Received: 11 April 2017 | Revised: 17 January 2018 | Accepted: 5 March 2018

DOI: 10.1111/peps.12266

WILEY PERSONNEL PSYCHOLOGY

ORIGINAL ARTICLE

Gender and leadership emergence: A meta-analysis and explanatory model

Katie L. Badura[1] | Emily Grijalva[1] | Daniel A. Newman[2] | Thomas Taiyi Yan[3] | Gahyun Jeon[4]

[1]University at Buffalo

[2]University of Illinois at Urbana-Champaign

[3]University of Maryland

[4]Northwestern University

Correspondence
Katie Badura, Department of Organization and Human Resources, University at Buffalo, School of Management, 261 Jacobs Management Center, Buffalo, NY 14260-4000.
Email: katienib@buffalo.edu

Abstract

Research has shown that men tend to emerge as leaders more frequently than women. However, societal role expectations for both women and leaders have changed in the decades since the last empirical review of the gender gap in leader emergence (Eagly & Karau, 1991). We leverage meta-analytic evidence to demonstrate that the gender gap has decreased over time, but a contemporary gap remains. To understand why this gap in leader emergence occurs, we draw on social role theory to develop a Gender-Agency/Communion-Participation (GAP) Model—an integrative theoretical model that includes both trait and behavioral mechanisms. Specifically, we examine a sequence of effects: from gender to agentic and communal personality traits, from these traits to behavioral participation in group activities, and ultimately from participation to leader emergence. The model is tested using original meta-analyses of the personality and behavioral mechanisms (coding 1,632 effect sizes total). Gender differences in leadership emergence are predominately explained by agentic traits (positive) and communal traits (negative), both directly and through the mechanism of participation in group discussions. In addition, several paths in the theoretical model are moderated by situational contingencies. Our study enhances knowledge of the mechanisms and boundary conditions underlying the gender gap in leader emergence.

1 | INTRODUCTION

In a comprehensive review of the gender-leadership emergence literature, Eagly and Karau (1991) found that men emerged as leaders more often than did women, yet evidence suggests that this gender gap in leader emergence is not due to gender differences in leader effectiveness (Lanaj & Hollenbeck, 2015; Paustian-Underdahl, Walker, & Woehr,

2014). Because leadership emergence refers to "whether" and "to what degree" an individual is perceived as a leader by others (Judge, Bono, Ilies, & Gerhardt, 2002), the gender difference in leader emergence paints a troubling picture of the preconceptions women must overcome to obtain leadership roles. Being perceived as leaderlike in organizational contexts, such as committees and self-managing teams, is associated with higher ratings of job performance (Porath, Gerbasi, & Schorch, 2015; Zhang, Waldman, & Wang, 2012) and promotability (Staw & Barsade, 1993). As a consequence, the gender disparity in leader emergence has a potentially dramatic cumulative impact on women's career trajectories (Heilman, 2012).

However, over 25 years have elapsed since the last comprehensive review of the gender difference in leader emergence—a time period during which an increasing proportion of women joined the workforce and attained management roles (U.S. Bureau of Labor Statistics, 2004, 2015). Given this change in the composition of the workforce, we investigate whether the gender gap in leader emergence has also changed over time by performing an updated meta-analysis of the gender–leadership emergence relationship. Although we expect the gender gap to have declined over time, we also predict that a contemporary gender difference remains in leadership emergence. We make this prediction because gender-related stereotypes are generally slow to change (Lueptow, Garovich-Szabo, & Lueptow, 2001), and workforce statistics indicate a continuing gender disparity in leadership roles (i.e., women comprised 32% of managers in 1983 vs. 39% of managers in 2014; U.S. Bureau of Labor Statistics, 2004, 2015). Given the likelihood of a continuing gender gap, in this paper we seek to address both the explanatory mechanisms and boundary conditions for the gender difference in leader emergence.

First, drawing from social role theory (Eagly, 1987; Wood & Eagly, 2012), we utilize both trait and behavioral mechanisms to offer a possible explanation for *why* men and women emerge into leader roles at disparate rates. Altogether, we examine a sequence of effects: from gender to agentic traits (e.g., dominance and assertiveness) and communal traits (e.g., warmth and nurturance; Eagly, 1987), from these traits to behavioral disparities in men's and women's participation in group activities, and ultimately from participatory behavior to leader emergence. Estimating how much of the gender difference in leader emergence is uniquely accounted for by agentic versus communal personality traits permits us to untangle the extent to which it is women's lack of agency, abundance of communion, or both that lead to the gender gap. Further, incorporating *participation in group discussion*—which is considered to be one of the dominant and most proximal predictors of leadership emergence (Mullen, Salas, & Driskell, 1989)—provides insight into the behavioral differences between men and women that might contribute to the gender disparity. Through the development of this integrative theoretical model [which we label The Gender-Agency/Communion-Participation (GAP) Model], this paper answers calls to better unify trait and behavioral perspectives to explain leadership outcomes (Avolio, 2007; DeRue, Nahrgang, Wellman, & Humphrey, 2011).

In introducing this theoretical model, however, we are acutely aware that prior reviews have established that gender differences in various leadership outcomes are highly contingent on contextual cues (Eagly & Karau, 1991; Paustian-Underdahl et al., 2014). To take into account how possible boundary conditions affect the gender gap in leadership emergence, we investigate four focal moderators, including: study setting, gender egalitarian culture, length of interaction time, and social complexity of a task (each of these is elaborated below). Consistent with social role theory, we generally predict that the gender difference in leader emergence will be smaller when there is greater compatibility between women's perceived roles and leaders' perceived roles, as well as when there is greater opportunity to get to know women and thus base leadership judgments on individuals' qualifications instead of gender stereotypes.

To reconcile our theoretical model with the fact that there exist powerful moderators in the gender literature, we next assess the extent to which our model explains the gender gap in leader emergence across different levels of the aforementioned contingency factors. Incorporating moderators into our framework allows us to offer a more rigorous test of the proposed model, and enables readers to make better informed generalizations of the results across a variety of settings. In sum, this paper contributes to theory by: (a) assessing whether the gender difference in leader emergence has decreased over time, (b) identifying a process through which the gender gap arises, and (c) examining moderators of the gender gap in leader emergence, as well as how these moderators effect the paths in our theoretical model of the gender gap.

1.1 | Gender differences in leadership emergence across time

Social role theory advances the idea that individuals use implicit theories about what is generally expected in a work role to determine whether men or women are better suited for that particular role (Eagly & Karau, 2002). A fundamental tenet of this theory is that gender differences derive from the gendered division of labor within societies, which observers then attribute to intrinsic variations between men and women (Eagly & Steffen, 1984; Gawronski, 2004; Gilbert, 1998). In other words, because men historically were more likely to occupy "breadwinner and higher status roles," they were assumed to be inherently more agentic than women (e.g., assertive and dominant); and because women more frequently served in "homemaker and lower status roles" (Eagly & Karau, 2002, p. 574), they were assumed to be inherently more communal than men (e.g., nurturing, warm, and emotionally expressive; Eagly, 1987; Williams & Best, 1990). Of particular relevance to this paper, the agentic characteristics associated with men tend to align more closely with the characteristics expected of leaders, in comparison to the communal characteristics associated with women (Koenig, Eagly, Mitchell, & Ristikari, 2011[1]; also see Lord, De Vader, & Alliger, 1986). Indeed, this greater perceived compatibility between stereotypes about men and leadership roles is theorized to be the root cause of men's favored status when it comes to attaining leadership positions (Eagly & Karau, 2002; Heilman, 2001; Schein, 1973).

However, some evidence indicates that both perceptions of the leadership role and perceptions of gender may have changed over time (see Donnelly & Twenge, 2016; Duehr & Bono, 2006; Koenig et al., 2011; Powell, Butterfield, & Parent, 2002; Twenge, 1997). As a result of these changes, there may be repercussions that impact secular trends in the gender gap in leader emergence. In the quarter century since the publication of Eagly and Karau's (1991) paper, many organizations have spearheaded efforts to combat gender inequality in leadership attainment (e.g., through mentoring and other programs designed to curtail structural gender biases against women; Ibarra, Carter, & Silva, 2010). These and other efforts appear to be having some impact, as women have obtained a greater percentage of managerial positions over the last few decades (U.S. Bureau of Labor Statistics, 2004, 2015). According to social role theory, these changes in the distribution of the labor force should be accompanied by changes in societal stereotypes of women (Koenig & Eagly, 2014; Weber & Crocker, 1983; Wood & Eagly, 2012). Indeed, research confirms that stereotypes of women appear to be shifting to reflect more agentic (stereotypically masculine) content (Diekman & Eagly, 2000), and women's agentic trait levels increased between 1974 and 2012 (Donnelly & Twenge, 2016).

Beyond changing stereotypes of women, there is also evidence that peoples' conceptions of leadership itself have become less agentic and more communal over time (Koenig et al., 2011; Powell et al., 2002). For example, Duehr and Bono (2006) updated prior work on Schein's (1973; 1975) *think manager–think male* paradigm to show that current managers in their study viewed "successful managers in general" as less agentic and more communal than a sample of managers in a 1989 comparison study. To contextualize these results, the authors stated that, "This movement toward greater balance in the stereotypically male (agentic) and female (communal) characteristics of successful middle managers is in line with arguments that modern leadership paradigms are moving toward the feminine" (e.g., Eagly & Carli, 2003; Fondas, 1997; p. 834). In sum, with the increasing prevalence of women in leadership positions, the ascendance of agentic traits in modern women's gender roles, as well as increasing communal and decreasing agentic content in leadership prototypes, we expect the gender gap in leader emergence to have diminished over recent decades. Thus, we examine the following hypothesis:

Hypothesis 1: Time moderates the gender-leadership emergence relationship, such that there is a smaller gender gap in newer studies as compared to older studies.

2 | A MODEL OF GENDER AND LEADERSHIP EMERGENCE

Although we hypothesize that the gender gap has diminished over time, we also anticipate that there will remain a contemporary gender difference in leader emergence (for the reasons provided above). Thus, we next address the

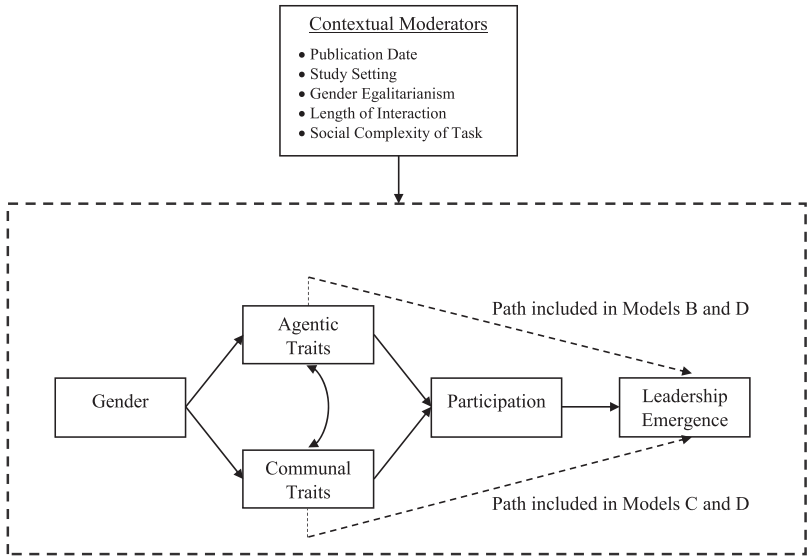

FIGURE 1 A Gender-Agency/Communion-Participation (GAP) Model of Leader Emergence. Model A: Does not include any direct paths from agentic or communal traits to leadership emergence. Model B: Includes a direct path from agentic traits to leadership emergence. Model C: Includes a direct path from communal traits to leadership emergence. Model D: Includes direct paths from both agentic and communal traits to leadership emergence

question: *Why* do men emerge into leadership roles more frequently than do women? In particular, our model includes the traits of agency (e.g., assertiveness and dominance) and communion (e.g., kindness and nurturance), as well as participatory behavior in group discussions, as three mechanisms through which the gender gap in leadership emergence arises.

2.1 | Gender and agentic/communal traits

Previously, we noted that social role theory serves as a basis for understanding why men typically have higher agency and women have higher communion (Eagly, 1987; Wood & Eagly, 2012). Gender differences in these traits are thought to be robust because they are maintained by a confluence of forces, including external pressure through socialization (Block, 1973; Leaper & Friedman, 2007; Lytton & Romney, 1991), backlash directed at individuals who deviate from gender roles (Carli, 2001; Rudman, 1998; Rudman & Glick, 1999; 2001), as well as internal pressure to possess sex-typed traits (e.g., self-regulatory processes aimed at adhering to one's gender identity; Wood, Christensen, Hebl, & Rothgerber, 1997). Because gender's association with agentic and communal traits has been well established, these relationships are incorporated into our model (Figure 1), but we do not advance separate formal hypotheses. Nevertheless, providing the first meta-analytic estimates of the gender–agency and gender–communion relationships are further contributions of the current study.

2.2 | Agency and participation

In building our explanatory model, we next seek to understand how the traits of agency and communion subsequently relate to participatory behavior. *Participatory behavior* refers to the amount that people contribute during group discussions and includes behaviors such as providing suggestions to the group and offering opinions (Bales, 1950). Bass (1949) was one of the first researchers to single out participation as a predictor of who tends to ascend to leadership positions. Since that time, participation has gone on to receive extensive attention in the study of leaderless groups and is considered to be the key driver of leader emergence (Bass, 1954; Kennedy, Anderson, & Moore, 2013; Mullen et al., 1989). In fact, Willard and Strodtbeck (1972) suggested that "one of the most consistent findings in the small

group literature is the strong relation between an individual's amount of participation in the group discussion and his leadership or influence in that group" (p. 161).

To establish the linkages between agency/communion and participation, we draw on classic personality theory, which proposes that traits tend to predict trait-relevant behavior (Allport, 1966; Murray, 1938). This is theorized to be the case because expressing traits is intrinsically rewarding, leading to pleasure and a sense of fulfillment, whereas acting in ways that are incongruous with traits produces anxiety and dissatisfaction (Allport, 1951; Tett & Burnett, 2003, p. 504). As mentioned above, agency is a trait characterized by attributes such as being dominant and assertive (Bakan, 1966; Wiggins, 1991), and we argue that participation is a corresponding behavior that logically follows from these attributes. That is, many of the specific behaviors related to speaking up during group discussions are agentic in nature, such as taking charge of the situation (Klein & Willerman, 1979; Leaper & Ayres, 2007), initiating structure (Abele & Wojciszke, 2007), and solving problems (Bass, 1954).

Thus, from a trait-based interactionist perspective (Tett & Burnett, 2003), it is likely that group settings will activate agentic propensities that are expressed through participation. Consistent with this assertion, Weiss and colleagues (2014) note that agentic individuals are more willing to voice suggestions and opinions within groups. Likewise, in a study examining the emergence of communication networks, dominance (a classic agentic trait; Wiggins, 1991) was associated with a greater amount of participation, including both sending and receiving more messages (Brown & Miller, 2000; see Schmid-Mast, 2002). Consequently, we posit that agentic traits are a critical factor encouraging the choice to actively participate in group discussions.

Hypothesis 2: Agentic traits have a positive relationship with participation.

2.3 | Communion and participation

The tendency to behave in trait-consistent ways is also expected to produce an association between communion and participation. In particular, we expect that within a group context, communal individuals will engage in social behaviors that promote affiliation and harmony, such as being friendly and smiling (Cashdan, 1998; Seibert & Gruenfeld, 1992). Interestingly, scholars have used terms to define communion that directly refer to participation (e.g., "communion entails group participation and cooperation with others"; Buss, 1990, p. 556). Of value in understanding why communal traits may be linked to participation is the idea that participation in group discussions often has a sociability component (Jose & McCarthy, 1988). In general, people who have higher levels of communion tend to have more interpersonal involvement in groups (Ickes, 1981; Leaper, 1987). This connotes that communal individuals may participate in groups because participation is one mechanism through which they can fulfill their sociability needs. Further, people possessing high levels of communion tend to focus on the collective well-being of the group (Bakan, 1966); thus, it follows that communal individuals may participate more often due to their perceived obligation to support and sustain the group.

