Neuropsychology

Intraindividual Variability in Reaction Time Predicts Cognitive Outcomes 5 Years Later

Allison A. M. Bielak, David F. Hultsch, Esther Strauss, Stuart W. S. MacDonald, and Michael A. Hunter

Online First Publication, September 20, 2010. doi: 10.1037/a0019802

Neuropsychology

0894-4105/10/$12.00 DOI: 10.1037/a0019802

Intraindividual Variability in Reaction Time Predicts Cognitive Outcomes 5 Years Later

Allison A. M. Bielak
Australian National University

David F. Hultsch, Esther Strauss,
Stuart W. S. MacDonald, and Michael A. Hunter
University of Victoria

Objective: Building on results suggesting that intraindividual variability in reaction time (inconsistency) is highly sensitive to even subtle changes in cognitive ability, this study addressed the capacity of inconsistency to predict change in cognitive status (i.e., cognitive impairment, no dementia [CIND] classification) and attrition 5 years later. ***Method:*** Two hundred twelve community-dwelling older adults, initially aged 64–92 years, remained in the study after 5 years. Inconsistency was calculated from baseline reaction time performance. Participants were assigned to groups on the basis of their fluctuations in CIND classification over time. Logistic and Cox regressions were used. ***Results:*** Baseline inconsistency significantly distinguished among those who remained or transitioned into CIND over the 5 years and those who were consistently intact (e.g., stable intact vs. stable CIND, Wald (1) = 7.91, $p < .01$, Exp(β) = 1.49). Average level of inconsistency over time was also predictive of study attrition, for example, Wald (1) = 11.31, $p < .01$, Exp(β) = 1.24. ***Conclusions:*** For both outcomes, greater inconsistency was associated with a greater likelihood of being in a maladaptive group 5 years later. Variability based on moderately cognitively challenging tasks appeared to be particularly sensitive to longitudinal changes in cognitive ability. Mean rate of responding was a comparable predictor of change in most instances, but individuals were at greater relative risk of being in a maladaptive outcome group if they were more inconsistent rather than if they were slower in responding. Implications for the potential utility of intraindividual variability in reaction time as an early marker of cognitive decline are discussed.

Keywords: inconsistency, cognition, attrition, CIND, mild cognitive impairment

Intraindividual variability in reaction time (RT), or rapid yet reversible changes in performance, has become an intriguing topic for aging researchers (e.g., Hultsch, Strauss, Hunter, & MacDonald, 2008). The magnitude of moment-to-moment variability (also termed *inconsistency*; Hultsch, MacDonald, Hunter, Levy-Bencheton, & Strauss, 2000) follows a U-shaped curve across the life span (e.g., 6–89 years, Li et al., 2004; 6–81 years, Williams, Hultsch, Strauss, Hunter, & Tannock, 2005), and markedly increases in the early to mid-70s (e.g., 75–92 years, Bielak, Hultsch, Strauss, MacDonald, & Hunter, in press; 70–102 years, Lövdén, Li, Shing, & Lindenberger, 2007; 75–89 years, MacDonald, Hultsch, & Dixon, 2003). Furthermore, inconsistency appears to be a relatively stable characteristic of an individual, correlating highly across time points, RT tasks (e.g., Fuentes, Hunter, Strauss, & Hultsch, 2001; Hultsch et al., 2000; Rabbitt, Osman, Moore, & Stollery, 2001), and even within trial (e.g., odd vs. even trials; Jensen, 1992).

More importantly, intraindividual variability is related to a number of other individual indicators, reiterating the trait-like nature of this measure. There is substantial evidence that increased moment-to-moment intraindividual variability is associated with maladaptive behaviors such as poorer cognitive ability (e.g., Hultsch, MacDonald, & Dixon, 2002), poorer physical performance (e.g., Li, Aggen, Nesselroade, & Baltes, 2001), and less lifestyle engagement (e.g., Bielak, Hughes, Small, & Dixon, 2007). It is interesting that the relationship with RT intraindividual variability also extends to health conditions, including various brain disorders (e.g., traumatic brain injury; Stuss, Murphy, Binns, & Alexander, 2003), neurodegenerative diseases (e.g., dementia, Hultsch et al., 2000; Parkinson's disease, Burton, Strauss, Hultsch, Moll, & Hunter, 2006), and symptoms consistent with mild cognitive impairment (MCI; Christensen et al., 2005) where individuals with general neurological dysfunction are more variable than healthy controls. Insults to the frontal lobes appear to be particularly

Allison A. M. Bielak, Ageing Research Unit, Centre for Mental Health Research, Australian National University; David F. Hultsch, Esther Strauss, Stuart W. S. MacDonald, and Michael A. Hunter, Department of Psychology, University of Victoria.

This research was supported by grants from the Canadian Institutes of Health Research (CIHR) to David F. Hultsch, the Alzheimer Society of Canada to Esther Strauss, and the Natural Sciences and Engineering Research Council of Canada to David F. Hultsch and Esther Strauss. Allison A. M. Bielak was supported by a postdoctoral research fellowship from CIHR and doctoral scholarships from the CIHR–Institute of Aging, the Michael Smith Foundation for Health Research (MSFHR), and the British Columbia Medical Services Foundation. Stuart W. S. MacDonald was supported by a Scholar Award from MSFHR. Portions of this investigation were conducted as part of Allison A. M. Bielak's doctoral dissertation. We thank Roger Dixon of the Department of Psychology, University of Alberta, for the use of data from the Victoria Longitudinal Study to establish norms for the classification of cognitive status of the present sample. Esther Strauss passed away in June 2009 prior to preparation of the final version of this article.

Correspondence concerning this article should be addressed to Allison A. M. Bielak, Ageing Research Unit, Centre for Mental Health Research, Building 63, Eggleston Road, Australian National University, Canberra ACT 0200, Australia. E-mail: allison.bielak@anu.edu.au

related to increased intraindividual variability, as frontal lobe dementia patients and those with frontal lesions have been shown to exhibit greater inconsistency on cognitive tasks than respective patients with Alzheimer's disease (Murtha, Cismaru, Waechter, & Chertkow, 2002) and nonfrontal lesions (Stuss et al., 2003). Although there have been fewer investigations into direct links between intraindividual variability and the brain, similar relationships have been found in the cognitive neuroscience literature, where inferior structural (e.g., Anstey et al., 2007), neuromodulatory (e.g., MacDonald, Cervenka, Farde, Nyberg, & Bäckman, 2009), and functional brain characteristics (e.g., Kelly, Uddin, Biswal, Castellanos, & Milham, 2008; MacDonald, Nyberg, Sandblom, Fischer, & Bäckman, 2008) were associated with greater intraindividual variability in performance.

Overall, the similarity of findings across the cognitive, neuropsychological, and neurobiological research domains demonstrates the likelihood that behavioral within-trial intraindividual variability has neural origins (see MacDonald, Nyberg, & Bäckman, 2006, for a review). Although the precise neurological cause of inconsistency is still unknown, intraindividual variability in response speed is believed to be a behavioral indicator of neurological integrity, where increased fluctuation in performance is indicative of brain disturbance or dysfunction (e.g., Hultsch & MacDonald, 2004; Li & Lindenberger, 1999). Consequently, intraindividual variability may be an early marker of impending disease and behavioral impairment.

A number of recent studies have advanced this predictive possibility by demonstrating that short-term intraindividual variability significantly covaries with long-term cognitive change in healthy older adulthood (Bielak et al., in press; Lövdén et al., 2007; MacDonald et al., 2003). Each study found that this relationship followed the expected negative relationship; on occasions when trial-to-trial inconsistency on an RT task was high, cognitive ability was correspondingly low. Therefore, performance variability and various cognitive outcomes appear to track together over time, a critical precursor to establishing any predictive relationships. Furthermore, because the association was found across a range of time intervals (3 years, MacDonald et al., 2003; 2 years, Lövdén et al., 2007; 1 year, Bielak et al., in press), intraindividual variability appears to be highly sensitive to even subtle changes in cognitive ability.

