The Leipzig Leadership Model: Measuring leadership orientations

Timo Meynhardt[1,2] · Josephina Steuber[1] · Maximilian Feser[1]

Abstract

The Leipzig Leadership Model (LLM) connects theory and practice. Building on actor-world relations and insights from motivational psychology (actor-action relations), the LLM proposes a holistic framework able to integrate the existing plethora of leadership theories and styles. This opens up a new perspective on a comprehensive understanding of leadership roles. With its four leadership orientation dimensions of purpose, entrepreneurial spirit, responsibility, and effectiveness, the LLM enables leaders to identify and reflect on relevant leadership competencies. In order to facilitate future research on the LLM and its dimensions, we report on two studies that are developing and validating a 32-item LLM-based scale. We applied oblique bifactor target rotation in a bifactor Exploratory Structural Equation Model within CFA approach in a German sample ($N = 309$) with robust WLSMV-estimates to fit an LLM-based model to the data. The results suggest a good fit. Furthermore, as ECV, PUC and ARPB support the multidimensional nature of the scale, we report the appropriate bifactor statistical indices. After parallel back translation, an English version of the scale was tested in a second sample ($N = 311$) to replicate our earlier findings. This study facilitates future empirical research by providing a concise and integrative self-rating measure of leadership orientations. We further strengthen the scientific foundation of the LLM by empirically testing its conceptually developed four-factor structure. The scale provides a starting point for further research into leadership orientations (also as standalone subscales), and offers an applicable guideline for self-reflection and decision-making.

Keywords Leadership · Competencies · Purpose · Common good · Value contribution

"Leadership means accepting responsibility for oneself, others and the future" (Kirchgeorg et al., 2017). Today, this future is fundamentally shaped by a multitude of grand challenges (GCs), such as climate change, globalization, geopolitics, or digital transformation, all of which have serious implications for leadership (Meynhardt et al., 2022; Brammer et al., 2019; van Tulder, 2018). In general, GCs are characterized by non-linear system dynamics and unknown solutions with intertwined social and technical elements (Eisenhardt et al., 2016). These characteristics result in complexity and ambiguity; and even though technical knowledge is still important, value judgments

become indispensable in the face of uncertainty without predefined solutions (Brammer et al., 2019). Or, as Meynhardt (2004) put it: "Knowing is more than feeling, but often feeling is the only way of knowing" (p. 11).

In the context of GCs, successful leadership becomes a question of a leader's ability to act creatively and pro-actively in the face of uncertainty and to cope with the unknown. In other words, successful leadership becomes a matter of a person's leadership competencies encompassing the specific dispositions that enable self-organized performance in highly complex contexts (Erpenbeck et al., 2017). Consequently, leadership competency models have gained a strong foothold in (HR) practice, given these models' benefits in both allowing organizations to communicate underlying values and expectations of leadership, and simultaneously providing the necessary orientation for leaders in the form of concise lists of expected behaviors (Boyatzis, 2008).

At the same time, theory-practice gaps continue to exist, as the implementation of competency models has mostly

✉ Timo Meynhardt
timo.meynhardt@hhl.de

[1] Dr. Arend Oetker Chair of Business Psychology and Leadership, HHL Leipzig Graduate School of Management, Leipzig, Germany

[2] Center for Leadership and Values in Society, University of St. Gallen, St. Gallen, Switzerland

restricted itself to non-academically founded "best practices". As critics argue, this approach can lead to highly reductionist results that lack generalizability and a future orientation, while also often neglecting ethical, emotional, and social aspects of leadership (Bolden & Gosling, 2006; Hollenbeck et al., 2006). Given that GCs constantly evolve along with technical and societal progress, leadership competencies themselves must be revised continuously and therefore need orientation regarding the reasons why they should be applied. In order to bridge theory-practice gaps, leadership frameworks that holistically structure the revision of leadership competencies, while allowing individual degrees of freedom in diverse practical contexts, are needed (Yukl & Gardner, 2020; Antonakis, 2017). This is why the Leipzig Leadership Model (LLM; Kirchgeorg et al., 2017) aims to familiarize leaders with ways in which to organize experience, direct attention, and legitimize decisions and actions.

The LLM encourages critical and continuous self-reflection along its four dimensions of leadership (purpose, entrepreneurial spirit, responsibility and effectiveness), while measuring them against their value contribution (Kirchgeorg et al., 2017). It reduces leadership complexity to "the barest minimum possible" (Meynhardt et al., 2019, p. 33) by structuring reflections according to its guiding questions. This results in a dynamic framework that tasks leaders with autonomously identifying and reflecting on relevant leadership competencies, while simultaneously providing them with a holistic grasp of their leadership role. However, while Kirchgeorg et al. (2017) and Meynhardt et al. (2019) introduced the LLM conceptually and elaborated on the model's theoretical foundation, research on the LLM is still missing a scale to empirically test the model's assumptions. In this paper, we therefore develop and validate a 32-item LLM-based scale in a German and English sample in order to enable future empirical research within the LLM-framework.

Theoretical background

The LLM measures leadership against its value contribution (Kirchgeorg et al., 2017), thereby building especially on the early works of Chester Barnard (1938) and Drucker (1973), and work focusing on servant leadership (Liden et al., 2008). In contrast to a mere focus on results, the idea of contribution acknowledges that no person can directly control the outcome of an action in a complex situation, but can contribute to it and exercise leadership in that way. The ability to provide a value contribution is determined on a micro level by a person's leadership competencies as "dispositions of self-organized action" (Erpenbeck et al., 2017, p. XXIII).

Leadership in practice is significantly impacted by GCs and their resulting complexity, ambiguity, and non-linear system dynamics (Meynhardt et al., 2019; Brammer et al., 2019; Eisenhardt et al., 2016). Even on a smaller, day-to-day scale, researchers find that leaders are increasingly confronted with adaptive challenges (Heifetz, 2006), for which solutions cannot be derived from existing technical knowledge. To overcome adaptive challenges, Heifetz concludes that leaders need a "retooling, in a sense, of people's own ways of thinking and operating" (p. 77). Developmental psychologists picked up on this notion when formulating the concept of transformational learning (e.g. Kegan, 2009), manifesting an epistemological change where intrinsic value-based knowledge becomes central to decision making.

To facilitate such transformational learning, the four LLM dimensions provide a holistic basis for orienting and framing leadership competencies: Where should I focus when I want to become a "good" leader? How do I best use my competencies without getting unbalanced and paralyzed? It is important to note that the LLM's four leadership dimensions themselves do not represent competencies in the sense of 'being able to perform a certain task.' There is neither a purpose competency nor a responsibility competency. Instead, the LLM dimensions focus on the very core value of competencies, orienting a capability toward an action. For example, strategic orientation as a leadership competency may be motivated by various different purposes and oriented to a diverse set of behaviors. Without elaborating on the why, what and how, strategic orientation becomes a competency without a direction. As competencies ultimately find expression in leadership behavior (Boyatzis, 2008; Erpenbeck et al., 2017), the LLM roots its leadership dimensions epistemologically in the actor-world relations framework inherent in any (leadership) behavior (Habermas, 1987). Accordingly, behavior can be self-referential (focused on the actor, "purpose"), social-communicative (focused on interactions, "responsibility"), or task-referential (focused on the object of an action, "effectiveness"). However, insights from motivational psychology (e.g. Heckhausen & Gollwitzer, 1987) expand this framework by identifying a fourth dimension in the actor-action relation ("entrepreneurial spirit"). Consequently, actions can differ not only depending on the action's target, but also according to the level of activation and energy toward action-implementation and the intensity of outcome-orientation. In other words: while the initial three relations encompass the different actor-world relations, the fourth relation pertains to the general underlying quality of relations an actor displays across the initial three. Any orientations for purpose, responsibility and effectiveness requires action-orientation to turn internal states of mind into overt behavior (Meynhardt et al., 2019). At the same time, a strong action-orientation does not necessarily imply equal strength in the other dimensions, as entrepreneurial spirit may lack a sense

of direction, for example. Therefore, this fourth dimension can be understood as both a stand-alone feature and as being instrumental in the first three dimensions. Multidimensional concepts integrating both specific aspects (here: purpose, responsibility and effectiveness) and general aspects (in this case entrepreneurial spirit) that partially influence the specific aspects, play prominent roles in various fields of psychology, ranging from clinical to organizational, personality and motivational theories (Marsh et al., 2014; Eid et al., 2018; Epstein, 1985; Heckhausen & Gollwitzer, 1987), including the literature on competencies (Erpenbeck et al., 2017).

Taken together, the Leipzig Leadership Model conceptualizes leadership orientations from an epistemological viewpoint of different actor relations. This allows for a justification of four leadership dimensions. The following paragraphs outline in more detail the core aspects of the LLM's four leadership dimensions as they relate to the underlying actor-world or actor-action relations.

Purpose — why and what for?

Kirchgeorg et al. (2017) define *purpose* as the intention of a person or organization to contribute to the common good, thereby creating public value. This line of thinking follows a long tradition of viewing organizations' purpose as transcending profit making (e.g. Barnard, 1938; Drucker, 1973). Most recently, Mayer (2021) elaborates by directing purpose towards "enhancing the wellbeing and prosperity of shareholders, society and the natural world" (p. 889). Similarly, Damon et al. (2003) conceptualize *purpose* as "a stable and generalizable intention to accomplish something that is at once meaningful to the self and leads to productive engagement with some aspect of the world beyond the self" (p. 121). On the micro level, these contributions have a profound impact on one's own identity (Epstein, 1985; Ryan & Deci, 2000), rendering purpose a highly motivating force (Hamel, 2009). Continuous reflection, by asking *why* and *what for*, unlocks and deepens leaders' understanding of themselves in terms of values and norms, enabling them to inspire their followers. As expectations directed at organizations to fulfil their social function increase, purpose-related questions of *why* and *what for* ultimately become questions of legitimacy and long-term survivability, to which leaders need to provide answers (Drucker, 1973; Collins & Porras, 1994). Researchers agree that uncovering and aligning oneself with one's purpose is key in the context of GCs, given the motivational, orientational, and legitimizing properties of such a purpose (Hamel, 2009; Kirchgeorg et al., 2017; van Tulder, 2018). Consequently, key purpose aspects, such as contributing and inspiring through visions, have been incorporated in various leadership conceptualisations,

such as transformational leadership (Avolio & Bass, 1995), shared leadership (Gockel & Werth, 2010; Pearce & Conger, 2010), or servant leadership (Liden et al., 2008).

Anchoring the LLM's leadership dimensions with purpose at its core reflects the importance of value-based judgments as orientational knowledge in the context of GCs. By guiding reflection on leadership competencies, this dimension highlights the self-referential importance of turning inwards as the source of lasting and reliable orientation in a complex, uncertain, and fast-changing environment. With purpose at its center, the LLM focuses leadership on its intended (public) value contributions (Kirchgeorg et al., 2017) and sets the questions of *why* and *what for* as the beginning of any (self-)reflection.

Entrepreneurial spirit – how?

Having raised the questions of *why* and *what for*, the LLM poses *how* as the subsequent guiding question. While an organization's purpose remains relatively constant (Collins & Porras, 1994; Damon et al., 2003), the ways in which to provide value contributions are subject to technological and societal changes, among others (Brammer et al., 2019; Selznick, 1975). The concept of *entrepreneurial spirit* captures this drive for continuous organizational renewal in the sense of identifying innovative ways of value contribution. Next to creativity, research has identified especially proactivity and the willingness to take risks as being at the core of entrepreneurial spirit (Morris & Jones, 1999). Representing the actor-action relation, this proactivity and drive capture the motivational activation of a leader required in the context of GC (Brammer et al., 2019). However, previous conceptualizations of e.g. entrepreneurial leadership (Gupta et al., 2004) solely focused on competitiveness and the tendency to challenge oneself (Gorostiaga et al., 2019). Entrepreneurial spirit expands this limited view, following the lead of e.g. empowering leadership (Ahearne et al., 2005) to shift the attention inwards, and encompasses a fundamental drive for individual learning and growth. Accordingly, this drive can promote not only leaders' own creativity, but extend to facilitate organizational creativity as "the process by which new ideas, objects or processes are introduced" (Csikszentmihalyi, 2014, p. 194). By shaping a supportive and empowering organizational environment, leaders can therefore ensure both individual creativity and their organization's capacity for renewal.