Hypothesis 3a: Communal traits have a positive relationship with participation.

At the same time, we would like to make clear that we expect communal traits to have a lesser impact than agentic traits on participation. Substantial research has shown that leaderless group discussions are typically task oriented (e.g., solving a dilemma or completing a task; Eagly & Karau, 1991; McGrath, 1984), and scholars have proposed that task-oriented behaviors are more strongly associated with agentic traits than with communal traits (e.g., Jose & McCarthy, 1988; Lanaj & Hollenbeck, 2015). Accordingly, there is likely to be greater congruence between the tendency to participate in group discussions and agentic characteristics, versus communal characteristics.

Hypothesis 3b: The relationship between communal traits and participation is weaker than the relationship between agentic traits and participation.

2.4 | Participation and leadership emergence

Having provided arguments for why agency and communion are related to participation, we now address the relation of participatory behavior with leader emergence. Consistent with the "babble hypothesis," which theorizes that those who speak more are viewed as leaderlike (Bass, 1990, p. 93; Riggio, Riggio, Salinas, & Cole, 2003), past research indicates that participation is a concrete action that prompts others to infer an individual is well-suited for the leader role (Bales, Strodtbeck, Mills, & Roseborough, 1951; Bass, 1949; Leaper & Ayres, 2007; Mullen et al., 1989; Zeleny, 1939; for a review see Paunova, 2015). Scholars have theorized that participation affects emergent leadership perceptions by increasing members' relative salience within the group (Mullen et al., 1989; Shaw, 1959). Salience subsequently leads to the impression that high participators are more important and prominent than other members who participate less frequently (Mullen et al., 1989).

In addition, greater participation may lead group members to conclude that active participators have more expertise and knowledge of group tasks (Littlepage, Schmidt, Whisler, & Frost, 1995), and are more motivated and interested in performing leadership roles (Mullen et al., 1989; Riggio et al., 2003). Thus, an individual's quantity of participation may signal to others not only an ability, but also a willingness to contribute to the group (Littlepage et al., 1995; Sorrentino & Boutiller, 1975). In this way, the importance of participation in groups is consistent with a functional leadership theory perspective in which people are expected to emerge as leaders when they appear to be instrumental to completing critical tasks needed to sustain the group (McGrath, 1962; Zaccaro, Rittman, & Marks, 2001). Indeed, it has been proposed that, "election to the position of leadership is a reward for the group member who moves the group toward task completion" (Mullen et al., 1989, p. 547). For these reasons, we expect that frequent participators will tend to emerge as leaders within groups.

Hypothesis 4: Participation is positively related to leadership emergence.

2.5 | An integrated model of gender and leadership emergence: The Gender-Agency/Communion-Participation (GAP) Model

Based on the aforementioned theoretical rationale, we advance the model depicted in Figure 1 to explain *why* there is a gender gap in leader emergence. Altogether, the integrated model implies the following two mediation hypotheses.

Hypothesis 5a: Agentic traits and participation behavior will serially mediate the effect of gender on leadership emergence.

Hypothesis 5b: Communal traits and participation behavior will serially mediate the effect of gender on leadership emergence.

To test the theoretical model, we use a large-scale meta-analytic database. We examine four variations on the basic mediation model. As depicted in Figure 1, Model A is a full mediation model (i.e., gender to traits to participation to leader emergence), whereas Models B, C, and D reflect variations on the mediation model (partial mediation models), in which the agentic and communal traits are specified to have direct effects on leader emergence, over and above participatory behavior.

3 | CRITICAL CONTINGENCIES

To acknowledge complexities associated with the gender gap in leader emergence, we examine several moderators that we expect to magnify (or attenuate) the gap: study setting, gender egalitarianism, length of interaction, and social complexity. Our choice of moderator variables is grounded in social role theory and leads us to predict that the gender gap will be *smaller* when (a) tasks offer contextual cues that reduce the perceived incompatibility between the role expectations for women and leaders, and (b) people are able to base leader emergence ratings on attributes other than

gender (such as knowledge, abilities, and performance). We discuss the moderating role of each variable in turn and then pose a research question designed to address how these moderators influence our integrated theoretical model.

3.1 | Study setting

Study setting is a potentially important contextual factor in leadership emergence. The original gender-leadership emergence meta-analysis by Eagly and Karau (1991) purposefully excluded studies of managers working in organizations due to legitimate concerns that available studies used real-world promotions as an index of leader emergence, and these promotions were likely based on factors (e.g., seniority) other than those thought to underlie the interpersonal phenomenon of leader emergence. Since that time, more studies have been conducted on the emergence of leaders in real work contexts, making it possible for the current paper to investigate the age old question of whether results based primarily on lab studies translate to work contexts (cf., Dipboye & Flanagan, 1979).

In lab and classroom settings, participants do not typically have information about group members' past performance or behaviors and are therefore forced to rely more heavily on superficial or surface-level characteristics (Harrison, Price, Gavin, & Florey, 2002), such as gender, when making judgments about leader selection. Alternatively, in business settings, individuals rating a focal member's emergent leadership often have access to information regarding past performance and behavior. Koch, D'Mello, and Sackett (2015) assert that, "the more a decision maker has access to information about credentials, skills, relevant experience, and the like, the less the decision maker relies on gender as the basis for decision" (p. 130; see also Singletary & Hebl, 2009). Because gender differences in both actual job performance and leadership effectiveness are close to nil, on average (Paustian-Underdahl et al., 2014; Roth, Purvis, & Bobko, 2012), we expect that access to performance-related information would shrink the gender gap in leader emergence in business settings. Thus, we hypothesize the following:

Hypothesis 6: Study setting moderates the relationship between gender and leadership emergence. There is a smaller gender gap in leader emergence in business settings, compared to (a) lab settings and (b) classroom settings.

3.2 | Gender egalitarianism

Role expectations for men and women may vary across cultures, which has implications for the generalizability of the gender gap in leader emergence. Gender egalitarianism is a particularly relevant cross-cultural difference with respect to gender and leader emergence, as it reflects the extent to which a culture uses biological sex to determine a person's societal role (Emrich, Denmark, & Den Hartog, 2004). There is likely greater alignment between gender role expectations for women and leaders within highly gender egalitarian cultures, because these societies seek to minimize role differences between men and women (House & Javidan, 2004; Lyness & Judiesch, 2014). In support of this idea, Koenig and colleagues (2011) recently demonstrated that there is a greater incongruity between the role expectations for women and leaders in Eastern societies (i.e., lower gender egalitarian cultures) as compared to Western societies (i.e., higher gender egalitarian cultures). Hence, we propose that within higher gender egalitarian nations, there will be a weaker gender difference in leader emergence.

Hypothesis 7: Gender egalitarianism moderates the relationship between gender and leadership emergence, such that there is a smaller gender gap when the culture is more gender egalitarian.

3.3 | Length of interaction

The gender gap in leader emergence may also depend upon how long group members have known each other prior to selecting a leader (Eagly & Karau, 1991).[2] Gender is one of the most salient surface-level characteristics (Ridgeway & Correll, 2004; Stangor, Lynch, Duan, & Glas, 1992), and stereotypes of men tend to align with leadership roles (Eagly & Karau, 2002; Schein, 1973). Thus, at initial acquaintance, people may tend to automatically categorize men as leaders more often than women. However, expectancy states theory posits that people tend to place greater emphasis on

specific status cues (e.g., skills and abilities), as compared to diffuse cues (e.g., gender), when the specific status cues are relevant to decision making (Berger, Rosenholtz, & Zelditch, 1980; Correll & Ridgeway, 2006). This is consistent with research showing that, as the length of group interaction time increases, people rely less on demographic stereotypes (Biesanz, West, & Millevoi, 2007; Harrison et al., 2002) and instead base leadership emergence ratings on group members' unique attributes and behaviors (Eagly & Karau, 1991; 2002). In this way, individuating information that amasses over time can override peoples' reflexive reliance on stereotypes. As a result, we predict that the gender gap in leadership emergence will be weakened when individuals interact for longer as opposed to shorter periods of time.

Hypothesis 8: Length of interaction moderates the relationship between gender and leadership emergence, such that as the length of interaction time increases, the size of the gender gap decreases.

3.4 | Social complexity

The nature of the task itself also potentially influences the extent to which men and women emerge into leadership roles. In particular, Eagly (2016) notes that women may have a leadership advantage in groups performing socially complex tasks (i.e., tasks that require a high degree of interpersonal coordination, such as identifying innovative solutions to problems and jury deliberation; Eagly & Karau, 1991).[3] This is because there are expected to be greater communal/relational demands inherent to tasks involving interpersonal coordination, and women tend to perform better at such tasks than do men (Eagly, 2016; Eagly & Karau, 1991). Following this logic, there is initial evidence demonstrating a weaker gender gap in leadership emergence when tasks entail a high degree of social complexity (Eagly & Karau, 1991). More recently, Eagly (2016) urged scholars to pay more attention to social complexity, noting that "existing findings are promising" but "require more exploration" (p. 209). This study allows us to answer this call for research by investigating whether social complexity remains an important contemporary boundary condition.

Hypothesis 9: Social complexity of the task moderates the relationship between gender and leadership emergence, such that there is a smaller gender gap when the task has a high degree of social complexity.

3.5 | Generalizing the theoretical model across moderator categories

Given our expectation that the gender gap in leadership emergence varies across the contextual variables outlined above, we are also interested in establishing whether our explanatory model (Figure 1) holds across these moderating conditions (i.e., publication date, study setting, gender egalitarianism, length of interaction, and social complexity). Accounting for the presence of moderators in path analysis allows researchers to more clearly represent how boundary conditions affect a theoretical model and in doing so increases a model's theoretical precision. We therefore test whether the entire integrated model shown in Figure 1 generalizes across different levels of the aforementioned moderators.

Research Question: Does the theoretical model of the gender gap in leader emergence (Figure 1) depend upon: (a) publication date, (b) study setting, (c) cultural gender egalitarianism, (d) length of interaction time, and/or (e) social complexity of task?

4 | METHODS

4.1 | Literature search

We performed a multistep literature search to identify primary data for inclusion in the current study. First, using ERIC, Google Scholar, Proquest dissertations, PsycINFO, and Web of Science, we conducted a literature search for primary studies using each of the keywords listed in Appendix A. Second, we searched the online programs from the annual

meetings of the Academy of Management, International Leadership Association, and the Society for Industrial and Organizational Psychology for all of the years in which the programs were electronically available. Third, we conducted a forward search for more recent articles that cited papers included in the current meta-analysis. Fourth, we conducted a backward search in which we examined the reference sections of papers included in our meta-analysis to identify older primary studies. Finally, we attempted to find all of the primary studies included in relevant past meta-analyses (Eagly & Karau, 1991; Ensari, Riggio, Christian, & Carslaw, 2011; Leaper & Ayres, 2007; Mullen et al., 1989; Twenge, 1997).

We adhered to a few general inclusion guidelines. First, primary studies needed to report enough information to calculate an effect size (i.e., *d* or *r*) for one of the proposed relationships in the current paper.[4] Next, we only included samples that comprised adults (e.g., undergraduates, working adults). Further, articles were excluded if they were written in a language other than English. Beyond these general guidelines, we established more fine-grained inclusion criteria for each variable in this meta-analysis, as explained below.

With regard to leadership emergence, primary studies were coded when there was a direct measure capturing the extent to which group members were perceived as leaders in leaderless groups. These direct measures included leader elections, Likert scale ratings of perceived leadership, and rank ordering of group members on emergent leadership. However, measures of group participation and indirect measures of leadership emergence (e.g., who sat at the head seat of the table; Nemeth, Endicott, & Wachtler, 1976) were not coded as measures of leader emergence in the current study (consistent with the conceptualization of leadership emergence used in other meta-analyses; Grijalva, Harms, Newman, Gaddis, & Fraley, 2015a; Judge et al., 2002).[5] This updated operationalization of leadership emergence resulted in the exclusion of 10 primary studies initially coded as leadership emergence by Eagly and Karau (1991; for a list of the excluded studies see Appendix B).

When coding agency and communion, we included data that were collected using the Bem Sex Role Inventory (BSRI; Bem, 1974) and the Personal Attributes Questionnaire (PAQ; Spence, Helmreich, & Stapp, 1975), consistent with past meta-analysis (Twenge, 1997).[6] In addition, based on strong consensus within the participation literature (Leaper & Ayres, 2007), we focused our participation search on studies that used a "quantitative observational measure" to capture the extent to which people talk within group discussions (p. 333). This resulted in the inclusion of the following measures of participation within group discussion: "(a) number of words or utterances, (b) rate or time sampling, (c) words per turn, (d) duration of talking, (e) total turns, and (f) total statements or speech acts" (Leaper & Ayres, 2007, p. 333).

In applying the above criteria, our search resulted in the inclusion of 1,632 effect sizes. The primary study effect sizes used in our original meta-analyses (reported in Tables 1–4) can be found in the Online Supplemental Appendix C. For a summary of the number of independent samples included in each of the original meta-analyses conducted for this paper, see Table 4 (e.g., 136 independent samples were included for the gender–leadership emergence relationship).

4.2 | Data coding

To start, we reviewed the primary studies identified in our literature search and coded the effect sizes, reliability estimates, and sample sizes. We also coded multiple hypothesized moderators for the gender–leadership emergence relationship. First, we coded the year of data collection. To be consistent with previous work (Grijalva et al., 2015b; Twenge, Konrath, Foster, Campbell, & Bushman, 2008), when available, we used the year in which the authors conducted the study. If this information was not reported, we used 2 years prior to the date of publication for published articles. For unpublished work, we used the year of the conference or dissertation/thesis completion. Second, we coded the gender egalitarianism of each primary sample using country-level scores reported in the GLOBE study (Emrich et al., 2004, p. 365). We also coded the setting in which the study was conducted—we examined three different settings: lab, classroom, and work. For coding length of interaction time, we based our codes on when the leadership emergence ratings occurred: (a) almost immediately after initial acquaintance; (b) after initial acquaintance, but within 20 minutes or fewer; (c) after more than 20 minutes in a single interaction; or (d) after more than one interaction. Finally, consistent with Eagly and Karau (1991), we coded the social complexity of a task as either low (e.g., "structured physical activities,

rank ordering of items given to subjects, problem solving with one or more possible solutions given to subjects") or high (e.g., "problem solving that requires the generation of novel solutions, interpersonal problem solving, projects that fulfill course requirements, jury deliberations, socially complex simulation games, purely social interactions," p. 690).

TABLE 1 Meta-analytic results for gender and leadership emergence

	k	*N*	*d*	δ	*SD* δ	95% CI	80% CV
Gender–leadership emergence	136	19,073	.21	.22	.38	[.13, .29]	[−.27, .71]
Study setting							
Lab	82	7,401	.42	.43	.49	[.29, .56]	[−.19, 1.06]
Classroom	39	7,827	.05	.06	.30	[−.06, .16]	[−.32, .44]
Business setting	11	3,117	.15	.17	.11	[.04, .25]	[.03, .31]
Length of interaction							
Immediately after initial acquaintance	36	2,206	.82	.83	.70	[.55, 1.09]	[−.06, 1.73]
Less than 20 minutes of interaction	25	2,036	.45	.48	.43	[.25, .66]	[−.07, 1.02]
More than 20 minutes in one interaction	15	1,790	.12	.12	.20	[−.02, .26]	[−.13, .37]
More than one interaction	50	10,965	.07	.09	.26	[−.01, .16]	[−.25, .42]
Social complexity of task							
High social complexity	78	14,436	.12	.14	.32	[.04, .21]	[−.27, .54]
Low social complexity	55	3,927	.57	.58	.51	[.39, .75]	[−.07, 1.23]
Measurement type							
Election	90	9,249	.29	.31	.49	[.16, .42]	[−.32, .94]
Questionnaire	48	10,067	.14	.16	.24	[.05, .22]	[−.15, .46]
Rank	9	1,220	−.01	−.08	.34	[−.27, .25]	[−.52, .35]
Publication type							
Published	104	14,190	.27	.28	.40	[.17, .37]	[−.23, .80]
Unpublished	32	4,883	.04	.04	.29	[−.09, .16]	[−.33, .41]
Dissertations	25	3,419	.04	.04	.35	[−.13, .21]	[−.40, .49]
Conference and working papers	7	1,464	.02	.02	.00	[−.07, .11]	[.02, .02]

Note. Positive *d* values denote that men scored higher than women. *k* = number of effect sizes in the meta-analysis; *N* = total sample size in the meta-analysis; *d* = uncorrected standardized mean difference between men and women; δ = standardized mean difference corrected for unequal sampling and attenuation in the criterion; SD δ = standard deviation of the corrected correlation; CI = confidence interval around *d*; CV = credibility interval around δ. Moderation tests for publication date and gender egalitarianism (both continuous moderators) appear in Table 7.

