PERSONNEL PSYCHOLOGY
2004, 57, 573–605

ASSESSING THE BURDENS OF LEADERSHIP: EFFECTS OF FORMAL LEADERSHIP ROLES ON INDIVIDUAL PERFORMANCE OVER TIME

DAVID V. DAY, HOCK-PENG SIN, TINA T. CHEN
Department of Psychology, The Pennsylvania State University

This study adopted a role-based perspective in examining whether changes in performance over time (i.e., dynamic criteria) were a function of changes in individual leadership role responsibilities. Longitudinal data from captains in the modern era of the National Hockey League ($N = 201$) were used to test a dynamic criterion hypothesis using multilevel growth modeling. Time ($k = 10$) was modeled as a random effect, whereas captain status (i.e., leadership role responsibility) was included as a time-varying covariate. Individual performance was measured as the adjusted points (goals scored plus assists adjusted for individual and historical effects). Results of a series of model building steps that included the examination of alternative complex error structures indicated an overall negative performance trend. Those seasons in which a player assumed formal leadership responsibilities (i.e., team captain) were associated with better performance compared to seasons in which the player had no leadership responsibilities. These results were found to be robust even after controlling for individual performance in the previous season. Results are discussed in terms of the possible positive implications for individual performance and the motivation to lead through developing a culture in which leadership positions are highly valued by the organization, visible to others on the team, and where leadership responsibilities do not interfere with task performance.

A topic of considerable interest and attention in the job performance literature is whether criteria are static (i.e., stable) or dynamic over time. The notion of *dynamic criteria* (Ghiselli, 1956) was challenged nearly 20 years ago as a "received doctrine" in the industrial-organizational (I-O) psychology literature (Barrett, Caldwell, & Alexander, 1985). This challenge initiated a number of responses and empirical examinations using a variety of methods and populations. An overall assessment of this

Tina T. Chen is now with Sempra Energy, San Diego, CA. We thank the National Hockey League Hall of Frame for providing the rosters of team captains. We also thank Kimberly Erickson for her help with data coding and Rick DeShon for providing comments on an earlier draft manuscript.

Correspondence and request for reprints should be addressed to David V. Day, 124 Bruce V. Moore Building, Department of Psychology, The Pennsylvania State University, University Park, PA, 16802; dvd1@psu.edu.

literature has led to the general conclusion that the notion of dynamic criteria is supported by the empirical literature (see Ployhart & Hakel, 1998, for a review).

Despite the accumulation of evidence in support of dynamic criteria, relatively little is known about the nature of performance changes or the factors that are associated with performance change over time. There has been a call in the I-O literature for over a decade to "move beyond demonstrations of the phenomenon (or nihilistic critiques) to developing and testing explanations" of dynamic criteria (Austin & Villanova, 1990, p. 856). This study answers this call by examining the effects of assuming formal team leadership role responsibilities on individual performance over time–a topic that has been widely overlooked in both the leadership and job performance literatures. As such, this study links dynamic criteria and leadership.

Background on Dynamic Criteria

The dynamic criteria literature distinguishes two basic forms of dynamism. One form involves changes in criterion dimensionality over time. A second form involves only changes in the level of performance. The focus of this study is on the latter: To what degree does overall performance change over time? To that end we modeled performance as a unidimensional variable based on objective (i.e., nonratings) data.

In understanding dynamic criteria issues, another distinction is drawn between intraindividual (within-person) and interindividual (between-person) change. Intraindividual change summarizes the individual growth trajectories that are represented in a sample or population of interest, whereas interindividual change refers to differences between people in their growth trajectories (slopes and intercepts). Evidence of either intraindividual change (i.e., nonzero performance slopes) or interindividual change (e.g., rank order differences in people) is taken as support for dynamic criteria or changes in performance over time. Sophisticated modeling techniques presently allow researchers to estimate both forms of change simultaneously. Techniques such as random coefficients modeling, multilevel growth modeling, hierarchical linear modeling, and latent growth modeling are various names for a family of approaches that estimate and compare growth curves and that can accommodate both intraindividual and interindividual change (Raudenbush, 2001). When it comes to testing explanations of changes in performance over time, most of the more recent studies have used some type of multilevel growth modeling, focusing on the nature of intraindividual change (i.e., linear or curvilinear) and then using individual difference measures to predict systematic interindividual differences in these trajectories.

Some of the earliest research that adopted sophisticated growth models to test a dynamic criteria hypothesis was reported by Hofmann and associates, who modeled the performance of major league baseball players (Hofmann, Jacobs, & Gerras, 1992) and insurance sales personnel (Hofmann, Jacobs, & Baratta, 1993). The Hofmann et al. studies focused mostly on identifying intraindividual change patterns and clustering the patterns into similar interindividual groupings. More recent research has extended this earlier work by also using individual difference variables as predictors of intraindividual change parameters. These studies include: a hierarchical linear modeling approach to predicting sewing machine operator performance over time and modeling the interindividual effects of ability and experience (Deadrick, Bennett, & Russell, 1997); latent growth curve modeling in the prediction of differences in quarterly sales performance trajectories using biodata and self-assessed persuasion and empathy ratings (Ployhart & Hakel, 1998); latent growth modeling in a longitudinal examination of newcomer adjustment across three time periods in predicting adjustment trends using attitudinal measures collected at the time of hire (Lance, Vandenberg, & Self, 2000); hierarchical linear modeling of a longitudinal study of top-level executive performance in which variance in the rates of change in performance was predicted by people-oriented competency ratings (Russell, 2001); and random coefficients modeling of simulated data regarding leader adaptability over time to illustrate how leader agreeableness (interindividual effect) could be shown to predict initial status in leader adaptability as well as individual change parameters (Ployhart, Holtz, & Bliese, 2002).

Role-based Perspective on Performance Change over Time

These previous studies have all reported results that generally support a dynamic criteria perspective and have furthered understanding of interindividual predictors of performance trends but have not contributed a great deal to understanding the intraindividual influences on performance changes. An explanation that is examined in this study is grounded in changes in fundamental role responsibilities. Over their careers, individuals are likely to be assigned to different projects with different levels of challenge or sometimes given "stretch" assignments to prompt development. Most individuals are promoted at some point in their careers to new positions, assigned to a formal leadership role, or experience other types of role changes. Concomitant changes in role requirements potentially affect an individual's ability to perform at an optimal level, at least initially. Although role changes are likely to have some effect on individual performance, there is disagreement in the literature as to whether multiple roles detract from optimal performance or might potentially enhance it.

Within a team context, an individual teammate might be designated as a formal leader. Along with that designation is the expectation to assume additional responsibilities such as dealing with personnel issues, communicating with upper management, coordinating team performance, and taking responsibility for team outcomes. However, relatively little is known about how these formal role responsibilities affect individual performance over time. It has been noted that "we know surprisingly little about how leaders create and manage effective teams" (Zaccaro, Rittman, & Marks, 2001, p. 452), and we know even less about the effect of assuming leadership responsibilities on individual leader performance. Being appointed as team leader does not typically absolve a leader of continuing as an individual contributor to the team in addition to fulfilling key leadership obligations. Leadership usually involves the expansion of individual responsibilities over and above the fundamental responsibility of maintaining effective individual performance. A focal question of this study is whether leadership status is reliably associated with changes in performance over time. If so, does becoming a team leader help or hurt individual performance (or neither)? As mentioned, what happens to the task performance of leaders after they assume leadership responsibilities has been an issue that has been virtually unexplored in the literature.

In framing this study around leadership, it is important to point out that we examined a particular form of leadership. Leadership can be brought to a team by an external individual such as a supervisor, manager, or a coach. Thus, their primary role responsibilities involve managing and leading a team or multiple teams. Another form of leadership can be thought of as internal team leadership. In those cases, an individual contributor on the team is also assigned some type of formal leadership responsibilities such as serving as a liaison between the team and general management and in dealing with issues of morale, motivation, and conflict within the team. But the leader in those situations also remains very much a part of the team and is expected to contribute to team performance as an individual contributor as well as to provide internal leadership. The present focus was on the effects of this type of internal team leadership to the performance of the individual leader.

There are multiple theoretical explanations for the expected effects of role changes on individual performance. It has been argued that individuals can easily become overloaded with role responsibilities, contributing to role strain. An alternative explanation that has been offered in the literature is that multiple roles create more resources than they deplete. Each of these perspectives is briefly reviewed.

Role overload. A general assumption in the role theory literature is that individuals take on (or are assigned) more roles than they have the time, energy, or resources to handle, which results in *role overload* (Turner, 2002). The gist of this perspective is that individuals have only a finite amount of resources to devote to role performance. Resources become taxed when an individual's total role obligations are overdemanding, with the resulting anxiety, tension, and frustration creating a condition known as *role strain* (Goode, 1960). Because of the demands associated with maintaining maximal individual performance, any additional role requirements may detract from the cognitive and physical resources needed to perform at one's peak. This explanation is consistent with the view of some leadership researchers that taking on a formal leadership role usually involves some costs to individual leaders (Chan & Drasgow, 2001).

