The Leadership Quarterly 32 (2021) 101458

Contents lists available at ScienceDirect

The Leadership Quarterly

journal homepage: www.elsevier.com/locate/leaqua

Country differences in the relationship between leadership and employee engagement: A meta-analysis

Peikai Li [a], Jian-Min Sun [b,*], Toon W. Taris [a], Lu Xing [b], Maria C.W. Peeters [a,c]

[a] Social, Health and Organizational Psychology, Utrecht University, the Netherlands
[b] School of Labor and Human Resources, Renmin University of China, China
[c] Eindhoven University of Technology, Human Performance Management Group, The Netherlands

ARTICLE INFO

Article history:
Received 14 November 2018
Received in revised form 11 August 2020
Accepted 30 August 2020
Available online 2 October 2020

Keywords:
Leadership
Work engagement
National culture
Cultural tightness-looseness
Meta-analysis

ABSTRACT

Leadership is frequently related to important organizational outcomes such as follower engagement. However, to date we have little insight into the degree to which this relation is contingent upon (a) types of leadership style and (b) national culture. These two issues are addressed in a meta-analysis of 209 independent (257 effect sizes), mainly cross-sectional studies (79%), involving 82,386 participants from 45 countries. The findings show that whereas abusive supervision was negatively associated with work engagement, several leadership styles (e.g., servant, empowering, ethical, and charismatic leadership) have positive correlations with subordinate engagement; some dimensions of national culture (e.g., gender egalitarianism, human orientation, performance orientation, future orientation, and power distance) moderate the leadership–employee engagement relationship. However, the correlations between servant, ethical, and transactional leadership and subordinate engagement are less likely to vary across national cultural characteristics. Notwithstanding the proliferation of leadership–employee engagement literature with more than 200 published articles, a strong reliance on cross-sectional designs have impeded it to gain any solid conclusions about causality due to endogeneity biases. We conclude by providing a detailed future research agenda and discussing how our results can stimulate future leadership research and inform practices with regards to leader development.

Introduction

Work engagement is a positive, fulfilling, work-related state of mind that is characterized by vigor, dedication, and absorption (Schaufeli, Salanova, González-Romá, & Bakker, 2002). Engaged employees experience more positive emotions (Bakker, Demerouti, & Sanz-Vergel, 2014), higher levels of in-role and extra-role performance (Christian, Garza, & Slaughter, 2011), and better psychological and physical health (Bakker, Schaufeli, Leiter, & Taris, 2008). However, the Gallup Employee Engagement Report indicated that in 2016, only 34% of the employees in the United States were engaged (Harter, 2018), and that disengaged employees cost the USA somewhere between $450 and $550 billion each year (O'Boyle & Harter, 2013). Similarly, the European Working Conditions 2015 Survey revealed that in Sweden, Greece, and Germany, respectively, only 6.1%, 4.8%, and 4.3% of the employees were highly engaged, and that there is a great deal of variability in employee engagement across thirty-five European countries (Schaufeli, 2018). These high disengagement rates as well as the importance of employee engagement for employee health and productivity have generated numerous studies that investigated its antecedents (Rich, Lepine, & Crawford, 2010).

In this respect, although longitudinal and (quasi-)experimental work is scarce, an increasing number of studies suggests that leaders play an important role in employee work engagement (Bakker & Albrecht, 2018). The Gallup organizational research indicates that at least 70% of the variance in team engagement can be explained by the quality of the leader (Harter, 2018). Unfortunately, so far, most empirical studies assess a single leadership style and employee engagement and have mainly taken single-country samples into account. Comparisons of the relationship between leadership styles and engagement across countries have been largely ignored. Both the fit logic of national culture research (i.e., leader's practices are "consistent" with employees' expectations; Newman & Nollen, 1996; Rabl, Jayasinghe, Gerhart, & Kühlmann, 2014) and the implicit leadership theory, which states that individuals have their implicit beliefs, convictions, and assumptions concerning attributes and behaviors of effective leaders (Offermann, Kennedy, & Wirtz, 1994), suggest that leadership is culturally contingent (House et al., 2004). For example, the GLOBE research program, as well as a growing number of empirical studies (e.g., Lian, Ferris, & Brown, 2012) have shown that the influence of leaders varies

* Corresponding author at: Center for Leadership Studies, Renmin University of China, Beijing 100872, China
E-mail address: chinajms@126.com (J.-M. Sun).

https://doi.org/10.1016/j.leaqua.2020.101458
1048-9843/

considerably as a consequence of the cultural forces in the countries in which the leaders function. Yet, the extent to which the relationship of leadership and employee engagement is culturally contingent remains unclear.

To fill this gap, this study investigates whether the strength of the associations between leadership and employee engagement varies across national cultures. We aim to contribute to the literature in three ways: First, by systematically *meta-analyzing* the moderating role of national culture, our study sheds light on cross-cultural differences in the relationship between perceived leadership and employee engagement. As suggested by Schyns and Schilling (2013), only the perception of leader behavior by employees can affect employees' outcomes. Therefore, in this study leadership refers to follower's perception of leader's behavior. Second, in addition to transformational leadership, we include different types of leadership constructs in our meta-analysis, such as authentic, servant, engaging, and ethical leadership that were only to a limited extent covered in previous engagement studies (e.g., Bedi, Alpaslan, & Green, 2016; Kim, Beehr, & Prewett, 2018). Finally, we examine a set of potential methodological moderators of the leadership–employee engagement relationship. Following these discussions, we draw attention to major limitations that currently apply to this field (especially a strong reliance on cross-sectional designs which cannot establish causal effects due to *endogeneity biases*), and further provide pragmatic suggestions for future research.

Leadership and employee engagement: A moderating effect of national culture?

Although Bass (1997) argued that transactional leadership and transformational leadership are culturally universal, numerous studies have shown that culture moderates the relationship between leadership and employee outcomes (Gelfand, Erez, & Aycan, 2007, for a review). One might assume that cultural differences will also exist in the leadership–employee engagement relationship, meaning that the relationships between leadership and engagement may be contingent on certain cultural values (Kirkman, Chen, Farh, Chen, & Lowe, 2009; Lian et al., 2012).

To address the question whether and to what extent the relation between leadership and employee engagement differs across countries, in line with previous cross-cultural meta-analyses (e.g., Rabl et al., 2014) we rely primarily on national culture perspectives (House et al., 2004). These perspectives emphasize that national cultural differences thwart or even dictate whether management or leadership practices will be effective in different countries (Rabl et al., 2014). In addition, Gelfand, Nishii, and Raver (2006) suggested that the variance in individual attributes will be lower in countries with tight national cultures, because in such countries strong norms clearly prescribe the behaviors that are appropriate in particular situations, and there is a lower tolerance of deviant behavior than in culturally looser countries (Rabl et al., 2014). Therefore, we include tightness-looseness as an important country-level moderator.

National culture

Hofstede (1980) defined national culture as the collective programming of one's mind that distinguishes the members of one group or category of people from another. To investigate national culture as a moderator of the leadership–employee engagement relationship, we decided to first focus on power distance, uncertainty avoidance, collectivism, assertiveness, and performance orientation. These are well-established and well-tested dimensions of national culture (Hofstede, 1980; Rabl et al., 2014). Meta-analytic reviews suggest that collectivism (Rockstuhl, Dulebohn, Ang, & Shore, 2012), power distance (Mackey, Frieder, Brees, & Martinko, 2015), uncertainty avoidance (House et al., 2004; Yan & Hunt, 2005), and masculinity vs. femininity (Zhang, Liu, Xu, Yang, & Bednall, 2019) moderate the relationship between leadership and employee outcomes. Further, we include future orientation, gender egalitarianism, and human orientation as important cultural dimensions. The GLOBE leadership project demonstrated that these dimensions significantly differentiate among societies and organizations, and that there is significant respondent agreement within cultures for all these dimensions (House et al., 2004). It should be mentioned that in the current study we examine the between-study moderating effect of national culture.

The fit logic of national culture research (Rabl et al., 2014) suggests that when leadership fits with a national culture, employees are likely to feel satisfied, engaged, and committed. As a result, they may be able or willing to perform well (Newman & Nollen, 1996). Therefore, we argue that when leadership fits well with a national culture, the relationship between leadership and employee engagement will be stronger. Previous studies have convincingly demonstrated that transformational leadership (e.g., Hoch, Bommer, Dulebohn, & Wu, 2016), servant leadership (Hunter et al., 2013), ethical leadership (Bedi et al., 2016), authentic leadership (Hoch et al., 2016; Neider & Schriesheim, 2011), transactional leadership (Judge & Piccolo, 2004), engaging leadership (Nikolova, Schaufeli, & Notelaers, 2019), charismatic leadership (Chen & Huang, 2016), and empowering leadership (Kim et al., 2018) are positively related to employee engagement. These eight leadership styles represent dimensions of *positive leadership* (Hoch et al., 2016). Building on this argument, we propose that the best-fitting national culture for an optimal relationship between positive leadership style and subordinates' engagement is one high on gender egalitarianism, high on human orientation, low on power distance, low on in-group collectivism, high on performance orientation, high on future orientation, high on assertiveness, and low on uncertainty avoidance.

Gender egalitarianism. Gender egalitarianism is defined as the extent to which "an organization or a society minimizes gender role differences while promoting gender equality" (House et al., 2004). In a highly gender-egalitarian society, individuals care about others, about the quality of life and about interpersonal relationships (Hofstede, 2001). As a result, employees are more likely to react positively to positive leadership behaviors (as a kind of social exchange, Blau, 1964; Brown & Treviño, 2006). Specifically, when followers perceive a leader as caring and being concerned for their well-being or when they receive support and trust, they feel obliged to reciprocate that behavior (Bedi et al., 2016; Gouldner, 1960). Therefore, the positive leadership–employee engagement relationship should be stronger in a high gender-egalitarian country (vs. low) (*Hypothesis 1a*).

Human orientation. Human orientation is the extent to which an organization or society encourages and rewards individuals for being fair, altruistic, friendly, generous, caring, and kind to others (Den Hartog et al., 1999). In high human orientation societies, members are responsible for promoting the well-being of others and people are more strongly urged to provide social support to each other than in other societies (House et al., 2004). Therefore, cultures that emphasize social support and promote others' well-being fit well with the nature of positive leadership behaviors (e.g., caring and support; Bedi et al., 2016). Thus, the positive leadership–employee engagement relationship will be stronger in a high human orientation country (vs. low) (*Hypothesis 1b*).

Power distance. Power distance is "the extent to which a community accepts and endorses authority, power differences, and status privileges. In high power distance countries, power, authority, and information are unequally distributed" (House et al., 2004, p. 536), and reward allocation is based on other criteria than performance (Aycan, 2005). Since it is not one's capacities and performance that are rewarded, employee's motivation to engage in their work may be reduced. Thus, this cultural setting may reduce the effectiveness of leadership that aims to increase performance by enhancing employee engagement. Conversely, in low power distance cultures ability and performance play an important

role in promoting and motivating employees (Rabl et al., 2014). Thus, motivation related to leadership components such as trust and support are more congruent with low power distance countries. Therefore, the positive leadership–employee engagement relationship will be stronger in a low power distance country (vs. high) (*Hypothesis 1c*).

In-group collectivism. In-group collectivism measures the degree to which individuals are integrated into groups and "express pride, loyalty, and interdependence in their families" (House et al., 2004). In high in-group collectivistic cultures, individual achievement is not valued; and even though individuals play an important role, rewards are often given to the collective (Yan & Hunt, 2005). Selection and training/development decisions are often based on personal connections, in-group status, and social obligations (House et al., 2004). Contrarily, in low in-group collectivistic cultures, greater weight is given to decisions based on individual differences in ability, skills, and performance. In such cultures, employee engagement-oriented leadership behaviors fit well with this culture. Consequently, the positive leadership–employee engagement relationship will be stronger in a low in-group collectivism country (vs. high) (*Hypothesis 1d*).

