Conceptualizing and Measuring Transformational and Transactional Leadership

Abstract

Existing conceptualizations and measures of transformational and transactional leadership have unclear theoretical bases, confound leadership and its effects, and are not necessarily suitable for public organizations. Overcoming these problems is necessary to test how leadership affects performance. Many public administration scholars apply the concepts, emphasizing the need to ensure that the concepts are applicable in both public and private organizations. The article re-conceptualizes transformational and transactional leadership and develops and tests revised measures that can be employed on employees and leaders, are robust in terms of repeated use by the same respondents, and applicable to public and private organizations alike.

Keywords

Transformational leadership; transactional leadership; measurement validity; public organizations; private organizations.

Introduction

Public administration research often suggests that improving leadership in the public sector is a key to increasing organizational performance (Moynihan, Pandey, & Wright, 2012; Rainey, 2014; Van Wart, 2013). Boyne (2003) finds that managerial variables are a stronger source of performance improvement than resources, regulation, market structure, and organization, a finding that is also supported by other studies (Fernandez, 2005; Hassan & Hatmaker, 2015; Moynihan & Ingraham, 2004; Moynihan et al., 2012; Parry & Sinha, 2005; Trottier, Van Wart, & Wang, 2008; Van Wart, 2013). However, in order to comprehend the full potential of leadership in public organizations, we need to identify relevant leadership strategies for this sector.

There are several indications that transformational and transactional leadership are relevant. First, transformational leadership is the most researched leadership theory in both generic leadership literature (Judge & Piccolo, 2004) and public administration research (Vogel & Masal, 2015). This raises the question whether this great interest also reflects best research practice. Second, the concepts of transformational and transactional leadership strategies have in multiple studies been related to employee well-being and performance (Lowe, Kroeck & Sivasubramaniam, 1996). Recent public administration studies (e.g., Bellé, 2014; Andersen & Pallesen, 2008) have confirmed that these strategies can indeed increase goal attainment in public organizations, but the extent to which the strategies affect many other relevant outcomes remains unexplored. Third, the proposition that these two leadership strategies should be more effective in private organizations has been challenged (e.g., Wright & Pandey, 2010; Wright, Moynihan, & Pandey, 2012, p. 207). This makes it particularly relevant to revise the conceptualizations and measures of these two leadership strategies with the purpose of applying them in future empirical research in both sectors.

However, generic leadership research has subjected the leadership strategies to a fundamental critique concerning conceptual and methodological problems with the generally applied "full-range leadership theory" and its MLQ measure (Van Knippenberg & Sitkin, 2013). In order to argue for and test these leadership strategies in public administration research and practice, we need to address these fundamental problems by revising their conceptualization and operationalization. While we respond to the problematic issues raised in the generic leadership literature in relation to the current conceptualization and operationalization of transformational and transactional leadership, we revise the concepts and measures to fit both public and private organizations, allowing for future comparisons between organizations with different degrees of publicness: Do the levels of transformational and transactional leadership differ between public and private organizations, are the effects of the leadership strategies the same, and do the same factors affect the use of leadership in the two types of organizations? The distinction between public and private organizations can be conceptualized as financial publicness, publicness based on the level of political authority and/or ownership status of the organization (Bozeman 1987; Rainey and Bozeman 2000), and the concepts and measures developed in this article should be applicable to all combinations of these important dimensions to allow for future public-private comparisons.

Further, we consider the applicability for different sources (leader and employee ratings), at different points in time/repeated measures and in an intervention study. We focus exclusively on the constructs and measures of transformational and transactional leadership, enabling later contributions to address their relations with outcomes.

After a more detailed discussion of the critiques of transformational and transactional leadership, we present our revised conceptualization and operationalization of these two types of leadership strategies. This is followed by a description of the methods used to test the operationalization, after which we present the results of the test. The article concludes

with a discussion of contributions and limitations, including how our revised

conceptualization and operationalization meet the critiques.

Unfolding the Conceptualization Critique of Transformational and Transactional Leadership

The discussion of transformational and transactional leadership in the generic leadership literature relies mainly on the work of Burns (1978) and Bass (1985). Transformational leadership refers to directing and inspiring individual efforts by transforming (and motivating) employees. This leadership strategy thus conceptualizes behaviors that seek to satisfy employees' higher order needs in order to engage them in attaining the organizational goals. Transactional leadership is based on transactions of pecuniary and non-pecuniary character (Antonakis, Avolio, & Sivasubramaniam, 2003; Bass, 1985). This leadership strategy refers to behaviors where the leader rewards employees for high effort and/or good performance or sanction them if their work effort or results are unsatisfactory (Bass, 1985). Thus, incentive structures are used to increase employees' attainment of organizational goals. Together, transactional and transformational leadership make up the active components of what is referred to as the "full-range leadership theory" (Antonakis, 2012), which also holds a passive component, namely laissez-faire leadership defined as the absence of active leadership behavior.

The assessment of the full range theory is widely conducted with the Multifactor Leadership Questionnaire (MLQ) (Van Knippenberg & Sitkin, 2013), which measures transformational leadership on four dimensions: i) idealized influence, also known as charisma, ii) inspirational motivation, iii) individualized consideration, and iv) intellectual stimulation. Transactional leadership includes three dimensions: i) contingent reward, ii) active management-by-exception, and iii) passive management-by-exception. Although other

models have been proposed (Carless, Wearing, & Mann, 2000; Podsakoff et al., 1990; Rafferty & Griffin, 2004; Conger & Kanungo, 1987; Shamir, House, & Arthur, 1993), the full-range theory and its MLQ is the best-known point of reference (Van Knippenberg & Sitkin, 2013).

Critical voices in generic leadership research have expressed concerns about the conceptualization of the full-range theory and its associated measure, MLQ. First, the conceptualization confounds the definitions of the leadership strategies with their effects (Van Knippenberg & Sitkin, 2013, Yukl, 1999; Judge & Piccolo, 2004). Transformational leadership has predominantly been described by its effects, e.g., that transformational leaders instill pride and respect, shift motivation from self-interest to collective interest, and inspire and motivate performance beyond expectations (Bass, 1985; Jung & Avolio, 2000). Defining a concept in terms of its effects is not an uncommon flaw, but it has highly undesirable consequences. The most important problem is that it prevents rigorous analysis of the particular leadership strategy and its effects (Van Knippenberg & Sitkin, 2013). For example, if transformational leadership per definition motivates employees, we cannot investigate the association between leadership and employee motivation, because the dependent variable becomes a defining part of the independent variable. Therefore, if no association between leadership and employee motivation is found, the leadership is – per definition – not transformational in this understanding of the concept. Second, the dimensions of transformational and transactional leadership are not exhaustively theorized (Van Knippenberg & Sitkin, 2013). The full-range theory does not answer, for example, why there are four transformational leadership dimensions, how they differ and relate, and what their common unifying factor is. Furthermore, the full-range theory does not distinguish between pecuniary and non-pecuniary transactional rewards, although social psychology research (e.g., Deci, Koestner, & Ryan, 1999) has convincingly demonstrated that these reward types

have completely different effects. Third, and following the lack of distinction between pecuniary and non-pecuniary transactional rewards, the full-range theory does not suggest boundary conditions for its application. We argue that the applicability to both private and public organizations demands that leadership concepts and measures are discussed in relation to Bozeman's (1987) aforementioned synthesis of ownership, funding and control into a dimensional model of publicness. Boyne (2002) for example identifies five studies that compare the managerial values of leaders in organizations with different publicness (ownership and funding measures), and he concludes that public managers are less materialistic than their private sector counterparts. If future studies should be able to analyze whether this leads to behavioral differences, we must develop leadership measures that are applicable in both types of organizations. In this respect, distinguishing between pecuniary and non-pecuniary rewards could be a big advantage for future public-private comparisons.

