Ethical Leadership on the Rise? A Cross-Temporal and Cross-Cultural Meta-Analysis of its Means, Variability, and Relationships with Follower Outcomes Across 15 Years

Justine Amory[1] · Bart Wille[1] · Brenton M. Wiernik[2] · Sofie Dupré[1]

Abstract

Scholars have suggested that leaders' ethical failures at the beginning of the twenty-first century have raised awareness about the importance of ethical leadership (EL). Yet, there has been no systematic effort to evaluate whether this awareness indeed led to changes in EL or how followers react to this leadership style over time. To address this gap, we examine the evolution in EL means, variability, and its associations with follower outcomes between 2004 and 2019. Our cross-temporal meta-analysis included 359 independent samples from 314 studies published between 2005 and 2020 and focused on followers' ratings of their leaders using the Ethical Leadership Scale (ELS; Brown et al., in Organizational Behavior and Human Decision Processes, 97(2), 117–134, 2005). Using cubic spline meta-regressions, our results indicated no global changes in EL mean levels and variability across the 15-year period. Nevertheless, country-level comparisons revealed different EL-trends in China (i.e. decrease) versus the United States (i.e. stable), and further moderation analyses highlighted the role of cultural value dimensions and national corruption rates. Finally, we also found that the relationship between EL and desirable follower outcomes (e.g. organizational citizenship behavior) became gradually stronger over time on a global level. These results provide a solid empirical basis to evaluate cross-temporal trends in EL and its (changing) impact on follower outcomes across the globe.

Keywords Ethical leadership · Cross-temporal meta-analysis · Cross-cultural meta-analysis · Corruption · Cultural dimensions · Follower outcomes

Introduction

Around the start of the twenty-first century, a series of business scandals (e.g. WorldCom, Enron, Tyco) were widely discussed in international media, drawing public attention to the theme of corporate ethics. Over the following years, the scope of such scandals broadened, and reflected more contemporary ethical issues, such as greenwashing (e.g. Volkswagen) and sexual misconduct in organizational settings

and beyond (e.g. the large number of #MeToo cases). A recurring theme in these global events has been the critical role of people in authority—leaders—being exposed and held accountable for their ethical misconduct. As a reaction to this increased public awareness, different stakeholders have put ethical leadership (EL) high(er) on the agenda. For instance, governments, public and private organizations have reinforced regulations in this area and started implementing codes of conduct at an increasing rate (Kaptein, 2010). In a similar vein, academic research on EL has boomed (Hoch et al., 2018), showing how this leadership style contributes to the well-being and performance of employees (Banks et al., 2021). But has this increased awareness and the actions undertaken to impose EL actually led to increases in EL over time? And to what extent have followers become more sensitive to EL, as indicated by their responses to such leader behaviors?

The current paper addresses these questions by, for the first time in the literature, taking a cross-temporal approach

---

Brenton Wiernik is currently an independent researcher and Research Scientist at Meta, Demography and Survey Science. The current research was conducted while he was at the University of South Florida.

Extended author information available on the last page of the article

to EL. Specifically, we use contextual leadership perspectives (Oc, 2018) and person-environment fit theory (Kristof, 1996) to investigate three ways in which time can shape EL; namely (a) by influencing its base rates or means, (b) by restricting its variability, and (c) by molding its relationships with outcome variables. First, we use cross-temporal meta-analysis (CTMA; Twenge, 2000) to provide an empirical analysis of changes in *mean* EL in studies published between 2005 and 2020 and using the Ethical Leadership Scale (Brown et al., 2005). Analyzing changes in mean EL provides a straightforward test of how time, as an omnibus context factor (Johns, 2006), influences the base rates of EL behaviors (e.g. "*Has ethical leadership increased over time?*"). Second, we rely on CTMA to analyze changes in the *variability* in EL ratings over these 15 years. If leaders are influenced by a shared temporal context in which the emphasis on ethics increases, then these shared experiences are also expected to reduce the variability between leaders across time (e.g. "*Have leaders become more similar over time?*"). Third, our cross-temporal analysis also examines temporal changes in the relationship between EL and follower outcomes. Whereas the first two analyses are helpful to understanding changes in leader ethics per se, focusing on the associations with outcome variables provides insight into how temporal context changes what followers value from their leaders and how this impacts the relevance of EL (e.g. "*Is EL becoming more important to followers?*"). Finally, we investigate whether temporal trends in EL are subject to cross-cultural differences by examining the role of national (i.e. China versus United States), cultural and structural factors in these temporal trends (e.g. "*Are temporal changes in EL dependent on the broader cultural context?*").

These questions are important for both theoretical and practical reasons. By examining changes over time, we answer calls to take contextual factors into account when studying organizational behavior (Johns, 2006) and—more specifically—leadership (Oc, 2018). Moreover, the inclusion of cultural factors and follower outcomes in this study also answers calls to further examine cross-cultural differences in EL and the importance attributed to it (Ahmad et al., 2020; Den Hartog, 2015; Resick et al., 2006). As a result, the findings in this study can be particularly relevant to policymakers and managers of multinational companies. Nonetheless, our results also inform any practitioner interested in understanding the impact of EL on important work outcomes and its evolution over time. Finally, on a methodological level, the current study presents an innovative way to meta-analyze complex subordinate ratings, which can represent both individual and aggregated ratings. The proposed methodology can therefore serve as a blueprint for future meta-analyses examining changes in leadership (or any other construct typically relying on multiple raters) over time, thereby also allowing for non-linear changes (Rudolph et al., 2020).

In sum, the current study contributes to the expanding literature on EL by examining to what extent and how EL has become more ingrained in organizational contexts over the 15-year period between 2004 and 2019. We chose 2004 as the starting point for this cross-temporal analysis because in this year the publication of the ethical leadership scale (ELS; Brown et al., 2005) marked the beginning of a surge in empirical research on this topic, generating EL data on a global scale. Interestingly, using data collected in the early 2000's also allowed us to cover the run-up to several potentially influential events and evolutions (e.g. the collapse of Lehman Brothers in 2008 and the worldwide financial crisis following from that). Further, because we wanted to avoid including data that had been collected during the COVID-19 period—i.e. an exceptional temporal context which could have introduced highly time-specific effects—2019 was chosen as the end date of our analysis. In doing so, this study can help evaluate empirically whether significant social (e.g. movements raising awareness), structural (e.g. newly implemented legislation) and cultural factors (e.g. power distance) have impacted the occurrence of EL and how it is perceived by followers.

Shifting Mean Levels of EL

EL is commonly defined as ''the demonstration of normatively appropriate conduct through personal actions and interpersonal relationships, and the promotion of such conduct to followers through two-way communication, reinforcement, and decision-making'' (Brown et al., 2005, p. 120). Theoretically, potential shifts in EL across time can be understood as the result of contextual influences on leadership (Oc, 2018). Indeed, in Johns' (2006) model of work context, the time period has been described as an informative omnibus context variable influencing organizational phenomena. In the past, time-based changes in the omnibus context have been invoked to predict changes in various worker characteristics (e.g. Myers & Sadaghiani, 2010; Twenge et al., 2010a, 2010b; Wegman et al., 2018). Similarly, it can be argued that leadership does not occur in a vacuum but is likely to be affected by time-based context (Garretsen et al., 2020). Compared to other moral leadership styles such as authentic, servant, or humble leadership (Lemoine et al., 2019), EL in particular can be expected to be shaped by the broader context in which it operates. Specifically, EL is based on deontology, a normative ethical theory which suggests that proper behavior conforms to established societal standards and norms (Lemoine et al., 2019). These standards and norms are deeply rooted in specific national and cultural environments. In contrast, other moral leadership styles such as authentic and humble leadership are rooted in a different philosophical perspective (i.e. virtue ethics) which primarily relies on personal and internal moral principles.

The time period can affect EL in different ways. First, time-based context captures opportunities and constraints that affect the occurrence (i.e. base rates) of organizational behaviors (Johns, 2006), including (ethical) leadership behavior (Oc, 2018). In other words, time-based context can affect EL by fostering or inhibiting the likelihood that leaders in organizations display these behaviors, thus affecting mean levels of EL. One way to think about this is from a demands-abilities (D-A) person-environment fit perspective (Dierdorff et al., 2009; Edwards, 1996; Kristof, 1996), where fit is a function of the worker's ability to meet contextual demands. Specifically, time-based context can change the nature of the leadership role, which in turn influences (changes in) the knowledge, skills, and abilities (KSAs) needed to perform effectively in this role (Livingstone et al., 1997). In this way, the overarching implication of the D-A fit perspective is that leaders with the requisite KSAs to meet changing contextual demands will be increasingly competitive (Stevens & Campion, 1994) and that organizations should actively target or reinforce leaders with these KSAs. When considering tangible contextual changes potentially affecting the nature of the leadership role and thus the base rates of EL, we distinguish between factors discouraging unethical behavior and factors encouraging ethical behavior.

Factors Discouraging Unethical Behavior

We identified two societal evolutions that could have contributed to discouraging unethical behavior during the fifteen years between 2004 and 2019. The first refers to the increased mediatization of corporate scandals in the past two decades. Specifically, it can be argued that through the different electronic news platforms, which can now react instantly and often have a global and population-wide outreach, organizations and their leaders run a higher risk of reputation damage, especially those operating internationally. As Levi (2006, p. 1050) suggests, "one would expect it to deter misconduct and crimes that are expected to have some chance of being discovered and publicized in the media, to some extent independent of prosecution, conviction, and imprisonment." Organizations' apprehension to appear in the media can be seen, for instance, in the use of headline tests or frontpage tests, encouraging employees to ask themselves "what would an organization's constituents think about the act or decision if it was reported in the media?" (Anand et al., 2004).

A second factor potentially discouraging unethical behavior involves the reinforcement of regulation. Indeed, a growing number of legal measures have been taken in the last two decades to prevent (or at least reduce) unethical conduct in organizations. Examples include the Sarbanes–Oxley Act of 2002, implemented in reaction to specific corporate scandals (e.g. Enron, WorldCom, Tyco) occurring in the early 2000s in the United States. This federal law aims to prevent ethical misconduct in public companies through a series of actions, ranging from the obligation to adopt a code of ethics (or justify its absence) to enhancing white-collar criminal penalties. Around the same time, European legislation also tightened the bolts by reinforcing punishments for corruption and ethical transgressions through its anti-fraud office (OLAF; Pujas, 2003). In sum, these (and potentially other) factors can create a leadership context in which the occurrence or base rate of ethical misconduct is restricted, which can then gradually become visible in higher mean levels of EL across time.

Factors Encouraging Ethical Behavior

Importantly, the time-based context impacting on leadership has arguably also changed under the influence of factors encouraging ethical behavior. Specifically, an increased focus on ethics in organizations not only serves to prevent misconduct, but can equally constitute a positive and deliberate strategy of organizations to attain or foster their competitive advantage (Trevino & Nelson, 2021). Indeed, research has for instance demonstrated that organizations perceived as socially responsible tend to perform better economically (Orlitzky & Benjamin, 2001) and are also more successful in attracting job candidates (Strobel et al., 2010), which represents a significant advantage in a context of tight labor markets. Because of such reasons, ethical behavior is actively encouraged by stakeholders, who increasingly invest in socially responsible businesses (Schroders Global Investor Study, 2017; US Forum for Sustainable and Responsible Investment [US SIF], 2020) and urge organizations to both raise their ethical standards and closely monitor their realization, also in leadership roles (McWilliams et al., 2006).

Taken together, external pressures discouraging unethical and encouraging ethical behavior in organizations have created a leadership context in which ethics occupy a more central position (Kaptein, 2010). Drawing on the aforementioned D-A fit perspective, it can therefore be expected that those people who better meet these evolving standards have a competitive advantage when it comes to the likelihood of being selected for leadership roles (Stevens & Campion, 1994). Similarly, once in such a role, those leaders who are more effective in installing ethics-based policies (and thus meeting the relevant contextual demands) should be more likely to be retained for these positions. Either way, as organizations will become more populated with ethical leaders, we expect this to be reflected in higher mean ratings of this leadership style across time. This translates into the following hypothesis:

> *Hypothesis 1: Mean levels of EL have globally increased between 2004 and 2019.*

Shifting Variability in EL

In addition to influencing base rates, the context has also been theorized to affect the observed *range* of organizational phenomena (Johns, 2006). In this regard, the impact of context can be understood from the perspective of situational strength (Mischel, 1968), whereby so-called strong situations, with obvious norms and rigid roles, tend to constrain the expression of interindividual differences. In contrast, weak situations permit more latitude or opportunity for expressing such diversity.

When applied to EL, as described above, the period between 2004 and 2019 was characterized by a series of impactful events warning against the danger of unethical behaviors while at the same time uncovering significant benefits of ethical behaviors. Because of this, the temporal context in which leadership operates can be hypothesized to have evolved in the direction of a strong situation in which explicit norms about ethics have been installed for and by leaders, both formally (e.g. through new or enhanced legislation) and informally (e.g. as embedded in the corporate culture). In such a context of increasingly established ethics policies and more widely shared norms and beliefs about the value of corporate ethics, it can be expected that leaders, as a group, will also become more similar (or less dissimilar) with respect to their levels of ethicality. We thus propose the following hypothesis:

Hypothesis 2: Variability in EL has globally decreased between 2004 and 2019.

Note that it is crucial to examine variability changes beyond potential mean levels because both approaches provide unique insights into the contextual effects at play. Although examinations of changed variability are still uncommon in organizational contexts there are good examples in other disciplines attesting to this approach. Research on climate change, for instance, increasingly focuses on (evolutions of) temperature and precipitation means *and* variability across countries. It is one thing to know whether, on average, a particular region (e.g. Europe) has warmed; it is equally relevant to know whether climate has changed in such a way that potential differences between European countries have diminished (or not). A similar case can be made for EL, where, for instance, an absence of changes in mean levels could potentially obscure a tendency towards greater homogeneity across time. To our knowledge, the current study is the first cross-temporal meta-analysis that looks at changes in leader behavior by tracking both variability (or similarity) and mean levels.

Changing EL–Outcome Relationships

Changes in EL can not only manifest via changes in mean levels and variability but also via altering effects of this leadership style on outcome variables (Oc, 2018). One of the most critical findings in leadership concerns the idea that the impact of leadership on effectiveness depends on the context (Yukl, 2012). Hence, changes in (temporal) context can also affect the associations between this leadership style and relevant criteria. Put differently, one final way to test whether EL has become more ingrained in organizational contexts involves estimating its effects on important follower outcomes and comparing these effects across time. Although previous research has already meta-analytically summarized the criterion-related validity of EL (Bedi et al., 2016; Ng & Feldman, 2015), the current study is the first to do this by adopting a temporal lens.

Theoretically, it can be argued that the nature and strength of the relationship between a particular leadership style and follower outcomes will depend on how followers value the behaviors that define this style (Ehrhart & Klein, 2001). If ethical behavior is valued as something positive by followers, then EL should be accompanied by positive outcomes. Moreover, the more a positively valued leadership style is expected by followers, the stronger the association between that style and follower outcomes should be. The concept of implicit leadership theories (ILTs; Offermann et al., 1994) has been put forward to specifically describe those expectations that followers have vis-à-vis their leaders and the behaviors that these followers value. Indeed, characteristics and behaviors such as "ethical", "fair", "honest", and "trustworthy", have emerged in people's ILTs (Offermann & Coats, 2018), which also explains the positive effect of EL on follower outcomes. Importantly, Offermann and Coats (2018) proposed that changes in ILTs can occur when the context evolves such that "major changes in the environmental demands on leaders may change people's perceptions of the characteristics necessary for successful leadership over time" (p. 514). In their own research, however, no substantial changes in the ethical aspects of ILTs were observed in the period between 1994 and 2014.

Taken together, in addition to studying mean levels and variability in EL over time, it is important to also investigate whether the impact of this leadership style on follower outcomes has shifted, as this could reflect potential changes in followers ILTs in response to the societal evolutions described earlier. If this is indeed the case, and based on the principle of leader–follower value congruence (Brown & Treviño, 2009; Edwards & Cable, 2009), it can be expected that, over time, followers have started to respond more strongly to their leaders' ethical behavior, resulting in stronger EL–outcome relationships. This translates into the following hypothesis:

Hypothesis 3: The relationship between EL and follower outcomes has globally grown stronger between 2004 and 2019.

Country-Level and Cultural Moderators

Leadership research has increasingly acknowledged the importance of sociocultural context, as shown by the steep rise of cross-cultural leadership studies (Smith & Peterson, 2017). For EL in particular, research has pointed towards cultural differences in the conceptualization of this concept (Resick et al., 2006), as well as cultural differences in how this style influences follower outcomes (Bedi et al., 2016). In line with this, it is important to take country-level/cultural factors into account when investigating EL through a cross-temporal lens. Specifically, EL relies on external standards and norms which are deeply rooted in national and/or cultural contexts which, from the D-A fit perspective, constitute particular macro-level leadership demands that can evolve over time. However, as can be expected with such high-level context effects, the impact of these national/cultural differences on change trends in EL can take many forms. For example, a new contextual demand could be so country-specific (e.g. the 2002 implementation of the Sarbanes–Oxley Act in the United States) such that its impact on EL potentially also remains restricted to that particular country. Indeed, nation-specific changes in regulation (or a lack thereof) can be particularly impactful for the manifestation of EL, since ethical leaders tend to be morally guided by such external norms and procedures (Lemoine et al., 2019). Alternatively, certain events and/or contextual demands can be global in nature while their societal impact manifests at a different pace and/or to a different degree in various countries. For instance, one study examining the impact of the 2008 financial crisis in 58 countries found that the timing and magnitude of the crisis differed significantly across countries, with countries directly linked to the United States facing the earliest and countries vulnerable before the crisis facing the most severe consequences (Claessens et al., 2010).