In the context of GCs, an organization's capability for renewal becomes essential to ensure lasting future value contributions. With concepts of creativity, risk affinity and proactivity at its core, the *entrepreneurial spirit* highlights actors' action-oriented attitudes and mindsets. Within the LLM framework, *entrepreneurial spirit* captures this attitude and the openness with which leaders

and followers alike face the uncertainty and complexity of their contexts, and thus has significant implications for all other actor-world relations as represented by the other LLM-dimensions.

Responsibility – how?

Following any narrative without continuously challenging its status quo creates room for distortion of previously benign motives. The LLM therefore asks *how*, not only in the sense of 'how to achieve future value contributions', but in the sense of how one's entrepreneurial pursuit of purpose might be restricted. Consequently, responsibility is defined as the respect and fulfillment of legitimate trust-based expectations (Kirchgeorg et al., 2017).

The implications of respect for organizational sciences have been researched extensively (e.g. Blader & Yu, 2017; Dunning et al., 2016), since it is part of most leadership conceptualizations (Rudolph et al., 2021). In the context of leadership, respect does not only extend to a leader's immediate followers, but also to external stakeholders, including the societies surrounding an organization (Maak & Pless, 2006). The LLM underscores that respect requires leaders to consider their decisions' justifiability and fairness (Ciulla, 2006). Simultaneously, trust has been widely recognized as one of the central factors that connect organizational relationships and their outcomes (Braun et al., 2013; Dirks & Ferrin, 2002). Trust is commonly defined as the willingness to be vulnerable to someone based on the perceived trustworthiness of that person. This trustworthiness can stem from two separate processes (Dirks & Ferrin, 2002): Cognitive trust is the result of evaluating a person in terms of competence, integrity, reliability, and dependability; while affective trust emerges over time as an emotional tie based on mutual exhibition of care and concern. As trust is reciprocal, a leader's ability to fulfill a role model function has been established as the most significant antecedent of trust (Avolio & Bass, 1995; Kirkpatrick & Locke, 1996). Consequently, the notion of leadership on the basis of providing an (ethical) role model has been incorporated by various conceptualizations, most prominently ethical leadership (Brown & Treviño, 2006) and responsible leadership (Maak & Pless, 2006).

With trust and respect highlighting the social contexts of leadership, *responsibility* encompasses the fundamental aspect of leadership as embodying role model behavior. When guiding leaders' reflections, this dimension turns their attention to the socio-communicative aspects of leadership, highlighting their interactions and social bonds with various stakeholders. It refers to the willingness and ability to question the extent to which leadership actions are justifiable to themselves and others, but also to the (social) environment and to nature itself.

Effectiveness – what?

"There is usually nothing quite so useless as doing with great efficiency what should not be done at all" (Drucker, 1963, p. 54). In much the same way that Drucker (1963) subordinated matters of *effectiveness* to the guiding properties of *purpose*, the LLM posits *what* (are we doing) after the initial questions of *why* and *how*. Consequently, asking *what* aims to translate purpose-based, responsible entrepreneurial strategies into tangible processes and solutions. In line with e.g. upper echelon theory and strategic leadership (e.g. Hambrick and Mason, 1984), *effectiveness* focuses on the task-referential aspects of leadership within the LLM framework. It identifies leadership behaviors that facilitate the successful implementation of concrete measures and processes. Upon reviewing decades of (effective) leadership research, Yukl and Gardner (2020) identified three central components of these task-referential behaviors: planning, clarification, and monitoring.

Planning describes a mostly cognitive activity that ensures efficiency through sensibly structuring and coordinating activities, and utilizing resources. Its importance to effective leadership has long been established and empirically proven (Boyatzis, 1982; Drucker, 1973). On the other hand, clarification encompasses defining job requirements, setting prioritized and clear goals, and assigning tasks in a sensible manner. Finally, monitoring involves having feedback processes to continuously evaluate implemented measures and realign them with their intended purpose.

The present studies

The goal of our studies is to develop and validate a 32-item LLM-based scale as a concise self-rating tool of leadership orientations, thereby advancing not only the previously solely conceptually developed LLM, but also the research on leadership and competencies, and the relevant strands of literature regarding the LLM's four dimensions. Furthermore, the scale developed in this study provides an applicable guideline for leaders' self-reflection in practice, constituting an initial step toward bridging prevalent theory-practice gaps regarding leadership competency models. Following the recommended process of scale development (e.g. Worthington & Whittaker, 2006) to generate an LLM-based item pool, we rely on two independent samples and proceed in eight steps: we conducted item development (1), confirmation of the factor structure (2), reliability assessment (3), convergent validity assessment (4), multidimensionality assessment (5), discriminant validity assessment (6), measurement invariance tests (7), and concurrent validity assessment within the

German sample (study 1); while on the basis of parallel back translation of the German scale into English, we replicated and tested the factor structure in a second English sample in study 2. We expect both scales to represent the multidimensional three-plus-one structure as postulated by the LLM (Kirchgeorg et al., 2017) in both samples.

Developing an initial item pool and content validity

In order to develop an initial item pool, we systematically reviewed the relevant literature on the LLM and its four dimensions, as well as existing scales showing conceptual proximity as elaborated on (and given examples for) in the theoretical background. Next, we captured as many statements as possible that could best describe the LLM's four dimensions. Together with an expert panel ($N = 23$), consisting of five professors (all of whom contributed to developing and advancing the model), three advanced leadership doctoral students, 10 executives and five business school students, all of whom were native German speakers, we then proceeded to formulate items (each expert brainstorming separately). This approach resulted in a total of 179 items, which we then presented to the whole panel to rate eachitem for its clarity and face and content validity (Moosbrugger & Kelava, 2020). This way, the expert panel selected 67 items to be included in study 1, following established standards for pre-test item generation and selection regarding the avoidance of e.g. ambiguity, double-barreled items, or negative wordings while also ensuring to extensively represent the underlying constructs (MacKenzie et al., 2011; Ford & Scandura, 2018; Moosbrugger & Kelava, 2020).

Study 1

In this study, we intended to establish and test a scale based on the 67 selected items that can represent the LLM's multidimensional three-plus-one structure. Historically, such tests would have involved a combination of exploratory (EFA) and confirmatory factor analyses (CFA) (Marsh et al., 2014). However, Asparouhov et al. (2015), among others, were able to show that exploratory structural equation modelling (ESEM) is able to overcome the main drawbacks of both methods, namely that EFA does not allow for the modelling of latent structures, while CFA tends to inflate factor correlations and parameter estimations due to cross-loadings of factors being constrained to zero (Eid et al., 2018; Marsh et al., 2014). To capture and model multidimensionality, researchers have increasingly been applying bifactor models (Reise, 2012; Eid et al., 2018) in various areas of psychological research. In bifactor models, each variance is split into a general factor, specific residual factors, and measurement error (Moosbrugger & Kelava, 2020; Eid, 2020).

In order to retain meaningfully interpretable g-factors, the so-called S-1 bifactor model was introduced (Heinrich et al., 2020). Instead of modelling a g-factor parallel to the existing theoretical dimensions, one of these existing dimensions is extended to be the g-factor rendering all other dimensions residual factors of g. Figure 1 shows a structural example of a S-1 bifactor ESEM model.

From a S-1 bifactor perspective, *entrepreneurial spirit* is considered the g-factor within the LLM framework, as it represents the "plus-one" aspect of the proposed multidimensional structure, referring to the quality of activation underlying the three other actor-world relations. ESEM is not only able to represent a more realistic representation of potential cross-loading of items and factors, it also allows for these a-priori theoretical considerations to guide initial exploratory rotations in the form of oblique bifactor target rotations (Marsh et al., 2014; Jennrich & Bentler, 2012). We selected the final subset of items based on the results of these rotations. Therefore, statistical analyses in study 1 applied a S-1 bifactor ESEM approach (Asparouhov et al., 2015; Marsh et al., 2014).

Method

Sample

A total of 431 German executives participated in our survey in exchange for financial remuneration, through the services of the independent German market research institute Respondi (respondi.com). To ensure comparability of the results, we excluded participants who worked in companies with fewer than five employees and with less than one year of leadership experience, resulting in a final sample of $N = 309$ German executives. The respondents were between 21 and 67 years old (M = 49.62; SD = 10.38). Approximately two fifths of the sample consisted of female executives (38.5%), and 64.7% had a college or higher education level. Almost half (45.60%) of the participants had a monthly household net income of more than 4,000 euros.

Measures and concurrent validity

In the online survey, participants evaluated their leadership behavior by rating the 67 provisory items of the LLM scale. Following the prevalent literature on the inclusion of midpoints in Likert scales (Nadler et al., 2015; Chyung et al., 2017) we decided on forced-choice format with six categories ranging from 1 (strongly disagree) to 6 (strongly agree), thereby omitting the "neutral" midpoint of our scale. This omission allows us to control both for the sometimes abstract nature especially of the items in our purpose-subscale, as well as the likely bias due to the social desirability of the items especially in the responsibility-subscale (see Tables 2 and 5).
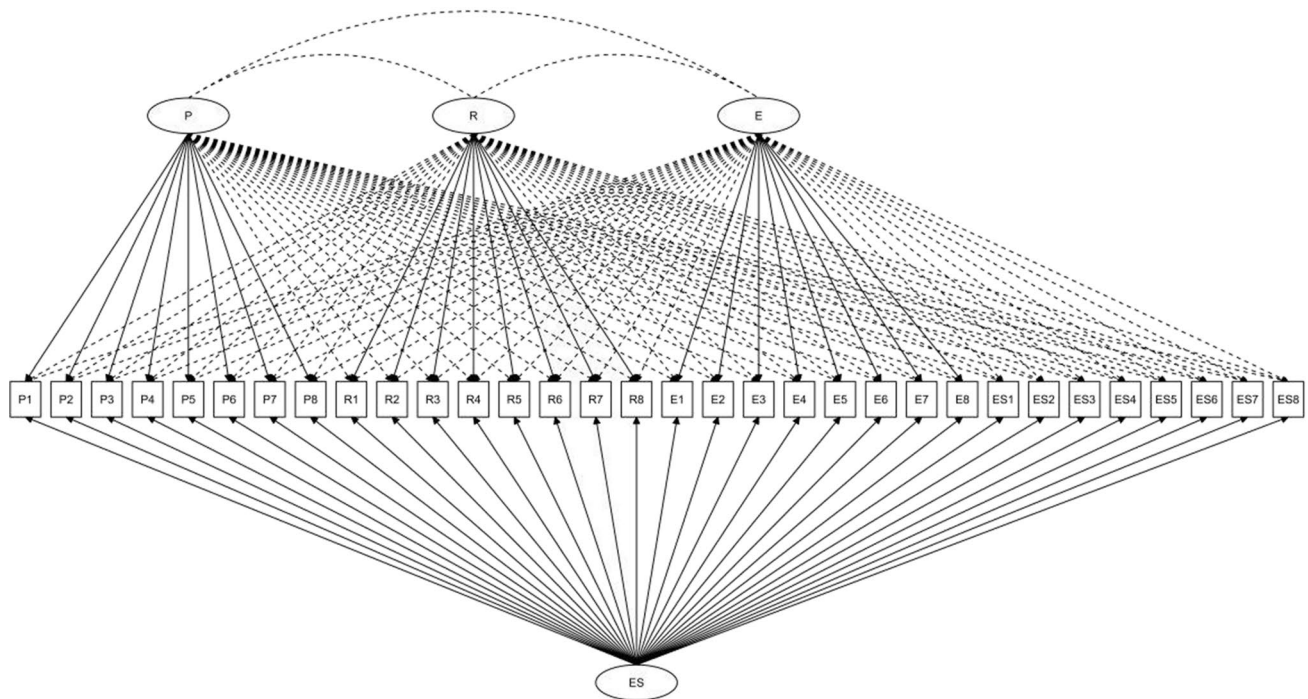
Fig. 1 Bifactor ESEM model of the LLM scale; P = purpose; R = responsibility; E = effectiveness; ES = entrepreneurial spirit; P1 to P8 = items for purpose; R1 to R8 = items for responsibility; E1 to E8 = items for effectiveness; ES1 to ES8 = items for entrepreneurial spirit. Full unidirectional arrows represent factor loadings; dotted unidirectional arrows represent cross-loadings; dotted curved lines represent correlations

In order to estimate concurrent validity of the LLM's subscales (Moosbrugger & Kelava, 2020), scales with conceptual proximity were included in the survey. The four LLM dimensions frame competencies, which ultimately find expression in a person's behavior, emotions and cognitions (Erpenbeck et al., 2017). Therefore, these competencies are closely related to that person's personality (Epstein, 1985). While personality traits may serve as a disposition for leadership orientations, the latter cannot be reduced to them. For example, openness to new experience can be a characteristic for entrepreneurial spirit. However, a felt sense to change the status quo may also be motivated by other emotional and cognitive processes, such as anxiety and frustration.