These important theoretical and methodological shortcomings emphasize that it is essential to revise the conceptualization and measurement of transformational and transactional leadership. We aim to do so by separating the leadership concepts from their expected outcomes, constructing meaningful dimensions and ensuring applicability in both private and public organizations. In the conclusion, we assess our re-conceptualization and re-operationalization of transformational and transactional leadership in relation to these three problems in the existing literature.

Transformational Leadership: Working Towards Sharing the Vision

Conceptualizing Transformational Leadership

When conceptualizing transformational leadership, it is import to identify the core behaviors of that particular leadership strategy. The existing multi-dimensional conceptualization in the full-range theory does not specify such core leadership behaviors, their relevance and

6

combination to form transformational leadership and the inclusion and exclusion criteria for dimensions are unclear (see Van Knippenberg & Sitkin, 2013, for an in-depth discussion of this problem): Why should, for example, idealized influence and inspirational motivation be seen as two separate dimensions although they arehighly correlated empirically? Conceptual overlap and empirical correlation question the entire multi-dimensionality of transformational leadership.

We argue that the distinctive theoretical aspect of transformational leadership is the leader's *intent* to activate employees' higher order needs. In this respect, the core ambition of transformational leaders is to induce employees to transcend their own self-interest for the sake of the organization (Bass, 1990; Antonakis, Avolio, & Sivasubramaniam, 2003; Podsakoff et al., 1990; Wright & Pandey, 2010; Wright, Moynihan, & Pandey, 2012). We therefore argue that the transformational leadership concept should capture leaders' systematic effort to transform employees to share the organizational goals because they are desirable in themselves. Transformational leaders may not succeed in transforming the employees, but transformational leadership behaviors are characterized by an ambition to foster a shared understanding among employees of how the organization should contribute to what is seen as desirable outcomes. Again, this does not imply that employees of transformational leaders necessarily share organizational goals or transcend their self-interest (because these are effects rather than constituent parts of the leadership strategy itself), but it clarifies the theoretical basis for arguing why some leadership behaviors should be grouped together and termed transformational. We argue that three behaviors are relevant: The leader's *attempts* to (1) formulate the organizational goals as a desirable future (a vision), (2) share this understanding with employees and (3) sustain the vision in the long run. This set of behaviors, that is, behaviors aimed at developing, sharing and sustaining an organizational vision, are theoretically seen as logical parts of the same latent concept capturing the *efforts*

to make employees share organizational goals and transcend their own self-interest. Leaders are expected to see all three behaviors as necessary for employee achievement of organizational goals through self-interest transcendence. Below, we explain in more detail below why each behavior should be characterized as transformational.

The first behavioral element concerns the leader's attempt to clarify the organization's vision. It is characterized as transformational behavior, because leaders are expected to see the existence of a clear vision as an important driver of unselfish employee action. Scientific evidence suggests that this can be correct (Latham & Yukl, 1975; Locke & Latham, 2002; Wright, 2007), but the key theoretical argument is still that leaders expect the development of a clear vision for the organization as necessary when they aim to make employees achieve organizational goals for other reasons than self-interest. This is about concretizing the organizational vision, thus trying to create an appealing vision seen as desirable by the employees.

The second behavioral element is attempts to share the vision with those employees who are ultimately supposed to execute it. When trying to share the vision, leaders with a transformational leadership strategy try to establish a clear understanding of the relation between actions and goals reflected in the vision. Again, the key argument is that leaders will see these sharing efforts as necessary for making the employees want to contribute to realizing the vision, and the importance of such sense-giving for informing and constraining employees' actions has been demonstrated in organizational psychology research (Weick, Sutcliffe, & Obstfeld, 2005). Public administration research has connected awareness of the vision to employee motivation (Paarlberg & Perry, 2007), and vision-sharing behaviors can be seen as attempts to articulate the direction in which the organization is heading and explicate how the day-to-day activities and actions of employees support achievement of

goals and vision (Paarlberg & Perry, 2007; Paarlberg & Lavigna, 2010). As such, we classify them as transformational leadership behavior.

The third behavioral element in our conceptualization of transformational leadership is actions intended to sustain a shared vision in the short and long run. Leaders are expected to see such actions as having potential for facilitating enduring acceptance of, and collaboration to achieve, the vision. By continuously emphasizing how employees' work contributes to the organization and its vision, leaders can make an effort to reinforce employees' perceptions of task significance and energy to pursue certain actions in the short as well as the long run (Wright, Moynihan, & Pandey, 2012), and this makes it an integral part of the transformational leadership concept.

The generic leadership literature emphasizes visions as a core element of transformational leadership (Van Knippenberg & Sitkin, 2013), and this applies to both public and private organizations. When leaders share the vision, the message is expected to establish a shared sense of purpose and give multiple employees the same understanding of the purpose of work (Carton, Murphy, & Clark 2014). This shared understanding can potentially alter employee perceptions of goal priorities and encourage them to devote effort toward the vision (Wright, Moynihan, & Pandey, 2012, p. 207). All types of leaders can face conditions that make it very important to develop, share and sustain a vision. Organizations with a high level of political control can have multiple, ambiguous and potentially conflicting goals and they often face conflicting expectations and involvement of many different stakeholders. A leadership behavior focused on visions can give this type of organization a coherent direction. If organizations have public ownership, politicians are the ultimate principals, and all societal groups are legitimate stakeholders. This can make it more difficult – but also more important – for public managers to give direction through leadership strategies with focus on the organizational vision. Funding considerations can also make the

visionary element of leadership very relevant for both public and private organizations. In organizations with high levels of public funding, the leader must be able to align the organizational vision with the demands of the sponsor body, while new market conditions in the private sector can demand strong emphasis on the vision to make organizations adapt to a changed context. The latter scenario can be exemplified by Nokia leader Stephen Elop's burning platform memo where he urged the company to change to be able to compete with the iPhone (Alcacer, Khanna & Snively 2014). Based on these arguments, our revised conceptualization of transformational leadership focuses on the leaders' efforts to establish a strong vision, and this conceptualization is relevant for both private and public leaders.

In short, we theoretically define transformational leadership as behaviors that seek to develop, share, and sustain a vision, and the key theoretical reason for categorizing these behaviors as transformational is that we see the intention behind these behaviors as attempts to encourage employees to transcend their own self-interest and achieve organizational goals. Although there are three types of transformational behavior, we argue that they are intertwined in the sense that they all reflect the same latent ambition to transform the employees to share and act on the vision and that the behaviors are only theoretically meaningful if used together.

Operationalizing Transformational Leadership

One of the key criticisms of transformational leadership is, as mentioned, that it conceptually confounds leadership actions and their effects (Van Knippenberg & Sitkin, 2013, p. 43), and this conceptual problem leads to an operationalization problem where leadership effects are measured rather than leadership behavior. Van Knippenberg and Sitkin (2013, p. 40) thus argue "the definitional problems have their parallels in measurement problems". This threatens the validity of the many effect studies based on these measures. In addressing this

issue, we base our operationalization on a review of the literature to identify items that clearly reflect a leader's *actions* to develop, share, and sustain a vision. In generic management literature, we relied on inspiration from Podsakoff et al. (1990), MacKenzie et al. (2001) and House (1998), and in public administration literature, we focused on Moynihan et al. (2012). Seven items were selected and carefully rephrased to focus on leader behaviors to avoid confounding of leadership actions and leadership effects. The wording can be seen in Table 1 below, while Table A-1 in the online appendix explicates the link to existing items from the literature ([www address to online appendix will appear here after the review – the appendix is now part of the article]). In the results section, we assess the psychometric properties of our transformational leadership scale including convergent and discriminant validity. All items were adapted to fit both leader and employee ratings. The items were framed for leaders by stating: "As a leader I …", and for employees the introductory text was: "My leader …".

