Psychology and Aging
1999, Vol. 14, No. 2, 245-263

0882-7974/99/$3.00

Use It or Lose It: Engaged Lifestyle as a Buffer of Cognitive Decline in Aging?

David F. Hultsch
University of Victoria

Christopher Hertzog
Georgia Institute of Technology

Brent J. Small
University of South Florida

Roger A. Dixon
University of Victoria

Data from the Victoria Longitudinal Study were used to examine the hypothesis that maintaining intellectual engagement through participation in everyday activities buffers individuals against cognitive decline in later life. The sample consisted of 250 middle-aged and older adults tested 3 times over 6 years. Structural equation modeling techniques were used to examine the relationships among changes in lifestyle variables and an array of cognitive variables. There was a relationship between changes in intellectually related activities and changes in cognitive functioning. These results are consistent with the hypothesis that intellectually engaging activities serve to buffer individuals against decline. However, an alternative model suggested the findings were also consistent with the hypothesis that high-ability individuals lead intellectually active lives until cognitive decline in old age limits their activities.

Cognitive performance depends both on the demand characteristics of the task and the abilities of the individuals performing the task (Jenkins, 1979). This latter dimension—the characteristics of persons—has received considerable attention from cognitive gerontologists in recent years. Broadly speaking, researchers have focused on two classes of individual-difference variables as potential influences on age differences and changes in performing memory and other complex cognitive tasks (Dixon & Hertzog, 1996). One strand of research is rooted in the view that performance on a wide variety of tasks is dependent on a smaller set of basic processing components, such as information processing speed (Salthouse, 1996). These basic processing components are hypothesized to decline with age. Individual differences in the decline of these resources will, in turn, lead to differential trajectories of decline on more complex cognitive tasks. A second strand of research has suggested that performance on complex cognitive tasks may be moderated by a relatively broad array of contextual variables reflective of the individual's exposure to various events and environments (Schaie, 1996; Schooler, 1990). A central theme of this work has been examination of the proposition that favorable life experiences or conditions may forestall or attenuate the declines typically seen in a variety of cognitive processes in later adulthood.

The "disuse" perspective on cognitive aging (Salthouse, 1991) is one way to characterize the impact of favorable experiences or conditions on cognitive performance in later adulthood. This view attempts to account for age changes in cognitive performance in terms of changes in the nature of activities performed by people across the life span. Essentially, it is suggested that changes in activity patterns result in disuse and consequent atrophy of cognitive processes and skills. The view is often captured in the adage "use it or lose it."

If disuse of cognitive skills exacerbates age-related cognitive decline, then one might expect that deliberate practice of such skills would at least result in stable performance either through maintenance or enhancement of the original skill components (Ericsson & Charness, 1994) or through a variety of potential compensatory mechanisms (Dixon & Bâckman, 1995). A more optimistic perspective on practice would predict that it may even reverse age-related declines (Schaie & Willis, 1986). Indeed, researchers have argued that older adults possess considerable reserve capacity that permits them to benefit from exposure to performance-enhancing environments (Baltes & Baltes, 1990). Studies that have manipulated experience by deliberate training or practice have revealed that the cognitive abilities of older adults show considerable plasticity, although there appear to be declines in reserve capacity with increasing age as well (e.g., Baltes & Willis, 1982; Kliegl, Smith, & Baltes, 1990; Schaie & Willis, 1986; Verhaeghen, 1993).

Research on plasticity typically pertains to domain-specific experience and effects. Does the stimulation provided by typical everyday activities facilitate the maintenance and improvement of general cognitive skills in a manner that is analogous to exposure

David F. Hultsch and Roger A. Dixon, Department of Psychology, University of Victoria, Victoria, British Columbia, Canada; Christopher Hertzog, School of Psychology, Georgia Institute of Technology; Brent J. Small, Department of Gerontology, University of South Florida.

This research was supported by Grant AG08235 from the National Institute on Aging. We thank the volunteer participants of the Victoria Longitudinal Study for their time and effort and research staff members for their assistance with data collection and scoring. In particular we thank Scott Maitland for assistance in statistical analysis of the health and activity measures.

Correspondence concerning this article should be addressed to David F. Hultsch, Department of Psychology, University of Victoria, Victoria, British Columbia, Canada, V8W 3P5. Electronic mail may be sent to dfh@uvic.ca.

to cognitive training? Participation in cognitively demanding activities may serve to maintain learning strategies or other mechanisms that could support performance on cognitive tests and tasks typically administered in the laboratory. This notion is expressed formally in Schooler's (1987) environmental complexity hypothesis. This hypothesis proposes that the complexity of an individual's environment is defined by its stimulus and demand characteristics. Complex environments are characterized by diverse stimuli, the requirement for multiple and complex decisions, and ill-defined and apparently contradictory contingencies. To the extent that patterns of reinforcement within such an environment reward cognitive effort, individuals should be motivated to develop their intellectual capacity and to generalize the resulting cognitive processes to other situations. From this perspective it might be expected that individuals who engage in activities that make significant demands on their cognitive skills will show greater maintenance or improvement of their abilities than individuals who are exposed to less complex environments with more minimal cognitive demands (Schaie, 1984; Schooler, 1987). It may be possible, then, that different patterns of participation in everyday activities may be associated with different trajectories of cognitive change later in life.

Many of the extant studies examining the relationship between level of participation in activities and performance on various cognitive tasks have used cross-sectional designs (e.g., Arbuckle, Gold, & Andres, 1986; Christensen et al., 1996; Craik, Byrd, & Swanson, 1987; Erber & Szuchman, 1996; Hill, Wahlin, Winblad, & Bäckman, 1995; Hultsch, Hammer, & Small, 1993; Luszcz, Bryan, & Kent, 1997; Schaie, 1996; van Boxtel, Langerak, Houx, & Jolles, 1996). In general, these studies have reported that greater participation in physical, social, and intellectual activities is associated with higher levels of cognitive performance on a wide range of cognitive tasks. It has also been suggested that this relationship may become stronger with increasing age in later life (Hultsch et al., 1993).

Although the results of these cross-sectional studies are intriguing, one must be concerned with issues of causal directionality. Does participation in cognitively demanding activities promote the development and maintenance of cognitive abilities, or do capable people tend to participate in environments that are cognitively demanding? This conundrum can probably never be completely resolved. However, longitudinal data that permit examination of the sequencing of events are potentially useful in exploring the issue. Some longitudinal data are available (e.g., Arbuckle, Gold, Andres, Schwartzman, & Chaikelson, 1992; Arbuckle, Gold, Chaikelson, & Lapidus, 1994; Gold et al., 1995; Gribbin, Schaie, & Parham, 1980; Schaie, 1984).

Arbuckle et al. (1992) found evidence for a link between leisure activities and intellectual performance in a sample of World War II veterans tested twice over a 40-year period. Consistent with the cross-sectional findings reviewed above, they reported that health, education, and participation in intellectual activity were related to maintenance of intellectual performance. Gold et al. (1995) recently used structural equation models to examine relationships among these variables in the same sample and argued that individuals with higher levels of intellectual ability, education, and socioeconomic status are more likely to develop an engaged lifestyle, which in turn contributes to the maintenance of verbal intelligence in later life. Additional longitudinal evidence is available from Schaie's Seattle Longitudinal Study (SLS; Schaie, 1984, 1996). Schaie examined the relationship between types of persons defined by activity patterns and intellectual change over 7 and 14 years (Schaie, 1984). These analyses revealed the least decline was exhibited by individuals with high socioeconomic status who were fully engaged with their environment, whereas the greatest decline was shown by widowed women who had never been in the workforce, and who exhibited a current disengaged lifestyle. Two other groups, those with average status who were fully engaged and those who were semi-engaged, showed intermediate decline.

The evidence reviewed above suggests that an active, engaged lifestyle is related to the maintenance of cognitive performance with increasing age. However, the relationship between individual differences in lifestyle and cognitive performance is complicated by a number of factors. In particular, it is obvious that various other individual characteristics may moderate the relationship between activity level and cognitive performance. For example, age, gender, race, education, socioeconomic status, and personality may all influence the types of activities selected by the individual (e.g., Verbrugge, Gruber-Baldini, & Fozard, 1996). Moreover, these characteristics may themselves influence cognitive change directly. However, a number of studies have used measures of engaged lifestyle that appear to confound such indicators of initial ability, socioeconomic status, and demographic status and current environmental engagement. For example, Gold et al.'s (1995) latent lifestyle variable consisted of indicators of socioeconomic status, locus of control, and intellectual activity. Similarly, a number of Schaie's (1996) environmental factors (e.g., Prestige, Social Status) that show the strongest relationships to intellectual functioning appear to reflect substantial educational or socioeconomic dimensions and, therefore, raise questions of the role of differences in initial level of ability in producing these differences.

It is also important to distinguish between two classes of hypotheses regarding the relationships of lifestyle variables to cognitive change. The first class of hypotheses argues that levels of lifestyle variables are related to subsequent cognitive change because of accumulated influences (positive or negative), lagged effects,[1] or the like. For example, stable behavioral dispositions, such as depression or low openness to experience, may play a role in limiting exposure to optimal learning environments. On the positive side, a life history of intellectual activities or vigorous physical activities may lead to preservation of cognitive abilities. The second class of hypotheses involves concomitant change of contextual variables and cognitive abilities. Such hypotheses stipulate that changes in the contextual variables are likely to influence changes in cognitive functioning. For example, an increase in physical activity and concomitant change in cerebral blood flow might be expected to lead to improvements in cognitive performance. In contrast, reduced participation in cognitively demanding activities may contribute to cognitive declines through a variety of mechanisms (e.g., loss of strategic skills, reduced feelings of self-efficacy). We investigated these two classes of hypotheses

[1]These are also known as sleeper effects in which some time is required before the influence of a variable is evident. In such instances, earlier measures of a behavior may correlate more highly with the criterion than a concurrent measure.

with respect to the relationship between level and changes in activities and changes in cognition in later life.

The present data set has a number of advantages for examining the relationships between activity and cognition. First, the analyses are based on a multicohort longitudinal study of middle-aged and older adults, thus enabling us to examine longitudinal as well as cross-sectional linkages between the various predictor variables and cognitive performance. Second, in addition to measures of cognitive performance and activity, the data set also contained indicators of other variables that may be related to changes in either cognition or activity patterns (e.g., health status, personality). The availability of these measures enabled us to examine complex linkages among selected individual-differences predictors and cognitive variables. Finally, we obtained multiple measures of most constructs. This permits us to develop a measurement model at the latent variable level, thus ameliorating some of the problems associated with measuring change. Accordingly, we applied structural equation modeling techniques to examine cross-sectional and longitudinal relationships between lifestyle variables and cognitive functioning, particularly participation in everyday activities and memory performance.

Method

This article is based on data from the Victoria Longitudinal Study (VLS). This study used longitudinal sequences consisting of multiple cohorts of middle-aged and older adults retested at intervals of 3 years with new samples added at intervals of 6 years. Data from the first three waves of Sample 1 were available for the present analyses. The design, participants, measures, and procedures of the VLS have been described extensively elsewhere (see Hultsch, Hertzog, Dixon, & Small, 1998), and therefore, these components of the method will only be summarized here.

Participants

Participants in the VLS were community-dwelling adults living in a medium-size metropolitan area (Victoria, British Columbia) and were recruited through advertisements in the public media and appeals to community groups. Sample 1 consisted initially of 487 community-dwelling adults (290 women and 197 men) aged 55–86 years at the first occasion of measurement (1986). The sample attrited to 335 participants after 3 years and to 250 participants after 6 years. These numbers include all participants, even those with substantial amounts of missing data on specific variables. Analyses examining the impact of attrition from the sample have been reported elsewhere (Hultsch et al., 1998). Briefly, there was selective attrition across the first two occasions of measurement resulting in a positively biased sample with respect to both demographic characteristics and cognitive performance. There was less evidence for selective attrition from Time 2 to Time 3. Moreover, there did not appear to be much indication that the selective attrition from the sample resulted in significant underestimation of decline over the 6-year period.

Examination of the characteristics of the sample indicated that the participants constituted a relatively select group compared with the general population. Within the sample, 83% had completed at least 11 years of education, and 44% had completed at least 14 years of education. Comparable figures for the province of British Columbia are 51% and 18%, respectively (Statistics Canada, 1989b). The average level of education of the 6-year longitudinal sample was 13.42 years. Although most of the participants were retired, they also showed substantially higher levels of occupational attainment than the general population. Within the sample, 53% held, or previously held, professional or semiprofessional jobs, whereas only 7% were classified as unskilled. This compared with 27% professional and semiprofessional and 9% unskilled for the population of the province (Statistics Canada, 1989a). In addition, over 80% of the participants rated their health as very good or good. Further details of the demographic characteristics of the sample may be found in Hultsch et al. (1998).

