2007

The Past, Present, and Future of Dynamic Performance Research

Michael C. Sturman
Cornell University, mcs5@cornell.edu

The Past, Present, and Future of Dynamic Performance Research

Abstract
This article reviews the extensive history of dynamic performance research, with the goal of providing a clear picture of where the field has been, where it is now, and where it needs to go. Past research has established that job performance does indeed change, but the implications of this dynamism and the predictability of performance trends remain unresolved. Theories are available to help explain dynamic performance, and although far from providing an unambiguous understanding of the phenomenon, they offer direction for future theoretical development. Dynamic performance research does suffer from a number of methodological difficulties, but new techniques have emerged that present even more opportunities to advance knowledge in this area. From this review, I propose research questions to bridge the theoretical and methodological gaps of this area. Answering these questions can advance both research involving job performance prediction and our understanding of the effects of human resource interventions.

Keywords
dynamic performance research, job performance

Disciplines
Human Resources Management

Comments

The Past, Present, and Future of Dynamic Performance Research

Michael C. Sturman

This article reviews the extensive history of dynamic performance research, with the goal of providing a clear picture of where the field has been, where it is now, and where it needs to go. Past research has established that job performance does indeed change, but the implications of this dynamism and the predictability of performance trends remain unresolved. Theories are available to help explain dynamic performance, and although far from providing an unambiguous understanding of the phenomenon, they offer direction for future theoretical development. Dynamic performance research does suffer from a number of methodological difficulties, but new techniques have emerged that present even more opportunities to advance knowledge in this area. From this review, I propose research questions to bridge the theoretical and methodological gaps of this area. Answering these questions can advance both research involving job performance prediction and our understanding of the effects of human resource interventions.

Introduction

The extensive history to the study of employee job performance is filled with research that is predominantly static in nature. That is, most of this research examines the correlates of various sorts of job performance ratings, with the often implicit assumption that the results would generalize to the same population of subjects at any other point in time. Yet there is abundant, and as I will argue in this paper overwhelming, evidence that indeed individuals' job performance does change with time. Accepting that an individual's job performance changes, also known as dynamic performance or dynamic criteria, requires research on job performance to consider the effects associated with the passage of time (Hulin, Henry, & Noon, 1990). Sometimes, time-related issues seem to receive some acknowledgment, such as demonstrated by the extensive use of variables such as age, organizational tenure, or job experience as controls (Sturman, 2003), yet job performance research still primarily focuses on the cross-sectional prediction of what is commonly called the criterion (Austin & Villanova, 1992; Campbell, 1990; Dunnette, 1963).

Even though the examination of job performance at a given point in time is most common, job performance, perhaps more than any other individual-level variable in organizational research, has been examined in conjunction with time. Research has considered job performance longitudinally since at least the 1940s (e.g., Kunst, 1941; Rothe, 1946, 1947), and effects associated with time have long been recognized as important when measuring job performance (e.g., Ghiselli, 1956; Ghiselli & Haire, 1960). Research on dynamic performance has most often been framed as a critical issue for selection (e.g., Ghiselli & Haire, 1960; Henry & Hulin, 1989; Prien, 1966; Steele-Johnson, Osburn, & Pieper, 2000). If performance changes over time, then the validity of selection devices for predicting job performance obtained from an original validation study may not be stable over time. But the impact of a dynamic criteria reaches far beyond just selection. Human resource researchers investigate how employees are selected, placed, developed, trained, appraised, and compensated within their organizations, processes all intended to affect job performance and all inherently involving the passage of time. For our field to understand employee job performance we require an understanding of what happens to this performance with the passage of time. Nevertheless, despite the time that has passed since issues related to dynamic performance were first raised, there has been less progress in this area than its long history might suggest.

The study of dynamic performance is also complicated by a wide array of methodological issues. First, although time is inherently a longitudinal issue, cross-sectional research can be used to address questions related to the effects of time. It is important to understand where cross-sectional research can and cannot help explain the effects on performance associated with time. Second, recent advances in analytical methods are providing new means to analyze longitudinal data. These methods open up a wide range of possibilities for examining job performance as it relates to time, with each method possessing different assumptions, advantages, and weaknesses. Third, the very nature of studying job performance over time gives rise to a variety of methodological problems that will confront all research on the topic. These issues cause any longitudinal analysis of performance ratings to be at least somewhat flawed, and so it is important to understand the implications of these necessary data limitations. To advance our knowledge about dynamic performance, it is important to have a good understanding of the methodological issues facing those who study the effects associated with time on job performance.

The purpose of this article is to review the current state of knowledge about dynamic performance, discuss the relevant analytical methods and issues, and provide some structure to emerging research in this area. The alignment of past work, past theories, new theories, and methodological advances provides an exciting opportunity for research of both applied and theoretical value. It is my hope that not only will this review clarify the current state of knowledge regarding job performance considered within the context of time, but also will inspire more research on the phenomenon.

Defining Job Performance and Dynamic Performance

Before reviewing research and theory on dynamic performance, it is critical first to articulate both how job performance itself and the phenomenon of dynamic performance have been defined. Past research in this area has used a variety of measures and definitions of both. My goal here is to review past practices and provide clear definitions that I will employ for the rest of this paper.

Defining Job Performance

Past research on job performance has most commonly defined the construct as behaviors that are under the control of the individual and that contribute to the goals of the organization (Dunnette, 1963; Campbell, 1990; Campbell, McCloy, Oppler, & Sager, 1993; Motowidlo, Borman, & Schmitt, 1997; Rotundo & Sackett, 2002). A key issue here is that job performance is defined as behavior, and hence distinguishable from the results of such behavior. While this definition is applicable to the performance of work in any role within any form of organization (Campbell, 1990), I will assume here that this behavior is within the context of an employment relationship. The employing organization is also assumed to have goals, with the job performance in some way (directly or indirectly) being able to contribute to those goals.

Past research focused on understanding the definition and conceptualization of job performance has usually employed this definition. Other research, and particularly research involving performance over time, has also considered job performance in different ways. Previous longitudinal job performance research often uses results-based measures, such as sales or output rates (Sturman, Cheramie, & Cashen, 2005). It is also important to note that job performance has been considered as the organizational value associated with employees' behaviors (Brogden & Taylor, 1950; Schmidt & Kaplan, 1971). Papers taking this perspective are based on the idea that employees' behaviors and the results of their behaviors have a direct or indirect association with organizational value, and this value can be approximated and studied as a substantive outcome. This performance value, often referred to as utility, relates to the particular monetary value associated job performance behaviors (Boudreau, 1991). From this perspective, research has estimated the value of various human resource programs by considering the stream of costs and benefits associated with employee performance, often in a longitudinal context (e.g., Boudreau & Berger, 1985; Sturman, 2000; Sturman, Trevor, Boudreau, & Gerhart, 2003).

Any review of the research on job performance and time must therefore be careful to distinguish between, but still consider, the various forms of performance that have been examined. It is important (1) to note that these conceptualizations of job performance are very different and (2) to make a distinction between the theoretical and methodological issues relevant to understanding each. For this paper, I specifically distinguish between job $\text{performance}_{\text{(behaviors)}}$, job $\text{performance}_{\text{(results)}}$, and job $\text{performance}_{\text{(utility)}}$. Unless otherwise noted, for simplicity and space, references to ‘‘job performance’’ will refer to job $\text{performance}_{\text{(behaviors)}}$.

Defining Dynamic Performance

Past Definitions

Research considering job performance over time has also devoted energy to the definition of what it means for performance to be dynamic. For much of the history of the literature, three definitions generally prevailed. Two definitions involve evidence from the individual-level of analysis; the third about changes at the group-level.

First, performance dynamism may be defined as occurring when the rank-ordering of scores on the criterion change over time (Barrett, Caldwell, & Alexander, 1985; Deadrick & Madigan, 1990; Hanges, Schneider, & Niles, 1990). This definition of dynamic performance has most often led to the examination of correlations between criterion scores at multiple points in time. Such studies have been framed as considering the test–retest reliability or the stability of performance ratings.

Second, performance dynamism has been defined as occurring when predictor validities change over time (Austin, Humphreys, & Hulin, 1989; Barrett et al., 1985; Ghiselli, 1956; MacKinney, 1967; Prien, 1966; Smith, 1976; Steele-Johnson et al., 2000). Research using this definition has focused on examinations of the validity of selection devices for predicting job performance of employees over multiple time periods. Some have argued that a dynamic criterion would lead to a decrease in validity over time (Austin et al., 1989); others have suggested that simply the fluctuation of validity is evidence of a dynamic criterion (Barrett et al., 1985); still others have argued that dynamic criteria could lead to predictors becoming more valid with time (Ackerman, 1987; Murphy, 1989).

Third, performance dynamism has been defined as changes over time in the average level of group performance (Barrett et al., 1985; Hanges et al., 1990). This definition has been criticized as the weakest conceptually and operationally (Barrett et al., 1985). In part, average performance curves may not reflect the shape of the individual performance curves comprising them. Group-level performance could even change when individuals' performance remains constant if the performance level of those leaving the organization were different than the performance level of those entering (Boudreau & Berger, 1985).

Proposed Definition of Dynamic Performance

The three definitions for dynamic performance present an interesting divergent set of ways of considering performance over time, and the use of any single definition has often led to very different research tasks. The problem with these definitions is that they do not present a logically consistent set of classifications. That is, it is possible for the first condition to be met without meeting the second or third definitions. Similarly, the third definition could be met without meeting the first two. The reason this occurs is that the second and third definitions consider the potential consequences of dynamic The Past, Present, and Future of Dynamic Performance Research 53 performance. For this reason, the definition of dynamic performance should be based on the first definition – changes in the rank-order (or correlations) of job performance over multiple time periods – because it is the only definition that directly addresses the issue of stability (Hanges et al., 1990). Moreover, this dynamism should occur specifically for job performance$_{\text{(behaviors)}}$. While the same definition (i.e., changes in the rank-order) can be applied to any outcome, the definition of job performance should be consistent with the view that job performance connotes behaviors. Environmental changes that affect performance results or utility (such as changes in situational constraints), while potentially related to job performance$_{\text{(behaviors)}}$, should be recognized as a different phenomenon and not direct evidence of dynamic performance. With this perspective, research can easily distinguish between dynamic performance and the consequences of this dynamism, such as changes in the validity of selection devices (i.e., the second definition), changes in job performance ratings aggregated to a group-level (i.e., the third definition), changes in job performance$_{\text{(results)}}$, or changes in job performance$_{\text{(utility)}}$.

The first definition, though, needs to be considered carefully. A correlation less than one between performance measures is not necessarily indicative of performance dynamism. Rather, correlations between performance measures over time may be affected by measurement error rather than actual changes in job performance (Barrett et al., 1985; Hanges et al., 1990; Sturman et al., 2005), and it is important to distinguish between temporal consistency, stability, and test–retest reliability (Sturman et al., 2005). Temporal consistency is the correlation between performance measures at different points of time (Heise, 1969; Sturman et al., 2005). It captures the relationship between measures of job performance but not necessarily of the true construct of performance. Test–retest reliability refers to the amount of transient error that affects ratings of job performance at different points in time (Sturman et al., 2005). For performance to be dynamic, changes must occur to the construct of performance. This has been defined as stability: the extent to which the true value of a construct remains constant over time (Carmines & Zeller, 1979; Sturman et al., 2005). I thus define dynamic performance as a lack of stability in job performance$_{\text{(behaviors)}}$ over time.

The Past: Three Streams of Dynamic Performance Research

When one looks at the body of research related to performance and time, three streams of work emerge. All three lines of research involve the prediction of job performance, yet the way these goals are pursued are markedly differently, thereby involving notably different theoretical and methodological issues. These three areas of research are (1) the search for evidence of the dynamic performance phenomenon, (2) the prediction of changes in job performance, and (3) the prediction of job performance trends.

Evidence of Dynamic Performance

The earliest work on dynamic performance primarily focused on measuring job performance over time and the implications of any inconsistency for the validation of selection devices. Much of this early research addressed the question of ‘‘is performance dynamic?’’ That is, does job performance satisfy the earlier definitions of dynamic performance articulated above.

This work on dynamic performance was concerned with describing the nature of performance consistency. Essentially, this research challenged the assumption of a criterion that is reliable across time. While psychological research often insists on a highly reliable measure of job performance (or for that matter, any criterion) assessed at a point in time, scant attention was paid to whether the criterion had reliability from one time-period to the next. Consequently, a body of research emerged examining the reliability of performance ratings at various time lags (e.g., Ghiselli, 1956; Prien, 1966; Rambo, Chomiak, & Price, 1983; Rambo, Chomiak, & Rountree, 1987; Rothe, 1946, 1978; Rothe & Nye, 1958, 1959, 1961).

Other work in this area sought to determine the prevalence of simplex (or quasi-simplex) patterns in measures of job performance (e.g., Bass, 1962; Deadrick & Madigan, 1990; Dennis, 1954, 1956; Ghiselli & Haire, 1960; Hanges et al., 1990; Henry & Hulin, 1987). For job performance, the simplex pattern of correlations (Guttman, 1955; Humphreys, 1960) is a systematic decrease in the magnitude of correlations between measures of job performance as the time-span between performance measures increases. A perfect simplex is based on a model with no or negligible measurement error; a quasi-simplex model includes a measurement model (Jöreskog, 1970). If job performance follows a simplex or quasi-simplex pattern, and especially if that led to correlations between measures of performance approaching zero, then this would suggest that the validity of selection devices could not be generalized across time. If true, then the utility (economic and practical) of selection devices would be much lower than cross-sectional research has suggested (Henry & Hulin, 1987).

In all, this body of research resulted in strong, arguably undeniable, support of a lack of performance consistency. Empirically reviewing the research, Sturman et al. (2005) attempted to partial out unreliability from stability, and thus present information on the extent to which performance truly is dynamic. The results of their study showed that while there is evidence of test–retest unreliability (and other measurement error) causing some of the observed inconsistency in job performance ratings over time, job performance ratings (both job performance(behaviors) and job performance(results)) are dynamic. While there remains debate in the literature as to the pervasiveness and extent of performance changes (e.g., Austin et al., 1989; Barrett & Alexander, 1989; Barrett et al., 1985), and there does appear to be at least some portion of job performance that is stable over time (Hanges et al., 1990; Sturman et al., 2005), there is now abundant research and general consensus that job performance does change over time (Deadrick &Madigan, 1990; Deadrick, Bennett, & Russell, 1997; Henry & Hulin, 1987; Hofmann, Jacobs, & Baratta, 1993; Hofmann, Jacobs, & Gerras, 1992; Hulin et al., 1990; Ployhart & Hakel, 1998; Sturman & Trevor, 2001; Sturman et al., 2005). Still in question, though, are the implications of performance dynamism and the causes and correlates of individual job performance changes over time.

Changing Predictability of Job Performance

Explicitly stated in some research on dynamic performance, and implicit in others, is that the presence of dynamic criteria poses a significant problem for the prediction of job performance over time (i.e., reviewed earlier as the formerly second definition of dynamic criteria). Indeed, this was a concern raised by a number of researchers examining dynamic criteria (e.g., Ghiselli, 1956; Hanges et al., 1990; Henry & Hulin, 1987; MacKinney, 1967; Prien, 1966; Rambo et al., 1983; Smith, 1976). Some argued that the existence of dynamic criteria does not necessarily mean a lack of predictability (Ackerman, 1988, 1989; Barrett, Caldwell, & Alexander, 1989, 1992; Hanges et al., 1990). This led to an extensive debate in the literature on the effect of time on the validity of job performance predictors.

In their examination of validities examined longitudinally, Barrett et al. (1985) found examples of both stable and instable validities. They concluded that, ‘‘factors such as temporal unreliability and restriction of range serve as viable explanations in the few instances where significant change over time was found'' (Barrett et al., 1985, p. 53). Overall, they argued that evidence of dynamic criteria (as specified by the second definition) was relatively rare. Other researchers took an opposing view, arguing that the same evidence reviewed in Barrett et al. (1985) was not as dismissive of a dynamic criterion as Barrett et al. suggest (Austin et al., 1989). A similar debate emerged soon thereafter. A paper by Henry and Hulin (1987) argued that ''instability and change in nearly all areas of human performance, skills, and measures of general ability are more to be expected than is stability'' (p. 461) and therefore the long-term predictability of performance is questionable. This paper was criticized by Ackerman (1989), who argued that while job performance ratings may follow a simplex patter, ''ability measures can maintain levels of predictive validity over time and, when chosen properly, may actually increase'' (Ackerman, 1989, p. 364), followed by a rejoinder by Henry and Hulin (1989) countering some of the criticisms. The point here is not specifically to weigh in on these debates, but their review shows that there are divergent opinions on the matter, and the evidence had not yielded definitive conclusions for the field.