TABLE 1 HERE

A valid insight from the full-range theory is that it is not enough to conceptualize and measure transformational leadership in isolation. It should, as a minimum, also function in relation to a consistent conceptualization of transactional leadership. This position is supported by existing studies, which suggest that transformational leadership is important both in its own right and in combination with transactional leadership (Bass & Riggio, 2006; Rowold, 2011; Hargis, Watt, & Piotrowski, 2011; O'Shea et al., 2009; Bass et al., 2003). Below, we discuss the conceptualization of transactional leadership.

Transactional Leadership: Working Towards Aligning Organizational Goals with Employees' Self-interest

Conceptualizing Transactional Leadership

Transactional leadership has transactions between leader and employees at its conceptual core (Podsakoff et al., 2006), and we accordingly define it as the use of contingent rewards and sanctions. While both transactional and transformational leadership are directed towards achieving organizational goals, the key difference is that we see transactional leadership behavior as being intended to create employee self-interest in achieving the goals, while transformational leadership theoretically is based on an intention to encourage employees to transcend their own self-interest. Transactional leadership thus entails the use of contingent rewards and sanctions to make individual employees pursue their own self-interest while contributing to organizational goal attainment. This rests on the assumption that through appropriate incentives the self-interest of individual employees may align with the interest of the organization. Only contingent rewards and sanctions are relevant: Whenever employees are rewarded or sanctioned, these transactions should relate directly to employees' specific effort or performance. Otherwise, the transactions cannot be expected to be effective.

We agree with the differentiation in the full-range theory between rewards and sanctions, but in line with House (1998) we argue that it is important to refine the differentiation by also distinguishing between pecuniary and non-pecuniary rewards as these have different effects. Perry, Engbers, and Jun (2009) and Weibel, Rost, and Osterloh (2010) for example demonstrate that the effects of pecuniary rewards in public organizations can be negative. Especially if financial incentives are seen as controlling, they can crowd out intrinsic motivation (Weibel et al., 2009 Andersen & Pallesen, 2008), and this is not the case with non-pecuniary (e.g., verbal) rewards (e.g., Deci et al., 1999). Consequently, we conceptualize transactional leadership as entailing the use of three types of performance- or

effort-contingent types of behavior: use of contingent non-pecuniary rewards, contingent pecuniary rewards, and contingent sanctions. Specifically related to the applicability to both private and public organizations, the distinction between different types of rewards is highly warranted. Given that public managers are less materialistic (Boyne 2002), they could very well have a different reward-related behavior compared with private leaders, and it is an empirical question whether they use rewards less compared with private managers or whether they simply substitute material rewards with non-material rewards. The revised conceptualization of transactional leadership will allow future studies to find out not only whether public leaders use rewards less than private leaders, but also whether they use different (non-pecuniary) rewards. If valued by the employees, all three behaviors would be theoretically expected to have a potential effort-inducing effect, but the perception of sanctions and the two types of rewards can be very different (Frey 1997). Pecuniary rewards can for example be bonus pay and perks, while non-pecuniary rewards can be praise. The last type of performance/effort-contingent behavior is sanctions, e.g., punishment of errors, negative effort, and performance deviances.

We argue that transactional leadership should be seen as a formative construct where the use of pecuniary and non-pecuniary rewards and sanctions jointly construe the conceptual and empirical significance of transactional leadership. These three types of behaviors can be – but are not necessarily – alternatives. For example, if a leader uses both contingent pecuniary rewards and contingent sanctions, the leader is rated to have a higher level of transactional leadership, because the leader uses two types of transactions. This means that it is not necessary for the three types of behavior to co-vary, given that they can be seen as alternative transaction types, that is, different ways to perform transactional leadership, while it would alter the conceptual domain of transactional leadership if one of three types of behavior is excluded (Jarvis, MacKenzie, & Podsakoff, 2003).

Operationalizing Transactional Leadership

Similar to the process applied for our operationalization of transformational leadership, we reviewed the transactional leadership literature for items that reflect our three conceptual elements. Again, we carefully selected and re-phrased items in order for them not to confound leadership with its effects. Table 2 presents the 12 items reflecting the three components of transactional leadership, and Table A-2 in the Online Appendix shows how they are based on existing literature.

TABLE 2 ABOUT HERE

Methods and Validation Procedures

Our validation procedure includes two steps to ensure validity and reliability of the transformational and transactional leadership scales. In the first step, we assess the psychometric properties of a four-factor model, compare it to alternative factor structures, and examine whether the measurement model is consistent across sectors, in repeated surveys, and in an intervention study. The four factors reflect: 1) transformational leadership, 2) contingent non-pecuniary rewards, 3) contingent pecuniary rewards, and 4) contingent sanctions. In the second step, we investigate the correlations between the transformational leadership and transactional leadership and test whether the scales discriminate from each other. While the main results are presented for employee ratings of transformational and transactional leadership, we test whether the measurement model is equally applicable for leaders' self-ratings. Before we discuss each step in greater detail, the sample and data collection is briefly described.

Sample and Data Collection

Our data stem from four independent surveys collected in relation to an experimental test of a leadership training program in Denmark. Surveys were distributed among leaders and employees before the beginning of the training program (April and August 2014), and follow-up surveys were administered to leaders and their employees after the training program (August 2015). The leadership training program ran from September 2014 to May 2015. Surveys were internet-based, but paper-based invitations including a link to the online questionnaire were distributed to employees who did not have a valid e-mail account. As discussed in Boye et al. (2015; 2016), we engaged in a number of procedures to obtain high response rates and make sure that respondents prioritized answering the survey carefully. Five different sectors were included in the study: high schools (only public organizations), schools (public and private organizations), day care (public and private organizations), tax offices (only public organizations), and banks (only private organizations). From these sectors, 672 leaders volunteered to participate in the experimental training program and completed the pre-treatment survey (i.e., a response rate of 100 %). A pre-treatment survey was distributed to their 19,552 employees with an overall response rate of 45.3 %. The follow-up survey was completed by 451 leaders in August 2015 (corresponding to a response rate of 87 %[1]), and 7,538 employees fully or partially completed the post-treatment employee survey (response rate 49.8 % of all recipients of the second questionaire). The survey data represents an unbalanced panel, because it includes employees working in the organizations throughout the study, as well as employees leaving and entering the organizations during the study. All surveys included identical items on transformational leadership and transactional leadership (cf. Table 1 and 2), and this allows us to assess the psychometric properties of a

---

[1] The post-treatment questionnaire was distributed to the 521 leaders who remained part of the experiment throughout the leadership training program. Dropouts were primarily due to job changes.

measurement model across time and respondents (employees' other-ratings and leaders' self-ratings of transformational and transactional leadership).

A stratified random sampling method was used to assign the leaders to one of four groups (three treatment groups and a control group). Strata ensure an even distribution of leaders from the different types of organizations in treatment and control groups, and the random assignment prevents selection bias of participants on treatments (Angrist and Pischke, 2009, p. 15). The three separate treatments encompass interventions designed to train the leaders in transformational leadership, transactional leadership, or a combination of the two. Leaders in the control group were not assigned to any of the treatment groups. The leadership training consisted of four full days of instruction and exercises. The training was set up to support participants' work on applying the leadership strategy in their organizations in-between instruction days. Elaborate information on the assignment-to-treatment procedure, learning principles behind the training programs, contents of individual teaching sessions, and design of the experiment can be found in Jacobsen, Bøllingtoft and Andersen (2015) and Holten, Bøllingtoft and Wilms (2015).

