Neuropsychology
1997, Vol. 11, No. 3, 428–436

0894-4105/97/$3.00

Examination of Age-Related Deficits on the Wisconsin Card Sorting Test

Nathanael M. Fristoe and Timothy A. Salthouse
Georgia Institute of Technology

John L. Woodard
Emory University School of Medicine

Adult age differences in Wisconsin Card Sorting Test (WCST) measures were examined before and after statistical control of age-related differences in measures of feedback usage, working memory, and perceptual-comparison speed. The proportion of age-related variance associated with a summary measure of WCST performance was greatly reduced after controlling for measures of feedback usage, working memory, and perceptual-comparison speed. Furthermore, the age-related variance associated with the feedback-usage measure was reduced after controlling for working memory and perceptual-comparison speed measures. These results are consistent with the idea that age-related performance differences in the WCST are partially mediated by adult age differences in feedback usage and that age differences in feedback usage are mediated by age differences in working memory, which are in turn mediated by age-related reductions in processing speed, indexed by measures of perceptual-comparison speed.

This study was designed to analyze which cognitive processes mediate age-related differences in performance on a widely used neuropsychological test, the Wisconsin Card Sorting Test (WCST; Grant & Berg, 1948; Heaton, Chelune, Talley, Kay, & Curtiss, 1993). Correlational analyses were used, so that performance differences on a given task could be interpreted in terms of the relations among potential mediating variables and a specific criterion measure. When applied to the analysis of age-related performance differences on neuropsychological tasks, correlational analyses can provide valuable information regarding the contribution an assessment measure makes toward describing potentially clinically significant age-related performance differences (see Salthouse & Fristoe, 1995). This approach focuses on determining which aspects of task performance are mediated by age-related differences in presumably simpler tasks.

The WCST is a concept-identification task designed to measure abstraction and planning abilities, as well as the tendency to perseverate to a given response pattern. Originally developed by Grant and Berg (1948), the current WCST consists of four stimulus cards and 128 response cards. The task is to sort the response cards according to color, form, and number. The examiner is allowed to give only positive or negative feedback after the examinee's placement of each response card beneath one of the stimulus cards. On the basis of this information, the participant must derive the correct sorting principle. After 10 consecutive correctly sorted cards, the sorting category changes without warning from color, to form, to number, and then to each category again in the same order.

Measures of performance on the WCST have been shown to be sensitive both to brain damage (Drewe, 1974; Milner, 1963; Robinson, Heaton, Lehman, & Stilson, 1980) and to effects associated with increased age (Anderson, Damasio, Jones, & Tranel, 1991; Axelrod & Henry, 1992; Crockett, Blisker, Hurwitz, & Kozak, 1986; Daigneault, Braun, & Whitaker, 1992; Heaton et al., 1993; Libon et al., 1994; Nelson, 1976; Parkin & Walter, 1991).

The brain-damaged groups tested in several of these studies showed difficulty shifting from one category to another, and the same perseverative tendency has also been observed in older adults. For example, Axelrod and Henry (1992) tested 20 adults in each of four age groups (50-year-olds, 60-year-olds, 70-year-olds, and 80-year-olds) on the WCST and found significant age-related increases in the number of perseverative errors and the number of perseverative responses, as well as significant age-related decreases in the number of categories achieved. However, they found no significant age differences in total correct responses or in the number of nonperseverative errors. Daigneault et al. (1992) replicated the results of Axelrod and Henry (1992) in a comparison of the performance of adults age 20–35 years ($n = 70$) with those age 45–65 years ($n = 58$), finding significant age-related increases in the number of perserverative errors and perseverative responses and decreases in the number of categories achieved. Boone, Ghaffarian, Lesser, Hill-Gutierrez, and Berman (1993) observed a somewhat different pattern of age-related deficits in WCST performance. These researchers tested a sample of 91 men and women ranging in age from 45 years to 83 years and examined the WCST number of categories, total errors,

---

Nathanael M. Fristoe and Timothy A. Salthouse, School of Psychology, Georgia Institute of Technology; John L. Woodard, Department of Neurology, Emory University School of Medicine.

