



High Performers = Better Leaders? Evidence From 55 Years of Professional Soccer on the Validity of Performance-based Promotion to Leader Positions

Joyce Elena Schleu¹ · Stefan Krumm² · Alfred Zerres³ · Joachim Hüffmeier⁴

Accepted: 1 June 2023 / Published online: 15 August 2023
© The Author(s) 2023

Abstract

Promoting high-performing employees to leadership positions is a pervasive practice and has high face validity. However, little is known about the actual link between employee and subsequent leader performance as prior results are inconsistent. Given the prevalence of this performance-based promotion strategy, we conducted a study to address this inconsistency. To account for prior diverging results, we (a) competitively tested predictions from different theoretical perspectives (i.e., the performance requirements perspective, the follower-centric perspective, and the Theory of Expert Leadership), (b) considered possible changes in the predictive validity of this strategy over time, and (c) included job complexity as potential moderator of the link between employee and subsequent leader performance. In a high stakes context (i.e., the first German soccer league), we tested the predictive validity of employee performance for leader performance. Our results suggest a low validity of performance-based promotion, as we could not find evidence for a link between employee performance and leader performance—neither initially following the promotion nor over time, which is most in line with the performance requirements perspective. We, thus, caution against the (sole) application of performance-based promotion principles.

Keyword Leader selection · Peter Principle · Performance-based promotion · Employee promotion · Expert leadership · Soccer

“Dr. Peter observed that one reason so many employees are incompetent is that the skills required to get a job often have nothing to do with what is required to do the job itself” (Peter & Hull, 2011, xi).

It is common practice in organizations to promote high-performing employees to leader positions: A recent study revealed that performance is the strongest predictor for subsequent promotion (Church et al., 2021) and therefore has a gatekeeping-function when filling leadership positions (see Gallup, 2014). While this prevalent HR-strategy seems

face-valid (i.e., rewarding high-performing employees with a more senior position; see Kim, 2019), empirical findings on the actual validity of employee performance (EP) for leader performance (LP) have been inconclusive: While some studies reported a positive link between EP and subsequent LP (e.g., Dawson & Dobson, 2002; Goodall & Pogrebná, 2015), others found a negative association between EP and LP (e.g., Benson et al., 2019; Muehlheusser et al., 2018). Hence, the validity of performance-based promotion is unclear which is troubling given its preponderance in current organizational practice (see Church et al., 2015). As invalid selection decisions for leadership positions result in particularly high costs (Brogden, 1949; Cronbach & Gleser, 1965; see also Schmidt & Hunter, 1998), learning about the validity of this performance-based promotion strategy is crucial.

To examine the validity of the prevalent strategy to promote high-performing employees to leader positions, we conducted a study in a professional sports context—the Bundesliga (i.e., Germany’s first soccer league). Specifically, we examined the transition of former professional players to a head coach position in the Bundesliga. The Bundesliga is a relevant occupational context: In the 2020/21 season,

✉ Joyce Elena Schleu
joyce.schleu@ru.nl

¹ Behavioural Science Institute, Radboud University, Nijmegen, Netherlands

² Institute of Psychology, FU Berlin, Berlin, Germany

³ Faculty of Economics and Business, University of Amsterdam, Amsterdam, Netherlands

⁴ Institute of Psychology, TU Dortmund University, Dortmund, Germany

the Bundesliga generated a revenue of €3.5 billion (Statista, n.d.). Moreover, the Bundesliga employs around 127,000 people (McKinsey, 2020) and is a highly visible organization (see Türck, 2019). The many advantages of this context (e.g., clear and standardized rules, relatively high number of performance episodes, and objective performance data) ensure a well-controlled setting to examine organizational research questions (see Gentry et al., 2017; Rogers et al., 2017; Wolfe et al., 2005).

In this study, we consider three theoretical perspectives to potentially explain the validity of performance-based promotion (following Schleu & Hüffmeier, 2021): (1) The performance requirements perspective (see Zaccaro, 2012), (2) the follower-centric perspective (see Uhl-Bien et al., 2014), and (3) the Theory of Expert Leadership (TEL; Goodall & Bäker, 2015). Based on these theoretical perspectives, we derive contrasting hypotheses on the initial predictive validity of EP for LP as well as on potential changes in the predictive validity over time. As a first contribution of our study, we empirically test the contrasting hypotheses derived from the three theoretical perspectives mentioned above. This integrative approach will further the understanding of the predictive validity of performance-based promotion and facilitate an interdisciplinary discussion (e.g., between personnel selection and management research) on the question of whether high-performing employees truly make successful leaders.

To date, the validity of performance-based promotion has not received a lot of attention in organizational research. Even though the number of studies on performance-based promotion increased during the last 20 years, the resulting insights are still limited (see Schleu & Hüffmeier, 2021). While the overall results pattern ranges from positive (e.g., Dawson & Dobson, 2002; Goodall & Pogrebna, 2015) to negative associations between EP and LP (e.g., Benson et al., 2019; Muehlheusser et al., 2018), the majority of prior studies on performance-based promotion found negative, non-significant, or mixed associations between EP and LP (for an overview, see Schleu & Hüffmeier, 2021). Currently, this inconsistency in results cannot be explained. However, a recent review identified two promising moderators to explain the range of results: temporal changes and relevance of performance requirements in employee positions for leader positions (Schleu & Hüffmeier, 2021). As the second contribution, our study thus provides the first systematic empirical test of temporal changes of the validity of EP for LP (following recent calls, e.g., Fischer et al., 2017). More specifically, we tested hypotheses on the predictive validity of EP for LP immediately after the promotion and on temporal changes in predictive validity over time. Learning about potential short-term effects and changes across time could facilitate evidence-based decisions concerning performance-based promotion strategies. Furthermore, we test whether

EP is associated with overall LP (i.e., aggregated over time), which allows an assessment of the legitimacy of using high EP as a gatekeeper to leader positions.

As our third contribution, we examine the relevance of the performance requirements in employee positions (specifically, the complexity of the player positions in soccer) for leader positions as a potential moderator. One may argue that EP is more predictive of LP when the previous employee position mirrors the subsequent performance requirements of a leader position (Zaccaro et al., 2018), which should be the case for more complex employee positions (Hunter et al., 1990). Examining potential boundary conditions for the relation between EP and LP provides the opportunity to identify conditions with higher (and lower) predictive validity to explain the hitherto inconsistent findings, and thus, potentially enable a situation-contingent use of performance-based promotion strategies. As a fourth contribution, we derive practical implications from our results. We thereby point out potential improvements of this pervasive strategy.

Theoretical Background and Hypotheses

Despite the inconsistent evidence, the prevalence of performance-based promotion is high in practice (Benson et al., 2019; Church et al., 2015). In this section, we consider three relevant theoretical perspectives to explain the relationship between EP and LP. The theoretical perspectives yield contrasting hypotheses concerning the predictive validity of EP for initial LP (i.e., immediately following the promotion to a leadership position), changes in predictive validity over time, and the predictive validity of EP for LP aggregated over time. In this study, we conceptualize EP as “behaviors or actions that are relevant to the goals of the organization” (Campbell, 1990; p. 704; see also Rich et al., 2010) and are exhibited in a subordinate position; further, we understand LP as the effect that a leader has on the performance of the led team (see also Fischer et al., 2017).

I. The Performance Requirements Perspective

Following the logic of the performance requirements perspective (see Zaccaro et al., 2018), which relies on individual differences to predict LP, performance-based promotion to leader positions should be a valid strategy only to the extent that the employee and the leader position have matching performance requirements (Asher & Sciarrino, 1974; Zaccaro et al., 2018). General taxonomies for the workforce (e.g., Bartram, 2005) as compared to managerial taxonomies (e.g., Borman & Brush, 1993; Tett et al., 2000), however indicate a low match between employee and leader tasks (Schleu & Hüffmeier, 2021): Many dimensions of the managerial taxonomies are rather specific to leader positions (such as

guiding, directing, coordinating, and motivating subordinates; Borman & Brush, 1993; Tett et al., 2000), indicating a low match of performance requirements. Moreover, leader positions as compared to employee positions are more complex (Hunter et al., 1990) and require “a variety of different activities in carrying out the work, which involve the use of a number of different skills and talents” (Hackman & Oldham, 1975, p. 161).

Further, to compare the performance requirements of employee and leader positions in more detail, it is instructive to focus on commonly accepted predictors of leadership success (for an overview, see Zaccaro et al., 2018), such as personality, job knowledge, and motivation. Concerning personality predictors, the link to leadership success seems to be context-dependent (see De Hoogh et al., 2005) and contingent on employed performance criteria (e.g., rated leader performance versus team performance; for meta-analytic evidence, see DeRue et al., 2011). Taking into account the variability of required personality traits (i.e., for different leader positions), the performance requirements of employee and leader positions concerning personality may match only in some specific contexts. That is, while certain attributes influence EP positively (e.g., conscientiousness or agreeableness) they might not necessarily translate into high LP—but could in some cases even translate into low LP. For instance, a successful, dutiful employee might become an ineffective micromanager or a conflict-avoiding supervisor (see Smith et al., 2018).

Job knowledge acquired as an employee might be helpful when becoming a leader, for instance, for structuring the employees’ work (see Day et al., 2009). At the same time, previous experience might not translate to new positions or contexts (see Salomon & Perkins, 1989; van Iddekinge et al., 2019) and expertise might reduce a leader’s cognitive flexibility facing changes (see Dane, 2010). Furthermore, the need for specific job knowledge acquired in an employee position should decrease with higher hierarchy levels (see Day et al., 2009; Katz & Kahn, 1978). Thus, the performance requirements concerning job knowledge again may match for some contexts (e.g., lower-level leadership positions), but not in general.

While the importance of motivation for performance is obvious and was recently reaffirmed (for meta-analytic evidence, see van Iddekinge et al., 2018), the manifestation of employee motivation could either be stable (i.e., more trait-like) or context-dependent (i.e., more state-like). Consequently, it is unclear to what extent motivation in an employee position transfers to leader positions (i.e., *motivation to lead*; Kark & van Dijk, 2007) due to the different tasks and responsibilities (Porter et al., 2016). To sum up, the performance requirements perspective allows predictions of positive (i.e., for highly overlapping performance requirements), null (i.e., for little overlap of performance

requirements), and negative relations (i.e., if performance requirements of the employee position hinder high performance as a leader) between EP and LP when considering the specific positions. On average, the explanatory value of EP for LP should be low, as the performance requirements of employee and leader positions match only to a limited degree.

II. The Follower-centric Perspective

We further consider the follower-centric perspective to understand potential links between EP and LP (see Steffens et al., 2021; Uhl-Bien et al., 2014). According to this perspective, leadership is a social construction of followers and it focuses on the requirements and processes to convince a team to follow (Lord & Maher, 2002; Lord et al., 1984) or to gain credibility in a team (Kouzes & Posner, 2011). Followers evaluate their leaders (see Lord & Dinh, 2014) based on past experiences and their socialization concerning typical characteristics of leaders (i.e., implicit leadership theories [ILTs]; Phillips & Lord, 1981; Schyns et al., 2005). Based on their experiences, followers make sense of organizational processes (Weick, 1995): Prior high performance as an employee should increase the degree to which leaders are perceived to fulfill ILTs (Ensari & Murphy, 2003; Lord & Maher, 2002). This might facilitate the attribution of positive outcomes to them (e.g., Meindl, 1995).

Leaders who were promoted based on previous EP in particular should be perceived to embody core attributes of the team (i.e., be prototypical; Schleu & Hüffmeier, 2021). As a result, the leader has informative value for the team and might be perceived as a role model (Hogg, 2001; Uhl-Bien et al., 2014). Consequently, followers attribute higher credibility to prototypical leaders and support them more (Uhl-Bien et al., 2014), which should result in more successful leaders (i.e., higher LP; Steffens et al., 2021). Thus, the follower-centric perspective suggests a moderate, positive link between EP and LP.

III. Theory of Expert Leadership

The propositions of TEL (Goodall & Bäker, 2015) emphasize the importance of expert knowledge (i.e., acquired as an employee) to be a good leader, at least for knowledge-intensive organizations. Expert knowledge is assumed to be acquired through technical education, practice, and working experience in a particular sector. Goodall and Bäker (2015) assume that expert knowledge is beneficial for LP due to the following main reasons: First, it is assumed to provide a particularly solid base for decision making (see Goodall & Bäker, 2015). Thus, expert leaders as compared to non-expert leaders profit from representing information holistically and rely on abstract concepts to solve problems

(for a related overview, see Ericsson et al., 2006). Second, expert leaders should have a good understanding of the core workers' values, motivations, and typical challenges, due to a shared knowledge base with these workers. As they have a common background with their subordinates, they are potentially perceived "as being the first among equals" (Goodall & Bäker, 2015, p. 57, note that this aspect is somewhat comparable to being prototypical). This constellation should enable them to create a productivity-enhancing working environment, to set realistic goals, and to evaluate the performance of employees appropriately. Third, Goodall and Bäker (2015) emphasize the advantages of expert leaders in personnel selection due to the homophily principle (cf. Rogers & Bhowmik, 1970): When expert leaders hire new personnel, they are assumed to hire applicants with a high potential (similar to themselves), instead of feeling threatened by capable applicants. This should improve the quality of personnel selection compared to non-expert leaders. In sum, TEL would assume that newly promoted leaders profit from their expert knowledge and, hence, that there is at least a moderate, positive link between EP and LP.