TABLE 2 Meta-analytic results for gender's relationship with agency, communion, and participation

	k	*N*	*d*	δ	*SD* δ	95% CI	80% CV
Gender							
Agentic traits	409	100,915	.41	.48	.25	[.38, .44]	[.16, .80]
Communal traits	416	101,055	−.62	−.73	.28	[−.59, −.65]	[−.37, −1.09]
Participation	45	4,477	.16	.16	.34	[.04, .29]	[−.28, .60]

Note. Positive *d* values denote that men scored higher than women. *k* = number of effect sizes in the meta-analysis; *N* = total sample size in the meta-analysis; *d* = uncorrected standardized mean difference between men and women; δ = standardized mean difference corrected for unequal sampling and attenuation in the criterion; SD δ = standard deviation of the corrected correlation; CI = confidence interval around *d*; CV = credibility interval around δ.

TABLE 3 Meta-analytic results for participation and leadership emergence

	k	*N*	*r*	ρ	SD ρ	95% CI	80% CV
Participation							
Agentic traits	11	825	.31	.34	.05	[.24, .38]	[.28, .40]
Communal traits	11	906	−.02	−.02	.15	[−.12, .08]	[−.22, .17]
Leadership emergence							
Agentic traits	12	1,256	.34	.39	.14	[.26, .43]	[.21, .57]
Communal traits	11	1,185	−.12	−.13	.05	[−.05, −.18]	[−.06, −.20]
Participation	27	2,550	.51	.52	.14	[.45, .57]	[.35, .70]

Note. k = number of effect sizes in the meta-analysis; N = total sample size in the meta-analysis; r = uncorrected correlation; ρ = correlation corrected for attenuation in the predictor and criterion; SD ρ = standard deviation of the corrected correlation; CI = confidence interval around r; CV = credibility interval around ρ.

In addition to the above information, we examined four methodological moderators (i.e., publication status, leadership emergence measurement type, average subject age, and average group size). Publication status included published articles and unpublished works (e.g., dissertations/theses and conference presentations). We also coded leadership emergence measurement type (i.e., questionnaire, election, or rank order). Finally, subject age and average group size were coded as continuous variables.

Consistent with past meta-analytic work, we made multiple efforts to maintain the independence of the data included in this study. In situations where primary studies reported multiple effect sizes for multiple measures of the same construct, we used composite formulas (when sufficient information was available; Ghiselli, Campbell, & Zedeck, 1981; Sackett & Ellingson, 1997), or we used the average correlation (when sufficient information was not available to compute composite correlations). In addition, when multiple primary studies were based on the same or overlapping data (e.g., Hong, 2005; Hong, Catano, & Liao, 2011), we included only one of the primary studies in our analysis (i.e., we used the primary study with the bigger sample size; or when two reports of the same data used the same sample size, then we relied upon the published primary study). To assess the accuracy of the above data coding procedures, two authors and five trained research assistants initially coded 45% of the articles, and there was high agreement among coders (99%). Given the high agreement among raters, the first author coded the remaining articles. As a precautionary measure, all articles that were ambiguous or difficult to code were flagged and discussed among the first three authors until resolution was reached.

4.3 | Data analysis

Using the meta-analytic procedures outlined in Schmidt and Hunter (2015), we corrected for two statistical artifacts. First, we used the correction for unequal sample sizes (p. 287) in instances where gender was a predictor variable. This correction is used to address underestimation of effect sizes that results from uneven sampling from two groups of interest (e.g., men and women; Schmidt & Hunter, 2015). In addition, we corrected for measurement error in both the predictor and criterion variables. To complete this step, we coded all available reliability information from primary studies. For missing reliability estimates, we imputed a sample-weighted average reliability derived from available information in the included primary studies (e.g., for leadership emergence effects, we imputed $\alpha = .92$ for missing reliability estimates). Separate reliability estimates were calculated for each measure of agentic and communal traits. The Cronbach's α values for agentic traits were as follows: .84 (BSRI), .73 (PAQ), and .79 (BSRI-PAQ composite). The following estimates were used for communal traits: .83 (BSRI), .76 (PAQ), and .80 (BSRI-PAQ composite). Finally, based on the objective nature of measurement for both gender and participation (i.e., amount of time spent talking), we used conservative reliability estimates of 1.00 for these two variables (so as not to overcorrect).

To test our hypothesized mediated relationships, we used the meta-analytic correlation matrix shown in Table 4. Consistent with past meta-analytic path analyses, we elected to use effect sizes corrected for statistical artifacts (i.e.,

TABLE 4 Meta-analytic correlation matrix

Variables	Gender	Agency	Communion	Participation	Leader emergence
1. Gender	–				
2. Agency	.22	–			
	409/100,915				
Old studies (1947–1995)	.25	–			
	185/39,318				
New studies (1996–2017)	.20	–			
	224/61,597				
3. Communion	−.32	.09	–		
	416/101,055	554/110,243			
Old studies (1947–1995)	−.35	.06	–		
	185/38,868	278/43,738			
New studies (1996–2017)	−.30	.11	–		
	231/62,187	276/66,505			
4. Participation	.08	.34	−.02	–	
	45/4,477	11/825	11/906		
Old studies (1947–1995)	.12	.35	.09	–	
	32/2,492	2/178	2/259		
New studies (1996–2017)	.02	.34	−.07	–	
	12/1,931	9/647	9/647		
Lab settings only	.13	.34	−.02	–	
	37/2,802	11/825	11/906		
Short-term interaction	.17	.30	.10	–	
	15/1,208	7/284	8/436		
Long-term interaction	.03	.36	−.14	–	
	21/2,700	3/470	3/470		
High social complexity	.05	.37	−.10	–	
	31/3,550	8/447	8/447		
Low social complexity	.17	.30	.05	–	
	9/495	3/378	3/459		
5. Leadership emergence	.09	.39	−.13	.52	–
	136/19,073	12/1,256	11/1,185	27/2,550	
Old studies (1947–1995)	.17	.43	−.16	.53	–
	71/6,982	5/553	4/482	13/781	
New studies (1996–2017)	.05	.36	−.11	.52	–
	65/12,091	7/703	7/703	14/1,769	
Lab settings only	.18	.38	−.13	.52	–
	82/7,401	8/719	7/648	24/2,405	
Short-term interaction	.28	.47	−.16	.50	–
	61/4,242	4/256	4/256	10/845	
Long-term interaction	.04	.36	−.12	.53	–
	65/12,755	7/929	7/929	16/1,537	

(Continues)

TABLE 4 (Continued)

Variables	Gender	Agency	Communion	Participation	Leader emergence
High social complexity	.06	.37	−.12	.54	–
	78/14,436	9/978	9/978	15/1,392	
Low social complexity	.24	.48	−.19	.50	–
	55/3,927	3/278	2/207	9/950	

Note. Positive correlations mean that men scored higher than women. Each cell contains the correlation corrected for attenuation in the predictor and criterion, and gender correlations are also corrected for unequal sampling; followed by the number of effect sizes (*k*) and the total sample size (*N*).

estimates of the true relationships among latent theoretical constructs). We based our overall path analysis on the harmonic mean sample size (Viswesvaran & Ones, 1995), which was $N = 2{,}149$. We tested four alternative theoretical models (using Lisrel 9.2), which differed only in the specification of direct paths from agentic and communal traits to leadership emergence. These four models were as follows: (Model A) no direct paths from agentic and communal traits to leadership emergence, (Model B) direct path from agentic traits to leadership emergence, (Model C) direct path from communal traits to leadership emergence, and (Model D) direct paths from both agentic and communal traits to leadership emergence (see Figure 1).

To investigate the continuous moderators (i.e., year of data collection and gender egalitarianism), sample-weighted least squares regression was used (Aguinis, Gottfredson, & Wright, 2011; Schmidt & Hunter, 2015; Steel & Kammeyer-Mueller, 2002) with year of data collection or gender egalitarianism predicting the primary study's gender–leadership emergence effect size. Categorical moderators (e.g., study setting) were evaluated using procedures outlined by Raju and Brand (2003) to test whether the meta-analytic effect sizes differed significantly across conditions.

Finally, we analyzed the entire integrated theoretical model (Figure 1) to assess how the model paths changed across moderator conditions. For example, we tested the theoretical model using only studies involving short-term interactions, and then tested the model again using only studies with longer term interactions. To perform these analyses, a metacorrelation matrix was created for each of seven moderator categories (i.e., older studies, newer studies, short-term interactions, long-term interactions, high social complexity, low social complexity, and lab only), which included effect sizes for that particular moderator category only. Using these meta-correlation matrices, we conducted multigroup structural equation modeling (SEM) analyses in Lisrel to test for statistically significant differences in path coefficients across moderator categories. To illustrate, we tested whether participation is more or less important in predicting leader emergence in situations that differ in length of interaction among participants. For these moderator analyses, we first set all model path estimates to be free and then fixed a single path to be equal across groups (Stark, Chernyshenko, & Drasgow, 2006). A chi-square difference test was used to compare path coefficients between groups (this step was performed separately for each of the seven paths included in our model).

When testing moderators of the entire theoretical model, there were sufficient data to code the following moderators: publication date (i.e., older studies vs. newer studies),[7] length of interaction time (i.e., short-term interactions vs. long-term interactions),[8] and social complexity of a task (i.e., high vs. low). With respect to study setting, we did find a sufficient number of primary studies to enable us to report the bivariate gender–leader emergence effect size broken down by lab, classroom, and business settings. However, for the overall model (Figure 1), there were not enough primary studies to separately examine agency, communion, and participation effects in classroom or business settings—multiple cells in the matrix contained zero studies. In addition, there were not sufficient primary data to fill in the meta-correlation matrix for different levels of gender egalitarianism culture. Thus, we were unable to test the research question for the moderators of study setting and gender egalitarianism. Finally, for three of the bivariate effects (i.e., gender–agency, gender–communion, and agency–communion), the only moderator we coded was publication date, because there was no theoretical reason to expect that the other moderators influenced these proposed associations (i.e., these relationships did not inherently involve a task, a particular setting, or a dyadic interaction).

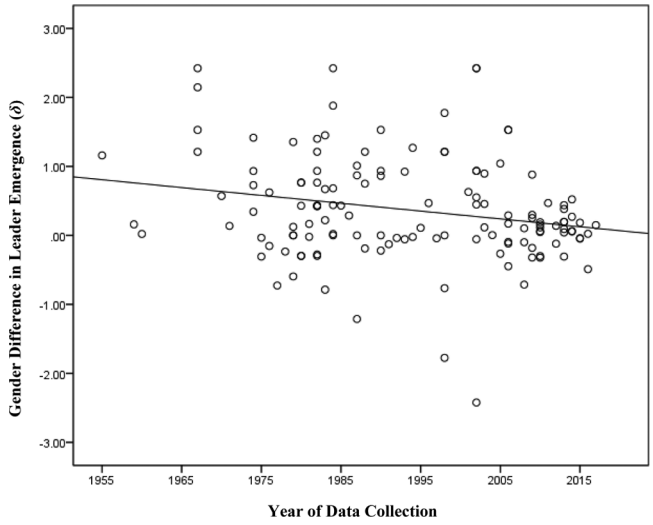

FIGURE 2 Gender differences in leadership emergence from 1955 to 2017
Note: δ represents the standardized mean difference corrected for uneven sampling and attenuation in the criterion. Positive δ values denote that men scored higher on leadership emergence than women.

5 | RESULTS

Results for the gender–leadership emergence relationship are reported in Table 1. Consistent with our prediction, men tended to emerge in leadership roles more often than did women ($d = .21$; $\delta = .22$; $k = 136$; $N = 19{,}073$; 95% CI [.13, .29]).[9] To test whether the gender difference in leadership emergence changed over time, we performed a sample-size weighted regression analysis in which the year of data collection was used to predict the gender–leadership emergence effect size (δ). As depicted in Figure 2, the gender gap in leadership emergence decreased over time ($b = -.011$, standardized $\beta = -.34$, $t = -4.24$, $p < .05$), providing support for *Hypothesis 1*. When we restricted our analysis to studies conducted within the recent 5-year period (2012–2017), there continued to be a significant gender difference in leadership emergence favoring men ($d = .09$; $\delta = .10$; $k = 19$; $N = 5{,}356$; 95% CI [.02, .16]); this suggests that although the gender gap in leader emergence is shrinking, it has not disappeared.

As expected, our original meta-analyses showed that, on average, men tended to possess higher levels of agentic traits ($d = .41$; $\delta = .48$; $k = 409$; $N = 100{,}915$; 95% CI [.38, .44]), women tended to possess higher levels of communal traits ($d = -.62$; $\delta = -.73$; $k = 416$; $N = 101{,}055$; 95% CI [−.59, −.65]), and men tended to participate more often in group discussions ($d = .16$; $\delta = .16$; $k = 45$; $N = 4{,}477$; 95% CI [.04, .29]; see Table 2). We hypothesized that agentic and communal traits would be influential factors driving participation. Agentic traits had a positive relationship with participation ($r = .31$; $\rho = .34$; $k = 11$; N = 825; 95% CI [.24, .38]; supporting *Hypothesis 2*), but communal traits were not related to participation ($r = -.02$; $\rho = -.02$; $k = 11$; $N = 906$; 95% CI [−.12, .08]; failing to support *Hypothesis 3a*). In comparing the magnitude of the effect sizes, agentic traits had a stronger positive association with participation in group discussions than did communal traits ($z = 6.25$, $p < .05$; supporting *Hypothesis 3b*). There was also a positive association between participatory behavior and leadership emergence ($r = .51$; $\rho = .52$; $k = 27$; $N = 2{,}550$; 95% CI [.45, .57]; supporting *Hypothesis 4*; see Table 3).

5.1 | Mediator analyses

We next tested our model's indirect effects using the meta-analytic correlation matrix shown in Table 4. We compared four theoretical models (Models A through D, as described in Figure 1), using both chi-square difference tests and change in CFI (see Table 5). Using these goodness-of-fit indices, we retained Model D as the best-fitting model, which

TABLE 5 Fit statistics for alternative theoretical models

Models	χ^2	*df*	$\Delta\chi^2$	RMSEA	CFI	Δ CFI	NNFI	SRMR
Model A[1]	232.60*	4	223.45*	.163	.855	.140	.637	.064
Model B[2]	75.62*	3	66.47*	.106	.954	.041	.846	.037
Model C[3]	190.03*	3	180.88*	.170	.881	.114	.604	.058
Model D[4]	9.15*	2		.041	.995		.977	.014

$N = 2{,}149$; * $p < .05$.
[1]Does not include any direct paths from agentic or communal traits to leadership emergence.
[2]Includes a direct path from agentic traits to leadership emergence.
[3]Includes a direct path from communal traits to leadership emergence.
[4]Includes direct paths from both agentic and communal traits to leadership emergence.