Portions of this research were presented at the 1996 Cognitive Aging Conference, Atlanta, Georgia, April 1996; and the 104th Annual Convention of the American Psychological Association, Toronto, Ontario, Canada, August 1996. This research was supported by National Institute on Aging Training Grant AG00175 and Grant 6826. We would like to acknowledge the valuable assistance of S. H. Rhee in data management and J. Thomas in recruiting and testing participants.

Correspondence concerning this article should be addressed to Nathanael M. Fristoe, School of Psychology, Georgia Institute of Technology, Atlanta, Georgia 30332-0170. Electronic mail may be sent via Internet to gt4360b@acme.gatech.edu.

perseverative responses, trials to first category, percentage of perseverative errors, percentage of conceptual-level responses, and failure-to-maintain-set measures. They found age-related deficits only in the number of errors and in the percentage of conceptual-level response measures.

With a sample of 259 adults between 18 and 94 years of age, Salthouse, Fristoe, and Rhee (1996) found significant age effects for every WCST measure, with the exception of number of trials to first category. An exploratory factor analysis on the measures from the WCST revealed that despite initial variations in the relations of specific measures to age, two factors accounted for a sizable proportion of variance (R^2 = .661 and .114 for Factors 1 and 2, respectively) and were moderately correlated with age (r = .36 and .29 for Factors 1 and 2, respectively). This high degree of intercorrelation among the WCST measures suggests that all WCST measures may be assessing similar aspects of performance, at least with respect to age-related influences. However, the question of which cognitive processes are most important for successful WCST performance and which processes might be differentially sensitive to age-related effects remains unanswered.

One possible source of the age-related deficits on the WCST is the failure of older adults to induce the correct sorting category on the basis of the available feedback. That is, the "right–wrong" feedback provided by the administrator of the WCST may not be remembered or may not be interpreted by the older adult in a way that facilitates the selection of the appropriate response. Effective use of feedback information should result in a *win–stay* and *lose–shift* response pattern. That is, after positive "win" feedback, one should retain or *stay* with a hypothesis, and after negative "lose" feedback, one should change or *shift* hypotheses. If older adults are not using feedback information effectively, then they should display significantly less win–stay and lose–shift behavior than young adults. The WCST failure-to-maintain-set performance measure can be viewed as a gross measure of feedback usage in that it is a record of nonoptimal win–shift behavior, but because an occurrence is noted only after a break from five or more correct sorts, it provides no information about win–shift behavior for each item. A potentially more sensitive measure of effective usage of feedback could be derived from the number of times a participant stayed with a hypothesis after positive feedback (win–stay) and shifts from a hypothesis after negative feedback (lose–shift).

Using this approach, Offenbach (1974) examined the win–stay and lose–shift behavior of older adults on a concept-identification task similar to the WCST. Inspired by the work of Levine (1963) on the use of feedback information in a discrimination learning paradigm, the task developed by Offenbach required individuals to choose between two alternatives differing along several dimensions and then to select from eight possible hypotheses the hypothesis on which they had based their previous decision. After these two choices, the experimenter provided the participant with feedback indicating whether his or her response had been correct or incorrect based on the response criterion in effect. Offenbach found that after positive feedback for a hypothesis, young adults were very likely to repeat that hypothesis on subsequent choices (proportion of responses = .884), whereas older adults were as likely to use a different hypothesis as they were to use the previous hypothesis (proportion of responses = .495). Similarly, after negative feedback, the young adults were unlikely to use the previous hypothesis (proportion of responses = .094), whereas the older adults were significantly more likely to do so (proportion of responses = .200).

Offenbach (1974) interpreted these findings as indicating that older adults had less reliance on memory for previous outcomes than young adults. In fact, using a similar paradigm, Kellogg (1983) found that young adults were more accurate than older adults in recalling their previously used hypotheses. Because effective use of feedback depends on working memory ability, in that previously presented information must be retained to correctly process new information, age-related declines in effectiveness of feedback usage might be related to age-related declines in working memory capacity.

To summarize, the results of Offenbach (1974) and Kellogg (1983) suggest that older adults often fail to use the feedback information provided in concept-identification tasks such as the WCST, and it is suggested that these failures may be related to working memory limitations. Furthermore, on the basis of the results of Salthouse (1991, 1996) and others, it is possible that many age-related working memory limitations may be mediated by decreases in speed of processing. Salthouse (1996) proposed that even on untimed tasks such as the WCST, slower processing speed may adversely affect performance. For instance, slower processing speed may result in a decrease in the amount of information that is simultaneously active, which may correspond to less working memory capacity.