Measures and Procedure

The measurement battery consisted of multiple questionnaires, tests, and tasks designed to measure both cognitive variables (including experimentally based information processing components and psychometrically based abilities) and noncognitive variables (including self-reported health, activity participation, and personality). The test battery was administered across multiple sessions scheduled over a period of about 1 to 1.5 months. There were three testing sessions (2 group and 1 individual) at the first and second occasions of measurement and four testing sessions (2 group and 2 individual) at the third occasion of measurement. Each session was approximately 2 hr long, with a rest break in the middle. Tasks were administered in the same order to all participants. For the recall measures of fact recall, word recall, and story recall (see below), six alternate forms were created, with two forms administered at each of the three longitudinal occasions. Participants were randomly divided into three order groups, and a partial Latin square was used to counterbalance the order of the three sets across the three occasions. Additional information on the procedures of the study is available in Hultsch et al. (1998).

Cognitive Variables

The various cognitive measures coalesced as indicators of nine hypothesized latent variables (Hultsch et al., 1998; Hultsch, Hertzog, Small, McDonald-Miszczak, & Dixon, 1992).

Fact recall. This latent variable was measured by two sets of 40 questions that tested individuals' recall of world knowledge. The questions were drawn from the domains of science, history, literature, sports, geography, and entertainment (Nelson & Narens, 1980). The questions were presented in booklets, and participants wrote their answers under self-paced timing conditions. The number of correct items from each of the two sets was used as the measures.

Word recall. This variable consisted of immediate free recall of two lists of 30 English words selected from the total set of six lists (Hultsch, Hertzog, & Dixon, 1990). Each list consisted of six words from each of five taxonomic categories (e.g., birds, flowers) typed on a single page in unblocked order. Participants were given 2 minutes to study each list and 5 minutes to write their recall. The number of correctly recalled words from each of the two lists were used as the measures.

Story recall. This latent variable was measured by immediate gist recall of two narrative stories about an event in the life (or lives) of an older adult (or couple). The total set of six stories was selected from a larger set of 25 structurally equivalent texts developed by Dixon, Hultsch, and Hertzog (1989). Each story was approximately 300 words and 160 propositions long. The stories were presented in typed booklets for study followed by written recall. Participants were given 4 minutes to read each story and 10 minutes to write their recall. Recall protocols were scored for gist recall using criteria described in Dixon et al. (1989). Reliability estimates of the scoring system across all possible pairs of scorers at all waves of measurement exceeded 90%. The total numbers of gist propositions recalled from each of the two stories were used as the measures.

Vocabulary. The latent indicator of vocabulary consisted of performance on a 54-item multiple-choice (recognition) vocabulary test composed by concatenating three 18-item tests from the Educational Testing Service Kit of Factor Referenced Tests (Ekstrom, French, Harman, & Dermen, 1976). For structural equation models, the test was divided into two split halves, using the number of correct responses for the set of items included in each half.

Verbal fluency. This latent variable was indicated by performance on three tasks from Ekstrom et al. (1976): Similarities, Opposites, and Figures of Speech. Under timed conditions, participants were asked to (a) write as many words as possible that have the same meaning, or nearly the same meaning, as a set of target words (6 minutes); (b) write as many words as possible that have the opposite meaning, or nearly the opposite, as a set of target words (5 minutes); and (c) write as many words or phrases as possible that can be used to complete a figure of speech (5 minutes). The total number of correct responses from each task was used.

Reading comprehension. This latent variable consisted of responses to a memory-loaded comprehension task. Participants read short passages (approximately eight sentences and 100 words) presented one sentence at a time on the computer screen. The rate of presentation was self-paced, although participants were asked to read the passages at their normal reading rate. Immediately after reading the passage, participants were asked to answer factual questions about the content of the passage. A total of six passages were read and tested by five questions each. For structural equation models, the six passages were divided at random into three groups of two passages, and the correct responses from these three groups of passages were used as the measures.

Working memory. This latent variable was indexed by two measures at Time 1 and four measures at Times 2 and 3. The sentence construction task was administered at all three occasions. Participants read aloud a series of sentences while keeping in memory a key word from each. When combined at recall, the key words from the series formed a new sentence. Twelve series of increasing length (three trials at series lengths 3, 4, 5, and 6 sentences) were presented. The measure consisted of the number of new sentences correctly reported. The number tracking task was also administered at all three occasions. Participants were required to keep track of a number in the last, next-to-last, or third-to-last position in a string of indeterminate length presented one digit at a time. There were 10 trials at each level of difficulty. The number of correctly reported digits in the final set (third-to-last) was used as the measure.

Given concerns about the measurement properties of these two tasks as convergent measures of working memory, two widely used working memory tasks developed by Salthouse and Babcock (1991) were added at Time 2 and Time 3: the computation span and listening span tasks. Both tasks require storage of information and simultaneous processing of that information. In computation span, participants must solve arithmetic problems while holding one number from each problem in memory for later recall. In listening span, participants listen to orally presented sentences and choose answers to simple questions about each sentence while retaining the last word of each sentence for later recall. In each task, the number of items (problems, sentences) increases from one to seven, with three trials at each series length. For each task, the score used was the highest span (1–7) correctly recalled on two out of three trials.

Comprehension speed. This latent variable was defined by time taken to process more complex verbal materials. The first indicator consisted of reading speed associated with reading short passages of text at the participant's "normal" reading rate. Six passages were presented one sentence at a time on the computer screen, and the timing of the task was self-paced. The measure was computed as time per proposition, and a trimmed mean across all six passages was used. Scores equal to or greater than three standard deviations above the individual's initial mean score were deleted, and the individual's mean score was recomputed. The second indicator asked participants to read a short passage of text as rapidly as possible to find the answer to a specific question stated prior to reading the passage. Sixteen passages were presented one sentence at a time, and the timing of the task was self-paced. Total time across all passages was the measure used.

Semantic speed. This latent variable was indexed by two tasks requiring speeded processing of verbal materials. In the lexical decision task, participants were presented with a string of letters on the computer screen and were asked to indicate as quickly as possible whether they formed an English word (e.g., island vs. nabion). A total of 60 items were presented, and trimmed mean latencies (scores equal to or greater than three standard deviations above the individual's initial mean score were deleted) of correct responses were used as the measure. In the semantic decision task, participants were asked to judge as rapidly as possible the plausibility of sentences presented on the computer screen (e.g., The tree fell to the ground with a loud crash vs. The pig gave birth to a litter of kittens this morning). A total of 50 sentences was presented, and trimmed mean latencies of correct responses were used as the measure.

Activity Lifestyle

Level of engagement with the environment was measured by participants' self-report of the frequency with which they engaged in a variety of everyday activities. The list incorporated 70 activities, and participants rated their typical frequency of participation over the last year on a 9-point scale (*never, less than once a year, about once a year, 2 or 3 times a year, about once a month, 2 or 3 times a month, about once a week, 2 or 3 times a week, daily*). The item pool did not represent all possible types of everyday activities, nor did it sample equally from all domains. Efforts were made to include a disproportionate number of activities that placed demands on the use of knowledge or cognitive skills. We classified 64 of the items into six categories as listed below. Six items were dropped because of difficulty in classifying them into one of the six groups.

1. Physical activities such as jogging or walking ($n = 4$)
2. Self-maintenance activities such as preparing a meal or shopping ($n = 6$)
3. Social activities such as visiting friends or attending a party ($n = 7$)
4. Hobbies and home maintenance activities such as playing a musical instrument or repairing mechanical items ($n = 12$)
5. Passive information processing such as listening to the radio or watching a sporting event ($n = 8$)
6. Novel information processing activities such as learning a language or playing bridge ($n = 27$)

All variables were scaled so that higher scores were associated with greater levels of activity, with zero indicating no activity in that domain. Items within each of the six categories were summed to produce composite activity measures.

Self-Reported Health

We measured four of the five components of self-reported physical health identified by Liang (1986): Chronic illness, number of illness episodes, instrumental health, and self-rated health. Liang's fifth component, physical self-maintenance, was not included because of the intact nature of the sample. In addition, we measured medication use. The variables were scaled so that higher scores indicated more illness, poorer health, and greater medication use. The specific measures were as follows.

Chronic illness. This measure consisted of self-reports of the presence and severity of 26 chronic conditions from eight broad categories (infirmities [e.g., arthritis], respiratory conditions [e.g., emphysema], circulatory disorders [e.g., stroke], digestive disorders [e.g., colitis], nervous disorders [e.g., Parkinson's disease], glandular disorders [e.g., diabetes], anemia, and cancer). Participants were asked to indicate whether they had experienced each of the 26 problems during the past 2 years, and if so, whether it was "fairly serious" or "not too serious." These reports were combined to yield scores ranging from 0 to 2 and summed to yield an overall chronic illness score.

Illness episodes. Participants reported the number of illness episodes they had experienced during the past year. Three items reporting the number of nights confined in a hospital, number of visits to a physician, and number of days home sick in bed all or most of the day were used.

Participants were asked to indicate the appropriate frequency in each instance using a 5-point frequency scale, and the scores were summed across all three items to yield an overall illness episode score.

Instrumental health. Participants rated the extent to which their health had caused them to change their level or pattern of daily activity in eight domains (employment, household chores, getting around town, mental activities, recreational activities, hobbies, socializing, and travel) over the past 2 years. Participants were asked to rate each domain on a 5-point scale, ranging from 1 (*no change*) to 5 (*give up activity*). Two of the items were excluded from the analysis. The domain of employment was not used due to the high frequency of retirement in this sample; the domain of mental activities was not used to avoid possible construct contamination with self-reported intellectual activities. Scores were summed across the six remaining items to yield an overall instrumental health score.

Self-rated health. This measure consisted of two items that asked participants to rate their own health on a 5-point scale (*very good, good, fair, poor, very poor*) compared with a perfect state of health and compared with other people their age. Scores on the two items were summed to yield an overall score.

Medication use. Participants were asked to indicate which prescription and nonprescription medications they were currently taking from a list of 20 drugs or types of drugs. The total number of drugs reported was used as the measure.

Personality

This domain was indexed by Costa and McCrae's NEO Personality Inventory (NEO–PI; Costa & McCrae, 1985). This inventory provides scores for five personality traits: Neuroticism, Extraversion, Openness to Experience, Agreeableness, and Conscientiousness. The inventory consists of 181 statements (e.g., I really like most people I meet). Participants are asked to indicate their agreement with each statement on a 5-point Likert scale, ranging from 0 (*strongly disagree*) to 4 (*strongly agree*). In the 1985 version of the instrument, the N, E, and O dimensions were second-order factors, each reflective of specific facets at the primary level. Although the A and C traits are considered to be higher-order dimensions as well, these were global scales in the 1985 version of the instrument (the recent NEO–PI–R [Costa & McCrae, 1992] is slightly longer and incorporates facets for these two factors as well). Because the personality traits identified by the NEO–PI have been found to be highly stable over time during adulthood (McCrae & Costa, 1990), the instrument was administered only at the first and third occasions of measurement.

Missing Data Treatment

For all cognitive and personality variables, no missing values substitution was used. However, to maintain adequate sample size, we used item mean substitution for the health and activity indicators. For the six categories of instrumental health, sample means were substituted prior to summing items into the composite scale. Item mean substitution was used only if two or fewer of the six items were missing. For the four indicators of health, including the composite instrumental health scale, missing values were recoded to the indicator mean at that longitudinal occasion, provided that only one of the indicators was missing. For activities, missing responses were assumed to indicate no activity, and were recoded to zero, provided that 6 or fewer of the 70 activity-item responses had been omitted.

After this limited recoding of missing data, all analyses excluded cases with missing data on a listwise basis. Thus, effective sample sizes varied across the statistical analyses, and these *N*s are reported for each instance in the Results section.

Statistical Procedure

The longitudinal data were analyzed using structural equation models (Bollen, 1989), which are appropriate for assessing prediction of cognitive change. The VLS was designed for this approach, and has multiple indicators (observed variables) for each latent variable at each time of measurement (see Hultsch et al., 1998, for additional exposition). Given the large number of indicators and the limited sample size for structural modeling, we included only the data from Time 1 and Time 3 in our structural equation models.

Because fact recall, word recall, and story recall consisted of six alternative forms, administered under a partial counterbalancing scheme, we sought to minimize the effect of the counterbalancing procedure on estimated individual differences in change. This was done by computing the overall mean for each alternative form of the task using all longitudinal data and then subtracting that mean from each person's score on that particular alternative form. Preliminary evaluation of the longitudinal factor structure of the cognitive measures showed that this approach did not affect the correlations of the latent variables, but did slightly increase the factor loadings for the recall measures and also slightly improved the fit of the factor model to the data (Hultsch et al., 1998).