Performance orientation. House et al. (2004) defined performance orientation as the extent to which a society "encourages and rewards innovation, high standards, and performance improvement". High performance-oriented societies emphasize results more than people, reward performance, and have a "can-do" attitude, whereas low performance-oriented societies emphasize loyalty and belongingness (House et al., 2004). Given the focus on results and performance in high performance-oriented cultures, leadership that highlights performance through employee engagement will fit well with such a culture. Thus, the positive leadership–employee engagement relationship will be stronger in a high performance-oriented country (vs. low) (*Hypothesis 1e*).

Future orientation. Future orientation is the extent to which "members of a society or an organization believe that their current actions will influence their future, focus on investment in their future, believe that they will have a future that matters…" (House et al., 2004, p.285). Societies high on future orientation tend to have employees who are intrinsically motivated, achieve economic success, and value long-term success more than low future orientation societies (House et al., 2004). Thus, a high future-oriented culture could be a setting that enhances the effectiveness of leadership styles (e.g., trust, fairness and support) that increase employee motivation. Therefore, the positive leadership–employee engagement relation will be stronger in a high future-oriented country (vs. low) (*Hypothesis 1f*).

Assertiveness. The concept of assertiveness reflects "beliefs as to whether people are or should be encouraged to be assertive, aggressive, and tough or nonassertive, nonaggressive, and tender in social relationships" (House et al., 2004, p.395). High assertiveness cultures value success and progress, hold just-world beliefs, and expect demanding and challenging targets, whereas low assertiveness countries tend to value people and warm relationships, and hold unjust-world beliefs (House et al., 2004, p.405). The emphasis on "just world" in high assertiveness countries has a better fit with positive leadership, such as social exchange and its associated norms of reciprocity and trust (Bedi et al., 2016). The positive leadership–employee engagement relation will be stronger in a high assertiveness country (vs. low) (*Hypothesis 1 g*).

Uncertainty avoidance. Uncertainty avoidance (UA) refers to the degree to which "ambiguous situations are threating to individuals, to which rules and order are preferred, and to which uncertainty is tolerated in a society" (House et al., 2004, p.602). High uncertainty avoidance societies tend to take more moderate calculated risks, show stronger resistance to change, and show less tolerance for breaking rules (House et al., 2004, p.618). To reduce uncertainty, leaders in high UA societies often adopt structural formalization and centralization policies, which reduces the degree to which important information and decision-making are shared with subordinates. Conversely, in low uncertainty avoidance cultures information is shared with employees. Leadership components such as trust and support fits better with low uncertainty avoidance countries. Therefore, the positive leadership–employee engagement relationship will be stronger in a low uncertainty avoidance country (vs. high) (*Hypothesis 1 h*).

Abusive supervision. Leaders' abusive behavior damages the quality of social exchange between the victim and their leader (Peng, Schaubroeck, & Li, 2014), and subordinates may decrease their work effort when they feel treated unfairly by their supervisors (Liu & Wang, 2013). As a result, abusive supervision is negatively related to employee well-being (Tepper, Simon, & Park, 2017). Building on the cultural fit logic, that is, when leadership does not fit with its national culture, the relationship between abusive supervision (*negative leadership*; Mackey et al., 2015) and employee engagement will be strongly negative. For instance, employees in a low power distance country are more sensitive to negative leadership behavior, because they believe that such negative leadership behavior violates norms (i.e., does not fit with the national culture) for interpersonal interactions (Hofstede, 1980; Zhang & Liao, 2015). Based on the above reasoning about how the national culture dimensions influence the positive leadership – employee engagement relationship, we expect that.

Hypothesis 2. The relationship between abusive supervision and employee engagement will be more strongly negative in countries characterized by high gender equalitarianism (H2a), high human orientation (H2b), low power distance (H2c), low collectivism (H2d), high performance orientation (H2e), high future orientation (H2f), high assertiveness (H2g), and low uncertainty avoidance (H2h) than in other countries.

National culture's tightness-looseness. Cultural tightness-looseness is defined as the strength of social norms and the degree of sanctioning within societies (Gelfand et al., 2006), and refers to how external norms and constraints relate to cross-cultural differences in behavior. Gelfand et al. (2006) predicted that the variance in individual attributes is less likely to occur in societies with tight national cultures because norms are stronger in such societies and they have a lower tolerance of deviant behavior than loose societies. Taras, Kirkman, and Steel (2010) have found that the predictive power of national culture on employee individual workplace attitudes and behavior is stronger for tighter, rather than looser cultures. Thus, the fit between leadership and national culture will be more important in tight cultures: leaders' practices that are congruent with employee expectations will have stronger positive correlations with employee engagement, whereas leadership that is inconsistent with employee expectations will have stronger negative correlations with employee engagement for tight national cultures.

Hypothesis 3. Cultural tightness-looseness will moderate the positive leadership–employee engagement relationship, such that the relationship between positive leadership styles and employee engagement will be more positive in countries characterized by a tight national culture as well as high gender egalitarianism (H3a), high human orientation (H3b), low power distance (H3c), low in-group collectivism (H3d), high performance orientation (H3e), high future orientation (H3f), high assertiveness (H3g), and low uncertainty avoidance (H3h), compared to looser national cultures.

Hypothesis 4. Cultural tightness-looseness will moderate the abusive supervision –employee engagement relationship, such that the

relationship of abusive supervision–employee engagement will be strongly negative in countries characterized by both a tight national culture and high gender equalitarianism (H4a), high human orientation (H4b), low power distance (H4c), low in-group collectivism (H4d), high performance orientation (H4e), high future orientation (H4f), high assertiveness (H4g), and low uncertainty avoidance (H4h), compared to loose national cultures.

Potential moderators

Publication status

Publication bias refers to researchers' and editors' inclination to publish only significant results (Rosenthal, 1979). There are mixed findings regarding the effect of this "file drawer problem" on the relationship between antecedents and employee engagement. Whereas Bedi et al. (2016) found stronger mean-corrected correlations for published rather than for unpublished studies, Christian et al. (2011) found a stronger correlation between engagement and job characteristics (e.g., social support and autonomy) for unpublished studies than for published work. Overall, the effect of publication status on leadership and employee engagement is unclear.

Study design

It has been suggested that longitudinal studies should report lower correlations than cross-sectional studies (Christian et al., 2011). In addition, as diary studies often account for more sources of variation, Christian et al. (2011) suggested a stronger correlation for within-person studies than for between-person designs. Thus, we will investigate whether the relations between leadership and employee engagement differ across different study designs.

Measurement type

The Utrecht Work Engagement Scale (UWES) (Schaufeli, Bakker, & Salanova, 2006) is the most frequently used measure of employee engagement (Christian et al., 2011). However, other measures are available and it is possible that the magnitude of a correlation will be influenced by the measure that is used. For example, Christian et al. (2011) found that the contextual performance–engagement relation was significantly stronger for other measures than the UWES. Thus, we will examine whether there are differences among the UWES and other measures of engagement.

Rating source of leadership

In the leadership literature, the self-other rating agreement (SOA) issue has been discussed by researchers (Zhang & Liao, 2015). It is highly possible that a self-serving bias will affect the relationship between leadership and employee engagement. As a result, leader-assessed leadership may differ from employee-perceived leadership behavior, which may influence the leadership employee engagement relationship.

In summary, the present study provides a meta-analytic estimate of the degree to which the association between leadership and employee engagement differs across countries and to what degree national culture, cultural tightness-looseness, and methodological moderators may help to explain any such differences.

Method

Literature search

We conducted an extensive search to identify as many published and unpublished studies as possible. Databases utilized in the search were PsycINFO, Web of Science, EBSCO, ProQuest Dissertation, Scopus, and Google Scholar. The search included terms related to (1) leadership and (2) engagement. For leadership, we used the terms *leadership, leader, and supervisor*; selected leadership terms were *transformational leadership, authentic leadership, ethical leadership, servant leadership, abusive supervision, paternalistic leadership, charismatic leadership,* and *transactional leadership.* For engagement, we used the term *work engagement,* and combinations of *job, work, employee, physical, emotional, cognitive, vigor, dedication*, and *absorption* with the keyword *engagement* (Christian et al., 2011). We also conducted a manual search in major journals (i.e., *Journal of Applied Psychology, Academy of Management Journal, Personnel Psychology, Journal of Organizational Behavior,* and *The Leadership Quarterly*) and checked the reference lists of articles on work engagement and leadership, including theoretical reviews (e.g., Bakker & Albrecht, 2018; Carasco-Saul, Kim, & Kim, 2015) and meta-analytic reviews (e.g., Banks, McCauley, Gardner, & Guler, 2016; Hoch et al., 2016; Mackey et al., 2015). We obtained 4913 records from searching the databases and 59 hits from manual search records. The primary search was conducted in March 2018. In March 2020 we updated our search (i.e., a search of published studies from 2018 to March of 2020), resulting in an additional 2225 articles. After removing 1960 duplicated articles that due to searching in different databases, a total of 5237 articles (3791 from primary search, and 1446 from updating search) were advanced to title and abstract screening.

Primary inclusion criteria and coding procedures

For inclusion, each primary study had to (a) present a quantitative field study; (b) contain measures of leadership and work engagement at the individual level; (c) report sample sizes along with correlations or statistical results that allowed us to adequately calculate effect size, and (d) be written in English. Studies that did not meet these standards were excluded.

The first screen of the primary searched abstracts was initially double-screened for relevance by two of the authors with an initial agreement of 94.5% (210 conflicts among 3791 studies); disagreements were resolved to 100% agreement after discussion. The updating searched 1446 articles were screened by the first author only. This produced 334 potentially relevant articles. In the second round, two of the authors independently double-coded 100 articles. For these papers the inter-rater observed agreement (e.g., sample sizes, reliabilities, effect size, etc.) was 96.75%, and the initial Cohen's kappa was 0.70, which indicates good agreement (Cohen, 1960). All discrepancies were resolved after discussion, resulting in 100% agreement for inclusion. Then the first author coded the remaining 234 articles, independently coding the effect sizes twice after which differences were checked and corrected. To deal with studies using the same data set, following suggestions of Wood (2008), we recorded the names of all authors of each study and then arranged these in alphabetic order to detect duplicate studies. If studies had authors in common, we further checked the study characteristics and samples. Ultimately we found 6 duplicate studies (i.e., same author(s), same data). These were eliminated from further analysis. Further, there are some studies only reporting regression coefficient. After contacting the primary authors to request the correlation table, we deleted studies of which the authors failed to provide the correlation table ($n = 14$). Because researchers have suggested that using standardized regression weights (i.e., beta coefficients) to impute missing correlations is associated with potentially large biases when estimating aggregated effect sizes (Roth, Le, Oh, Van Iddekinge, & Bobko, 2018). This process produced 209 studies and 257 effect sizes (Online supplementary Appendix A1 reports the included studies along with sample sizes and effect sizes; Appendix A2 shows a flowchart of our literature search).

Finally, we included 9 leadership styles (i.e., transformational, authentic, ethical, servant, charismatic, transactional, engaging, and empowering leadership, and abusive supervision). Table 1 presents the study characteristics for the nine leadership styles.

Meta-analysis procedures

Random effect meta-analytic procedures were applied using the R metafor packages (Viechtbauer, 2017). The sample-weighted mean correlations and their variances were corrected for sampling error. For

Table 1
Study characteristics for the leadership styles.

Leadership Styles	Average year of publication	Total number of studies	Total number of samples	N	Number of published samples	Same-time samples
Overall	2016	209	217	82,386	170(78%)	172(79%)
Transformational	2015	99	100	39,482	77(77%)	89(89%)
Authentic	2016	46	46	15,223	34(74%)	31(67%)
Ethical	2016	22	23	6940	22(96%)	14(61%)
Servant	2016	25	26	8639	17(65%)	17(65%)
Abusive Supervison	2016	11	11	2958	9(82%)	5(45%)
Transactional	2016	23	23	6664	13(57%)	21(91%)
Empowering	2017	10	10	8846	8(80%)	7(70%)
Charismatic	2015	4	4	2182	3(75%)	3(75%)
Engaging	2018	5	6	2087	4(67%)	4(67%)

Note: The total number of effect sizes are 257 (30 studies have included more than one leadership, for details see online supplementary Appendix A1). The included studies have been published in more than 100 different journals, and there are 14 journals published more than three papers.

those studies that only reported the correlation between leadership dimensions and engagement, we used Hunter and Schmidt's (2004) formula to integrate effect sizes.

