	.27	.53	.18	.27	.21	.10	.24	.10	.16	.23	.24
30. RO.3.S.1	3.98	.77	.15	.15	.10	.20	.21	.19	.12	.12	.16	.41	.17	.19	.15	.20	.20	.21	.17	.19
31. BU.1.B.1	3.88	.82	.37	−.02	.10	.00	.08	.13	.15	.13	.03	.03	.39	.00	.05	.05	.13	.19	.12	.15
32. BU.1.BM.1	3.89	.80	.06	.11	.05	.14	.10	.01	.28	.12	.09	.10	.22	.21	.03	.13	.18	.07	.26	.09
33. BU.1.DR.1	4.08	.81	.12	−.01	.46	.01	.02	.09	.17	.16	−.01	.02	.14	.04	.42	.06	.08	.22	.12	.15
34. BU.1.DR.2	4.1	.86	.10	−.09	.18	.43	.14	.21	.22	.17	.21	.14	.14	.03	.17	.40	.20	.24	.17	.22
35. BU.1.DR.3	4.04	.87	.14	−.09	.10	.18	.48	.12	.07	.19	.06	.18	.15	.01	.11	.21	.47	.14	.08	.17
36. BU.1.DR.4	4.05	.83	.11	.00	.06	.17	.11	.42	.23	.17	.16	.05	.08	.03	.06	.22	.21	.51	.22	.19
37. BU.1.P.1	3.98	.81	.17	−.05	.02	.10	.08	.14	.51	.15	.12	.10	.15	.04	.02	.10	.12	.17	.42	.22
38. BU.1.P.2	3.98	.84	.16	.03	.15	.08	.13	.15	.24	.49	.15	.02	.14	.03	.10	.07	.17	.10	.21	.51
39. BU.1.P.3	3.96	.88	.11	.05	.11	.08	.08	.08	.31	.24	.41	.03	.10	.11	.05	.10	.17	.14	.23	.21
40. BU.1.S.1	3.54	.91	.07	−.01	.09	.03	.13	.12	.09	.07	.02	.26	.00	.03	.07	.08	.17	.13	.08	.13
41. BU.2.B.1	3.87	.82	.40	.05	.11	.05	−.02	.08	.18	.18	.15	.01	.47	.07	.05	.07	.07	.11	.16	.16
42. BU.2.BM.1	3.97	.82	.02	.32	.01	−.04	.01	.06	.18	.06	.10	−.04	.12	.39	.01	.08	.10	.08	.20	.11
43. BU.2.DR.1	4.05	.83	.11	.09	.48	.12	.10	.14	.20	.18	.10	.02	.11	.07	.51	.09	.11	.17	.18	.18
44. BU.2.DR.2	4.11	.83	.18	−.04	.20	.48	.11	.13	.20	.15	.17	.06	.19	.09	.20	.49	.15	.20	.19	.21
45. BU.2.DR.3	4.1	.81	.15	−.02	.17	.13	.51	.19	.09	.22	.10	.12	.19	.05	.17	.25	.52	.19	.08	.17
46. BU.2.DR.4	4.06	.85	.11	−.02	.11	.20	.11	.42	.15	.21	.09	.08	.12	−.05	.11	.19	.18	.50	.11	.21
47. BU.2.P.1	3.96	.83	.20	.00	.12	.15	.08	.16	.55	.20	.14	.09	.22	.04	.12	.15	.17	.24	.49	.24
48. BU.2.P.2	3.96	.84	.13	−.10	.10	.12	.11	.17	.23	.57	.17	.04	.10	−.05	.07	.10	.14	.12	.14	.56

(*Appendices continue*)

Appendix B (*continued*)

Variable	*M*	*SD*	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
49. BU.2.P.3	3.91	.85	.12	.14	.09	.13	.12	.09	.29	.21	.46	.00	.12	.19	.03	.15	.15	.12	.22	.16
50. BU.2.S.1	3.62	.77	.03	−.02	.06	.05	.15	.03	.03	.02	.12	.23	.01	.02	.04	.08	.13	.07	.06	.03
51. BU.3.B.1	3.76	.88	.36	.06	.14	.02	.11	.06	.12	.18	.07	.05	.40	.09	.07	.10	.22	.11	.11	.16
52. BU.3.BM.1	3.9	.82	.05	.24	.07	.09	.06	−.08	.30	.12	.23	.03	.15	.18	.05	.08	.09	.09	.30	.12
53. BU.3.DR.1	4.07	.84	.07	.08	.46	.09	.08	.14	.20	.17	.03	.04	.10	.11	.44	.10	.11	.26	.19	.22
54. BU.3.DR.2	4.1	.84	.08	−.05	.15	.41	.16	.15	.29	.22	.17	.10	.12	.07	.15	.37	.21	.20	.29	.31
55. BU.3.DR.3	4.04	.91	.12	.03	.10	.17	.44	.11	.09	.19	.05	.18	.17	.09	.11	.20	.43	.11	.07	.21
56. BU.3.DR.4	4.05	.88	.13	.00	.10	.15	.08	.34	.11	.11	.10	.07	.08	.01	.12	.22	.18	.46	.16	.17
57. BU.3.P.1	3.96	.86	.16	.06	.12	.12	.08	.22	.53	.17	.12	.07	.17	.15	.11	.12	.11	.23	.46	.24
58. BU.3.P.2	3.96	.91	.13	.02	.10	.05	.06	.13	.18	.47	.16	.08	.16	.06	.04	.05	.10	.10	.15	.39
59. BU.3.P.3	3.85	.93	.16	.12	.13	.11	.21	.18	.33	.24	.42	.14	.08	.22	.07	.13	.24	.17	.29	.23
60. BU.3.S.1	3.3	.94	.04	−.07	.08	.03	.11	.07	.12	.07	.02	.26	.02	.04	.00	.07	.11	.10	.10	.11
61. FS.1.B.1	3.81	.84	.40	.09	.10	.11	.04	.15	.17	.18	.17	.04	.39	.05	.11	.08	.01	.23	.16	.19
62. FS.1.BM.1	3.84	.73	−.01	.43	.08	.03	.08	.25	.20	.19	.20	.14	.14	.48	.14	.09	.01	.16	.22	.08
63. FS.1.DR.1	3.76	.88	.15	.05	.52	.16	.04	.14	.19	.14	.14	.03	.13	.14	.47	.07	.07	.21	.12	.14
64. FS.1.DR.2	3.85	.94	.08	.05	.11	.50	.19	.22	.19	.11	.21	.07	.12	.13	.21	.48	.21	.23	.15	.17
65. FS.1.DR.3	3.84	.87	.17	.15	.12	.15	.51	.16	.17	.19	.11	.10	.12	.16	.15	.25	.52	.18	.13	.15
66. FS.1.DR.4	3.83	.89	.10	.08	.11	.16	.12	.49	.13	.07	.10	.01	.10	.01	.16	.11	.11	.53	.12	.15
67. FS.1.P.1	3.76	.91	.08	.01	−.02	.09	.00	.15	.49	.17	.18	.01	.12	.12	.00	.06	.01	.11	.43	.20
68. FS.1.P.2	3.68	.95	.17	.15	.08	.13	.04	.21	.26	.54	.25	.05	.16	.12	.03	.09	.10	.20	.18	.48
69. FS.1.P.3	3.65	.95	.09	.17	.05	.14	.03	.16	.26	.25	.56	.05	.13	.20	.09	.17	.04	.15	.23	.20
70. FS.1.S.1	3.66	.77	.06	−.03	.00	.06	.02	.19	.07	.07	.08	.28	.05	−.01	.03	.11	.00	.20	.07	.14
71. FS.2.B.1	3.91	.89	.29	.21	.10	.15	.02	.18	.14	.16	.22	−.02	.36	.07	.09	.08	.05	.14	.19	.12
72. FS.2.BM.1	3.95	.86	.08	.38	.04	−.01	.00	.17	.12	.00	.18	.06	.11	.46	.07	−.07	−.10	.13	.15	−.04
73. FS.2.DR.1	3.96	.90	.02	.15	.37	.17	.03	.17	.20	.18	.18	.02	.07	.18	.38	.07	.06	.21	.18	.11
74. FS.2.DR.2	4.01	.91	.13	−.02	.10	.51	.11	.24	.19	.14	.20	−.02	.13	.10	.13	.50	.17	.22	.21	.15
75. FS.2.DR.3	4.04	.89	.12	−.01	.14	.20	.47	.19	.15	.18	.11	.01	.14	.07	.10	.19	.51	.21	.15	.14
76. FS.2.DR.4	4.06	.87	.12	.12	.05	.18	.04	.44	.14	.16	.18	.03	.12	.04	.13	.13	.08	.43	.12	.20
77. FS.2.P.1	3.84	.98	.05	.04	−.06	.11	−.02	.18	.42	.16	.23	−.02	.08	.13	.01	.08	−.01	.17	.44	.22
78. FS.2.P.2	3.85	.95	.08	.11	.04	.14	.05	.21	.20	.52	.25	.05	.12	.08	.04	.13	.08	.15	.15	.48
79. FS.2.P.3	3.74	1.02	.06	.22	.09	.16	.05	.21	.25	.25	.49	.04	.09	.25	.06	.13	.09	.20	.30	.16
80. FS.2.S.1	3.68	.79	.01	.01	.02	.07	.09	.09	.15	.11	.15	.28	.07	−.01	0	.07	.03	.12	.11	.11
81. FS.3.B.1	3.94	.84	.36	.06	.12	.13	.09	.22	.17	.23	.30	.06	.41	.04	.15	.09	.07	.23	.23	.16
82. FS.3.BM.1	4.04	.84	−.02	.45	.09	.10	−.09	.20	.18	.02	.18	.06	.05	.50	.12	.02	−.12	.11	.19	.04
83. FS.3.DR.1	3.97	.88	.06	.07	.48	.17	.06	.24	.21	.16	.18	.02	.08	.14	.48	.13	.09	.28	.16	.13
84. FS.3.DR.2	4.03	.93	.11	.00	.14	.53	.16	.24	.18	.17	.20	.06	.14	.10	.16	.54	.18	.27	.17	.19
85. FS.3.DR.3	4.02	.90	.16	.09	.11	.18	.51	.24	.11	.20	.10	.05	.11	.10	.08	.20	.54	.25	.12	.16
86. FS.3.DR.4	4.06	.81	.12	.05	.14	.22	.00	.51	.15	.16	.17	.04	.13	.07	.16	.20	.07	.56	.19	.20
87. FS.3.P.1	3.93	.91	.12	.13	−.01	.14	.04	.23	.53	.22	.27	−.01	.20	.21	.05	.11	.06	.17	.52	.25
88. FS.3.P.2	3.87	.91	.20	.09	.13	.15	.11	.26	.25	.54	.28	.06	.18	.07	.11	.14	.16	.21	.19	.54
89. FS.3.P.3	3.81	.93	.10	.18	.13	.20	.08	.11	.25	.27	.53	.06	.14	.18	.12	.18	.10	.15	.31	.21
90. FS.3.S.1	3.77	.65	.02	.04	.02	.15	.07	.13	.17	.13	.17	.34	.03	.10	.07	.16	.12	.17	.12	.08
91. Div. perf.	4.14	.36	.14	.09	.24	.28	.24	.28	.17	.18	.20	.13	.17	.13	.22	.25	.28	.26	.17	.11
92. Psych. safe.	4.09	.36	.12	.05	.16	.19	.20	.30	.13	.16	.19	.11	.17	.22	.14	.18	.22	.26	.10	.12
93. Inform. inf.	.57	.16	−.09	.13	−.02	.09	.06	.01	.02	−.05	.02	−.06	−.08	.07	.08	.03	.09	.02	−.02	.00
94. Well-being	4.19	.63	.08	.03	−.02	.00	.09	.03	.01	.10	.08	.19	.00	−.04	.00	.03	.06	.08	.05	.10
95. Male	.75	.43	−.03	−.04	−.06	.00	.01	.03	−.01	−.11	.03	.02	−.13	−.15	.02	.04	−.02	.06	.01	−.06
96. Org. tenure	13.43	1.85	.19	−.01	.10	.10	.06	.12	.16	.07	.22	.02	.18	.02	.13	.08	.00	.12	.15	.12
97. Team. Dist.	53.58	34.06	.04	−.01	.06	.05	.08	.05	.01	.04	.01	−.03	.07	.07	.05	.02	.09	.07	−.03	.04