We used structural equation models that directly estimate longitudinal change in relevant constructs, rather than the type of autoregressive models more common for panel data (e.g., Alwin, 1988). The method is based on latent change models developed by McArdle and associates (McArdle & Anderson, 1990; McArdle & Nesselroade, 1994). The approach can be succinctly described as a type of longitudinal factor analysis, specifically set up to measure changes as latent variables, combined with standard structural regression modeling approaches. The longitudinal factor analysis involves specifying occasion-specific factors (e.g., Working Memory at Time 1, Working Memory at Time 3). In most cases, the same variables were available at each time of measurement, and the model then estimated two ways in which a variable could correlate with itself over time—through the stability of the latent variable and through a stable source of variance specific to the observed variable itself. In standard longitudinal factor analysis (e.g., Hertzog & Schaie, 1986), change is estimated indirectly through the stability of the latent variables. In latent change analysis, the model is reformulated so that initial level and change are estimated for each latent variable. For example, given Working Memory at Time 1 (WM1) and Time 3 (WM3), the latent change model would specify two equations that map these occasion-specific factors into WMLevel and WMChange latent variables:

$$\mathrm{WM1} = \mathrm{WMLevel} \tag{1}$$

$$\mathrm{WM3} = \mathrm{WMLevel} + \mathrm{WMChange} \tag{2}$$

Differencing these two equations makes the latent change specification more obvious:

$$\mathrm{WMChange} = \mathrm{WM3} - \mathrm{WM1} \tag{3}$$

The full latent change model specifies the unobserved level and change variables for each latent variable (Word Recall, Fact Recall, etc.), and in the case of two occasions of measurement, these parameters are just-identified in Equations 1 and 2. Instead of estimating covariances of all occasion-specific factors, as in the standard longitudinal factor analysis, the model estimates covariances of level and change latent variables. It is a simple reparameterization and has an identical fit to the data as the standard longitudinal factor model (see McArdle & Nesselroade, 1994, for further explanation).

The latent change model benefits from the structural equation model in the usual way (i.e., correction for attenuation due to measurement error, adjustment of estimated longitudinal relationships for specific residual covariances; see Alwin, 1988). The model has several additional advantages over the standard approach. First, one does not need to use autoregressive models (Alwin, 1988) to indirectly infer prediction of change; instead, variables can be specified to have direct effects on latent change. For example, if Neuroticism is thought to cause change in Self-Reported Health, then this effect would be added to the model. Second, it is possible to directly specify and estimate relationships between latent change variables. Therefore, if change in Self-Reported Health is specified to predict change in Word Recall, this path can be easily and directly specified. It cannot be specified in standard autoregressive models. The advantage of this approach over the use of change scores is that the difference scores are never directly calculated, avoiding some of the problems associated with the use of change scores (e.g., negative bias, vanishing variance due to high stability; see Rogosa, Brandt, & Zimowski, 1982). The model we used differs from so-called latent growth curve analysis because the means are not included in the model (see McArdle & Anderson, 1990).

We used the latent change models for cognitive variables reported by Hultsch et al. (1998) as a starting point for all models including contextual variables. As reported in that monograph, the cognitive variables were modeled by imposing equality constraints on all estimated factor loadings (regressions of observed variables on latent variables). Due to a slight ceiling effect for the sentence construction variable at Time 1, its residual variance was fixed to a value determined by a specification search described in Hultsch et al. (1998). At Time 3, all four measures of Working Memory, including the Salthouse and Babcock (1991) measures, were used in the model. Measurement equivalence was obtained by fixing the factor loading of sentence construction to 1 at both Time 1 and Time 3 and constraining the loading of number tracking to be equal across occasions (see Hultsch et al., 1998, for more information about the behavior of the measures of Working Memory in these data).

We estimated all models using Bentler's EQS 5.4 program (Bentler, 1995). Model fit was evaluated by use of the Comparative Fit Index (CFI), Lagrangian Multiplier tests, and residual correlations. Due to an apparent bug in the program, we were unable to obtain Lagrangian Multiplier tests for the complex models involving activities, health, and cognitive change, perhaps due to high covariances of estimate involved in longitudinal data with a high degree of covariance stability. Hence, any modifications for these models were not based on this information. Evaluation of models was not merely based on fit to the data, but also on plausibility and interpretability of parameter estimates. Our approach was exploratory in nature, despite the use of structural equation models (Hertzog, 1990), and a number of alternative models were typically tried before settling on the final model that is reported in the Results section. Wherever the alternative models are especially theoretically relevant, the alternatives are discussed.

Results

The data analysis consisted of three major parts. The first two parts consisted of preliminary analyses conducted to (a) examine the issue of mean level change and cross-occasion stability and (b) confirm the measurement model for the activity, health, and personality measures (these issues have been addressed in previous reports for the cognitive variables; Hultsch et al., 1992; Hultsch et al., 1998). The main analyses use structural equation modeling techniques as outlined above to examine relationships among the contextual and cognitive variables.

Age Differences and Changes in Contextual Predictors

We have reported previously that the cognitive measures, with only a few exceptions (Vocabulary, Story Recall), show significant average decline and substantial individual differences in change across the 6-year interval (Hultsch et al., 1998). However, an additional question is stability of the contextual variables of activity lifestyle, self-reported health, and personality. As a preliminary step, then, we examined whether there was mean level change and cross-occasion stability in these various contextual variables. This question is important in the case of the second class of hypotheses outlined above in which changes in the contextual domain are predictive of changes in cognition. To qualify as a candidate explanation, variables should yield evidence of longitudinal change to be considered a promising explanation of cognitive change.[2]

We examined mean effects for the variables using a multivariate analysis of variance (MANOVA), followed by a univariate analysis of variance (ANOVA) in cases where the overall tests were significant. We used data from all three times of measurement for these analyses to provide the best estimate of average change for these variables, despite the fact, as noted earlier, that the structural equation models will only include data from Time 1 and Time 3. Polynomial contrasts (linear and quadratic) were used to examine the time of testing effect. Unless otherwise specified, we used an alpha level of .05 to assess the significance of statistical effects generated by these analyses. For the between-subjects univariate tests of significance, we used a sequential Bonferroni adjustment to maintain a familywise alpha of .05 (Ramsey, 1982). We estimated the proportion of variance associated with significant effects by computing eta squared (η^2). In addition, an initial estimate of the cross-occasion stability of these variables was obtained by computing 6-year auto-correlations. For these analyses participants were divided into young-old (55–70 years) and old-old (71–86 years) age/cohort groups based on their age at the initial occasion of measurement.

Activity

A total of 236 participants (155 young-old [95 women and 54 men] and 82 old-old [42 women and 40 men]) were available for this analysis. A 2 × 2 × 3 (Gender × Age/Cohort Group × Time of Testing) MANOVA conducted on the six activity groupings revealed significant overall main effects for gender, Wilks' λ = .776, $F(6, 227) = 10.95$, $p < .001$; age/cohort group, Wilks' λ = .860, $F(6, 227) = 6.16$, $p < .001$; and time of testing, Wilks' λ = .775, $F(12, 221) = 5.35$, $p < .001$.

Univariate tests related to gender revealed that women engaged in more self-maintenance activities, $F(1, 232) = 38.05$, $p < .001$, $\eta^2 = .141$ ($M_{\text{Women}} = 30.17$; $M_{\text{Men}} = 25.99$), and more social activities, $F(1, 232) = 4.25$, $p < .05$, $\eta^2 = .018$ ($M_{\text{Women}} = 25.21$; $M_{\text{Men}} = 23.52$) than men. In contrast, men participated in more physical activities, $F(1, 232) = 4.73$, $p < .05$, $\eta^2 = .020$ (M_{Women} = 14.82; $M_{\text{Men}} = 16.09$), and more novel information processing, $F(1, 232) = 10.44$, $p < .01$, $\eta^2 = .043$ ($M_{\text{Women}} = 66.94$; M_{Men} = 73.53) than women.

In terms of age/cohort differences, the univariate tests indicated that the young-old age/cohort group participated in more physical activities, $F(1, 232) = 29.28$, $p < .001$, $\eta^2 = .112$ ($M_{\text{Young-Old}}$ = 17.03; $M_{\text{Old-Old}} = 13.88$); more self-maintenance activities, $F(1, 232) = 4.03$, $p < .05$, $\eta^2 = .017$ ($M_{\text{Young-Old}} = 28.76$; $M_{\text{Old-Old}} = 27.40$); more hobbies, $F(1, 232) = 10.83$, $p < .001$,

[2]Technically, individual differences in change are statistically independent of mean change, and hence, one could find that a variable like health status would have no mean change, but still covary with cognitive change. However, given that Hultsch et al. (1998) already demonstrated average age declines in cognition in this sample, it would simplify and strengthen interpretation of predictor variables' influences if they displayed average age changes while also predicting individual differences in cognitive change.

η^2 = .045 ($M_{\text{Young-Old}}$ = 21.20; $M_{\text{Old-Old}}$ = 17.71); and more novel information processing, $F(1, 232) = 6.36$, $p < .05$, η^2 = .027 ($M_{\text{Young-Old}}$ = 72.81; $M_{\text{Old-Old}}$ = 67.66) than the old-old age/cohort group.

There were significant longitudinal effects reflecting linear decreases in activity over the 6 years on physical activities, $F(1, 232) = 5.04$, $p < .05$, $\eta^2 = .021$; self-maintenance activities, $F(1, 232) = 4.58$, $p < .05$, $\eta^2 = .019$; hobbies, $F(1, 232) = 26.84$, $p < .001$, $\eta^2 = .104$; and novel information processing, $F(1, 232) = 21.35$, $p < .001$, $\eta^2 = .084$. The quadratic trends for hobbies, $F(1, 232) = 3.95$, $p < .05$, $\eta^2 = .017$, and novel information processing, $F(1, 232) = 4.24$, $p < .05$, $\eta^2 = .018$, were also significant. The means and standard deviations for the activity groupings at the three times of measurement are shown in Table 1.

In addition to the significant main effects, there were significant interactions of Age/Cohort × Time of Testing, Wilks' $\lambda = .904$, $F(12, 221) = 1.95$, $p < .05$, and Gender × Age/Cohort × Time of Testing, Wilks' $\lambda = .901$, $F(12, 221) = 2.03$, $p < .05$. In the case of the two-way interaction, univariate analyses indicated a significant linear effect for hobbies, $F(1, 232) = 4.56$, $p < .05$, η^2 = .019. Participants in the young-old group declined less (T_1 = 21.71, T_2 = 21.49, T_3 = 20.40) than participants in the old-old group (T_1 = 18.95, T_2 = 18.36, T_3 = 15.81). There was also a significant Age/Cohort Group × Time of Testing interaction for the quadratic effect for novel information processing, $F(1, 232) = 5.65$, $p < .05$, $\eta^2 = .024$. Participants in the old-old group showed little change from Time 1 to Time 2, but experienced a steeper drop from Time 2 to Time 3 (T_1 = 69.06, T_2 = 69.44, T_3 = 64.48), whereas participants in the young-old group showed gradual decline over the period (T_1 = 74.14, T_2 = 72.68, T_3 = 71.60). There was also an Age/Cohort Group × Time interaction for the linear effect for social activities. This effect was further modified by a triple interaction, $F(1, 232) = 10.27$, $p < .01$, η^2 = .042. There were age differences in the linear trend for social activities for men, but not women. Specifically, younger men showed a tendency toward increased social activity (T_1 = 23.19, T_2 = 24.58, T_3 = 24.86), whereas older men showed a trend toward decreased social activity (T_1 = 24.73, T_2 = 22.70, T_3 = 21.05). In contrast, women from both age/cohort groups showed similar trends across time (young-old group: T_1 = 26.10, T_2 = 26.15, T_3 = 25.74; old-old group: T_1 = 24.14, T_2 = 25.19, T_3 = 23.93).

Table 1
Means and Standard Deviations for Activity Groupings at Three Times of Measurement

Variable	Time of measurement: Time 1	Time 2	Time 3
Physical activity			
M	16.02	15.81	15.46
SD	4.86	5.14	5.27
Self-maintenance activity			
M	28.94	28.64	28.29
SD	5.67	6.46	6.49
Social activity			
M	24.79	25.00	24.40
SD	6.77	6.81	7.09
Hobbies			
M	20.64	20.42	18.69
SD	8.83	8.47	8.50
Passive information processing			
M	33.95	34.31	33.76
SD	7.21	7.24	7.84
Novel information processing			
M	71.96	71.10	68.67
SD	16.11	16.90	16.10

Table 2
Six-Year Stability Coefficients for Contextual and Cognitive Variables

Variable	Stability coefficient
Activity	
Physical activity	.634
Self-maintenance	.543
Social activity	.574
Hobbies	.737
Passive information processing	.677
Novel information processing	.761
Self-reported health	
Chronic illness	.523
Illness episodes	.314
Instrumental health	.461
Self-rated health	.569
Medication use	.616
Personality	
Neuroticism	.841
Extraversion	.833
Openness to Experience	.857
Agreeableness	.692
Conscientiousness	.734
Cognitive variables	
Fact Recall	.793
Word Recall	.736
Story Recall	.727
Vocabulary	.850
Verbal Fluency	.753
Reading Comprehension	.631
Working Memory	.634
Comprehension Speed	.674
Semantic Speed	.672

Consistent with previous research, then (e.g., Verbrugge et al., 1996), there is a trend toward less participation in a variety of activities over time in later life. Of particular interest for the present analyses is the finding of a significant average decline in activities that appear to place significant demands on various aspects of cognitive functioning (novel information processing).