This study was designed to investigate which factors might be responsible for the age-related differences in WCST performance. The primary hypothesis was that many of the age-related performance deficits observed on the WCST could be accounted for by failures of two related factors: (a) working memory deficits and (b) reduced speed of processing. If these factors are important, then controlling for measures of these hypothesized mediators with statistical procedures will reduce the magnitude of the age-related effects on the measures of WCST performance. We also predicted that age-related variation in measures of effectiveness of feedback usage would be substantially attenuated after control of working-memory measures and that the age-related variation in measures of working memory would be significantly attenuated after control of measures of speed of processing.

Method

Participants

Young and older adults were recruited through advertisements posted in a large southeastern city. All participants were compensated with $20 for their participation in a single session lasting between 1.5 and 2 hr. The young adult group consisted of 48 persons, ranging from 18 to 38 years of age, and the older adult group consisted of 49 persons, ranging from 60 to 86 years of age.

Very few participants in either group were enrolled as full-time students at the time of testing. Demographic characteristics of the research participants are reported in Table 1. Note that significant group differences were found only for self-reported indications of health satisfaction and the reported use of hypertension medications.

Design and Procedure

The WCST was administered on computers, with a program developed by Woodard (1994). In addition to this computer-administered version of the WCST, a special version was developed and administered. This hypothesis-generation WCST version required participants to indicate on what basis they intended to make their subsequent response (i.e., according to number, form, or color). This hypothesis information was used to assess the probability of a change in hypothesis after negative or positive feedback (win–stay and lose–shift behavior).

The computer-administered version of the WCST used displays of stimulus cards identical to those in the traditional version. We administered it as described by Heaton et al. (1993). The four stimulus "cards" were continuously present in the upper portion of a color monitor. For each trial, a single response card was presented in the lower half of the screen. To sort the response cards below the stimulus cards, the participant pressed a number key between 1 and 4 (these numbers corresponded to the four stimulus cards in left-to-right order). If the pairing matched the sorting category in effect, a brief, high-pitched tone was presented, and the word RIGHT appeared in the lower left corner of the screen. If the pairing was incorrect, a low-pitched tone was presented, and the word WRONG appeared in the lower left corner of the screen. The previous response remained in view while the participant made a new response. A total of 128 response cards were presented, with the examinee controlling the pace of presentation.

Table 1
Demographic Characteristics of Research Participants

Variable	Young adults	Older adults	Group difference
N	48	49	
Gender			$\chi^2(1, N = 97) = 1.09$
% women	75	65	
% men	25	35	
Age			
M	26.7	70.1	
SD	5.7	7.2	
Education			$t(95) = -1.68$
M	13.3	13.9	
SD	1.3	2.0	
Health satisfaction			$t(93) = 2.62$*
M	2.5	2.1[a]	
SD	0.8	0.7	
Health rating			$t(95) = 1.43$
M	2.1	1.9	
SD	0.8	0.7	
Health-related activity limitations			$t(93) = -.29$
M	1.7	1.7[a]	
SD	1.0	1.0	
Cardiovascular surgery			$t(93) = -2.37$
M	0.0	0.1[a]	
SD	0.0	0.3	
Hypertension medications			$t(93) = -4.44$*
M	0.04	0.4[a]	
SD	0.2	0.5	
Head injury			$t(93) = 1.67$
M	0.1	0.02[a]	
SD	0.3	0.1	
Neurological treatment			$t(92) = 1.42$
M	0.1	0.04[a]	
SD	0.3	0.2	

Note. Age is in years, education is self-reported number of years of formal education completed, and health satisfaction and health rating are self-ratings on a scale ranging from 1 (*excellent*) to 5 (*poor*). Participants indicated *yes* (1) or *no* (0) for reports of health-related activity limitations, cardiovascular surgery, hypertension medications, head injury, and neurological treatment.
[a] One or more missing observations.
*$p \leq .01$.

The computer-administered WCST with hypothesis generation was similar to the standard computer-administered WCST, with one important modification. Before each response, the participant was required to indicate on what basis he or she intended to make his or her subsequent response, according to color, form, or number (by pressing either C, F, or N on the keyboard). This information was used in the analyses to determine the proportion of trials for which participants changed their hypothesis after either negative or positive feedback.