(*Appendices continue*)

Appendix B (*continued*)

Variable	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38
20. RO.2.S.1	.22	—																		
21. RO.3.B.1	.08	.07	—																	
22. RO.3.BM.1	.13	.13	.16	—																
23. RO.3.DR.1	.06	.14	.17	.06	—															
24. RO.3.DR.2	.14	.18	.22	.12	.23	—														
25. RO.3.DR.3	.06	.26	.16	.12	.27	.25	—													
26. RO.3.DR.4	.09	.21	.19	.19	.16	.28	.24	—												
27. RO.3.P.1	.20	.04	.26	.09	.09	.25	.15	.14	—											
28. RO.3.P.2	.27	.15	.22	.18	.13	.15	.19	.23	.29	—										
29. RO.3.P.3	.55	.20	.27	.25	.12	.22	.17	.19	.25	.30	—									
30. RO.3.S.1	.16	.50	.20	.22	.18	.26	.25	.30	.18	.20	.28	—								
31. BU.1.B.1	.06	.07	.34	.00	.10	.14	.16	.26	.16	.16	.19	.16	—							
32. BU.1.BM.1	.01	.02	.24	.23	.14	.25	.11	.10	.16	.10	.05	.14	.16	—						
33. BU.1.DR.1	.00	.04	.16	.04	.47	.15	.07	.07	.13	.13	.09	.11	.23	.11	—					
34. BU.1.DR.2	.09	.10	.16	.04	.18	.48	.12	.23	.23	.13	.18	.16	.25	.33	.27	—				
35. BU.1.DR.3	.04	.18	.13	−.05	.17	.19	.40	.13	.10	.19	.08	.18	.28	.21	.22	.33	—			
36. BU.1.DR.4	.06	−.01	.06	.04	.06	.14	.15	.41	.14	.23	.09	.12	.22	.12	.18	.34	.21	—		
37. BU.1.P.1	.06	.03	.13	−.02	.11	.09	.07	.11	.41	.12	.11	.04	.29	.25	.18	.18	.14	.17	—	
38. BU.1.P.2	.16	.03	.13	−.04	.10	.08	.09	.14	.25	.48	.14	.09	.25	.02	.18	.20	.23	.20	.28	—
39. BU.1.P.3	.44	−.02	.07	−.02	.05	.12	.08	.18	.22	.23	.39	.09	.27	.13	.15	.25	.12	.33	.32	.30
40. BU.1.S.1	−.01	.23	.01	.06	.08	.10	.15	.19	.09	.13	.05	.29	.33	.13	.22	.29	.33	.27	.23	.24
41. BU.2.B.1	.18	−.01	.40	.05	.12	.11	.09	.13	.21	.13	.20	.08	.62	.16	.17	.15	.12	.12	.19	.21
42. BU.2.BM.1	.14	−.06	.08	.39	.12	.20	.08	.11	.09	.12	.11	.06	.05	.51	.10	.15	.07	.20	.14	.05
43. BU.2.DR.1	.11	.06	.12	.08	.51	.20	.15	.09	.13	.13	.18	.12	.17	.13	.66	.29	.18	.18	.17	.19
44. BU.2.DR.2	.14	.04	.14	.11	.20	.49	.07	.19	.25	.16	.15	.16	.18	.19	.28	.71	.27	.29	.21	.18
45. BU.2.DR.3	.09	.14	.13	.05	.19	.17	.44	.14	.09	.21	.14	.13	.29	.11	.19	.22	.67	.17	.14	.20
46. BU.2.DR.4	−.01	.03	.08	−.03	.11	.18	.23	.46	.18	.22	.11	.17	.25	.13	.18	.32	.18	.65	.24	.22
47. BU.2.P.1	.10	.01	.13	.06	.13	.11	.06	.16	.42	.16	.13	.10	.29	.16	.19	.19	.16	.26	.64	.25
48. BU.2.P.2	.18	.07	.11	−.08	.09	.14	.13	.18	.25	.53	.18	.10	.20	.05	.12	.15	.22	.16	.22	.66
49. BU.2.P.3	.50	.03	.06	.00	.08	.07	.05	.08	.19	.20	.49	.12	.19	−.02	.05	.08	.04	.16	.21	.23
50. BU.2.S.1	.09	.19	−.06	.09	.04	.04	.11	.06	.06	.05	.07	.30	.16	.17	.10	.22	.16	.14	.17	.15
51. BU.3.B.1	.13	.13	.33	.06	.11	.14	.16	.24	.15	.19	.13	.18	.69	.32	.21	.24	.35	.17	.23	.29
52. BU.3.BM.1	.20	.08	.12	.22	.13	.18	.15	.05	.20	.11	.09	.11	.05	.63	.14	.20	.15	.19	.23	.14
53. BU.3.DR.1	.06	.07	.08	.09	.52	.18	.13	.10	.16	.15	.14	.13	.20	.09	.68	.30	.17	.24	.18	.20
54. BU.3.DR.2	.10	.09	.13	.06	.14	.43	.13	.24	.33	.23	.18	.14	.20	.27	.28	.71	.33	.32	.28	.30
55. BU.3.DR.3	.03	.18	.12	.08	.12	.17	.41	.13	.13	.22	.08	.21	.31	.30	.20	.31	.70	.21	.24	.34
56. BU.3.DR.4	.01	.03	.09	−.06	.10	.16	.14	.41	.08	.19	.11	.19	.28	.08	.20	.30	.19	.75	.16	.18
57. BU.3.P.1	.10	.01	.10	.01	.13	.14	.10	.14	.44	.17	.13	.11	.39	.20	.22	.21	.20	.26	.74	.31
58. BU.3.P.2	.13	.08	.17	−.03	.12	.04	.13	.17	.23	.45	.16	.07	.31	.09	.18	.18	.26	.15	.32	.72
59. BU.3.P.3	.41	.08	.03	.13	.09	.11	.12	.14	.20	.24	.42	.15	.31	.11	.16	.25	.19	.30	.33	.31
60. BU.3.S.1	.00	.15	−.01	.04	.11	.13	.12	.12	.09	.11	.04	.25	.30	.23	.21	.33	.24	.25	.28	.25
61. FS.1.B.1	.11	.01	.20	−.09	.01	.10	−.02	.21	.11	.10	.08	−.03	.28	.07	.05	.09	.06	.09	.15	.12
62. FS.1.BM.1	.12	.05	−.07	.49	.09	.10	.09	.13	.08	.17	.18	.01	.04	.19	.14	.20	.16	.10	.12	.02
