2010

How Supervisors Influence Performance: A Multilevel Study of Coaching and Group Management in Technology-Mediated Services

Xiangmin Liu
Pennsylvania State University

Rosemary Batt
Cornell University, rb41@cornell.edu

How Supervisors Influence Performance: A Multilevel Study of Coaching and Group Management in Technology-Mediated Services

Abstract

This multilevel study examines the role of supervisors in improving employee performance through the use of coaching and group management practices. It examines the individual and synergistic effects of these management practices. The research subjects are call center agents in highly standardized jobs, and the organizational context is one in which calls, or task assignments, are randomly distributed via automated technology, providing a quasi-experimental approach in a real-world context. Results show that the amount of coaching that an employee received each month predicted objective performance improvements over time. Moreover, workers exhibited higher performance where their supervisor emphasized group assignments and group incentives and where technology was more automated. Finally, the positive relationship between coaching and performance was stronger where supervisors made greater use of group incentives, where technology was less automated, and where technological changes were less frequent. Implications and potential limitations of the present study are discussed.

Disciplines

Labor Relations | Organizational Behavior and Theory | Performance Management | Technology and Innovation

Comments

How Supervisors Influence Performance:

A multilevel study of coaching and group management in technology-mediated services

Xiangmin Liu

Pennsylvania State University

Rosemary Batt

Cornell University

This multilevel study examines the role of supervisors in improving employee performance through the use of coaching and group management practices. It examines the individual and synergistic effects of these management practices. The research subjects are call center agents in highly standardized jobs, and the organizational context is one in which calls, or task assignments, are randomly distributed via automated technology, providing a quasi-experimental approach in a real-world context. Results show that the amount of coaching that an employee received each month predicted objective performance improvements over time. Moreover, workers exhibited higher performance where their supervisor emphasized group assignments and group incentives and where technology was more automated. Finally, the positive relationship between coaching and performance was stronger where supervisors made greater use of group incentives, where technology was less automated, and where technological changes were less frequent. Implications and potential limitations of the present study are discussed.

In response to evolving customer demands, many companies are adopting competitive strategies that emphasize innovation in products, processes, and technologies. These strategies, in turn, have enhanced the demand for workplace learning because employees need to absorb new skills and routines to perform their jobs (Salas & Cannon-Bowers, 2001). U.S. organizations invested $134.9 billion in learning and development in 2007, with two-thirds of the total spent on internal developmental activities (American Society for Training & Development, 2007).

Along with the increased emphasis on workplace learning, evidence also is accumulating that organizations are devolving human resource management (HR) responsibilities to supervisors and line managers in order to enhance employee performance (Hall & Torrington, 1998; McGovern, Gratton, Hope-Hailey, Stiles, & Truss, 1997). This decentralization of tasks broadens the core responsibilities of first-line supervision—from traditional duties of monitoring and administration to a set of performance-oriented tasks that identify, assess, and develop the competencies of subordinates and align their performance with the strategic goals of the organization (Hales, 2005; Purcell & Hutchinson, 2007). Thus, our subject of study is the HR role of supervisors in skill development and performance improvement.

One approach to performance improvement is for supervisors to provide individualized instruction and guidance to employees in the context of daily work. This activity is generally referred to as informal training, but it is more accurately described as coaching, which the literature defines as an unstructured,

developmental process in which managers provide one-on-one feedback and guidance to employees in order to enhance their performance (Heslin, VandeWalle, & Latham, 2006). Coaching has advantages over formal training because it is considerably less expensive and more closely fits the current need for ongoing learning and continuous improvement in the context of firm-specific workplace processes and technologies.

However, supervisors may combine individualized coaching with other strategies to improve performance. Although they have little control over such HR policies as recruitment, selection, or compensation, they have primary responsibility for coaching and managing the working relationships among employees in their work groups. They can, for example, create a work environment that enhances group processes of communication, motivates cooperation and learning (Argote & McGrath, 1993), and reinforces their one-on-one coaching interactions with employees. We refer to practices that enhance working relationships among peers as "group management practices." Our assumption is that these practices may be effective for work that is individualized or loosely organized into groups—they do not depend on high levels of interdependence in teams (Hackman, 1987; Hackman & Wageman, 2005).

Our approach to understanding employee performance brings together two sets of literatures: the training literature and the strategic HR management literature. We draw on the training literature to test a multilevel model of coaching in relationship to other organizational factors that influence performance. Although many have called for this type of approach to training, few studies have actually adopted it (Blanchard & Thacker, 2007; Kozlowski & Salas, 1997; Salas & Cannon-Bowers, 2001). We draw on the strategic HR management literature to conceptualize "other organizational factors" in terms of the role of HR management. That literature has shown that HR practices, in combination, may lead to better performance than if they are implemented in isolation (Combs, Liu, Hall, & Ketchen, 2006).

In particular, the HR literature has identified three dimensions of the HR system that enhance performance: investment in training, work designed to allow employees to interact and develop their skills and problem-solving abilities, and incentives to motivate effort (Appelbaum, Bailey, Berg, & Kalleberg, 2000; Batt, 2002; Delery, 1998). Although the strategic HR literature has found significant relationships between these dimensions and performance at the organizational level (Combs et al., 2006), some have called for studies that illuminate how these relationships are effectively implemented at lower levels of the organization (Wright & Boswell, 2002; Wright & Nishii, 2009). We contribute to the HR literature by providing a context-specific example of how supervisors implement these three dimensions of the HR system to improve employee performance.

We contribute to the training literature by showing the link between coaching and other HR management activities that, taken together, should improve performance. This emphasis on management practices departs from the training literature, which often treats training as primary and other organizational factors as "context," or "environment." We theorize that supervisory variation in individual coaching and group management practices has both direct and synergistic effects on individual performance improvement. The synergies depend on whether these practices are congruent, or consistent, among themselves (Kozlowski & Salas, 1997).

Third, we theorize that management practices designed to improve performance should be understood in the context of workplace technologies that enable and constrain those practices and their outcomes. Most of the literature on training, as well as that on HR management, has failed to take workplace technologies into account, except as a means for implementing training itself. In sum, by conceptualizing coaching in terms of HR management, we focus on managers' actions rather than employee perceptions of climate or environment or training transfer. We believe this approach can enhance the training literature by highlighting what managers can do and by linking the research results more directly to their practical

implications for managers. At the same time, the theoretical framing from the training literature can strengthen the HR management literature by better theorizing what factors explain individual performance.

Our methodological approach also differs from prior research on coaching (Smither & Reilly, 2001) and informal training more generally (Salas & Cannon-Bowers, 2001). Although the coaching research has tended to focus on newly hired employees (e.g., Lefkowitz, 1970; Tews & Tracey, 2008) or executive coaching (Olivero, Bane, & Kopelman, 1997; Smither, London, Flautt, Vargas, & Kucine, 2003), with studies often using managers in MBA courses as subjects (Hall, Otazo, & Hollenbeck, 1999; Hollenbeck & McCall, 1999), the subjects of our study are incumbent workers doing standardized, routine service work. In addition we examine individual performance over time rather than cross sectionally or as a relationship between training and different individuals' behavior or perceptions. Although most research on coaching uses perceptual and cross-sectional measures of coaching and performance (e.g., Agarwal, Angst, & Magni, forthcoming), our applied setting—with random assignment of tasks; longitudinal, hierarchically structured data; real-time measures of coaching; and objective measures of performance—provides a stronger methodological approach. It also responds to some calls for training research to be operationalized in more context-specific ways (Kozlowski & Salas, 1997, p. 267; Rousseau, 1985).

Theory and Hypotheses

There is a general recognition that training research needs to move beyond the individual level approach and incorporate organizational phenomenon, but building multilevel theories and testing them has only begun to take shape. One series of studies has conceptualized the work environment as influencing individual perceptions and beliefs, such as training motivation (Quinones, 1995), opportunities to perform (Ford, Quinones, Sego, & Sorra, 1992), and support from supervisors and coworkers (Smith-Jentsch, Salas, & Brannick, 2001). Although these approaches have found empirical support for their arguments, they have conceptualized the work environment at the individual level, thus measuring individual perceptions more than the actual work, organizational features, or management practices at higher levels of analysis. A second stream of research has viewed the work environment in terms of employee perceptions of training climate or culture. Here, researchers have found that shared perceptions of training climate or learning culture are positively related to posttraining behavior (Rouiller & Goldstein, 1993; Tracey, Tannenbaum, & Kavanaugh, 1995). However, empirical studies have found little support for a moderating relationship of training climate (Tracey et al., 1995). Neither studies of individual perceptions nor workplace climate of training highlight what managers can do.

One attempt to construct a more integrated approach to training and development in organizations has come from Eduardo Salas, Kevin Kozlowski, and colleagues (Kozlowski & Salas, 1997; Salas & Cannon-Bowers, 2001). We use this as a starting point in our paper as it provides several distinct advantages over prior conceptualizations. In a critical review that highlighted the limitations of prior training research, Kozlowski and Salas (1997) developed what they refer to as a "systems" approach that incorporates insights from the training literature and organization theory. The "systems" concept captures the idea that there are moderating or synergistic effects—rather than independent or additive effects—operating between different factors in the organization. Their approach moves beyond prior frameworks in three ways: it develops a *multilevel* framework that recognizes that training outcomes at the individual level depend on organizational factors that operate at higher levels of analysis; it specifies the *content* of two types of factors that are theorized to influence the exercise and transfer of training: "enabling process" and "techno-structural" factors; and it specifies that the extent of *congruence* or consistency in variables—both across levels and

content areas—is a key theoretical explanation for training effectiveness. Enabling process factors refer to social processes that shape attitudes and behavior at work, whereas the technostructural factors refer to the concrete, tangible, or visible aspects of the work system. The incorporation of technical features is reminiscent of the sociotechnical systems approach but distinct because that literature emphasized the need to fit technology to the needs of human beings, and most of the actual research focused on self-managed teams, to the exclusion of technology (Cohen & Bailey, 1997; Pasmore, Francis, & Haldeman, 1982).