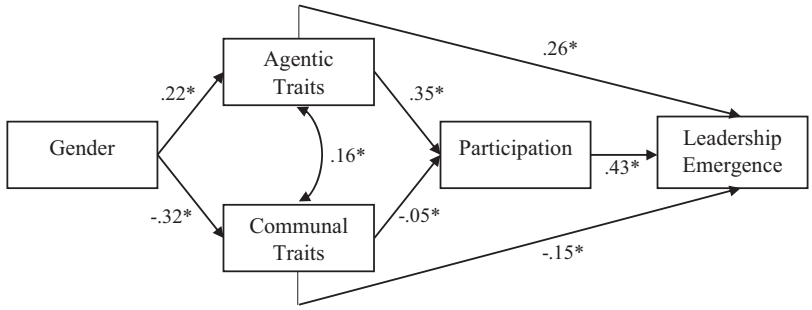

FIGURE 3 Path model results for The Gender-Agency/Communion-Participation (GAP) Model (Model D)
Note: Standardized path coefficients (β's) are presented. $N = 2{,}149$; χ^2 (2) = 9.15, $p < .05$; RMSEA = .041, CFI = .995; NNFI = .977; SRMR = .014; * $p < .05$.

included direct paths from both agentic and communal traits to leadership emergence (see Figure 3). This model displayed good overall fit [$\chi^2_{(df=2)} = 9.15$; $p < .05$; RMSEA = .041; CFI = .995; NNFI = .977; SRMR = .014].[10] Path coefficients for Model D appear in Figure 3. All hypothesized paths were statistically significant in the expected directions, with the exception of the *negative* path from communal traits to participation ($\beta = -.05$, $p < .05$; i.e., we had hypothesized a positive effect, but we found a small, negative effect of communion on participation). Although not hypothesized, we also found that communal traits were detrimental to leader emergence ($\beta = -.15$, $p < .05$).

In order to test the mediation hypotheses (*H5a* and *H5b*), we used two formal tests: (a) a test of joint significance (i.e., both paths X→M and M→Y are statistically significant; Mackinnon, Lockwood, Hoffman, West, & Sheets, 2002), and (b) a test of the indirect effects (i.e., the product of path X→M multiplied by path M→Y is statistically significant; Hayes & Scharkow, 2013). When testing the statistical significance of the indirect effects, we used Monte Carlo confidence intervals (CI). Monte Carlo CI outperforms the Sobel test due to the accommodation of potential asymmetry in the CI for the indirect effect and is equally as accurate as bootstrap methods, which are impossible to implement in the case of meta-analysis (Preacher & Selig, 2012). Our mediation model (Figure 1) entailed *six indirect effects* in all (i.e., 2 three-path mediation indirect effects [from gender to trait to participation to leader emergence] and 4 two-path mediation indirect effects [two from gender to trait to leader emergence; and two from gender to trait to participation]; see Table 6). In addition to testing these six indirect effects implied by our model, we also conducted a seventh, overall test of the *total indirect effect* from gender to leader emergence (testing the four component indirect effects from gender to leader emergence together, as a set).

The mediation tests based on joint significance and the mediation tests based on Monte Carlo CI for indirect effects both provided the exact same pattern of statistical significance. For the *six indirect effects* implied by our hypothesized model (see Table 6), all were statistically significant ($p < .05$): (a) gender→ agency→ participation→ leadership emergence (*supporting Hypothesis 5a*), (b) gender→ communion→ participation→ leadership emergence (*supporting Hypothesis 5b*), (c) gender→ agency→ leadership emergence, (d) gender→ communion→ leadership emergence, (e) gender→

TABLE 6 Test of indirect effects for the gender-leadership emergence relationship

Path	Product of coefficients	Indirect effect	95% Monte Carlo CI	Statistically significant
Gender→ Agency → Participation→ Leader Emergence	(.22)*(.35)*(.43)	.03	[.03, .04]	Yes
Gender→ Communion→ Participation→ Leader Emergence	(−.32)*(−.05)*(.43)	.01	[.001, .01]	Yes
Gender→ Agency→ Leader Emergence	(.22)*(.26)	.06	[.04, .07]	Yes
Gender→ Communion→ Leader Emergence	(−.32)*(−.15)	.05	[.04, .06]	Yes
Total Indirect Effect (Combining all 4 paths from Gender to Leader Emergence)		.15	[.12, .17]	Yes
Gender→ Agentic Traits→ Participation	(.22)*(.35)	.08	[.06, .09]	Yes
Gender→ Communal Traits→ Participation	(−.32)*(−.05)	.02	[.004, .03]	Yes

agency→ participation, and (f) gender→ communion→ participation. Further, the *total indirect effect* (including all four indirect effects from gender to leader emergence as a set) was statistically significant (see Table 6), suggesting that agentic traits, communal traits, and participation jointly mediate the relationship between gender and leadership emergence.

5.2 | Moderator analyses

We next examined four moderating factors expected to influence the strength of the gender–leader emergence association (results are reported in Table 1). Notably, the gender gap in leader emergence was supported when looking exclusively at business settings ($\delta = .17$; 95% CI [.04, .25])—a finding that had not been investigated in previous meta-analyses. As compared to business settings, the gender gap was even stronger in lab settings ($\delta = .43$; 95% CI [.29, .56]; $z_{business\ vs.\ lab} = -5.00$, $p < .05$; supporting *H6a*), but surprisingly was weaker in classroom settings ($\delta = .06$; 95% CI [−.06, .16]; $z_{business\ vs.\ classroom} = 2.38$, $p < .05$; contrary to *H6b*).

We further proposed that cultures high in gender egalitarianism would experience a weaker gender gap in leadership emergence. Our results indicate that gender egalitarianism does not moderate the gender–leader emergence association ($b = .10$, $\beta = .03$; $t = .33$, $p > .05$; *n.s.*). Therefore, we fail to support *Hypothesis 7*. To be comprehensive, we also examined a related cross-cultural moderator, gender parity, which captures "national gender gaps on economic, political, education and health criteria" (World Economic Forum, 2015, p. 3; see also Post & Byron, 2015). Gender parity also does not moderate the gender gap in leadership emergence ($b = 1.24$, $\beta = .04$; $t = .42$, $p > .05$; *n.s.*).

In addition, we found that the length of interaction time participants spent together moderated the gender gap in leadership emergence, similar to the findings of Eagly and Karau (1991). The gender difference was large and significant when leader emergence ratings were made upon initial acquaintance ($\delta = .83$; 95% CI [.55, 1.09]) and within the first 20 minutes of interaction ($\delta = .48$; 95% CI [.25, .66])—e.g., similar to a job interview or assessment center exercise. Alternatively, there were not significant gender differences in leadership emergence when ratings occurred after more than 20 minutes in a single interaction ($\delta = .12$; 95% CI [−.02, .26]) or after more than one interaction ($\delta = .09$; 95%

TABLE 7 Multiple regression analysis for the gender gap in leadership emergence

Variable	β	β	β	β	β	β	β
Publication date	−.34*						−.28*
Study setting							
Lab (vs. business settings)		.24*				−.69*	−1.17*
Classroom (vs. business settings)		−.11				−.11	−.26*
Gender egalitarianism			.03				.08
Length of interaction							
Initial acquaintance *and* < 20 minutes vs. > 20 minutes *and* Multiple sessions				.44*		.71*	.63*
Initial acquaintance vs. < 20 minutes				.17*		.17*	.14
> 20 minutes vs. multiple sessions				.02		.44	.66*
Social complexity of task					−.35*		−.22
R^2	.12	.11	.00	.24	.12	.26	.33

Note. Regression coefficients are standardized. * $p < .05$.

CI [−.01, .16]). There was a statistically significant drop in the size of the gender gap from initial acquaintance to the first 20 minutes of interaction ($z = -4.60$, $p < .05$), and from the first 20 minutes to more than 20 minutes in a single interaction ($z = -4.65$, $p < .05$); but the gender gap in leader emergence did not shrink significantly as interaction time increased from more than 20 minutes in a single interaction to multiple sessions ($z = -.82$; n.s.). Overall, these results provide support for *Hypothesis 8*, confirming that longer interaction time shrinks the gender gap in leader emergence (especially when interaction time exceeds 20 minutes).

Finally, we proposed that the social complexity of the task moderates the gender–leadership emergence relationship. Consistent with our prediction, there was a smaller gap observed when the task had a high level of social complexity ($\delta = .14$; 95% CI [.04, .21]), as compared to a low level of social complexity ($\delta = .58$; 95% CI [.39, .75]; $z = -10.77$, $p < .05$), supporting *Hypothesis 9*.

5.2.1 | Multiple regression analysis for hypothesized moderators

To better understand the relative importance of each moderator, we next performed a weighted least squares regression analysis on all moderators simultaneously. As can be seen in Table 7, *publication date* was a statistically significant moderator of the gender gap in leadership emergence (after controlling for all other moderators; $\beta = -.28$, $p < .05$). In addition, results indicated that *study setting* was a significant moderator [(lab settings vs. business settings: $\beta = -1.17$, $p < .05$) and (classroom settings vs. business settings: $\beta = -.26$, $p < .05$)]. Furthermore, *length of interaction time* remained an important moderator (i.e., 20 minutes or less vs. more than 20 minutes; $\beta = .63$, $p < .05$). After controlling for the set of moderators, *social complexity of task* ($\beta = -.22$, $p > .05$, *n.s.*) was no longer a statistically significant moderator. Similarly, *gender egalitarianism* remained a nonsignificant moderator ($\beta = .08$, $p > .05$, *n.s.*). To summarize, publication date, study setting, and length of interaction time offer uniquely important insight into the conditions under which the gender gap in leadership emergence is magnified (whereas gender egalitarianism and social complexity of a task do not). Together, these moderators accounted for 33% of the variance in the gender difference in leadership emergence across studies.

5.3 | Generalizing the theoretical model across moderator categories

We also sought to establish whether our moderators impact the overall theoretical model (Figure 1). We had sufficient data to evaluate whether our path model explained the gender gap in leadership emergence in seven specific contexts (i.e., older studies, newer studies, lab studies, shorter interaction, longer interaction, high social complexity of task, and low social complexity of task). For each of the seven models, the fit statistics were within acceptable ranges (see

TABLE 8 Fit statistics for the theoretical model across moderator categories

Models	Harmonic *N*	χ^2	*df*	RMSEA	CFI	NNFI	SRMR
Old studies	655	4.47	2	.044	.996	.978	.015
New studies	1,399	15.87*	2	.070	.986	.929	.028
Lab only	1,616	4.80	2	.029	.998	.988	.014
Short-term interaction	629	17.87*	2	.112	.970	.850	.043
Long-term interaction	1,327	26.79*	2	.097	.975	.877	.037
Low social complexity	602	10.13*	2	.082	.984	.920	.032
High social complexity	1,313	14.73*	2	.070	.987	.937	.028

Note. Results are reported for Model D. * $p < .05$.

TABLE 9 Test of indirect effects for the gender-leadership emergence relationship across moderator categories

Models	G→A→P→LE	G→C→P→LE	G→A→LE	G→C→LE	Total indirect effect
Old studies	.04*	−.01	.07*	.08*	.18*
New studies	.03*	.01*	.04*	.03*	.11*
Lab only	.03*	.01*	.05*	.05*	.14*
Short-term interaction	.03*	−.01	.08*	.07*	.17*
Long-term interaction	.04*	.02*	.05*	.03*	.14*
Low social complexity	.03*	−.00	.08*	.08*	.19*
High social complexity	.04*	.02*	.05*	.03*	.14*

Note. Monte Carlo confidence intervals were used to test the significance of the indirect effects. * $p < .05$.

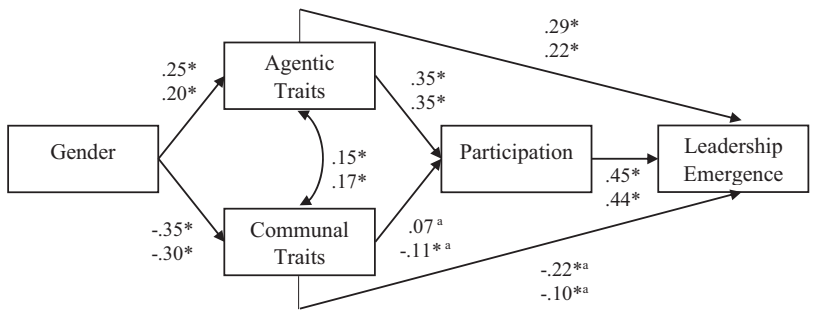

FIGURE 4 Publication date as a moderator of path model results [older studies (1947–1995) vs. newer studies (1996–2017)]
Note: The path coefficient on the top represents older studies; the path coefficient on the bottom represents newer studies. [a]Statistically significant difference between path coefficients. Standardized path coefficients (β's) are presented.

Table 8) and the total indirect effects were statistically significant (see Table 9)—which provides support that agency, communion, and participation work in concert to explain the gender gap in leadership emergence across a variety of situations.

We next evaluated whether the strength of model paths differed across moderator categories (see Figures 4–6). In the theoretical model, three path coefficients in particular appear to be moderated: (a) the communion–participation path, (b) the communion–leadership direct path, and (c) the agency–leadership direct path. For the *communion-participation path*, we observed that the effect from communion to participation was only significant [and was negative] for tasks with high social complexity ($\beta_{\text{high social}} = -.13$, $p < .05$; $\beta_{\text{low social}} = .02$, n.s.; $\Delta\ \chi^2\ (1) = 11.37$, $p < .05$), tasks involving longer duration of interaction ($\beta_{\text{longer}} = -.17$, $p < .05$; $\beta_{\text{short}} = .07$, n.s.; $\Delta\ \chi^2\ (1) = 28.93$, $p < .05$), and for newer studies ($\beta_{\text{newer}} = -.11$, $p < .05$; $\beta_{\text{older}} = .07$, n.s.; $\Delta\ \chi^2\ (1) = 15.98$, $p < .05$).

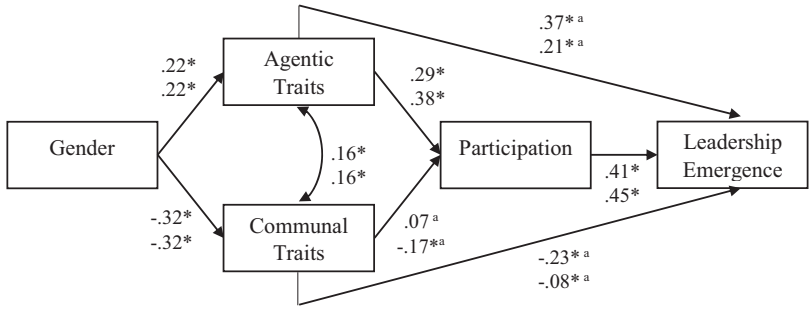

FIGURE 5 Length of interaction as a moderator of path model results

Note: The path coefficient on the top represents short-term interaction; the path coefficient on the bottom represents long-term interaction. [a]Statistically significant difference between path coefficients. Standardized path coefficients (β's) are presented.

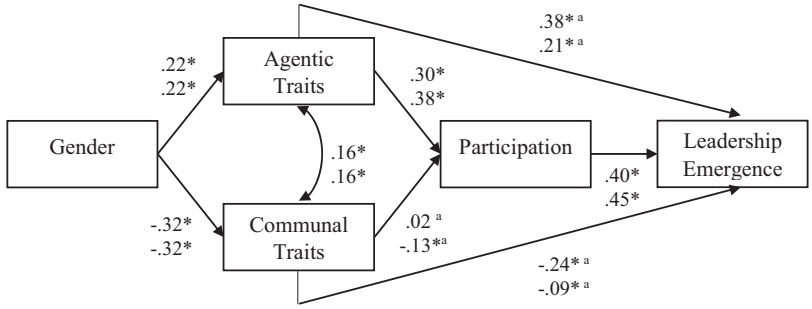

FIGURE 6 Social complexity of task as a moderator of path model results

Note: The path coefficient on the top represents low social complexity; the path coefficient on the bottom represents high social complexity. [a]Statistically significant difference between path coefficients. Standardized path coefficients (β's) are presented.