To test our hypotheses, we used a five-stage process. First, we tested how each leadership style related to employee engagement. Second, we conducted a 7 (leadership styles) × 8 (country level scores of culture) moderation analysis to test the moderation effect of national culture. We conducted meta-regression for leadership styles that were included in at least ten studies (Borenstein, Hedges, Higgins, & Rothstein, 2011), so engaging leadership ($k = 6$) and charismatic leadership ($k = 4$) were excluded from meta-regression analysis. An exception is abusive supervision as it is the only negative leadership in our study; and following suggestions of Higgins and Thompson (2004), to control the risk of spurious findings from our meta-regression models, permutation tests (which is a specific form of resampling methods) were conducted if significant moderation effect was found for small samples (i.e., $k < 10$). The permutation tests have been suggested as a well-established mean to calculate significance levels in meta-regression analysis (Higgins & Thompson, 2004), and we conducted it in R 'metafor' package by using the "permutest" function (Viechtbauer, 2017, p.150–152). For each study, we identified the country in which the survey had been conducted. Unfortunately, studies did not usually include national culture as a variable. Therefore, in line with previous meta-analyses (Rabl et al., 2014; Rockstuhl et al., 2012), we used the GLOBE research scores of each dimension (i.e., gender egalitarianism, human orientation, power distance, in-group collectivism, performance orientation, future orientation, assertiveness, and uncertainty avoidance; House et al., 2004) as a measure of national culture. Note that some countries in our meta-analysis were not included in the GLOBE research. Cultural dimensions of these countries were treated as missing value, which makes the number of included studies in moderation analysis differ from that in calculating the pooled effect sizes (i.e., Table 2). Third, to test the joint moderation effect of national cultural dimensions and cultural tight-looseness, we used Gelfand et al.'s (2011) scores to place 45 countries along this dimension (Rabl et al., 2014; Taras et al., 2010). Fourth, we tested other potential moderators (e.g., publication status, study design, and measurement of engagement) and tested for publication bias (Del Re, 2015). Finally, we performed several sensitivity analyses and diagnostics by inspecting the outliers, funnel plot (trim and fill), and P-curve; and correcting for measurement error. Specifically, we conducted specific-sample removed sensitivity analyses by removing studies with effect sizes that exceeded the 95% CI of the overall effect size (Harrer, Cuijpers, Furukawa, & Ebert, 2019; Hunter & Schmidt, 2004). In addition, we used Cronbach's alpha coefficient of internal consistency to correct correlations for artifact distributions of measurement error for perceptions of leadership and work engagement. In line with previous meta-analysis (e.g., Mackey et al., 2015), for studies do not report Cronbach's alpha, we used a mean internal consistency value from other studies included in our meta-analysis. We reported these results in the aggregated effect sizes in Table 2.

Results

Table 2 presents a summary of the meta-analytic results for the associations between employee engagement and perceptions of different leadership styles. The sample-size weighted average correlations were positive for employee engagement and servant ($k = 26$, $\rho = 0.474$), ethical ($k = 23$, $\rho = 0.457$), transformational ($k = 100$, $\rho = 0.430$), charismatic ($k = 4$, $\rho = 0.455$), authentic ($k = 46$, $\rho = 0.419$), empowering ($k = 10$, $\rho = 0.460$), engaging ($k = 6$, $\rho = 0.345$), and transactional leadership ($k = 23$, $\rho = 0.275$). In addition, follower-perceived abusive supervision was negatively associated with employee

Table 2
Meta-analysis results for leadership and engagement.

Variable	*k*	*N*	*r*	*ρ*	*SEρ*	Q	Lower	Upper	I^2	H^2	$\rho_{Correct\ Measurement\ error}$	$\rho_{sensitivity\ analysis}$	*Trim-and fill*
Transformational	100	39,482	0.418	0.430	0.022	1508.106***	0.395	0.465	94.19%	17.2	0.473	0.432	0.468
Authentic	46	15,223	0.407	0.419	0.029	590.584***	0.371	0.465	91.95%	12.43	0.464	0.414	0.471
Ethical	23	6940	0.440	0.457	0.047	398.612***	0.381	0.528	93.29%	14.91	0.499	0.426	0.558
Servant	26	8639	0.465	0.474	0.035	206.250***	0.420	0.525	89.48%	9.51	0.519	0.462	0.474
Abusive Supervison	9	2662	−0.232	−0.233	0.026	13.539	−0.280	−0.185	40.57%	1.68	−0.253	−0.233	−0.244
Transactional	23	6664	0.270	0.275	0.034	146.315***	0.212	0.336	85.65%	6.97	0.326	0.266	0.343
Empowering	10	8846	0.453	0.460	0.059	144.304***	0.365	0.546	93.50%	15.39	0.495	0.469	0.571
Charismatic	4	2182	0.437	0.455	0.140	139.721***	0.213	0.644	96.36%	27.48	0.508	0.455	0.641
Engaging	6	2087	0.342	0.345	0.051	24.655***	0.254	0.430	80.52%	5.13	0.389	0.345	0.366

Note: N = total number of respondents; k = number of independent samples included; r = weighted mean correlation; ρ = sample-size-weighted mean observed correlation; $SE\rho$ = standard error for population estimate; I^2 is an index of heterogeneity computed as the percentage of variability in effects sizes that are due to true differences among the studies; Q provides information on whether there is statistically significant heterogeneity (i.e., yes or no heterogeneity). $\rho_{sensitivity\ analysis}$ = specific-sample removed sensitivity analyses; $\rho_{correct\ measurement\ error}$ = mean score correlation (corrected for unreliability for both variables and sampling error variance).

engagement ($k = 9, \rho = -0.233$). It should be mentioned that in the current study, all the "effect sizes" between different leadership styles and employee engagement are correlations and cannot be interpreted as causal effects.

Although the leadership–employee engagement effect sizes were statistically significantly different from zero, the size of the underlying correlations varied considerably (from −0.233 to 0.474). In addition, as indicated in Table 2, we found a significant Q statistic and high I^2 (applying the 75% rule described in Hunter & Schmidt, 2004), which indicates there is sufficient heterogeneity for potential moderators to be investigated (D'Innocenzo, Mathieu, & Kukenberger, 2016). Next, we report the test of moderation effects of national culture and the results of other moderators.

Moderation effects: National culture

Hypothesis 1 stated that the correlations of positive leadership style and subordinates' engagement will be stronger for members in countries of high gender egalitarianism, high human orientation, low power distance, low collectivism, high performance orientation, high future orientation, high assertiveness, and low uncertainty avoidance. Out of 56 tested interactions (8 National characteristics × 7 Leadership styles), we found 10 significant moderation effects (i.e., 17.86%) (for a summary, see Table 3). We describe these moderation results below (for detailed results see online supplementary Appendix B).

Gender egalitarianism. We found that gender egalitarianism negatively moderates the relationships between several leadership styles and employee engagement (for transformational leadership, $k = 83$, $B = -0.260$, $p = .021$; and empowering leadership, $k = 8$, $B = -0.487$, $p = .019$, $p_{permutation\ test} = 0.044$). Contrary to our prediction, the effect size (i.e., mean corrected correlation) was higher ($k = 23, \rho = 0.500$; and $k = 1, \rho = 0.678$ for transformational leadership and empowering leadership, respectively) in countries with low gender egalitarianism (i.e., 1 standard deviation below the mean) than in high gender egalitarianism countries ($k = 14, \rho = 0.324$; and $k = 1, \rho = 0.181$ for low transformational leadership and empowering leadership, respectively). The relationships between leadership styles of authentic ($k = 36$, $B = -0.194$, $p = .109$), ethical ($k = 17$, $B = 0.419$, $p = .085$), servant ($k = 20$, $B = -0.048$, $p = .825$), and transactional ($k = 20$, $B = 0.099$, $p = .546$), and engagement were not different between high and low gender egalitarian countries (H1a was not supported).

Human orientation. We found that human orientation negatively moderates the relationship between empowering leadership and employee engagement ($k = 8, B = -0.244, p = .007, p_{permutation\ test} = 0.040$). Contrary to our prediction, the mean corrected correlation was lower ($k = 1, \rho = 0.182$) in countries with high human orientation than in countries with low human orientation ($k = 2, \rho = 0.678$). Other moderation effects of human orientation on leadership–employee engagement effect sizes were insignificant (for transformational, $k = 83$, $B = 0.073$, $p = .238$; authentic, $k = 36$, $B = 0.045$, $p = .590$; ethical, $k = 17$, $B = -0.028$, $p = .812$; servant, $k = 20$, $B = -0.162$, $p = .365$; transactional, $k = 20$, $B = 0.094$, $p = .236$) (Hypothesis 1b was not supported).

Power distance. The results showed that power distance positively moderates the relationship between authentic leadership and employee engagement ($k = 36, B = 0.165, p = .046$). Contrary to our prediction, the mean corrected correlation was higher ($k = 10, \rho = 0.558$) in countries with high power distance (i.e., 1 standard deviation above the mean) than in low power distance countries ($k = 8, \rho = 0.390$). No other significant moderation effects for power distance were found (for transformational, $k = 83$, $B = -0.007$, $p = .910$; servant, $k = 20$, $B = 0.084$, $p = .484$; transactional, $k = 20$, $B = 0.027$, $p = .752$; empowering leadership, $k = 8$, $B = 0.483$, $p = .103$, $p_{permutation\ test} = 0.152$; and ethical leadership, $k = 17$, $B = -0.153$, $p = .219$) (Hypothesis 1c was not supported).

In-group collectivism. The results showed that in-group collectivism did not statistically significantly moderate any leadership–employee engagement relationship (for transformational, $k = 83$, $B = 0.006$, $p = .858$; authentic, $k = 36$, $B = 0.049$, $p = .310$; ethical, $k = 17$, $B = -0.057$, $p = .403$; servant, $k = 20$, $B = 0.054$, $p = .330$; transactional, $k = 20$, $B = -0.019$, $p = .658$; and empowering leadership, $k = 8$, $B = -0.307$, $p = .574$). Hypothesis 1d was not supported.

Performance orientation. The results showed that performance orientation moderates the relationship between leadership styles (for transformational leadership, $k = 83, B = 0.225, p = .009$; authentic leadership, $k = 36, B = -0.413, p = .005$) and employee engagement. Specifically, contrary to our hypothesis, the mean corrected correlation was higher (for transformational leadership, $k = 5, \rho = 0.488$; authentic leadership, $k = 7, \rho = 0.566$) in countries with low performance orientation (i.e., 1 standard deviation below the mean) than in countries with high performance orientation (transformational leadership, $k = 43, \rho = 0.478$; authentic leadership, $k = 3, \rho = 0.223$). Performance orientation did not statistically significantly moderate other leadership–employee engagement relationships (for ethical, $k = 17, B = -0.004, p = .988$; servant, $k = 20, B = -0.0106, p = .456$; transactional, $k = 20, B = 0.010, p = .928$; and empowering leadership, $k = 8, B = -0.293, p = .437$). Hypothesis 1e was not supported.

Future orientation. We found future orientation positively moderates the relationship between leadership styles of ethical ($k = 17, B = 0.348, p = .019$) and empowering ($k = 8, B = -0.420, p = .021, p_{permutation\ test} = 0.042$) and employee engagement. The mean corrected correlation was higher (for ethical leadership, $k = 3, \rho = 0.723$; empowering leadership, $k = 7, \rho = 0.564$) in countries with high future orientation than in average future orientated countries (for ethical leadership,

Table 3
Summary of analyses of moderation effect of national cultures.