Given the sensitive nature of the coupling relationship, is intraindividual variability in RT also sensitive to change in other meaningful outcomes, such as incipient disease? As noted above, intraindividual variability has been linked to a number of maladaptive outcomes, but these associations have been based on cross-sectional research. Many have echoed Nesselroade and Salthouse's (2004) claim that the literature needs to move "toward the building of predictive relationships" (e.g., Hultsch & MacDonald, 2004; Martin & Hofer, 2004, p. P54), but only one study has investigated the possible predictive links. MacDonald, Hultsch, and Dixon (2008) found that intraindividual variability in cognitive performance gradually increased with each additional year closer to death and could predict impending death up to 15 years later. Together with these novel results, the significant cross-sectional associations with various conditions, and the demonstration that baseline inconsistency precedes and predicts cognitive decline but baseline cognitive ability has little influence on changes in inconsistency (Lövdén et al., 2007), the potential for inconsistency to predict other meaningful outcomes is promising.

With specific regard to cognitive aging, a key interest concerns whether inconsistency can predict changes related to cognitive decline, and thereby possibly serve as an early indicator of mild cognitive impairment or outcomes associated with declining health. Two classification schemes were evaluated in the present study. First, the MCI classification remains controversial as the stability of the classification is poor (e.g., see Tuokko & Hultsch, 2006), making it difficult to know which individuals truly are displaying symptoms of MCI. However, given that individuals with various subtypes of MCI tend to be more variable than healthy older adults at one point in time (e.g., Dixon et al., 2007), and also the possibility that intraindividual variability is a sensitive marker of neurological integrity, inconsistency may also be attuned to poorer yet reliable change patterns over 5 years, thus identifying individuals with valid MCI classifications. Next, Sliwinski, Hofer, Hall, Buschke, and Lipton (2003) noted that because deleterious events and pathological changes are more likely to occur in older adulthood, there is an increased likelihood of individuals dropping out of longitudinal studies. Rabbitt, Watson, Donlan, Bent, and McInnes (1994) found that the onset and progress of pathologies were the main reasons for withdrawal from longitudinal studies in older age. Moreover, Rabbitt, Lunn, and Wong (2005) found the effects of impending death and dropout on prior cognitive performance to be identical. Consequently, attrition may be a reflection of nonnormative influences on cognitive decline in older adulthood, such as disease and cognitive impairment (MacDonald et al., 2003). Therefore, just as attrition appears to be an indicator of underlying pathology, intraindividual variability in response speed may be predictive of attrition.

There has also been much discussion in the intraindividual variability literature about whether level of speeded performance (i.e., mean RT) provides the same information as inconsistency. Because the measures are typically highly related to one another (i.e., a wider response range increases the mean and standard deviation), some have argued that short-term variability may not offer any unique predictive power beyond the mean (e.g., Salthouse, Nesselroade, & Berish, 2006). Although many studies have found that inconsistency does predict outcome measures independent of mean level (e.g., Burton, Strauss, Hunter, & Hultsch, 2009; Li et al., 2001; MacDonald, Hultsch, et al., 2008), some studies have found negligible predictive differences between the two measures (e.g., Bielak, Hughes, et al., 2007; Christensen et al., 2005). However, in possibly the strongest investigation to date, Lövdén et al. (2007) found higher inconsistency reliably preceded and predicted subsequent decline in cognition, but mean performance did not significantly predict later cognitive decline. Therefore, inconsistency may be stronger at reliably predicting meaningful change outcomes than the mean level of speeded performance.

Finally, a number of research studies have found that differences in intraindividual variability are most apparent on speeded tasks that challenge cognitive functioning, such as those that place large demands on executive functioning or working memory. For example, across three RT tasks of varying difficulty, the largest age effects occurred on the task that required ignoring the present stimulus and instead responding to the stimulus from the previous trial (Dixon et al., 2007). Similar findings have been found for neurological conditions, where the greatest distinctions among

groups were on cognitively complex tasks (e.g., dementia vs. healthy, Hultsch et al., 2000; dementia subtypes, Murtha et al., 2002; cognitive impairment, no dementia [CIND] status, Strauss, Bielak, Bunce, Hunter, & Hultsch, 2007). Bielak et al. (in press) found that the complexity of the inconsistency measure also affected the covariation relationships with cognition, as inconsistency derived from moderately and highly complex tasks shared consistently stronger change relationships with cognition than those derived from the simplest RT tasks. It may be that inconsistency in completing cognitively challenging tasks is most attuned to the integrity of the neurological system and thus also serves as the best predictor of meaningful change.

The present study addressed these issues by investigating whether meaningful 5-year change outcomes could be predicted by intraindividual variability in RT, and compared these results with those using mean level of performance. We focused on three research questions for each of two outcomes: (a) cognitive status change and (b) attrition. First, does inconsistency predict change 5 years later? Generally speaking, we expected inconsistency to significantly differentiate among individuals who (a) remained classified as CIND, became CIND over time, fluctuated in status classification, or maintained an intact classification; and (b) remained in the study from those who dropped out. In each instance we expected those with maladaptive change patterns (e.g., transitioned into CIND, dropped out) to show greater inconsistency in responding.[1] Next, is inconsistency a better predictor of change group than mean level of performance? The finding that inconsistency, but not mean level, reliably preceded and predicted subsequent cognitive decline (Lövdén et al., 2007) led us to expect that inconsistency would similarly be a better predictor of the present outcomes. Third, are the predictive strengths of inconsistency different based on the cognitive effort of the tasks from which they are derived? We expected inconsistency based on the moderately and highly complex tasks to have better predictive ability than estimates derived from the simpler RT tasks.

Method

Participants

The sample began with 304 community-dwelling older adults initially 64 to 92 years of age (208 women, 96 men). Participants resided in the region of Victoria, British Columbia, Canada, and were originally recruited through advertisements in the local media requesting healthy volunteers who were concerned about their mental functioning. Initial exclusionary criteria included a diagnosis of dementia by a physician or a Mini-Mental State Examination (Folstein, Folstein, & McHugh, 1975) score less than or equal to 24, a history of significant head injury (i.e., loss of consciousness for more than 5 min), other neurological or major medical illness (e.g., Parkinson's disease, cancer), severe sensory impairment even with corrective aids (e.g., difficulty reading newspaper-size print while wearing glasses), drug or alcohol abuse, a current psychiatric diagnosis, psychotropic drug use, and lack of fluency in English.

Two hundred twelve participants completed all of the relevant tests 5 years later.[2] By this time, the majority of participants were now in the old–old age group classification: young–old (65–74), $n = 65$, $M = 72.38$ years, $SD = 1.35$; old–old (75+), $n = 147$, $M = 81.31$ years, $SD = 4.67$. Women made up the majority of the sample (69.8%). The participants were well educated ($M = 15.26$ years, $SD = 3.02$), ranging from 7 to 24 years of education, with only 9.4% having fewer than 12 years of schooling. All of the participants were Caucasian. The participants were relatively healthy, with few reported chronic disorders ($M = 3.22$, $SD = 1.94$). The participants' perceptions of their relative health were also high, with 80.7% rating themselves good or very good compared with a perfect state of health, and 94.3% rating themselves good or very good compared with the health of others their own age. Finally, most participants reported that they were very capable of completing instrumental activities of daily living (Lawton & Brody, 1969); more than 88% were able to shop, prepare food, complete laundry and housekeeping, oversee their own medications and finances, and drive or travel independently.