For the *communion–leadership emergence path*, we observed that the negative effect of communion on leadership was weaker under the same conditions where the indirect effect via participation was stronger (i.e., when the indirect effect of communion on leadership through participation is stronger, then the direct effect of communion on leadership becomes weaker). In particular, the negative direct effect of communion on leadership was stronger for tasks with low social complexity ($\beta_{\text{low social}} = -.24$, $p < .05$; $\beta_{\text{high social}} = -.09$, $p < .05$; $\Delta\ \chi^2\ (1) = 15.22$, $p < .05$), tasks involving short duration of interaction ($\beta_{\text{short}} = -.23$, $p < .05$; $\beta_{\text{longer}} = -.08$, $p < .05$; $\Delta\ \chi^2\ (1) = 16.81$, $p < .05$), and for older studies ($\beta_{\text{older}} = -.22$, $p < .05$; $\beta_{\text{newer}} = -.10$, $p < .05$; $\Delta\ \chi^2\ (1) = 9.00$, $p < .05$). Finally, for the *agency–leadership emergence path*, the positive direct effect of agency on leadership was weaker for tasks with high social complexity ($\beta_{\text{high social}} = .21$, $p < .05$; $\beta_{\text{low social}} = .38$, $p < .05$; $\Delta\ \chi^2\ (1) = 17.77$, $p < .05$) and for tasks requiring longer interaction ($\beta_{\text{longer}} = .21$, $p < .05$; $\beta_{\text{short}} = .37$, $p < .05$; $\Delta\ \chi^2\ (1) = 15.72$, $p < .05$). Thus, with respect to our research question, we found that specific paths in the theoretical model depended upon publication date, social complexity of a task, and the length of interaction time required for a task. Together, these findings underscore the importance of considering context when examining the mechanisms giving rise to the gender gap in leadership emergence.

5.4 | Methodological moderators

We also investigated four methodological moderators. First, we examined various approaches to measuring leader emergence. Election measures ($\delta = .31$; 95% CI [.16, .42]) and questionnaires ($\delta = .16$; 95% CI [.05, .22]) both produced statistically significant gender differences, but rank order measures ($\delta = -.08$; 95% CI [−.27, .25]) did not. In addition, our results indicated that average subject age ($b = .00$, $\beta = .01$, $t = .06$, $p > .05$, *n.s.*) and average group size

($b = .02$, $\beta = .09$, $t = .96$, $p > .05$, *n.s.*) were not statistically significant methodological moderators. Finally, we examined the impact of publication type on the gender–leadership emergence association. The gender–leadership emergence relationship was larger in published studies (published: $\delta = .28$; 95% CI [.17, .37]; unpublished: $\delta = .04$; 95% CI [−.09, .16]; $z = 6.06$, $p < .05$). To further examine publication bias, we performed the trim-and-fill procedure outlined by Duval and Tweedie (2000; implemented in R using the metafor package; Viechtbauer, 2010). The publication-bias adjusted effect is smaller (publication-bias adjusted $d = .14$; vs. observed $d = .21$), but remains statistically significant.

6 | DISCUSSION

Over 25 years have passed since Eagly and Karau's (1991) groundbreaking review of the gender and leadership emergence literature. In the years following this review, extensive work has documented a narrowing gap between the prototypes of leaders and gender role expectations for women (Duehr & Bono, 2006; Eagly, & Sczesny, 2009; Koenig et al., 2011; Powell et al., 2002; Schein, 2001). Although these societal changes indicate that women are increasingly being viewed as leaderlike, it is unclear whether this change has affected actual emergence into leadership roles. Thus, one main objective for our study was to assess change in the gender gap in leadership emergence. To do so, we aggregated 59 years of research—including nearly twice as many studies as the previous meta-analysis—to confirm that men on average tended to emerge more often into leadership roles than did women. Our results for publication date show that whereas the gender difference has been slowly decreasing over time, it continues to exist in the most recent time period. The persistence of a gender gap is consistent with the struggle many organizations still face to reduce gender inequality (U.S. Bureau of Labor Statistics, 2004, 2015).

6.1 | Theoretical contributions

6.1.1 | Theoretical model

Another primary contribution of this paper was examining why a gender difference arises in leader emergence. Social role theory suggests that women may have a leadership disadvantage because they possess characteristics that are perceived to be less consistent with the leadership role (as compared to men; Eagly & Karau, 2002). However, there has been some ambiguity regarding the degree to which the gender gap results from women's lack of agency, abundance of communion, or both. Our comprehensive review of the literature suggests that the answer is both, as agency and communion each exhibited unique incremental effects (directly and indirectly through the mechanism of participatory behavior) that explained the gender gap in emergent leadership (see Figure 3 and Table 6). These findings help to clarify the roles that agency and communion have in the leader emergence process.[11]

Notably, one unexpected result from our theoretical model was that communal traits (which women tend to possess at higher levels) exerted a *negative* impact on leader emergence—rather than a positive effect. Despite substantial evidence establishing that leader effectiveness is composed of both task-oriented and socially-oriented components (Judge, Piccolo, & Ilies, 2004; Stogdill, 1950), the current results suggest that exhibiting concern for others, sensitivity, and other communal proclivities makes one less likely to emerge as a leader, on average. Our finding that communion is devalued when making leader emergence judgments helps to explain recent research showing that women tend to underemerge as leaders given their levels of leadership effectiveness (Lanaj & Hollenbeck, 2015). These authors speculated that although individuals generally recognize communal behaviors as contributing to group success, these same behaviors do not, "trigger automatic perceptions of leadership emergence" (p. 1488). The discrepancy between how communion is perceived to misfit with the leadership role, versus how it actually enhances success in that role, means that organizations may often select the wrong people to fill vacant leadership positions. That is, the negative effect of communal traits on leader emergence in the current study indicates a cognitive bias toward promoting individuals who match stereotypic beliefs about what leaders should be like, but who also may lack some of the necessary interpersonal skills required to be successful in a leadership role.

6.1.2 | Moderators

Beyond our theoretical model, we also examined several substantive moderators of the gender gap in leader emergence. First, regarding the generalizability across different study settings, we observed a significant gender gap in both lab and business settings. At the same time, we found that the gender gap was weaker in business settings than in lab settings. This finding is consistent with past theory suggesting that in business settings, one's coworkers typically have greater access to information about past performance, knowledge, and skills that can be used as a basis for leader selection, reducing the need to rely upon gender when choosing a leader (Harrison et al., 2002; Koch et al., 2015).

A related explanation for the weaker business setting effect could involve a second moderator we examined: length of interaction time. Results for interaction time demonstrated a smaller gender gap when participants had worked together for longer as opposed to shorter lengths of time. This is relevant to the business setting results, because people working together in the business context tend to engage in longer term interactions (in comparison to lab settings). It follows that one reason we found a weaker effect in business settings than lab settings could be that participants in the workplace were more familiar with one another and, therefore, had to rely less heavily on gender stereotypes to make judgments about leader emergence.

We conducted a supplementary analysis of our data to examine this explanation. When we regressed the gender gap onto both study setting (business vs. lab) and length of interaction time (see Table 7), we found that the effect of study setting on the gender gap flipped direction, so that—once length of interaction time was controlled—the gender gap became *larger* in business settings than in lab settings. This finding suggests that business settings can be important in maintaining the gender gap, beyond the fact that they involve longer interaction times than do lab settings. Further, these results imply that in work settings there may be additional incentives to vie for leadership (e.g., promotions and pay raises) that disadvantage women, or perhaps that the work setting has stronger norms to adhere to gender stereotypes than are present in lab or educational contexts. This being said, we urge readers to bear in mind that there were fewer available primary studies for the gender gap in business settings compared to lab and classroom settings (i.e., there were a total of 3,117 participants across 11 primary studies for the gender gap in business settings vs. 39 primary studies for classroom settings and 82 primary studies for lab settings). This highlights that additional research is needed to examine the gender gap in the business context.

Further, we were surprised to find that the gender difference in leader emergence did not vary due to cross-cultural differences in gender egalitarianism. This finding failed to support our prediction derived from social role theory (Eagly & Karau, 2002) that the gender gap would be weaker in higher gender egalitarian cultures. However, it is important to note that there was a limited number of countries represented in our meta-analysis (i.e., there were nine different countries). For this reason, we urge readers to interpret our null results with caution and encourage continued research to examine whether our model extends to countries with different traditional gender roles.

Finally, we found that the gender difference in leader emergence was larger when individuals were performing tasks requiring a low degree of social complexity (tasks requiring little interpersonal coordination). However, when we examined our substantive moderators simultaneously (Table 7), social complexity no longer exerted a statistically significant influence on the gender gap, after controlling for other moderators. Based on these results, we concluded that publication date, study setting, and length of interaction time were the primary contextual factors influencing the strength of the gender gap in leader emergence.

6.1.3 | Generalizing the theoretical model across contexts

We also examined whether our integrated theoretical model (Figure 1) was impacted by the aforementioned moderators. These moderator analyses for the theoretical model were limited to the three moderator variables for which there were enough primary studies to perform the analyses (i.e., publication date, length of interaction time, and social complexity of a task). We found that our model generalized across all of the contexts examined, providing support that agentic traits, communal traits, and participatory behavior work in concert to explain the gender gap in leader emergence across a variety of situations.

At the same time, we observed that three of the specific paths in our model became stronger/weaker depending on context. The three model paths that were moderated are: (a) the direct effect from communion to leader emergence, (b) the direct effect from agency to leader emergence, and (c) the effect from communion to participation. In particular, the two direct effects of personality traits on leader emergence (i.e., the *negative* communion–leader emergence link, and the *positive* agency–leader emergence link), were both weaker for longer interaction times and higher social complexity tasks. These results are consistent with social role theory, because more stereotypically feminine tasks (requiring longer durations of social interaction and interpersonal coordination) were found to simultaneously weaken both the agency *advantage* and the communion *disadvantage* (Eagly & Karau, 2002; Koenig et al., 2011). Next, with regard to time period (i.e., publication date), the direct agency–leader emergence path did not significantly differ between older and newer studies, whereas the negative communion–leader emergence association was weaker in newer studies. This suggests that agentic traits have been considered prototypical leader attributes to a similar extent over time periods (the agentic advantage has remained consistent), but communal traits have become less detrimental over time (the communal disadvantage has waned).

The third path coefficient that was moderated is the communion-participation path. Surprisingly, the communion-participation path was *negative* (indicating communal individuals participated *less*), under three conditions: (a) for tasks requiring a high degree of social complexity (interpersonal coordination), (b) for tasks involving a longer length of interaction among participants, and (c) in newer studies. One potential explanation for the negative communion-participation path may be that communal individuals attempt to boost group harmony by adopting more of a supportive, listening role when situations call for greater interpersonal coordination, or when the longer duration of a task enables the development of a closer interpersonal relationship or rapport. Consistent with this idea, James and Drakich (1993) observed that whereas, "taking and holding the floor for long periods" is indicative of an agentic strategy, communal individuals are, "careful not to take up a disproportionate amount of talking time" (p. 285). In addition, newer studies tended to employ more socially complex and longer duration tasks, so the idea that communal individuals may adopt more facilitative (rather than active) roles in these situations might also help to explain the time period/publication date moderator effect. We hope these results encourage additional investigation into whether communal individuals are more likely to take on facilitative, listening roles in group situations that require socially complex, longer term interpersonal interactions. Although communal traits have a modest negative effect on participation in studies that use longer term interaction and higher social complexity tasks, we should note that the *overall* negative effect (ignoring the moderators) of communal traits on participation is relatively weak (and the zero-order correlation is not statistically significant). That is, overall, participation behavior plays a stronger role in the agency–leader emergence link than in the communion-leadership emergence link (Figure 3).

6.2 | Practical implications

Gender differences in behavioral participation in the workplace have recently received increased attention, due in part to the controversial best-selling book *Lean In* (2013) by Sheryl Sandberg (Chief Operating Officer of Facebook), in which she argues: "If we want a world with greater equality, we need to acknowledge that women are less likely to keep their hands up. We need institutions and individuals to notice and correct for this behavior…and women have to learn to keep their hands up" (p. 36). These provocative remarks have garnered both praise for their candor and criticism for victim blaming—that is, some perceive that this advice blames women for gender inequality in the workplace (e.g., Brooks, 2014; Hoye, 2013). Despite this controversy, the above quote implicates the role of participation in group discussions as a potentially important factor in the emergence of women into leadership positions. We would like to make clear, however, that we do not recommend that women who want to be regarded as leaders should simply engage in more agentic, participatory behaviors. Participating in group discussion is a complex phenomenon, and there are likely many reasons why men are participating more than women (Rudman, Moss-Racusin, Glick, & Phelan, 2012). For example, there is a well-documented backlash effect in which women are perceived more negatively for displaying the same agentic behaviors as men (see review by Williams & Tiedens, 2016). To us, therefore, the participation mechanism for

the gender gap in leader emergence reflects a complex interweaving of motivation to speak out and oppressive barriers that operate together to curb women's participation in group discussions.

We believe that organizational interventions stemming from the current findings should be targeted at educating the perceivers of agentic and communal behaviors rather than targeted at changing women's behavior. We prefer this approach because we do not presume that advocating for women to suppress their traits and behaviors or change their behaviors to accommodate gender stereotypes is an advisable long-term strategy to achieve gender equality. As such, one way an organization could apply our results (i.e., the finding that agency helps and communal traits hurt individuals' chances of emerging as leaders) would be to provide training to make employees aware of the natural tendency to overemphasize agency and undervalue communion when selecting leaders. This tendency not only relates to gender inequality, but potentially also prevents organizations from making the best possible hiring/promotion decision (regardless of gender).

In addition, organizations can implement training programs designed to help women overcome subtle biases in the workplace, such as being interrupted more frequently and not receiving credit for their suggestions (Bennett, 2015; McMillan, Clifton, McGrath, & Gale, 1977). For instance, organizations may teach employees to "amplify" or repeat the contributions of others within meetings (always citing the individual who initially offered the suggestion; Kantor, 2016). In doing so, this may ensure that both men's and women's participation in group discussion is not only heard, but attributed to the correct individual.

Finally, the current study suggested that gender differences in leader emergence were attenuated when raters had the opportunity to interact with the target individual for a longer length of time or over multiple occasions. One reason for this may be that raters rely less on stereotypes after longer opportunities to interact with others, because they are able to instead base ratings on observed behaviors (as suggested by Harrison, Price, & Bell, 1998). Thus, a final piece of advice aimed at reducing gender bias in hiring and promotion decisions is that organizations should conduct multiple employment or promotion interviews with each candidate (to provide sufficient opportunity for evaluators to become familiar with applicants for leadership positions).

6.3 | Limitations and future research directions

First, because a portion of our primary studies collected data using cross-sectional, nonexperimental research designs, we were unable to make definitive claims regarding causality. That being said, the ordering of our causal model is consistent with existing theory—societal gender norms are thought to precede personality development (Eagly & Wood, 2012; Ruble, Martin, & Berenbaum, 1998), personality is traditionally considered to consist of relatively stable tendencies that drive behavior (Funder, 1991; Ozer & Benet-Martínez, 2006), and leadership emergence is conceptualized as the result of both personality traits (Judge et al., 2004) and behavior (Bass, 1954).

Further, every methodological design, including meta-analysis, has strengths and limitations. Typically, meta-analysis is a highly powered design that allows researchers to study moderators which vary across studies, but it does not usually permit studying individual-level interaction effects. Thus, although we would have liked to examine the possible consequences for women of participating more frequently in group discussions (participation × gender interaction) and whether agency and communion interact to predict leader emergence (agency × communion interaction), it was not possible using the meta-analytic method. Whereas this is a limitation of all meta-analyses, primary studies have investigated such interaction effects, confirming (as discussed above) that men and women are at times viewed differently when performing the same behaviors (Rudman, 1998) and that individuals are more likely to emerge as leaders when they simultaneously possess high levels of agency and communion (Kent & Moss, 1994). Nonetheless, the current paper makes a contribution by teasing apart the unique effects of agency and communion—helping to better understand traits that are often conflated in theoretical discussions of the underlying causes of leader emergence.