Adding further to this complexity, these country-level/cultural variables can be operationalized in substantially different ways (e.g. Dickson et al., 2003). Therefore, to provide a comprehensive examination of these factors with regard to time-related trends in EL, country-level/cultural variables will be examined in three complementary ways in the current study, namely (i) by directly comparing two prototypical countries (i.e. United States and China), (ii) by examining the moderating role of specific cultural dimensions at the country-level (i.e. power distance and uncertainty avoidance), and (iii) by examining the moderating role of one relevant structural country-level variable (i.e. corruption). Although these country-level/cultural variables will

be considered for all three focal change trends examined in this study (i.e. EL means, EL variability, and EL-outcome associations over time), only for the mean-level trends in EL specific hypotheses are formulated. Specifically, our general expectation is that, globally, EL ratings will have increased over time (see Hypothesis 1), but that a number of significant country-level/cultural factors can theoretically and/or empirically be expected to either enhance or diminish this increasing trend. In the sections below, we describe these factors in greater detail along with their anticipated effects on EL means over time (see Hypotheses 4–7).

Conversely, as will be explained in greater detail below, it is much less evident to generate a priori expectations about the role of these country-level/cultural factors in shaping the temporal trends in EL variability and EL-outcome relationships. Instead of formal hypotheses, more exploratory research questions will therefore be used to guide these aspects of our investigation.

Country-Level Comparison: Mean EL Trends in the United States versus China

First, we will directly compare two large countries—i.e. China and the United States—which are often considered prototypes of Eastern and Western cultures, respectively. Although it is challenging to formulate strong a priori expectations, there are a number of indications that lead us to expect a stronger increase in mean EL levels in the United States as compared to China. First, as mentioned earlier, the United States in particular has been a frontrunner in deploying new and large-scale initiatives to prevent unethical and foster ethical behaviors in organizations, across different industries. One significant example is the federal Sarbanes–Oxley Act which was passed in 2002 with bipartisan congressional support to improve auditing and public disclosure in response to several accounting scandals in the early-2000s. To give another example from a different industry, in 2006, the Pentagon announced that all servicemen in Iraq were to undergo additional military ethics training, including lessons in "core warrior values".

However, although perhaps less visible to the Western world, corporate ethics has also become a notable theme in China over the past two decades. Lu (2009) provides an in-depth analysis of the state and evolution of business ethics in China, identifying two separate periods relative to China's entry into the World Trade Organization (WTO) in 2001. In the first pre-WTO period, business ethics was primarily an academic discipline studied by scholars in ethics, philosophy, sociology, and few in economics. The impact on business life was, hence, limited. However, this changed when China entered the WTO, which required a substantially altered legal environment for business operations and a refocus on ethical requirements. In some regions,

these economic changes also made Chinese workers more aware of their rights; an awareness that was also bolstered by the internet. Nevertheless, although business ethics in China has expanded from being an academic discipline to a consideration of central market agents (e.g. legislators, corporate leadership), as noted by Lu (2009), there are many challenges that hindered this consideration to become really tangible: "Here is China's current situation: we have many laws but Chinese corporations still need to build their own mature ideology and ethics" (p. 457). Besides ethics in organizations, there are other illustrative examples of China requiring more time to fully absorb Western ethical awareness. For instance, the #MeToo movement instigated in the United States in 2017 spread internationally and arrived in China (#MiTu) one year later, but structural (e.g. censorship) and cultural factors (e.g. Confucian values associated with respecting seniority) have limited or at least delayed its impact there (Zeng, 2019). Drawing on these analyses and examples, we therefore expect the mean increase in EL to be stronger in the United States compared to China or, more formally:

Hypothesis 4: Mean levels of EL have increased more strongly between 2004 and 2019 in the United States than in China.

The Moderating Role of Culture Dimensions in Mean EL Trends

To delve deeper into the mechanisms underlying potential country-level differences, specific cultural dimensions are also examined in the current study. Specifically, two culture dimensions are selected that are most strongly connected to ethics on a conceptual level and that have also been related to (un)ethical behavior empirically (Seleim & Bontis, 2009; Taras et al., 2010), namely power distance and uncertainty avoidance.

Power distance refers to the extent to which people accept and expect power to be distributed unequally (Hofstede, 1997) and has been linked to the occurrence of unethical behavior: the higher a country scores on power-distance, the higher corruption rates tend to be (Connelly & Ones, 2008; Husted, 1999). One potential explanation can be found in the leadership mechanisms that are supported by high power climates. Specifically, the larger the distance between the persons in authority and their followers, the fewer checks and balances are available to prevent ethical misconduct. Put differently, unethical leadership can be facilitated in cultures scoring high on power distance, because, there, leaders have to face fewer individuals who participate in or even challenge the decision-making process (Javidan et al., 2006). In contrast, when power distance is low, it is easier and probably also more psychologically comfortable for subordinates

to monitor leadership decision-making, and to intervene when there is a risk of transgressions. Remember that 'two-way communication' is indeed a behavioral component that was explicitly incorporated in the definition of EL (Brown et al., 2005). Following this logic, cultures scoring low on power distance can be seen as a more favorable environment for EL to flourish and grow in, compared to cultures high on power distance. This leads to the following hypothesis:

Hypothesis 5: Mean levels of EL have increased more strongly between 2004 and 2019 in countries with low compared to countries with high power distance.

The second cultural dimension investigated in the current study is uncertainty avoidance, which refers to the extent to which people are anxious about unpredictable situations and try to avoid these using policies and procedures (Hofstede, 1997). Although research on the link between this cultural dimension and the prevalence of (un)ethical behavior has yielded mixed findings (e.g. Hofstede, 1997; Husted, 1999; Vitell et al., 1993), there is an obvious link with EL in particular, at least at the conceptual level. Specifically, a defining feature of EL is the deontological focus on compliance and alignment with standards and normative expectations (Lemoine et al., 2019). In this regard, EL can actually be conceived as an instrument in high uncertainty avoiding cultures to realize predictability. The more people strive to avoid uncertainty, the greater their inclination to embrace consistency, structure, formal procedures, and legislation to govern their daily lives. A culture high on uncertainty avoidance thus arguably provides fertile ground for a leadership style which is, among other aspects, centered around establishing and complying to normative standards. This leads us to propose the following hypothesis:

Hypothesis 6: Mean levels of EL have increased more strongly between 2004 and 2019 in countries with high compared to low uncertainty avoidance.

The Moderating Role of National Corruption in Mean EL Trends

Finally, we also explore the role of one country-level structural characteristic, namely national corruption, which is defined as an "abuse of entrusted power for private gain" (Transparency International, 2022). Specifically, the Corruption Perceptions Index (CPI) used in the current study covers a broad range of manifestations of public sector corruption including (but not limited to) bribery, diversion of public funds, and nepotistic appointments in the civil service. Corruption has been identified as a serious impediment to good governance, not only undermining the perceived legitimacy of public institutions (Pollitt, 2016), but also impairing the "moral fabric" (Bashir & Hassan, 2020, p. 674) of a society

as a whole by lowering incentives for people to value honesty and uphold the rule of law. As such, national corruption provides an estimate of the broader ethical context in which organizations and their leaders are embedded, with high corruption levels being a breeding ground for a range of other unethical behaviors, also in leadership roles. Indeed, prior research has shown that EL levels are significantly lower in countries where corruption is more prevalent (Resick et al., 2009), and also within one country, research has demonstrated a negative relationship between corruption and EL in organizations (e.g. in Pakistan; Bashir & Hassan, 2020). Building on these findings suggesting that EL is less likely to flourish and grow in more corrupted contexts, the following hypothesis is proposed:

> *Hypothesis 7: Mean levels of EL have increased more strongly between 2004 and 2019 in countries with low compared to high corruption.*

Country-Level/Cultural Differences in Trends in EL Variability and Outcome Relationships

Whereas previous findings and theory can guide the formulation of a priori hypotheses for country-level/cultural differences in trends in EL means, this is much less the case for trends in EL variability and outcome relationships.

First, when studying changes in variability in leadership-related phenomena across time, scholars have typically focused on the potential evolution of differences between countries. van Hoorn (2019), for instance, examined how differences in managerial values evolved over time, and reported how for some of these values the national differences across the 32 included countries diminished, whereas for others these differences widened. However, to the best of our knowledge, no prior research has investigated the effect of country-level/cultural factors on the variability of leadership perceptions *within* these countries or cultures. In one scenario, when a culture scores higher on a specific culture value which aligns well with the concept of EL (e.g. uncertainty avoidance—see Hypothesis 6), this could potentially create a situational context (Mischel, 1968) in which the range of (perceived) EL behaviors gradually becomes more constrained over time (i.e. variability diminishes). However, it is unclear to what extent cultural factors can effectively influence the (within-culture) dispersion of such behaviors/perceptions. For example, in their influential study, Mastroianni and Gilbert (2023) showed that global consensus on the evolution of morality was unaffected by the cultural context of people, suggesting that the potential of cultural factors to influence the distribution of such perceptions is limited. The following research question is therefore proposed:

> *Research question 1: Are temporal trends in EL variability dependent on country-level/cultural factors?*

Similarly, it is quite unclear what to expect about the impact of country-level/cultural factors on the evolution of the relationship between EL and relevant outcomes across time. As a matter of fact, studies have already yielded mixed findings on whether and how such factors may influence EL-outcome associations at one point in time—thus without considering potential cross-temporal evolutions (e.g. Bedi et al., 2016; Resick et al., 2011). For instance, in their archival study including 59 countries, Resick et al. (2011) found that while the degree to which EL components (i.e. character/integrity, altruism, collective motivation, and encouragement) were endorsed and emphasized indeed varied between cultures, these components were nevertheless universally recognized as important factors contributing to effective leadership across all cultures. In the light of this complexity, we formulate the following research question:

> *Research question 2: Are temporal trends in the relationship between EL and follower outcomes dependent on country-level/cultural factors?*

Non-Linear Shifts

A final aspect of our cross-temporal analysis of EL involves the exact nature of these temporal trends. As described earlier, the period spanning 2004 and 2019 has seen several impactful events, each instigating a stream of initiatives contributing to ethical awareness and enforcement. In addition to the dispersed timing of these events, it is also unclear what the expected time lag can be between an event and the potential changes described above; in other words, how long it takes before the event affects EL in organizations and the way this is experienced by followers. To account for this volatility, the current study will refrain from using strict linear models and will instead turn to a modeling approach that allows flexible estimation of non-linear trends (cf. Costanza et al., 2021—see further in the Methods section).

Method

Transparency and Openness

The sampling plan, data exclusions, manipulations, and measures in the study are described below, and we adhered to the *Journal Article Reporting Standards (JARS)*. Data, code, and research materials are available at the following link: https://osf.io/j7fbu/?view_only=8cbf26d557ea440 fab6e66161d47951e. Data were analyzed using R, version 4.0.0 (R Core Team, 2020), and the specific packages used for our analyses are described below.

Search Strategies and Inclusion Criteria

To enable cross-temporal comparisons, data are used from studies using the same instrument to measure EL. Specifically, we relied on Brown et al. (2005) Ethical Leadership Scale (ELS), given that this scale has the longest history in EL research and is used globally. Because we were interested in analyzing trends in EL between 2004 and (pre-covid) 2019, we searched the Social Sciences Citation Index of the Web of Science for studies using the ELS published between 2005 (i.e. the publication year of the ELS) and 2020.

As we mainly meta-analyze non-focal statistics (i.e. means and standard deviations), publication bias should be less of an issue (Mathur & VanderWeele, 2021) and this allowed us to focus exclusively on published articles. In addition, when analyzing the EL-outcome associations, our focus was on the temporal trends in these relationships rather than on their absolute strength (or statistical significance). Studies were included that met the following criteria: (1) EL mean, standard deviation and/or correlation with an outcome were available, (2) EL was measured with the ELS, (3) EL was rated by subordinates, (4) EL ratings were not manipulated (e.g. in a scenario study), and (5) EL was measured in the workplace context. A flowchart depicting the different identification, inclusion, and exclusion phases can be found in Appendix A in the supplementary information.

Coding Procedure

The first author coded all studies, after which a trained research intern coded a random sample of the studies (50%). The average inter-rater agreement on the study variables (mean, SD, country, and year of data collection) was 93%, and disagreements were discussed until a consensus was reached. Whenever (1) EL mean, (2) EL standard deviation, (3) year of data collection, or (4) the number of leaders rated in the study was missing, we contacted the first authors of the study by email. The authors provided information on at least one of the missing variables for 119 samples (33.1%). When the year of data collection was not available in the article nor provided by the authors (226 samples), we approximated the data collection year by subtracting two years from the date of publication, which is a common approach in CTMA (e.g. Costanza et al., 2021; Wegman et al., 2018). Studies for which no means, standard deviations, or correlation could be obtained (N = 25) were excluded[1]. The final sample consisted of 314 studies, 359 independent samples, and more than 34,518 leaders (based on 63% of the studies for which the number of leaders was available).

In addition to analyzing (trends in) means and standard deviations, we also looked at relationships between EL and follower outcomes in the studies that were retrieved using the procedures described above. A text-search procedure

identified the most frequently considered outcomes across all samples, and six outcomes were selected, meeting a cutoff of a minimum of 15 samples: organizational citizenship behavior (OCB; 47 studies, 51 samples, $k = 65$ effect sizes), employee job performance (25 studies, 26 samples, $k = 24$ effect sizes), job satisfaction (22 studies, 23 samples, $k = 24$ effect sizes), turnover intentions (15 studies, 17 samples, $k = 19$ effect sizes), leader–member exchange (15 studies, 16 samples, $k = 16$ effect sizes), and trust in the leader (14 studies, 17 samples, $k = 19$ effect sizes). An overview of all studies included in this CTMA is provided in Appendix B in the supplementary information.

Statistical Considerations and Procedures

Level of Analysis

For the analyses examining change in EL means and standard deviations, the level of analysis of interest was *leaders*—we aimed to investigate how leader ethical behavior has changed over time. Accordingly, we transformed statistics to reflect the group level across studies consistently. The level of analysis does not affect the mean, so no transformation was needed for this statistic. However, the individual-level *SD* reflects a combination of between-leader variance (the quantity of interest) and within-leader variance. Accordingly, we adjusted individual-level *SD* estimates to reflect only between-leader variance by multiplying the reported *SD* by the square root of the ICC2 (Bartko, 1976). For studies that did not report an ICC2, we imputed the meta-analytic mean ICC2. Focusing exclusively on changes between leaders allowed us to shed light on macro-level trends (e.g. national changes in homogeneity) since changes in homogeneity at the team level were already accounted for. Nonetheless, we were not able to account for changes in homogeneity at the organizational level because single-organization samples were underrepresented ($k = 58$, 16%) to conduct such analyses. From a theoretical perspective, however, societal pressures are likely to trickle down into smaller units, leading organization/team-level groups to become more internally homogeneous as well.

Meta-Regression Models

To examine changes in EL means, standard deviations, and correlations with outcomes over time, we fit meta-regression

---

[1] We also identified several studies which have already been criticized (and some of them retracted) due to erroneous statistical results and inability to provide raw data (https://retractionwatch.com/category/by-author/walumbwa/). Checking for retractions and expressions of concern (using https://pubpeer.com/) led to an additional exclusion of five papers.

models with the year of data collection as a predictor. Rather than a strictly linear model, we specified year effects using restricted cubic splines with knots in 2012, 2014, and 2016 (yielding a 5-degree of freedom basis matrix for the spline; Durrleman & Simon, 1989). This approach allows for a very flexible estimation of nonlinear time trends (cf. Costanza et al., 2021). Compared to standard cubic *b*-splines, restricted cubic splines constrain trends in the variable tails to be linear, yielding better tail performance. For each meta-regression model, we weighted effect sizes by their inverse variance and estimated random effects heterogeneity using the REML estimator (Viechtbauer, 2007). When reporting model predictions in figures and tables, we report both confidence intervals (CI) for the estimated mean effect size and prediction intervals (PI; incorporating both the estimated heterogeneity and uncertainty in the mean; IntHout et al., 2016).

Means and Standard Deviations

Means and *SD*s were converted to percentage of maximum possible (POMP) scores to put the different Likert scales on the same metric (Cohen et al., 1999). For the meta-regression models predicting EL means and *SD*s over time, we analyzed the two statistics together in a multivariate meta-regression model, with the two parameters nested within samples. We meta-analyzed means directly, using $SE_{\widehat{mean}} = \widehat{SD}/\sqrt{N_{leaders}}$ to compute the standard error. We meta-analyzed standard deviations with a log transformation using the methods by Nakagawa et al. (2015): $\widehat{lnSD} = \log\left(\widehat{SD}\right) + 1/\left[2 \times \left(N_{leaders} - 1\right)\right]$ and $SE_{\widehat{lnSD}} = 1/\left[2 \times \left(N_{leaders} - 1\right)\right]$. We freely estimated random effects standard deviations for the two effect sizes, as well as their random effects correlation.