Our multilevel model includes activities at the level of the work group, the individual, and the individual over time. We consider how individual coaching affects individual performance trajectories; how management practices at the work group level affect individual performance levels and the relationship between individual coaching and performance outcomes; and how technical processes affect individual performance as well as the relationship between coaching and performance. Our approach differs somewhat from the Kozlowski and Salas framework because we conceptualize supervisors as key actors with discretion in both their coaching and group management practices, and we focus specifically on objective performance outcomes rather than training transfer. In the sections below, we review the specific literature on coaching and then hypothesize how group management practices and process technologies are likely to affect performance and interact with coaching effectiveness and individual performance.

Coaching

Coaching is a process through which supervisors may communicate clear expectations to employees, provide feedback and suggestions for improving performance, and facilitate employees' efforts to solve problems or take on new challenges (Heslin et al., 2006). It consists of regular interactions that help employees adopt effective work skills and behaviors. The literature has differentiated coaching from other types of informal training, such as mentoring and tutoring (Chao, 1997; D'Abate, Eddy, & Tannenbaum, 2003). Although coaching focuses on specific, short-term performance improvements, mentoring provides individuals with psychological support and social resources in order to reach long-term career goals. Tutoring typically involves an expert who passes on domain-specific knowledge to novices. In coaching, however, supervisors may not necessarily be domain experts but may help individuals gain greater competence and overcome barriers to performance. Examples of coaching activities include helping employees set specific goals, providing constructive feedback on specific tasks, offering resources and suggestions to adopt new techniques, and helping employees understand the broader goals of the organization (Ellinger, Ellinger, & Keller, 2003).

Coaching may affect individual performance through three mechanisms: the acquisition of job-related knowledge and skills, the enhancement of motivation and effort, and process of social learning. Coaching is an effective source of skill acquisition because supervisors can observe specific employee behaviors and performance and provide constructive feedback and guidelines for improvement (Heslin et al., 2006). This type of timely and individualized instruction contributes to the construction and recall of an individual's declarative and procedural knowledge (Kraiger, Ford, & Salas, 1993). Proximity between the learning task during coaching and its practical application at work reduces the loss associated with transfer of training, which is problematic for structured, off-site training activities (Baldwin & Ford, 1988). Coaching helps employees develop and maintain knowledge of a firm's products, customers, and work processes; and skills to effectively communicate with customers, respond to their requests, and deliver prompt service.

Coaching also may enhance an individual's motivation to improve or take personal initiative. It may allay goal ambiguity and stimulate a process of "spontaneous goal-setting" by clarifying performance expectations (Locke & Latham, 1990). Smither et al. (2003) found that managers who worked with an external coach were more likely than other managers to set specific (rather than vague) goals and to solicit ideas for improvement from supervisors. Finally, emerging perspectives on socially constructed learning, or dialogical approaches, stress that knowledge and learning are socially embedded in power relationships and cultural values (Burke, Scheuer, & Meredith, 2007; Holman, 2000). Coaching consists of a sequence of ongoing conversations and actions that promote continuous exchange of experience, feedback, and encouragement (Heslin et al., 2006). Thus, it may serve as an important vehicle through which situation-specific knowledge and organizational norms are formed, articulated, and dispersed among supervisors and subordinates. Studies have shown that a dialogue-based coaching intervention leads to successful performance (efficiency, creativity, and work climate) by enhancing peer relations and enabling employees to develop and use collective knowledge (Mulec & Roth, 2005).

Some studies have suggested a positive relationship between coaching and job performance (Agarwal et al., 2009; Ellinger et al., 2003); but empirical evidence remains weak because these studies only used perceptual measures and estimated performance differences between individuals as a result of differential treatments of coaching. Yet the literature's prediction, that coaching leads to better performance, pertains to within-individual differences as well. That is, coaching stimulates a positive, development-oriented process that should result in an individual's performance improvement over time. This line of argument suggests the following hypothesis.

Hypothesis 1: The amount of supervisor coaching an employee receives is positively related to individual performance over time.

Group Management Practices: Direct and Synergistic Effects

Beyond individual coaching activities, supervisors may influence performance by how they shape the working relationships among the employees they oversee. One approach is to create an environment of individual competition based on the assumption that such an environment motivates all employees to perform better than they otherwise would because they want to outperform their peers. Alternatively, supervisors may adopt group management practices that foster a cooperative environment based on the assumption that group interaction provides social support or opportunities for mutual learning that enhances the performance of all employees.

Much recent theory and empirical work has supported the performance benefits of group- based work and incentives over individualized ones. One argument draws on group process theory, which emphasizes the role of effective communication and coordination (Argote & McGrath, 1993). If supervisors implement practices that enhance social interactions and information sharing, then they create an environment in which workers are able and motivated to solve problems together, and this group interaction leads to better individual performance. For example, pairing up novice employees with more experienced ones may be a vehicle for handling idiosyncratic work systems or peer training (Stajkovic & Luthans, 1997), or peers may help each other engage in self-disclosure and reflection (Lankau & Scandura, 2002). Supervisors also may emphasize team-based work or group rewards, both of which are particularly effective where monitoring and performance metrics are visible to all workers (Sewell, 1998), as is the case in this study.

Although research has demonstrated a significant relationship between better performance and group-based forms of work (Cohen & Bailey, 1997; Guzzo & Dickson, 1996; Kozlowski & Bell, 2003) and group incentives (Hansen, 1997; Weitzman & Kruse, 1990), most of the literature has viewed task interdependence as a critical condition for the benefits of group processes to be realized (Hackman, 1987). Individualized work settings (as in this study) would not necessarily benefit from group-based approaches. However, if group activities or peer collaborations are sources of learning or motivation, then they may be effective tools for performance improvement even where task interdependence is low. For example, in a cross-level study of call center workers, Batt (1999) found that objective sales performance was higher for workers in self-directed groups compared to those in traditionally supervised groups, in part because the former solved technical problems more effectively. Similarly, studies of "communities of practice" (Brown & Duguid, 1991; Lave & Wenger, 1991) describe how learning occurs between peers in the context of every day work. Kunda (1992) found that the performance of technicians working individually in remote sites depended importantly on regular informal meetings among technicians to exchange ideas and share results.

In our multilevel model of supervisor coaching and group management, we also are interested in whether there are synergies between these two approaches to performance improvement. In the terms of Kozlowski and Salas (1997), is there congruence between content areas such that, in combination, they produce higher performance than would otherwise be the case? We argue that practices that foster group interactions should also enhance coaching because, according to social information processing theory, "people learn what their needs, values, and requirements should be in part from their interactions with others" (Salancik & Pfeffer, 1978, p. 230). In the context of training, group norms and culture define the accepted patterns of employee interaction and work practices and thus affect posttraining work behaviors (Rouiller & Goldstein, 1993; Tracey et al., 1995). Some empirical results are consistent with this argument: Mathieu, Tannenbaum, and Salas (1992) found that when trainees lacked coworker support they were less likely to apply newly acquired skills to the job; and Smith-Jentsch, Salas, and Brannick (2001) showed that team leader supportive attitudes moderated the relationship between training and behavioral outcomes in a simulated laboratory setting. Pairing with experienced peers, for example, may encourage, remind, and reinforce the learning goals and behaviors of trainees, whereas the use of group incentives may encourage group members to look out for the interests of others and support performance improvement of the whole group (DeMatteo, Eby, & Sundstrom, 1998). Thus, we expect that the relationship between coaching and performance will be stronger when supervisors also use other practices to enhance group interaction and cooperation.

Hypothesis 2a: Where supervisors make greater use of group management practices, individuals will demonstrate higher levels of performance.

Hypothesis 2b: Group management practices will moderate the relationship between coaching and performance. Specifically, the positive relationship between coaching and performance trajectories will be stronger where group management practices are more frequently used.

Technical Processes: Direct and Synergistic Effects

Supervisors typically have little control over the design of technical systems that enable or constrain opportunities for individual learning and performance, but these systems set the physiological and psychological requirements of tasks and shape individual performance. In this study, we consider two types

of technologies that are central to call center performance (as well as that of many manufacturing and service operations)—the level of process automation and the extent of process change. Process automation refers to the extent to which certain tasks can be performed by minimizing human contact, for example, through the use of automated information systems. The level of process automation directly affects overall levels of performance by increasing efficiencies, not only in manufacturing settings but also in service operations that rely on information and computer technology. Call centers use automated call distribution systems that set the pace of work and voice recognition systems that answer some inquiries without an operator's intervention. However, the level of automation is rarely similar across establishments. Where information in the databases is less accurate, where place names are more idiosyncratic, or where customers provide inaccurate information, operators spend more time manually searching databases. Thus, the greater the automation, the less time is needed per call and the higher the performance of individuals working in this system.

Beyond the direct effects of automation, how does it influence the relationship between coaching and individual performance? Are there positive or negative synergies? Arguably, differences in the technical features of work present different levels of opportunities for individuals to apply acquired skills (Ford et al., 1992). As individuals acquire knowledge and skills needed to complete a variety of job duties, the performance benefits of training are greater when employees have the opportunity to perform many or all of the tasks they were trained to do. As process automation increases, by contrast, the role of human intervention is narrower, and individuals have limited opportunities to use acquired skills or influence process outcomes. Coaching may be less important where process automation is high because of the limited contribution that individual skills can contribute to performance. Thus, we believe there are negative synergies, or a lack of congruence, between process automation and coaching: The relationship between coaching and productivity will be lower where process automation is higher.

> *Hypothesis 3a*: Process automation will be positively related to performance, such that when process automation is higher, individuals will have higher levels of performance.
>
> *Hypothesis 3b*: Process automation will moderate the relationship between coaching and performance. Specifically, the positive relationship between coaching and performance will be lower when process automation is higher.