In one of the most comprehensive examinations of performance predictors over time, Keil and Cortina (2001) examined the validity of cognitive ability, perceptual speed ability, and psychomotor ability to predict job performance. They found that the validities deteriorate with time. This deterioration occurred for all three predictors, and for both consistent and inconsistent tasks. Although they argued that this deterioration is pervasive, there are still examples from other research of selection devices maintaining their predictability over time.

Published in the same year, Farrell and McDaniel (2001) examined how well general mental ability, perceptual speed, and psychomotor ability predicted job performance at various experience levels in a large cross-sectional sample of employees. They found that the correlation between the various abilities and performance did differ when the model was divided by job consistency. They also found some instances of correlations increasing with experience, decreasing with experience, and fluctuating (decreasing and then increasing again) with experience.

Other studies can be found that also show that there is no definitive answer to this research question. For example, McEvoy and Beatty (1989) examined the validity of an assessment center over seven years. While the correlations varied (from 0.19 to 0.41 for supervisory ratings of performance), the authors demonstrated that the selection device had validity over an extended period of time. Tziner, Ronen, and Hacohen (1993) also demonstrated the long-term validity of an assessment center. Similarly, Deadrick and Madigan (1990) examined the validity of The Past, Present, and Future of Dynamic

Performance Research 57 psychomotor ability, cognitive ability, and experience over a six-month period. They found that psychomotor ability predicted initial performance, and cognitive ability predicted performance growth. In conclusion, while it appears that Keil and Cortina (2001) provide strong evidence that the validities of ability measures decrease with greater time lags, conflicting (but albeit cross-sectional) findings of Farrell and McDaniel (2001), along with the presence of some other exceptions from longitudinal studies and the contradictory findings with regard to assessment centers, keeps alive the question about what happens to the validity of selection devices over time.

The long history of research and debate in this area might suggest that issues of dynamic performance would be salient in the staffing literature. However, selection research still often ignores the dimension of time. For example, in recent studies on staffing and selection tools, time is not considered (e.g., Behling, 1998; Carlson, Connerley, & Mecham, 2002; Chait, Carraher, & Buckley, 2000; Stevens & Campion, 1999; Ryan, McFarland, Baron, & Page, 1999; Ryan & Tippings, 2004). Similarly, current texts on selection tend to devote little space to the role of time. Time may be mentioned with regard to estimating test–retest reliability (e.g., Heneman, Heneman, & Judge, 1997). The text by Gatewood and Feild (2001) does briefly mention Hulin et al.'s (1990) conclusion that the validity of some measures decay with time, although the concern is quickly dismissed and there is no real discussion of the implications of performance changes. In an exception, Ployhart, Schneider, and Schmidt (2006) explicitly state that performance is dynamic, and that this has implications for validity, but that there is enough stability in performance for it to be predicted. Nonetheless, the attention devoted to the role of time remains minimal, despite that the goals of staffing are to ''improve organizational functioning and effectiveness by attracting, selecting, and retaining people who will facilitate the accomplishments of organizational goals. ...'' (Ployhart et al., 2006, p. 2, emphasis added). Inherent in this definition is the passage of time. If performance is defined as the behaviors that contribute to the goals of organizations, and if these behaviors change over time, then the passage of time is critical to the issue of selection.

Researchers studying job performance and time have often disagreed as to the proper interpretation of past evidence, but they all seem to agree that more research is needed into the implications and consequences of its instability (Ackerman, 1989; Barrett et al., 1985, 1989; Hanges et al., 1990) or inconsistency (Austin et al., 1989). This research should involve better conceptualizations of the outcomes being predicted (Austin et al., 1989) and the identification and removal of intrinsic and extrinsic source of criterion unreliability in measures of job performance over time (Barrett et al., 1985; Sturman et al., 2005).

Nonetheless, the question of ‘‘is performance dynamic?’’ as defined in this paper, is resolved. There is no need for future research specifically to address this question. However, the field is still far off from a clear understanding of what happens to job performance over time, what causes it to be dynamic, how effectively selection devices work over time, and how human resource interventions can be used to affect job performance when considered in a longitudinal context.

Predicting Performance Trends

The most recent development in the dynamic performance literature has been the examination of employee performance trends. This new line of research presents a shift in the focus of dynamic performance research to investigations of individual change patterns (Hofmann et al., 1993).

The goal of this line of research is to model within-person patterns of performance and to understand what affects these patterns. By its nature, this research is interested in the prediction of job performance at more than one point in time (Ployhart, Holtz, & Bliese, 2002). The early work in this area simply demonstrated that modeling individual performance trends was possible. Performance trends were shown to be systematic, and hence predictable (Hofmann et al., 1992, 1993). Research in this area then expanded on this finding by similarly modeling performance trends, but also considering individual-level characteristics which predict the level and slope of these trends (Deadrick et al., 1997). Later research expanded both the complexity of the performance trend model (to non-linear patterns) and the types of predictors used to explain the trends (Day, Sin, & Chen, 2004; Ployhart & Hakel, 1998; Stewart & Nandkeolyar, 2006; Thoresen, Bradley, Bliese, & Thoresen, 2004).

Related to this stream of research has been the work examining the consequences of performance trends. Some research has examined how performance trends influence the likelihood of turnover (Harrison, Virick, & William, 1996; Sturman & Trevor, 2001). Both studies found that changes in subsequent performance scores affected the likelihood of turnover, and Sturman and Trevor (2001) showed that even after controlling for this change, long-term performance trends also predicted turnover. This research shows that performance changes (short-term and long-term) may be valuable The Past, Present, and Future of Dynamic Performance Research 59 predictors in other longitudinal phenomena, even though work on the implications of performance trends is still in its early stages.

Investigations in this third stream of dynamic performance research show that performance trends are both predictable and related to outcomes of interest. As a result, it seems clear that, consistent with the conclusions of many in the second stream of research reviewed above, selection devices need to be considered not just in terms of their validity at a point in time (or even validity at several points in time), but in terms of how well they predict performance trends. Furthermore, the dynamic nature of performance and the predictability of performance trends suggest that for essentially all areas of human resource research interested in the prediction of job performance ratings, it would provide a more accurate understanding of performance to consider the predictability of performance levels and trajectories.

The Present: Current Theory Relevant to Dynamic Performance

Over the history of research on dynamic performance, repeated calls have been made for more theoretical development (e.g., Austin et al., 1989; Campbell, 1990; Deadrick & Madigan, 1990; Deadrick et al., 1997; Henry & Hulin, 1987; Hofmann et al., 1992, 1993; Hulin et al., 1990; Steele-Johnson et al., 2000). While more theoretical progress is certainly desired, there are some notable works that have considered the issue of what happens to performance over time that provide a useful framework for research in this area. Some of this research has been widely cited in the dynamic performance literature. Others are relevant but have not been extensively applied or developed within the literature. The purpose of this section is to review theoretical perspectives that are applicable for studying performance over time, hopefully presenting opportunities for greater clarification and demonstrating where future research is most needed.

Changing Subjects and Changing Tasks Models

Two models have emerged directly from the literature on dynamic performance to help explain why the relationship between predictors and performance changes over time: the changing tasks model and the changing subjects model. While, as reviewed above, there is debate as to the extensiveness and rapidity of decreases in validity, both models are valuable for understanding why performance changes over time.

The Changing Subjects Model

The changing subjects model (also referred to as the changing-person model; e.g., Keil & Cortina, 2001) posits that individuals possess various characteristics which result in (i.e., cause) performance (be it performance on a task or performance on a job; I will be focusing exclusively on job performance). While most uses of this model have considered abilities, it may also refer to characteristics such as motivation and job knowledge. Because these performance-causing characteristics change over time, performance levels change even if the contribution of these characteristics to performance remains constant (Adams, 1957; Alvares & Hulin, 1973; Deadrick & Madigan, 1990; Humphreys, 1960; Keil & Cortina, 2001).

While employed in an organization, a multitude of changes occur to an individual than can affect performance. For example, while holding a given job, an employee accumulates experience that then affects performance levels (e.g., McDaniel, Schmidt, & Hunter, 1988; Schmidt, Hunter, & Outerbridge, 1986; Sturman, 2003). Simultaneously, aging may affect performance (e.g., Lawrence, 1988; Rhodes, 1983; Salthouse, 1979; Sterns & Doverspike, 1989; Sturman, 2003; Waldman & Avolio, 1993). Because experience and age are related to job performance, the changes in the individual's characteristics cause job performance to change with the passage of time. Although other models (discussed below) provide additional explanations as to why performance changes over time, research supports the changing subjects model as at least a partial explanation as to why performance changes over time (Deadrick & Madigan, 1990; Dunham, 1974). Indeed, much of the field of training is based on the idea that individual characteristics which cause performance can be changed, so performance can be improved through effective training (e.g., Noe, 2005). The logic behind changing compensation plans is also similarly based on the idea that incentives can cause individuals to change in ways that affects their performance levels (e.g., Milkovich & Newman, 2005).

Certainly, there has been debate regarding the validity of the changing subjects model, but much of this debate was caused by issues related to the definition of abilities. If one considers an ability to be a relatively static trait, then there are definitional flaws with the changing subjects model if one defines it as changes in abilities causing performance changes. However, by broadening the model to consider abilities and The Past, Present, and Future of Dynamic Performance Research 61 skills (e.g., Keil & Cortina, 2001), or broadening it even further as I do above by considering all individual-level performance-causing characteristics (including abilities, skills, knowledge, and motivation), then the focus of the model is wider and more consistent with static models of job performance (e.g., Campbell, 1990; Motowidlo et al., 1997; Schmidt et al., 1986). Additionally, with a broader focus, the changing-subjects model can be seen as complementary to the changing tasks model rather than as a competing alternative explanation (Keil & Cortina, 2001).

The Changing Tasks Model

In addition to changes in individual characteristics, performance changes may be attributable to job changes, new job roles, or revised organizational requirements. The changing tasks model predicts that an individual's performance changes because the determinants of performance change (Alvares & Hulin, 1972; Deadrick & Madigan, 1990; Fleishman & Hempel, 1954). Changes in job requirements – such as after a promotion, transfer, the introduction of new technology, or other change in job duties – may lead to the need for new sets of abilities while reducing the impact of current abilities on job performance (Alvares & Hulin, 1972; Fleishman, 1953, 1960; Fleishman & Hempel, 1954; Fleishman & Rich, 1963; Murphy, 1989; Steele-Johnson et al., 2000). For example, a scientist may be promoted to a management position (e.g., Boudreau & Berger, 1985). In this circumstance, the company might lose a high performing scientist while gaining a poor performing manager. When an employee changes jobs, individual characteristics may remain the same, but the changes in the job duties may cause different individual characteristics (e.g., managerial experience and knowledge rather than scientific experience and knowledge) to become determinants of job performance.

By drawing attention to how the requirements of individual performance change over time, the changing tasks model may also explain variance in performance remaining after controlling for prior performance (e.g., Sturman et al., 2005). The effect of task changes on performance dynamism depends on the similarity between the old job and the new job. The greater the similarity, the more that past performance should be able to predict future performance.

The logic behind the changing tasks model is consistent with the underlying assumption behind such organizational actions such as work redesign (e.g., Hackman & Oldham, 1976) and empowerment (e.g., Lawler, 1986). That is, by changing the nature of the job, environment, or organization, employee job performance can be improved.

Static Models of Job Performance and their Implications for Time

A common approach for conceptualizing the determinants of job performance is some form of static model, such as

$$\text{Performance} = f\ (\text{motivation}, \text{ability}),$$

$$\text{Performance} = f\ (\text{motivation}, \text{ability}, \text{opportunity}),$$

$$\text{Performance} = f\ (\text{declarative knowledge}, \text{procedural knowledge and skill}, \text{motivation}).$$

As noted by Campbell (1990), determining the precise functional form of such a model is likely impossible; however, considering this sort of model does provide insights for understanding the nature of performance changes over time, particularly if one builds upon the changing subjects and changing tasks models discussed above.

Even if not determining all the specific causes of performance, one can generally specify that performance is a function of certain characteristics, some of whom are stable and some of which change with time. For example, cognitive ability is generally shown to be highly correlated with job performance (e.g., Hunter, 1986; Hunter & Hunter, 1984; Ree & Earles, 1992; Schmitt, Gooding, Noe, & Kirsch, 1984; Schmidt & Hunter, 1998) and is also shown to be relatively stable for adults over their careers (Bayley, 1949, 1955; Charles, 1953; Hertzog & Schaie, 1986, 1988; Jensen, 1980; Owens, 1953; Schaie, 1994;

Thorndike, 1940). Similarly, personality has been shown to be related to performance (e.g., Barrick & Mount, 1991; Dudley, Orvis, Lebiecki, & Cortina, 2006; Hurtz & Donovan, 2000; Salgado, 1997; Tett, Jackson, & Rothstein, 1991) and also is a relatively stable individual characteristic (Costa & McCrae, 1988, 1992). Other individual characteristics related to performance change with time, such as job knowledge (e.g., Schmidt et al., 1986), job experience (e.g., Schmidt et al., 1986; Sturman, 2003), leadership (e.g., Day et al., 2004), and motivation (e.g., Kanfer, 1991, 1992). Based on any static performance model, at any point in time, job performance is at least partly determined by a function of these characteristics. This perspective incorporates aspects of the changing subjects model, as any longitudinal application of this model will show that performance over time changes because some of the causes of performance change with time. At the same time, the changing tasks model suggests that some predictors of performance change over time.

For both stable and dynamic characteristics, the functional relationship of predictors of performance can be either stable or dynamic. This leads to the following general model:

$$\mathrm{Pt} = \mathrm{B}_0 + \mathrm{B}_1 \times (\text{Stable Characteristics}) + \mathrm{B}_{2(\mathrm{t})} \times (\text{Stable characteristics}) + \mathrm{B}_3 \times (\text{Dynamic characteristicst}_{\mathrm{t}}) + \mathrm{B}_{4(\mathrm{t})} \times (\text{Dynamic characteristicst}_{\mathrm{t}}) + \mathrm{e}$$

Where B_1 and B_3 are stable coefficients over time, and B_2 and B_4 change over time (as signified above with the addition of (t) in the subscript; the bold indicates matrices of characteristics and coefficients).

The above model shows the problem associated with using cross-sectional data to consider longitudinal phenomena. Specifically, in any sort of cross-sectional analysis, one cannot observe changes within subjects across time. This means that when examining the results of any analysis, one will be unable to distinguish between the stable and dynamic betas or between stable and dynamic characteristics. As such, the coefficients derived from any model may be accurate, but may not generalize to even the same subjects at a different point in time. This condition limits the potential value of cross-sectional analyses when considering longitudinal phenomena. As will be discussed later, though, this does not mean that cross-sectional research is of no value. Nonetheless, the above model cannot be explicitly tested as shown. One may be able to employ other research to distinguish between the stable and dynamic characteristics, but it remains impossible to know the functional form of the dynamic betas.

To help understand where time plays a role in the prediction of job performance, perhaps the easiest adjustment to the above model is to consider lagged measures of job performance. By using a lagged measure of job performance, one can derive a simpler model that can help focus attention on the dynamic factors associated with job performance. That is, if one is modeling $P_{(t)}$, and subtracts $P_{(t-1)}$ from each side of the equation, one gets the following:

$$P_t - P_{(t-1)} = B_{0t} + B_1 \times S + B_{2(t)} \times S + B_3 \times D_{(t)} + B_{4(t)} \times D_{(t)} - P_{(t-1)} + e_{(t)}$$

With substitution, this becomes

$$P_t - P_{(t-1)} = B_{0t} + B_1 \times S + B_{2(t)} \times S + B_3 \times D_{(t)} + B_{4(t)} \times D_{(t)} + e_{(t)} - (B_{0(t-1)} + B_1 \times S + B_{2(t-1)} \times S + B_3 \times D_{(t-1)} + B_{4(t-1)} \times D_{(t-1)} + e_{(t-1)}$$

or

$$\Delta P = \grave{B}_0 + B_{2(t-1)} \times S + B_{2(t)} \times S + B_3 \times \left(D_t - D_{(t-1)}\right) + B_{4(t)} \times D_{(t)} - B_4 \times D_{(t-1)} + \grave{e}$$

To help with explanation, this can be simplified as follows:

$$\Delta P = B_I + B_A \times S + B_B \times \Delta D + B_C \times D_t + B_D \times D_{(t-1)} + e$$

Note that by using a lagged variable, the stable effects associated with stable characteristics are eliminated from the model. Also in this model, B_A represents evidence of the changing tasks model. If stable characteristics are related to performance after controlling for prior performance, then this can only occur if it is because the way in which the stable characteristics relate to performance change with time. Significant coefficients within B_c presents evidence of the changing subjects model, and significant coefficients of B_D presents evidence of the simultaneous effects of both changing subjects and changing tasks.