Psychometric Properties of Leadership Scales

To test the validity of our transformational and transactional leadership measures, we performed confirmatory factor analysis (CFA). CFA formally tests whether a set of indicators converges on latent factors as specified a priori by theory (Acock, 2013). CFA is based on an asymptotic distribution free estimator to account for the ordinal nature of our data (5-point Likert-scale items). Consistent with our conceptualizations of transformational leadership and transactional leadership, a four-factor measurement model was specified from our pool of 19 items (cf. Tables 1 and 2) to reflect: 1) transformational leadership, 2) contingent non-pecuniary rewards, 3) contingent pecuniary rewards, and 4) contingent sanctions. The latter

three factors can be seen as a formative construct where each factor constitutes unique and non-interchangeable components of the transactional leadership construct. Contingent non-pecuniary rewards, contingent pecuniary rewards and contingent sanction are modelled as three independent first-order factors because they can have different antecedents and consequences, and existing literature accordingly treats them separately (this is also the case in the full-range theory and its associated measure, MLQ, see for example Antonakis, Avolio, and Sivasubramaniam 2003). Convergent validity is demonstrated when the average variance extracted for an indicator is above 0.5. To decide on overall fit of the measurement model – that is, its ability to reproduce the observed covariance matrix (Vandenberg & Lance, 2000, p. 43) – we rely on the chi-square test and three of the most common approximate fit measures: the root mean square of approximation (RMSEA), the comparative fit index (CFI), and the standardized root mean square residual (SRMR). We apply the model fit thresholds proposed by Williams, Vandenberg, and Edwards (2009). To evaluate the robustness of our four-factor model, we compared it to two alternative models with simpler factor structures. We tested our model against: 1) A two-factor model in which all three transactional leadership factors were combined into a single factor and 2) a three-factor model in which contingent pecuniary rewards and contingent non-pecuniary rewards were combined into a single factor.

The structure of our data in multiple groups (five sectors, public-private, treatment-control) and time points (baseline-follow-up) allows us to validate our leadership scales across these contexts (Williams, Vandenberg, & Edwards, 2009). We thereby explore measurement invariance, testing whether the properties of the underlying measurement model are consistent across groups and time. Specifically, we test for "configural" and "metric" invariance. Configural invariance refers to equivalence of our measurement model configuration, that is, the pattern of factors and indicators is the same across groups and time.

Metric invariance concerns equivalence of factor loadings (i.e., the relationship between individual items and factors) across groups and time (Horn & McArdle, 1992; Vandenberg & Lance, 2000). The measurement invariance tests are performed in sequential steps imposing still stricter restrictions on the measurement model. For each step it is evaluated whether data supports the preceding invariance test. We compare the nested models by chi-square statistics (change in the chi-square value, $\Delta\chi^2$, given the change in degrees of freedom between models) and differences in RMSEA (ΔRMSEA), CFI (ΔCFI) and SRMR (ΔSRMR). Following the recommendations of Cheung and Rensvold (2002) and Chen (2007), invariance is demonstrated for a difference of less than 0.015 in RMSEA, 0.01 in CFI, and 0.030 in SRMR.

Inter-factor Correlations and Discriminant Validity

Often-voiced criticisms of the MLQ are that transformational and transactional leadership factors correlate highly and discriminate poorly (Van Knippenberg & Sitkin, 2013), and this makes it important to address these issues for our measures. The four factors in our measurement model are therefore allowed to correlate in the CFA, enabling us to investigate the inter-factor correlations. A related question is whether factors discriminate. Discriminant validity exists when a latent factor (e.g., the transformational leadership factor in our model) accounts for more variance in the indicators/items related to this factor than in other factors (e.g., the contingent non-pecuniary rewards factor) or measurement error (Farrell & Rudd, 2009). According to Fornell and Larcker (1981), discriminant validity is established when the average variance extracted for any two factors exceeds the shared variance between these factors. The average variance extracted consists of the average of the squared correlations (or factor loadings) between individual indicators and the associated factor. Shared variance between any two factors equals the squared correlation between these factors. Using this

approach, we test discriminant validity by comparing all combinations of our four factors: transformational leadership, contingent non-pecuniary rewards, contingent pecuniary rewards, and contingent sanctions.

Results

In this section, we present the results of a series of confirmatory factor analyses to assess the validity and reliability of our transformational and transactional leadership scales. The main results are presented for employee data, but we also test the measurement model using leaders' self-ratings, demonstrating equal applicability across sources.

A Four-Factor Measurement Model: Psychometric Properties

According to the chi-square test of exact fit: $\chi^2(146) = 3567.29$, $p < 0.001$ and the CFI = 0.858 (although acceptable for RMSEA = 0.039 and SRMR = 0.076), our 19-item target model (cf. Tables 1 and 2) did not fit data well. Since our sample includes data from multiple groups (e.g., different sectors), we follow the procedure adopted by Antonakis and House (2014) to re-estimate a trimmed model based on a homogenous subsample of public school participants, which is also the largest subsample in our data. Based on a number of iterations, we inspected the modification indexes and retained items that clearly reflected our conceptualizations, loaded highly on their respective factors, and showed discriminant properties. On the basis, we replicated the test of the trimmed four-factor model with 13 items on the full data. The fit to data on the full sample is good: $\chi^2(59) = 1006.77$, $p < 0.001$, RMSEA = 0.032, CFI = 0.956, and SRMR = 0.029. Mean standardized factor loadings were high for all factors ($\lambda_{\text{Transformational}} = 0.82$, $\lambda_{\text{Non-pecuniary rewards}} = 0.91$, $\lambda_{\text{Pecuniary rewards}} = 0.85$, $\lambda_{\text{Sanctions}} = 0.85$) with no individual loadings below 0.5, suggesting convergent validity of our model. Next, we compared the model with two alternatives: 1) A two-factor model where all

transactional leadership items were constrained to load on the same factor and 2) a three-factor model in which the items affiliated with the contingent reward (i.e., non-pecuniary and pecuniary) factors were constrained to load on the same common factor. Results from CFA on our main model and the alternative model specifications are presented in Table 3. The two-factor and three-factor models failed to fit data well and performed significantly worse than our four-factor model based on difference tests: two-factor model: $\Delta\chi^2(5) = 8{,}914$, $p < 0.001$ and three-factor model: $\Delta\chi^2(3) = 5{,}047$, $p < 0.001$. Differences in RMSEA, CFI, and SRMR clearly support this pattern. Thus, the four-factor model was retained for further analyses.

TABLE 3 HERE

A Four-Factor Measurement Model: Testing Multiple Group and Time Invariance

To assess the psychometric properties of our four-factor model across multiple groups and time, we tested for configural and metric invariance across 1) time, 2) treatment/control group, and 3) sector. Three independent invariance tests were performed, each entailing two sequential steps: First, a model with all parameters constrained to be equal was compared to a model where only the factorial structure and pattern of loadings were constrained to be equal across the grouping variable (i.e., test for configural invariance). Second, the latter model was compared to a model where factor loadings were also required to be equal across the grouping variable (i.e., test for metric invariance). Table 4 presents factor loadings and fit indices for the four-factor model.