The essence of a role overload hypothesis is that being appointed team leader is likely to be detrimental to individual performance. Individuals are thought to have finite cognitive or physical resources and adding role responsibilities has the potential to overtax these resources, resulting in a strain on performance. In this manner, leadership can be a burden to an individual in terms of depleting resources that would otherwise be available for task performance. Taken to the extreme, however, this would seem to argue that team leaders are generally poor performers compared to other team members. This does not match with experience (Peter Principle notwithstanding) because it would be expected that a leader would be historically among the top performers on a team. Indeed, one of the reasons for being chosen as leader is thought to be because of superior individual performance. A key component of the implicit leadership theories of most people is that a leader is associated with effective performance and success (Lord & Maher, 1991). Thus, a leader who did not perform at an effective level probably would not be a leader for very long and certainly would not be likely to be perceived by others on the team as a leader.

The role overload perspective also fails to account for individual learning. Roles that are difficult at first become easier through practice and experience. With most complex skill acquisition, there is a learning curve in that the relationship between time and performance generally follows a negatively accelerating power function (Ployhart et al., 2002). Performance tends to be low initially but gradually improves over time until some maximum point is reached at which the time-performance slope flattens to zero or close to it. This assumes that individuals learn over time, provided some minimum level of role preparation. A lack of preparation (i.e., role readiness) could result in being overwhelmed and immobilized by a new role, with little learning or positive change in performance over time.

Role accumulation. Contrary to the position that managing multiple roles is a difficult and sometimes overwhelming task (Goode, 1960), others have argued that the benefits of *role accumulation* outweigh the challenges or the strain associated with role overload. One means by which role accumulation can lead to positive outcomes is by providing "privileges, buffers and resources [that] enhance one's self conception" (Sieber, 1974, p. 576). In this manner, appointment as team leader can serve as a self-fulfilling prophecy or promote a Pygmalion effect in which raising expectations regarding an individual's performance actually boosts that performance (Eden, 1990; Rosenthal & Jacobson, 1968).

From a follower's perspective, it has been proposed that leadership is mainly an attribution made by others (Calder, 1977; Pfeffer, 1977) and that appointing someone as a team leader would likely prime leadership attributions as well as other corresponding positive attributes to that formal leader. In a similar vein, others have argued that leadership is typically conceptualized in overly heroic terms and, thus, followers are likely to believe that leaders have more influence over team and organizational outcomes than might be true (Meindl, Ehrlich, & Dukerich, 1985). Both of these follower-centric perspectives on leadership provide additional theoretical support for why assuming a leadership role might create additional social resources for a leader, which is consistent with the role accumulation literature.

Another possible theoretical mechanism behind the role accumulation perspective is that the more roles that are assumed, the more resources are generated for performing in a variety of capacities. Instead of a *scarcity* approach to resources and human energy, it has been postulated that there is an energy-creation or an *expansion* approach in which activity produces more energy than it consumes (Marks, 1977). In this way, multiple roles provide some people with more energy and resources than those roles deplete. At the core of the expansion approach to the role accumulation perspective is the view that resources are considered to be an expandable commodity rather than a fixed entity. Empirical support for the resource expansion approach was reported recently in a study of managerial women, in which multiple role commitment was positively related to life satisfaction, self-esteem, self-acceptance, as well as to interpersonal and task-related managerial performance (Ruderman, Ohlott, Panzer, & King, 2002).

A third possibility is that an internal team leadership role has no impact on individual performance. In other words, it is possible that there is no positive or negative role spillover. This seems less plausible than either the role overload or role accumulation perspectives. Nonetheless, it offers a theoretically feasible null hypothesis that can be adopted in conducting the various model tests with regard to leadership role responsibilities.

Study Background

Leadership is a dynamic process and a formal leadership role designation can vary across individuals and time periods. For example, a team leader role designation might be rotated among team members or key leadership role responsibilities may vary across different projects. A leader in one team at one time period might be just a “regular” team member in a different team or subsequent time period. To examine the central research issue regarding the possibility of a time-varying and naturally occurring intervention (i.e., team leader role status) on individual performance, we modeled changes in individual performance of team captains from the modern era of the National Hockey League (NHL). In order to better understand the relevance of this sample to the issue at hand, we next provide background and an overview on the importance of team captains in the NHL.

There is a tradition beginning in the 1947–48 season of NHL captains wearing the designation of the letter “C” for captain on their sweaters (i.e., game jerseys). The C is a large and highly visible sign of a player’s formal leadership status. According to 1997 NHL Rulebook (http://www.icehockeyrules.8m.com/rulebook/rule14fr.htm), Team Rule 14(a) states:

> One Captain shall be appointed by each team, and he alone shall have the privilege of discussing with the Referee any questions relating to interpretation of rules which may arise during the progress of a game. He shall wear the letter “C,” approximately three inches (3″) in height and in contrasting color, in a conspicuous position on the front of his sweater.

Although it is an honor to wear hockey’s badge of leadership—it has been called “the most prestigious player leadership honor in pro sports” (Kreiser, 2001, para. 1)—this designation is more than mere formality. Team captains have many on-ice and off-ice responsibilities. As noted in the official NHL rules, captains have sole authority to speak for the team with referees regarding disputed calls and rule interpretations. More informally, they often have the authority to call team meetings; are the bridge of communication to the coach, general manager, news media, and fans; and have also been known to influence important personnel decisions (LaPointe, 1997). The role of NHL captain appears in many ways to be similar to that of other types of team leaders, but there are also some unique aspects of this type of leadership role that should be acknowledged. One such difference is the public nature of performance in professional sports. In very few domains is there such keen interest and attention given to individual and team performance, especially regarding the recognized leaders on NHL teams. Another difference is the pressure associated with such a heavy performance orientation. Players that do not perform to expected levels are typically released or traded. Of course, NHL players are compensated very well for this pressure and the public display of performance. In the 2003–2004 season the average NHL player salary was $1.79 million

(Erwin, 2004), which is safe to conclude is substantially higher than the average "team player" salary in most other business organizations.

An assumption among some players and coaches in the league is that the role of formal team leader comes with a heavy cost to individual performance. The captain designation can seem like a weight that pulls a player down. As one NHL team coach put it: "Give them the 'C' and watch them falter" (LaPointe, 1997, p. C6). Others have noted: "The captaincy is a burden that weighs heavier on some players than others" (Kreiser, 2001, para. 26). Of central concern is whether this implicit role overload perspective is supported by objective data. If so, we would expect that the years in which a player was a captain would be associated with worse performance than those years that he was not a captain (i.e., intraindividual differences).

An alternative perspective is offered by the role accumulation literature (Marks, 1977; Sieber, 1974) in which the raised expectations associated with being named as the formal team leader results in a type of self-fulfilling prophecy. It is also possible that multiple roles expand a leader's potential resources with regard to performance. In particular, salience, status, and opportunity are important resources that could be gained by becoming team captain. Being a salient player on the ice by benefit of the captain role (and visible sweater designation) means that other players on the team possibly are more aware of the captain and provide him with greater opportunities to score and to assist others in scoring, which are the two primary components of the most widely accepted criterion for assessing NHL player performance (i.e., points). In this way followers help to complete the self-fulfilling prophecy circle, guided by their attributional processes and heroic conceptualizations of leadership (Meindl et al., 1985).

It is also important to acknowledge that whether team leadership creates a burden or a benefit may depend a great deal on the role context. In one context, a team leader might be provided with greater knowledge of problem-solving techniques, obtain beneficial information, or receive actual assistance from others that would enhance individual performance. In hockey, this might take the form of greater opportunities to score or assist and being the recipient of other forms of on-ice help, such as blocking and general protection from overly physical players on the opposing team. In some contexts, however, the formal leadership role may provide no direct benefit for performing one's job duties. Indeed, it may add role responsibilities that can create conflict. The conventional wisdom in the context of professional hockey is that these leadership role responsibilities create some extra burden that can interfere with the individual performance of those team leaders (Kreiser, 2001; LaPointe, 1997). This perspective has been echoed by leadership researchers who have argued that taking on a formal leadership role usually involves some costs to individual leaders, and, therefore, there must be some individual motivation to lead to overcome these perceived costs (Chan & Drasgow, 2001). We put these contrary perspectives to an empirical test by examining the effects of formal leadership status on individual performance among NHL captains.

