Published by
American Psychological Association
750 First Street, NE
Washington, DC 20002
www.apa.org

To order
APA Order Department
P.O. Box 92984
Washington, DC 20090-2984
Tel: (800) 374-2721;
Direct: (202) 336-5510
Fax: (202) 336-5502;
TDD/TTY: (202) 336-6123
Online: www.apa.org/books/
E-mail: order@apa.org

In the U.K., Europe, Africa, and the Middle East, copies may be ordered from
American Psychological Association
3 Henrietta Street
Covent Garden, London
WC2E 8LU England

Typeset in New Century Schoolbook by Circle Graphics, Inc., Columbia, MD

Printer: United Book Press, Baltimore, MD
Cover Designer: Mercury Publishing Services, Rockville, MD

Library of Congress Cataloging-in-Publication Data

Individual pathways of change : statistical models for analyzing learning and development / edited by Peter C.M. Molenaar and Karl M. Newell.
p. cm.
Papers presented at a conference held at The Pennsylvania State University, University Park in Sept. 2008.
Includes bibliographical references and index.
ISBN-13: 978-1-4338-0772-5
ISBN-10: 1-4338-0772-6
ISBN-13: 978-1-4338-0773-2 (e-book)
ISBN-10: 1-4338-0773-4 (e-book)
1. Developmental psychology—Congresses. I. Molenaar, Peter C. M. II. Newell, Karl M., 1945-
BF713.5.I54 2010
155—dc22 2009046232

British Library Cataloguing-in-Publication Data
A CIP record is available from the British Library.

Printed in the United States of America
First Edition

3

Modeling Retest and Aging Effects in a Measurement Burst Design

Martin Sliwinski, Lesa Hoffman, and Scott Hofer

Researchers who study human development are interested in how psychological, physiological, and behavioral phenomena change over time in aging individuals. In fact, Baltes and Nesselroade (1979) identified the primary objective of longitudinal developmental research as the "direct identification of intraindividual change" (p. 23). However, this goal is complicated by the possibility that observable change in any given individual may reflect the joint influences of multiple processes. For example, observable decreases in memory performance over time (i.e., with increasing age) may reflect the complementary effects of declining vascular health and the progression of Alzheimer's dementia (Sliwinski, Hofer, Hall, Buschke, & Lipton, 2003; Sliwinski, Lipton, Buschke, & Stewart, 1996). In contrast, observable change in cognitive performance may reflect a mixture of competing influences, such as aging-related declines that are partially or completely offset by performance gains attributable to repeated testing (i.e., retest or practice effects).

The purpose of this chapter is to examine a novel approach to decompose age (decline) and retest (gains) effects in longitudinal data. Specifically, we argue that conventional longitudinal designs consisting of repeated and widely spaced single measurements are significantly limited in their ability to disentangle multiple time-dependent processes, such as practice gains and age-related declines in cognition. We present an alternative approach that relies on the longitudinal measurement burst design (Nesselroade, 1991) and a nonlinear measurement model that represents cognitive performance as a function of previous experience and latent potential (i.e., asymptotic performance).

Retest Effects in Aging Research

The term *retest* (or *practice*) *effects* refers to performance gains that result from repeated exposure to testing procedures or materials. There is considerable evidence to indicate that repeated administration of the same or similar cognitive

This research was supported in part by Grants AG12448, AG026453, and AG026728 from the National Institute on Aging.

tests results in improved performance (e.g., Horn, 1972), and several studies have demonstrated significant retest gains after testing intervals of 5 to 10 years (e.g., Salthouse, Schroeder, & Ferrer, 2004; Schaie, 1996; Thorndike, Bregman, Tilton, & Woodyard, 1928). Because many longitudinal studies have retest intervals well less than 10 years, the potential durability of retest-related practice gains complicates the statistical analysis of developmental and maturational influences on cognitive performance. Failure to consider the influence of retest effects can lead to the inaccurate characterization of the rate and pattern of cognitive change (Salthouse et al., 2004) as well as confound attempts to study predictors of and correlations among estimates of change (Ferrer, Salthouse, McArdle, Stewart, & Schwartz, 2005; Wilson, Bienias, & Bennett, 2006). These rather serious consequences have prompted researchers to examine and try to correct for possible retest effects in longitudinal data.

Traditional longitudinal designs consist of widely spaced measurement occasions, often separated by long intervals (typically between 1 and 7 years). This type of design confounds the influence of repeated testing (i.e., practice) and aging on performance because the difference between any two scores from adjacent occasions reflects both the passage of time (and, presumably, aging) and increased exposure to testing (i.e., Occasion 1, Occasion 2, Occasion 3, etc.). If performance were to improve as a function of retest (because of practice) but decrease as a function of time (because of aging), then the observed performance would reflect the combined influence of these two competing latent processes. One approach to disentangle retest from aging has been to compare a control group (e.g., Schaie, 1965; Thorvaldsson, Hofer, Berg, & Johansson, 2006) that was tested only one time but at the same age as a comparison group that was tested multiple times as part of a longitudinal study. This experimental approach works well for quantifying the average retest effects in the population, but it cannot distinguish between retest and age effects in any given individual, which complicates both intraindividual and interindividual analyses of change.

