Are processing speed tasks biomarkers of cognitive aging?

Ian J. Deary, Wendy Johnson, John M. Starr

University of Edinburgh, UK
Abstract

We examined the association between five processing speed measures and general cognitive ability in a large (> 900) sample of relatively healthy men and women at age 70. The processing speed tests were Wechsler Digit Symbol-Coding and Symbol Search, simple reaction time, 4-choice reaction time, and inspection time. To inquire whether the processing speed tasks might be biomarkers of cognitive aging, the attenuations in their associations with general cognitive ability were examined after adjusting for cognitive ability measured almost 60 years earlier. With the exception of inspection time, the attenuations were substantial. Inspection time was the only processing speed measure—all of which were measured at age 70—whose correlation with cognitive ability at age 70 was significantly greater than the correlation with cognitive ability at age 11. In old age, individual differences in most commonly-used measures of processing speed are largely dependent on childhood cognitive ability. For all processing speed tasks a little variance is left that appears to be related to aging differences. Inspection time, the marker that was least dependent on childhood intelligence, should be considered further as one biomarker of cognitive aging.
Introduction

Processing speed and cognitive abilities

The speed with which mental processes can be carried out has featured as an important aspect of the psychology of cognitive differences from the inception of scientific study within the field (Cattell, 1890; Galton, 1890). Whereas some mental tests measure how much people know, how much they can remember, or how much information they can both retain and transform, mental speed tasks typically examine how efficiently people can complete series of items with simple cognitive content. Sometimes called mental speed, speed of information processing, or just processing speed, this construct features as a group factor in models and batteries of intelligence. Carroll’s (1993, p. 626) three stratum model of the psychometric structure of intelligence is widely accepted. Stratum I comprises specific abilities, Stratum II comprises major domains of cognitive function, and Stratum III is general mental ability (usually called g), the variance shared by the major domains. “Processing Speed” is a Stratum II ability, one level below general intelligence at Stratum III. The Wechsler Adult Intelligence Scale-III has a Processing Speed factor at the same level, and it has a high loading on the general intelligence (g) factor (Deary, 2001, p. 128; Wechsler, 1997).

The measurement of processing speed in humans takes place on at least three levels of description (Deary, 2000). First, at a relatively molar level, it is assessed using psychometric tests such as the Digit Symbol-Coding and Symbol Search subtests of the Wechsler (1997) intelligence scales. In these tests, each item requires a simple decision. If there were unlimited time, most people would complete all items correctly. Digit Symbol-Coding, for example, requires the subject to enter a simple symbol below each number in a long series, and the look-up code for each number is printed on the same page as the test’s items. Scores are based on the numbers of items completed correctly in the time given.
Second, and third, at arguably lower levels of description, processing speed is assessed using tasks derived not from psychometric tests, but from cognitive-experimental psychology and psychophysics, respectively. In both of these latter cases, the stimulus-response contingencies are more tightly controlled, the item content is even simpler, and the relevant completion times for items are much shorter than for psychometric test items (Deary, 2000). From experimental psychology, simple and choice reaction time tasks are used to assess processing speed. From psychophysics, inspection time is used to assess processing speed. All three are correlated significantly with higher-level cognitive functions (Deary, Der, & Ford, 2001; Grudnik and Kranzler, 2001). There is, arguably, an even lower level of processing speed, involving the efficiency of neural transmission and neural systems. This is more problematic to measure in vivo in humans, and is not measured or addressed further in detail in the present study.

Processing speed—however it is assessed—holds a disputed position within the psychology of individual differences. As described above, some position it as a group factor within the multi-level structure of cognitive differences. A group factor represents a major domain of cognitive function—more general than a specific task, and more specific than general intelligence—such as may be found at Stratum II in Carroll’s (1993) model. At the same time, processing speed has always had its champions as a possible cause of individual differences in cognitive abilities (see Deary, 2000, for a discussion of this). That is, rather than merely being one of the group factors of intelligence that is loaded on \( g \), it might be a cause of individual differences in the other group factors, and of \( g \) too. Those who argue along these lines have suggested that processing speed could be an index of one fundamental capacity of the central nervous system and, as such, its variance might be shared with those of higher-level cognitive tasks because the speed with which so-called
elementary cognitive operations can occur dictates the efficiency of more complex mental operations (Deary, 2000).

Processing speed and aging

Processing speed declines (slows), on average, as people grow older. This occurs with whichever type of processing speed task is used: psychometric (Hedden & Gabrieli, 2004; Salthouse, 2004), cognitive-experimental (Der & Deary, 2006), or psychophysical (Gregory, Nettelbeck, Howard, & Wilson, 2008; Nettelbeck & Rabbitt, 1992). However, many other types of mental function also show average declines as people grow older, such as aspects of memory, reasoning, and executive functions (Hedden & Gabrieli, 2004; Salthouse, 2004; Schaie, 2005). We described above that processing speed has a debated position within the structure of intelligence. Researchers argue whether processing speed is merely one of a number of domains of cognitive ability—all of which load on general cognitive ability—or whether it is a more fundamental construct, which might be one cause of individual differences in general cognitive ability. There is a similar debate about the place and importance of processing speed within the field of cognitive aging. That is, is processing speed just one of a number of mental capabilities that show age-related declines; or is processing speed one fundamental cause of age-related declines in other cognitive capabilities? Madden (2001, p. 288) described speed of information processing as a “fundamental property of the nervous system,” one which provides a foundation for the efficient implementation of other cognitive functions. Salthouse (1996) championed the idea that processing speed might be one major cause of age-related declines in other cognitive functions. He marshalled much—mostly cross-sectional—data to show that, when associations between mental test scores (e.g. on memory and reasoning tests) and chronological age were statistically adjusted for processing speed, the associations were
severely attenuated (Salthouse, 1996; Verhaegen & Salthouse, 1997). The use of cross-sectional data for identifying biomarkers of cognitive aging has been criticised on methodological grounds because there is no way to establish the direction of causation (Hofer, Flaherty, & Hoffmann, 2006; Lindenberger & Pötter, 1998). Despite this, some of Salthouse’s more recent models show processing speed as one domain of cognition that ages alongside others, with no presumption that it is an underlying cause of cognitive aging (Salthouse, 2001, 2004; Salthouse & Ferrer-Caja, 2003). Others also take this position (e.g., Finkel, Reynolds, McArdle, & Pedersen, 2005; Wilson et al., 2002). Moreover, when longitudinal data are used, processing speed has a much less strongly attenuating influence on the age-cognitive ability association (Lemke and Zimprich, 2005; Zimprich & Martin, 2002).