Of interest in this study was modeling the performance trajectories of individuals in which there was a naturally occurring intervention that all of them experienced—being named as team captain. What differentiates this study from others in the dynamic criteria literature is this naturally occurring intervention or "treatment" (Raudenbush, 2001, p. 520) that occurs at least once during every career in this sample of team captains. Examining intraindividual change factors (e.g., leadership status as a time-varying covariate) and interindividual predictors of change (e.g., individual difference variables) allowed us to test whether being a team leader was a detriment or benefit to individual performance, or if it had no reliable effect. The primary research questions that were addressed in the study were: (a) What was the overall individual performance trend (positive, negative, or zero) following the years after being named as a captain for the first time? (b) What was the effect of formal leadership role responsibilities on individual performance (positive, negative, or zero)? (c) Were there individual differences that predicted interindividual performance, changes in performance (if any), or the effect (if any) of being a leader?

Methods

Sample

Data for this study were obtained on 201 players in the NHL who also met two additional criteria. First, the player was a captain for a period of 1 year or more in the course of his NHL career. Second, the player must have played in the modern era of the NHL (beginning with 1967–68 season). The modern era corresponds to the first league expansion as well as when more precise player statistics were recorded. In order to enhance the comparability of statistics across players and years, we chose to focus on those captains solely from the modern era. Other steps were taken to enhance comparability on the primary dependent variable (described in the section titled *Dependent Variable*). A list of team captains by season was obtained from the NHL Hall of Fame. Information on each captain including yearly performance statistics was compiled from *Total Hockey: The Official Encyclopedia of the National Hockey League* (Diamond, 1998).

All players were male. At the time that this dataset was compiled, 64 of these players had completed their careers, whereas 137 were still active in the NHL (though not all were still captains). Players held diverse positions on the team: center ($n = 57$), right/left wings ($n = 70$); defense ($n = 62$); and those playing in more than one scoring position ($n = 12$). Thus, approximately 70% of the players were in offensive positions. There were no goalie captains because NHL rules prohibit this. The analyses for this study used only the data beginning with the first time period that each player was appointed as team captain (with time coded as 0) with subsequent seasons coded as time = 1 through k. The total number of nonindependent seasons was 1,148. The number of seasons that an individual served as a team captain ranged from 1 to 10 in this sample.

Players were in the NHL during different years, ranging from the 1960s to 1998. Individual players also differed in terms of the number of years spent in the NHL. In addition, they attained captain status at different times in their careers and had a broad range of total career seasons, both of which led to a wide range of seasons prior ($M = 7.13$, $SD = 3.85$, range = 0–21) and subsequent ($M = 2.71$, $SD = 2.15$, range 0–9) to the first captaincy. Because of the small number of players who were still in the league 10 or more years after first being named captain ($n = 23$) and that very few were still captains in years 10 or beyond ($n = 2$), estimating models with complex error structures was not possible because the results would not converge on a stable solution with these data included. Thus, we chose to include only the initial captain year and 9 subsequent years in our model tests.

According to published announcements of NHL team captains, there is no apparent consistency across the league in terms of how or why particular players are chosen as captains. There is no standard set of criteria nor is the process of selecting a team captain universal across all NHL teams. What appears to be the norm, however, is that the decision is made by the team's management and usually in some combination of the coach, general manager, vice president, and team president. The captain does not have to be the best player on the team, although there appears to be a trend toward naming the teams' best young player as captain early on in his career "usually because team management wants to establish a new order in the locker room" (Kreiser, 2001, para. 23). But according to one NHL franchise president and general manager, "The one thing that jumps out about a captain is that you want guys who really care about your team, care about the players, and care about the organization" (Kreiser, para. 9).

Dependent Variable

The sole dependent variable was each player's adjusted points for the season. Points are computed as the sum of player goals plus assists and are considered to be the most important statistic of individual player performance (Diamond, 1998).[1] Critical issues associated with estimating intraindividual and interindividual performance trends were that the number of games in a season has increased since 1967, the average number of goals scored varied across seasons, and there were both within- and between-player differences in the number of games played in a given season. Thus, to enhance performance comparability across players and time periods, player points were adjusted for: (a) number of total games in the respective season; (b) average number of goals per game in the respective season; and (c) number of games that the player appeared in the season.[2] The adjusted points statistics reported in *Total Hockey* (Diamond, 1998) were already adjusted for (a) and (b). We further adjusted these data by dividing the *Total Hockey* adjusted points statistic for each player by the number of games that he appeared in each respective season (c). The resulting statistic was interpretable as the adjusted average number of points earned by a player per game for a given season. By adjusting gross points by these factors, a fairer comparison could be made between and within players as they appeared in different numbers of games over different seasons. Given that all performance and adjustment data were taken from NHL records, the criterion was highly reliable.

Analyses

Important to this study was the expectation that individual performance changed over time; thus, a central focus of this study was modeling change with time as an independent variable. Given the repeated measures nature of the data, there were likely to be heterogeneous errors across time periods as well as correlated errors between times (i.e., autocorrelation). Thus, standard regression procedures were inadequate and inappropriate for these data. For these reasons, multilevel growth modeling offered a preferred approach. Because of nonbalanced and missing data (i.e., captains played different numbers of seasons and were appointed captain at

[1]Some have proposed that a player's plus/minus ratio is the most important individual statistic in professional hockey. It is operationalized as the ratio of goals scored to those given up by a team when a particular player is on the ice. Whereas some have argued that this is a key statistic of player impact, others have countered that it is unduly biased by team quality and team performance. Analyses that we conducted with plus/minus ratios as the dependent variable indicated no significant effects once team performance was considered as a covariate. The evidence, therefore, suggests that plus/minus is biased by overall team quality (performance). This was not the case for individual performance as measured by adjusted points. Including team performance in those analyses did not substantively change those results.

[2]The results were substantively the same if the adjusted points criterion was not corrected for the number of games played in a season.

different times and different numbers of times), the particular growth modeling technique we used was random coefficients modeling (RCM).

The RCM technique is considered to be more robust than structural equation modeling (SEM) techniques in applications where there are missing data, although recent advances in SEM applications have "narrowed the gap" (Schnabel, Little, & Baumert, 2000, p. 12). RCM models are also more flexible in terms of modeling time. Most approaches based on SEM (e.g., latent growth curve modeling) require completely balanced and nonmissing data, which was a requirement that was not met in this study. Furthermore, it is a requirement that would not likely be met in many studies using real-world data across numerous time periods. In addition, captain status (i.e., captain or not a captain in a given season) was modeled as a naturally occurring time-varying covariate, which could not be incorporated easily into SEM. We followed the specific model-building steps outlined elsewhere (Bliese & Ployhart, 2002), with particular attention to modeling and interpreting the effects of complex error structures on the parameter estimates for time and captain status.

Level-1 Predictors

Level-1 analyses were conducted at the intraindividual (within-person) level. The predictors that were modeled in the analyses were time and captain status. All were included in the model as uncentered predictors.

Time was coded as 0 for each player commencing with the first year that the player was appointed captain. All players did not appear in all seasons. There was also great variability in terms of which players were captains in which particular years. The time-varying covariate in the model (captain status) captured whether or not a player was a captain in a given season. All players in the sample ($n = 201$; 100%) were captains at some point in their respective careers (modeled as time = 0); however, captain status varied substantially after that. For Seasons 1 through 9 following initial captain appointment, the number and proportion of captains in the sample for the respective seasons were as follows: $n = 124$ (67%), 86 (51%), 62 (43%), 46 (39%), 36 (36%), 23 (28%), 22 (34%), 16 (32%), and 11 (26%). Descriptive statistics and performance intercorrelations across the 10 seasons are reported in Table 1, along with correlations with captain status.[3]

[3] Although some of the interseason performance correlations reported in Table 1 appear smaller for those years involving Season 9 than the intercorrelations for other seasons, supplementary analyses (available from the first author) demonstrated that there were no overall differences in model test results when Season 9 data were excluded. We chose to include data from this season to be as inclusive as possible in terms of modeling the available career performance data.

TABLE 1

Performance Descriptive Statistics and Intercorrelations by Season

Season	*N*	*M*	*SD*	0	1	2	3	4	5	6	7	8	9
0[a]	201	.61	.35	–									
1	185	.58	.35	.82**	–								
2	167	.59	.37	.77**	.83**	–							
3	145	.56	.35	.75**	.79**	.83**	–						
4	119	.55	.36	.74**	.77**	.79**	.75**	–					
5	101	.56	.36	.76**	.83**	.81**	.79**	.86**	–				
6	83	.55	.37	.79**	.83**	.82**	.77**	.76**	.83**	–			
7	64	.57	.41	.80**	.81**	.79**	.77**	.75**	.85**	.85**	–		
8	50	.52	.36	.73**	.76**	.77**	.74**	.72**	.83**	.85**	.88**	–	
9	43	.57	.37	.60**	.68**	.61**	.63**	.67**	.70**	.71**	.77**	.77**	–
Captain status (*n*)	NA[b]	–	–	NA[c]	.15* (124)	.29** (86)	.33** (62)	.33** (46)	.19† (36)	.39** (23)	.34** (22)	.33* (16)	.24 (11)

Note. Performance = adjusted points for season. Captain status coded as 0 = *not a captain*, 1 = *captain*. Captain status n = number of captains in a season.