An alternative approach (McArdle & Anderson, 1990; McArdle & Woodcock, 1997) involves statistically partialing the effects of age and retest occasion. This statistical approach has been used in numerous studies and is an increasingly popular analytic model for separating age and retest effects in longitudinal data (e.g., Ferrer, Salthouse, Stewart, & Schwartz, 2004; Ferrer et al., 2005; Ghisletta, McArdle, & Lindenberger, 2006; McArdle, Ferrer-Caja, Hamagami, & Woodcock, 2002; Rabbitt, Diggle, Smith, Holland, & McInnes, 2001; Salthouse et al., 2004; Wilson et al., 2006). The idea behind this approach is to include separate terms that capture both maturational influences (i.e., aging) and practice effects (i.e., retest occasion) in the analytic model of intraindividual change:

$$Y_{ti} = b_{0i} + b_{1i}(\text{age}_{ti}) = b_{2i}(\text{occasion}_{ti}) + r_{ti}, \qquad (3.1)$$

where Y_{ti} is the cognitive performance for person i at time t, b_{0i} is the intercept for person i, b_{1i} is the linear age (or time in study) slope, and b_{2i} is the linear retest (practice) effect for person i. The estimate b_{1i} is the age effect partialed or statistically controlled for the effect of occasion (retest). We refer to this model as the *age + occasion retest model.* This approach will accurately recover population parameters for retest and age effects if its assumptions are reasonably met.

One set of assumptions is general to the underlying statistical model used for estimating the parameters of interest in longitudinal settings (e.g., the mixed model). The other set of assumptions pertains to the underlying conceptual model that allows interpretation of the estimated parameters as reflecting separable retest (practice) and age effects. It is this second set of assumptions that we examine in this chapter.

Retest effects in longitudinal studies could reflect several types of influences. Performance gains across repeated testing might reflect habituation to a type of "white coat" phenomenon resulting in a relief of general testing anxiety. Individuals might also become better at test taking, which would produce a generalized improvement across all cognitive tests, or they might become more proficient only at taking the specific tests in the longitudinal assessment battery. Also, as is the case for many longitudinal studies, individuals might learn the specific content of tests that repeat the same items across repeated assessments. For simplicity, we assume that retest effects in longitudinal studies reflect some type of learning of general test-taking skill, specific testing procedures, or testing material.

The age + occasion retest model requires that there be some variability in the interval between testing occasions, or else the longitudinal effects for age (time) and retest would be completely collinear. By introducing variability in the duration of follow-up it becomes possible to estimate the longitudinal effect of retest, which depends solely on the occasion variable (without respect to the actual interval), and the longitudinal effect of age, which depends solely on the interval between testing occasions. Thus, the age + occasion approach assumes that practice effects are invariant across different retest intervals. This raises the question of whether it is plausible to assume equivalent performance gains between two testing occasions that were separated by either a few days or a few decades.

It seems likely that retest effects would to some extent depend on the interval between testing occasions. At least one study, which had retest intervals ranging from a few days to a few decades, suggested that the magnitude of retest effects does diminish as the interval between testing occasions increases (Salthouse et al., 2004). One reason for diminishing practice effects with increasing retest intervals is that they are offset by the influence of aging. Another reason might be that the benefits of practice dissipate over time, independent of the influence of aging. Also, as Salthouse et al. (2004) suggested, this loss of performance gains across time may be an important component of retest effects in longitudinal studies. However, if retest effects do diminish as a function of time, then the age effect in the age + occasion retest model becomes difficult to interpret. This effect (b_{1i} from Equation 3.1) could reflect aging (as is its usual interpretation), but it could also reflect a type of forgetting (i.e., the time-dependent loss of retest gains) or some combination of both aging and forgetting. In skill acquisition studies forgetting, as evidenced by loss of previously demonstrated performance gains, is observed over intervals as short as 1 day (e.g. Newell, Mayer-Kress, & Lui, 2006; Rickard, 2007), and the magnitude of this loss may depend on the interval between assessments (Anderson, Fincham, & Douglass, 1999). A further complication is that there might be individual differences in the amount and rate of time-dependent forgetting (MacDonald, Stigsdotter-Neely, Derwinger, & Backman, 2006) that would complicate analysis of individual differences in estimates of age-related change obtained from the age + occasion retest model.

This discussion of the conceptual assumptions of the age + occasion retest model is not meant to imply that the results from any given application of this model are incorrect; instead, it is intended to highlight the complexity of disentangling multiple time-dependent processes that drive intraindividual cognitive change. The age + occasion retest model addresses this complexity by representing intraindividual change as a function of two competing processes: (a) aging, which is measured by the passage of time, and (b) retest, which is measured by the number of testing occasions. We now present an alternative approach that relies on the longitudinal measurement burst design (Nesselroade, 1991) and a formal measurement model that represents cognitive performance as a nonlinear function of both testing experience and latent potential (i.e., asymptotic performance).