In summary, processing speed is an important domain of cognitive ability, and one which shows dramatic mean decline with age. Moreover, within the hierarchical structure of cognitive abilities, and within cognitive aging, some have suggested that processing speed might be a fundamental component of individual differences and age-related change in other mental capabilities. In this vein, it has recently been suggested that processing speed—especially inspection time—might be useful as a biological marker of cognitive aging (we discuss the characteristics of biological markers in the next section). The reasons for doing this include the findings that inspection time is stable across the generations, and that relatively short-term changes in inspection time predict changes in other cognitive functions (Gregory et al., 2008; Nettelbeck & Wilson, 2004, 2005). Applying the term biological marker, or ‘biomarker’, to processing speed implies that processing speed assessments tap into some fundamental capacities of the central nervous system (Madden,
2001), and that the level of the biomarker is useful as an indicator of at least a more general cognitive system state.

*What properties should a biomarker have?*

There is no consensus about what qualities a biomarker of aging in general or a biomarker of cognitive aging in particular should have. Ingram (2006) provided a useful overview, as had McClearn (1997), with several overlapping definitions and suggestions. A simple idea is that, “a biomarker of aging is a biological parameter intended as a quantitative measure of the rate of aging in an organism that represents a more accurate index than can be provided by the organism’s chronological age” (Ingram, 2006, p. 137). This is similar to Baker and Sprott’s (1988) much-cited definition, as follows, “A biological parameter of an organism that either alone or in some multivariate composite will, in the absence of disease, better predict functional capacity at some later age than will chronological age” (p. 223). To these ideas Miller (2001) added that a biomarker of aging, “should predict the outcome of a wide range of age-sensitive tests in multiple physiological domains, in an age-coherent way, and do so better than chronological age… [It] should predict remaining longevity at an age at which 90% of the population is still alive”.

McClearn (1997) usefully discusses the problem of there being no agreed upon, standard aging criterion measure against which putative biomarkers can be evaluated. One problem in translating biomarker concepts into the psychological domain is that much of the work with them has been conducted in the context of short-lived, non-human species, and has been relatively uninformed by psychometric criteria. Ingram (2006) attempted partially to correct this and provided three additional criteria for a biomarker of aging, as follows, “significant cross-sectional correlation with age… significant longitudinal change with age consistent with cross-sectional correlation… significant stability of individual differences”
(p. 138). These will resonate with psychologists, especially, and may be applied to processing speed measures. Examples of variables that have been used as biomarkers of aging include handgrip strength, blood pressure and visual acuity (as used by Gregory et al., 2008), and many more including indices of lung function and blood markers such as glycated hemoglobin.

Against the above criteria, how well do measures of processing speed fare? For example, does choice reaction time meet these statistical biomarker criteria? It shows stable individual differences across several years in large, population-representative, adult samples (Deary & Der, 2005b). It shows a clear pattern of age-related deterioration from young adulthood to old age in population-representative samples (Der & Deary, 2006). It correlates significantly, after correction for chronological age, with more complex mental tasks (Deary, Der, & Ford, 2001). Its baseline measurement, adjusted for chronological age, is associated with morbidity and mortality in longitudinal studies (Deary & Der, 2005a; Shipley, Der, Taylor, & Deary, 2006). Its changes from baseline measurement across seven years, adjusted for chronological age, are independently related to morbidity and mortality (Shipley, Der, Taylor, & Deary, 2007). In simple terms, therefore, knowing about a person’s choice reaction time—and perhaps any of the other processing speed measures (Gregory et al., 2008)—could offer useful information about people’s states and/or rates of mental aging—and perhaps broader physical aging—that would not be available from their chronological ages.

To the extent that individual differences have become important in the discussion of biomarkers of aging (Ingram, 2006), there is an additional matter that must be considered: the possibility that a measure which shows strong age-related mean change does not necessarily provide information about individual differences in aging. Again, this is discussed here—merely for illustration—using choice reaction time. First, choice reaction
time correlates significantly with higher cognitive abilities in youth and in old age (Deary, 2000; Deary, Der, & Ford, 2001; Jensen, 2006). Second, choice reaction time and fluid higher cognitive abilities show strong age-related mean deteriorations in healthy population samples (Deary & Der, 2005b; Der & Deary, 2006; Schaie, 2005). Among other considerations, choice reaction time might be accepted statistically as a biomarker of the state of cognitive aging if it accounted for variance in cognitive ability in old age above and beyond that accounted for by: (1) individual differences in cognitive ability in youth that presumably reflect the lifetime stable trait of cognitive ability before aging effects set in, and (2) chronological age. If so, choice reaction time would possibly be capturing differential states of cognitive aging by indicating the changes in rank order of the cognitive scores that took place between the two ages. However, it is also possible that choice reaction time provides information only about cognitive variation in old age that has been stable throughout the lifespan. Although there is mean age-related slowing in choice reaction time, people might slow in their choice reaction times at the same rate, preserving individuals’ rank orders over time. Therefore, choice reaction time in old age could only be considered a biomarker of having aged if it provided additional information about cognitive function in old age beyond that provided by some measure of cognitive function taken in youth.