Another limitation of the meta-analytic method is that researchers can only examine relationships that have been studied by past researchers. In the current meta-analysis, there were not enough primary studies to permit us to examine our theoretical model across all of our moderators. In particular, there were not enough studies to investigate the full model in the business setting. Nonetheless, by definition, personality traits are relatively stable across situations

(Funder, 1991), indicating that agency and communion would manifest to a similar degree in the workplace (as compared to other situations); and there is further evidence that participation in group discussions is linked to leader emergence in business contexts (Gerpott, Lehmann-Willenbrock, & van Vugt, 2016). These findings suggest that agency, communion, and participation likely mediate the gender gap in the business setting; however, future research is needed to confirm this proposition.

Finally, although agentic traits, communal traits, and participatory behavior operate in concert to mediate the gender effect on leader emergence, there exist other potential mediators of the gender gap. Based on the recommendation of a reviewer, we performed a broad literature search to investigate other possible mediators. This search revealed two additional possible mechanisms of the gender gap that had received sufficient research attention to meta-analyze (i.e., there were at least three primary studies linking these variables to leader emergence). These potential mediators were leader–member exchange (LMX) and past job performance. Whereas LMX was positively related to leader emergence ($\rho = .21$, $k = 3$, $N = 794$, 95 % CI [.12, .24]), there was not a statistically significant gender–LMX association ($\delta = .01$, $k = 241$, $N = 72{,}066$, 95 % CI [−.01, .05]). Given this null finding, LMX cannot serve as a mediator for the gender gap in leader emergence (Mackinnon et al., 2002). Likewise, although past job performance was positively related to leader emergence ($\rho = .21$, $k = 3$, $N = 508$, 95 % CI [.08, .29]), existing research has shown that on average women tend to have *higher* levels of job performance than men ($\delta = -.11$; $k = 61$; $N = 45{,}733$; Roth et al., 2012). Because women exhibit higher average job performance, we also ruled out past job performance as an alternative mechanism for the gender gap in leader emergence. Together, these findings further clarify our understanding of the gender gap in leader emergence by eliminating LMX and past job performance as explanations for why men are more likely to emerge into leadership roles than women.

7 | CONCLUSION

The current study shows that the gender gap in leader emergence is slowly shrinking over time, but has not disappeared. We also find support for a theoretical model that offers an explanation for why a gender gap exists in leadership emergence. In particular, agentic traits are beneficial whereas communal traits are detrimental to emerging into leadership roles, and these traits operate in part through behavioral participation in group discussions. Further, we demonstrate that the theoretical model generalizes across a variety of conditions. At the same time, several contextual moderators (i.e., duration of interaction, social complexity, and publication date) influence the relative strengths of the model paths. It is our hope that empirical evidence of a continued gender difference in leadership emergence (as well as insight into both the trait and behavioral factors that explain this difference) will increase efforts to understand women's pathways into leadership.

ACKNOWLEDGMENTS

This work was partially supported by a grant awarded to Emily Grijalva and Katie Badura from the University at Buffalo Gender Institute. We thank Alice Eagly, Terry Blum, G. James Lemoine, and Timothy Maynes for their helpful comments on earlier versions of this manuscript.

NOTES

[1] Koenig, Eagly, Mitchell, and Ristikari (2011) empirically summarize research across three experimental paradigms, including: Schein's (1973) think manager-think male paradigm, Powell and Butterfield's (1979) agency–communion paradigm, and Shinar's (1975) masculinity-femininity paradigm. Studies from each of these paradigms investigate the degree of correspondence between gender stereotypes and stereotypes held about leaders, by having participants rate the characteristics of a fictitious leader (e.g., "a good manager" or "a successful middle manager"). The current study differs from Koenig, Eagly, Mitchell, and Ristikari's (2011) focus on stereotypes, by instead examining the gender difference in a specific leadership outcome (i.e., leader emergence).

[2] Notably, two moderators previously investigated by Eagly and Karau (1991)—interaction time and social complexity—are investigated for two reasons. First, it is unknown whether these variables *uniquely* moderate the gender gap, after controlling

for other potential moderators (i.e., publication date, study setting, and cultural gender egalitarianism) (see Table 7). Second, it is unknown how these moderators might influence our integrated theoretical model (Figure 1) that explains the gender gap (see Figures 5 and 6).

[3] We note that social complexity of task differs from some other well-known task taxonomies. Wood (1987) observed that, "although a number of schemes are available for classifying tasks (e.g., Laughlin, 1980; McGrath, 1984; Shaw, 1973), none of these are specifically focused on the amount of task versus social activity required" (p. 55). As such, when studying the gender gap, we follow Eagly and Karau (1991) by focusing here on the social complexity of tasks rather than using another taxonomy. We want to emphasize that Eagly and Karau's (1991) concept of *social complexity* is not the same thing as general *job complexity* (Dunham, 1976; Hunter & Hunter, 1984). In particular, the social complexity of tasks describes whether the task involves interpersonal coordination, whereas job complexity refers to the information processing demands of a job (Hunter & Hunter, 1984; Judge, Bono, & Locke, 2000; Roth, Bevier, Bobko, Switzer, & Tyler, 2001).

[4] Notably, for the relationship between gender and participation, Leaper and Ayres' (2007) meta-analysis had included 33 samples for which there was insufficient data to calculate a primary study effect size (e.g., if an unreported effect was not statistically significant, Leaper and Ayres coded it as zero). For a list of primary studies from Leaper and Ayers (2007) that we excluded, see Online Supplemental Appendix E.

[5] We note that Eagly and Karau's (1991) meta-analysis also included participation in group discussion, but coded participation as a form of leadership emergence itself, rather than as a behavioral mechanism. A considerable amount of research supports the conceptual and empirical distinctiveness of the two constructs of participation and leader emergence, including a meta-analysis showing that the two are related, but not synonymous ($\rho = .55$, $k = 33$, $N = 3{,}611$; Mullen, Salas, & Driskell, 1989).

[6] Donnelly and Twenge (2016) studied changes in men's and women's agency and communion over time but did not report the overall gender difference in agency or communion (see also Twenge, 1997).

[7] For publication date, we examined two categories: older studies (1947–1995) and newer studies (1996–present). Our category selection is consistent with Paustian-Underdahl, Walker, and Woehr's (2014) meta-analysis of gender and leadership effectiveness. These authors combined labor force statistics and Kanter's (1977) research on token status percentages to identify theoretically meaningful categories. Due to limited availability of primary studies, we collapsed the date ranges into two categories (rather than four date categories). For a more thorough description, please see Paustian-Underdahl, Walker, & Woehr (2014, p. 1135).

[8] Shorter interactions (20 minutes or less) versus longer interactions (more than 20 minutes).

[9] Eagly and Karau's (1991) meta-analysis included 74 separate samples, and the gender–leader emergence effect size was $d = .32$ (uncorrected). In comparison, the current paper included 136 separate samples and the uncorrected gender difference was $d = .21$ ($\delta = .22$, corrected for unreliability).

[10] At the request of a reviewer, we estimated two additional post hoc models that each extend Model D: (a) the first model (Model E) added a direct path from gender to participation ($\chi^2_{(1)} = 8.72$, RMSEA = .060, CFI = .995, NNFI = .951, SRMR = .012), direct path $\beta = -.01$ ($p > .05$, *n.s.*), and (b) the second model (Model F) added a direct path from gender to leader emergence ($\chi^2_{(1)} = 0.43$, RMSEA = .000, CFI = 1.00, NNFI = 1.00, SRMR = .003), direct path $\beta = -.06$ ($p < .05$). The small negative direct path coefficient in Model F ($\beta = -.06$ from gender to leader emergence) is indicative of a suppression effect (Tzelgov & Henik, 1991), which suggests that, after controlling for our set of mediators (i.e., agentic traits, communal traits, and participation), women are very slightly *more* likely to emerge as leaders (i.e., the gender gap in leadership emergence is *more than* completely explained by the specified trait and behavioral mechanisms).

[11] In related work, Fiske, Cuddy, and Glick (2007) have cited the agentic and communal gender stereotypes as examples of two fundamental underlying dimensions in the domain of person perception: competence and warmth. Future research should extend our theoretical model to observers' perceptions of agency/competence and communion/warmth.

ORCID


Katie L. Badura http://orcid.org/0000-0002-6323-2568

REFERENCES

[*References for meta-analytic primary studies can be found in online supplemental Appendix D.*]

Abele, A. E., & Wojciszke, B. (2007). Agency and communion from the perspective of self versus others. *Journal of Personality and Social Psychology*, *93*, 751–763. https://doi.org/10.1037/0022-3514.93.5.751

Aguinis, H., Gottfredson, R. K., & Wright, T. A. (2011). Best-practice recommendations for estimating interaction effects using meta-analysis. *Journal of Organizational Behavior*, *32*, 1033–1043. Retrieved from https://doi.org/10.1002/job.719

Allport, G. W. (1951). Personality–A psychological interpretation. London: Constable, pp. xiv + 588, price 16s. *British Journal of Educational Pyschology*, *13*, 48–50. Retrieved from https://doi.org/10.1111/j.2044-8279.1943.tb02719.x

Allport, G. W. (1966). Traits revisited. *American Psychologist*, *21*, 1–10. Retrieved from https://doi.org/10.1037/h0023295

Aries, E. J. (1976). Interaction patterns and themes of male, female, and mixed sex groups. *Small Group Behavior*, *7*, 7–18. Retrieved from https://doi.org/10.1177/104649647600700102

Aries, E. J. (1982). Verbal and nonverbal behavior in single sex and mixed-sex groups: Are traditional sex roles changing? *Psychological Reports*, *51*, 127–134. Retrieved from https://doi.org/10.2466/pr0.1982.51.1.127

Askinas, B. E. (1971). *The impact of coeducational living on peer interaction*. Unpublished doctoral dissertation, Stanford University, Stanford, CA.

Avolio, B. J. (2007). Promoting more integrative strategies for leadership theory-building. *American Psychologist*, *62*, 25–33. https://doi.org/10.1037/0003-066X.62.1.25

Bakan, D. (1966). *The duality of human existence: Isolation and communion in Western man*. Chicago, IL: Rand McNally.

Bales, R. F. (1950). *Interaction process analysis: A method for the study of small groups*. Oxford, England: Addison-Wesley.

Bales, R. F., Strodtbeck, F. L., Mills, T. M., & Roseborough, M. E. (1951). Channels of communication in small groups. *American Sociological Review*, *16*, 461–468. https://doi.org/10.2307/2088276

Bass, B. M. (1949). An analysis of the leaderless group discussion. *Journal of Applied Psychology*, *33*, 527–533. https://doi.org/10.1037/h0058164

Bass, B. M. (1954). The leaderless group discussion. *Psychological Bulletin*, *51*, 465–492.

Bass, B. M. (1990). *Bass & Stogdill's handbook of leadership: Theory, research, and managerial applications* (3rd ed.). New York, NY: Free Press.

Bem, S. L. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology*, *42*, 155–162. https://doi.org/10.1037/h0056881

Bennett, J. (2015). How not to be "manterrupted" in meetings. *Time*. Retrieved from https://time.com/3666135/sheryl-sandberg-talking-while-female-manterruptions/

Berger, J., Rosenholtz, S. J., & Zelditch, M. (1980). Status organizing processes. *Annual Review of Sociology*, *6*, 479–508. Retrieved from https://doi.org/10.1146/annurev.so.06.080180.002403

Biesanz, J. C., West, S. G., & Millevoi, A. (2007). What do you learn about someone over time? The relationship between length of acquaintance and consensus and self-other agreement in judgments of personality. *Journal of Personality and Social Psychology*, *92*, 119–135. Retrieved from https://doi.org/10.1037/0022-3514.92.1.119

Block, J. H. (1973). Conceptions of sex role: Some cross-cultural and longitudinal perspectives. *American Psychologist*, *28*, 512–526. https://doi.org/10.1037/h0035094

Brooks, R. (2014). Recline, don't 'Lean In' (why I hate Sheryl Sandberg). *The Washington Post*. Retrieved from https://www.washingtonpost.com/blogs/she-the-people/wp/2014/02/25/recline-dont-lean-in-why-i-hate-sheryl-sandberg/

Brown, T. M., & Miller, C. E. (2000). Communication networks in task-performing groups: Effects of task complexity, time pressure, and interpersonal dominance. *Small Group Research*, *31*, 131–157. Retrieved from https://doi.org/10.1177/104649640003100201

Buss, D. M. (1990). Unmitigated agency and unmitigated communion: An analysis of the negative components of masculinity and femininity. *Sex Roles*, *22*, 555–568. Retrieved from https://doi.org/10.1007/BF00288234

Carli, L. L. (2001). Gender and social influence. *Journal of Social Issues*, *57*, 725–741. Retrieved from https://doi.org/10.1111/0022-4537.00238

Cashdan, E. (1998). Smiles, speech, and body posture: How women and men display sociometric status and power. *Journal of Nonverbal Behavior*, *22*, 209–228. Retrieved from 10.1023/A:1022967721884

Correll, S. J., & Ridgeway, C. L. (2006). Expectation states theory. In J. Delamater (Ed.), *Handbook of social psychology* (pp. 29–51). New York, NY: Springer.

Crosbie, P. V. (1979). The effects of sex and size on status ranking. *Social Psychology Quarterly*, *42*, 340–354. https://doi.org/10.2307/3033804

Davis, M., & Songer-Nocks, E. (1978). *Comparing leader-group relations in male and female groups*. Paper presented at the meeting of the Southeastern Psychological Association, Atlanta.

DeRue, D. S., Nahrgang, J. D., Wellman, N. E. D., & Humphrey, S. E. (2011). Trait and behavioral theories of leadership: An integration and meta-analytic test of their relative validity. *Personnel Psychology*, *64*, 7–52. Retrieved from https://doi.org/10.1111/j.1744-6570.2010.01201.x

Diekman, A. B., & Eagly, A. H. (2000). Stereotypes as dynamic constructs: Women and men of the past, present, and future. *Personality and Social Psychology Bulletin*, *26*, 1171–1188. Retrieved from https://doi.org/10.1177/0146167200262001

Dipboye, R. L., & Flanagan, M. F. (1979). Research settings in industrial and organizational psychology: Are findings in the field more generalizable than in the laboratory? *American Psychologist*, *34*, 141–150. https://doi.org/10.1037/0003-066X.34.2.141

Donnelly, K., & Twenge, J. M. (2016). Masculine and feminine traits on the Bem Sex-Role inventory, 1993–2012: A cross-temporal meta-analysis. *Sex Roles*, 1–10. Retrieved from https://doi.org/10.1007/s11199-016-0625-y

Duehr, E. E., & Bono, J. E. (2006). Men, women, and managers: Are stereotypes finally changing? *Personnel Psychology*, *59*, 815–846. Retrieved from https://doi.org/10.1111/j.1744-6570.2006.00055.x

Dunham, R. B. (1976). The measurement and dimensionality of job characteristics. *Journal of Applied Psychology*, *61*, 404–409. Retrieved from https://doi.org/10.1037/0021-9010.61.4.404

Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot–based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, *56*, 455–463. Retrieved from https://doi.org/10.1111/j.0006-341X.2000.00455.x

Eagly, A. H. (1987). *Sex differences in social behavior: A social-role interpretation*. Hillsdale, NJ, England: Erlbaum, Inc.