Temporal Changes of Predictive Validity

The three theoretical perspectives outlined above make contrasting predictions concerning the predictive validity of EP for LP initially and over time. According to the performance requirements perspective, EP should on average have limited explanatory value for predicting LP, as the performance requirements of employee and leader positions differ. Consequently, the predictive validity of EP for initial LP (directly following the promotion [at t1]) should be small at most (*Hypothesis 1a*).¹ As we report standardized path coefficients, we used conventions for correlation coefficients as a yardstick (i.e., small: 0.10; moderate: 0.30; large: 0.50; Acock, 2014; Cohen, 1988); thus, the proposed relation in H1a corresponds to a standardized path coefficient ranging from -0.10 to 0.10. Further, we consider the effect of time on the relationship between EP and LP: Based on the performance requirements perspective, LP may change over time due to learning and adaptation to the requirements of the leader position. However, these processes are not necessarily related to previous EP. According to this perspective, the

predictive validity of EP for LP should remain small over time (*Hypothesis 1b*).

Based on the follower-centric perspective, prior high EP should increase the degree, to which leaders are perceived as prototypical and to which they meet ILTs (Ensari & Murphy, 2003; Lord & Maher, 2002). Hence, followers should support promoted leaders more (Uhl-Bien et al., 2014), which should result in more leadership success (i.e., higher LP; Steffens et al., 2021). Thus, the predictive validity of EP for initial LP (directly following the promotion [at t1]) should at least be moderate (*Hypothesis 2a*), corresponding to a standardized path coefficient ≥ 0.30 . Over time, leaders do not only need to embody core attributes of the team to be perceived as prototypical, but also need to prove their prototypicality through their engagement for the team (for instance by empowering team members; for pertinent overviews, see Haslam et al., 2011; Steffens et al., 2014). However, neither the follower-centric perspective nor empirical evidence (that we are aware of) suggest a link between former EP and leaders' engagement for their team. Hence, while the follower-centric perspective proposes a link between EP and leaders' initial perceived prototypicality (and thus initial LP), the effect of EP should diminish over time: Leader engagement for the team becomes more important for leaders to maintain their perceived prototypicality over time, which consequently affects the experienced support by their team, and, thus, their success as a leader (Haslam et al., 2011; Steffens et al., 2014). In addition, neither higher prototypicality nor identity leadership propose a link to the quality of decision making (e.g., developing strategies); thus, over time, the team may also be confronted with bad decisions. Consequently, the predictive validity of EP for LP should decrease over time (*Hypothesis 2b*).

Following the logic of TEL, leaders should benefit from their expert knowledge when getting promoted (e.g., by making better decisions and creating a productivity-enhancing working environment). Thus, the predictive validity of expert knowledge—as indicated through EP—for initial LP (directly following the promotion [at t1]) should at least be moderate (*Hypothesis 3a*), corresponding to a standardized path coefficient ≥ 0.30 . Despite one similar mechanism between the follower-centric perspective and TEL (i.e., having a common background with subordinates is somewhat comparable to being prototypical), both theoretical perspectives suggest different developments of the EP-LP link over time. While we argued above that the follower-centric perspective assumes that mechanisms unrelated to EP will become more relevant over time (i.e., engagement for the team), TEL incorporates mechanisms, which are supposedly related to EP and should unfold over time. That is, expertise acquired as an employee will help leaders, for instance, to hire applicants with high potential and create a productivity-enhancing working environment (Goodall &

¹ Feedback on an earlier version of this manuscript suggested to analyze temporal changes in the validity of EP for LP and to account for the quality of the led team. To address this need, we derived and pre-registered the outlined hypotheses (see <https://aspredicted.org/blind.php?x=cx9yu8>). The wording of the hypotheses differs from the pre-registration due to space limitations for the preregistration. While the predictions did not change, we modified the wording and order of the hypotheses to improve the readability.

Bäker, 2015). With time, the overall effect of the proposed mechanisms thus should increase. According to TEL, the predictive validity of EP for LP should increase over time (*Hypothesis 3b*).

Moderating Effect of Relevance due to Job Complexity

As explained above, the performance requirements of employee and leader positions generally do not match well (cf. Bartram, 2005; Borman & Brush, 1993; Tett et al., 2000). Among other aspects, leader positions are more complex than employee positions (Hunter et al., 1990). Following the logic of the performance requirements perspective, more complex employee positions may approximate the complexity of later leader positions (Hunter et al., 1990). Consequently, (high) EP in complex positions (i.e., high variability in tasks) should indicate the ability to handle various activities and a diverse skill set, which should then predict (high) LP, as this is a relevant leader characteristic. Therefore, the predictive validity of EP for LP should be higher for more complex employee positions (see Quiñones et al., 1995; Tesluk & Jacobs, 1998; *Hypothesis 4*), corresponding to a standardized path coefficient > 0.10 . In comparison, the other two outlined theoretical perspectives are rather mute about the proposed moderation (i.e., complexity of the employee position).

Contrasting Expert Leaders to Non-expert Leaders

To provide another test of the outlined theoretical perspectives, we also derive contrasting predictions on having expertise as an employee versus not having such pertinent expertise. Specifically, we also test whether our previous predictions (i.e., H1a-b, H2a-b, and H3a-b) will hold if experience at a pertinent employee position (yes vs. no) instead of EP is used as a predictor of LP. As the overlap of performance requirements of employee and leader positions is on average rather limited (see Bartram, 2005; Borman & Brush, 1993), the performance requirements perspective predicts the following: Leaders, who have pertinent experience as an employee, should not show better LP than leaders without such experience and this should not change over time (*Hypothesis 1c*).

According to the follower-centric perspective, a leader's previous experience as an employee increases perceived prototypicality and follower performance (Steffens et al., 2021; Uhl-Bien et al., 2014). Experience as an employee, however, should not relate to other aspects of perceived prototypicality (e.g., empowering team members; see Haslam et al., 2011), thus its effect should decrease over time: Leaders, who have pertinent experience as an employee, should initially show better LP than leaders without such experience. This difference should, however, decrease over time (*Hypothesis 2c*).

TEL (Goodall & Bäker, 2015), in contrast, predicts that leaders with employee experience should have a better start (e.g., by creating a productivity-enhancing working environment) and this difference should become even more pronounced over time (e.g., due to the lagged consequences of better personnel selection decisions). Leaders, who have pertinent experience as an employee should, thus, initially show better LP than leaders who have no such experience. This difference should further increase over time (*Hypothesis 3c*).

Method

We tested our hypotheses on performance-based promotion strategies in the context of professional sport (i.e., the Bundesliga). Due to clear and standardized rules in this sport, the relatively high number of performance episodes, objective performance measures, and the reliable documentation of performance data, this context provides a rather controlled setting to examine organizational research questions (see Gentry et al., 2017; Maynard et al., 2017; Rogers et al., 2017; Wolfe et al., 2005). In addition, this context allows examining actual leaders (compared to artificial set-ups in the lab). Hence, this context helps to draw reliable conclusions about the initial predictive validity, changes over time, and the predictive validity aggregated over time. We collected the longitudinal data from the websites www.transfermarkt.de and www.kicker.de. To ensure the quality of the extracted data, two trained research assistants collected the data.²

Participants

We included *all* head coaches (i.e., the population) of clubs competing in Germany's first male soccer league (i.e., Bundesliga). Thus, we went through every season of the Bundesliga from 1963/64 (i.e., its first season) to 2018/19 on the homepage www.transfermarkt.de, identified all head coaches ($N = 407$) of the competing clubs, and extracted their data. We attempted to obtain comprehensive data on their career paths. To do so, we included player and coach performance data from the first and second German soccer league, as well as data from every first and second foreign league included in the Union of European Football Associations (UEFA), if the data were available on the respective websites (for further details,

² Data collected during the revision of this manuscript (i.e., on coaching licenses, relevant university degrees, injuries, and first coaching performances in lower leagues) were collected by one trained research assistant.

please see below).³ About $\frac{3}{4}$ of the coaches were former professional soccer players ($n = 309$; 75.92%). Among those, 222 played in the Bundesliga or a first soccer league of the UEFA, 17 played in Germany's second soccer league, and 70 coaches played only in lower soccer leagues or detailed information about their professional player career could not be obtained. The remaining 98 coaches (24.08%) had no professional player experience. Coaches with professional player experience participated on average in 143.18 matches of the Bundesliga ($SD = 140.97$) during their player careers.

Operationalization and Measures

Employee Performance

Employee performance is defined as “behaviors or actions that are relevant to the goals of the organization” (Campbell, 1990; p. 704) and includes both task performance and contextual performance aspects (see Koopmans et al., 2014; van Scotter et al., 2000). Our two indicators for EP—*number of played matches* and *average player ratings*—reflect this definition and incorporate task and contextual performance.

The number of played matches was obtained for each soccer player (including foreign and lower league matches as described above). Typically, a player only takes part in a match if he is the strongest player available in the team for his position at the point of time, since the Bundesliga (and professional soccer in general) is a highly competitive environment (Frick, 2010). Thus, a player is nominated for a match if he is the best to contribute to the overall goal: to win the game. Moreover, a player will mostly be seen as the strongest for a position if he exceeds others in both, task (e.g., scoring or defending) and contextual (i.e., assisting or cooperating with others, for instance when a striker is helping out on defense) performance dimensions. Finally, prior studies also relied on this indicator to assess player performance (e.g., Hall & Pedace, 2016; Martinez & Caudill, 2013).

Average player ratings reflect the player's individual performance (Baumann et al., 2011) as well as his contribution to the team performance (Frick, 2010). Hence, they

are well in line with our definition of EP and also incorporate task and contextual performance. We relied on player ratings, which are assigned by two expert sport journalists out of a trained, stable expert team from a German Soccer magazine called “Kicker”. The experts evaluate the players' performance in a match based on their impression, while also taking into account the statistics of the match, such as running distance, goals, and assists (according to personal communication with N. Peer, who is responsible for the internal and external communication at Kicker; for more information on the Kicker player ratings, see also Kroemer, 2010). Player ratings are available for each match of the Bundesliga, in which a player participated for at least 30 min. The 6-point-scaling (from 1.0 = very good to 6.0 = insufficient) was inverted prior to the data analysis to ease the interpretation of the results such that high values represent good performance ratings. In our analysis, we included the averaged player rating (over the entire player career).

Employee Experience

We operationalized employee experience via a dummy variable (yes vs. no): All coaches who had a player profile on the respective homepage (i.e., www.transfermarkt.de) and played in (at least) one of the first or second soccer leagues of the UEFA were coded as having employee experience.

Leader Performance

To operationalize LP, we relied on the three indicators *points per game*, *goal ratio*, and *number of coached matches*, which resemble established leadership measures such as team performance and satisfaction with the leader (for an overview, see De Rue et al., 2011). As the coach is responsible for high team performance, we assessed his performance by evaluating team performance directly. In choosing this indicator, we build on research that supports the assumption of a direct link between the behavior of a coach and the teams' performance (e.g., Crust & Lawrence, 2006), which is also confirmed by a recent review (Gammelsäter, 2013). Thus, we operationalized LP via the formal performance criterion in the Bundesliga—the achieved *points per game* (PPG; wins = 3 points, ties = 1 point, and defeats = 0 points). To complement PPG with a more fine-grained measure, we also relied on the *goal ratio* (i.e., the ratio of scored to conceded goals). Previous studies (e.g., Dawson & Dobson, 2002) have operationalized coach performance similarly (i.e., winning percentage while accounting for team quality). We believe that our team performance measures closely resemble this coach performance measure, but are even more fine-grained (i.e., the PPG also account

³ To account for quality differences between leagues we consulted the UEFA Coefficient (retrieved from www.5-jahres-wertung.de). As our analyses are based on a German sample, we set the first German soccer league as a reference. Consequently, we relativized all UEFA Coefficients (x) against the German UEFA Coefficient (z) separately for every season. Thereby, we obtained a weighting factor to put player and coach performance into context (e.g., $\text{Weighted Games} = \text{Number of Games} * x / z$). As there is no official coefficient to weight second soccer leagues relative to first soccer leagues, we weighted all matches of the second league by half of the original weighting factor.

for draws and the goal ratio allows for more nuance regarding the game outcomes).⁴

As a third indicator, we used the *number of coached matches* (including foreign and lower league matches as outlined above). In the highly competitive environment of professional soccer, bad match results are not tolerated, but result in coach succession (Cannella & Rowe, 1995). Hence, the number of coached matches reflects the performance of the coach in relation to the performance expectations of the club: As the decision makers will usually consider available resources and circumstances when evaluating coach performance and deciding on coach succession, the number of coached matches should be a valid measure of LP, specifically the satisfaction with the leader (see DeRue et al., 2011). Prior studies relied on similar measures (e.g., Audas et al., 1999). As outlined above (see Footnote 3), all coach performance measures were weighted to account for quality differences of the soccer leagues in the UEFA.

Further, all operationalizations of coach performance were sampled for three coached seasons.⁵ When we refer to t2 and t3, this represents the second or third season (after the initial season, t1) in the Bundesliga. We use the term “across time” when comparing results of t1 with t2 and t3, but please note that actual time between coached seasons may vary. As the number of available coach performance episodes drops drastically across time (i.e., only few coaches coached many seasons), we restricted our analyses to the first three measurement points (t1–t3) for the longitudinal analyses (i.e., Hypotheses 1–3).

Job Complexity of the Employee Position

To operationalize job complexity of the employee position, we relied on the *Group Structure Model* (Chelladurai, 2013; Chelladurai & Carron, 1977; Grusky, 1963) to capture the complexity of different positions in sports. Across different team sport contexts, player positions vary in their propinquity (i.e., observability and visibility) and their degree of task dependence (Chelladurai, 2013; Chelladurai & Carron, 1977; Grusky, 1963; for an overview, see Carron & Eys, 2012). In this study, we applied the group structure model to the context of soccer. With increasing propinquity and task dependence, a player receives more information about game-related processes and interacts more frequently with

other players (Carron & Eys, 2012). Thus, we considered high manifestations of propinquity and task dependence for a position as operationalization of the complexity of the employee position.