The present study

Here we inquired to what extent processing speed differences measured in old age reflect cognitive ability differences beyond those observable in youth, and thus can perhaps serve at least as statistical biomarkers of cognitive age. We examined this using data from the Lothian Birth Cohort 1936 (Deary et al., 2007), and the most commonly used measures of processing speed. The participants in the study had the following characteristics, unusual
in combination, which facilitated the investigation: the sample was large (N > 900) and relatively healthy at age 70, with very little chronological age variance. The participants completed the same validated general mental test at age 11 and 70 years, and they completed five psychometric, experimental, and psychophysical processing speed tasks at age 70. The extent to which the processing speed measures contributed variance to mental test scores at age 70 that was not captured by the same mental test at age 11 provided important indications of the extent to which measures of processing speed can be considered at least as statistical biomarkers of cognitive function orthogonal to individual differences in cognitive status in youth. Because of the variety of processing speed measures available, we were also able to compare their relative effectiveness in this regard.

Method

Participants

The participants were members of the Lothian Birth Cohort 1936 (LBC1936). This is a sample of 1091 people (548 men and 543 women) who were born in 1936 and attended school in Scotland in 1947. On June 4th 1947 they took part in the Scottish Mental Survey 1947 (Scottish Council for Research in Education, 1949; Deary, Whalley, & Starr, 2009). This was a nationwide survey in which almost all Scottish schoolchildren born in 1936 (N = 70,805, which was approximately 95% of the 1936-born population) completed a validated mental test called the Moray House Test No. 12. The recruitment of the LBC1936, the sample characteristics, and all assessments and procedures were described in detail in a free-access, full protocol of the study (Deary et al., 2007). Briefly, the Lothian area Chief Medical Officer informed potential participants about the study by writing to people living in the Edinburgh area who were born in 1936 and were registered with a general medical practitioner on the Community Health Index, and by local media
advertisements. Participants were relatively healthy, lived independently and travelled to a clinical research facility at the Western General Hospital in Edinburgh for medical and psychological testing. The mean (SD) age of the participants when they took part in the Scottish Mental Survey test on 4\textsuperscript{th} June 1947 was 10.9 years (0.28). When they returned for re-testing between 2004 and 2007 their mean (SD) age was 69.6 (0.83). Therefore, there was very little age variation either in youth or old age. No participant had any history of dementia, and those examined here scored 24 or better on the Mini-Mental State Examination (Folstein, Folstein, & McHugh, 1975), which is widely used as a screening test for dementia.

\textit{Cognitive and processing speed tests}

Details of all measures taken by the participants and the procedures applied are given in the LBC1936 protocol article (Deary et al., 2007). Here, we describe the subset of tasks that are analysed in the present study.

\textit{Moray House Test No. 12}. This was administered at school, in groups, at about age 11 years in the Scottish Mental Survey 1947 (Scottish Council for Research in Education, 1949; Deary, Whalley, & Starr, 2009). Participants took the same test again at age about 70, at which time the test was administered individually. On both occasions a 45 minute time limit was used. The test has a variety of item types: following directions (14 items), same-opposites (11), word classification (10), analogies (8), practical items (6), reasoning (5), proverbs (4), arithmetic (4), spatial items (4), mixed sentences (3), cypher decoding (2), and other items (4). The maximum score is 76. It is described in even more detail elsewhere (Deary et al., 2004). The test scores from childhood have good validity credentials. In childhood, the test correlated .81 with the individually-administered Form L Terman-Merrill
revision of the Stanford Binet test (Deary, Whalley, & Starr, 2009). Test scores from childhood were strongly associated with educational, occupational, and health outcomes in adulthood and later life (Deary, Whalley, & Starr, 2009).

Assessment of general cognitive ability. Participants undertook the following five tests from the Wechsler Adult Intelligence Scale III\textsuperscript{UK} (Wechsler, 1997): Matrix Reasoning, Letter Number Sequencing, Block Design, Digit Symbol-Coding, and Symbol Search. Using principal components analysis, the scores were saved from a first unrotated principal component, which provided an assessment of general cognitive ability (\(g\)). A second principal components analysis was applied to only the Matrix Reasoning, Letter Number Sequencing, Block Design tests, which provided an assessment of general cognitive ability (\(g'\)) without the two processing speed tests.

Psychometric assessment of processing speed: Digit Symbol-Coding and Symbol Search. These two tests are subtests of the Wechsler Adult Intelligence Scale III\textsuperscript{UK} (Wechsler, 1997). Thus, they are standard psychometric-type (paper and pencil) tests and considered tests of processing speed. They had a dual function in this study. They provided marker tasks for the extraction of a \(g\) component as well as psychometric assessments of processing speed. In Digit Symbol-Coding, the participant writes in symbols below rows of numbers 1 to 9, inclusive, according to an explicit code. The score is the number of correct symbols which are written in 120 seconds. In Symbol Search, the participant indicates whether one of two target symbols on the left of a row also appears among five symbols printed to the right. The score is the number of correct answers provided in 120 seconds. The items in both tasks require such simple decisions that, given adequate time, it would be unlikely that adults with unimpaired intelligence would make errors.

Cognitive-experimental assessment of processing speed: Simple and 4-choice reaction time. We name these tasks cognitive-experimental-level assessments of processing
speed because reaction time tasks are used widely in cognitive-experimental psychology. The content of the items is simpler than that of the psychometric assessments. A self-contained reaction time box was used. It has been described in detail by us previously, and has been employed in large population studies in the United Kingdom (Deary & Der, 2001, 2005c; Der & Deary, 2006). For simple reaction time there were 8 practice trials and 20 test trials. Simple reaction time testing was followed by 4-choice reaction time. For 4-choice reaction time there were 8 practice trials and 40 test trials. The inter-trial interval varied between 1 and 3 seconds for both simple and 4-choice reaction time. For simple reaction time, participants pressed a key when a 0 appeared in an LCD window. For 4-choice reaction time participants pressed the appropriate key when 1, 2, 3 or 4 appeared in the window; participants kept the index and middle fingers of their left and right hands lightly on the keys between trials. Participants made few errors on the 4-choice reaction time test. Scores for these tests were means for simple and 4-choice correct response reaction times.