63. FS.1.DR.1	.15	.00	.04	.00	.43	.13	.03	.06	.09	.09	.07	.04	.01	.09	.37	.07	.05	.04	.06	.12
64. FS.1.DR.2	.10	.04	.02	.08	.09	.42	.11	.16	.14	.07	.09	.05	.01	.22	.03	.41	.21	.17	.04	.12
65. FS.1.DR.3	.07	.14	.06	.01	.20	.12	.39	.03	.05	.14	.03	.08	.13	.07	.01	.12	.43	.10	.10	.15
66. FS.1.DR.4	.04	−.05	−.02	−.07	.02	.09	.09	.37	.13	.05	−.02	.04	.01	−.02	.02	.12	.03	.42	.16	.07
67. FS.1.P.1	.12	−.01	.01	−.01	.00	.04	−.05	.05	.33	.18	.04	−.03	.08	.15	.00	.11	.02	.19	.34	.19
68. FS.1.P.2	.18	.00	.06	.06	.04	.06	−.01	.13	.14	.45	.13	.02	.11	−.03	.04	.09	.03	.20	.12	.42
69. FS.1.P.3	.48	−.02	−.03	.04	.02	.06	−.03	.08	.14	.20	.37	−.04	−.01	−.03	−.04	.04	−.01	.14	.15	.12
70. FS.1.S.1	.13	.18	−.1	−.03	.04	.09	−.02	.08	−.02	.08	.04	.17	.05	−.04	.02	.08	.07	.12	.09	.11
71. FS.2.B.1	.11	−.07	.17	−.04	.05	.05	.02	.14	.15	.06	.10	−.02	.17	.05	.03	.05	.03	.09	.12	.13

(*Appendices continue*)

Appendix B (*continued*)

Variable	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38
72. FS.2.BM.1	.09	.01	−.11	.35	.01	−.02	−.07	.17	−.04	.06	.06	−.02	−.04	.03	.03	.06	−.03	.15	−.01	−.02
73. FS.2.DR.1	.12	.02	−.09	.02	.27	.10	−.03	.02	.11	.08	.03	−.01	.00	.08	.28	.11	.05	.15	.08	.15
74. FS.2.DR.2	.12	.00	.04	.07	.06	.35	.03	.16	.15	.09	.11	.00	.07	.10	.01	.39	.16	.18	.14	.10
75. FS.2.DR.3	.05	.09	.08	.03	.16	.11	.37	.14	.09	.19	.05	.06	.13	.07	.07	.15	.42	.14	.12	.14
76. FS.2.DR.4	.06	−.02	−.09	−.12	.06	.08	−.01	.32	.10	.10	.03	.02	.00	−.02	.02	.10	.02	.35	.15	.13
77. FS.2.P.1	.16	−.08	−.03	−.02	−.03	.01	−.10	.09	.29	.09	.05	−.08	.04	.00	.05	.06	.01	.21	.36	.13
78. FS.2.P.2	.16	.00	.03	.03	.02	.12	−.04	.08	.10	.41	.15	−.01	.06	−.02	.04	.08	.09	.14	.16	.37
79. FS.2.P.3	.43	−.01	−.02	.06	.03	.08	−.04	.05	.15	.16	.37	−.03	−.03	.05	−.04	.08	.10	.14	.12	.12
80. FS.2.S.1	.13	.24	−.10	.04	−.01	.12	.04	.08	.02	.08	.03	.19	−.01	.07	−.01	.11	.10	.05	.04	.10
81. FS.3.B.1	.14	−.02	.20	−.02	.11	.11	.07	.19	.14	.10	.12	.00	.18	.05	.00	.13	.10	.13	.14	.12
82. FS.3.BM.1	.13	.06	−.17	.45	.10	.14	−.06	.10	.06	.06	.06	.05	−.16	.13	−.01	.06	−.02	.08	.01	−.17
83. FS.3.DR.1	.14	.04	−.07	−.02	.41	.15	.04	.12	.12	.13	.05	.01	.00	−.02	.39	.19	.09	.17	.08	.13
84. FS.3.DR.2	.13	.08	.03	.07	.11	.43	.11	.19	.09	.12	.13	.03	.02	.18	.04	.43	.19	.15	.09	.09
85. FS.3.DR.3	.04	.07	.06	.00	.12	.12	.36	.11	.06	.16	.02	.06	.12	.12	.01	.14	.46	.17	.10	.17
86. FS.3.DR.4	.13	−.02	.00	.06	.05	.12	.04	.43	.17	.13	.09	.08	.05	.08	.15	.15	.02	.50	.11	.13
87. FS.3.P.1	.19	−.03	.04	.14	.01	.08	−.05	.13	.36	.14	.10	.01	.08	.09	.03	.10	.05	.27	.37	.18
88. FS.3.P.2	.21	.02	.09	−.05	.08	.13	.01	.09	.17	.43	.14	.02	.07	.03	.09	.12	.12	.15	.19	.47
89. FS.3.P.3	.50	.05	.01	.00	.01	.13	−.01	.04	.16	.18	.39	.02	.09	−.02	−.02	.13	.04	.15	.10	.14
90. FS.3.S.1	.15	.28	−.10	−.01	.01	.16	.04	.11	.07	.05	.04	.27	−.03	.06	−.01	.17	.16	.19	.07	.09
91. Div. perf.	.13	.14	.07	.11	.12	.17	.21	.23	.16	.07	.15	.12	.02	.17	.13	.19	.17	.21	.12	.09
92. Psych. safe.	.13	.10	.07	.09	.11	.15	.17	.19	.11	.08	.11	.07	.11	.14	.07	.17	.15	.19	.11	.05
93. Inform. inf.	.07	−.01	−.05	−.04	.13	.03	.03	−.02	−.01	−.04	.02	.01	−.04	.00	.02	.02	−.06	−.07	.01	.00
94. Well-being	.10	.18	.02	−.12	.00	.01	.06	.10	.05	.07	.03	.14	.02	−.13	.01	.06	.10	.01	.06	.12
95. Male	.02	.03	−.07	−.16	−.02	−.08	−.04	.02	−.01	−.02	.09	−.01	.04	.05	.00	.04	.00	.10	.10	.06
96. Org. tenure	.17	−.04	.16	−.03	.09	.03	.06	.16	.18	.11	.19	−.03	.18	.10	.19	.14	.14	.19	.17	.17
97. Team. Dist.	.09	−.04	−.01	.04	.08	−.03	.08	.15	.01	.01	.05	−.04	.01	.07	.08	−.06	.10	.00	.06	.07