Inspection of the stability coefficients in Table 2 suggests that there may also be substantial individual differences in change on the activity measures as well. The correlations are, of course, significant, but they are generally lower than those seen for the cognitive[3] (range = .63–.85) or personality (range = .73–.86)

[3]In order to provide comparable estimates of stability for the cognitive measures, each indicator was separately standardized at each occasion of measurement and then used to form nine latent composite measures, one for each latent variable.

variables. However, standard stability coefficients do not permit us to distinguish between instability related to lack of measurement reliability and instability related to individual differences in change. This issue is clarified later when we consider the latent variable analyses.

Self-Reported Physical Health

A total of 228 participants (149 young-old [94 women and 55 men] and 79 old-old [43 women and 36 men]) were available for this analysis. A 2 × 2 × 3 (Gender × Age/Cohort Group × Time of Testing) MANOVA conducted on the health variables revealed significant overall main effects for gender, Wilks' λ = .884, $F(5, 220) = 5.79$, $p < .001$; age/cohort group, Wilks' λ = .926, $F(5, 220) = 3.50$, $p < .01$; and time of testing, Wilks' λ = .838, $F(10, 215) = 4.16$, $p < .001$. None of the interactions were significant.

The univariate ANOVAs pointed to gender differences in instrumental health, $F(1, 224) = 6.07$, $p < .05$, $\eta^2 = .026$, and medication use, $F(1, 224) = 9.97$, $p < .01$, $\eta^2 = .043$. Women reported significantly poorer instrumental health and more medication use than men (instrumental health: $M_{\text{Women}} = 3.66$, $M_{\text{Men}} = 2.55$; medication use: $M_{\text{Women}} = 2.09$, $M_{\text{Men}} = 1.44$). Univariate analyses of the overall age/cohort group effect revealed that the younger age/cohort group reported better instrumental health than the older age/cohort group, $F(1, 224) = 11.90$, $p < .01$, $\eta^2 = .050$ ($M_{\text{Young-Old}} = 2.33$, $M_{\text{Old-Old}} = 3.88$). Finally, the univariate tests suggested there was significant change over the 6 years on several of the health measures. The means for the health measures are shown in Table 3. Significant univariate linear decline in health was observed on the measures of chronic illness, $F(1, 224) = 15.34$, $p < .001$, $\eta^2 = .064$; instrumental health, $F(1, 224) = 19.72$, $p < .001$, $\eta^2 = .081$; and self-rated health, $F(1, 224) = 13.91$, $p < .001$, $\eta^2 = .058$. There was also a significant linear increment in medication use, $F(1, 224) = 12.00$, $p < .01$, $\eta^2 = .051$. None of the quadratic trends were significant.

The health measures, then, show some evidence of average change over time, generally reflecting increasingly poorer average health across the 6-year interval. Examination of the 6-year autocorrelations for the health measures shown in Table 2 suggests there may be individual differences in change on these variables as well.

Table 3
Means and Standard Deviations for Self-Reported Health Variables at Three Times of Measurement

Variable	Time 1	Time 2	Time 3
Chronic illness			
M	3.08	3.40	3.88
SD	2.72	2.99	3.04
Illness episodes			
M	2.94	2.94	3.17
SD	2.42	2.28	2.45
Instrumental health			
M	2.39	2.78	3.67
SD	3.86	3.80	4.04
Self-rated health			
M	1.34	1.43	1.65
SD	1.36	1.35	1.39
Medication use			
M	1.60	1.81	2.01
SD	1.64	1.68	1.76

Note: Time 1, Time 2 and Time 3 appear under the spanning header "Time of measurement".

Personality

A total of 229 participants (151 young-old [96 women and 55 men] and 78 old-old [40 women and 38 men]) were available for this analysis. On the basis of previous literature, we expected to find little evidence of change either at the level of means or individual differences for any of these personality traits. Thus, we view personality as a stable marker of individual differences.

A 2 × 2 × 2 (Gender × Age × Time of Testing) MANOVA was computed using the five personality factors. Raw scores were used in this analysis. The MANOVA revealed significant overall main effects for gender, Wilks' η = .899, $F(5, 221) = 4.96$, $p < .001$, and time of testing, Wilks' λ = .931, $F(5, 221) = 3.26$, $p < .01$. None of the interactions were significant.

In the case of the gender effect, the univariate tests indicated that women scored significantly higher on Openness to Experience, $F(1, 225) = 8.44$, $p < .01$, $\eta^2 = .036$ ($M_{\text{Women}} = 159.46$, $M_{\text{Men}} = 152.25$), and Agreeableness, $F(1, 225) = 13.02$, $p < .001$, $\eta^2 = .055$ ($M_{\text{Women}} = 69.63$, $M_{\text{Men}} = 66.89$) than men. In the case of the time of testing effect, the univariate ANOVAs revealed a slight decrease in Openness to Experience scores, $F(1, 225) = 11.12$, $p < .001$, $\eta^2 = .047$ ($T_1 = 157.03$, $T_3 = 154.67$). It is clear, however, from examining the means and variance estimate associated with this result that the magnitude of the decline is modest. Given that the items measuring Openness to Experience include questions about participation in various activities, this slight decrease may reflect age changes in capacity (e.g., increased physical disability and lower energy level), rather than a decline in interest in trying new activities. Consistent with previous reports (McCrae & Costa, 1990), then, these results suggest a picture of mean level stability of personality in middle and late adulthood, except perhaps for the dimension of Openness to Experience.

An examination of the 6-year stability coefficients reported in Table 2 also indicates that there were relatively few individual differences in personality change. The correlations are uniformly high and consistent with those observed in other samples (Costa & McCrae, 1988). The disattenuated stability coefficients yielded by a longitudinal confirmatory factor analysis on these data are even higher, ranging from .87 to .93.

As expected, then, personality as measured by the NEO shows substantial, although not perfect, stability both at the level of means and individual differences in change. This is important for the analyses to come because, in combination with theoretical perspectives indicating stability of adult personality (Costa & McCrae, 1994), it suggests there would be little utility in using change in personality as a predictor of change in cognition. In the analyses to follow, then, our models use initial status (Time 1) personality in the model rather than considering change in this variable over time.

Measurement Model for Contextual Variables

Models for Time 1 Cross-Sectional Data

In order to set the stage for analyses relating cognitive change to the contextual variables, we began by examining correlations between indicators of the contextual variables. In the activity domain, preliminary confirmatory factor analyses suggested that the two dimensions used by Hultsch et al. (1993) on the basis of principal components analysis were not well reproduced. In particular, the self-maintenance and passive information processing measures did not correlate highly with other activity scales and had no correlations with cognitive variables. Hence our measurement model used the activity scales of physical activity, social activity, hobbies, and novel information processing. The first three variables were used to measure a latent variable of Active Lifestyle. Novel Information Processing was treated as a single indicator latent variable, given its theoretical importance as a measure of intellectual engagement that could predict maintenance of cognitive functioning.

In the case of the health measures, number of illness episodes and medication use had low correlations with other measures of subjective health and were therefore not used as indicators of Self-Reported Health in the model. The two indicators of self-rated health, the absolute and relative self-rated health variables, had a high correlation with each other and similar correlations with other variables. Hence these two ratings were summed to form a single indicator of self-rated health. The Self-Reported Health latent variable was also measured by the indicators of instrumental health and chronic illness.

In the domain of personality, preliminary analyses of Time 1 cross-sectional correlations showed that the NEO scales of Agreeableness and Conscientiousness had no significant correlations with cognitive variables, hence they were not included in our models. We also used a reduced set of NEO facet variables to define the latent variables of Neuroticism, Extraversion, and Openness to Experience. Variables were selected to conform to a simple structure for these latent variables and to minimize the total number of variables (with a goal of having at least three indicators for each of the three personality latent variables). Ultimately the NEO scales of Anxiety, Depression, and Vulnerability were chosen to measure Neuroticism; the Warmth, Gregariousness, and Positive Emotions scales were chosen to measure Extraversion; and the Fantasy, Aesthetics, Feelings, Values, and Ideas scales were chosen to measure Openness to Experience.

We then ran a final confirmatory factor analysis that included these indicators and the cognitive latent variables. Complete data were available for 431 persons at Time 1 for these variables. Some modifications to the initially hypothesized model were required. It was necessary to add two additional loadings for NEO variables: Positive Emotions on Openness to Experience and Feelings on Extraversion. This model fared well, but there were still residual correlations with the NEO Ideas scale and cognitive factors. A final model specified Ideas and Novel Information Processing as single-indicator latent variables, along with Neuroticism, Extraversion, Openness to Experience, Self-Reported Health, Active Lifestyle, and the nine cognitive factors. Ideas no longer loaded on Openness to Experience. This model fit the data relatively well ($\chi^2 = 920.88$, $df = 546$, CFI = .940). Residual correlations of NEO, activity, and health indicators with cognitive indicators were all relatively low (less than .2, and typically less than .1). Moreover, there was no systematic pattern to the residual correlations that suggested a misspecification for the contextual variables.

Table 4 reports the standardized factor loadings for the measurement model. All the latent variables were defined by moderate to large factor loadings, with the smallest loadings associated with the Active Lifestyle factor. Table 5 reports the factor correlations among the contextual latent variables. Many of the correlations were quite low, although the pattern of significant correlations was consistent with expectations from the literature. For example, Neuroticism correlated significantly with Self-Reported Health. Active Lifestyle was significantly related to the three major personality factors, but the modest magnitude of the factor correlations (all < .3) supported the discriminant validity of this dimension from personality per se.

As would be expected, the Ideas scale correlated moderately with the Openness factor, and the Novel Information Processing scale correlated moderately with the Active Lifestyle factor. However, the pattern of correlations with other factors was sufficiently divergent to reveal why separating these scales improved the

Table 4
Factor Loadings for Measurement Model of NEO, Self-Reported Health, and Activity Variables

Variable	Factor	Loading	Factor	Loading
Chronic illness	Self-Reported Health	.69	—	—
Instrumental health	Self-Reported Health	.65	—	—
Self-Rated health	Self-Reported Health	.73	—	—
NEO Anxiety	Neuroticism	.80	—	—
NEO Depression	Neuroticism	.82	—	—
NEO Vulnerability	Neuroticism	.78	—	—
NEO Warmth	Extraversion	.72	—	—
NEO Gregariousness	Extraversion	.57	—	—
NEO Positive Emotions	Extraversion	.63	Openness to Experience	.21
NEO Fantasy	Openness to Experience	.66	—	—
NEO Aesthetics	Openness to Experience	.62	—	—
NEO Feelings	Openness to Experience	.60	Extraversion	.21
NEO Values	Openness to Experience	.42	—	—
Physical activities	Active Lifestyle	.46	—	—
Social activities	Active Lifestyle	.38	—	—
Hobbies	Active Lifestyle	.66	—	—

Note. NEO Ideas and Novel Information Processing were treated as single-item latent variables with 1.0 loadings. Em-dashes indicate that there was no secondary loading.

Table 5
Factor Correlations for NEO, Self-Reported Health, and Activities (Time 1)

Factor	1	2	3	4	5	6	7
Self-Reported Health	—						
Neuroticism	.32***	—					
Extraversion	−.22**	−.37***	—				
Openness to Experience	.09	.09	.27***	—			
Openness to Ideas	.06	−.17**	.09	.64***	—		
Active Lifestyle	−.16	−.23**	.28***	.27***	.18	—	
Novel Information Processing	−.06	−.21***	.01	.34***	.44***	.55***	—

**$p < .01$. ***$p < .001$.

model fit. In particular, the Ideas scale correlated more highly with Novel Information Processing than did Openness to Experience. This pattern is sensible, because a number of items from the Ideas scale measure preference for intellectual pursuits that are similar to the kinds of activities measured by the Novel Information Processing scale. One can think of the Ideas scale as measuring a behavioral disposition toward the kinds of activities measured by the Novel Information Processing scale.