Procedure

Potential participants were initially screened for inclusion and exclusion criteria by a telephone interview. The measures were administered during seven sessions (one group and six individual) scheduled over approximately 3 months. The entire testing battery was repeated yearly three times, totaling four waves of data. The fifth testing wave involved only a brief telephone interview, but Wave 6 involved two group testing sessions during which a number of relevant tasks were readministered. Data from Waves 1, 3, 4, and 6 were used to construct the cognitive status change outcome groups; baseline inconsistency and mean level were derived from Wave 1 data; and inconsistency and mean level values from Waves 2, 3, and 4 were also used to contribute to the attrition analyses.

Construction of Outcome Groups

The outcome *cognitive status change* was based on potential variations in the participants' cognitive status classification over the 5 years. First, cognitive status at each testing wave was determined by participants' performance on five tests assessing different cognitive domains: perceptual speed (Wechsler Adult Intelligence Scale—Revised Digit Symbol Substitution; Wechsler, 1981), inductive reasoning (Letter Series; Thurstone, 1962), episodic memory (immediate free word recall; Hultsch, Hertzog, & Dixon, 1990), verbal fluency (Controlled Associations; Ekstrom, French, Harman, & Dermen, 1976), and vocabulary (Extended Range Vocabulary; Ekstrom et al., 1976). These tasks were administered only once per testing wave in a group format. Each participant's performance was compared with norms obtained from the Victoria Longitudinal Study (Dixon & de Frias, 2004), which included an independent sample of 445 adults 65 to 94 years

[1] In light of the lack of consensus surrounding preclinical cognitive impairment, we have chosen to use the term *cognitive impairment, no dementia* (CIND), which is a more general, inclusive term that encompasses many of the various definitions (Tuokko & Frerichs, 2000), including MCI (Petersen et al., 1999).

[2] Six additional participants completed various components of the test battery, but did not complete all relevant measures, and were excluded from the cognitive status change analyses. They were, however, included in the attrition analyses.

of age drawn from the same population.[3] Although published norms are available for most of these measures, they are derived from a variety of different samples with varying demographic characteristics. The use of local norms is preferred given the close demographic match to the current sample and the ability to make more accurate comparisons across tasks. Furthermore, the tests were conormed, providing an additional advantage. The normative sample was partitioned into four age by education groups (age groups = 65–74 years and 75+ years; education groups = 0–12 years and 13+ years), and means and standard deviations were computed for these groups for the five measures. These normative values were then used to classify the present sample.

There were two ways an individual could be classified as possible CIND. First, participants whose performance was more than 1.5 *SD* below the mean of their age- and education-matched peers on only one cognitive task fit the criteria for CIND-single. The 1.5 *SD* criterion has been widely used in the clinical literature (e.g., Petersen et al., 1999; Tuokko, Gabriel, & the Canadian Study of Health and Aging Neuropsychology Working Group, 2006) and represents a stricter criterion than those previously used with this data set (e.g., 1.0 *SD*; Bielak, Hultsch, Kadlec, & Strauss, 2007). Second, participants who scored more than 1.0 *SD* below the normative sample on two or more cognitive tasks fit the criteria for CIND-multiple. The CIND-multiple cutoff was 1.0 *SD* as previous studies with this data set have demonstrated significant cognitive differences between the CIND-multiple group and healthy older adults (e.g., Strauss et al., 2007), and the 1.0 *SD* CIND-multiple guideline has been used alongside the 1.5 *SD* CIND-single classification in population-based studies (e.g., Canadian Study of Health and Aging). All remaining participants who did not fit the criteria for either CIND group were classified as cognitively intact (Intact). Classifications were completed for each wave.

Next, for the purposes of detecting change in classification over time, the CIND-single and CIND-multiple groups were collapsed into one CIND group. This procedure follows that employed by the Canadian Longitudinal Study of Health and Aging (see Tuokko et al., 2006), and was necessary given the possibility in our study to be both CIND-single and CIND-multiple at the same wave of testing (i.e., a participant could show greater than 1.5 *SD* impairment in only one cognitive domain and also show greater than 1.0 *SD* impairment in two or more cognitive domains). Change in cognitive status was assessed using classifications from Waves 1, 3, 4, and 6. Because the classification of cognitive impairment is imprecise and thus prone to fluctuation, at least four change patterns warranted comparison (see Figure 1). First, individuals who consistently showed cognitive ability on par with or better than their peers (stable intact, $n = 118$) and those who consistently showed poorer cognitive ability relative to their peers (stable CIND, $n = 21$) likely represented two distinct groups. Next, there were two possibilities of fluctuating classification patterns between these two extremes: individuals whose classifications changed across the waves (fluctuating, $n = 58$), and those who became CIND over time and continued to be classified this way (stable decline, $n = 15$).

Attrition was based on whether participants remained in the study and completed Wave 6 testing. There were 86 participants who did not complete the Wave 6 test battery, leaving 218 continuing participants.

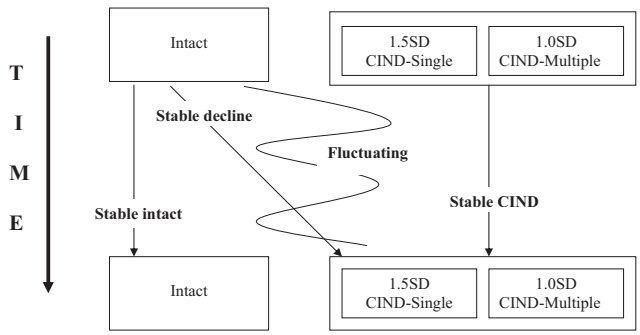

Figure 1. Possible change trajectories in cognitive status. CIND = cognitive impairment, no dementia.

Intraindividual Variability and Intraindividual Mean

Intraindividual variability and mean were calculated from RT latencies from several multitrial computer-based tasks. For all tasks except finger tapping, participants were instructed to emphasize speed in responding while minimizing errors. These tasks were administered five times per testing Waves 1–4, approximately 2 weeks apart, in individually administered sessions.

Finger tapping. Participants were instructed to tap a response key as rapidly as possible, first with their left hand and then with their right hand (47 taps/hand).

Four-choice reaction time (CRT). Participants were presented with a horizontal row of four plus signs, with a matching arrangement of keys on the response console. One plus sign changed into a box, and the participant was required to press the key corresponding to its location (60 trials).

Four-choice reaction time 1-back (BRT). The same stimulus display and response keyboard as CRT were used, but participants were instructed to press the key corresponding to the location of the box on the previous trial (60 trials).

Shape, color, and task switching. Figures varying in shape (square, circle) and color (red, green) were presented in a white frame with a cue indicating the currently relevant stimulus dimension (shape, color) above it. Participants were required to press the right-hand key for circles and red objects and the left-hand key for squares and green objects. Three conditions of 52 trials each were presented: (a) respond to the *shape* of the figure while ignoring the color; (b) respond to the *color* of the figure while ignoring the shape; (c) the relevant response dimension varied randomly (*task-switching*).

Data Preparation

The distributions of raw latency scores were first examined for outliers. Extremely fast or slow responses may represent various

[3] Data on all 445 participants were available for the measures of perceptual speed, inductive reasoning, verbal fluency, and vocabulary, but because of a counterbalancing procedure, information on the episodic memory task was available for only 194 of the 445 participants. Individuals participating in the longitudinal study from which the normative sample was drawn were not accepted into the current sample.

sources of measurement error (e.g., accidental key press, distraction). Lower bounds for legitimate responses were suggested by prior research (150 ms; Hultsch et al., 2002), and were employed for all tasks except finger tapping, which had no lower limit for responding. Upper boundaries involved computing the mean and standard deviation for each task and occasion for each age group and removing any trials that exceeded the mean by 3 or more *SD*s.

Missing value estimates were imputed using a regression substitution procedure that forms individual equations of RTs across all trials, which were then used to predict the missing RT entry (Hultsch et al., 2000). The procedures for eliminating outlying trials and imputing missing values decreases within-subject variation, thus representing a conservative approach to examining inconsistency.