Ongoing changes in technical systems also are a common feature in organizations today as employers regularly update technologies or as companies merge, restructure, or introduce new products and services. Even though technical changes are made to improve efficiency, they also are likely to disrupt work routines (McAfee, 2002) and lead to lower performance when they are initially introduced. Therefore, in contrast to automation, process upgrades are likely to be associated with lower individual performance in the weeks or months after they are introduced.

How process change influences the relationship between coaching and performance is a more complex question. Positive synergies could emerge if supervisors are able to rapidly learn the new processes themselves and impart new techniques to employees. However, this is an unlikely scenario because it is the employees themselves who are spending the most time directly involved with new technologies. Supervisors are likely to have greater difficulty keeping up with ongoing changes, so their coaching of employees under these conditions is likely to be less effective than it otherwise would. Therefore, we expect that the relationship between coaching and performance will be lower where process changes are more frequent.
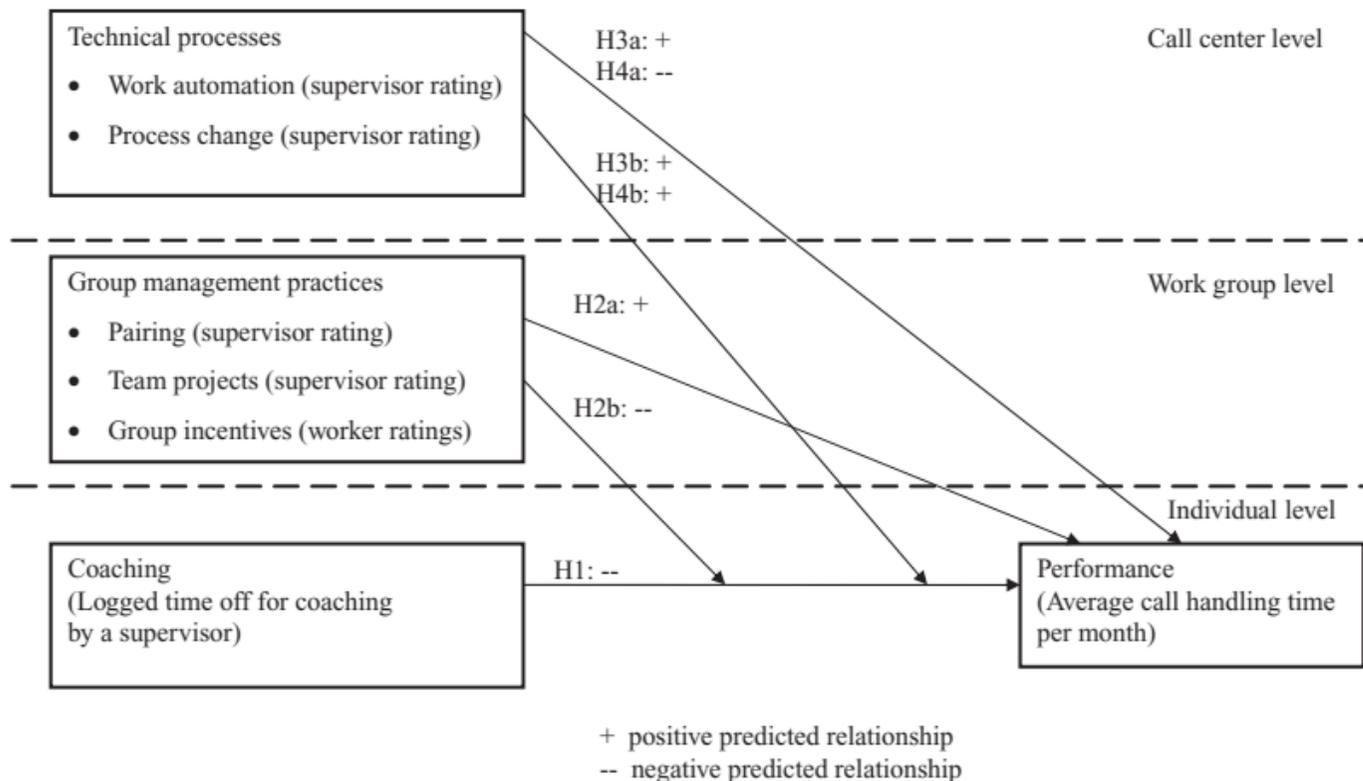
Figure 1: **The Hypothesized Model: Coaching, Group Management, and Technical Processes.**

Hypothesis 4a: Technical process changes will be negatively related to individual performance in the period when they are initially implemented such that when processes change more frequently, individuals will demonstrate lower levels of performance.

Hypothesis 4b: Technical process changes will moderate the relationship between coaching and performance. Specifically, the relationship between coaching and performance will be lower when processes change more frequently.

Figure 1 depicts the hypothesized relationships between coaching, group management, technical processes, and performance. This model reflects a contextualized organizational approach to this research, consistent with Rousseau and Fried's (2001) suggestions, in which we focus on a set of salient features based on our understanding of the work activities and organizational setting.

Methods

Research Setting

The research setting is the telephone operator services division of a unionized telecommunications company operating in a multistate region. Telephone operators are the core occupational group—the largest group of nonmanagerial employees in the business unit (Batt, 2002). The strategy of focusing on one occupational group in one business unit limits the confounding effects of unmeasured factors such as business and HR strategy. This site also has the advantage of offering a real-world setting in which work tasks are randomly assigned: The automatic call distribution system sends calls to the next available operator in each center. As soon as one call has ended, a second one enters the operator's headset.

Our field research provided background on competitive pressures, business operations, the nature of tasks and technology, and how and why coaching is important in this context. Operators handle directory assistance inquiries from anywhere in the United States. Government-mandated service levels require the company to answer 97.5% of calls in 6 seconds. Cost competition is intense in this commodity business, and companies can save millions of dollars by reducing call handling time by fractions of seconds. They accomplish this by adopting new technologies (e.g., voice recognition systems that process portions of calls) or training workers to use new technologies or procedures, to communicate more effectively, or to develop more efficient database search strategies. The company also requires an 85% customer satisfaction rating, as measured by an outside vendor survey. Initial training includes basic keyboarding and technical/procedural knowledge, ensuring that new hires have accurate and efficient keyboarding skills and know the procedures for retrieving information from a variety of databases. The company provides an average of 2.1 weeks of initial training (according to our surveys), and it takes employees about 6 months to become proficient on the job. For purposes of this study, we focused on incumbent workers whose job tenure exceeded 6 months.

The company in this case viewed supervisors as the primary providers of coaching, and the information system categorized supervisory coaching into five domains: general feedback, methods training (new procedures), customer satisfaction (ways to improve service quality), district issues (business-specific information), ergonomics, and performance improvement activities. The company policy required all supervisors to observe and provide feedback to at least 70% of their employees each month, and coaching was initiated by the supervisor not the employee. The majority of coaching consisted of individualized feedback based on monitoring of calls, behaviors, and keystrokes. Other types of coaching occurred when

new procedures, systems, or services were being initiated. Overall, considerable variation existed in coaching activities because they varied by supervisory staffing levels, supervisory competency, and workplace-specific conditions. Based on our survey, supervisors were spending an average of 12.27 hours on individualized feedback and continuous coaching of employees each week, but the 10th percentile did only 5 hours per week, and the 90th percentile did 20 hours.

Supervisors also were responsible for managing their work groups. In this research setting, most HR policies were set at the business unit level or by union contract. Social interaction among peers was limited because work rules required employees to stay in their seats and answer individual inquiries at least 85% of their work time. However, supervisors were encouraged to find creative ways to motivate employees through individual or group activities or incentives. In our site visits, we observed some supervisors using creative tools to foster interaction, even in such a standardized work environment. One tactic was to use a "buddy system" to pair novice employees with more experienced ones. The experienced worker would offer tacit know-how for handling the information system as well as guidance and emotional support for dealing with difficult customers. A second approach was to use ad hoc team projects, which allowed groups of workers time off the phones to discuss work-related problems or challenges. A third practice was to use cash and noncash group incentives for meeting group performance targets, such as call answering time, call handling time, and absenteeism.

The work environment of call centers is highly structured and automated. Based on our archival data, the average operator handled over 1,000 calls per day. The level of automation varied across centers located in different states due to differences in inherited systems from different companies that were now part of a merged entity. The company had not yet standardized the information system across all centers; thus, some variation existed in the extent of change or updates in systems across the geographic footprint of the company. The company also was introducing new changes to enhance revenue generation: just prior to our fieldwork, for example, it had begun to offer national 411 service (as opposed to regional service only). An important source of new revenues, it required operators to shift from a regional database—where they had tacit knowledge of local terminology or names of businesses that diverged from official listings—to a national one, where they had no such knowledge. Supervisors reported that operators received an average of 6.7 e-mails per day on updates or new procedures. In sum, in what is often considered a relatively low-skilled routine clerical job, ongoing changes in information systems and work processes required regular attention to informal training.

Sample, Data, and Data Construction

The company had a population of 6,937 telephone operators organized into 168 supervisor-led work groups in 64 centers. Data came from three sources: company archives and supervisor and worker surveys. We merged two data archives: demographic data from the human resource information system (HRIS) and monthly data on training and performance from the electronic monitoring system. Surveys of supervisors and workers provide data on group management practices and technology.

We sampled 16% of workers and all of the supervisors at each center (to ensure an adequate sample size in the latter case). We received 666 completed worker surveys (72% response rate) and 110 supervisor surveys (40% response rate). The lower response rate among supervisors reflects the fact that workers received time away from work to complete the survey, but supervisors did not.

