Journal of Management
1997, Vol. 23, No. 6, 745-757

Using Hierarchical Linear Modeling to Examine Dynamic Performance Criteria Over Time

Diana L. Deadrick
Old Dominion University
Nathan Bennett
Louisiana State University
Craig J. Russell
University of Oklahoma

The selection literature has long debated the theoretical and practical significance of dynamic criteria. Recent research has begun to explore the nature of individual performance over time. This study contributes to this body of research through a hierarchical linear modeling analysis of dynamic criteria. The purpose of this study was to investigate the role of ability in explaining initial job performance, as well as the rate of improvement—or performance trend—among a sample of 408 sewing machine operators over a 24 week period. The results of a hierarchical linear modeling analysis suggest that ability measures are differentially related to initial performance and performance improvement trend.

Variability in employee performance over time, as evidenced by dynamic criteria, has been a source of concern in personnel selection. Early research suggested the existence of dynamic criteria (e.g., Bass, 1962; Ghiselli & Haire, 1960), which sparked a debate about the theoretical and practical significance of such a phenomenon (cf. Austin, Humphreys, & Hulin, 1989; Barrett & Alexander, 1989; Barrett, Caldwell, & Alexander, 1985; Henry & Hulin, 1987). Recent evidence indicates that the relative (rank-ordered) performance of individuals changes systematically over time (e.g., Deadrick & Madigan, 1990; Hanges, Schneider, & Niles, 1990; Hofmann, Jacobs, & Baratta, 1993; Hofmann, Jacobs, & Gerras, 1992; Henry & Hulin, 1987; Hulin, Henry, & Noon, 1990; Rambo, Chomiak & Price, 1983; Rambo, Chomiak, & Rountree, 1987; Rothe, 1978).

The above-mentioned research, together with evidence of changes in ability-performance relationships over time (see Ackerman, 1989; Murphy, 1989),

Direct all correspondence to: Diana L. Deadrick, Department of Management, Old Dominion University, Norfolk, VA 23529.

suggests the need for further research on dynamic criteria and intra-individual performance changes over time (Ackerman, 1989; Austin et al., 1989; Austin & Villanova, 1992; Austin, Villanova, Kane, & Bernardin, 1991; Borman, 1991; Murphy, 1989). This type of research should increase our understanding of what causes dynamic criteria (systematic versus random within-person performance change) and what causes inter-individual differences in performance changes (e.g., individual difference variables). The purposes of this article were to: (a) examine possible determinants of dynamic criteria; (b) describe an analytic technique that is appropriate for investigating determinants of dynamic criteria; and (c) present evidence pertaining to determinants of dynamic criteria.

Theoretical Development

Hofmann and associates (Hofmann, Jacobs, & Baratta, 1993; Hofmann, Jacobs, & Gerras, 1992) examined individual growth curves as a means of better understanding dynamic criteria. Hofmann et al. (1992) estimated growth curves for major-league baseball players and found different patterns of intra-individual performance. Specifically, they found changes in rank-ordered performance (dynamic criteria) were systematic and that there were subgroups of players with different patterns of change. In a later study, Hofmann et al. (1993) used hierarchical linear modeling (HLM) to further investigate inter-individual differences in performance change patterns among life insurance salespeople. HLM procedures identified three clusters of systematic change patterns. Although they were unable to examine possible determinants of inter-individual differences in performance change patterns, Hofmann et al. (1993) speculated that individual differences in abilities and/or goal orientations might explain the different change patterns. Based on their findings, Hofmann and associates argued that future research needs to be conducted that examines the determinants of inter-individual differences in performance trajectories.

Determinants of Performance Changes

In order to better understand the nature and causes of dynamic criteria, two related issues need to be examined. First, what accounts for the simplex-like pattern of performance intercorrelations associated with dynamic criteria? Although the simplex pattern indicates that individuals continually change their rank-ordered performance over time, is this change systematic or random? Second, what factors explain inter-individual differences in performance change patterns over time? Although performance change patterns imply systematic change, to what extent is the rate of change (i.e., performance improvement) predictable from knowledge about individual difference variables? Clearly, situational variables affect performance over time; however, this discussion focuses on the relative influence of individual differences in abilities and dispositional (e.g., motivation) variables. Hofmann and associates suggest underlying changes are systematic and vary across individuals. Conceptual models put forth by Murphy (1989) and Kanfer and Ackerman (1989) provide a basis for explaining possible causes of those performance changes.