Consequently, we identified and chose three of the five Short Scales for the Big Five Dimensions of Personality (Lang, 2005) for this study, namely openness (three items per dimension, e.g. "I see myself as someone who is creative and comes up with new ideas"), which relates to the proactive and change-oriented dimension of entrepreneurial spirit (ES); agreeableness (three items per dimension, e.g. "I see myself as someone who is gentle and kind to others"), which is closely related to the social-communicative dimension of responsibility (R); and conscientiousness (three items per dimension, e.g. "I see myself as someone who works thoroughly"), which is closely related to the task-related dimension of effectiveness (E). Furthermore, as there is no

Big Five-related dimension reflective of the self-referential aspects embodied by purpose (P), the participants filled out the Short Servant Leadership Scale (Liden et al., 2015; seven items, e.g. "I emphasize the importance of giving back to the community"). Since the g-factor in (S-1) bifactor models binds shared variance of each residual factor (Eid et al., 2018), we would only expect strong correlations for the openness scale and entrepreneurial spirit. The remaining "residual correlations" are expected to show moderate to strong values (Rodriguez et al., 2015).

Analyses

We conducted analyses in version 4.2.2 of R (R Core Team, 2022), using packages dplyr (Wickham et al., 2022), lavaan (Rosseel, 2012), BifactorIndicesCalculator (Dueber, 2021), EFA.dimensions (O'Connor, 2023) and esemComp (Silvestrin & de Beer, 2023). In a first step, we extracted factor loadings using oblique bifactor target rotation. Apart from considerations of the theoretical saturation of the remaining item pool, we eliminated items if they showed communalities lower than 0.4, insubstantial factor loadings (< 0.3), misplaced factor loadings, and cross-loading differences smaller than 0.2 (Costello & Osborne, 2005; Brown, 2015). We proceeded in a stepwise fashion with new bifactor target rotations after each elimination, until a final scale with

eight items per factor and satisfactory loading patterns was achieved. We also conducted a parallel analysis to test the four-factorial structure for the first as well as the final rotation (Moosbrugger & Kelava, 2020). In a second step, an ESEM-model is estimated with the exploratory loadings as anchors for the estimation of model parameters (Marsh et al., 2014). This model is then compared against an oblique bifactor CFA solution to assess the impact of the constrained cross-loadings.

Model evaluation

Despite research indicating acceptable performance of robust continuous estimators, such as robust Maximum-Likelihood for scales with five or more anchors (Rhemtulla et al., 2012), all confirmatory analyses apply the robust WLSMV estimator controlling for the ordinal nature of Likert-scales (Moosbrugger & Kelava, 2020). Model fit was further assessed on the basis of the model fit indices proposed by Sun (2005), with the Comparative Fit Index (CFI) and Tucker-Lewis Index (TLI) above 0.90 indicating acceptable (and above 0.95 good) fit, similar to a Root-Mean-Square-Error of Approximation (RMSEA) below 0.05 and a Standardized Root-Mean-Square-Residual (SRMR) below 0.08. Furthermore, we follow the recommendations of Rodriguez et al. (2016) for evaluating bifactor models in a SEM context according to (1) McDonalds omega (ω) (1970) as a model-based reliability estimate similar to Cronbach's alpha; (2) omega hierarchical (ωH) as the amount of true variance explained by the g-factor; (3) omega hierarchical subscale (ωHS) as the amount of true variance explained by a residual factor; (4) construct replicability H (Hancock & Mueller, 2001) indicating "how well a latent variable is represented by a given set of items" (Rodriguez et al., 2016, p.143); and finally (5) explained common variance (ECV), percent uncontaminated correlations (PUC), and average relative parameter bias (ARPB) to assess the multidimensionality of the model. While cutoffs for the various ω's (especially ωHS) could not yet be established (Rodriguez et al., 2015), Hancock and Mueller (2001) require H $\geq$ 0.70 for well-represented factors, while multidimensionality is indicated by a combination of ωH $\leq$ 0.80, either ECV $\leq$ 0.70 or PUC $\leq$ 0.70 and an ARPB $\leq$ 0.15 (Rodriguez et al., 2016).

Model comparison and measurement invariance

Invariance of the final LLM scale was assessed across age and gender, following the common recommendations, according to Moosbrugger and Kelava (2020), of configural, metric, scalar and strict invariance. For all model comparisons, both robust χ^2-difference tests for nested models

(Satorra & Bentler, 2010) and ΔCFI-values are reported, with |ΔCFI| $\geq$ 0.01 between two models signifying a significant difference in model fit (Chen, 2007). In order to assess the measurement invariance across age, the sample is split along the median of the age distribution.

Convergent and discriminant validity

For the evaluation of convergent validity, primary factor loadings > 0.3 are considered significant (Moosbrugger & Kelava, 2020). Additionally, we follow Fornell and Larcker (1981) by assessing the amount of variance in latent variables explained by their respective factors. In bifactor models, this is achieved by utilizing ωHS instead of the average variance extracted (AVE; Moosbrugger and Kelava, 2020). In order to further evaluate discriminant validity, we compare ωHS for each factor with the maximum squared factor correlation of that factor (maximum shared variance, MSV).

Results

Exploratory item selection

Results of the exploratory oblique target rotation are displayed in Table 1. The final 32 items show substantial and theory-equivalent loadings across all four factors, as well as communalities > 0.40. Overall, the four factors explain 62% of total variance. The total Kaiser-Meyer-Olkin criterion (KMO) value is 0.95, indicating "marvelous" suitability of the sample (Kaiser, 1974), with the KMO for each item ranging between 0.92 and 0.97 as well. However, while a significant Bartlett's test (χ^2_{496}=7206.00; p $\leq$ .001) indicates the correlation matrix to be substantially different from an identity matrix, the absolute determinant R² of the matrix is $\leq$ 0,00001, suggesting multicollinearity between the items. Further analysis of the correlation matrix revealed 55 (out of 496) bivariate correlations between r = .60 and r = .80. To control for the complex potential impact of multicollinearity within SEM-frameworks (Marsh et al., 2004), we added further tests in subsequent analyses.

Confirmation of the factor structure, reliability and convergent validity

In a next step, we estimated the bifactor ESEM model parameters using the exploratory loadings as anchors (Marsh et al., 2014). Following established guidelines (Morin et al., 2016), we simultaneously estimated a bifactor CFA model, restraining cross loadings for the residual factors to zero (Fig. 2). The resulting factor loadings for both models are reported in Table 2, with substantial factor loadings without significant cross-loadings in combination with ωHS > 0.50 across all three residual factors suggesting convergent validity of the

Table 1 Factor loadings, communalities (h^2), Eigenvalues, % variance explained and multiple R^2 for each factor after initial oblique bifactor target rotation

		ES	P	R	E	h²
P1	Ich möchte, dass meine Arbeit der Gesellschaft nutzt.	**0.53**	**0.68**	0.04	0.07	0.76
P2	Ich möchte, dass meine Arbeit zu einer positiven Veränderung in der Welt beiträgt.	**0.58**	**0.63**	-0.06	0.02	0.72
P3	Ich richte mein Handeln am Gemeinwohl aus.	**0.56**	**0.70**	-0.03	0.07	0.80
P4	Ich verzichte auf eigene Vorteile, wenn es dem Gemeinwohl schadet.	**0.40**	**0.66**	0.08	-0.12	0.61
P5	Bei Konflikten berücksichtige ich das Gemeinwohl.	**0.48**	**0.68**	0.06	0.02	0.70
P6	Ich begeistere mich für Arbeiten mit gesellschaftlicher Relevanz.	**0.56**	**0.65**	-0.02	-0.08	0.74
P7	Ich bin stolz, wenn ich durch meine Arbeit einen Beitrag zur Gesellschaft leisten kann.	**0.56**	**0.59**	0.02	0.01	0.67
P8	Organisationen sollten sich primär durch ihre gesellschaftliche Aufgabe leiten lassen.	**0.46**	**0.64**	-0.13	0.05	0.61
R1	Auf mich kann man sich verlassen.	**0.39**	-0.01	**0.51**	0.19	0.57
R2	Ich halte Vereinbarungen ein.	**0.40**	-0.10	**0.61**	0.17	0.68
R3	Ich bin aufrichtig.	**0.49**	0.08	**0.60**	-0.06	0.57
R4	Ich verhalte mich fair.	**0.52**	0.06	**0.67**	-0.05	0.70
R5	Ich verhalte mich respektvoll.	**0.51**	0.08	**0.59**	0.00	0.62
R6	Ich bin mir der Vorbildrolle von Führungskräften bewusst.	**0.48**	-0.03	**0.53**	0.10	0.58
R7	Ich bin bereit, die Folgen meiner Handlungen zu tragen.	**0.51**	-0.06	**0.62**	-0.09	0.59
R8	Ich kann eigene Fehler eingestehen.	**0.48**	-0.04	**0.55**	0.06	0.57
E1	Ich arbeite zielorientiert.	**0.51**	-0.02	0.11	**0.56**	0.66
E2	Ich verschaffe mir einen Überblick, bevor ich entscheide.	**0.46**	-0.02	0.14	**0.56**	0.65
E3	Ich wäge die Vor- und Nachteile einer Lösung ab.	**0.41**	-0.05	0.13	**0.49**	0.50
E4	Ich vergleiche das gewünschte Ergebnis mit dem erforderlichen Aufwand.	**0.45**	0.07	-0.02	**0.48**	0.42
E5	Ich plane voraus.	**0.42**	-0.03	0.01	**0.57**	0.51
E6	Ich unterscheide zwischen Wichtigem und Unwichtigem.	**0.39**	0.01	-0.01	**0.66**	0.59
E7	Ich setze Prioritäten und halte mich daran.	**0.47**	0.05	0.04	**0.63**	0.64
E8	Ich strukturiere Arbeitsprozesse sinnvoll.	**0.49**	0.02	-0.14	**0.66**	0.58
ES1	Ich möchte Neues auf den Weg bringen.	**0.84**	0.02	-0.04	-0.04	0.71
ES2	Ich kann mich für neue Ideen begeistern.	**0.84**	-0.05	0.07	-0.08	0.72
ES3	Ich löse Probleme auf kreative Art.	**0.71**	-0.06	-0.08	-0.03	0.52
ES4	Ich ermutige Mitarbeitende, Neues auszuprobieren.	**0.72**	0.05	0.10	-0.04	0.53
ES5	Ich sehe Wandel als etwas Positives.	**0.68**	0.09	0.11	-0.09	0.48
ES6	Ich sprühe vor Ideen.	**0.69**	0.08	0.08	0.03	0.49
ES7	Mir fällt es leicht, Gewohntes loszulassen.	**0.74**	-0.10	-0.17	0.01	0.59
ES8	Ich suche Herausforderungen.	**0.78**	-0.03	-0.17	0.05	0.64
	Eigenvalues	13.27	4.37	2.12	1.45	
	% of variance (cummulative)	32	43	52	62	
	Mult. R²	0.94	0.89	0.87	0.86	

Factor loadings in bold indicate significant loadings at $p \leq .001$

32 items of the LLM scale. Both the bifactor ESEM model (RMSEA = 0.01; SRMR = 0.03; CFI = 0.99; TLI = 0.99) and the bifactor CFA model (RMSEA = 0.02; SRMR = 0.04; CFI = 0.99; TLI = 0.99) indicated good model fit, with composite reliability reaching high levels for the bifactor CFA solution ($\omega = 0.97$) (McDonald, 1970). However, a comparison of the two models revealed that, by restraining the cross-loadings to zero in the bifactor CFA solution, model fit did not decrease significantly ($\Delta \chi^2_{\Delta df=60} = 3.59$; $p = 1$; ΔCFI=-0.00). Therefore, the impact of restricted cross-loadings in our sample is deemed negligible, and in the name of parsimony the bifactor CFA model is preferred for the LLM-scale (Moosbrugger & Kelava, 2020; Morin et al., 2016).