For both WCST versions, the number of perseverative errors, the percentage of perseverative errors, the number of conceptual-level responses, the percentage of conceptual-level responses, and the number of categories achieved were selected as measures of performance. These particular measures were chosen because they are widely used (e.g., Axelrod & Henry, 1992; Ragland, Gur, Deutsch, Censits, & Gur, 1995) and are considered to be most sensitive to cognitive flexibility (e.g., Ozonoff, 1995).

Computer-administered numerical and verbal working memory tasks were included to assess the relationship between failure to use feedback information and working memory limitations. Computerized digit–digit and digit–symbol comparison tasks and several paper-and-pencil speed measures were also included, to test mediational hypotheses involving the role of processing speed in determining age-related differences in WCST performance. Measures of sensorimotor speed involved primarily a motor response output component (e.g., rapidly drawing lines in designated locations), whereas perceptual-comparison speed measures involved an additional comparison of two or more stimuli (e.g., comparison of two strings of letters with a judgment of similarity).

All participants performed the same battery of tasks in the same order. Everyone began by filling out a health questionnaire, followed by several paper-and-pencil tests, the computer-administered WCST, the computer-administered digit–digit and digit–symbol reaction time tasks, the computer-administered WCST with hypothesis generation, and the verbal and numerical working memory tests. Only the working memory tasks were administered in counterbalanced order: verbal, numerical, numerical again, then verbal. The WCST conditions were not counterbalanced because of the concern that administering the standard WCST after the hypothesis-generation WCST would bias performance in the standard condition.

The paper-and-pencil tests consisted of the boxes, letter-comparison, pattern-comparison, digit-copying, and synonym and antonym vocabulary tests used by Salthouse (1993). With the exception of the vocabulary tests, these tasks each contained an instruction page with several examples, followed by two pages of test items.

The stimuli in the boxes task consisted of three sides of an incomplete square. For each three-sided box, the participant was to draw a line forming a fourth side, to make a square. The boxes test spanned two pages, with 100 boxes to a page. There was a 30-s time limit per page, and the boxes score was the average number of items completed across the two pages.

Letter-comparison items consisted of 21 pairs of three, six, or nine letters. Participants were to write an *S* (for same) if both members of the pair were the same or a *D* (for different) if they were different. One half of the letter pairs were different because of a difference in the identity of one letter in the pair member. The letter-comparison score was the number of correct pairs minus the number of incorrect pairs completed in 30 s. Two separately timed pages of the letter-comparison items were administered.

Items in the pattern-comparison task were 30 pairs of line segments composed of either three, six, or nine segments. The research participant was to write an *S* (for same) between the two patterns if they were identical or a *D* (for different) if they were not. One half of the pairs differed because of a shift in the position of one line segment in a pair member. The score was the number of patterns correctly completed minus the number incorrectly completed in 30 s. Two separately timed pages of the pattern-comparison task were administered.

The digit-copying task consisted of 100 pairs of boxes, one on top of the other, with a digit in the upper box and a blank lower box. The participant was to copy the upper digit in the blank lower box. Two pages of the digit-copying task were administered. The digit-copying score was the number of items completed in a 30-s period.

Vocabulary questions were taken from those used by Salthouse (1993). Both antonym and synonym vocabulary tests contained 10 five-alternative multiple-choice questions. Two minutes were allowed for each test, and the score was the number of correct answers on each test.

The boxes and digit-copying tests have been hypothesized to represent the construct of sensorimotor speed because they both involve perception of the stimuli followed by a simple motor response. The letter-comparison and pattern-comparison tests are thought to represent a perceptual-comparison speed construct (Salthouse, 1994) because both tasks require not only perception of the stimuli and a simple motor response but also an additional judgment of physical identity.

The digit–symbol reaction time test was a modification of the Wechsler Digit Symbol Substitution Test (see Salthouse, 1992), designed to assess perceptual comparison speed. Eighteen practice trials were followed by a block of 90 test trials. The digit–digit test was also used by Salthouse (1992) and was designed to assess sensorimotor speed. Eighteen practice trials were followed by a block of 90 test trials. For both the digit–digit and digit–symbol tests, individual trial latencies and accuracy percentages were recorded and summarized with measures of median reaction time and percentage correct. Because accuracy rates for both the young adults and older adults were greater than 95%, only reaction times are reported.