A flaw with the above model, though, is that the analysis of change scores is not always desirable (e.g., Edwards, 1994, 2001). While there are a number of issues related to difference in scores, most salient here is that modeling the change score above is equivalent to the following:

$$P_t = B_I + B_A \times S + B_B \times \Delta D \quad - B_C \times D_t - B_D \times D_{(t-1)} + (1.0) \times P_{(t-1)} + e$$

That is, it assumes that the effect associated with lagged performance is 1.0. If one is able to model performance longitudinally, there is little advantage to making this assumption. Rather, one should allow the model to estimate the effect of the lagged performance measure, as its interpretation can be quite useful. Hence, one should model

$$P_t = B_I + B_A \times S + B_B \times \Delta D \quad - B_C \times D_t - B_D \times D_{(t-1)} + B_{Lag} \times P_{(t-1)} + e$$

If the performance model being used is fully specified, then B_{Lag} should be equal to one, an assumption that can be tested empirically.

The theoretical value of such a test, though, is not likely to be large. This is because research has already shown that performance trends tend to be non-linear (Deadrick et al., 1997; Ployhart & Hakel, 1998; Sturman et al., 2005). This means that, because job performance tends to follow a negatively accelerating curve, controlling for the linear effects of prior performance will not fully capture the extent to which past performance can predict future performance.

The non-linear trends of job performance will cause the above model to be under specified in a way that could affect the interpretation of the independent variables of interest. That is, although lagged performance may be in the model, this coefficient does not partial out all of the effects associated with past performance. The potential for this problem can be seen by comparing the results of Harrison et al. (1996) and Sturman and Trevor (2001). In their studies of turnover, both papers showed that changes in performance were associated with the probability of turnover. However, Sturman and Trevor (2001) extended the work by Harrison et al. (1996) by demonstrating that even after controlling for the most recent change in turnover, long-term trends of performance also affected the probability of turnover. Hypothetically, in another study of turnover that controlled only for the most recent change in performance, it is possible that some independent variable under investigation would be correlated with these performance trends. If so, the variable may falsely appear to relate to the dependent variable because the effects of the long-term trends were not controlled for in the model.

At the present, this is only a hypothetical. With few exceptions (Harrison et al., 1996; Sturman & Trevor, 2001), there is little research looking at the consequences of performance trends. When using some sort of lagged performance model, researchers need to give careful consideration to the potential effects associated with performance trends. This is a potential concern in many areas of human resource research, but perhaps most so in the areas of compensation and training. In compensation, rewards are commonly associated with more than just the most recent performance evaluation; if estimating the effects of a training program, controlling for the trajectory of performance scores may be essential for isolating the effects associated with the training intervention. Ideally, one should control for multiple measures of prior job performance to more fully specify any performance model, thereby gaining confidence in any potential effects associated with an independent variable of interest (e.g., earning a certain bonus in time [t_1], participating in a given training program at [t-1]).

The question thus arises: how many prior performance scores should be controlled for in longitudinal analytical models? Unfortunately, the existing research on dynamic performance does not have a definitive answer. Simply put, more is better. If one can control for one instance of prior performance, the analyses will control the linear effects associated with job performance on the dependent variable, but it will not partial out the known non-linear effects on the dependent variable of interest. If one controls for two instances of prior performance, one is then controlling for the linear effects of prior performance and the effects of performance changes. Controlling for three prior measures of performance controls for the linear effects associated with performance in addition to the effects associated with change and the effects associated with which the rate of change is itself changing. Obviously, using more measures of performance is more specified. More measures also presents a more conservative test if trying to show that some other independent variable (e.g., a bonus, a training program) has an effect on performance over time.

Research examining performance trends has generally examined no more than cubic trends (i.e., the rate in which the change in performance is changing) (e.g., Hofmann et al., 1992, 1993; Keil & Cortina, 2001; Ployhart & Hakel, 1998). Although it would be desirable to test the assumption that performance slopes can be described with up to cubic parameters, controlling for three measures of prior job performance may fully capture the effects associated with performance trends. Of course, longitudinal research is often rare, and adding the extra demand that longitudinal studies of job performance have at least four waves of data would only make such research less feasible. Researchers should be aware of the potential effects of nonlinear (cubic) performance trends when considering longitudinal models of performance. However, given the current lack of research in this area, potential specification error of effects associated with prior performance is a limitation that is worth accepting until the quantity of knowledge in this area is sufficiently expanded.

Employment Stage Models

Another key theoretical development in the area of dynamic performance research is the employment stage models developed by Ackerman (1987, 1988, 1992) and Murphy (1989). Ackerman's work is focused on skill development; Murphy's model applied Ackerman's work to job performance. Both are obviously related, and both are pertinent and have been applied to the study of job performance over time.

Ackerman proposed a theory of skill acquisition, predicting that task performance becomes automatic through practice. His theory posits that the importance of certain abilities to task performance change during skill acquisition. Furthermore, the nature of this change depends on the particular ability and the nature of the task.

The theory postulates that individuals proceed through three phases of skill acquisition. The first phase (the cognitive phase) involves a strong demand on the general mental ability of the performer as performance strategies are formulated and tested. During this phase, performance speed and accuracy increase quickly, and the demands on cognitive ability are reduced. The second phase (the associative phase) involves the refinement of the stimulus–response connections developed in the first phase. Here, ‘‘perceptual speed ability’’ refers to abilities that are associated with ‘‘an individual’s facility in solving items of increasing complexity y[and] the speed of processing’’ (Ackerman, 1988, p. 290). Ultimately, the third and final stage of performance is characterized as automatic. In this phase, tasks can be completed competently even without the full attention of the performer. Performance in this phase is less dependent on perceptual speed ability and more so on psychomotor ability (defined as processing speed and accuracy independent of information processing per se; Ackerman, 1988).

Although described as distinct, Ackerman postulates that individuals proceed through the three phases in a continuous manner. The effects of the various abilities on task performance change continuously with practice. The effect of general mental ability begins high and decreases; the effect of perceptual speed ability begins low, increases to a peak in Phase two, and ultimately decreases again; the effect of psychomotor ability begins low and increases with practice.

Ackerman also predicts that progression through stages is affected by the complexity and consistency of the task. Complexity refers to the cognitive demands of the task, including memory load, amount of information to process, number of responses, amount of information to be learned, and the amount of stimulus–response compatibility. Greater complexity changes the importance of the various abilities on performance. For example, in tasks with a weak compatibility between stimulus and response, the task will place higher cognitive demands on the learner to determine and execute the appropriate response. This places a greater emphasis on perceptual speed ability and delays the emphasis on general mental ability until further into the skill acquisition process (Ackerman, 1988).

Task consistency effects the rate in which tasks can be mastered. Initially, it has no effect on skill acquisition, as the task being learned is novel to all performers. However, the inconsistency in the task slows the rate in which practice allows for skill acquisition. As a result, inconsistent tasks require performers to depend on cognitive processing (i.e., Phase one) for longer periods of time.

A limitation of Ackerman’s work is that the theory is focused on task performance. While task performance is obviously related to job performance, the contexts and constructs are distinct. For example, Ackerman's work generally examined tasks where skills acquisition can be completed in fewer than 20 hours of training. For job performance, even the simplest of jobs generally involves performance of multiple tasks (Borman, 1991). Second, while tests of task performance employ relatively simple measures of performance, the construct of job performance is recognized to be far more difficult to measure, complex, and multidimensional (e.g., Motowidlo et al., 1997; Murphy & Schiarella, 1997; Rotundo & Sackett, 2002; Viswesvaran, Ones, & Schmidt, 2005; Welbourne, Johnson, & Erez, 1998). Third, tests of Ackerman's model generally frame time in terms of minutes or hours (e.g., Ackerman, 1988, 1992; Kanfer & Ackerman, 1989) whereas studies of job performance over time generally describe performance in terms of months or years (e.g., Deadrick et al., 1997; Ployhart & Hakel, 1998; Sturman, 2003; Sturman et al., 2005; Sturman & Trevor, 2001).

Recognizing the substantive differences between task performance and job performance, Murphy (1989) applied Ackerman's theory to the job performance context. Murphy's application is in many ways similar to Ackerman's work, but with some notable differences. First, because of the different nature of tasks (as examined by Ackerman) and the elements comprising job performance, Murphy does not distinguish between complexity and consistency. Rather, complexity ''is used as a gross index of a job's cognitive demands'' (Murphy, 1989, p. 195).

Murphy's application also results in a model with two phases instead of three: the transition and maintenance stages. The transition stage occurs when an employee is new to a job or when the major duties associated with a job change. Similar to the first phase of Ackerman's model, this phase places high cognitive demands on workers who must acquire new information and cannot rely on past experience. The maintenance stage occurs when jobs are well-learned. Murphy predicts that in this stage, cognitive ability is less important and personality and motivational factors play a more important role in the determination of job performance (note that Ackerman's model does not consider personality or motivational factors). Although Murphy recognizes that ''predicting the length of transition stages may be particularly difficult'' (p. 191), his model predicts that there are ''distinct stages that characterize a worker's performance on the job'' (p. 192).

Subsequent relevant empirical tests using job performance as a dependent variable have not distinguished between Ackerman's and Murphy's models. For example, papers by both Farrell and McDaniel (2001) and Keil and Cortina (2001) present tests of Ackerman's model but examine job performance ratings over time frames that are consistent with research on jobs but not the earlier work on task performance. Specifically, Keil and Cortina (2001) considered both short-term (o1 day) and long-term outcome measures; Farrell and McDaniel (2001) exclusively employed supervisor ratings of job performance.

Although these studies purport to investigate Ackerman's model, they actually provide useful evidence for comparing and contrasting the predictions of Murphy with Ackerman. The results from Keil and Cortina (2001) actually provide support of Murphy's model as it is differentiated from Ackerman's. For example, Keil and Cortina (having examined cognitive ability, perpetual speed ability, and psychomotor ability) concluded that ''the most pervasive finding of the present study was that validities deteriorate over time'' (p. 687). While Ackerman had predicted different functional forms for the relationships between these abilities and task performance over time, Keil and Cortina found that the relationship between these abilities and performance began to deteriorate in the early stages of task performance for both consistent and inconsistent tasks. Note that while this is inconsistent with Ackerman's predictions, they are perfectly consistent with Murphy's prediction that abilities (in general, and including all three studied here) would initially predict performance and then decrease in validity in the transition stage.

Keil and Cortina also support what they labeled a ''Eureka effect'' which is increases in experience tend to produce insights that lead to sudden jumps in performance. Recall that Ackerman predicted continuous development through stages, whereas Murphy called for ''distinct'' stages. This finding is also consistent with Murphy but contrary to Ackerman's task-based model.

On the other hand, Farrell and McDaniel's study is in some ways more consistent with the predictions of Ackerman's model, and in other ways contradictory to both models. The effects of general mental ability, perceptual speed ability, and psychomotor speed were of different functional forms. There were also notable differences in these relationships for consistent and inconsistent jobs. However, contrary to both models, the effect of abilities on performance for consistent jobs appeared to increase with experience. It should be noted, though, that Farrell and McDaniel's study suffer from the limitations of cross-sectional research of longitudinal phenomena. That is, the authors did show a significant (negative) correlation between their temporal variable (experience) and all of their ability measures. This suggests that the cohorts in their sample were not equivalent. This could mean that hiring standards have increased over time (and therefore newer people have entered the population who have higher levels of the abilities), that higher performers (or at least those with higher levels of abilities) were more likely to leave the organization, that abilities under consideration were not static as has been assumed, or some combination of these explanations. It is simply not possible to determine the cause of these correlations or to know the consequences of them in a cross-sectional study. Nonetheless, their findings do have value when considering the validity of stage models.

In all, both Ackerman's and Murphy's models are valuable in that they provide a theory for understanding why performance changes with time, and how the relationship between predictors and performance should change over this time. Although it has not received much direct explicit attention, Murphy's approach to differentiating task performance from individual job performance is a critical theoretical advance for understanding individual job performance. The contradictory predictions of Ackerman and Murphy, and the contrary findings of Keil and Cortina (2001) and Farrell and McDaniel (2001) demonstrate that more work in this area is needed. Even if not building upon Murphy's contribution, other work considering Ackerman's model of task performance needs to be specifically adapted to understanding the construct of job performance. Murphy's work is a good demonstration of this, but the model is far from resolved.

Learning Curve Theory

A potentially fruitful theoretical perspective generally unexplored in the dynamic performance literature for modeling job performance over time comes from Learning Curve Theory. While the dynamic performance literature has described performance trends as following a learning curve (e.g., Farrell & McDaniel, 2001; Ployhart & Hakel, 1998), little use has been made of Learning Curve Theory to make specific predictions. Despite its origins in psychological research (e.g., Kjerstad, 1919; Thurstone, 1919), Learning Curve Theory has received scant research attention focused on individual employees. Instead, Learning Curve Theory remains a staple of operations research, and has been dominated by a macro-organizational perspective, describing the collective efforts of many employees (Hirschmann, 1964).

Learning Curve Theory predicts that organizational productivity improves based on the accumulation of experience. The learning curve phenomenon is the graphical representation of the learning-by-doing phenomenon observed in people performing manual tasks. The theory, supported by empirical evidence, suggests that with the repetition of a task, the performance of that task improves predictably. The advances to this domain of literature have been in the analytical methods used to represent the functional form of the relationship, methods to estimate the required parameters, and applications of these methods in industrial settings.

The value of Learning Curve Theory for the dynamic performance literature is that it posits a specific functional form that performance should follow. Essentially, the literature related to Learning Curve theory describes the nature of aggregated employee performance over time. Although based on considering performance at an aggregated level, my purpose of reviewing Learning Curve Theory is to demonstrate its value of (re)applying it to the individual-level of analysis.

At its core, Learning Curve Theory stipulates that people learn by doing (Teplitz, 1991). It was originally developed in the psychological literature to describe the rate in which individuals learn how to perform a repetitive, manual task (Kjerstad, 1919; Thurstone, 1919), but it was soon applied to the task of predicting production rates and production costs in manufacturing settings (e.g., Wright, 1936). The theory posits, with repeated examples of empirical support, that as experience with a task increases, the resources required to complete the task (usually labor hours or price per unit) decreases. More specifically, it is held that as the quantity of production doubles, the resources needed to complete the production will be reduced by a constant percentage (called the learning rate; Yelle, 1979). Additionally, this learning rate is expected to be consistent for every doubling of production, a phenomenon known as the ‘‘doubling effect.’’ Much of the research and use of Learning Curve Theory has been involved in the estimation of this learning rate, or the development of methods to estimate the learning rate more accurately.

Because of the strong emphasis on the mathematics of the theory’s premise, methodological representations of the learning curve have played a very important role in the Learning Curve Theory literature. The most common representation of the learning curve is the log-linear model, also called the Wright model, represented as follows:

$$P_x = I \times X^N$$

P_x=measure of performance (usually either the number of labor hours required to produce the Xth unit or the cost to produce the Xth unit, see Yelle, 1979).

- I=The number of units required to produce the first unit (e.g., labor hours), or the expected initial performance level.
- X=The cumulative unit number.
- N=The learning index; log Φ /log 2.
- Φ=The learning rate.
- 1– Φ=The progress ratio.

Other learning curve models include the plateau model (Guibert, 1945), the Sigmoid S Curve (Carr, 1946), the Prior-Learning Model (Stanford Research Institute, 1949), the asymptotic model (DeJong, 1957), the adaptation function (Levy, 1965), the exponential function (Pegels, 1969); time-rate models (Bemis, 1981), and cost-rate models (Smith, 1981). It is not my purpose here to provide a comprehensive review of all modifications of the learning curve model; this is a task best left for texts specifically on this topic (see, Belkaoui, 1986 and Teplitz, 1991). Nonetheless, it is important to at least bring up that a variety of mathematical functions exist, as the use of one of these specific models may be more appropriate when modeling individual job performance scores.