*Psychophysical assessment of processing speed: inspection time.* We name this a psychophysical assessment of processing speed because it eliminates the motor response time that is part of the reaction times in those tasks. The task is a forced choice, two-alternative visual discrimination. The inspection time task we used here has been illustrated and described in detail elsewhere (Deary et al., 2004, 2007). All stimuli were presented on a computer with a vertical refresh rate of 160 Hz. The two stimuli that must be discriminated were pi-type shapes with markedly different parallel, vertical lines. The participant’s task was to decide whether the longer line was on the right or the left. They indicated the choice after each trial by pressing the 1 or 2 key on the number pad of a computer keyboard. Responses were made at leisure; no reaction times were recorded. The stimuli were presented ten times at each of 15 durations, ranging from 6 ms to 200 ms,
which occurred at random. The 36 participants who failed to score 17 or more out of 20 in the sum of the two easiest durations (150 ms and 200 ms) were excluded. As expected, group mean responses increased from chance (50% correct) at the shortest exposure durations to almost perfect responding (about 100% correct) at the longest durations. The score used was the total number correct out of 150 responses.

Results

Descriptive statistics and Pearson correlations are shown in Table 1. The Moray House Test at age 11 was strongly correlated with the Moray House Test and with \( g \) at age 70. Squaring this coefficient \((r = .692)\) would indicate that just under 50% of the variance was shared between age 11 and age 70 years, even without correction for period-free reliability and attenuation of range in the sample when compared with the background population. Another interpretation of this coefficient would suggest that the shared variance is actually 69.2% (Ozer, 1985). This is not a principal issue in the present study, so will not be discussed further. The mean raw scores on the Moray House Test were lower at age 11 than at age 70, reflecting the cognitive development that took place after age 11. The measures derived from the processing speed tests all correlated in the expected directions: speed in one measure was associated with speed in the others. Some tests assessed the time taken for a correct response, and some documented the number correct, but speed was implied in both.

The associations of principal interest were those between the speed measures and the more general cognitive measures (Table 1). All of the speed measures showed significant correlations with cognitive ability scores in the expected direction: greater processing speed was associated with higher scores on cognitive ability tests at age 70. Because of the large sample size, even correlations at about .1 were significant at about \( p = .001 \). All of the
processing speed measures had correlations greater than .2 with Moray House Test score at age 70. The two WAIS subtests had correlations greater than 0.4, and 4-choice reaction time was correlated greater than 0.3. Correlations with g and g' at age 70 were higher, with inspection time’s being .32 and .29, and 4-choice reaction time’s at -.46 and -.38. Therefore, the cross-sectional findings were universally supportive of processing speed measures as markers of more general cognition.

All of the speed measures taken at age 70 were significantly associated with Moray House Test scores at age 11, taken almost 60 years earlier (Table 1). The correlations between each speed measure and Moray House Test scores at ages 11 and 70 were similar. For the WAIS subtest scores and the simple and 4-choice reaction time measures, the largest difference between the correlations with age 11 and age 70 Moray House Test scores was just over .04. The differences between the correlations of these processing speed tasks with age 11 and age 70 Moray House Test scores were all non-significant. For inspection time the difference was more marked: the correlation with Moray House Test score at age 70 was .09 higher than that with the Moray House Test score at age 11. The difference between inspection time’s correlations with Moray House test at age 11 and 70 was highly significant, $t(939) = 3.71, p < .001$. Therefore, all processing speed measures taken at age 70 were associated with general cognitive ability measured about 60 years earlier; and all correlations except inspection time’s were non-significantly different in magnitude to the cross-sectional speed-cognitive ability associations.

We used several quantitative estimates—partial correlations and changes to the parameters from linear regression models—to convey the attenuation in the age 70-based, cross-sectional processing speed-cognitive ability associations after prior cognitive ability at age 11 was taken into account. It is important to show these attenuations: much research in this field is based upon cross-sectional correlations and we wished to demonstrate the
extent to which these can be dependent upon general cognitive ability differences measured almost 60 years previously.

First, the partial correlations (Table 1, upper right) between the speed measures and the Moray House Test scores at age 70 were substantially lower than the zero-order correlations (Table 1). An apparent exception was inspection time, which showed the least attenuation, and had a partial correlation with Moray House Test score at age 70 that was greater than the simple reaction time measure, and similar to that of 4-choice reaction time. The WAIS subtests’ partial correlations were substantially lower than the zero-order associations too. The same pattern of results was seen between the reaction time and inspection time measures and \( g \) at age 70, and between all of the processing speed tests and \( g' \) at age 70. Here, the greatest partial correlations were between 4-choice reaction time mean and \( g \) and between Symbol Search and \( g' \) (Table 1, upper right), but the association that showed the least attenuation after partialling out age 11 IQ was that between inspection time and \( g \), which remained at .32.

We then re-computed the correlations and partial correlations adjusting for estimates of the reliabilities of the processing speed measures. Correction of the zero-order correlations was performed according to McNemar (1969, p. 171 equation 10.20) and Cohen, Cohen, West, & Aiken (2003, p. 119, equation 2.10.5). Partial correlations were corrected for unreliability according to Cohen, Cohen, West, & Aiken (2003, p.120, equation 4.3.5). The resulting coefficients are shown in parentheses below the uncorrected correlations in Table 1. We computed two sets of corrected correlations, as described below.