Modeling Changes in Performance and Latent Potential

Performance differences across long as compared with short retest intervals may reflect both increased aging influences and diminished retest influences (e.g., due to forgetting). One approach to addressing this confound is to use a mix of very closely spaced retest intervals to model practice effects and longer intervals to model age-related changes. This type of longitudinal design is referred to as a *measurement burst* (Nesselroade, 1991; Sliwinski, 2008), and it consists of repeated bursts of closely spaced measurements. This measurement burst design is in contrast to conventional multiwave longitudinal designs, which consist of widely spaced single measurements. In the present study, a "burst" consisted of six measurement occasions that occurred within a 10-day period. A day or two separated occasions within each burst, and each burst was repeated every 6 months for 2 years, yielding up to 30 observations for each individual. We hypothesized that overt performance on a speeded cognitive task would improve across sessions within bursts, because of the benefit of practice, but that estimates of individuals' latent potential (i.e., their asymptotic or best level of performance) would reveal slowing across bursts, because of aging (i.e., senescence, involution).

To represent this hypothesis mathematically, we begin with a model that represents response time (RT) as a negative exponential function of practice occasions:

$$\mathrm{RT}_{ti} = a_i + g_i \exp[-r_i(\mathrm{occasion}_{ti})] + e_{ti}. \qquad (3.2)$$

The first part of this equation, a_i, refers to a person's *asymptotic* response time, which is his or her fastest RT (or latent potential) given unlimited practice. The second part, $g_i \exp[-r_i(\mathrm{occasion}_{ti})]$, reflects that portion of the observed response time, RT_{ti}, that is attributable to his or her experience. The r_i parameter is the rate of learning or improvement across repeated measurement occasions, and the g_i, or *gain* parameter, refers to the difference between an individual's initial performance with no practice and his or her estimated asymptotic performance. Other functions (e.g., power, hyperbolic) could also have been used, but the negative exponential has provided consistently better fits to our data than alternative functions.

Figure 3.1A shows what such a learning function might look like for data collected from a measurement burst design. The points on the graph connected by a line are from the same burst and separated by 1 day, whereas adjacent points that are not connected by a line come from different measurement bursts and are separated by approximately 6 months. The function in Figure 3.1A depicts a situation in which learning is not disrupted by the interval between bursts because an individual picks up on the first session of a burst exactly where he or she should be given where he or she left off on the last session of the previous burst. There is also a common asymptote across bursts, signifying that an individual's latent potential remains constant across the study's duration.

Figure 3.1B shows a slightly more complicated but perhaps more realistic situation in which individuals exhibit some forgetting (i.e., slowing) from the last session of the previous burst to the first session of the current burst. The practice gains on bursts after the first session reflect a recovery of what was lost during the interburst interval as well as performance gains that reflect a continuation of learning that occurred during earlier bursts. One way to model this situation would be to fit separate learning functions to data from each burst (see Rickard 2007). This would imply that there is a single learning process that transpires across bursts but would allow the learning rate to vary from burst to burst. Another approach would assume that the learning function for all bursts after the first reflects two processes: (a) continuous learning and (b) a recovery or warm-up effect. The result would be a rate of improvement during follow-up bursts that is faster than could be predicted by a single exponential learning function. This can be represented mathematically by a double negative exponential function:

$$\mathrm{RT}_{ti} = a_i + g_i \exp[-r_i(\mathrm{occasion}_{ti})] + (\mathrm{Burst}_{kj} > 1) \times g_i^* \exp[-ri_i^*(\mathrm{occasion}_{tki})] + e_{ti}. \qquad (3.3)$$

This equation stipulates that RT is a function of a person's asymptote plus two different learning/retest processes. The first, conveyed by the term g_i $\exp[-r_i(\mathrm{occasion}_{tki})]$, reflects how a person's RT decreases as a function of the total amount of practice he or she has received on that task. The second process, conveyed by the term$(\mathrm{Burst}_{kj} > 1) \times g_i^*$ $\exp[-r_i^*(\mathrm{occasion}_{tki})]$, is a warm-up' process, which operates only during follow-up bursts (i.e., bursts > 1) and indicates that a person's RT starts off higher on the first session of a new burst and then decreases rapidly (i.e., a warm-up effect).

For a real life example of such dual-process learning consider a middle-aged adult who takes up cross-country skiing. During her first season, she displays considerable and rapid improvement in her skiing ability, perhaps indexed by the time taken to complete a local trail. Then spring arrives, the snow melts, and 9 months pass before she can resume skiing. When she resumes skiing the following winter, she is a bit rusty and not quite as fast as she was at the end of the