Murphy (1989) presented a two-stage model of dynamic job performance. Transition stages of performance occur when employees are new to the job or when any of the major duties or responsibilities of the job change. Occasional changes in job demands and the job environment can trigger additional transition stages even though the job title and description have not changed. Because transition stages require learning new tasks and/or considerable judgement, performance should depend largely on cognitive ability. Maintenance stages occur when job tasks are well-learned and employees are no longer confronted with novel or unpredictable job demands. Performance during maintenance stages should, therefore, depend largely on dispositional (e.g., motivation, personality) variables versus cognitive ability factors (Murphy, 1989). Although all employees new to a job would start off in a transition stage, regardless of prior experience, the duration and frequency of transition stages will vary depending on the job (e.g., routinization), the individual (e.g., ability to learn and adapt), and the situation (e.g., quality of training and supervision).

Individual differences in performance stage (content and duration) may explain both changes in rank-order of performance over time and possible determinants of these performance changes. At any given time, some employees will be in transition stages and others will be in maintenance stages, regardless of job tenure. Therefore, the average performance trend will depend on the proportion of employees in transition and maintenance stages. However, inter-individual differences in performance change patterns are expected due to individual differences in performance stages and individual differences in ability and dispositional variables. The relative impact of ability and dispositional variables on performance over time will vary depending on the proportion of employees in transition or maintenance stages of performance. If the majority of employees are in transition stages of performance, individual differences in cognitive ability are expected to determine inter-individual differences in performance and rates of performance improvement. Individual differences in dispositional factors should determine inter-individual differences in performance change patterns if the majority of employees are in maintenance stages. The work of Kanfer and Ackerman (1989) provides a similar explanation for performance change patterns. Their model focuses on performance during different stages of skill acquisition and suggests that individual change patterns will vary depending on the stage of skill acquisition (i.e., declarative knowledge, knowledge compilation, or procedural knowledge) and individual differences in ability and motivation (i.e., the amount and allocation of attentional resources). During the declarative knowledge stage, employees are learning the job, and performance is slow and error-prone. Performance is "resource-dependent" at this stage, due to the high demand for cognitive capacity and attentional effort (motivation). During the procedural stage, employees have "automatized" task-related knowledge, rules, and procedures, which results in fast and accurate task performance that requires less attention. At this stage, performance is "resource-insensitive"; there is less demand for cognitive resources and attentional effort focused on the task. Therefore, psychomotor ability should be a primary determinant of performance during the procedural knowledge stage.

Kanfer and Ackerman's model (1989) is similar to Murphy's model (1989) in that each posits that performance changes over time are due to individual differences in performance (skill acquisition) stages and individual differences in ability and dispositional variables (motivation). Therefore, the overall performance trend among a group of employees will depend on the proportion of employees in the different stages of skill acquisition, and inter-individual differences in performance trends (change patterns) will reflect individual differences in skill acquisition, ability, and motivation.

While Murphy's model addresses general cognitive abilities, Kanfer and Ackerman's model makes predictions about the relative importance of specific ability factors over time. When the majority of employees are in the declarative knowledge stage of skill acquisition, individual differences in cognitive ability should be a primary determinant of inter-individual differences in performance improvement. If performance at this stage is in fact "resource-dependent", then you would expect that: (a) motivation will be high; (b) performance changes due to individual differences in ability will be more pronounced; and (c) individual differences in cognitive ability will determine inter-individual differences in performance improvement more so than individual differences in psychomotor ability (cf. Kanfer & Ackerman, 1989). When the majority of employees are in the procedural knowledge stage, individual differences in psychomotor ability should determine inter-individual differences in performance improvement.

Using Hierarchical Linear Modeling to Study Change

One technique that can be applied to the study of within-person, time-related phenomena is hierarchical linear modeling (HLM). As noted elsewhere (cf. Bryk & Raudenbush, 1987, 1992), longitudinal data are implicitly multilevel and nested, though they are rarely treated as such. HLM provides what Bryk and Raudenbush (1992) call an "integrated approach for studying the structure and predictors of individual growth" (p. 131). In the context of dynamic criteria, it provides a means of examining the existence, nature, and causes of within-person performance changes over time. As a result, HLM allows for a more complete analysis of dynamic criteria: (a) it explicitly recognizes and investigates systematic individual change patterns over time, (b) it provides for the estimation of both static and longitudinal performance parameters (i.e., intercept and slope), and (c) it enables analyses of both within- and between-person performance change patterns.

