NIH Public Access
Author Manuscript
Intelligence. Author manuscript; available in PMC 2012 January 1.

Published in final edited form as:
Intelligence. 2011 ; 39(4): 222–232. doi:10.1016/j.intell.2011.03.001.

What cognitive abilities are involved in trail-making performance?

Timothy A. Salthouse
Department of Psychology, University of Virginia, Charlottesville, VA 29904

Abstract

The cognitive abilities involved in the Connections (Salthouse, et al., 2000) version of the trail making test were investigated by administering the test, along with a battery of cognitive tests and tests of complex span and updating conceptualizations of working memory, to a sample of over 3,600 adults. The results indicate that this variant of the trail making test largely reflects individual differences in speed and fluid cognitive abilities, with the relative contributions of the two abilities varying according to particular measure of performance considered (e.g., difference, ratio, residual). Relations of age on trail making performance were also examined. Although strong age differences were evident in the Connections and working memory measures, with both sets of variables there was nearly complete overlap of the age differences with individual differences in speed and fluid cognitive abilities.

Keywords

neuropsychological assessment; meaning of tests; working memory; fluid ability; age differences

Trail making tests have been extensively used in neuropsychological assessment (e.g., Butler, Retzlaff & Vanderploeg, 1991; Rabin, Barr & Burton, 2005; Sellers & Nadler, 1992). Most variants of this test, which was apparently introduced in 1938 by Partington (Partington & Leiter, 1949), have at least two conditions. In condition *A* the participant is to draw lines to connect circled numbers in a numerical sequence (i.e., 1-2-3, etc.) as rapidly in possible. In condition *B* the participant is to draw lines to connect circled numbers and letters in an alternating numeric and alphabetic sequence (i.e., 1-A-2-B, etc.) as rapidly as possible.

Although trail making tests are very simple, they have been hypothesized to reflect a wide variety of cognitive processes including attention, visual search and scanning, sequencing and shifting, psychomotor speed, abstraction, flexibility, ability to execute and modify a plan of action, and ability to maintain two trains of thought simultaneously (for citations see Lezak, Howieson & Loring, 2004; Salthouse & Fristoe, 1995, and Strauss, Sherman & Spreen, 2006). A number of studies have examined the patterns of relations between trail-making performance and other variables, with recent reviews by Sanchez-Cubillo, Perianez, Adrover-Roig, Rodrigues-Sanchez, Rios-Lago, Tirapu and Barcelo (2009) and Misdraji and Gass (2010). However, most of the studies had limitations such as weak statistical procedures, a restricted set of other variables, and relatively small samples, many of which involved mixtures of different patient groups.

Correlational analyses are most informative when the predictor variables are latent constructs rather than manifest (observed) variables, and when several predictors are examined simultaneously. Because they focus on the systematic variance shared by different observed variables representing the relevant ability, latent constructs minimize task-specific influences and measurement error. Moreover, the availability of multiple predictors in the same analysis allows unique (or direct) and shared (or indirect) relations to be distinguished, instead of relying on an undifferentiated mixture of the two types of influences as is the case with simple correlations. The current project therefore involved a relatively large sample of participants to obtain precise estimates of the relations, and multiple measures of cognitive abilities were acquired to examine relations at the construct level instead of the level of individual variables. In addition, relations of four different cognitive abilities were examined simultaneously to allow unique relations to be identified.

The target variables were derived from a variant of the trail making test known as the Connections Test. The test, which was illustrated in Figure 1 of Salthouse, Toth, Daniels, Parks, Pak, Wolbrette and Hocking (2000), consists of a set of pages containing 49 circles, with either a number or a letter with in each circle. The task for the participant is to draw lines as quickly as possible to connect the elements in sequence, with different pages involving numeric, alphabetic, or alternating numeric and alphabetic sequences. A recent confirmatory factor analysis by Atkinson and Ryan (2008) revealed that the measures from the Connections Test had loadings on the same factors as measures from two other trail making variants (i.e., the Delis-Kaplan Trail Making Test and the Comprehensive Trail Making Test). Among the differences between the Connections Test and the standard trail making test are: (a) irrelevant influences of visual search and hand movements are minimized because successive target elements are in adjacent locations rather than scattered around the page; (b) a counterbalanced presentation order of simple (*A*) and alternating (*B*) trials eliminates a confounding in which the alternating condition always follows the simple condition; (c) both letters and numbers are presented in the simple condition instead of only numbers, which eliminates a confounding of type of material and condition; (d) multiple trials are presented in each condition to increase the precision of the performance measures and allow reliability to be assessed; and (e) administration is efficient because a limit of 20 seconds is allowed to work on each page instead of monitoring the time until all items are completed.

The cognitive abilities involved in performance on this version of the trail making test were investigated with the contextual analysis model portrayed in Figure 1. The term contextual analysis refers to analyses in which the effects on a target variable are examined in the context of effects on other variables (e.g., Salthouse, 2005; 2010a; Salthouse, Pink & Tucker-Drob, 2008). The model is portrayed with age as a key predictor variable, but it is important to recognize that the underlying logic is identical when other variables, such as severity of a neuropathological condition, presence or absence of particular genes, psychosocial status, etc., replace age in the model.

Two types of information are available from contextual analyses. One type concerns the relations of the predictor variables to the target variable. When predictor variables correspond to different cognitive abilities, the results are informative about the relative influence of each ability on the target variable, independent of any age-related influences on the predictor and target variables. That is, because relations of age are postulated to all variables, the influences of age are effectively partialled out, such that the relations among other variables correspond to what would be expected at the average age in the sample. In some cases there may be interest in whether the relations between the cognitive abilities and the target variables differ as a function of age, and this question can be examined by interaction analyses or by repeating the analyses with subsamples of participants within narrow age ranges.

The second type of information available from contextual analysis concerns the relations of age (or any individual difference variable of interest) on the target variable after taking individual differences in the predictor variables into account. The rationale is that age relations on the reference cognitive abilities are already well established, and therefore the primary interest is in the magnitude of the relation of age on the target variable after these influences have been taken into consideration. If there is little or no direct relation (i.e., independent of indirect relations through the cognitive abilities), then the results would be consistent with an interpretation that most of the age-related influences on the target variable are shared with influences on the predictor variables. In contrast, a discovery of a unique relation of age on the target variable would imply that an explanation of the effects of age on the predictor variables would not be sufficient to account for the effects of age on the target variable. Information of this type is relevant to the question of how many distinct dimensions are involved in cross-sectional age differences in cognitive functioning.

Although many different variables or abilities could be used as predictors in contextual analyses, a considerable amount of research has been conducted with the reference variables used in the current project. The variables represent four broad cognitive abilities which have been found to have statistically independent age-related influences (e.g., Salthouse, 2004; Salthouse & Ferrer-Caja, 2003). The Gf ability consists of three reasoning tests and three spatial visualization tests which have been combined because the correlation between the reasoning and spatial visualization constructs was .84 (Salthouse, 2005). These variables have good reliability (i.e., the median coefficient alpha was .84) and validity (i.e., standardized factor loadings greater than .65), with similar patterns in different samples (Salthouse, Pink & Tucker-Drob, 2008), across different age ranges (Salthouse, 2004), and when age is partialled from all variables (Salthouse, et al., 2008).