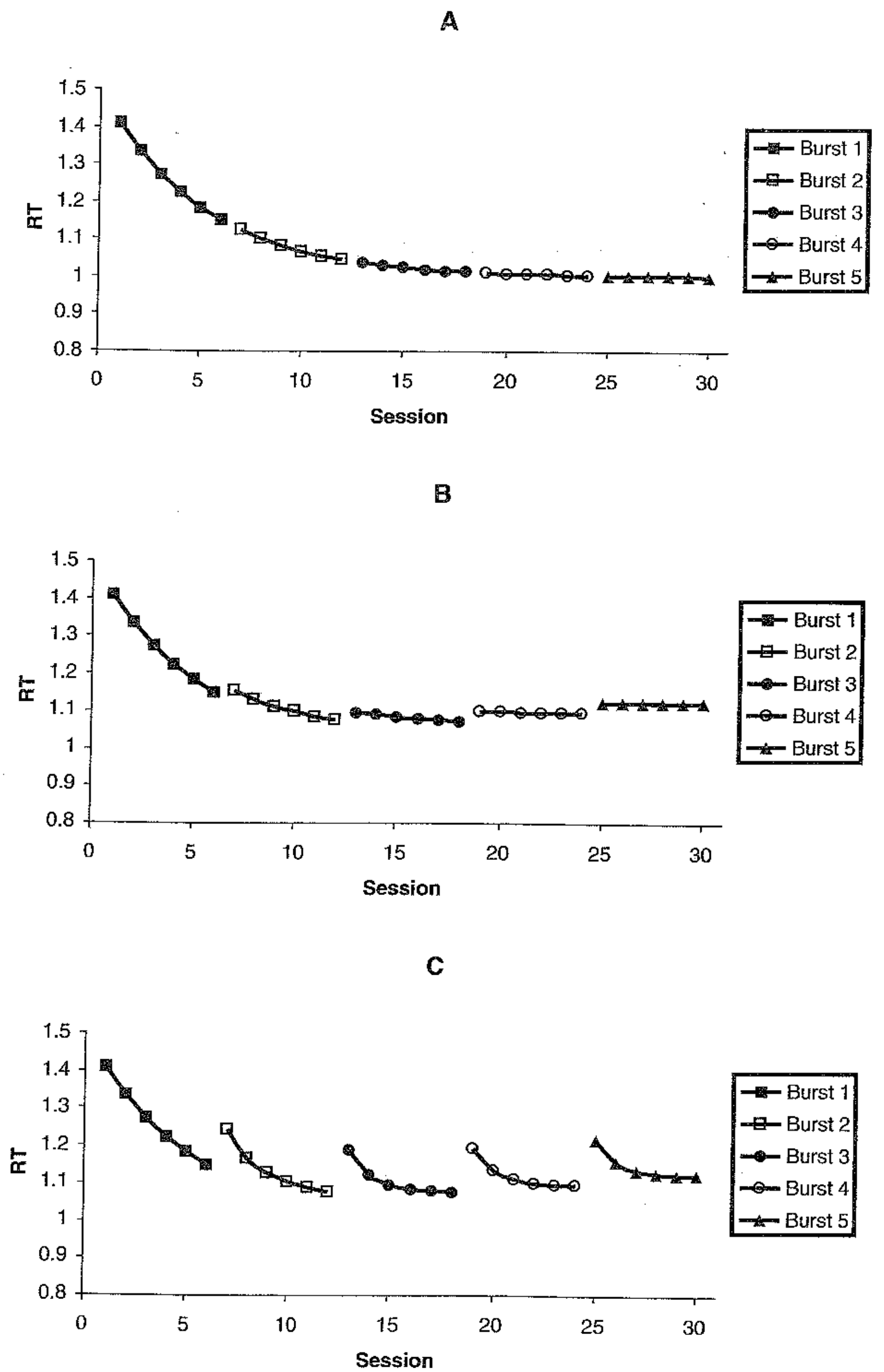

Figure 3.1. Hypothetical practice functions for a measurement burst design. RT = response time.

previous season. With practice, however, she quickly recovers the skill that was "lost" during the off-season and then continues to improve upon her best time. After a substantial temporal disruption in practice, performance becomes a function not only of the total amount practice (i.e., the cumulative practice) but also of how much practice has recently occurred (i.e., local practice).

Another complication may be overlaid on performance gains attributable to cumulative and local practice, namely that another process operates during the interval that separates measurement bursts. To follow our example, although our novice skier is becoming more skilled every season, as reflected in her performance, her potential or maximal speed might be decreasing across seasons because she is aging. Figure 3.1C shows the expected pattern of RTs if there were dual-process learning (i.e., cumulative learning and local warm-up effects) along with an upward drift in asymptote to signify age-related slowing that manifests across bursts. Assuming change in asymptote is a linear function of aging, the function would be:

$$\mathrm{RT}_{ti} = a_i + \Delta a_i(\mathrm{burst}_{kj}) + g_i \exp[-r_i(\mathrm{occasion}_{ti})] + (\mathrm{Burst}_{kj} > 1) \times g_i^* \exp[-r_i^*(\mathrm{occasion}_{tki})] + e_{ti}. \quad (3.4)$$

This equation adds the term $\Delta a_i(\mathrm{burst}_{kj})$, which conveys the amount by which the asymptote changes from one burst to the next. The "fast" change that occurs across sessions within bursts conveys information about retest learning and relearning (the r and r^* parameters, respectively), whereas the "slow" change that occurs across bursts (Δa_i) reflects the effects of aging, senescence or involution.