Several applications of HLM to examine multilevel models have recently appeared in the management literature (cf. Hofmann et al., 1993; Kidwell, Mossholder & Bennett, 1997; Vancouver, Millsap, & Peters, 1994). When applying HLM to the study of individual performance over time, the focus of the level-1 and level-2 analysis shifts, as does some of the notation. First, the level-1 model is a within-person model that examines the nature of intra-individual performance over time. The independent variable at level-1 is performance and the dependent variable is a time vector. Estimated in this level-1 model are intercept and slope parameters for each individual. The intercept parameter (π_{0i}) represents initial performance; the slope parameter (π_{1i}) represents the performance trend. Second, the level-2 model is a between-person model that examines the relationships between individual characteristics (e.g., ability, experience) and the intercept and slope parameters estimated in level-1.

The Present Study

The purpose of this study was to replicate the study by Hofmann et al. (1993) and to extend it through an examination of individual difference variables that might explain inter-individual differences in performance change patterns over time. This study also builds upon a previously-published study on dynamic criteria (Deadrick and Madigan, 1990) in that it re-analyzes the data in terms of the underlying individual performance change patterns. In their study, Deadrick and Madigan (1990) found evidence for dynamic criteria; there was a steady decline in performance intercorrelations over time for a sample of sewing machine operators. In the present study, we used HLM to extend their findings about dynamic criteria to include an examination of performance change patterns.

There were two questions of interest in this study. First, *are there inter-individual differences in performance change over time?* We know that a portion of variation in performance change over time is due to random influences (cf. Alexander, Barrett & Doverspike, 1991). Our question, however, addresses that portion due to non-random influences. The participants in this study consisted of newly-hired sewing machine operators, and most of them had no previous sewing experience. The only training provided by the organization was on-the-job training, and the company expected inexperienced (as well as experienced) operators to achieve production proficiency by the end of the first 12 weeks on the job. This situation gave us the opportunity to examine learning curves across the sample of operators. To the degree that most operators were in a transition (declarative knowledge) stage of performance, we expect to find an overall linear trend in performance that reflects an aggregated learning curve. Because operators varied in terms of their ability and prior sewing experience, thus their stage of performance (skill acquisition), we expect to find inter-individual differences in performance change patterns.

Second, *do individual differences in ability account for any observed inter-individual differences in performance change patterns?* The organization under study was characterized by a stable environment; i.e., stable and routine technology, standardized work procedures and job requirements, constant piece-rate standards and production goals. In addition, the piece-rate system that operators worked under provided an additional control for differences in motivation and goals; this tightly-linked pay-for-performance system should produce constant (high) levels of motivation and constant (maximization) performance goals across all operators. Based on these features of the organization and job context, we had the opportunity to somewhat isolate the effect of individual differences in ability as a means of explaining inter-individual differences in performance change patterns. We, therefore, expect individual differences in ability to be a significant determinant of any inter-individual differences in change patterns. To the extent that most operators were in fact in a transition (declarative knowledge) stage of performance, we expect cognitive (versus psychomotor) ability to be the primary determinant of inter-individual differences in performance change over time (i.e., rate of improvement).

Method

Data

The data were collected as part of a predictive validation study conducted for the Virginia Employment Commission to evaluate the General Aptitude Test Battery (GATB). Additional analyses of the data are reported in Deadrick and Madigan (1990).

Participants

The sample consisted of 408 sewing machine operators employed at five garment manufacturing plants operating in the southeastern United States. The five plants were owned by the same company, manufactured the same kind of garment, used similar equipment and production procedures, and operated under a uniform system of management policies, practices, and record-keeping procedures. Due to these similarities, the analyses reported here include employees from all five plants. Demographically, the sample can be described as follows: 100% were female, 32% were black, the average age was 26 years with a standard deviation of 8.47 years.

Measures

Performance. Operator performance was measured using average hourly production (piece-rate) earnings per week (i.e., total production earnings per week, divided by the number of hours actually worked during the week). These data were obtained from company records and did not include the minimum-wage guarantee, rework time, or time-not-worked. Earnings were determined by piece rate and actual production; piece rate was determined by industrial engineering studies conducted by the company. The data were coded without decimal points, so this measure should be interpreted as the mean cents per hour for the focal week.