Variables	Gender Egalitarianism	Human orientation	Power Distance	In-group Collectivism	Performance Orientation	Future Orientation	Assertiveness	Uncertainty Avoidance	Ratios of significant interactions/ interactions tested (%)
Transformational	1	0	0	0	1	0	0	0	0.25
Authentic	0	0	1	0	1	0	0	0	0.25
Ethical	0	0	0	0	0	1	0	0	0.13
Servant	0	0	0	0	0	0	0	1	0.13
Abusive Supervison	0	1	0	0	0	0	0	0	0.13
Transactional	0	0	0	0	0	0	0	0	0.00
Empowering	1	1	0	0	0	1	0	0	0.38

Note: 1 = significant interaction; 0 = insignificant.

$k = 14$, $\rho = 0.420$; empowering leadership, $k = 1$, $\rho = 0.182$); No further statistically significant moderation effects of future orientation on the relationship between leadership and employee engagement were found (transformational, $k = 83$, $B = 0.030$, $p = .638$; servant, $k = 20$, $B = -0.013$, $p = .927$; authentic, $k = 36$, $B = -0.029$, $p = .748$; transactional, $k = 20$, $B = -0.025$, $p = .750$). Hypothesis 1f was partially supported.

Assertiveness. Assertiveness did not statistically significantly moderate any leadership–employee engagement relationship (for leadership styles of transformational, $k = 83$, $B = 0.072$, $p = .307$; authentic, $k = 36$, $B = 0.114$, $p = .261$; ethical, $k = 17$, $B = 0.161$, $p = .292$; servant, $k = 20$, $B = 0.107$, $p = .353$; transactional, $k = 20$, $B = -0.074$, $p = .296$; empowering, $k = 8$, $B = 0.393$, $p = .064$, $p_{permutation\ test} = 0.136$). Thus, Hypothesis 1 g was not supported.

Uncertainty avoidance. We found uncertainty avoidance negatively moderates the servant leadership–employee engagement relationship ($k = 20$, $B = -0.198$, $p = .014$). As expected, the mean corrected correlation was lower ($k = 6$, $\rho = 0.343$) in countries with high UA than in countries with low UA ($k = 4$, $\rho = 0.678$). Other moderation effects of UA were statistically not significant (for transformational, $k = 83$, $B = 0.064$, $p = .290$; authentic, $k = 36$, $B = -0.051$, $p = .602$; ethical, $k = 17$, $B = -0.096$, $p = .322$; transactional leadership, $k = 20$, $B = 0.015$, $p = .876$; and empowering leadership, $k = 8$, $B = -0.183$, $p = .191$). Hypothesis 1 h was partially supported.

Hypothesis 2 stated that national culture would moderate the abusive supervision–employee engagement relationship. Human orientation significantly moderates the abusive supervision–employee engagement relationship ($k = 8$, $B = -0.176$, $p = .012$, $p_{permutation\ test} = 0.006$). As expected, the negative correlation was higher in countries with high human orientation ($k = 7$, $\rho = -0.254$) than in countries with low human orientation ($k = 1$, $\rho = -0.131$). The moderation effects of gender egalitarianism ($k = 8$, $B = 0.248$, $p = .235$), power distance ($k = 8$, $B = 0.165$, $p = .344$), ingroup-collectivism ($k = 8$, $B = -0.061$, $p = .067$, $p_{permutation\ test} = 0.142$), performance orientation ($k = 8$, $B = -0.192$, $p = .493$), future orientation ($k = 8$, $B = 0.275$, $p = .037$, $p_{permutation\ test} = 0.101$), and assertiveness ($k = 8$, $B = 0.123$, $p = .039$, $p_{permutation\ test} = 0.114$), and uncertainty avoidance ($k = 8$, $B = -0.056$, $p = .425$) on the relationship between abusive supervision and engagement were insignificant (only Hypothesis 2b was supported). Note that the subgroup estimates were based on a small number of primary studies.

Tightness-looseness of national culture

Hypotheses 3 proposed that the fit between positive leadership and national culture will be more important in tight cultures. Out of 48 interactions, we did not find any significant interaction effects (Hypotheses 3a-h were not supported, detailed results can be found in online supplementary Appendix C).

Similarly, no significant results were obtained for the joint effects of cultural tightness-looseness and national culture on abusive supervision-employee engagement. Thus, Hypotheses 4 was not supported, detailed results can be found in online supplementary Appendix C.

Other moderators

Publication status. Table 4 moderation analyses revealed no significant differences between published ($k = 77$, $\rho = 0.456$; $k = 34$, $\rho = 0.430$; $k = 17$, $\rho = 0.490$; $k = 13$, $\rho = 0.264$, respectively) and unpublished studies ($k = 23$, $\rho = 0.474$; $k = 12$, $\rho = 0.493$; $k = 9$, $\rho = 0.566$; $k = 10$, $\rho = 0.311$, respectively) for transformational ($t = 0.173$, $p = .678$), authentic ($t = 0.592$, $p = .442$), servant ($t = 1.285$, $p = .257$), and transactional leadership ($t = 0.498$, $p = .480$), and employee engagement.

Study design. Table 5 reports the moderation analysis of study design. Ethical ($t = 8.063$, $p = .005$) and servant leadership ($t = 5.115$, $p = .024$) and employee engagement showed stronger mean corrected correlations for cross-sectional studies ($k = 15$, $\rho = 0.571$; $k = 17$, $\rho = 0.570$) than for longitudinal ($k = 8$, $\rho = 0.354$; $k = 9$, $\rho = 0.417$) studies. Although transformational ($t = 0.487$, $p = .485$), authentic leadership ($t = 0.178$, $p = .673$) and employee engagement showed stronger mean corrected correlations for cross-sectional studies ($k = 90$, $\rho = 0.459$; $k = 42$, $\rho = 0.455$, respectively) than for longitudinal studies ($k = 6$, $\rho = 0.404$; $k = 3$, $\rho = 0.389$), the difference were insignificant. Finally, we compared the effect sizes of within-person level (e.g., a given employee's engagement fluctuates daily in response to his or her boss's behavior that day) correlations and between-person level (e.g., employees of abusive supervision have lower engagement than employees of positive leadership) correlations (McCormick, Reeves, Downes, Li, & Ilies, 2020). Transformational leadership showed stronger mean corrected correlations for between-person level correlations ($k = 98$, $\rho = 0.460$, 95% CI [0.417, 0.503]) than for within-person level correlations ($k = 5$, $\rho = 0.315$, 95% CI [0.126, 0.504]), but the difference was not significant ($t = 2.146$, $p = .143$). Interestingly, for abusive supervision, the mean corrected correlation was stronger for within-person level studies ($k = 3$, $\rho = -0.268$, 95% CI [−0.509, −0.028]) than for between-person level studies ($k = 8$, $\rho = -0.252$, 95% CI [−0.307, −0.197]), but the difference was not significant ($t = 0.017$, $p = .895$).

Measurement of engagement and leadership. Transformational leadership ($t = 4.055$, $p = .044$) showed stronger mean corrected correlations for studies using other measurements of work engagement ($k = 20$, $\rho = 0.535$) than UWES ($k = 78$, $\rho = 0.442$). Authentic ($t = 2.244$, $p = .134$), ethical ($t = 1.900$, $p = .168$), servant leadership ($t = 3.100$, $p = .078$), transactional leadership ($t = 0.702$, $p = .402$) and employee engagement showed no significant mean corrected correlations

Table 4
Meta-analysis results for leadership and engagement: the role of publication status.

Leadership	Subgroup	k	ρ	SEρ	95%CI		Q	I^2	*t*	*p*
					Lower	Upper				
Transformational	published	77	0.456	0.026	0.404	0.507	1381.94	0.95	0.173	0.678
	unpublished	23	0.474	0.035	0.405	0.543	125.919	0.83		
Authentic	published	34	0.430	0.028	0.375	0.485	255.528	0.87	0.592	0.442
	unpublished	12	0.493	0.078	0.341	0.645	308.031	0.96		
Servant	published	17	0.490	0.047	0.399	0.582	142.699	0.89	1.285	0.257
	unpublished	9	0.566	0.048	0.472	0.66	45.5675	0.82		
Transactional	published	13	0.264	0.050	0.165	0.362	116.981	0.90	0.498	0.480
	unpublished	10	0.311	0.043	0.226	0.396	28.4129	0.68		

Note: k = number of independent samples included. ρ = correlation for population estimate corrected for attenuation due to sampling error variance; $SE\rho$ = standard error for population estimate; I^2 is an index of heterogeneity computed as the percentage of variability in effects sizes that are due to true differences among the studies; Q provides information on whether there is statistically significant heterogeneity (i.e., yes or no heterogeneity).

Table 5
Meta-analysis results for leadership and employee engagement: the role of study design.

Leadership	Subgroup	k	ρ	SEρ	95%CI		Q	I^2	*t*	*p*
					Lower	Upper				
Tansformational	Cross-sectional	90	0.459	0.023	0.414	0.503	1308.22	0.93	0.487	0.485
	Longitudinal	6	0.404	0.075	0.257	0.551	131.919	0.96		
Authentic	Cross-sectional	42	0.455	0.030	0.396	0.514	526.618	0.92	0.178	0.673
	Longitudinal	3	0.389	0.154	0.087	0.691	38.6415	0.95		
Ethical	Cross-sectional	15	0.571	0.061	0.452	0.69	246.09	0.94	8.063**	0.005
	Longitudinal	8	0.354	0.046	0.264	0.445	41.1839	0.83		
Servant	Cross-sectional	9	0.417	0.052	0.315	0.518	46.4211	0.83	5.115*	0.024
	Longitudinal	17	0.570	0.041	0.49	0.65	120.01	0.87		

Note: *, $p < .05$; **, $p < .01$; k = number of independent samples included. ρ = correlation for population estimate corrected for attenuation due to sampling error variance; $SE\rho$ = standard error for population estimate; I^2 is an index of heterogeneity computed as the percentage of variability in effects sizes that are due to true differences among the studies; Q provides information on whether there is statistically significant heterogeneity (i.e., yes or no heterogeneity).

Table 6
Meta-analysis results for leadership and employee engagement: the role of measurement of engagement.

Leadership	Subgroup	k	ρ	SEρ	95%CI		Q	I^2	*t*	*p*
					Lower	Upper				
Tansformational	UWES	78	0.442	0.026	0.391	0.492	1221.23	0.94	4.055*	0.044
	Others	20	0.535	0.039	0.459	0.611	199.121	0.90		
Authentic	UWES	39	0.436	0.032	0.373	0.499	514.985	0.93	2.244	0.134
	Others	6	0.543	0.063	0.418	0.667	42.7477	0.88		
Ethical	UWES	19	0.514	0.055	0.406	0.621	357.245	0.95	1.900	0.168
	Others	3	0.370	0.089	0.196	0.544	13.6491	0.85		
Servant	UWES	23	0.499	0.037	0.426	0.572	189.472	0.88	3.100†	0.078
	Others	3	0.648	0.076	0.499	0.797	7.2796	0.73		
Trasactional	UWES	15	0.310	0.030	0.251	0.368	48.9366	0.71	0.702	0.402
	Others	8	0.237	0.081	0.078	0.396	86.1617	0.92		

Note: for subgroup sample size <3, we did not test the subgroup analysis. †, $p < .10$, *, $p < .05$; UWES = Utrecht work engagement scale; k = number of independent samples included. ρ = correlation for population estimate corrected for attenuation due to sampling error variance; $SE\rho$ = standard error for population estimate; I^2 is an index of heterogeneity computed as the percentage of variability in effects sizes that are due to true differences among the studies; Q provides information on whether there is statistically significant heterogeneity (i.e., yes or no heterogeneity.

difference for other measures (ρs = 0.543, 0.370, 0.648, 0.237) than for the UWES (ρs =0.436, 0.514, 0.499, 0.310, respectively) (see Table 6). Finally, regarding the moderation effect of rating sources on the relationship between transformational leadership and employee engagement, the results showed that the effect size was stronger for employee perceived transformational leadership ($k = 95$, $\rho = 0.470$, 95% CI [0.427, 513]) than for leader-reported ($k = 3$, $\rho = 0.154$, 95% CI [−0.117, 425]), $t = 5.10$, $p = .023$.