To estimate the job complexity of the different player positions, we invited soccer experts, such as (former) players, coaches, soccer managers, and sport journalists to participate in an online survey. Fifty-nine experts completed our questionnaire. We excluded five participants as they expressed doubts about their expertise. Thus, 54 experts ($M=42.5$ years, 98.1% males) remained in the sample. On average, our experts were engaged in competitive soccer for a duration of $M=18.17$ ($SD=8.92$) years as a player and of $M=6.47$ years ($SD=6.83$) as a coach. After a short introduction to the concepts of propinquity and task dependence, the experts were asked to rate each position in soccer (i.e., goalkeeper, defender, midfielder, and forward; see Table 1). All items were rated on a 4-point scale ranging from 1 (very poor) to 4 (very well). The experts rated the goalkeeper position to be the least complex ($M=2.35$), followed by the defender ($M=2.54$), the forward position ($M=2.9$), and the midfielder position ($M=3.25$; see Table 1). We included the rated job complexity as a moderator in our analyses (see Hypothesis 4). We obtained information on the former player position for almost all coaches with a prior player career ($n=301$, 97.41%; midfielders: 130, defenders: 96, forwards: 62, and goalkeepers: 13). For players who held more than one position, we recorded the most frequently played position (as listed on transfermarkt.de).

Control Variables

We preregistered prior team quality as a central control variable. We further considered the temporal mid-point of a player’s career and coaches’ continued employment as further control variables. Finally, we considered the following variables as potential control variables as they were suggested by anonymous reviewers: player age and player injuries.

The number of played matches (i.e., our first operationalization of EP) might be affected by the number of injuries a player experienced throughout his career. Hence, we gathered respective data, which was available from 1968 onwards. Furthermore, the number of played matches might be influenced by the age of the player when they ended their player career. Hence, we also gathered the respective data. Our second operationalization of EP (i.e., player ratings) decreased on average over time (i.e., worse grades nowadays as compared to 1963). The correlation between date of recorded data (in years) and the mean grades for each year showed a strong link ($r=0.94$, $p<0.001$). Thus, we considered this potentially confounding effect by including the temporal mid-point of a player’s career (i.e., the mean of the years the respective player got graded).

⁴ While we consider our operationalizations to be more fine grained as the win ratio, we acknowledge that the win-ratio is more commonly used in prior studies (e.g., Dawson & Dobson, 2002; Martinez & Caudill, 2013). Thus, we repeated our analyses with this operationalization. Results were consistent with our reported findings.

⁵ If, however, a coach was fired during the season and if he was re-employed by another club in the same season, the coach performance for the second club was gathered as a second measurement point.

Table 1 Experts ($N = 54$) Rated the Complexity of Each Player Position Based on the Group Structure Model (Chelladurai & Carron, 1977; Grusky, 1963)

	Goalkeeper (SD)	Defender (SD)	Midfielder (SD)	Forward (SD)
Propinquity				
How well could a player at the respective position oversee the build-up of the match?	3.26 (0.85)	3.04 (0.70)	3.07 (0.67)	2.37 (0.76)
How central was a player on this position to the game for his team?	2.20 (1.07)	2.48 (0.75)	3.54 (0.61)	2.67 (0.82)
Averaged	2.73 (0.77)	2.76 (0.61)	3.31 (0.54)	2.52 (0.61)
Task Dependence				
To what extent were the players in all other positions dependent on the work of a player in this position?	1.96 (0.99)	2.31 (0.72)	3.20 (0.68)	3.28 (0.98)
Complexity	2.35 (0.77)	2.54 (0.58)	3.25 (0.49)	2.90 (0.62)

In soccer, the team performance is not only influenced by the coach, but also by the team quality. We therefore assessed prior team quality (i.e., the *teams' ranking* in the league) *before* the start of a coach's appointment and included it as a control variable in our analyses (see also Dawson et al., 2000).⁶ To receive a continuous ranking for the first and the second league we transformed the variable *teams' ranking*: After the transformation higher numbers represent a better ranking (i.e., more specifically, values between 2 and 1 represent the ranking for the first league and values between 1 and 0 represent the ranking for the second league). As the number of competing teams in the first and second league changed over time, the relative value of the ranks changed as well (e.g., the 10th rank in a league of 16 teams vs. 20 teams). Hence, this transformation allows for comparisons over time. Furthermore, to account for the effects of a coach's continued employment with a club (i.e., continuous coaching) as compared to a change in employment, we included a dummy variable (i.e., 1 = continued employment since the previous season; 0 = change in employment since the previous season).

Analytical Strategy

As we gathered data on the whole population of head coaches of the Bundesliga between 1963/64 to 2018/19, we examined in a first step the mere effect size of the EP-LP relationship without statistical inference. Note that we nevertheless go beyond mere effect sizes in subsequent steps (Alexander, 2015) since we intend to make inferences

on the EP-LP relationship outside the Bundesliga context and to assess the EP-LP link while accounting for potential control variables. In a second step, we therefore gathered data on potential control variables (i.e., players' injuries, age, the average time of the player rating, team quality [before t1, t2, and t3], and continuous employment [t1-t2 and t2-t3]) and tested their relationship with our indicators of LP (i.e., the dependent variable). Based on the recommendations of Becker (2005), we only added variables as controls to our main analyses if they showed a significant relationship with the dependent variable (i.e., we added team quality [before t1, t2, and t3], continuous employment [t1-t2 and t2-t3] for our main analyses; for bivariate correlations, see Table 2). Next, we analyzed our panel data with manifest path models (including relevant control variables) and conducted six separate analyses for the different operationalizations of player and coach performance.⁷ In particular, we included the hypothesized paths from EP to LP (at t1, t2, and t3) and added complexity and the respective interaction term (i.e., EP × complexity) to our model (see Fig. 1). Furthermore, we included paths from LP at t1 to t2 and from t2 to t3. We also assumed that LP will be further predicted by our control variables (i.e., team quality [before t1, t2, and t3], and continuous employment [t1-t2 and t2-t3]) and specified our model accordingly. We also allowed for covariances (see Fig. 1,

⁶ To demonstrate the robustness of our results we repeated our analyses with the average rank of the last three seasons (instead of only one season) before the coach took over. Results were consistent with our reported findings (see Table S3).

⁷ We initially planned to analyze data with structural equation modeling (SEM). In a preliminary confirmatory factor analysis (CFA), we modeled player performance and coach performance as latent factors for all measurement points. The CFA revealed negative residual variances and was therefore deemed inadmissible. Potentially, different measures of player or coach performance reflect unique performance aspects and not a shared latent variable. Thus, we continued our analyses with manifest path models.

and LP were in a similar range (t2: $r_s \geq -.03$ to $r_s \leq .18$, with an average correlation of $r = .09$; t3: $r_s \geq -.06$ to $r_s \leq .07$, with an average correlation of $r = -.01$). Across t2 and t3, only three out of 12 coefficients exceeded the threshold for small correlations (for all correlations, see Table 2). Both the complexity of the player position ($r_s \geq -.06$ to $r_s \leq .07$, with an average correlation of $r = .00$), as well as the interaction-terms of complexity and EP indicated at most a small to moderate relation ($r_s \geq -.12$ to $r_s \leq .17$, with an average correlation of $r = .04$) with LP initially and over time (see Table 2). Finally, we also looked at the relationships between player experience and LP, initially ($r_s \geq -.04$ to $r_s \leq .07$, with an average correlation of $r = .02$) and over time (t2: $r_s \geq -.04$ to $r_s \leq -.01$, with an average correlation of $r = -.03$; t3: $r_s \geq -.04$ to $r_s \leq -.01$, with an average correlation of $r = -.03$).

Overall, these correlations are mostly in line with the performance requirements perspective (Hypothesis 1a-c). As the results of the different analyses include some variance (see Table 2), we will rely on the following, more complex analyses (see the next paragraph) to draw firmer conclusions. In particular, we will include control variables (e.g., prior team quality of the coached team) to provide more robust tests.

Manifest Path Models

We tested six different path models (for model fit statistics, see Table 3) to realize all combinations of different player performance measures (i.e., number of played matches vs. average player rating) with coach performance measures (i.e., PPG, goal ratio, and the number of coached matches).⁸ As our analyses revealed a consistent pattern of results (for more information see Table 3), we exemplarily describe the findings of the first model in more detail: The first model (see Fig. 1) indicated no significant link between the number of played matches and coach performance (i.e., operationalized as PPG) directly after the promotion to the coach position (t1: $\beta = -0.00$, $SE = 0.06$, $p = .986$). Over time, the relationship remained non-significant (t2: $\beta = 0.02$, $SE = 0.06$, $p = .965$; t3: $\beta = -0.04$, $SE = 0.06$, $p = .810$). To test for changes in the predictive validity, we restricted the path coefficients of player performance to be equal over time and conducted a χ^2 difference test between restricted and unrestricted models (see Table S1 in the supplemental file for all restricted models). The results indicated no difference ($\Delta\chi^2 = 0.48$, $\Delta df = 2$, $p = .788$). The path models with other combination of EP and LP indicators (see Table 3 and Table S1, Models 2–6) mirror those findings. Overall, the results pattern, thus, supports Hypothesis 1a and 1b, as the relationship between player performance and coach performance was at most small and non-significant initially (at t1)

⁸ To test the robustness of our results, we also conducted hierarchical regression analyses for all operationalizations and time points. The results remain virtually unchanged to the reported path analyses.

as well as over time (at t2 and t3). Further, our results did not indicate a significant change of the relationship between player and coach performance over time.

Further, we also tested the relation between player performance and overall coach performance (i.e., aggregated over the whole coach career) with six models without the FIML estimator (for model fit statistics, see Table 4). Model 1–6 indicated no significant link between player performance (operationalized via the number of played matches or the average player rating) and coach performance (operationalized via PPG, goal ratio, or the number of coached matches). Overall, those aggregated analyses indicate a small-to-moderate, non-significant relation between player and coach performance.

Testing the Moderating Effect of Job Complexity

We inspected the hypothesized interaction of player performance and job complexity in our prior analyses (see Table 3 and 4). The main effects of job complexity (except for one analysis; see Table 4) and the interaction effects (i.e., player performance \times job complexity) were insignificant, both initially and over time. Thus, our analyses did not support Hypothesis 4.

Contrasting Expert Leaders With Non-expert Leaders

Finally, we tested the link between player experience (professional soccer player: yes vs. no) and all three operationalizations of coach performance over time with three models (for model fit statistics, see Table 5). The first model did not indicate a significant link between player experience and coach performance (i.e., operationalized via PPG) initially after the promotion to the coach position and over time (t1: $\beta = 0.07$, $SE = 0.11$, $p = .137$; t2: $\beta = 0.02$, $SE = 0.12$, $p = .756$; t3: $\beta = -0.03$, $SE = 0.11$, $p = .756$). To test for changes in the predictive validity, we restricted the path coefficients of player experience to be equal over time and conducted χ^2 difference tests. The results indicated no difference ($\Delta\chi^2 = 1.98$, $\Delta df = 2$, $p = .371$). The second model indicated a small significant relationship between player experience and coach performance (i.e., operationalized via goal ratio) directly following the promotion to the coach position, t1: $\beta = 0.09$, $SE = 0.12$, $p = .038$. Over time, the relationship became non-significant (t2: $\beta = 0.03$, $SE = 0.09$, $p = .591$; t3: $\beta = -0.05$, $SE = 0.12$, $p = .433$). Testing for changes in the predictive validity, the results indicated no change ($\Delta\chi^2 = 4.27$, $\Delta df = 2$, $p = .118$). The third model neither indicated a significant relationship between player experience and the number of coached matches directly after the promotion to the coach position (t1: $\beta = -0.03$, $SE = 0.10$, $p = .561$) nor later (t2: $\beta = 0.08$, $SE = 0.11$, $p = .136$; t3: $\beta = 0.01$, $SE = 0.13$, $p = .819$). Accordingly, the predictive validity did not change over time ($\Delta\chi^2 = 3.02$, $\Delta df = 2$, $p = .221$). Overall, those aggregated analyses indicate support for Hypothesis 1c.

Robustness Checks and Control Analyses

We conducted various robustness checks and additional analyses to either study the influence of our methodological choices (e.g., operationalizations or treatment of outliers) or of additional variables (e.g., different career trajectories). We report these analyses in the following.

Triangulation as an Approach to Control for the Robustness of our Results

Hypotheses 1a-c as derived from the performance requirements perspective come close to hypothesizing a null effect. We therefore followed Cortina and Folger's (1998) suggestions to rule out potential alternative explanations of potentially observing a null effect (e.g., invalid operationalizations). In particular, we incorporated the triangulation procedure recommended by the authors as we, first, included various measurements for both EP and LP across different situations (i.e., initially following promotion, over time, and aggregated over the whole career). Doing so increases the chance that the observed association differs from a null effect—especially if liberal p -values (i.e., $p=0.1$) are adopted. Even after adopting liberal p -values,⁹ we did not observe a relationship between EP and LP initially and over time. Second, we investigated additional variables that should—from a theoretical point of view—affect the dependent variables (i.e., prior team quality for the dependent variables points per game and goal ratio; prior coach performance [the previous number of coached matches, indicating satisfaction with the leader] for the current number of coached games), which overall showed the expected substantial relationships (see Table 3). Finally, we calculated and reported effect sizes and confidence intervals to facilitate interpretation (see Cortina & Folger, 1998). Overall, results obtained from the triangulation procedure did not indicate that our previously observed null results are invalid.

