Psychology and Aging
2005, Vol. 20, No. 4, 623–633

0882-7974/05/$12.00 DOI: 10.1037/0882-7974.20.4.623

Cognitive Performance Inconsistency: Intraindividual Change and Variability

Nilam Ram
University of Virginia

Patrick Rabbitt
University of Oxford

Brian Stollery
University of Bristol

John R. Nesselroade
University of Virginia

Although many studies have examined inconsistency of cognitive performance, few have examined how inconsistency changes over time. 91 older adults (age 52 to 79) were tested weekly for 36 consecutive weeks on a series of multitrial memory speed (i.e., letter recognition) tasks. A number of multivariate techniques were used to examine how individuals' level of inconsistency changed across weeks and how this change was related to interindividual differences in age and intelligence. Results indicated that (a) inconsistency of performance is a construct separate from the underlying performance ability (i.e., memory speed); (b) inconsistency reduces exponentially with practice; (c) individuals with higher scores on tests of fluid general intelligence (G_f) reached lower asymptotic levels of inconsistency compared to lower scorers; and (d) after controlling for the systematic effects of practice, variability in inconsistency from week-to-week was more pronounced for individuals with lower G_f scores compared to individuals with higher scores.

Keywords: multivariate, memory speed, aging, growth curves, nonlinear

Intraindividual variability in cognitive performance has sometimes been conceptualized as inconsistency (e.g., Fuentes, Hunter, Strauss, & Hultsch, 2001; Hultsch, MacDonald, Hunter, Levy-Bencheton, & Strauss, 2000; Li & Lindenberger, 1999; MacDonald, Hultsch, & Dixon, 2003; Stuss, Pogue, Buckle, & Bondar, 1994; Stuss, Stethem, Hugenholtz, Picton, & Richard, 1989). Defined as the variability of performance across occasions (e.g., Hultsch & MacDonald, 2004), inconsistency can be indexed or measured using the intraindividual standard deviation (ISD) of performance computed across occasions (Slifkin & Newell, 1998). Individuals whose performance varies little from occasion to occasion, regardless of level, are *consistent* (i.e., they have relatively low ISDs), whereas those individuals whose performance varies dramatically from occasion to occasion are *inconsistent* (i.e., they have relatively high ISDs). In this manner, intraindividual variability across occasions can be indexed with a single score. This inconsistency score can then be examined in the same manner as any other intraindividual or interindividual difference variable.

Nilam Ram and John R. Nesselroade, Department of Psychology, University of Virginia; Patrick Rabbitt, Department of Experimental Psychology, University of Oxford; Brian Stollery, Department of Experimental Psychology, University of Bristol.

Nilam Ram gratefully acknowledges the support provided by Grant T32 AG20500 from the National Institute on Aging in the preparation of this article. Special thanks to those at the Institute for Developmental and Health Research Methodology at the University of Virginia for helpful comments on earlier versions of this work.

Correspondence concerning this article should be addressed to Nilam Ram, Department of Psychology, P.O. Box 400400, University of Virginia, Charlottesville, VA 22904-4400. E-mail: nilam@virginia.edu

At a theoretical level, inconsistency of performance on relatively simple cognitive tasks has been considered a measure of variability in central nervous system functioning (Hendrickson, 1982; Hultsch & MacDonald, 2004; Li & Lindenberger, 1999). Consistent with this hypothesis, inconsistency (or intraindividual variability) in cognitive performance has been found to be related to age, injury, health, and intelligence (Hultsch & MacDonald, 2004). Generally, older adults tend to be more inconsistent in their performances than younger adults (e.g., Anstey, 1999; Fozard, Vercruyssen, Reynolds, Hancock, & Quilter, 1994; Hertzog, Dixon, & Hultsch, 1992; Hultsch, MacDonald, Hunter, Maitland, & Dixon, 2002; West, Murphy, Armilio, Craik, & Stuss, 2002), unhealthy or dysfunctional persons tend to be more inconsistent than healthy or functional persons (e.g., Fuentes et al., 2001; Spieler, Balota, & Faust, 1996; Stuss et al., 1994, 1989), and persons with lower levels of cognitive task performances (e.g., measures of G_f and G_c) tend to be more inconsistent than persons with higher levels of performance (e.g., Hultsch et al., 2002; Li, Aggen, Nesselroade, & Baltes, 2001; Rabbitt, Osman, Moore, & Stollery, 2001). In sum, on a variety of fronts, greater inconsistency seems to be a marker of impending decline or low functionality (Hendrickson, 1982; Li & Lindenberger, 1999; Rowe & Kahn, 1985, 1997).

In most studies, inconsistency has been studied as the day-to-day or week-to-week variability in performance. However, inconsistency can be measured across any time frame (Hultsch & MacDonald, 2004; Slifkin & Newell, 1998). For instance, an ISD calculated across multiple trials can be used to measure intraindividual variability or inconsistency in the moment-to-moment fluctuations in performance over a short time period. Similarly an ISD calculated across observations gathered days or weeks apart can be used to measure inconsistency (or intraindividual variability) in performance over a longer time period.

In this study, we examined inconsistency in memory speed performance in older adults using a series of multitrial letter recognition tasks (Rabbitt et al., 2001). Inconsistency was operationalized as individuals' intraindividual variability in performance across trials (i.e., moment-to-moment or within-session variability). We then examined within-person changes in inconsistency across weeks (i.e., between-session changes in inconsistency).

A number of writers (e.g., Cattell, 1957; Horn, 1972; Hultsch & MacDonald, 2004; Nesselroade & Featherman, 1997; Wohlwill, 1973) have discriminated between types of within-person changes. Fiske and Rice (1955), for example, distinguished between reactive or adaptive variability — those changes that are ordered in some fashion (e.g., cycles and oscillations) and spontaneous variability — those changes that do not show any systematic trends over time (i.e., the ordering of occasions is immaterial). Nesselroade (1991) distinguished between *intraindividual change* — those changes that may or may not be reversible and that may or may not be synchronous across individuals — and *intraindividual variability*–those changes that are more or less reversible and that may or may not be synchronous across individuals. Generally, intraindividual change is conceptualized as long-term lasting change, whereas intraindividual variability is conceptualized as short-term or transient fluctuation.

Intraindividual Change

Intraindividual change refers to within-person change that is enduring and characterizes the types of changes seen in a construct (e.g., cognitive ability) as a result of learning, development, or aging (Nesselroade, 1991). For example, the development and decline of fluid intelligence (G_f) over an individual's life span, characterized by a rapid increase in abilities during the first 20 years followed by a steady decline over the remainder of life, would constitute a pattern of intraindividual change. Similarly, on a shorter time scale, the learning that occurs with practice can also be characterized as intraindividual change, which is usually best described by exponential or power functions (e.g., Thurstone, 1919; Heathcote, Brown, & Mewhort, 2000).

One common method for studying intraindividual change is growth curve analysis (e.g., Bryck & Raudenbush, 1987, 1992; Rogosa & Willet, 1985; Singer & Willett, 2003; Wishart, 1938). Generally, this type of analysis is used to describe, test hypotheses, and make inferences about time-related phenomena, that is, change. By allowing specific parameters in the growth expression to vary between individuals, we can examine differences in persons' initial levels of performance (intercept), rates of improvement over time, asymptotic levels of performance, and so forth. In such manner, many researchers have fruitfully examined interindividual differences in intraindividual change across a wide variety of domains (see McArdle & Nesselroade, 2003, for a history).