The first set of unreliability-corrected correlations was based upon an as-yet incomplete subsequent wave of testing of the Lothian Birth Cohort 1936. There were 410 subjects with complete data on the five processing speed measures at wave 1 and wave 2.
Their mean age at wave 1 was 68.9 (SD = 0.5), and at wave 2 was 72.0 (SD = 0.5). The mean time between the two tests was 3.1 years (SD = 0.2). The test-retest correlations were: Digit Symbol-Coding = .832; Symbol Search = .648; simple reaction time = .560; 4-choice reaction time = .775; inspection time = .574. These were used as the reliability coefficients. For the other variables—Moray House Test at age 11 and 70 years, and g and g’—we used 0.900 as the estimated reliability coefficient. The principal effects of these corrections were relatively to elevate the correlations of simple reaction time and inspection time, the measures with relatively low reliabilities. A good example is that, corrected for unreliability, the age 11 Moray House Test-adjusted coefficients between g and inspection time and 4-choice reaction time were almost identical.

The second set of unreliability-corrected correlations was computed for all of the processing speed tests except Symbol Search. This second set of corrections for unreliability had the advantage over the Lothian Birth Cohort wave 2 data that they were tested over a much shorter retest period. For simple and 4-choice reaction time mean the reliability coefficients were obtained from period-free reliability estimates provided in Deary & Der (2005c, p. 119). They applied the same reaction time procedure as used in the present study to 49 healthy adults (18 men, 31 women) with a mean (SD) age of 37.1 (11.4) years. The maximum re-test period was one day. The reliability coefficient for simple reaction time mean was .67, and for 4-choice reaction time mean was .92. For Digit Symbol and inspection time, the reliability estimates were obtained from unpublished data from a PhD student of one of the authors (see Author Note). The same Digit Symbol and inspection time procedures as used in the present study were applied to 80 healthy adults (31 men, 49 women) with a mean (SD) age of 49.0 (16.4) years. The mean (SD) number of days between tests was 9.2 (7.3), with a range of 1 to 37. The reliability coefficient for Digit Symbol was .934, and for inspection time was .805. As expected, these have the
effect (Table 1) of producing correlations that fall between the uncorrected correlations and those that were corrected using the three-year retest of the Lothian Birth Cohort 1936. The corrections for unreliability do not alter the pattern of results seen with the uncorrected correlations; in particular, they do not alter the general pattern of differences in attenuation among the processing speed measures' correlations with cognitive ability at age 70 after adjusting for Moray House Test score at age 11.

Second, we demonstrated the attenuation of the age 70 processing speed-cognitive ability associations after adjusting for cognitive ability at age 11 using linear regression to compare two models. In the first model, each speed measure was entered individually as an independent variable with Moray House Test and $g$ at age 70 as the dependent variables. The $R^2$ values for these models are shown in Table 2. In the second model, Moray House Test score at age 11 was added to the speed measure as a second independent variable. Again, the $R^2$ values are shown in Table 2. Also shown are the percentage attenuations of the $R^2$ values after adding age 11 Moray House Test score. For the models in which Moray House Test score at age 70 was the dependent variable, the attenuations of $R^2$ values for the WAIS and reaction time variables were at or greater than 85%. The attenuation of the inspection time $R^2$ was less, at 64%. For the models in which $g$ and $g'$ at age 70 was the dependent variable, the attenuations of $R^2$ values for the reaction time measures were between 57% and 60%, and that for inspection time was only 42% and 14%, respectively. The $t$ statistics from the regression models were compared in the same way. Again, there were substantial attenuations to their values when Moray House Test at age 11 was added to the models, and the attenuations were much less for inspection time than for the other speed variables. This applied whether Moray House Test age 70 or $g$ or $g'$ was the dependent variable.
The foregoing analyses were used to discover the extent to which individual differences in processing speed and associations between processing speed and more general cognitive ability in old age were the legacies of individual differences in Moray House Test ability in youth. The third analyses examined the extent to which the processing speed measures differed in their correlations with general cognitive ability after partialling out cognitive ability at age 11. This was important in order to establish which had the better predictive power, independent of original mental ability. This was achieved by conducting formal significance tests of the differences between these partial correlations. With respect to partial correlations with Moray House Test at age 70, the two WAIS-III processing speed tests were almost universally stronger than the reaction time and inspection time variables (Table 3). There were no significant differences in the partial correlations with Moray House Test at age 70 among the reaction time and inspection time variables. With respect to the partial correlations with $g$, the 4-choice reaction time and inspection time measures had stronger partial correlations than simple reaction time (Table 3). The inspection time and 4-choice reaction time partial correlations did not differ significantly. With respect to the partial correlations with $g'$, the Symbol Search and Digit Symbol correlations were almost all larger than those with the reaction and inspection time tests. The inspection time and 4-choice reaction time correlations were larger than those with simple reaction time.

Discussion

The dramatic age-related decline in processing speed measures and the correlations of these measures with more general cognitive ability make processing speed measures attractive as potential biomarkers of cognitive age, and perhaps bodily age more generally. Here we found that, despite these features, individual differences in most processing speed
measures at age 70 still substantially reflected the individual differences in general cognitive ability that existed 60 years previously. The speed measures explained only a small amount of additional variance in age 70 cognitive function after taking age 11 cognitive function into consideration. Still, most of them did provide some significant additional explanatory power. They did not appear to be strong statistical biomarkers, but they did act in this manner to some degree. Inspection time performed the best of the non-psychometric tasks in this regard; though its overall correlations with cognitive ability were not the highest, it had the greatest amount of predictive power independent of age 11 cognitive function.

Strengths of the study were: the large sample size; the availability of childhood cognitive ability data and follow-up cognitive test data (including the same test that had been taken in childhood) almost 60 years later; the testing of five processing speed tasks at three different levels of description; and the very narrow range of chronological age in the sample. Of course, the results are reported from a particular birth cohort in a particular geographical location, and tested on specific processing speed and cognitive tasks, and one cannot assume generalisability beyond that. The estimates of reliability obtained from the retest of the Lothian Birth Cohort 1936 were not period-free. The three-year ‘reliability’ measures we used also included some differences in age-related change between measurement occasions, which might or might not be trivial across such a short period in healthy subjects. However, we did not assume this and, for four of the five processing speed tests, we had nearly period-free reliability estimates based on healthy adult subjects. McClearn (1997) discussed how to assess the reliability of a biomarker. This latter point also prompts mention of the fact that we did not have information on the relative sensitivity of the different speed measures to individual differences in change. Also, even if the constructs assessed by the different speed measures showed the same amounts of change
in nature, the tests or tasks by which these changes were measured may differ in their sensitivity to capture those changes. This cannot be addressed empirically because the speed measures were not applied at age 11, and is a methodological limitation of the study.