Finally, we re-assessed the content validity of the items using a confirmatory Q-sort method. We presented the items to 41 MBA students after a thematic introduction to the LLM and its four dimensions, and asked them to match the items to the definitions. For all items, the hit ratio was above 95%, indicating a high degree of face validity (Nahm et al., 2002).

Table 2 Factor loadings (sd), factor correlations, model-based reliability estimates, explained common variance (ECV), percentage of uncontaminated correlations (PUC), construct replicability (H), factor determinacy (FD) and average relative parameter bias (ARPB) of the bifactor CFA model in study 1

		ES	P	R	E
P1	Ich möchte, dass meine Arbeit der Gesellschaft nutzt.	0.76 (0.07)	0.81 (0.08)		
P2	Ich möchte, dass meine Arbeit zu einer positiven Veränderung in der Welt beiträgt.	0.77 (0.08)	0.85 (0.08)		
P3	Ich richte mein Handeln am Gemeinwohl aus.	0.75 (0.07)	0.82 (0.07)		
P4	Ich verzichte auf eigene Vorteile, wenn es dem Gemeinwohl schadet.	0.51 (0.09)	0.88 (0.07)		
P5	Bei Konflikten berücksichtige ich das Gemeinwohl.	0.66 (0.08)	0.81 (0.09)		
P6	Ich begeistere mich für Arbeiten mit gesellschaftlicher Relevanz.	0.69 (0.08)	0.88 (0.07)		
P7	Ich bin stolz, wenn ich durch meine Arbeit einen Beitrag zur Gesellschaft leisten kann.	0.78 (0.08)	0.77 (0.09)		
P8	Organisationen sollten sich primär durch ihre gesellschaftliche Aufgabe leiten lassen.	0.59 (0.08)	0.82 (0.08)		
R1	Auf mich kann man sich verlassen.	0.29 (0.05)		0.53 (0.06)	
R2	Ich halte Vereinbarungen ein.	0.28 (0.07)		0.66 (0.06)	
R3	Ich bin aufrichtig.	0.47 (0.06)		0.45 (0.07)	
R4	Ich verhalte mich fair.	0.45 (0.06)		0.50 (0.05)	
R5	Ich verhalte mich respektvoll.	0.45 (0.06)		0.48 (0.05)	
R6	Ich bin mir der Vorbildrolle von Führungskräften bewusst.	0.38 (0.07)		0.53 (0.06)	
R7	Ich bin bereit, die Folgen meiner Handlungen zu tragen.	0.41 (0.06)		0.45 (0.05)	
R8	Ich kann eigene Fehler eingestehen.	0.43 (0.07)		0.60 (0.06)	
E1	Ich arbeite zielorientiert.	0.39 (0.05)			0.53 (0.05)
E2	Ich verschaffe mir einen Überblick, bevor ich entscheide.	0.40 (0.06)			0.63 (0.05)
E3	Ich wäge die Vor- und Nachteile einer Lösung ab.	0.33 (0.06)			0.55 (0.05)
E4	Ich vergleiche das gewünschte Ergebnis mit dem erforderlichen Aufwand.	0.44 (0.07)			0.43 (0.07)
E5	Ich plane voraus.	0.34 (0.06)			0.52 (0.05)
E6	Ich unterscheide zwischen Wichtigem und Unwichtigem.	0.33 (0.06)			0.57 (0.06)
E7	Ich setze Prioritäten und halte mich daran.	0.40 (0.05)			0.55 (0.06)
E8	Ich strukturiere Arbeitsprozesse sinnvoll.	0.42 (0.05)			0.46 (0.06)
ES1	Ich möchte Neues auf den Weg bringen.	0.88 (0.05)			
ES2	Ich kann mich für neue Ideen begeistern.	0.83 (0.05)			
ES3	Ich löse Probleme auf kreative Art.	0.71 (0.06)			
ES4	Ich ermutige Mitarbeitende, Neues auszuprobieren.	0.78 (0.07)			
ES5	Ich sehe Wandel als etwas Positives.	0.73 (0.05)			
ES6	Ich sprühe vor Ideen.	0.80 (0.06)			
ES7	Mir fällt es leicht, Gewohntes loszulassen.	0.76 (0.06)			
ES8	Ich suche Herausforderungen.	0.90 (0.06)			
		ES	P	R	E
Factor correlations: ES		0.97[a]			
P		-[b]	0.95[a]		
R		-	0.06	0.92[a]	
E		-	-0.03	0.68[***]	0.91[a]
Bifactor statistical indices: ω_{total}		0.97			
$\omega H/ \omega HS$		0.77	0.56	0.59	0.60
H		0.95	0.85	0.85	0.83
ECV		0.52			
PUC		0.83			
ARPB		0.23			

All factor loadings are significant at $\rho \leq 0.001$; [a] the diagonal values are replaced with the internal reliability ω of the factor; [b] the g-factor (ES) is defined to be orthogonal to the residual factors; [***] $\rho \leq 0.001$
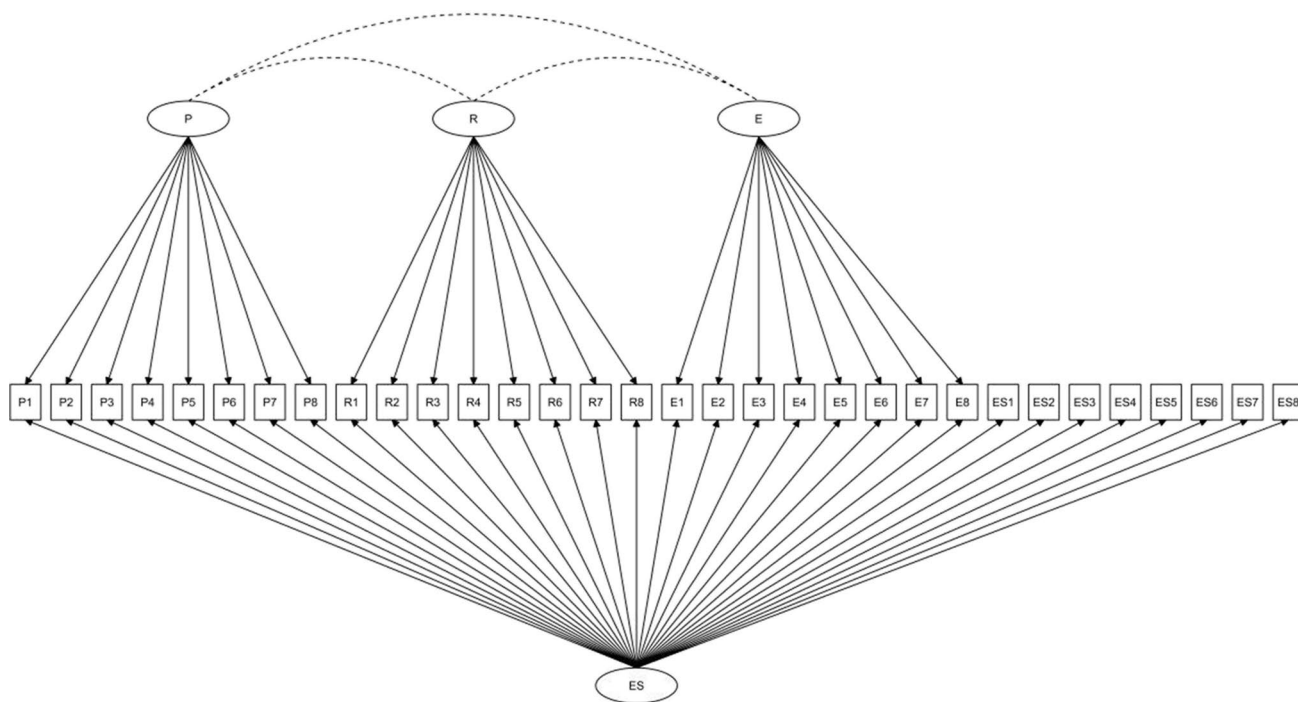
Fig. 2 Bifactor CFA model of the LLM scale; P = purpose; R = responsibility; E = effectiveness; ES = entrepreneurial spirit; P1 to P8 = items for purpose; R1 to R8 = items for responsibility; E1 to E8 = items for effectiveness; ES1 to ES8 = items for entrepreneurial spirit. Full unidirectional arrows represent factor loadings; dotted curved lines represent correlations

Multidimensionality of the LLM scale

Overall, 77% of true variance is explained by the g-factor of entrepreneurial spirit (ωH ES). In combination with a rather low ECV = 0.52, the data support the multidimensionality of the scale, despite the high PUC = 0.83. An ARPB = 0.23 signifies that, by assuming unidimensionality, the relative bias in estimating parameters would amount to 23%. To further test the model against the assumption of unidimensionality, a one-factor solution was modeled (RMSEA = 0.11; SRMR = 0.14; CFI = 0.90; TLI = 0.89), but provided a significantly worse model fit ($\Delta \chi^2_{27}$=69.22; $p \leq .001$; ΔCFI=-0.10). To further test the robustness of our multidimensional bifactor CFA model, a correlated factor model was fit to the data and provided a surprisingly good fit as well (RMSEA = 0.03; SRMR = 0.05; CFI = 0.97; TLI = 0.97), with the reduction in model fit not being significant ($\Delta \chi^2_{21}$=7.03; $p = .99$; ΔCFI=-0.00). However, an analysis of the parameter estimates revealed that, due to the excessive factor correlations between responsibility (R) and effectiveness (E) (r = .782), their MSV = 0.61 exceeded their respective AVE_R=0.55 and AVE_E=0.56. This provides further support of the superiority of bifactor modelling in complex psychological constructs (Eid et al., 2018). To ensure sufficient discriminant validity of the factors, the bifactor CFA was retained, despite having fewer degrees of freedom.

Discriminant validity and multicollinearity

As seen in Table 2, all residual factors in the bifactor CFA model explain more than 50% of their residual variance and are well-defined by their items, as indicated by H_{ES}=0.95, H_P=0.85, H_R=0.85, and H_E=0.83. Furthermore, as only R and E correlate significantly (r = .68, $p \leq .001$), the resulting MSV = 0.46 is lower than their respective ωHS_R=0.59 and ωHS_E=0.60, indicating sufficient discriminant validity. While the g-factor is per definition orthogonal to the residual factors in bifactor modelling, as it is allowed to bind all shared variance across the items (Eid et al., 2018), residual factors can be allowed to correlate with each other, as was done by applying the oblique bifactor target rotation. However, as Marsh et al. (2004) show, this pattern of one strong factor correlation in combination with other non-significant correlations can be the result of multicollinearity in SEM. To control for this effect, we followed Marsh et al. (2004), by fitting a bifactor CFA model that restricted all factor correlations to be equal, while another model was estimated with orthogonal residual factors. The oblique solution with equal correlations (RMSEA = 0.06; SRMR = 0.08; CFI = 0.97; TLI = 0.97) provided significantly worse model fit compared to the bifactor CFA model ($\Delta \chi^2_2$=233.71; $p \leq .001$; ΔCFI=-0.03), as did the orthogonal solution ($\Delta \chi^2_3$=171.75;

Table 3 Robust fit statistics of the bifactor ESEM- and CFA models as well as invariance tests across gender and age for the final bifactor CFA solution of the LLM scale (study 1)