TABLE 4 HERE

20

Tables 5-7 summarize the results of the invariance tests. The results indicate that our measurement model is invariant across repeated surveys (time). The baseline model in which all parameters were constrained to be equal across time fits data well: $\chi^2(163) = 1211.69$, $p < 0.001$, RMSEA = 0.028, CFI = 0.951, and SRMR = 0.036. Almost identical results appear for the baseline model in which all parameters are constrained to be equal across treatment and control group, $\chi^2(163) = 1211.31$, $p < 0.001$, RMSEA = 0.028, CFI = 0.951, and SRMR = 0.042, indicating not only configural and metric invariance, but also equivalence of error-variances and intercepts. Results indicate configural and metric invariance across sectors with differences in RMSEA less than 0.015, in CFI less than 0.01, and in SRMR less than 0.030 between models 2 and 3 (cf. table 7). Across study sectors, our four-factor measurement model generally seems equally applicable in people-changing (service) and people-processing (administration) organizations. Importantly, it also seems equally applicable to public and private organizations. Factor loadings, model fit indices, and reliability scores for each "sector" subsample split by time are listed in Table A-3 in the online appendix. It should be acknowledged that SRMR is high in our small subsamples (e.g., private lower secondary school, private day care, and banks). Simulation studies show that SRMR is sensitive to sample size with a greater positive bias in small samples (Anderson & Gerbing, 1984). If we compare CFI values, which are not sensitive to sample size, we see that they are practically unchanged between small and large subsamples (e.g., banks versus tax offices in Table A-3).. Moreover, the four-factor model displays acceptable psychometric properties for leader self-rated transformational and transactional leadership (cf. Table A-4 in the online appendix). For the full sample of 982 observations (597 unique leaders with valid answers on the relevant items in at least one survey), all standardized factor loadings are above 0.5 with RMSEA and CFI values of 0.035 and 0.933, respectively. Despite the small number of observations in some of the configurations (e.g., the treatment versus control group

comparison), our measurement model generally performs satisfactorily for leaders' self-ratings.

A Four-Factor Measurement Model: Inter-factor Correlations and Discriminant Validity

Table 8 presents: 1) inter-factor correlations, 2) average variance extracted and shared variance for assessing discriminant validity, and 3) composite reliability scores in terms of Cronbach's alpha and Jöreskog's rho. Correlations between the four factors range between 0.589 and 0.135 and any two model factors thus share less than half of their variance with each other. Our leadership scales also discriminate as the average variance extracted for any two model factors well exceeds the shared variance between these factors. Finally, composite reliability scores suggest internal consistency among the leadership scales. Cronbach's alpha is well above the recommended lower threshold of 0.7 for all composite constructs, and Jöreskog's rho (which is not sensitive to the number of items) supports this pattern with values far exceeding the 0.6 threshold. Inter-factor correlations, evidence of discriminant validity, and reliability scores for leader ratings all support discriminant properties of the four-factor model using leaders' self-ratings and internal consistency among items associated with individual factors (see Table A-5 in the online appendix).

Conclusion

This article aims to contribute to solving three problems of earlier conceptualizations and measurements of transformational and transactional leadership, and the obvious question being how successful our re-conceptualization and re-operationalization efforts have been?

Concerning the need to separate the conceptualizations and operationalization of leadership behavior from their effects, we argue that the conceptualizations do not confound the leadership strategies with their proposed effects, because they focus on behavior (behaviors that seek to develop, share, and sustain a vision and actual use of rewards and sanctions). Correspondingly, the items used in the new operationalizations also ask about behavior (e.g., whether the leaders seek to make employees accept common goals for the organization).

Whether we have succeeded in constructing theoretically and empirically meaningful dimensions is more difficult to assess. Theoretically, we distinguish between sanctions, pecuniary rewards and non-pecuniary rewards, because the theoretic dynamics are fundamentally different. Existing research also shows that these types of transactions have different consequences. Consistent with our conceptualization of a unidimensional transformational leadership construct and three transactional leadership components (i.e., contingent non-pecuniary, contingent pecuniary, and contingent sanctions), we find empirical support for a 13-item measurement model of transformational leadership and transactional leadership consisting of four first-order factors. This model demonstrates convergent validity, discriminant validity, and measurement invariance across groups, sources, sectors and time.

Concerning applicability in both private and public organizations, we have presented theoretical arguments for the relevance of the same concepts, and our tests demonstrated that the same items can be used in both types of organizations. We fully acknowledge that some types of leadership behavior need to be conceptualized and measured differently in different

types of organizations, but when public and private organizations can be compared using the same concepts and measures, it is a huge advantage. Although we would expect different levels of transformational and transactional leadership in organizations with varying degree of publicness (that is, different levels of public funding, political control and public ownership), nothing indicates that transformational and transactional leadership cannot be conceptualized and measured similarly for private and public organizations. The conceptualizations and measures of transformational and transactional leadership set forth in this article thus allow future research to continue to make comparisons between public and private organizations, contributing to important questions on the antecedents and consequences of these leadership strategies in both types of organizations. On this basis, our proposed model seems to offer a good point of departure for reconsolidating the field. To further explore the empirical applicability of the proposed measures of transformational and transactional leadership, we encourage scholars to use the scales in other organizational, national and/or cultural contexts. For example, it would be very relevant to assess whether our conceptualization and operationalization can also be used in non-profit organizations.

Our proposed measures of transformational and transactional leadership are available to the research community and we invite scholars to apply and use these measures in their future research to help build sound and cumulative knowledge on the effects of transformational and transactional leadership in the private, nonprofit, and public sectors.

References

Acock, A. C. (2013). *Discovering structural equation modeling using Stata* (1st ed.). College Station, TX: Stata Press.

Alcacer, J., Khanna, T. & Snively, C. (2014). *The Rise and Fall of Nokia*. Harvard Business School http://www.essaybaywriters.com/blog/wp-content/uploads/2016/03/714428-PDF-ENG.pdf (accessed 5 August 2016).

Andersen, L. B., & Pallesen, T. (2008). "Not just for the money?" How financial incentives affect the number of publications at Danish research institutions. *International Public Management Journal*, *11*(1), 28–47.

Anderson, J. C., & Gerbing, D. W. (1984). The Effect of Sampling Error on Convergence, Improper Solutions, and Goodness-of-Fit Indices for maximum Likelihood Confirmatory Factor Analysis. *Psychometrika*, *49*(2): 155–173.

Angrist, J. D., & Pischke, J.-S. (2009). *Mostly harmless econometrics: an empiricist's companion*. Princeton: Princeton University Press.

Antonakis, J. (2012). Transformational and charismatic leadership. In D. V. Day, & J. Antonakis (Eds.), *The Nature of Leadership* (2nd ed., pp. 256-288). Los Angeles: SAGE.

Antonakis, J., Avolio, B. J., & Sivasubramaniam, N. (2003). Context and leadership: an examination of the nine-factor full-range leadership theory using the Multifactor Leadership Questionnaire. *The Leadership Quarterly*, *14*(3), 261–295.

Antonakis, J., & House, R. J. (2014). Instrumental leadership: Measurement and extension of transformational–transactional leadership theory. *The Leadership Quarterly, 25*(4), 746-771.

Avolio, B. J., Reichard, R. J., Hannah, S. T., Walumbwa, F. O., & Chan, A. (2009). A meta-analytic review of leadership impact research: experimental and quasi-experimental

studies. *The Leadership Quarterly*, *20*(5), 764–784.

Bass, B. M. (1985). *Leadership and performance beyond expectations*. New York : London: Free Press, Collier Macmillan.

Bass, B. M. (1990). From transactional to transformational leadership: learning to share the vision. *Organizational Dynamics*, *18*(3), 19–31.

Bass, B. M., & Riggio, R. E. (2006). *Transformational leadership* (2nd ed.). Mahwah, N.J: L. Erlbaum Associates.