Variable	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58
40. BU.1.S.1	.25	—																		
41. BU.2.B.1	.21	.12	—																	
42. BU.2.BM.1	.13	.11	.19	—																
43. BU.2.DR.1	.14	.19	.12	.18	—															
44. BU.2.DR.2	.19	.18	.14	.18	.31	—														
45. BU.2.DR.3	.12	.22	.16	.07	.23	.22	—													
46. BU.2.DR.4	.25	.26	.08	.18	.18	.31	.23	—												
47. BU.2.P.1	.34	.23	.24	.17	.23	.23	.13	.25	—											
48. BU.2.P.2	.22	.20	.17	.12	.18	.16	.22	.21	.22	—										
49. BU.2.P.3	.63	.08	.25	.14	.13	.16	.09	.16	.21	.24	—									
50. BU.2.S.1	.23	.51	.08	.19	.17	.16	.21	.20	.17	.18	.13	—								
51. BU.3.B.1	.21	.28	.62	.22	.11	.18	.28	.14	.22	.21	.18	.19	—							
52. BU.3.BM.1	.20	.17	.08	.58	.19	.15	.14	.17	.21	.22	.10	.20	.13	—						
53. BU.3.DR.1	.20	.24	.13	.13	.68	.28	.21	.26	.21	.16	.08	.13	.13	.14	—					
54. BU.3.DR.2	.24	.26	.14	.21	.25	.64	.18	.30	.25	.25	.08	.19	.25	.27	.32	—				
55. BU.3.DR.3	.18	.30	.17	.16	.20	.24	.68	.18	.16	.28	.07	.22	.35	.35	.23	.33	—			
56. BU.3.DR.4	.27	.35	.12	.21	.17	.27	.20	.71	.21	.13	.09	.20	.26	.27	.29	.33	.19	—		
57. BU.3.P.1	.38	.32	.20	.19	.23	.22	.20	.24	.67	.28	.28	.20	.32	.31	.23	.33	.26	.27	—	
58. BU.3.P.2	.28	.24	.27	.16	.18	.17	.21	.19	.20	.68	.18	.19	.29	.27	.19	.31	.30	.13	.37	—
59. BU.3.P.3	.70	.33	.19	.26	.21	.19	.20	.24	.34	.24	.65	.29	.26	.19	.26	.25	.26	.26	.39	.27
60. BU.3.S.1	.28	.70	.08	.17	.19	.21	.16	.25	.23	.22	.12	.54	.31	.23	.27	.30	.30	.30	.36	.30
61. FS.1.B.1	.08	.03	.29	.00	.08	.10	.11	.15	.17	.09	.11	.03	.17	.07	.01	.09	.11	.03	.14	.10

(*Appendices continue*)

Appendix B (*continued*)

Variable	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58
62. FS.1.BM.1	.10	.10	−.01	.40	.21	.21	.11	.18	.11	.04	.10	.06	.03	.23	.09	.21	.19	.01	.14	.09
63. FS.1.DR.1	.07	.00	.06	.07	.43	.13	.08	.11	.17	.09	.06	.00	.01	.04	.39	.14	.06	.06	.13	.10
64. FS.1.DR.2	.10	.05	.05	.12	.12	.43	.18	.19	.07	.12	.08	.06	.00	.20	.07	.38	.24	.09	.05	.07
65. FS.1.DR.3	.04	.10	.05	.06	.14	.14	.48	.12	.18	.17	.13	.03	.11	.01	.11	.16	.37	.07	.20	.07
66. FS.1.DR.4	.07	.04	.05	.06	.06	.08	.07	.42	.19	.07	−.01	−.01	.00	.02	.06	.18	.02	.35	.13	.06
67. FS.1.P.1	.21	.04	.08	.16	.02	.10	.03	.11	.38	.15	.10	.05	.03	.29	.03	.17	.07	.13	.38	.15
68. FS.1.P.2	.15	.07	.11	.07	.11	.08	.09	.23	.19	.43	.15	.07	−.01	.05	.08	.16	.07	.12	.12	.35
69. FS.1.P.3	.37	−.03	.12	.14	.07	.09	.02	.11	.19	.15	.47	.00	−.05	.13	.02	.12	.03	.10	.16	.12
70. FS.1.S.1	.06	.21	.06	.02	.03	.03	.06	.17	.12	.06	.03	.23	.05	.07	.09	.09	.06	.17	.11	.01
71. FS.2.B.1	.10	.00	.28	.01	.06	.09	.10	.13	.13	.05	.10	.05	.18	.13	.03	.10	.09	.03	.07	.10
72. FS.2.BM.1	.06	.06	−.03	.22	.10	−.01	.00	.09	.07	−.06	.07	.02	.01	.07	.14	.03	.00	.11	.01	.04
73. FS.2.DR.1	.11	.04	.07	.08	.36	.11	.06	.15	.17	.05	.08	.03	−.01	.04	.32	.10	.02	.08	.09	.10
74. FS.2.DR.2	.12	.08	.11	.10	.08	.47	.14	.27	.18	.10	.10	.09	.04	.14	.06	.43	.15	.17	.14	.07
75. FS.2.DR.3	.01	.07	.06	.03	.14	.15	.52	.24	.20	.12	.07	.04	.12	−.01	.13	.19	.42	.11	.14	.09
76. FS.2.DR.4	.03	.04	.02	.00	.07	.13	.04	.42	.17	.14	.10	−.01	.02	−.07	.05	.12	−.01	.31	.12	.11
77. FS.2.P.1	.25	.03	.06	.03	.05	.09	.01	.19	.36	.13	.17	.05	−.05	.13	.12	.14	.07	.13	.37	.13
78. FS.2.P.2	.20	0	.02	−.06	.07	.11	.12	.19	.18	.42	.17	.06	.02	−.01	.06	.18	.14	.1	.16	.32
79. FS.2.P.3	.37	.02	.08	.19	.05	.14	.09	.09	.20	.12	.44	.01	−.03	.14	.04	.14	.06	.09	.14	.15
80. FS.2.S.1	.06	.20	.01	.04	.00	.06	.06	.11	.07	.12	.07	.28	.04	.07	.05	.13	.12	.04	.08	.05
81. FS.3.B.1	.10	−.03	.29	.03	.08	.11	.15	.09	.18	.11	.14	.03	.18	.11	.04	.12	.13	.05	.10	.09
82. FS.3.BM.1	.01	.01	−.03	.34	.12	.10	−.08	.01	.13	−.11	−.01	.04	−.08	.12	.09	.11	−.05	.00	.00	.01
83. FS.3.DR.1	.08	.01	.02	.07	.48	.24	.11	.14	.22	.07	.11	.02	.00	.00	.45	.20	.09	.10	.12	.09
84. FS.3.DR.2	.05	.00	.08	.08	.12	.48	.17	.23	.12	.11	.10	.00	.02	.15	.08	.43	.16	.14	.08	.08
85. FS.3.DR.3	.01	.04	.05	.10	.10	.17	.53	.21	.14	.14	.07	.03	.17	.08	.06	.17	.46	.13	.17	.06
86. FS.3.DR.4	.15	.13	.05	.03	.15	.15	.05	.47	.20	.11	.11	.01	.02	.02	.19	.22	.01	.41	.15	.09
87. FS.3.P.1	.26	.08	.10	.12	.08	.13	.04	.18	.44	.14	.23	.05	.03	.14	.07	.18	.06	.21	.44	.14
88. FS.3.P.2	.18	.02	.12	−.02	.10	.13	.15	.15	.18	.46	.16	.02	.05	.07	.13	.25	.17	.08	.18	.42
89. FS.3.P.3	.39	.02	.17	.15	.06	.20	.07	.08	.22	.15	.51	.05	.05	.11	.02	.21	.03	.08	.17	.13
90. FS.3.S.1	.10	.21	−.01	.11	.05	.12	.11	.25	.14	.07	.05	.28	.04	.07	.05	.15	.12	.19	.11	.02
91. Div. perf.	.16	.13	.07	.09	.20	.24	.24	.32	.13	.11	.11	.14	.08	−.01	.24	.21	.16	.31	.14	.08
92. Psych. safe.	.19	.10	.13	.06	.12	.16	.24	.30	.11	.10	.16	.14	.08	−.06	.18	.17	.20	.24	.12	.03
93. Inform. inf.	.03	−.11	.02	−.03	.06	.01	.02	−.04	−.07	−.02	.03	−.09	−.11	−.01	.08	−.05	−.03	−.09	−.03	.02
94. Well-being	.11	.14	.03	−.2	.01	.00	.09	.04	.02	.06	−.05	.15	.09	−.12	.03	.09	.14	.07	.08	.09
95. Male	.10	.14	.03	.03	.01	−.01	.01	.01	.02	.01	.11	.08	.06	.14	.03	.01	.04	.05	.10	.09
96. Org. tenure	.18	.10	.16	.05	.21	.11	.07	.13	.17	.19	.21	.04	.13	.08	.18	.15	.06	.13	.19	.15
97. Team. Dist.	.09	.03	.04	.14	.11	.05	.07	−.01	.02	.06	−.03	−.04	.05	.02	.11	.03	.04	−.02	.00	.07