	χ^2	df	CFI	TLI	RMSEA [90% CI]	SRMR	$\Delta \chi^{2a}$	ΔCFI[a]
Full-ESEM	413.790	377	0.999	0.998	0.012 [0.000–0.018]	0.027	-	-
Bifactor CFA (oblique)	514.103	437	0.996	0.996	0.021 [0.012–0.028]	0.040	3.589	−0.003
Further Comparisons								
Bifactor CFA (oblique)	514.103	437	0.996	0.996	0.021 [0.012–0.028]	0.040	-	-
Bifactor CFA (orthogonal)	815.826	440	0.972	0.969	0.056 [0.050–0.062]	0.079	171.747[***b]	−0.027
One-factor CFA	1540.104	464	0.896	0.889	0.105 [0.099–0.111]	0.139	69.221[***]	−0.100
Correlated factors	564.776	458	0.993	0.993	0.027 [0.018–0.034]	0.049	7.031	−0.003
Bifactor CFA (oblique – equal factor correlations)	845.561	439	0.973	0.969	0.055 [0.050–0.061]	0.079	233.708[***]	−0.026
Invariance: Gender								
Configural	922.396	874	0.997	0.996	0.019 [0.000–0.031]	0.048	-	-
Metric	974.953	926	0.994	0.994	0.025 [0.000–0.040]	0.063	15.566	−0.003
Scalar	1002.690	954	0.994	0.994	0.024 [0.000–0.039]	0.063	1.586	−0.000
Strict	1031.191	986	0.994	0.995	0.023 [0.000–0.039]	0.065	2.968	−0.000
Invariance: Age								
Configural	931.259	874	0.996	0.996	0.021 [0.000–0.032]	0.050	-	-
Metric	975.648	926	0.993	0.993	0.028 [0.000–0.044]	0.066	22.095	−0.003
Scalar	1005.756	954	0.993	0.993	0.028 [0.000–0.044]	0.067	1.910	−0.000
Strict	1040.594	986	0.992	0.992	0.028 [0.000–0.044]	0.070	3.268	−0.001

[***] $p \leq 0.001$; [a] for *Further Comparisons*, $\Delta \chi^2$ and ΔCFI are stated in relation to the bifactor CFA (oblique) model; for *invariance* tests, each level of invariance is tested against the previous level. [b] reported here is the corrected $\Delta \chi^2$ (Moosbrugger & Kelava, 2020) to achieve χ^2-distributed $\Delta \chi^2$-values after robust χ^2 estimation

$p \leq .001$; ΔCFI=-0.03), despite its overall acceptable fit (RMSEA = 0.06; SRMR = 0.08; CFI = 0.97; TLI = 0.97). According to Marsh et al. (2004), this indicates that (a) the residual factors cannot be considered orthogonal, and (b) that the correlations between the residual factors are not equal but differ significantly from each other. This suggests that the impact of multicollinearity in our sample is rather small, which would be in line with simulation studies by Grewal et al. (2004), who found that in samples with a large multiple correlation of R²≥0.75 (see Table 1), a large sample-to-item ratio of at least 6:1, and a high reliability

of at least 0.80 (as is fulfilled in our sample), the effects of multicollinearity between r = .60 and r = .80 (as found in exploratory analyses) are negligible.

Measurement invariance tests

As can be seen in Table 3, our sample supported strict measurement invariance across both age and gender. For age, the sample was split along the median of 51 years into two groups to be compared. Table 3 further provides an overview of all model comparisons reported in study 1.

Table 4 Correlation matrix of regression-based factor scores, including (sub)scale specific reliability (study 1)

	1	2	3	4	5	6	7	8
1. Entrepreneurial Spirit	0.97[1]							
2. Purpose	0.08	0.95[1]						
3. Responsibility	0.01	0.01	0.92[1]					
4. Effectiveness	0.04	0.00	0.70	0.91[1]				
5. Servant Leader	0.70[***]	0.46[***]	0.11	0.09	0.79[2]			
6. Agreeableness	0.46[***]	0.21[***]	0.32[***]	0.25[***]	0.55[***]	0.66[2]		
7. Conscientiousness	0.41[***]	-0.01	0.43[***]	0.50[***]	0.35[***]	0.44[***]	0.65[2]	
8. Openness	0.66[***]	0.04	-0.06	-0.02	0.53[***]	0.36[***]	0.35[***]	0.77[2]

[***]$\rho \leq .001$; [*] $\rho \leq .01$; [*] $\rho \leq .05$; Note: [1] for multidimensional subscales, ω is provided, [2] for unidimensional subscales, Cronbach's α is provided

Concurrent validity

As the construct replicability H for all subscales is above 0.70, factor scores can be calculated using the regression method (Rodriguez et al., 2016). In order to assess our hypotheses regarding the concurrent validity of our subscales, Table 4 provides the spearman correlation matrix of the respective factor scores. The overall medium to strong correlations of the Big Five and servant leadership scales with ES supports the finding from other bifactor ESEM-research (Rodriguez et al., 2015), namely that the g-factor in bifactor models seems to bind substantial variance and reduces correlations of residual factors with external criterions. Further in line with our hypotheses, the openness-score only correlates significantly with the ES-score ($r = .66$, $p \leq .001$). Furthermore, within the residual factors, servant leadership-scores correlate only significantly with P-scores ($r = .46$, $p \leq .001$), agreeableness-scores correlate most strongly with R-scores ($r = .32$, $p \leq .001$), and conscientiousness-scores correlate most strongly with E-scores ($r = .50$), a value which even exceeds its correlation with ES-scores ($r = .46$, $p \leq .001$). Both the correlations of agreeableness-scores and conscientiousness-scores are significantly larger than their next highest correlation at $p \leq .05$.

Study 2

In preparation for our second study, we translated the LLM scale into English using the parallel back-translation procedure (Brislin, 1986). Two bilingual English-speaking professional linguists translated the 32-item LLM scale into English, after which a bilingual German-speaking professional linguist retranslated it back into German to control for any major deviations from the original item pool. A bilingual panel with in-depth knowledge of the research constructs in question reviewed the results to perform minor adjustments and create a consolidated and localized English translation of the scale. The 32-item English LLM scale was then used in a second sample to assess whether we would be able to replicate the bifactor CFA model using the English version.

Method

We recruited British employees via the online recruiting platform Prolific (prolific.co; see Peer et al., 2017 for an evaluation of the platform). After applying the same criteria as in study 1, the final sample consisted of 311 British employees with supervisory responsibilities, 81.0% of whom are currently in a leadership role (11.2% in top management). The respondents were between 18 and 71 years old ($M = 32.52$; $SD = 10.30$). Approximately two fifths of the sample consisted of female respondents (38.2%), and 70.6% had a college education or a higher qualification.

In order to assess the replicability of the bifactor CFA model fit in the English sample, participants evaluated their leadership behavior by rating the 32 translated items of the English LLM scale on a six-point Likert scale ranging from 1 (strongly disagree) to 6 (strongly agree). Additionally, as in the first study, the participants completed the Short Servant Leadership Scale (Liden et al., 2015). Because our first study revealed weak internal consistencies ($\alpha < 0.77$) of the scales suggested by Lang (2005), we alternatively applied the more comprehensive Big Five Inventory (John & Srivastava, 1999) to assess conscientiousness (nine items per dimension, e.g. "I see myself as someone who works thoroughly"), agreeableness (nine items per dimension, e.g. "I see myself as someone who is generally trusting"), and openness to experience (ten items per dimension, e.g. "I see myself as someone who is creative and comes up with new ideas"). In a next step, the bifactor CFA was fit to the data and evaluated according to the criteria formulated by Sun (2005) and Rodriguez et al. (2016), as well as convergent and discriminant validity. Finally, measurement invariance tests were conducted in the same manner as in study 1. Analyses were conducted in version 4.2.2 of R (R Core Team, 2022), using packages dplyr (Wickham et al., 2022), lavaan (Rosseel, 2012) and BifactorIndicesCalculator (Dueber, 2021).

Results

After data collection, a bifactor CFA model was fit to the data of the English sample. Results are presented in Table 5. The results replicate the good fit achieved in the German sample (RMSEA = 0.03; SRMR = 0.05; CFI = 0.99; TLI = 0.99). Item loadings reach substantial level (Moosbrugger & Kelava, 2020), and composite reliability of the scale remains high at $\omega = 0.93$. Overall, 72% of variance is explained by the ES-factor, which, in combination with the low ECV = 0.48 and high ARPB = 0.43, again suggests the multidimensionality of the scale, despite a rather high PUC = 0.83. Moreover, H reaches acceptable levels ≥ 0.73 for all factors, indicating good representation of the factors by the items, and allowing for factor scores to be calculated. As Table 6 shows, strict measurement invariance was replicated across both gender and age (split at the sample's median of 29 years) with the models for strict invariance across age (RMSEA = 0.02; SRMR = 0.07; CFI = 0.99; TLI = 0.99) and gender (RMSEA = 0.03; SRMR = 0.07; CFI = 0.99; TLI = 0.99) showing good model fit. With substantial factor loadings (> 0.30) and a $\omega HS \geq 0.50$ across all factors indicating good convergent validity, discriminant validity following Fornell and Larcker (1981) is sufficient given the largest factor correlation of $r_{RE} = 0.60$, resulting in a MSV = 0.36. However, unlike the German sample, both the factor correlation between P and R ($r = .36$, $p \leq .001$)

Table 5 Factor loadings (sd), factor correlations, model-based reliability estimates, explained common variance (ECV), percentage of uncontaminated correlations (PUC), construct replicability (H), factor determinacy (FD) and average relative parameter bias (ARPB) of the bifactor CFA model in study 2

		ES	P	R	E
P1	I want my work to benefit society.	0.46 (0.07)	0.75 (0.07)		
P2	I want my work to contribute to changing the world for the better.	0.49 (0.06)	0.70 (0.07)		
P3	I focus what I do on the common good.	0.42 (0.06)	0.58 (0.07)		
P4	I check whether my decisions are in line with social expectations.	0.35 (0.07)	0.52 (0.07)		
P5	I forgo personal gain if it harms the common good.	0.42 (0.07)	0.43 (0.08)		
P6	I am passionate about work that has a social relevance.	0.51 (0.06)	0.62 (0.07)		
P7	I am proud if I can make a contribution to society with my work.	0.41 (0.06)	0.67 (0.06)		
P8	I think that organisations should be guided mainly by their social function.	0.29 (0.07)	0.49 (0.07)		
R1	I can be relied upon.	0.35 (0.06)		0.58 (0.05)	
R2	I honor agreements.	0.27 (0.04)		0.45 (0.04)	
R3	I am sincere.	0.37 (0.06)		0.50 (0.05)	
R4	I behave fairly.	0.32 (0.06)		0.54 (0.05)	
R5	I behave respectfully.	0.26 (0.05)		0.54 (0.04)	
R6	I am mindful that leaders have a role model function.	0.30 (0.05)		0.51 (0.05)	
R7	I am ready to accept the consequences of my actions.	0.37 (0.05)		0.47 (0.05)	
R8	I can admit to mistakes.	0.38 (0.06)		0.52 (0.06)	
E1	I am goal-oriented in my work.	0.48 (0.06)			0.52 (0.06)
E2	I get an overview before I make decisions.	0.28 (0.06)			0.62 (0.05)
E3	I weigh the pros and cons of a solution.	0.31 (0.07)			0.55 (0.06)
E4	I compare the intended result with the effort needed to achieve it.	0.43 (0.07)			0.48 (0.06)
E5	I plan ahead.	0.27 (0.08)			0.43 (0.08)
E6	I differentiate between what is important and what is not.	0.32 (0.06)			0.49 (0.05)
E7	I set priorities and stick to them.	0.39 (0.06)			0.46 (0.06)
E8	I structure work processes intelligently.	0.40 (0.05)			0.33 (0.05)
ES1	I want to help new things get off the ground.	0.66 (0.05)			
ES2	I am enthusiastic about new ideas.	0.77 (0.05)			
ES3	I solve problems creatively.	0.60 (0.05)			
ES4	I encourage employees to try new things.	0.71 (0.05)			
ES5	I see change as positive.	0.68 (0.06)			
ES6	I am a fountain of ideas.	0.54 (0.06)			
ES7	I find it easy to let go of routines.	0.38 (0.08)			
ES8	I look for challenges.	0.81 (0.06)			
		ES	P	R	E
	Factor correlations: ES	0.93[a]			
	P	-[b]	0.87[a]		
	R	-	0.36***	0.89[a]	
	E	-	0.40***	0.60***	0.83[a]
	Bifactor statistical indices: ω_{Total}	0.93			
	ωH/ ωHS	0.72	0.58	0.64	0.54
	H	0.92	0.79	0.83	0.74
	ECV	0.48			
	PUC	0.83			
	ARPB	0.43			