As we were interested in modeling performance trends, the performance measures analyzed here were weekly production earnings of new hires in their first 24 weeks. The company allowed a 12-week on-the-job training period for all new hires, after which time they expected all operators to be proficient. Thus, the period included in these data encompass the 12 week learning period, as well as an additional 12 weeks of experienced performance. The 408 employees are the subset of those considered in earlier work (Deadrick & Madigan, 1990) for whom complete performance data over their first 24 weeks on the job were available. Because the time vector has been coded 0 to 23, the intercept can be interpreted as an individual's initial performance.

Ability. We included three measures of ability in our analysis. First, *cognitive ability* was computed as the raw score composite of the General, Verbal, and Numerical (G, V, N) aptitude scales of the GATB. Second, *psychomotor ability* was the raw-score composite of the Coordination, Finger Dexterity, and Manual

Dexterity (K, F, M) scales of the GATB. Third, previous sewing *experience* was included as a control variable. Experience was dummy coded (0,1) such that a '1' indicated the employee had previous sewing experience.

Analysis

Analyses were performed using the HLM2L statistical package (Bryk, Raudenbush, & Congdon, 1992). Here, the level-1 model portrayed each individual's performance over a 24 week period as a linear trend. An individual's performance trend can be described using estimated intercept and slope parameters (Bryk & Raudenbush, 1992). These parameters were then treated as dependent variables in the level-2 analysis, where they were predicted using the ability measures and experience.

Results

The correlation matrix and descriptive statistics for the level-2 predictor variables, and a sampling of the weekly performance measures are reported in Table 1. The data on job performance over time revealed a simplex pattern, as was true for the full sample reported in Deadrick and Madigan (1990). Alexander, Barrett, and Doverspike (1991) demonstrated that simplex patterns (i.e., decreasing $x \rightarrow y$ correlations as y is measured further out in time) occur by definition when y_t is a function of y_{t-1} and random error.

Assessing the Appropriateness of the Linear Trend

The literature contains support for the contention that a linear performance trend should describe the data well (cf. Dawes, 1979; Deadrick & Madigan, 1990). At the same time, there may be other viable models of the performance (e.g., quadratic, cubic) (Hofmann et al., 1992, 1993; Rambo et al., 1983). To provide an estimation of how well a linear trend captured within-person variance in performance, we estimated a null model — that is, a model with no level-1 predictors. Then, we estimated a model with a linear time vector. By comparing

Table 1. Descriptive Statistics

	x	*sd*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*
1. Experience	.34	.47	1.00							
2. Cognitive ability	272.28	34.35	.01	1.00						
3. Psychomotor ability	324.57	45.74	−.02	.26	1.00					
4. Performance, week 1	199.27	102.58	.36	.08	.14	1.00				
5. Performance, week 6	321.99	116.69	.27	.07	.19	.79	1.00			
6. Performance, week 12	384.30	113.85	.21	.12	.19	.70	.82	1.00		
7. Performance, week 18	418.02	113.85	.21	.16	.19	.60	.73	.83	1.00	
8. Performance, week 24	447.69	118.66	.15	.14	.19	.55	.67	.75	.81	1.00

Notes: $n = 408$. A correlation of ± .11 is significant at $p < .05$, of ± .18 at $p < .01$, of ± .26 at $p < .001$. A correlation matrix with complete weekly performance data is available from the authors on request.

Table 2. Results of Hierarchical Linear Modeling Analysis—Unconditional Model

Fixed effects	*Coefficient*	*Standard error*	*t*	*p*
Mean initial performance, β_{00}	266.21	5.61	46.33	.000
Mean performance trend, β_{10}	9.30	.23	40.03	.000
Random effects	*Variance Component*	*df*	χ^2	*p*
Initial performance, r_{0i}	12477.87	407	13286.13	.000
Performance trend, r_{1i}	19.81	407	4091.54	.000
Level-1 error, e_{ti}	2516.93			
Reliability of OLS estimates				
Initial performance, π_{0i}	.97			
Performance trend, π_{1i}	.90			

Note: $n = 408$

these two models, it was possible to compute the R^2 for the reduction in within-person variance for a linear trend (Bryk & Raudenbush, 1992). In these data, 64% of the within-person variance was explained by the linear trend. This suggests, as we expected, that a linear trend accounted for a substantial amount of within-person variance in these data. Based on this result, we continued our analysis using the linear trend.