The fact that Moray House Test mean scores were higher at age 70 than at age 11 indicates that the age 11 scores represented a mixture of lifetime stable individual differences, mean maturation level, and individual differences in relative maturation. Therefore, by controlling for age-11 variance, we controlled for that portion of the variance in the biomarkers of age that was collinear with maturation. At the same time, there is little question that age 11 Moray House Test scores’ correlations with score at age 70 largely represent the continuity of the stable intelligence trait across the lifespan. Thus, these points should not detract from the rare ability of the present study to examine correlations between general mental ability and a battery of speed measures in old age, and to take into account how much these were affected by childhood intelligence.

The additional variance—beyond childhood cognitive ability—that processing speed measures added to general cognitive outcomes was not large in absolute terms. But it was not insubstantial. With the Moray House Test the contributions ranged from less than 1% to 3.5%, and with the $g$ score the contributions ranged from 1.5% to 8.6%. In relative terms, these are important. For example, $APOE$ genotype, which is much studied with respect to cognitive aging (Luciano et al., in press), contributes approximately 1%, as does smoking (Whalley, Fox, Starr, & Deary, 2005). Therefore, recognising that explaining individual differences in cognitive aging is a multivariate problem, we should not expect individual contributions to be large.

How should one choose among the processing speed measures with regard to utility as possible biomarkers of cognitive age? The absolute size of the association after adjustment for childhood cognitive ability is one possibility. The WAIS-III processing
speed subtests performed best in this regard. However, their stimulus-response characteristics were relatively complex, more like the complex outcome tests than either reaction time or inspection time tests. In fact, this might be why they were more correlated. Another possible way of choosing the best biomarker is theoretical tractability, the possibility that some of the processing speed tests might be better understood in neural terms than others.

The original theoretical interpretation of inspection time was that individual differences may be relatively—compared with the other measures of processing speed—direct indicators of individual differences in the rate of evidence accumulation through the senses from stimuli (Vickers, Nettelbeck, and Willson, 1972; Vickers and Smith, 1986). That is, people with better inspection times were thought of as being able to make better representations of incoming stimuli in a given time. Thus, especially given a stimulus with a short duration, the person with better inspection time would have a better representation of that stimulus on which to make subsequent decisions. Further theoretical development of inspection time, specifically, has been sparse. However, the rate of accumulation of evidence in tasks that ask participants to make a forced choice between two responses has been characterised as the parameter called ‘drift rate’ in Ratcliff’s (2008) diffusion model, which has considerable theoretical development and empirical evidence. That is, part of Ratcliff’s model of how people make decisions addresses itself to how efficiently a person processes and accumulates information from the stimulus, so that a decision can be made. Ratcliff (2008) suggests that aging tends to affect conservativeness of response in two-alternative decision tasks rather than accumulation of evidence from the stimuli (drift rate). This would imply that it is the later stages of decision making in his tasks—rather than the earlier stages which would appear to be more relevant to inspection time—which are especially affected by age. Ratcliff’s decision-making tasks are relatively complex
compared with inspection time, which was conceived as information processing exclusive of response speed. Therefore, a direct comparison of the two models—inspection time and Ratcliff’s model of decision making—is not straightforward and would be better conducted by employing both tasks in the same study.

A further possible way of choosing among the processing speed tests is relative independence from early life cognitive ability. On both these latter counts inspection time performed relatively well. It had the least dependence on age 11 cognitive ability. And the neural correlates of inspection time performance are becoming understood. In functional MRI studies in both younger and older subjects the neural networks of inspection time performance were consistent (Deary, Simonotto, et al., 2004; Waiter et al., 2008). Among older people, those who had more successful cognitive aging from age 11 to their mid-60s retained the same pattern of cortical activation and deactivation as younger adults when performing the inspection time task. These neural activation-deactivation networks were not apparent in those with less successful cognitive aging; there was less overall activation and deactivation (Waiter et al., 2008). Added to these strengths, inspection time apparently does not show the Flynn effect (Nettelbeck & Wilson, 2004) and change in inspection time has been reported to predict some change in more general cognitive abilities (Gregory et al., 2008).

With regard to inspection time as a possible biomarker of age, the present study and that of Gregory et al. (2008) provide different, complementary and congruent data. In the study by Gregory et al. there was no measure of early life intelligence, but they did find that short term inspection time changes within old age were able to predict reasoning, perceptual speed and working memory after adjusting for baseline inspection time. They did not compare inspection time with other measures of processing speed. Instead, they compared inspection time with other, more general potential biomarkers: grip strength,
systolic blood pressure, and visual acuity (see Gregory et al., 2008, Table 3). Two of these might be unpromising as biomarkers. Systolic blood pressure does increase with age, but little such change would be expected within the short time period that Gregory et al.’s (2008) study covered, and many older people’s blood pressure is treated with medication. Moreover, a slight continued rise in systolic blood pressure within the age range studied by Gregory et al. (2008) has been found in healthy older people, while those who experience a drop in systolic blood pressure within the same period are more likely to have developed disease (Starr, Inch, Cross, MacLennan, & Deary, 1998). Therefore, a rise in systolic blood pressure at this period in old age may not be considered a biomarker of deleterious age effects. Visual acuity, similarly, would not be expected to change much over such a short period. The present study additionally finds that inspection time’s association with cognitive ability within old age is mostly independent of prior cognitive ability, from childhood. Added to the suggestions of Gregory et al. (2008) that inspection time might be practically useful, the present study would also suggest that 4-choice reaction time be explored further as a practical biomarker of aging, as we now discuss.