Bass, B. M., Avolio, B. J., Jung, D. I., & Berson, Y. (2003). Predicting unit performance by assessing transformational and transactional leadership. *Journal of Applied Psychology*, *88*(2), 207–218.

Bellé, N. (2014). Leading to make a difference: a field experiment on the performance effects of transformational leadership, perceived social impact, and public service motivation. *Journal of Public Administration Research and Theory*, *24*(1), 109–136.

Bettencourt, L. A. (2004). Change-oriented organizational citizenship behaviors: the direct and moderating influence of goal orientation. *Journal of Retailing*, *80*(3), 165–180.

Boye, S., Christensen, J., Jensen, U.T., Bro, L.L., Bøllingtoft, A., Eriksen, T.M., Jacobsen, C.B., Ladenburg, J., Nielsen, P.A., Salomonsen, H.H., Westergaard-Nielsen, N., Andersen, L.B. (2015) *Technical report. Survey of leaders and employees, pre-treatment*

http://ps.au.dk/fileadmin/Statskundskab/Billeder/Forskning/Forskningsprojekter/LEAP/Dokumenter/LEAP_technical_report_pretreatment_final2.pdf (accessed 5 August 2016).

Boye, S., Christensen, J., Jensen, U.T., Bro, L.L., Bøllingtoft, A., Eriksen, T.M., Jacobsen, C.B., Ladenburg, J., Nielsen, P.A., Salomonsen, H.H., Westergaard-Nielsen, N., Würtz, A., Andersen, L.B. (2016) *Technical report. Survey of leaders and employees,*

post-treatment

http://ps.au.dk/fileadmin/Statskundskab/Billeder/Forskning/Forskningsprojekter/LEAP/Dokumenter/LEAP_techrepport_posttreatment.pdf (accessed 5 August 2016).

Boyne, G. A. (2002). Public and private management: What is the difference? *Journal of Management Studies*, *39*(1), 97–122.

Boyne, G. A. (2003). Sources of public service improvement: a critical review and research agenda. *Journal of Public Administration Research and Theory*, *13*(3), 367–394.

Bozeman, B. (1987). *All Organizations Are Public*. San Francisco: Jossey-Bass.

Burns, J. M. (1978). *Leadership*. New York: Harper & Row.

Carless, S., Wearing, A., & Mann, L. (2000). A short measure of transformational leadership. *Journal of Business and Psychology*, *14*(3), 389–405.

Carton, A. M., Murphy, C., & Clark, J. R. (2014). A (blurry) vision of the future: How leader rhetoric about ultimate goals influences performance. *Academy of Management Journal*, *57*(6), 1544–1570.

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.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *9*(2), 233–255.

Conger, J. A., & Kanungo, R. N. (1987). Toward a behavioral theory of charismatic leadership in organizational settings. *Academy of Management Review*, *12*(4), 637–647.

Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin*, *125*(6), 627–668.

Farrell, A. M., & Rudd, J. M. (2009). Factor analysis and discriminant validity: a brief review of some practical issues. Presented at the Australia and New Zealand Marketing Academy Conference (ANZMAC).

Fernandez, S. (2005). Developing and testing an integrative framework of public sector leadership: evidence from the public education arena. *Journal of Public Administration Research and Theory*, *15*(2), 197–217.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, *18*(1), 39.

Frey, B. S. (1997). *Not Just for the Money: An Economic Theory of Personal Motivation*. Edward Elgar Publishing: Books

Hargis, M. B., Watt, J. D., & Piotrowski, C. (2011). Developing leaders: examining the role of transactional and transformational leadership across business contexts. *Organization Development Journal*, *29*(3), 51–66.

Hassan, S., & Hatmaker, D. M. (2015). Leadership and performance of public employees: effects of the quality and characteristics of manager-employee relationships. *Journal of Public Administration Research and Theory*, *25*(4), 1127–1155.

Holten, A.-L., Bøllingtoft, A., & Wilms, I. (2015). Leadership in a changing world: developing managers through a teaching and learning programme. *Management Decision*, *53*(5), 1107–1124.

Horn, J. L., & McArdle, J. J. (1992). A practical and theoretical guide to measurement invariance in aging research. *Experimental Aging Research*, *18*(3), 117–144.

House, R. J. (1998). Appendix: measures and assessments for the charismatic leadership approach: scales, latent constructs loadings, Cronbach alphas, interclass correlations. In F. Dansereau, & F. J. Yammarino (Eds.), *Leadership: the multiple-level*

approaches. Contemporary and alternative (pp. 23-30). London: JAI Press.

Jacobsen, C. B., & Andersen, L. B. (2015) Is Leadership in the Eye of the Beholder? A Study

of Intended and Perceived Leadership Practices and Organizational Performance.

Public Administration Review 75, (6): 829–841.

Jacobsen, C. B., Bøllingtoft, A., & Andersen, L. B. (2015). Can leadership training teach

leaders to LEAP? Experimenting with leadership training and leadership strategies.

Presented at the Association for Public Policy Analysis and Management Conference,

November 12-14 2015, Miami, FL.

Jarvis, C. B., MacKenzie, S. B., & Podsakoff, P. M. (2003). A critical review of construct

indicators and measurement model misspecification in marketing and consumer

research. *Journal of Consumer Research*, *30*(2), 199–218.

Judge, T. A., & Piccolo, R. F. (2004). Transformational and transactional leadership: a meta-

analytic test of their relative validity. *Journal of Applied Psychology*, *89*(5), 755–768.

Jung, D. I., & Avolio, B. J. (2000). Opening the black box: an experimental investigation of

the mediating effects of trust and value congruence on transformational and

transactional leadership. *Journal of Organizational Behavior*, *21*(8), 949–964.

Latham, G. P., & Yukl, G. (1975). A review of research on the application of goal setting in

organizations. *Academy of Management Journal*, *18*(4), 824–845.

Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and

task motivation: a 35-year odyssey. *American Psychologist*, *57*(9), 705–717.

Lowe, K. B., Kroeck, K. G., & Sivasubramaniam, N. (1996). Effectiveness correlates of

transformational and transactional leadership: a meta-analytic review of the MLQ

literature. *The Leadership Quarterly*, *7*(3), 385–425.

MacKenzie, S. B., Podsakoff, P. M., & Rich, G. A. (2001). Transformational and

transactional leadership and salesperson performance. *Journal of the Academy of*

Marketing Science, *29*(2), 115–134.

Moynihan, D. P., & Ingraham, P. W. (2004). Integrative leadership in the public sector: a model of performance-information use. *Administration & Society*, *36*(4), 427–453.

Moynihan, D. P., Pandey, S. K., & Wright, B. E. (2012). Setting the table: how transformational leadership fosters performance information use. *Journal of Public Administration Research and Theory*, *22*(1), 143–164.

O'Shea, P. G., Foti, R. J., Hauenstein, N. M. A., & Bycio, P. (2009). Are the best leaders both transformational and transactional? A pattern-oriented analysis. *Leadership*, *5*(2), 237–259.

Paarlberg, L. E., & Lavigna, B. (2010). Transformational leadership and public service motivation: driving individual and organizational performance. *Public Administration Review*, *70*(5), 710–718.

Paarlberg, L. E., & Perry, J. L. (2007). Values management: aligning employee values and organization goals. *The American Review of Public Administration*, *37*(4), 387–408.

Parry, K. W., & Sinha, P. N. (2005). Researching the trainability of transformational organizational leadership. *Human Resource Development International*, *8*(2), 165–183.

Perry, J. L., Engbers, T. A., & Jun, S. Y. (2009). Back to the future? Performance-related pay, empirical research, and the perils of persistence. *Public Administration Review*, *69*(1), 39–51.