All factor loadings are significant at $\rho \leq 0.001$; [a] the diagonal values are replaced with the internal reliability ω of the factor; [b] the g-factor (ES) is defined to be orthogonal to the residual factors; *** $\rho \leq 0.001$

Table 6 Invariance tests across gender and age of the English version of the bifactor CFA model (study 2)

	χ^2	df	CFI	TLI	RMSEA	SRMR	$\Delta \chi^{2*}$	ΔCFI
Gender								
Configural	953.595	874	0.991	0.990	0.029	0.058	-	-
Metric	995.586	926	0.988	0.987	0.032	0.069	13.801	−0.003
Scalar	1027.981	954	0.987	0.987	0.032	0.070	2.292	−0.000
Strict	1060.957	986	0.987	0.987	0.032	0.074	4.240	−0.000
Age								
Configural	941.503	874	0.992	0.991	0.026	0.057	-	-
Metric	952.139	926	0.995	0.995	0.019	0.064	11.508	−0.003
Scalar	981.953	954	0.995	0.995	0.019	0.065	2.140	−0.000
Strict	1021.454	986	0.994	0.994	0.022	0.070	4.435	−0.001

[***] $p \leq 0.001$; [*]reported here is the corrected $\Delta \chi^2$ (Moosbrugger & Kelava, 2020) to achieve χ^2-distributed $\Delta \chi^2$-values after robust χ^2 estimation

and P and E (r = .40, $p \leq .001$) were significant. Concurrent Validity was estimated to be sufficient, based on the strong correlation between openness and ES (r = .55, $p \leq .001$), and the medium correlations between servant leadership and P (r = .38, $p \leq .001$), agreeableness and R (r = .33, $p \leq .001$), and conscientiousness and E (r = .43, $p \leq .001$), which where all in line with our hypotheses.

Discussion

As GCs continuously evolve alongside societal and technological progress, among other factors, leadership is required to regularly revise and re-identify relevant competencies in order to adapt to its contexts (Brammer et al., 2019; Selznick, 1975). Meynhardt et al. (2019) argue that approaching leadership competencies through the lens of leadership orientations allows leaders to reflect continuously and holistically on relevant leadership behaviors, even as their context and the concomitant requirements of leadership change continuously. Integrating actor-world (Habermas, 1987) and actor-action relations (Heckhausen & Gollwitzer, 1987), the LLM structures leadership orientations along the dimensions of purpose, entrepreneurial spirit, responsibility, and effectiveness.

With the present studies, we contribute by developing and validating a scale able to represent the complex multidimensional structure postulated by the LLM. This study not only continues the trend of recent research that reiterates the superior modelling capabilities of (S-1) bifactor ESEM in complex psychological contexts (Eid et al., 2018; Heinrich et al., 2020), but as far as we know it also provides the first applications of this methodology to the field of leadership research, after it was initially applied to non-clinical contexts by e.g. Howard et al. (2016) or Litalien et al. (2017). Through our study's transdisciplinary approach, we contribute not only to the research on leadership and competencies, but also the relevant strands of literature regarding the

LLM's dimensions. Finally, the scale developed in this study provides an applicable guideline for leadership self-reflection in practice, constituting an initial step toward bridging the prevalent theory-practice gap regarding leadership competency models.

Theoretical contributions

Our study is the first to develop and test a scale representing the three-plus-one-factor structure of the leadership orientation dimensions postulated by the LLM (Kirchgeorg et al., 2017). Drawing from a transdiciplinarily developed conceptualization of a holistic leadership role, leadership orientations are understood as the interpersonal tendency to place varying emphasis on the actor-world or actor-action relations inherent in that leadership role. While historically, the term 'leadership orientations' mostly referred to the intrinsic preference of leadership styles (e.g. dominant/friendly, social-oriented/result-oriented) (Kuehl et al., 1975), a few recent studies have adapted a role-oriented view of leadership orientations (e.g. Bergman et al., 2014). Despite its somewhat neglected "niche" status within leadership research, we argue, in line with Meynhardt et al. (2019), that this approach holds substantial potential for adressing the wide-spread criticism of leadership literature as being characterized by what Antonakis (2017) calls "theorrhea" – a plethora of highly reductionist leadership theories and theoretically unfounded "best practices". Building on meta-analytic results suggesting significant redundancy between these various theories (Hoch et al., 2018; Banks et al., 2016), calls for integration and conceptualization of holistic leadership roles continue to increase (Yukl & Gardner, 2020). Furthermore, while these studies find e.g. ethical leadership to not explain further variance over transformational leadership, the latter still faces considerable criticism in its lacking acknowledgement of ethical aspects of leadership and the resulting potential

for abuse of (charismatic and inspirational) power (Price, 2003). By building on the fundamental actor-world and actor-action relations underlying all (leadership) behavior, the LLM contributes by offering a new approach to the conceptualization of leadership orientations that balances theoretically exhaustive dimensions of leadership, thus providing a holistic and concise understanding of the leadership role. The LLM as a framework integrates various theories, to identify common denominators across the barrage of individual leadership styles and best practices, and not to get lost in the process by neglecting either the why, how, or what of leadership.

Furthermore, through the comparison of the correlated factor and bifactor CFA model in study 1, we find additional support for previous research arguing for the superior modelling qualities of (S-1) bifactor models, both in CFA and ESEM approaches, (Asparouhov et al., 2015; Morin et al., 2016; Marsh et al., 2014). As our data show, restrained factor loadings in the correlated factor model led to inflation of the existing factor correlations and, despite good model fit, undermined the model's discriminant validity. In line with e.g. Eid et al. (2018) and Heinrich et al. (2020), we promote the use of these more sophisticated statistical approaches for modelling complex psychological constructs. While previous works by e.g. Howard et al. (2016) and Litalien et al. (2017) expanded the application of (S-1) bifactor models to non-clinical settings, our study further expands on their work by applying it to leadership research.

Based on our results, we were able to validate the 32-item LLM scale for both the German and English sample. Comparing the bifactor ESEM and the bifactor CFA model suggests that the impact of the additional restrained factor loadings was negligible, which was why the bifactor CFA model was applied for subsequent analyses. Furthermore, Fornell and Larcker's (1981) criteria for convergent and discriminant validity could be met for both samples. However, at this point the results in both studies deviate in one noteworthy aspect from each other: While in the German sample, only the correlation between responsibility and entrepreneurial spirit turned significant, all residual factors correlated significantly in the English sample. This latter pattern, in fact, is closer to our a priori expectations, which is why we applied oblique bifactor target rotation to our exploratory analyses. To control for e.g. reported impacts of multicollinearity (e.g. Marsh et al., 2004), we tested the bifactor CFA model against orthogonal assumptions and a bifactor CFA model with equal factor correlations. Yet, based on model fit comparisons we concluded that (a) the oblique bifactor CFA model with free estimates of factor correlations fit the data best, and (b) the impact of multicollinearity, especially in the light of findings from e.g. Grewal et al. (2004), could be deemed insignificant. Yet, the precise relationship between the residual factors within the LLM will need further study.

At this point, we can only conclude that the strong correlation between responsibility and effectiveness is theoretically plausible, given their respective theoretical underpinnings. As elaborated during the introduction of the dimensions, trust as a vital part of responsibility emerges in part due to a leader being perceived as reliable and consistent, for example (Dirks & Ferrin, 2002; Braun et al., 2013). At the same time, leadership effectiveness can be conceptualized along the components identified by Yukl and Gardner (2020) of e.g. planning and monitoring. It seems logical that leaders who are adept at planning and monitoring would be perceived by others as being more consistent and reliable. From relational identity perspectives (Epstein, 1985) we would further assume that these perceptions are communicated by others to the leader and internalized in their own self-conceptualization, which we measured with our self-rating tool. This hypothesis, however, will need further evaluation in multi-trait multi-method designs to cross-validate self and expert ratings of the LLM scale.

Our study's results suggest that the LLM scale's sub-dimensions can function as standalone scales. This has important implications, especially for the *purpose* and *entrepreneurial spirit* subscales. So far, empirical studies on purposeful leadership have been scarce, due to a lack of measurement tools. The few existing studies in the field have tried to overcome the definitional issue by applying measures of related yet theoretically different concepts, such as perceived organizational purpose (Jasinenko & Steuber, 2022), meaningfulness of work (Gartenberg et al., 2019), common good-oriented job characteristics (Allan et al., 2018), or workplace spirituality (Kolodinsky et al., 2008). By applying the purpose subscale, we aim to encourage future research to further develop this rather novel field of research. Similarly, despite being partially addressed in prior scales, the aspect of self-development has only recently been included in measurement tools relating to entrepreneurial orientation in an educational setting with students as the target group (Gorostiaga et al., 2019). Our subscale therefore transfers the notion of self-development as a fundamental need (Epstein, 1985) and a part of entrepreneurial orientation/spirit to measurement tools in the leadership context.

Implications for practice

In light of non-linear system dynamics with complex social and technical solutions which GCs pose to leadership, value judgements become indispensable (Eisenhardt et al., 2016; Brammer et al., 2019). Additionally, GCs are continuously evolving along societal and technical progress. Consequently, leadership development has undergone a paradigm-shift from episodic programs (e.g. workshops) to focusing on self-directed learning (Nesbit, 2012). Building on critical (self-)reflection leaders not only develop value-based

decision-making, but are further able to continuously evaluate and adapt their own behavior. As self-reflection ("inner work") not only provides orientation, but is shown to e.g. increase a leader's psychological capital (Luthans et al., 2008), regular practice not only decreases e.g. leader's stress but has further positive spill-over effects on individual and team job performance and health (Walumbwa et al., 2010; Branson, 2007).

The LLM translates its four dimensions of leadership orientations into applicable guiding questions. This makes its dimensions tangible to lay people, and enables them to measure their own leadership behavior and reflect on it. It incorporates the paradigm shift in both leadership and developmental psychology research signified by transformational learning by anchoring its holistic conceptualization of leadership in the self-referential *purpose* dimension. With the help of the newly developed scale, it becomes possible to identify one's own dispositions of over- or underemphasizing any leadership orientation dimensions as a first step to balancing one's leadership profile.

We are currently developing an app (Leipzig Leadership Profile) to enable quick and simple self-assessment and effect a comparison through external assessment in order to facilitate applicability and feedback. Both assessments will have detailed reports and practical implications. By using the LLM scale, leaders can, for example, identify the discrepancies between self-assessment and employees' perceptions. Organizations can use the questionnaire for recruiting, leadership coaching, and performance appraisal or feedback. Given its nature of providing four stable leadership orientations, the LLM scale will be of advantage if applied to reflect on company-specific competency models.

Limitations and future research avenues

While applying a cross-sectional design might be beneficial because it allows us to generate large samples, this approach has limitations that should be considered when the results are interpreted. For example, Podsakoff et al. (2003) highlight the possibility of a systematic error variance in the form of a common method bias. However, we followed their recommendations to randomize the order in which the scales are presented. Consequently, mitigation of this limitation can be assumed.

Additionally, to reduce the impact of social desirability biases, we assured participants that there were no right or wrong answers regarding their estimates and that their data would be processed anonymously. Additionally, we followed the recommendations of e.g. Chyung et al. (2017) to utilize a forced-choice format for our Likert scale. Given the underlying controversial literature, we encourage further studies to test the effects of the inclusion of midpoints in the LLM scale. Finally, our use of an asymmetric six-point scale could be seen as controversial, as it is sometimes associated with

measurement errors due to a missing middle point (Weijters et al., 2010). However, recent studies find beneficial effects of asymmetric Likert scales when participants are unfamiliar with the rated subject (Nadler et al., 2015) or to reduce social desirability biases (e.g. Chyung et al., 2017). Our study provides the first empirical results on the distinctiveness of the model's dimensions compared to the adjacent servant leadership, conscientiousness, agreeableness, and openness to experience, although we need further theoretical and empirical research to understand the constructs' distinctiveness in more depth.