In order to ensure an adequate number of employee responses to aggregate to the work group level, we randomly chose a limited number of work groups at each center (1–2) and randomly selected at least 10

workers per group. This resulted in at least five responses per work group, usually more. In addition, we limited surveys to centers with 40 employees or more because, to get a meaningful sample, we would have had to survey a much larger proportion of the workforce than the employer was willing to allow. We constructed a three-level data set—months, individuals, work groups—but there were not enough work groups per center to create a fourth level (See Figure 1). To create the cross-level data set, we aggregated worker and supervisor surveys to the work-group level (in some cases groups had a supervisor and assistant supervisor); we matched the aggregated surveys to individual archival data via administrative codes. The matching process was limited by errors in the administrative codes, missing supervisor surveys, and missing archival data. The final sample included 9,918 observations from 2,327 telephone operators in 42 work groups in 31 centers (327 worker surveys and 58 supervisor surveys). The study sample was primarily White (78%) and female (86%), with an average age of 40 and company tenure of 10 years. The average group size was 55. Although the HRIS system did not provide educational data, our survey of employees showed that variance in formal education was low: Most employees had some postsecondary education, and 8% had a college degree.

Response and Attrition Bias

Because nonrandom loss of observations may create estimation bias and reduce external validity of study conclusions, we conducted two sets of analyses to address these concerns. The first was a nonresponse analysis to test whether supervisors' decisions to respond to a survey created differences between the respondents and nonrespondents, and whether such a decision resulted in bias in ratings (Werner, Praxedes, & Kim, 2007). A comparison of mean values from supervisors in the population on archival data to mean values from supervisors who responded to surveys indicated that these two groups were not significantly different in race, sex, and salary levels. However, age and organizational tenure of respondents were higher. We investigated this issue further by testing whether these factors (especially age and tenure) related to the scores of survey items. As expected, age and tenure were not significant predictors of ratings in each of the reported variables.

A second concern was attrition bias when one matches data from different sources. As we had to retrieve data from each establishment and use the company's administrative codes as identification to merge data, some observations were lost due to inconsistency in these administrative codes. For example, an operator who has a group identification confirmed in the survey may fail to also have her training and performance information incorporated. To explore this issue, we compared mean values from all returned surveys and mean values in the matched data. We found no significant difference in the ratings of pairing, team projects, group incentives, and process changes. However, the score of automation was slightly lower in the matched data. Moreover, we compared mean values from operators in the population and those in the matched data. We found that the final sample was younger and more likely to be White and male. But in all of these cases, differences were small to moderate. Therefore, we found little evidence that loss of observations due to nonresponse and attrition will bias the study findings.

Variables

Our measure of performance comes from the electronic monitoring system, which continuously records the work activities of each operator, including time online with customers and offline for coaching or other activities. It is measured by call handling time, the average number of seconds an operator spends on a

customer call for a given month. This is the most important performance metric used in operator services. Lower call handling time equals higher productivity. The monthly data cover the period of January 2001 to May 2001. The average call handling time was 21.09 seconds.

Coaching is the length of time that a worker received coaching from a supervisor. Each time an employee logged off the computer for coaching, the minutes of coaching were recorded. The percentage of all operators in the company who received coaching each month ranged from 93.2% to 94.9%, with an average coaching intensity that ranged from 54 to 71 minutes each month. In the analyses, we used the accumulated amount of coaching in previous months to predict call handling time.

We measured group management practices in three ways: Pairing, team projects, and group incentives. "Pairing" is the extent to which new employees are paired up with experienced workers, as reported by supervisors, on Likert scale ranging from 1 = *not at all* to 5 = *completely*. For the use of team projects, we asked supervisor whether their subordinates were currently participating in any special project teams or task forces (*yes* = 1, *no* = 0). To measure group incentives, we used a 5-point Likert frequency scale and asked workers how often their supervisor used group-based rewards. Items included "When your work group does its job well, how often are you rewarded with noncash rewards (e.g., free lunch or dinner, public recognition, or small gifts)?" and "When your work group does its job well, how often are you rewarded with cash rewards (e.g., gift certificates, cash bonus)?" We used the worker reports of this measure because it provides a more objective evaluation than supervisors' self-reported measure.

The level of automation is captured by a 3-item scale based on supervisors' reports of how often their employees needed to resort to paper methods (reverse coded). The items were rated on a 5-point Likert scale, ranging from 1 = *rarely* to 5 = *extremely often* and included, "workers have to look something up in a manual," "workers have to fill out pen and paper form," and "workers have to do calculations by hand or calculator." Scale scores were created by taking the average of the three items (α = 0.85). To measure technical process change, we used a three-item index based on supervisor reports of how often their employees received updates regarding (a) product features, (b) pricing, and (c) service options. The items used a 5-point Likert-type scale ranging from 1 = *rarely* to 5 = *extremely often*, with high scores representing more rapid information changes (α = 0.78). Because technical architecture is typically set at the establishment level, we aggregated these scores to the call-center level and then applied them to each work group in the center.

We controlled for initial performance in order to improve our causal model. We measured proficiency in the first month by the percentage of objectives achieved for each operator, based on the company's archival data. Each local call center specified minimum performance requirements or workers at the site, depending on customer characteristics. This measure is calculated as the proportion of expected call handling time over actual call handling time. The measure usually ranged from 94% to 107%, with a high score indicating high performance. We were able to retrieve these data for 1,975 operators, with missing values for 372 operators. We used the single imputation technique to handle incompleteness. That is, each missing value is imputed from the variable mean of the complete cases, whereas a dummy variable is generated to indicate nonresponse. We then use standard statistical procedures for the "complete" data set. Compared to list wise deletion, the imputation approach avoids a substantial reduction in sample size and the possibility that the remaining data set is biased due to nonrandom missing values (Little & Rubin, 1987).

Finally, we controlled for variation in the size of work groups and organizational tenure of workers using archival data. Group size is often used as a proxy of span of supervisory control. Employees in larger groups may find they receive less personalized attention from supervisors than those in smaller groups. We

controlled for organizational tenure because experienced workers may accumulate more tacit skills and knowledge.

Data Aggregation

Supervisors reported on technology variables, and their reports were averaged to the center level and applied to the work groups in their centers. (There were not enough groups per center to compute aggregation statistics.) The supervisor reported on pairing and team project activities, as they are the most accurate source on these subjects. Workers reported on whether they received group incentives, which were aggregated to the group level, because we believed supervisors might be more prone to report positively on this question. For the group incentives variable, we followed James (1982), James, Demaree, and Wolf (1984, 1993) to assess interrater agreement $r_{wg(j)}$ within each of the 42 groups. $r_{wg(j)}$ ranges between 0.5 and 0.95, with 93% of the estimates suggesting moderate to
strong within-group agreement. The mean value of 0.89 indicates a high level of agreement on this measure at the group level. We further calculated the average deviation (AD) indices, which provide direct assessments of interrater agreement in the units of the original measurement scale (Burke & Dunlap, 2002). The overall mean AD was 0.48, suggesting a high level of agreement (cutoff point is 0.80). The interpretation of this AD value is that, over average, the subordinates deviated from the mean of their ratings by 0.48 units of the 5-point scales. We then conducted one-way analyses of variance and found significant between-group variance ($p < .08$). The intraclass correlation (ICC1) was 0.05 and reliability of group mean (ICC2) was 0.27. This represents a small to medium effect, suggesting group membership influenced employees' ratings on group rewards (LeBreton & Senter, 2008). Further analysis suggested that low ICCs values are not due to lack of rating similarity but rather due to an artifact of the distribution of ratings. That is, although ratings on group incentives were made on a 5-point scale, in over 85% of total responses, only three of the scale points were actually used. In this case, ICCs are low because inconsistencies in rank orders mask strong levels of interrater agreement (LeBreton, Burgess, Kaiser, Atchley, & James, 2003). Therefore, aggregation is justified by theory and supported by $r_{wg(j)}$ value and AD indices (Chen & Bliese, 2002; LeBreton & Senter, 2008).

Analytical Strategy

To model the relationships among coaching and performance within individuals and to examine the effects of group management and technical features between individuals across work groups, we used three-level hierarchical linear modeling (HLM; Byrk & Raudenbush, 1992). In HLM, each level is represented by its own equation. In this study, the Level 1 analysis estimated the growth trajectory of each operator's performance over time by including monthly observations of coaching and call handling time at five time points. The Level 2 analysis introduced worker characteristics and estimated individual variation in the trajectory of performance gains across operators in the same work group. The Level 3 analysis included the higher level measures of group management and technical features and examined systematic variation in levels and trajectories of performance improvement across work groups. Thus, Level 1 variables are at the within-person level of analysis, Level 2 variables at the between-person and within-group level, and Level 3 variables at the between-group level of analysis. Following prior discussions on cross-level models, we tested the direct effects of higher level variables (e.g., group management and technical processes) on lower level variables (individual performance) through direct effects on intercepts and tested the synergistic effects through cross-level moderation of slopes (Klein & Kozlowski, 2000). The Level 1, Level 2, Level 3, and

TABLE 1
Descriptive Statistics and Correlation Matrix

Variables	Mean	SD	1	2	3	4	5	6	7	8	9	10
1. Call handling time	21.09	4.64										
2. Coaching	2.82	2.06	−0.047*									
3. Pairing	2.48	1.23	−0.062*	−0.013								
4. Team projects	0.50	0.51	−0.118*	0.026	0.075*							
5. Group incentives	2.15	0.48	−0.126*	−0.010	0.368*	0.450*						
6. Automation	3.03	0.80	−0.199*	0.017	−0.126*	−0.140*	−0.066*					
7. Process change	2.02	1.05	−0.077*	0.118*	0.054*	−0.017	0.162*	−0.024				
8. Initial performance	1.03	0.13	−0.569*	−0.051*	−0.009	−0.002	−0.043*	0.012	−0.017			
9. Initial performance dummy (= 1 if missing)	0.11	0.32	0.034*	0.005	−0.038*	−0.112*	−0.027*	0.050*	−0.076*			
10. Group size	56.67	39.16	0.009	0.021	−0.391*	−0.236*	−0.604*	−0.024	−0.178*	0.017	0.132*	
11. Organizational tenure	10.20	9.59	0.261*	−0.038*	−0.077*	−0.054*	0.129*	−0.002	0.081*	−0.000	0.632*	−0.186*

Notes: Sample size: 9,918 observations (Level 1), 2,327 individuals (Level 2), and 42 work groups (Level 3).
*Significant at .05 level; Bonferroni adjusted.

combined models that are tested here are provided in the appendix. To reduce multicollinearity problems and aid the interpretation of variables, we followed Kreft and De Leeuw (1998) and centered all independent variables to grand mean in the model.