Table 6 reports the cross-sectional (Time 1) factor correlations of the contextual predictors with cognition. There were a number of significant relationships involving Openness to Experience, Ideas, Active Lifestyle, Novel Information Processing, and Self-Reported Health. Neuroticism and Extraversion produced little correlation with cognitive functioning in this sample. The most impressive correlations were associated with Novel Information Processing. All these correlations were reliable at $p < .001$, with .4 or higher correlations for Fact Recall, Verbal Fluency, and Comprehension Speed. The Openness to Experience factor and the Ideas scale correlated most highly with Verbal Fluency. These correlations are broadly consistent with the findings of Hultsch et al. (1993) that the self-reported health and activity variables are related to cognitive functioning (despite differences in the variables used to define health and activity). In the activity domain, however, it was clear that Novel Information Processing was by far the strongest predictor of cognition; all nine correlations were higher for that scale, despite the fact that the correlations with the Active Lifestyle factor were corrected for attenuation due to measurement error.

Longitudinal Factor Model for Health and Activity

We shifted attention to producing a longitudinal measurement model for the contextual variables. A sample of 224 individuals produced complete data for the personality and contextual variables. The three personality factors were treated as static background variables, measured only at Time 1. The NEO Idea indicator was treated as a latent variable that was predicted by Neuroticism and Openness to Experience, and the residual Idea factor was allowed to covary with the contextual latent variables. This specification allowed us to distinguish between the Big-Five factors of Neuroticism and Openness to Experience, on the one hand, and the variance specific to the idea variable on the other hand, as sources of correlation with contextual variables. The NEO Values indicator was dropped from the model because preliminary models showed that it had a few substantial residual correlations with other indicators, and we were uninterested in isolating and understanding that specific source of variance.

The self-reported health and activity measures were measured at both Time 1 and Time 3. The longitudinal measurement model specified occasion-specific Self-Reported Health, Activity Lifestyle, and Novel Information Processing factors as in the Time 1 cross-sectional model. Correlated error terms were allowed for all

Table 6
Correlations of NEO, Self-Reported Health, and Activity Variables With Cognition (Time 1)

Latent variable	Self-Reported Health	Neuroticism	Extraversion	Openness to Experience	Openness to Ideas	Active Lifestyle	Novel Information
Fact Recall	−.13	−.16**	−.18**	.26***	.34***	.15	.41***
Word Recall	−.09	−.05	.03	.18**	.23	.20**	.30***
Story Recall	−.23***	−.13	−.07	.15	.13**	.13	.29***
Vocabulary	−.06	−.11	−.23***	.26***	.33***	.02	.37***
Verbal Fluency	−.08	−.05	−.03	.37***	.41***	.26***	.40***
Reading Comprehension	−.23**	−.09	.04	.07	.12	.05	.29***
Working Memory	−.26***	−.12	.03	.22**	.24**	.16	.26***
Comprehension Speed	.23***	.14	.04	−.29***	−.30**	−.25**	−.45***
Semantic Speed	.26***	.11	−.01	−.22***	−.12	−.18**	−.29***

**$p < .01$. ***$p < .001$.

Table 7
Factor Loadings for Longitudinal Model of Self-Reported Health and Activity Variables

Variable	Latent variable	Metric coefficients T1 & T3	Standardized T1	Standardized T3
Chronic illness	Self-Reported Health	1 (−)	.79	.78
Instrumental health	Self-Reported Health	1.224 (0.126)	.67	.74
Self-rated health	Self-Reported Health	0.338 (0.038)	.60	.60
Physical activity	Active Lifestyle	0.390 (0.086)	.40	.41
Social activity	Active Lifestyle	0.542 (0.117)	.41	.42
Hobbies	Active Lifestyle	1 (−)	.59	.70
Novel information processing	Novel Information Processing	1 (−)	1.00	1.00
Anxiety	Neuroticism	1 (−)	.77	—
Depression	Neuroticism	1.025 (0.089)	.84	—
Vulnerability	Neuroticism	0.712 (0.064)	.78	—
Warmth	Extraversion	1 (−)	.78	—
Gregariousness	Extraversion	0.997 (0.138)	.65	—
Positive emotions	Extraversion	0.816 (0.123)	.55[a]	—
Fantasy	Openness to Experience	1 (−)	.66	—
Aesthetics	Openness to Experience	1.139 (0.153)	.68	—
Feelings	Openness to Experience	0.843 (0.114)	.63[b]	—
Ideas	Openness to Experience	1.022 (0.153)	.65[c]	—

Note. Em-dashes indicate that NEO treated as static background variables measured only at Time 1.
[a] Positive emotions loading on Openness to Experience = .24; [b] Feelings loading on Extraversion = .24; [c] Ideas loading on Neuroticism = −.24.

indicators except Novel Information Processing, which was treated as a single-indicator latent variable at both occasions of measurement. Factor loadings for the Self-Reported Health and Active Lifestyle factors were constrained equally across the two occasions of measurement. A multivariate Lagrangian Multiplier test of the hypothesis of longitudinal measurement equivalence (equal factor loadings) was not significant, $\chi^2 = 5.26$, $df = 4$, justifying these equality constraints. All correlated errors were significantly different from zero. Table 7 reports the metric and rescaled factor loadings for this model. The standardized loadings were similar to the estimated loadings from the cross-sectional model.

Table 8 reports the latent standard deviations and stability coefficients for the three repeated latent variables. The model had been specified with extended standardized latent variables for activity and health, so that standard errors of estimate could be obtained for the standard deviations and standardized stability coefficients (see Hultsch et al., 1998). Self-Reported Health, Active Lifestyle, and Novel Information Processing all showed less than perfect stability, with 95% confidence intervals for the stability coefficients that did not include 1.0. The stability of Novel Information Processing was relatively high considering it was measured by a single indicator. Nevertheless, there appeared to be individual differences in 6-year changes in all facets of activities and subjective health. Unlike the cognitive variables, however (see Hultsch et al., 1998), the Active Lifestyle and Self-Reported Health factors showed no statistically reliable increases in latent standard deviations over the 6-year period ($z = 0.26$ and 1.03, respectively, $p > .25$). In contrast, the latent standard deviation for Novel Information Processing actually decreased from Time 1 to Time 3 ($z = -3.60$, $p < .001$).

Table 8
Standard Deviations and Stability Coefficients for Self-Reported Health and Activity Latent Variables

Latent variable	*SD*, Time 1	*SD*, Time 3	Stability coefficient
Self-Reported Health	2.444 (.197)	2.486 (.201)	.805 (.042)
Active Lifestyle	5.183 (.738)	5.745 (.737)	.862 (.060)
Novel Information Processing	17.910 (.833)	15.296 (.718)	.761 (.028)

Note. Values in parentheses are standard errors.

Residual covariances for all repeated indicators were statistically reliable, and in the case of activities, were quite substantial in magnitude. When standardized, the residual correlations for physical activity, social activity, and hobbies were .66, .55, and .60, respectively. These findings suggest relative stability in preferred specific activity patterns over the 6-year interval.

Longitudinal Relationships of Contextual and Cognitive Variables

Latent Change Correlations

We next turned to the issue of specifying latent change factors for the contextual variables and correlating initial level and longitudinal changes in these variables with latent cognitive change. Complete data on personality, contextual, and cognitive variables were available for 214 persons. The latent change model was closely based on the latent change model for cognitive variables reported by Hultsch et al. (1998), expanded to include the variables used in the longitudinal measurement model for contextual variables and personality. The two occasion-specific factors for Self-Reported Health, Activity Lifestyle, and Novel Information Processing were each reconfigured into two latent variables (a latent

Table 9
Correlations of Personality and Initial Level of Self-Reported Health and Activity With Initial Level of Cognitive Variables

Latent variable	Neuroticism	Extraversion	Openness to Experience	Ideas	Self-Reported Health	Active Lifestyle	Novel Information
Fact Recall	−.06	−.28**	.20*	.25**	−.09	.05	.33**
Word Recall	−.06	−.03	.03	.16*	−.22*	.28**	.34**
Story Recall	−.15	−.07	.06	.08	−.33**	.18	.26**
Vocabulary	.00	−.27**	.20*	.20**	−.09	.02	.35**
Verbal Fluency	.03	−.01	.27**	.24**	−.14	.22*	.35**
Reading Comprehension	−.10	.07	−.02	.10	−.28**	.11	.30**
Working Memory	−.03	.04	.06	.24**	−.31**	.23	.23*
Comprehension Speed	.16*	.06	−.18	−.14	.37**	−.31**	−.45**
Semantic Speed	.10	.07	−.05	−.03	.30**	−.17	−.14

*$p < .05$. **$p < .01$.

initial level or status factor and a latent change over the 6-year longitudinal interval; see Method section). We also included a latent initial level and change factors for the cognitive variables. All latent level and latent change factors were allowed to covary. The overall fit of the model was acceptable, $\chi^2 = 2231.15$, $df = 1778$, CFI = .941.

The model produced interesting results for the contextual variables that expanded on the understanding gained from the occasion-specific longitudinal measurement model. All three contextual latent change factors produced estimated standard deviations that were significantly greater than zero, indicating true individual differences in latent change. These outcomes corroborated inferences about individual differences in change based on the stability coefficients from the longitudinal factor model (see also Hultsch et al., 1998). New outcomes included the fact that there were statistically reliable ($p < .01$) negative covariances of initial status with change in Self-Reported Health, Activity Lifestyle, and Novel Information Processing (rescaled as correlations, they were −.37, −.34, and −.49, respectively). Individuals who were healthy were more likely to develop chronic and other diseases, and individuals initially high in self-reported activities were more likely to decrease their level of these activities. Despite the fact that this pattern held for all three latent variables, latent changes in the three variables were not significantly correlated ($p > .05$). Change in Active Lifestyle correlated −.22 with change in Health (decreases in activity were associated with decreases in health). Change in Novel Information Processing correlated .13 with changes in Self-Reported Health and .22 with changes in Active Lifestyle. These outcomes suggest that the mean patterns of decline in subjective health and activity patterns are most likely to affect those with initially good health and high levels of activity.

The estimated covariance of contextual and cognitive initial status and change latent variables demonstrated that cross-sectional relationships do not necessarily imply relationships in longitudinal changes. For ease of understanding, we report the relevant correlations, rescaled from the estimated covariance parameters, in three tables. Table 9 reports the correlations of initial status (level) of all variables. Table 10 reports the correlations of initial level of these variables with latent changes. Table 11 reports the correlations of the latent change variables.

There are several major points of interest. First, the initial level correlations in Table 9 are broadly consistent with the findings from the Time 1 cross-sectional model. Attrition from the longitudinal sample did not radically alter the patterns of relationships, although the magnitude of the correlations appeared, in aggregate, to be slightly attenuated relative to the full sample at Time 1. The self-reported health and activity variables correlated significantly with cognitive variables, as did Openness to Experience and the residualized Ideas scale (controlling for variance related to Neuroticism and Openness to Experience). Second, there were relatively few significant relationships of cognitive change to either initial status or to latent change in health and activity. Personality variables did not predict cognitive change. Change in Self-Reported Health also had no appreciable association with change in cognition. This finding was surprising, given the cross-sectional

Table 10
Correlations of Latent Predictors With Changes in Cognitive Variables

Cognitive change factor	Neuroticism	Extraversion	Openness to Experience	Self-Reported Health	Active Lifestyle	Novel Information
Fact Recall	−.02	.06	.11	−.22	.10	.11
Word Recall	.11	.03	.12	−.08	.03	−.01
Story Recall	.10	.14	.17	.05	.05	.06
Vocabulary	.21	.07	.12	.10	−.08	−.06
Verbal Fluency	−.03	.08	.07	−.04	.08	.08
Reading Comprehension	.00	−.17	.06	−.09	.11	.11
Working Memory	−.05	.05	.30	−.11	−.03	.31*
Comprehension Speed	−.08	−.23*	−.04	.12	−.17	−.07
Semantic Speed	−.02	−.19*	−.14	.12	−.06	−.11

*$p < .05$.

Table 11
Correlations of Changes in Latent Predictors With Changes in Cognitive Variables

Latent variable	Self-Reported Health	Active Lifestyle	Novel Information Processing
Fact Recall	−.01	.12	.32**
Word Recall	−.02	.27	.19
Story Recall	−.14	.33	.09
Vocabulary	−.08	.16	.16
Verbal Fluency	.07	.09	.12
Reading Comprehension	−.01	−.16	.01
Working Memory	.00	.49*	.10
Comprehension Speed	−.05	.00	−.11
Semantic Speed	.05	−.13	−.03

$^{*}p < .05$. $^{**}p < .01$.

relationships observed in Table 9. There were at least some correlations of activity change with cognitive change. However, in general, the magnitude of these relationships was modest. The largest correlations were associated with Working Memory.