Both the numerical and verbal working memory tasks were computer administered, as described by Salthouse and Coon (1994). Briefly, the numerical working memory task presented a simple two-term arithmetic problem (e.g., "5 + 2 = ?") in the upper portion of the computer screen. The participant indicated the correct answer by using the up and down arrow keys to position a large arrow adjacent to one of three alternative answers (e.g., "5," "7," or "9") presented beneath the problem. Consecutive sequences of two or more problems were presented, followed by a request for the participant to recall the last digit from the previously presented arithmetic problems. For the numerical working memory task to continue, participants were required to be correct on both the arithmetic and memory tasks on two of the three trials at a given length.

For the verbal working memory task, a sentence was presented in the upper portion of the computer screen (e.g., "Tom and Mary went to the party last night"), and a brief question was asked about the sentence (e.g., "Where did they go?"). Again, the participant indicated the correct answer by using the up and down arrow keys to position a large arrow beside one of three answers (e.g., "school," "shopping," or "party") presented beneath the question. After two or more sentences, the participant was asked to recall the last word from each sentence. The test was discontinued when the participant could not recall the words from two sentence sequences of the same length.

Two blocks of each working memory task were administered, each with different items. For both the verbal and numerical working memory tasks, a participant's span was determined by the longest sequence of correctly recalled stimuli, with the requirement of correct processing of the arithmetic or comprehension questions on at least two of the three trials.

Results

Because of the large number of comparisons made in these analyses, a significance criterion of $\alpha = .01$ was used. Given the numerous variables not relevant for specific sets of comparisons (i.e., demographic variables), adoption of this more conservative alpha level seemed preferable to more drastic familywise error rate corrections.

Table 2 contains means, standard deviations, and results of independent *t* tests for age group differences on the paper-and-pencil measures, the computer-administered digit–digit and digit–symbol tests, and the computer-administered verbal and numerical working memory measures. Missing values on the vocabulary measures are attributable to the first 10 participants not receiving these tests, due to a procedural error. It can be seen that the mean differences between the two groups were significantly different from zero for all measures. Age trends were as expected, with older adults scoring lower than young adults on the speed and working memory measures but showing significantly better performance than young adults on the vocabulary measures.

Composite variables were formed by averaging the sample *z* scores of different sets of measures to represent the constructs of working memory (verbal and numerical working memory span scores; $r = .52$), perceptual-comparison speed (pattern- and letter-comparison measures; $r = .65$), and sensorimotor speed (digit-copying and boxes measures; $r = .69$). For each composite, reliabilities were estimated by means of the Spearman–Brown formula, and the split-half correlations between composites were formed from the first and second administrations of each measure. These reliability estimates were .89 for working memory, .89 for perceptual-comparison speed, and .97 for sensorimotor speed.

Measures of WCST performance consisted of the number of categories achieved, the number of conceptual level responses, the percentage of conceptual-level responses, the number of perseverative errors, and the percentage of perseverative errors. Means and standard deviations for these five WCST performance measures are reported in Table 3.

To evaluate a possible Age Group × WCST Version interaction, we conducted an Age Group × WCST Version multivariate analysis of variance (MANOVA) on the data summarized in Table 3. The number of conceptual-level responses and number of perseverative errors measures were

Table 2
Means and Standard Deviations for Speed, Vocabulary, and Working Memory Measures

Variable	Young adults (n = 48) *M*	*SD*	Older adults (n = 49) *M*	*SD*	*t*
Box completion	54.7	11.4	46.9	11.5	$t(95) = 3.39$
Digit copying	55.6	9.1	46.4	9.0	$t(95) = 5.05$
Pattern comparison	17.6	3.0	12.6	3.4	$t(95) = 7.66$
Letter comparison	10.4	2.6	6.8	3.0	$t(95) = 6.21$
Antonym vocabulary	2.1	1.6	3.7	3.2	$t(85) = -2.97$
Synonym vocabulary	3.2	2.2	5.4	3.2	$t(85) = -3.68$
Digit–digit	712	182.6	881	137.6	$t(95) = -5.15$
Digit–symbol	1,330	240.7	1,806	429.8	$t(95) = -6.72$
Working memory span					
Verbal	2.5	1.3	1.8	1.3	$t(95) = 2.75$
Numerical	3.8	2.1	2.0	2.2	$t(95) = 4.06$