Results

Table 1 presents the means, standard deviations, reliabilities, and intercorrelations of the variables in the study. An examination of Table 1 reveals that coaching, group management practices, and technical processes are significantly related to call handling time. Demographic characteristics also are associated with call handling time.

Before proceeding to test our hypotheses with HLM, we investigated whether systematic within-individual, between-individual/within-group, and between-group variance existed in the dependent variable (call handling time) by estimating a null model. Results of the null model (not shown) indicate that variation of the means over the 42 work groups was 3.61 ($p < .00$), variation of the means over the 2,327 operators was 17.92 ($p < .00$), and the error variance was 2.11. That is, 15% of the total variance in call handling time resides between groups whereas 76% of the variance is between individuals within the same work group. Partitioning of variance components suggested the existence of sufficient variability of call handing time across each level. This finding provides a basis for examining individual-level and group-level predictors of job performance, as well as time-variant predictors (i.e., coaching) of it.

Table 2 presents our results using a hierarchical regression format: control variables in the first column, coaching added in the second, work group characteristics in the third, and moderators in the fourth. Coaching explains considerable unique variance in call handling time beyond that explained by the control variables. As predicted, coaching has a negative effect on call handling time (-0.09, $p < .01$) and therefore increases performance. This result indicates a strong positive performance growth trajectory, thus providing support for Hypothesis 1.

Coaching also remains significant when we add the main effects for group management practices and technical processes, as reported in the third column. Hypothesis 2a predicted that group management practices would increase performance. Results indicate that the use of project teams (-0.86, $p < .05$) and group rewards (-1.86, $p < .01$) are negatively associated with mean changes in call handling time. The effect of pairing is not significant. Hypothesis 2a is partially supported. In addition, automation is significantly and negatively related to call handling time, or higher performance (-1.20, $p < .01$), as predicted by Hypothesis 3a. However, frequency of information updates is not significantly related to call handling time. Hypothesis 4a is not supported.

Finally, Hypotheses 2b, 3b, and 4b predicted that group level characteristics would have a cross-level moderating effect on the relationship between coaching and job performance. Column 4 of Table 2 presents these results. Hypothesis 2b predicted that group management practices would moderate the relationship between coaching and performance in such a way that the more group interaction, the stronger the relationship. We found that the interaction of group incentives is significant as predicted (-0.10, $p < .01$). Using points one standard deviation above and one standard deviation below the means of each variable, we plotted the interaction in Figure 2. The performance effect of coaching is stronger among operators whose supervisors emphasized group-based rewards. The moderating effect of pairing is significant but not in the expected direction (0.05, $p < .01$). Pairing with experienced peers appears to attenuate the relationship between coaching and job performance. This finding may be indicative of the distinct content focus and domain of supervisor coaching and peer coaching (Sisson, 2001). If peers make suggestions that are contrary

TABLE 2
Results of Hierarchical Linear Modeling Analyses

	Model 1	Model 2	Model 3	Model 4
Individual/time level predictor				
Coaching		-0.092^{***}	-0.092^{***}	-0.089^{***}
		(0.034)	(0.034)	(0.016)
Work group level predictors				
Pairing			-0.155	-0.054
			(0.160)	(0.201)
Team projects			-0.859^{**}	-0.886^{**}
			(0.425)	(0.429)
Group incentives			-1.864^{***}	-1.906^{***}
			(0.530)	(0.540)
Automation			-1.200^{***}	-1.217^{***}
			(0.256)	(0.255)
Process change			-0.119	-0.112
			(0.356)	(0.361)
Coaching × pairing				0.053^{***}
				(0.019)
Coaching × team projects				0.017
				(0.032)
Coaching × group incentives				-0.095^{***}
				(0.035)
Coaching × automation				0.113^{***}
				(0.027)
Coaching × process change				0.075^{***}
				(0.026)
Control variables				
Initial performance	-21.943^{***}	-21.992^{***}	-21.963^{***}	-21.964^{***}
	(2.752)	(2.748)	(2.738)	(2.746)
Initial performance dummy (= 1 if missing)	0.345	0.340	0.371	-0.365
	(0.471)	(0.471)	(0.467)	(0.469)
Group size	-0.002	-0.002	-0.013^{*}	-0.014^{*}
	(0.009)	(0.009)	(0.008)	(0.007)
Org. tenure	0.059^{**}	0.053^{**}	0.054^{**}	0.054^{**}
	(0.024)	(0.022)	(0.022)	(0.022)
Constant	43.341	43.388	43.989	44.035

Notes: Sample size: 9,918 observations (Level 1), 2,327 individuals (Level 2), and 42 work groups (Level 3).
*Significant at .10 level; **Significant at .05 level; ***Significant at .01 level.
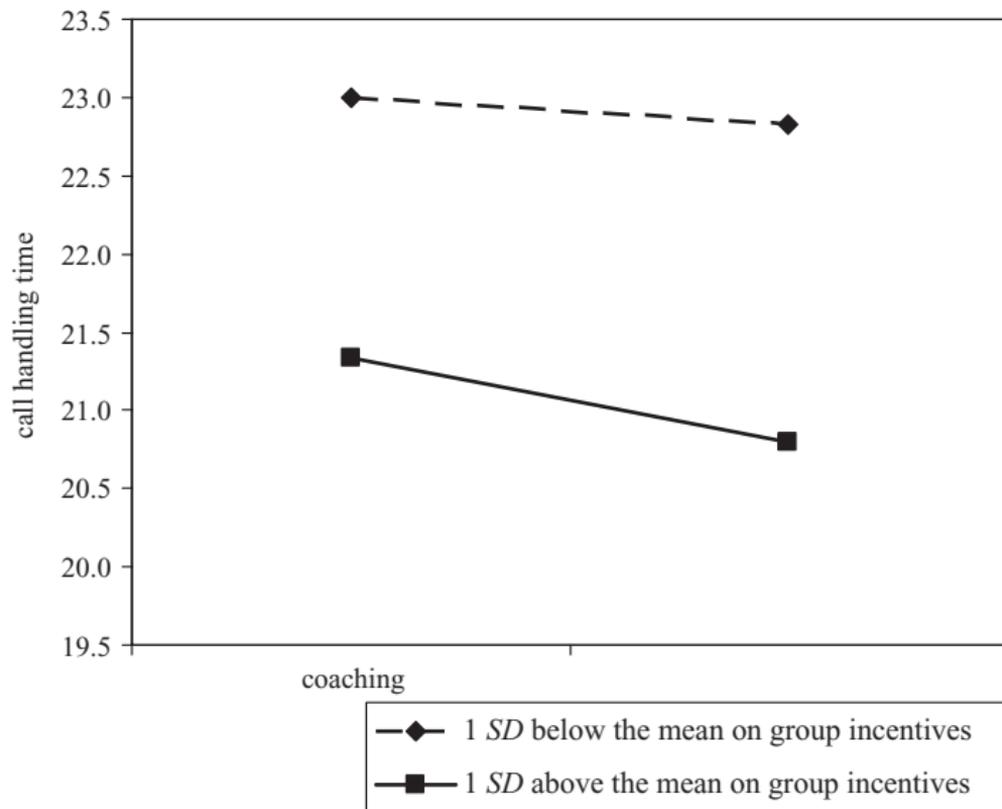
Figure 2: **Interaction Between Group Incentives and Coaching in Predicting Performance.**

to those of supervisors, pairing an individual with an experienced peer may inhibit the application of skills acquired from a supervisor. Use of project teams has no significant moderating effect. These results partially support Hypothesis 2b.

Hypothesis 3b predicted that process automation would moderate the relationship between coaching and performance in such a way that the coaching–performance link is weaker when automation is high. As Figure 3 shows, this hypothesis is fully supported (0.11, $p < .01$). Finally, the relationship between coaching and call handling time is lower when frequency of information updates is high (0.08, $p < .01$), providing support for Hypothesis 4b. Figure 4 illustrates the interactive effect.

Discussion

In this paper, we focused on the role of supervisors in influencing employee performance among incumbent workers in routine service jobs—an important subject and a setting that have been relatively understudied. Using a cross-level, longitudinal approach and hierarchical linear modeling, we sought to develop and test a multilevel model of how supervisors influence individual performance over time by integrating individual coaching and work group management activities and incentives. Our study produced three central findings. First, we confirmed the economic benefits of coaching, which had a strong and significant impact on improving individual performance over time. Second, how supervisors manage their work groups has a direct impact on individual performance, with the use of team activities and group incentives associated with significantly higher individual performance. In addition, technical processes influence performance, with greater automation associated with higher performance. Finally, we found that group incentives and technical processes moderated the relationship between coaching and performance. Specifically, the performance effect of coaching was stronger when supervisors made greater use of group rewards, when work automation was lower, and when process changes were less frequent.

Potential Limitations

There are a number of limitations to take into account when interpreting our findings. The generalizability of findings may be limited by the unique setting of the study with its highly standardized work processes and low levels of social interaction. However, in this study we have taken a "critical case" approach by choosing an environment where we would be less likely to find positive or synergistic effects of coaching and group management practices. Compared to many other settings, the degrees of freedom for supervisors to influence performance are relatively small due to high levels of process automation and routinized work tasks. Similarly, this setting of highly individualized work is an unlikely one in which to find that group management practices are effective. If coaching and supervisory efforts matter in this context, our findings should generalize to settings with more complex tasks and more opportunities for creativity and knowledge sharing. In fact, our findings of the interactive effects between coaching and automation support this argument. That is, even in this highly standardized environment, we found that coaching is more effective where automation is lower and group management practices are more frequent; therefore, coaching should be more effective in the many other types of occupations and organizations where processes are less standardized and opportunities for group interaction are higher.