Computation of intraindividual variability. Inconsistency was computed as the across-trial intraindividual standard deviation (*ISD*) about each individual's mean RT. Other techniques to calculate inconsistency have been statistically criticized, and our approach follows an alternative logic (see Hultsch et al., 2008). Potential confounding influences (e.g., age differences in mean RT, practice effects) and their higher order interactions were partialed out using a split-plot regression:

$$Y = a + b(\text{age group}) + c(\text{trial}) + d(\text{Age Group} \times \text{Trial}) + e.$$

By using the residuals produced from this regression, we effectively removed any systematic within- (i.e., trial) and between-subjects (i.e., age group) sources of variance in mean RT, leaving only the possibility of evaluating each individual's unsystematic portion (i.e., inconsistency). Although prior investigations into dementia and cognitive classification have also purified by cognitive status (e.g., Strauss et al., 2007), the present outcomes were based on changes over time, precluding the ability to use them for baseline *ISD* calculations. Furthermore, the largest systematic effects have been found in relation to trial and interactions with trial (Hunter & Bielak, 2005), indicating that purification by group causes minimal changes to the resulting residuals. The residuals were converted to standardized *T* scores ($M = 50$, $SD = 10$) to enable comparisons across tasks, and each individual's standard deviation was calculated. *ISD* values were computed for each task, for each session, for each wave.

To obtain the most reliable estimate of inconsistency at each wave, we individually averaged the *ISD*s across the five sessions for each task, producing one *ISD* score per task per wave for each individual. Session effects were not investigated because week-to-week variability is likely the result of different influences (e.g., stress, fatigue) than those observed moment-to-moment (i.e., brain-based influences; Hultsch et al., 2008).

Next, to reduce the large number of *ISD* values that would be investigated across time (i.e., one for each of the 10 RT tasks), we calculated *ISD* composites across the tasks. The composites were based on factor structures identified by Strauss et al. (2007) using the same sample of participants: motor *ISD*, right and left finger tapping; basic *ISD*, color and shape RTs and CRT; and complex *ISD*, BRT and two-choice switch RT.

Computation of intraindividual mean. Following data preparation procedures, we converted the individual RT trials for each task at each session for each wave to standardized *T* scores to enable comparisons across tasks and with the *ISD* values. An intraindividual mean (*IM*) was then computed in the traditional manner as the mean RT of each individual's standardized latencies across all trials for each task for each session at each wave. Next, composites were calculated according to the same three factors described for intraindividual variability (i.e., motor, basic, and complex).

Statistical Analyses

A series of multinomial logistic regression analyses were used to evaluate cognitive status change. Given the sensitivity of logistic regression to strong multicollinearity among the predictor variables and the range of correlations between the *ISD* and *IM* composites (motor, $r = .68$; basic, $r = .85$, complex, $r = .90$), there were no models that involved both of the corresponding *IM* and *ISD* composites (e.g., complex *IM*, complex *ISD*). Rather, the overall models of all *ISD* composites and age group were compared with the overall models of the *IM* composites and age group,[4] and each composite was also evaluated individually using multinomial logistic regression. Cox regression was used to evaluate attrition. This analysis has the advantage of not only providing the knowledge of *whether* an outcome is likely to occur, but also insight into *when* the outcome might occur. The proportional-hazards assumption was violated for a number of predictors, requiring Cox regression with time-dependent covariates.

Results

The results are presented in two main parts as follows for the two outcomes of cognitive status change and attrition. For cognitive status change, we first investigated the ability of Wave 1 intraindividual variability to predict the outcome groups and contrasted that with the ability of the Wave 1 *IM*. Second, we examined possible differences in prediction strength due to task complexity. The same plan of analysis was followed for the attrition outcome, but intraindividual variability and mean indices were based on Waves 1–4 to allow time variation of these measures in Cox regression. Tests of significance were adjusted for Type I error ($\alpha = .013$) for the four predictor *ISD* and *IM* models.

Cognitive Status Change

As a group, Wave 1 motor, basic, and complex *ISD*s and age group were able to significantly predict cognitive status change group, $\chi^2(12, N = 212) = 37.62$, $p < .001$, Nagelkerke's $R^2 = .18$. Age group and complex *ISD* were noted as being particularly important to the model, as removal of either significantly reduced the model's power: age group, $\chi^2(3, N = 212) = 14.73$, $p < .005$; complex *ISD*, $\chi^2(3, N = 212) = 11.55$, $p < .01$. Given that age group was not the main focus of the present analyses, its significance and various group comparisons are only noted in the individual model comparisons. Complex *ISD* was able to significantly distinguish between the stable intact and stable CIND groups, $\beta = 0.40$, $SE = 0.14$, Wald (1) = 7.91, $p < .01$, Exp(β) = 1.49, where each 0.1 *SD* increase in complex *ISD* increased the likeli-

[4] The assumption of an absence of high multicollinearity was met for the correlations between the composites.

hood of being in the stable CIND group compared with the stable intact group by 49%. A similar increase in risk was seen between the stable intact and fluctuating groups, $\beta = 0.25$, $SE = 0.10$, Wald (1) = 6.02, $p = .014$, Exp(β) = 1.28.

In comparison, the *IM* model was also significant, $\chi^2(12, N = 212) = 34.79$, $p < .01$, Nagelkerke's $R^2 = .17$, and removing either complex *IM* or age group caused significant disruptions to the overall model, $\chi^2(3, N = 212) = 18.78$, $p < .001$, and $\chi^2(3, N = 212) = 11.59$, $p < .01$, respectively. Complex *IM* was able to distinguish between stable intact and stable CIND, $\beta = 0.24$, $SE = 0.07$, Wald (1) = 12.15, $p < .001$, Exp(β) = 1.28, and stable intact and fluctuating groups, $\beta = 0.17$, $SE = 0.05$, Wald (1) = 10.39, $p < .001$, Exp(β) = 1.18. Therefore, for each unit increase in complex *IM*, there was a 28% increased likelihood of being in the stable CIND group compared with the stable intact group, and an individual was 18% more likely to be in the fluctuating group compared with the stable intact group. Overall, there was a trend toward the *ISD* model doing a slightly superior job at predicting cognitive status change group ($\chi^2 = 37.62$, $R^2 = .18$) than the *IM* model ($\chi^2 = 34.79$, $R^2 = .17$), and the related odds ratios were stronger for the *ISD* measures.

Individually, the basic *ISD*, $\chi^2(3, N = 212) = 15.24$, $p < .01$, Nagelkerke's $R^2 = .08$, and complex *ISD*, $\chi^2(3, N = 212) = 18.62$, $p < .001$, Nagelkerke's $R^2 = .10$, composites were significant. Although the complex *ISD* model had the largest effect size, the basic *ISD* model had stronger odds ratios in distinguishing the groups. Specifically, basic *ISD* was able to significantly distinguish the stable intact group from the three other groups: fluctuating, $\beta = 0.26$, $SE = 0.10$, Wald (1) = 6.66, $p < .05$, Exp(β) = 1.29; stable decline, $\beta = 0.48$, $SE = 0.16$, Wald (1) = 8.50, $p < .01$, Exp(β) = 1.61; and stable CIND, $\beta = 0.34$, $SE = 0.14$, Wald (1) = 5.63, $p < .05$, Exp(β) = 1.40. Therefore, a 1-point increase in basic *ISD* score at Wave 1 increased an individual's likelihood of being in the fluctuating group by 29%, the stable decline group by 61%, and the stable CIND group by 40%. The Wald statistics and odds ratios for the other *ISD* measures were in the same direction and accounted for similar group comparisons (see Table 1).