Even with the variety of possible approaches to modeling learning curves that have been created, the Wright model is the basis for all other developments, and remains the most popular approach. While abundant research has worked at improving, extending, and applying Learning Curve Theory and learning curve models, it is unquestionable that the basic premise of the theory has received extensive support, and the application of the theory has proven extremely valuable to many different industries (Belkaoui, 1986; Muth, 1986; Teplitz, 1991; Yelle, 1979).

The work on Learning Curve Theory from the operations literature has shown that organizational productivity tends to follow a specific functional form over time. As reviewed earlier, the dynamic performance literature similarly has suggested that there exist systematic and predictable relationships between individual-level job performance measures and time (e.g., Day et al., 2004; Deadrick et al., 1997; Ployhart & Hakel, 1998; Stewart & Nandkeolyar, 2006; Thoresen et al., 2004), but has not sought to propose a basic theoretical form for this relationship. My purpose of reviewing Learning Curve Theory in this article is that it may prove useful for the dynamic performance literature. Although individual-level performance curves of interest to dynamic performance researchers and the aggregated results-based performance curves from the operations literature are not the same, if the two are related such that evaluative performance is an unbiased indicator of results-based performance, then aggregation of evaluative performance measures should follow the same functional form as the aggregation of results-based performance. Common to such macro perspectives, the key assumption behind a model of aggregated job performance is that there are substantial consistencies in the behavior of individuals, hence making it is possible to focus on aggregate responses and ignore variation across individuals (Klein & Kozlowski, 2000; Kozlowski & Klein, 2000). As Learning Curve Theory describes the aggregated productivity of an organization, the applicability of the theory to evaluative performance measures would be confirmed when evaluative measures of job performance are well-modeled by Learning Curve Theory functions.

While so far I have argued that Learning Curve Theory, in principle, may be valuable for understanding individual evaluative measures of job performance over time, specific adaptation of the theory is necessary before it is possible to confirm or falsify this proposition. This application will require three issues to adapt Learning Curve Theory (back) to understanding individual job performance.

First, the correct functional form needs to be identified. The Wright model (Eq. (8)) does not specifically lend itself to modeling job performance ratings. The Wright model is most applicable for modeling costs, with the functional form showing a decrease in costs with the accumulation of experience. With performance ratings, we expect job performance to increase with the accumulation of experience (Sturman, 2003), and other research has suggested a quadratic form to performance trends (e.g., Hofmann et al., 1992, 1993; Keil & Cortina, 2001; Ployhart & Hakel, 1998). The literature on performance trends and Learning Curve Theory can be combined to yield an appropriate functional form for modeling job performance over time.

Second, another assumption of Learning Curve Theory is that it is used to model repetitive tasks, and its applications have been predominantly in the manufacturing sector. Although this is not a flaw in the theory, it does limit its potential generalizability to a wider array of jobs and the modeling of individual performance scores. Job complexity has been shown to moderate relationships with job performance, particularly with regard to temporal issues (Farrell & McDaniel, 2001; Sturman, 2003; Sturman et al., 2005). Because job complexity affects the relationship between time and performance (Sturman, 2003) and the stability of job performance ratings over time (Sturman et al., 2005), it will affect the functional form of performance over time. The resultant functional form of learning to represent job performance over time should be explicitly capable of incorporating the effects associated with job complexity.

Third, a fundamental problem with the premise of the Learning Curve Theory, particularly in light of the SHRM literature, is that the theory seems to imply that managers can simply sit back and await guaranteed productivity gains (Teplitz, 1991). This is inconsistent with the fundamental premise of the SHRM literature, which is based on the idea that HR programs can influence organizational performance through effects on individuals within the organization (Wright & Boswell, 2002). It is also inconsistent with the notable different ''learning rates'' found in many applications of Learning Curve Theory (Adler & Clark, 1991). Therefore, just as Learning Curve Theory can be of value to the dynamic performance literature, the theory itself can incorporate the potential effects of human resource interventions (e.g., Adler & Clark, 1991) which can then affect organizational performance. In other words, the methods behind the functional form must be capable of representing organizational-level effects on the individual-level phenomenon (i.e., job performance) being modeled.

In sum, Learning Curve Theory has potential applicability to modeling performance over time. The success of the theory in the operations literature suggests it has merit, but its applicability to helping understand behaviorally based performance measures needs explicit testing. Before such testing can occur, though, the functional forms must be adapted to this sort of measure, identifying the necessary parameters, and developing the model to be tested.

The primary benefit of applying Learning Curve Theory to the dynamic performance literature is that it provides a theory and related methodology for explicitly modeling the shape of individual job performance over time. Specific functional forms can be hypothesized and tested. With a theoretical basis for predicting the shape of individual job performance, research could then focus more attention on the factors influencing this shape. Currently, research on dynamic performance is ad hoc in terms of specifying its functional form. Some studies consider linear effects, others quadratic, and others cubic. Little rationale is given for any form, and there is certainly no consensus regarding which functional form is most appropriate. Learning Curve Theory provides the opportunity to develop theory regarding the nature of this functional form.

The Present: Analytical Techniques and Methodological Issues

Dynamic performance research has employed a wide variety of methodologies. These have included the use of cross-sectional methodologies with The Past, Present, and Future of Dynamic Performance Research 75 temporal variables (e.g., Farrell & McDaniel, 2001; Sturman, 2003), the examination of correlations between performance measures for various time lags (e.g., Deadrick & Madigan, 1990; Keil & Cortina, 2001; Sturman et al., 2005), and the application of new advances in research methods – latent growth curve modeling (LGCM) and hierarchical linear modeling – to analyze performance trends (Deadrick et al., 1997; Hofmann et al., 1992, 1993; Ployhart & Hakel, 1998; Sturman & Trevor, 2001). While research methodologies play an important role in any field, there are particular conditions in the study of job performance and time that make it crucial to discuss certain methodological issues and problems. These include research design, analytical techniques, and the sort of statistical problems that necessarily accompany research in this area.

Research Design

A critical issue for research on performance that must be addressed early in the research process is the development of a study's methodology. Most basically, the study must employ either a cross-sectional or longitudinal design. While calls for longitudinal research are the norm, it is often difficult to get extensive longitudinal data. What's more, studies involving time do not necessarily need longitudinal data to test their hypotheses. Although it may at first seem that any sort of hypothesis involving time must be studied with longitudinal data, one must first consider the nature of what is being studied to best determine the study design.

The nature of the research design dictates the sort of research question that can be tested. Cross-sectional research provides no opportunity for examining within-person changes (Hulin et al., 1990), whereas longitudinal studies provide the potential to consider both within-person and across-person effects. Below I will discuss issues related to both research designs.

Cross-Sectional Designs

Cross-sectional data are still useful when considering performance over time if one is interested in modeling across-person relationships. Essentially, if one is examining how the relationship between some predictor (X) and some rating of job performance (Y) changes with time, then cross-sectional data can yield useful hypothesis tests.

This design is useful for applications (explicitly or implicitly) of the changing tasks model. In such models, cross-sectional data can be used to consider if the relationship between X and Y changes by testing if time moderates the relationship between X and Y. Fortunately, time-related variables are frequently available in organizational research. Specifically, age, organizational tenure, and job experience are often available and used as control variables, but seldom are the complexities of the dynamic performance literature drawn upon to appropriately integrate such temporal variables in meaningful ways (Sturman, 2003). One way to consider potential dynamic effects is by looking for interactions associated with temporal variables. For example, if a given characteristic, say cognitive ability, is interacted with job experience, and this interaction is shown to be associated with job performance, a logical conclusion is that the effect of cognitive ability on job performance changes with time. Similarly, dividing employees into cohorts by experience (e.g., Farrell & McDaniel, 2001; Schmidt et al., 1986) allows one to consider longitudinal issues with cross-sectional data. For example, if the coefficients associated with an independent variable varies systematically across cohorts (like for cognitive ability as shown in Farrell & McDaniel, 2001), this evidence would support the conclusion that the effect of cognitive ability changes with the passage of time. Given we have prior reason to believe that cognitive ability is a stable characteristic, the presence of the type of significant effects described above would suggest support for the changing tasks model. In terms of the static and dynamic model reviewed earlier, the effect of cognitive ability (B) is contained within $B_{2(t)}$.

Certainly, cross-sectional analysis has its limitations when considering job performance over time. The disadvantage of this approach is that conclusions must be based on two critical assumptions. First, one must assume that the mean level of the characteristic does not vary with time. For the example above, cognitive ability should not be lower for individuals with less experience than for those with more experience. Unfortunately, in the study of job performance, it is likely that this assumption will be violated. One reason this may occur is because of the demonstrated relationship between job performance and turnover (Trevor, Gerhart, & Boudreau, 1997; Salamin & Hom, 2005; Williams & Livingstone, 1994). If the characteristics under investigation are associated with performance, and if performance is related to turnover, then the distribution of the characteristic will likely change over time. Because of the negative relationship between performance and turnover (Williams & Livingstone, 1994), those with low levels of the characteristic would be expected to leave the company over time, and a cross-sectional representation would have a wider range of the characteristic for newer employees and a restricted distribution of the characteristics for those with more tenure. If high performers are also leaving (Trevor et al., 1997; Salamin & Hom, 2005), this would further restrict the distribution. Methodologically, this creates heteroskedasticity. More generally, it limits any potential causal implications that can be made when interpreting the interaction or differences across cohorts. Of course, this assumption can be tested in any given sample, such as by examining the distribution of the characteristic at various experience levels (e.g., using a median split). If not statistically significantly different, one may have some confidence that this first assumption holds.

Even if the independent variable included in the model does not specifically cause turnover, a second reason that the mean levels of a predictor may change with time is that the distribution of an independent variable may become restricted because of selection processes. The Attraction-Selection-Attrition (ASA) model (Schneider, 1987) suggests that through the process of recruitment, selection, and turnover, individuals self-select and are selected in ways that increase the similarity of individuals within the organization. This filtering process may create the appearance of a relationship between these characteristics and performance. For example, even if a given characteristic, say a dimension of personality, is not related to performance, the simultaneous effects of learning for individuals who remain in the company in conjunction with the increase in homogeneity of individuals within the company over time could cause an interaction of the characteristic and time to appear significant. Again, this concern may be tested by examining the distribution of the characteristic in question for various levels of the temporal variable.

The second key assumption when using cross-sectional data to consider effects associated with time is that characteristics of the hiring process (or internal selection or turnover processes) remained stable over time. If this is not true, specification error may cause variables to appear to have an interaction with time. For example, if the hiring process has changed, resulting in employees with a different distribution of characteristics than those hired earlier, other correlated characteristics may falsely appear to be significant if the model is not appropriately specified. Researchers can try to minimize this risk by having well-developed theoretical models which identify the variables to be measured and by ruling out alternative explanations for their findings. Nonetheless, this is a concern that in all likelihood cannot be fully mitigated.

Despite these limitations, though, cross-sectional research can play an important role for considering dynamic phenomena. Any longitudinal model of job performance that can be used to make predictions about job performance over time should also have implications if a ''snap-shot'' of the employees is taken during any point of their development. Cross-sectional research, when using interactions with temporal variables or samples divided into cohorts by some temporal variable, can be of use when considering the implications of such models. While clearly not ideal for studying longitudinal phenomena and its limitations must be explicitly recognized, the information provided in cross-sectional research should not be dismissed or ignored. Furthermore, any study that examines job experience, job tenure, organizational tenure, or age is already implicitly involving the modeling of effects that are associated with time. Research on performance, even if intended to be static and using cross-sectional data, needs at least to recognize where the literature on dynamic performance may be relevant to their proposed models, and perhaps may even necessitate that variables of interest be interacted with temporal variables to provide at least some test of the stability of the beta-coefficient of interest.

Longitudinal Designs

For studying job performance over time, longitudinal designs have obvious benefits. With longitudinal data, one can look at both within-person and across-person differences. This allows the examination of how performance changes, how individual characteristics change, and how the effects of individual characteristics on performance change. With these benefits, though, come both practical and methodological problems.

The most practical problems are associated with data collection, but it is more than a convenience problem. Difficulties of longitudinal research in general are compounded by the specific needs of research on dynamic performance. First, the collection of data from multiple time periods already can be practically difficult, but for modeling within-person relationships this difficulty is compounded because more than two waves of data are highly preferable. As discussed earlier, we know job performance trends are nonlinear, and that controlling for simply the last instance of job performance will yield an underspecified model. Given the evidence (so far) that job performance actually follows (at least) a cubic form (e.g., Hofmann et al., 1992, 1993; Keil & Cortina, 2001; Ployhart & Hakel, 1998), this means that a highly specified model requires data from at least four waves, and more would be preferable to test this assumption (although applications of Learning Curve Theory could provide a model with a different specification).

Second, dynamic performance research has frequently used sales as a measure of performance. This is convenient because the data is often available on a monthly basis, thereby making the data demands of dynamic performance research more easily satisfied; however, this is counter to most theoretical work on job performance that defines individual job performance as behaviors. Most typically, behavioral measures of job performance are based on ratings by a supervisor. Sometimes specific research tools are created for that purpose, but other times researchers rely on the supervisory ratings employed by an organization. With such ratings typically performed annually, the desired longitudinal study would require at least four years of data. The practical difficulties of soliciting this sort of data from an organization only increases the difficulty of performing the sort of research on job performance over time that the existing literature suggests is preferable.

Third, as has been discussed, research on dynamic performance has moved beyond the simple question of ‘‘is performance dynamic?’’ Research is needed into the functional form of performance trends, and more importantly, on the causes and consequences of these trends. Not only does this necessitate multiple years of supervisory evaluations, but other variables of interest must also be available. While companies may possess records of employee performance, the data required to advance the literature on performance over time make studies involving attitudinal data very difficult.

Fourth, longitudinal research designs require the use of methodologies that are more complex than most cross-sectional analyses. These issues will be discussed below, but may require complex treatments of error terms, methods more complex than OLS regression, methods for handling missing data, corrections for range restriction, and more.

The nature of any research involving time makes it highly desirable to use longitudinal data. Unfortunately, it is far easier to call for such research than it is to perform. So, while longitudinal designs are preferable for studying dynamic performance, they are no panacea.

Analytical Tools

The research needs of the dynamic performance literature combined with the practical difficulties of collecting and analyzing the requisite data presents a daunting research problem. Many of these negatives, though, are counterbalanced by new methodological developments that provide exciting opportunities for research in this area. New (or relatively new) techniques that allow modeling within-person relationships are providing valuable opportunities for analyzing the trends of individual scores over time and the correlates of these trends. Adding some confusion to this area is the fact that there are even more names to represent these techniques.

The techniques to which I am referring have been called covariance components models, hierarchical linear models, hierarchical models, latent curve analysis, latent growth models, mixed models, mixed linear models, multilevel models, multilevel linear models, random effects models, and random coefficient models (Raudenbush, 2001). In actuality, these many names essentially represent two approaches, which I will review below. Furthermore, although there are different approaches to estimate these models, fortunately there are common characteristics that make a general framework applicable to understanding the modeling of individual performance trends (Raudenbush, 2001).

For all these approaches, job performance over time is now typically thought of as a multilevel problem, with individuals' performance scores over time nested within individuals. Typically, the within-person scores are referred to as the first-level (i.e., Level-1), whereas individual-characteristics are at the second-level (i.e., Level-2). Conceivable, more levels are possible, including dyads, teams, departments, organizations, industries, and nations (cf., Ployhart, 2004); however, research on dynamic performance has yet to expand to such domains.

A simple Level-1 model capturing the linear effects of time would appear as follows:

$$\text{Job Performance}_{\text{it}} = \text{B}_{0\text{i}} + \text{B}_{\text{li}} \times \text{time}_{\text{it}} + \text{e}_{\text{it}}$$

A more complex model, capturing demonstrated cubic effects would be

$$\text{Job Performance}_{\text{it}} = \text{B}_{0\text{i}} + \text{B}_{\text{li}} \times \text{time}_{\text{it}} + \text{B}_{2\text{i}} \times {\text{time}_{\text{it}}}^2 + \text{B}_{3\text{i}} \times {\text{time}_{\text{it}}}^3 + \text{e}_{\text{it}}$$

Each beta coefficient is computed for each individual. If the betas are modeled with error, it is a random effects model; otherwise, it is a fixed effects model.

Time may be entered simply as the raw variable (e.g., months, years), or may be transformed. As one employs a higher order model, centering becomes a greater issue (Hofmann & Gavin, 1998). When using polynomials, and specifically wanting to isolate linear from quadratic from cubic effects, orthogonal polynomials for time can be used because such a process allows one to decompose the various time elements (linear, quadratic, and cubic) and avoid potential multicollinearity (Ployhart & Hakel, 1998; Willett & Sayer, 1994).