It is important to note that the LLM scale's item list consists of self-rating items. In their fairly recent meta-analysis, Lee and Carpenter (2018) found that employees tend to rate a leader lower compared to the leader's self-rating, and that leaders tend to rate themselves too highly, especially on items regarding what they call "relation-oriented leadership dimensions" (p. 3). While we could validate the LLM as a tool for leadership research, self-assessment and self-reflection, a modified item list is needed in order to allow the combination of self-ratings and observer ratings. By marginally adapting the items, the questionnaire may also be used for external assessment, for example by employees. Future studies should, however, examine the validity of the adapted items. The specific epistemological logic (three plus one dimensions) did prove empirically appropriate in our studies. However, it requires further investigation regarding interaction effects, and also, more fundamentally, regarding the consequences for intervention.

Given the LLM's novel transdisciplinary conceptualization of holistic leadership, various fields for future research open up. First, studies are needed to expand and solidify the discriminant and convergent validity of the LLM scale. We call for and invite researchers to contribute to the understanding of leadership orientations, also in relation to outcomes, moderators, and mediators. The LLM's adequateness with regard to the continued call from leadership research for integrative models also needs to be examined further. We agree with those who claim that increasing the accessibility of research by filtering and integrating the existing plethora of leadership styles is perhaps the most important step in reducing, if not overcoming, the existing theory-practice gaps. Our proposal is not to add a new distinction (e.g. transformational vs. transactional, initiating structure vs. consideration, or the like), but rather to help structure different leadership dimensions within from an action perspective.

From a broader perspective, the LLM and its scale are an invitation to study the relationship between individual leadership styles and organizational leadership, for example by distinguishing between individual and organizational purpose. The LLM provides a common language to articulate common problems on different levels of analysis. The authors hope that this may motivate the exploration of the

next frontier of leadership research: from an ever more fine-grained differentiation leadership phenomenon to a next level of synthesis based on a common-sense perspective of the why, the what and the how of leadership.

Data availability The datasets and item lists generated and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee, and also with ESOMAR code of conduct. (https://www.esomar.org/uploads/public/knowledge-and-standards/codes-and-guidelines/ICCESOMAR_Code_English_.pdf). Our research is not a medical study conducted with the participation of patients; therefore, the 1964 Helsinki declaration and its later amendments are not applicable here.

Consent for publication No further consent for publication is warranted.

Informed consent This was a fully anonymous online study, and the participants were informed about it and agreed to participate by clicking on the link in the invitation e-mail. We did not collect any other form of informed consent.

Competing interests The authors have no competing interests to declare that are relevant to the content of this article.

Conflict of interest On behalf of all authors, the corresponding author states that there is no further conflict of interest.

References

Ahearne, M., Mathieu, J., & Rapp, A. (2005). To empower or not to empower your sales force? An empirical examination of the influence of Leadership empowerment behavior on customer satisfaction and performance. *Journal of Applied Psychology, 90*(5), 945–955. https://doi.org/10.1037/0021-9010.90.5.945

Allan, B. A., Duffy, R. D., & Collisson, B. (2018). Task significance and performance. *Journal of Career Assessment, 26*(1), 172–182. https://doi.org/10.1177/1069072716680047

Antonakis, J. (2017). On doing better science: from thrill of discovery to policy implications. *The Leadership Quarterly, 28*(1), 5–21. https://doi.org/10.1016/j.leaqua.2017.01.006

Asparouhov, T., Muthén, B., & Morin, A. J. S. (2015). Bayesian structural equation modeling with cross-loadings and residual covariances: comments on Stromeyer et al. *Journal of Management, 41*(6), 1561–1577. https://doi.org/10.1177/0149206315591075

Avolio, B. J., & Bass, B. M. (1995). Individual consideration viewed at multiple levels of analysis: a multi-level framework of transformational leadership. *The Leadership Quarterly, 6*(2), 199–218. https://doi.org/10.1016/1048-9843(95)90035-7

Banks, G. C., McCauley, K. D., Garnder, W. L., & Guler, C. E. (2016). A meta-analytic review of authentic and transformational leadership: a test for redundancy. *The Leadership Quarterly, 27*(4), 634–652. https://doi.org/10.1016/j.leaqua.2016.02.006

Barnard, C. I. (1938). *The functions of the executive*. Harvard University Press.

Bergman, D., Lornudd, C., Sjöberg, L., & Von Schwarz, T. (2014). Leader personality and 360-degree assessments of leader behavior. *Scandinavian Journal of Psychology, 55*(4), 389–397. https://doi.org/10.1111/sjop.12130

Blader, S. L., & Yu, S. (2017). Are status and respect different or two sides of the same coin? *Academy of Management Annals, 11*(2), 800–824.

Bolden, R., & Gosling, J. (2006). Leadership competencies: time to change the tune? *Leadership, 2*(2), 147–163. https://doi.org/10.1177/1742715006062932

Boyatzis, R. E. (1982). *The competent manager*. John Wiley.

Boyatzis, R. E. (2008). Competencies in the 21st century. *Journal of Management Development, 27*(1), 5–12. https://doi.org/10.1108/02621710810840730

Brammer, S., Branicki, L., Linnenluecke, M., & Smith, T. (2019). Grand challenges in management research: attributes, achievements, and advancement. *Australian Journal of Management, 44*(4), 517–533. https://doi.org/10.1177/0312896219871337

Branson, C. M. (2007). Improving leadership by nurturing moral consciousness through structured self-reflection. *Journal of Educational Administration, 45*(4), 471–495. https://doi.org/10.1108/09578230710762463

Braun, S., Peus, C., Weisweiler, S., & Frey, D. (2013). Transformational leadership, job satisfaction, and team performance: a multilevel mediation model of trust. *The Leadership Quarterly, 24*(1), 270–283. https://doi.org/10.1016/j.leaqua.2012.11.006

Brislin, R. W. (1986). The wording and translation of research instruments. In W. J. Lonner, & J. W. Berry (Eds.), *Field methods in cross-cultural research* (pp. 137–164). Sage Publications, Inc.

Brown, M. E., & Treviño, L. K. (2006). Ethical leadership: a review and future directions. *The Leadership Quarterly, 17*(6), 595–616. https://doi.org/10.1016/j.leaqua.2006.10.004

Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). Guilford Press.

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of Measurement Invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 14*(3), 464–504. https://doi.org/10.1080/10705510701301834

Chyung, S. Y. Y., Roberts, K., Swanson, I., & Hankinson, A. (2017). Evidence-based survey design: the use of a midpoint on the Likert scale. *Performance Improvement, 56*(10), 15–23. https://doi.org/10.1002/pfi.21727

Ciulla, J. B. (2006). Ethics: The heart of leadership. In T. Maak, & N. M. Pless (Eds.), *Responsible leadership in business* (pp. 37–52). Routledge.

Collins, J. C., & Porras, J. I. (1994). *Built to last: Successful habits of visionary companies* (1st ed.). HarperBusiness.

Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. *Practical Assessment, 10*(7), 1–9. https://doi.org/10.7275/jyj1-4868

Csikszentmihalyi, M. (2014). The systems model of creativity. *Springer Netherlands*. https://doi.org/10.1007/978-94-017-9085-7

Damon, W., Menon, J., & Bronk, K. C. (2003). The development of purpose during adolescence. *Applied Developmental Science, 7*(3), 119–128.

Dirks, K. T., & Ferrin, D. L. (2002). Trust in leadership: meta-analytic findings and implications for research and practice. *Journal of Applied Psychology, 87*(4), 611–628. https://doi.org/10.1037//0021-9010.87.4.611

Drucker, P. F. (1973). *Management: tasks, responsibilities, practices.* Harper colophon books: CN/1207. HarperBusiness.

Drucker, P. F. (1963). Managing for business effectiveness. *Harvard Business Review, 41*(3), 53–60.

Dueber, D. (2021). *BifactorIndicesCalculator: Bifactor Indices Calculator.* R package version 0.2.2. Retrieved July 1, 2023, from https://CRAN.R-project.org/package=BifactorIndicesCalculator

Dunning, D., Fetchenhauer, D., & Schlösser, T. (2016). Chapter one The psychology of respect: a case study of how behavioral norms regulate human action. In A. J. Elliot (Ed.), *Advances in motivation science* Academic Press. https://doi.org/10.1016/bs.adms.2015.12.003

Eid, M. (2020). Multi-faceted constructs in abnormal psychology: implications of the Bifactor S – 1 model for individual clinical assessment. *Journal of Abnormal Child Psychology, 48*, 895–900. https://doi.org/10.1007/s10802-020-00624-9

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

Eisenhardt, K. M., Graebner, M. E., & Sonenshein, S. (2016). Grand challenges and inductive methods: rigor without rigor mortis. *Academy of Management Journal, 59*(4), 1113–1123. https://doi.org/10.5465/amj.2016.4004

Epstein, S. (1985). The implications of cognitive-experiential self-theory for research in social psychology and personality. *Journal for the Theory of Social Behaviour, 15*(3), 283–310. https://doi.org/10.1111/j.1468-5914.1985.tb00057.x

Erpenbeck, J., Grote, S., & Sauter, W. (2017). Einführung. In J. Erpenbeck, von L. Rosenstiel, S. Grote, & W. Sauter (Eds.), *Handbuch Kompetenzmessung: Erkennen, verstehen und bewerten von Kompetenzen in der betrieblichen, pädagogischen und psychologischen Praxis* (3rd ed., pp. IX–XXXVIII). Schäffer-Poeschel.

Ford, L. R., & Scandura, T. A. (2018). A typology of threats to construct validity in item generation. *American Journal of Management, 18*(2), 132–142. https://doi.org/10.33423/ajm.v18i2.298

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*(1), 39–50. https://doi.org/10.1177/002224378101800104

Gartenberg, C. M., Prat, A., & Serafeim, G. (2019). Corporate purpose and financial performance. *Organization Science, 30*(1), 1–18.

Gockel, C., & Werth, L. (2010). Measuring and modeling shared leadership. *Journal of Personnel Psychology, 9*(4), 172–180. https://doi.org/10.1027/1866-5888/a000023

Gorostiaga, A., Aliri, J., Ulacia, I., Soroa, G., Balluerka, N., Aritzeta, A., & Muela, A. (2019). Assessment of entrepreneurial orientation in vocational training students: development of a new scale and relationships with self-efficacy and personal initiative. *Frontiers in Psychology, 10,.* https://doi.org/10.3389/fpsyg.2019.01125

Grewal, R., Cote, J. A., & Baumgartner, H. (2004). Multicollinearity and measurement error in structural equation models: Implications for theory testing. *Marketing Science, 23*(4), 519–529. https://doi.org/10.1287/mksc.1040.0070

Gupta, V., MacMillan, I. C., & Surie, G. (2004). Entrepreneurial leadership: Developing and measuring a cross-cultural construct. *Journal of Business Venturing, 19*(2), 241–260. https://doi.org/10.1016/S0883-9026(03)00040-5

Habermas, J. (1987). *Theory of communicative action, Vol. 2: Lifeworld and System: a critique of functionalist reason*. Beacon Press.

Hambrick, D. C., & Mason, P. A. (1984). Upper echelons: the organization as a reflection of its top managers. *Academy of Management Review, 9*(2), 193–206. https://doi.org/10.5465/amr.1984.4277628

Hamel, G. (2009). Moon shots for management. *Harvard Business Review, 87*(2), 91–98.

Hancock, G. R., & Mueller, R. O. (2001). Rethinking construct reliability within latent variable systems. In R. Cudeck, du S. Toit, & D. Sorbom (Eds.), *Structural equation modeling: Present and future—A festschrift in honor of Karl Joreskog* (pp. 195–216). Scientific Software International.

Heckhausen, H., & Gollwitzer, P. M. (1987). Thought contents and cognitive functioning in motivational versus volitional states of mind. *Motivation and Emotion, 11*(2), 101–120. https://doi.org/10.1007/BF00992338

Heifetz, R. A. (2006). Anchoring leadership in the work of adaptive progress. In F. Hesselbein, & M. Goldsmith (Eds.), *The leader of the future 2: Visions, strategies, and practices for the new era (Drucker Foundation Future Series)* (pp. 73–84). Jossey-Bass.