Outliers

Although we relied on mostly objective data, our findings may be influenced by a few unorthodox data points. Thus, we carefully screened the data for outliers (on the basis of Mahalanobis distances; e.g., Meade & Craig, 2012). Since this method cannot handle missing data, we additionally identified univariate outliers via boxplots to follow a conservative approach. Across the different

operationalizations and time points, we detected 24 outliers for the coaches' points per game, 39 outliers for the coaches' goal ratio, and four outliers for the number of coached matches. We repeated our main analyses while excluding those outliers and found very similar results (see Table S2 in the supplemental file).

Additional Facets of the Coaches' Expertise

During the revision of the paper,¹⁰ we gathered additional data to capture further aspects of expert leadership beyond prior player performance (i.e., university degrees in relevant fields, such as sports sciences, and potential coaching licenses). Those variables have the potential to capture further facets of expertise and, thus, offer a more comprehensive test of TEL. Remarkably, only 47 coaches had a sports-related degree and former players (compared to coaches without that background) were significantly less likely to have obtained such a degree, $\chi^2(1, N=186)=13.88$, $p<.001$. Concerning the coaching license, only $n=19$ coaches of our sample did not have a coaching license, whereas $n=18$ were awarded with an A-license, $n=4$ with a B-license, and $n=240$ with a Pro-license (with the Pro-license being the most advanced license, which is nowadays a requirement in the Bundesliga). We found no data on coaching licenses for $n=127$ coaches. Our analyses overall did not indicate a systematic link between sports-related degrees or coaching licenses and LP, neither initially nor over time (see Table 2, for correlations and Table S4 in the supplemental file). That is, only two of the 18 empirical links between sports-related degrees and LP were significant (see Table S4).

Different Career Trajectories

Some coaches started their coaching career as head coaches of a club in a first or second professional league within the UEFA, while others started their career in lower leagues or with youth teams. For the majority of the head coaches ($n=258$; 63.39%), the head coach position in a first or second professional league within the UEFA (as a permanent or interim coach) was indeed their first coaching position, compared to $n=144$ (35.38%) who first worked as a coach in a lower league or with youth teams, and $n=5$ (1.23%) coaches with missing data. We tested the link between EP and LP for both groups of coaches (i.e., those who started their coaching career in a first or second professional league within the UEFA and those who started their coaching career in lower leagues or with youth teams). Due to the reduced sample size (i.e., only $n=101$ of those 144 coaches

⁹ Please note that we previously adjusted the p -values for each model through Benjamini and Hochberg's (1995) step-up Bonferroni procedure as recommended for structural equation models (Cribbie, 2007) to account for the alpha inflation associated with multiple tests.

¹⁰ We thank the anonymous Reviewer 1 for the related hint.

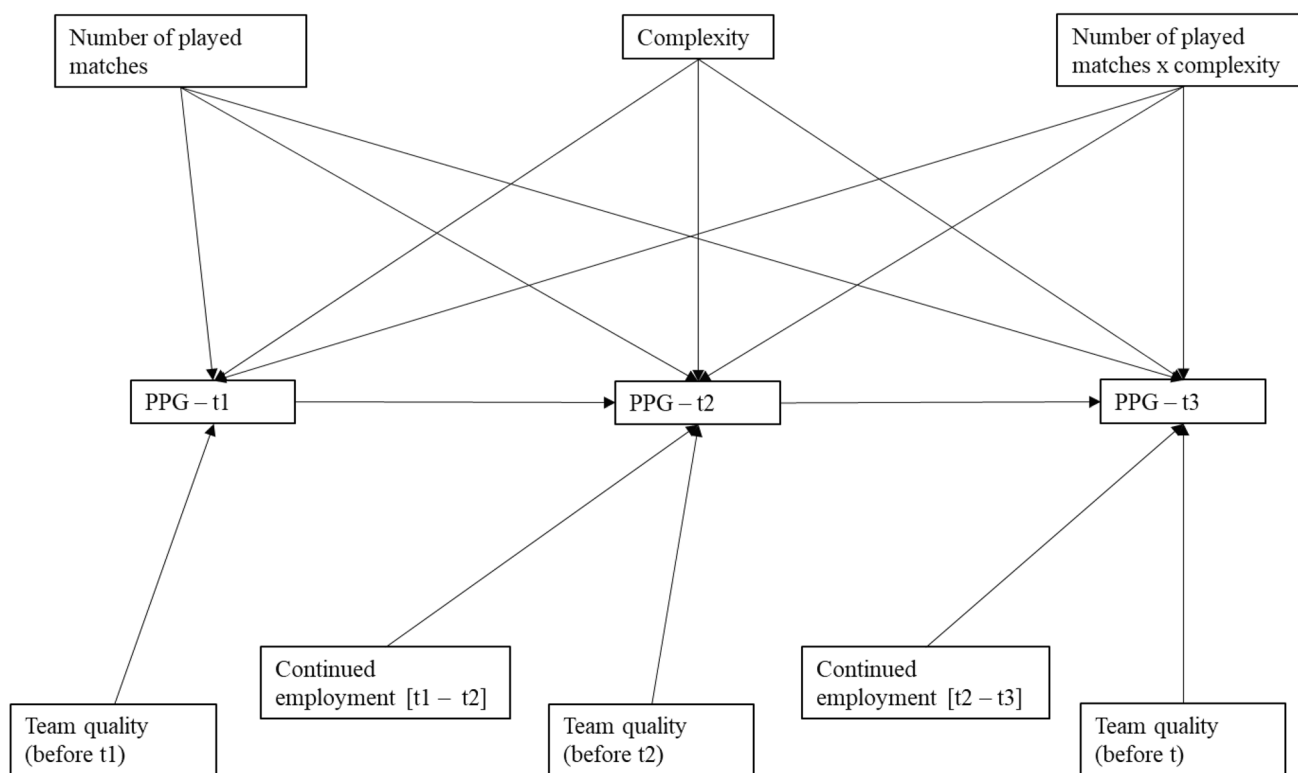


Fig. 1 Visualization of Model 1 (testing Hypotheses 1–4)

for illustrations of our model specifications in model 1). In a third step, we conducted additional analyses to check for the robustness of our results (please see below).

When analyzing the link between EP and LP (Hypotheses 1–4), we only considered coaches who had been soccer players before ($n = 309$). Similarly, we included only players in our analyses who were graded ($n = 181$) for the respective analyses (i.e., to ensure there is no missing data concerning the predictor variable; in this case, missing data for LP or control variables (e.g., prior team quality) were estimated using the FIML estimator; Enders, 2010). To test for changes in predictive validity of player performance for coach performance over time, we relied on model comparisons: We computed a restricted model where we restricted the path coefficients of the player performance to be equal over time and compared this model to the original model (i.e., with free path coefficients). Then we conducted χ^2 difference tests to investigate whether the restriction impaired the model fit. Furthermore, we also conducted analyses considering the aggregated career of the coach (i.e., by considering the aggregated coach performance) to test for the overall effect of performance-based promotion.

To account for different metrics, we standardized the predictor and criterion variables. More specifically, we standardized the coach performance from all measurement points simultaneously to allow for comparisons over time. As effect

size measure, we report standardized path coefficients and used the following yardstick for their interpretation: small effects correspond to path coefficients of 0.10, moderate effects to coefficients of 0.30, and large effects to coefficients of 0.50 (Acock, 2014; Cohen, 1988). To correct for alpha inflation associated with multiple tests, we adjusted the p -values for each model through Benjamini and Hochberg's (1995) step-up Bonferroni procedure as recommended for structural equation models (Cribbie, 2007). In the following, we report p -values that are adjusted in accordance with this procedure.

Results

Descriptive Analyses on the EP-LP Relationship

In a first step, we examined the link of the EP-LP relationship in the population of Bundesliga coaches without statistical inference (and without considering any control variables). Bivariate correlations indicated at most a small to moderate link between EP and LP, both initially (t1) and over time (t2 and t3). At t1, the correlations between EP and LP ranged from $r_s \geq .01$ to $r_s \leq .15$, with an average correlation of $r = .06$ and only one of six coefficients exceeding the threshold for small correlations. At t2 and t3, the correlations between EP

Table 2 Means, Standard Deviations, and Correlations with Confidence Intervals

Variable	<i>M</i>	<i>SD</i>	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Number of played matches	192.56	144.89														
2. Average player rating	3.94	0.46	.02													
3. Player Experience	0.76	0.43	NA	NA												
4. Complexity	2.91	0.33	-.06	-.25**	NA											
5. Interaction number of played matches×Complexity	557.58	420.94	-.10	-.05	NA	.01										
6. Interaction average player rating×Complexity	11.32	1.62	.04	-.18*	NA	.10	.03									
7. PPG t1	1.11	0.65	.02	.15	.04	-.06	-.03	-.12								
8. PPG t2	1.02	0.60	.03	.16*	-.04	-.06	.07	-.05	.37**							
9. PPG t3	0.99	0.59	-.00	-.06	-.03	.04	.15*	-.01	.25**	.40**						
10. Goal ratio t1	0.91	0.91	.04	.01	.07	.03	.04	-.11	.68**	.31**	.18**					
11. Goal ratio t2	0.83	0.57	.01	.16*	-.03	-.03	.05	-.07	.42**	.89**	.42**	.39**				
12. Goal ratio t3	0.83	0.63	-.04	.07	-.04	.06	.16*	-.01	.26**	.40**	.86**	.21**	.48**			
13. Number of coached matches t1	14.30	11.21	.05	.10	-.04	.07	.04	.13	.23**	.26**	.22**	.09	.25**	.21**		
14. Number of coached matches t2	17.67	12.43	-.03	.18*	-.01	-.03	.17*	.11	.28**	.53**	.36**	.13*	.49**	.23**	.48**	
15. Number of coached matches t3	17.93	13.08	-.03	-.02	-.01	.00	.15*	.01	.19**	.41**	.61**	.10	.38**	.45**	.31**	.50**
16. Overall PPG	1.01	0.42	-.02	.18*	.03	-.02	.02	-.11	.61**	.58**	.58**	.36**	.63**	.58**	.31**	.46**
17. Overall goal ratio	0.78	0.45	-.02	.17*	.05	.06	.05	-.09	.50**	.49**	.54**	.43**	.66**	.65**	.33**	.40**
18. Overall number of coached matches	134.29	157.80	-.06	.06	.12*	.13*	.11	.04	.14**	.24**	.39**	.16**	.29**	.33**	.27**	.34**
19. Initial PPG at lower leagues	1.46	0.64	.19	-.06	.10	-.01	-.17	.19	-.16	-.09	.08	-.09	-.04	-.02	.19	-.01
20. Initial goal ratio at lower leagues	1.38	0.95	.17	.12	.06	-.05	-.15	.09	-.03	.09	.23*	-.01	.15	.12	.26**	.09
21. Initial number of coached matches at lower leagues	20.40	11.95	.07	-.14	.08	.07	.09	-.04	-.13	.08	.18	.03	.07	.04	.15	.15
22. Coach license	0.93	0.25	-.01	-.41**	.07	.22**	-.06	.36**	-.05	-.01	-.01	.01	-.03	.01	.01	.03
23. University degree	0.25	0.44	-.32**	-.12	-.29**	-.02	.08	-.07	.02	.19*	-.01	.04	.11	.02	.03	.04
24. Player age	31.58	3.79	.34**	-.68**	NA	.04	-.04	.13	-.08	-.15	-.09	-.01	-.17*	-.10	-.09	-.21**
25. Injuries	4.16	12.06	.06	-.23*	.06	-.01	.00	.13	.09	-.02	.05	.00	-.03	.03	.03	-.05
26. Team quality (before t1)	1.21	0.44	.06	.25**	-.07	-.12	-.16*	-.16	.43**	.19**	.18*	.32**	.26**	.21**	.22**	.05
27. Team quality (before t2)	1.14	0.48	-.02	.30**	-.14*	-.06	-.03	-.08	.40**	.39**	.37**	.31**	.45**	.38**	.28**	.30**
28. Team quality (before t3)	1.13	0.57	.13	.12	-.06	-.10	-.03	-.02	.18*	.40**	.45**	.08	.40**	.40**	.18*	.34**
29. Continued employment (t1-t2)	0.56	0.50	-.06	-.00	.02	-.05	.13	.15	.18**	.22**	.20**	.10	.19**	.14*	.39**	.45**
30. Continued employment (t2-t3)	0.55	0.50	.05	.05	-.13*	-.01	.06	.06	.11	.29**	.11	.04	.29**	.07	-.01	.38**
31. Average time of player rating	1983.23	11.70	.19**	-.72**	NA	.06	-.05	.15*	-.09	-.15	-.07	-.01	-.17*	-.10	-.09	-.19*

Table 2 (continued)