[a]Initial captain season.

[b]Captain status varies by season.

[c]All players were captains (no variance).

$^{\dagger}p < .10$ $\quad ^{*}p < .05$ $\quad ^{**}p < .01$. All tests 2-tailed.

TABLE 2

Descriptive Statistics and Intercorrelations for Level-2 Predictors

Variable	*M*	*SD*	1	2	3	4	5
1. Height	71.90	2.24	–				
2. Weight	192.27	14.10	.65***	–			
3. Position	1.31	0.46	.19**	.29***	–		
4. Age	27.66	3.98	−.16*	−.23**	−.01	–	
5. Experience	7.13	3.85	−.08	−.14*	−.03	.88***	–
6. Captain frequency	3.17	2.44	−.01	.14*	−.09	−.27***	−.19**

Note. $N = 201$. Height measured in inches. Position coded as 1 = *offense* and 2 = *defense*. Experience measured as the number of seasons in NHL before becoming a captain. Captain frequency measured as the number of seasons as captain.

$^{*}p < .05$ $^{**}p < .01$ $^{***}p < .001$. All tests 2-tailed.

Level-2 Predictors

Level-2 analyses were conducted at the interindividual (between-persons) level. Individual difference variables at Level 2 were used to estimate intercept and slope differences at Level 1. The Level-2 predictors that were included in this study included two measures of physical stature (player weight and player height in inches), player position (offense or defense), number of years played before becoming a captain, age at the time of first captaincy, and the total number of times as captain during the player's career. Players who were captains only once ($n = 59$) comprised approximately 29% of the total sample, whereas those who were captains two times were approximately 22% of the sample ($n = 44$), those who were captains three times comprised 16% of the sample, and players who were captains four or more times in the course of their careers consisted of 33% of the sample. Descriptive statistics and intercorrelations among the Level-2 variables are presented in Table 2. All between-player variables were assumed to be time invariant (i.e., single value for each player) and were taken from the first season each individual became captain. One exception was the number of times appointed captain, which was a time-invariant variable when calculated across all seasons.

Results

Random coefficients modeling was implemented using the hierarchical linear modeling (HLM) statistical program to simultaneously estimate the intraindividual (Level-1) and interindividual (Level-2) effects. Bliese and Ployhart (2002) outlined a series of recommended steps in building a Level-1 model, beginning with a simple regression model and ending with a complete growth model by successively adding complexity to the model. The Level-2 model was estimated only after the most appropriate Level-1 model had been determined. We estimated all models using either two-level hierarchical linear model (HLM-2) or hierarchical multivariate linear model (HMLM) procedures (Raudenbush & Bryk, 2002). HMLM allows for the estimation of more complex Level-1 error structures than does HLM-2 and was used to estimate the effects of first-order autoregression and heterogeneous variances and covariances on the model effect estimates.

Determine the Level-1 (Intraindividual) Model

Estimating the ICC. As a preliminary step, it is recommended that an intraclass correlation (ICC) be estimated from a random-intercepts model to assess the strength of nonindependence in the data. In this case, an ICC estimates the proportion of variance in an outcome that is between individuals. A relatively large ICC indicates that there are likely to be between-person (interindividual) effects that can be modeled at a higher level with Level-2 analyses and that it is appropriate to use a random-intercepts model. For this study, the ICC estimate was .625 indicating that approximately 63% of the variance in the dependent variable (adjusted points) was attributable to interindividual differences. This suggested a fairly substantial degree of nonindependence in the data and that using a random-intercepts model was necessary.

Determine fixed functions for time and captain status. Given the relatively large ICC estimate, the next step was to estimate the fixed functions for time and captain status in a random-intercepts model (i.e., a different intercept term estimated for each captain in the sample). At this stage, Bliese and Ployhart (2002) also recommend determining whether a linear, quadratic, cubic, or other higher order function best models the time variable. Results (see Table 3) indicated that the only significant time function was linear and negative ($t = -6.44$, $p < .001$). There were no significant quadratic or cubic effects. We subsequently estimated a Level-1 model that added captain status (0 = *player*, 1 = *captain*) as a fixed effect. Results indicated a significant positive effect for captain status ($t = 2.63$, $p < .01$). These results suggested that despite the overall negative trend across time, being a captain was associated with better intraindividual performance. It is important to remember that this is a within-person effect: Relative to other seasons, those years in which a player served as captain were associated with better performance than those years in which he was not a captain. The correlations reported in Table 1 also suggested a robust interindividual effect for captain status. In seven of the seasons, there was a statistically significant ($p < .05$) bivariate relationship between captain

TABLE 3

Hierarchical Linear Modeling Level-1 Analysis: Unconditional Model for Season Performance

Effect	Coefficient	*SE*	*df*	*t*
Fixed effect				
Mean performance as first-time captain (β_{00})	.577	.028	200	20.79***
Mean performance trend (β_{10})	−.024	.004	200	−6.44***
Mean captain effect (β_{20})	.045	.017	200	2.63**
	Variance component	*SD*	Deviance	ΔDeviance
Random effect				
Initial performance (r_{0i})	.101	.317	−49.53 (2)	936.47[a]***
Performance trend (r_{1i})	.001	.023	−64.10 (4)	15.47**
Captain status (r_{2i})	.004	.065	−66.87 (7)	1.87
Level-1 error (e_{1i})	.032	.178		

Note. Reported coefficients and variance components based on sequential model steps (Bliese & Ployhart, 2002). Deviance = −2 log likelihood value.

[a]Comparison with simple model with all fixed effects.

$^{*}p < .05$ $^{**}p < .01$ $^{***}p < .001$. All tests 2-tailed.

status and performance and, in one of the other two seasons, the correlation was marginally significant ($p < .10$).

Determine variability in growth parameters. The next step involved testing whether there was significant interindividual variability in the Level-1 parameters to justify including Level-2 predictors. Tests were conducted by comparing likelihood ratios using chi-square difference tests among successively more complex models. The baseline model was the random-intercepts model, testing first for significant slope variability on the time variable and then on the captain status variable. Likelihood contrasts between increasingly complex models (random intercepts; random intercepts and time slope; random intercepts, time slope, and captain slope) indicated significant slope variability for time, $\chi^2(2) = 15.47$, $p < .01$ but no significant random variability for captain status, $\chi^2(3) = 1.87$, $p > .05$. These findings suggested that there were interindividual differences in the initial performance of team leaders and that there were interindividual differences in the performance slopes over time but that the significant effect for captain status did not vary significantly among players. Being appointed as formal team leader (i.e., captain) apparently had a similar, beneficial effect on performance regardless of the player or any individual differences. Hence, the unconditional RCM model for the present data can be expressed as:

Level 1 (intraindividual):

$$\text{Adjusted\_points}_{ij} = \beta_{0j} - \beta_{1j}\text{Time}_{ij} + \beta_{2j}\text{Captain\_status}_{ij} + r_{ij}, \quad [1]$$

Level 2 (interindividual):

$$\beta_{0j} = \gamma_{00} + \cdots + \gamma_{0k} + u_0 \quad [2]$$

$$\beta_{1j} = \gamma_{10} + \cdots + \gamma_{1k} + u_1 \quad [3]$$

$$\beta_{2j} = \gamma_{20} \quad [4]$$

In the preceding formulae, adjusted points is the Level-1 dependent variable with an intercept (β_{0j}) and two estimated slopes coefficients (β_{1j}, β_{2j}) plus an error term (r_{ij}). At Level 2, the intercept and one slope coefficient from Level 1 are used as dependent variables. Level-2 predictors are presented as γ_{00} and γ_{10} (to γ_{0k} and γ_{1k}, respectively, indicating that each Level-2 outcome could be modeled with multiple predictors). The Level-1 intercept term (β_{0j}) and the Level-1 time slope (β_{1j}) were estimated as random effects including error (u_0 and u_1, respectively). Because the captain-status effect (β_{2j}) was treated as a fixed effect (no error or variability in its effects) in the Level-1 analyses, it was not included as a Level-2 outcome.

Determine the error structure. In modeling alternative covariance structures, several growth models were estimated. An important assumption in conducting these analyses was that the "complete data" were balanced. That is, it was possible to have data missing at random such that the number of seasons per person varied (e.g., not all players had data for all 10 seasons); however, the captain status variable had different distributions across players. For that reason, the complete data were unbalanced. Because of this constraint, the captain status effect could only be considered as a fixed effect in the model estimations. However, the Level-1 analyses suggested that captain status was most appropriately modeled as a fixed effect so there was no subsequent loss of meaning by considering captain status as a fixed effect in the analyses that tested various complex error structures.