For the *IM*s, the complex model was significant, $\chi^2(3, N = 212) = 20.83$, $p < .001$, Nagelkerke's $R^2 = .11$, followed by the basic model, $\chi^2(3, N = 212) = 7.27$, $p = .06$, Nagelkerke's $R^2 = .04$, in predicting cognitive status change group. Complex *IM* was able to significantly distinguish the stable intact group from the fluctuating group, $\beta = 0.11$, $SE = 0.04$, Wald (1) = 8.32, $p < .01$, Exp(β) = 1.12, the stable decline group, $\beta = 0.14$, $SE = 0.06$, Wald (1) = 5.61, $p < .05$, Exp(β) = 1.15, and the stable CIND group, $\beta = 0.20$, $SE = 0.05$, Wald (1) = 14.43, $p < .001$, Exp(β) = 1.22. Therefore, compared with the stable intact group, each 0.1 *SD* millisecond increase in complex *IM* increased an individual's likelihood of being in the fluctuating group by 12%, the stable decline group by 15%, and the stable CIND group by 22%. Consequently, it appeared that the complexity of the RT tasks was important in providing the strongest prediction of later cognitive status change group. However, in contrast to *ISD* where moderately complex tasks were superior, highly complex tasks provided the best differentiation for *IM*.

As an aside, age group at Wave 1 significantly predicted cognitive status change group, $\chi^2(3, N = 212) = 8.01$, $p < .05$, Nagelkerke's $R^2 = .04$, specifically distinguishing the stable intact group from the stable decline group, $\beta = -1.21$, $SE = 0.58$, Wald (1) = 4.38, $p < .05$, Exp(β) = 0.30, and the fluctuating from the stable decline group, $\beta = -1.66$, $SE = 0.62$, Wald (1) = 7.12, $p < .01$, Exp(β) = 0.19. Compared with being in the stable intact group, a young–old participant was 70% less likely to be in the stable decline group, and compared with the fluctuating group, a young–old participant was 81% less likely to be in the stable decline group.

Attrition

Figure 2 shows the survival function for remaining in the study across the testing waves, divided according to age group at Wave 1. There was a significant age group effect ($p < .01$), with members of the old–old age group at baseline demonstrating an 87% higher rate of attriting than members of the young–old group. It is apparent that cumulative survival dropped more between Waves 1 and 2 (i.e., between 0 and 365 days) than the later waves of testing, indicating that most participants who did not continue chose to do so early on. However, both age groups appear to follow a similar rate of dropout after this time interval.

Table 1
Odds Ratios of Year 1 Intraindividual Standard Deviation (ISD) and Intraindividual Mean (IM) Composites Individually Predicting Cognitive Status Change Groups

	Reaction time measure	
Composite	*ISD* Exp(β)	*IM* Exp(β)
Motor	Stable intact vs. stable CIND = 1.22, $p = .05$	Stable intact vs. stable decline = 1.08, $p = .08$
Basic	Stable intact vs. fluctuating = 1.29* Stable intact vs. stable decline = 1.61** Stable intact vs. stable CIND = 1.40*	Stable intact vs. stable decline = 1.14, $p = .07$ Stable intact vs. stable CIND = 1.14*
Complex	Stable intact vs. fluctuating = 1.20** Stable intact vs. stable decline = 1.32** Stable intact vs. stable CIND = 1.38**	Stable intact vs. fluctuating = 1.12** Stable intact vs. stable decline = 1.15* Stable intact vs. stable CIND = 1.22*** Fluctuating vs. stable CIND = 1.10, $p = .09$

Note. 1 unit = 0.1 *SD*; CIND = cognitive impairment, no dementia. The reference group is listed first. Only significant or nearly significant comparisons are shown.
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

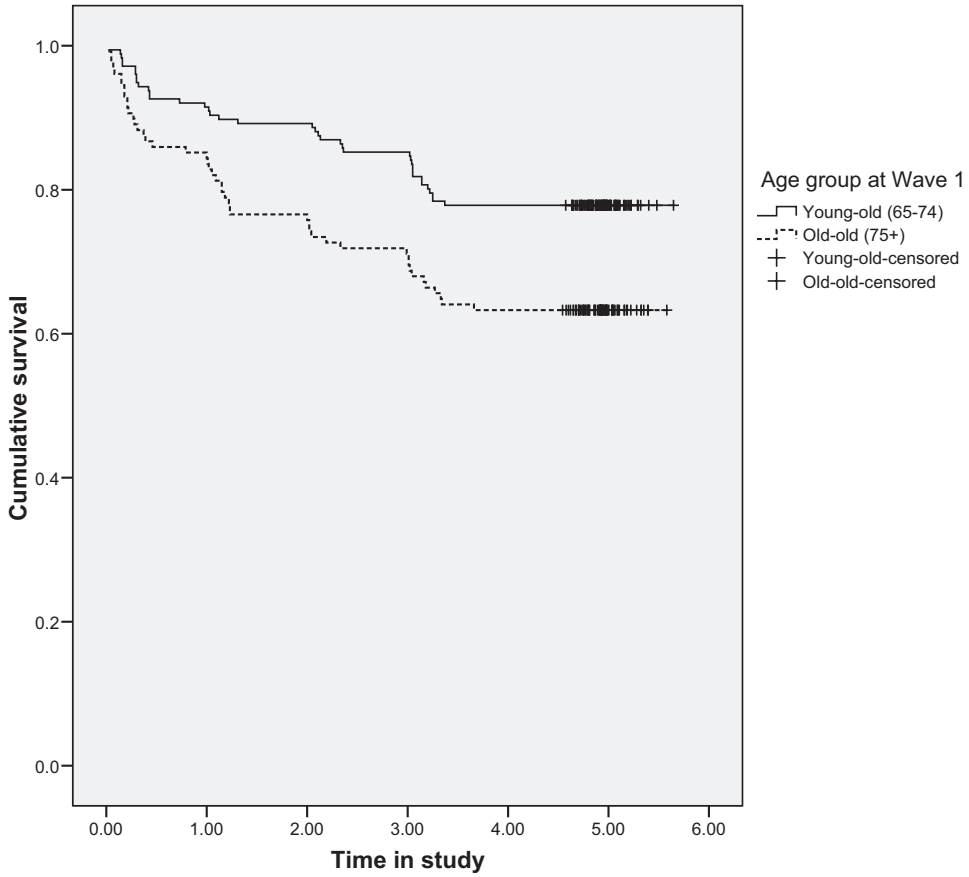

Figure 2. Cumulative survival function of attrition by time in study according to age group at Wave 1. Time in study "0" refers to Wave 1 testing. Participants did not complete the study at precise yearly intervals, resulting in some participants' time in study appearing to be past Wave 6 data collection. Censored refers to individuals who completed all waves of testing.

The Cox regression model with time-dependent covariates included all three *ISD* composites across the four waves and age group at Wave 1. The overall model was significant, $\chi^2(4, N = 304) = 25.17$, $p < .001$. Table 2 shows the statistics for each predictor, holding the other predictors constant. It is interesting that age group, motor *ISD*, and complex *ISD* did not significantly contribute any unique variance to predicting attrition. Rather, participants who showed greater average inconsistency over time on the basic *ISD* tasks were at a substantially greater risk of dropping out of the study; each 0.1 *SD* increase in basic *ISD* above the mean *ISD* (i.e., 7.51 in *T* score units) uniquely increased the rate of attriting by 24%.

In comparison, the Cox regression model including all three time-dependent *IM* composites and age group was also significant, $\chi^2(4, N = 304) = 25.20$, $p < .001$. Table 2 shows the unique effects of all predictors, but only basic *IM* was uniquely significant in predicting rate of attrition. For each 0.1 *SD* that a participant's average mean speed of responding on the basic tasks over the four waves was above the mean *IM* (i.e., 49.19 in *T* score units), the rate of attriting uniquely increased by 10%. Overall, although *IM* and *ISD* showed identical overall model effects, the associated hazard ratios were larger for the *ISD* measure.