In the present study, we hypothesized that as individuals gained more and more exposure to the tasks the quality of their performances would improve; they would become more consistent. Other studies of inconsistency have found average levels of inconsistency to decrease across multiple testing sessions (e.g., Hultsch et al., 2000; West et al., 2002). Our interest was in how inconsistency changes over time at the individual level. Using growth curve analysis methods, we examined the interindividual differences in intraindividual change in inconsistency and we investigated what other personal characteristics were related to such differences.

Intraindividual Variability

In addition to the systematic, long-term, lasting intraindividual changes noted previously, individuals' level of inconsistency may also exhibit intraindividual variability or systematic transient fluctuation. Intraindividual variability has been characterized in a number of ways, including error, wobble, lability, instability, inconsistency, and noise (Butler, Hokanson, & Flynn, 1994; Hultsch & MacDonald, 2004; Hultsch et al., 2000; Li & Lindenberger, 1999; Nesselroade, 1988; Nesselroade & Ford, 1985; Shammi, Bosman, & Stuss, 1998; Shoda, Mischel, & Wright, 1994; Slifkin & Newell, 1998). For example, Nesselroade and Ford (1987) suggested that the transient and relatively rapid changes that characterize the variation within an individual might represent the steady-state hum or base condition of daily functioning. Instead of being considered as random errors in performance, an individual's variability in performance across time is a useful and informative interindividual difference construct. For instance, levels of intraindividual variability in performance have been found to be predictive of impending cognitive developmental transitions (see, e.g., Siegler, 1994). Similarly, differences in the intraindividual variability in infants' heart rate are predictive of later differences in temperament (Fox & Porges, 1985; Kagan, 1994). Among adults, the amount of variability in self-esteem (i.e., self-esteem lability) is predictive of depression proneness (Butler et al., 1994) and, among older adults, the level of week-to-week variability in internality beliefs is a risk factor for mortality some 5 years later (Eizenman, Nesselroade, Featherman, & Rowe, 1997). This research indicates that patterns or amounts of intraindividual variability may be related to age, health, vulnerability, and ultimately death (e.g., Rowe & Kahn, 1985, 1997). Such findings highlight the importance of analyzing interindividual differences in intraindividual variability.

The study of intraindividual variability usually proceeds along two avenues. First, the intraindividual variability can be examined for short-term patterns of change (Nesselroade, 2002; Slifkin & Newell, 1998). In much the same manner that we extract patterns of long-term intraindividual change from occasion-to-occasion variance (e.g., using growth curve analysis), we can also extract meaningful (i.e., interpretable) patterns of more transient changes. For instance, Horn (1972) identified patterns of week-to-week intraindividual variability that reflect the fluid-crystallized intelligence distinction. Similarly, Hampson (1990) demonstrated the existence of short-term patterns in performance fluctuations that appear to be hormonally driven. Such findings illustrate that meaningful and informative short-term statelike patterns of intraindividual variability can and do exist alongside the more traditional long-term intraindividual change patterns. Second, after accounting for systematic short- and long-term changes, we can quantify the amount of within-person occasion-to-occasion variability. This can be done in a number of ways, such as calculating the ISD across occasions, the distance between high and low scores, coefficient of variation, and so forth (Hultsch & MacDonald, 2004; Slifkin & Newell, 1998). Such scores are then examined for interindividual differences.

In this study, intraindividual variability across weeks was examined along both of the avenues outlined previously. First, the structure of the intraindividual variability was examined for patterns. Exploratory time-series analysis techniques were used to obtain information about and to describe any systematic statelike patterns that existed in the data. Second, the gross amount of intraindividual variability across weeks was quantified and examined as an interindividual difference variable.

Note that our gross measure of intraindividual variability differs from our measures of inconsistency in time frame only (week-to-week vs. trial-to-trial). The mechanics of the calculations are identical. In fact, inconsistency is intraindividual variability. Here, however, we make a distinction between intraindividual variability across trials (inconsistency) and intraindividual variability across weeks (intraindividual variability) for the sake of readability. This distinction highlights a key study design feature. In this study, data were simultaneously collected along two time scales (Newell, Liu, & Mayer-Kress, 2001) in a compressed *measurement burst design* (Nesselroade, 1991). Within-burst assessments were obtained on a moment-to-moment (i.e., trial-to-trial) time scale and were used to construct a measure of inconsistency. Across-burst measurements were obtained on a week-to-week time scale and were used to examine how inconsistency changed and varied from occasion to occasion.

Summary of Purposes

In this article, we examine how individuals performed on a set of memory speed tasks over successive trials and weeks to identify and explain interindividual differences in the intraindividual change and intraindividual variability of inconsistency. Our specific goals were the following:

1. Establish a multivariate measurement model of inconsistency.
2. Identify intraindividual change in inconsistency (i.e., practice effects).
3. Identify interindividual differences in intraindividual change.
4. Identify intraindividual variability of inconsistency.

a. Identify any systematic statelike patterns in the intraindividual variability of inconsistency.

b. Quantify the gross amount of intraindividual variability in inconsistency (level of noise).

5. Identify interindividual differences in intraindividual variability (i.e., differences in level of noise).

Method

Participants

Ninety-one older adults (26 men, 65 women) who were 52 to 79 years of age ($M = 65.9$ years, $SD = 7.1$) were recruited from a larger sample of older adults in the Manchester Longitudinal Study of Cognitive Aging (Rabbitt, 1993) to take part in an intensive training study. Participants were selected on the basis of complete physical examinations by experienced geriatric physicians so that all participants were free of pathology that might impair their cognitive function and their ability to complete the long and taxing involvement required by the intense repeated measures study design. Furthermore, participants were matched on unadjusted AH4-1 Intelligence Test (Heim, 1970) scores ($M = 36.7$, $SD = 9.6$) such that there were no mean differences between the three age groups of 51–60 years, 61–70 years, and 71–80 years. Across all ages, the sample participants were in good health and had completed some higher education. Further information can be obtained from an earlier description (see Rabbitt et al., 2001).

Procedure

Participants took part in a total of 36 weekly sessions (after an initial familiarization session) and two follow-up sessions (3 and 6 months later). Sessions took place at one of four times of day: 10 a.m., 12 p.m., 2 p.m., and 4 p.m. To balance circadian effects, age groups were matched for time of day, and each individual attended at the same time and day each week. With rest periods, each session lasted between 60 and 90 min during which participants completed a substantial battery of cognitive measures. The present study focuses on a series of multitrial letter search tasks that were presented at each occasion.

Letter Search Tasks

Participants were presented with a set of 2, 4, or 6 letters to be memorized (target letters). Their ability to remember these letters was then tested by requiring them to select the presented letters from a larger set of 4, 8, or 12 letters shown on a computer screen. If the target letters were not selected correctly, the letters were presented again; the cycle continued until the participant was able to demonstrate recognition. After achieving 100% accuracy in target letter recognition, testing trials began. A sequence of letters was presented, one at a time, on the screen. Participants were asked to indicate if each letter was or was not one of the previously memorized letters. As soon as a response was made, whether correct or not, the next letter was presented. Letters were presented in 50 trial blocks, with a short break between blocks.