It should be noted that the null model also provides an estimate of the amount of variance that lies between individuals. This is computed as the between group variance divided by the total variance. In these data, 55% of the variance in performance was between individuals.

Random Coefficient Regression Model

Our next step in conducting the HLM analysis was to estimate an unconditional model; that is, a model containing no level-2 predictors. As Bryk and Raudenbush (1992) noted, this model provides several initial findings; in the context of our analyses it reports the average initial performance (intercept) and the average performance trend (slope) across individuals, as well as important baseline information for subsequent analyses. The results of this analysis are presented in Table 2.

The top portion of Table 2 presents the fixed effects results for the unconditional model. The estimated mean intercept, β_{00}, was 266.91; the estimated mean growth rate—or slope—β_{10}, was 9.30. This indicates that the HLM estimate of the average earnings in the first week was just under 267 cents per hour and that the estimated mean increase in earnings per week was just over 9 cents per hour. The significant *t*-ratios for each parameter suggest that each is necessary for describing the mean individual performance trend.

The next section of Table 2 reports the variance components for the random effects. These parameters were central to the investigation of the nature of the deviations of individual performance trends from the mean performance trend. The estimates of the variances for the initial performance and performance trend parameters, π_{0i} and π_{1i}, were 12477.87 and 19.81, respectively. For each of these parameters, the HLM analysis includes a χ^2 test for the null hypothesis that there is no true variation in these parameters. As we expected, both χ^2 tests were significant, suggesting that we can reject each null hypothesis and conclude that employees do, in fact, differ in both their initial performance levels and their performance trends. The variance estimate for the slope parameter (19.81) yields a standard deviation of 4.45. This indicates that an employee whose growth rate is one standard deviation above average is expected to increase his/her performance at the rate of 13.75 (9.30 + 4.45) cents per hour each week.

The fourth result from the unconditional model involves the reliabilities of the estimated slope and intercept parameters. These reliability estimates represent the proportion of between group variance that is systematic (i.e., that can be modeled in the level-2 equation using between-person variables). It is important that a reasonable amount of the variability in these parameters be reliable; if significant amounts of this variation are due to error, it would be difficult to model these parameters with between-person measures. The estimated reliabilities for the intercept and slope parameters were .97 and .90, suggesting that modeling these as a function of our level-2 variables is worthwhile.

Finally, the unconditional model produces an estimate of the correlation between true initial status (π_{0i}, intercept) and true performance trend (π_{1i}, slope). In this dataset, this correlation is –.43, suggesting that employees who produce at a lower level initially tend to increase their performance at a higher rate than those who initially produce at a higher level.

The Intercept- and Slopes-as-Outcomes Model

The model used to test our predictions about the impact of individual differences on inter-individual differences in performance change patterns is commonly referred to as the intercept- and slopes-as-outcomes model. Here, ability variables were used to model the intercept and slope parameters estimated at level-1. The results are presented in Table 3. Examination of the results for initial status (intercept, π_{0i}), reported in the top part of Table 3, indicate that experience and psychomotor ability were positively associated with an individual's initial performance on the job. Specifically, higher initial performance levels were associated with individuals who had previous sewing experience and who had greater psychomotor ability as measured by the GATB.

The results for performance trend (slope, π_{1i}) are presented in the lower portion of Table 3. The pattern of findings here suggests that individuals' rate of performance improvement was negatively associated with experience and positively associated with cognitive ability. Specifically, individuals with previous sewing experience improved more slowly than those without experience, and individuals with greater cognitive ability improved more quickly than individuals with less cognitive ability. This finding provides support for our prediction that cognitive ability would be a more significant determinant of performance than psychomotor ability.

Table 3. Results of Hierarchical Linear Modeling Analysis-Conditional Model

Fixed effects	*Coefficient*	*Standard error*
Model for initial status, π_{0i}		
Intercept, β_{00}	85.55	50.36
Experience, β_{01}	77.61***	11.18
Cognitive ability, β_{02}	.06	.16
Psychomotor ability, β_{03}	.41***	.12
Model for performance trend, π_{1i}		
Intercept, β_{10}	5.63**	2.17
Experience, β_{11}	−2.03***	.48
Cognitive ability, β_{12}	.02*	.00
Psychomotor ability, β_{13}	.00	.01

Notes: $n = 408$. * $p < .05$, ** $p < .01$, *** $p < .001$.