In addition to a baseline model including only the year splines as predictors, we also examined several additional models as moderator analyses. First, we estimated subgroup models separately for samples from the United States ($k = 51$) and China ($k = 62$); these countries had the largest numbers of samples and are quite distinct on many of the national context variables hypothesized to moderate EL time trends. Second, we fit a series of additional meta-regression models adding interactions between the time splines and a national context variable. We fit a moderator model for the Corruption Perception Index (CPI). We also fit moderator models using the hypothesized cultural dimensions (power distance and uncertainty avoidance) and the other dimensions in the Hofstede value system (individualism, masculinity, long-term orientation, and indulgence; Hofstede, 1991) for completeness (see Appendix D and E). For the CPI, we used scores for each country and year (Appendix F shows all countries represented in this study). For the Hofstede dimensions, scores were only available for one time point, so we used the same values for a country across all years.

Outcome Correlations

For the meta-regression models predicting EL–outcome correlations over time, we analyzed correlations with all six outcomes together in a multivariate meta-regression model, with correlations with multiple outcomes nested within samples. We meta-analyzed correlations directly, using $SE_{\hat{r}} = \left(1 - \bar{r}^2\right)/\sqrt{N_{rated} - 1}$, where $\bar{r}$ is the sample-size weighted mean correlation for the specific criterion and N_{rated} is the number of people or groups for which the outcome was measured, to compute the standard error (Schmidt & Hunter, 2015). We adjusted correlations and standard errors for measurement error using coefficient α values for the outcome variable (Wiernik & Dahlke, 2020). We freely estimated random effects standard deviations for each of the six outcome constructs and their random effects intercorrelations. In addition to year splines, we also included the measurement level of the outcome (individual versus group) as a control variable. We included separate intercepts for each outcome construct but estimated only a single time trend across constructs due to sample size concerns. We also estimated a model that included only OCB correlations as a sensitivity analysis.

Analysis Software

We estimated all meta-regression models using the *metafor* package (v. 3.1–3; Viechtbauer, 2010, 2021) in *R* (v. 4.1.2; R Core Team, 2021). We adjusted correlations and standard errors for outcome measurement error using the *psychmeta* package (v. 2.6.0.900; Dahlke & Wiernik, 2019, 2017/2021) in *R*.

Results

Mean Level Trends

Hypothesis 1 predicted an overall increase in mean levels of EL ratings between 2004 and 2019. As shown in Fig. 1 and Table 1, when all countries were included in the model, there were no meaningful mean increases over the 15-year interval (e.g. $\widehat{mean}_{2004} = 70.0$, 95% CI [61.8, 78.1], 95% PI [51.8, 88.2]; $\widehat{mean}_{2019} = 68.9$, 95% CI [64.0, 73.8], 95% PI [51.9, 85.9]). Hypothesis 1 was thus not supported.

Fig. 1 Overall ethical leadership trend over time (All Countries). Ethical leadership mean levels are shown on the upper (blue) line. Ethical leadership variability is shown on the lower (red) line
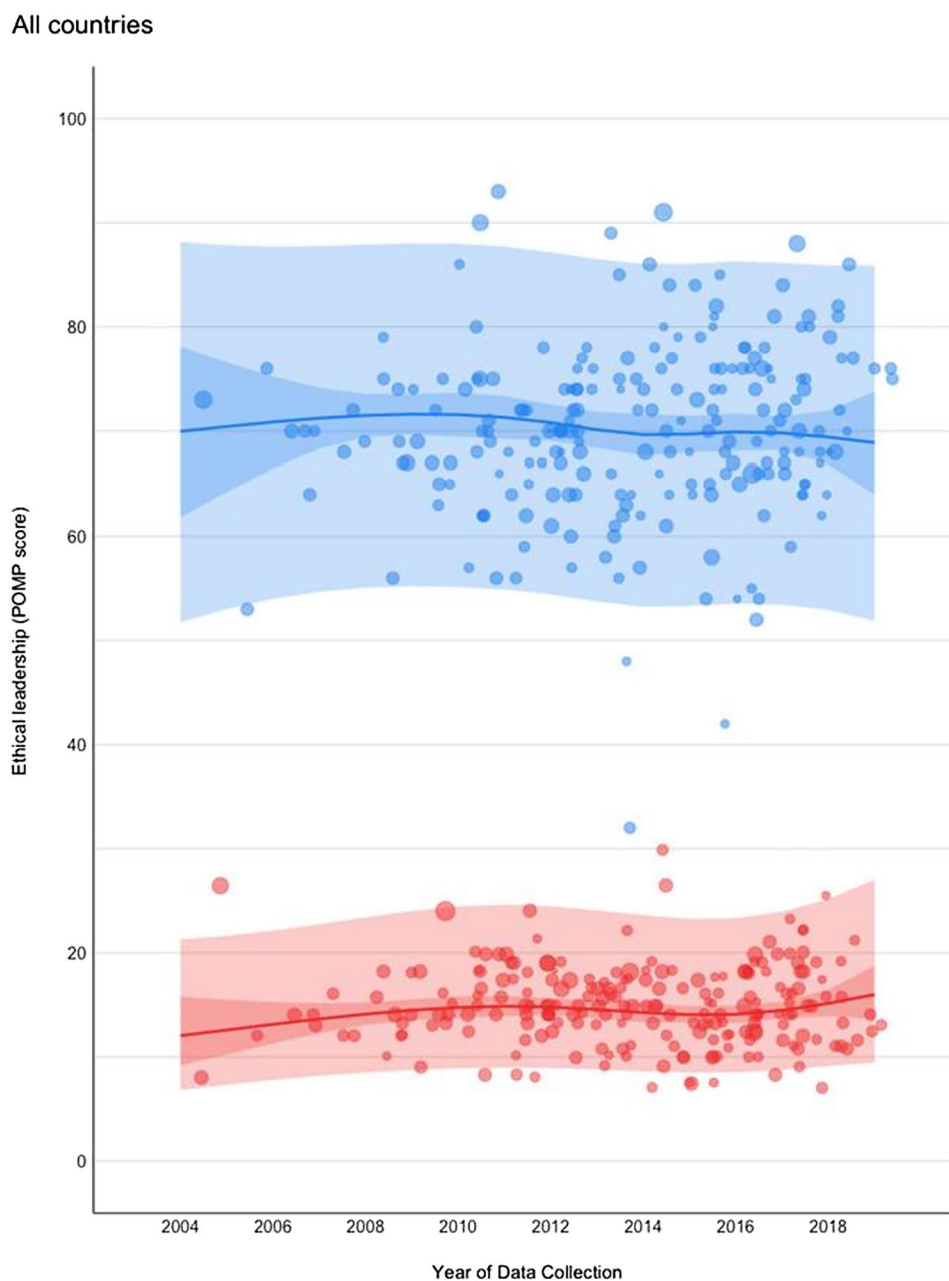
Variability Trends

Hypothesis 2 predicted an overall decrease in variability in EL ratings over the 15-year time interval. As shown in Fig. 1 and Table 1, when all countries were included in the model, there was no apparent change in EL variability over this period (between 2004 and 2019, $\widehat{SD}$ ranged 12.0–16.0, with substantial confidence interval overlaps). Thus, Hypothesis 2 was not supported.

Outcome Correlation Trends

Hypothesis 3 predicted that the relationship between EL and employee outcomes would become stronger between 2004 and 2019 on a global level. The results summarized in Table 2 and Fig. 2 show the predicted values for individual-level associations between EL and employee outcomes over time. These correlations demonstrate an upward trend over time, increasing from an average correlation of $\widehat{\bar{r}}_{2007} = 0.26$, 95% CI [0.10, 0.42], 95% PI [− 0.07, 0.59] in 2007 to $\widehat{\bar{r}}_{2019} = 0.53$, 95% CI [0.38, 0.67], 95% PI [0.20, 0.85] in 2019. For comparison, a linear meta-regression model showed an

Table 1 Overall ethical leadership trend (all countries)

Parameter	Year	Predicted	SE	τ	95% CI		95% PI	
Mean	2004	70.0	4.2	8.3	61.8	78.1	51.8	88.2
Mean	2005	70.4	3.1	8.3	64.3	76.6	53.0	87.8
Mean	2006	70.9	2.2	8.3	66.5	75.2	54.0	87.7
Mean	2007	71.2	1.5	8.3	68.3	74.1	54.7	87.8
Mean	2008	71.5	1.0	8.3	69.4	73.5	55.1	87.9
Mean	2009	71.6	1.0	8.3	69.7	73.6	55.2	88.0
Mean	2010	71.6	1.1	8.3	69.5	73.7	55.2	88.0
Mean	2011	71.3	1.0	8.3	69.3	73.4	54.9	87.7
Mean	2012	70.8	0.9	8.3	69.1	72.5	54.4	87.2
Mean	2013	70.1	0.9	8.3	68.4	71.8	53.8	86.5
Mean	2014	69.6	1.0	8.3	67.7	71.6	53.2	86.0
Mean	2015	69.7	0.9	8.3	68.0	71.4	53.3	86.1
Mean	2016	69.9	0.9	8.3	68.1	71.7	53.5	86.3
Mean	2017	69.8	0.8	8.3	68.2	71.4	53.4	86.2
Mean	2018	69.4	1.3	8.3	66.9	72.0	52.9	85.9
Mean	2019	68.9	2.5	8.3	64.0	73.8	51.9	85.9
SD	2004	12.0	1.1	3.1	9.2	15.8	6.8	21.3
SD	2005	12.6	1.1	3.2	10.2	15.5	7.3	21.6
SD	2006	13.1	1.1	3.4	11.3	15.3	7.8	22.1
SD	2007	13.6	1.1	3.5	12.3	15.2	8.2	22.8
SD	2008	14.1	1.0	3.6	13.1	15.2	8.5	23.4
SD	2009	14.5	1.0	3.7	13.5	15.5	8.7	24.0
SD	2010	14.8	1.0	3.8	13.8	15.8	8.9	24.5
SD	2011	14.9	1.0	3.8	13.9	15.9	9.0	24.6
SD	2012	14.8	1.0	3.8	14.0	15.6	8.9	24.5
SD	2013	14.6	1.0	3.7	13.7	15.4	8.8	24.1
SD	2014	14.3	1.0	3.6	13.3	15.3	8.6	23.6
SD	2015	14.1	1.0	3.6	13.3	14.9	8.5	23.3
SD	2016	14.1	1.0	3.6	13.3	15.0	8.5	23.3
SD	2017	14.5	1.0	3.7	13.7	15.3	8.7	23.9
SD	2018	15.1	1.0	3.9	13.9	16.5	9.1	25.2
SD	2019	16.0	1.1	4.1	13.6	18.8	9.5	27.0

estimated increase in EL–outcome correlations of $+0.11$ (95% CI [0.00, 0.21]) correlation points over ten years. Thus, Hypothesis 3 was generally supported. This trend was consistent whether estimated using all the modeled outcome constructs together (with outcome-specific intercepts) or modeling OCB correlations alone.

As pointed out by an anonymous reviewer, associations between EL and outcomes such as OCB could be inflated when both are rated by the same source (i.e. the follower). Therefore, additional analyses were performed in which those studies were excluded where the outcome variables were also provided by the followers themselves. As shown in Fig. 2b, a linear meta-regression model including only other-rated outcomes showed a similar and even stronger increase in EL–outcome correlations of $+0.19$ (95% CI [-0.06, 0.32]) correlation points over ten years.

Moderator Analyses

For each moderator, we estimated simple trends and predicted EL means and SDs for samples at approximately the quartiles of the moderator values. In the figures, effect sizes are plotted in the panel for the quartile nearest their moderator value.

China versus United States

Meta-regression models were estimated separately for these two countries to examine country-specific trends. Much like the across-countries results, change in EL means in the United States was modest (see Fig. 3 and Table 3). There was a slight upward trend ($\approx 9\%$ of the possible score range) between 2016 ($\widehat{mean}_{2016} = 71.2$, 95% CI [67.7, 74.7], 95% PI [57.0, 85.3]) and 2019 ($\widehat{mean}_{2019} = 80.4$, 95% CI [70.1,

Table 2 Ethical leadership-outcomes correlation trend over time

Parameter	Year	Predicted	SE	τ	95% CI		95% PI	
All outcomes								
r	2007	0.26	0.08	0	0.10	0.42	− 0.07	0.59
r	2008	0.25	0.05	0	0.15	0.36	− 0.06	0.56
r	2009	0.25	0.04	0	0.18	0.33	− 0.05	0.55
r	2010	0.26	0.03	0	0.19	0.32	− 0.04	0.56
r	2011	0.27	0.03	0	0.21	0.34	− 0.02	0.57
r	2012	0.30	0.03	0	0.25	0.36	8.03e-03	0.60
r	2013	0.34	0.03	0	0.28	0.40	0.04	0.64
r	2014	0.35	0.03	0	0.29	0.42	0.05	0.65
r	2015	0.32	0.03	0	0.26	0.37	0.02	0.61
r	2016	0.29	0.03	0	0.23	0.35	− 9.86e-03	0.58
r	2017	0.32	0.03	0	0.26	0.38	0.03	0.62
r	2018	0.41	0.04	0	0.33	0.49	0.11	0.71
r	2019	0.53	0.07	0	0.38	0.67	0.20	0.85
OCB alone								
r	2007	0.41	0.12	0.14	0.18	0.64	0.05	0.78
r	2008	0.36	0.08	0.14	0.19	0.52	0.03	0.69
r	2009	0.32	0.07	0.14	0.19	0.45	6.10e-03	0.64
r	2010	0.31	0.06	0.14	0.19	0.43	− 3.32e-03	0.62
r	2011	0.32	0.06	0.14	0.21	0.44	0.01	0.63
r	2012	0.37	0.05	0.14	0.26	0.47	0.06	0.67
r	2013	0.44	0.06	0.14	0.32	0.55	0.13	0.75
r	2014	0.47	0.06	0.14	0.34	0.59	0.16	0.78
r	2015	0.41	0.05	0.14	0.32	0.51	0.11	0.72
r	2016	0.36	0.05	0.14	0.26	0.46	0.06	0.67
r	2017	0.40	0.05	0.14	0.30	0.50	0.10	0.70
r	2018	0.52	0.07	0.14	0.38	0.65	0.20	0.83
r	2019	0.67	0.13	0.14	0.42	0.93	0.29	1.05

90.7], 95% PI [63.2, 97.6]), but confidence intervals for these predictions overlapped. In China, larger changes over time were apparent. There was a slight increase ($\approx 6\%$ of the possible range) in mean EL ratings from 2011 ($\widehat{mean}_{2011} = 67.9$, 95% CI [63.2, 72.6], 95% PI [48.2, 87.6]) to 2016 ($\widehat{mean}_{2016} = 73.6$, 95% CI [69.4, 77.7], 95% PI [54.0, 93.1]) but confidence intervals overlapped here as well. This increase was followed by a steeper decline ($\approx 16\%$ of the possible range) until 2019 ($\widehat{mean}_{2019} = 57.5$, 95% CI [48.2, 66.8], 95% PI [36.2, 78.8]). Hypothesis 4, predicting a stronger increase in the United States than China, was therefore not fully supported. As for changes in variability, the United States similarly showed little change in SDs over time (between 2007 and 2019, $\widehat{SD}$ ranged 12.6–12.8, with substantial confidence interval overlaps). For China, there appeared to be small fluctuations in EL variability coinciding with the rise and fall in EL means (more variability in 2011–2012 and 2018–2019 when means were lowest, less variability in 2015–2016 when means were highest), but even here confidence intervals overlapped substantially (Fig. 4, Table 4).