The simplest error structure was associated with a *compound symmetry* model. This is considered to be a special case of the random intercepts model in that it assumed that random effects were independent with homogenous Level-1variance (σ^2) and that all participants had the same linear slope. This was considered to be the baseline comparison model. The *first-order-autoregressive* or *AR(1)* model added an autocorrelation term to the compound symmetry model. The next two models estimated *random time slope with autocorrelation* and *random time slope with heterogeneous level-1 variance* (σ_t^2). The final model was an *unrestricted model* that included 58 unique variance–covariance parameters that were estimated. It was, therefore, considered to be the most complex error structure against which one could test the fit of more parsimonious submodels.

Table 4 presents the model estimates and fit statistics for the five estimated models. Looking across the five sets of Level-2 estimates (top half of the table), the overall intercept (β_{00}) was roughly similar (ranging from .573 to .588) with the standard errors approximately the same (rounded to .027). Estimates of the average growth rate (β_{10}) were also similar with very similar standard errors. There was an apparent trend in the estimates for the fixed effect for captain status (β_{20}) to become progressively smaller across models from the least complex (compound symmetry) to the most complex (unrestricted) error structures. These results suggest that coefficients for intercept and growth rate (time) were relatively insensitive to the choice of model, but the fixed Level-2 coefficient for captain status was somewhat overestimated in the simpler error structure models.

In terms of model fit comparisons (lower half, Table 4), the unrestricted model that allowed separate estimation of 58 parameters (10 variances, 45 covariances, and 3 Level-2 coefficients) provided the best overall fit to the data (-2 log likelihood ratio (deviance) $= -200.41$) as compared with any of the more parsimonious models (all chi-square differences were statistically significant, $p < .05$). The finding that the unrestricted model best fit the data was not surprising given the large number of parameters used to estimate the model.

Summary of the Level-1 Model Results

The results from the HLM-2 and HMLM analyses suggested that there was an overall significant decrease in individual performance over time (after being named a team captain initially); however, being the formal leader of a team was associated with better intraindividual performance. Team leaders (captains) differed in their individual performance at the time of the first captaincy (intercept differences), and there were between-player differences in terms of the underlying growth patterns (time). Results suggested that the beneficial effects associated with leadership on individual performance did not vary significantly across individual players.

Supplementary Level-1 Analyses

Prior to building the Level-2 (interindividual) model, an unresolved issue was examined in a supplementary Level-1 analysis. By conducting the additional analysis, we hoped to eliminate an alternative explanation

TABLE 4

Hierarchical Multivariate Linear Modeling Analyses: Comparison of Five Models with Different Error Structures

	Fixed slopes				Random slopes					
	Compound symmetry		AR(1)		AR(1)		Heterogeneous		Unrestricted	
	Coeff.	*SE*	Coeff.	*SE*	Coeff.	*SE*	Coeff.	*SE*	Coeff.	*SE*
Level-2 coefficients										
Average intercept, β_{00}	.573	.027	.575	.027	.580	.027	.588	.027	.577	.027
Average growth rate, β_{10}	−.022	.003	−.022	.003	−.024	.003	−.025	.003	−.024	.003
Average captain effect, β_{20}	.049	.016	.045	.016	.041	.016	.032	.015	.029	.015
Variance–covariance component										
Level 2	$\hat{\tau}_{00} = .095$		$\hat{\tau} = .092$		$\hat{\tau} = .096$	−.001	$\hat{\tau} = .100$	−.001		
						.0003		.0004		
Level 1	$\hat{\sigma}^2 = .034$		$\hat{\sigma}^2 = .036$		$\hat{\sigma}^2 = .033$		$\hat{\sigma}_1^2 = .027$		.110 .093 .097 .087 .092 .085 .086 .102 .069 .070	
			$\hat{\rho} = .142$		$\hat{\rho} = .075$		$\hat{\sigma}_2^2 = .025$		.130 .114 .098 .104 .109 .107 .106 .087 .093	
							$\hat{\sigma}_3^2 = .025$		.143 .107 .114 .114 .111 .116 .090 .084	
							$\hat{\sigma}_4^2 = .027$		.119 .099 .097 .091 .101 .068 .074	
							$\hat{\sigma}_5^2 = .064$		.174 .117 .105 .104 .069 .116	
							$\hat{\sigma}_6^2 = .021$		.133 .108 .114 .092 .104	
							$\hat{\sigma}_7^2 = .024$		.132 .112 .096 .105	
							$\hat{\sigma}_8^2 = .052$		.170 .097 .103	
							$\hat{\sigma}_9^2 = .027$		.110 .088	
							$\hat{\sigma}_{10}^2 = .042$		.150	
Model fit										
Deviance (*df*)	−71.92 (5)		−83.26 (6)		−89.28 (8)		−137.83 (16)		−200.41 (58)	

for the previously reported results demonstrating that captain status was positively related to individual performance. The issue pertains to the potentially confounding role of prior performance on those results. Without somehow taking into account a player's prior performance, it is impossible to rule out the possibility that the captain status variable was merely reflecting the level of previous performance. Those players (or captains) with superior performance possibly were appointed as captains in a subsequent season and those with inferior performance were not appointed or reappointed. In other words, captain status likely covaried with previous performance and the results reported previously mainly reflected the role of prior performance on subsequent performance. We undertook the supplementary analyses to address this possibility directly.

In order to address this issue, a follow-up analysis was conducted that also included the time-varying covariate of a player's performance (adjusted points) in the prior season, in addition to time and the captain status variable as used in the previous Level-1 analyses, to predict individual performance in the subsequent season. Put in more technical terms, prior performance was included in the analysis as a T-1 (where T = time) lagged covariate. Thus, the effects of captain status on individual performance could be examined while controlling for performance in the previous season.

Based on an unrestricted model estimation procedure, the results indicated a strong, positive effect of performance in the previous season on individual performance ($t = 33.53$, $p < .001$). The effect for time was no longer statistically significant once prior season's performance was considered ($t = -0.51$, $p > .10$), which makes sense in that controlling for previous performance accounts for much of the performance change over time. The key test was for captain status, and those results were in the same positive direction and statistically significant even after controlling for previous performance ($t = 3.51$, $p < .01$). Thus, the results from these supplementary analyses support the conclusion that captaincy is a positive predictor of performance above and beyond the effects of prior performance.

Building the Level-2 (Interindividual) Model

The final model step involved estimating the intercept- and slopes-as-outcomes models. Individual differences variables were included in the conditional model as predictors of intercepts differences as well time-related slope differences between players. Models were estimated using HMLM with an unrestricted error structure (see Table 5 for results).

Intercepts as outcomes. The Level-2 results for the intercept differences indicate that taller players had lower performance during the initial

TABLE 5

Hierarchical Multivariate Linear Model Level-2 Analyses: Conditional Model for Performance

	Initial status (π_{01})		Performance trend (π_{11})	
Fixed effect	β_{0j}	$SE\ \beta$	β_{1j}	$SE\ \beta$
Intercept-2	1.063***	.057	−0.06***	.009
Height[a]	−.021*	.01	−.001	.001
Weight[a]	.001	.002	.001	.0002
Position[b]	−.365***	.04	.021**	.006
Seasons played before captain[b]	.029**	.01	.002	.001
Age at first captain[a]	−.043***	.01	−.005**	.001
Number of times as captain[b]	.046***	.008	.001	.001

Note. $N = 201$. Height measured in inches. Position coded as 1 = *offense* and 2 = *defense*.

[a]Grand-mean centered.

[b]Uncentered.

$^{*}p < .05$ $^{**}p < .01$ $^{***}p < .001$. All tests 2-tailed.

captain year than shorter players ($t = -2.05$, $p < .05$), which may be a function of player speed and agility. Shorter players (who also tend to play offense; see Table 2) may be faster and more agile skaters, which allows them better get into position to score or assist. There was a stronger effect for player position, and as expected, offensive players had higher initial individual performance than defensive players ($t = -9.02$, $p < .001$). The number of seasons played in the NHL before becoming a captain (i.e., experience) was positively related to performance in the initial season as captain ($t = 2.95$, $p < .01$), whereas player age was negatively related to initial status ($t = -4.41$, $p < .001$). The only other significant Level-2 predictor of the intercepts-as-outcome analysis was found for the number of times a player was a captain throughout his career ($t = 5.68$, $p < .001$). Because of the temporal issues associated with these particular variables, the only interpretation that makes much sense is that captains with higher initial performance were more likely to be named as captains again in the future as compared with captains with lower initial performance. The physical weight of the player was unrelated to initial performance levels as captain.