The Present Study

The present study used a measurement burst design in which each burst consisted of six sessions that were repeated every six months for a period of 2 years. As depicted in Figure 3.2, the measurement burst design allows modeling of performance changes across different time scales. Performance change within each burst reflects fast practice gains, whereas change across bursts reflects the slow effects of aging. Performance change within follow-up bursts (i.e., bursts >1) reflects two processes: (a) cumulative learning and (b) local warm-up effects.

The first objective of this analysis was to determine whether a double negative learning function can describe the retest effects observed in an intensive measurement burst study. If the data follow the pattern depicted in Figure 3.1A, then a double learning function would not be necessary because there would be no forgetting or relearning processes. However, if the pattern of results resembles either Figure 3.1B or Figure 3.1C, that would imply that retest effects depend on the actual interval between occasions and would rule out use of the age + occasion retest model. The fit of a double negative exponential function will be compared with the fit of multiple single negative exponential functions fit to each measurement burst.

The second objective was to test the hypothesis that, despite retest-related improvements in observed performance, asymptotic performance shows age-

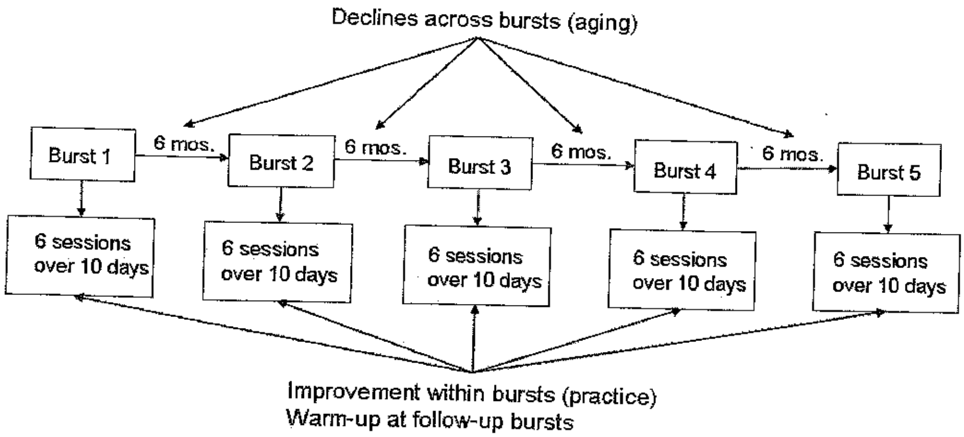

Figure 3.2. Different cadences of change in a measurement burst design. mos. = months.

related declines (i.e., slowing). This prediction implies results that follow the pattern depicted in Figure 3.1C and is grounded in research on cognitive training and testing-the-limits (Kliegl, Smith, & Baltes, 1989). Older adults maintain the ability to improve performance through practice; however, despite this preserved cognitive plasticity advancing age may result in diminished latent potential or reserve capacity. Detecting significant slowing in asymptotic speeded performance would be consistent with age-related reductions in latent potential.

Method

SAMPLE. One hundred sixteen older adults were recruited for participation in a longitudinal study of health and cognition through advertisements in local newspapers and flyers posted in senior centers. All older adults had intact mental status and were compensated $10 for each completed session. The average age was 80.23 (*SD* = 6.30, range: 66–95), the average years of education was 14.9 (*SD* = 2.40), and there was a higher percentage of women than men (72% vs. 28%, respectively).

PROCEDURE AND STIMULI. Participants were given a brief introduction to the study, and the experimenter obtained informed consent as approved by the Syracuse University institutional review board. Participants were told that they were taking part in a study that was examining health and cognition in adulthood. They were scheduled to visit the research site six times within a 14-day period. The research site was a rented apartment at a local senior residence. Half of the sessions (each lasting 1 hour) for each participant were scheduled before 11:00 a.m., and half were scheduled after 1:00 p.m. These bursts of daily measurements were repeated every 6 months, for a 2-year period, yielding up to five bursts of 30 daily assessments.

We examined performance on number comparison speed task, which required participants to compare two strings (three digits in length) to determine whether the same digits were in each string, regardless of their order. In the first session of each burst, sufficient practice trials for all tasks were provided until participants become comfortable with each procedure. Approximately 10 warm-up trials were given before commencement of each task during Sessions 2 through 6. Participants performed a block of 40 trials at each session. Participants were told to press the "/" key if the two digit strings were a match and to press the "z" key if there were a nonmatch. The next number string appeared 500 milliseconds after each response. Participants were instructed to be both fast and accurate. A high-resolution monitor controlled by a Pentium IV–based computer displayed the stimuli. The average RT from correct trials served as the dependent measure for this task. Accuracy was very high (mean proportion correct = .96) and did not significantly change across session within bursts or across bursts, so only RT data were analyzed. A computer-based vision check was administered to verify that all participants could identify test stimuli within video displays of 10.4° of visual angle.