There are different ways to estimate this sort of multilevel model. A criticism of research employing these methods, though, is that software choice often influences, or even precedes the analytical strategy. One common approach is hierarchical linear modeling (HLM), often associated with the package ''HLM'' or ''mixed models'' from SAS Proc Mixed. The other most common approach, often referred to as a latent growth curve model, or LGCM, is performed using structural equations modeling software such as AMOS, EQS, or LISREL (Raudenbush, 2001). In actuality, both approaches are special cases of the General Linear Mixed Model (GLMM) (Rovine & Molenaar, 2001), although they have divergent structures and assumptions which make them different in application. I will briefly review the principles behind the GLMM, highlighting the difficulties with this approach and why the other approaches ultimately become more practical alternatives. I will then discuss both HLM and LGCM in turn, including their relative advantages and disadvantages.

The General Linear Mixed Model

The GLMM (Laird & Ware, 1982) is an expanded case of the general linear model. The model is as follows:

$$y_i = \beta X_i + Z_{iy_i} + \varepsilon_i$$

where y_i is a 1 x n_i vector of outcomes (i.e., job performance) for individual i, X_i is a b x n_i matrix of fixed effects, Z_i is a g x n_i matrix for the random effects, g_i is a 1 x g vector of random effects, and b is a 1 x b vector of fixed effects parameters. Residuals between any two individuals are assumed to be uncorrelated, but residuals within an individual have a particular covariance structure. The fixed effects component (b) are constant across individuals; the random effects component (Z_i) are different across individuals, hence the indexing subscript i. The random effects (g_i) are assumed to be distributed independently across individuals, with the following distribution:

$$y_i \sim N(0, \sigma^2 D)$$

where D is an arbitrary ''between subjects'' covariance matrix (Rovine & Molenaar, 2001)

The within-subjects errors (e_i) have the distribution

$$\varepsilon_i \sim N(0, \sigma_\varepsilon{}^2 W_i$$

Note that the general static model of job performance discussed earlier is actually a form of the GLMM. Examining Eq. (1), the stable coefficients (B_1 and B_3) are both parameters within b from Eq. (11). Similarly, $B_{2(t)}$ and $B_{4(t)}$ are parameters within Z_i. The individual characteristics (both stable and dynamic) are individual observations, taken at each point in time, within X_i and c_i. Thus, one can immediately see that there is a high degree of potential synergy between the analytical methods emerging from the GLMM and the study of job performance over time.

The problem with such a general model is that it is statistically impossible to estimate in its most general form (Laird & Ware, 1982; Rovine & Molenaar, 2001). Constraints must be placed on the model. HLM and structural equations approaches (i.e., LGCM) use different constraints, giving each advantages and disadvantages.

Hierarchical Linear Modeling

HLM is a methodological technique designed to analyze multilevel data, where data is nested hierarchically in groups. For the study of performance over time, this methodology has immediate relevance, as individual job performance ratings are gathered on a set of people, and the repeated measures contain information about each individual's performance trends (Hofmann, 1997; Raudenbush & Bryk, 2002). The approach recognizes that an individual's performance scores (i.e., the within-individual data) may be more similar to each other than data from other individuals, which is in contrast to an OLS approach where within-and across-individual residuals are not estimated separately (Hofmann, 1997).

As is typically employed, HLM approaches model employee performance by using two levels of analysis. The within-person analysis, labeled Level-1, is modeled as specified above in Eq. (9), or with greater complexity as in Eq. (10). In HLM, the second-level of analysis is used to model the parameters from the first level. The Level-2 model (for Eq. (9)) may appear as follows:

$$\mathrm{B}_{0\mathrm{i}} = \mathrm{y}_{00} + \mathrm{y}_{01}\mathrm{X}_1 + \mathrm{y}_{02}\mathrm{X}_2 + \cdots + \mathrm{U}_{0\mathrm{i}}$$

$$\mathrm{B}_{1\mathrm{i}} = \mathrm{y}_{10} + \mathrm{y}_{11}\mathrm{X}_1 + \mathrm{y}_{12}\mathrm{X}_2 + \cdots + \mathrm{U}_{1\mathrm{i}}$$

Here, each individual's intercept (B_{0i}) is modeled as a function of an overall average (y_{00}), some covariates (X), and across-person error (U_{0i}). Simultaneously, the individual's performance slope (B_{1i}) is estimated by an intercept (g_{10}), some covariates (not necessarily the same ones as in Eq. (14)), and error ($\mathrm{U}_{1\mathrm{i}}$). If estimating a two-level model, the effects of the Level-2 covariates are fixed effects (i.e., they are estimated without estimating an error term specifically for those coefficients). However, the model can be expanded to possess more levels, where each parameter is estimated as a function of higher-order characteristics (e.g., job characteristics if studying employees in multiple jobs, organizational characteristic is studying employees from multiple organizations, and so on). Applications of HLM, though, have rarely gone beyond the second-level of analysis, and popular software like the ''HLM'' package only allow up to three-level models.

The primary advantages for using an HLM approach for modeling individual performance trends are that it does not require equal observations per person, it does not require that the observations be spaced in the same way across subjects (Raudenbush, 2001; Raudenbush & Bryk, 2002), and there are well-articulated series of tests to determine the adequacy of using the full model (called the slopes-as-outcomes models, shown above as Eqs. (14) and (15)) (Deadrick et al., 1997; Hofmann, 1997; Raudenbush, 2001; Raudenbush & Bryk, 2002). The primary disadvantage of HLM is that it places a number of restrictions on how the data is modeled (Raudenbush, 2001). Specifically, the nature of the error structures is fixed, making it impossible to model alternative error structures such as autocorrelation. HLM approaches have nonetheless proven to be a useful means for helping understand the nature of job performance over time (e.g., Deadrick et al., 1997; Hofmann et al., 1993; Sturman & Trevor, 2001).

Latent Growth Curve Modeling

LGCM provides another means for modeling development, represented as different latent factors which capture the growth function. The LGCM approach involves testing the effect of latent constructs representing the growth parameters that define the shape of the performance function over time. At its simplest, the model includes an intercept and a linear construct. With sufficient information, the model can be expanded to include quadratic, cubic, or higher order functions if desired.

Like HLM, LGCM requires restrictions of the GLMM. Whereas HLM required restrictions with regard to how error and covariates are modeled, LGCM has restrictions on the structure of the data to be analyzed. When using a structural equations approach (SEM) to modeling, a key constraint is that the covariance matrices of within subjects errors must be the same for each individual (Raudenbush, 2001). Expressed mathematically (and referring back to Eq. (13)),W_i=W. Note that this is not a limitation of SEM software, but is an inherent characteristic to the method (Raudenbush, 2001). It is this constraint that most differentiates LGCM from HLM. That is, this assumption requires that when performing LGCM there are the same number of observations per individual. It also requires that all the observations be spaced at the same temporal intervals. If the number of observations or the spacing between observations differ, such models cannot be estimated with LGCM (Raudenbush, 2001). (Note, however, that I will discuss below missing data techniques that may ultimately allow LGCM to use unbalanced data.)

While the requirement of equal within-subjects error covariance matrices limits the nature of the type of data that can be analyzed with LGCM, the approach allows for far greater sophistication with regard to modeling the error structure. The relationship between the errors associated with the separate observations of job performance can be modeled in a variety of ways.

The simplest approach is to assume that the residuals are independent. The residuals would then have the following pattern:

$$\sigma^2 \begin{bmatrix} 1 & & & \\ 0 & 1 & & \\ 0 & 0 & 1 & \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Note that this is the error covariance structure assumed by HLM (i.e., the variance is computed from the sum squared residuals from Level-1 of the analysis). The problem with this pattern, and the advantage of LGCM, is that when modeling longitudinal data, the residuals may not be independent. A common approach to modeling longitudinal data is to assume that the residuals are correlated, such that $\varepsilon_{it} = p \times \varepsilon_{i(t-1)} + v_i$, where r is the autocorrelation coefficient and $v_i \sim N(0, \sigma^2)$. This is a first-order autoregressive model (AR[1]), yielding the following structure:

$$\sigma^2 \begin{bmatrix} 1 & & & \\ p & 1 & & \\ p^2 & p & 1 & \\ p^3 & p^2 & p & 1 \end{bmatrix}$$

An even more flexible option would be a general autoregressive pattern:

$$\begin{bmatrix} \sigma_1 & & & \\ \sigma_2 & \sigma_1 & & \\ \sigma_3 & \sigma_2 & \sigma_1 & \\ \sigma_4 & \sigma_3 & \sigma_2 & \sigma_1 \end{bmatrix}$$

The advantage of this approach is that it allows non-linear error variances, which may be most appropriate for modeling job performance over time. The disadvantage is that such an approach requires the estimation of additional parameters, potentially decreasing the chance of being able to estimate the desired model. The flexibility of the LGCM approach is quite appealing, but at the cost of greater information demands which may not be feasible and the requirement of an equal data structure which may not reflect the realities of collecting job performance data.

Contrasting HLM and LGCM

As discussed earlier, HLM and LGCM are special cases of the GLMM. In many ways, the approaches are very similar. Looking at the HLM equations, the coefficients at Level-1 can be seen as latent variables: unobservable parameters that are approximated with error at Level-2 of the model (Raudenbush, 2001). For LGCM, these models are multilevel (or hierarchical) because they describe data that varies at two levels: within and across persons; they are random coefficients model because each within-person observation is modeled with error, and the latent growth variables (i.e., the Level-2 across-person parameters) are also modeled with error (Raudenbush, 2001). In fact, for certain instances, HLM and LGCM are equal. If the longitudinal data has the same number of observations per person, if all observations are spaced with the same temporal intervals, and if the individual-level residuals are assumed to be uncorrelated, then HLM and LGCM will yield the exact same parameter estimates.

Some have also suggested the utility of second-order latent growth models (Sayer & Cumsille, 2001). While the specifics of this analysis are best described in detail elsewhere (e.g., Duncan, Duncan, Strycker, Li, & Alpert, 1999; Sayer & Cumsille, 2001; Williams, Edwards, & Vandenberg, 2003), this approach allows one to distinguish between the variance attributable to potentially different factors (Sayer & Cumsille, 2001) like group membership (e.g., if different jobs cause different performance slopes). This approach is again similar to (or potentially equal to) an HLM model, but a three-level model where the first level represents within-person performance scores, the second level represents individual-level characteristics, and the third level captures representation within groups (e.g., jobs, and perhaps characteristics associated with specific jobs).

Power Issues for HLM and LCGM

A concern for all empirical research methods is that of power, and some research has begun to pay attention to power issues with regard to longitudinal modeling. One such work examining the power of LCGM to between two variable, showed its power to be quite low (Hertzog, Lindenberger, Ghisletta, & von Oertzen, 2006). Most salient for research on job performance over time, a simulation study revealed that power did not exceed .80 for a sample size of 200 until reliability was nearly perfect (e.g., W0.96) for designs of 6 or fewer occasions (Hertzog et al., 2006). While studies of performance results, like monthly sales over a multiyear period may appear well suited for LGCMs, realities of data collection may make them less appropriate when using annual ratings of job performance behaviors.

Power for HLM is somewhat harder to determine because of the likely potential of unbalanced data. If the data is balanced, then the power would be the same as in the equivalent LGCM, suggesting that HLM too suffers from the power concerns discussed by Hertzog et al. (2006). Other research also suggests that HLM has power concerns. Zhang and Willson (2006) showed that HLM needs large sample sizes to have adequate power – upwards of 35 observations at Level-1, a size unlikely to be reached in longitudinal studies of job performance. They also showed that HLM models are more sensitive to changes in the Level-2 coefficients than SEM approaches.

These empirical investigations into the power of HLM and LGCM give cause for some concern; however, research measuring performance trends has generally been quite successful in finding significant effects. In models of job performance trends using just an intercept and slope, significance has been found quite frequently (e.g., Deadrick et al., 1997; Stewart & Nandkeolyar, 2006; Sturman & Trevor, 2001). Hofmann et al. (1993) found significant effects for their hypothesized quadratic and cubic growth terms (although not linear). Therefore, power does not seem to be a major hindrance for modeling growth trends, although power for detecting moderators of these trends still remains in question. Applicants of LGCM are less common in management research (Williams et al., 2003), and particularly for modeling individual job performance over time. In one example, Ployhart and Hakel (1998) using LGCM found the linear, quadratic, and cubic effects that they hypothesized. Thus, while power is always a concern in empirical research and has the potential to be particularly limiting for studies of dynamic performance, it does not appear to have been a significant issue in applications of job performance trend modeling.

Choosing between HLM and LGCM

As specific applications of the GLMM, HLM and LGCM each have advantages and disadvantages. The choice of HLM and LGCM approaches depends on the nature of the data structure and the desired treatment of error terms (Raudenbush, 2001). In short, HLM allows for greater flexibility with regard to the form of the data, but limited flexibility with regard to the error structure. In contrast, LGCM allows for more choices for modeling the error, but less flexibility in terms of the data's structure (Raudenbush, 2001).

If one has unbalanced data, the LCGM is simply not an option (unless missing data techniques can make the data balanced, but this will be discussed later). HLM provides the flexibility to model this sort of data structure and determine the nature of employee trends. On the other hand, if one is testing or otherwise cannot accept the assumptions regarding the error structure (or other factors that can be manipulated in the SEM-base LCGM method that cannot be changed in HLM), then the flexibility of LCGM makes it a more desirable option. Ideally, this choice should not simply be guided by the convenience of data availability, but be driven by a strong theoretical rationale or the need for specific hypothesis testing.

Methodological Issues

While characteristics of the data and the nature of the desired analyses must be considered to choose a research design and a technique for analyzing the subsequent job performance data, there are also methodological issues that are inevitable when studying dynamic performance. All of these issues will influence research on job performance over time in at least some way regardless of analytical choice. I will briefly review the types of issues and their likely consequences for research on this topic.

The Measurement of Job Performance

Research examining performance over time has generally examined either job performance ratings from supervisors or measures of performance results (Sturman et al., 2005). While research on job performance has acknowledged that the two measures are different (e.g., Bommer, Johnson, Rich, Podsakoff, & Mackenzie, 1995), the dynamic performance literature has rarely made this distinction.

Measurement error is a concern for both behavior- and results-based performance measures. While there is extensive history to understanding reliability of a given measure at a point in time (Nunnally & Bernstein, 1994) and the reliability (intra-rater and inter-rater) of job performance specifically (Bommer et al., 1995; Viswesvaran, Ones, & Schmidt, 1996), less attention has been paid to estimating the test–retest reliability of job performance ratings. Sturman et al. (2005) examined the consistency of job performance ratings (objective and subjective) by separating the variance due to (a lack of) test–retest reliability and performance (in)consistency. They supported their hypothesis that the test–retest reliability of subjective (behavioral) measures of job performance would be higher than the test– retest reliability of objective (results-based) measures. They argued that the greater unreliability of objective measures was due to environmental constraints beyond the control of employees (i.e., beyond the influence of their behaviors). This perspective is consistent with the findings of Stewart and Nandkeolyar (2006) who showed that a measure of environmental constraints affected employee performance trends. Specifically, they showed that the environmental factor of sales referrals explained 60% of variation in salesperson weekly performance.

These findings reveal that all measures of job performance are subject to error. ‘‘Objective’’ measures may be unaffected by a lack of intra- or inter-rater reliability, but they suffer from more test–retest unreliability than do ‘‘subjective’’ measures (Sturman et al., 2005). The result is that, regardless of the type of measure employed, measurement error is a methodological problem for all job performance research.

The findings from Sturman et al. (2005) highlight the importance of distinguishing between job performance$_{\text{(behaviors)}}$ and job performance$_{\text{(results)}}$. By measuring results instead of behaviors, such research is considering a related but fundamentally different phenomenon than the research on job performance generally considers to be the focal construct. Furthermore, as most jobs do not possess an ‘‘objective’’ measure of performance like jobs with sales data, it is not apparent if results based on objective measures explain the nature and trend of job performance$_{(behaviors)}$ in other contexts. The data availability of ‘‘objective’’ data may make them at first to appear preferable, but such measures do not directly speak to the construct of job performance(behaviors) that theory on job performance is looking to advance. While results-based measures may proxy job performance$_{(behaviors)}$, they capture (at least) the additional effects of environmental constraints and thus cannot contribute as well to theoretical development in this area.