Variable	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78
60. BU.3.S.1	.35	—																		
61. FS.1.B.1	.01	.03	—																	
62. FS.1.BM.1	.15	.14	.13	—																
63. FS.1.DR.1	.04	.01	.21	.22	—															
64. FS.1.DR.2	.07	.04	.20	.32	.20	—														
65. FS.1.DR.3	.09	.05	.15	.13	.17	.26	—													
66. FS.1.DR.4	.06	.02	.25	.07	.22	.22	.17	—												
67. FS.1.P.1	.12	.05	.29	.26	.17	.21	.21	.22	—											
68. FS.1.P.2	.13	.04	.30	.19	.17	.19	.18	.12	.29	—										
69. FS.1.P.3	.31	−.05	.29	.27	.16	.25	.11	.12	.33	.38	—									
70. FS.1.S.1	.08	.21	.25	.18	.15	.17	.07	.24	.26	.2	.26	—								
71. FS.2.B.1	.02	.04	.62	.15	.16	.24	.14	.23	.30	.28	.3	.22	—							
72. FS.2.BM.1	.19	.13	.18	.6	.20	.14	.05	.12	.27	.16	.23	.20	.22	—						

(*Appendices continue*)

Appendix B (*continued*)

Variable	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78
73. FS.2.DR.1	.08	.02	.29	.35	.63	.28	.13	.19	.28	.26	.19	.19	.26	.33	—					
74. FS.2.DR.2	.12	.05	.23	.32	.17	.67	.26	.21	.25	.27	.21	.20	.28	.20	.28	—				
75. FS.2.DR.3	.08	.02	.22	.24	.20	.29	.63	.16	.21	.16	.07	.12	.20	.12	.23	.34	—			
76. FS.2.DR.4	.09	.05	.37	.24	.26	.30	.16	.62	.26	.29	.24	.32	.38	.27	.36	.37	.28	—		
77. FS.2.P.1	.17	.03	.31	.23	.17	.16	.14	.20	.67	.33	.39	.27	.31	.24	.28	.26	.20	.34	—	
78. FS.2.P.2	.13	.04	.24	.12	.15	.23	.16	.10	.27	.69	.37	.18	.29	.02	.20	.26	.13	.31	.32	—
79. FS.2.P.3	.36	−.02	.28	.36	.20	.25	.16	.15	.31	.32	.69	.24	.31	.32	.30	.28	.21	.33	.36	.36
80. FS.2.S.1	.09	.27	.24	.27	.10	.18	.15	.16	.19	.28	.26	.50	.31	.30	.20	.16	.15	.25	.20	.24
81. FS.3.B.1	.06	.01	.59	.24	.15	.23	.23	.17	.32	.28	.34	.23	.72	.25	.26	.32	.25	.31	.28	.33
82. FS.3.BM.1	.06	.03	.15	.63	.26	.23	.02	.18	.27	.20	.31	.21	.23	.74	.35	.30	.09	.34	.26	.12
83. FS.3.DR.1	.16	.03	.21	.26	.65	.29	.15	.25	.23	.23	.19	.19	.19	.33	.70	.30	.23	.37	.25	.16
84. FS.3.DR.2	.08	−.01	.20	.24	.23	.68	.23	.26	.15	.24	.17	.21	.20	.15	.29	.71	.30	.33	.17	.21
85. FS.3.DR.3	.09	.06	.17	.18	.16	.29	.68	.21	.21	.20	.08	.11	.17	.07	.19	.35	.71	.26	.18	.18
86. FS.3.DR.4	.16	.09	.23	.17	.25	.23	.09	.66	.18	.23	.22	.21	.26	.19	.30	.27	.10	.67	.22	.20
87. FS.3.P.1	.19	.08	.29	.38	.18	.20	.21	.20	.72	.35	.44	.28	.31	.30	.31	.28	.24	.35	.75	.34
88. FS.3.P.2	.12	.06	.27	.11	.20	.24	.21	.17	.28	.72	.32	.21	.34	.12	.25	.29	.20	.31	.30	.72
89. FS.3.P.3	.39	.02	.30	.26	.12	.26	.16	.08	.26	.33	.67	.20	.27	.21	.24	.27	.14	.23	.33	.38
90. FS.3.S.1	.07	.24	.21	.24	.16	.23	.14	.19	.19	.18	.20	.47	.27	.23	.26	.21	.20	.28	.16	.19
91. Div. perf.	.21	.07	.05	.13	.21	.19	.19	.20	.04	.05	.06	.01	.04	.11	.16	.24	.17	.22	.04	.06
92. Psych. safe.	.23	.11	.15	.22	.21	.23	.22	.27	.06	.12	.13	.10	.11	.18	.17	.25	.23	.31	.10	.09
93. Inform. inf.	.06	−.07	−.07	−.01	.03	.06	.06	−.01	−.03	.00	.00	−.08	.03	.05	.10	.02	.07	.03	−.03	.01
94. Well-being	.03	.10	.12	−.08	.06	.04	.09	.04	.06	.05	.04	.14	.11	.03	.05	.07	.08	.07	.11	.07
95. Male	.15	.17	−.12	−.13	−.06	−.05	−.05	−.09	−.06	−.09	.08	−.10	−.08	−.09	−.04	−.06	−.08	−.13	−.05	−.04
96. Org. tenure	.22	.10	.17	.14	.13	.06	.08	.15	.14	.14	.15	.07	.16	.01	.13	.12	.08	.13	.17	.15
97. Team. Dist.	.02	.04	.02	.06	.05	.06	.09	.03	.01	.06	.04	.06	.06	−.02	.04	.04	.00	.03	.07	.02