Table 4. Variance Explained in Initial Performance and Performance Trend

Model	*Initial performance, π_{0i}*	*Performance trend, π_{1i}*
Unconditional	12477.87	19.81
Conditional on Level-2 predictors	10881.51	18.78
Percent variance explained	12%	5%

Note: $n = 408$

Table 4 reports the variance in initial performance and performance trend that is explained by the person-characteristics considered in the analysis. These estimates were computed by comparing the random effects variance components in the unconditional model (Table 2) to those obtained in the model that includes person-characteristics as predictors. The variance explained statistic is the ratio of total parameter variance (e.g., the unconditional model) less residual parameter variance (e.g., the conditional model) over the total parameter variance. As Table 4 indicates, the ability measures explained 12% of the variance in initial performance and 5% of the variance in performance trend.

Discussion

Overall, our findings support the predictions drawn from the Murphy (1989) and Kanfer and Ackerman (1989) models. Analyses revealed, for this sample of newly-hired sewing machine operators, that: (a) there was a systematic linear change in performance over time; (b) there were inter-individual differences in performance trends (i.e., those operators with low levels of initial performance exhibited higher rates of performance improvement over time); and (c) individual differences in cognitive ability and experience were significant determinants of inter-individual differences in the rate of performance improvement over time. Although we made no predictions regarding the role of previous experience, the results show that it was clearly related to both initial performance and performance improvement. Taken together, these findings suggest that the majority of operators were in a transition (declarative knowledge) stage of performance and individual differences in cognitive ability influenced skill acquisition and rate of performance improvement.

From a practical perspective, our results suggest that the determinants of initial performance differ from the determinants of performance improvement. This finding supports the changing-subject model of dynamic performance: the abilities that determine performance over time change. Whereas psychomotor ability initially explained differences in the level of performance, cognitive ability was a relatively stronger predictor of differences in learning and performance improvement. Interestingly, prior experience was a significant determinant of both initial performance and performance improvement. Although we made no predictions about this variable, our results suggest that prior sewing experience had a stable influence on sewing machine operator performance over time. Experienced operators initially performed at higher levels of performance, and, over time, showed less performance change. This finding suggests that experienced operators were in a procedural knowledge (maintenance) stage of performance, characterized by diminishing performance improvements (cf. Kanfer & Ackerman, 1989).

From a theoretical perspective, the findings highlight a need for job performance theories and research that focus on the phenomenon of individual performance change. Although numerous researchers have made such pleas (Ackerman, 1989; Austin et al., 1989, 1991; Austin & Villanova, 1992; Borman, 1991; Deadrick & Madigan, 1990; Hofmann et al., 1992, 1993; Murphy, 1989), this aspect of the criterion domain has been largely ignored. The current findings underscore the importance of refocusing our efforts. Although individual differences in ability and experience were significant determinants of inter-individual differences in performance change patterns, these factors accounted for only 5% of the variance in the rate of performance change over time. Clearly, there are important yet unmeasured variables that moderate performance change.

The job environment in this study was relatively stable in terms of situational and motivational influences on performance. However, Murphy's model (1989) posits that transition stages of performance can be triggered by not only structural changes in the job but also perceptual changes in how employees perceive, thus interact with, the job environment. Future research needs to examine the impact of both job environment and dispositional factors, as well as person-environment interactions, on intra-individual and inter-individual performance change patterns (cf. Prien, 1966).

In sum, we agree with Hofmann and associates (1992, 1993) that studies of systematic performance changes over time will permit new insight into the traditional concerns with dynamic criteria. This study sheds some light on the determinants of inter-individual differences in performance patterns and trajectories for sewing machine operators. Future research should incorporate both changing-task and changing-subject models of dynamic criteria and examine the relative impact of individual differences variables and situational (job environment) variables.