I am not suggesting that research on job performance$_{(results)}$ has been for naught. Certainly, the results are useful for demonstrating how job performance$_{(behaviors)}$ and job performance$_{(results)}$ differ, illustrating how different methodologies can be used to study job performance over time, and providing a useful starting point for considering how we expect job performance$_{(behaviors)}$ to change with time. Studying job performance$_{(results)}$ also has advantages that it enables more within-person observations, and certainly job performance$_{(results)}$ is an outcome of interest to organizations. Nonetheless, to improve our understanding of job performance$_{(behaviors)}$ over time, the methodological conveniences of job performance$_{(results)}$ do not overcome the fundamental differences that exist between job performance$_{(behaviors)}$ and job performance$_{(results)}$. To advance our understanding of how the construct of job performance$_{(behaviors)}$ changes with time, dynamic performance research must place a greater emphasis on this criterion and forgo the conveniences associated with ‘‘objective’’ measures. Researchers will either need to deal with the difficulties associated with gathering subjective ratings of performance or accept and acknowledge the imprecision and unreliability of objective ratings of performance results.

Missing Data

In addition to measurement error, missing data is also a ubiquitous problem for research studying employee performance over time. While missing data issues are common for longitudinal research in general because of attrition in multiwave studies (Goodman & Blum, 1996; Newman, 2003), it is a particular problem in longitudinal studies of job performance because there are systematic relationships between job performance and attrition. First, extensive evidence reveals a relationship between job performance and voluntary turnover (e.g., Harrison et al., 1996; Sturman & Trevor, 2001; Trevor et al., 1997; Salamin & Hom, 2005; Williams & Livingstone, 1994). Second, companies use such mechanisms as probationary periods to fire low performers, creating a relationship between job performance and involuntary turnover (e.g., De Corte, 1994). Third, companies may promote high performers to other jobs, thereby creating another mechanism that can create a relationship between performance and the likelihood of missing data. The consequence of the systematic relationships between performance and attrition not only will cause data to be missing in longitudinal studies of job performance, but it will restrict the range in observed performance scores (Sturman & Trevor, 2001).

Research on missing data (e.g., Little & Rubin, 2002) has identified three types of missing data: missing completely at random (MCAR), missing at random (MAR), and not missing at random (NMAR). For data to be MCAR, it must both be ‘‘observed at random’’ (OAR) and ‘‘missing at random’’ (MAR). That is, (1) the pattern of missing data must not depend on the values of data that are observed (i.e., it is OAR), and (2) the likelihood that data is missing must not depend on the values of the data that are missing (i.e., it is MAR). If these conditions are met, then missing data will not likely bias population mean estimates (Little & Rubin, 2002). For predicting job performance (P_t), data is classified as MAR if the probability of missingness of P_t (i.e., the probability of not being able to observe job performance) depends on X (a predictor of job performance) but not after controlling for X. If this condition is not met (i.e., the probability that data is missing at time t depends on job performance at time t), then data is NMAR.

Because of the relationship between performance and turnover, MCAR is obviously not likely. Furthermore, if one were to observe a longitudinal sample of job performance where data was not missing, one would also have to question the generalizability of that data. Undoubtedly, one must consider whether data is MAR or NMAR. For data to be MAR, then the likelihood of missing data on Pt must only depend on characteristics from the prior time period(s) (X variables and/or $P_{(t-1)}$). For data to be NMAR, the likelihood of missing data on P_t depends on P_t. In many ways, by this definition, it seems that longitudinal performance data will be MAR. First, if someone is fired because of low performance, then data on P_t is missing because of the value of P_{t-1}. Similarly, if a high performer at P_{t-1} feels unrewarded and seeks new employment (i.e., leaves the company and is unobserved for Pt), then again missing data is MAR. On the other hand, the nature of performance measurement may cause data to be NMAR. For example, an employee may be performing badly during a given year (year t), and because of being self-aware of this performance, feedback from others, or being discharged, may leave the organization. Because supervisory ratings of job performance are measured annually, the measure for performance at time t would not occur. Nonetheless, the data is missing in this situation specifically because of performance in time t. In such a case, data would be NMAR.

To date, no research has specifically considered the issue of missing data with respect to longitudinal studies of job performance. It is not clear if missing data should be treated MAR or MCAR. What is clear, however, is that most research on dynamic performance has not directly addressed issues relevant to missing data. Specifically, most studies give the issue very little consideration, instead simply use list-wise deletion to address the missing data problem. That is, previous studies examining employee performance levels over time have most frequently eliminated from the sample those employees who leave the job before the full length of data collection (e.g., Henry & Hulin, 1987; Ployhart & Hakel, 1998; Rambo et al., 1983, 1987; Rothe, 1978). When interested in only predicting data in the final wave of a study (e.g., predicting performance in wave six of a six-wave study), then list-wise deletion does perform as well as other missing data techniques (Newman, 2003). However, when interested in predictors and information in earlier waves of the study (such as in trying to estimate coefficients explaining performance trends), then list-wise deletion generally performs worse than all other missing data techniques (Newman, 2003).

In his study of missing data techniques for longitudinal research, Newman (2003) concluded that list-wise deletion should be avoided, and instead one should employ maximum likelihood or multiple imputation approaches. Both of these approaches were shown to work best when generating parameter estimates; the full information maximum likelihood (a form of maximum likelihood estimation) and multiple imputation methods worked best for estimating standard errors. Based on this research, if one is studying longitudinal performance data and needs to employ a missing data technique, it appears that full information maximum likelihood or multiple imputation should be used. To date, this has not occurred for research on dynamic performance.

If one wants to employ LGCM, then missing data is a major concern because the technique requires the same number of observations per subject. Fortunately, programs like LISREL, AMOS, Mplus, and SAS provide routines for implementing FIML (e.g., in LISREL, one adds the command ‘‘mi=.’’ to the data step). On the other hand, one can avoid the missing data issue by using HLM, which does not require balanced data (although at least two points of data are needed to model a linear effect, three to model a quadratic effect, and so on, so missing data may still be an issue for analyses performed with HLM). This once again raises the debate as to whether one should use a LGCM or HLM approach. Currently, the issue cannot be resolved; each technique has advantages and disadvantages, and more methodological research is needed to specifically consider these sorts of issues for dynamic performance research. Given the importance of the contrast between HLM and LGCM approaches, research is needed on missing data techniques (or not using missing data techniques) for empirical work specifically on job performance over time. The field needs to know the consequences of choosing HLM over LGCM, and if job performance can be considered MAR or if it is NMAR.

In short, while advances in methodology present exciting opportunities for the analysis of longitudinal data, the nature of studying job performance over time creates specific problems that may influence the utility of these new techniques. Until the field has a better understanding of these specific issues, the interpretation of longitudinal results will always be open to some question.

The Future: New Directions for Dynamic Performance Research

The opportunities for future research to contribute to our understanding of job performance within the context of time are quite substantial. Yet, progress in this area will require both theoretical and methodological advances. Furthermore, it will be critical that the theoretical and methodological research have a reciprocal relationship, using findings in one area to guide the next steps of research in the other. In this section, I will identify a number of areas for future research and specific research questions that need to be addressed to better understand the dynamic performance phenomenon.

Further Specifying Longitudinal Models

The review of theory earlier in this article presented a number of different perspectives of dynamic performance: the changing-subjects and changing-tasks model; longitudinal extensions of static performance models; Ackerman's and Murphy's performance stage models; and Learning Curve Theory. These models, though, are surprisingly complementary, and taken together suggest that theory for modeling job performance over time is not as under developed as some have claimed.

The changing-subjects and changing-tasks models are more metaphorical than theoretical. They present two explanations as to why job performance changes with time, but otherwise do not provide the type of propositions requisite of a theory. However, when considering the longitudinal extension of static performance models, the changing-subjects and changing-tasks models facilitate the discussion of the types of effects that longitudinal models can detect, and they help present a structure for framing the discussion of any effects that are discovered.

The stage models provide clarification of the sort of variables that should affect performance change. Ackerman's and Murphy's works highlight that the effects associated with abilities should change with time. Previous research has already identified a number of factors that can be incorporated into these models (Steele-Johnson et al., 2000). Longitudinal applications of static performance models and approaches to modeling performance trends will reveal ways in which such effects should be observed and which additional variables should be considered. A stream of research, utilizing both cross-sectional and longitudinal designed, is therefore needed to answer this first set of research questions:

- What static variables have stable relationships with job performance?
- What static variables have changing relationships with job performance?
- What dynamic variables have stable relationships with job performance?
- What dynamic variables have changing relationships with job performance?

Searching for the predictors of performance and estimating the functional forms of their relationships with job performance require a variety of methodological approaches. Longitudinal analyses can examine within-person relationships to help answer the above questions. It is likely, though, that obtaining such data will be difficult, and any longitudinal analyses will suffer from methodological limitations (including missing data, multiple sources of error, and potentially low power). Therefore, I recommend that future research answering the questions above should be complemented with cross-sectional research. Research capturing a snap-shot of performance relationships will help identify the variables that relate to performance. Furthermore, any longitudinal model should include predictions as to what such a model implies for a point in time, and these hypotheses should be tested. Failure to support the point-in-time predictions from longitudinal models would falsify the model. Consequently, such tests are critical for theory development. Given the many difficulties associated with longitudinal studies of job performance, it would slow the potential progress of the field to ignore the value of appropriately designed cross-sectional research.

Further Refining Analytical Methods

Answering (at least to some degree) the questions above will clarify and specify the stage models of job performance. Nonetheless, confirmatory cross-sectional tests cannot conclusively prove any such longitudinal model, and longitudinal research is inevitably required to fully understand job performance within the context of time. The question remains as to how best to analyze these variables with a longitudinal design. HLM and LGCM are both potentially fruitful analytical techniques, but we still need guidance as to how these methods should be applied specifically to the issue of modeling job performance.

This is where Learning Theory can play an important role. A better understanding of the specific shape of the individual job performance learning curves will provide guidance into the structure of both LGCM and HLM approaches for modeling performance trends. It is often espoused that theory should drive analytical models, and given the difficulty in justifying the form of highly parameterized polynomial models, it would be desirable to have a tested theoretical rationale upon which to base model design. In the operations literature, Learning Curve Theory has provided this sort of insight at an aggregated level of performance. The theory has the opportunity to provide similar guidance at the individual-level of analysis.

Learning Curve Theory should also be able to shed light on to the nature of how error terms are related over time. One of the key advantages of the LGCM approach is the flexibility of its form. This flexibility comes at the cost of additional parameters needing to be estimated (Rovine & Molenaar, 2001). Learning Curve Theory can shed light on more than just the nature of job performance trends, but also on the form of the model's error structure. By having a better understanding of the functional form of job performance over time, including of and between its coefficients and for its error terms, future research will be better able to employ the LGCM approach by fixing certain parameters based on appropriate theoretical estimates. This leads to a second set of research questions:

- How can Learning Curve Theory be applied to modeling job performance over time?
- How should longitudinal models of job performance represent performance curves (and what are the implications of failing to model these curves correctly)?
- What is the nature of the error structure for longitudinal models of job performance (and the implications of failing to consider this structure)?

Answering these questions will provide guidance to researchers as to how to design their models and allow research to move beyond questions of model structure (e.g., should performance be measured with a linear term, or up to cubic terms?) and instead focus on other practical and theoretical questions (e.g., what predicts or moderates the growth curves?).

Addressing Methodological Problems

Continuing the stream of research combining the performance and turnover literatures (e.g., Harrison et al., 1996; Sturman & Trevor, 2001) also seem to have useful theoretical and methodological implications. Turnover affects data attrition, which influences the existence of missing data. The need for balanced data is the most obvious difference between HLM and LGCM approaches, and although we have some insights about handling missing data in longitudinal studies, the nature of the missing data (MAR or NMAR) for longitudinal studies of job performance, the comparison of missing data techniques with LGCM versus HLM, and the implications of not employing missing data techniques with HLM, are unaddressed.

For the prediction of job performance over time, turnover is also a key outcome. One cannot really study performance over time without considering when performance no longer exists. While much of the discussion so far has focused on explaining or modeling existing data, understanding turnover and job performance is important for any sort of prediction problem. This leads to a third set of theoretical and methodological research questions:

- Is missing data in longitudinal studies of job performance MAR or NMAR?
- How should turnover and missing data be incorporated into models of job performance over time?
- What are the implications of the performance/turnover relationship for modeling job performance over time?

Refining Stage Models of Job Performance

Hopefully, progress can be made in understanding the structure of job performance over time and the processes involved in its modeling, but the field needs to understand the predictors and moderators of performance trends to better achieve the goal of understanding and affecting job performance over time. The Murphy and Ackerman models suggest that contextual factors can influence the nature of performance over time. In particular, both models mention job complexity, which other research has shown to be an important moderator of job performance predictors (e.g., Schmitt et al., 1984; Sturman, 2003; Sturman et al., 2005; Tubre & Collins, 2000). Unfortunately, the stage models are still relatively undeveloped in terms of their specific predictions. Addressing the research questions already articulated above will help identify the variables that should be included in revisions of the models. As more variables are included in these models, theoretical development of these models should follow so as to address how the newly specified variables relate to job performance over time.

It will also be important to clearly distinguish between the Ackerman and Murphy models, or perhaps further adapt them for the specific purpose of predicting job performance. A weakness of Ackerman's and Murphy's models is that they both have simplistic treatments of performance (task performance for Ackerman's model, and job performance in general for Murphy's model). Developments in the understanding of job performance have shown job performance to be multidimensional (Motowidlo et al., 1997; Rotundo & Sackett, 2002; Viswesvaran et al., 2005; Welbourne et al., 1998). When Murphy's paper was written, the literature on job performance had not yet made this distinction. In his work, Murphy defines his criterion, job performance, as ''overall job performance'', taking into account performance on specific tasks but also ''variables such as success in maintaining good interpersonal relations, absenteeism and withdrawal behaviors, substance abuse, and other behaviors that increase hazards at the workplace'' (p. 185). As such, his model's focus on overall job performance is comparable with Rotundo and Sackett's definition of overall job performance, and is similarly comprised on job task performance (e.g., ‘‘performance on specific tasks’’), contextual performance (e.g., ‘‘maintaining good interpersonal relations’’ and ‘‘behaviors that contribute to…the achievement of goals associated with their jobs’’), and counter-productive behaviors (e.g., ‘‘withdrawal behaviors, substance abuse, and other behaviors that increase hazards at the work place’’). It is likely that there are different functional relationships for the predictors of the different dimensions of performance (Steele-Johnson et al., 2000). The current developments in understanding job performance as a multidimensional construct suggest that the stage models can be extended to consider the different dimensions of job performance. The need for future work developing and refining the performance stage models leads to this fourth set of research questions:

- How can the differences between the Ackerman and Murphy’s models be resolved to yield a single dynamic model of job performance?
- How can such a resultant stage model of job performance be modified and updated to better understand job performance over time?
- How should models of job performance over time be adapted to incorporate the different dimensions of job performance?

Determining the Effects of Human Resource Interventions on Job Performance Over Time

So far, the research questions I have identified are aimed at improving the understanding and prediction of job performance; yet this knowledge has limited direct applied value. The desire to affect performance curves highlights the need to understand job performance and time when considering any sort of human resource intervention. While certainly prediction is valuable for selection decisions, the purpose of many human resource interventions is to affect employee performance. Research on job performance over time has the potential to benefit many varied fields of human resources.

It should be recalled that much of the early work on dynamic performance stemmed from a concern about the prediction of job performance. In particular, researchers considered the implications of dynamic performance for the validity of various selection devices. While the question of whether performance is dynamic has been resolved, the implications of this dynamism are still unknown. Research is needed into the validity of selection devices for predicting both initial performance levels and performance curves. The nature of this predictability should also be evaluated to improve decision-making. When performance was assumed to be static, decision-making with selection devices was simple: higher scores were better. However, there are likely to be tradeoffs when considering selection devices in a longitudinal context. How does one compare the utility of a device with high initial predictability but poor predictability in terms of performance trends, with a device with poor initial predictability but high validity for predicting performance trends?

Beyond selection, other functional areas of human resources would benefit from considering job performance over time. The purpose of many human resource interventions is to improve employee performance. For example, pay-for-performance is supposed to affect motivation to yield better performance; training programs are supposed to affect motivation or abilities to elicit higher performance. All of these interventions implicitly involve the passage of time to achieve the desired results. Given what we know about performance trends, simply looking at before/after change scores is incomplete with regard to understanding job performance over time. Longitudinal designs are needed to control for current performance trends to determine if a human resource intervention truly has the intended effect. This leads to a fifth set of research questions that would behoove future research on job performance over time to address.