Although this study did not provide a direct test of causality, we have employed a lagged approach (viewing performance as a function of coaching accumulated in the previous months) in order to separate causal antecedents from their outcomes. The random assignment of tasks across employees also strengthens
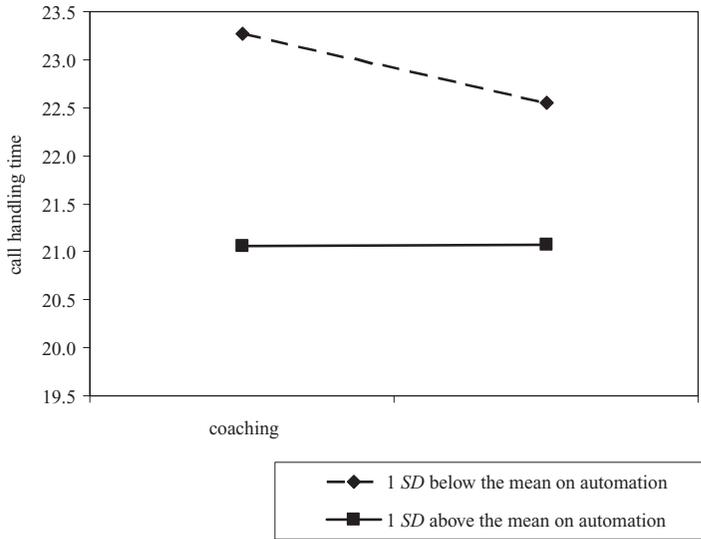
Figure 3: **Interaction Between Automation and Coaching in Predicting Performance.**
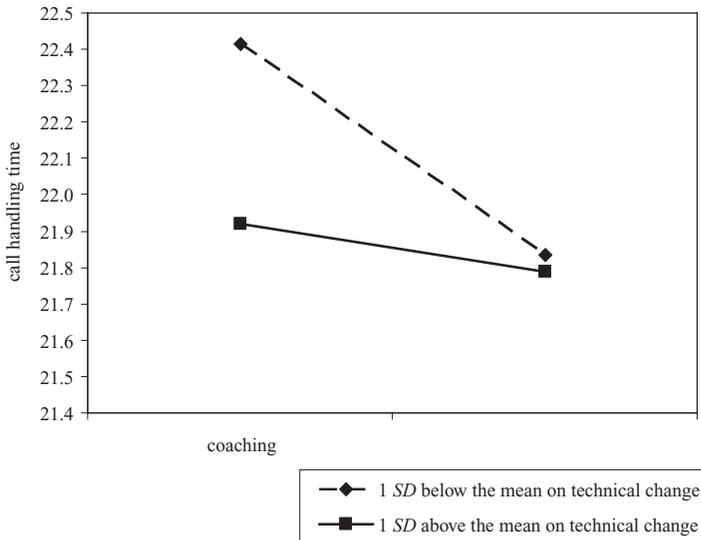
Figure 4: **Interaction Between Technical Change and Coaching in Predicting Performance.**

the research design, although it does not entirely mitigate problems of attributing causation. Moreover, although the data did not allow us to control for initial training, we mitigated this concern in several ways. First, we focused on a group of incumbent workers who have similar educational credentials. These workers have an average tenure of 10 years; initial training is probably not as important as overall company tenure—a proxy for firm-specific human capital, which we controlled for. Second, we controlled for the job proficiency of employees at time period one. Third, the use of random coefficients in HLM analysis partly mitigates this concern.

Finally, this study operationalizes coaching as the length of time that a supervisor provided individualized feedback and guidance. Although this measure improves upon previous measures, such as the incidence of coaching (e.g., Smither et al., 2003; whether or not any coaching occurred in the observation period), and objectively accesses the intensity of coaching, it does not capture how coaching was actually implemented in the workplace. Nevertheless, computerized records from the monitoring system indicated that the company categorized a majority of supervisor coaching as individualized feedback and performance improvement assistance (77% of total coaching time), followed by training about new procedures (10%) and region-specific business information (10%). In recent years, studies using behavioral measures of supervisor coaching have begun to emerge. For example, Heslin et al. (2006) developed a 10-item behavioral observation scale and asked subordinates to report the extent to which supervisors demonstrated those behaviors at work. This approach, however, may be subject to measurement bias. Future research may benefit from the combined use of objective measures and behavior-based instruments to fully explore the variation and complexity of supervisor coaching.

Theoretical Implications

Taking these limitations into account, this study makes several contributions to organizational approaches to the training and HR management literatures. First, we developed a multilevel model of coaching and individual performance, based on the Kozlowski and Salas' (1997) framework, and we tested the direct and synergistic relationships among coaching, group management activities, and performance. Second, we showed that a particular set of group management practices moderates the coaching performance relationship. In addition, by treating coaching and group management together as part of a set of HR management practices, we were able to emphasize the active role of supervisors in constructing complementary practices that reinforce the goals of learning. This moves us away from the idea of coaching as primary and organizational factors as context or secondary. The strength of these results is underscored by the fact that they are found in the unlikely environment of a service organization where work tasks are highly individualized.

These results also highlight the importance of incorporating informal training (i.e., coaching) into continuous improvement strategies. Although the literature often conceptualizes performance gains as a result of personal growth and development (Salas Cannon-Bowers, 2001), empirical studies typically have focused on differences in training and performance between individuals. Controlling for initial performance and using five waves of observations, we were, in effect, able to focus on variability around each individual's mean level of performance, control for some unmeasured individual characteristics, and model the residual effects attributable to changes in coaching over time. Future research may further strengthen our understanding of informal training by developing the nomological network of informal training and other interactional processes such as leader—member exchange at work (Scandura & Schriesheim, 1994).

The study also has implications for strategic HR management and the recent interest in the changing role of supervisors. We showed how supervisors influence performance via three dimensions of HR management— investment in training, group projects, and group incentives—providing an example of how the HR—performance link, which has been found to hold at the organizational level, operates among supervisors, work groups, and individual employees. The HR literature has noted the importance of decentralized HR systems (Purcell & Hutchinson, 2007) and has called for mesolevel studies and studies of implementation (Wright & Boswell, 2002; Wright & Nishii, 2009), but little research attention has focused on the roles of supervisors and line managers. This study indicates that it is not just the existence of formal HR policies but the informal implementation of practices by line managers that matter. The findings in this study strengthen the scientific basis for the role of supervisors in performance improvement and suggest the need for more HR studies that examine the sets of management practices that shape performance at this level of the organization.

Finally, we incorporated the direct and interactive effects of technology into the study of coaching and performance, moving beyond the current training literature that generally takes technology as a design feature (how technology can be an effective tool in learning and development; e.g., Brown, 2001). Similarly, the study signals the need for HR research to incorporate technology as a direct and moderating factor in studies of performance. In particular, the findings suggest the effectiveness of coaching but also identify limits in contexts in which technical change is high and supervisory knowledge is unable to keep pace with change.

From a methodological perspective, several features of this study may provide implications for future research. We reduced heterogeneity by focusing on one occupational group in one line of business in one company. We took a contextualized approach that captured a set of salient, proximal workplace practices and performance outcomes, consistent with recommendations by multilevel researchers (Kozlowski & Salas, 1997; Rousseau, 1985). Operators in this study learn and apply acquired skills to job duties in a natural setting. Unlike laboratory experiments that rely on student samples or simulated tasks or social relations, this study maximizes the "realism of context" (Scandura & Williams, 2000, p. 1251). In addition, because technical processes are often context specific, this approach is particularly important in studies that seek to incorporate the effects of technology into the analysis.

Furthermore, prior studies generally have relied on cross-sectional measures of performance to capture the benefits of training. Some scholars have shown the changing nature of performance across time and have criticized one-time measures that may introduce an unknown amount of measurement variance (Ployhart & Hakel, 1998). This can result in erroneous conclusions about the training– performance relationship. The longitudinal design in this study allows us to adopt a more dynamic view of performance and empirically examine within- and between-individual differences in performance growth trajectories. In addition, the large sample size provided sufficient power to adequately test out hypotheses. We also collected data from multiple sources (including the electronic monitoring system and surveys of workers and supervisors), which reduced the potential confound due to common method bias. Moreover, the random assignment of almost homogeneous tasks via call center technology provides a condition close to lab experiments, which reduces the possibility of statistical artifacts when individuals are selected into different assignments based on their competence.

Practical Implications

There are immediate practical implications of this research for call center operations but more general implications a broader set of occupations and management settings. For call centers, the findings are important because most corporations now make some use of these remote service delivery channels, and in many cases, they play a strategic role in managing the interface with customers. However, many firms continue to view these operations as cost centers, where the investment in training or HR practices should be minimized and where high turnover is viewed as inevitable—despite the fact that customer dissatisfaction is high. Effective use of coaching and group management practices is a cost efficient way to improve service quality and productivity. Call centers employ an estimated 3% of the U.S. labor force, or about 4 million employees, and despite the perceived popularity of offshoring, the comparable Indian call center workforce numbers less than 300,000 employees (Batt, Doellgast, & Kwon, 2006). In most countries around the world, including the U.S., the call center workforce is continuing to grow (Batt, Holman, & Holtgrewe, 2009), and the call center model of standardized, technology-mediated work organization has been adapted to a larger and larger swath of more complex jobs—from IT help desks to insurance agents and medical advisors.

In addition, the findings in this study are relevant to a broader set of low-skilled and semi-skilled service jobs where supervisors play an important role in the organizational hierarchy. Our study is meant to address the broad phenomenon of how supervisors manage employees who work individually or in loosely organized groups. A large portion of the labor market includes jobs that fit this description: clerical workers, bank workers, sales representatives, technicians, transport workers, postal workers, distributors, housekeepers, hotel workers, among others. Although companies may choose to organize these groups into interdependent teams, they often do not; rather, supervisors oversee employees, who are organized into administrative groups with varying levels of social interaction and group support for individual work.