Variable	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97
80. FS.2.S.1	.25	—																	
81. FS.3.B.1	.37	.28	—																
82. FS.3.BM.1	.44	.29	.22	—															
83. FS.3.DR.1	.26	.16	.21	.32	—														
84. FS.3.DR.2	.23	.15	.22	.23	.30	—													
85. FS.3.DR.3	.19	.18	.27	.01	.24	.30	—												
86. FS.3.DR.4	.24	.14	.21	.23	.34	.25	.13	—											
87. FS.3.P.1	.42	.28	.33	.37	.26	.18	.24	.25	—										
88. FS.3.P.2	.35	.22	.38	.15	.22	.28	.22	.24	.32	—									
89. FS.3.P.3	.77	.24	.38	.25	.22	.23	.18	.18	.38	.35	—								
90. FS.3.S.1	.22	.61	.31	.29	.21	.20	.18	.21	.21	.2	.23	—							
91. Div. perf.	.07	.03	.07	.03	.22	.29	.21	.25	.12	.08	.10	.11	—						
92. Psych. safe.	.16	.07	.11	.15	.22	.34	.25	.30	.14	.12	.16	.13	.63	—					
93. Inform. inf.	.05	−.08	−.03	.05	.08	.06	.02	.00	−.03	.06	.04	−.03	−.02	−.01	—				
94. Well-being	.08	.18	.07	−.04	.04	.03	.07	.07	.07	.10	.05	.24	.13	.13	.02	—			
95. Male	.03	−.11	−.03	−.08	−.03	−.05	−.03	−.07	−.05	−.08	.05	−.04	.00	−.05	.05	.03	—		
96. Org. tenure	.12	.01	.18	.09	.10	.08	.10	.18	.16	.15	.14	.06	.02	.13	.04	.00	.11	—	
97. Team. Dist.	−.05	.00	.01	−.06	.04	.05	.05	.05	.04	.11	.00	.01	.08	.04	−.04	.05	.01	.09	—

Note. Leadership dimensions are labeled as follows: RO = results orientation; BU = business understanding; FS = forging synergy. Sources are labeled as: B = boss; BM = board member; DR = direct report; P = peer; S = self. Male was dummy code (1 = *male*; 0 = *female*). Org. (Organization) tenure is represented in years. Team dist. (distance) refers to the percentage of time that team members, on average during a typical work week, work at the same physical worksite. Div. perf. = division performance; Psych. safe. = psychological safety; Inform. inf. = information influence.

(*Appendices continue*)

Appendix C

Interrater Agreement and Interrater Reliability for Each Dimension of Multisource Leadership Ratings

	Interrater agreement				Interrater reliability			
Leadership dimension	Board member	Superior	Direct report	Peer	Board member	Superior	Direct report	Peer
Results orientation	.83	.89	.84	.87	.70	.77	.70	.73
Business understanding	.86	.90	.85	.86	.71	.80	.71	.80
Forging synergy	.87	.90	.87	.88	.66	.68	.70	.72

Note. Interrater agreement was assessed using the $r_{wg(j)}$ index while assuming a rectangular null distribution and reflects the average across all participants (LeBreton & Senter, 2008). Interrater reliability was assessed using the ICC(2, *k*) index (McGraw & Wong, 1996). ICC = intraclass correlation coefficients.

Appendix D

Average Level of Observability Across Leadership Dimensions and Sources

Leadership dimension	Arena	Reputation	Identity	Direct report	Peers	Superior	Board members
Results orientation	3.67	3.75	3.33	4.67	3.33	4.00	3.00
Business understanding	3.80	3.75	4.00	3.67	4.00	4.33	3.00
Forging synergy	3.33	3.33	3.33	4.33	3.67	3.00	2.33

Note. Ratings of observability are averaged across raters. We conducted an analysis of variance where differences in observability ratings were considered across raters, sources (e.g., board vs. direct report), and leadership dimension. We found a Significant Source × Leadership Dimension interaction ($F = 5.16$, $p = .03$), which accounted for 42% of the variability in observability ratings. This indicates that, even while accounting for any potential differences across the raters, the average level of observability varies significantly across *both* sources and leadership dimensions.

Appendix E

Factor Loadings From LARI (S-1) Models for Each Leadership Dimension

		Results orientation			Business understanding			Forging synergy		
Factor	Item label	Std. estimate	*SE*	*p*	Std. estimate	*SE*	*p*	Std. estimate	*SE*	*p*
Arena	B.1.1	0.85	0.04	.00	0.70	0.04	.00	0.72	0.04	.00
Arena	BM.1.1	0.31	0.07	.00	0.15	0.07	.08	0.25	0.06	.00
Arena	DR.1.1	0.22	0.03	.00	0.26	0.03	.00	0.25	0.03	.00
Arena	P.1.1	0.32	0.03	.00	0.30	0.03	.00	0.39	0.03	.00
Arena	S.1.1	0.17	0.04	.00	0.25	0.04	.00	0.32	0.04	.00
Arena	B.1.2	0.75	0.04	.00	0.85	0.04	.00	0.84	0.04	.00
Arena	BM.1.2	0.26	0.07	.00	0.27	0.07	.00	0.28	0.07	.00
Arena	DR.1.2	0.20	0.03	.00	0.33	0.03	.00	0.26	0.03	.00
Arena	P.1.2	0.31	0.03	.00	0.35	0.03	.00	0.42	0.03	.00
Arena	S.1.2	0.13	0.04	.02	0.37	0.05	.00	0.36	0.03	.00
Arena	B.1.3	0.71	0.04	.00	0.83	0.04	.00	0.84	0.04	.00
Arena	BM.1.3	0.23	0.07	.01	0.17	0.07	.05	0.30	0.07	.00
Arena	DR.1.3	0.25	0.03	.00	0.32	0.03	.00	0.33	0.03	.00
Arena	P.1.3	0.34	0.03	.00	0.40	0.03	.00	0.39	0.04	.00
Arena	S.1.3	0.30	0.04	.00	0.37	0.05	.00	0.38	0.04	.00
Arena	DR.2.1	0.22	0.03	.00	0.27	0.03	.00	0.24	0.03	.00
Arena	DR.2.2	0.21	0.03	.00	0.32	0.03	.00	0.25	0.03	.00
Arena	DR.2.3	0.27	0.03	.00	0.32	0.03	.00	0.34	0.03	.00
Arena	DR.3.1	0.21	0.03	.00	0.26	0.03	.00	0.25	0.03	.00
Arena	DR.3.2	0.20	0.03	.00	0.32	0.03	.00	0.25	0.03	.00
Arena	DR.3.3	0.27	0.03	.00	0.30	0.03	.00	0.33	0.03	.00
Arena	DR.4.1	0.22	0.03	.00	0.26	0.03	.00	0.25	0.03	.00
Arena	DR.4.2	0.22	0.03	.00	0.33	0.03	.00	0.28	0.03	.00
Arena	DR.4.3	0.27	0.03	.00	0.31	0.03	.00	0.35	0.03	.00

(*Appendices continue*)

Appendix E (*continued*)