These correlations are not strongly supportive of arguments that health and activity patterns are important mediators and/or protective factors that buffer the effects of aging on cognitive change. Neither level of Self-Reported Health or change in Self-Reported Health correlated significantly with any cognitive change variable. The same was true of level of Active Lifestyle. Changes in Active Lifestyle did have significant relationships with Working Memory and Story Recall, and changes in Novel Information Processing were associated in a few cases with latent cognitive change. Nevertheless, the magnitude and consistency of these relationships was not particularly impressive. In the next section, we report results from models that structured these latent change correlations in structural equation models, examining whether changes in contextual variables are associated with cognitive changes.

Latent Change Structural Equation Model

We next adapted the latent cognitive change path model reported by Hultsch et al. (1998) to include the Self-Reported Health, Active Lifestyle, and Novel Information Processing latent variables. The personality variables were dropped, given their lack of systematic patterns of correlation with the cognitive change variables. We also included chronological age as a predictor. This allowed us to test the hypothesis that any significant correlations observed in Table 11 were causally spurious artifacts of age-associated decline (particularly after age 75). All initial level variables were allowed to freely correlate with each other and with chronological age. Working Memory and Semantic Speed were treated as general resources predicting other variables (Salthouse, 1991, 1996). Age predicted latent changes in health and activity, as well as Working Memory change, Semantic Speed change, and Vocabulary change (given its relative resistance to age effects). In addition, paths from Active Lifestyle change to Working Memory change, Novel Information Processing change to Fact Recall change, and Self-Reported Health change to Active Lifestyle change were specified, based on the observed correlations. Thus, this model represented an exploratory attempt to find a parsimonious regression model based on the correlations, rather than an a priori test of a theory of contextual influence.

The model fit the data well (CFI = .957) but the path from Active Lifestyle change to Working Memory change was not significant, and there were small but salient residual correlations of all Working Memory indicators with Novel Information Processing at Time 1 and Time 3. We therefore dropped the path from Active Lifestyle and added paths of both level and latent change in Novel Information Processing to change in Working Memory. As expected, this specification caused the previously significant path from Novel Information Processing change to Fact Recall change to become nonsignificant, and it was trimmed from the final model. Paths from Self-Reported Health change to Active Lifestyle change and age to Self-Reported Health change were also trimmed for the same reason.

The standardized solution for the final model is depicted in Figure 1. The fit was acceptable, $\chi^2(1445, N = 214) = 1745.99$, CFI = .957. As would be expected, the relationships among the cognitive variables are consistent with the latent change model reported in Hultsch et al. (1998). The paths from Novel Information Processing level and change to Working Memory change were both significant ($z = 2.27$ and $z = 3.70$, respectively, $p < .05$), although obviously small to moderate in magnitude (.21 and .37, respectively). Inspection of the residuals did not suggest any other candidate paths from contextual predictors to cognitive change; in particular, there continued to be no need to use Self-Reported Health change (or initial status) as a predictor of change in any cognitive variable.

The intellectual engagement hypothesis might stipulate that level of Novel Information Processing buffers individuals against cognitive change; this concept would suggest that level of Novel Information Processing (but not change) ought to be the vehicle for predicting change in Working Memory, Fact Recall, and the like. To examine this hypothesis, we recomputed the model dropping the path from change in Novel Information Processing to Working Memory. This model did not fit as well, $\chi^2(1446, N = 214) = 1763.87$, CFI = .954, and the path of level of Novel Information Processing to Working Memory change was not reliably different from zero (standardized effect of .02, $z = 0.21$, $p > .25$). Thus, it appeared that an association of changes in the two constructs was required to adequately fit the data. When this was done, both the level of Novel Information Processing and change related significantly to change in Working Memory.

Given that the relationship between change in Novel Information Processing and change in Working Memory is, in a sense, a concurrent correlation, it could be argued that the causal influence runs in the opposite direction (i.e., Working Memory decline causes constriction in the kinds of intellectually engaging activities measured by Novel Information Processing). The model depicted in Figure 1 cannot adjudicate the true direction of influence. A model reversing the direction of the arrow represents an equivalent respecification, which would produce approximately the same level of fit to the data (see MacCallum, Wegener, Uchino, & Fabrigar, 1993; Stelzl, 1986). We did not bother to estimate that model.

In order to illustrate this larger point, however, and also to frame the problem without reference to the cognitive resource perspective that guided the logic of the model of Figure 1, we estimated an alternative model. It specified a General Cognitive Change factor (as for one model of cognitive change reported in Hultsch et

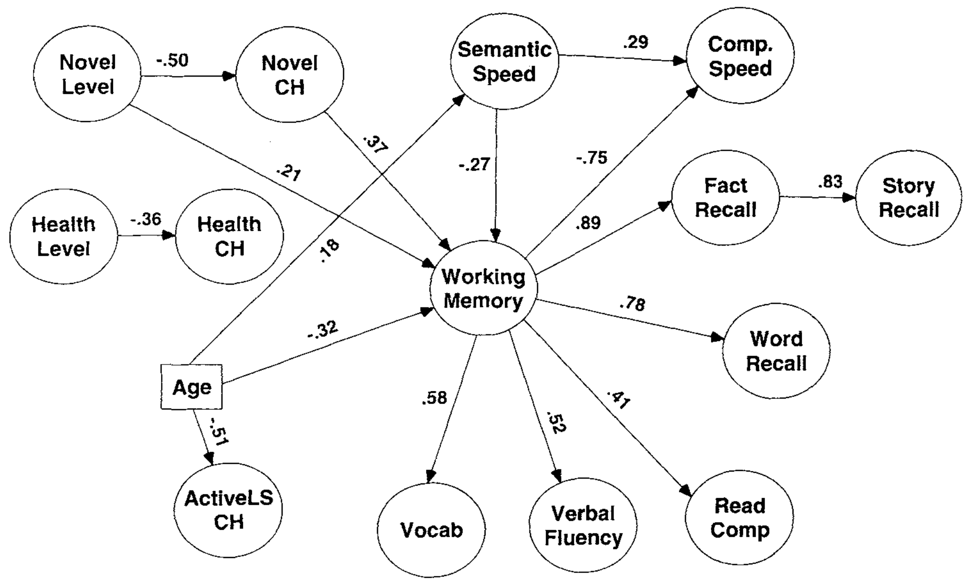

Figure 1. Structural equation model depicting relationships among latent changes in contextual variables and latent changes in cognitive variables. Novel = Novel Information Processing; Active LS = Active Lifestyle; Level = status of variable at Time 1; CH = latent change from Time 1 to Time 3; Comp = Comprehension; Vocab = Vocabulary.

al., 1998) instead of the cognitive resource path model. The General Cognitive Change factor was a hierarchical (or second-order) factor that accounts for much of the covariances in the individual cognitive change variables. It is analogous to a general intelligence factor in studies of psychometric intelligence (see Carroll, 1993), but it captures the extent to which individual differences in the rate of change are consistent across different cognitive variables. Hultsch et al. (1998) found that this General Cognitive Change factor was perfectly related to Working Memory change, and the same result was obtained in this analysis. Therefore, the alternative model is not much different from that of Figure 1, except that it postulates that cognitive decline actually causes decline in intellectually relevant activities, not the other way around (contrary to the use-it-or-lose-it hypothesis).

This alternative model fit well—indeed, as well as the resource path model, $\chi^2(1445, N = 214) = 1744.75$, CFI = .957. The standardized solution is shown in Figure 2. The model estimates that General Cognitive Change has a .30 direct effect on change in Novel Information Processing ($z = 3.90$, $p < .001$).

Discussion

Our initial expectation that the variables representing personality, self-reported health, and an active lifestyle would mediate changes in cognitive functioning was not strongly supported. We did find a relationship between intellectually related activities and change in cognitive functioning. However, we found minimal relationships between self-reported social and physical activity and cognitive change. We also found no relationships of self-reported health and cognitive change. This outcome does not appear to rest on fundamental inadequacies in the prediction measures themselves. In particular, the measures of self-reported health and active lifestyle showed both average change over the 6-year interval and sufficient degrees of individual differences in change to potentially operate as influences on cognitive change. Although we observed change in health, activity, and cognition, changes in the contextual variables, other than intellectually engaging activities, were not correlated with changes in cognition.

The results with self-reported health were particularly surprising. Despite mixed results in the literature, we expected to observe some significant relationships between changes in subjective health and changes in cognitive performance. This expectation was based, in part, on the fact that our measures of self-reported health were more extensive than those of some earlier studies. It was also based, in part, on our earlier cross-sectional findings showing significant cross-sectional relationships between health and these cognitive measures (Hultsch et al., 1993). The present analyses replicate the finding of significant correlations of initial level of health with initial level of cognitive performance with a reduced longitudinal sample. However, there was no evidence of relationships of longitudinal changes in health with changes in cognition. How are we to interpret these findings?

One potential explanation is selective attrition from the sample. If those who drop out of the sample are more likely to be in poor health, then perhaps the changes observed in health are insufficient to reach a threshold that impacts negatively on cognitive functioning. For example, Hertzog, Schaie, and Gribbin (1978) found that persons with evidence of cardiovascular disease that nevertheless remained in the SLS sample for 14 years showed smaller intellectual decline than individuals whose cardiovascular disease was related to morbidity, mortality, and attrition from the longitudinal sample after 7 years. Although this argument cannot be completely rejected, there are two considerations that tend to discount it. First, an attrition analyses reported previously by Hultsch et al. (1998) revealed that although

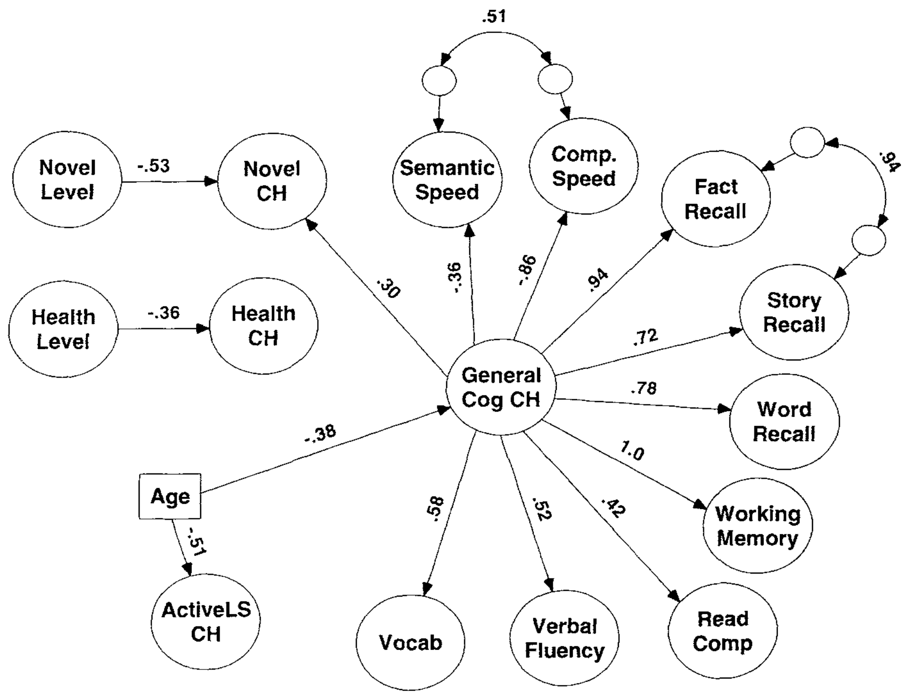

Figure 2. Structural equation model depicting relationships among latent changes in contextual variables and general cognitive change. Novel = Novel Information Processing; Active LS = Active Lifestyle; Level = status of variable at Time 1; CH = latent change from Time 1 to Time 3; Comp = Comprehension; Vocab = Vocabulary.

dropouts had less education and lower occupational status than returnees at the initial test, these two groups did not differ significantly on any of the health measures. Second, as noted earlier, we did observe significant evidence of change on the health measures over the interval. Moreover, initial health status and changes in health were negatively related; individuals in better health at the outset of the study were more likely to develop chronic illnesses. Thus, it would seem less likely that attrition from the sample had restricted the changes in health to such a degree that they no longer were predictive of cognition.

Alternatively, it could be the case that the recruitment of a select inception sample (volunteer retirees from an affluent region of Canada), and the fact that participants were required to attend multiple testing sessions at the University of Victoria, tended to prevent more impaired persons from participating in the study. In pursuit of this question of external validity, we recently examined cross-sectional relationships among self-reported health, activity lifestyle, and cognitive performance variables in a representative sample of 1,278 community-dwelling adults aged 65 to 100 years (Ball, 1998). One might expect that the greater variability in health, activity, and cognition in this more diverse sample would result in stronger relationships than those observed in a more select sample. However, the relationships observed in the representative sample were of similar magnitude to those reported in the original cross-sectional analysis of the VLS data (Hultsch et al., 1993). These results, although cross-sectional, do not suggest a substantial divergence of the pattern of results between a representative sample and the more select VLS sample. Nevertheless, given a sample with a larger proportion of individuals showing declines in health to more debilitating levels, it is possible that the relationships between changes in health and changes in cognition would emerge.