Sensitivity analyses and diagnostics

The sensitivity analyses revealed that after removing outliers (i.e., the study's confidence interval does not overlap with the confidence interval of the pooled effect, Harrer et al., 2019), most results did not differ much from the overall meta-results (see Table 2). However, the I^2 (i.e., total heterogeneity/total variability) has been decreased, with mostly below the 75% described in Hunter and Schmidt (2004). Specifically, we found positive correlations between employee engagement and leadership styles of transformational (with 28 studies removed, $k = 72$, $\rho = 0.432$, $I^2 = 42.39\%$), authentic (with 16 studies removed, $k = 30$, $\rho = 0.414$, $I^2 = 68.69\%$), ethical (with 8 studies removed, $k = 15$, $\rho = 0.426$, $I^2 = 63.77\%$), servant (with 8 studies removed, $k = 18$, $\rho = 0.462$, $I^2 = 52.92\%$), empowering (with 2 studies removed, $k = 8$, $\rho = 0.469$, $I^2 = 86.11\%$), charismatic (with 1 study removed, $k = 3$, $\rho = 0.455$, $I^2 = 17.89\%$), engaging (with 1studies removed, $k = 5$, $\rho = 0.345$, $I^2 = 46.23\%$), and transactional (with 4 studies removed, $k = 19$, $\rho = 0.266$, $I^2 = 45.13\%$). Lastly, the effect sizes of follower-perceived abusive supervision with employee engagement was same ($k = 9$, $\rho = -0.233$).

Publication bias

A common issue that has been discussed in meta-analysis research is the file-drawer or publication bias problem, which assumes that a study with a low effect size is less likely to be published than a study with high effect sizes (Harrer et al., 2019; Rosenthal, 1979). To examine this kind of bias, we used funnel plot, Eggers test, and trim-and fill approach. For seven leadership styles included in our study (publication bias analysis were not conducted for engaging leadership and charismatic leadership due to small sample size), the funnel plots show that effect sizes in our review are symmetrically distributed around the aggregated effect size (see online supplementary Appendix D), suggesting that publication bias issue is not severe (Card, 2012). Since the funnel plot only provides a subjective assessment of publication bias, we further used Egger, Smith, Schneider, and Minder's (1997) regression test. Specifically, the *p*-value of Egger's test was statistically not significant (for leadership styles of transformational, $k = 100$, $t = -0.47$, $p = .636$; authentic, $k = 46$, $t = -0.709$, $p = .482$; servant, $k = 26$, $t = -0.007$, $p = .994$; abusive, $k = 9$, $t = 0.174$, $p = .867$; transactional, $k = 23$, $t = -0.988$, $p = .334$; empowering, $k = 10$, $t = -2.129$, $p = .07$; for an exception is ethical leadership, $k = 23$, $t = -0.231$, $p = .031$). We should be cautious to interpret Egger's results when the number of studies is small (i.e., $k < 10$). Finally, we used Duval's (2005) trim and fill method. A test of the null hypothesis that the number of missing studies (on the chosen side) is zero was retained only for servant leadership. The adjusted effect size for servant leadership was same as the overall pooled effect size (see Table 2). However, for transformational leadership (with 11 added studies, $k = 111$, adjusted $\rho = 0.468$, 95% CI [0.429, 0.505]), authentic (with 10 added studies, $k = 56$, adjusted $\rho = 0.471$, 95% CI [0.424, 0.516]), ethical leadership (with 8 added studies, $k = 31$,

adjusted $\rho = 0.558$, 95% CI [0.479, 0.627]), abusive supervision (with 1 added studies, $k = 10$, adjusted $\rho = -0.244$, 95% CI [−0.296, −0.191]), transactional leadership (with 7 added studies, $k = 30$, adjusted $\rho = 0.343$, 95% CI [0.275, 0.408]), and empowering leadership (with 4 added studies, $k = 14$, adjusted $\rho = 0.571$, 95% CI [0.455, 0.669]), the trim-and-fill procedure shows that our initial results were underestimated due to publication bias, and the "true" effect when controlling for selective publication might be higher than the original pooled effect sizes.

Recently, researchers argued that publication bias is mostly due to *P*-hacking (i.e., by selectively removing outliers, choosing different outcomes, controlling for different variables, researchers make a non-significant finding becoming significant, Harrer et al., 2019). A model selection method called *P*-Curve has been suggested to examine such a publication bias (McShane, Böckenholt, & Hansen, 2016). Accordingly, using R package "dmetar" (Harrer et al., 2019) we tested such kind of publication bias. The results showed that our meta-analysis has a quite high power estimation. Specifically, for leadership styles of transformational ($k = 96$, 96% included, power estimate is 99% (95% CI: 99%–99%), authentic ($k = 46$, 100% included, power estimate is 99% (95% CI: 99%–99%), ethical ($k = 23$, 100% included, power estimate is 99% (95% CI: 99%–99%), servant ($k = 26$, 100% included, power estimate is 99% (95% CI: 99%–99%), abusive ($k = 9$, 100% included, power estimate is 97% (95% CI: 90%–99%), transactional ($k = 20$, 86.96% included, power estimate is 99% (95% CI: 99%–99%), and empowering ($k = 9$, 90% included, power estimate is 99% (95% CI: 99%–99%). These results suggest that for the leadership styles included in our study evidential values are present, and that they are not absent or inadequate, so the *P*-Curve estimates that there is a "true" effect size behind our findings, and that the results are not the product of publication bias and *P*-hacking alone. The "true effect sizes" for authentic, ethical, servant, empowering, and abusive supervision were same as the initial pooled effect sizes given all studies were included when we estimate the *p*-hacking bias. To estimate the "true effect sizes" for transformational and transactional leadership, we used the Hedges (1984) model selection method, as McShane et al. (2016) found that the original Hedges approach performed better than *P*-Curve and *P*-uniform approaches in more realistic settings. The Hedges (1984) adjusted effect sizes showed that for transactional leadership ($\rho = 0.306$, 95% CI: 0.231, 0.380) the adjusted effect sizes were higher than the initial pooled effect sizes, whereas they were lower for transformational leadership ($\rho = 0.429$, 95% CI: 0.360, 0.499). Because the Hedges (1984) and trim-and-fill approaches are based on different assumptions about publication bias (i.e., small sample bias and p-hacking bias), the results were not exactly consistent, and therefore we reported both results (Harrer et al., 2019). Although using conventional tests we find no severe publication bias, we cannot rule out this possibility. Most of the primary field studies we included are endogeneity plagued, which may increase the likelihood of any result, both published and unpublished, to be statistically significant. Thus, no method, even p-curving, will be able to detect this bias. Again, one should be cautious in interpreting the publication bias results when the number of studies is small (i.e., $k < 15$; Kepes, Banks, & Oh, 2014).

Discussion

This study aimed to examine how the leadership–employee engagement relationship varies across national cultures. The results showed that 10 out of 56 national cultures × leadership styles moderating effects were significant, which to some extent supports the claim that leadership is culturally contingent (House et al., 2004). Specifically, the leadership–employee engagement relationship is stronger in countries high on future orientation (for ethical leadership and empowering leadership), and low on uncertainty avoidance (servant leadership).

Although we found other significant moderating effects of national culture as well, most of these went against our hypotheses. Specifically, contrary to national culture-based logic, the leadership–employee engagement relation was stronger in countries low on gender egalitarianism (for transformational and empowering leadership), low on human orientation (empowering leadership), low on performance orientation (for transformational and authentic leadership), high on power distance (for authentic leadership). Similarly, Rabl et al. (2014) found that the high performance work system (HPWs)-business performance relationship was more strongly positive in high power distance countries. The moderating effect of assertiveness was not supported. To some extent, these results are in line with previous meta-analytic results that on average, HPWS do not work better when they fit the national culture (Rabl et al., 2014). Similarly, Rockstuhl et al. (2012) found that national culture did not affect the relationships between leader-member exchange with task performance and organizational commitment.

Further, the results indicated that the relation between abusive supervision and employee engagement is stronger in countries high on human orientation. Whereas Zhang and Liao (2015) found that power distance moderates the relationships between abusive supervision and subordinates' performance, we did not replicate this effect for engagement. One explanation is that the influence of culture may depend on the study outcomes (cf. Lian et al., 2012, who found that whether power distance exacerbated or mitigated the effect of abusive supervision depended on the outcome).

Third, our results suggest that in general, national cultural factor does *not* constrain the correlations of transactional leadership, ethical leadership, and servant leadership with subordinates' work engagement. The relationship between ethical, servant, and transactional leadership and employee engagement appears to be stable across national cultural factors (compared to transformational, authentic, and empowering leadership). These results are in line with the cultural universal argument. For instance, previous study compared leadership in Western and Asian countries and showed cultural universality for supportive (e.g., servant), and contingent reward (e.g., transactional) leader behaviors (Dorfman et al., 1997). Interestingly, while House et al. (2004) found that the effects of transformational leadership did not depend on culture, in our study we found that the relation between transformational leadership and employee engagement was stronger in countries low on gender egalitarianism and low on performance orientation. Thus, these results supported both the cultural contingency and cultural universal arguments.

Finally, although our study includes more than 209 studies involving 45 countries, the hypothesized interaction effects between national culture and cultural tightness-looseness were not supported. Out of 56 interactions, we did not find any significant interaction effect. These results are not in line with previous meta-analytic studies. For instance, Taras et al. (2010) found that cultural values have significantly stronger effects on outcomes in culturally tighter countries than in looser countries. Similarly, Rabl et al. (2014) found that in tight cultures, the relationship between HPWS and business performance was more positive in national cultures that were relatively low on power distance and/or high on performance orientation (for similar findings see Liu, Jiang, Shalley, Keem, & Zhou, 2016). However, our results did not replicate these interaction effects on engagement. A tentative explanation is that both national culture values and cultural tightness–looseness scores (Gelfand et al., 2011) were taken from other studies and matched with the data in our meta-analytic data set, which is a conservative way to test these interaction effects. Taken together, our findings indicate that the relationship between national culture and leadership–employee engagement is more complex and nuanced than suggested in the literature, as it varies not only with leadership style but also with national culture.

Implications for leadership and employee engagement research

Our study contributes to the leadership literature by systematically investigating how different leadership styles relate to employee

engagement. Although the findings that leadership styles that are typically considered 'positive' (e.g., ethical and servant leadership) are positively associated with employee engagement while abusive supervision was negatively related to employee engagement may not be surprising, this further emphasizes the importance of leadership in the workplace. Importantly, we comprehensively examined how national culture relates to the leadership–employee engagement relationship. The results supported both the cultural contingency and cultural universality arguments, in that the leadership–employee engagement correlations in some cases depend on national cultural characteristics (e.g., human orientation, gender egalitarianism, and UA), whereas the correlations between ethical, servant, transactional leadership and employee engagement appear to be stable across national cultural factors. Finally, the current study supports recent calls from researchers to incorporate a wider array of cultural values beyond individualism–collectivism to understand how cultural values work (Gelfand, Leslie, & Fehr, 2008). Accordingly, we have offered some initial evidence showing that for leadership, not only collectivism and power distance, but also gender egalitarianism (2/7 significant interactions), future orientation (2/8 significant interactions), and performance orientation (2/8 significant interactions) are important cultural dimensions (for a summary see Table 3). Although we used the permutation tests to control the risk of spurious findings from meta-regression as suggested by Higgins and Thompson (2004), note that the significant findings reported here were not Bonferroni corrected for multiple testing. After Bonferroni adjustment of the alpha level (i.e., $p = .00089$), none of our tested interactions were significant, meaning that there is a risk of increased Type I error in our findings. Future investigations are necessary to validate the results that were found in our meta-analysis. In addition, our meta-analysis is limited by including only studies written in English, which is especially problematic when focusing on country differences. Future research can apply a multi-language meta-analysis when comparing leadership effectiveness across countries (e.g., Hiller, Sin, Ponnapalli, & Ozgen, 2019).

Finally, the majority of the literature included in our meta-analysis are based on traditional survey data with cross-sectional research design (79%), which has a strong potential of endogeneity bias in their findings and makes it difficult to conclude any causal relationships between leadership and employee engagement. We strongly recommend and believe that future research should pay more attention to the endogeneity issues in research models and try to address this issue in research designs. Below, we provide a research agenda for future research, with a special emphasis on various methodological issues. We hope this could shed some lights for researchers about how to better advance studies on the relationship between leadership and employee outcomes in the future.