Cox regressions with individual continuous time-varying predictors were run to investigate whether the complexity of the RT tasks played a role in the strength of predicting the rate of attrition. All of the *ISD* composites were significant predictors in the expected direction (i.e., ps $< .01$; increased *ISD*, increased risk of attrition), but the results corresponded to those reported in the multivariable analyses: Motor and complex *ISD* had smaller hazard ratios, $\text{Exp}(\beta) = 1.11$ for both, compared with the basic measure, $\text{Exp}(\beta) = 1.26$. Overall, basic *ISD* offered the best prediction of rate of attrition, and based on the multivariable analysis, accounted for the largest amount of unique variance. The results for the *IM* measures were not as strong, as only the basic composite model was significant, $p < .001$, $\text{Exp}(\beta) = 1.09$. Therefore, for both measures, the strongest predictor of rate of attrition involved the moderately complex tasks.

Discussion

Through a series of 5-year prospective analyses, the present study evaluated the validity of the hypothesis that intraindividual variability in RT is indicative of change in cognitive status and attrition from the study. We also compared this prediction with that

Table 2
Model Statistics of Intraindividual Standard Deviation (ISD) and Intraindividual Mean (IM) Composites and Wave 1 Age Group Predicting Rate of Attrition, Controlling for Other Predictors

Model	β	*SE*	Wald	Exp(β)	95% CI
ISD					
Motor ISD	.03	.04	0.44	1.03	[0.95, 1.12]
Basic ISD	.22	.06	11.31**	1.24	[1.09, 1.41]
Complex ISD	−.04	.05	0.55	0.97	[0.88, 1.06]
Age group[a]	.30	.25	1.41	1.35	[0.83, 2.19]
IM					
Motor IM	−.02	.02	1.05	0.98	[0.95, 1.02]
Basic IM	.09	.02	16.29***	1.10	[1.05, 1.15]
Complex IM	−.02	.02	1.18	0.99	[0.96, 1.01]
Age group[a]	.43	.23	3.45	1.54	[0.98, 2.43]

Note. 1 unit = 0.1 *SD*; Exp(β) = hazard ratio.
[a] Reference category for age group = young–old.
** $p < .01$. *** $p < .001$.

attained by intraindividual mean level of performance, and potential differences in the predictive relationship due to task complexity were evaluated.

Cognitive Status Change

Our hypothesis that baseline inconsistency may be able to identify individuals with reliable CIND classifications (i.e., those who maintain CIND status or become CIND over time) was supported. For example, compared with the stable intact group, each 0.1 *SD* increase in baseline basic *ISD* score increased an individual's likelihood of being in the fluctuating group by 29%, the stable decline group by 61%, and the stable CIND group by 40%.

These findings demonstrate that within-trial intraindividual variability may be a valuable tool in predicting preclinical dementia. This finding is particularly noteworthy because, at the moment, clinicians are unable to reliably classify individuals as CIND (e.g., Tuokko & McDowell, 2006). However, because initial inconsistency did not completely distinguish the groups (i.e., there were only three significant distinctions among the four change groups for even the strongest *ISD* composite), the predictive power of the findings must be tempered. Rather than this indicating unreliability of the link between inconsistency and later CIND status, however, this more likely represents difficulties in determining the cognitive status change groups. For example, the change groups were expected to follow a continuum of severity in cognitive impairment (i.e., stable intact < fluctuating < stable decline < stable CIND). However, the largest implication for greater baseline variability was an increased likelihood of being in the stable decline group (i.e., becoming CIND over the 5 years) rather than being at the end of the continuum as stable CIND. Furthermore, greater baseline inconsistency did not differentiate among the three "impaired" change groups, suggesting superior groupings may exist. However, the finding that initial inconsistency significantly differentiated each of these groups from the one "healthy" group (i.e., stable intact) verifies the predictive utility of intraindividual variability over longer time periods. Although it remains to be seen whether all of the CIND change groups go on to develop dementia, inconsistency in cognitive speed could be used as an indicator of those "at risk" for later cognitive decline.

Another interesting finding was the significant odds ratio found between the stable intact and fluctuating groups. The fluctuating group included individuals who changed from intact to CIND and back to intact (or vice versa) at least once over the four time points. Given the poor stability of CIND status, fluctuation in cognitive status classification was expected, but it was uncertain whether this type of pattern reflected anything more than individuals having random fluctuations in their performance. Clearly, instability in cognitive status was also meaningful and potentially indicative of the initial stages of neurological disturbance. Because CIND status is believed to encompass symptoms that are risk factors for later dementia, the present results are congruent with the hypothesis that moment-to-moment fluctuations in cognitive performance are the behavioral manifestations of neurological dysfunction.

Attrition

Intraindividual variability in response speed across the waves was able to significantly differentiate those who dropped out of the study from those who remained in the study. For example, being 1 *SD* above the sample mean on basic *ISD* across the waves uniquely increased the risk of attriting by 240%. Given these impressive findings, the considerable evidence showing links between inconsistency and various neurological problems (e.g., traumatic brain injury, Stuss et al., 2003; dementia, Hultsch et al., 2000) and brain characteristics (e.g., corpus callosum size, Anstey et al., 2007; brain activation, Bellgrove, Hester, & Garavan, 2004; regulation of competing neural processes, Kelly et al., 2008), and the findings that all maladaptive change patterns of cognitive status in the present study showed higher baseline inconsistency, attrition did appear to be a reasonable proxy for impending health problems (e.g., Sliwinski et al., 2003). If participants were dropping out of the study for normative reasons such as lack of interest or time constraints, we would not have found such a strong prediction of attrition by inconsistency.

Table 3 summarizes the self-reported reasons for participant dropout at each wave. Although a number of participants provided normative justification (i.e., busy or not interested, family health problems, moved, other) for not continuing in the study after the first testing wave, this number was not grossly different from those reporting nonnormative reasons across the testing waves (i.e., died,

Table 3
Number of Reported Reasons for Attrition

		Attrited	
Wave	Returned	Normative	Nonnormative
1	304	–	–
2	270	25	9
3	256	7	7
4	239	7	10
5	234	1	4
6	218	11	5

Note. Normative = busy or not interested, family health problems, moved, other; Nonnormative = died, memory or health problems.

memory or health problems). We investigated the possibility of a discrepancy between participants' reported reasons for dropout and potential actual explanations by looking more closely at 21 of the participants who dropped out for normative reasons after Wave 1.[5] Although these participants described themselves as busy or not interested in further participating, one third demonstrated poorer cognitive ability than their peers (i.e., CIND-multiple at Wave 1), almost half had three or more chronic health conditions, and nearly one third completed fewer than 13 years of education. Therefore, some of the participants who reported normative dropout had potentially poorer health and cognitive abilities. Together with the strong demonstrated relationship between attrition and inconsistency, these data support the possibility that not all participants reported the true reasons for removing themselves from the study.

The present findings are consistent with those by MacDonald and colleagues (2003), who found that individuals who dropped out of a 6-year longitudinal study showed greater fluctuations in their cognitive performance at baseline. Furthermore, these results are in line with findings that inconsistency significantly increased per additional year closer to death, and that intraindividual variability could predict impending death up to 15 years later (MacDonald, Hultsch, et al., 2008). Clearly, intraindividual variability in RT is a valid early indicator of maladaptive outcomes, supporting the hypothesis that inconsistency is a trait-like characteristic that reflects neurological integrity. Consequently, the ability to predict other outcomes via inconsistency, such as dementia, is promising.

Across the Outcomes

It is important to note that stronger prediction might be found in a more impaired sample. The present sample was relatively healthy and well educated, and represented a more select group of older adults than might be randomly found in the population. Furthermore, only those participants who completed all six waves of testing were included in the cognitive status change analyses. Individuals who remained in the sample at Wave 6 were significantly younger ($M = 73.41$ years; nonreturning, $M = 75.42$ years), had fewer chronic conditions ($M = 2.73$; nonreturning, $M = 3.37$), and viewed themselves to be in better health than others their own age ($M = 4.36$; nonreturning, $M = 3.95$).