Eagly, A. H. (2016). When passionate advocates meet research on diversity, does the honest broker stand a chance? *Journal of Social Issues*, *72*, 199–222. Retrieved from https://doi.org/10.1111/josi.12163

Eagly, A. H., & Carli, L. L. (2003). The female leadership advantage: An evaluation of the evidence. *The Leadership Quarterly*, *14*, 807–834. Retrieved from https://doi.org/10.1016/j.leaqua.2003.09.004

Eagly, A. H., & Karau, S. J. (1991). Gender and the emergence of leaders: A meta-analysis. *Journal of Personality and Social Psychology*, *60*, 685–710. https://doi.org/10.1037/0022-3514.60.5.685

Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review*, *109*, 573–598. https://doi.org/10.1037/0033-295X.109.3.573

Eagly, A. H., & Sczesny, S. (2009). Stereotypes about women, men, and leaders: Have times changed? In M. Barreto, M. K. Ryan, & M. T. Schmitt (Eds.), *The glass ceiling in the 21st century: Understanding barriers to gender equality* (pp. 21–47). Washington, DC: American Psychological Association Books.

Eagly, A. H., & Steffen, V. J. (1984). Gender stereotypes stem from the distribution of women and men into social roles. *Journal of Personality and Social Psychology*, *46*, 735–754. Retrieved from https://doi.org/10.1037/0022-3514.46.4.735

Eagly, A. H., & Wood, W. (2012). Social role theory. In P. A. M. Van Lange, A. W. Kruglanski, & E. T. Higgins (Eds.), *Handbook of theories of social psychology* (pp. 458–476). London, UK: Sage.

Emrich, C. G., Denmark, F. L., & Den Hartog, D. N. (2004). Cross-cultural differences in gender egalitarianism: Implications for societies, organizations and leaders. In R. J. House, P. J. Hanges, M. Javidan, P. W. Dorfman, & V. Gupta (Eds.), *Culture, leadership, and organizations: The GLOBE study of 62 societies* (pp. 343–394). Thousand Oaks, CA: Sage Publications Inc.

Ensari, N., Riggio, R. E., Christian, J., & Carslaw, G. (2011). Who emerges as a leader? Meta-analyses of individual differences as predictors of leadership emergence. *Personality and Individual Differences*, *51*, 532–536. https://doi.org/10.1016/j.paid.2011.05.017

Fiske, S. T., Cuddy, A. J., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences*, *11*, 77–83. Retrieved from https://doi.org/10.1016/j.tics.2006.11.005

Fondas, N. (1997). Feminization unveiled: Management qualities in contemporary writings. *Academy of Management Review*, *22*, 257–282. Retrieved from https://doi.org/10.5465/AMR.1997.9707180266

Fulton, R. F. (1981). *The influence of the emergent informal group leader on unanimous group decisions*. Unpublished master's thesis, Oklahoma State University, Stillwater, OK.

Funder, D. C. (1991). Global traits: A neo-Allportian approach to personality. *Psychological Science*, *2*, 31–39. Retrieved from https://doi.org/10.1111/j.1467-9280.1991.tb00093.x

Gawronski, B. (2004). Theory-based bias correction in dispositional inference: The fundamental attribution error is dead, long live the correspondence bias. *European Review of Social Psychology*, *15*, 183–217. https://doi.org/10.1080/10463280440000026

Gerpott, F. H., Lehmann-Willenbrock, N., & van Vugt, M. (2016). *Emergent leadership in self-directed teams: Communication shifts over time*. Paper presented at the 76th annual meeting of the Academy of Management, Los Angeles, CA.

Ghiselli, E. E., Campbell, J. P., & Zedeck, S. (1981). *Measurement theory for the behavioral sciences: Origin & evolution*. San Francisco, CA: WH Freeman & Company.

Gilbert, D. T. (1998). Ordinary personology. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vol. *2*, pp. 89–150). Boston, MA: McGraw-Hill.

Grijalva, E., Harms, P. D., Newman, D. A., Gaddis, B. H., & Fraley, R. C. (2015a). Narcissism and leadership: A meta-analytic review of linear and nonlinear relationships. *Personnel Psychology*, *68*, 1–47. Retrieved from https://doi.org/10.1111/peps.12072

Grijalva, E., Newman, D. A., Tay, L., Donnellan, M. B., Harms, P. D., Robins, R. W., & Yan, T. (2015b). Gender differences in narcissism: A meta-analytic review. *Psychological Bulletin*, *141*, 261–310. https://doi.org/10.1037/a0038231

Hall, K. E. P. (1973). *Sex differences in initiation and influence in decision-making among prospective teachers*. Unpublished doctoral dissertation, Stanford University, Stanford, CA.

Harrison, D. A., Price, K. H., & Bell, M. P. (1998). Beyond relational demography: Time and the effects of surface- and deep-level diversity on work group cohesion. *Academy of Management Journal*, *41*, 96–107. Retrieved from https://doi.org/10.2307/256901

Harrison, D. A., Price, K. H., Gavin, J. H., & Florey, A. T. (2002). Time, teams, and task performance: Changing effects of surface-and deep-level diversity on group functioning. *Academy of Management Journal*, *45*, 1029–1045. Retrieved from https://doi.org/10.2307/3069328

Hayes, A. F., & Scharkow, M. (2013). The relative trustworthiness of inferential tests of the indirect effect in statistical mediation analysis: Does method really matter? *Psychological Science*, *24*, 1918–1927. Retrieved from https://doi.org/10.1177/0956797613480187

Heilman, M. E. (2001). Description and prescription: How gender stereotypes prevent women's ascent up the organizational ladder. *Journal of Social Issues*, *57*, 657–674. Retrieved from https://doi.org/10.1111/0022-4537.00234

Heilman, M. E. (2012). Gender stereotypes and workplace bias. *Research in Organizational Behavior*, *32*, 113–135. https://doi.org/10.1016/j.riob.2012.11.003

Hong, Y. (2005). *Motivation to lead: Antecedents and resulting leader emergence*. Unpublished master's thesis, St. Mary's University, Canada.

Hong, Y., Catano, V. M., & Liao, H. (2011). Leader emergence: The role of emotional intelligence and motivation to lead. *Leadership & Organization Development Journal*, *32*, 320–343. https://doi.org/10.1108/01437731111134625

House, R. J., & Javidan, M. (2004). Overview of GLOBE. In R. J. House, P. J. Hanges, M. Javidan, P. W. Dorfman, & V. Gupta (Eds.), *Culture, leadership, and organizations: The GLOBE study of 62 societies* (pp. 9–28). Thousand Oaks, CA: Sage Publications Inc.

Hoye, S. (2013). The myth of balancing motherhood and a successful career. *CNN*. Retrieved from https://www.cnn.com/2013/03/11/living/sandberg-advice-working-mothers/index.html

Hunter, J. E., & Hunter, R. F. (1984). Validity and utility of alternative predictors of job performance. *Psychological Bulletin*, *96*, 72–98. Retrieved from https://doi.org/10.1037/0033-2909.96.1.72

Ibarra, H., Carter, N. M., & Silva, C. (2010). Why men still get more promotions than women. *Harvard Business Review*, *88*, 80–85.

Ickes, W. (1981). Sex-role influences in dyadic interaction: A theoretical model. In C. Mayo & N. M. Henley (Eds.), *Gender and nonverbal behavior* (pp. 95–128). New York, NY: Springer.

James, D., & Drakich, J. (1993). Understanding gender differences in amount of talk: A critical review of research. In D. Tannen (Ed.), *Gender and conversational interaction* (pp. 281–306). New York, NY: Oxford University Press.

James, R. M. (1956). Status and competence of jurors. *American Journal of Sociology*, *64*, 563–570. https://doi.org/10.1086/222589

Jose, P. E., & McCarthy, W. J. (1988). Perceived agentic and communal behavior in mixed-sex group interactions. *Personality and Social Psychology Bulletin*, *14*, 57–67. Retrieved from https://doi.org/10.1177/0146167288141006

Judge, T. A., Bono, J. E., Ilies, R., & Gerhardt, M. W. (2002). Personality and leadership: A qualitative and quantitative review. *Journal of Applied Psychology*, *87*, 765–780.

Judge, T. A., Bono, J. E., & Locke, E. A. (2000). Personality and job satisfaction: The mediating role of job characteristics. *Journal of Applied Psychology*, *85*, 237–249. Retrieved from https://doi.org/10.1037/0021-9010

Judge, T. A., Piccolo, R. F., & Ilies, R. (2004). The forgotten ones? The validity of consideration and initiating structure in leadership research. *Journal of Applied Psychology*, *89*, 36–51. https://doi.org/10.1037/0021-9010.87.4.765

Kanter, R. (1977). Some effects of proportions on group life: Skewed sex ratios and responses to token women. *American Journal of Sociology*, *82*, 965–990. REtrieved from 10.1086/226425

Kantor, D. (2016). Bringing "shine theory" to the corporate world. *The Huffington Post*. Retrieved from https://www.huffingtonpost.co.uk/daria-kantor/bringing-shine-theory-to-_b_12261134.html

Kennedy, J. A., Anderson, C., & Moore, D. A. (2013). When overconfidence is revealed to others: Testing the status-enhancement theory of overconfidence. *Organizational Behavior and Human Decision Processes*, *122*, 266–279. Retrieved from https://doi.org/10.1016/j.obhdp.2013.08.005

Kent, R. L., & Moss, S. E. (1994). Effects of sex and gender role on leader emergence. *Academy of Management Journal*, *37*, 1335–1346. Retrieved from https://doi.org/10.2307/256675

Klein, H. M., & Willerman, L. (1979). Psychological masculinity and femininity and typical and maximal dominance expression in women. *Journal of Personality and Social Psychology*, *37*, 2059–2070. Retrieved from https://psycnet.apa.org/doi/10.1037/0022-3514.37.11.2059

Koch, A. J., D'Mello, S. D., & Sackett, P. R. (2015). A meta-analysis of gender stereotypes and bias in experimental simulations of employment decision making. *Journal of Applied Psychology*, *100*, 128–161. https://doi.org/10.1037/a0036734

Koenig, A. M., & Eagly, A. H. (2014). Evidence for the social role theory of stereotype content: Observations of groups' roles shape stereotypes. *Journal of Personality and Social Psychology*, *107*, 371–392. Retrieved from https://doi.org/10.1037/a0037215

Koenig, A. M., Eagly, A. H., Mitchell, A. A., & Ristikari, T. (2011). Are leader stereotypes masculine? A meta-analysis of three research paradigms. *Psychological Bulletin*, *137*, 616–642. Retrieved from https://psycnet.apa.org/doi/10.1037/a0023557

Lanaj, K., & Hollenbeck, J. R. (2015). Leadership over-emergence in self-managing teams: The role of gender and countervailing biases. *Academy of Management Journal*, *58*, 1476–1494. https://doi.org/10.5465/amj.2013.0303

Laughlin, P. R. (1980). Social combination processes of cooperative, problem-solving groups at verbal, intellective tasks. In M. Fishbein (Ed.), *Progress in social psychology* (Vol. *1*, pp. 127–155). Hillsdale, NJ: Erlbaum.

Leaper, C. (1987). Agency, communion, and gender as predictors of communication style and being liked in adult male-female dyads. *Sex Roles*, *16*, 137–149. Retrieved from https://doi.org/10.1007/BF00289645

Leaper, C., & Ayres, M. M. (2007). A meta-analytic review of gender variations in adults' language use: Talkativeness, affiliative speech, and assertive speech. *Personality and Social Psychology Review*, *11*, 328–363. https://doi.org/10.1177/1088868307302221

Leaper, C., & Friedman, C. K. (2007). The socialization of gender. In J. Grusec & P. Hastings (Eds.), *Handbook of socialization: Theory and research* (pp. 561–587). New York, NY: Guilford.

Littlepage, G. E., Schmidt, G. W., Whisler, E. W., & Frost, A. G. (1995). An input-process-output analysis of influence and performance in problem-solving groups. *Journal of Personality and Social Psychology*, *69*, 877–889. Retrieved from https://doi.org/10.1037/0022-3514.69.5.877

Lockheed, M. E. (1977). Cognitive style effects on sex status in student work groups. *Journal of Educational Psychology*, *69*, 158–165. https://doi.org/10.1037/0022-0663.69.2.158

Lockheed, M. E., & Hall, K. P. (1976). Conceptualizing sex as a status characteristic: Applications to leadership training strategies. *Journal of Social Issues*, *32*, 111–124. https://doi.org/10.1111/j.1540-4560.1976.tb02600.x

Lord, R. G., De Vader, C. L., & Alliger, G. M. (1986). A meta-analysis of the relation between personality traits and leadership perceptions: An application of validity generalization procedures. *Journal of Applied Psychology*, *71*, 402–410. https://doi.org/10.1037/0021-9010.71.3.402

Lueptow, L. B., Garovich-Szabo, L., & Lueptow, M. B. (2001). Social change and the persistence of sex typing: 1974–1997. *Social Forces*, *80*, 1–36. Retrieved from https://doi.org/10.1353/sof.2001.0077

Lyness, K. S., & Judiesch, M. K. (2014). Gender egalitarianism and work–life balance for managers: Multisource perspectives in 36 countries. *Applied Psychology*, *63*, 96–129. Retrieved from https://doi.org/10.1111/apps.12011

Lytton, H., & Romney, D. M. (1991). Parents' differential socialization of boys and girls: A meta-analysis. *Psychological Bulletin*, *109*, 267–296. https://doi.org/10.1037/0033-2909.109.2.267

Mackinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods*, *7*, 83–104. https://doi.org/10.1037/1082-989X.7.1.83

McGrath, J. E. (1962). *Leadership behavior: Some requirements for leadership training*. Washington, DC: U.S. Civil Service Commission, Office of Career Development.

McGrath, J. E. (1984). *Groups: Interaction and performance*. Englewood Cliffs, NJ: Prentice-Hall.

McMillan, J. R., Clifton, A. K., McGrath, D., & Gale, W. S. (1977). Women's language: Uncertainty or interpersonal sensitivity and emotionality? *Sex Roles*, *3*, 545–559. Retrieved from https://doi.org/10.1007/BF00287838

Mullen, B., Salas, E., & Driskell, J. E. (1989). Salience, motivation, and artifact as contributions to the relation between participation rate and leadership. *Journal of Experimental Social Psychology*, *25*, 545–559. https://doi.org/10.1016/0022-1031(89)90005-X

Murray, H. A. (1938). *Explorations in personality*. New York, NY: Oxford University Press.