Note. For the vocabulary measures, $n = 42$ young adults and 45 older adults. The digit–digit and digit–symbol tests were measured in milliseconds. All t values were significant at $p \leq .01$.

not included in this MANOVA because of their redundancy with the corresponding percent measures. The main effects for both age group, $F(3, 93) = 9.97, p \leq .001$, and WCST version, $F(3, 93) = 16.14, p \leq .001$, were significant, but the Age Group × WCST Version interaction was not significant, $F(3, 93) = 2.50$, $p = .06$. Univariate F tests, reported in Table 3, revealed significant age differences for all measures. Furthermore, the proportion of age-related variance associated with each WCST measure was still significantly greater than zero after statistical control of variables representing years of education, self-rated health satisfaction, reported cardiovascular problems, reported neurological treatment, reported use of hypertension medications, and reported loss of consciousness. The only exception to this trend was the number-of-conceptual-level-responses measure from the standard WCST. For this measure, the age-related proportion of variance was no longer significant after control of either the reported neurological treatment or hypertension medications variables.

We then conducted a principal-components analysis on 6 of the 10 measures listed in Table 3. The number of conceptual-level responses and the number of perseverative errors were excluded because of their redundancy with the corresponding percent measures. The first principal component explained 74.9% of the variance in the relevant variables and was significantly correlated with age ($r = -.52$). The factor loadings on this component were negative for the WCST perseverative errors measure (a measure of poor WCST performance) and positive for the WCST measures related to good performance (i.e., number of categories and percentage conceptual-level responses). The first principal component was therefore used as the index of card-sorting performance in subsequent analyses, with a high score representing better performance. Note that this measure is an aggregate index of several aspects of performance in the two versions of the WCST administered in this study and does not directly correspond to any single measure derived from the WCST.

Variables representing effectiveness of feedback usage were derived from the participant's hypothesis choices made before each response in the hypothesis-generation WCST. For optimum performance, a hypothesis would be retained after positive feedback (win–stay) and discarded after negative feedback (lose–shift). For example, if a participant chose the *color* hypothesis and then matched a response card with two red stars to the stimulus card with one red triangle, the match might have been followed by a brief, high-pitched tone with the word RIGHT presented. The positive feedback

Table 3
Means and Standard Deviations for WCST Measures

Measure	Young adults (n = 48) *M*	*SD*	Older adults (n = 49) *M*	*SD*	$F(1, 95)$
Standard WCST					
No. categories	4.8	1.9	3.2	2.0	16.04
No. conceptual-level responses	65.5	17.5	55.2	20.3	7.14
% conceptual-level responses	61.3	18.7	47.2	20.7	12.22
No. perseverative errors	16.4	10.0	25.2	12.1	15.37
% perseverative errors	14.1	7.1	20.2	8.8	13.95
Hypothesis WCST					
No. categories	5.6	1.0	3.7	2.1	31.39
No. conceptual-level responses	70.9	11.0	60.1	18.8	11.78
% conceptual-level responses	77.1	12.6	55.2	21.7	36.71
No. perseverative errors	9.1	5.4	20.3	13.0	30.01
% perseverative errors	9.2	3.7	16.6	9.4	25.49

Note. WCST = Wisconsin Card Sorting Test. All F values were significant at $p \leq .01$.

produced by this response would encourage the participant to retain the *color* hypothesis (i.e., win–stay) rather than switching to either *form* or *number.* On the other hand, if in the preceding example, the participant had received negative feedback for the response, lose–shift behavior would dictate that a new hypothesis (e.g., *form* or *number*) be selected on the next sort. Individuals not using feedback information appropriately would display nonoptimal win–stay and lose–shift behavior.

An analysis of win–stay and lose–shift behavior is only meaningful if both young and older adult participants show consistency between reported hypotheses and card sorts (e.g., after reporting a hypothesis of *form,* the participant matches the card on the basis of form). Young adults showed a mean percentage consistency between reported hypothesis and actual response of 94% ($SD = 6.5\%$), whereas the older adults were consistent, on average, for 85% ($SD = 15.7\%$) of their responses. A t test indicated that the age group difference in hypothesis–response consistency was significant, $t(95) = 3.65$, but because the mean consistency between reported hypothesis and actual response was high for both groups, the probability of repeating a hypothesis after positive or negative feedback was calculated.