Prior research with contextual analysis methods has revealed that few significant unique relations of age on the target variables are evident when those variables are considered in the context of age-related influences on the reference cognitive abilities. This is true for a variety of memory variables (e.g., Salthouse & Siedlecki, 2007a; Salthouse, Siedlecki & Krueger, 2006), including prospective memory (Salthouse, Berish & Siedlecki, 2004), source memory (Siedlecki, Salthouse & Berish, 2005), and several measures of working memory (Salthouse, Pink & Tucker-Drob, 2008). Contextual analyses have also revealed very strong relations of fluid cognitive ability (Gf) to constructs hypothesized to represent executive functioning (e.g., Salthouse, 2005; Salthouse, Atkinson & Berish, 2003; Salthouse & Davis, 2006), and to assorted variables postulated to represent specific aspects of executive functioning (e.g., Salthouse, 2005; 2010a; Salthouse, Atkinson & Berish, 2003; Salthouse & Siedlecki, 2007b). These patterns suggest that Gf may reflect nearly the same dimension of individual differences as executive functioning (also see Decker, Hill & Dean; 2007; Obonsawin, Crawford, Page, Chalmers, Cochrane & Low, 2002), and therefore to the extent that trail making performance reflects executive functioning, a unique influence of Gf on trail making measures would be predicted in the contextual analyses.

If trail making performance is found to be significantly related to Gf, there are at least two reasons to ask whether the Gf relation is primarily attributable to relations shared with working memory (WM). First, WM has been hypothesized to be involved in trail making performance because of the need to keep track of encountered letters and numbers (e.g., Crowe, 1998; Sanchez-Cubillo, et al., 2009; Zakzanis, Mraz & Graham, 2005). And second, a considerable amount of research has documented strong correlations between Gf and WM (e.g., Ackerman, Beier & Boyle, 2004; Blair, 2006; Colom, Rebollo, Palacios, Juan-Espinosa, & Kyllonen, 2004; Salthouse, et al., 2008).

The role of WM in the relations of Gf on trail making performance was examined by adding WM as another predictor in the contextual analysis. The rationale is that if one predictor variable is primarily responsible for the variance in the target variable, then little influence of the other variable would be expected when the predictors are examined simultaneously. In other words, if individual differences in WM are responsible for the Gf relations, then unique relations of WM with trail making performance should be found when WM is considered simultaneously with Gf, and inclusion of WM in the analyses should reduce, or possibly even eliminate, the relation of Gf with trail making performance.

An important issue to be considered when analyzing trail making performance is how to represent the contrast between performance in the simple (*A*) and alternating (*B*) conditions. A number of derived indices have been proposed (e.g., Corrigan & Hinkeldey, 1987; Drane, Yspeh, Huthwaite & Klinger, 2002; Lange, Iverson, Zakrzewski, Ethel-King & Franzen, 2005; Misdraji & Gass, 2010), with the simple difference (*B*–*A*) and ratio (*B*/*A*) indices the most common. With moderately large sample sizes a potentially more informative index, because it is empirically-based, is the residual derived from regression equations relating performance in the *B* condition to performance in the A condition. That is, if the predicted *B* performance is designated as B^*, then $B^* = a + bA$ + error, and the residual is determined by subtracting the predicted *B* from the observed *B* (i.e., $B - B^*$). Although a variety of methods have been used to attempt to derive purer measures of the theoretically most interesting processes, it is important to recognize that they each involve somewhat different assumptions about the relations between the *A* and *B* variables, and thus they could have different relations with cognitive abilities and with age.

To summarize, four major questions were investigated in the current project. First, what cognitive abilities are associated with performance in the Connections version of the trail making test? Second, is the pattern of cognitive ability relations similar for performance in the simple (*A*) and alternating (*B*) versions, and for the different indices that can be used to express the contrast between them? Third, what are the relative roles of Gf and WM on performance in the simple (*A*) and alternating (*B*) conditions? And fourth, are there unique age-related influences in these tasks? The relevant data were obtained from a large sample of adults between 18 and 98 years of age who performed the Connections version of the trail making test and a battery of cognitive ability tests. Working memory influences were investigated in two subsamples from the larger sample in which data were also available on different types of WM tasks.

Methods

Participants

The total sample consisted of 3,665 adults ranging from 18 to 98 years of age, and included two partially overlapping subsamples who also performed WM tests. Participants were recruited from advertisements and referrals from other participants. Various demographic characteristics and their relations with age are presented in Table 1, where it can be noted that most participants reported themselves to be in good to excellent health, and had completed several years of college. The table also contains age-adjusted scaled scores for four variables from a commercial test battery to assess representativeness of the sample. It can be seen that the participants in this sample performed at higher levels than the nationally representative normative samples, and that there were slightly higher scaled scores on the digit symbol and logical memory variables at older ages.

NIH-PA Author Manuscript
NIH-PA Author Manuscript

Procedure

The Connections test began with brief instructions indicating that the task consisted of using a pencil to draw lines connecting circles in a specified sequence as rapidly as possible. Different conditions within the test involved different sequences for connecting the circles, with either numeric, alphabetic, or alternating numeric and alphabetic sequences. The eight separately timed pages were presented in the order N (Numbers), L (Letters), NL (Numbers-Letters), LN (Letters-Numbers), LN, NL, L, and N.

Participants were allowed to work on each test page for 20 seconds, and performance was assessed in terms of the number of correct connections minus the number of incorrect connections. Unlike some versions of trail making tests, errors were not pointed out by the examiner during the test because of a concern that differences in how quickly the examiner noted and reported the error, and how quickly the examinee responded to the error, would introduce additional variability in the primary performance measure of number of correct responses in the designated time. Furthermore, preliminary analyses of the errors revealed that they were relatively infrequent, and thus were unlikely to distort the patterns based on the number correct scores.

Higher scores correspond to better performance when performance is assessed in terms of the number of items completed in a fixed period of time. Scores in the simpler *A* condition will therefore usually be larger than those in the *B* condition, which means that the *B*–*A* difference will be negative and the B/*A* ratio will be less than 1, with values of the former farther from 0, and of the latter farther from 1, indicating larger discrepancies between performance in the simple and the alternating conditions. It is worth noting that performance in the A and B versions could be expressed in units of seconds per item, in which case the B values would typically be larger than the A values, and larger discrepancies between conditions would be evident in B–A differences greater than zero and B/A ratios greater than 1.)

The 16 cognitive tests are briefly described in the appendix, and more details are provided in other publications (e.g., Salthouse, 2005; Salthouse, Atkinson & Berish, 2003; Salthouse & Ferrer-Caja, 2003; Salthouse, et al., 2008). The variables represent the following cognitive abilities, fluid ability (Gf), verbal episodic memory, perceptual speed, and vocabulary.

A total of 830 participants performed the two complex span WM tasks illustrated in Figure 2 (see Conway, Kane, Bunting, Hambrick, Wilhelm & Engle, 2005, for more detailed descriptions). Both tasks involved the presentation of information that was to be remembered, either positions in a matrix or identities of letters, and intervening processing requiring a decision. Processing in the Operation Span (OSpan) task involved answering arithmetic problems, and processing in the Symmetry Span (SymSpan) task involved making vertical symmetry decisions (i.e., was the display symmetric along the vertical axis?). The number of to-be-remembered items increased across successive trials, and at the end of each sequence within a given trial the participant attempted to recall the presented items from a set of 12 letters (OSpan) or 16 locations (SymSpan) alternatives. Separate measures of processing (accuracy on the processing component) and storage (number of items recalled) were obtained in each span task.

A total of 1,056 participants performed two running memory tests designed to require updating of continuously changing information (see Salthouse, et al., 2008, for more detailed description of the tasks). As illustrated in Figure 3, one updating task involved sequences of letters and the other involved sequences of matrix positions. The requirement in both versions was to report the last 4 items from a sequence which varied unpredictably in length from 4 to 12 items. Performance in both tasks was assessed as the number of items recalled in the correct order from the last four items in the lists.

Results

Figure 4 portrays the means, with standard errors, of the number of items correct in the four types of trials in the Connections test as a function of age decade. It can be seen that, with the exception of the youngest age, performance was slightly faster with numbers than with letters, and was faster when the alternating sequence started with numbers than when it started with letters, despite the fact that the two sequences are identical after the first item. In all remaining analyses the N and L variables were combined for the simple (*A*) condition, and the NL and LN variables were combined for the alternating (*B*) condition.