Cultural Dimensions

Power Distance Countries scoring high on power distance showed a steep downward trend ($\approx 70\%$ of the possible range) in EL mean levels over the 15-year interval, as shown in Appendix E ($\widehat{mean}_{2004} = 129.0$, 95% CI [97.3, 160.7], 95% PI [92.5, 165.4]; $\widehat{mean}_{2019} = 58.8$, 95% CI [52.5, 65.0], 95% PI [39.8, 77.7]). In contrast, countries scoring low on this value showed an upward trend ($\approx 34\%$ of the possible range) in EL ($\widehat{mean}_{2004} = 51.3$, 95% CI [34.5, 68.1], 95% PI [26.7, 75.8]; $\widehat{mean}_{2019} = 85.4$, 95% CI [72.4, 98.4], 95% PI [63.3, 107.6]). Hypothesis 5, predicting a stronger increase in low compared to high power-distance cultures, was thus supported. There was no significant change in variability among low power distance countries, as shown by the overlapping confidence intervals ($\widehat{SD}_{2004} = 18.0$, 95% CI [10.5, 30.8], 95% PI [8.7, 37.5]; $\widehat{SD}_{2019} = 12.0$, 95% CI [8.2, 17.6], 95% PI [6.4, 22.5]). However, there was a substantial increase in variability among high power distance countries over time ($\widehat{SD}_{2004} = 4.6$, 95% CI [1.5, 14.7], 95% PI [1.3,
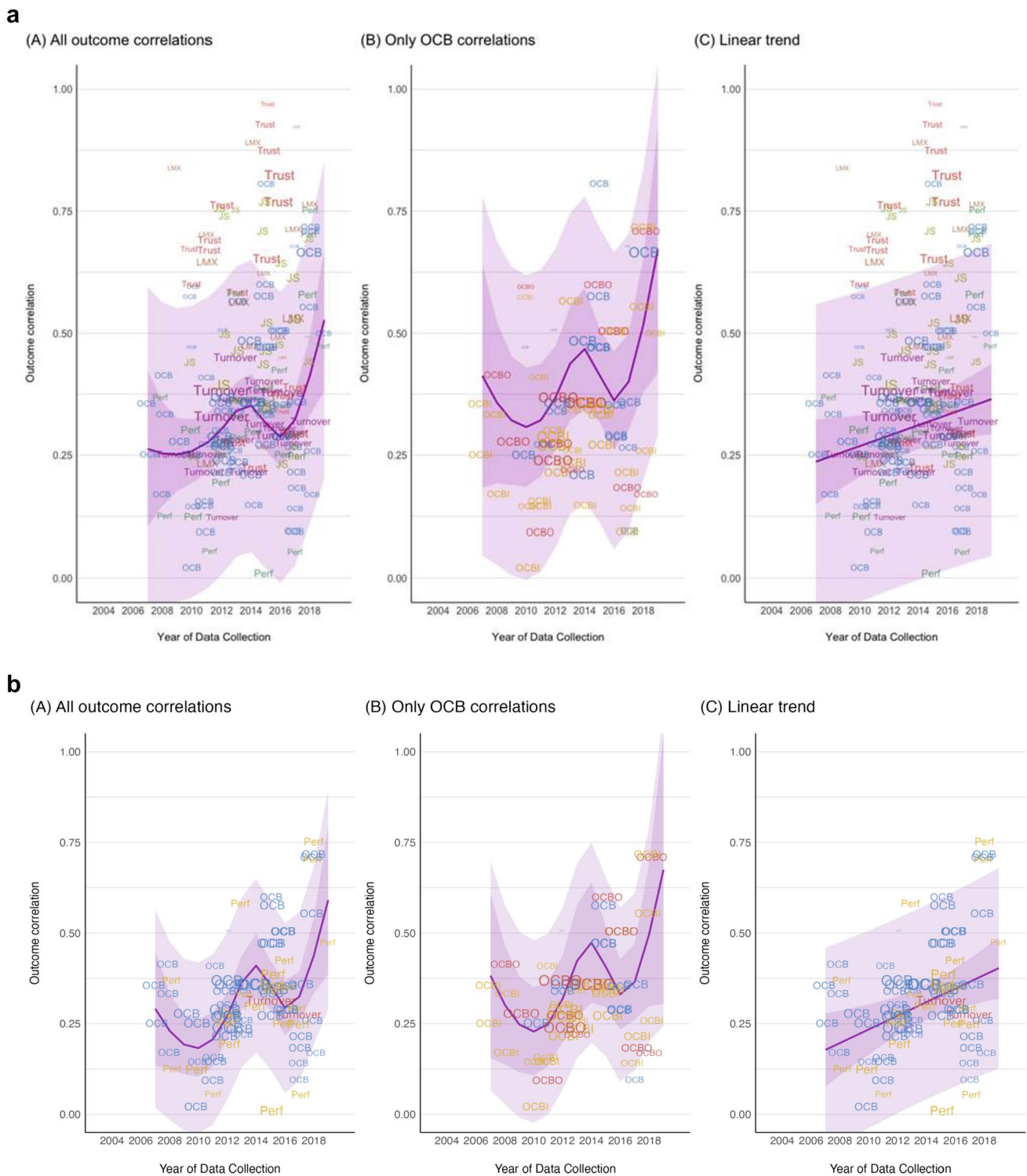
Fig. 2 **a** Outcome-ethical leadership correlation trends over time (all outcomes, OCB, Linear Trend). **b** other-rated outcome-ethical leadership correlation trends over time (all outcomes, OCB, Linear Trend)

16.3]; $\widehat{SD}_{2019} = 19.6$, 95% CI [16.1, 23.8], 95% PI [11.4, 33.5]). Lastly, power distance was not a significant moderator of the relationship between EL and follower outcomes: a

similar increasing trend was observed in both high and low power distance countries (Figure C1).
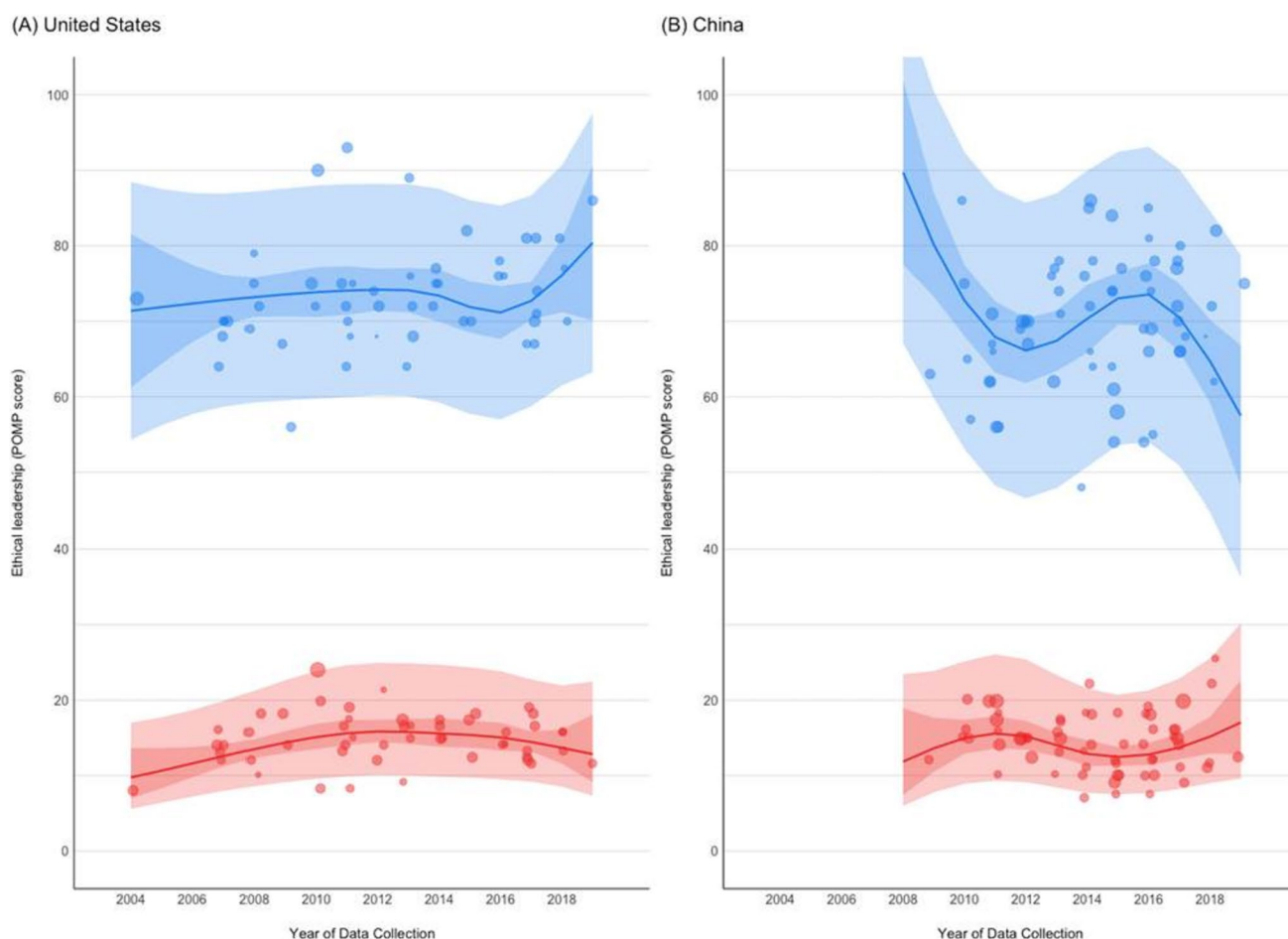
Fig. 3 *Moderation of Ethical Leadership Trends by Geographical Location (United States and China).* Ethical leadership mean levels are shown on the upper (blue) line. Ethical leadership variability is shown on the lower (red) line

Uncertainty Avoidance Similarly, countries scoring high on uncertainty avoidance showed an increase ($\approx 57\%$ of the possible range) in EL, as seen in Fig. 5 and Table 5 ($\widehat{mean}_{2004}$ = 12.9, 95% CI [-29.6, 55.5], 95% PI [-13.5, 64.3]; $\widehat{mean}_{2019}$ = 69.6, 95% CI [60.5, 78.6], 95% PI [48.3, 90.8]), while countries scoring low on uncertainty avoidance showed a strong decreasing trend ($\approx 50\%$ of the possible range) in EL ($\widehat{mean}_{2004}$ = 105.7, 95% CI [75.7, 135.7], 95% PI [70.0, 141.3]; $\widehat{mean}_{2019}$ = 55.5, 95% CI [47.2, 63.7], 95% PI [34.6, 76.4]). Hypothesis 6, predicting a stronger increase in high compared to low uncertainty-avoidance cultures, was thus supported. There was an increase in variability ($\widehat{SD}_{2004}$ = 7.6, 95% CI [3.1, 18.7], 95% PI [2.7, 21.5]; $\widehat{SD}_{2019}$ = 20.1, 95% CI [15.7, 25.8], 95% PI [11.4, 35.6]) among low uncertainty avoidance countries, although confidence intervals overlapped slightly. However, there was no significant decrease in variability ($\widehat{SD}_{2004}$ = 26.2, 95% CI [7.5, 91.8], 95% PI [6.8, 101.6]; $\widehat{SD}_{2019}$ = 16.6, 95% CI [12.8, 21.7], 95%

PI [9.3, 29.7]) among high uncertainty avoidance countries over time. Finally, the relationship between EL and follower outcomes became stronger in both high and low uncertainty avoidance countries over time (Figure C2).

Corruption

As shown in Fig. 6 and Table 6, CPI results show a similar pattern as the comparison of the United States and China. At low levels of corruption (CPI = 0.25), there was little change in EL means over time as shown by the overlapping confidence intervals ($\widehat{mean}_{2004}$ = 58.2, 95% CI [47.1, 69.4], 95% PI [35.9, 80.6]; $\widehat{mean}_{2019}$ = 73.8, 95% CI [61.0, 86.5], 95% PI [50.6, 96.9]). In contrast, at high levels of corruption (CPI = 0.65), a strong decreasing trend ($\approx 95\%$ of the possible range) is observed (e.g. $\widehat{mean}_{2004}$ = 160.1, 95% CI [118.1, 202.0], 95% PI [113.9, 206.2]; $\widehat{mean}_{2019}$ = 65.2, 95% CI [58.5, 71.8], 95% PI [44.7, 85.6]). Hypothesis 7, which predicted a stronger increase in low compared to

Table 3 Moderation of ethical leadership trends by geographical location (United States and China)

United States

Parameter	Year	Predicted	SE	τ	95% CI		95% PI	
Mean	2004	71.4	5.2	7.0	61.2	81.5	54.3	88.5
Mean	2005	71.9	3.8	0.2	64.4	79.4	56.2	87.5
Mean	2006	72.3	2.6	7.0	67.2	77.5	57.7	87.0
Mean	2007	72.8	1.7	0.2	69.5	76.1	58.7	86.9
Mean	2008	73.2	1.3	7.0	70.6	75.8	59.2	87.2
Mean	2009	73.6	1.5	0.2	70.7	76.5	59.5	87.6
Mean	2010	73.9	1.7	7.0	70.6	77.1	59.7	88.0
Mean	2011	74.1	1.6	0.2	70.9	77.3	60.0	88.2
Mean	2012	74.2	1.4	7.0	71.4	77.0	60.2	88.2
Mean	2013	74.1	1.5	0.2	71.2	77.1	60.1	88.2
Mean	2014	73.4	1.8	7.0	69.9	76.9	59.2	87.6
Mean	2015	71.9	1.7	0.2	68.5	75.3	57.7	86.0
Mean	2016	71.2	1.8	7.0	67.7	74.7	57.0	85.3
Mean	2017	72.7	1.3	0.2	70.2	75.2	58.8	86.7
Mean	2018	76.1	2.5	7.0	71.2	81.0	61.5	90.7
Mean	2019	80.4	5.2	0.2	70.1	90.7	63.2	97.6
SD	2004	9.7	1.2	68.3	7.0	13.6	5.6	17.0
SD	2005	10.6	1.1	2.4	8.3	13.7	6.4	17.7
SD	2006	11.6	1.1	81.3	9.8	13.8	7.2	18.7
SD	2007	12.6	1.1	2.8	11.2	14.1	7.9	19.9
SD	2008	13.5	1.0	94.7	12.3	14.8	8.6	21.3
SD	2009	14.4	1.1	3.3	13.0	15.9	9.1	22.7
SD	2010	15.1	1.1	105.8	13.5	16.9	9.6	23.8
SD	2011	15.6	1.1	3.5	14.0	17.4	9.9	24.6
SD	2012	15.8	1.0	111.0	14.4	17.4	10.1	24.9
SD	2013	15.8	1.1	3.6	14.2	17.5	10.0	24.9
SD	2014	15.6	1.1	109.2	13.8	17.6	9.8	24.7
SD	2015	15.4	1.1	3.5	13.7	17.3	9.7	24.3
SD	2016	15.0	1.1	105.4	13.3	17.0	9.5	23.8
SD	2017	14.4	1.0	3.3	13.2	15.8	9.2	22.7
SD	2018	13.7	1.1	95.7	11.6	16.1	8.5	21.9
SD	2019	12.8	1.2	2.9	9.1	18.1	7.3	22.5

China

Parameter	Year	Predicted	SE	τ	95% CI		95% PI	
Mean	2008	89.7	6.2	9.7	77.5	101.9	67.1	112.4
Mean	2009	80.1	3.5	0.3	73.3	87.0	59.8	100.4
Mean	2010	72.7	2.4	9.7	68.0	77.4	53.0	92.4
Mean	2011	67.9	2.4	0.3	63.2	72.6	48.2	87.6
Mean	2012	66.1	2.2	9.7	61.8	70.4	46.5	85.7
Mean	2013	67.4	2.0	0.3	63.4	71.4	47.9	86.9
Mean	2014	70.3	2.3	9.7	65.8	74.8	50.7	90.0
Mean	2015	73.0	1.8	0.3	69.6	76.5	53.6	92.4
Mean	2016	73.6	2.1	9.7	69.4	77.7	54.0	93.1
Mean	2017	70.5	2.3	0.3	66.0	75.0	50.9	90.1
Mean	2018	64.7	2.8	9.7	59.2	70.2	44.8	84.6
Mean	2019	57.5	4.7	0.3	48.2	66.8	36.2	78.8
SD	2008	11.8	1.3	115.4	7.4	18.9	6.0	23.4
SD	2009	13.6	1.1	3.4	10.5	17.7	7.8	23.8

Table 3 (continued)

China

Parameter	Year	Predicted	SE	τ	95% CI		95% PI	
SD	2010	15.0	1.1	145.8	12.9	17.4	8.9	25.1
SD	2011	15.6	1.1	3.9	13.6	17.9	9.3	26.1
SD	2012	15.2	1.1	148.3	13.4	17.2	9.1	25.4
SD	2013	14.0	1.1	3.6	12.5	15.7	8.4	23.3
SD	2014	12.9	1.1	125.5	11.3	14.7	7.7	21.5
SD	2015	12.5	1.1	3.2	11.3	13.8	7.5	20.7
SD	2016	12.8	1.1	124.6	11.4	14.3	7.7	21.3
SD	2017	13.7	1.1	3.5	12.1	15.5	8.2	22.9
SD	2018	15.2	1.1	147.8	12.9	17.8	9.0	25.5
SD	2019	17.0	1.2	4.3	12.8	22.6	9.6	30.2
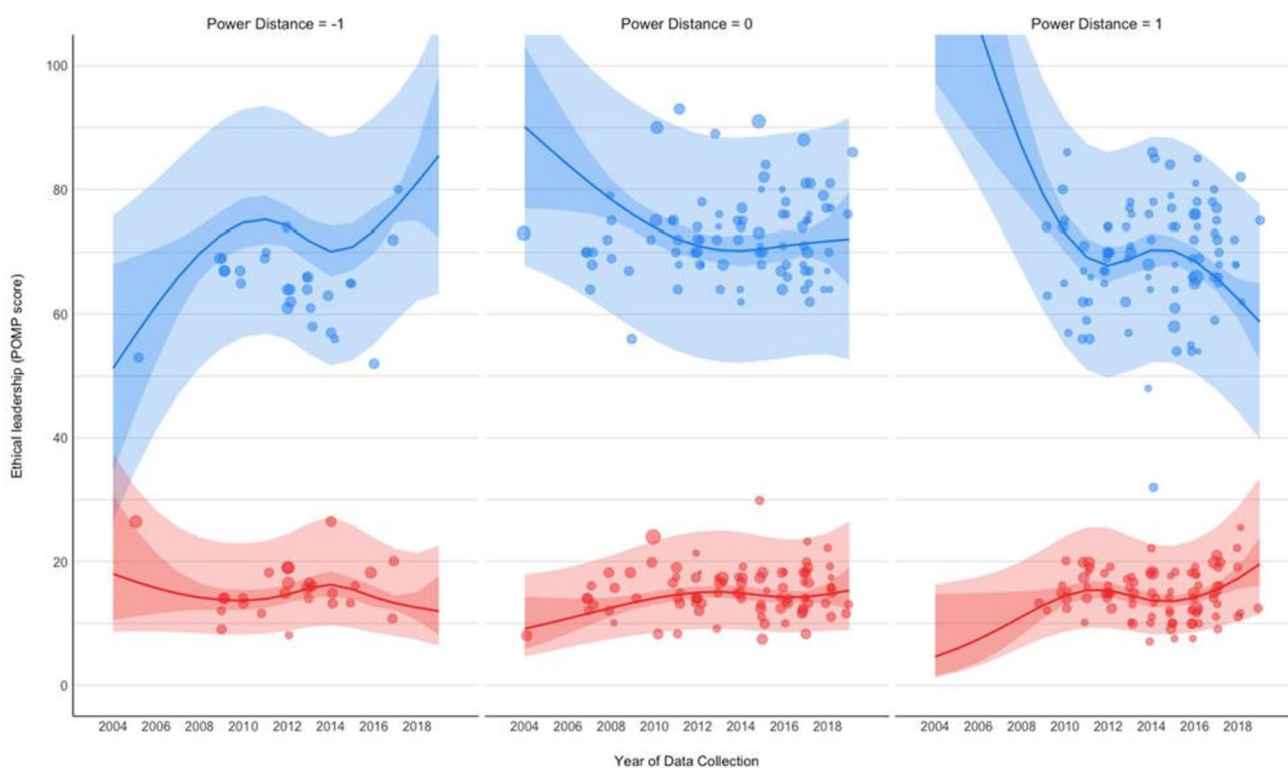
Fig. 4 *Moderation of ethical leadership trends by power distance. Note*. Ethical leadership mean levels are shown on the upper (blue) line. Ethical leadership variability is shown on the lower (red) line

highly corrupted countries, was thus not supported. At low levels of corruption (CPI = 0.25), there was also little change in EL variability over time ($\widehat{SD}_{2004}$ = 15.4, 95% CI [11.3, 20.9], 95% PI [8.7, 27.3]; $\widehat{SD}_{2019}$ = 13.2, 95% CI [9.2, 18.8], 95% PI [7.2, 24.1]), while at high levels (CPI = 0.65), variability tended to increase over time (e.g. $\widehat{SD}_{2004}$ = 1.5, 95% CI [0.3, 6.3], 95% PI [0.3, 6.8]; $\widehat{SD}_{2019}$ = 17.2, 95% CI [14.4, 20.7], 95% PI [10.3, 29.0]). Finally, the relationship between EL and follower outcomes has become stronger in both high and low corruption countries (Figure C3).