Podsakoff, P. M., Bommer, W. H., Podsakoff, N. P., & MacKenzie, S. B. (2006). Relationships between leader reward and punishment behavior and subordinate attitudes, perceptions, and behaviors: a meta-analytic review of existing and new research. *Organizational Behavior and Human Decision Processes*, *99*(2), 113–142.

Podsakoff, P. M., MacKenzie, S. B., Moorman, R. H., & Fetter, R. (1990). Transformational

leader behaviors and their effects on followers' trust in leader, satisfaction, and

organizational citizenship behaviors. *The Leadership Quarterly*, *1*(2), 107–142.

Rafferty, A. E., & Griffin, M. A. (2004). Dimensions of transformational leadership:

conceptual and empirical extensions. *The Leadership Quarterly*, *15*(3), 329–354.

Rainey H.G. & Bozeman, B. (2000). Comparing Public and Private Organizations: Empirical

Research and the Power of the A Priori *Journal of Public Administration Research

and Theory, 10*(2), 447-470.

Rainey, H. G. (2014). *Understanding and managing public organizations* (5th Edition). San

Francisco: Jossey-Bass & Pfeiffer Imprints, Wiley.

Rowold, J. (2011). Relationship between leadership behaviors and performance: the

moderating role of a work team's level of age, gender, and cultural heterogeneity.

Leadership & Organization Development Journal, *32*(6), 628–647.

Shamir, B., House, R. J., & Arthur, M. B. (1993). The motivational effects of charismatic

leadership: a self-concept based theory. *Organization Science*, *4*(4), 577–594.

Trottier, T., Van Wart, M., & Wang, X. (2008). Examining the nature and significance of

leadership in government organizations. *Public Administration Review*, *68*(2), 319–

333.

Van Knippenberg, D., & Sitkin, S. B. (2013). A critical assessment of charismatic—

transformational leadership research: back to the drawing board? *The Academy of

Management Annals*, *7*(1), 1–60.

Van Wart, M. (2013). Administrative leadership theory: a reassessment after 10 years. *Public

Administration*, *91*(3), 521–543.

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement

invariance literature: suggestions, practices, and recommendations for organizational

research. *Organizational Research Methods*, *3*(1), 4–70.

Vigoda-Gadot, E., & Beeri, I. (2012). Change-oriented organizational citizenship behavior in public administration: the power of leadership and the cost of organizational politics. *Journal of Public Administration Research and Theory*, *22*(3), 573–596.

Vogel, R., & Masal, D. (2015). Public leadership: a review of the literature and framework for future research. *Public Management Review*, *17*(8), 1165–1189.

Weibel, A., Rost, K., & Osterloh, M. (2009). Pay for performance in the public sector: benefits and (hidden) costs. *Journal of Public Administration Research and Theory*, *20*(2), 387–412.

Weick, K. E., Sutcliffe, K. M., & Obstfeld, D. (2005). Organizing and the process of sensemaking. *Organization Science*, *16*(4), 409–421.

Williams, L. J., Vandenberg, R. J., & Edwards, J. R. (2009). 12 structural equation modeling in management research: a guide for improved analysis. *The Academy of Management Annals*, *3*(1), 543–604.

Wright, B. E. (2007). Public service and motivation: Does mission matter? *Public Administration Review*, *67*(1), 54–64.

Wright, B. E., & Pandey, S. K. (2010). Transformational leadership in the public sector: Does structure matter? *Journal of Public Administration Research and Theory*, *20*(1), 75–89.

Wright, B. E., Moynihan, D. P., & Pandey, S. K. (2012). Pulling the levers: transformational leadership, public service motivation, and mission valence. *Public Administration Review*, *72*(2), 206–215.

Yukl, G. (1999). An evaluation of conceptual weaknesses in transformational and charismatic leadership theories. *The Leadership Quarterly*, *10*(2), 285–305.

Table 1. Operationalization of measurement instrument for transformational leadership,

Item no.	Item wording: A. Leader version	Item wording: B. Employee version
	As a leader I…	*My leader…*
1.	Concretize a clear vision for the organization's future	Concretizes a clear vision for the organization's future
2.	Communicate my vision of the organization's future	Communicates a clear vision of the organization's future
3.	Make a continuous effort to generate enthusiasm for the organization's vision	Makes a continuous effort to generate enthusiasm for the organization's vision
4.	Have a clear sense of where I believe our organization should be in 5 years	Has a clear sense of where he/she believes our organization should be in 5 years
5.	Seek to make employees accept common goals for the organization	Seeks to make employees accept common goals for the organization
6.	Strive to get the organization to work together in the direction of the vision	Strives to get the organization to work together in the direction of the vision
7.	Strive to clarify for the employees how they can contribute to achieve the organization's goals	Strives to clarify for the employees how they can contribute to achieve the organization's goals

Note: In the questionnaire organization is replaced by the specific sector organization, e.g. "school" for the school sector. (Likert format): (1 = strongly disagree; 2 = somewhat disagree; 3 = neither agree nor disagree; 4 = somewhat agree; 5 = strongly agree).

Table 2. Operationalization of measurement instrument for transactional leadership

Item no.	Item wording: A. Leader version	Item wording: B. Employee version
	Pecuniary reward – As a leader I...	*Pecuniary reward – My leader ...*
8.	Reward the employees' performance when they live up to my requirements.	Rewards the employees' performance when they live up to the leader's requirements.
9.	Reward the employees' dependent on how well they perform their jobs.	Rewards the employees' dependent on how well they perform their jobs.
10.	Point out what employees will receive if they do what is required.	Points out what employees will receive if they do what is required.
11.	Let employees' effort determine received rewards.	Lets employees' effort determine received rewards.
	Non-pecuniary rewards – As a leader I...	*Non-pecuniary rewards – My leader ...*
12.	Give individual employees positive feedback when they perform well.	Gives individual employees positive feedback when they perform well.
13.	Actively show my appreciation of employees who do their jobs better than expected	Actively shows his/her appreciation of employees who do their jobs better than expected
14.	Generally do not acknowledge individual employees' even though they perform as required (R).	Generally does not acknowledge individual employees' even though they perform as required (R).
15.	Personally compliment employees when they do outstanding work.	Personally compliments employees when they do outstanding work.
	Contingent sanctions – As a leader I...	*Contingent sanctions – My leader ...*
16.	Give negative consequences to the employees if they perform worse than their colleagues.	Gives negative consequences to the employees if they perform worse than their colleagues.
17.	Make sure that it has consequences for the employees if they do not consistently perform as required.	Makes sure that it has consequences for the employees if they do not consistently perform as required.
18.	Take steps to deal with poor performers who do not improve.	Takes steps to deal with poor performers who do not improve.
19.	Give negative consequences to my employees if they do not perform as I require.	Gives negative consequences to his/her employees if they do not perform as the leader requires.

Note: No pretext was offered for item 14.