Blocks of trials were constructed in two ways. In a varied mapping condition, target and distractor letters were both selected from a single pool of letters (i.e., C, W, Q, G, M, H, K, N, V, A, L, X). In a constant mapping condition, target and distractor letters came from two separate pools of letters (e.g., target letters selected from B, Z, E, R, F, T and distractor letters selected from D, Y, P, U, J, S). In both conditions, letters were chosen at random from the relevant pool such that half of the trials (letters) were target letters and half were distractor letters.

In sum, there were six letter search task conditions: 3 (target load sizes: 2, 4, or 6 target letter memory sets) $\times$ 2 (block conditions: variable or constant mapping). All participants received all six conditions. Thus, in each session participants completed six blocks (presented in random order) of 50 trials each. Feedback (% accuracy) was given after every block and at the end of the session (overall accuracy and mean reaction time).

Measures

Memory speed. Reaction time (RT) was recorded for each trial as the number of milliseconds between letter presentation and the participant's indication on the keyboard of whether the letter was a target or nontarget (i.e., distractor) letter. Two mean RTs were computed for each block by separately averaging response times for target letters and for nontarget letters. These measures, 12 for each session, were used as indicators of memory speed.

Inconsistency. In the same manner, two ISDs were computed across trials on the within-block RTs, that is, for target letters and for nontarget letters. These ISDs, 12 for each session, were used as indicators of individuals' within-occasion inconsistency.

Accuracy. Accuracy of response (i.e., correctly identifying letters as target or nontarget letters) was computed as the percentage of correct responses on target letter trials and, separately, the percentage of correct responses on nontarget letter trials. Thus, matched to the memory speed (RT) and inconsistency (ISD) measures, there were 12 accuracy measures for each of 36 sessions. Participants were highly accurate with within-person mean accuracy ranging from 96.9% to 99.4% across tasks. Because the accuracy measures showed so little variation both within and between participants, they provided little intraindividual or interindividual information and so were not used in any of the following analyses.

Explanatory variables. In addition to the repeated measures, a number of demographic and individual difference characteristics were obtained. These included age (number of years since birth to first occasion of measurement) and Cattell Culture Fair Intelligence Test scores (Cattell, 1973), an established measure of fluid intelligence (G_f).

Data Analysis

Multivariate techniques were used to build a true score measure of inconsistency and to model interindividual differences in intraindividual change and intraindividual variability. First, factor analysis was used to develop a parsimonious measurement model allowing for the precise, error-free quantification of inconsistency. Second, growth models were used to extract the intraindividual change in inconsistency that occurs with practice across the 36 weekly sessions. Third, canonical correlation was used to determine if and how interindividual differences in intraindividual change were related to age and intelligence. Fourth, weekly measures of practice-adjusted inconsistency were derived by removing the systematic intraindividual changes.[1] The remaining intraindividual variability (residual) series were examined for evidence of systematic statelike patterns using exploratory time-series methods (e.g., Ljung-Box test for random noise and ARMA model fitting), and the gross amount of noise within them was quantified. Finally, canonical correlation was again used to examine how interindividual differences in the amount of intraindividual variability were related to other available interindividual difference variables (i.e., age, intelligence). All models were fit to the data using either LISREL (Version 8.54, Jöreskog & Sörbom, 2003) or SAS. Incomplete data ($<7\%$) were treated as missing at random (Little & Rubin, 1987). Further description of the analysis techniques and results are presented in the next section in stepwise fashion, with later analyses having been informed by earlier ones.

Results

Multivariate Measurement of Inconsistency

Individuals who perform inconsistently on one task may also perform inconsistently on other tasks. That is, individuals may have relatively inconsistent days or relatively consistent days. On a "bad" day, they may be inconsistent on all the different tasks they attempt. Alternatively, on a "good" day, they might perform consistently on all tasks.

To examine these notions of task performance covariation across occasions, the 12 inconsistency (ISD) measures were used as multiple indicators of a single within-person inconsistency process. Likewise, the 12 memory speed (RT) measures were used as multiple indicators of a single within-person memory speed process. We built and tested an appropriate multivariate measurement model via a series of chain P-technique factor analyses.

Chain P-technique factor model. In the commonly used R-technique factor analysis, a Persons × Variables data matrix for a single occasion of measurement is analyzed to identify patterns in the relationships between variables that are defined across persons. In P-technique factor analysis (Cattell, Cattell, & Rhymer, 1947), an Occasions × Variables matrix for a single person is analyzed to identify patterns in the relationships between variables that are defined across occasions for one individual. Chain P-technique (Cattell, 1963) combines these approaches by literally chaining together the Multioccasion × Variable data matrices of two or more persons. This composite matrix is then analyzed to examine the relationships between variables that are defined across both persons and occasions, thereby defining a common (or invariant) measurement space.

By design, chain P-technique analyses confound interindividual and intraindividual differences by including multiple individuals and multiple occasions in the same data matrix. Thus, to be sure that the factor structure obtained in our analysis represented measurement of within-person process, we took care to remove interindividual differences before combining persons into a single chain P matrix. Specifically, the 24 observed measures (12 mean RTs and 12 ISDs) were first standardized within person, thereby removing all interindividual differences in mean and amount of variation. Between-persons differences due to mean and variance differences were thus minimized or eliminated.

On a first look at the data, exploratory chain P-technique factor analyses indicated that one latent factor accounted for the common variation between the 12 memory speed (RT) measures and that one latent factor accounted for the common variation between the 12 inconsistency (ISD) measures. Then, in a confirmatory structural equation modeling framework, the 24 measures were modeled simultaneously, with the 12 memory speed (RT) measures indicating a single memory speed factor and the 12 inconsistency (ISD) measures indicating a single inconsistency factor. This measurement (factor) model, graphically presented in Figure 1, allowed us to separate the "true" memory speed and inconsistency variance from residual and error variance (u_1 to u_{24}), thus affording us more reliable measurements of the phenomenon of interest, inconsistency.

The multivariate measurement model fit the data well (χ^2 = 3746, *df* = 239; RMSEA = .067, CFI = .95, Tucker-Lewis Index [TLI] = .94; see Table 1), providing support for our notions that the 12 RT and 12 ISD measures could be represented parsimoniously by two underlying factors. Although significant improvements in fit could be achieved by expanding the model to include a number of multimethod factors for the different task conditions (e.g., Campbell & Fiske, 1959; Widaman, 1985, 1992), such elaboration seemed only to detract from the current questions regarding intraindividual change and variability. Thus the parsimonious model was retained.

Separability of memory and inconsistency factors. With RT tasks, the computed means (speed) and standard deviations (inconsistency) are often highly related, with some argument as to if and how they measure different processes (see, e.g., Hultsch & MacDonald, 2004). Here, the separability of the two processes (factors) was tested using hierarchically nested factor models (see

[1] Hultsch and colleagues have suggested that the calculation of intraindividual variability be done after controlling for potential confounds such as differing materials (tasks), mean differences in speed, and time-related effects such as practice or learning to learn (Hultsch & MacDonald, 2004; Hultsch et al., 2000). We also have considered these effects, but in a different manner. Differing task effects were accommodated through the measurement model. Our multivariate measure of inconsistency represents only that portion of within-session intraindividual variability that was common across all tasks. Mean differences and practice effects were modeled using the exponential growth model. These effects were then partialed (subtracted) from our subsequently derived measures of intraindividual variability (across weeks).