More generally, in delineating and supporting the linkages among coaching, group management practices, and performance in a field setting, our study shows that supervisory coaching has clear economic benefits. In this case, the monthly wage of the average operator was $2,764, but that of a supervisor was $4,944. The wages spent on coaching equaled $48 per hour). Our results showed that 1 hour of coaching was associated with a 0.09 second reduction in handling time per phone call. This translates into a monthly return of $18 over the cost of the $48 investment in coaching. This study suggests that business practitioners should capitalize on the benefits of supervisory coaching and incorporate it as a valuable component in the learning system of organizations. However, research suggests that supervisors differ substantially in their inclination to coach their subordinates (Heslin et al., 2006). Supervisors often are reluctant to openly communicate or provide guidance because they do not have the time or they lack confidence when put in the position of "playing God" (Wexley & Latham, 2001). This is especially true when they do not have the skills or resources needed for coaching. Therefore, a practical implication of this research is for employers to equip supervisors with sufficient resources, as well as coaching and guidance skills, and encourage them to share work-related knowledge through group management practices.

Beyond the economic implications of coaching, this study illustrates the practical importance of group management practices. With other variables held constant, the use of team projects led to a 0.89 second reduction in call handling time, which means a 4.4% increase in performance, or labor savings of $180 every month. In terms of group incentives, a one standard deviation increase was related to 5.0% increase in performance, or $207 in monthly labor savings.

Although some have speculated that first-line supervisors might lose their importance due to the flattening of organizational structures and the use of information technologies (e.g., Kerr, Hill, & Broedling, 1986), this study adds to the emerging evidence that supervisors have a central role to play in functional HR practices such as employee development and performance management (Gittell, 2001; Hales, 2005;

McGovern et al., 1997; Purcell & Hutchinson, 2007). In the process of implementing formal organizational policies, supervisors interpret and enact these policies in different ways. This suggests that management has an important interest in designing effective training and management systems for frontline supervisors as well.

REFERENCES

Agarwal R, Angst CM, Magni M. (2009). The performance impacts of coaching: A multilevel analysis using hierarchical linear modeling. *International Journal of Human Resource Management, 20*, 2110–2134.

American Society for Training & Development. (2007). *State of the industry report*. Alexandria, VA: American Society for Training & Development.

Appelbaum E, Bailey T, Berg P, Kalleberg A. (2000). *Manufacturing advantage: Why high-performance work systems pay off*. Ithaca, NY: Cornell University Press.

Argote L, McGrath JE. (1993). Group processes in organizations: Continuity and change. In Cooper CL, Robertson IT (Eds.), *International review of industrial and organizational psychology* (Vol. 8, pp. 333–389). New York: John Wiley & Sons.

Baldwin TT, Ford JK. (1988). Transfer of training: A review and directions for future research. PERSONNEL PSYCHOLOGY, *41*, 63–105.

Batt R. (1999). Work organization, technology, and performance in customer service and sales. *Industrial and Labor Relations Review, 52*, 539–564.

Batt R. (2002). Managing customer services: Human resource practices, quit rates, and sales growth. *Academy of Management Journal, 45*, 587–597.

Batt R, Doellgast V, Kwon H. (2006). Service management and employment systems in U.S. and Indian call centers. In Collins S, Brainard L (Eds.), *Brookings trade forum 2005: Offshoring white-collar work – the issues and implications* (pp. 335–372). Washington, DC: The Brookings Institution.

Batt R, Holman D, Holtgrewe U. (2009). The globalization of service work: Comparative institutional perspectives on call centers (Introduction to a Special Issue). *Industrial and Labor Relations Review, 62*(4), 453–488.

Blanchard NP, Thacker J. (2007). *Effective training: Systems, strategies and practices (3rd edition).* Englewood Cliffs, NJ: Prentice-Hall.

Brown K. (2001). Using computers to deliver training: Which employees learn and why? PERSONNEL PSYCHOLOGY, *54*, 271–296.

Brown JS, Duguid P. (1991). Organizational learning and communities-of-practice: Toward a unified View of working, learning and innovation. *Organization Science, 12*, 40–57.

Burke MJ, Dunlap WP. (2002). Estimating interrater agreement with the average deviation index: A user's guide. *Organizational Research Methods, 5*, 159–172.

Burke MJ, Scheuer ML, Meredith RJ. (2007). A dialogical approach to skill development: The case of safety skills. *Human Resource Management Review, 17*, 235–250.

Byrk A, Raudenbush S. (1992). *Hierarchical linear models*. Thousand Oaks, CA: Sage.

Chao GT. (1997). Unstructured training and development: The role of organizational socialization. In Ford K, Kozlowski SWJ, Kraiger K, Salas E, Teachout MS (Eds.), *Improving training effectiveness in work organizations* (pp. 129–151). Hillsdale, NJ: Erlbaum.

Chen G, Bliese PD. (2002). The role of different levels of leadership in predicting self- and collective efficacy: Evidence for discontinuity. *Journal of Applied Psychology, 87*, 549–556.

Cohen SG, Bailey DE. (1997).What makes teams work: Group effectiveness research from the shop floor to

the executive suite. *Journal of Management, 23*, 239–290.

Combs J, Liu Y, Hall A, Ketchen D. (2006).How much do high-performance work practices matter? A meta-analysis of their effects on organizational performance. PERSONNEL PSYCHOLOGY, *59*, 501–528.

D'Abate CP, Eddy ER, Tannenbaum SI. (2003). What's in the name? A literature-based approach to understanding mentoring, coaching, and other constructs that describe developmental interactions. *Human Resource Development Review, 2*, 360–384.

Delery J. (1998). Issues of fit in strategic human resource management: Implications for research. *Human Resource Management Review,* 8(3), 289–310.

DeMatteo JS, Eby GT, Sundstrom E. (1998). Team-based rewards: Current empirical evidence and directions for future research. In Cummings LL, Staw B (Eds.), *Research in organizational behavior* (Vol. 20, pp. 141–183). Greenwich, CT: JAI Press.

Ellinger AD, Ellinger AE, Keller SB. (2003). Supervisory coaching behavior, employee satisfaction, and warehouse employee performance: A dyadic perspective in the distribution industry. *Human Resource Development Quarterly, 14*, 435–458.

Ford JK, Quinones MA, Sego DJ, Sorra JP. (1992). Factors affecting the opportunity to perform trained tasks on the job. PERSONNEL PSYCHOLOGY, *45*, 511–527.

Gittell JH. 2001. Supervisory span, relational coordination, and flight departure performance: A reassessment of postbureaucracy theory. *Organization Science, 12*(4), 468–483.

Guzzo RA, Dickson MW. (1996). Teams in organization: Recent research on performance and effectiveness. *Annual review of psychology* (Vol. 47, pp. 307–338). Palo Alto, CA: Annual Reviews.

Hackman JR. (1987). The design of work teams. In Lorsch JW (Ed.), *Handbook of organizational behavior* (pp. 315–342). Englewood Cliffs, NJ: Prentice-Hall.

Hackman JR, Wageman R. (2005). A theory of team coaching. *Academy of Management Review,* 30(2), 269–287.

Hales C. (2005). Rooted in supervision, branching into management: Continuity and change in the role of first-line manager. *Journal of Management Studies, 42*, 471–506.

Hall DT, Otazo KL, Hollenbeck GP. (1999). Behind closed doors: What really happens in executive coaching. *Organizational Dynamics, 27*, 39–52.

Hall L, Torrington D. (1998). Letting go or holding on–the devolution of operational personnel activities. *Human Resource Management Journal, 8*, 41–55.

Hansen DG. (1997). Worker performance and group incentives: A case study. *Industrial and Labor Relations Review, 51*, 37–49.

Heslin PA, VandeWalle D, Latham GP. (2006). Keen to help? Managers' implicit person theories and their subsequent employee coaching. PERSONNEL PSYCHOLOGY, *59*, 871–902.

Hollenbeck GP, McCall MW. (1999). Leadership development: Contemporary practices. In Kraut AI, Korman AK (Eds.), *Evolving practices in human resource management* (pp. 172–200). San Francisco: Jossey-Bass.

Holman D. (2000). A dialogical approach to skill and skilled activity. *Human Relations, 53*, 957–980.

James LR. (1982). Aggregation bias in estimates of perceptual agreement. *Journal of Applied Psychology, 6*, 219–229.

James LR, Demaree RG,Wolf G. (1984). Estimating within-group interrater reliability with and without response bias. *Journal of Applied Psychology, 69*, 85–98.

James LJ, Demaree RG, Wolf G. (1993). r_{WG}: An assessment of within-group interrater agreement. *Journal of Applied Psychology, 78*, 306–309.

Kerr S, Hill KD, Broedling L. (1986). The first-line supervisor: Phasing out or here to stay? *Academy*

of Management Review, 11, 103–117.

Klein KJ, Kozlowski SWJ. (2000). From micro to meso: Critical steps in conceptualizing and conducting multilevel research. *Organizational Research Methods, 3*, 211–236.

Kozlowski SWJ, Bell BS. (2003).Work groups and teams in organizations. In Borman WC, Ilgen DR, Klimoski RJ (Eds.), *Handbook of psychology: Industrial and organizational psychology* (Vol. 12, pp. 333–375). London: Wiley.

Kozlowski SWJ, Salas E. (1997). An organizational systems approach for the implementation and transfer of training. In Ford JK, Associates (Ed.), *Improving training effectiveness in work organizations* (pp. 247–287). Hillsdale, NJ: Erlbaum.

Kraiger K, Ford JK, Salas E. (1993). Integration of cognitive, behavioral, and affective theories of learning into new methods of training evaluation. *Journal of Applied Psychology, 78*, 311–328.

Kreft I, De Leeuw J. (1998). *Introducing multilevel modeling*. Thousand Oaks, CA: Sage.

Kunda G. (1992). *Engineering culture: Control and commitment in a high-tech corporation.* Philadelphia, PA: Temple University Press.

Lankau MJ, Scandura TA. (2002). An investigation of personal learning in mentoring relationships: Content, antecedents, and consequences. *Academy of Management Journal, 45*, 779–790.