Results

The RTs for each session and burst averaged across all individuals are displayed in Figure 3.3. Connected points belong to sessions obtained within the same burst, and there is approximately a 6-month gap between the last session of one burst and the first session of the following burst. The pattern of average RTs shows a decelerating rate of improvement across sessions within bursts

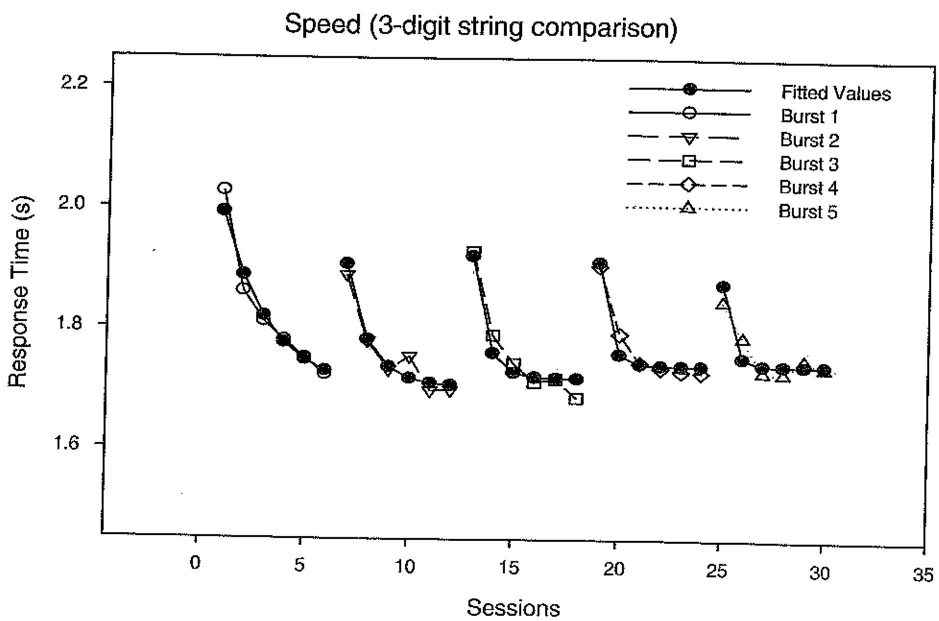

Figure 3.3. Observed and fitted values for average response time for the number comparison.

but exhibits slowing at the first session of each new burst, producing a scalloped retest pattern across bursts. This pattern demonstrates that the duration between occasions influences retest effects, making the age + occasion retest model unsuitable for these data.

We next compared the fit of the double negative exponential model (Equation 3.4) with that of single negative exponential (Equation 3.2). Preliminary model fitting indicated that the best-fitting single negative exponential model allowed all parameters (a, g, and r) to vary across bursts. The best-fitting double negative exponential constrained the parameters r, g, and r^* to be constant across bursts and allowed g^* and a to vary linearly across bursts.

We fit both models using nonlinear mixed modeling implemented by the PROC NLMIXED routine in SAS. Parameters that entered the model linearly (a, g, and g^*, and Δa) were specified as random, but the rate parameters (r and r^*) were constrained to be fixed to facilitate estimation and convergence of the mixed model. Although there is no formal significance test for comparing nonnested models, the double negative exponential fit better than the single according to both the Akaike information criterion (−1,222 vs. −1,207) and Bayesian information criterion (−1,190 vs. −1,174). There are no established guidelines for assessing meaningful differences in comparative fit indices, but Raftery (1995) suggested that a 10-point difference in the Bayesian information criterion constitutes "very strong" evidence in favor of the model with the more negative value.

Figure 3.3 represents the average of the fitted values obtained from the double negative exponential, which closely approximates the observed average values across all 30 sessions. Two different time scales are conveyed in this plot: (a) data points that are connected by a line occur on adjacent days, and (b) adjacent data points not connected by a line are separated by approximately 6 months. Fast change (within bursts) is described by the two rate parameters (r and r^*), and slow change (across bursts) is described by the Δa parameter. The estimated parameters for the double negative exponential model are shown in Table 3.1. The two rate parameters ($r = 0.17$, $SE = 0.011$, and $r^* = 3.44$,

Table 3.1. Fixed and Random Effects Estimates from Dual-Exponential Fit

Parameter	Estimate	*SE*
a	1.500	0.048
Δa	0.054	0.007
g	0.560	0.043
r	0.172	0.011
g^*	1.060	0.007
r^*	3.450	0.141
Var(a)	0.200	0.028
Var(g)	0.115	0.023
Var(g^*)	0.187	0.056
Var(Δa)	0.054	0.001

Note. For all values, $p < .01$.

$SE = 0.144$) reflect the cumulative learning rate across all sessions and the rate of warm-up related improvement at follow-up bursts, respectively. The asymptote was estimated to be 1.500 seconds ($SE = 0.048$) at the first burst and significantly increased ($\Delta a = 0.053$, $SE = 0.007$) across the follow-up bursts. These results indicate that, on average, asymptotic speed was slowing by about 53 milliseconds across every 6-month interval separating the measurement bursts. However, the significant variance component for the Δa parameter (var[Δa] = .003, $p < .001$) implies significant individual differences in the rate at which asymptotic speed slowed across the bursts. All values of Δa greater than 0 indicate slowing, so dividing the difference between 0 and the average Δa value of 0.053 by the square root of the variance component yields a z score of –0.92, implying that approximately 82% of individuals would be expected to exhibit asymptotic slowing, assuming a normal distribution of the random Δa effects.