Nemeth, C., Endicott, J., & Wachtler, J. (1976). From the'50s to the'70s: Women in jury deliberations. *Sociometry*, *39*, 293–304. Retrieved from https://doi.org/10.2307/3033495

Ozer, D. J., & Benet-Martinez, V. (2006). Personality and the prediction of consequential outcomes. *Annual Review of Psychology*, *57*, 401–421. Retrieved from https://doi.org/10.1146/annurev.psych.57.102904.190127

Paunova, M. (2015). The emergence of individual and collective leadership in task groups: A matter of achievement and ascription. *The Leadership Quarterly*, *26*, 935–957. https://doi.org/10.1016/j.leaqua.2015.10.002

Paustian-Underdahl, S. C., Walker, L. S., & Woehr, D. J. (2014). Gender and perceptions of leadership effectiveness: A meta-analysis of contextual moderators. *Journal of Applied Psychology*, *99*, 1129–1145. https://doi.org/10.1037/a0036751

Porath, C. L., Gerbasi, A., & Schorch, S. L. (2015). The effects of civility on advice, leadership, and performance. *Journal of Applied Psychology*, *5*, 1527–1541. https://doi.org/10.1037/apl0000016

Post, C., & Byron, K. (2015). Women on boards and firm financial performance: A meta-analysis. *Academy of Management Journal*, *58*, 1546–1571. Retrieved from https://doi.org/10.5465/amj.2013.0319

Powell, G. N., & Butterfield, D. A. (1979). The "good manager": Masculine or androgynous? *Academy of Management Journal*, *22*, 395–403. Retrieved from https://doi.org/10.2307/255597

Powell, G. N., Butterfield, D. A., & Parent, J. D. (2002). Gender and managerial stereotypes: Have the times changed? *Journal of Management*, *28*, 177–193. Retrieved from https://doi.org/10.1177/014920630202800203

Preacher, K. J., & Selig, J. P. (2012). Advantages of Monte Carlo confidence intervals for indirect effects. *Communication Methods and Measures*, *6*, 77–98. https://doi.org/10.1080/19312458.2012.679848

Raju, N. S., & Brand, P. A. (2003). Determining the significance of correlations corrected for unreliability and range restriction. *Applied Psychological Measurement*, *27*, 52–71. Retrieved from https://doi.org/10.1177/0146621602239476

Ridgeway, C. L., & Correll, S. J. (2004). Unpacking the gender system a theoretical perspective on gender beliefs and social relations. *Gender & Society*, *18*, 510–531. Retrieved from https://doi.org/10.1177/0891243204265269

Riggio, R. E., Riggio, H. R., Salinas, C., & Cole, E. J. (2003). The role of social and emotional communication skills in leader emergence and effectiveness. *Group Dynamics: Theory, Research, and Practice*, *7*, 83–103. https://doi.org/10.1037/1089-2699.7.2.83

Roth, P. L., Bevier, C. A., Bobko, P., Switzer, F. S., & Tyler, P. (2001). Ethnic group differences in cognitive ability in employment and educational settings: A meta analysis. *Personnel Psychology*, *54*, 297–330. Retrieved from https://doi.org/10.1111/j.1744-6570.2001.tb00094.x

Roth, P. L., Purvis, K. L., & Bobko, P. (2012). A meta-analysis of gender group differences for measures of job performance in field studies. *Journal of Management*, *38*, 719–739. Retrieved from https://doi.org/10.1177/0149206310374774

Ruble, D. N., Martin, C. L., & Berenbaum, S. A. (1998). Gender development. In N. Eisenberg, W. Damon, & R. M. Lerner (Eds.), *Handbook of child psychology: Social, emotional, and personality development* (Vol. *3*, pp. 858–932). Hoboken, NJ: John Wiley & Sons, Inc.

Rudman, L. A. (1998). Self-promotion as a risk factor for women: The costs and benefits of counterstereotypical impression management. *Journal of Personality and Social Psychology*, *74*, 629–645. https://doi.org/10.1037/0022-3514.74.3.629

Rudman, L. A., & Glick, P. (1999). Feminized management and backlash toward agentic women: The hidden costs to women of a kinder, gentler image of middle managers. *Journal of Personality and Social Psychology*, *77*, 1004–1010. https://doi.org/10.1037/0022-3514.77.5.1004

Rudman, L. A., & Glick, P. (2001). Prescriptive gender stereotypes and backlash toward agentic women. *Journal of Social Issues*, *57*, 743–762. Retrieved from https://doi.org/10.1111/0022-4537.00239

Rudman, L. A., Moss-Racusin, C., Glick, P., & Phelan, J. E. (2012). Reactions to vanguards: Advances in backlash theory. In M. P. Zanna & J. M. Olson (Eds.), *Advances in experimental social psychology* (pp. 167–227). San Diego, CA: Academic Press.

Sackett, P. R., & Ellingson, J. E. (1997). The effects of forming multi-predictor composites on group differences and adverse impact. *Personnel Psychology*, *50*, 707–721. Retrieved from https://doi.org/10.1111/j.1744-6570.1997.tb00711.x

Sandberg, S. (2013). *Lean In: Women, work, and the will to lead*. New York: Random House Inc.

Schein, V. E. (1973). The relationship between sex role stereotypes and requisite management characteristics. *Journal of Applied Psychology*, *57*(2)95–100. https://doi.org/10.1037/h0037128

Schein, V. E. (1975). Relationships between sex role stereotypes and requisite management characteristics among female managers. *Journal of Applied Psychology*, *60*(3), 340–344. https://doi.org/10.1037/h0076637

Schein, V. E. (2001). A global look at psychological barriers to women's progress in management. *Journal of Social Issues*, *57*, 675–688. Retrieved from https://doi.org/10.1111/0022-4537.00235

Schmid-Mast, M. (2002). Dominance as expressed and inferred through speaking time: A meta-analysis. *Human Communication Research*, *28*, 420–450. Retrieved from https://doi.org/10.1111/j.1468-2958.2002.tb00814.x

Schmidt, F. L., & Hunter, J. E. (2015). *Methods of meta-analysis: Correcting for error and bias in research findings*. Newbury Park, CA: Sage.

Seibert, S., & Gruenfeld, L. (1992). Masculinity, femininity, and behavior in groups. *Small Group Research*, *23*, 95–112. Retrieved from https://doi.org/10.1111/j.1468-2958.2002.tb00814.x

Shaw, M. E. (1959). Acceptance of authority, group structure, and the effectiveness of small groups. *Journal of Personality*, *27*, 196–210. Retrieved from https://doi.org/10.1111/j.1467-6494.1959.tb01829.x

Shaw, M. E. (1973). Scaling group tasks: A method for dimensional analysis. *JSAS Catalog of Selected Documents in Psychology*, *3*, 8.

Shaw, M. E., & Sadler, O. W. (1965). Interaction patterns in heterosexual dyads varying in degree of intimacy. *Journal of Social Psychology*, *66*, 345–351. Retrieved from https://doi.org/10.1080/00224545.1965.9919654

Shinar, E. H. (1975). Sexual stereotypes of occupations. *Journal of Vocational Behavior*, *7*, 99–111. Retrieved from https://doi.org/10.1016/0001-8791(75)90037-8

Singletary, S. L., & Hebl, M. R. (2009). Compensatory strategies for reducing interpersonal discrimination: The effectiveness of acknowledgments, increased positivity, and individuating information. *Journal of Applied Psychology*, *94*, 797–805. Retrieved from https://doi.org/10.1037/a0014185

Sorrentino, R. M., & Boutillier, R. G. (1975). The effect of quantity and quality of verbal interaction on ratings of leadership ability. *Journal of Experimental Social Psychology*, *11*, 403–411. Retrieved from https://doi.org/10.1016/0022-1031(75)90044-X

Spence, J. T., Helmreich, R., & Stapp, J. (1975). Ratings of self and peers on sex role attributes and their relation to self-esteem and conceptions of masculinity and femininity. *Journal of Personality and Social Psychology*, *32*, 29–39. https://doi.org/10.1037/h0076857

Stake, J. E., & Stake, M. N. (1979). Performance-self-esteem and dominance behavior in mixed-sexed dyads. *Journal of Personality*, *47*, 71–84. Retrieved from https://doi.org/10.1111/j.1467-6494.1979.tb00615.x

Stangor, C., Lynch, L., Duan, C., & Glas, B. (1992). Categorization of individuals on the basis of multiple social features. *Journal of Personality and Social Psychology*, *62*, 207–218. Retrieved from https://doi.org/10.1037/0022-3514.62.2.207

Stark, S., Chernyshenko, O. S., & Drasgow, F. (2006). Detecting differential item functioning with confirmatory factor analysis and item response theory: Toward a unified strategy. *Journal of Applied Psychology*, *91*, 1292–1306. Retrieved from https://doi.org/10.1037/0021-9010.91.6.1292

Staw, B. M., & Barsade, S. G. (1993). Affect and managerial performance: A test of the sadder-but-wiser vs. happier-and-smarter hypotheses. *Administrative Science Quarterly*, *38*, 304–331. https://doi.org/10.2307/2393415

Steel, P. D., & Kammeyer-Mueller, J. D. (2002). Comparing meta-analytic moderator estimation techniques under realistic conditions. *Journal of Applied Psychology*, *87*, 96–111. Retrieved from https://doi.org/10.1037//0021-9010.87.1.96

Stogdill, R. M. (1950). Leadership, membership, and organization. *Psychological Bulletin*, *47*, 1–14. https://doi.org/10.1037/h0053857

Strodtbeck, F. L., & Mann, R. D. (1956). Sex role differentiation in jury deliberations. *Sociometry*, *19*, 3–11. Retrieved from https://doi.org/10.2307/2786099

Tett, R. P., & Burnett, D. D. (2003). A personality trait-based interactionist model of job performance. *Journal of Applied Psychology*, *88*, 500–517. Retrieved from https://doi.org/10.1037/0021-9010.88.3.500

Twenge, J. M. (1997). Changes in masculine and feminine traits over time: A meta-analysis. *Sex Roles*, *36*, 305–325. Retrieved from https://doi.org/10.1007/BF02766650

Twenge, J. M., Konrath, S., Foster, J. D., Campbell, W. K., & Bushman, B. J. (2008). Egos inflating over time: A cross-temporal meta-analysis of the Narcissistic Personality Inventory. *Journal of Personality*, *76*, 875–902. https://doi.org/10.1111/j.1467-6494.2008.00507.x

Tzelgov, J., & Henik, A. (1991). Suppression situations in psychological research: Definitions, implications, and applications. *Psychological Bulletin*, *109*, 524–536. Retrieved from https://doi.org/10.1037/0033-2909.109.3.524

U.S. Bureau of Labor Statistics. (2004). *Women in the labor force: A databook*. Retrieved from https://www.bls.gov/cps/wlf-databook.pdf

U.S. Bureau of Labor Statistics. (2015). *Women in the labor force: A databook*. Retrieved from https://www.bls.gov/opub/reports/cps/women-in-the-labor-force-a-databook-2015.pdf

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*, 1–48.

Viswesvaran, C., & Ones, D. S. (1995). Theory testing: Combining psychometric meta-analysis and structural equations modeling. *Personnel Psychology*, *48*, 865–885. Retrieved from https://doi.org/10.1111/j.1744-6570.1995.tb01784.x

Walsh, D. H. (1975). *Sex-role differentiation in problem-solving groups*. Unpublished doctoral dissertation, New York University, New York.

Weber, R., & Crocker, J. (1983). Cognitive processes in the revision of stereotypic beliefs. *Journal of Personality and Social Psychology*, *45*, 961–977. https://doi.org/10.1037/0022-3514.45.5.961

Weiss, M., Kolbe, M., Grote, G., Dambach, M., Marty, A., Spahn, D. R., & Grande, B. (2014). Agency and communion predict speaking up in acute care teams. *Small Group Research*, *45*, 290–313. Retrieved from https://doi.org/10.1177/1046496414531495

Wiggins, J. S. (1991). Agency and communion as conceptual coordinates for the understanding and measurement of interpersonal behavior. In D. Cicchetti & W. M. Grove (Eds.), *Thinking clearly about psychology: Essays in honor of Paul E. Meehl* (pp. 89–113). Minneapolis, MN: University of Minnesota Press.

Willard, D., & Strodtbeck, F. L. (1972). Latency of verbal response and participation in small groups. *Sociometry*, *35*, 161–175. Retrieved from https://doi.org/10.2307/2786556

Williams, J. E., & Best, D. L. (1990). *Measuring sex stereotypes: A multination study*. THousand Oaks, CA: Sage Publications, Inc.

Williams, M. J., & Tiedens, L. Z. (2016). The subtle suspension of backlash: A meta-analysis of penalties for women's implicit and explicit dominance behavior. *Psychological Bulletin*, *142*, 165–197. Retrieved from https://doi.org/10.1037/bul0000039

Wood, W. (1987). Meta-analytic review of sex differences in group performance. *Psychological Bulletin*, *102*, 53–71. Retrieved from https://doi.org/10.1037/0033-2909.102.1.53

Wood, W., Christensen, P. N., Hebl, M. R., & Rothgerber, H. (1997). Conformity to sex-typed norms, affect, and the self-concept. *Journal of Personality and Social Psychology*, *73*, 523–535. https://doi.org/10.1037/0022-3514.73.3.523

Wood, W., & Eagly, A. H. (2012). Biosocial construction of sex differences and similarities in behavior. In J. M. Olson & M. P. Zanna (Eds.), *Advances in experimental social psychology* (Vol. *46*, pp. 55–123). Oxford, UK: Elsevier.

Wood, W., & Karten, S. J. (1986). Sex differences in interaction style as a product of perceived sex differences in competence. *Journal of Personality and Social Psychology*, *50*, 341–347. https://doi.org/10.1037/0022-3514.50.2.341

World Economic Forum. (2015). The global gender gap report. Retrieved from https://www3.weforum.org/docs/GGGR2015/cover.pdf

Zaccaro, S. J., Rittman, A. L., & Marks, M. A. (2001). Team leadership. *The Leadership Quarterly*, *12*, 451–483. Retrieved from 10.1016/S1048-9843(01)00093-5

Zeleny, L. D. (1939). Sociometry of morale. *American Sociological Review*, *4*, 799–808. https://doi.org/10.2307/2083757

Zhang, Z., Waldman, D. A., & Wang, Z. (2012). A multilevel investigation of leader–member exchange, informal leader emergence, and individual and team performance. *Personnel Psychology*, *65*, 49–78. Retrieved from https://doi.org/10.1111/j.1744-6570.2011.01238.x

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Badura KL, Grijalva E, Newman DA, Yan TT, Jeon G. Gender and leadership emergence: A meta-analysis and explanatory model. *Personnel Psychology*. 2018;1–33. https://doi.org/10.1111/peps.12266

APPENDIX A

Keywords Used for the Primary Study Literature Search

Relationship	Keywords Used
Gender and Leadership Emergence	The following keywords: *leadership emergence, leadership, leaderless, leaderless group, leader, emergent leader, informal leader, preferred leader, leadership potential, leadership performance, or leader performance* were paired with one of the following keywords: *gender or sex*
Gender and Agency; Gender and Communion; Agency and Communion	The following keywords were used: *Bem Sex-Role Inventory, Personal Attributes Questionnaire*
Gender and Participation; Agency and Participation; Communion and Participation	The following keywords: *talkativeness, talking time, or speech duration* were paired with one of the following keywords: *Sex, gender, masculinity, masculine, agentic, femininity, feminine, or communal*
Participation and Leadership Emergence	The following keywords: *leadership emergence, leadership, leaderless, leaderless group, leader, emergent leader, informal leader, preferred leader, leadership potential, leadership performance, or leader performance* were paired with one of the following keywords: *talkativeness, talking time, or speech duration*
Agency and Leadership Emergence; Communion and Leadership Emergence	The following keywords: *leadership emergence, leadership, leaderless, leaderless group, leader, emergent leader, informal leader, preferred leader, leadership potential, leadership performance, or leader performance* were paired with one of the following keywords: *Masculinity, masculine, agentic, femininity, feminine, or communal*

APPENDIX B

Leadership Emergence Articles Excluded From Current Meta-Analysis That Were Included in Eagly and Karau (1991)

Authors	Reason For Exclusion
Aries (1976)	Did not report enough information to calculate an effect size
Aries (1982)	Did not report enough information to calculate an effect size
Askinas (1971)	Did not use a direct measure of leadership emergence
Crosbie (1979)	Did not use a direct measure of leadership emergence
Davis and Songer-Nocks (1978)	Could not locate study
Fulton (1981)	Could not locate study
Hall (1973)	Did not use a direct measure of leadership emergence
James (1956)	Did not use a direct measure of leadership emergence
Jose and McCarthy (1988)	Did not use a direct measure of leadership emergence
Lockheed (1977)	Included a non-adult sample
Lockheed and Hall (1976)	Included a non-adult sample
McMillan, Clifton, McGrath, and Gale (1977)	Did not use a direct measure of leadership emergence
Nemeth, Endicott, and Wachtler (1976)	Did not use a direct measure of leadership emergence
Shaw and Sadler (1965)	Did not report enough information to calculate an effect size
Stake and Stake (1979)	Did not report enough information to calculate an effect size
Strodtbeck and Mann (1956)	Did not use a direct measure of leadership emergence
Walsh (1975)	Did not use a direct measure of leadership emergence
Wood and Karten (1986)	Did not use a direct measure of leadership emergence

Note: Direct measures of leadership emergence include election, rank-order, or Likert scale ratings of leader emergence. Studies often were excluded because they used an indirect form of leadership emergence (e.g., who sat at the head seat of the table; Nemeth, Endicott, & Wachtler, 1976) or used participation as a proxy for leadership emergence.