Factor	Item label	Results orientation			Business understanding			Forging synergy		
		Std. estimate	*SE*	*p*	Std. estimate	*SE*	*p*	Std. estimate	*SE*	*p*
Arena	P.2.1	0.32	0.03	.00	0.29	0.03	.00	0.38	0.03	.00
Arena	P.2.2	0.31	0.03	.00	0.35	0.03	.00	0.41	0.03	.00
Arena	P.2.3	0.34	0.03	.00	0.37	0.03	.00	0.40	0.04	.00
Arena	P.3.1	0.32	0.03	.00	0.28	0.03	.00	0.38	0.03	.00
Arena	P.3.2	0.32	0.03	.00	0.33	0.03	.00	0.41	0.03	.00
Arena	P.3.3	0.35	0.03	.00	0.37	0.03	.00	0.39	0.04	.00
Identity	S.1.1	0.83	0.04	.00	0.58	0.04	.00	0.53	0.04	.00
Identity	S.1.2	0.70	0.04	.00	0.72	0.04	.00	0.67	0.03	.00
Identity	S.1.3	0.54	0.04	.00	0.77	0.04	.00	0.70	0.04	.00
Board Member	BM.1.1	0.81	0.06	.00	0.61	0.07	.00	0.54	0.06	.00
Board Member	BM.1.2	0.71	0.06	.00	0.71	0.08	.00	0.75	0.06	.00
Board Member	BM.1.3	0.76	0.06	.00	0.72	0.07	.00	0.71	0.07	.00
Direct Report	DR.1.1	0.36	0.06	.00	0.27	0.05	.00	0.30	0.08	.00
Direct Report	DR.1.2	0.26	0.05	.00	0.20	0.05	.00	0.18	0.09	.08
Direct Report	DR.1.3	0.40	0.03	.00	0.18	0.05	.00	0.06	0.09	.50
Direct Report	DR.2.1	0.33	0.05	.00	0.36	0.04	.00	0.21	0.07	.00
Direct Report	DR.2.2	0.26	0.05	.00	0.20	0.05	.00	0.17	0.09	.08
Direct Report	DR.2.3	0.41	0.03	.00	0.19	0.05	.00	0.07	0.09	.50
Direct Report	DR.3.1	0.33	0.05	.00	0.36	0.04	.00	0.22	0.07	.00
Direct Report	DR.3.2	0.26	0.05	.00	0.19	0.05	.00	0.17	0.09	.08
Direct Report	DR.3.3	0.42	0.03	.00	0.17	0.05	.00	0.06	0.09	.50
Direct Report	DR.4.1	0.34	0.05	.00	0.36	0.04	.00	0.22	0.07	.00
Direct Report	DR.4.2	0.28	0.05	.00	0.20	0.05	.00	0.19	0.09	.08
Direct Report	DR.4.3	0.42	0.03	.00	0.18	0.05	.00	0.07	0.09	.50
Peers	P.1.1	0.26	0.07	.00	0.01	0.06	.87	0.26	0.06	.00
Peers	P.1.2	0.29	0.06	.00	0.16	0.06	.03	0.32	0.04	.00
Peers	P.1.3	0.39	0.04	.00	0.17	0.06	.02	0.35	0.04	.00
Peers	P.2.1	0.30	0.05	.00	0.00			0.44	0.05	.00
Peers	P.2.2	0.29	0.06	.00	0.16	0.06	.03	0.31	0.04	.00
Peers	P.2.3	0.39	0.04	.00	0.15	0.06	.02	0.36	0.04	.00
Peers	P.3.1	0.30	0.05	.00	0.00			0.45	0.05	.00
Peers	P.3.2	0.30	0.06	.00	0.15	0.06	.03	0.31	0.04	.00
Peers	P.3.3	0.41	0.04	.00	0.15	0.06	.02	0.35	0.04	.00
Reputation	DR.1.1	0.21	0.06	.00	0.31	0.05	.00	0.24	0.07	.00
Reputation	DR.1.2	0.28	0.05	.00	0.26	0.04	.00	0.41	0.05	.00
Reputation	DR.1.3	0.02	0.05	.72	0.38	0.04	.00	0.40	0.03	.00
Reputation	DR.2.1	0.27	0.06	.00	0.26	0.04	.00	0.29	0.05	.00
Reputation	DR.2.2	0.29	0.05	.00	0.26	0.04	.00	0.40	0.05	.00
Reputation	DR.2.3	0.02	0.05	.72	0.39	0.04	.00	0.41	0.03	.00
Reputation	DR.3.1	0.27	0.06	.00	0.26	0.04	.00	0.30	0.05	.00
Reputation	DR.3.2	0.29	0.05	.00	0.25	0.04	.00	0.39	0.05	.00
Reputation	DR.3.3	0.02	0.05	.72	0.36	0.04	.00	0.39	0.03	.00
Reputation	DR.4.1	0.28	0.06	.00	0.26	0.04	.00	0.30	0.05	.00
Reputation	DR.4.2	0.31	0.05	.00	0.26	0.04	.00	0.43	0.05	.00
Reputation	DR.4.3	0.02	0.05	.72	0.37	0.04	.00	0.43	0.03	.00
Reputation	BM.1.1	0.29	0.07	.00	0.30	0.10	.01	0.42	0.07	.00
Reputation	BM.1.2	0.06	0.08	.57	0.18	0.10	.14	0.37	0.08	.00
Reputation	BM.1.3	0.00			0.52	0.09	.00	0.31	0.08	.00
Reputation	P.1.1	0.29	0.05	.00	0.32	0.04	.00	0.21	0.04	.00
Reputation	P.1.2	0.32	0.05	.00	0.38	0.04	.00	0.28	0.04	.00
Reputation	P.1.3	0.07	0.06	.24	0.42	0.04	.00	0.24	0.05	.00
Reputation	P.2.1	0.29	0.05	.00	0.31	0.04	.00	0.20	0.04	.00
Reputation	P.2.2	0.32	0.05	.00	0.37	0.04	.00	0.27	0.04	.00
Reputation	P.2.3	0.07	0.06	.24	0.39	0.04	.00	0.25	0.05	.00
Reputation	P.3.1	0.29	0.05	.00	0.30	0.04	.00	0.20	0.04	.00
Reputation	P.3.2	0.32	0.05	.00	0.36	0.04	.00	0.27	0.04	.00
Reputation	P.3.3	0.08	0.06	.24	0.38	0.04	.00	0.24	0.05	.00

Note. $n = 491$. Models were estimated using full maximum likelihood estimation to account for missing data (Newman, 2014). LARI (S-1) model is an extension of the LARI model that excludes the specific superior factor and defines the Arena factor with the superior's ratings. LARI (S-1) model with criteria references extends this model to include predictive pathways between the LARI factors and measures of leader effectiveness (i.e., division performance, psychological safety, informal influence, and well-being). These models treat multiple raters within a particular source as interchangeable by including equality constraints across the raters within this group. Small negative factor loadings were constrained to zero for all models (B. S. Connelly et al., 2022). LARI = leadership Arena–Reputation–Identity; Std. = standard; *SE* = standard error; B = boss; BM = board member; DR = direct report; P = peer; S = self.

(*Appendices continue*)

Appendix F

Summary of Robustness Checks

Along with the analyses reported in the article, we conducted several robustness checks. They include: (a) replicating the relative contributions of the LARI (S-1) model factors using a random-effects model, (b) ensuring that the moderating effects of the observability of leadership dimensions are distinct from their reliability (Lance et al., 2010; Oh et al., 2011), and (c) conducting a small-scale simulation study to ensure the statistical validity of our models. Below we describe each of these additional analyses.

Random-Effects Models

Along with treating the leadership dimensions as a fixed effect, we estimated the models testing the relative contributions of the LARI (S-1) factors using a random-effects model. We again found that including the interactions between the observability of a leadership dimension and the relative predictive validity of LARI (S-1) factor improved the model fit (Δpseudo $R^2 = .10$). The pattern of the interaction effects was also consistent with what we observed using a fixed-effects modeling approach.

Controlling for Reliability

The LARI (S-1) model presumes that differential predictive effects across sources can be attributed to observability (B. S. Connelly et al., 2022; Human & Biesanz, 2011; Vazire, 2010). However, an alternative, psychometric, perspective contends that

Table F1

Moderating Effects of Observability on Predictive Relationships Between LARI (S-1) Factors and Leader Effectiveness (Random Effects)

	Model 1		
	Dependent variable: β weights		
Predictor	β	*SE*	*p*
Intercept	.58	.18	<.001
Observability of leadership dimension	−.13	.05	.01
Observability × Arena	.13	.07	.07
Observability × Identity	.12	.06	.06
Observability × Direct Reports	.07	.06	.22
Observability × Peers	.23	.06	<.001
Observability × Board Members	.14	.06	.02
LARI factors (reference—Reputation)			
Arena	−.57	.26	.03
Identity	−.54	.22	.02
Direct Reports	−.17	.21	.43
Peers	−.94	.23	<.001
Board Members	−.61	.20	.01
Criterion (reference—Division performance)			
Informal leadership	−.11	.03	.00
Psychological safety	−.02	.03	.47
Well-being	−.11	.03	<.001
Factor × Criterion			
Arena × Informal Leadership	.09	.04	.03
Identity × Informal Leadership	.12	.04	.01
Direct Reports × Informal Leadership	−.03	.04	.40
Peers × Informal Leadership	.17	.04	<.001
Board Members × Informal Leadership	.11	.04	.01
Arena × Psychological Safety	.03	.04	.43
Identity × Psychological Safety	.02	.04	.61
Direct Reports × Psychological Safety	−.03	.04	.49
Peers × Psychological Safety	.03	.04	.48
Board Members × Psychological Safety	.02	.04	.64
Arena × Well-Being	.11	.04	.01
Identity × Well-Being	.15	.04	<.001
Direct Reports × Well-Being	−.05	.04	.25
Peers × Well-Being	.12	.04	.01
Board Members × Well-Being	.15	.04	<.001
Pseudo R^2		.66	
ΔPseudo R^2		.10	
n		72	

Note. Change statistics compare a model that excludes the interaction terms for observability and LARI factors to those presented here. LARI (S-1) = leadership Arena–Reputation–Identity extension model; *SE* = standard error; *n* = number of parameters estimated across the three LARI models.