- What are the temporal validities of common selection devices (e.g., unstructured interview, structured interview, cognitive ability tests, personality tests, assessment centers, integrity tests)?
- How do compensation systems (e.g., pay policy, pay hierarchy, bonuses, raises, pay-for-performance linkages, group-based incentives) influence job performance trends?
- How does training (e.g., training types, training delivery methods, trainee characteristics, trainer characteristics) influence job performance trends?
- What other human resource interventions affect the modeling of employee performance trends?

Introducing the Need to Predict Employee Performance Vectors

It is now clear that job performance is dynamic, multidimensional, and constrained by turnover. For these reasons, I argue that the performance prediction problem needs to evolve beyond predicting a single performance score to the prediction of what I will label job performance vectors. While such a data structure is not novel from a statistical point of view, the information that such a metric contains presents a new approach to human resource research and human resource decision-making.

I define a job performance vector as a $C \times N \times 3$ matrix of information on a given employee (or applicant). This matrix includes C dimensions of performance (e.g., task performance, citizenship behaviors, counterproductive behavior), projected for N time periods (e.g., annual performance for up to 10 year). The predicted performance level is one piece of information contained in the matrix (the first component of the third dimension), the estimated accuracy of this estimate (e.g., standard error) is the second component of the third dimension, and the third component provides the estimated probability of the performance being observed (i.e., the probability that the individual remains employed by the organization). Each individual's matrix can contain information on both past performance and the predicted levels of performance and turnover likelihood.

All human resource decisions that involve predicting performance (e.g., who to hire, who to promote, who to reward, who to train) can be based on the information contained in this matrix. Similarly, human resource interventions can be evaluated based on the their predicted effects on data contained in these matrices (e.g., what are the expected effects of implementing a new selection system, a new pay plan, a new training program, a new feedback system?).

Estimating performance vectors will require the combination of theory, empirical research, individual-specific information, and company-specific information. Existing theory and empirical evidence helps establish expected patterns. For example, Learning Curve Theory or past evidence from research predicting performance trends can provide a baseline of expected values. That is, with no other information, instead of assuming all performance has an expected value of the mean (0 if expressed in standardized scores), expected performance levels should follow some sort of learning curve. General company information can provide further information, such as the probability of turnover for any given position.

As new information is acquired, the vectors can be updated and refined, either based on company-specific investigations or from existing research. For example, once performance has been observed, subsequent expected performance levels and turnover probabilities can be updated. Information on job candidates, used in conjunction with the results of validation studies or existing research, can also refine this information. As more information is collected, both within the company and from research advances, the quality of information contained in the matrix can be improved. The methods used to derive the necessary information will also advance as companies perform their own research (e.g., validation studies) and as new studies emerge with relevant findings.

The idea of building, refining, and using performance vectors in human resource practice is new, and would certainly require new advances in methodology and decision making to implement successfully. First, tools would need to be developed that can combine information from a company's human resource information system with varied and complex research findings. Second, methods for empirically reviewing existing research findings would need to be applicable to studying many variables simultaneously rather than a single relationship in isolation. Third, theory would need to provide specific information on functional forms rather than just general information on whether an effect is positive or not. Fourth, all of this information would need to be able to be combined to yield specific point estimates (of performance levels, the accuracy of these estimates, and the likelihood of turnover) for job applicants and incumbents. Finally, the methods used to derive these estimates would have to be capable of ‘‘learning’’ and updating these values as new data is acquired and new research findings emerge. In short, this is no small task.

It is my belief that this task of performance vectoring presents a new but valuable approach to the applied prediction problem, and a new way to build a connection between research and practice. Fundamentally, the task is that of predicting performance over time; yet, the requirements of the task reveal how all realms of human resource research related to the prediction of job performance need to be combined to provide any hope of being able to make this task feasible. With performance vectoring being just introduced, any research trying to contribute to this area must begin by attempting to address the following three fundamental research questions:

- How can performance vectors be modeled?
- How can performance vectors be used to make human resource decisions?
- How can performance vectors be used to evaluate human resource programs?

Conclusion

The study of dynamic performance has a long history, but understanding the nature of job performance over time has had only limited development. I argue, based on what I see as clear and convincing evidence, that the answer to the question ‘‘is performance dynamic?’’ has been resolved. The answer is a resounding ‘‘yes’’; job performance does change over time. The field has thus moved beyond this simple question to trying to understand the nature of job performance over time and its implications for human resource practice. Theoretical models are available that provide general information as to why performance changes or what performance trends may look like, but there has been no clear direction as to what variables to study, what questions to ask, what methods to employ, how to employ those methods, and how to interpret their results. Although greatly limited by the general difficulty of getting sufficiently large longitudinal datasets, the limitations of various methodological designs can be well understood and less-than-ideal datasets (including cross-sectional ones) can utilize complementary methods to make significant progress along this research path.

An employment relationship, by its very nature, connotes events, reactions, behaviors, and perceptions that occur over time. From an organization’s point of view, a primary (if not the primary) outcome of this relationship is the employee’s job performance. As such, what happens to performance over time is central to the employment relationship, but it is frequently ignored and far from well-understood. If simply the study of the criterion (i.e., job performance) has been cited as one of the most neglected elements in the applied prediction problem (Dunnette, 1963; Campbell, 1990; Motowidlo et al., 1997), performance within the context of time has received even less attention and is even less-well understood. And yet, between the available empirical examples, models and theories of learning, and methodological advances, there is genuine opportunity for our understanding of job performance over time to make significant strides in the future. Will future research perform these needed steps to make these contributions? Ironically, time will tell.

References

Ackerman, P. L. (1987). Individual differences in skill learning: An integration of psychometrics and information processing perspectives. *Psychological Bulletin, 102*, 3–27.

Ackerman, P. L. (1988). Determinants of individual differences during skill acquisition: Cognitive abilities and information processing. *Journal of Experimental Psychology: General, 177,* 288–318.

Ackerman, P. L. (1989). Within-task intercorrelations of skilled performance: Implications for predicting individual differences? *Journal of Applied Psychology, 74*, 360–364.

Ackerman, P. L. (1992). Predicting individual differences in complex skill acquisition: Dynamics of ability determinants. *Journal of Applied Psychology, 77*, 598–614.

Adams, J. A. (1957). The relationship between certain measures of ability and the acquisition of a psychomotor response. *Journal of General Psychology, 56,* 121–134.

Adler, P. S., & Clark, K. B. (1991). Behind the learning curve: A sketch of the learning process. *Management Science, 37*, 267–281.

Alvares, K. M., & Hulin, C. L. (1972). Two explanations of temporal changes in ability-skill relationships: A literature review and theoretical analysis. *Human Factors, 14*, 295–308.

Alvares, K. M., & Hulin, C. L. (1973). An experimental evaluation of a temporal decay in the prediction of performance. *Organizational Behavior and Human Performance, 9*, 169–185.

Austin, J. T., Humphreys, L. G., & Hulin, C. L. (1989). Another view of dynamic criteria: A critical reanalysis of Barrett, Caldwell, and Alexander. *Personnel Psychology, 42*, 583–596.

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

Barrett, G. V., & Alexander, R. A. (1989). Rejoinder to Austin, Humphreys, and Hulin: Critical reanalysis of Barrett, Caldwell, and Alexander. *Personnel Psychology, 42*, 597–612.

Barrett, G. V., Alexander, R. A., & Doverspike, D. (1992). The implications for personnel selection of apparent declines in predictive validities over time: A critique of Hulin, Henry, and Noon. *Personnel Psychology, 45,* 601–617.

Barrett, G. V., Caldwell, M. S., & Alexander, R. A. (1985). The concept of dynamic criteria: A critical reanalysis. *Personnel Psychology, 38*, 41–56.

Barrett, G. V., Caldwell, M. S., & Alexander, R. A. (1989). The predictive stability of ability requirements for task performance: A critical reanalysis. *Human Performance, 2*, 167–181.

Barrick, M. R., & Mount, M. K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology, 44*, 1–26.

Bass, B. M. (1962). Further evidence on the dynamic character of criteria. *Personnel Psychology, 15,* 93–97.

Bayley, N. (1949). Consistency and variability in the growth of intelligence from birth to eighteen years. *Journal of General Psychology, 75*, 165–196.

Bayley, N. (1955). On the growth of intelligence. *American Psychologist, 10*, 805–818.

Behling, O. (1998). Employee selection: Will intelligence and conscientiousness do the job? *Academy of Management Executive, 12,* 77–86.

Belkaoui, A. (1986). *The Learning Curve: A management accounting tool*. Westport, CT: Quarum Books.

Bemis, J. (1981). A model for examining cost implications of production rate. *Concepts: The Journal of Defense Systems Acquisition Management, 4,* 63–76.

Bommer, W. H., Johnson, J. L., Rich, G. A., Podsakoff, P. M., & Mackenzie, S. B. (1995). On the interchangeability of objective and subjective measures of employee performance: A meta-analysis. *Personnel Psychology, 48*, 587–605.

Borman, W. C. (1991). Job behavior, performance, and effectiveness. In: M. D. Dunnette & L.M. Hough (Eds), *Handbook of industrial and organizational psychology* (2nd ed., Vol. 2, pp. 271–326). Palo Alto, CA: Consulting Psychologists Press.

Boudreau, J. W. (1991). Utility analysis. In: M. D. Dunnette & L. M. Hough (Eds), *Handbook of industrial and organizational psychology* (Vol. 2). Palo Alto, CA: Consulting Psychologists Press.

Boudreau, J. W., & Berger, C. J. (1985). Decision-theoretic utility analysis applied to external employee movement. *Journal of Applied Psychology, 70*, 581–612.

Brogden, H. E., & Taylor, E. K. (1950). The dollar criterion: Applying the cost accounting concept to criterion construction. *Personnel Psychology, 3*, 133–167.

Campbell, J. P. (1990). Modeling the performance prediction problem in industrial and organizational psychology. In: M. D. Dunnette & L. M. Hough (Eds), *Handbook of industrial and organizational psychology* (Vol. 1). Palo Alto, CA: Consulting Psychologists Press.

Campbell, J. P., McCloy, R. A., Oppler, S. H., & Sager, C. E. (1993). A theory of performance. In: N. Schmitt & W. Borman (Eds), *Personnel selection in organizations* (pp. 35–70). San Francisco, CA: Jossey Bass.

Carlson, K. D., Connerley, M. L., & Mecham, R. L., III. (2002). Recruitment evaluation: The case for assessing the quality of applicants attracted. *Personnel Psychology, 55,* 461–490.

Carmines, E. G., & Zeller, R. A. (1979). *Reliability and validity assessment.* Newbury Park, CA: Sage.

Carr, G.W. (1946). Peacetime cost estimating requires new learning curves. *Aviation*, April, 76–77.

Chait, H. N., Carraher, S. M., & Buckley, M. R. (2000). Measuring service orientation with biodata. *Journal of Managerial Issues, 12*, 109–120.

Charles, D. C. (1953). Ability and accomplishment of persons earlier judged mentally deficient. *General Psychological Monographs, 47*, 3–71.

Costa, P. T., Jr., & McCrae, R. R. (1988). Personality in adulthood: A six-year longitudinal study of self-reports and spouse ratings on the NEO Personality Inventory. *Journal of Personality and Social Psychology, 54*, 853–863.

Costa, P. T., Jr., & McCrae, R. R. (1992). Four ways five factors are basic. *Personality and Individual Differences, 13*, 653–665.

Day, D. V., Sin, H., & Chen, T. T. (2004). Assessing the burdens of leadership: Effects of formal leadership roles on individual performance over time. *Personnel Psychology, 57*, 573–605.

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

Deadrick, D. L., & Madigan, R. M. (1990). Dynamic criteria revisited: A longitudinal study of performance stability and predictive validity. *Personnel Psychology, 43*, 717–744.

De Corte, W. (1994). Utility analysis for the one-cohort selection-retention decision with a probationary period. *Journal of Applied Psychology, 79*, 402–411.

DeJong, J. R. (1957). The effects of increasing skill on cycle time and its consequences for time standards. *Ergonomics, 1,* 51–60.

Dennis, W. (1954). Predicting scientific productivity in later maturity from records of earlier decades. *Journal of Gerontology, 9*, 465–467.

Dennis, W. (1956). Age and productivity among scientists. *Science, 123*, 724–725.

Dudley, N. M., Orvis, K. A., Lebiecki, J. E., & Cortina, J. M. (2006). Meta-analytic investigation of conscientiousness in the prediction of job performance. Examining the intercorrelations and the incremental validity of narrow traits. *Journal of Applied Psychology, 91*, 40–57.

Duncan, T. E., Duncan, S. C., Strycker, L. A., Li, F., & Alpert, A. (1999). *An introduction to latent variable growth curve modeling: Concepts, issues, and applications*. Mahwah, NJ: Lawrence Erlbaum.

Dunham, R. (1974). Ability-skill relationships. An empirical explanation of change over time. *Organizational Behavior and Human Performance, 12*, 372–382.

Dunnette, M. D. (1963). A note on the criterion. *Journal of Applied Psychology, 47*, 251–254.

Edwards, J. R. (1994). The study of congruence in organizational behavior research: Critique and a proposed alternative. *Organizational Behavior and Human Decision Processes, 58*, 51–100.

Edwards, J. R. (2001). Ten difference score myths. *Organizational Research Methods, 4*, 265–287.

Farrell, J. N., & McDaniel, M. A. (2001). The stability of validity coefficients over time: Ackerman's (1988) model and the general aptitude battery. *Journal of Applied Psychology, 86,* 60–79.

Fleishman, E. A. (1953). A factor analysis of intra-task performance on two psychomotor tests. *Psychometrika, 18*, 45–55.

Fleishman, E. A. (1960). Abilities at different stages of practice in rotary pursuit performance. *Journal of Experimental Psychology, 60,* 162–171.

Fleishman, E. A., & Hempel, W. E., Jr. (1954). Changes in factor structure of a complex psychomotor test as a function of practice. *Psychometrika, 19*, 239–252.

Fleishman, E. A., & Rich, S. (1963). Role of kinesthetic and spatial-visual abilities in perceptual-motor learning. *Journal of Experimental Psychology, 66*, 6–11.

Gatewood, R. D., & Feild, H. S. (2001). *Human resource selection* (5th ed.). New York: Harcourt.

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

Ghiselli, E. E., & Haire, M. (1960). The validation of selection tests in the light of the dynamic character of criteria. *Personnel Psychology, 13,* 225–231.

Goodman, J. S., & Blum, T. C. (1996). Assessing the non-random sampling effects of subject attrition in longitudinal research. *Journal of Management, 22*, 627–652.

Guibert, P. (1945). Le Plan de Fabrication Aeronautique. Paris: Dunod. (English translation ''*Mathematical Studies of Aircraft Construction''*) available through Central Air Documents Office, Wright-Patterson AFB, Dayton, OH.

Guttman, L. (1955). A generalized simplex for factor analysis. *Psychometrika, 20*, 173–192.

Hackman, J. R., & Oldham, G. R. (1976). Motivation through the design of work: Test of a theory. *Organizational Behavior and Human Performance, 16*, 250–279.

Hanges, P. J., Schneider, B., & Niles, K. (1990). Stability of performance: An interactionist perspective. *Journal of Applied Psychology, 75*, 658–667.

Harrison, D. A., Virick, M., & William, S. (1996). Working without a net: Time, performance, and turnover under maximally contingent rewards. *Journal of Applied Psychology, 81*, 331–345.

Heise, D. R. (1969). Separating reliability and stability in test–retest correlation. *American Sociological Review, 34*, 93–101.

Heneman, H. G., Heneman, R. L., & Judge, T. A. (1997). *Staffing organizations*. Chicago, IL: Irwin.

Henry, R. A., & Hulin, C. L. (1987). Stability of skilled performance across time: Some generalizations and limitations on utilities. *Journal of Applied Psychology, 72*, 457–462.

Henry, R. A., & Hulin, C. L. (1989). Changing validities: Ability-performance relations and utilities. Journal of *Applied Psychology, 74*, 365–367.

Hertzog, C., Lindenberger, U., Ghisletta, P., & von Oertzen, T. (2006). On the power of multivariate latent growth curve models to detect correlated change. *Psychological Methods, 11*, 244–252.