The probability of repeating a hypothesis after positive feedback (mean percentage win–stay) was significantly higher for the young adults ($M = .973$, $SD = .055$) than for the older adults ($M = .914$, $SD = .117$), $t(95) = 3.18$. After negative feedback, the probability of staying with a hypothesis (percentage lose–stay) was significantly lower for young adults ($M = .268$, $SD = .157$) than for older adults ($M = .378$, $SD = .248$), $t(95) = -2.60$. The correlation between percentage win–stay and percentage lose–shift was low and not significantly different from zero ($r = .13$), perhaps because of the small variance in the percentage win–stay measure relative to that in the percentage lose–shift measure. Because both the percentage lose–stay and percentage win–shift measures are indicative of poor use of feedback information, we formed a composite feedback-usage variable by averaging the z scores of these measures.

We conducted a series of hierarchical regression analyses using the first principal component as the criterion variable. Table 4 contains the results. Initially, only the proportion of variance accounted for by age was calculated, then the age-related variance was examined after controlling for the feedback-usage index, working memory, and speed of processing. Controlling for the feedback-usage index resulted in a 77.4% attenuation of the age-related variance (i.e., from .266 to .060), and controlling for measures of working memory resulted in a 45.5% attenuation of the age-related variance (i.e., from .266 to .145). Simultaneously controlling for both the working memory index and the feedback-usage index resulted in a greater attenuation (85.0%) of the age-related variance than when controlling for either measure in isolation. Additional analyses revealed that the overall pattern of results was similar when individual measures of WCST performance were used as the criterion variables.

Table 4

Results of Hierarchical Regression Analyses With First Principal Component From the WCST Measures as the Criterion Variable

Predictor	Age R^2	β_{age}
Age	.266	−.516
Age–feedback	.060	−.267
Age–WM	.145	−.427
Age–Dig-Dig	.161	−.461
Age–Dig-Sym	.058	−.314
Age–MSPD	.238	−.558
Age–PSPD	.045	−.294
Age–MSPD, PSPD	.057	−.337
Age–feedback, WM	.040	−.235
Age–feedback, WM, PSPD	.022	−.213

Note. WCST = Wisconsin Card Sorting Test; Feedback = feedback-usage index; WM = working memory index; Dig-Dig = digit–digit; Dig-Sym = digit–symbol; MSPD = sensorimotor speed index; PSPD = perceptual-comparison speed index. All values except those for age–feedback, WM, PSPD were significant at $p \leq .01$.

From the results presented in Table 4, it can be seen that large attenuations in age-related variance were found, with control of the variables representing the construct of perceptual-comparison speed (the digit–symbol and the perceptual-speed index). However, large attenuations in age-related variance were also found with individual and combined statistical control of the working memory index, the feedback-usage index, and the perceptual-comparison speed index (e.g., controlling for all three resulted in a 91.7% attenuation). This result is consistent with the hypothesis that age-related variation in WCST performance is partially mediated by feedback-usage and working memory components. The results reported in Table 4 also show that controlling for the sensorimotor speed indexes resulted in little additional attenuation of age-related variance in the first principal component, and if anything, the sensorimotor speed index acted as a suppressor variable.

We also conducted hierarchical regression analyses, using the feedback-usage index as the criterion variable. Table 5 contains the results of these analyses, focusing on the attenuations in the age-related variance after controlling for working memory and speed measures. Control of working memory, perceptual-comparison speed, or digit–symbol speed resulted in large attenuations of the age-related propor-

Table 5

Results of Hierarchical Regression Analyses With Feedback Composite Measure as the Criterion Variable

Predictor	Age R^2	β_{age}
Age	.156*	.395*
Age–WM	.077*	.311*
Age–MSPD	.096*	.354*
Age–PSPD	.006	.103
Age–Dig-Dig	.052	.261
Age–Dig-Sym	.019	.180
Age–MSPD, PSPD	.006	.108
Age–WM, PSPD	.003	.075

Note. WM = working memory index; Dig-Dig = digit–digit; Dig-Sym = digit–symbol; MSPD = sensorimotor speed index; PSPD = perceptual-comparison speed index.
$^*p \leq .01$.