Strong age relations were evident in both the simple (*A*) and alternating (*B*) measures as the correlations with age were −.61 for *A* and −.48 for *B*. Nonlinear age trends were also examined, but the increment in R^2 associated with the quadratic age term was only .01 for both the *A* and *B* measures. Examination of Figure 4 reveals that the difference between the mean number of correct items in the simple (*A*, solid lines) and alternating (*B*, dotted lines) conditions was smaller at older ages. This pattern is also apparent in the cross-sectional age slopes (i.e., number of items/year) which were more negative for performance in the *A* (−.28) condition than in the *B* (−.14) condition.

There was no significant age relation on the residual (*B.A*) index (r = .00), and only a small relation on the ratio (*B*/*A*) index (r = .07), but a large positive correlation of age with the *B*–*A* difference (r = .49). This latter correlation indicates that increased age was associated with a smaller absolute discrepancy (i.e., the difference was less negative, and closer to 0) between performance in the simple (*A*) and alternating (*B*) conditions.

The contextual analysis model in Figure 1 was next used to determine which cognitive abilities were involved in this task, and the extent to which the cross-sectional age differences in the simple and alternating conditions were independent of age differences in other cognitive variables.[1] The structural equation models were investigated with the AMOS (Arbuckle, 2007) program including the full-information maximum likelihood estimation algorithm to deal with missing data by using all of the available information in the estimates. Standardized factor loadings and factor correlations for the complete sample are contained in Table 2, and means (and standard errors) of the composite cognitive ability constructs used in the contextual analyses are portrayed as a function of decade in Figure 5. It should be noted that the age trends in Figure 5 are similar to those obtained from nationally representative samples (cf., Salthouse, 2009; Figures 1.6 to 1.8 in Salthouse, 2010b).

The results of the contextual analyses in the total sample are presented in the top of Table 3. Entries in the column labeled “total age” are simple correlations, and entries in the other columns are standardized regression coefficients derived from a model like that portrayed in Figure 1. Note that performance in the simple (*A*) condition was strongly related to speed,

[1]It is important to note that the contextual analysis results are not an artifact of the use of statistical procedures to control the variation in the reference cognitive abilities because a similar reduction of the age relations was evident when the analyses were conducted with subsamples within a narrow range of scores on the reference abilities, thereby relying on matching instead of statistical adjustment. To illustrate, the correlation in the total sample between age and performance in the simple (A) condition was −.61, and after statistical control of the variation in both the speed and Gf composite scores it was −.12. In the sample of participants within the middle 20% of the distribution on both the speed and Gf composites (N = 135), the age correlation was −.19. Although the two adjusted correlations were quite comparable, this degree of matching required elimination of over 90% of the data, and thus it is much less efficient than statistical control which is based on all of the data. (See Salthouse, 2010b, for another example of the advantage of statistical control rather than matching).

and was also related to Gf. In addition, there were unique relations of Gf on performance in the alternating (*B*) condition after controlling influences through the simple condition. The entries in the bottom of Table 3 indicate that the patterns were also similar in separate analyses conducted on data from participants in each age decade. The only notable differences across age groups were more negative total age relations for adults over 60 years of age.

The next set of analyses was conducted with composite target variables formed by averaging the z-scores for the variables in the simple (*A*) and alternating (*B*) conditions, and then computing the difference, ratio, and residuals from the composite scores. Contextual analysis results on these composite variables are reported in the top of Table 4.

It can be seen that there were strong relations of speed and a moderate relation of Gf on the measure of performance in the simple (*A*) condition, and a slightly weaker speed relation but a stronger Gf relation on the measure of performance in the alternating (*B*) condition. Of particular interest is the absence of a Gf relation on the difference (*B*–*A*) measure, which was primarily related to speed. Higher speed was associated with a smaller *B*/*A* ratio (reflecting a greater discrepancy between performance in the alternating and simple conditions among people with high levels of speed), and higher Gf was associated with a larger *B*/*A* ratio (reflecting a smaller discrepancy between performance in the *B* and *A* conditions for people with high levels of Gf). A similar pattern was evident in the analysis with the residuals representing deviations of *B* from the prediction based on *A*, as the deviations were smaller (more negative) with higher speed, and were larger (more positive) with higher Gf.

The results in the top of Table 4 indicate that inferences about relations of cognitive abilities on Connections performance will vary depending on the particular measure of performance examined. For example, a moderate to large relation of speed was apparent in all measures, but there was no relation of Gf with the difference score (*B*–*A*) measure. If Gf is assumed to reflect nearly the same dimension of individual differences as executive functioning, these results imply that the difference score eliminates the influence of executive functioning in trail making performance as assessed in the Connections Test.

Relations of sex, self-rated health, and amount of education were also examined with each of the measures of Connections performance. Although some correlations were significantly different from 0 (e.g., .20 between health and performance in the simple (*A*) condition), none of the relations was independent of individual differences in the reference abilities. That is, when these variables replaced age in the model in Figure 1, their standardized regression coefficients on the *A* and *B* target variables ranged from −.02 to .03, and none was significantly different from 0.

All analyses except that with the difference score measure revealed a moderate to large relation of Connections performance with Gf, in addition to a unique relation of Gf on performance in the alternating (*B*) condition after control of the variance in performance in the simple (*A*) condition (cf. Table 3). Because some of the participants also performed WM tasks, relations on the Connections scores could be examined with Gf and WM as simultaneous predictors. As noted above, the rationale is that when both predictors are considered simultaneously, most of the predictive relations can be inferred to overlap if only one of the predictors is found to have unique relations. For example, if most of the Gf relations on performance are attributable to aspects of WM, then when both are considered simultaneously only WM might be expected to have significant relations with the measures of trail making performance.

The (storage) recall and (processing) error measures in each complex span task were converted to z-scores, and the means and standard errors plotted by decade in Figure 6. It can be seen that performance was significantly lower with increased age in each measure. The age correlations were −.35 for OSpan storage accuracy, .24 for OSpan processing errors, −.56 for SymSpan storage accuracy, and .22 for SymSpan processing errors.

The initial analysis with the complex span WM construct used it as the target variable in a contextual analysis with the reference cognitive abilities as predictors. A latent complex span construct was created with the two processing and two storage variables as indicators, and loadings for this construct, which were all significantly ($p<.01$) different from zero, are reported in the bottom of Table 2. The contextual analysis results in which the WM construct was predicted from the reference cognitive abilities are presented in the bottom of Table 4. It can be seen that the largest relation with the complex span WM construct was with Gf ability. The absence of a unique age relation indicates that when the age variation in the reference abilities was controlled, there were no significant relations of age on the WM measures.

Next the contextual analysis was repeated on the two Connections measures after adding the complex span WM construct as another predictor. The results of these analyses, presented in the middle panel of Table 5, indicate that there was no significant relation of the complex span WM construct on performance in this version of the trail making test when it was examined in the context of other abilities. Furthermore, the pattern of reference cognitive ability relations on the Connections trail making variables was very similar to that without inclusion of the WM predictor (i.e., compare the values in the top two rows with those in the fifth and sixth rows).

It is possible that relations of the Gf construct were stronger than those with the WM construct because the Gf construct was based on six variables instead of only two, such that its scope may have been broader than the WM construct. The contextual analysis was therefore repeated with a narrower Gf construct based on only the Matrix Reasoning and Letter Sets variables. As can be seen in the third and fourth, and seventh and eighth, rows of Table 5, the results in this analysis were very similar to those from the analysis with the broad Gf construct, and to the analysis without the WM construct.

Complex span tasks are not the only way to assess WM, and they may not be the most sensitive method of assessing sequencing and updating processes which may be involved in the Connections task. Because it is possible that somewhat different results would be found with an updating conceptualization of WM, similar analyses were carried out with the updating WM measures. As with the complex span measures, the number correct scores in the letters and positions versions of the running memory tasks were converted to z-scores, and the means and standard errors of these updating z-scores plotted in Figure 7 as a function of age. It is apparent in the figure that there were nearly linear age trends in the updating tasks, with age correlations of −.27 for running memory letters and −.39 for running memory positions.