Slopes as outcomes. The Level-2 results in predicting interindividual differences performance trends indicated a significant effect for position ($t = 2.47$, $p < .05$), suggesting that defensive players had steeper negative performance trends over time than offensive players. The only other significant effect was found for age at the time of initial captain appointment ($t = -3.11$, $p < .01$), suggesting that older players had a flatter performance trajectory than did younger players. There were no significant effects for height, weight, experience, or number of times as captain in predicting variability in performance trends.

Hence, the final Level-2 (interindividual) model can be expressed as:

$$\beta_{0j} = \gamma_{00} - \gamma_{01}\text{Height} - \gamma_{02}\text{Position} + \gamma_{03}\text{Experience}$$
$$- \gamma_{04}\text{Age} + \gamma_{04}\text{Captain Frequency} + u_0 \quad [5]$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}\text{Position} - \gamma_{11}\text{Age} + u_1 \quad [6]$$

where β_{0j} is the intercepts-as-outcome and β_{1j} is the slopes-as-outcome for time.

Discussion

This study addressed three particular research questions pertaining to leadership and longitudinal performance trends within a dynamic criteria framework: (a) What was the overall average performance trend in the years following initial leadership role responsibilities? (b) What were the effects of formal leadership role responsibilities on intraindividual (within-person) performance? (c) Were there interindividual (between-person) factors that predicted performance, changes in performance, or the effect (if any) of being a team leader? These questions were addressed using data obtained from all team captains in the modern era of the National Hockey League (NHL).

Performance Trends

Results of the multilevel growth modeling analyses were clear that performance changed over time (i.e., evidence of dynamic criteria) and that the overall performance trend was negative. Performance was measured beginning with the first time each player was appointed as team captain (i.e., formal leader) through as many as nine subsequent seasons. Rather than suggesting that this negative performance trend was due to the so-called burdens of leadership that created a detriment to individual performance over time, a more reasonable interpretation is that players tended to be appointed as captain at or near the zenith of their respective careers. The negative performance trend is therefore likely a function of aging. Hockey is a physical sport and age-related decrements in physical abilities over time affect individual player careers. Supplementary analyses of precaptain performance (available from the first author) indicated a general positive slope for player performance prior to being appointed as a formal team leader for the first time. Taken together, the overall career performance trajectory of the population of NHL captains from the modern era of the league can be characterized as curvilinear (specifically, an inverted "U"), with the asymptote centered on or near the season of initial captaincy. Performance was dynamic over time and significant overall trends were identified. We are unaware of any published studies that have examined career trajectories in other nonsports domains. Thus, we cannot say with any certainty whether business leaders or corporate executives attain top leadership positions at or near the crest of their respective "games" or whether retirement comes after a gradual performance decline over time. This would be an interesting topic for future research.

The "Burdens" of Leadership

Formal leadership role responsibilities were modeled as a time-varying covariate in a model that also included random intercepts and random time slopes. Leadership in the form of being appointed as team captain was a naturally occurring intervention or "treatment" (Raudenbush, 2001) that affected each player at least once and as many as 10 times. At initial status (time = 0), all players were captains; however, the distribution of captains varied across subsequent years. Random coefficients modeling allowed us to analyze captain status as an intraindividual effect across as many as 10 consecutive seasons. Results indicated that the time-varying covariate was significant and positive, suggesting that those seasons in which a player had formal leadership responsibilities were associated with better individual performance than those seasons without leadership responsibilities. These results held up even when the prior season's performance was included in the model, suggesting that the captain status variable was not merely modeling the effects of previous performance. The proper interpretation of these results is at the intraindividual level.

Correlational analyses indicated that there was an interindividual component to this effect as well. Correlations between performance and captain status were statistically significant in seven of nine time periods and marginally significant in one of the other two seasons (there was no variance on captain status in the initial year because everyone was a captain). These results support the inference that taking on formal leadership responsibilities in a team is associated with better individual performance relative to other noncaptain seasons (within-person) and is associated with better performance relative to other players without formal leadership responsibilities (between-person).

The results of model comparisons also indicated that the beneficial effects of leadership on individual performance were fixed. This means that the data did not support a slopes-as-outcomes model for captain status. Assuming a formal leadership role on a team was associated with better individual performance regardless of any interindividual differences such as physical stature, position, age, experience, or frequency of being a captain. This effect was robust even when the model was estimated with complex error structures associated with first-order autoregression and heterogeneous variances and covariances.

Although we are ultimately left to speculate on the underlying processes for this effect, it is feasible that being a formal leader makes one's presence salient to teammates, which translates into greater opportunities to contribute to the team's success (and enhances individual performance). It also is possible that because of the historical distinction accorded to NHL captains by the league, that self- and others' expectations are enhanced, creating a self-fulfilling prophecy (Eden, 1990; Eden & Shani, 1982; Rosenthal & Jacobson, 1968). Teammates might perceive and interact with a formal leader in primarily positive ways because of the heroic and romanticized notions of leadership that are part of the social culture (Meindl et al., 1985), which is emphasized in the NHL in ways that accord high status and visibility to team captains. These findings failed to support the perspective that taking on additional leadership roles resulted in poorer individual performance. Nonetheless, perceptions or feelings of role overload were not measured directly in this study, and without such data, we cannot rule out the possibility that role overload was experienced. Without such data it also could not be concluded that a role accumulation hypothesis was unequivocally supported. What can be concluded with certainty is that being appointed as team leader in the NHL was found to be associated with better and not worse individual performance, regardless of interindividual differences or performance in the previous season.

Despite the limitations surrounding the lack of process measures in this study, these findings are at least consistent with the results of recent research on the beneficial effects of multiple roles for managerial women. Ruderman et al. (2002) reported that managerial resources (psychological resources and social support) could be enhanced through commitment to multiple roles. Consistent with the role accumulation literature (Marks, 1977; Sieber, 1974), resources that accumulate through one role enhanced coping in a different life role. Future research is needed to replicate the present results in other samples of team leaders and to take a closer look at the underlying processes that might better explain such beneficial effects on individual leader performance in team contexts.

Interindividual Differences in Performance Trends

In addition to the intraindividual effect for time that suggested an overall negative performance trend, results also indicated significant random effects for the initial status (intercept) as well as the performance trend (slope). These findings provided evidence of possible interindividual differences that were subsequently examined using higher level intercepts- and slopes-as-outcomes models. Results suggested several interindividual factors associated with initial status differences (height, position, experience, age, number of times as captain), yet, there were only two significant effects associated with performance trend differences. Defensive players—comprising approximately 30% of the captain population—tended to have steeper drops in performance than offensive players, which is at least partially due to the nature of the criterion. Although defensive players also can accumulate points through scoring and assists, the initial status data suggests relatively large differences between offensive and defensive players in the adjusted-points criterion. The steeper decline in performance over time among defensive players may be a function of their ability to contribute to the team in ways other than scoring and assists, such as preventing the other team from scoring. Offensive players, however, make their team contributions mainly through scoring goals or assisting others. If offensive performance drops too much then they are unlikely to remain in the league for very long.

It is worth noting again that the beneficial effect associated with captain status on the criterion of adjusted points was applicable for both offensive and defensive players. Even those players whose primary performance roles do not involve scoring or assisting performed better in those areas when serving as the team leader. The interindividual results also indicated that older players had flatter overall performance trajectories than did younger players. Players who were older at initial status perhaps had already begun their performance decline resulting in negative performance trends that were not as steep as for younger players just reaching their peak.

Perhaps the most important aspect of the interindividual results is that few individual difference variables contributed to the slopes-as-outcomes models. This may be a function of the lack of access to individual difference data. Cognitive and psychological individual difference variables (e.g., personality data) might be more robust predictors of interindividual slopes than the physical characteristics, player position, and other historical data that were available from archival sources. Another important point is that the time-varying covariate of captain status was robust in its positive effects on individual performance. The one factor that appeared to be consistently related to reducing the noted performance declines over time was serving as the formal team leader. The advantages of leadership apparently outweigh its costs when it comes to individual performance in the NHL. Given the longitudinal nature of this study and the consistently positive effects of leadership status on performance across numerous time periods, it might be tempting to conclude that leadership status caused better individual performance. That would be premature. The nature of causal inferences with time-varying covariates is a relatively new area of methodological inquiry. According to Raudenbush (2001):

> The problem of causal inference for time-varying treatments, assuming no randomization, is extraordinarily challenging. It is a cutting-edge issue on which progress is being made, but methodological workers in this area have not yet achieved a consensus on the best ways to proceed. (p. 522)

For these reasons, it is best to be conservative with regard to any possible causal inferences and conclude that there is a positive relationship between leadership status and individual performance over time. The specific causal nature of that relationship awaits further scrutiny.