Table 1). In addition to the model shown in Figure 1, the data were also modeled using a single factor indicated by all 24 measures (χ^2 = 4281, *df* = 240, RMSEA = .072, CFI = .93, TLI = .92). Such a model would fit better if there were a single process driving both speed and inconsistency. A chi-square difference test comparing the one- and two-factor models indicated that the two-factor model fit the data significantly better ($\Delta\chi^2/\Delta df$ = 535/1). Thus, we concluded that memory speed and inconsistency represent different and separable constructs. Higher scores on the memory speed factor indicate slower memory search, and lower scores indicate faster memory search. Similarly, higher scores on the inconsistency factor indicate greater variability in performances of multiple tasks, that is, an erratic day; lower scores indicate consistency of performance, that is, a steady day.

Inconsistency factor scores. Having established a reasonable structure and theoretical basis for the latent factors, LISREL was used to compute estimates of inconsistency factor scores (Anderson & Rubin, 1956) for use in the subsequent analyses. These factor scores are considered to be unbiased estimates of the factors. Their sample covariance matrix is exactly equal to the estimated covariance matrix of the reference variables. In other words, the estimated factor scores are individuals' scores on the error-free latent construct to the extent that the factor model fits the data. With issues concerning the indeterminacy of factor scores in mind (e.g., Guttman, 1955), we examined the correlation between the estimated factor scores and the unobserved latent factor for an indication of the reliability of the estimates. A correlation of .86 indicated that factor indeterminacy was not problematic (i.e., $r >$ .70; Grice, 2001). Thus, here, individuals' estimated factor scores are considered to be a reasonable (though by no means perfect) representation of how the latent process underlying performance inconsistency develops over time.

Note that the factor scores were obtained using the raw data, not the standardized data used previously. Thus, these factor scores contain the interindividual and intraindividual difference information, some portions of which we will examine in the following analyses. The 91 participants' estimated scores for the 36 weekly sessions are plotted in Figure 2. Each line represents the trajectory of an individual's inconsistency over time.

Table 1
Factor Models of Memory Speed and Inconsistency

Model	FIML χ^2	*df*	$\Delta\chi^2/\Delta df$	RMSEA	CFI/TLI
One factor	4281	240	—	0.072	.93/.92
Two factor	3746	239	535/1	0.067	.95/.94

Note. FIML = full information maximum likelihood; *df* = degree of freedom; RMSEA = root-mean-square-error of approximation; CFI = comparative fit index; TLI = Tucker-Lewis index.

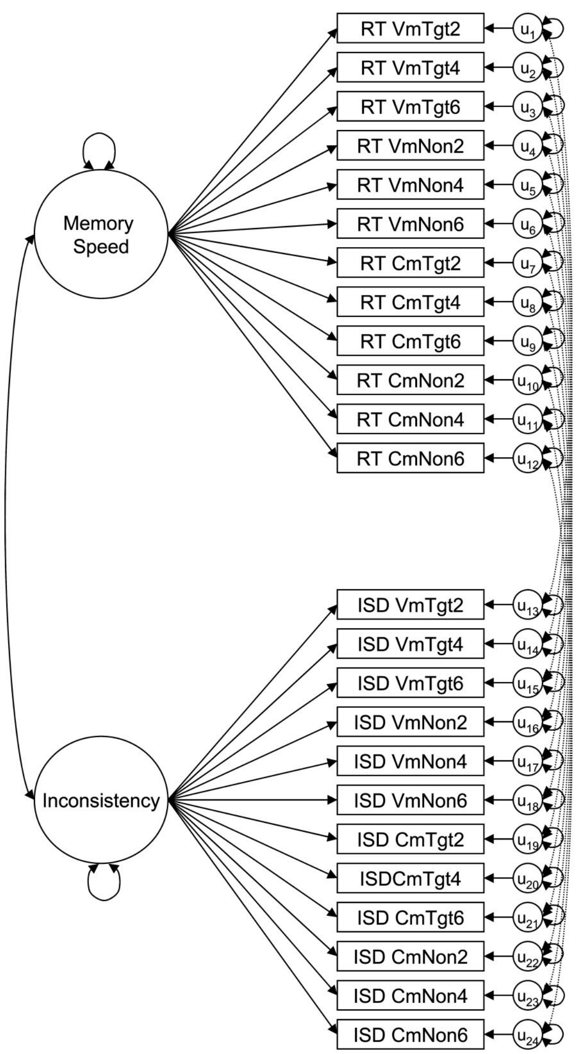

Figure 1. Factor (multivariate measurement) model of memory speed and inconsistency. The memory speed factor is indicated by 12 mean reaction time (RT) measures and the inconsistency factor is indicated by 12 intraindividual standard deviation (ISD) measures.

Intraindividual Change in Inconsistency (Practice Effects)

We hypothesized that the quality of individuals' performances would improve with increased exposure to the tasks. More specifically, individuals' performances would become more consistent as they progressed through the 36 weeks of training. A series of linear and nonlinear growth curve analyses were used to test this hypothesis and to find the best description of such changes.

No-growth, linear, quadratic, and exponential growth models were used to systematically model intraindividual change in inconsistency. In line with our hypothesis, the models of change (linear, quadratic, etc.) fit the data better than a no-growth model (e.g., a linear model fit significantly better than a no-growth model; Δ -2 log likelihood [-2LL]/Δdf = 1670/3; see Table 2). As individuals became increasingly familiar with the task, they became, on average, more consistent in their within-session performances. As seen in Table 2, a comparison between a number of different models of change (i.e., different mathematical expressions of growth) indicated that, of the models fitted, a model of exponential change provided the best overall representation of how individuals' level of inconsistency changed with practice.

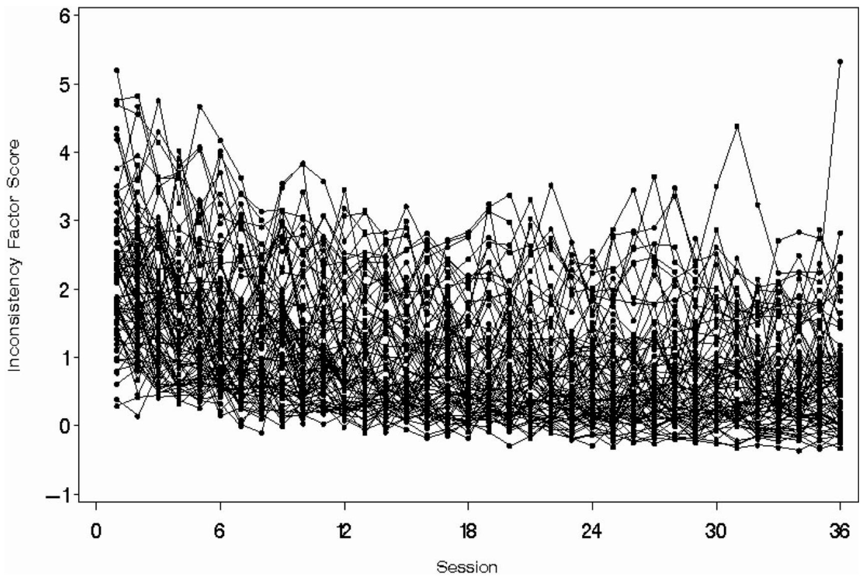

Figure 2. Estimated inconsistency factor scores across 36 weeks. Each individual is represented by a single line connecting their week-to-week scores.