Contextual analysis results with the updating WM construct as the target variable are reported in the bottom of Table 4. As with the complex span WM construct, there was a strong Gf relation and no unique relation of age on WM, suggesting that this WM construct also largely reflects the same dimension of age-related individual differences as the Gf construct. The bottom of Table 5 contains the results of the contextual analyses of trail making performance with the addition of the updating WM construct. Notice that there was little relation of the updating WM construct on the Connections trail making measures beyond Gf, and that the pattern of reference construct relations was similar with and without

WM. Furthermore, this was also true in the analysis with a narrower Gf construct defined by only two variables.

A final analysis involved a combined WM construct based on the four complex span variables in addition to the two updating variables for the 234 participants with data on all six variables. The results in the bottom of Table 4 indicate that, as with the separate WM constructs, the combined WM construct was strongly related to Gf, and the results in Table 5 indicate that inclusion of combined WM construct had little effect on the coefficients for other abilities, and that the broader WM construct was not significantly related to Connections trail making performance when it was considered simultaneously with the other cognitive ability predictors.

Discussion

The results of this project provide an unambiguous answer to the question of what cognitive abilities are involved in the Connections version of the trail making test. That is, the discovery of large relations of speed and Gf abilities in both conditions of the task implies that these abilities contribute to successful trail making performance. The speed influence likely occurs because of the requirement for rapid responding in both conditions. The Gf relation may be due to the cognitive demands of maintaining the current sequence position while searching for next element in the sequence, which are required in both the simple (*A*) and alternating (*B*) versions of the task. Furthermore, the existence of a unique Gf relation in the alternating version suggests that there are additional requirements related to fluid abilities when there is a need to alternate between two sequences.

Trail making tests such as the Connections Test are very sensitive to individual differences, and particularly to age-related differences. Surprisingly, however, the age relations in the current project were greater in the simple (*A*) condition than in the more complex alternating (*B*) condition. This may be because multiple processes are involved in both conditions of the test, but age-related individual differences in processes specific to the more complex version were small, and highly variable, relative to individual differences in processes common to both versions.

Another important finding in the study is that the nature of the contrast between performance in the simple (*A*) and alternating (*B*) conditions affects the relations with other cognitive abilities and with age. For example, the discovery that the general pattern of ability relations was similar in the simple (*A*) and alternating (*B*) conditions suggests that there may be little advantage of using one versus the other as a single measure of trail making performance. The availability of both measures obviously yields greater information than any one measure, but both measures appear to assess a mixture of speed and Gf abilities. Furthermore, the contextual analysis results indicate that the simple difference (B–A) primarily reflects speed, and thus it may not be sensitive to other influences of interest in the test.

Contrasts expressed as ratios (*B*/*A*) or residuals (*B.A*) were found to have similar patterns. In both cases there were weak direct age relations, but significant Gf and speed relations. The ratio measure does not completely eliminate the influence of speed, but it does so to a much greater extent than the simple difference, and for this reason it may be the preferable derived index in trail making tasks when sample sizes are too small to permit meaningful regression analyses to derive residuals.

Both WM conceptualizations had strong correlations with the Connections variables, but because they are highly correlated with Gf, they need to be examined simultaneously with Gf to determine whether they have unique, or incremental, prediction. That is, when two variables are both related to a criterion, the most relevant question is not necessarily their total (i.e., unique plus shared) relations, but rather their unique contributions.

The analyses with WM and Gf as simultaneous predictors revealed no unique prediction of WM, and this was true even when the Gf construct was defined narrowly based on only two variables and the WM construct was defined broadly with variables from both the complex span and updating conceptualizations of WM. There was also very little shift in the magnitude of the Gf influence on Connections performance when the WM construct was included in the model. These results are consistent with an interpretation that WM, at least as operationalized with these particular complex span and updating tasks, may reflect nearly the same dimension of individual differences as Gf, but from a narrower perspective and consequently with more limited predictive power (also see Salthouse, et al., 2008, for a similar interpretation). Many of the prior studies in which relations of WM were investigated have not included other predictor variables in the analyses, and therefore any variance that WM shared with other predictors was necessarily attributed to WM. However, the availability of multiple predictors allows their relative contributions to be assessed, and in the current case, even the narrow Gf construct appears to have broader predictive power than WM. Similar findings were recently reported by Shamosh, DeYoung, Green, Reis, Johnson, Conway, Engle, Braver and Gray (2008) as there was little unique contribution of WM in the prediction of delay discounting performance when it was examined simultaneously with a measure of general intelligence.

In conclusion, the current article has both substantive and methodological contributions. The substantive results indicate that the Connections version of the trail making test primarily reflects cognitive abilities of speed and fluid intelligence (Gf), that the pattern of influences depends on the measures used to summarize trail making performance, and that there are no relations of working memory with Connections trail making performance beyond those shared with fluid ability. Among the methodological contributions are a set of procedures to investigate the meaning of a neuropsychological or cognitive test by determining its unique relations to established cognitive abilities, to estimate the magnitude of unique age-related influences on a target variable, and to determine whether additional constructs, such as working memory, represent the same dimensions of individual differences as established cognitive ability constructs.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Preparation of this manuscript was supported by a grant from the National Institute on Aging (R37 024270). The information in this manuscript has not been previously published, and the author has no conflicts of interest.

References

Ackerman PL, Beier ME, Boyle MO. Working memory and intelligence: The same or different constructs? Psychological Bulletin. 2004; 131:30–60. [PubMed: 15631550]

Arbuckle, JL. AMOS (Version 7) [Computer Program]. Chicago: SPSS; 2007.

Atkinson TM, Ryan JP. The use of variants of the trail making test in serial assessment: A construct validity study. Journal of Psychoeducational Assessment. 2008; 26:42–53.

Blair C. How similar are fluid cognition and general intelligence? A developmental neuroscience perspective on fluid cognition as an aspect of human cognitive ability. Behavioral and Brain Sciences. 2006; 29:109–160. [PubMed: 16606477]

Butler M, Retzlaff P, Vanderploeg R. Neuropsychological test usage. Professional Psychology: Research and Practice. 1991; 6:510–512.

Colom R, Rebollo I, Palacios A, Juan-Espinosa M, Kyllonen PC. Working memory is (almost) perfectly predicted by g. Intelligence. 2004; 32:277–296.

Conway ARA, Kane MJ, Bunting MF, Hambrick DZ, Wilhelm O, Engle RW. Working memory span tasks: A methodological review and user's guide. Psychonomic Bulletin & Review. 2005; 12:769–786. [PubMed: 16523997]

Corrigan JD, Hinkeldey NS. Relationships between parts A and B of the Trail Making Test. Journal of Clinical Psychology. 1987; 43:402–409. [PubMed: 3611374]

Crowe SF. The differential contribution of mental tracking, cognitive flexibility, visual search, and motor speed to performance on Parts A and B of the Trail Making Test. Journal of Clinical Psychology. 1998; 54:585–591. [PubMed: 9696108]

Decker SL, Hill SK, Dean RS. Evidence of construct similarity in executive functions and fluid reasoning abilities. International Journal of Neuroscience. 2007; 117:736–748.

Drane DL, Yuspeh RL, Huthwaite JS, Klinger LK. Demographic characteristics and normative observations for derived trail making indices. Neuropsychiatry, Neuropsychology and Behavioral Neurology. 2002; 15:39–43.

Kline, RB. Principles and practice of structural equation modeling. 2. New York: Guilford Press; 2005.

Lange RT, Iverson GL, Zakrzewski MJ, Ethel-King PE, Franzen MD. Interpreting the trail making test following traumatic brain injury: Comparison of traditional time scores and derived indices. Journal of Clinical and Experimental Neuropsychology. 2005; 27:897–906. [PubMed: 16183622]

Lezak, MD.; Howieson, DB.; Loring, DW. Neuropsychological Assessment. 4. Oxford University Press; New York, NY: 2004.