Potential Study Limitations and Implications

This study contributes to a better understanding of the longitudinal relationship between formal leadership role responsibilities on individual performance. It is a question that has not been examined previously in the leadership, job performance, or role literatures, in that theory and research have focused mainly on the relationships between leadership and team or organizational performance. There are compelling reasons to examine the links between leadership responsibility and individual performance given the assumption on the part of scholars (Chan & Drasgow, 2001) and practitioners (LaPointe, 1997) that leadership responsibilities come with some cost to individual performance. That assumption, however, was not supported by these results. Nonetheless, there are certain study limitations and corresponding implications that should be addressed in conjunction with those conclusions.

Generalizability of results. Questions might be raised about the generalizability of these results to other types of teams. Although NHL teams are certainly a special type of organizational team and NHL athletes represent elite sports professionals, there is a rich history of using data from the domain of sports to examine organizationally relevant issues in the management sciences and sociology (e.g., Brown, 1982; Gamson & Scotch, 1964; Grusky, 1963; Pfeffer & Davis-Blake, 1986), as well as applied psychology (e.g., Dirks, 2000; Hofmann et al., 1992; Landis, 2001; Lord & Hohenfeld, 1979). Given the important roles that captains play with regard to team leadership in the NHL, the choice of that particular sporting domain makes sense in terms of the research questions that were studied. There are captains in other major team sports (e.g., football, baseball, and basketball), "but in no other is the Captaincy as important as it is in hockey" (Kreiser, 2001, para. 6).

Other advantages associated with using the data from the NHL include the important concern that individual player (and captain) performance was well measured and highly reliable across time periods given that it was based on objective data taken from NHL records. Another aspect that enhances potential generalizability is that individual performance is a function of task interdependence, which is characteristic of many organizational teams (Landis, 2001). Scoring a goal involves the contributions of teammates in terms of blocking, passing, and positioning. Earning an assist involves getting the puck to a teammate who subsequently scores a goal. Thus, even at the individual performance level, success involves the contributions of teammates. Creating an assist involves the individual ability of recognizing a player who is in a good scoring position, getting the puck to that player, and having that player actually score. For these reasons, this measure of individual performance could be construed as parallel to objective performance measures from many other types of work teams engaged in highly interrelated tasks. The advantage of using data from professional sports is that individual player statistics are meticulously recorded and could be adjusted for potential confounds. In short, the quality of the performance data is likely to be superior to what could be obtained from most nonsports domains.

A related concern is whether "real world" teams experience leadership changes that are similar to those in the NHL. In many ways the answer is yes. Teams in all kinds of organizations continuously form, perform, and disband in a cyclical fashion. The formal leaders of project teams change all the time such that a leader in one team may be just a "player" in another team. Leaders are transferred from units, departments, and projects creating similar kinds of leadership flows that were modeled in this study. In some other ways, however, the generalizability of these results is questionable. Hockey players are elite professional athletes who command top salaries even for those in the compensation stratosphere of pro sports. Their triumphs and failures are played out in a very public way: Their performance is viewed by many thousands of fans in person and considered by many more through various media. In addition, the internal leadership that team captains provide is a form of peer leadership. Others in the organization, most notably the coach and general manager, are also charged with leadership responsibilities of the team.

Another consideration is that that we studied leadership status and individual performance in action teams; thus, it is unclear how generalizable these results are to different kinds of teams (e.g., service, management, or parallel teams; Sundstrom, 1999). Furthermore, whether team leadership turns out to be a benefit or a burden is likely to depend on the role context. If the leader role is complementary to that of an individual contributor, then there is likely to be a greater benefit to performance than if those roles conflict. If a leader is burdened with extra responsibilities that take them away from or interfere with the individual contributor role (e.g., administrative duties), then there may not be the same kinds of benefits to performance that were found in this study. For these reasons, additional research is needed before any firm conclusions can be drawn regarding the generalizability of these findings, but there are reasons to be optimistic. Critical to the goal of replicating these findings in other domains and with other teams is developing psychometrically sound criterion measures of individual performance and to identify those contexts in which a team leader might draw resources that are beneficial for his or her individual performance. There also may be more specific points of comparison to work environments in which internal team leaders have highly specialized skills, such as engineering teams or high technology work teams.

Missing data. There might be expressed concerns about the effects of missing data on the results. Players did not all have equal numbers of seasons played, and the captain status variable was distributed differently across players. An important advantage of random coefficients modeling in studying change is the flexibility in its approach to handling missing data. Nonetheless, results are most robust when the data are assumed to be missing at random or missing completely at random. Neither of these missing data assumptions was likely met completely in this study.

It is typical that superior individual performers are selected as team leaders (regardless of context). Thus, prior performance and captain status probably are not completely independent, which we have attempted to address through supplementary analyses. Furthermore, players retire or are released when their individual performance no longer meets their or others' standards. As noted, hockey is a physical sport and as a player ages he becomes generally less able to tolerate the physical stresses of the game. Injuries also result in player retirement, but those might be considered as more random occurrences given that younger players also experience career-ending injuries. In cases of *nonignorable missingness,* results are robust to the extent that all data are efficiently used and the amount of missing data is relatively small (Schafer, 1997). Although missing data remains a potential issue with these results, the notion that leadership roles change over time and that people have different career lengths argues for the realistic nature of our data.

Leadership performance. Although individual performance was well measured in this study, the quality of leadership brought by the various captains remains unknown. Leadership is a difficult construct to measure, primarily because of challenges associated with prevalent approaches that are based on survey questionnaires, which are susceptible to various perceptual and schematic biases that people hold about the construct (Lord & Maher, 1991). Researchers have also tended to equate different leadership criteria such as leader emergence and leadership effectiveness

(Lord & Hall, 1992). Nonetheless, changes in captain status were likely a joint function of both individual on-ice performance as well as the more intangible leadership qualities of an individual. Although it was beyond the scope of this study, future research that examines possible predictors of leadership changes in a team (including the leadership skills and abilities of individual leaders) would be interesting. Latent transition analysis (Collins, Hyatt, & Graham, 2000), which expresses change in the form of transition probabilities, could provide a powerful tool toward this end.

Along these lines, there is an important set of unanswered questions regarding what are the specific factors that predict performance as an internal team leader. Specifically, it would be valuable to know the kinds of behaviors that enhance others' or overall team performance. As noted previously, there is "surprisingly little" known about how leaders contribute to team effectiveness (Zaccaro et al., 2001, p. 452). It would also be interesting to ascertain whether there was any relationship between leadership effectiveness and individual leader task performance. It may be that the best internal team leaders are those individuals who can maintain excellent individual productivity while also engaging in behaviors that enhance teamwork, team cohesion, and team performance.

Alternative analyses. It might be of interest to compare this sample with a matched group of players who were never formal leaders. The focal question would be addressing what are the performance trends of matched players who were never captains and do they differ systematically from the captains? In addressing this question, an important consideration would be the ability to effectively match captains with players who were never captains on important parameters that might affect performance (e.g., career length, team strength, mobility between teams). In addition, because these hypothetically matched players were never team captains, there would be no variability on the captain status variable to model as a time-varying covariate. Thus, an analytical procedure other than multilevel growth modeling would be needed to compare players across the two samples. A particular strength of this approach is that we have included all players who were a captain at least once in the modern era of the NHL. In any given year (except the initial status year in which all players were captain), the sample consisted of a mix of captains and regular players who were captains in a previous season(s), who served as interindividual comparisons. Thus, we believe that this provides the best and soundest test for the effects of formal team leadership responsibilities on individual performance over time.

Practical Implications

There are certain practical implications associated with these results that may be relevant to more general organizational teams. One of the possible reasons why the NHL captaincy is associated with a performance boost rather than a performance drain may have to do with the culture of captaincy or the culture of leadership that has been nurtured in the league. There is a common perception that being named team captain is a high honor as well as a big responsibility—"the 'C' is a symbol of their teammates' respect and their coaches' trust" (Kreiser, 2001, para. 6). Other organizations could learn from the NHL in terms of building a similar culture of trust and respect around their own team leadership roles. But, it is unlikely that this type of culture change develops quickly. The leadership culture in the NHL has been built over a period of more than 50 years, and it is quite possible that the beneficial captain effects found in this study did not occur immediately after the rule change in 1947. Nonetheless, an investment in developing a culture of leadership similar to that nurtured in the NHL may be a distinct benefit over time to other types of organizations.

Another potential practical implication is in terms of individuals' willingness to serve as a leader—what Chan and Drasgow (2001) have termed the *motivation to lead*. They point out that if the costs of leading are perceived to be high relative to the benefits, then individuals will not want to lead. This could be detrimental to the team and the organization in the long run and may even have important sociopolitical implications because of the personal demands that are placed on public leaders. But even in the very public world of professional hockey, these results indicate that the benefits might actually outweigh the costs to an individual leader. With the right kind of leadership culture in an organization, in conjunction with expectations of high levels of team member interdependence and that the performance of the overall team is seen as paramount, the motivation to lead could be enhanced among employees. A second-order effect of such a culture is in creating a deeper "bench strength" of potential leaders in an organization who are willing to take on formal team leadership responsibilities when needed.