Heinrich, M., Zagorscak, P., Eid, M., & Knaevelsrud, C. (2020). Giving G a meaning: an application of the bifactor-(S-1) approach to realize a more symptom-oriented modeling of the Beck depression inventory–II. *Assessment, 27*(7), 1429–1447. https://doi.org/10.1177/1073191118803738

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

Hollenbeck, G. P., McCall, M. W., & Silzer, R. F. (2006). Leadership competency models. *The Leadership Quarterly, 17*(4), 398–413. https://doi.org/10.1016/j.leaqua.2006.04.003

Howard, J. L., Gagné, M., Morin, A. J. S., & Forest, J. (2016). Using bifactor exploratorystructural equation modeling to test for a continuum structure of motivation. *Journal of Management, 44*(7), 2638–2664. https://doi.org/10.1177/0149206316645653

Jasinenko, A., & Steuber, J. (2022). Perceived organizational purpose: systematic literature review, construct definition, measurement and potential employee outcomes. *Journal of Management Studies*. https://doi.org/10.1111/joms.12852. (conditionally accepted).

Jennrich, R. I., & Bentler, P. M. (2012). Exploratory Bi-factor analysis: the Oblique Case. *Psychometrika, 77*(3), 442–454. https://doi.org/10.1007/s11336-012-9269-1

John, O. P., & Srivastava, S. (1999). The big five trait taxonomy: history, measurement, and theoretical perspectives. In O. P. John, & L. A. Pervin (Eds.), *Handbook of personality: theory and research* (pp. 102–138). Guilford Press.

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika, 39*(1), 31–36. https://doi.org/10.1007/BF02291575

Kegan, R. (2009). What "form" transforms? A constructive-developmental approach to transformative learning. In K. Illeris (Ed.), *Contemporary theories of learning* (2nd ed., pp. 35–52). Routledge.

Kirchgeorg, M., Meynhardt, T., Pinkwart, A., Suchanek, A., & Zülch, H. (2017). *Das Leipziger Führungsmodell: Führung neu denken* (2nd ed.). HHL Academic Press.

Kirkpatrick, S. A., & Locke, E. A. (1996). Direct and indirect effects of three core charismatic leadership components on performance and attitudes. *Journal of Applied Psychology, 81*(1), 36–51. https://doi.org/10.1037/0021-9010.81.1.36

Kolodinsky, R. W., Giacalone, R. A., & Jurkiewicz, C. L. (2008). Workplace values and outcomes: exploring personal, organizational, and interactive workplace spirituality. *Journal of Business Ethics, 81*(2), 465–480. https://doi.org/10.1007/s10551-007-9507-0

Kuehl, C. R., DiMarco, N., & Wims, E. W. (1975). Leadership orientation as a function of interpersonal need structure. *Journal of Applied Psychology, 60*(1), 143–145. https://doi.org/10.1037/h0076361

Lang, F. R. (2005). Zur Reliabilität und Validität zweier ultrakurzer tests und des BFI-S. *DWI Research Notes*, (9), 1–32. Retrieved July 1, 2023, from https://www.diw.de/de/diw_01.c.449734.de/publikationen/research_notes/2005_0009/erfassung_des_kogni tiven_leistungspotenzials_und_der__big_fi___litaet_und_valid itaet_zweier_ultrakurzer_tests_und_des_bfi-s.html

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

Liden, R. C., Wayne, S. J., Meuser, J. D., Hu, J., Wu, J., & Liao, C. (2015). Servant leadership: validation of a short form of the SL-28. *The Leadership Quarterly, 26*(2), 254–269. https://doi.org/10.1016/j.leaqua.2014.12.002

Liden, R. C., Wayne, S. J., Zhao, H., & Henderson, D. (2008). Servant leadership: development of a multidimensional measure and multi-level assessment. *The Leadership Quarterly, 19*(2), 161–177. https://doi.org/10.1016/j.leaqua.2008.01.006

Litalien, D., Morin, A. J. S., Gagné, M., Vallerand, R. J., Losier, G. F., & Ryan, R. M. (2017). Evidence of a continuum structure of academic self-determination: A two study test using bifactor ESEM representation of academic motivation. *Contemporary Educational Psychology, 51*, 67–82. https://doi.org/10.1016/j.cedpsych.2017.06.010

Luthans, F., Avey, J. B., & Patera, J. L. (2008). Experimental analysis of a web-based training intervention to develop positive psychological capital. *Academy of Management Learning and Education, 7*(2), 209–221. https://doi.org/10.5465/amle.2008.32712618

Maak, T., & Pless, N. M. (2006). Responsible leadership in a stakeholder society – a relational perspective. *Journal of Business Ethics, 66*(1), 99–115. https://doi.org/10.1007/s10551-006-9047-z

MacKenzie, S. B., Podsakoff, P. M., & Podsakoff, N. P. (2011). Construct measurement and validation procedures in MIS and behavioral research: integrating new and existing techniques. *MIS Quarterly, 35*(2), 293–334. https://doi.org/10.2307/23044045

Marsh, H. W., Dowson, M., Pietsch, J., & Walker, R. (2004). Why Multicollinearity matters: a reexamination of relations between self-efficacy, self-concept, and achievement. *Journal of Educational Psychology, 96*(3), 518–522. https://doi.org/10.1037/0022-0663.96.3.518

Marsh, H. W., Morin, A. J. S., Parker, P. D., & Kaur, G. (2014). Exploratory structural equation modeling: an integration of the best features of exploratory and confirmatory factor analysis. *Annual Review of Clinical Psychology, 10*(1), 85–110. https://doi.org/10.1146/annurev-clinpsy-032813-153700

Mayer, C. (2021). The future of the corporation and the economics of purpose. *Journal of Management Studies, 58*(3), 887–901. https://doi.org/10.1111/joms.12660

McDonald, R. P. (1970). The theoretical foundations of principal factor analysis, canonical factor analysis, and alpha factor analysis.

British Journal of Mathematical and Statistical Psychology, 23, 1–21. https://doi.org/10.1111/j.2044-8317.1970.tb00432.x

Meynhardt, T. (2004). *Wertwissen: was organisation wirklich bewegt. Internationale Hochschulschriften: Vol. 432*. Waxmann.

Meynhardt, T., Kirchgeorg, M., Pinkwart, A., Suchanek, A., & Zülch, H. (Eds.). (2022). *Führen in der Krise: Herausforderungen an das Leipziger Führungsmodell*. HHL Academic Press.

Meynhardt, T., Kirchgeorg, M., Suchanek, A., & Zülch, H. (2019). Powerful or powerless? Beyond power and powerlessness: the leipzig leadership model provides some answers. *Leadership Education Personality: An Interdisciplinary Journal, 1*(1–2), 29–33. https://doi.org/10.1365/s42681-020-00007-0

Moosbrugger, H., & Kelava, A. (Eds.). (2020). *Testtheorie und Fragebogenkonstruktion* (3rd ed). Springer.

Morin, A. J. S., Arens, A. K., & Marsh, H. W. (2016). A bifactor exploratory structural equation modeling framework for the identification of distinct sources of construct-relevant psychometric multidimensionality. *Structural Equation Modeling: A Multidisciplinary Journal, 23*(1), 116–139. https://doi.org/10.1080/10705511.2014.961800

Morris, M. H., & Jones, F. F. (1999). Entrepreneurship in established organizations: the case of the public sector. *Entrepreneurship Theory and Practice, 24*(1), 71–91. https://doi.org/10.1177/104225879902400105

Nadler, J. T., Weston, R., & Voyles, E. C. (2015). Stuck in the middle: the use and interpretation of mid-points in items on questionnaires. *Journal of General Psychology, 142*(2), 71–89. https://doi.org/10.1080/00221309.2014.994590

Nahm, A. Y., Rao, S. S., Solis-Galvan, L. E., & Ragu-Nathan, T. S. (2002). The Q-Sort method: assessing reliability and construct validity of questionnaire items at a pre-testing stage. *Journal of Modern Applied Statistical Methods, 1*(1), 114–125. https://doi.org/10.22237/jmasm/1020255360. Article jmasm.eP.

Nesbit, P. L. (2012). The role of self-reflection, emotional management of feedback, and self-regulation processes in self-directed leadership development. *Human Resource Development Review, 11*(2), 203–226. https://doi.org/10.1177/1534484312439196

O'Connor, B. P. (2023). *EFA.dimensions: Exploratory factor analysis functions for assessing dimensionality*. R package version 0.1.7.6. Retrieved July 1, 2023, from https://CRAN.R-project.org/package=EFA.dimensions

Pearce, C. L., & Conger, J. A. (2010). *Shared leadership: Reframing the hows and whys of leadership*. SAGE Publications.

Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017). Beyond the Turk: alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology, 70*, 153–163. https://doi.org/10.1016/j.jesp.2017.01.006

Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: a critical review of the literature and recommended remedies. *Journal of Applied Psychology, 88*(5), 879–903. https://doi.org/10.1037/0021-9010.88.5.879

Price, T. L. (2003). The ethics of authentic transformational leadership. *The Leadership Quarterly, 14*(1), 67–81. https://doi.org/10.1016/S1048-9843(02)00187-X

R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Retrieved July 1, 2023, from https://www.R-project.org/

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

Rhemtulla, M., Brosseau-Liard, P., & Savalei, V. (2012). When can categorical variables be treated as continuous? A comparison of robust continuous and categorical SEM estimation methods

under suboptimal conditions. *Psychological Methods, 17*(3), 354–373. https://doi.org/10.1037/a0029315

Rodriguez, A., Reise, S. P., & Haviland, M. G. (2015). Applying bifactor statistical indices in the evaluation of psychological measures. *Journal of Personality Assessment, 98*(3), 223–237. https://doi.org/10.1080/00223891.2015.1089249

Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016). Evaluating bifactor models: calculating and interpreting statistical indices. *Psychological Methods, 21*(2), 137–150. https://doi.org/10.1037/met0000045

Rosseel, Y. (2012). Iavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1–36. Retrieved July 1, 2023, from http://www.jstatsoft.org/v48/i02/

Rudolph, C. W., Katz, I. M., Ruppel, R., & Zacher, H. (2021). A systematic and critical review of research on respect in leadership. *Leadership Quarterly, 32*(1). https://doi.org/10.1016/j.leaqua.2020.101492

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*(1), 68–78. https://doi.org/10.1037/0003-066X.55.1.68

Satorra, A., & Bentler, P. M. (2010). Ensuring positiveness of the scaled difference chi-square test statistic. *Psychometrika, 75*(2), 243–248. https://doi.org/10.1007/s11336-009-9135-y

Selznick, P. (1975). *Leadership in administration: A sociological interpretation*. Quid Pro Books.

Silvestrin, M. T., & de Beer, L. (2023). *esemComp: ESEM-within-CFA syntax composer*. R package version 0.2. Retrieved July 1, 2023, from https://mateuspsi.github.io/esemComp

Sun, J. (2005). Assessing goodness of fit in confirmatory factor analysis. *Measurement and Evaluation in Counseling and Development, 37*(4), 240–256. https://doi.org/10.1080/07481756.2005.11909764

van Tulder, R. (2018). *Business & the sustainable development goals: A framework for effective corporate involvement*. RSM Positive Change Series. Retrieved July 1, 2023, from http://hdl.handle.net/1765/110689

Walumbwa, F. O., Peterson, S. J., Avolio, B. J., & Hartnell, C. A. (2010). An investigation of the relationships among leader and follower psychological capital, service climate, and job performance. *Personnel Psychology, 63*(4), 937–963. https://doi.org/10.1111/j.1744-6570.2010.01193.x

Weijters, B., Cabooter, E., & Schillewaert, N. (2010). The effect of rating scale format on response styles: The number of response categories and response category labels. *International Journal of Research in Marketing, 27*(3), 236–247. https://doi.org/10.1016/j.ijresmar.2010.02.004

Wickham, H., François, R., Henry, L., & Müller, K. (2022). *dplyr: a grammar of data manipulation*. R package version 1.0.10. Retrieved July 1, 2023, from https://CRAN.R-project.org/package=dplyr

Worthington, R. L., & Whittaker, T. A. (2006). Scale development research. *The Counseling Psychologist, 34*(6), 806–838. https://doi.org/10.1177/0011000006288127

Yukl, G. A., & Gardner, W. L. (2020). W. L. *Leadership in organizations* (9th ed). Pearson.