Hertzog, C., & Schaie, K. W. (1986). Stability and change in adult intelligence: 1. Analysis of longitudinal means and covariance structures. *Psychology and Aging, 1*, 159–171.

Hertzog, C., & Schaie, K. W. (1988). Stability and change in adult intelligence: 2. Simultaneous analysis of longitudinal means and covariance structures. *Psychology and Aging, 3*, 122–130.

Hirschmann, W. B. (1964). Profit from the learning curve. *Harvard Business Review*, January– February, 128.

Hofmann, D. A. (1997). An overview of the logic and rationale of hierarchical linear models. *Journal of Management, 23*, 723–744.

Hofmann, D. A., & Gavin, M. B. (1998). Centering decisions in hierarchical linear models: Implications for research in organizations. *Journal of Management, 24*, 623–641.

Hofmann, D. A., Jacobs, R., & Baratta, J. E. (1993). Dynamic criteria and the measurement of change. *Journal of Applied Psychology, 78*, 194–204.

Hofmann, D. A., Jacobs, R., & Gerras, S. J. (1992). Mapping individual performance over time. *Journal of Applied Psychology, 77*, 185–195.

Hulin, C. L., Henry, R. A., & Noon, S. L. (1990). Adding a dimension: Time as a factor in the generalizability of predictive relationships. *Psychological Bulletin, 107*, 328–340.

Humphreys, L. G. (1960). Investigations of the simplex. *Psychometrika, 25*, 313–323.

Hunter, J. E. (1986). Cognitive ability, cognitive aptitudes, job knowledge, and job performance. *Journal of Vocational Behavior, 29*, 340–362.

Hunter, J. E., & Hunter, R. F. (1984). Validity and utility of alternate predictors of job performance. *Psychological Bulletin, 96,* 72–98.

Hurtz, G. M., & Donovan, J. J. (2000). Personality and job-performance: The Big Five revisited. *Journal of Applied Psychology, 85*, 869–879.

Jensen, A. R. (1980). *Bias in mental testing*. New York: Free Press.

Jöreskog, K. G. (1970). Estimation and testing of simplex models. *British Journal of Mathematical and Statistical Psychology, 23,* 121–145.

Kanfer, R. (1991). Motivation theory and industrial and organizational psychology. In: M. D. Dunnette & L. M. Hough (Eds), *Handbook of industrial and organizational psychology* (Vol. 1, pp. 75–170). Palo Alto, CA: Consulting Psychologists Press.

Kanfer, R. (1992). Work motivation: New directions in theory and research. In: C. L. Cooper & I. T. Robertson (Eds), *International review of industrial and organizational psychology* (Vol. 7, pp. 1–53). Chichester, England: Wiley.

Kanfer, R., & Ackerman, P. L. (1989). Motivation and cognitive abilities: An integrative/ aptitude-treatment interaction approach to skill acquisition. *Journal of Applied Psychology, 74*, 657–690.

Keil, C. T., & Cortina, J. M. (2001). Degradation of validity over time: A test and extension of Ackerman's model. *Psychological Bulletin, 127*, 673–697.

Kjerstad, C. L. (1919). Form of the learning curves for memory. *Psychological Monographs, 26*, 1–89.

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

Kozlowski, S. W. J., & Klein, K. J. (2000). A multilevel approach to theory and research in organizations: Contextual, temporal and emergent processes. In: K. J. Klein & S. W. J. Kozlowski (Eds), *Multilevel theory, research and methods in organizations: Foundations, extensions, and new directions* (pp. 3–90). San Francisco, CA: Jossey-Bass.

Kunst, E. J. (1941). Variations in work performance under normal industrial conditions. *Psychological Bulletin, 38,* 530.

Laird, N. M., & Ware, J. H. (1982). Random effects models for longitudinal data. *Biometrics, 38*, 963–974.

Lawler, E. E. (1986). *High involvement management.* San Francisco, CA: Jossey-Bass.

Lawrence, B. S. (1988). New wrinkles in the theory of age: Demography, norms, and performance ratings. *Academy of Management Journal, 31*, 309–337.

Levy, F. K. (1965). Adaptation in the production process. *Management Science, 11*, B136–B154.

Little, R. J. A., & Rubin, D. B. (2002). *Statistical analysis with missing data* (2nd ed.). Hoboken, NJ: Wiley.

MacKinney, A. C. (1967). The assessment of performance change: An inductive example. *Organizational Behavior and Human Performance, 2*, 56–72.

McDaniel, M. A., Schmidt, F. L., & Hunter, J. E. (1988). Job experience correlates of job performance. *Journal of Applied Psychology, 73*, 327–330.

McEvoy, G. M., & Beatty, R. W. (1989). Assessment centers and subordinate appraisals of managers: A seven-year examination of predictive validity. *Personnel Psychology, 42*, 37–52.

Milkovich, G. T., & Newman, J. M. (2005). *Compensation*. New York: McGraw Hill.

Motowidlo, S. J., Borman, W. C., & Schmitt, M. J. (1997). A theory of individual differences in task and contextual performance. *Human Performance, 10,* 71–83.

Murphy, K. R. (1989). Is the relationship between cognitive ability and job performance stable over time? *Human Performance, 2,* 183–200.

Murphy, K. R., & Schiarella, A. H. (1997). Implications of the multidimensional nature of job performance for the validity of selection tests: Multivariate framework for studying test validity. *Personnel Psychology, 50,* 823–854.

Muth, J. F. (1986). Search theory and the manufacturing progress function. *Management Science, 32,* 948–962.

Newman, D. A. (2003). Longitudinal modeling with randomly and systematically missing data: A simulation of ad hoc, maximum likelihood, and multiple imputation techniques. *Organizational Research Methods, 6*, 328–362.

Noe, R. A. (2005). *Employee training and development* (3rd ed.). New York: McGraw-Hill.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw Hill.

Owens, W. A. (1953). Age and mental abilities: A longitudinal study. *General Psychological Monographs, 48*, 3–54.

Pegels, C. C. (1969). On startup of learning curves: An expanded view. *AIIR Transactions, 1(3)*, 216–222.

Ployhart, R. E. (2004). Organizational staffing: A multilevel review, synthesis, and model. *Research in Personnel and Human Resources Management, 23*, 121–176.

Ployhart, R. E., & Hakel, M. D. (1998). The substantive nature of performance variability: Predicting interindividual differences in intraindividual performance. *Personnel Psychology, 51*, 859–901.

Ployhart, R. E., Holtz, B. C., & Bliese, P. D. (2002). Longitudinal data analysis: Applications of random coefficient modeling to leadership research. *Leadership Quarterly, 13*, 455–486.

Ployhart, R. E., Schneider, B., & Schmidt, N. (2006). *Staffing organizations: Contemporary practice and theory* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum.

Prien, E. P. (1966). Dynamic character of criteria: Organizational change. *Journal of Applied Psychology, 50*, 501–504.

Rambo, W. W., Chomiak, A. M., & Price, J. M. (1983). Consistency of performance under stable conditions of work. *Journal of Applied Psychology, 68*, 78–87.

Rambo, W. W., Chomiak, A. M., & Rountree, R. J. (1987). Temporal interval and the estimation of the reliability of work performance data. *Perceptual and Motor Skills, 64*, 791–798.

Raudenbush, S. W. (2001). Toward a coherent framework for comparing trajectories of individual change. In: L. M. Collins & A. G. Sayer (Eds), *New methods for the analysis of change* (pp. 35–64). Washington, DC: American Psychological Association.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models* (2nd ed.). Thousand Oaks, CA: Sage.

Ree, M. J., & Earles, J. A. (1992). Intelligence is the best predictor of job performance. *Current Directions in Psychological Science, 1*, 86–89.

Rhodes, S. R. (1983). Age-related differences in work attitudes and behavior: A review ad conceptual analysis. *Psychological Bulletin, 93*, 328–367.

Rothe, H. F. (1946). Output rates among butter wrappers: I. Work curves and their stability. *Journal of Applied Psychology, 30*, 199–211.

Rothe, H. F. (1947). Output rates among machine operators: Distributions and their reliability. *Journal of Applied Psychology, 31*, 484–489.

Rothe, H. F. (1978). Output rates among industrial employees. *Journal of Applied Psychology, 63*, 40–46.

Rothe, H. F., & Nye, C. T. (1958). Output rates among coil winders. *Journal of Applied Psychology, 42,* 182–186.

Rothe, H. F., & Nye, C. T. (1959). Output rates among machine operators: II. Consistency related to methods of pay. *Journal of Applied Psychology, 43*, 417–420.

Rothe, H. F., & Nye, C. T. (1961). Output rates among machine operators: III. A nonincentive situation in two levels of business activity. *Journal of Applied Psychology, 45*, 50–54.

Rotundo, M., & Sackett, P. R. (2002). The relative importance of task, citizenship, and counterproductive performance to global ratings of job performance: A policy-capturing approach. *Journal of Applied Psychology, 87*, 66–80.

Rovine, M. J., & Molenaar, P. C. M. (2001). A structural equations modeling approach to the general linear mixed model. In: L. M. Collins & A. G. Sayer (Eds), *New methods for the analysis of change* (pp. 67–96). Washington, DC: American Psychological Association.

Ryan, A. M., McFarland, L., Baron, H., & Page, R. (1999). An international look at selection practices: Nation and culture as explanations for variability in practice. *Personnel Psychology, 52*, 359–391.

Ryan, A. M., & Tippings, N. T. (2004). Attracting and selecting: What psychological research tells us. *Human Resource Management, 43*, 305–318.

Salgado, J. F. (1997). The five factor model of personality and job performance in the European community. *Journal of Applied Psychology, 82*, 30–43.

Salamin, A., & Hom, P. W. (2005). In search of the elusive U-shaped performance-turnover relationship: Are high performing Swiss bankers more liable to quit? *Journal of Applied Psychology, 90*, 1204–1216.

Salthouse, T. (1979). Adult age and speed accuracy trade-off. *Ergonomics, 22,* 811–821.

Sayer, A. G., & Cumsille, P. W. (2001). Second-order latent growth models. In: L. M. Collins & A. G. Sayer (Eds), *New methods for the analysis of change* (pp. 179–200). Washington, DC: American Psychological Association.

Schaie, K. W. (1994). The course of adult intellectual development. *American Psychologist, 49*, 304–313.

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin, 124*, 262–274.

Schmidt, F. L., Hunter, J. E., & Outerbridge, A. (1986). The impact of job experience and ability on job knowledge work sample performance, and supervisory rating of performance. *Journal of Applied Psychology, 71*, 432–439.

Schmidt, F. L., & Kaplan, L. B. (1971). Compositive vs. multiple criteria: A review and resolution of the controversy. *Personnel Psychology, 24*, 419–434.

Schmitt, N., Gooding, R. Z., Noe, R. A., & Kirsch, M. (1984). Meta-analyses of validity studies published between 1964 and 1982 and the investigation of study characteristics. *Personnel Psychology, 37,* 407–422.

Schneider, B. (1987). The people make the place. *Personnel Psychology, 40*, 437–453.

Smith, C. (1981). Effect of production rate on weapon system cost. *Concepts: The Journal of Defense Systems Acquisitions Management, 4*, 77–83.

Smith, P. C. (1976). Behaviors, results, and organizational effectiveness: The problems of criteria. In: M. D. Dunnette (Ed.), *Handbook of industrial and organizational psychology* (1st ed.). Chicago, IL: Rand McNally.

Stanford Research Institute. (1949). *Relationships for determining the optimum expansibility of the elements of a Peacetime Aircraft Procurement Program*. Palo Alto, CA: Stanford Research Institute.

Steele-Johnson, D., Osburn, H. G., & Pieper, K. F. (2000). A review and extension of current models of dynamic criteria. *International Journal of Selection and Assessment, 8*, 110–136.

Sterns, H. L., & Doverspike, D. (1989). Aging and the training and learning process. In: I. L. Goldstein (Ed.), *Training and development in organizations* (pp. 299–332). San Francisco, CA: Jossey-Bass.

Stevens, M. J., & Campion, M. A. (1999). Staffing work teams: Development and validation of a selection test for teamwork settings. *Journal of Management, 25*, 207–228.

Stewart, G. L., & Nandkeolyar, A. K. (2006). Adaptation and intraindividual variation in sales outcomes: Exploring the interactive effects of personality and environmental opportunity. *Personnel Psychology, 59*, 307–332.

Sturman, M. C. (2000). Implications of utility analysis adjustments for estimates of human resource intervention value. *Journal of Management, 26*, 281–299.

Sturman, M. C. (2003). Searching for the inverted U-shaped relationship between time and performance: Meta-analyses of the experience/performance, tenure/performance, and age/performance relationships. *Journal of Management, 29,* 609–640.

Sturman, M. C., Cheramie, R. A., & Cashen, L. H. (2005). The impact of job complexity and performance measurement on the temporal consistency, stability, and test-retest reliability of employee job performance ratings. *Journal of Applied Psychology, 90*, 269–283.

Sturman, M. C., & Trevor, C. O. (2001). The implications of linking the dynamic performance and employee turnover literatures. *Journal of Applied Psychology, 86*, 684–696.

Sturman, M. C., Trevor, C. O., Boudreau, J. W., & Gerhart, B. (2003). Is it worth it to win the talent war? Evaluating the utility of performance-based pay. *Personnel Psychology, 56*, 997–1035.

Teplitz, C. J. (1991). *The Learning Curve Deskbook: A reference guide to theory, calculations, and applications*. New York: Quorum Books.

Tett, R. P., Jackson, D. N., & Rothstein, M. (1991). Personality measures as predictors of job performance: A meta-analytic review. *Personnel Psychology, 44,* 703–742.

Thoresen, C. J., Bradley, J. C., Bliese, P. D., & Thoresen, J. D. (2004). The Big-Five personality traits and individual job performance growth trajectories in maintenance and transitional stages. Journal of *Applied Psychology, 89*, 835–863.

Thorndike, R. L. (1940). Constancy of the I.Q. *Psychological Bulletin, 37*, 167–186.

Thurstone, L. L. (1919). The learning curve equation. *Psychological Monographs, 26*, 1–51.

Trevor, C. O., Gerhart, B., & Boudreau, J.W. (1997). Voluntary turnover and job performance: Curvilinearity and the moderating influences of salary growth and promotions. *Journal of Applied Psychology, 82*, 44–61.

Tubre, T. C., & Collins, J. M. (2000). Jackson and Schuler (1985) revisited: A meta-analysis of the relationships between role ambiguity, role conflict, and job performance. *Journal of Management, 26,* 155–169.

Tziner, A., Ronen, S., & Hacohen, D. (1993). A four-year validation study of an assessment center in a financial corporation. *Journal of Organizational Behavior, 14*, 225–237.

Viswesvaran, C., Ones, D. S., & Schmidt, F. L. (2005). Is there a general factor in ratings of job performance? A meta-analytic framework for disentangling substantive and error influences. *Journal of Applied Psychology, 90*, 108–131.

Viswesvaran, C., Schmidt, F. L., & Ones, D. S. (1996). Comparative analysis of the reliability of job performance ratings. *Journal of Applied Psychology, 81*, 557–574.

Waldman, D. A., & Avolio, B. J. (1993). Aging and work performance in perspective: Contextual and developmental considerations. In: K. M. Rowland & G. R. Ferris (Eds), *Research in Personnel and Human Resources Management* (Vol. 11, pp. 133–162). Greenwich, CT: JAI Press.

Welbourne, T. M., Johnson, D. E., & Erez, A. (1998). The role-based performance scale: Validity analysis of a theory-based measure. *Academy of Management Journal, 41*, 540–555.

Willett, J. B., & Sayer, A. G. (1994). Using covariance structure analysis to detect correlates and predictors of individual change over time. *Psychological Bulletin, 116,* 363–381.

Williams, L. J., Edwards, J. R., & Vandenberg, R. J. (2003). Recent advances in causal modeling methods for organizational and management research. *Journal of Management, 29*, 903–936.

Williams, C. R., & Livingstone, L. P. (1994). Another look at the relationship between performance and voluntary turnover. *Academy of Management Journal, 37*, 269–298.

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

Wright, T. P. (1936). Factors affecting the cost of airplanes. *Journal of Aeronautical Science, 3*, 122–128.

Yelle, L. (1979). The learning curve: Historical review and comprehensive survey. *Decisions Sciences, 10*, 302–324.

Zhang, D., & Willson, V. L. (2006). Comparing empirical power of multilevel structural equation models and hierarchical linear models: Understanding cross-level interactions. *Structural Equation Modeling, 13,* 615–630.