Taken as a whole, these results may have value in terms of practical recommendations to managers in other kinds of organizations. In particular, there appear to be three interrelated recommendations to consider: (a) Make sure that the leadership position on the team is valued by the organization, (b) make sure the leadership position is salient to others on the team, and (c) maximize to the extent possible that leadership role requirements do not interfere with task requirements. Given that the causal linkages with regard to these results are speculative at this time, the usual caveats apply. Certainly more research is needed on the effects of leadership role responsibilities on the individual performance of team leaders, which as a topic has been virtually unexplored in the leadership literature.

Conclusions

In conclusion, this study adds to the literature demonstrating that leadership matters. Based on properly interpreted leadership succession studies, previous research has shown that leadership matters in enhancing organizational performance (Day & Lord, 1988; Thomas, 1988). This study shows that leadership responsibilities can be an important factor with regard to individual performance as well. The results of this multilevel growth modeling analyses were consistent with a role accumulation perspective in which taking on additional roles creates more resources than are depleted. Additional research that examines leadership as a role responsibility, as well as research that directly measures the role perceptions and feelings of leaders, will help to better understand the mechanisms associated with the apparent beneficial relationship with individual performance. Although causal inferences from this non randomized, time-varying treatment design cannot be made with much certainty at this point (Raudenbush, 2001), additional methodological and substantive research on leadership role responsibilities might contribute to understanding how to bolster the performance of team leaders through the enhancement of leadership role value, visibility, and attention to potentially conflicting role responsibilities.

REFERENCES

Austin JT, Villanova P. (1990). The criterion problem: 1917–1992. *Journal of Applied Psychology, 77*, 836–874.

Barrett GV, Caldwell MS, Alexander RA. (1985). The concept of dynamic criteria: A critical reanalysis. PERSONNEL PSYCHOLOGY, *38*, 41–56.

Bliese PD, Ployhart RE. (2002). Growth modeling using random coefficient models: Model building, testing and illustrations. *Organizational Research Methods, 5*, 362–388.

Brown MC. (1982). Administrative succession and organizational performance: The succession effect. *Administrative Science Quarterly, 27*, 1–16.

Calder BJ. (1977). An attribution theory of leadership. In Staw BM, Salancik GR (Eds.), *New directions in organizational behavior* (pp. 179–204). Chicago: St. Clair Press.

Chan K-Y, Drasgow F. (2001). Toward a theory of individual differences and leadership: Understanding motivation to lead. *Journal of Applied Psychology, 86*, 481–498.

Collins LM, Hyatt SL, Graham JW. (2000). Latent transition analysis as a way of testing models of stage-sequential change in longitudinal data. In Little TD, Schnabel KU, Baumert J (Eds.), *Modeling longitudinal and multilevel data: Practical issues, applied approaches, and specific examples* (pp. 147–161). Mahwah, NJ: Erlbaum.

Day DV, Lord RG. (1988). Executive leadership and organizational performance: Suggestions for a new theory and methodology. *Journal of Management, 14*, 453–464.

Deadrick DL, Bennett N, Russell CJ. (1997). Using hierarchical linear modeling to examine dynamic performance criteria over time. *Journal of Management, 23*, 745–757.

Diamond D. (Ed.). (1998). *Total hockey: The official encyclopedia of the National Hockey League*. Kansas City, MO: Andrews McMeel.

Dirks KT. (2000). Trust in leadership and team performance: Evidence from NCAA basketball. *Journal of Applied Psychology, 85*, 1004–1012.

Eden D. (1990). *Pygmalion in management: Productivity as a self-fulfilling prophecy*. Lexington, MA: Lexington Books.

Eden D, Shani AB. (1982). Pygmalion goes to bootcamp: Expectancy, leadership, and trainee performance. *Journal of Applied Psychology, 67*, 194–199.

Erwin S. (2004, February 12). Study: NHL player salaries "inconsistent" with sound business practice. *National Post.* Retrieved February 16, 2004, from http://www.thescore.ca/new_site/content/Hockey.asp?target=467011&storyType=

Gamson WA, Scotch NA. (1964). Scapegoating in baseball. *American Journal of Sociology, 70*, 69–72.

Ghiselli EE. (1956). Dimensional problems of criteria. *Journal of Applied Psychology, 40*, 1–4.

Goode WJ. (1960). A theory of role strain. *American Sociological Review, 25*, 483–496.

Grusky O. (1963). Managerial succession and organizational effectiveness. *American Journal of Sociology, 69*, 21–31.

Hofmann DA, Jacobs R, Baratta JE. (1993). Dynamic criteria and the measurement of change. *Journal of Applied Psychology, 78*, 194–204.

Hofmann DA, Jacobs R, Gerras SJ. (1992). Mapping individual performance over time. *Journal of Applied Psychology, 77*, 185–195.

Kreiser J. (2001, February). If the captain fits. *Hockey Digest*. Retrieved February 13, 2004, from http://www.findarticles.com/cf_dls/m0FCM/4_29/68620642/p1/article.jhtml?term=

Lance CE, Vandenberg RJ, Self RM. (2000). Latent growth models of individual change: The case of newcomer adjustment. *Organizational Behavior and Human Decision Processes, 83*, 107–140.

Landis RS. (2001). A note on the stability of team performance. *Journal of Applied Psychology, 86,* 446–450.

LaPointe J. (1997, September 28). Wearing hockey's badge of leadership. *The New York Times,* pp. C1, C6.

Lord RG, Hall RJ. (1992). Contemporary views of leadership and individual differences. *Leadership Quarterly, 3*, 137–157.

Lord RG, Hohenfeld JA. (1979). Longitudinal field assessment of equity effects on the performance of major league baseball players. *Journal of Applied Psychology, 64*, 19–26.

Lord RG, Maher KJ. (1991). *Leadership and information processing: Linking perceptions and performance*. Boston: Unwin Hyman.

Marks SR. (1977). Multiple roles and role strain: Some notes on human energy, time and commitment. *American Sociological Review, 42*, 921–936.

Meindl JR, Ehrlich SB, Dukerich JM. (1985). The romance of leadership. *Administrative Science Quarterly, 30*, 78–102.

Pfeffer J. (1977). The ambiguity of leadership. *Academy of Management Review, 2*, 104–112.

Pfeffer J, Davis-Blake A. (1986). Administrative succession and organizational performance: How administrator experience mediates the succession effect. *Academy of Management Journal, 29*, 72–83.

Ployhart RE, Hakel MD. (1998). The substantive nature of performance variability: Predicting interindividual differences in intraindividual performance. PERSONNEL PSYCHOLOGY, *51*, 859–901.

Ployhart RE, Holtz BC, Bliese PD. (2002). Longitudinal data analysis: Applications of random coefficients modeling to leadership research. *Leadership Quarterly, 13*, 455–486.

Raudenbush SW. (2001). Comparing personal trajectories and drawing causal inferences from longitudinal data. *Annual Review of Psychology, 52*, 501–525.

Raudenbush SW, Bryk AS. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage.

Rosenthal R, Jacobson L. (1968). *Pygmalion in the classroom*. New York: Holt, Rinehart & Winston.

Ruderman MN, Ohlott PJ, Panzer K, King SN. (2002). Benefits of multiple roles for managerial women. *Academy of Management Journal, 45*, 369–386.

Russell CJ. (2001). A longitudinal study of top-level executive performance. *Journal of Applied Psychology, 86*, 560–573.

Schafer J. (1997). *Analysis of incomplete multivariate data*. London: Chapman & Hall.

Schnabel KU, Little TD, Baumert J. (2000). Modeling longitudinal and multilevel data. In Little TD, Schnabel KU, Baumert J (Eds.), *Modeling longitudinal and multilevel data: Practical issues, applied approaches, and specific examples* (pp. 9–13). Mahwah, NJ: Erlbaum.

Sieber SD. (1974). Toward a theory of role accumulation. *American Sociological Review, 39*, 567–578.

Sundstrom E. (1999). The challenges of supporting work team effectiveness. In Sundstrom E (Ed.), *The ecology of work group effectiveness: Design guidelines for organizations, facilities, and information system for teams* (pp. 3–23). San Francisco: Jossey-Bass.

Thomas AB. (1988). Does leadership make a difference to organizational performance? *Administrative Science Quarterly, 33*, 388–400.

Turner RH. (2002). Role theory. In Turner JH (Ed.), *Handbook of sociological theory* (pp. 233–254). New York: Kluwer Academic/Plenum.

Zaccaro SJ, Rittman AL, Marks MA. (2001). Team leadership. *Leadership Quarterly, 12*, 451–483.