The exponential growth model was specified as:

$$\text{Inconsistency}_{ij} = \pi_{0i} + \pi_{1i}\, e^{(-\pi 2i * \text{SESSION}ij)} + \varepsilon_{ij}$$

The model uses three parameters to describe intraindividual change. The π_{0i} parameter describes individuals' asymptotic level of performance and is interpreted as the limit of an individual's capability (i.e., the lowest level of inconsistency he or she can attain). When combined with the intercept parameter, $\pi_{0i} + \pi_{1i}$ represents the level of inconsistency at which the person began the training (i.e., level of inconsistency at time = 0). Thus, π_{1i} on its own represents an individual's potential for improvement from his or her initial level. Finally, π_{2i} indicates the rate at which an individual's consistency improved.

In sum, with week-to-week repetition, individuals generally became exponentially more consistent. However, these improvements eventually leveled off as an asymptotic level of consistency was reached. From the (multilevel) exponential growth curve analysis, we were able to derive predicted curves describing how each individual's inconsistency developed over the course of the 36 weeks of testing. These predicted growth curves of inconsistency are shown in Figure 3.

Interindividual Differences in Intraindividual Change (Age and Intelligence Effects)

In the plots (see Figure 3), it is clear that some individuals start out more inconsistent than others, some individuals improve more quickly than others, and some individuals reach a final level of inconsistency that is higher than the level of others. In other words, there are clear interindividual differences in intraindividual change. We examined how these differences in intraindividual change were related to age and intelligence.

A significant canonical correlation indicated that interindividual differences in the parameters π_{0i}, π_{1i}, and π_{2i} were related to measures of age and intelligence (i.e., age in years and Cattell Culture Fair Intelligence test scores; $R^2 = .62$), $F(6, 156) = 7.416$, $p < .0001$. In follow-up tests, only the asymptotic level of inconsistency, π_{0i}, was significantly related to intelligence, $t(89) = 6.36$, $p < .0001$, $\beta = -.61$. To summarize, ages and intelligence scores were not systematically related either to the total amount by which individuals' consistency improved over the 36 weeks of training or to the rates at which this improvement occurred. However, individuals with higher intelligence scores were more likely

Table 2
Growth Curve Models of Intraindividual Change in Inconsistency

Model	General equation	No. of parm	-2LL	Δ-2LL/Δ*df*
No growth	$Y_{ij} = \pi_{0i} + \varepsilon_{ij}$	3	5516	—
Linear	$Y_{ij} = \pi_{0i} + \pi_{1i}(t_{ij}) + \varepsilon_{ij}$	6	3846	1670/3[a]
Quadratic	$Y_{ij} = \pi_{0i} + \pi_{1i}(t_{ij}) + \pi_{2i}(t_{ij}^2) + \varepsilon_{ij}$	10	3261	585/4[b]
Exponential	$Y_{ij} = \pi_{0i} + \pi_{1i}\exp(-\pi_{2i}(t_{ij})) + \varepsilon_{ij}$	10	3118	728/4[b]

Note. parm = estimated parameters; -2LL = -2 log likelihood; $\Delta-2\text{LL}/\Delta df$ = relative fit for nested models.
[a] Improvement in fit relative to no-growth model. [b] Improvement in fit relative to linear model.

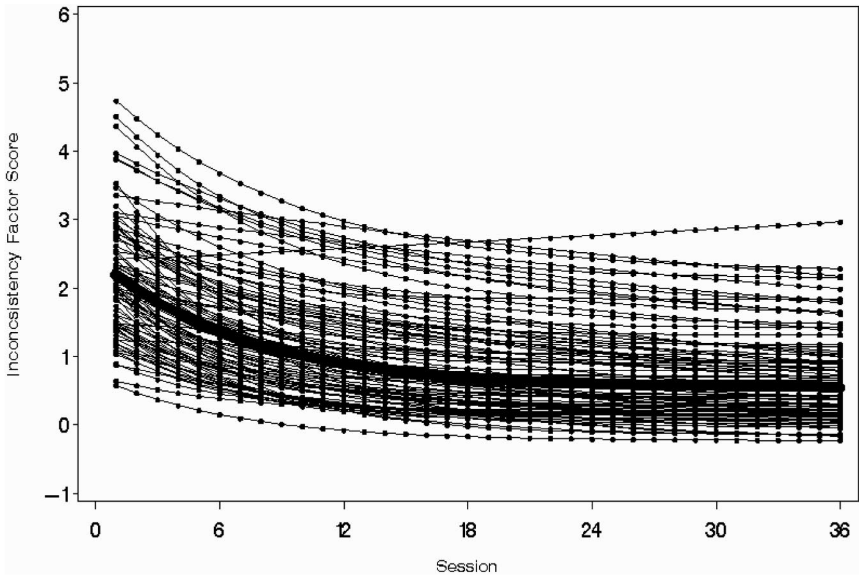

Figure 3. Predicted exponential growth curves representing intraindividual change in inconsistency. Each individual is represented by a single line connecting the week-to-week scores, illustrating practice effects. The prototypical trajectory is indicated as thick bold line.

to attain greater consistency than individuals with lower intelligence scores.

Intraindividual Variability of Inconsistency

To separate intraindividual change from intraindividual variability, the predicted inconsistency scores derived from the exponential growth model (intraindividual change) were subtracted from actual inconsistency factor scores. This effectively removes the systematic effects associated with practice or learning (Hultsch et al., 2000). Residuals are plotted in Figure 4. Each residual noise series represents individuals' week-to-week fluctuations in inconsistency (intraindividual variability), after we controlled for systematic intraindividual changes. We proceeded to analyze these residuals along two avenues.

Structure of intraindividual variability. First, we examined the noted intraindividual variability in practice-adjusted inconsistency for evidence of systematic statelike patterns. We began by using Ljung-Box tests (Ljung & Box, 1978) to test whether individual

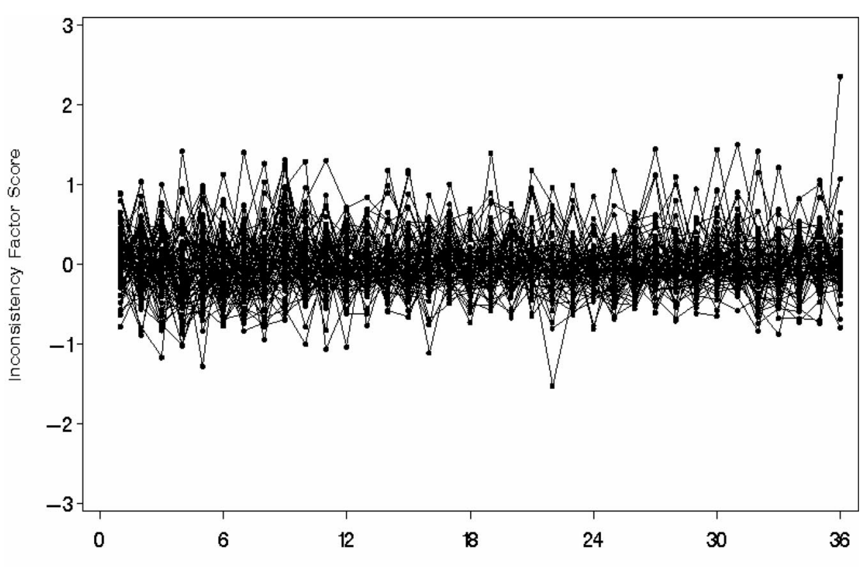

Figure 4. Residual (practice adjusted) inconsistency factor scores across 36 weeks. Each individual's intraindividual variability in inconsistency is represented by a single line.

time series were white noise. This statistical test indicates whether or not the given time series can be considered a random white-noise series. All but 3 of the 91 individuals' intraindividual variability series were white noise (Ljung-Box Q $<$ critical χ^2, $p > .01$ for up to 12 lags).