Variable	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
1. Number of played matches																
2. Average player rating																
3. Player Experience																
4. Complexity																
5. Interaction number of played matches × Complexity																
6. Interaction average player rating × Complexity																
7. PPG t1																
8. PPG t2																
9. PPG t3																
10. Goal ratio t1																
11. Goal ratio t2																
12. Goal ratio t3																
13. Number of coached matches t1																
14. Number of coached matches t2																
15. Number of coached matches t3																
16. Overall PPG	.45**															
17. Overall goal ratio	.38**	.87**														
18. Overall number of coached matches	.39**	.41**	.52**													
19. Initial PPG at lower leagues	.03	-.00	.02	.12												
20. Initial goal ratio at lower leagues	.11	.07	.07	.14	.80**											
21. Initial number of coached matches at lower leagues	.22	.11	.27**	.23*	.39**	.28**										
22. Coach license	.09	-.07	-.02	.11	.11	.05	.21									
23. University degree	.00	.10	.10	-.02	-.19	-.22	-.08	.07								
24. Player age	-.02	-.10	-.15*	-.22**	.14	-.09	.02	.24**	-.01							
25. Injuries	.04	.03	-.05	-.04	.09	.01	-.25	.11	-.17	.42**						
26. Team quality (before t1)	.06	.25**	.24**	.02	-.12	.01	-.23*	-.10	-.03	-.16	-.01					
27. Team quality (before t2)	.22**	.41**	.39**	.14*	.07	.18	-.05	-.09	.04	-.23*	.02	.60**				
28. Team quality (before t3)	.32**	.39**	.32**	.14	.03	.14	-.06	-.08	-.08	-.07	.07	.34**	.66**			
29. Continued employment (t1-t2)	.26**	.22**	.20**	.25**	-.13	-.00	-.06	-.02	.15	-.07	.06	-.00	.11	.05		
30. Continued employment (t2-t3)	.36**	.19**	.15**	.14*	-.02	-.04	.05	.16*	.06	-.07	-.07	-.01	-.04	.08	.05	
31. Average time of player rating	-.02	-.11	-.14*	-.22**	.04	-.15	-.07	.26**	.07	.98**	.43**	-.15	-.21*	-.09	-.06	-.07

M and *SD* are used to represent mean and standard deviation (not standardized), respectively. PPG = points per game. Variable 3, 22, 23, 29, and 30 are dichotomous variables (i.e., 1 = yes, 0 = no) and respective correlations are point-biserial. All variables were z-standardized. * indicates $p < .05$. ** indicates $p < .01$

Table 3 Model 1-6 Testing Hypotheses 1–4

Model 1 - Criterion: PPG

Model fit: $\chi^2 = 49.94, df = 29, p = .009, CFI = 0.94, RMSEA = .05, SRMR = .05, n = 309$

	t1			t2			t3		
	β	SE	95% CI	β	SE	95% CI	β	SE	95% CI
Number of played matches	-0.00	0.06	[-0.12; 0.12]	0.02	0.06	[-0.10; 0.14]	-0.04	0.06	[-0.15; 0.09]
Complexity	-0.01	0.06	[-0.12; 0.10]	-0.00	0.05	[-0.10; 0.09]	0.06	0.06	[-0.05; 0.18]
Number of played matches x Complexity	-0.00	0.07	[-0.14; 0.13]	0.05	0.06	[-0.06; 0.16]	0.13	0.06	[0.01; 0.25]
Team quality (before t1)	0.47***	0.06	[0.39; 0.61]						
Team quality (before t2)				0.35***	0.06	[0.22; 0.44]			
Team quality (before t3)							0.40*	0.13	[0.17; 0.66]
Leader performance (t1)				0.19*	0.06	[0.05; 0.29]			
Continuous engagement with club (t1-t2)				0.16*	0.10	[0.08; 0.49]			
Leader performance (t2)							0.20	0.11	[-0.01; 0.41]
Continuous engagement with club (t2-t3)							0.04	0.12	[-0.15; 0.30]

Model 2 - Criterion: Goal ratio

Model fit: $\chi^2 = 37.18, df = 30, p = .172, CFI = 0.98, RMSEA = .03, SRMR = .05, n = 309$

	t1			t2			t3		
	β	SE	95% CI	β	SE	95% CI	β	SE	95% CI
Number of played matches	0.04	0.09	[-0.12; 0.24]	-0.02	0.04	[-0.09; 0.07]	-0.05	0.05	[-0.14; 0.05]
Complexity	0.06	0.08	[-0.05; 0.26]	0.00	0.04	[-0.07; 0.08]	0.07	0.05	[-0.03; 0.15]
Number of played matches x Complexity	0.07	0.10	[-0.07; 0.29]	0.01	0.04	[-0.06; 0.10]	<i>0.14°</i>	0.05	[0.02; 0.21]
Team quality (before t1)	0.34***	0.09	[0.30; 0.67]						
Team quality (before t2)				0.44***	0.05	[0.20; 0.41]			
Team quality (before t3)							0.38*	0.11	[0.12; 0.54]
Leader performance (t1)				0.25	0.07	[0.08; 0.36]			
Continuous engagement with club (t1-t2)				<i>0.12°</i>	0.08	[0.01; 0.32]			
Leader performance (t2)							<i>0.33°</i>	0.15	[0.08; 0.66]
Continuous engagement with club (t2-t3)							-0.01	0.09	[-0.19; 0.15]

Table 3 (continued)

Model 3 - Criterion: Number of coached matches

Model fit: $\chi^2 = 38.40$, $df = 29$, $p = .114$, $CFI = 0.98$, $RMSEA = .03$, $SRMR = .06$, $n = 309$

	t1			t2			t3		
	β	SE	95% CI	β	SE	95% CI	β	SE	95% CI
Number of played matches	0.07	0.06	[-0.06; 0.18]	-0.01	0.08	[-0.14; 0.18]	-0.06	0.07	[-0.20; 0.08]
Complexity	0.11	0.05	[0.00; 0.19]	0.02	0.07	[-0.09; 0.19]	0.04	0.07	[-0.08; 0.17]
Number of played matches x Complexity	0.03	0.06	[-0.09; 0.14]	0.12	0.07	[-0.01; 0.25]	0.06	0.08	[-0.09; 0.21]
Team quality (before t1)	0.23***	0.05	[0.11; 0.31]						
Team quality (before t2)				0.28**	0.08	[0.14; 0.44]			
Team quality (before t3)							0.18	0.10	[0.03; 0.43]
Leader performance (t1)				-0.16	0.48	[-1.31; 0.52]			
Continuous engagement with club (t1-t2)				0.47*	0.32	[0.45; 1.71]			
Leader performance (t2)							0.56***	0.14	[0.34; 0.88]
Continuous engagement with club (t2-t3)							0.12	0.16	[-0.05; 0.57]

Model 4 - Criterion: PPG

Model fit: $\chi^2 = 37.95$, $df = 28$, $p = .099$, $CFI = 0.96$, $RMSEA = .04$, $SRMR = .06$, $n = 181$

	t1			t2			t3		
	β	SE	95% CI	β	SE	95% CI	β	SE	95% CI
Average player rating	0.02	0.09	[-0.16; 0.18]	0.04	0.08	[-0.11; 0.18]	-0.17	0.11	[-0.36; 0.05]
Complexity	0.02	0.07	[-0.12; 0.16]	0.03	0.06	[-0.10; 0.15]	0.05	0.08	[-0.11; 0.19]
Average player rating x Complexity	-0.09	0.07	[-0.22; 0.05]	-0.11	0.07	[-0.24; 0.03]	0.01	0.08	[-0.14; 0.16]
Team quality (before t1)	0.43***	0.08	[0.28; 0.58]						
Team quality (before t2)				0.46***	0.07	[0.26; 0.53]			
Team quality (before t3)							0.26	0.13	[0.06; 0.55]
Leader performance (t1)				-0.21	0.15	[-0.51; 0.07]			
Continuous engagement with club (t1-t2)				0.29**	0.14	[0.24; 0.79]			
Leader performance (t2)							0.68*	0.26	[0.14; 1.15]
Continuous engagement with club (t2-t3)							-0.14	0.22	[-0.67; 0.15]

Table 3 (continued)

Model 5 - Criterion: Goal ratio									
Model fit: $\chi^2 = 39.11$, $df = 29$, $p = .099$, $CFI = 0.95$, $RMSEA = .04$, $SRMR = .06$, $n = 181$									
	t1			t2			t3		
	β	SE	95% CI	β	SE	95% CI	β	SE	95% CI
Average player rating	-0.06	0.15	[-0.43; 0.13]	0.04	0.04	[-0.07; 0.10]	-0.00	0.04	[-0.08; 0.09]
Complexity	0.10	0.11	[-0.04; 0.39]	-0.02	0.04	[-0.09; 0.07]	0.06	0.04	[-0.05; 0.12]
Average player rating x Complexity	-0.10	0.11	[-0.38; 0.03]	-0.05	0.04	[-0.10; 0.06]	0.04	0.04	[-0.06; 0.11]
Team quality (before t1)	0.30*	0.16	[0.22; 0.82]						
Team quality (before t2)				0.34***	0.05	[0.12; 0.29]			
Team quality (before t3)							0.29	0.08	[0.05; 0.33]
Leader performance (t1)				0.29	0.08	[0.07; 0.36]			
Continuous engagement with club (t1-t2)				<i>0.17°</i>	0.09	[0.01; 0.35]			
Leader performance (t2)							0.40***	0.09	[0.19; 0.54]
Continuous engagement with club (t2-t3)							0.01	0.08	[-0.15; 0.17]
Model 6 – Criterion: Number of coached matches									
Model fit: $\chi^2 = 31.77$, $df = 28$, $p = .284$, $CFI = 0.98$, $RMSEA = .03$, $SRMR = .06$, $n = 181$									
	t1			t2			t3		
	β	SE	95% CI	β	SE	95% CI	β	SE	95% CI
Average player rating	0.09	0.06	[-0.04; 0.20]	0.07	0.07	[-0.06; 0.20]	-0.11	0.08	[-0.27; 0.05]
Complexity	0.05	0.06	[-0.08; 0.15]	-0.04	0.06	[-0.16; 0.08]	0.05	0.08	[-0.11; 0.21]
Average player rating x Complexity	0.13	0.06	[0.00; 0.22]	0.02	0.05	[-0.07; 0.11]	-0.07	0.08	[-0.21; 0.09]
Team quality (before t1)	0.18*	0.06	[0.04; 0.26]						
Team quality (before t2)				<i>0.16°</i>	0.07	[0.01; 0.29]			
Team quality (before t3)							0.15	0.12	[-0.01; 0.42]
Leader performance (t1)				0.28***	0.08	[0.18; 0.48]			
Continuous engagement with club (t1-t2)				0.36***	0.13	[0.43; 0.94]			
Leader performance (t2)							0.39***	0.09	[0.22; 0.56]
Continuous engagement with club (t2-t3)							0.22*	0.16	[0.10; 0.74]

All p -Values were adjusted with Benjamini and Hochberg's (1995) false discovery rate controlling step-up Bonferroni procedure as recommended for structural equation models. According to recommendations for triangulation (Cortina & Folger, 1998), we adopted liberal p -values of $p \leq .10$ (emphasized in italics).

° $p \leq .10$; * $p \leq .05$; ** $p \leq .01$; *** $p \leq .001$

Table 4 Model 1–6 Testing the Predictive Validity of Player Performance for Aggregated Coach Performance (i.e., Aggregated Over the Entire Career)

	β	SE	95% CI
Model 1 – Criterion: PPG			
Model fit: $\chi^2 = 0.67, df = 2, p = .714, CFI = 1.00, RMSEA = .00, SRMR = .02, n = 244$			
Number of played matches	-0.02	0.07	[-0.15; 0.12]
Complexity	-0.01	0.07	[-0.14; 0.13]
Number of played matches x Complexity	0.02	0.07	[-0.12; 0.15]
Model 2 – Criterion: Goal ratio			
Model fit: $\chi^2 = 0.84, df = 2, p = .657, CFI = 1.00, RMSEA = .00, SRMR = .02, n = 247$			
Number of played matches	-0.01	0.07	[-0.14; 0.12]
Complexity	0.06	0.07	[-0.07; 0.20]
Number of played matches x Complexity	0.05	0.07	[-0.08; 0.18]
Model 3 – Criterion: Number of coached matches			
Model fit: $\chi^2 = 0.84, df = 2, p = .657, CFI = 1.00, RMSEA = .00, SRMR = .02, n = 247$			
Number of played matches	-0.05	0.06	[-0.18; 0.08]
Complexity	0.12°	0.06	[-0.01; 0.25]
Number of played matches x Complexity	0.10°	0.06	[-0.01; 0.22]
Model 4 – Criterion: PPG			
Model fit: $\chi^2 = 2.05, df = 1, p = .152, CFI = 0.95, RMSEA = .08, SRMR = .04, n = 178$			
Average player rating	0.17	0.07	[0.01; 0.29]
Complexity	0.04	0.06	[-0.09; 0.15]
Average player rating x Complexity	-0.09	0.07	[-0.22; 0.06]
Model 5 – Criterion: Goal ratio			
Model fit: $\chi^2 = 1.85, df = 1, p = .174, CFI = 0.96, RMSEA = .07, SRMR = .03, n = 181$			
Average player rating	0.18	0.07	[0.02; 0.27]
Complexity	0.09	0.06	[-0.04; 0.18]
Average player rating x Complexity	-0.07	0.06	[-0.17; 0.07]
Model 6 – Criterion: Number of coached matches			
Model fit: $\chi^2 = 1.85, df = 1, p = .174, CFI = 0.96, RMSEA = .07, SRMR = .03, n = 181$			
Average player rating	0.12	0.07	[-0.03; 0.25]
Complexity	0.18*	0.06	[0.04; 0.28]
Average player rating x Complexity	0.05	0.06	[-0.08; 0.17]

All *p*-Values were adjusted with Benjamini and Hochberg's (1995) false discovery rate controlling step-up Bonferroni procedure as recommended for structural equation models. According to recommendations for triangulation (Cortina & Folger, 1998), we adopted liberal *p*-values of $p \leq .10$ (emphasized in italics)

° $p \leq .10$; * $p \leq .05$

Table 5 (Continued)

Model 3 – Criterion: Number of coached matches

Model fit: $\chi^2 = 47.66$, $df = 19$, $p \leq .001$, $CFI = 0.95$, $RMSEA = .06$, $SRMR = .06$, $n = 407$

	t1			t2			t3		
	β	SE	95% CI	β	SE	95% CI	β	SE	95% CI
Player experience (dummy)	-0.03	0.10	[-0.25; 0.12]	0.08	0.11	[-0.04; 0.39]	0.01	0.13	[-0.21; 0.29]
Team quality (before t1)	0.25***	0.04	[0.14; 0.31]	0.21***	0.05	[0.11; 0.31]	0.18*	0.09	[0.05; 0.40]
Team quality (before t2)									
Team quality (before t3)				0.29***	0.06	[0.21; 0.43]			
Leader performance (t1)				0.31***	0.10	[0.43; 0.81]			
Continuous engagement with club (t1-t2)									
Leader performance (t2)							0.36***	0.07	[0.25; 0.52]
Continuous engagement with club (t2-t3)							0.21***	0.11	[0.23; 0.67]

All *p*-Values were adjusted with Benjamini and Hochberg's (1995) false discovery rate controlling step-up Bonferroni procedure as recommended for structural equation models. According to recommendations for triangulation (Cortina & Folger, 1998), we adopted liberal *p*-values of $p \leq .10$ (emphasized in italics).