Recommendations for future research

Study designs

Although we conducted an exhaustive search for relevant studies, we are disappointed that we found only one experimental study among working populations (Kovjanic, Schuh, & Jonas, 2013), and the majority of articles included in this meta-analysis employed a cross-sectional design (i.e., 79%); excluding nine experience sampling designs, most "longitudinal designs" are time-lagged measurement designs (the mean time lag is 8.47 weeks, and 19/30 used time intervals within one month) with only one used repeated measurement (i.e., Nikolova et al., 2019), which might lead to endogeneity bias and that precludes drawing strong causal inferences (Antonakis, Bendahan, Jacquart, & Lalive, 2010). Endogeneity bias refers to an instance when a predictor variable (whether categorized as independent variable, mediator, or moderator) is associated with the error term of the outcome variable (Antonakis et al., 2010). Therefore, an important direction for future research on leadership and employee engagement is to advance causal identification.

Several strategies have been identified to improve causal identification. For instance, better measures, data and sample, and statistical approaches (for a review see Shaver, 2019). In the current study, we highlight two ways that have often been proposed to deal with the issue of endogeneity and make valid causal inferences possible: experimental designs and the use of instrumental variables (Antonakis et al., 2010; Fischer, Dietz, & Antonakis, 2017). One obvious reason for the use of experimental designs is that it can provide evidence for causality (Podsakoff & Podsakoff, 2019). The logic of an experiment is that the origin of the change in dependent variable (e.g., work engagement) stems from only the manipulated variable (i.e., leadership) (Antonakis et al., 2010). When designing an experiment, the randomized experiment is always the golden standard, in which several issues need to be carefully considered by researchers: manipulation check, control group, sample size, and ethical issues (for other recommendations for conducting experimental research, see Eden, 2017; Lonati, Quiroga, Zehnder, & Antonakis, 2018). Below, we explain these issues in detail.

Manipulation check

When conducting an experimental study in the laboratory, one must make sure that "the manipulations of leadership phenomena are valid, representative, fair, and powerful enough to produce the intended effects" (Podsakoff & Podsakoff, 2019). This is because manipulation checks can minimize the risk of "demand effects", which refers to "changes in behavior by experimental subjects due to cues about what constitutes appropriate behavior" (Zizzo, 2010). Studies on abusive supervision (e.g., Porath & Erez, 2007), servant leadership (Podsakoff, Podsakoff, MacKenzie, & Klinger, 2013), charismatic (Howell & Frost, 1989), and transformational leadership (Kovjanic et al., 2013) provide insights into how successfully manipulating leadership in an experimental setting could be conducted. For example, researchers have used video (Podsakoff et al., 2013), and pen and paper methods (Van Dierendonck, Stam, Boersma, De Windt, & Alkema, 2014) to manipulate servant leadership (Eva, Robin, Sendjaya, van Dierendonck, & Liden, 2019).

Control group

When conducting an experiment, researchers should always consider to add a control group and randomly assign participants to a control group or an experimental group. Makin and de Xivry (2019) listed "absence of an adequate control condition/group" to assess the effect of intervention (or manipulation) as one of the ten common statistical mistakes. This is because changes in outcome measures can be caused by other elements of the study that do not directly relate to the manipulation. Therefore, for any experimental studies in the future that intends to examine the effect of a leadership manipulation on employee outcomes, it is crucial to compare the effect of this experimental manipulation with the effect of a control condition (e.g., transformational leadership (TFL) versus non-TFL; Kovjanic et al., 2013).

Sample size

It is also important to have enough respondents for detecting a desired effect. With a small sample size, the effect size of the false positives is large and it is also more susceptible for missing an effect that exists in the data (Type II error) (Makin & de Xivry, 2019). Hence, for leadership research with large samples, researchers can reduce the possibility of not detecting an influence when in fact leadership has an influence on outcomes. Accordingly, Makin and de Xivry (2019) proposed two suggestions: (a) present evidence that study has sufficient power to detect an effect (e.g., a priori power analysis); (b) or perform a replication of study.

Ethical issues

Finally, when conducting an experiment, ethical issues should always be considered for treating human participants. Especially, researchers should "minimize ethical concerns about harm to

participants, inequity, paternalism, and deception" (Podsakoff & Podsakoff, 2019). The study should be approved by the local Ethical Committee and in accordance with some general guidelines (e.g., APA's Ethical Principles of Psychologists and Code of Conduct, American Psychological Association, 2010).

A second way to combat the endogeneity issue is to use instrumental variable estimation (Maydeu-Olivares, Shi, & Fairchild, 2020). The underlying idea is that using only the exogenous part of the variation in the independent variable *x* (the part not associated with the error term *e*) to estimate its effect on the dependent variable *y* (Sajons, 2020). Specifically, cross-sectional and longitudinal field studies could be extended with instrumental variable models (IVs, that is, another variable *z* which causes variation in *x*, but which is not influenced by simultaneity or omitted variables, Sajons, 2020) to separate the effects of the endogenous variable from method bias (see Antonakis et al., 2010; Daryanto, 2020, for a tutorial in SPSS; Sajons, 2020, for technical details). Typically, stable individual differences such as demographic information, personality, and cognitive ability could be used as instrumental variables (Antonakis et al., 2010; Hughes, Lee, Tian, Newman, & Legood, 2018). Larcker and Rusticus (2010) suggested that cognitive ability is more suitable as an instrumental variable than personality (Hughes et al., 2018).

In addition to using instrumental variable estimation, stronger survey designs should be employed which allows us for causal identification (Shaver, 2019). A strong reliance on cross-sectional designs and self-report within the leadership – employee engagement research has impeded us to draw meaningful causality conclusions from these findings. Especially, none of these articles has dealt with the endogeneity issue within their design. This issue in leadership studies has been raised by several researchers (e.g., Antonakis et al., 2010; Eva et al., 2019; Shaver, 2019). Therefore, future research should employ stronger survey designs. For instance, using longitudinal designs to investigate possible reverse causation between leadership and its subordinate-related "outcomes", because leader behaviors and employee reactions might mutually affect each other (Kim et al., 2018; Nielsen & Taris, 2019). We consistently found that concurrent study designs result in stronger correlations than time-lagged designs (exceptions are servant leadership and abusive supervision). Similar results were reported by Christian et al. (2011) who found that longitudinal studies usually reported lower correlations than cross-sectional studies. However, it is possible that the influence between variables may be reversed or even reciprocal. We thus recommend leadership researchers to employ longitudinal designs (and use instrumental variable models) in their research in order to test the causal direction (e.g., Nikolova et al., 2019).

Measurement of leadership and employee engagement

Our subgroup analysis showed that the correlation between transformational leadership and employee engagement was weaker when leader-reported scores were used to measure leadership styles than if employee perceptions of leadership were used, but the moderation test showed no significant difference. To some extent this agrees with Kim et al. (2018) who reported that the association between leadership and contextual performance was stronger for self-reported than for other-reported criteria. Future research can use multisource ratings of leadership and investigate whether "a seeing eye in eye effect" (Matta, Scott, Koopman, & Conlon, 2015) influences employee engagement.

Measurement of cultural values

Although we examined the effects of national culture on leadership and employee engagement relations, these cultural values were not directly assessed in the studies included in this meta-analysis. This approach might underestimate the true relationship between culture and leadership–employee engagement relationships (Rockstuhl et al., 2012). Echoing Rockstuhl et al.'s recommendation, future research should consider including subordinates' cultural values and test whether effects at the individual level are similar to what we found at the national level. This especially applies to the tightness-looseness dimension. Several meta-analyses (Rabl et al., 2014; Taras et al., 2010) and empirical studies (e.g., Aktas, Gelfand, & Hanges, 2016) have demonstrated the merits of including this dimension, but in our study we did not find the hypothesized interaction effects. Therefore, more empirical studies are needed in future research. It should be noted some moderation analyses of national culture in our study were based on small sample sizes. Thus, we should be cautious in interpreting these results as Type-I errors are likely to occur when using 15 or fewer studies in a meta-analysis (Field, 2001). Only one article we reviewed has included samples from multiple countries (i.e., Rahmadani, Schaufeli, Ivanova, & Osin, 2019). Future leadership research could try to include multiple national samples within the same study, and to include different cultural dimensions that may help clarify the effectiveness of leadership across cultures (Dickson, Castaño, Magomaeva, & Den Hartog, 2012).

In summary, we offer three overall suggestions for future leadership – employee engagement research regarding the research design and measurement in Table 7.

Practical implications

From a practical perspective, our findings offer a guide for practitioners to better understand how different leadership styles relate to employee engagement. Consistent with previous research (e.g., Hoch et al., 2016), our study confirms the positive association of employing positive leadership styles and employee engagement, and a negative association of abusive supervision and employee engagement across cultures. The strongest relations with engagement were found for servant leadership, ethical, and empowering leadership. Managers who wish to increase employee engagement may consider to broadcast these positive leadership styles and avoid abusive supervision behaviors. Note that our results cannot be interpreted as causal due to the fact that the included studies suffer from endogeneity bias.

Notably, our results indicate that servant, transactional, and ethical leadership are desirable, independent from the cultural context. Especially the correlations with employee engagement are relatively stable

Table 7
Suggestions for future leadership and employee engagement designs.

1. Where it possible, using randomized experiment to establish causality.
a. Adding a control group and randomly assign participants to the manipulation group or control group.
b. Sample size: present evidence that a study has sufficient power to detect an effect (e.g., a priori-power analysis) or perform a replication of study.
c. In addition to use students employees to conduct laboratory study, researchers are encouraged to conduct more field experiments and using work populations as participants (Antonakis, Bastardoz, Jacquart, & Shamir, 2016).
d. Always test whether manipulation of leadership works (preferably with external samples, Lonati et al., 2018) to minimize "demand effects". For an example of video manipulation of servant leadership, see Podsakoff et al. (2013).
e. Ethical issues ("Minimizes ethical concerns about harm to participants, inequity, paternalism, and deception", Podsakoff & Podsakoff, 2019)
2. If employing a survey design:
a. Using instrumental variable models (e.g., cognitive ability) for leadership study to combat endogeneity bias.
b. Employing panel designs (or intensive longitudinal designs) such as experience sampling method.
c. Using validated measurement of leadership.
3. When considering country difference of leadership effectiveness.
a. Directly measure culture values of participants (e.g., cultural tightness-looseness).
b. Including participants from multiple countries.

when accounting for other national cultural variables. Insofar as current findings on the positive associations between leadership styles and employee engagement, this suggests that organizations aiming to enhance employee engagement *may* benefit from developing leadership training programs (Knight, Patterson, & Dawson, 2017) to promote servant and ethical leader behaviors.

In addition, our findings highlight the role of cultural differences in the relationships between some leadership styles (e.g., transformational leadership, authentic leadership, empowering leadership, and abusive supervision) and employee engagement in different cultural context. To improve cultural fit, leaders may consider a country's national culture (House et al., 2004) when making decisions or when interacting with employees from different cultural backgrounds.

Accordingly, leadership development programs (especially for those who work in a multi-cultural context) would benefit from integrating leader, follower, and national characteristics (Aktas et al., 2016; Padilla, Hogan, & Kaiser, 2007). When designing a training program, mangers need to consider in what kind of national culture a new leader will be going to work, and make sure that leadership is contingent with employees' culture values. Especially, they should use the knowledge that national culture and leadership styles jointly affect some leadership–employee engagement relationships (e.g., the empowering – employee engagement relationship is more strongly positive for national cultures with a low performance orientation); a greater focus on ethical standards and followers (i.e., ethical and servant leadership, Hoch et al., 2016) in leadership training and education is not a guarantee, but it could promote employee engagement as these two leadership styles are desirable globally. However, causal assertions are not warranted since the current review is mainly based on cross-sectional studies.