Post hoc examinations of these three individuals' performances indicated that they were idiosyncratic in one way or another. For two individuals, their intraindividual change with practice would have been better identified by a linear rather than an exponential pattern of change. The misidentification of the intraindividual change model for these individuals left a systematic pattern in their residual noise series. The other individual's intraindividual change exhibited a 4-week period of higher than expected inconsistency, consecutive "outliers," so to speak. These observations also led to a misidentification of his exponential pattern of change and thus a systematic pattern in his intraindividual variability series.

For all other individuals, the residual intraindividual variability series were white noise. Thus, we concluded that, overall, there was very little evidence of systematic patterns within the intraindividual variability of inconsistency. There was no systematic structure in the intraindividual variability that we could try to model. We could, however, attempt to characterize the noise.

Level of intraindividual variability (noise). The gross amount of intraindividual variability in each individual's residual noise series was quantified by calculating the ISD across weeks (Slifkin & Newell, 1998). This ISD was used as an index of the level of noise or "hum" in individuals' week-to-week inconsistency. Some individuals exhibited large fluctuations, that is, lots of noise; others exhibited only small fluctuations, that is, only a little noise.

Interindividual Differences in Intraindividual Variability

Across the 91 participants, the gross amount of intraindividual variability (ISD) ranged from 0.135 to 0.751 (M = 0.339, SD = .118). There were clear, relatively normally distributed, interindividual differences in amount of intraindividual variability (i.e., level of noise). Using multiple regression, we also found that level of noise was significantly related to age and intelligence scores, $F(2, 79) = 17.97$, $p < .0001$. Particularly, level of noise was negatively related to Cattell Culture Fair Test scores, $t = 5.44$, $p < .0001$, $\beta = -.51$, and not related to age, $t(89) = 0.28$, $p > .05$. In sum, individuals with lower levels of intelligence tended to exhibit greater variability in consistency. On some occasions, they were much more consistent than expected, and on others they were much less consistent than expected. Individuals with higher intelligence scores did not vary as much in their week-to-week levels of inconsistency.

Discussion

Our main purpose was to analyze a set of multivariable, multioccasion, multiperson, multi–time-scale data to extract and understand interindividual differences in both intraindividual change and intraindividual variability of inconsistency in memory speed performance. Through five steps of data analysis, inconsistency was shown to manifest systematic intraindividual changes, intraindividual variability, and interindividual differences in these two types of within-person change.

We used chain P-technique factor analysis to derive a model for the multivariate measurement of inconsistency. We found good convergence across multiple measures of inconsistency that allowed for the reliable measurement of inconsistency. To our knowledge, this is the first time that the advantages of multivariate measurement models to nullify the intrusion of error have been applied in the study of inconsistency.

Second, we identified intraindividual change in inconsistency using growth curve analysis techniques. Exponential curves provided the best description of how an individual's inconsistency changes with practice and supported our hypothesis that inconsistency changes with practice. These findings mirror a plethora of research wherein learning processes have often been found to be best described by curves of the exponential family (Heathcote et al., 2000; Newell & Rosenbloom, 1981; Newell, et al. 2001; Thorndike, 1913; Thurstone, 1919). Here, the evidence indicates that consistency may be a product of learning as well.

The exponential model of change is particularly useful as a descriptor of learning, because the estimated parameters can be articulated in relation to theories of intellectual ability. The model parameters, π_{0i}, π_{1i}, and π_{2i}, describe the final level of inconsistency reached, the increment in improvement between initial level and final level, and the rate of improvement over time. Provided that the length of assessment is long enough for individuals to reach an asymptotic level of performance, as appears to have been the case in this study, the final level of inconsistency can be interpreted as the limit of an individual's capacity (e.g., testing the limits, Kliegl, Smith, & Baltes, 1989; biological substrate, Thurstone, 1919). The individual cannot get any better. We have, in our analysis of intraindividual change, not only described the learning process, but also have been able to extract a measure of an individual's "true" ability level, a meaningful measure on which to compare individuals.

Third, we examined interindividual differences in intraindividual change. We found that individuals' intelligence scores were systematically related to their true ability levels (i.e., capacities) but were not systematically related to either their potential for improvement or the rate of their improvement over time. We only found that individuals with higher intelligence scores were able to, in the end, achieve lower levels of inconsistency. From a sampling perspective, such results are as we would expect. Individuals' previous experience with the task, familiarity with test taking, and so forth are not known at the first occasion of measurement. We expect such individual characteristics — characteristics that affect task performance or inconsistency — to be randomly distributed when participants enter into training. Therefore, we would not expect these characteristics to be systematically related to intelligence or age. Here they were not. However, after individuals have progressed in their training and achieved a substantial level of common experiences and familiarity with the tasks at hand, we would expect those initial differences to be greatly reduced if not erased. Any remaining interindividual differences would be expected to be related to true ability (i.e., intelligence). This was indeed the case.

Additionally, age was not systematically related to any of these three parameters. Within the context of these data and particular sample, such a finding is to be expected. The participants were specifically selected so that age was not related to intelligence level. Participants across three decades of life were matched on intelligence scores (i.e., AH4-1). Thus, this sample does not exhibit the usual negative age trends in intelligence scores that many older samples exhibit (Rabbitt et al., 2001; Craik & Salthouse, 2000). We also found that the participants' age is not related to their pattern of intraindividual change in inconsistency.

Fourth, we identified intraindividual variability in inconsistency. We examined the intraindividual variability in practice-adjusted inconsistency for evidence of systematic statelike patterns. Except for a few isolated cases, we found none. Individuals' across-week variability in inconsistency was random noise. Although we had not expected to arrive at this point so quickly, this is what we had hoped to find at the end of the analysis. There were no systematic patterns left to explain in the data. Thus, we found within-person changes that were completely transient (i.e., true intraindividual variability; Hultsh & Macdonald, 2004). Noise, although qualitatively similar across all persons (i.e., all noise series have no inherent structure), differed in quantity across individuals.

Finally, we examined these interindividual differences in intraindividual variability, that is, noise. In line with previous research (e.g., Li & Lindenberger, 1999), we found that level of noise was related to level of intelligence. Individuals who exhibited higher levels of noise in inconsistency tended to score lower on intelligence tests. Thus, we provide new evidence that noise is a marker for some inherent characteristic that is associated with a wide variety of functional decrements. Recent research suggests that noisy brain activation patterns are related to decrements in memory, intelligence, and so forth (see Raz, 2000, for a review). Our findings suggest that future research in this area should also include investigation of how neural activation relates to cognitive performance inconsistency (see also Li & Lindenberger, 1999).