Reaction time, especially 4-choice reaction time mean, also fared quite well. Although its individual differences in old age were more influenced by age 11 cognitive ability than were those of inspection time, its independent effects had relatively large effect sizes. Moreover, it is widely classed as an ‘elementary cognitive task’ (ECT) by psychologists interested in the foundations of cognitive differences (Deary, 2000, Chapter 6). The term ECT is somewhat disingenuous, because the neural foundations of reaction time performance are not yet well understood, though brain white matter integrity might be one biological foundation of its individual differences (Deary et al., 2006). However, it may be accepted that it is simpler than the WAIS-III processing speed tests. For convenience of testing with older people and others, it is quick and straightforward to
perform; it takes less time than inspection time, and does not appear to be affected by differences in ambient light levels. Also, reaction time and change in reaction time from baseline have validity as predictors—perhaps biomarkers?—of mortality (Shipley et al., 2006, 2007). This predictive power beyond cognitive outcomes is a special strength for a putative biomarker.

In conclusion, we should be wary of accepting that processing speed measures are relatively pure assessments of the information processing ‘state’ of the organism: how it has deviated from a cognitive baseline. Their individual differences are often, to a large extent, legacies of individual differences in childhood general cognitive ability. Still, they do appear to have some incremental value as statistical biomarkers. Among the measures, inspection time may show the best promise as a biomarker, though choice reaction time should be considered too.

Author Note

Ian J. Deary, Wendy Johnson, John M. Starr, Centre for Cognitive Ageing and Cognitive Epidemiology, Departments of Psychology (IJD, WJ) and Geriatric Medicine (JMS), University of Edinburgh. Wendy Johnson holds an RCUK Fellowship.

We thank the following for collecting and collating data: Alan Gow, Michelle Taylor, Janie Corley, Caroline Brett, Paula Davies, and Caroline Cameron. The study was supported by a programme grant from Research Into Ageing. The study is supported by Help The Aged as part of the Disconnected Mind project. The work was undertaken by The University of Edinburgh Centre for Cognitive Ageing and Cognitive Epidemiology, part of the cross council Lifelong Health and Wellbeing Initiative. Funding from the BBSRC, EPSRC, ESRC and MRC is gratefully acknowledged.
We are grateful to Jennifer Scotland for providing unpublished data from her PhD studies on the test-retest reliability of Digit Symbol and inspection time.

Correspondence concerning this article should be addressed to Professor Ian J. Deary, Centre for Cognitive Ageing and Cognitive Epidemiology, Department of Psychology, University of Edinburgh, 7 George Square, Edinburgh EH8 9JZ, Scotland, UK. Email I.Deary@ed.ac.uk.


Deary, I. J., Gow, A. J., Taylor, M. D., Corley, J., Brett, C., Wilson, V., Campbell, H.,


Miller, R. A. (2001). Biomarkers of aging. SAGE KE. sageke.sciencemag.org/cgi/content/full/2001/1/pe2


Table 1

Descriptive statistics. Pearson Zero-Order Correlations (Lower Diagonal) Between Processing Speed Measures and Moray House Test (MHT) Scores at Age 11 and 70, and g. Partial Correlations (Upper Diagonal) Between Processing Speed Measures and MHT and g Scores, Controlling for MHT Scores at Age 11. Coefficients in Parenthesis are Corrected for Unreliability: those in the Second Row Within Each Cell are Corrected Based on a Three-year Retest of Over 400 of the Lothian Birth Cohort 1936; and Those in the Third Row are Corrected Based on Period-free Estimates (see Results Section)

<table>
<thead>
<tr>
<th>Moray House Test age 11</th>
<th>Moray House Test age 70</th>
<th>g</th>
<th>g'</th>
<th>Inspection time</th>
<th>Simple reaction time mean</th>
<th>4-choice reaction time mean</th>
<th>WAIS-III symbol search</th>
<th>WAIS-III digit symbol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moray House Test age 70</td>
<td>.692</td>
<td>-</td>
<td>.180</td>
<td>-.116</td>
<td>-.177</td>
<td>.264</td>
<td>.215</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.265)</td>
<td>(.144)</td>
<td>(.200)</td>
<td>(.341)</td>
<td>(.225)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.223)</td>
<td>(.130)</td>
<td>(.182)</td>
<td>N/A</td>
<td>(.209)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g</td>
<td>.618</td>
<td>.669</td>
<td>-</td>
<td>.321</td>
<td>-.173</td>
<td>-.380</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(.472)</td>
<td>(.239)</td>
<td>(.469)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.397)</td>
<td>(.214)</td>
<td>(.427)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g'</td>
<td>.564</td>
<td>.663</td>
<td></td>
<td>.265</td>
<td>-.126</td>
<td>-.259</td>
<td>.410</td>
<td>.308</td>
</tr>
<tr>
<td></td>
<td>(.390)</td>
<td>(.223)</td>
<td>(.343)</td>
<td>(.595)</td>
<td>(.369)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.328)</td>
<td>(.201)</td>
<td>(.312)</td>
<td>N/A</td>
<td>(.341)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Inspection time</td>
<td>Simple reaction time mean</td>
<td>4-choice reaction time mean</td>
<td>WAIS-III symbol search</td>
<td>WAIS-III digit symbol</td>
<td>Mean</td>
<td>(SD)</td>
<td>N</td>
</tr>
<tr>
<td>---------------------------</td>
<td>-----------------</td>
<td>---------------------------</td>
<td>---------------------------</td>
<td>-----------------------</td>
<td>-----------------------</td>
<td>------</td>
<td>-----</td>
<td>----</td>
</tr>
<tr>
<td></td>
<td>.117 (.163)</td>
<td>-.251 (-.353)</td>
<td>-.274 (-.328)</td>
<td>.462 (.605)</td>
<td>.423 (.489)</td>
<td>49.3</td>
<td>11.7</td>
<td>942</td>
</tr>
<tr>
<td></td>
<td>.210 (.292)</td>
<td>-.255 (-.359)</td>
<td>-.312 (-.373)</td>
<td>.488 (.639)</td>
<td>.434 (.501)</td>
<td>64.9</td>
<td>8.0</td>
<td>988</td>
</tr>
<tr>
<td></td>
<td>.323 (.449)</td>
<td>-.287 (-.404)</td>
<td>-.457 (-.547)</td>
<td>- (.747)</td>
<td>- (.556)</td>
<td>0.066</td>
<td>0.96</td>
<td>986</td>
</tr>
<tr>
<td></td>
<td>.289 (.402)</td>
<td>-.272 (-.383)</td>
<td>-.382 (-.457)</td>
<td>.571 (.747)</td>
<td>.481 (.524)</td>
<td>.027</td>
<td>9.8</td>
<td>996</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-.172 (-.323)</td>
<td>-.328 (-.502)</td>
<td>.294 (-.524)</td>
<td>.264 (-.524)</td>
<td>113.1</td>
<td>(52.4)</td>
<td>995</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>.456 (-.420)</td>
<td>-.235 (-.485)</td>
<td>-.265 (-.485)</td>
<td>273.3</td>
<td>(79.9)</td>
<td>994</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-.431 (-.593)</td>
<td>.593 (-.576)</td>
<td>635.8</td>
<td>(6.2)</td>
<td>993</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>25.1</td>
<td>(12.4)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>57.6</td>
<td>(12.4)</td>
<td></td>
</tr>
</tbody>
</table>