* $p \leq .10$, ** $p \leq .05$, *** $p \leq .01$, **** $p \leq .001$

whad a previous player career), path models were not suitable. Hence, we conducted six multiple regression analyses using our different operationalizations of EP and LP for both coach groups. This approach allows for comparisons of the EP-LP link between both groups of coaches (i.e., those who started their coaching career in a first or second professional league within the UEFA and those who started their coaching career in lower leagues or with youth teams). Our analyses showed non-significant EP-LP links for both groups of coaches with one exception (i.e., the regression with the operationalizations player rating and the goal ratio at lower leagues indicated a negative relationship; for further information, see the supplemental file, Table S5).

Improved Test of TEL within Organizations

As TEL proposed processes that should unfold within an organization over time (e.g., positive effects of hiring applicants with high potential), those effects cannot be tested conclusively with leaders who change teams relatively frequently (as in our study). When restricting our sample to coaches that did not change their club for three seasons, our sample was again too small to conduct path analyses ($n = 77$; see Kline, 2011). Thus, in an exploratory analysis, we computed partial correlations (i.e., controlling for prior team quality) and Steiger's *z*-test (Diedenhofen & Musch, 2015; Steiger, 1980) to test whether the relationship between prior player and later coach performance differed over time (e.g., comparing the link between the average player rating and PPG at t1 and t2). The link between EP and LP increased descriptively over time for only two of the six combinations of operationalizations (i.e., average player rating with PPG, t1: $r = -.25$, $p = .151$; t2: $r = .01$, $p = .974$; t3: $r = .21$, $p = .244$; average player rating with goal ratio, t1: $r = -.19$, $p = .296$; t2: $r = .04$, $p = .838$; and t3: $r = .15$, $p = .388$). However, despite the descriptive increase the correlations did not change significantly over time ($|z| < 1.95$, $ps > .05$).

Discussion

We derived and tested contrasting hypotheses on performance-based promotion from three different theoretical perspectives (i.e., the performance requirements perspective, the follower-centric perspective, and the TEL) to resolve the inconsistency of prior findings. To do so, we focused on the role of time and job complexity of the employee position. The underlying assumption of performance-based promotions is that successful employees will make successful leaders. However, our various analyses did not support this assumption as the range of effect sizes for the relationship between EP and LP mostly ranged from small negative to small positive effects in our sample (i.e.,

the population of Bundesliga coaches). Likewise, across time, the relation between EP and LP mostly ranged from at most small negative to small positive effects. Thus, overall, we found no systematic link between EP and LP. Even when aggregating LP over the whole career, our analyses indicated no significant link with EP. Further, for professional player experience (yes vs. no), only one of three analyses indicated a small positive link with initial LP. However, this association appears to depend on the operationalization of the criterion (i.e., LP operationalized as goal ratio) and overall there was no systematic link between player experience and LP.

Concerning positions with a higher overlap of performance requirements due to the complexity of the employee position, our analyses did not confirm the proposed moderation. However, a bigger sample size than available in this study might be required to reliably detect a moderation of an already small main effect as proposed in our study. The descriptive findings nevertheless do not indicate the proposed moderation. While we could not identify conditions with higher (lower) predictive validity, our results suggest that the overall association between EP and LP—at least in the context of our study—is at most small, both initially and over time. Thus, there is no empirical reason to expect a head start from previously high-performing employees.

Theoretical Implications and Future Research

With this study, we aim to initiate an exchange between different disciplines (i.e., personnel selection and management research) on the question of whether high-performing employees may later truly make successful leaders. To do so, we relied on the performance requirements perspective (Zaccaro, 2012), the follower-centric perspective (Uhl-Bien et al., 2014), and TEL (Goodall & Bäker, 2015) to derive our hypotheses: In particular, we specified hypotheses on the initial validity of performance-based promotion as well as on its validity over time. Due to different assumptions on mediating processes (such as performance requirements vs. perceived leader prototypicality vs. better strategic decision making) those perspectives led to contrasting hypotheses on the predictive validity of EP for LP. We provided the first empirical test on these contrasting hypotheses. Thereby, we systematically examined how the link between EP and LP unfolds over time (i.e., initially after a promotion and across consecutive seasons). Our study results (i.e., no systematic associations between EP and LP, neither initially nor over time) are consistent with the performance requirements perspective (see Zaccaro et al., 2018), but not with the follower-centric perspective or with TEL. Moreover, our findings are mostly in line with the conclusion of a recent review (Schleu & Hüffmeier, 2021).

The performance requirements perspective specifies the predictive validity of EP for later LP by comparing the performance requirements of the employee and the leader position. Due to the different tasks of employee and leader positions (cf. Bartram, 2005; Borman & Brush, 1993), employee performance requirements match only to a small degree with the performance requirements for the leader. Our research supports this general reasoning and the importance of this theoretical perspective when predicting the overall EP-LP link.

With an increasing overlap of performance requirements of EP and LP, the validity of performance-based promotion should increase. We tested job complexity (as one way to increase the overlap of performance requirements) of the employee position as a potential moderator of the relationship between EP and LP. Our results, however, indicated no support for the proposed moderation in our study context. To further clarify whether the theoretical rationale concerning an increasing overlap in the respective performance requirements is correct, further studies are needed to examine job complexity (and other proposed moderators) more comprehensively. According to the assumptions of the performance requirements perspective, complexity moderates the small link between EP and LP. Consequently, a comprehensive test might require a bigger sample size in the future. This might allow a more informed use of performance-based promotion in the future.

Our results did not support the follower-centric perspective concerning performance-based promotion in the context of our study. Future research could clarify under which conditions these findings hold (within and beyond sports) and thereby also improve the current understanding of the follower-centric perspective. In particular, it might be promising in more conventional organizational contexts, where subjective performance evaluation is more common, to consider several sources before and after a promotion for a comprehensive analysis of the perceived prototypicality, EP, and LP (while avoiding the same rater bias; Hoyt, 2000). Potentially, such biases could inflate the perceived validity of performance-based promotion and therefore ensure the ongoing prevalence of this strategy in practice.

Our results provided no support for TEL (Goodall & Bäker, 2015) in the context of our study (i.e., professional soccer), as the analyses indicated overall no systematic link between EP and LP. As the processes proposed by TEL should unfold within an organization over time (e.g., positive effects of hiring applicants with high potential), it is conceivable that it may not be possible to test these effects conclusively with leaders who change teams relatively frequently (as in our study). When restricting our sample to coaches who did not change their club over the course of the least three seasons (in a post-hoc analysis), our findings were inconsistent concerning

the hypothesized effects over time in this substantially smaller subsample (i.e., two of the six analyses indicated a descriptive increase of the EP-LP link over time). To adequately test the proposed effects of TEL in the future, a rather stable study context (i.e., avoiding power problems) or shorter time lags (i.e., coach performance per week or match day) might be beneficial, as the coaches oftentimes worked only one season for a club.

As the majority of prior studies on performance-based promotion reported negative, non-significant, and mixed findings (for an overview, see Schleu & Hüffmeier, 2021) and our results did not indicate a systematic link between EP and LP, the empirical evidence questions the general validity of the logic underlying performance-based promotion. Some variability in the results of extant studies could potentially be explained by the varying importance of the different theoretical perspectives for explaining the EP-LP link across different contexts. Hence, it might be promising to theorize on boundary conditions for the different theoretical perspectives to understand when, for instance, the follower-centric perspective is equally or even more important than the performance requirement perspective.

Limitations

Acknowledging the limitations of our research, we would like to point out, first, that we only examined the association between EP and LP and tested the three theoretical perspectives (on how EP relates to LP) independently. While we based our theorizing concerning the three perspectives on the respective assumptions about underlying mechanisms, our data did not allow testing these mechanisms. Obviously, research on performance-based promotion would profit from examining the proposed mediators. This approach would also provide relevant insights about the explained variance of the different proposed mechanisms (see Schleu & Hüffmeier, 2021). Hence, we suggest that future research employs designs allowing for a comparative and more comprehensive evaluation of the underlying mechanisms of the three perspectives.

Second, as our study design is non-experimental (i.e., an observational field study), we could face endogeneity-related issues (see Antonakis, 2017; Antonakis et al., 2010). To limit the related risks, we relied on various objective—rather than subjective—performance measures that provided us with data of high content validity (Quiñones et al., 1995; Sturman, 2007) to avoid confounding effects due to biases (see Ciancetta & Roch, 2021; Kossek & Buzzanell, 2018). As the variety of measures produced consistent results across different time frames (see also our additional analyses in the Online Supplement), our results can be considered to be reliable (Cortina & Folger, 1998). Further, we collected panel data (see

Mackey, 2008) and included relevant (control) variables in our analyses to avoid omitted variable bias affecting the relation between EP and LP (see Antonakis et al., 2010), such as team-level variables (i.e., prior team quality measures for the team-related LP measures).

Third, our sample is range-restricted in the criterion, which is common in personnel selection research (see Sackett & Yang, 2000): While our data entail all head coaches of Germany's first soccer league, obviously those who tried to obtain a head coach position but failed to do so are not represented. However, the common range-restriction scenarios (cf. Sackett & Yang, 2000) should not apply to our study, as we gathered data from several organizations (i.e., soccer clubs of the Bundesliga) and the whole population of Bundesliga coaches. Hence, our analyses should not be affected severely by range-restriction.

Finally, the generalizability of our findings to leadership contexts outside the domain of professional team sports might be limited. The characteristics of our study context, professional soccer (e.g., a highly competitive context with rather short leadership tenure, a coaching license as a basic requirement to enter this career path, etc.), might be different to other work contexts. Also, while we deliberately chose the soccer context to examine the link between EP and LP due to the availability of objective performance data over whole career courses, many other contexts cannot rely on (comparable) objective performance measures. Consequently, decision makers base their promotion decisions on more subjective and, therefore, likely (more) biased performance ratings (Ciancetta & Roch, 2021; Kossek & Buzzanell, 2018), which may affect the EP-LP link in such contexts.

Practical Implications

Our results do not provide support for the validity of performance-based promotion, at least in the high-stakes context of our study. Considering the importance of leaders for an organization (see Montano et al., 2017), selection decisions based on low-validity methods are very costly (Cronbach & Gleser, 1965; Schmidt & Hunter, 1998). While we are cautious not to overstretch the implications of our results—as prior empirical evidence is inconclusive (Schleu & Hüffmeier, 2021)—many practitioners would likely benefit from questioning and adapting their routine promotion procedures. Therefore, we would like to point out two ways to improve the validity of performance-based promotion. First, as our results indicated initial support for the performance requirements perspective, practitioners might focus on the performance requirements of the vacant leader position to improve the validity of the performance-based promotion strategy. To identify relevant performance requirements, conducting a systematic job analysis for the vacant leader position should be helpful. An additional job analysis of the respective employee position allows to identify

the performance requirements indicated by EP (see Hesketh & Robertson, 1993). This approach would allow identifying positions with a high overlap of performance requirements of both (employee and leader) positions—and thereby ensure a preferably high validity of performance-based promotion. Further, this approach would allow considering only performance aspects of prior EP that are relevant for the later leader position. Thereby, this approach should improve the selection of candidates (see Wernimont & Campbell, 1968).

Second, organizations might review—and if necessary adapt—their performance assessment concerning the following criteria to maximize the validity: Does the performance evaluation follow a standardized approach (i.e., are relevant criteria defined and examples provided; see Schleicher et al., 2018)? Are the evaluating leaders trained in the procedure, but also in recognizing and avoiding potential biases (see Amis et al., 2020; van Dijk et al., 2020)? Could the performance be assessed by more than one rater (see Harari & Viswesvaran, 2018)? Could EP be complemented by other indicators (e.g., structured interviews or assessment centers targeting central characteristics of the vacant position) to increase the validity (see Sackett et al., 2022)? All those adaptations have the potential to optimize the validity of the performance assessment and thereby performance-based promotion.

Beyond its validity, practitioners need to account for effects on the organizational level (Wright & Boswell, 2002) when evaluating the overall merit of performance-based promotion. As performance-based promotion rewards good performance (Kim, 2019; Scully, 2002), this promotion strategy should influence the organization's work culture to focus on performance (see Bagdadli et al., 2006). Further, performance-based promotion should increase perceived career opportunities in the organization (Webster & Beehr, 2013), as employees perceive this approach as fairer than more arbitrary approaches (Beehr et al., 2004). Hence, performance-based promotion strategies should attract talented applicants and motivate employees (Gruman & Saks, 2011). Thus, when practitioners adapt performance-based promotion to increase its validity, they should be careful not to undermine the positive side effects of this strategy.