Previous research has indicated that inconsistency is related to age, injury, health, and intelligence (Hultsch & MacDonald, 2004) and is a marker of impending decline or of already low functionality (e.g., Li & Lindenberger, 1999; Rowe & Kahn, 1985). This investigation indicates that inconsistency, itself, exhibits distinct and meaningful patterns of change, some enduring and some transient. In other words, inconsistency is not a stable interindividual difference construct. It changes from moment to moment. Furthermore, these changes in inconsistency appear to contain meaningful interindividual difference information (e.g., correlated to intelligence). Implications of these results are that repeated measurements of intraindividual variability (e.g., inconsistency) may be warranted and that further examinations of the change and variability of intraindividual variability (and change) may also prove fruitful.

Rabbitt et al. (2001) examined the data used here for individual differences in performance variability. They found that within-session variability (what we have termed *inconsistency*) in task performance was positively related to between-sessions variability in mean task performance (which we did not examine). Similar results have also been noted by Hultsch et al. (2000) and West et al. (2002). To help place the results from this examination in relation to this set of previous results, we noted a positive relationship (r = .61) between individuals' (asymptotic) within-session variability, π_{0i}, and between-sessions variability in (practice-adjusted) within-session variability (intraindividual variability). These results provide further evidence that there may be, as Li and Lindenberger (1999) propose, a single neurobioloical substrate underlying all kinds of intraindividual variability.

These previous studies have also noted that inconsistency makes independent contributions to the prediction of other individual differences (e.g., mental status, intelligence) in addition to those made by level of performance. We noted a similar type of finding. Both the (asymptotic) level of inconsistency and intraindividual variability of inconsistency (i.e., level of noise) made independent contributions to the prediction of intelligence scores (R^2 = .42, ΔR^2 with addition of level of noise variable = .06), $F(2, 79)$ = 29.08, $p < .0001$. In sum, this study builds upon the previous research by digging one level deeper into the characteristics of inconsistency, with both the intraindividual changes and variability of inconsistency seeming to hold some utility for furthering our understanding of cognitive functioning.

In conclusion, we used a number of multivariate techniques to analyze a set of longitudinal data that included two time scales of measurement. Week-to-week measurements were taken across-occasions, while second-to-second measurements were taken within-session. Such a burst design provided a number of opportunities for analysis. Along the shorter time frame, we studied intraindividual variability orinconsistency. Along the longer time frame, we were able to examine both intraindividual change and intraindividual variability in this construct. Without multiple time scales of measurement, such an investigation would not have been possible. Furthermore, if individuals' actual trial-to-trial performance information had been available (rather than only the within-session summary statistics), we could have examined within-inconsistency structures. We would encourage other researchers to examine such design elaborations (see also Newell, Liu, & Mayer-Kress, 2001). In the same way that much knowledge has been gained over the years with the inclusion of multiple variables into study designs, much might also be learned with the inclusion of multiple time frames.

References

Anderson, R. D., & Rubin, H. (1956). Statistical inference in factor analysis. *Proceedings of the Third Berkeley Symposium of Mathematical Statistics and Probability, 5,* 111–150.

Anstey, K. J. (1999). Sensorimotor variables and forced expiratory volume as correlates of speed, accuracy, and variability in reaction time performance in late adulthood. *Aging, Neuropsychology, and Cognition, 6,* 84–95.

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

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

Butler, A. C., Hokanson, J. E., & Flynn, H. A. (1994). A comparison of self-esteem lability and low trait self-esteem as vulnerability factors for depression. *Journal of Personality and Social Psychology, 66,* 166–177.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56,* 81–105.

Cattell, R. B. (1957). *Personality and motivation: Structure and measurement.* New York: World Book.

Cattell, R. B. (1963). The structuring of change by P-technique and incremental R-technique. In C. W. Harris (Ed.), *Problems in measuring change* (pp. 167–198). Madison: University of Wisconsin Press.

Cattell, R. B. (1973). *Cattell Culture Fair Intelligence Test*. Champaign, IL: Institute for Personality and Ability Testing.

Cattell, R. B., Cattell, A. K. S., & Rhymer, R. M. (1947). P-technique demonstrated in determining psychophysical source traits in a normal individual. *Psychometrika, 12,* 267–288.

Craik, F. I. M., & Salthouse, T. A. (Eds.) (2000). *The handbook of aging and cognition* (2nd ed.). Mahwah, NJ: Erlbaum.

Eizenman, D. R., Nesselroade, J. R., Featherman, D. L., & Rowe, J. W. (1997). Intra-individual variability in perceived control in an elderly sample: The MacArthur Successful Aging Studies. *Psychology and Aging, 12,* 489–502.

Fiske, D. W., & Rice, L. (1955). Intra-individual response variability. *Psychological Bulletin, 52,* 217–250.

Fox, N. A., & Porges, S. W. (1985). The relationship between neonatal heart period patterns and developmental outcome. *Child Development, 56,* 28–37.

Fozard, J. L., Vercruyssen, M., Reynolds, S. L., Hancock, P. A., & Quilter, R. E. (1994). Age differences and changes in reaction time: The Baltimore Longitudinal Study of Aging. *Journal of Gerontology: Psychological Sciences, 49,* P179–P189.

Fuentes, K., Hunter, M. A., Strauss, E., & Hultsch, D. F. (2001). Intraindividual variability in cognitive performance in persons with chronic fatigue syndrome. *The Clinical Neuropsychologist, 15,* 210–227.

Grice, J. W. (2001). Computing and evaluating factor scores. *Psychological Methods, 6,* 430–450.

Guttman, L. (1955). The determinacy of factor score matrices with applications for five other problems of common factor theory. *British Journal of Statistical Psychology, 8,* 65–82.

Hampson, E. (1990). Variations in sex related cognitive abilities across the menstrual cycle. *Brain and Cognition, 14,* 26–43.

Heathcote, A. Brown, S., & Mewhort, D. J. K. (2000). The power law repealed: The case for an exponential law of practice. *Psychonomic Bulletin & Review, 7,* 185–207.

Hendrickson, A. E., (1982). The biological basis of intelligence: I. Theory. In H. J. Eysenck (Ed.), *A model for intelligence* (pp. 151–196). Berlin, Germany: Springer-Verlag.

Hertzog, C. Dixon, R. A., & Hultsch, D. F. (1992). Intraindividual change in text recall of the elderly. *Brain and Language, 42,* 248–269.

Horn, J. L. (1972). State, trait, and change dimensions of intelligence. *The British Journal of Educational Psychology, 42,* 159–185.

Hultsch, D. F., & MacDonald, S. W. S. (2004). Intraindividual variability in performance as a theoretical window onto cognitive aging. In R. A. Dixon, L. Bäckman, & L.-G. Nilsson (Eds.), *New frontiers in cognitive aging* (pp. 65–88). Oxford, England: Oxford University Press.

Hultsch, D. F., MacDonald, S. W. S., Hunter, M. A., Levy-Bencheton, J., & Strauss, E. (2000). Intraindividual variability in cognitive performance in older adults: Comparison of adults with mild dementia, adults with arthritis, and healthy adults. *Neuropsychology, 14,* 588–598.