However, even with the selectivity of our sample, for both outcomes, an individual's likelihood of being in a poorer outcome group increased with each unit increase in baseline intraindividual variability. Therefore, whatever the cause of the underlying impairment (i.e., CIND, health concerns), there is clearly a link between greater inconsistency at baseline and being a member of one of these maladaptive groups 5 years later. Furthermore, given that CIND status is based on potentially early behavioral characteristics of dementia, and attrition from longitudinal studies is believed to be indicative of underlying influences such as disease and cognitive impairment (e.g., MacDonald et al., 2003; Sliwinski et al., 2003), the link between intraindividual variability and these specific deleterious outcomes supports hypotheses that fluctuations in behavioral performance are the result of neurological mechanisms (e.g., Li & Lindenberger, 1999).

Despite the remarkable strength between an individual's initial variability in responding and the changes over time, age group was also a reliable predictor. For each outcome, being in the old–old age group (i.e., 75–92 years) at the initial wave of testing resulted in a greater likelihood of showing cognitive decline and attrition 5 years later. The substantial influence of age group on predicting 5-year change outcomes was not unexpected given the greater risk of decline and disease with older age, and reiterates the importance of including biological age in estimating prospective outcomes.

Intraindividual Variability Versus Intraindividual Mean

Generally speaking, intraindividual variability and intraindividual mean were comparable in differentiating among the various change outcomes. The two measures appeared to mirror one another in terms of overall effect size (e.g., cognitive status change, *ISD*: Nagelkerke's $R^2 = .18$; *IM*: Nagelkerke's $R^2 = .17$), and there were similarities in terms of which groups were significantly distinguished (e.g., stable intact and stable CIND by the basic composite scores). However, the related Wald statistics often revealed that a 1-unit increase in inconsistency had a greater impact on the likelihood of being in a maladaptive group than did a 1-unit increase in average speed of responding. For example, the individual complex composite models predicting cognitive status change showed that each additional unit of inconsistency increased the likelihood of being in the stable CIND group compared with the stable intact group by 38%, whereas the related odds only increased by 22% for a unit increase of the mean. Similar findings were found for the attrition outcome (see Table 2). Therefore, individuals were at greater relative risk of being in a maladaptive outcome group if they were more inconsistent rather than if they were slower in responding. The present results corroborate recent research showing the strength of intraindividual variability in comparison to the mean, particularly in predicting subsequent cognitive decline (Lövdén et al., 2007) and impending death (MacDonald, Hultsch, et al., 2008).

Although it has recently been shown that the reliability of *IM*s tends to be higher than the reliability of *ISD*s (Schmiedek, Lövdén, & Lindenberger, 2009), it is also the case that the impact of this difference is reduced with increasing numbers of occasions. Given that the present *ISD* scores were based on anywhere from 47 to 60 trials, depending on the RT task, any related influence on the results was likely minimal.

Differences Due to Task Complexity

As expected, the complexity of the RT tasks enhanced differentiation among the change groups. For inconsistency, the basic composite provided the most insight in predicting cognitive status group and rate of attrition. These results are consistent with findings that inconsistency based on cognitively demanding tasks provided greater sensitivity to various cross-sectional outcomes (e.g., CIND; Strauss et al., 2007), but suggests that moderate cognitive challenge, rather than one that is highly demanding, may be particularly sensitive to longitudinal changes in cognitive abil-

[5] Two participants could not be located because they moved, and two participants dropped out because of family health problems. We chose not to include these individuals as we were primarily interested in those who gave busy or not interested as a reason for not continuing in the study.

ity. In fact, Bielak et al. (in press) suggested that a threshold may exist in the optimal task complexity, as they found that inconsistency derived from moderately challenging cognitive tasks showed similar coupling links to cognition as inconsistency derived from highly challenging cognitive tasks. Tasks that require some mental effort and judgment presumably stimulate the frontal regions of the brain, an area where intraindividual variability appears to be particularly reactive to any injuries or disease (e.g., Stuss et al., 2003). There appears to be no such threshold for mean level of performance, however, at least in predicting change in cognitive status, where the most cognitively challenging tasks were generally the best predictors of the later change group. However, this characteristic is not clear, as the basic version of mean responding was the most informative in predicting attrition. Given inconsistency's hypothesized greater sensitivity to neurological integrity, less cognitive demand may be required to obtain an accurate measurement than for mean response speed. Thus, another potential distinction may exist between the related measures. These results suggest particular attention should be given to the type of RT tasks used to calculate inconsistency and mean rate of responding.

Conclusions

Overall, the present study showed strong support for the sensitivity of intraindividual variability in RT to cognitive change over 5 years. The initial level of inconsistency was particularly sensitive to changes associated with the early behavioral symptoms of dementia (i.e., CIND), and the average level of inconsistency over time was predictive of attrition. In each case, greater inconsistency was associated with a greater likelihood of being in a maladaptive group 5 years later. Mean rate of responding was a comparable predictor of change in most instances, but the resulting relative risks of being in a maladaptive outcome group were greater for inconsistency than response speed. Intraindividual variability based on moderately challenging tasks appeared to be the most sensitive to longitudinal changes in cognitive ability, whereas both highly and moderately demanding tasks tended to provide the most insight for mean speed of responding.

Given the recent longitudinal findings (i.e., Bielak et al., in press; MacDonald et al., 2003; Lövdén et al., 2007), researchers can infer with some degree of confidence that the longitudinal link between inconsistency and cognition is legitimate. Furthermore, with the promising finding that inconsistency is sensitive to changes in cognition and other meaningful change outcomes in this and other studies (e.g., Lövdén et al., 2007), it is clear that the future of this field is in longitudinal and predictive relationships. Specifically, further study is needed across longer time frames, prospective studies involving various neurological conditions, regarding the dose–response effects of the relationship (e.g., the magnitude of the associated risk of decline with each increase in intraindividual variability), and the mechanisms underlying performance variability (see MacDonald et al., 2006).

References

Anstey, K. J., Mack, H. A., Christensen, H., Li, S.-C., Reglade-Meslin, C., Maller, J., . . . Sachdev, P. (2007). Corpus callosum size, reaction time speed and variability in mild cognitive disorders and in a normative sample. *Neuropsychologia, 45,* 1911–1920.

Bellgrove, M. A., Hester, R., & Garavan, H. (2004). The functional neuroanatomical correlates of response variability: Evidence from a response inhibition task. *Neuropsychologia, 42,* 1910–1916.

Bielak, A. A. M., Hughes, T. F., Small, B. J., & Dixon, R. A. (2007). It's never too late to engage in lifestyle activities: Significant concurrent but not change relationships between lifestyle activities and cognitive speed. *Journal of Gerontology: Psychological Sciences, 62*(B)*,* P331–P339.

Bielak, A. A. M., Hultsch, D. F., Kadlec, H., & Strauss, E. (2007). Was that part of the story or did I just think so? Age and cognitive status differences in inference and story recognition. *Experimental Aging Research, 33,* 295–322.

Bielak, A. A. M., Hultsch, D. F., Strauss, E., MacDonald, S. W. S., & Hunter, M. A. (in press). Intraindividual variability is related to cognitive change in older adults: Evidence for within-person coupling. Manuscript submitted for publication.

Burton, C. L., Strauss, E., Hultsch, D. F., Moll, A., & Hunter, M. A. (2006). Intraindividual variability as a marker of neurological dysfunction: A comparison of Alzheimer's disease and Parkinson's disease. *Journal of Clinical and Experimental Neuropsychology, 28,* 67–83.

Burton, C. L., Strauss, E., Hunter, M. A., & Hultsch, D. F. (2009). The relationship between everyday problem solving and inconsistency in reaction time in older adults. *Aging, Neuropsychology, and Cognition, 16,* 607–632.