Lave J, Wenger E. (1991). *Situated learning: Legitimate peripheral participation*. New York: Cambridge University Press.

LeBreton JM, Burgess JRD, Kaiser RB, Atchley EKP, James LR. (2003). The restriction of variance hypothesis and interrater reliability and agreement: Are ratings from multiple sources really dissimilar? *Organizational Research Methods, 6*, 80–128.

LeBreton JM, Senter JL. (2008). Answers to 20 questions about interrater reliability and interrater agreement. *Organizational Research Methods, 11*, 815–852.

Lefkowitz J. (1970). Effect of training on the productivity and tenure of sewing machine operators. *Journal of Applied Psychology, 54*, 81–86.

Little RJA, Rubin DB. (1987). *Statistical analysis with missing data*. New York: John Wiley & Sons.

Locke EA, Latham GP. (1990). *A theory of goal setting and task performance.* Englewood Cliffs, NJ: Prentice-Hall.

Mathieu JE, Tannenbaum SI, Salas E. (1992). Influences of individual and situational characteristics on measure of training effectiveness. *Academy of Management Journal, 35*, 828–847.

McAfee A. (2002). The impact of enterprise information technology adoption on operational performance: An empirical investigation. *Production and Operations Management, 11*, 33–53.

McGovern F, Gratton L, Hope-Hailey V, Stiles P, Truss C. (1997). Human resource management on the line? *Human Resource Management Journal, 7*, 12–29.

Mulec K, Roth J. (2005). Action, reflection, and learning–coaching in order to enhance the performance of drug development project management teams. *R&D Management, 35*, 483–491.

Olivero G, Bane KD, Kopelman RE. (1997). Executive coaching as a transfer of training tool: Effects on productivity in a public agency. *Public Personnel Management, 26*, 461–469.

Pasmore WA, Francis C, Haldeman J. (1982). Sociotechnical systems: A North American reflection on empirical studies of the seventies. *Human Relations,* 35, 1179–1204.

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

Purcell J, Hutchinson S. (2007). Front-line managers as agents in the HRM-performance causal chain: Theory, analysis and evidence. *Human Resource Management Journal, 17*, 3–20.

Quinones MA. (1995). Pretraining context effects: Training assignment as feedback. *Journal of Applied Psychology, 80*, 226–238.

Rouiller JZ, Goldstein IL. (1993). The relationship between organizational transfer climate and positive

transfer of training. *Human Resource Development Quarterly, 4*, 377–390.

Rousseau DM. (1985). Issues of level in organizational research: Multi-level and cross-level perspectives. In Cummings LL, Staw BM (Eds.), *Research in organizational behavior* (pp. 1–37). Greenwich, CT: JAI Press.

Rousseau DM, Fried Y. (2001). Location, location, location: Contextualizing organizational research. *Journal of Organizational Behavior, 22*, 1–13.

Salancik GR, Pfeffer J. (1978). A social information processing approach to job attitudes and task design. *Administrative Science Quarterly, 23*, 224–253.

Salas E, Cannon-Bowers JA. (2001). The science of training: A decade of progress. *Annual review of psychology* (Vol. 52, pp. 471–499). Palo Alto, CA: Annual Reviews.

Scandura TA, Schriesheim CA. (1994). Leader-member exchange and supervisor career mentoring as complementary constructs in leadership research. *Academy of Management Journal, 37*, 1588–1602.

Scandura TA, Williams EA. (2000). Research methodology in management: Current practices, trends, and implications for future research. *Academy of Management Journal, 43*, 1248–1264.

Sewell G. (1998). The discipline of teams: The control of team-based industrial work through electronic and peer surveillance. *Administrative Science Quarterly, 43*, 397–428.

Sisson GR. (2001). *Hands-on training: A simple and effective method for on-the-job training*. San Francisco: Berrett-Koehler.

Smith-Jentsch KA, Salas E, Brannick MT. (2001). To transfer or not to transfer? Investigating the combined effects of trainee characteristics, team leader support, and team climate. *Journal of Applied Psychology, 86*, 279–292.

Smither JW, London M, Flautt R,Vargas Y,Kucine I. (2003).Can working with an executive coach improve multisource feedback ratings over time? A quasi-experimental field study. PERSONNEL PSYCHOLOGY, *56*, 23–44.

Smither JW, Reilly SP. (2001). Coaching in organizations. In LondonM(Ed.), *How people evaluate others in organizations* (pp. 221–252). Mahwah, NJ: Erlbaum.

Stajkovic A, Luthans F. (1997). A meta-analysis of the effects of organizational behavior modification on task performance. *Academy of Management Journal, 40*, 1122–1149.

Tews MJ, Tracey JB. (2008). An empirical examination of posttraining on-the-job supplements for enhancing the effectiveness of interpersonal skills training. PERSONNEL PSYCHOLOGY, *61*, 375–401.

Tracey JB, Tannenbaum SI, Kavanaugh MJ. (1995). Applying trained skills on the job: The importance of the work environment. *Journal of Applied Psychology, 80*, 239–252.

Weitzman ML, Kruse DL. (1990). Profit sharing and productivity. In Blinder AS (Ed.), *Paying for productivity*. Washington, DC: Brookings Institution.

Werner S, Praxedes M, Kim H. (2007). The reporting of nonresponse analyses in survey research. *Organizational Research Methods, 10*, 287–295.

Wexley KN, Latham GP. (2001). *Developing and training human resources in organizations*. Englewood Cliffs, NJ: Prentice Hall.

Wright PM, Boswell W. (2002). Desegregating HRM: A review and synthesis of micro and macro human resource management research. *Journal of Management, 28*, 247–276.

Wright PM, Nishii LH. (2009). Strategic HRM and organizational behavior: Integrating multiple levels of analysis. In Guest D (Ed.), *Innovations in HR*. Oxford: Blackwell.

APPENDIX I

The Null Model

We hypothesized that job performance would be associated with coaching and characteristics of supervisor-led groups. Therefore, a precondition for the support of these hypotheses is that there must be significant within-person, between-person/within-group, and between-group variance in job performance. As such, the null model specifies job performance as the outcome variable without including any predictors. It can be described in HLM equation form as follows:

$$\text{Level 1} \quad Perf_{mig} = \pi_{0ig} + e_{mig} \qquad \text{(A1)}$$
$$\text{Level 2} \quad \pi_{0ig} = \beta_{00g} + \gamma_{0ig} \qquad \text{(A2)}$$
$$\text{Level 3} \quad \beta_{00g} = \gamma_{000} + u_{00g} \qquad \text{(A3)}$$

Where m represents time periods, i denotes individuals, and g denotes groups. $Perf_{mig}$ refers to job performance of the i_{th} operator in work group g in the m_{th} month as measured by call handling time. As described by Byrk and Raudenbush (1992), this model forces all of the within-person variance in performance over time into the Level 1 residual term (i.e., variance in e_{mig}), all of the between- person/within-group variance into the Level 2 residual term (i.e., variance in γ_{0ig}), and all of the between-group variance into the Level 3 residual term (i.e., variance in u_{00g}). In other words, this three-level model partitions the variance in job performance into its within-person, between-person/within-group, and between-group components.

Random Coefficient Regression Models

When there is significant variance across each of these three levels, we can turn to testing the hypotheses. In the Level-1 model, we hypothesize that job performance ($Perf_{mig}$) could be predicted by accumulated coaching ($Coaching_{mig}$) over time (Hypothesis 1). Hence, the Level-1 model will have two coefficients for each operator: the intercept and the *Coaching* slope.

$$\text{Level 1} \quad Perf_{mig} = \pi_{0ig} + \pi_{1ig} \cdot Coaching_{mig} + e_{mig} \qquad \text{(A4)}$$

Where $Coaching_{mig}$ is the accumulated informal training that individual i in work group g have received from Month 1 to Month m, π_{0ig} is the Level-1 intercept and π_{1ig} is the slope of $Coaching_{mig}$, and e_{mig} is the Level-1 random effect.

Hypotheses 2a, 3a, and 4a predict that group characteristics have cross-level main effects on individual performance. We can use *intercepts-as-outcomes* models to test these hypotheses.

Level 2

$$\pi_{0ig} = \beta_{00g} + \beta_{01g} \cdot OrgTen_{ig} + \beta02g \cdot Intial\ performance_{ig} + \beta_{03g} \cdot Intial\ performance\ dummy_{ig} + \gamma_{0ig} \quad \text{(A5)}$$

$$\pi1_{ig} = \beta_{10g} + \gamma_{1ig} \qquad \text{(A6)}$$

Level 3

$$\beta_{00g} = \gamma_{000} + \gamma001 \cdot Pairing_g + \gamma_{002} \cdot Team\ project_g$$
$$+ \gamma_{003} \cdot Group\ incentives_g + \gamma_{004} \cdot Automation_g \qquad (A7)$$
$$+ \gamma_{005} \cdot Process\ change_g + \gamma_{005} \cdot Group\ size_g + u_{00g}$$

$$\beta_{01g} = \gamma_{010} + u_{01g} \qquad\qquad (A8)$$
$$\beta_{02g} = \gamma_{020} + u_{02g} \qquad\qquad (A9)$$
$$\beta_{03g} = \gamma_{030} + u_{03g} \qquad\qquad (A10)$$
$$\beta_{10g} = \gamma_{100} + u_{10g} \qquad\qquad (A11)$$

Hypotheses 2b, 3b, and 4b predict that supervisory practices moderate the relationship between informal training and performance. The hypothesized cross-level interaction can be specified as a *slopes-as-outcomes* model by substituting equation (A11) with equation (A12) as follows.

$$\beta10g = \eta_{100} + \eta_{101} \cdot Pairing_g + \eta_{102} \cdot Team\ project_g$$
$$+ \eta_{103} \cdot Group\ incentives_g + \eta_{104} \cdot Automation_g \qquad (A12)$$
$$+ \eta_{105} \cdot Process\ chang_{eg} + v_{10g}$$