Table 4 Moderation of ethical leadership trends by power distance

Parameter	Power distance	Year	Predicted	SE	τ	95% CI		95% PI	
Mean	− 1	2004	51.3	8.6	9.1	34.5	68.1	26.7	75.8
Mean	− 1	2005	56.5	6.5	9.1	43.6	69.3	34.4	78.5
Mean	− 1	2006	61.4	4.7	9.1	52.2	70.6	41.3	81.6
Mean	− 1	2007	65.9	3.1	9.1	59.8	72.1	47.0	84.9
Mean	− 1	2008	69.8	2.2	9.1	65.6	74.0	51.4	88.2
Mean	− 1	2009	72.8	1.9	9.1	69.0	76.5	54.5	91.1
Mean	− 1	2010	74.6	2.0	9.1	70.7	78.6	56.3	93.0
Mean	− 1	2011	75.2	2.0	9.1	71.4	79.0	56.8	93.5
Mean	− 1	2012	74.2	1.7	9.1	70.9	77.4	55.9	92.4
Mean	− 1	2013	71.8	1.7	9.1	68.4	75.3	53.6	90.1
Mean	− 1	2014	70.1	2.1	9.1	66.0	74.2	51.7	88.5
Mean	− 1	2015	70.8	1.9	9.1	67.1	74.6	52.5	89.1
Mean	− 1	2016	73.4	1.9	9.1	69.8	77.1	55.1	91.7
Mean	− 1	2017	77.0	1.2	9.1	74.6	79.3	58.9	95.1
Mean	− 1	2018	81.1	3.1	9.1	75.0	87.1	62.2	100.0
Mean	− 1	2019	85.4	6.6	9.1	72.4	98.4	63.3	107.6
Mean	0	2004	90.1	6.7	9.1	76.9	103.3	67.9	112.4
Mean	0	2005	87.0	5.2	9.1	76.8	97.2	66.4	107.6
Mean	0	2006	84.0	3.8	9.1	76.5	91.4	64.6	103.4
Mean	0	2007	81.1	2.6	9.1	76.0	86.1	62.4	99.7
Mean	0	2008	78.4	1.7	9.1	75.1	81.6	60.1	96.6
Mean	0	2009	75.9	1.3	9.1	73.4	78.4	57.8	94.0
Mean	0	2010	73.9	1.3	9.1	71.4	76.3	55.8	92.0
Mean	0	2011	72.2	1.2	9.1	69.8	74.6	54.1	90.3
Mean	0	2012	71.0	1.0	9.1	69.0	73.1	53.0	89.1
Mean	0	2013	70.4	1.1	9.1	68.3	72.5	52.3	88.4
Mean	0	2014	70.2	1.3	9.1	67.7	72.7	52.1	88.3
Mean	0	2015	70.5	1.2	9.1	68.2	72.9	52.5	88.6
Mean	0	2016	71.0	1.2	9.1	68.6	73.5	52.9	89.1
Mean	0	2017	71.4	1.0	9.1	69.5	73.4	53.4	89.5
Mean	0	2018	71.8	1.9	9.1	68.1	75.4	53.5	90.1
Mean	0	2019	72.1	3.8	9.1	64.6	79.5	52.7	91.5
Mean	1	2004	129.0	16.2	9.1	97.3	160.7	92.5	165.4
Mean	1	2005	117.6	12.8	9.1	92.4	142.7	86.7	148.4
Mean	1	2006	106.5	9.6	9.1	87.7	125.4	80.5	132.5
Mean	1	2007	96.2	6.7	9.1	83.1	109.3	74.0	118.4
Mean	1	2008	86.9	4.2	9.1	78.7	95.2	67.2	106.7
Mean	1	2009	79.1	2.4	9.1	74.4	83.8	60.6	97.6
Mean	1	2010	73.1	1.6	9.1	70.0	76.2	54.9	91.3
Mean	1	2011	69.2	1.6	9.1	66.2	72.3	51.0	87.4
Mean	1	2012	67.9	1.4	9.1	65.1	70.6	49.7	86.0
Mean	1	2013	68.9	1.2	9.1	66.5	71.3	50.8	87.0
Mean	1	2014	70.4	1.4	9.1	67.5	73.2	52.2	88.5
Mean	1	2015	70.2	1.1	9.1	68.1	72.4	52.2	88.3
Mean	1	2016	68.6	1.1	9.1	66.4	70.8	50.5	86.7
Mean	1	2017	65.9	1.1	9.1	63.8	68.0	47.8	83.9
Mean	1	2018	62.5	1.7	9.1	59.2	65.8	44.3	80.7
Mean	1	2019	58.8	3.2	9.1	52.5	65.0	39.8	77.7
SD	− 1	2004	18.0	1.3	4.6	10.5	30.8	8.7	37.5
SD	− 1	2005	16.8	1.2	4.3	11.1	25.5	8.8	32.2
SD	− 1	2006	15.7	1.2	4.0	11.6	21.3	8.7	28.3

Table 4 (continued)

Parameter	Power distance	Year	Predicted	SE	τ	95% CI		95% PI	
SD	− 1	2007	14.8	1.1	3.8	12.0	18.4	8.6	25.6
SD	− 1	2008	14.2	1.1	3.6	12.2	16.5	8.4	23.9
SD	− 1	2009	13.8	1.1	3.5	12.2	15.6	8.2	23.1
SD	− 1	2010	13.7	1.1	3.5	12.2	15.4	8.2	22.9
SD	− 1	2011	14.0	1.1	3.6	12.5	15.6	8.4	23.3
SD	− 1	2012	14.6	1.1	3.7	13.3	16.1	8.8	24.4
SD	− 1	2013	15.7	1.1	4.0	14.1	17.5	9.4	26.2
SD	− 1	2014	16.3	1.1	4.2	14.3	18.4	9.7	27.3
SD	− 1	2015	15.5	1.1	4.0	13.9	17.4	9.3	25.9
SD	− 1	2016	14.2	1.1	3.6	12.7	16.0	8.5	23.8
SD	− 1	2017	13.3	1.0	3.4	12.1	14.5	8.0	22.0
SD	− 1	2018	12.5	1.1	3.2	10.4	15.1	7.4	21.4
SD	− 1	2019	12.0	1.2	3.1	8.2	17.6	6.4	22.5
SD	0	2004	9.1	1.3	2.3	5.8	14.3	4.7	17.9
SD	0	2005	9.9	1.2	2.5	7.0	14.2	5.4	18.4
SD	0	2006	10.8	1.1	2.8	8.3	14.1	6.1	19.0
SD	0	2007	11.7	1.1	3.0	9.7	14.0	6.8	19.9
SD	0	2008	12.5	1.1	3.2	11.1	14.2	7.5	21.0
SD	0	2009	13.4	1.0	3.4	12.2	14.6	8.0	22.2
SD	0	2010	14.1	1.0	3.6	13.0	15.2	8.5	23.3
SD	0	2011	14.6	1.0	3.7	13.6	15.7	8.8	24.3
SD	0	2012	15.0	1.0	3.8	14.1	15.9	9.1	24.8
SD	0	2013	15.1	1.0	3.9	14.1	16.1	9.1	25.0
SD	0	2014	14.9	1.0	3.8	13.8	16.1	9.0	24.8
SD	0	2015	14.5	1.0	3.7	13.5	15.6	8.8	24.1
SD	0	2016	14.2	1.0	3.6	13.2	15.3	8.6	23.6
SD	0	2017	14.3	1.0	3.7	13.4	15.2	8.6	23.7
SD	0	2018	14.7	1.1	3.8	13.2	16.4	8.8	24.6
SD	0	2019	15.3	1.1	3.9	12.3	19.1	8.9	26.5
SD	1	2004	4.6	1.8	1.2	1.5	14.7	1.3	16.3
SD	1	2005	5.9	1.6	1.5	2.3	14.8	2.1	16.8
SD	1	2006	7.4	1.4	1.9	3.7	14.9	3.1	17.5
SD	1	2007	9.2	1.3	2.3	5.6	15.0	4.5	18.5
SD	1	2008	11.1	1.2	2.8	8.1	15.2	6.1	20.0
SD	1	2009	12.9	1.1	3.3	10.8	15.4	7.6	22.0
SD	1	2010	14.5	1.1	3.7	13.1	16.0	8.7	24.1
SD	1	2011	15.4	1.0	3.9	14.0	16.8	9.2	25.5
SD	1	2012	15.4	1.0	3.9	14.1	16.7	9.2	25.5
SD	1	2013	14.5	1.0	3.7	13.5	15.6	8.8	24.1
SD	1	2014	13.7	1.0	3.5	12.6	14.9	8.2	22.8
SD	1	2015	13.6	1.0	3.5	12.7	14.5	8.2	22.5
SD	1	2016	14.2	1.0	3.6	13.2	15.2	8.5	23.5
SD	1	2017	15.4	1.0	3.9	14.4	16.5	9.3	25.5
SD	1	2018	17.3	1.1	4.4	15.6	19.1	10.3	28.8
SD	1	2019	19.6	1.1	5.0	16.1	23.8	11.4	33.5

Discussion

Since the publication of the Ethical Leadership Scale (Brown et al., 2005), academic literature has witnessed a proliferation of studies investigating the nature, the antecedents, and the consequences of this leadership style (Bedi et al., 2016; Hoch et al., 2018; Ng & Feldman, 2015). Similarly, in international business life, corporate ethics
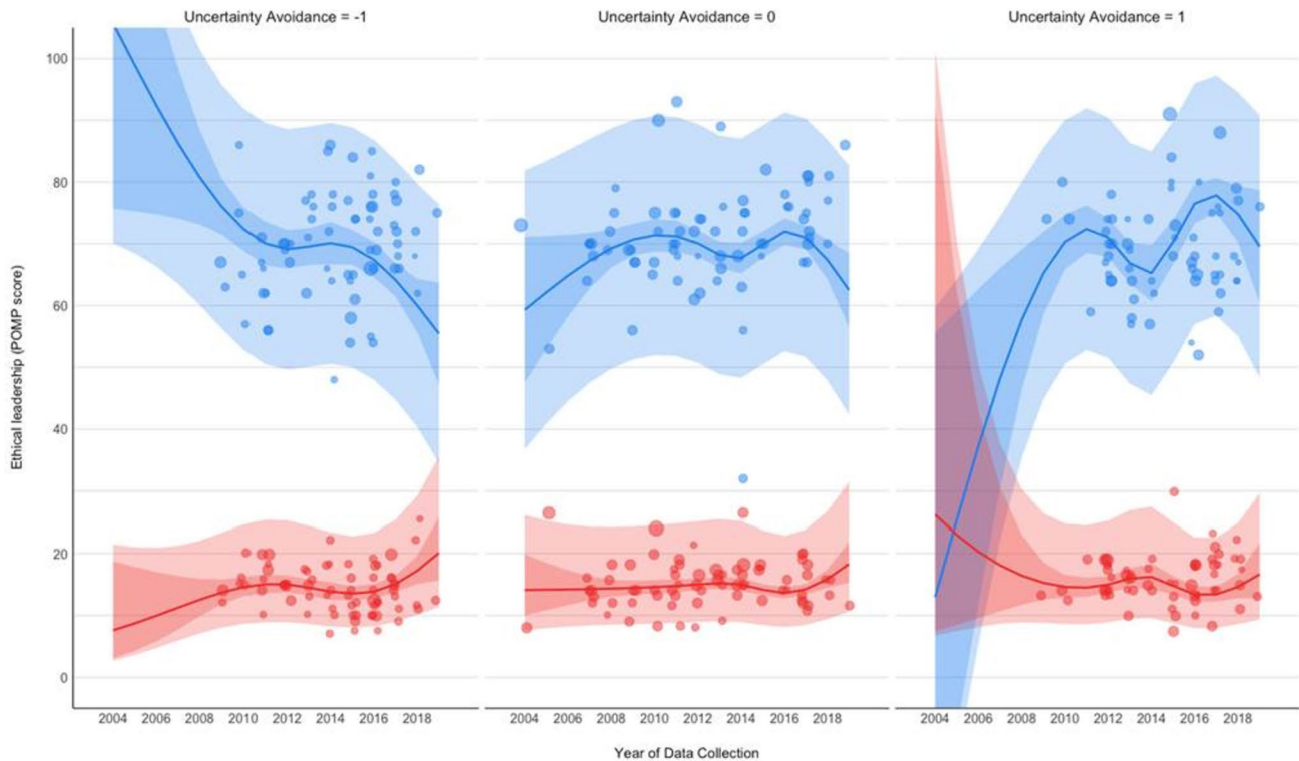
Fig. 5 *M*oderation of ethical leadership trends by uncertainty avoidance. Ethical leadership mean levels are shown on the upper (blue) line. Ethical leadership variability is shown on the lower (red) line

have been put high on the agenda, partly as a response to the public outrage following major ethics scandals around the start of the twenty-first century (Giessner et al., 2015). But to what extent, and in what way, has EL become more ingrained in organizations globally across time? The current study is, to the best of our knowledge, the first to address this question empirically by studying EL using a combination of a cross-temporal and cross-cultural approach.

Our analysis first focused on whether leaders in organization were, on average, rated higher on EL in 2019 as compared to 2004. Interestingly, our results showed no global increase in these mean EL scores, which contradicts the idea that societal evolutions (e.g. public awareness and pressure) and organizational trends (e.g. increased legislation) have universally driven leaders to become more ethical (or less unethical) over time. Similarly, no clear global trend could be observed with regard to the change in variability in EL scores over time, which indicates that leaders have not become more similar over time with respect to this leadership style, at least as rated by their followers. Combined, the large-scale quantitative analyses in this study contradict the belief that EL has globally become more ingrained in organizations during the 15 years between 2004 and 2019. This also suggests that an impactful global event such as

the 2008 financial crisis, with clear and widely communicated linkages with unethical behavior displayed by people in leadership positions, did not uniformly serve as an accelerator of EL worldwide.