Conclusion

Performance-based promotion is a face-valid approach to fill leader positions and currently a prevalent strategy (Church et al., 2021). However, extant empirical evidence has been inconsistent (Schleu & Hüffmeier, 2021) and our findings did not support this strategy's merit either. Our work has contrasted three theoretical perspectives on performance-based promotion and tested two proposed moderators (i.e., time and job complexity of the employee position) to potentially resolve previously inconsistent results. Our

results support the performance requirements perspective (Zaccaro et al., 2018) regarding the predictive validity of EP for LP. Further, we found no evidence supporting the follower-centric perspective (see Uhl-Bien et al., 2014) and TEL (Goodall & Bäker, 2015). Thus, practitioners may be well-advised to focus on the respective performance requirements of the leader (and the employee) position when evaluating the EP of potential candidates. Shifting the focus to required characteristics holds the potential to improve the validity of performance-based promotion decisions—and consequently leader selection.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10869-023-09893-9>.

Acknowledgements We would like to thank Ann-Kathrin Torka and Catherine Schittenhelm for their help with data collection.

Data Availability The data that support the findings of this study are available from the corresponding author upon reasonable request.

Declarations

Conflict of Interests The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Acock, A. C. (2014). *A gentle introduction to Stata* (4th ed.). Stata Press.
- Alexander, N. (2015). What's more general than a whole population? *Emerging Themes in Epidemiology*, 12(1), 1–5. <https://doi.org/10.1186/s12982-015-0029-4>
- Amis, J. M., Mair, J., & Munir, K. A. (2020). The organizational reproduction of inequality. *Academy of Management Annals*, 14(1), 195–230. <https://doi.org/10.5465/annals.2017.0033>
- 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.leafqua.2017.01.006>
- Antonakis, J., Bendahan, S., Jacquart, P., & Lalive, R. (2010). On making causal claims: A review and recommendations. *The Leadership Quarterly*, 21(6), 1086–1120. <https://doi.org/10.1016/j.leafqua.2010.10.010>
- Asher, J. J., & Sciarino, J. A. (1974). Realistic work sample tests: A review. *Personnel Psychology*, 27(4), 519–533. <https://doi.org/10.1111/j.1744-6570.1974.tb01173.x>
- Audas, R., Dobson, S., & Goddard, J. (1999). Organizational performance and managerial turnover. *Managerial and Decision*

- Economics*, 20(6), 305–318. [https://doi.org/10.1002/\(SICI\)1099-1468\(199909\)20:6%3c305::AID-MDE933%3e3.0.CO;2-O](https://doi.org/10.1002/(SICI)1099-1468(199909)20:6%3c305::AID-MDE933%3e3.0.CO;2-O)
- Bagdadli, S., Roberson, Q., & Paoletti, F. (2006). The mediating role of procedural justice in responses to promotion decisions. *Journal of Business and Psychology*, 21(1), 83–102. <https://doi.org/10.1007/s10869-005-9017-y>
- Bartram, D. (2005). The great eight competencies: A criterion-centric approach to validation. *Journal of Applied Psychology*, 90(6), 1185–1203. <https://doi.org/10.1037/0021-9010.90.6.1185>
- Baumann, F., Friehe, T., & Wedow, M. (2011). General ability and specialization: Evidence from penalty kicks in soccer. *Journal of Sports Economics*, 12(1), 81–105. <https://doi.org/10.1177/1527002510371194>
- Becker, T. E. (2005). Potential problems in the statistical control of variables in organizational research: A qualitative analysis with recommendations. *Organizational Research Methods*, 8(3), 274–289. <https://doi.org/10.1177/1094428105278021>
- Beehr, T. A., Nair, V. N., Gudanowski, D. M., & Such, M. (2004). Perceptions of reasons for promotion of self and others. *Human Relations*, 57(4), 413–438. <https://doi.org/10.1177/0018726704043894>
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57(1), 289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- Benson, A., Li, D., & Shue, K. (2019). Promotions and the Peter Principle. *The Quarterly Journal of Economics*, 134(4), 2085–2134. <https://doi.org/10.1093/qje/qjz022>
- Borman, W. C., & Brush, D. H. (1993). More progress toward a taxonomy of managerial performance requirements. *Human Performance*, 6(1), 1–21. https://doi.org/10.1207/s15327043hup0601_1
- Brogden, H. E. (1949). When testing pays off. *Personnel Psychology*, 2(2), 171–183. <https://doi.org/10.1111/j.1744-6570.1949.tb01397.x>
- Campbell, J. P. (1990). Modeling the performance prediction problem in industrial and organizational psychology. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (pp. 687–732). Consulting Psychologists Press.
- Cannella, A. A., & Rowe, W. G. (1995). Leader capabilities, succession, and competitive context: A study of professional baseball teams. *The Leadership Quarterly*, 6(1), 69–88. [https://doi.org/10.1016/1048-9843\(95\)90005-5](https://doi.org/10.1016/1048-9843(95)90005-5)
- Carron, A. V., & Eys, M. A. (2012). *Group dynamics in sport* (4th ed.). Fitness Information Technology.
- Chelladurai, P. (2013). A personal journey in theorizing in sport management. *Sport Management Review*, 16(1), 22–28. <https://doi.org/10.1016/j.smr.2011.12.003>
- Chelladurai, P., & Carron, A. V. (1977). A reanalysis of formal structure in sport. *Canadian Journal of Applied Sport Sciences*, 2, 9–14.
- Church, A. H., Rotolo, C. T., Ginther, N. M., & Levine, R. (2015). How are top companies designing and managing their high-potential programs? A follow-up talent management benchmark study. *Consulting Psychology Journal: Practice and Research*, 67(1), 17–47.
- Church, A. H., Guidry, B. W., Dickey, J. A., & Scrivani, J. A. (2021). Is there potential in assessing for high-potential? Evaluating the relationships between performance ratings, leadership assessment data, designated high-potential status and promotion outcomes in a global organization. *The Leadership Quarterly*, 32(5), 101516. <https://doi.org/10.1016/j.leaqua.2021.101516>
- Ciancetta, L. M., & Roch, S. G. (2021). Backlash in performance feedback: Deepening the understanding of the role of gender in performance appraisal. *Human Resource Management*, 60(4), 641–657. <https://doi.org/10.1002/hrm.22059>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates. <https://doi.org/10.1037/cpb0000030>
- Cortina, J. M., & Folger, R. G. (1998). When is it acceptable to accept a null hypothesis: No way, Jose? *Organizational Research Methods*, 1(3), 334–350. <https://doi.org/10.1177/109442819813004>
- Cribbie, R. A. (2007). Multiplicity control in structural equation modeling. *Structural Equation Modeling*, 14(1), 98–112. <https://doi.org/10.1080/10705510709336738>
- Cronbach, L. J., & Gleser, G. C. (1965). *Psychological tests and personnel decisions* (2nd ed.). University of Illinois Press Psychological tests and personnel decisions.
- Crust, L., & Lawrence, I. (2006). A review of leadership in sport: Implications for football management. *Athletic Insight: The Online Journal of Sport Psychology*, 8(4), 28–48.
- Dane, E. (2010). Reconsidering the trade-off between expertise and flexibility: A cognitive entrenchment perspective. *Academy of Management Review*, 35(4), 579–603. <https://doi.org/10.5465/AMR.2010.53502832>
- Dawson, P., & Dobson, S. (2002). Managerial efficiency and human capital: An application to English association football. *Managerial and Decision Economics*, 23(8), 471–486. <https://doi.org/10.1002/mde.1098>
- Dawson, P., Dobson, S., & Gerrard, B. (2000). Estimating coaching efficiency in professional team sports: Evidence from English association football. *Scottish Journal of Political Economy*, 47(4), 399–421. <https://doi.org/10.1111/1467-9485.00170>
- Day, D. V., Harrison, M. M., & Halpin, S. M. (2009). An integrative approach to leader development: Connecting adult development, identity, and expertise. *Psychology Press*. <https://doi.org/10.4324/9780203809525>
- De Hoogh, A. H. B., Den Hartog, D. N., & Koopman, P. L. (2005). Linking the Big Five-Factors of personality to charismatic and transactional leadership: Perceived dynamic work environment as a moderator. *Journal of Organizational Behavior*, 26(7), 839–865. <https://doi.org/10.1002/job.344>
- DeRue, D. S., Nahrgang, J. D., Wellman, N. E. D., & Humphrey, S. E. (2011). Trait and behavioral theories of leadership: An integration and meta-analytic test of their relative validity. *Personnel Psychology*, 64(1), 7–52. <https://doi.org/10.1111/j.1744-6570.2010.01201.x>
- Diedenhofen, B., & Musch, J. (2015). cocor: A comprehensive solution for the statistical comparison of correlations. *PLoS one*, 10(4), e0121945. <https://doi.org/10.1371/journal.pone.0121945>
- Enders, C. K. (2010). *Applied missing data analysis*. Guilford Press.
- Ensari, N., & Murphy, S. E. (2003). Cross-cultural variations in leadership perceptions and attribution of charisma to the leader. *Organizational Behavior and Human Decision Processes*, 92(1–2), 52–66. [https://doi.org/10.1016/S0749-5978\(03\)00066-9](https://doi.org/10.1016/S0749-5978(03)00066-9)
- Ericsson, K. A., Charness, N., Feltovich, P. J., & Hoffman, R. R. (Eds.). (2006). *The Cambridge handbook of expertise and expert performance*: Cambridge University Press. <https://doi.org/10.1017/CBO9780511816796>
- Fischer, T., Dietz, J., & Antonakis, J. (2017). Leadership process models: A review and synthesis. *Journal of Management*, 43(6), 1726–1753. <https://doi.org/10.1177/0149206316682830>
- Frick, B. (2010). The football player's labor market: Recent developments and economic evidence. In J. Galí & X. Vives (Eds.), *Football Economics. Proceedings of the Armand Carabén Workshop on Sports Economics* (Vol. 4, pp. 103–140). Barcelona: IESE Business School, University of Navarra. Retrieved from <http://www.iese.edu/research/pdfs/ESTUDIO-135.pdf#page=105>
- Gallup. (2014). *Engagement Index Deutschland 2014 [Engagement Index Germany 2014]*.
- Gammelsæter, H. (2013). Leader succession and effectiveness in team sport. A critical review of the coach succession literature. *Sport, Business and Management: An International Journal*, 3(4), 285–296. <https://doi.org/10.1108/SBM-06-2013-0015>
- Gentry, W. A., Hoffman, B. J., & Lyons, B. D. (2017). Box scores and bottom lines: Sports data can inform research and practice in organizations. *Journal of Business and Psychology*, 32, 509–512. <https://doi.org/10.1007/s10869-017-9513-x>