(*Appendices continue*)

Table F2

Moderating Effects of Observability and Reliability on Predictive Relationships Between LARI (S-1) Factors and Leader Effectiveness

Predictor	Model 1		
	Dependent variable: β weights[2]		
	β	*SE*	*p*
Intercept	.64	.11	<.001
Observability of leadership dimension	−.14	.03	<.001
Observability × Arena	.14	.03	<.001
Observability × Identity	.14	.03	<.001
Observability × Direct Reports	.08	.03	.02
Observability × Peers	.25	.03	<.001
Observability × Board Members	.15	.03	<.001
Reliability	.00	.05	1.00
Reliability ×Arena	.00	.20	1.00
Reliability × Identity	.00	.11	1.00
Reliability × Direct Reports	.00	.07	1.00
Reliability × Peers	.00	.07	1.00
Reliability × Board Members	.00	.08	1.00
LARI factors (reference—Reputation)			
Arena	−.59	.15	<.001
Identity	−.61	.13	<.001
Direct Reports	−.20	.13	.11
Peers	−1.03	.12	<.001
Board Members	−.64	.11	<.001
Leadership dimension (reference—Business understanding)			
Strategic planning	−.01	.01	.30
Leading change	−.01	.01	.10
Criterion (reference—Division performance)			
Informal leadership	−.11	.01	<.001
Psychological safety	−.01	.01	.22
Well-being	−.11	.01	<.001
Factor × Criterion			
Arena × Informal Leadership	.09	.02	<.001
Identity × Informal Leadership	.12	.02	<.001
Direct Reports × Informal Leadership	−.03	.02	.08
Peers × Informal Leadership	.17	.02	<.001
Board Members × Informal Leadership	.11	.02	<.001
Arena × Psychological Safety	.03	.02	.10
Identity × Psychological Safety	.02	.02	.29
Direct Reports × Psychological Safety	−.03	.02	.15
Peers × Psychological Safety	.03	.02	.14
Board Members × Psychological Safety	.02	.02	.33
Arena × Well-Being	.11	.02	<.001
Identity × Well-Being	.15	.02	<.001
Direct Reports × Well-Being	−.05	.02	.02
Peers × Well-Being	.12	.02	<.001
Board Members × Well-Being	.15	.02	<.001
R^2		.77	
ΔR^2		.09	
n		72	

Note. Change statistics (ΔF and ΔR) compare a model that excludes the interaction terms for observability and LARI factors to those presented here. LARI (S-1) = leadership Arena–Reputation–Identity extension model; *SE* = standard error; *n* = number of parameters estimated across the three LARI models.

differential prediction reflects differences in the reliability of the sources. Because certain sources (e.g., Arena) are defined by a greater number of indicators, they are apt to exhibit greater reliability. Because reliability establishes the upper limit of validity, sources that are more reliable may demonstrate stronger associations with other variables. A psychometric perspective contends that if the variability cannot be consistently attributed to a particular source it must be error that will limit the validity of the sources within the LARI (S-1) model (Lance et al., 2010).

To test this alternative perspective, we reestimated our models testing the relative predictive validity of the LARI (S-1) factors while controlling for their reliability. We operationalized the reliability of the sources for each leadership dimension by extracting the ω values for each source (grand mean = .41, *SD* = 0.23). We then included these reliability estimates and two-way interactions in our model. These results indicate that, while controlling for observability, the reliability of the LARI (S-1) factors has limited effect on their relative predictive validity.

(*Appendices continue*)

Table F3
Results of Small-Scale Simulation Study for LARI (S-1) Models

Measure	Factor loadings			β estimates		
	M	*SD*	*Mdn*	*M*	*SD*	*Mdn*
Bias	.00	.04	.00	.03	.21	.00
Standardized bias	−.04	.11	−.03	.00	.06	.00
Coverage	.93	.04	.94	.94	.02	.94
Relative bias in standard error	−.05	.11	−.03	−.13	.30	−.04
Statistical power	.94	.19	1.00	.46	.37	.31
Statistical power (Est. > .25)	.94	.19	1.00	.72	.38	.97

Note. n = 243 factor loadings; 72 regression weights. Results are based on 3,000 replications. Bias = simulated parameter estimates − parameter value; standardized bias = average simulated parameter estimate − parameter value)/standard deviation of simulated parameter estimate); coverage = percentage of 95% confidence intervals from simulations that include underlying parameter; relative bias in standard error = (average simulated standard error − standard deviation of simulated parameter estimate)/standard deviation of bias across all replications; power = proportion of significant replications when testing whether the parameter estimates are different from 0 and assuming a critical value of .05. LARI = leadership arena–reputation–identity; Est. = estimate.

This null finding may be due, in part, to the nature of reliability of the LARI (S-1) factors. Specifically, it does not appear that the factors defined by more indicators are inherently more reliable (Lance et al., 2010). Our findings suggest the reliability of a particular factor (e.g., Arena, Reputation, Self) varies across specific leadership dimensions. This finding is consistent with studies of bifactor models that documented variability across constructs in the reliability of general and specific factors (Rodriguez et al., 2016). These results provide a useful complement to a psychometric perspective by demonstrating that the predictive validity of multisource leadership ratings depends upon the joint effects of what's being rated (i.e., the leadership dimension) and who's providing that rating (i.e., a specific source).

Small-Scale Simulation Study

To strengthen the statistical validity of our findings, we conducted three separate, small-scale simulation studies. Specifically, using the simsem package in *R* (Pornprasertmanit et al., 2021), we estimated 1,000 replications of the LARI (S-1) model for each of the leadership dimensions (see Table F3). For both the factor loadings and β estimates, we found, on average, relatively low levels of bias and standardized bias and adequate levels of coverage. For example, Collins et al. (2001) concluded that biases only become noticeable when standardized bias is greater than 0.40 in magnitude, which is far greater than what we observed. We also found acceptable levels of bias in the standard error estimates (Hoogland & Boomsma, 1998), which is often a particular concern with bifactor models (Eid et al., 2017; Zhang et al., 2021). We also found adequate levels of power for the factor loadings within the model and power levels for the regression parameters that were comparable to previous studies of predictive bifactor models, especially when examining those relationships that were moderate in size (β = .25; Zhang et al., 2021). Taken as a whole, these findings indicate that the parameters within the LARI S-1 model can be recovered with reasonable accuracy.

We also carefully considered the results of the small-scale simulation study by carefully reviewing several of the parameters generated within each simulation. First, we considered the degree to which the models produced negative factor loadings. Across the 1,000 simulations and 243 factor loadings (i.e., 81 loadings × 3 dimensions), we found nine parameters whose 90% confidence interval extended below 0. Of these, though, eight were relatively small deviations (−.03), which prior studies have typically constrained to be greater than 0 (B. S. Connelly et al., 2022). Second, we also considered the pattern of loadings between the general factors (i.e., Arena and Reputation) and specific factors. Specifically, for each item from a specific source (e.g., Item No. 1 rated by the leader), we examined the correlation between the loading for this item on its own specific factor (e.g., Identity) and the two general factors (e.g., Arena and Reputation) across the 1,000 simulation that were generated. Across the 153 pairings, we found the average correlation was −.22 (minimum = −.69, maximum = .08, 90% CI [−.45, .01]). Thus, across 3,000 simulations (1,000 simulations × 3 leadership dimensions), it appears that as the typical item becomes a stronger indicator of a general factor, it typically is a weaker indicator of a specific factor (and vice versa). This is consistent with the pattern of proportionality that Zhang et al. (2021) found to be a key factor for why augmented bifactor models perform better than a general bifactor model.

Received March 27, 2023
Revision received July 15, 2024
Accepted July 17, 2024 ■