Instead, our study revealed that these trends in EL means and variability were heavily contingent on sociocultural factors, which may cancel each other out at the global level. The role of sociocultural factors already became apparent in the specific country-level comparison where we found means in EL to follow substantially different change patterns in the United States (i.e. a stable trajectory) versus in China (i.e. decreasing trend). The question that now emerges is how to interpret these trends exactly? Specifically, a fundamental assumption when developing our expectations with regard to these trends was that the EL ratings made by followers—which were the basis of our analyses—are reflecting the actual ethical behaviors of the leaders they are rating. In this regard, the stable trend in EL means in the United States could be illustrative of a context in which the increasing attention to corporate ethics, as culminated in and reinforced by new legislation, did not translate into behavioral changes of the leaders in this context (e.g. Kaptein, 2010). Moreover, for China this would imply that, across the time period of this study, leaders started to behave in less

Table 5 Moderation of ethical leadership trends by uncertainty avoidance

Parameter	Uncertainty avoidance	Year	Predicted	SE	τ	95% CI		95% PI	
Mean	− 1	2004	105.7	15.3	9.8	75.7	135.7	70.0	141.3
Mean	− 1	2005	98.9	12.1	9.8	75.2	122.6	68.4	129.4
Mean	− 1	2006	92.4	9.0	9.8	74.7	110.0	66.3	118.4
Mean	− 1	2007	86.2	6.2	9.8	74.1	98.4	63.5	109.0
Mean	− 1	2008	80.7	3.9	9.8	73.1	88.4	60.1	101.4
Mean	− 1	2009	76.1	2.3	9.8	71.5	80.6	56.3	95.8
Mean	− 1	2010	72.4	1.8	9.8	68.9	75.9	52.9	91.9
Mean	− 1	2011	70.0	1.8	9.8	66.5	73.6	50.5	89.6
Mean	− 1	2012	69.1	1.6	9.8	66.0	72.2	49.6	88.6
Mean	− 1	2013	69.5	1.5	9.8	66.6	72.5	50.1	89.0
Mean	− 1	2014	70.1	1.7	9.8	66.7	73.5	50.6	89.6
Mean	− 1	2015	69.4	1.4	9.8	66.6	72.2	50.0	88.8
Mean	− 1	2016	67.4	1.5	9.8	64.6	70.3	48.0	86.8
Mean	− 1	2017	64.1	1.4	9.8	61.5	66.8	44.8	83.5
Mean	− 1	2018	60.0	2.2	9.8	55.7	64.3	40.3	79.7
Mean	− 1	2019	55.5	4.2	9.8	47.2	63.7	34.6	76.4
Mean	0	2004	59.3	6.0	9.8	47.5	71.1	36.8	81.9
Mean	0	2005	62.1	4.6	9.8	53.0	71.2	40.9	83.4
Mean	0	2006	64.8	3.4	9.8	58.3	71.4	44.5	85.1
Mean	0	2007	67.3	2.3	9.8	62.8	71.7	47.5	87.0
Mean	0	2008	69.2	1.6	9.8	66.2	72.3	49.8	88.7
Mean	0	2009	70.6	1.3	9.8	68.1	73.2	51.3	90.0
Mean	0	2010	71.4	1.3	9.8	68.7	74.0	52.0	90.7
Mean	0	2011	71.2	1.3	9.8	68.7	73.7	51.8	90.6
Mean	0	2012	70.1	1.1	9.8	67.9	72.2	50.7	89.4
Mean	0	2013	68.2	1.1	9.8	66.1	70.3	48.9	87.5
Mean	0	2014	67.7	1.3	9.8	65.2	70.1	48.3	87.0
Mean	0	2015	69.8	1.1	9.8	67.7	72.0	50.5	89.2
Mean	0	2016	72.0	1.1	9.8	69.8	74.1	52.6	91.3
Mean	0	2017	71.0	1.0	9.8	69.1	72.9	51.7	90.3
Mean	0	2018	67.4	1.6	9.8	64.3	70.5	47.9	86.9
Mean	0	2019	62.5	3.1	9.8	56.5	68.5	42.4	82.6
Mean	1	2004	12.9	21.7	9.8	-29.6	55.5	-33.8	59.7
Mean	1	2005	25.4	17.2	9.8	-8.4	59.2	-13.5	64.3
Mean	1	2006	37.3	13.0	9.8	11.9	62.7	5.4	69.2
Mean	1	2007	48.3	9.1	9.8	30.5	66.1	22.1	74.5
Mean	1	2008	57.7	5.8	9.8	46.3	69.1	35.4	80.1
Mean	1	2009	65.2	3.4	9.8	58.6	71.9	44.9	85.6
Mean	1	2010	70.3	2.2	9.8	65.9	74.7	50.6	90.0
Mean	1	2011	72.4	2.0	9.8	68.5	76.2	52.8	92.0
Mean	1	2012	71.0	1.6	9.8	67.9	74.1	51.6	90.5
Mean	1	2013	66.9	1.7	9.8	63.5	70.2	47.4	86.4
Mean	1	2014	65.3	2.3	9.8	60.7	69.8	45.5	85.0
Mean	1	2015	70.3	1.8	9.8	66.7	73.8	50.7	89.8
Mean	1	2016	76.5	1.6	9.8	73.3	79.7	57.0	96.0
Mean	1	2017	77.8	1.4	9.8	75.0	80.7	58.4	97.2
Mean	1	2018	74.8	2.3	9.8	70.2	79.4	55.0	94.5
Mean	1	2019	69.6	4.6	9.8	60.5	78.6	48.3	90.8
SD	− 1	2004	7.6	1.6	2.0	3.1	18.7	2.7	21.5
SD	− 1	2005	8.7	1.4	2.3	4.3	17.8	3.6	21.0

Table 5 (continued)

| Parameter	Uncertainty avoidance	Year	Predicted	SE	τ	95% CI		95% PI	
SD	− 1	2006	10.0	1.3	2.6	5.8	17.0	4.7	20.9
SD	− 1	2007	11.3	1.2	2.9	7.7	16.4	6.0	21.2
SD	− 1	2008	12.5	1.1	3.3	9.9	15.9	7.1	22.1
SD	− 1	2009	13.7	1.1	3.6	11.8	15.8	8.0	23.3
SD	− 1	2010	14.6	1.1	3.8	13.1	16.2	8.6	24.6
SD	− 1	2011	15.1	1.1	3.9	13.6	16.7	8.9	25.4
SD	− 1	2012	15.0	1.0	3.9	13.7	16.4	8.9	25.3
SD	− 1	2013	14.5	1.0	3.8	13.3	15.8	8.6	24.4
SD	− 1	2014	13.9	1.1	3.6	12.5	15.4	8.2	23.4
SD	− 1	2015	13.5	1.0	3.5	12.5	14.7	8.1	22.8
SD	− 1	2016	13.8	1.0	3.6	12.7	15.1	8.2	23.3
SD	− 1	2017	15.1	1.0	3.9	13.8	16.4	9.0	25.4
SD	− 1	2018	17.2	1.1	4.5	15.1	19.6	10.1	29.2
SD	− 1	2019	20.1	1.1	5.3	15.7	25.8	11.4	35.6
SD	0	2004	14.1	1.2	3.7	10.0	19.9	7.6	26.2
SD	0	2005	14.2	1.1	3.7	10.9	18.5	7.9	25.2
SD	0	2006	14.2	1.1	3.7	11.7	17.3	8.2	24.6
SD	0	2007	14.3	1.1	3.7	12.5	16.3	8.4	24.3
SD	0	2008	14.3	1.0	3.8	13.0	15.8	8.5	24.2
SD	0	2009	14.5	1.0	3.8	13.4	15.6	8.6	24.3
SD	0	2010	14.6	1.0	3.8	13.5	15.7	8.7	24.5
SD	0	2011	14.8	1.0	3.9	13.8	15.9	8.8	24.8
SD	0	2012	15.0	1.0	3.9	14.1	15.9	9.0	25.2
SD	0	2013	15.2	1.0	4.0	14.3	16.2	9.1	25.5
SD	0	2014	15.0	1.0	3.9	13.9	16.1	8.9	25.2
SD	0	2015	14.2	1.0	3.7	13.3	15.0	8.4	23.7
SD	0	2016	13.6	1.0	3.6	12.8	14.5	8.1	22.9
SD	0	2017	14.2	1.0	3.7	13.4	15.1	8.5	23.8
SD	0	2018	15.9	1.0	4.2	14.5	17.4	9.4	26.7
SD	0	2019	18.3	1.1	4.8	15.3	21.9	10.6	31.5
SD	1	2004	26.2	1.9	6.9	7.5	91.8	6.8	101.6
SD	1	2005	23.0	1.7	6.0	8.5	62.2	7.5	70.5
SD	1	2006	20.3	1.5	5.3	9.6	42.9	8.2	50.3
SD	1	2007	18.1	1.3	4.7	10.7	30.6	8.7	37.7
SD	1	2008	16.4	1.2	4.3	11.7	23.0	8.9	30.3
SD	1	2009	15.3	1.1	4.0	12.6	18.5	8.8	26.4
SD	1	2010	14.6	1.1	3.8	12.9	16.5	8.6	24.8
SD	1	2011	14.5	1.1	3.8	13.0	16.1	8.6	24.5
SD	1	2012	15.0	1.0	3.9	13.7	16.4	8.9	25.2
SD	1	2013	16.0	1.0	4.2	14.5	17.6	9.5	26.9
SD	1	2014	16.2	1.1	4.2	14.3	18.4	9.6	27.5
SD	1	2015	14.8	1.1	3.9	13.4	16.4	8.8	25.0
SD	1	2016	13.4	1.1	3.5	12.2	14.7	7.9	22.6
SD	1	2017	13.4	1.0	3.5	12.3	14.6	8.0	22.6
SD	1	2018	14.6	1.1	3.8	12.8	16.7	8.6	24.8
SD	1	2019	16.6	1.1	4.4	12.8	21.7	9.3	29.7
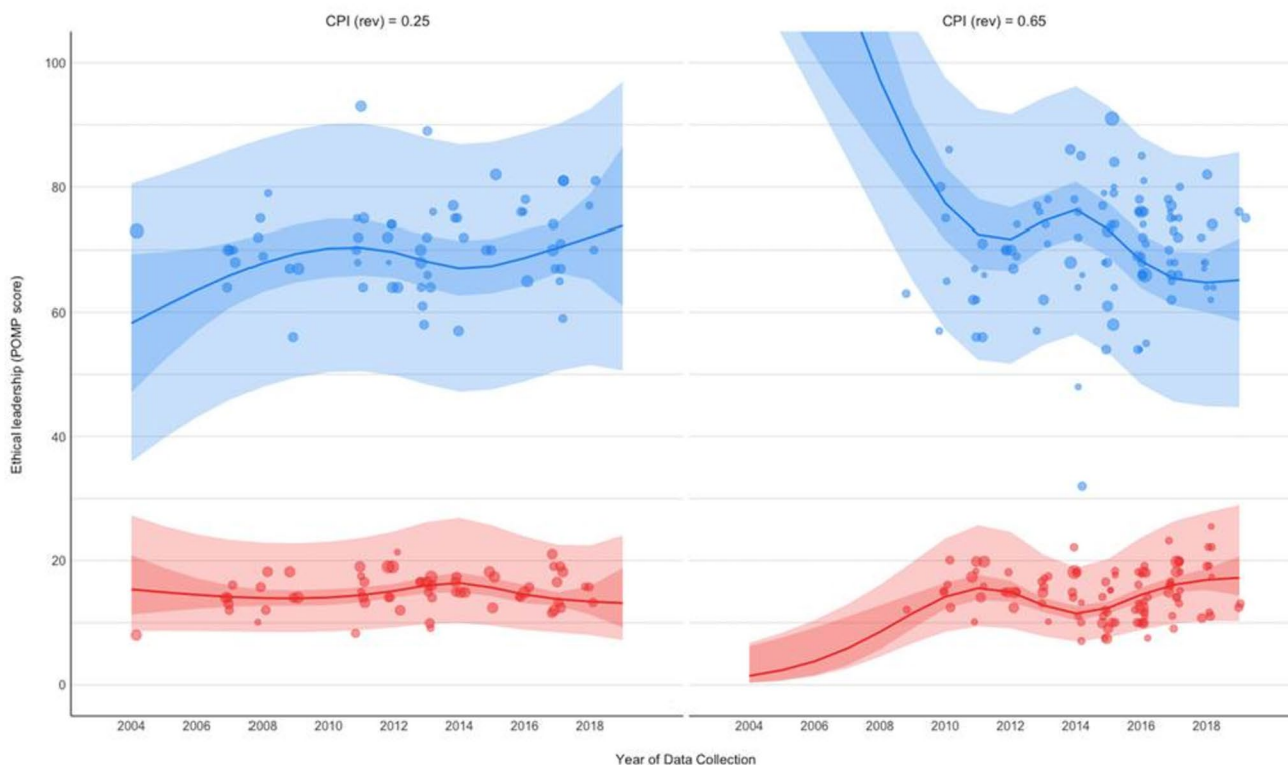
Fig. 6 Moderation of ethical leadership trends by Corruption Perception Index (CPI). Ethical leadership mean levels are shown on the upper (blue) line. Ethical leadership variability is shown on the lower (red) line

ethical ways in the eyes of their followers. In the introduction it was mentioned that China's entrance into the WTO meant the beginning of a new era in which corporate ethics were given significantly more attention. However, as noted by Lu (2009), an emphasis on the profit motive might still outweigh the impact of the slowly developing ethical culture in Chinese business.

However, as was pointed out by an anonymous reviewer, there are also other factors that could have contributed to these observed trends. Specifically, one aspect complicating the interpretation of contextual (e.g. country-level) differences in EL is the inherent contextual nature of the construct itself. As noted by others (Eisenbeiss, 2012), EL is always evaluated (e.g. by followers) *relative to the context* within which leaders are embedded. This could mean that, if the context changes, then (relative) ratings/evaluations of the ethical behaviors of leaders could also shift. For instance, when in general more leaders comply to ethical standards and pay greater attention to business ethics due to increased training and/or enhanced legislation, then leaders following this trend may actually stand out less (and be evaluated less 'favorably') then if only a minority acts in this way. In other words, such 'relative ratings' could make absolute shifts go unnoticed, as was potentially the case for the (lack

of) change in the mean level of EL in the United States. Such a mechanism would also be in line with influential work showing that (ethical) leadership is partly in the eye of the beholder (e.g. Giessner & Van Quaquebeke, 2010; Keck et al., 2020), which implies that leadership ratings are partly the result of the—contextually influenced—standards held by those making these ratings. Note that this mechanism might also serve as an alternative explanation for why EL decreased in China: followers' standards might have risen such that similar leader behavior is actually evaluated more harshly now, leading to an overall decrease in EL ratings there.

Investigating more specific cultural and structural differences between the countries included in the meta-analysis provided further insight into the role that these sociocultural factors play in shaping mean EL over time. Specifically, our results provided evidence for increases in mean EL, but only in countries characterized by lower power distance and higher on uncertainty avoidance. This was in line with our expectation that these specific cultural characteristics constitute a more favorable context for EL to flourish and grow in. Interestingly, for countries characterized by high power distance and low uncertainty avoidance, we did not observe a weaker increase, but instead a *decrease* in mean EL levels.

Table 6 Moderation of ethical leadership trends by corruption perception index (CPI)

Parameter	CPI	Year	Predicted	SE	τ	95% CI		95% PI	
Mean	0.25	2004	58.2	5.7	8.3	47.1	69.4	35.9	80.6
Mean	0.25	2005	61.0	4.4	8.3	52.2	69.7	39.7	82.2
Mean	0.25	2006	63.6	3.4	8.3	56.9	70.2	43.1	84.0
Mean	0.25	2007	65.9	2.7	8.3	60.7	71.2	45.9	86.0
Mean	0.25	2008	67.9	2.4	8.3	63.3	72.5	48.0	87.8
Mean	0.25	2009	69.4	2.3	8.3	64.8	74.0	49.5	89.3
Mean	0.25	2010	70.3	2.4	8.3	65.6	74.9	50.4	90.2
Mean	0.25	2011	70.4	2.3	8.3	65.9	74.9	50.5	90.2
Mean	0.25	2012	69.6	2.1	8.3	65.5	73.7	49.9	89.4
Mean	0.25	2013	68.1	2.1	8.3	64.0	72.3	48.4	87.9
Mean	0.25	2014	67.1	2.3	8.3	62.6	71.5	47.2	86.9
Mean	0.25	2015	67.4	2.2	8.3	63.1	71.8	47.6	87.2
Mean	0.25	2016	68.7	2.3	8.3	64.2	73.2	48.9	88.6
Mean	0.25	2017	70.3	2.0	8.3	66.3	74.3	50.6	90.1
Mean	0.25	2018	72.0	3.5	8.3	65.2	78.9	51.5	92.5
Mean	0.25	2019	73.8	6.5	8.3	61.0	86.5	50.6	96.9
Mean	0.65	2004	160.1	21.4	8.3	118.1	202.0	113.9	206.2
Mean	0.65	2005	142.9	17.0	8.3	109.5	176.2	104.3	181.4
Mean	0.65	2006	126.2	12.9	8.3	101.0	151.5	94.5	158.0
Mean	0.65	2007	110.8	9.1	8.3	93.0	128.7	84.5	137.1
Mean	0.65	2008	97.1	6.0	8.3	85.4	108.9	74.5	119.8
Mean	0.65	2009	85.8	3.8	8.3	78.3	93.3	65.0	106.6
Mean	0.65	2010	77.4	3.0	8.3	71.5	83.2	57.2	97.6
Mean	0.65	2011	72.5	2.9	8.3	66.9	78.1	52.3	92.6
Mean	0.65	2012	71.7	2.6	8.3	66.6	76.7	51.7	91.7
Mean	0.65	2013	74.5	2.2	8.3	70.3	78.8	54.7	94.4
Mean	0.65	2014	76.3	2.3	8.3	71.8	80.9	56.5	96.2
Mean	0.65	2015	73.2	2.2	8.3	68.9	77.5	53.4	93.0
Mean	0.65	2016	68.2	2.2	8.3	63.9	72.6	48.4	88.0
Mean	0.65	2017	65.4	2.2	8.3	61.1	69.8	45.6	85.3
Mean	0.65	2018	64.8	2.4	8.3	60.0	69.5	44.9	84.7
Mean	0.65	2019	65.2	3.4	8.3	58.5	71.8	44.7	85.6
SD	0.25	2004	15.4	1.2	3.8	11.3	20.9	8.7	27.3
SD	0.25	2005	14.9	1.1	3.7	11.8	18.8	8.7	25.6
SD	0.25	2006	14.5	1.1	3.6	12.2	17.2	8.7	24.3
SD	0.25	2007	14.2	1.1	3.5	12.6	16.0	8.6	23.4
SD	0.25	2008	14.0	1.0	3.4	12.7	15.3	8.5	22.9
SD	0.25	2009	13.9	1.0	3.4	12.7	15.2	8.5	22.8
SD	0.25	2010	14.1	1.0	3.5	12.9	15.4	8.6	23.1
SD	0.25	2011	14.5	1.0	3.6	13.3	15.7	8.8	23.7
SD	0.25	2012	15.1	1.0	3.7	14.1	16.3	9.3	24.7
SD	0.25	2013	16.0	1.0	3.9	14.8	17.4	9.8	26.2
SD	0.25	2014	16.4	1.1	4.0	14.9	18.1	10.0	26.9
SD	0.25	2015	15.7	1.0	3.9	14.3	17.2	9.6	25.7
SD	0.25	2016	14.6	1.1	3.6	13.1	16.1	8.9	23.9
SD	0.25	2017	13.8	1.0	3.4	12.8	14.9	8.4	22.6
SD	0.25	2018	13.4	1.1	3.3	11.3	15.9	8.0	22.5
SD	0.25	2019	13.2	1.2	3.2	9.2	18.8	7.2	24.1
SD	0.65	2004	1.5	2.1	0.4	0.3	6.3	0.3	6.8
SD	0.65	2005	2.4	1.8	0.6	0.7	7.6	0.7	8.4
SD	0.65	2006	3.8	1.6	0.9	1.6	9.2	1.4	10.4