Misdraji EL, Gass CS. The Trail Making Test and its neurobehavioral components. Journal of Clinical and Experimental Neuropsychology. 2010; 32:159–163. [PubMed: 19459077]

Obonsawin MC, Crawford JR, Page J, Chalmers P, Cochrane R, Low G. Performance on tests of frontal function reflect general intellectual ability. Neuropsychologia. 2002; 40:970–977. [PubMed: 11900749]

Parkington JE, Leiter RG. Partington's Pathway Test. The Psychological Service Center Bulletin. 1949; 1:9–20.

Rabin LA, Barr WH, Burton LA. Assessment practices of clinical neuropsychologists in the United States and Canada: A survey of INS, NAN, and APA Division 40 members. Archives of Clinical Neuropsychology. 2005; 20:33–65. [PubMed: 15620813]

Salthouse TA. Localizing age-related individual differences in a hierarchical structure. Intelligence. 2004; 32:541–561.

Salthouse TA. Relations between cognitive abilities and measures of executive functioning. Neuropsychology. 2005; 19:532–545. [PubMed: 16060828]

Salthouse TA. Decomposing age correlations on neuropsychological and cognitive variables. Journal of the International Neuropsychological Society. 2009; 15:650–661. [PubMed: 19570312]

Salthouse TA. Is flanker-based inhibition related to age? Identifying specific influences of individual differences on neurocognitive variables. Brain and Cognition. 2010a; 73:51–61. [PubMed: 20303636]

Salthouse, TA. Major issues in cognitive aging. Oxford University Press; New York, NY: 2010b.

Salthouse TA, Atkinson TM, Berish DE. Executive functioning as a potential mediator of age-related cognitive decline in normal adults. Journal of Experimental Psychology: General. 2003; 132:566–594. [PubMed: 14640849]

Salthouse TA, Berish DE, Siedlecki KL. Construct validity and age sensitivity of prospective memory. Memory & Cognition. 2004; 32:1133–1148.

Salthouse TA, Davis HP. Organization of cognitive abilities and neuropsychological variables across the lifespan. Developmental Review. 2006; 26:31–54.

Salthouse TA, Ferrer-Caja E. What needs to be explained to account for age-related effects on multiple cognitive variables? Psychology and Aging. 2003; 18:91–110. [PubMed: 12641315]

Salthouse TA, Fristoe NM. A process analysis of adult age effects on a computer-administered trail making test. Neuropsychology. 1995; 9:518–528.

Salthouse TA, Pink JE, Tucker-Drob EM. Contextual analysis of fluid intelligence. Intelligence. 2008; 36:464–486. [PubMed: 19137074]

Salthouse TA, Siedlecki KL. An individual differences analysis of false recognition. American Journal of Psychology. 2007a; 120:429–458. [PubMed: 17892087]

Salthouse TA, Siedlecki KL. Efficiency of route selection as a function of adult age. Brain and Cognition. 2007b; 63:279–287. [PubMed: 17079064]

Salthouse TA, Siedlecki KL, Krueger LE. An individual differences analysis of memory control. Journal of Memory and Language. 2006; 55:102–125.

Salthouse TA, Toth J, Daniels K, Parks C, Pak R, Wolbrette M, Hocking K. Effects of aging on the efficiency of task switching in a variant of the Trail Making Test. Neuropsychology. 2000; 14:102–111. [PubMed: 10674802]

Sanchez-Cubillo I, Perianez JA, Adrover-Roig D, Rodriguez-Sanchez JM, Rios-Lago M, Tirapu J, Barcelo F. Construct validity of the Trail Making Test: Role of task-switching, working memory, inhibition/interference control, and visuomotor abilities. Journal of the International Neuropsychological Society. 2009; 15:438–450. [PubMed: 19402930]

Sellers AH, Nadler JD. A survey of current neuropsychological assessment procedures used for different age groups. Psychotherapy in Private Practice. 1992; 11:47–57.

Shamosh NA, DeYoung CG, Green AE, Reis DL, Johnson MR, Conway ARA, Engle RW, Braver TS, Gray JR. Individual differences in delay discounting: Relation to intelligence, working memory, and anterior prefrontal cortex. Psychological Science. 2008; 19:904–911. [PubMed: 18947356]

Siedlecki KL, Salthouse TA, Berish DE. Is there anything special about the aging of source memory? Psychology and Aging. 2005; 20:19–32. [PubMed: 15769211]

Strauss, E.; Sherman, EMS.; Spreen, O. A compendium of neuropsychological tests: Administration, norms, and commentary. 3. Oxford University Press; New York, NY: 2006.

Wechsler, D. Wechsler Adult Intelligence Scale. 3. San Antonio, TX: The Psychological Corporation; 1997a.

Wechsler, D. Wechsler Memory Scale: Third Edition. San Antonio, TX: Psychological Corporation; 1997b.

Zakzanis KK, Mraz R, Graham SJ. An fMRI study of the Trail Making Test. Neuropsychologia. 2005; 43:1878–1886. [PubMed: 16168730]

Appendix A

Description of variables

Fluid cognitive ability (Gf)

Matrix Reasoning: Determine which pattern best completes the missing cell in a matrix.

Shipley Abstraction: Determine the words or numbers that are the best continuation of a sequence.

Letter Sets: Identify which of five groups of letters is different from the others.

Spatial Relations: Determine the correspondence between a 3-D figure and alternative 2-D figures.

Paper Folding: Determine the pattern of holes that would result from a sequence of folds and a punch through the folded paper.

Form Boards: Determine which combinations of shapes are needed to fill a larger shape.

Memory

Logical Memory: Number of idea units recalled across three stories.

Word Recall: Number of words recalled across trials 1 to 4 of the same word list.

Paired Associates: Number of response terms recalled when presented with a stimulus item.

Speed

Digit Symbol: Use a code table to write the correct symbol below each digit.

Letter Comparison: Same/different comparison of pairs of letter strings.

Pattern Comparison: Same/different comparison of pairs of line patterns.

Vocabulary

WAIS Vocabulary: Provide definitions of words

WJ-R Picture Vocabulary: Name the pictured object

Antonym Vocabulary: Select the best antonym of the target word

Synonym Vocabulary: Select the best synonym of the target word

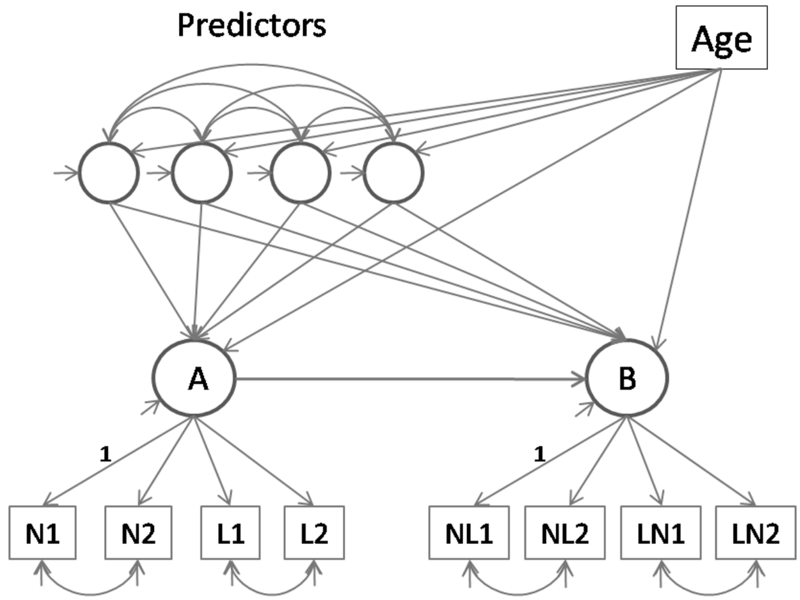

Figure 1.
Illustration of the structural equation model used to investigate the relations of reference cognitive abilities (predictors) on simple and alternating versions of trail making performance. Circles correspond to latent constructs representing the variance common to several observed (manifest) variables, which are portrayed as squares. Single-headed arrows represent directed (regression) relations, and double-headed arrows represent correlations. Arrows with numbers adjacent to them had their coefficients were fixed to 1, and arrows without a source represent variances. The cognitive variables used to define the latent predictor constructs are not portrayed in the figure, but are described in the appendix, with the factor loadings presented in Table 2.