- Goodall, A. H., & Bäker, A. (2015). A theory exploring how expert leaders influence performance in knowledge-intensive organizations. In I. Welpe, J. Wollersheim, S. Ringelhan, & M. Osterloh (Eds.), *Incentives and performance. Governance of research organizations* (pp. 49–67). Springer International Publishing Switzerland. https://doi.org/10.1007/978-3-319-09785-5_4
- Goodall, A. H., & Pogrebnia, G. (2015). Expert leaders in a fast-moving environment. *The Leadership Quarterly*, 26(2), 123–142. <https://doi.org/10.1016/j.leaqua.2014.07.009>
- Gruman, J. A., & Saks, A. M. (2011). Performance management and employee engagement. *Human Resource Management Review*, 21(2), 123–136. <https://doi.org/10.1016/j.hrmr.2010.09.004>
- Grusky, O. (1963). The effects of formal structure on managerial recruitment: A study of baseball organization. *Sociometry*, 26(3), 345–353. <https://doi.org/10.2307/2786074>
- Hackman, J. R., & Oldham, G. R. (1975). Development of the Job Diagnostic Survey. *Journal of Applied Psychology*, 60(2), 159–170. <https://doi.org/10.1037/h0076546>
- Hall, C. M., & Pedace, R. (2016). Do managers matter? Manager effects on organization performance. *Managerial and Decision Economics*, 37(8), 541–551. <https://doi.org/10.1002/mde.2743>
- Harari, M., & Viswesvaran, C. (2018). Individual job performance. In D. S. Ones, N. Anderson, & C. Viswesvaran (Eds.), *The sage handbook of industrial, work and organizational psychology* (55–72). SAGE Publications. <https://doi.org/10.4135/9781473914940.n4>
- Haslam, S. A., Reicher, S., & Platow, M. (2011). The new psychology of leadership: Identity, influence, and power. *Psychology Press*. <https://doi.org/10.4324/9780203833896>
- Hesketh, B., & Robertson, I. (1993). Validating personnel selection: A process model for research and practice. *International Journal of Selection and Assessment*, 1(1), 3–17. <https://doi.org/10.1111/j.1468-2389.1993.tb00079.x>
- Hinkin, T. R., & Tracey, J. B. (1999). An analysis of variance approach to content validation. *Organizational Research Methods*, 2(2), 175–186. <https://doi.org/10.1177/109442819922004>
- Hogg, M. A. (2001). A social identity theory of leadership. *Personality and Social Psychology Review*, 5(3), 184–200. https://doi.org/10.1207/S15327957PSPR0503_1
- Hoyt, W. T. (2000). Rater bias in psychological research: When is it a problem and what can we do about it? *Psychological Methods*, 5(1), 64–86. <https://doi.org/10.1037/1082-989X.5.1.64>
- Hunter, J. E., Schmidt, F. L., & Judiesch, M. K. (1990). Individual differences in output variability as a function of job complexity. *Journal of Applied Psychology*, 75(1), 28–42. <https://doi.org/10.1037/0021-9010.75.1.28>
- Kark, R., & Van Dijk, D. (2007). Motivation to lead, motivation to follow: The role of the self-regulatory focus in leadership processes. *Academy of Management Review*, 32(2), 500–528. <https://doi.org/10.5465/amr.2007.24351846>
- Katz, D., & Kahn, R. L. (1978). *The social psychology of organizations* (2nd ed.). Wiley.
- Kim, S. L. (2019). Enticing high performers to stay and share their knowledge: The importance of trust in leader. *Human Resource Management*, 58(4), 341–351. <https://doi.org/10.1002/hrm.21955>
- Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). Guilford Press.
- Koopmans, L., Bernaards, C. M., Hildebrandt, V. H., de Vet, H. C., & van der Beek, A. J. (2014). Construct validity of the Individual Work Performance Questionnaire. *Journal of Occupational and Environmental Medicine*, 56(3), 331–337. <https://doi.org/10.1097/JOM.000000000000113>
- Kossek, E. E., & Buzzanell, P. M. (2018). Women's career equality and leadership in organizations: Creating an evidence-based positive change. *Human Resource Management*, 57(4), 813–822. <https://doi.org/10.1002/hrm.21936>
- Kouzes, J. M., & Posner, B. Z. (2011). *Credibility: How leaders gain and lose it, why people demand it*. Jossey-Bass. <https://doi.org/10.1002/9781118983867>
- Kroemer, U. (2010, December 13). *Kicker-Noten: Die Zeugnisse der Profis*. Retrieved August 24, 2022, from <https://www.news.de/sport/855093240/die-zeugnisse-der-profis/1/>
- Lord, R. G., & Dinh, J. E. (2014). What Have We Learned That Is Critical in Understanding Leadership Perceptions and Leader-Performance Relations? *Industrial and Organizational Psychology*, 7(2), 158–177. <https://doi.org/10.1111/iops.12127>
- Lord, R. G., & Maher, K. J. (2002). *Leadership and information processing: Linking perceptions and performance*. Routledge. <https://doi.org/10.4324/9780203423950>
- Lord, R. G., Foti, R. J., & de Vader, C. L. (1984). A test of leadership categorization theory: Internal structure, information processing, and leadership perceptions. *Organizational Behavior and Human Performance*, 34(3), 343–378. [https://doi.org/10.1016/0030-5073\(84\)90043-6](https://doi.org/10.1016/0030-5073(84)90043-6)
- Mackey, A. (2008). The effect of CEOs on firm performance. *Strategic Management Journal*, 29(12), 1357–1367. <https://doi.org/10.1002/smj.708>
- Martinez, J. A., & Caudill, S. B. (2013). Does midseason change of coach improve team performance? Evidence from the NBA. *Journal of Sport Management*, 27(2), 108–113. <https://doi.org/10.1123/jism.27.2.108>
- Maynard, M. T., Resick, C. J., Cunningham, Q. W., & DiRenzo, M. S. (2017). Ch-Ch-Ch-changes: How action phase functional leadership, team human capital, and interim vs. permanent leader status impact post-transition team performance. *Journal of Business and Psychology*, 32(5), 575–593. <https://doi.org/10.1007/s10869-016-9482-5>
- McKinsey (2020). *Unternehmen Bundesliga: Die ökonomische Bedeutung des Profifußballs in Deutschland*. <https://www.dfl.de/de/ueber-uns/publikationen/wachstumsmotor-bundesliga/>
- Meade, A. W., & Craig, S. B. (2012). Identifying careless responses in survey data. *Psychological Methods*, 17(3), 437–455. <https://doi.org/10.1037/a0028085>
- Meindl, J. R. (1995). The romance of leadership as a follower-centric theory: A social constructionist approach. *The Leadership Quarterly*, 6(3), 329–341. [https://doi.org/10.1016/1048-9843\(95\)90012-8](https://doi.org/10.1016/1048-9843(95)90012-8)
- Montano, D., Reeske, A., Franke, F., & Hüffmeier, J. (2017). Leadership, followers' mental health and job performance in organizations: A comprehensive meta-analysis from an occupational health perspective. *Journal of Organizational Behavior*, 38(3), 327–350. <https://doi.org/10.1002/job.2124>
- Muehlheusser, G., Schneemann, S., Sliwka, D., & Wallmeier, N. (2018). The contribution of managers to organizational success: Evidence from German soccer. *Journal of Sports Economics*, 19(6), 786–819. <https://doi.org/10.1177/1527002516674760>
- Murphy, K. R., & Cleveland, J. N. (1995). *Understanding performance appraisal: Social, organizational and goal-based perspectives*. Sage Publications.
- Peter, L. J., & Hull, R. (Eds.). (2011). *The Peter Principle: Why things always go wrong* (1st ed.). Harper Business.
- Phillips, J. S., & Lord, R. G. (1981). Causal attributions and perceptions of leadership. *Organizational Behavior and Human Performance*, 28(2), 143–163. [https://doi.org/10.1016/0030-5073\(81\)90020-9](https://doi.org/10.1016/0030-5073(81)90020-9)
- Porter, T., Riesenmy, K., & Fields, D. (2016). Work environment and employee motivation to lead: Moderating effects of personal characteristics. *American Journal of Business*, 31(2), 66–84. <https://doi.org/10.1108/AJB-05-2015-0017>
- Quiñones, M. A., Ford, J. K., & Teachout, M. S. (1995). The relationship between work experience and job performance: A conceptual and meta-analytic review. *Personnel Psychology*, 48(4), 887–910. <https://doi.org/10.1111/j.1744-6570.1995.tb01785.x>

- Rich, B. L., Lepine, J. A., & Crawford, E. R. (2010). Job engagement: Antecedents and effects on job performance. *Academy of Management Journal*, 53(3), 617–635. <https://doi.org/10.5465/amj.2010.51468988>
- Rogers, E. M., & Bhowmik, D. K. (1970). Homophily-heterophily: Relational concepts for communication research. *Public Opinion Quarterly*, 34(4), 523–538. <https://doi.org/10.1086/267838>
- Rogers, B. L., Vardaman, J. M., Allen, D. G., Muslin, I. S., & Baskin, M. B. (2017). Turning up by turning over: The change of scenery effect in major league baseball. *Journal of Business and Psychology*, 32(5), 547–560. <https://doi.org/10.1007/s10869-016-9468-3>
- Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling and more Version 0.5–12 (BETA). *Journal of Statistical Software*, 48(2), 1–36. <https://doi.org/10.18637/jss.v048.i02>
- Sackett, P. R., & Yang, H. (2000). Correction for range restriction: An expanded typology. *Journal of Applied Psychology*, 85(1), 112–118. <https://doi.org/10.1037/0021-9010.85.1.112>
- Sackett, P. R., Zhang, C., Berry, C. M., & Lievens, F. (2022). Revisiting meta-analytic estimates of validity in personnel selection: Addressing systematic overcorrection for restriction of range. *Journal of Applied Psychology*, 107(11), 2040–2068. <https://doi.org/10.1037/apl0000994>
- Salomon, G., & Perkins, D. N. (1989). Rocky roads to transfer: Rethinking mechanism of a neglected phenomenon. *Educational Psychologist*, 24(2), 113–142. https://doi.org/10.1207/s15326985ep2402_1
- Schleicher, D. J., Baumann, H. M., Sullivan, D. W., Levy, P. E., Hargrove, D. C., & Barros-Rivera, B. A. (2018). Putting the system into performance management systems: A review and agenda for performance management research. *Journal of Management*, 44(6), 2209–2245. <https://doi.org/10.1177/0149206318755303>
- Schleu, J. E., & Hüffmeier, J. (2021). Simply the best? A systematic literature review on the predictive validity of employee performance for leader performance. *Human Resource Management Review*, 31(2), 100777. <https://doi.org/10.1016/j.hrmr.2020.100777>
- Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, 124(2), 262–274. <https://doi.org/10.1037/0033-2909.124.2.262>
- Schyns, B., & Meindl, J. R. (Eds.). (2005). *Implicit leadership theories: Essays and explorations*. Information Age Pub.
- Scully, M. A. (2002). Confronting errors in the meritocracy. *Organization*, 9(3), 396–401.
- Smith, M. B., Hill, A. D., Wallace, J. C., Recendes, T., & Judge, T. A. (2018). Upsides to dark and downsides to bright personality: A multidomain review and future research agenda. *Journal of Management*, 44(1), 191–217. <https://doi.org/10.1177/0149206317733511>
- Statista (nd). *Umsatz der 1. und 2. Fußball-Bundesliga von der Saison 2004/2005 bis zur Saison 2020/2021*. Retrieved July 8, 2022, from <https://de.statista.com/statistik/daten/studie/4867/umfrage/entwicklung-der-erluese-in-der-ersten-und-zweiten-fussballbundesliga/>
- Steffens, N. K., Haslam, S. A., Reicher, S. D., Platow, M. J., Fransen, K., Yang, J., & Boen, F. (2014). Leadership as social identity management: Introducing the Identity Leadership Inventory (ILI) to assess and validate a four-dimensional model. *The Leadership Quarterly*, 25(5), 1001–1024. <https://doi.org/10.1016/j.leaqua.2014.05.002>
- Steffens, N. K., Munt, K. A., van Knippenberg, D., Platow, M. J., & Haslam, S. A. (2021). Advancing the social identity theory of leadership: A meta-analytic review of leader group prototypicality. *Organizational Psychology Review*, 11(1), 35–72. <https://doi.org/10.1177/2041386620962569>
- Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin*, 87(2), 245–251. <https://doi.org/10.1037/0033-2909.87.2.245>
- Sturman, M. (2007). The past, present, and future of dynamic performance research. *Research in Personnel and Human Resources Management*, 26, 49–110. [https://doi.org/10.1016/S0742-7301\(07\)26002-5](https://doi.org/10.1016/S0742-7301(07)26002-5)
- Tesluk, P. E., & Jacobs, R. R. (1998). Toward an integrated model of work experience. *Personnel Psychology*, 51(2), 321–355. <https://doi.org/10.1111/j.1744-6570.1998.tb00728.x>
- Tett, R. P., Guterman, H. A., Bleier, A., & Murphy, P. J. (2000). Development and content validation of a “hyperdimensional” taxonomy of managerial competence. *Human Performance*, 13(3), 205–251. https://doi.org/10.1207/S15327043HUP1303_1
- Türk, M. (2019). *Das gesellschaftliche Engagement des Profifußballs*. In *CSR und Fußball* (85–99). Springer Gabler, Berlin, Heidelberg.
- Uhl-Bien, M., Riggio, R. E., Lowe, K. B., & Carsten, M. K. (2014). Followership theory: A review and research agenda. *The Leadership Quarterly*, 25(1), 83–104. <https://doi.org/10.1016/j.leaqua.2013.11.007>
- van Dijk, H., Kooij, D., Karanika-Murray, M., De Vos, A., & Meyer, B. (2020). Meritocracy a myth? A multilevel perspective of how social inequality accumulates through work. *Organizational Psychology Review*, 10(3–4), 240–269. <https://doi.org/10.1177/2041386620930063>
- Van Iddekinge, C. H., Aguinis, H., Mackey, J. D., & DeOrtentiis, P. S. (2018). A meta-analysis of the interactive, additive, and relative effects of cognitive ability and motivation on performance. *Journal of Management*, 44(1), 249–279. <https://doi.org/10.1177/0149206317702220>
- Van Iddekinge, C. H., Arnold, J. D., Frieder, R. E., & Roth, P. L. (2019). A meta-analysis of the criterion-related validity of pre-hire work experience. *Personnel Psychology*, 72(4), 571–598. <https://doi.org/10.1111/peps.12335>
- Van Scotter, J., Motowidlo, S. J., & Cross, T. C. (2000). Effects of task performance and contextual performance on systemic rewards. *Journal of Applied Psychology*, 85(4), 526–534. <https://doi.org/10.1037/0021-9010.85.4.526>
- Webster, J. R., & Beehr, T. A. (2013). Antecedents and outcomes of employee perceptions of intra-organizational mobility channels. *Journal of Organizational Behavior*, 34(7), 919–941. <https://doi.org/10.1002/job.1823>
- Weick, K. E. (1995). *Sensemaking in organizations* (2nd ed.). Sage Publications.
- Wernimont, P. F., & Campbell, J. P. (1968). Signs, samples, and criteria. *Journal of Applied Psychology*, 52(5), 372–376. <https://doi.org/10.1037/h0026244>
- Wolfe, R. A., Weick, K. E., Usher, J. M., Terborg, J. R., Poppo, L., Murrell, A. J., & Jourdan, J. S. (2005). Sport and organizational studies: Exploring synergy. *Journal of Management Inquiry*, 14(2), 182–210. <https://doi.org/10.1177/1056492605275245>
- Wright, P., & Boswell, W. R. (2002). Desegregating HRM: A review and synthesis of micro and macro human resource management research. *Journal of Management*, 28(3), 247–276. [https://doi.org/10.1016/S0149-2063\(02\)00128-9](https://doi.org/10.1016/S0149-2063(02)00128-9)
- Zaccaro, S. J. (2012). Individual differences and leadership: Contributions to a third tipping point. *The Leadership Quarterly*, 23(4), 718–728. <https://doi.org/10.1016/j.leaqua.2012.05.001>
- Zaccaro, S. J., Green, J. P., Dubrow, S., & Kolze, M. (2018). Leader individual differences, situational parameters, and leadership outcomes: A comprehensive review and integration. *The Leadership Quarterly*, 29(1), 2–43. <https://doi.org/10.1016/j.leaqua.2017.10>