Hierarchical linear modeling seems to have promise as a tool for management researchers interested in studying change over time. Recently, two articles appeared in this journal that together provide a useful critical review of the manner in which management researchers have previously treated (or mistreated) time in their work (Bergh, 1993a, 1993b). Several mistakes seemed fairly prevalent in Bergh's review of the management literature. For example, researchers often either employed inappropriate analytic tools or improperly employed appropriate analytic tools. Clearly, the onus is on the researcher to avoid these two pitfalls. Additionally, Bergh found that researchers often failed to consider changes in relationships among variables over time. As noted above and elsewhere (Hofmann, this volume), this is the particular appeal of HLM in longitudinal research. Finally, Bergh noted that researchers often failed to address violations in statistical assumptions underlying their data. No tool can itself address this shortcoming. As with other techniques for conducting analyses on longitudinal data, HLM is not without limitations (e.g., Bryk & Raudenbush, 1992; Hofmann, this volume). Concerns about issues common in longitudinal research (e.g., Kelly & McGrath, 1988; Podsakoff & Dalton, 1987), such as autocorrelation, are not solved simply by adopting a multilevel approach to the data.

Acknowledgment: An earlier version of this paper was presented at the 1995 Academy of Management Meeting as part of a symposium, titled 'Introduction, Exploration, and Illustrations of Hierarchical Linear Modeling as Management Research Tool,' chaired by Nathan Bennett. The authors would like to acknowledge the helpful comments of David Hofmann, Kevin Mossholder, and Larry James.

References

Ackerman, P.L. (1989). Within-task intercorrelations of skilled performance: Implications for predicting individual differences? (A comment on Henry & Hulin, 1987). *Journal of Applied Psychology, 74:* 360–364.

Alexander, R.A., Barrett, G.V. & Doverspike, D. (April,1991). Appropriate (and inappropriate) data and analyses relevant to the dynamic criteria question. In R.J. Hanges (Chair) Dynamic criteria: Are we shooting at a moving target? Symposium conducted at the Sixth Annual Conference for the Society of Industrial and Organizational Psychology, St. Louis, MO.

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.

Austin, J.T. & Villanova, P., Kane, J.S. & Bernardin, H.J. (1991). Construct validation of performance measures: Definitional issues, development, and evaluation of indicators. In Roland & Ferris (Eds.), *Research in Personnel and Human Resource Management, 9*: 159–233.

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., Caldwell, M.S. & Alexander, R.A. (1985). The concept of dynamic criteria: A critical reanalysis. *Personnel Psychology, 15:* 93–97.

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

Bergh, D.D. (1993a). Watch the time carefully: The use and misuses of time effects in management research. *Journal of Management, 19:* 683–705.

Bergh, D.D. (1993b). Don't 'waste' your time! The effects of time series errors in management research: The case of ownership concentration and research and development spending. *Journal of Management, 19:* 897–914.

Borman, W.C. (1991). Job behavior, performance, and effectiveness, Pp. 271–326. In MD Dunnette & LM Hough (Eds.), *Handbook of Industrial and Organizational Psychology, 2nd edition, Vol. 2.* Palo Alto: Consulting Psychologists Press, Inc.

Bryk, A.S. & Raudenbush, S.W. (1987). Application of hierarchical linear models to assessing change. *Psychological Bulletin, 101:* 147–158.

Bryk, A.S. & Raudenbush, S.W. (1992). *Hierarchical linear models: Applications and data analysis methods.* Newbury Park, CA: Sage.

Bryk, A.S., Raudenbush, S.W. & Congdon, R.J. (1992). *Hierarchical linear models: Applications and data analysis methods.* Newbury Park, CA: Sage.

Dawes, R.M. (1979). The robust beauty of improper linear models in decision making. *American Psychologist, 34:* 571–582.

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

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

Hanges, P.J., Schneider, B. & Niles, K. (1990). Stability of performance: An interactionist perspective. *Journal of Applied Psychology, 78:* 194–204.

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.

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

Hofmann, D.H., 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.

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.

Kelly, J.R. & McGrath, J.E. (1988). *On time and method.* Newbury Park, CA: Sage.

Kidwell, R.E., Mossholder, K.W. & Bennett, N. (1997). Cohesiveness and organizational citizenship behavior: A multi-level analysis using work groups and individuals. *Journal of Management, 23:*6.

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

Podsakoff, P.M. & Dalton, D. R. (1987). Research methodology in organizational studies. *Journal of Management, 13:* 419–441.

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.I. (1987). Temporal intervals and the estimation of the reliability of work performance data. *Perceptual and Motor Skills, 64:* 791–798.

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

Vancouver, J.B., Millsap, R.E. & Peters, P.A. (1994). Multilevel analysis of organizational goal congruence. *Journal of Applied Psychology, 79:* 666–679.