Table 6 (continued)

Parameter	CPI	Year	Predicted	SE	τ	95% CI		95% PI	
SD	0.65	2007	5.9	1.4	1.4	3.1	11.0	2.6	13.0
SD	0.65	2008	8.5	1.2	2.1	5.7	12.8	4.5	16.1
SD	0.65	2009	11.5	1.1	2.8	9.1	14.6	6.7	19.8
SD	0.65	2010	14.2	1.1	3.5	12.2	16.4	8.5	23.6
SD	0.65	2011	15.6	1.1	3.8	13.7	17.7	9.4	25.7
SD	0.65	2012	15.0	1.1	3.7	13.4	16.7	9.1	24.6
SD	0.65	2013	12.8	1.0	3.2	11.7	14.0	7.8	21.0
SD	0.65	2014	11.5	1.1	2.8	10.4	12.7	7.0	18.9
SD	0.65	2015	12.4	1.0	3.1	11.5	13.5	7.6	20.3
SD	0.65	2016	14.5	1.0	3.6	13.5	15.7	8.9	23.8
SD	0.65	2017	16.1	1.0	4.0	14.9	17.4	9.9	26.4
SD	0.65	2018	16.9	1.1	4.2	15.3	18.7	10.3	27.8
SD	0.65	2019	17.2	1.1	4.3	14.4	20.7	10.3	29.0

Similarly, those countries with the highest corruption indices demonstrated a decrease in mean EL levels between 2004 and 2019, illustrating the potentially large impact of sociocultural factors on temporal evolutions in leadership. With regard to corruption in particular, our findings are in line with Olken and Pande (2012) who argued that unethical leaders are resilient and adapt to their environment by finding strategies to offset anticorruption policies, ultimately instigating a downward spiral of EL in corrupt countries.

Interestingly, the importance of these country-level/cultural factors in shaping temporal trends in EL also extended to the observed evolutions in the variability in EL across time. Specifically, countries scoring high on power distance, low on uncertainty avoidance, and high on corruption not only demonstrated a decrease in mean levels, but also an increase in variability in EL over time. One interpretation is that in countries/cultures where EL tends to decrease, these contexts simultaneously constitute a weak situation where more interindividual variability in EL is tolerated, or at least perceived. Such moderating effects of cultural characteristics on variability across time challenge the idea of 'universal' trends in shared perceptions, for instance with regard to morality (e.g. Mastroianni & Gilbert, 2023).

Finally, although previous research has meta-analyzed the associations between EL and follower outcomes (e.g. Bedi et al., 2016), the current study was the first to do so while adopting a temporal lens. Importantly, our results confirmed our expectation that, in general, these EL-outcome associations would become stronger over time. So, although mean levels of EL have not changed on average, the positive impact of this leadership style on followers' work outcomes *has* increased over time, which can be interpreted as a signal that followers are becoming more sensitive to this type of leader behavior compared to the past (where potentially other expectations were held towards leaders). Note that this finding is even more remarkable knowing that cross-temporal research is often confronted with the *decline effect*, i.e. the phenomenon whereby the strength of observed effect sizes typically *decreases* as evidence accumulates over time (e.g. due to more stringent methodological procedures; Pietschnig et al., 2019). Finally, in contrast with the observed changes in EL means and variability, this evolution in EL-outcome associations did not depend on the country-level/cultural aspects considered in this study. Previous cross-cultural research had already shown that EL is universally endorsed as important for effective leadership (Resick et al., 2006), as shown by data from the Global Leadership and Organizational Effectiveness (GLOBE) collected during the mid-1990s. The current findings corroborate this idea and extend previous work by showing how the relevance of this leadership style to followers, as shown in EL-outcome associations, has globally become even stronger since the beginning of the twenty-first century.

Theoretical and Methodological Contributions

Our findings can be interpreted in the context of Johns' model of work (2006), which advocates for a simultaneous investigation of different elements of the omnibus context. While most meta-analyses are either cross-temporal (Twenge, 2000; Wegman et al., 2018) or cross-cultural (Eisenbeiss & Brodbeck, 2014; Resick et al., 2006), the current study is one of the very few in the leadership domain to simultaneously examine the interaction between time and culture. Importantly, the current study shows that focusing solely on the *when* or the *where* can indeed obscure important differential trends. By jointly considering the temporal and sociocultural context, the current study provided some unique insights into how the omnibus context is associated with EL.

Specifically, our findings feed into Oc's contextual leadership theory (2018), by documenting the impact of context on

EL for two of the three theorized effects, namely for the base rates or mean levels of EL and for the nature of the relationships between EL and outcomes. Indeed, our research demonstrated that changes in EL base rates over time were heavily dependent on the sociocultural environment, and in the current study these differential trends in EL were interpreted from the perspective of demands-abilities (D-A) person-environment fit (Kristof, 1996) according to which leaders who meet contextual demands become gradually more competitive in and considered by organizations.

In addition, our results revealed that the context impacted EL such that the relationship between EL and positive follower outcomes became stronger over time. Based on the notion of implicit leadership theories (ILTs), this could indicate that, while ethical features might not (yet) have infiltrated followers' prototypical view of leaders (Offermann & Coats, 2018), ethical behavior seems to be increasingly relevant to followers. Such results align with a leader–follower value congruence perspective (Brown & Treviño, 2009; Edwards & Cable, 2009), which predicts positive outcomes when the values endorsed and promoted by ethical leaders coincide better with followers' contemporary values. Importantly, this trend did not depend on country-level/cultural factors, which might reflect a tendency toward a more universal value system (van Hoorn, 2019).

Finally, in order to be able to address these theoretical questions, the current study also adopted two important methodological innovations when meta-analyzing EL scores over time. First, these leadership ratings typically represent a complex combination of individual-level and aggregated data, which the statistical procedures adopted in the current study were able to handle well. In addition, a significant complexity related to our research question concerned the exact timing and/or duration of contextual effects, which were largely unknown parameters. To address this, non-linear modeling techniques were adopted to investigate the role of time (and its interactions with sociocultural factors) in a more flexible and exploratory manner. Both these methodological innovations have already been successfully applied in other disciplines such as in medicine (e.g. Durrleman & Simon, 1989; Sutton et al., 2008), and their use in the current study can serve as a blueprint for future research in leadership and/or management aimed at addressing similar questions.

Practical Implications

The findings of this study have several practical implications. First, on a more general level, our results can inform international policymakers. For instance, our findings suggest that countries defining and relying heavily on standardized procedures (higher uncertainty avoidance) and those promoting equality (lower power distance) are more likely to witness a growth in EL. As such, countries scoring higher on these cultural dimensions (e.g. Germany) may serve as an inspiration to further promote EL in organizations worldwide.

At the organizational level, the strengthening relationship between EL and follower outcomes indicates that employees increasingly value and benefit from EL. As a result, promoting EL seems to be an effective strategy aimed at motivating employees and keeping them engaged (Tu & Lu, 2016). This could involve a wide range of initiatives, such as the selection, monitoring, and development of ethical leaders. Moreover, given that our results showed that these strengthening EL-outcome relationships are a cross-cultural phenomenon, multinational organizations can implement such EL-enhancing programs across their different national branches.

Limitations and Future Directions

The findings reported in this study should be evaluated in light of several limitations. A first limitation relates to the operationalization of EL in this study. Specifically, the scope of the ELS might be relatively narrow considering the broader societal trends described in the introduction. For instance, the ELS might miss out on more specific aspects of EL that have gained importance over the last decade, such as a greater concern for diversity issues and a rising intolerance for sexual misconduct. Future research can therefore evaluate potential changes in EL using alternative measures (e.g. the ELW; Kalshoven et al., 2011) or measures reflecting current ethical concerns, such as the Organizational Tolerance of Sexual Harassment Inventory (OTSHI; Fitzgerald et al., 1997). Nonetheless, cross-temporal meta-analysis typically requires using the same scale in order to make meaningful quantitative comparisons (e.g. Twenge et al. 2008, 2010a, 2010b). We chose the ELS for this purpose because it is the oldest, and most widespread EL instrument to date, which enabled us to go back 15 years and compile EL data based on a consistent set of items. Importantly, this scale was also used in this study because it exhibits notable cross-cultural applicability (e.g. Resick et al., 2011).

A second limitation of our study relates to the concept of EL itself. As explained in Banks et al. (2021) conceptual review, a significant limitation shared across different EL measures is that they often mix together EL behaviors with followers' cognitions and evaluations of leaders' traits, characteristics, and values. As a result, it is difficult to assess to what extent the (lack of) changes in EL are explained by leaders' actual behavior or followers' appraisal of it (see also our earlier discussion of ethical leadership being in the eye of the beholder). For instance, certain items of the ELS require some evaluation

of leaders' intentions and are not directly observable by their followers (e.g. "My leader has the best interests of employees in mind"). Hence, disentangling behavioral and evaluative components of EL might provide a more complete understanding of what underlies potential changes in EL ratings (Banks et al., 2021; Fischer & Sitkin, 2023). Future research could accomplish this by using leader behavioral schemes and frequency scales together with explicitly subjective measures, such as leaders' intentions (Fischer et al., 2021), followers' beliefs about ethical behavior (Peterson et al., 2001), or tolerance for ethical misconduct (Froelich & Kottke, 1991).

Third, it needs to be acknowledged that the Hofstede culture dimensions which were used for the moderator analyses in the current study are not free of limitations. One study has shown, for instance, that the dimension 'uncertainty avoidance' is also closely tied to stress levels, and suggested that true rule orientation aspiration might be better captured by what is included in the GLOBE questionnaire (Venaik & Brewer, 2010). However, important advantages associated with the Hofstede dimensions are their broader societal scope (reflecting not only leaders' ratings as is typical for GLOBE) and their significant potential for cross-disciplinary comparisons, beyond the leadership field. Nevertheless, future research can consider and compare alternative operationalizations of cultural characteristics in cross-temporal research on leadership styles, hereby for instance also considering the role of national (aggregate) personality traits (e.g. Connelly & Ones, 2008).

Finally, although we were able to track changes in EL across a relatively long and highly relevant period of time (i.e. between 2004 and 2019), this approach also came with some limitations. Specifically, one could argue that even longer timeframes (e.g. spanning 20 years) would allow greater insights into how societal movements eventually seep into organizations and the leadership styles that are enacted there. The current study explicitly chose to analyze changes in EL between 2004 (i.e. the year in which the ELS was started to being used) and the end of the pre-COVID era, given that researchers have raised concerns about the validity and generalizability of data collected through (mostly online) questionnaires during COVID times (e.g. Singh & Sagar, 2021). Future research seeking to extend the duration of comparable cross-temporal analyses will need to account for unwanted or random variability associated with (conducting research within) the exceptional COVID context.

Conclusion

Despite a booming concern for ethics worldwide, our findings suggest that EL has generally not become more ingrained in organizations on a global scale in the time period between 2004 and 2019. Instead, our results show that some countries experienced an increase while others observed a decrease in mean EL over time and that cultural factors (e.g. corruption, power distance, uncertainty avoidance) play a critical role in these temporal trends. Next to these mean-level changes, however, EL seems to increasingly pay off for organizations across the globe in the sense that this leadership style has gradually become an even stronger predictor of desirable follower outcomes. The present findings thus provide a solid empirical basis to evaluate cross-temporal trends in EL and its changing impact on organizational behavior worldwide.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

References

Ahmad, S., Fazal-e-hasan, S., & Kaleem, A. (2020). Is the meaning of ethical leadership constant across cultures? A test of cross-cultural measurement invariance. *International Journal of Manpower, 41*(8), 1323–1340. https://doi.org/10.1108/IJM-02-2019-0079

Anand, V., Ashforth, B. E., & Joshi, M. (2004). Business as usual: The acceptance and perpetuation of corruption in organizations. *Academy of Management Executive, 18*(2), 39–53. https://doi.org/10.5465/AME.2004.13837437

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

Bartko, J. J. (1976). On various intraclass correlation reliability coefficients. *Psychological Bulletin, 83*(5), 762–765. https://doi.org/10.1037/0033-2909.83.5.762

Bashir, M., & Hassan, S. (2020). The need for ethical leadership in combating corruption. *International Review of Administrative Sciences, 86*(4), 673–690. https://doi.org/10.1177/0020852318825386

Bedi, A., Alpaslan, C. M., & Green, S. (2016). A meta-analytic review of ethical leadership outcomes and moderators. *Journal of Business Ethics, 139*(3), 517–536. https://doi.org/10.1007/s10551-015-2625-1

Brown, M. E., & Treviño, L. K. (2009). Leader-follower values congruence: Are socialized charismatic leaders better able to achieve it? *Journal of Applied Psychology, 94*(2), 478–490. https://doi.org/10.1037/a0014069

Brown, M. E., Treviño, L. K., & Harrison, D. A. (2005). Ethical leadership: A social learning perspective for construct development and testing. *Organizational Behavior and Human Decision Processes, 97*(2), 117–134. https://doi.org/10.1016/j.obhdp.2005.03.002

Claessens, S., Dell'Ariccia, G., Igan, D., & Laeven, L. (2010). Cross-country experiences and policy implications from the global financial crisis. *Economic Policy, 25*(62), 267–293. https://doi.org/10.1111/j.1468-0327.2010.00244.x

Cohen, P., Cohen, J., Aiken, L. S., & West, S. G. (1999). The problem of units and the circumstance for POMP. *Multivariate Behavioral Research, 34*(3), 315–346. https://doi.org/10.1207/S15327906MBR3403_2

Connelly, B. S., & Ones, D. S. (2008). The personality of corruption: A national-level analysis. *Cross-Cultural Research, 42*(4), 353–385. https://doi.org/10.1177/1069397108321904

Costanza, D. P., Ravid, D. M., & Slaughter, A. J. (2021). A distributional approach to understanding generational differences: What do you mean they vary? *Journal of Vocational Behavior, 127*(1), 103585. https://doi.org/10.1016/j.jvb.2021.103585

Dahlke, J. A., & Wiernik, B. M. (2021). *psychmeta: Psychometric meta-analysis toolkit* (2.6.0.9000). [R package]. https://github.com/psychmeta/psychmeta

Dahlke, J. A., & Wiernik, B. M. (2019). psychmeta: An R package for psychometric meta-analysis. *Applied Psychological Measurement, 43*(5), 415–416. https://doi.org/10.1177/0146621618795933

Den Hartog, D. N. (2015). Ethical Leadership. *Annual Review of Organizational Psychology and Organizational Behavior, 2*(1), 409–434. https://doi.org/10.1146/annurev-orgpsych-032414-111237

Dickson, M. W., Den Hartog, D. N., & Mitchelson, J. K. (2003). Research on leadership in a cross-cultural context: Making progress, and raising new questions. *The Leadership Quarterly, 14*(6), 729–768. https://doi.org/10.1016/j.leaqua.2003.09.002

Dierdorff, E. C., Rubin, R. S., & Morgeson, F. P. (2009). The milieu of managerial work: An integrative framework linking work context to role requirements. *Journal of Applied Psychology, 94*(4), 972–988. https://doi.org/10.1037/a0015456

Durrleman, S., & Simon, R. (1989). Flexible regression models with cubic splines. *Statistics in Medicine, 8*(5), 551–561. https://doi.org/10.1002/sim.4780080504

Edwards, J. R. (1996). An examination of competing versions of the person-environment fit approach to stress. *Academy of Management Journal, 39*(2), 292–339. https://doi.org/10.5465/256782

Edwards, J. R., & Cable, D. M. (2009). The value of value congruence. *Journal of Applied Psychology, 94*(3), 654–677. https://doi.org/10.1037/a0014891

Ehrhart, M. G., & Klein, K. J. (2001). Predicting followers' preferences for charismatic leadership: The influence of follower values and personality. *Leadership Quarterly, 12*(2), 153–179. https://doi.org/10.1016/S1048-9843(01)00074-1

Eisenbeiss, S. A. (2012). Re-thinking ethical leadership: An interdisciplinary integrative approach. *The Leadership Quarterly, 23*(5), 791–808.