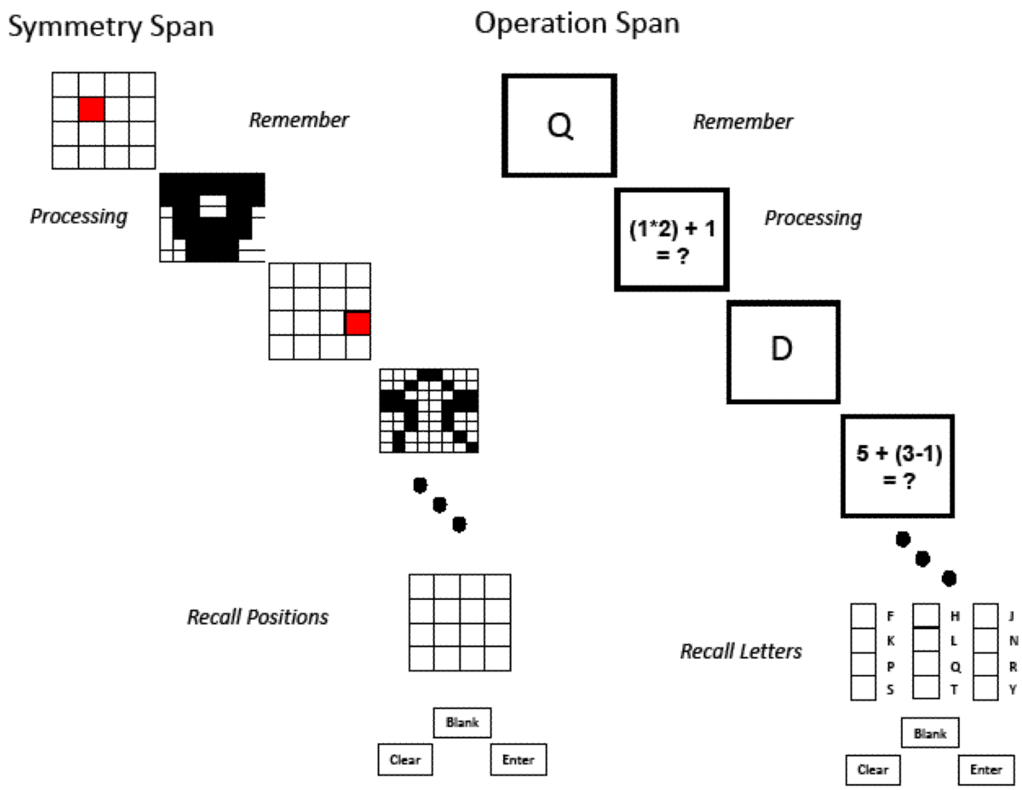

Figure 2.
Illustration of the sequence of displays in the Symmetry Span and Operation Span tasks used to assess working memory with complex span procedures.

NIH-PA Author Manuscript
NIH-PA Author Manuscript

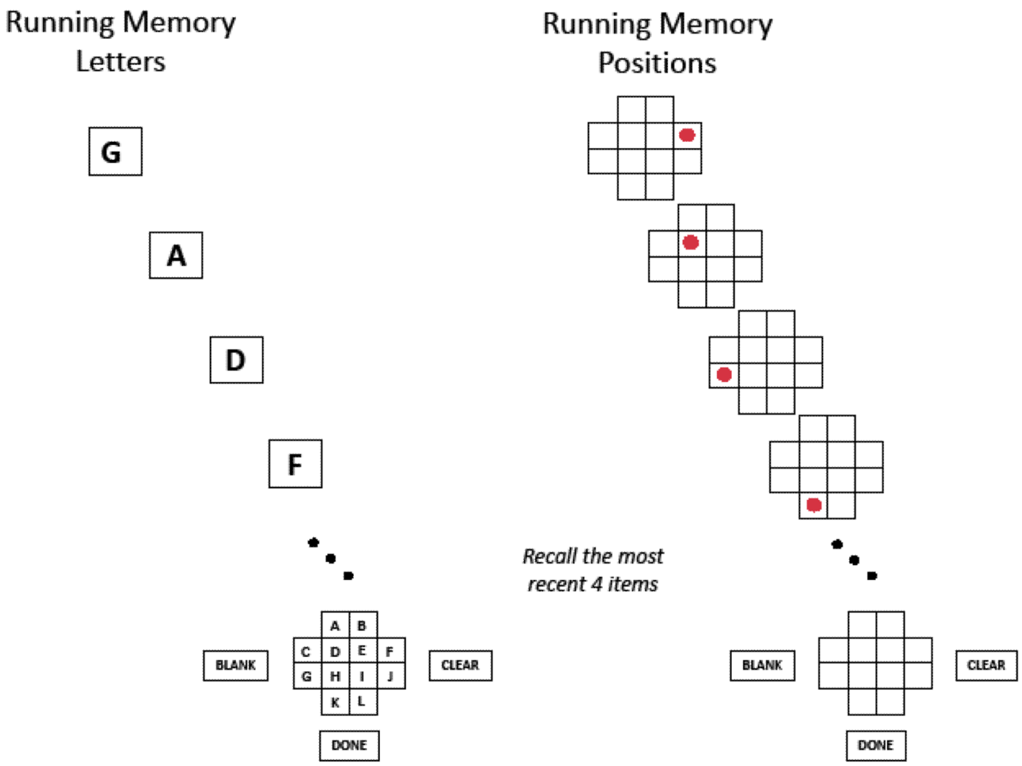

Figure 3.
Illustration of the sequence of displays in the two running memory tasks used to assess working memory with updating procedures.

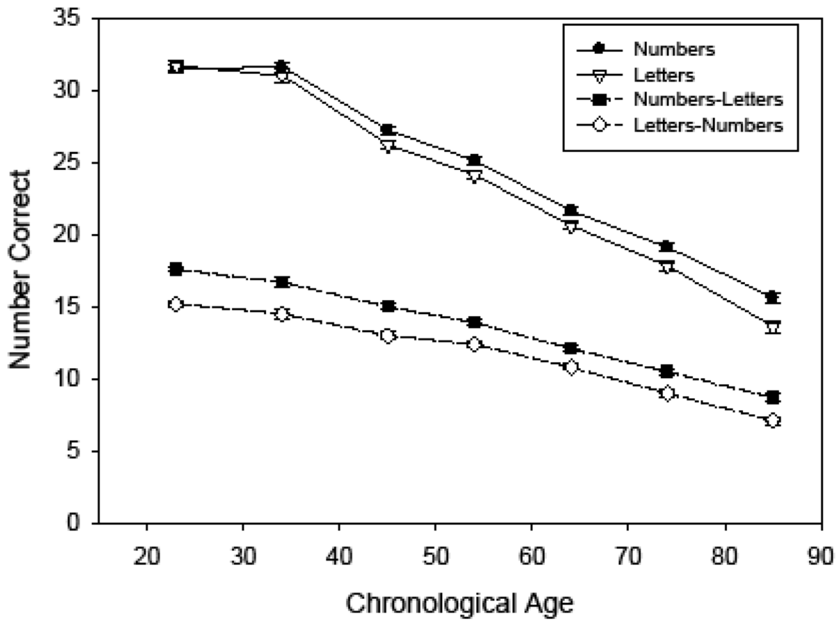

Figure 4.
Mean number of items correct when connecting elements in numerical, alphabetic, and alternating numeric and alphabetic order starting with either numbers or letters as a function of age decades. Bars around the means are standard errors. Sample sizes in each decade ranged from 220 to 830.