Table 3. Employee Ratings. Four-Factor versus Alternative Measurement Models of Transformational and Transactional Leadership. Confirmatory Factor Analysis with Standardized Factor Loadings

	Four-Factor Model	Two-Factor Model	Three-Factor Model
Transformational Leadership			
"Concretizes a clear vision for the [ORGANIZATION'S] future"	.797	.797	.799
"Seeks to make employees accept common goals for the [ORGANIZATION]"	.775	.775	.772
"Strives to get the [ORGANIZATION'S] employees to work together in the direction of the vision"	.871	.862	.865
"Strives to clarify for the employees how they can contribute to achieving the [ORGANIZATION'S] goals"	.854	.855	.856
Transactional Leadership: Non-Pecuniary Rewards			
"Gives individual employees positive feedback when they perform well"	.914	.882	.896
"Actively shows his/her appreciation of employees who do their jobs better than expected"	.899	.873	.888
"Personally compliments employees when they do outstanding work"	.932	.900	.914
Transactional Leadership: Pecuniary Rewards			
"Rewards the employees' performance when they live up to his/her requirements"	.907	.837	.850
"Rewards the employees' dependent on how well they perform their jobs"	.880	.802	.818
"Points out what employees will receive if they do what is required"	.750	.668	.678
Transactional Leadership: Sanctions			
"Gives negative consequences to the employees if they perform worse than their colleagues"	.789	.573	.773
"Makes sure that it has consequences for the employees if they do not consistently perform as required"	.878	.708	.877
"Gives negative consequences to employees if they do not perform as he/she requires"	.875	.677	.867
N (Employees)	15,971	15,971	15,971
N (Organizations)	605	605	605
Chi-squared	1006.77	9920.93	6053.99
Degrees of Freedom	59	64	62
Root Mean Squared Error of Approximation (RMSEA)	.032	.098	.078
Comparative Fit Index (CFI)	.956	.539	.720
Standardized Root Mean Square of Approximation (SRMR)	.029	.332	.155

Note: Confirmatory factor analysis based on asymptotic distribution free estimator. All standardized factor loadings are statistically significant at the .001-level.

Table 4: Employee Ratings. Measurement Models of Transformational and Transactional Leadership. Confirmatory Factor Analysis with Standardized Factor Loadings

	Full Sample	Pre	Post	Post (Treatment)	Post (Control)
Transformational Leadership					
"Concretizes a clear vision for the [ORGANIZATION'S] future"	.797	.795	.799	.793	.810
"Seeks to make employees accept common goals for the [ORGANIZATION]"	.775	.774	.776	.776	.779
"Strives to get the [ORGANIZATION'S] employees to work together in the direction of the vision"	.871	.865	.878	.881	.875
"Strives to clarify for the employees how they can contribute to achieving the [ORGANIZATION'S] goals"	.854	.855	.851	.847	.860
Transactional Leadership: Non-Pecuniary Rewards					
"Gives individual employees positive feedback when they perform well"	.914	.912	.918	.918	.922
"Actively shows his/her appreciation of employees who do their jobs better than expected"	.899	.896	.904	.903	.906
"Personally compliments employees when they do outstanding work"	.932	.933	.931	.931	.939
Transactional Leadership: Pecuniary Rewards					
"Rewards the employees' performance when they live up to his/her requirements"	.907	.905	.910	.907	.927
"Rewards the employees' dependent on how well they perform their jobs"	.880	.889	.874	.880	.868
"Points out what employees will receive if they do what is required"	.750	.748	.753	.753	.760
Transactional Leadership: Sanctions					
"Gives negative consequences to the employees if they perform worse than their colleagues"	.789	.787	.795	.799	.801
"Makes sure that it has consequences for the employees if they do not consistently perform as required"	.878	.880	.877	.874	.885
"Gives negative consequences to employees if they do not perform as he/she requires"	.875	.872	.882	.887	.872
N (Employees)	15,971	9,309	6,662	4,866	1,796
N (Organizations)	605	601	460	328	132
Chi-squared	1006.77	613.99	468.19	347.06	182.94
Degrees of Freedom	59	59	59	59	59
Root Mean Squared Error of Approximation (RMSEA)	.032	.032	.032	.032	.034
Comparative Fit Index (CFI)	.956	.957	.953	.953	.951
Standardized Root Mean Square of Approximation (SRMR)	.029	.030	.032	.033	.040

Note: Confirmatory factor analysis based on asymptotic distribution free estimator. All standardized factor loadings are statistically significant at the .001-level.

Table 5. Test for Measurement Invariance Across Multiple Time Periods (2014 and 2015). Employee Ratings.

Model	Comparison	Chi-squared (df)	ΔChi-squared (df)	RMSEA	ΔRMSEA	CFI	ΔCFI	SRMR	ΔSRMR
1.Baseline model, equal parameters		1211.69(163), $p < .001$		.028		.951		.036	
2.Same form model (Configural invariance)	2 vs 1	1082.19(118), $p < .001$	129.50(45), $p < .001$	.032	.004	.955	.004	.031	.005
3.Equal loadings model (Metric invariance)	3 vs 2	1092.44(127), $p < .001$	10.25(9), $p > .1$	.031	.001	.955	.000	.031	.000

Note: CFA based on asymptotic distribution free estimator. N = 15,971. Groups = 2. Values on grouping variable: 0 "Survey 2014" 1 "Survey 2015". Likelihood-Ratio Test performed for comparison of models.

Table 6. Test for Measurement Invariance Across Multiple Groups: Treatment vs. Control Group in Leadership Intervention Study. Employee Ratings.

Model	Comparison	Chi-squared (df)	ΔChi-squared (df)	RMSEA	ΔRMSEA	CFI	ΔCFI	SRMR	ΔSRMR
1.Baseline model, equal parameters		1211.31(163), $p < .001$		.028		.951		.042	
2.Same form model (Configural invariance)	2 vs 1	1083.56(118), $p < .001$	127.75(45), $p < .001$	.032	.004	.955	.004	.030	.012
3.Equal loadings model (Metric invariance)	3 vs 2	1093.35(127), $p < .001$	9.79(9), $p > .1$	.031	.001	.955	.000	.030	.000

Note: CFA based on asymptotic distribution free estimator. N = 15,971. Groups = 2. Values on grouping variable: 0 "Control Group" 1 "Treatment Group". Likelihood-Ratio Test performed for comparison of models.

Table 7. Test for Measurement Invariance Across Multiple Groups: Sectors. Employee Ratings.

Model	Comparison	Chi-squared (df)	ΔChi-squared (df)	RMSEA	ΔRMSEA	CFI	ΔCFI	SRMR	ΔSRMR
1. Baseline model, equal parameters		7287.69(683), $p < .001$		.065		.711		.176	
2. Same form model (Configural invariance)	2 vs 1	1388.79(413), $p < .001$	5898.90(270), $p < .001$	.032	.033	.957	.246	.049	.127
3. Equal loadings model (Metric invariance)	3 vs 2	1550.07(467), $p < .001$	161.28(54), $p < .001$	.032	.000	.953	.004	.056	.007

Note: CFA based on asymptotic distribution free estimator. N = Confirmatory factor analyses based on maximum likelihood estimator. N = 15,971. Groups = 7. Values on grouping variable: 1 "High school" 2 "Public school" 3 "Private school" 4 "Public day care" 5 "Private day care" 6 "Tax office" 7 "Bank branch". Likelihood-Ratio Test performed for comparison of models.

Table 8. Intercorrelations and Estimates for Discriminant Validity and Reliability. Employee Ratings.

	1	2	3	4
1.Transformational Leadership	(.895)/(.681)	.347	.118	.034
2.Contingent Non-pecuniary Rewards	.589 ***	(.937)/(.837)	.188	.019
3.Contingent Pecuniary Rewards	.344 ***	.434 ***	(.877)/(.720)	.079
4.Contingent Sanctions	.184 ***	.137 ***	.281 ***	(.874)/(.719)

Note: *** $p < .001$. N = 15,971. Subdiagonal entries are correlations between latent constructs. Entries above the diagonal are the squared correlation estimates (shared variance). The first entry on the diagonal is Cronbach's alpha for composite reliability. The second entry in the diagonal is the average variance extracted (average of squared factor loadings) for each latent construct. Jöreskog's Rho for reliability is: Transformational leadership = .895, Contingent non-pecuniary rewards = .939, Contingent pecuniary rewards = .884, and Contingent sanctions = .885.