Conclusion

Researchers as well as practitioners have since long argued that leadership affects employee functioning, including their levels of engagement. The present study confirmed the assumed positive relationship between several leadership styles and employee engagement. An investigation of the moderating effects of national culture revealed that culture may moderate the leadership–employee engagement relationship, whereas leadership styles like servant leadership and ethical leadership are seen as desirable everywhere (i.e., the relationship do not vary across cultural factors). And the negative correlation between abusive supervision and work engagement seems less likely to be influenced by national cultural characteristics (1/8 significant interaction). Thus, our study supports both culture consistency and cultural contingency for leadership effectiveness. Since most of the included studies in this meta-analysis were cross-sectional, the methodology of leadership–employee engagement research needs to be improved to strengthen the plausibility of causal claims regarding the effects of leadership; especially experimental designs should be conducted in the future.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.leaqua.2020.101458.

Acknowledgments

The authors would like to express their gratitude to the China Scholarship Council for supporting of the first author's study at Utrecht University. We would also like to thank Yayouk Willems, Ajla Mujcic and Susanne Schulz for their kind help during the literature search, coding, and data analysis processes. The second author was partially supported by National Natural Science Foundation of China (Grant number 71272158).

References*

Aktas, M., Gelfand, M. J., & Hanges, P. J. (2016). Cultural tightness-looseness and perceptions of effective leadership. *Journal of Cross-Cultural Psychology*, *47*, 294–309. https://doi.org/10.1177/0022022115606802.

American Psychological Association (2010). *Publication manual of the APA* (6th ed.). Washington, DC: American Psychological Association.

Antonakis, J., Bastardoz, N., Jacquart, P., & Shamir, B. (2016). Charisma: An ill-defined and ill-measured gift. *Annual Review of Organizational Psychology and Organizational Behavior*, *3*, 293–319. https://doi.org/10.1146/annurev-orgpsych-041015-062305.

Antonakis, J., Bendahan, S., Jacquart, P., & Lalive, R. (2010). On making causal claims: A review and recommendations. *The Leadership Quarterly*, *21*, 1086–1120. https://doi.org/10.1016/j.leaqua.2010.10.010.

Aycan, Z. (2005). The interplay between cultural and institutional/structural contingencies in human resource management practices. *International Journal of Human Resource Management*, *16*, 1083–1119. https://doi.org/10.1080/09585190500143956.

Bakker, A. B., & Albrecht, S. (2018). Work engagement: current trends. *Career Development International*, *23*, 4–11. https://doi.org/10.1108/CDI-11-2017-0207.

Bakker, A. B., Demerouti, E., & Sanz-Vergel, A. I. (2014). Burnout and work engagement: The JD–R approach. *Annual Review of Organizational Psychology and Organizational Behavior*, *1*, 389–411. https://doi.org/10.1146/annurev-orgpsych-031413-091235.

Bakker, A. B., Schaufeli, W. B., Leiter, M. P., & Taris, T. W. (2008). Work engagement: An emerging concept in occupational health psychology. *Work and Stress*, *22*, 187–200. https://doi.org/10.1080/02678370802393649.

Banks, G. C., McCauley, K. D., Gardner, W. L., & Guler, C. E. (2016). A meta-analytic review of authentic and transformational leadership: A test for redundancy. *The Leadership Quarterly*, *27*, 634–652. https://doi.org/10.1016/j.leaqua.2016.02.006.

Bass, B. M. (1997). Does the transactional–transformational leadership paradigm transcend organizational and national boundaries? *American Psychologist*, *52*, 130–139.

Bedi, A., Alpaslan, C. M., & Green, S. (2016). A meta-analytic review of ethical leadership outcomes and moderators. *Journal of Business Ethics*, *139*, 517–536. https://doi.org/10.1007/s10551-015-2625-1.

Blau, P. (1964). *Power and exchange in social life*. New York: J Wiley & Sons.

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2011). *Introduction to meta-analysis*. John Wiley & Sons.

Brown, M. E., & Treviño, L. K. (2006). Ethical leadership: A review and future directions. *The Leadership Quarterly*, *17*, 595–616. https://doi.org/10.1016/j.leaqua.2006.10.004.

Carasco-Saul, M., Kim, W., & Kim, T. (2015). Leadership and employee engagement. *Human Resource Development Review*, *14*, 38–63. https://doi.org/10.1177/1534484314560406.

Card, N. A. (2012). *Applied meta-analysis for social science research*. New York, NY: Guilford Press.

*Chen, Y. S., & Huang, S. Y. (2016). A conservation of resources view of personal engagement in the development of innovative behavior and work-family conflict. *Journal of Organizational Change Management*, *29*(6), 1030–1040. https://doi.org/10.1108/JOCM-11-2015-0213.

Christian, M. S., Garza, A. S., & Slaughter, J. E. (2011). Work engagement: A quantitative review and test of its relations with task and contextual performance. *Personnel Psychology*, *64*, 89–136. https://doi.org/10.1111/j.1744-6570.2010.01203.x.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*, 37–46.

Daryanto, A. (2020). EndoS: An SPSS macro to assess endogeneity. *The Quantitative Methods for Psychology*, *16*(1), 56–70. https://doi.org/10.20982/tqmp.16.1.p056.

Del Re, A. C. (2015). A practical tutorial on conducting meta-analysis in R. *Quantitative Methods for Psychology*, *11*, 37–50. https://doi.org/10.20982/tqmp.11.1.p037.

Den Hartog, D. N., House, R. J., Hanges, P. J., Ruiz-Quintanilla, S. A., Dorfman, P. W., Abdalla, I. A., ... Akande, B. E. (1999). Culture specific and cross-culturally generalizable implicit leadership theories: Are attributes of charismatic/transformational leadership universally endorsed? *The Leadership Quarterly*, *10*, 219–256. https://doi.org/10.1016/S1048-9843(99)00018-1.

Dickson, M. W., Castaño, N., Magomaeva, A., & Den Hartog, D. N. (2012). Conceptualizing leadership across cultures. *Journal of World Business*, *47*(4), 483–492. https://doi.org/10.1016/j.jwb.2012.01.002.

D'Innocenzo, L., Mathieu, J. E., & Kukenberger, M. R. (2016). A meta-analysis of different forms of shared leadership–team performance relations. *Journal of Management*, *42*(7), 1964–1991. https://doi.org/10.1177/0149206314525205.

Dorfman, P. W., Howell, J. P., Hibino, S., Lee, J. K., Tate, U., & Bautista, A. (1997). Leadership in Western and Asian countries: Commonalities and differences in Effec- tive leadership processes across cultures. *The Leadership Quarterly*, *8*(3), 233–274. https://doi.org/10.1016/S1048-9843(97)90003-5.

Duval, S. J. (Ed.). (2005). *The trim and fill method*. Chichester, UK: Wiley.

Eden, D. (2017). Field experiments in organizations. *Annual Review of Organizational Psychology and Organizational Behavior*, *4*, 91–122. https://doi.org/10.1146/annurev-orgpsych-041015-062400.

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, *315*(7109), 629–634. https://doi.org/10.1136/bmj.315.7109.629.

Eva, N., Robin, M., Sendjaya, S., van Dierendonck, D., & Liden, R. C. (2019). Servant leadership: A systematic review and call for future research. *The Leadership Quarterly*, *30*, 111–132. https://doi.org/10.1016/j.leaqua.2018.07.004.

* References marked with an asterisk indicate studies included in the meta-analysis that are discussed in the text. For a complete list of studies see online supplementary Appendix A1.

Field, A. P. (2001). Meta-analysis of correlation coefficients: A Monte Carlo comparison of fixed-and random-effects methods. *Psychological Methods*, *6*, 161–180. https://doi.org/10.1037/1082-989X.6.2.161.

Fischer, T., Dietz, J., & Antonakis, J. (2017). Leadership process models: A review and synthesis. *Journal of Management*, *43*, 1726–1753. https://doi.org/10.1177/0149206316682830.

Gelfand, M. J., Erez, M., & Aycan, Z. (2007). Cross-cultural organizational behavior. *Annual Review of Psychology*, *58*, 479–514. https://doi.org/10.1146/annurev.psych.58.110405.085559.

Gelfand, M. J., Leslie, L. M., & Fehr, R. (2008). To prosper, organizational psychology should… adopt a global perspective. *Journal of Organizational Behavior*, *29*, 493–517. https://doi.org/10.1002/job.530.

Gelfand, M. J., Nishii, L. H., & Raver, J. L. (2006). On the nature and importance of cultural tightness-looseness. *Journal of Applied Psychology*, *91*, 1225–1244. https://doi.org/10.1037/0021-9010.91.6.1225.

Gelfand, M. J., Raver, J. L., Nishii, L., Leslie, L. M., Lun, J., Lim, B. C., & Yamaguchi, S. (2011, May 27). Differences between tight and loose cultures: A 33-nation study. *Science*, *332*, 1100–1104. https://doi.org/10.1126/science.1197754.

Gouldner, A. W. (1960). The norm of reciprocity: A preliminary statement. *American Sociological Review*, *25*, 161–178https://www.jstor.org/stable/2092623.

Harrer, M., Cuijpers, P., Furukawa, T. A., & Ebert, D. D. (2019). *Doing meta-analysis in R: A hands-on guide*. . https://doi.org/10.5281/zenodo.2551803 https://bookdown.org/MathiasHarrer/Doing_Meta_Analysis_in_R/.

Harter, J. (2018). *Employee engagement on the rise in the U.S.* Gallup News August 26. Retrieved from https://news.gallup.com/poll/241649/employee-engagement-rise.aspx.

Hedges, L. V. (1984). Estimation of effect size under nonrandom sampling: The effects of censoring studies yielding statisti- cally insignificant mean differences. *Journal of Educational and Behavioral Statistics*, *9*, 61–85. https://doi.org/10.3102/10769986009001061.

Higgins, J. P., & Thompson, S. G. (2004). Controlling the risk of spurious findings from meta-regression. *Statistics in Medicine*, *23*(11), 1663–1682. https://doi.org/10.1002/sim.1752.

Hiller, N. J., Sin, H. P., Ponnapalli, A. R., & Ozgen, S. (2019). Benevolence and authority as WEIRDly unfamiliar: A multi-language meta-analysis of paternalistic leadership behaviors from 152 studies. *The Leadership Quarterly*, *30*(1), 165–184. https://doi.org/10.1016/j.leaqua.2018.11.003.

Hoch, J. E., Bommer, W. H., Dulebohn, J. H., & Wu, D. (2016). Do ethical, authentic, and servant leadership explain variance above and beyond transformational leadership? A meta-analysis. *Journal of Management*, *44*, Article 014920631666546. https://doi.org/10.1177/0149206316665461.

Hofstede, G. (1980). Culture and organizations. *International Studies of Management and Organization*, *10*, 15–41. https://doi.org/10.1080/00208825.1980.11656300.

Hofstede, G. (2001). *Culture's consequences: Comparing values, behaviors, institutions and organizations across nations* (2nd ed.). London: Sage.

House, R. J., Hanges, P. J., Javidan, M., Dorfman, P. W., Gupta, V., & GLOBE Associates (2004). *Culture, leadership, and organizations: The GLOBE study of 62 societies*. Thousand Oaks, CA: Sage.

Howell, J. M., & Frost, P. J. (1989). A laboratory study of charismatic leadership. *Organizational Behavior and Human Decision Processes*, *43*, 243–269. https://doi.org/10.1016/0749-5978(89)90052-6.

Hughes, D. J., Lee, A., Tian, A. W., Newman, A., & Legood, A. (2018). Leadership, creativity, and innovation: A critical review and practical recommendations. *The Leadership Quarterly*, *29*, 549–569. https://doi.org/10.1016/j.leaqua.2018.03.001.

Hunter, E. M., Neubert, M. J., Perry, S. J., Witt, L. A., Penney, L. M., & Weinberger, E. (2013). Servant leaders inspire servant followers: Antecedents and outcomes for employees and the organization. *The Leadership Quarterly*, *24*, 316–331. https://doi.org/10.1016/j.leaqua.2012.12.001.

Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings* (2nd ed.). Newbury Park, CA: Sage.