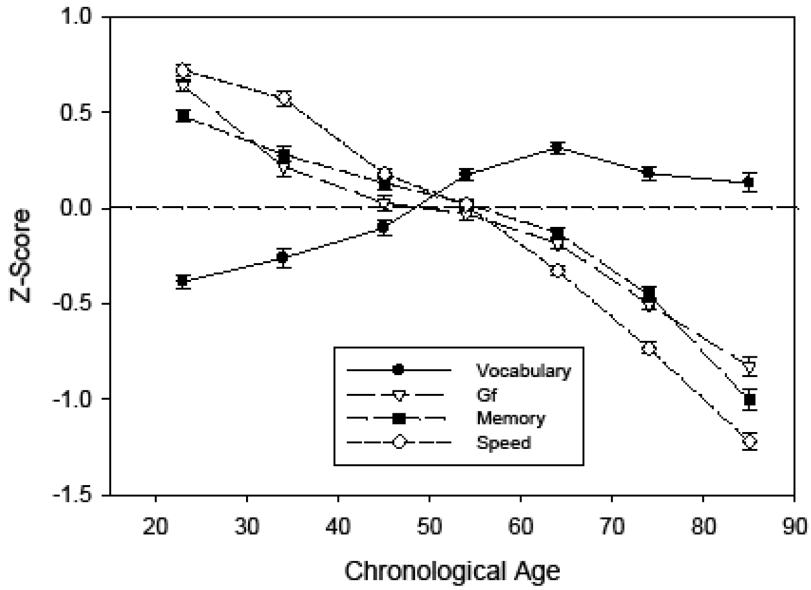

Figure 5.
Mean z-scores as a function of age decade on four reference cognitive abilities in the complete sample. Bars around the means are standard errors. Sample sizes in each decade ranged from 220 to 830.

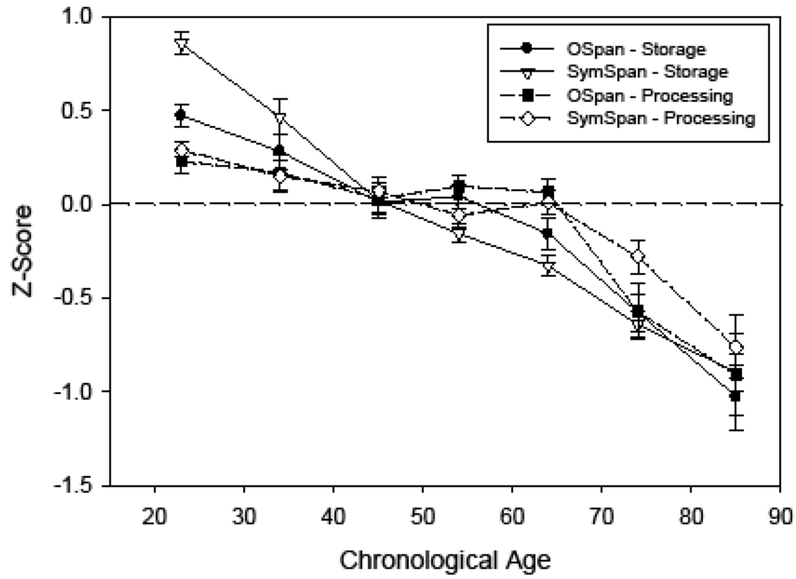

Figure 6.

Mean z-scores as a function of age decade for storage and processing measures from the complex span tasks. Bars around the means are standard errors. Sample sizes in each decade ranged from 25 to 165.

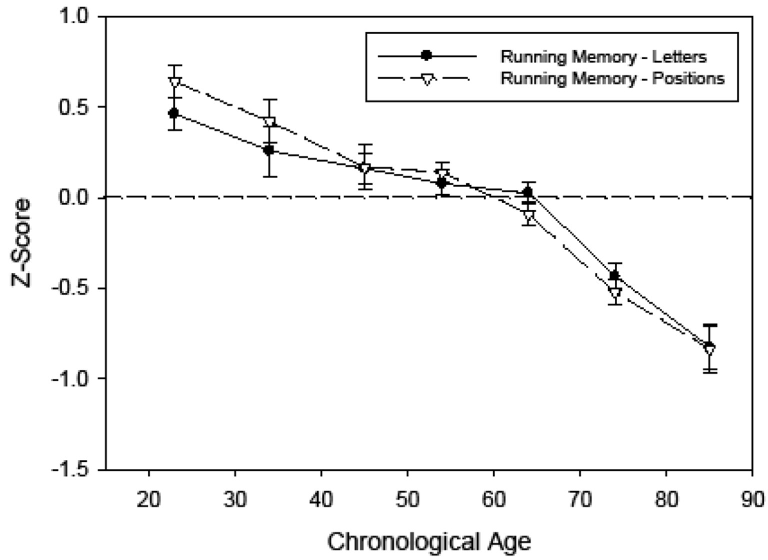

Figure 7.
Mean z-scores as a function of age decade for the running memory measures from the updating tasks. Bars around the means are standard errors. Sample sizes in each decade ranged from 23 to 114.

Table 1

Sample characteristics

Sample 1		
	Mean (SD)	**Age Correlation**
N	3,665	
Age	51.0 (18.5)	
Sex (% Female)	.65	−.03
Years of Education	15.6 (2.7)	.17*
Self-Rated Health	2.2 (0.9)	.15*
Scaled Scores		
Vocabulary	12.6 (3.0)	.04
Digit Symbol	11.3 (2.9)	.07*
Word Recall	12.2 (3.4)	.00
Logical Memory	11.8 (2.9)	.09*
Sample 2 (Complex Span WM)		
N	830	
Age	49.6 (18.1)	
Sex (% Female)	.66	.03
Years of Education	15.9 (2.6)	.24*
Self-Rated Health	1.7 (0.9)	.11*
Scaled Scores		
Vocabulary	13.2 (2.7)	.05
Digit Symbol	11.7 (2.7)	.12*
Word Recall	12.8 (3.2)	.04
Logical Memory	12.2 (2.7)	.13*
Sample 3 (Updating WM)		
N	1,056	
Age	56.1 (16.5)	
Sex (% Female)	.66	−.05
Years of Education	15.9 (2.7)	.18*
Self-Rated Health	2.1 (0.8)	.07
Scaled Scores		
Vocabulary	12.2 (2.8)	.17*
Digit Symbol	11.8 (3.2)	.10*
Word Recall	12.1 (3.1)	.05
Logical Memory	11.9 (2.9)	.12*

* $p<.01$

Note: Samples 2 and 3 are partially overlapping subsets of Sample 1. Health was rated on a scale ranging from 1 for “excellent” to 5 for “poor”. Scaled scores are age-adjusted scores from the Wechsler (1997a, 1997b) test batteries. In the nationally representative samples used to establish the norms for these tests, the age-adjusted scaled scores were designed to have means of 10 and standard deviations of 3.

Table 2

Standardized coefficients

Construct	Variable	Coefficient
Gf	Matrix Reasoning	.86
	Shipley Abstraction	.83
	Letter Sets	.69
	Spatial Relations	.78
	Paper Folding	.77
	Form Boards	.70
Memory	Word Recall	.78
	Paired Associates	.76
	Logical Memory	.70
Speed	Digit Symbol	.85
	Pattern Comparison	.77
	Letter Comparison	.78
Vocabulary	WAIS Vocabulary	.84
	Picture Vocabulary	.81
	Synonym Vocabulary	.88
	Antonym Vocabulary	.83
Connections		
Simple (*A*)	Numbers 1	.87
	Numbers 2	.90
	Letters 1	.89
	Letters 2	.89
Alternating (*B*)	Numbers-Letters 1	.78
	Numbers-Letters 2	.71
	Letters-Numbers 1	.68
	Letters-Numbers 2	.65
Complex Span		
	OSpan Storage	.78
	OSpan Processing	.51
	SymSpan Storage	.65
	SymSpan Processing	.50
Updating		
	Running Memory Letters	.72
	Running Memory Positions	.84

Construct Correlations

Gf: Memory = .65; Speed = .56; Vocabulary = .74; Complex Span = .87; Updating = .80

Memory: Speed = .50; Vocabulary = .66; Complex Span = .59; Updating = .59

Speed: Vocabulary = .48; Complex Span = .54; Updating = .54

Vocabulary: Complex Span = .62; Updating = .56

Complex Span: Updating = .81.