*Note.* aTotal number of items correct. bStandard score. cMilliseconds. All correlations are significant at \( p \leq .001 \). N/A = not available.
Table 2

Linear Regression Models with Processing Speed Measures as Independent Variables and Moray House Test (MHT) and g and g' Scores as Dependent Variables, With and Without MHT Scores at Age 11 as Independent Variables

<table>
<thead>
<tr>
<th></th>
<th>( R^2 (t) ) with MHT age 70</th>
<th>% attenuation of ( R^2 ) and ( t )</th>
<th>( R^2 (t) ) with g, adjusted for MHT age 11</th>
<th>% attenuation of ( R^2 ) and ( t )</th>
<th>( R^2 (t) ) with g', adjusted for MHT age 11</th>
<th>% attenuation of ( R^2 ) and ( t )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inspection time</td>
<td>.044</td>
<td>.015</td>
<td>.103</td>
<td>.060</td>
<td>.083</td>
<td>.071</td>
</tr>
<tr>
<td></td>
<td>(6.74)</td>
<td>(5.22)</td>
<td>(10.70)</td>
<td>(10.00)</td>
<td>(9.67)</td>
<td>(8.58)</td>
</tr>
<tr>
<td>Simple reaction time mean</td>
<td>.064</td>
<td>.007</td>
<td>.082</td>
<td>.017</td>
<td>.074</td>
<td>.015</td>
</tr>
<tr>
<td></td>
<td>(8.30)</td>
<td>(3.58)</td>
<td>(9.39)</td>
<td>(5.16)</td>
<td>(9.25)</td>
<td>(3.89)</td>
</tr>
<tr>
<td>4-choice reaction time mean</td>
<td>.098</td>
<td>.014</td>
<td>.208</td>
<td>.085</td>
<td>.146</td>
<td>.063</td>
</tr>
<tr>
<td></td>
<td>(10.32)</td>
<td>(5.07)</td>
<td>(16.12)</td>
<td>(12.17)</td>
<td>(13.51)</td>
<td>(8.21)</td>
</tr>
<tr>
<td>WAIS-III symbol search</td>
<td>.238</td>
<td>.035</td>
<td>.208</td>
<td>.085</td>
<td>.326</td>
<td>.135</td>
</tr>
<tr>
<td>WAIS-III digit symbol</td>
<td>.188</td>
<td>.023</td>
<td>.188</td>
<td>.023</td>
<td>.231</td>
<td>.078</td>
</tr>
<tr>
<td></td>
<td>(15.08)</td>
<td>(6.62)</td>
<td>(15.08)</td>
<td>(6.62)</td>
<td>(17.94)</td>
<td>(9.25)</td>
</tr>
</tbody>
</table>
Table 3

Tests of Differences (t Statistics) of Partial Correlations Between Processing Speed Markers and Moray House Test (MHT; above diagonal) and $g$ and $g'$ (Below Diagonal) at Age 70 After Controlling for MHT score at age 11

<table>
<thead>
<tr>
<th>Moray House Test age 70</th>
<th>Inspection time</th>
<th>Simple reaction time mean</th>
<th>4-choice reaction time mean</th>
<th>WAIS-III symbol search</th>
<th>WAIS-III digit symbol</th>
</tr>
</thead>
<tbody>
<tr>
<td>$g$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inspection time</td>
<td>-</td>
<td>1.32</td>
<td>.29</td>
<td>3.35</td>
<td>2.29</td>
</tr>
<tr>
<td>Simple reaction time mean</td>
<td>3.47</td>
<td>-</td>
<td>1.28</td>
<td>4.42</td>
<td>3.56</td>
</tr>
<tr>
<td>4-choice reaction time mean</td>
<td>1.72</td>
<td>6.11</td>
<td>-</td>
<td>3.79</td>
<td>2.91</td>
</tr>
<tr>
<td>WAIS-III symbol search</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1.23</td>
</tr>
<tr>
<td>$g'$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simple reaction time mean</td>
<td>3.47</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-choice reaction time mean</td>
<td>0.11</td>
<td>6.86</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WAIS-III symbol search</td>
<td>4.36</td>
<td>7.54</td>
<td>4.99</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WAIS-III digit symbol</td>
<td>1.20</td>
<td>4.72</td>
<td>1.57</td>
<td>3.98</td>
<td></td>
</tr>
</tbody>
</table>

Note. Listwise $N = 926$. Values greater than 2.33 are significant at $p < .01$ (not adjusted for multiple testing).