Hultsch, D. F., MacDonald, S. W. S., Hunter, M. A., Maitland, S. B., & Dixon, R. A. (2002). Sampling and generalisability in developmental research: Comparison of random and convenience samples of older adults. *International Journal of Behavioral Development, 26,* 345–359.

Jöreskog, K. G., & Sörbom, D. LISREL (Version 8.54) [Computer program]. Hillsdale, NJ: Scientific Software International.

Kagan, J. (1994). *Galen's prophecy.* New York: Basic Books.

Kliegl, R., Smith, J., & Baltes, P. B. (1989). Testing-the-limits and the study of adult age differences in cognitive plasticity of a mnemonic skill. *Developmental Psychology, 25,* 247–256.

Li, S. C., Aggen, S. H., Nesselroade, J. R., & Baltes, P. B. (2001). Short-term fluctuations in elderly people's sensorimotor functioning predict text and spatial memory performance: The MacArthur Successful Aging Studies. *Gerontology, 47,* 100–116.

Li, S. C., & Lindenberger, U. (1999). Cross-level unification: A computational exploration of the link between deterioration of neurotransmitter systems and dedifferentiation of cognitive abilitites in old age. In L. G. Nilsson & H. J. Markowitsch (Eds.), *Cognitive neuroscience of memory* (pp. 103–146). Seattle, WA: Hogrefe & Huber.

Little, R., & Rubin, D. (1987). *Statistical analysis with missing data.* New York: Wiley.

Ljung, G. M., & Box, G. E. P. (1978). On a measure of lack of fit in time series models. *Biometrika, 65,* 297–303.

MacDonald, S. W. S., Hultsch, D. F., & Dixon, R. A. (2003). Performance variability is related to change in cognition: Evidence from the Victoria Longitudinal Study. *Psychology and Aging, 18,* 510–523.

McArdle, J. J., & Nesselroade, J. R. (2003). Growth curve analysis in contemporary psychological research. In J. Shinka, & W. Velicer (Eds.), *Comprehensive handbook of psychology: Vol. 2. Research methods in psychology* (pp. 447–480). New York: Wiley.

Nesselroade, J. R. (1988). Some implications of the trait-state distinction for the study of development across the life span: The case of personality research. In P. B. Baltes, D. L. Featherman, & R. M. Lerner (Eds.), *Life-span development and behavior* (Vol. 8, pp. 163–189). Hillsdale, NJ: Erlbaum.

Nesselroade, J. R. (1991). The warp and woof of the developmental fabric. In R. Downs, L. Liben, & D. Palermo (Eds.), *Visions of development, the environment, and aesthetics: The legacy of Joachim F. Wohlwill* (pp. 213–240). Hillsdale, NJ: Erlbaum.

Nesselroade, J. R. (2002). Elaborating the different in differential psychology. *Multivariate Behavioral Research, 37,* 543–561.

Nesselroade, J. R., & Featherman, D. L. (1997). Establishing a reference frame against which to chart age-related changes. In M. A. Hardy (Ed.), *Studying aging and social change: Conceptual and methodological issues* (pp. 191–205). Newbury Park, CA: Sage.

Nesselroade, J. R., & Ford, D. H. (1985). P-technique comes of age: Multivariate, replicated, single-subject designs for research on older adults. *Research on Aging, 7,* 46–80.

Newell, A., & Rosenbloom, P. S. (1981). Mechanisms of skill acquisition and the law of practice. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 1–55). Hillsdale, NJ: Erlbaum.

Newell, K. M., Liu, Y. T., & Mayer-Kress, G. (2001). Time scales in motor learning and development. *Psychological Review, 108,* 57–82.

Rabbitt, P. (1993). Crystal quest: An examination of the concepts of "fluid" and "crystallised" intelligence as explanation for cognitive changes in old age. In A. D. Baddeley & L. Weiskrantz (Eds.), *Attention, selection, awareness, and control.* Oxford, England: Oxford University Press.

Rabbitt, P., Osman, P., Moore, B., & Stollery, B. (2001). There are stable differences in performance variability, both from moment to moment and from day to day. *The Quarterly Journal of Experimental Psychology, 54A,* 981–1003.

Raz, N. (2000). Aging of the brain and its impact on cognitive performance: Integration of structural and functional findings. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (2nd ed., pp. 1–90). Mahwah, NJ: Erlbaum.

Rogosa, D., & Willet, J. B. (1985). Understanding correlates of change by modeling individual differences in growth. *Psychometrika, 50,* 203–228.

Rowe, J. W., & Kahn, R. L. (1985). Human aging: Usual and successful. *Science, 237,* 143–149.

Rowe, J. W., & Kahn, R. L. (1997). Successful aging. *The Gerontologist, 37,* 433–440.

Shammi, P., Bosman, E., & Stuss, D. T. (1998). Aging and variability in performance. *Aging, Neuropsychology, and Cognition, 5,* 1–13.

Shoda, Y., Mischel, W., & Wright, J. C. (1994). Intraindividual stability in the organization and patterning of behavior: Incorporating psychological situations into the idiographic analysis of behavior. *Journal of Personality and Social Psychology, 67,* 674–687.

Siegler, R. S. (1994). Cognitive variability: A key to understanding cognitive development. *Current Directions in Psychological Science, 3,* 1–5.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis:*

Modeling change and event occurrence. New York: Oxford University Press.

Slifkin, A. B., & Newell, K. M. (1998). Is variability in human performance a reflection of system noise? *Current Directions in Psychological Science, 7,* 170–177.

Spieler, D. H., Balota, D. A., & Faust, M. E. (1996). Stroop performance in healthy younger and older adults and in individuals with dementia of the Alzheimer's type. *Journal of Experimental Psychology: Human Perceptions and Performance, 22,* 461–479.

Stuss, D. T., Pogue, J. Buckle, L., & Bondar, J. (1994). Characterization of stability of performance in patients with traumatic brain injury: Variability and consistency on reaction time tests. *Neuropsychology, 8,* 316–324.

Stuss, D. T., Stethem, L. L., Hugenholtz, H. Picton, T., & Richard, M. T. (1989). Reaction time after head injury: Fatigue, divided and focused attention, and consistency of performance. *Journal of Neurology, Neurosurgery, and Psychiatry, 52,* 742–748.

Thorndike, E. L. (1913). *Educational psychology: The psychology of learning* (Vol. 2). New York: Teachers College Press.

Thurstone, L. L. (1919). The learning curve equation. *Psychological Monographs, 26*(3), 1–51.

West, R., Murphy, K. J., Armilio, M. L., Craik, F. I. M., & Stuss, D. T. (2002). Lapses of intention and performance variability reveal age-related increases in fluctuations of executive control. *Brain and Cognition, 49,* 402–419.

Widaman, K. F. (1985). Hierarchically nested covariance structure models for multitrait-multimethod data. *Applied Psychological Measurement, 9,* 1–26.

Widaman, K. F. (1992). Multitrait-multimethod models in aging research. *Experimental Aging Research, 18,* 185–201.

Wishart, J. (1938). Growth rate determinations in nutrition studies with the bacon pig, and their analyses. *Biometrika, 30,* 16–28.

Wohlwill, J. F. (1973). *The study of behavioral development.* New York: Academic Press.

Received November 3, 2004
Revision received May 31, 2005
Accepted June 10, 20005 ■