NIH-PA Author Manuscript
NIH-PA Author Manuscript

Table 3

Contextual analysis results based on Figure 1 for total sample and for each decade

	Age								
	Total	**Unique**	**Simple (A)**	**Gf**	**Mem**	**Speed**	**Vocab**	**CFI**	**RMSEA**
Total Sample (N = 3,665)									
Simple (*A*)	−.63*	.03	NA	.31*	−.01	.69*	−.09*	.95	.06
Alternating (*B*)	−.55*	.09*	.73*	.35*	.01	−.00	.01		
20–29 (N = 667)									
Simple (*A*)	.11*	.08	NA	.30*	−.04	.69*	−.07	.95	.05
Alternating (*B*)	−.00	−.06	.61*	.32*	−.01	.15	−.01		
30–39 (N = 341)									
Simple (*A*)	−.11	.06	NA	.17	.10	.68*	.03	.95	.05
Alternating (*B*)	−.14	.01	.66*	.29*	.14	−.06	.04		
40–49 (N = 582)									
Simple (*A*)	−.10	−.05	NA	.41*	−.03	.61*	−.13	.95	.05
Alternating (*B*)	−.07	−.03	.56*	.44*	−.03	.09	−.06		
50–59 (N = 829)									
Simple (*A*)	−.07	.01	NA	.40*	−.02	.60*	−.11	.95	.05
Alternating (*B*)	−.03	.01	.77*	.39*	.01	−.08	.01		
60–69 (N = 587)									
Simple (*A*)	−.22*	.03	NA	.43*	−.03	.67*	−.12	.94	.05
Alternating (*B*)	−.24*	−.03	.67*	.40*	−.06	−.02	−.03		
70–79 (N = 439)									
Simple (*A*)	−.23*	−.02	NA	.36*	.07	.61*	−.23*	.94	.05
Alternating (*B*)	−.24*	.00	.74*	.32*	.02	−.01	.09		
80–89 (N = 199)									
Simple (*A*)	−.23*	−.01	NA	.40*	.07	.64*	−.19	.95	.05
Alternating (*B*)	−.08	.13	.57*	.17	.16	.05	.17		

* $p<.01$

NIH-PA Author Manuscript
NIH-PA Author Manuscript

Note: Total age refers to the simple correlation of age and the variable. Unique age is the standardized regression coefficient for age when the reference cognitive abilities are included as simultaneous predictors of the target variable as portrayed in Figure 1. NA means that the parameter was not estimated because it was not applicable in the model. Entries in the Gf, Mem (Memory), Speed, and Vocab (Vocabulary) columns are standardized regression coefficients obtained from a model similar to that in Figure 1. CFI is the Comparative Fit Index and RMSEA is the Root Mean Squared Error of Approximation. CFI values greater than .92 and RMSEA values less than .10 are often considered to represent a good fit (Kline, 2005).

Table 4

Contextual analysis results on composite variables in the Connections test and on latent constructs in working memory tasks

	Age							
	Total	**Unique**	**Gf**	**Mem**	**Speed**	**Vocab**	**CFI**	**RMSEA**
			Connections					
Simple (*A*)	−.61*	.03	.30*	−.00	.67*	−.08	.93	.08
Alternating (*B*)	−.48*	.10*	.52*	−.00	.42*	−.04	.93	.08
Difference (*B*–*A*)	.49*	.06	.04	.01	−.62*	.09*	.93	.08
Ratio (*B*/*A*)	.07*	.10*	.36*	−.02	−.20*	.08	.93	.08
Residual (*B*.*A*)	−.00	.14*	.46*	−.01	−.15*	.04	.93	.08
			Working Memory					
Complex Span	−.61*	−.08	.80*	.06	.07	−.05	.92	.07
Updating	−.45*	.15	.79*	.14	.15	−.13	.93	.08
Combined	−.58*	−.02	.81*	.08	.01	−.05	.92	.06

* p<.01

Note: Connections variables are composites formed by averaging z-scores from the relevant variables. Total age refers to the simple correlation of age and the variable. Unique age is the standardized regression coefficient for age when the reference cognitive abilities are included as simultaneous predictors of the target variable as portrayed in Figure 1. Entries in the Gf, Mem (Memory), Speed, and Vocab (Vocabulary) columns are standardized regression coefficients obtained from a model similar to that in Figure 1. With some variables the coefficient for speed is negative because higher speed is associated with a smaller (less negative) difference, ratio, or residual (e.g., people with higher speed had smaller difference between performance in the alternating condition and performance in the simple condition). CFI is the Comparative Fit Index and RMSEA is the Root Mean Squared Error of Approximation.

Table 5

Contextual analysis results on Connections performance with the addition of working memory variables

	Age							
	Total	**Unique**	**Simple (*A*)**	**Gf**	**Mem**	**Speed**	**Vocab**	**WM**
No WM								
Simple (*A*)	−.63*	.03	NA	.31*	−.01	.69*	−.09*	NA
Alternating (*B*)	−.55*	.09*	.73*	.35*	.01	−.00	.01	NA
Gf = Matrix Reasoning & Letter Sets								
Simple (*A*)	−.63*	.03	NA	.34*	−.02	.67*	−.09*	NA
Alternating (*B*)	−.55*	.12*	.72*	.46*	−.03	−.04	−.03	NA
WM = Complex Span								
Simple (*A*)	−.63*	.04	NA	.24	−.01	.68*	−.08*	.10
Alternating (*B*)	−.55*	.09*	.73*	.38*	.01	−.00	.01	−.04
Gf = Matrix Reasoning & Letter Sets								
Simple (*A*)	−.63*	.04	NA	.21	−.02	.67*	−.08	.13
Alternating (*B*)	−.55*	.12*	.73*	.63*	−.03	−.04	−.05	−.18
WM = Updating								
Simple (*A*)	−.63*	.06	NA	.44*	.02	.71*	−.11*	−.16
Alternating (*B*)	−.55*	.08*	.74*	.30*	.00	−.02	.02	.06
Gf = Matrix Reasoning & Letter Sets								
Simple (*A*)	−.63*	.08	NA	.56*	.00	.68*	−.14	−.22
Alternating (*B*)	−.55*	.14*	.70*	.54*	−.02	−.03	−.04	−.07
WM = Combined Complex Span and Updating								
Simple (*A*)	−.63*	.03	NA	.34*	−.00	.69*	−.09	−.03
Alternating (*B*)	−.55*	.09*	.73*	.32*	.01	−.01	.01	.03
Gf = Matrix Reasoning & Letter Sets								
Simple (*A*)	−.63*	.04	NA	.35	−.02	.67*	−.10	−.10
Alternating (*B*)	−.55*	.13*	.72*	.66*	−.03	−.03	−.05	−.20

* $p<.01$

Note: Total age refers to the simple correlation of age and the variable. Unique age is the standardized regression coefficient for age when the reference cognitive abilities are included as simultaneous predictors of the target variable as portrayed in Figure 1. NA means that the parameter was not estimated because it was not applicable in the model. Entries in the Gf, Mem (Memory), Speed, and Vocab (Vocabulary) columns are standardized regression coefficients. All CFI values were equal to or greater than .95 and all RMSEA values were less than .07.