As a final step in the analyses we correlated the random Δa effects with age and sensorimotor functioning, both of which are thought to relate to the rate of slowing in older adults. Age was positively associated with Δa ($r = .29$, $p < .01$), indicating that older individuals exhibited more rapid slowing. There was a strong negative correlation ($r = -.40$, $p < .01$) with a composite sensorimotor variable (a z-score average of visual acuity and grip strength). These correlations are in the expected direction and consistent with the a priori expectation that cognitive slowing would be accelerated in older individuals and in those with poorer sensorimotor function (MacDonald, Dixon, Cohen, & Hazlitt, 2004).

Discussion

The present results question the validity of a key assumption underlying age + occasion retest models, namely, that retest (i.e., practice) effects are invariant across different time intervals. They also support the utility of both measurement burst designs and the dual-exponential learning function as a model of clustered-practice effects that occur in measurement burst designs. The data summarized in Figure 3.2 clearly show within-burst speedup, across-burst loss of practice gains, and warm-up effects at follow-up bursts. The double-exponential learning model represents such clustered-practice effects by specifying two learning functions, one that describes cumulative learning and a second that describes local warm-up effects that result after a temporal delay between practice opportunities. These data are also similar in form to multisession learning data presented by Rickard (2007), which tend to show within-session learning, across-session forgetting, and a rapid warm-up effect at follow-up sessions. Additional research is required to determine whether a double-exponential learning model can also describe learning and warm-up effects in other contexts (e.g., multisession skill acquisition paradigms).

There are several noteworthy limitations of the present analyses. First, only six sessions per burst might not have been sufficient to bring each individual close to his or her asymptotic performance. Inspection of individual plots indicated that this was the case. Therefore, the present results depend on the accuracy of the extrapolation the model made for each individual asymptotic performance. Second, we aggregated across trials within each session and examined practice effects for the average RT across sessions, ignoring microlearning across trials within sessions. Consequently, the present analyses did not characterize microlevel (within-session) learning effects, which could have distorted the characterization of learning across sessions, within a given burst. For example, the warm-up effect might have been fully contained within the first dozen trials on the first session of follow-up bursts.

As intensive measurement designs become more prevalent, researchers will need to incorporate more dynamic measurement models that explicitly represent the role of time and repeated measurements. If cognitive performance is variable and does change across time, then a useful measurement model must specify how performance changes (e.g., exponentially with repeated measures), which aspects of cognitive function remain invariant (e.g., asymptote or latent potential), and the time scale over which change and invariance obtains (e.g., Newell et al., 2006). Most studies of learning and development have examined changes across fixed intervals and thus have not considered the importance of time scale for characterizing changes that are due to learning and development. One exception Newell, Mayer-Kress, and Lui's (2001) work, which provided a general theoretical framework for understanding motor learning and development across different time scales. Although consistent with Newell et al.'s argument for the importance of time scale, the present approach is purely descriptive. Equations 3.2 through 3.4 provide a simple measurement model that may be usefully applied to performance data that exhibit long-term (e.g., development) changes in parameters that are invariant in the short term (e.g., asymptotic performance). Future research is required to develop integrated theoretical accounts of short-term learning and long-term development (or aging) along the lines described by Newell et al.

Intensive measurement designs offer advantages over conventional single-shot prospective longitudinal designs, such as improved precision for tracking intraindividual change and estimation of changes in asymptotic performance. Processing capacity or latent potential, as indexed by asymptotic performance, might prove to be an especially sensitive marker of aging effects. However, a noteworthy limitation of the present study was the need to constrain r and r^* as fixed to obtain model convergence. In practice, learning rates will likely vary across individuals and within individuals across time. Future studies should consider including more within-burst sessions to facilitate estimation of person-specific learning rates as well as across-burst changes in rates. Despite this limitation, the analyses described in this chapter illustrate the utility of measurement burst designs and the dual-exponential learning function for separating retest performance gains, warm-up effects, and aging declines in asymptotic performance. Thus, the pairing of multiburst designs and informative measurement models may offer an especially useful approach for separating local and global developmental processes that operate across very different time scales.

References

Anderson, J. R., Fincham, J. M., & Douglass, S. (1999). Practice and retention: A unifying analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 1120–1136. doi: 10.1037/0278-7393.25.5.1120

Baltes, P. B., & Nesselroade, J. R. (1979). History and rationale of longitudinal research. In J. R. Nesselroade & P. B. Baltes (Eds.), *Longitudinal research in the study of behavior and development* (pp. 1–39). New York, NY: Academic Press.