Christensen, H., Dear, K. B. G., Anstey, K. J., Parslow, R. A., Sachdev, P., & Jorm, A. F. (2005). Within-occasion intraindividual variability and preclinical diagnostic status: Is intraindividual variability an indicator of mild cognitive impairment? *Neuropsychology, 19,* 309–317.

Dixon, R. A., & de Frias, C. (2004). The Victoria Longitudinal Study: From characterizing cognitive aging to illustrating changes in memory compensation. *Aging, Neuropsychology, and Cognition, 11,* 346–376.

Dixon, R. A., Garrett, D. D., Lentz, T. L., MacDonald, S. W. S., Strauss, E., & Hultsch, D. F. (2007). Neurocognitive markers of cognitive impairment: Exploring the roles of speed and inconsistency. *Neuropsychology, 21,* 381–399.

Ekstrom, R. B., French, J. W., Harman, H. H., & Dermen, D. (1976). *Manual for kit of factor-referenced cognitive tests.* Princeton, NJ: Educational Testing Service.

Folstein, M., Folstein, S., & McHugh, P. R. (1975). Mini-Mental State: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12,* 189–198.

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.

Hultsch, D. F., Hertzog, C., & Dixon, R. A. (1990). Ability correlates of memory performance in adulthood and aging. *Psychology and Aging, 5,* 356–368.

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). New York: Oxford University Press.

Hultsch, D. F., MacDonald, S. W. S., & Dixon, R. A. (2002). Variability in reaction time performance of younger and older adults. *Journal of Gerontology: Psychological Sciences, 57*(B)*,* P101–P115.

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., Strauss, E., Hunter, M. A., & MacDonald, S. W. S. (2008). Intraindividual variability, cognition, and aging. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (3rd ed., pp. 491–556). New York: Psychology Press.

Hunter, M. A., & Bielak, A. A. M. (2005). *Comparison of purification methods.* Unpublished manuscript.

Jensen, A. R. (1992). The importance of intraindividual variation in reaction time. *Personality and Individual Differences, 13,* 869–881.

Kelly, A. M. C., Uddin, L. Q., Biswal, B. B., Castellanos, F. X., & Milham, M. P. (2008). Competition between functional brain networks mediates behavioural variability. *NeuroImage, 39,* 527–537.

Lawton, M. P., & Brody, E. M. (1969). Assessment of older people: Self-maintaining and instrumental activities of daily living. *The Gerontologist, 9,* 179–186.

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 neurotransmitter systems and dedifferentiation of cognitive abilities in old age. In L.-G. Nilsson & H. J. Markowitsch (Eds.), *Cognitive neuroscience of memory* (pp. 103–146). Seattle: Hogrefe & Huber.

Li, S.-C., Lindenberger, U., Hommel, B., Aschersleben, G., Prinz, W., & Baltes, P. B. (2004). Transformations in the couplings among intellectual abilities and constituent cognitive processes across the life span. *Psychological Science, 15,* 155–163.

Lövdén, M., Li, S.-C., Shing, Y. L., & Lindenberger, U. (2007). Within-person trial-to-trial variability precedes and predicts cognitive decline in old and very old age: Longitudinal data from the Berlin Aging Study. *Neuropsychologia, 45,* 2827–2838.

MacDonald, S. W. S., Cervenka, S., Farde, L., Nyberg, L., & Bäckman, L. (2009). Extrastriatal dopamine D2 receptor binding modulates intraindividual variability in episodic recognition and executive functioning. *Neuropsychologia, 47,* 2299–2304.

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.

MacDonald, S. W. S., Hultsch, D. F., & Dixon, R. A. (2008). Predicting impending death: Inconsistency in speed is a selective and early marker. *Psychology and Aging, 23,* 595–607.

MacDonald, S. W. S., Nyberg, L., & Bäckman, L. (2006). Intra-individual variability in behaviour: Links to brain structure, neurotransmission and neuronal activity. *Trends in Neurosciences, 29,* 474–480.

MacDonald, S. W. S., Nyberg, L., Sandblom, J., Fischer, H., & Bäckman, L. (2008). Increased response-time variability is associated with reduced inferior parietal activation during episodic recognition in aging. *Journal of Cognitive Neuroscience, 20,* 779–786.

Martin, M., & Hofer, S. (2004). Intraindividual variability, change, and aging: Conceptual and analytical issues. *Gerontology, 50,* 7–11.

Murtha, S., Cismaru, R., Waechter, R., & Chertkow, H. (2002). Increased variability accompanies frontal lobe damage in dementia. *Journal of the International Neuropsychological Society, 8,* 360–372.

Nesselroade, J. R., & Salthouse, T. A. (2004). Methodological and theoretical implications of intraindividual variability in perceptual-motor performance. *Journal of Gerontology: Psychological Sciences, 59*(B), P49–P55.

Petersen, R. C., Smith, G. E., Waring, S. C., Ivnik, R. J., Tangalos, E. G., & Kokman, E. (1999). Mild cognitive impairment: Clinical characterization and outcome. *Archives of Neurology, 56,* 303–308.

Rabbitt, P., Lunn, M., & Wong, D. (2005). Neglect of dropout underestimates effects of death in longitudinal studies. *Journal of Gerontology: Psychological Sciences, 60*(B), P106–P109.

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

Rabbitt, P. M., Watson, P., Donlan, C., Bent, L., & McInnes, L. (1994). Subject attrition in a longitudinal study of cognitive performance in community-based elderly people. *Facts and Research in Gerontology, 14,* 203–207.

Salthouse, T. A., Nesselroade, J. R., & Berish, D. E. (2006). Short-term variability in cognitive performance and the calibration of longitudinal change. *Journal of Gerontology: Psychological Sciences, 61*(B), P144–P151.

Schmiedek, F., Lövdén, M., & Lindenberger, U. (2009). On the relation of mean reaction time and intraindividual reaction time variability. *Psychology and Aging, 24,* 841–857.

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

Strauss, E., Bielak, A. A. M., Bunce, D., Hunter, M. A., & Hultsch, D. F. (2007). Within-person variability in response speed as an indicator of mild cognitive impairment in older adults. *Aging, Neuropsychology, and Cognition, 14,* 608–630.

Stuss, D. T., Murphy, K. J., Binns, M. A., & Alexander, M. P. (2003). Staying on the job: The frontal lobes control individual performance variability. *Brain, 126,* 2363–2380.

Thurstone, T. G. (1962). *Primary mental abilities: Grades 9–12, 1962 revision*. Chicago: Science Research Associates.

Tuokko, H., & Frerichs, R. J. (2000). Cognitive impairment with no dementia (CIND): Longitudinal studies, the findings, and the issues. *The Clinical Neuropsychologist, 14,* 504–525.

Tuokko, H. A., Gabriel, G., & the Canadian Study of Health and Aging Neuropsychology Working Group. (2006). Neuropsychological detection of cognitive impairment: Inter-rater agreement and factors affecting clinical decision-making. *Journal of the International Neuropsychological Society, 12,* 72–79.

Tuokko, H. A., & Hultsch, D. F. (2006). *Mild cognitive impairment: International perspectives.* New York: Taylor & Francis.

Tuokko, H. A., & McDowell, I. (2006). An overview of mild cognitive impairment. In H. A. Tuokko & D. F. Hultsch (Eds.), *Mild cognitive impairment: International perspectives* (pp. 3–28). New York: Taylor & Francis.

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale—Revised.* New York: Psychological Corporation.

Williams, B. R., Hultsch, D. F., Strauss, E. H., Hunter, M. A., & Tannock, R. (2005). Inconsistency in reaction time across the life span. *Neuropsychology, 19,* 88–96.

Received September 28, 2009
Revision received March 3, 2010
Accepted March 8, 2010 ■