It is important to reiterate that we observed significant cross-sectional correlations between health and cognitive performance in the VLS longitudinal sample as well. The absence of a relationship between changes in health and changes in cognition suggests that these initial cross-sectional relationships reflect something other than a link between age-related changes in health and cognitive performance. One possibility is that the health measures may be more valid for more cognitively able participants. For example, brighter people may be less susceptible to anchoring effects such as adjusting their criterion for "good" health downward as they age. Resistance to this effect might result in them providing more accurate assessments of their health. This kind of influence could produce cross-sectional correlations of health and ability, but would not be expected to produce longitudinal relationships among changes in the two domains.

The present findings provide little support for a significant relationship between changes in subjective health and changes in cognitive performance. Nevertheless, it is important to differentiate between measures of subjective health, however reliable and valid, and alternative objective measures of physical health status and fitness. Previous research has shown a number of linkages between specific indicators of disease (e.g., Hertzog et al., 1978) and direct measures of physiological fitness (e.g., Rogers, Meyer, & Mortel, 1990) and cognitive performance. Although some of this work is cross-sectional, some of it is also longitudinal or experimental. Thus, our results do not necessarily suggest a lack of relationship between physiological functioning and cognitive performance. However, they do not provide much support for a relationship between changes in global self-reports of physical health as measured in social science research and changes in cognition.

In some respects the findings in the domain of social and physical activity were even more surprising to us than the findings for health. Our previous cross-sectional results (Hultsch et al., 1993) had suggested significant relationships between activity and performance that became stronger with age. Others have reported similar results (Christensen et al., 1996). The effects of age on activity variables suggested the possibility of a relationship between changes in activity and changes in performance. Moreover, at least two major longitudinal studies have reported links between patterns of everyday activity and cognitive change in later life (Gold et al., 1995; Schaie, 1984).

In contrast to these expectations and previous findings, we found no relationship between changes in our active lifestyle factor and changes in cognitive performance. Our active lifestyle factor was defined by participation in physical activities, social activities, and hobbies. Although there were significant correlations between initial status on this variable and initial cognitive status, changes in activity showed no link to changes in cognition. In part, this lack of relationship may be related to measurement inadequacies. There were a limited number of questions in these domains, particularly physical activity. Moreover, we did not have measures of the more direct impact of an active lifestyle on physical fitness such as vital capacity, cerebral blood flow, or the like. Other studies using such measures have provided positive evidence for a link between changes in activity and changes in cognition. For example, Albert et al. (1995) found that physiological function measures predicted differential cognitive change in old age. However, the present results offer no such connection.

Our structural equation model did identify a relationship of level and change in Novel Information Processing—our measure of intellectually engaging activities—to changes in working memory, and through working memory, to other cognitive variables. Individuals who participate in intellectually challenging activities are less likely to show cognitive decline, and those individuals who maintain their participation in such activities are also less likely to show cognitive change over time. These findings are consistent with the hypothesis that complex environments, and maintenance of engagement with such environments, contribute to enhancement and maintenance of cognitive functioning (Schaie, 1996; Schooler, 1987). However, although these outcomes support the intellectual engagement hypothesis, the evidence is far from definitive. First, there were strong relationships between initial status on Novel Information Processing and initial status on all cognitive variables. These cross-sectional relationships may indicate simply that brighter and more educationally advantaged people are more likely to participate in intellectually demanding activities. A similar interpretation of relationships of Openness to Experience with level of cognitive performance is also plausible. Moreover, our alternative model specifying that general cognitive decline predicts declines in intellectual activity fit as well as the model specifying that intellectual activity acted to maintain cognitive performance. This outcome certainly does not mean that the intellectual engagement hypothesis is wrong; on the other hand, it does serve to remind us that the kind of evidence used to argue for the hypothesis is also amenable to alternative interpretations.

What, then, should be made of the limited support of these data, relative to other longitudinal data sets, for the importance of activity, subjective health, and personality as predictors of individual differences in cognitive decline in old age? It is possible that some of the discrepancies with other studies can be attributed to differences in measures and design. First, our longitudinal study covered 6 years of change at present, and other studies have examined change over longer periods of time. If the benefits of social and physical activities on rates of cognitive change are small, then perhaps longer periods of time are needed to observe differential effects. Second, we only measured individuals in late middle age through old age, and it is possible that the changes in activity patterns that are crucial for late-life cognition occur earlier in the life cycle. Finally, of course, the select nature of the VLS sample may constrain results by limiting heterogeneity in health and activity, as noted earlier.

It is interesting, then, to compare results of our analyses with other longitudinal studies to identify possible sources of differences in outcomes and inferences regarding the influences of activity and self-reported health on cognitive change. The most salient discrepancy is with the results reported by Gold et al. (1995) from their long-term follow up of World War II Canadian military conscripts. These authors reported structural equation models they interpreted as demonstrating that (a) subjective health was a very strong predictor of change in nonverbal intelligence (defined by variables such as block design), and (b) an active lifestyle predicted changes in verbal intelligence. Differences in characteristics of the samples may be one potential source of the discrepancy between our results and those reported by Gold et al. Their sample is less select than the VLS sample and may contain greater heterogeneity in the variables of interest, particularly at the lower end of the distributions.

However, several apparent anomalies in the Gold et al. (1995) model led us to conduct a careful examination and reanalysis of their data. We noted that their zero-order correlations of subjective health (in this case, reported illnesses) with cognitive functioning were all $-.11$ or less (across six different intelligence tests, replicated at two points in time). Yet their model estimated the standardized direct effect of subjective health on nonverbal intelligence at Time 2 (controlling for Time 1 through an autoregressive coefficient) to be $-.78$. In their article, Gold et al. attributed this difference to a suppressor effect, but they have subsequently acknowledged that the coefficient was not statistically significant and should have been trimmed from the model (Gold, Andres, Etezadi, & Arbuckle, 1998). Hence it appears that subjective health does not predict cognitive change in their data either, a conclusion consistent with our results as well.

We also examined more closely the Gold et al. (1995) results supporting the hypothesis that an active lifestyle leads to maintenance of verbal intelligence. Their activities measure is a composite of items representing the habitual frequency with which social, recreational, physical, intellectual, and other activities are performed, with each item weighted with respect to degree of intellectual effort required. This weighting system may serve to make their activity measure behave empirically like our measure of novel information processing, which measures the frequency of putatively intellectually demanding activities. In this sense, both studies could be considered to show evidence that intellectually engaging activities are positively related to cognitive change. We

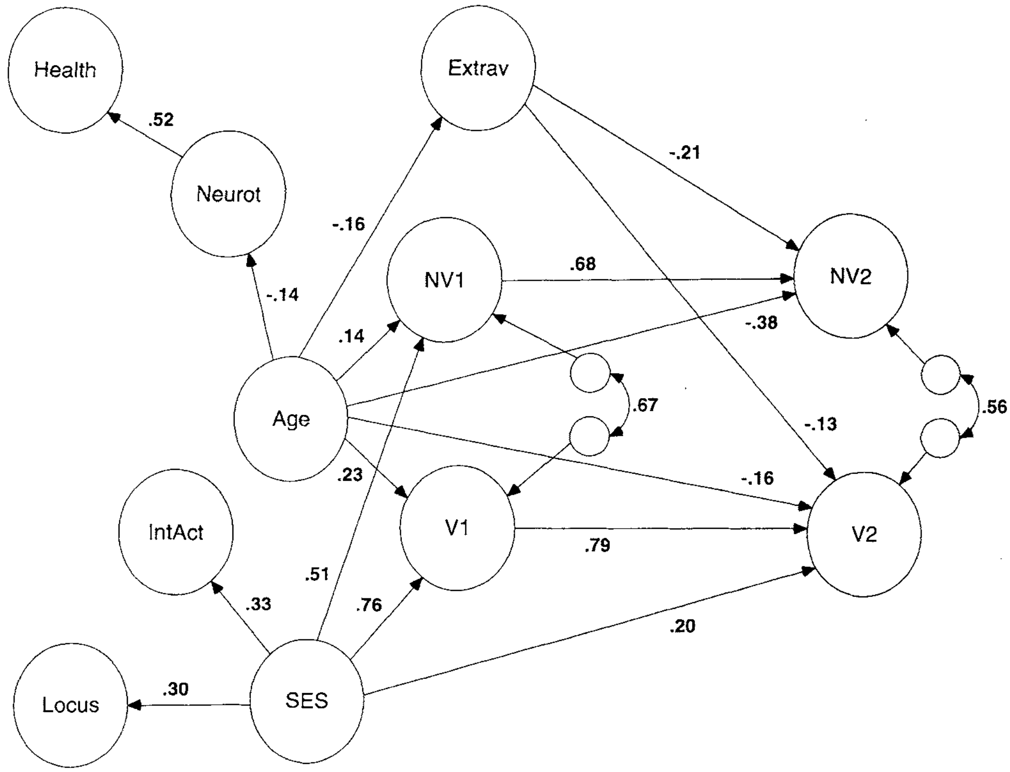

Figure 3. Alternative model for the data reported by Gold et al. (1995). Locus = Locus of Control; SES = Socioeconomic Status; IntAct = Intellectual Activities; Age = Chronological Age at testing; Neurot = Neuroticism; Extrav = Extraversion; Health = Self-Reported Health; NV1 = Nonverbal Intelligence, Time 1; NV2 = Nonverbal Intelligence, Time 2; V1 = Verbal Intelligence, Time 1; V2 = Verbal Intelligence, Time 2.

had several concerns about their model, however, one of which is addressed here. Specifically, Gold et al. estimated a latent variable of lifestyle using measures of socioeconomic status, locus of control, and intellectually engaging activities—rather than using activities itself as a predictor of cognitive change. However, the zero-order correlations of these variables were only .15, .17, and .20 (magnitudes conventionally considered too low for defining a latent variable). The low correlations among these variables may reflect not only a lack of empirical association but also a construct that is theoretically imprecise. For example, the authors do not present an argument for why global locus of control should be considered a measure of lifestyle. Some theorists have suggested that locus of control is a cause of activity, not an outcome of it. Moreover, socioeconomic status is at least as likely to be a cause of, rather than a measure of an active lifestyle (see Bollen, 1989, for more on the cause indicator vs. effect indicator distinction).

Figure 3 graphs the latent variable relations for an alternative model for the Gold et al. (1995) data.[4] Although we could have estimated a latent change model, we used an autoregressive specification similar to that of Gold et al. so that our model can be compared with their original results. Briefly, the model has no direct effect of intellectual activities or subjective health on Time 2 intelligence and fits their data well, $\chi^2(154, N = 316) = 297.01$, CFI = .946. We confirmed their finding that age and extraversion predicted change in verbal and nonverbal intelligence over the 40-year longitudinal period. Socioeconomic status strongly predicted both intelligence factors at Time 1, with a higher relationship to verbal intelligence, as would be expected from the literature on predictors of psychometric intelligence. More important, socioeconomic status predicted verbal intelligence at Time 2, indicating that change in verbal intelligence was related to social status. In contrast, there was no hint of a path from health or intellectual activities to intelligence at Time 2. Indeed, a model adding a path from intellectual activities to verbal intelligence at Time 2 resulted in a nonsignificant standardized effect of .04 ($z = 1.27, p > .25$). Thus it appears that social status, not activities, accounts for the changes in verbal intelligence in the Gold et al. data. In summary, we believe that their model produced inflated estimates of the relationships of activity and health to cognitive change.

[4]We used correlations provided by David Andres, augmented by standard deviations reported by Gold et al. (1995) to construct an input covariance matrix. We thank Dr. Andres and the other collaborators on Gold et al. (1995) for providing the data. We do not show several relationships in this model to simplify the presentation. All factor loadings on intelligence factors were constrained equal over time (and were highly significant). We also specified correlated residuals of intelligence tests over time. The single indicators of age, extraversion, neuroticism, locus of control, and subjective health were modeled with fixed-zero residual variances. Age had a specific (positive) path to Vocabulary at Times 1 and 2. A LISREL8 program estimating this model is available on request.

The second set of findings that have been used to argue for lifestyle and activity effects on cognitive change come from Schaie's (1996) SLS. These results are harder to evaluate and compare with the results of the present study. There are certainly differences in approach that could account for the differences in findings. In particular, Schaie and associates (Gribbin et al., 1980; Schaie, 1984) used indicators of socioeconomic status to help define lifestyle clusters, and it is possible that membership in these clusters was more influenced by social status than by lifestyle and intellectually demanding activities per se.