Ferrer, E., Salthouse, T. A., McArdle, J. J., Stewart, W. F., & Schwartz, B. S. (2005). Multivariate modeling of age and retest in longitudinal studies of cognitive abilities. *Psychology and Aging, 20,* 412–422. doi:10.1037/0882-7974.20.3.412

Ferrer, E., Salthouse, T. A., Stewart, W. F., & Schwartz, B. S. (2004). Modeling age and retest processes in longitudinal studies of cognitive abilities. *Psychology and Aging, 19,* 243–259. doi: 10.1037/0882-7974.19.2.243

Ghisletta, P., McArdle, J. J., & Lindenberger, U. (2006). Longitudinal cognition–survival relations in old and very old age: 13-year data from the Berlin Aging Study. *European Psychologist, 11,* 204–223. doi:10.1027/1016-9040.11.3.204

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

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. doi:10.1037/0012-1649.25.2.247

MacDonald, S. W. S., Dixon, R. A., Cohen, A., & Hazlitt, J. E. (2004). Biological age and 12-year cognitive change in older adults: Findings from the Victoria Longitudinal Study. *Gerontology, 50,* 64–81. doi:10.1159/000075557

MacDonald, S. W. S., Stigsdotter-Neely, A., Derwinger, A., & Backman, L. (2006). Rate of acquisition, adult age, and basic cognitive abilities predict forgetting: New views on a classic problem. *Journal of Experimental Psychology: General, 135,* 368–390.

McArdle, J. J., & Anderson, E. (1990). Latent variable growth models for research on aging. In J. E. Birren & K. W. Schaie (Eds.), *The handbook of the psychology of aging* (3rd ed., pp. 21–43). New York, NY: Plenum Press.

McArdle, J. J., Ferrer-Caja, E., Hamagami, F., & Woodcock, R. W. (2002). Comparative longitudinal structural analyses of the growth and decline of multiple intellectual abilities over the life span. *Developmental Psychology, 38,* 115–142.

McArdle, J. J., & Woodcock, R. W. (1997). Expanding test–retest designs to include developmental time-lag components. *Psychological Methods, 2,* 403–435. doi:10.1037/1082-989X.2.4.403

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

Newell, K. M., Mayer-Kress, G., & Lui, Y. (2001). Time scales in motor learning and development. *Psychological Review, 108,* 57–82. doi:10.1037/0033-295X.108.1.57

Newell, K. M., Mayer-Kress, G., & Lui, Y. (2006). Human learning: Power laws or multiple characteristic time scale? *Tutorials in Quantitative Methods for Psychology, 2,* 29–39.

Rabbitt, P., Diggle, P., Smith, D., Holland, F., & McInnes, L. (2001). Identifying and separating the effects of practice and of cognitive ageing during a large longitudinal study of elderly community residents. *Neuropsychologia, 39,* 532–543. doi:10.1016/S0028-3932(00)00099-3

Raftery, A. E. (1995). Bayesian model selection in social research (with discussion). *Sociological Methodology, 25,* 111–196. doi:10.2307/271063

Rickard, T. C. (2007). Forgetting and learning potentiation: Dual consequences of between-session delays in cognitive skill learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33,* 297–304. doi:10.1037/0278-7393.33.2.297

Salthouse, T. A., Schroeder, D. H., & Ferrer, E. (2004). Estimating retest effects in longitudinal assessments of cognitive functioning in adults between 18 and 60 years of age. *Developmental Psychology, 40,* 813–822. doi:10.1037/0012-1649.40.5.813

Schaie, K. W. (1965). A general model for the study of developmental problems. *Psychological Bulletin, 64,* 92–107. doi:10.1037/h0022371

Schaie, K. W. (1996). *Intellectual development in adulthood: The Seattle Longitudinal Study*. New York, NY: Cambridge University Press.

Sliwinski, M. J. (2008). Measurement-burst designs for social health research. *Social and Personality Psychology Compass, 2,* 245–261. doi:10.1111/j.1751-9004.2007.00043.x

Sliwinski, M. J., Hofer, S., Hall, C. B., Buschke, H., & Lipton, R. (2003). Modeling memory decline in older adults: The importance of preclinical dementia, attrition and chronological age. *Psychology and Aging, 18,* 658–671. doi:10.1037/0882-7974.18.4.658

Sliwinski, M., Lipton, R. B., Buschke, H., & Stewart, W. F. (1996). The effect of pre-clinical dementia on estimates of normal cognitive function in aging. *Journal of Gerontology: Psychological Sciences, 51B,* P217–P225.

Thorndike, E. L., Bregman, E. O., Tilton, J. W., & Woodyard, E. (1928). *Adult learning*. New York, NY: Macmillan.

Thorvaldsson, V., Hofer, S. M., Berg, S., & Johansson, B. (2006). Effects of repeated testing in a longitudinal age-homogeneous study of cognitive aging. *Journal of Gerontology: Psychological Sciences, 61B,* P348–P354.

Wilson, R. S., Li, Y., Bienias, J. L., & Bennett, D. A. (2006). Cognitive decline in old age: Separating retest effects from the effects of growing older. *Psychology and Aging, 21,* 774–789. doi:10.1037/0882-7974.21.4.774