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

Kepes, S., Banks, G. C., & Oh, I. S. (2014). Avoiding bias in publication bias research: The value of "null" findings. *Journal of Business and Psychology*, *29*(2), 183–203. https://doi.org/10.1007/s10869-012-9279-0.

Kim, M., Beehr, T. A., & Prewett, M. S. (2018). Employee responses to empowering leadership: A meta-analysis. *Journal of Leadership and Organizational Studies*, *25*, 257–276. https://doi.org/10.1177/1548051817750538.

Kirkman, B. L., Chen, G., Farh, J. L., Chen, Z. X., & Lowe, K. B. (2009). Individual power distance orientation and follower reactions to transformational leaders: A cross-level, cross-cultural examination. *Academy of Management Journal*, *52*, 744–764. https://doi.org/10.5465/amj.2009.43669971.

Knight, C., Patterson, M., & Dawson, J. (2017). Building work engagement: A systematic review and meta-analysis investigating the effectiveness of work engagement interventions. *Journal of Organizational Behavior*, *38*, 792–812. https://doi.org/10.1002/job.2167.

*Kovjanic, S., Schuh, S. C., & Jonas, K. (2013). Transformational leadership and performance: An experimental investigation of the mediating effects of basic needs satisfaction and work engagement. *Journal of Occupational and Organizational Psychology*, *86*(4), 543–555. https://doi.org/10.1111/joop.12022.

Larcker, D. F., & Rusticus, T. O. (2010). On the use of instrumental variables in accounting research. *Journal of Accounting and Economics*, *49*, 186–205. https://doi.org/10.1016/j.jacceco.2009.11.004.

Lian, H., Ferris, D. L., & Brown, D. J. (2012). Does power distance exacerbate or mitigate the effects of abusive supervision? It depends on the outcome. *Journal of Applied Psychology*, *97*, 107–123. https://doi.org/10.1037/a0024610.

Liu, D., Jiang, K., Shalley, C. E., Keem, S., & Zhou, J. (2016). Motivational mechanisms of employee creativity: A meta-analytic examination and theoretical extension of the creativity literature. *Organizational Behavior and Human Decision Processes*, *137*, 236–263. https://doi.org/10.1016/j.obhdp.2016.08.001.

Liu, X. Y., & Wang, J. (2013). Abusive supervision and organizational citizenship behaviour: Is supervisor-subordinate guanxi a mediator? *International Journal of Human Resource Management*, *24*, 1471–1489. https://doi.org/10.1080/09585192.2012.725082.

Lonati, S., Quiroga, B. F., Zehnder, C., & Antonakis, J. (2018). On doing relevant and rigorous experiments: Review and recommendations. *Journal of Operations Management*, *64*, 19–40. https://doi.org/10.1016/j.jom.2018.10.003.

Mackey, J. D., Frieder, R. E., Brees, J. R., & Martinko, M. J. (2015). Abusive supervision: A meta-analysis and empirical review. *Journal of Management*, *43*, 1940–1965. https://doi.org/10.1177/0149206315573997.

Makin, T. R., & de Xivry, J. J. O. (2019). Ten common statistical mistakes to watch out for when writing or reviewing a manuscript. *Elife*, *8*. https://doi.org/10.7554/eLife.48175.

Matta, F. K., Scott, B. A., Koopman, J., & Conlon, D. E. (2015). Does seeing "eye to eye" affect work engagement and organizational citizenship behavior? A role theory perspective on lmx agreement. *Academy of Management Journal*, *58*, 1686–1708. https://doi.org/10.5465/amj.2014.0106.

Maydeu-Olivares, A., Shi, D., & Fairchild, A. J. (2020). Estimating causal effects in linear regression models with observational data: The instrumental variables regression model. *Psychological Methods*, *25*(2), 243–258. https://doi.org/10.1037/met0000226.

McCormick, B. W., Reeves, C. J., Downes, P. E., Li, N., & Ilies, R. (2020). Scientific contributions of within-person research in management: Making the juice worth the squeeze. *Journal of Management*, *46*(2), 321–350. https://doi.org/10.1177/0149206318788435.

McShane, B. B., Böckenholt, U., & Hansen, K. T. (2016). Adjusting for publication Bias in meta-analysis: An evaluation of selection methods and some cautionary notes. *Perspectives on Psychological Science*, *11*(5), 730–749. https://doi.org/10.1177/1745691616662243.

Neider, L. L., & Schriesheim, C. A. (2011). The authentic leadership inventory (ALI): Development and empirical tests. *The Leadership Quarterly*, *22*, 1146–1164. https://doi.org/10.1016/j.leaqua.2011.09.008.

Newman, K. L., & Nollen, S. D. (1996). Culture and congruence: The fit between management practices and national culture. *Journal of International Business Studies*, *27*, 753–779. https://doi.org/10.1057/palgrave.jibs.8490152.

Nielsen, K. M., & Taris, T. W. (2019). Leading well: Challenges to researching leadership in occupational health psychology – And some ways forward. *Work and Stress*, *33*, 107–118. https://doi.org/10.1080/02678373.2019.1592263.

*Nikolova, I., Schaufeli, W., & Notelaers, G. (2019). Engaging leader–engaged employees? A cross-lagged study on employee engagement. *European Management Journal*, *37*(6), 772–783. https://doi.org/10.1016/j.emj.2019.02.004.

O'Boyle, E., & Harter, J. (2013). *State of the American workplace: Employee engagement insights for US business leaders*. Washington, DC: Gallup.

Offermann, L. R., Kennedy, J. K., & Wirtz, P. W. (1994). Implicit leadership theories: Content, structure, and generalizability. *The Leadership Quarterly*, *5*, 43–58. https://doi.org/10.1016/1048-9843(94)90005-1.

Padilla, A., Hogan, R., & Kaiser, R. B. (2007). The toxic triangle: Destructive leaders, susceptible followers, and conducive environments. *The Leadership Quarterly*, *18*, 176–194. https://doi.org/10.1016/j.leaqua.2007.03.001.

Peng, A. C., Schaubroeck, J., & Li, Y. (2014). Social exchange implications of own and coworkers' experiences of supervisory abuse. *Academy of Management Journal*, *57*, 1385–1405. https://doi.org/10.5465/amj.2012.0080.

Podsakoff, N. P., Podsakoff, P. M., MacKenzie, S. B., & Klinger, R. L. (2013). Are we really measuring what we say we're measuring? Using video techniques to supplement traditional construct validation procedures. *Journal of Applied Psychology*, *98*(1), 99–113. https://doi.org/10.1037/a0029570.

Podsakoff, P. M., & Podsakoff, N. P. (2019). Experimental designs in management and leadership research: Strengths, limitations, and recommendations for improving publishability. *The Leadership Quarterly*, *30*, 11–33. https://doi.org/10.1016/j.leaqua.2018.11.002.

Porath, C. L., & Erez, A. (2007). Does rudeness really matter? The effects of rudeness on task performance and helpfulness. *Academy of Management Journal*, *50*, 1181–1197. https://doi.org/10.5465/amj.2007.20159919.

Rabl, T., Jayasinghe, M., Gerhart, B., & Kühlmann, T. M. (2014). A meta-analysis of country differences in the high-performance work system-business performance relationship: The roles of national culture and managerial discretion. *Journal of Applied Psychology*, *99*, 1011–1041. https://doi.org/10.1037/a0037712.

*Rahmadani, V. G., Schaufeli, W. B., Ivanova, T. Y., & Osin, E. N. (2019). Basic psychological need satisfaction mediates the relationship between engaging leadership and work engagement: A cross-national study. *Human Resource Development Quarterly*, *30*(4), 453–471. https://doi.org/10.1002/hrdq.21366.

Rich, B. L., Lepine, J. A., & Crawford, E. R. (2010). Job engagement: Antecedents and effects on job performance. *Academy of Management Journal*, *53*, 617–635. https://doi.org/10.5465/amj.2010.51468988.

Rockstuhl, T., Dulebohn, J. H., Ang, S., & Shore, L. M. (2012). Leader-member exchange (LMX) and culture: A meta-analysis of correlates of LMX across 23 countries. *Journal of Applied Psychology*, *97*, 1097–1130. https://doi.org/10.1037/a0029978.

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, *86*, 638–641. https://doi.org/10.1037/0033-2909.86.3.638.

Roth, P. L., Le, H., Oh, I. S., Van Iddekinge, C. H., & Bobko, P. (2018). Using beta coefficients to impute missing correlations in meta-analysis research: Reasons for caution. *Journal of Applied Psychology*, *103*, 644–658. https://doi.org/10.1037/apl0000293.

Sajons, G. (2020). Estimating the causal effect of measured endogenous variables: A tutorial on experimentally randomized instrumental variables. *The Leadership Quaterly*, *55*(11) (Forthcoming).

Schaufeli, W. B. (2018). Work engagement in Europe: Relations with national economy, governance and culture. *Organizational Dynamics*, *47*, 99–106. https://doi.org/10.1016/j.orgdyn.2018.01.003.

Schaufeli, W. B., Bakker, A. B., & Salanova, M. (2006). The measurement of work engagmement with a short questionnaire: A cross-national study. *Educational and Psychological Measurement*, *66*, 701–716. https://doi.org/10.1177/0013164405282471.

Schaufeli, W. B., Salanova, M., González-Romá, V., & Bakker, A. B. (2002). The measurement of engagement and burnout: A two sample confirmatory factor analytic approach. *Journal of Happiness Studies*, *3*, 71–92. https://doi.org/10.1023/A:1015630930326.

Schyns, B., & Schilling, J. (2013). How bad are the effects of bad leaders? A meta-analysis of destructive leadership and its outcomes. *The Leadership Quarterly*, *24*(1), 138–158. https://doi.org/10.1016/j.leaqua.2012.09.001.

Shaver, J. M. (2019). Causal identification through a cumulative body of research in the study of strategy and organizations. *Journal of Management*. https://doi.org/10.1177/0149206319846272.

Taras, V., Kirkman, B. L., & Steel, P. (2010). Examining the impact of culture's consequences: A three-decade, multilevel, meta-analytic review of Hofstede's cultural value dimensions. *Journal of Applied Psychology*, *95*, 405–439. https://doi.org/10.1037/a0018938.

Tepper, B. J., Simon, L., & Park, H. M. (2017). Abusive supervision. *Annual Review of Organizational Psychology and Organizational Behavior*, *4*, 123–152. https://doi.org/10.1146/annurev-orgpsych-041015-062539.

Van Dierendonck, D., Stam, D., Boersma, P., De Windt, N., & Alkema, J. (2014). Same difference? Exploring the differential mechanisms linking servant leadership and transformational leadership to follower outcomes. *The Leadership Quarterly*, *25*(3), 544–562. https://doi.org/10.1016/j.leaqua.2013.11.014.

Viechtbauer, W. (2017). *Package "metafor". R package version 2.0–0. Vol. 1*. (pp. 1–262), 1–262 Retrieved from http://www.metafor-project.org/doku.php.

Wood, J. (2008). Methodology for dealing with duplicate study effects in a meta-analysis. *Organizational Research Methods*, *11*, 79–95. https://doi.org/10.1177/1094428106296638.

Yan, J., & Hunt, J. G. J. (2005). A cross-cultural perspective on perceived leadership effectiveness. *International Journal of Cross Cultural Management*, *5*, 49–66. https://doi.org/10.1177/1470595805050824.

Zhang, Y., & Liao, Z. (2015). Consequences of abusive supervision: A meta-analytic review. *Asia Pacific Journal of Management*, *32*, 959–987. https://doi.org/10.1007/s10490-015-9425-0.

Zhang, Y., Liu, X., Xu, S., Yang, L. -Q., & Bednall, T. C. (2019). Why abusive supervision impacts employee OCB and CWB: A meta-analytic review of competing mediating mechanisms. *Journal of Management*, *45*(6), 2474–2497. https://doi.org/10.1177/0149206318823935 X.

Zizzo, D. J. (2010). Experimenter demand effects in economic experiments. *Experimental Economics*, *13*(1), 75–98. https://doi.org/10.1007/s10683-009-9230-z.