Thus, it may be the case that the differences in results among these three longitudinal studies are largely a function of whether socioeconomic status is treated as an indicator of an active lifestyle. If so, it is important to consider whether operational definitions of the lifestyle construct should include or exclude indicators of social status. Clearly, individuals' activity patterns are shaped by educational background and available economic resources. From this perspective, such indicators are integral to the lifestyle construct. A problem, however, is that social status variables are also likely to affect cognitive abilities independently of their relationships to activity lifestyle. We argue that it is more important to define activity patterns independent of socioeconomic factors, treating these latter factors as influences on activity. It remains to be seen whether the use of socioeconomic indicators of activity patterns accounts for the apparent discrepancies with the SLS. Many cross-sectional and longitudinal studies have indicated that developmental curves for cognition are not moderated by the variables of income, occupational status, and education (e.g., Hultsch et al., 1998; Lindenberger & Baltes, 1997; Owens, 1959), but such effects have been reported in the literature (see Blum & Jarvik, 1974).

In summary, the VLS produced empirical evidence that is consistent with the hypothesis that intellectually engaging activities buffer against longitudinally measured cognitive decline. However, these findings are also broadly consistent with the hypothesis that high-ability adults lead intellectually active lives until cognitive decline in old age begins to limit their intellectually related activities. In general, the evidence in favor of positive lifestyle effects on cognitive change appears to us to be less compelling than recent research reports and some of the secondary reviews of the literature might suggest. Of course, it could be the case that the VLS will produce more compelling evidence of lifestyle relationships to cognitive change when we have longer term longitudinal data, or when we can separate terminal decline individuals from other older adults using death records and other data after sufficient incidence of mortality occurs in this sample.

References

Albert, M. S., Jones, K., Savage, C. R., Berkman, L., Seeman, T., Blazer, D., & Rowe, J. W. (1995). Predictors of cognitive change in older persons: MacArthur studies of successful aging. *Psychology and Aging, 10,* 578–589.

Alwin, D. F. (1988). Structural equation models in research on human development and aging. In K. W. Schaie, R. T. Campbell, W. Meredith, & S. C. Rawlings (Eds.), *Methodological issues in aging research* (pp. 71–170). New York: Springer.

Arbuckle, T. Y., Gold, D. P., & Andres, D. (1986). Cognitive functioning of older people in relation to social and personality variables. *Psychology and Aging, 1,* 55–62.

Arbuckle, T. Y., Gold, D. P., Andres, D., Schwartzman, A. E., & Chaikelson, J. (1992). The role of psychosocial context, age, and intelligence in memory performance of older men. *Psychology and Aging, 7,* 25–36.

Arbuckle, T. Y., Gold, D. P., Chaikelson, J., & Lapidus, S. (1994). Measurement of activity in the elderly: The Activities Checklist. *Canadian Journal on Aging, 13,* 550–565.

Ball, D. (1998, April). *The relationship between activity lifestyle, self-reported health, and cognition in a representative sample of older adults.* Paper presented at Cognitive Aging Conference, Atlanta, GA.

Baltes, P. B., & Baltes, M. M. (1990). Psychological perspectives on successful aging: The model of selective optimization with compensation. In P. B. Baltes & M. M. Baltes (Eds.), *Successful aging: Perspectives from the behavioral sciences* (pp. 1–34). New York: Cambridge University Press.

Baltes, P. B., & Willis, S. L. (1982). Plasticity and enhancement of intellectual functioning in old age: Penn State's Adult Development and Enrichment Project (ADEPT). In F. I. M. Craik & S. Trehub (Eds.), *Aging and cognitive processes* (pp. 353–389). New York: Plenum Press.

Bentler, P. M. (1995). *EQS Structural Equations Program Manual.* Encino, CA: Multivariate Software.

Blum, J. E., & Jarvik, L. F. (1974). Intellectual performance in octogenarians as a function of education and initial ability. *Human Development, 17,* 364–375.

Bollen, K. A. (1989). *Structural equations with latent variables.* New York: Wiley.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies.* New York: Cambridge University Press.

Christensen, H., Korten, A. E., Jorm, A. F., Henderson, A. S., Scott, R., & Mackinnon, A. J. (1996). Activity levels and cognitive functioning in an elderly community sample. *Age and Ageing, 25,* 72–80.

Costa, P. T., Jr., & McCrae, R. R. (1985). *The NEO personality inventory manual: Form S and Form R.* Odessa, FL: Psychological Assessment Resources.

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

Costa, P. T., Jr., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual.* Odessa, FL: Psychological Assessment Resources.

Costa, P. T., Jr., & McCrae, R. R. (1994). Set like plaster: Evidence for the stability of adult personality. In T. F. Heatherton & J. L. Weinberger (Eds.), *Can personality change?* (pp. 21–40). Washington, DC: American Psychological Association.

Craik, F. I. M., Byrd, M., & Swanson, J. M. (1987). Patterns of memory loss in three elderly samples. *Psychology and Aging, 2,* 79–86.

Dixon, R. A., & Bäckman, L. (1995). Concepts of compensation: Integrated, differentiated, and Janus-faced. In R. A. Dixon and L. Bäckman (Eds.), *Compensating for psychological deficits and declines: Managing losses and promoting gains* (pp. 3–19). Hillsdale, NJ: Erlbaum.

Dixon, R. A., & Hertzog, C. (1996). Theoretical issues in cognition and aging. In F. Blanchard-Fields & T. M. Hess (Eds.), *Perspectives on cognitive change in adulthood and aging* (pp. 25–65). New York: McGraw-Hill.

Dixon, R. A., Hultsch, D. F., & Hertzog, C. (1989). *A manual of 25 three-tiered structurally equivalent texts for use in aging research.* (CRGCA Technical Report No.2). Victoria, British Columbia, Canada: University of Victoria, Department of Psychology.

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

Erber, J. R., & Szuchman, L. T. (1996). Memory performance in relation to age, verbal ability, and activity. *Experimental Aging Research, 22,* 59–72.

Ericsson, K. A., & Charness, N. (1994). Expert performance: Its structure and acquisition. *American Psychologist, 49,* 725–747.

Gold, D. P., Andres, D. Etezadi, J., & Arbuckle, T. (1998). "Structural equation model of intellectual change and continuity and predictors of intelligence in older men": Erratum. *Psychology and Aging, 13,* 434.

Gold, D. P., Andres, D., Etezadi, J., Arbuckle, T. Y., Schwartzman, A. E., & Chaikelson, J. (1995). Structural equation model of intellectual change and continuity and predictors of intelligence in older men. *Psychology and Aging, 10,* 294–303.

Gribbin, K., Schaie, K. W., & Parham, I. A. (1980). Complexity of lifestyle and maintenance of intellectual abilities. *Journal of Social Issues, 36,* 47–61.

Hertzog, C. (1990). On the utility of structural equation models for developmental research. In P. B. Baltes, D. L. Featherman, & R. M. Lerner (Eds.), *Life-span development and behavior* (Vol. 10, pp. 257–290). Hillsdale, NJ: Erlbaum.

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

Hertzog, C., Schaie, K. W., & Gribbin, K. (1978). Cardiovascular disease and changes in intellectual functioning from middle to old age. *Journal of Gerontology, 33,* 872–883.

Hill, R. D., Wahlin, A., Winblad, B., & Bäckman, L. (1995). The role of demographic and lifestyle variables in utilizing cognitive support for episodic remembering among very old adults. *Journals of Gerontology: Psychological Sciences, 50B,* P219–P227.

Hultsch, D. F., Hammer, M., & Small, B. (1993). Age differences in cognitive performance in later life: Relationships to self-reported health and activity life style. *Journals of Gerontology: Psychological Sciences, 48,* 1–11.

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., Hertzog, C., Dixon, R. A., & Small, B. J. (1998). *Memory change in the aged.* New York: Cambridge University Press.

Hultsch, D. F., Hertzog, C., Small, B., McDonald-Miszczak, L., & Dixon, R. A. (1992). Short-term longitudinal change in cognitive performance in later life. *Psychology and Aging, 7,* 571–584.

Jenkins, J. J. (1979). Four points to remember: A tetrahedral model of memory experiments. In L. S. Cermak & F. I. M. Craik (Eds.), *Levels of processing in human memory* (pp. 429–446). Hillsdale, NJ: Erlbaum.

Kliegl, R., Smith, J., & Baltes, P. B. (1990). On the locus and process of magnification of age differences during mnemonic training. *Developmental Psychology, 26,* 894–904.

Liang, J. (1986). Self-reported physical health among aged adults. *Journal of Gerontology, 41,* 248–260.

Lindenberger, U., & Baltes, P. B. (1997). Intellectual functioning in old and very old age: Cross-sectional results from the Berlin Aging Study. *Psychology and Aging, 12,* 410–432.

Luszcz, M. A., Bryan, J., & Kent, P. (1997). Predicting episodic memory performance of very old men and women: Contributions from age, depression, activity, cognitive ability, and speed. *Psychology and Aging, 12,* 340–351.

MacCallum, R. C., Wegener, D. T., Uchino, B. N., & Fabrigar, L. R. (1993). The problem of equivalent models in applications of covariance structure analysis. *Psychological Bulletin, 114,* 185–199.

McArdle, J. J., & Anderson, E. (1990). Latent variable growth models for research on aging. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (pp. 21–44). San Diego, CA: Academic Press.

McArdle, J. J., & Nesselroade, J. R. (1994). Using multivariate data to structure developmental change. In S. H. Cohen & H. W. Reese (Eds.), *Life-span developmental psychology: Methodological contributions* (pp. 223–267). Hillsdale, NJ: Erlbaum.

McCrae, R. R., & Costa, P. T., Jr. (1990). *Personality in adulthood.* New York: Guilford Press.

Nelson, T. O., & Narens, L. (1980). Norms of 300 general-information questions: Accuracy of recall, latency of recall, and feeling-of-knowing ratings. *Journal of Verbal Learning and Verbal Behavior, 19,* 338–368.

Owens, W. A., Jr. (1959). Is age kinder to the initially more able? *Journal of Gerontology, 14,* 334–337.

Ramsey, P. H. (1982). Empirical power of procedures for comparing two groups on *p* variables. *Journal of Educational Statistics, 7,* 139–156.

Rogers, R. L., Meyer, J. S., & Mortel, K. F. (1990). After reaching retirement age physical activity sustains cerebral perfusion and cognition. *Journal of the American Geriatrics Society, 38,* 123–128.

Rogosa, D., Brandt, D., & Zimowski, M. (1982). A growth curve approach to the measurement of change. *Psychological Bulletin, 92,* 726–748.

Salthouse, T. A. (1991). *Theoretical perspectives on cognitive aging.* Hillsdale, NJ: Erlbaum.

Salthouse, T. A. (1996). General and specific speed mediation of adult age differences in memory. *Journals of Gerontology: Psychological Sciences and Social Sciences, 51B,* P30–P42.

Salthouse, T. A., & Babcock, R. L. (1991). Decomposing adult age differences in working memory. *Developmental Psychology, 27,* 763–776.

Schaie, K. W. (1984). Midlife influences upon intellectual functioning in old age. *International Journal of Behavioral Development, 7,* 463–478.

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

Schaie, K. W., & Willis, S. L. (1986). Can decline in adult intellectual functioning be reversed? *Developmental Psychology, 22,* 223–232.

Schooler, C. (1987). Psychological effects of complex environments during the life span: A review and theory. In C. Schooler & K. W. Schaie (Eds.), *Cognitive functioning and social structure over the life course* (pp. 24–49). Norwood, NJ: Ablex.

Schooler, C. (1990). Psychosocial factors and effective cognitive functioning in adulthood. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (pp. 347–358). San Diego, CA: Academic Press.

Statistics Canada. (1989a). *Census of Canada, 1986: Population and dwelling characteristics. Occupations* (Catalogue # 93-112). Ottawa, Ontario, Canada: Statistics Canada.

Statistics Canada. (1989b). *Census of Canada, 1986: Population and dwelling characteristics. Schooling and major field of study* (Catalogue # 93-110). Ottawa, Ontario, Canada: Statistics Canada.

Stelzl, I. (1986). Changing a causal hypothesis without changing the fit: Some rules for generating equivalent path models. *Multivariate Behavioral Research, 21,* 309–331.

van Boxtel, M. P. J., Langerak, K., Houx, P. J., & Jolles, J. (1996). Self-reported physical activity, subjective health, and cognitive performance in older adults. *Experimental Aging Research, 22,* 363–380.

Verbrugge, L. M., Gruber-Baldini, A. L., & Fozard, J. L. (1996). Age differences and age changes in activities: Baltimore Longitudinal Study of Aging. *Journals of Gerontology: Social Sciences, 51B,* S30–S41.

Verhaeghen, P. (1993). *Teaching old dogs new tricks: Plasticity in episodic memory performance in old age.* Leuven, Belgium: Catholic University of Leuven.

Received July 23, 1997
Revision received June 29, 1998
Accepted July 14, 1998 ■