Eisenbeiss, S. A., & Brodbeck, F. (2014). Ethical and unethical leadership: A cross-cultural and cross-sectoral analysis. *Journal of Business Ethics, 122*(2), 343–359. https://doi.org/10.1007/s10551-013-1740-0

Fischer, T., & Sitkin, S. B. (2023). Leadership styles: A comprehensive assessment and way forward. *Academy of Management Annals, 17*(1), 331–372. https://doi.org/10.5465/annals.2020.0340

Fischer, T., Tian, A. W., Lee, A., & Hughes, D. J. (2021). Abusive supervision: A systematic review and fundamental rethink. *The Leadership Quarterly, 32*(6), 101540. https://doi.org/10.1016/j.leaqua.2021.101540

Fitzgerald, L. F., Drasgow, F., Hulin, C. L., Gelfand, M. J., & Magley, V. J. (1997). Antecedents and consequences of sexual harassment in organizations: A test of an integrated model. *Journal of Applied Psychology, 82*(4), 578–589. https://doi.org/10.1037/0021-9010.82.4.578

Froelich, K. S., & Kottke, J. L. (1991). Measuring individual beliefs about organizational ethics. *Educational and Psychological Measurement, 51*(2), 377–383. https://doi.org/10.1177/0013164491512011

Garretsen, H., Stoker, J. I., & Weber, R. A. (2020). Economic perspectives on leadership: Concepts, causality, and context in leadership research. *Leadership Quarterly, 31*(3), 101410. https://doi.org/10.1016/j.leaqua.2020.101410

Giessner, S., & Van Quaquebeke, N. (2010). Using a relational models perspective to understand normatively appropriate conduct in ethical leadership. *Journal of Business Ethics, 95*, 43–55. https://doi.org/10.1007/s10551-011-0790-4

Giessner, S. R., Van Quaquebeke, N., van Gils, S., van Knippenberg, D., & Kollée, J. A. J. M. (2015). In the moral eye of the beholder: The interactive effects of leader and follower moral identity on perceptions of ethical leadership and LMX quality. *Frontiers in Psychology, 6*, 1–11. https://doi.org/10.3389/fpsyg.2015.01126

Hoch, J. E., Bommer, W. H., Dulebohn, J. H., & Wu, D. (2018). Do ethical, authentic, and servant leadership explain variance above and beyond transformational leadership? *A Meta-Analysis. Journal of Management, 44*(2), 501–529. https://doi.org/10.1177/0149206316665461

Hofstede, G. (1997). *Cultures and organizations: Software of the mind.* McGraw-Hill.

Husted, B. W. (1999). Wealth, culture, and corruption. *Journal of International Business Studies, 30*(2), 339–360. https://doi.org/10.1057/palgrave.jibs.8490073

IntHout, J., Ioannidis, J. P. A., Rovers, M. M., & Goeman, J. J. (2016). Plea for routinely presenting prediction intervals in meta-analysis. *British Medical Journal Open, 6*(7), e010247. https://doi.org/10.1136/bmjopen-2015-010247

Javidan, M., House, R. J., Dorfman, P. W., Hanges, P. J., & De Luque, M. S. (2006). Conceptualizing and measuring cultures and their consequences: A comparative review of GLOBE's and Hofstede's approaches. *Journal of International Business Studies, 37*(6), 897–914. https://doi.org/10.1057/palgrave.jibs.8400234

Johns, G. (2006). The essential impact of context on organizational behavior. *Academy of Management Review, 31*(2), 386–408. https://doi.org/10.5465/AMR.2006.20208687

Kalshoven, K., Den Hartog, D. N., & De Hoogh, A. H. B. (2011). Ethical leadership at work questionnaire (ELW): Development and validation of a multidimensional measure. *Leadership Quarterly, 22*(1), 51–69. https://doi.org/10.1016/j.leaqua.2010.12.007

Kaptein, M. (2010). The ethics of organizations: A longitudinal study of the U.S. working population. *Journal of Business Ethics, 92*(4), 601–618. https://doi.org/10.1007/s10551-009-0175-0

Keck, N., Giessner, S. R., Van Quaquebeke, N., & Kruijff, E. (2020). When do followers perceive their leaders as ethical? A relational models perspective of normatively appropriate conduct. *Journal of Business Ethics, 164*(3), 477–493. https://doi.org/10.1007/s10551-018-4055-3

Kristof, A. L. (1996). Person-organization fit: An integrative review of its conceptualizations, measurement, and implications. *Personnel Psychology, 49*(1), 1–49. https://doi.org/10.1111/j.1744-6570.1996.tb01790.x

Lemoine, G. J., Hartnell, C. A., & Leroy, H. (2019). Taking stock of moral approaches to leadership: An integrative review of ethical, authentic, and servant leadership. *Academy of Management Annals, 13*(1), 148–187. https://doi.org/10.5465/annals.2016.0121

Levi, M. (2006). The media construction of financial white-collar crimes. *British Journal of Criminology, 46*(6), 1037–1057. https://doi.org/10.1093/bjc/azl079

Livingstone, L. P., Nelson, D. L., & Barr, S. H. (1997). Person-environment fit and creativity: An examination of supply-value and demand-ability versions of fit. *Journal of Management, 23*(2), 119–146. https://doi.org/10.1016/s0149-2063(97)90040-4

Lu, X. (2009). A Chinese perspective: Business ethics in China now and in the future. *Journal of Business Ethics, 86*, 451–461. https://doi.org/10.1007/s10551-008-9857-2

Mastroianni, A. M., & Gilbert, D. T. (2023). The illusion of moral decline. *Nature, 618*(7966), 782–789. https://doi.org/10.1038/s41586-023-06137-x

Mathur, M. B., & VanderWeele, T. J. (2021). Estimating publication bias in meta-analyses of peer-reviewed studies: A meta-meta-analysis across disciplines and journal tiers. *Research Synthesis Methods, 12*(2), 176–191. https://doi.org/10.1002/jrsm.1464

McWilliams, A., Siegel, D. S., & Wright, P. M. (2006). Corporate social responsibility: Strategic implications. *Journal of Management Studies, 43*(1), 1–18. https://doi.org/10.1111/j.1467-6486.2006.00580.x

Mischel, W. (1968). *Personality and assessment*. John Wiley & Sons Inc.

Myers, K. K., & Sadaghiani, K. (2010). Millennials in the workplace: A communication perspective on millennials' organizational relationships and performance. *Journal of Business and Psychology, 25*(2), 225–238. https://doi.org/10.1007/s10869-010-9172-7

Nakagawa, S., Poulin, R., Mengersen, K., Reinhold, K., Engqvist, L., Lagisz, M., & Senior, A. M. (2015). Meta-analysis of variation: Ecological and evolutionary applications and beyond. *Methods in Ecology and Evolution, 6*(2), 143–152. https://doi.org/10.1111/2041-210X.12309

Ng, T. W. H., & Feldman, D. C. (2015). Ethical leadership: Meta-analytic evidence of criterion-related and incremental validity. *Journal of Applied Psychology, 100*(3), 948–965. https://doi.org/10.1037/a0038246

Oc, B. (2018). Contextual leadership: A systematic review of how contextual factors shape leadership and its outcomes. *Leadership Quarterly, 29*(1), 218–235. https://doi.org/10.1016/j.leaqua.2017.12.004

Offermann, L. R., & Coats, M. R. (2018). Implicit theories of leadership: Stability and change over two decades. *Leadership Quarterly, 29*(4), 513–522. https://doi.org/10.1016/j.leaqua.2017.12.003

Offermann, L. R., Kennedy, J. K., Jr., & Wirtz, P. W. (1994). Implicit leadership theories: Content, structure and generalizability. *Leadership Quarterly, 5*, 43–58.

Olken, B. A., & Pande, R. (2012). Corruption in developing countries. *Annual Review of Economics, 4*, 479–509. https://doi.org/10.1146/annurev-economics-080511-110917

Orlitzky, M., & Benjamin, J. D. (2001). Corporate social performance and firm risk: A meta-analytic review. *Business & Society, 40*(4), 369–396. https://doi.org/10.1177/000765030104000402

Peterson, D., Rhoads, A., & Vaught, B. C. (2001). Ethical beliefs of business professionals: A study of gender, age and external factors. *Journal of Business Ethics, 31*(3), 225–232. https://doi.org/10.1023/A:1010744927551

Pietschnig, J., Siegel, M., Eder, J. S. N., & Gittler, G. (2019). Effect declines are systematic, strong, and ubiquitous: A meta-meta-analysis of the decline effect in intelligence research. *Frontiers in Psychology, 10*(1), 1–13. https://doi.org/10.3389/fpsyg.2019.02874

Pollitt, C. (2016). The quest for good governance: How societies develop control of corruption. *International Review of Administrative Sciences, 82*, 617–618.

Pujas, V. (2003). The European Anti-Fraud Office (OLAF): A European policy to fight against economic and financial fraud? *Journal of European Public Policy, 10*(5), 778–797. https://doi.org/10.1080/1350176032000124087

R Core Team. (2020). *R: A language and environment for statistical computing* (v. 4.0.0). R Foundation for Statistical Computing. https://www.r-project.org/.

R Core Team. (2021). *R: A language and environment for statistical computing* (v. 4.1.2). R Foundation for Statistical Computing. https://www.r-project.org/.

Resick, C. J., Hanges, P. J., Dickson, M. W., & Mitchelson, J. K. (2006). A cross-cultural examination of the endorsement of ethical leadership. *Journal of Business Ethics, 63*(4), 345–359. https://doi.org/10.1007/s10551-005-3242-1

Resick, C. J., Martin, G. S., Keating, M. A., Dickson, M. W., Kwan, H. K., & Peng, C. (2011). What ethical leadership means to me: Asian, American, and European perspectives. *Journal of Business Ethics, 101*(3), 435–457. https://doi.org/10.1007/s10551-010-0730-8

Resick, C. J., Mitchelson, J. K., Dickson, M. W., & Hanges, P. J. (2009). Culture, corruption, and the endorsement of ethical leadership. In W. Mobley, Y. Wang, & M. Li (Eds.), *Advances in Global Leadership* (pp. 113–144). Emerald Group Publishing.

Rudolph, C. W., Costanza, D. P., Wright, C., & Zacher, H. (2020). Cross-temporal meta-analysis: a conceptual and empirical critique. *Journal of Business and Psychology, 35*(6), 733–750. https://doi.org/10.1007/s10869-019-09659-2

Schmidt, F. L., & Hunter, J. E. (2015). Methods of meta-analysis: Correcting error and bias in research findings. *SAGE Publications, Ltd.* https://doi.org/10.4135/9781483398105

Schroders Global Investor Study. (2017). *The Schroders Global Investor Study 2017: The global rise of sustainable investing.* https://www.schroders.com/en/uk/private-investor/insights/global-investor-study/2017-findings/sustainability/

Seleim, A., & Bontis, N. (2009). The relationship between culture and corruption: A cross-national study. *Journal of Intellectual Capital, 10*(1), 165–184. https://doi.org/10.1108/14691930910922978

Singh, S., & Sagar, R. (2021). A critical look at online survey or questionnaire-based research studies during COVID-19. *Asian Journal of Psychiatry, 65*(August), 102850. https://doi.org/10.1016/j.ajp.2021.102850

Smith, P. B., & Peterson, M. F. (2017). Cross-Cultural Leadership. In M. J. Gannon & K. L. Newman (Eds.), *The Blackwell Handbook of Cross-Cultural Management* (pp. 217–235). Hoboken: Blackwell Publishing Ltd.

Stevens, M. J., & Campion, M. A. (1994). The knowledge, skill, and ability requirements for teamwork: Implications for human resource management. *Journal of Management, 20*(2), 503–530. https://doi.org/10.1177/014920639402000210

Strobel, M., Tumasjan, A., & Welpe, I. (2010). Do business ethics pay off?: The influence of ethical leadership on organizational

attractiveness. *Zeitschrift Für Psychologie/journal of Psychology, 218*(4), 213–224. https://doi.org/10.1027/0044-3409/a000031

Sutton, A. J., Kendrick, D., & Coupland, C. A. C. (2008). Meta-analysis of individual- and aggregate-level data. *Statistics in Medicine, 27*(5), 651–669. https://doi.org/10.1002/sim.2916

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*(3), 405–439. https://doi.org/10.1037/a0018938

Transparency International. (2022). What is corruption? https://www.transparency.org/en/what-is-corruption

Trevino, L. K., & Nelson, K. A. (2021). *Managing business ethics: Straight talk about how to do it right*. Hoboken: John Wiley & Sons.

Tu, Y., & Lu, X. (2016). Do ethical leaders give followers the confidence to go the extra mile? The moderating role of intrinsic motivation. *Journal of Business Ethics, 135*(1), 129–144. https://doi.org/10.1007/s10551-014-2463-6

Twenge, J. M. (2000). The age of anxiety? The birth cohort change in anxiety and neuroticism, 1952–1993. *Journal of Personality and Social Psychology, 79*(6), 1007–1021. https://doi.org/10.1037/0022-3514.79.6.1007

Twenge, J. M., Campbell, S. M., Hoffman, B. J., & Lance, C. E. (2010a). Generational differences in work values: Leisure and extrinsic values increasing, social and intrinsic values decreasing. *Journal of Management, 36*(5), 1117–1142. https://doi.org/10.1177/0149206309352246

Twenge, J. M., Gentile, B., DeWall, C. N., Ma, D., Lacefield, K., & Schurtz, D. R. (2010b). Birth cohort increases in psychopathology among young Americans, 1938–2007: A cross-temporal meta-analysis of the MMPI. *Clinical Psychology Review, 30*(2), 145–154. https://doi.org/10.1016/j.cpr.2009.10.005

Twenge, J. M., Konrath, S., Foster, J. D., Campbell, W. K., & Bushman, B. J. (2008). Egos inflating over time: A cross-temporal meta-analysis of the narcissistic personality inventory. *Journal of Personality, 76*(4), 875–902. https://doi.org/10.1111/j.1467-6494.2008.00507.x

US Forum for Sustainable and Responsible Investment [US SIF]. (2020). *Report on US Sustainable and Impact Investing Trends*. https://www.ussif.org/currentandpast

van Hoorn, A. (2019). Generational shifts in managerial values and the coming of a unified business culture: A cross-national analysis using european social survey data. *Journal of Business Ethics, 155*(2), 547–566. https://doi.org/10.1007/s10551-017-3488-4

Venaik, S., & Brewer, P. (2010). Avoiding uncertainty in Hofstede and GLOBE. *Journal of International Business Studies, 41*(8), 1294–1315. https://doi.org/10.1057/jibs.2009.96

Viechtbauer, W. (2021). *metafor: Meta-analysis package for R (3.1–3) [R package]* (v. 3.1–3). https://github.com/wviechtb/metafor

Viechtbauer, W. (2007). Confidence intervals for the amount of heterogeneity in meta-analysis. *Statistics in Medicine, 26*(1), 37–52. https://doi.org/10.1002/sim.2514

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36*(3), 1–48.

Vitell, S. J., Nwachukwu, S. L., & Barnes, J. H. (1993). The effects of culture on ethical decision-making: An application of Hofstede's typology. *Journal of Business Ethics, 12*(10), 753–760. https://doi.org/10.1007/BF00881307

Wegman, L. A., Hoffman, B. J., Carter, N. T., Twenge, J. M., & Guenole, N. (2018). Placing job characteristics in context: Cross-temporal meta-analysis of changes in job characteristics since 1975. *Journal of Management, 44*(1), 352–386. https://doi.org/10.1177/0149206316654545

Wiernik, B. M., & Dahlke, J. A. (2020). Obtaining unbiased results in meta-analysis: The importance of correcting for statistical artifacts. *Advances in Methods and Practices in Psychological Science, 3*(1), 94–123. https://doi.org/10.1177/2515245919885611

Yukl, G. (2012). Effective leadership behavior: What we know and what questions need more attention. *Academy of Management Perspectives, 26*(4), 66–85. https://doi.org/10.5465/amp.2012.0088

Zeng, J. (2019). You Say #MeToo, I Say #MiTu: China's Online Campaigns Against Sexual Abuse. *Bianca Fileborn, Rachel Loney-Howes #MeToo and the Politics of Social Change* (pp. 71–83). Cham: Springer International Publishing.

Authors and Affiliations

Justine Amory[1] · **Bart Wille[1]** · **Brenton M. Wiernik[2]** · **Sofie Dupré[1]**

✉ Bart Wille
  bart.wille@ugent.be

  Justine Amory
  justine.amory@ugent.be

  Brenton M. Wiernik
  brenton@wiernik.org

  Sofie Dupré
  sofie.dupre@ugent.be

[1] Department of Developmental, Personality and Social Psychology, Ghent University, Henri Dunantlaan 2, 9000 Ghent, Belgium

[2] Department of Psychology, University of South Florida, 4202 East Fowler Avenue, PCD 4118G, Tampa, FL 33620-7200, USA