Journal of Gerontology: PSYCHOLOGICAL SCIENCES
1997, Vol. 52B, No. 5, P216–P228
Copyright 1997 by The Gerontological Society of America

Controlled and Automatic Forms of Memory and Attention: Process Purity and the Uniqueness of Age-Related Influences

Timothy A. Salthouse,[1] Jeffrey P. Toth,[1] Holly E. Hancock,[1] and John L. Woodard[2]

[1]School of Psychology, Georgia Institute of Technology.
[2]Department of Neurology, Emory University.

Estimates of controlled and automatic processes hypothesized to underlie performance in a memory task and in an attention task were derived for 115 participants from 18 to 78 years of age using the process-dissociation procedure. Participants also performed speed and neuropsychological tests that were suspected to be negatively related to age. Process estimates showed good reliability (from .76 to .98), and the qualitative distinction between processes was supported by the overall pattern of correlations among measures. However, only estimated automatic processes exhibited unique variance, as they were either weakly related or unrelated both to performance on the other tests and to each other. Estimates of the control processes, in contrast, shared considerable variance with measures from other tests, and there were no unique, or independent, age-related effects on these measures. The results highlight the need to distinguish between process purity and the uniqueness of age-related influences in accounting for age differences in cognition.

MOST measures of cognitive performance can be assumed to reflect a mixture of several theoretical processes. This is a potential problem if one is interested in assessing the relations among measures, or between cognitive measures and other variables, because the observed relations could be attributable to only some of the processes or to all of them. For example, if Measure 1 is determined by theoretical processes A and B, then any relation between Measures 1 and 2 could be due to process A, to process B, or to both processes. Identifying the specific process(es) responsible for the relation is important, particularly if processes A and B are thought to be differentially affected by variables of interest, such as age.

In comparing relations among measures, it is also important to be sensitive to the issues of sample size and reliability. Sample size needs to be considered because the precision of an estimated relation (i.e., the narrowness of the confidence interval) is inversely related to the size of the sample. Reliability also cannot be neglected because if the reliability of a measure is low, then there is little systematic variance in that measure available to be associated with other measures. This is of particular concern with difference scores (such as priming measures) because the reliability of such scores is often lower than that of either of the constituent scores (e.g., Cohen & Cohen, 1983, p. 69).

A primary goal in this project was to examine theoretical process estimates from two different tasks to determine their relations with each other, with chronological age, and with a variety of other cognitive measures. The relations among process estimates are interesting because if some of the estimates are assumed to reflect a similar theoretical construct, then one should expect a positive correlation between them. In contrast, little or no correlation would be expected between theoretical estimates if they were postulated to reflect different constructs.

Relations with individual difference variables such as age are interesting because if it is assumed that most of the relations with age are attributable to one particular process (e.g., A and not B), then relations would be expected between age and estimates of process A, but little or no relation would be expected between age and estimates of process B. Finally, relations of the process estimates with other measures of performance are relevant to the issue of the distinctiveness of the hypothesized constructs. That is, if two process estimates reflect the same theoretical construct, they should have a similar pattern of relations with other measures; but if they reflects different constructs, they should exhibit different patterns of relations with other measures.

When the sample consists of a wide range of ages, and several different measures are available from each participant, it is also possible to investigate the uniqueness of age-related influences. That is, many cognitive measures have been found to be sensitive to age-related effects, but it is usually unclear whether these effects are independent of one another. This issue can be important because even if a measure was thought to provide a pure estimate of a specific theoretical process, the interpretation of the age-related effects on that measure would be quite different depending on whether those effects were unique or were shared with other measures. For example, discovering that most of the age-related variance in process A was shared with other cognitive measures would be consistent with the view that age-related effects on process A were simply one manifestation of a more general shared or common factor contributing to the age-related influences on many measures. If this were the case, then at least with respect to age-related influences, there might be nothing special or unique about a theoretically pure process estimate if very little of its age-related influence was independent of the age-related influences on other measures.

Note that the issue here concerns the uniqueness or independence of age-related influences on various measures, and not the purity or validity of the theoretical processes the measures are postulated to assess. That is, examining age-related effects on a measure in the context of age-related effects on other measures is informative as to the uniqueness of age-related influences, but it is not necessarily relevant to the issue of what the measure actually represents.

When a variety of measures is available from individuals spanning a wide age range, mediational relationships among the measures can also be investigated to determine which measures, or theoretical processes, are plausible mediators of the age relations on other measures. In the present study the California Verbal Learning Test (CVLT) was used as the primary criterion task for these analyses, with the goal of determining the extent to which different measures might mediate age-related effects on this task. In the past, statistically controlling measures of processing speed has been found to reduce the age-related variance on a wide range of cognitive tasks (e.g., Salthouse, 1996); hence, speed measures will be examined together with estimates postulated to reflect pure measures of different theoretical processes. In addition, the relative attenuation of the age-related variance in the theoretical process estimates and in the speed measures was examined before and after statistical control of the other type of measure. The reasoning was that the more fundamental measures should produce the greatest attenuation of the age-related effects in the other measures.

To obtain estimates of theoretical processes in both a memory and an attention task, we used the process-dissociation procedure (Jacoby, 1991; Jacoby, Toth, & Yonelinas, 1993), a general analytic technique designed to separate the contribution of controlled and automatic processes to task performance. A stem-completion task similar to that used by Jacoby et al. (1993) was used to assess the contribution of these processes to memory performance. Participants were presented with a list of words under incidental encoding instructions and later were presented with a series of word-stems (e.g., tru _ _) corresponding to the previously presented (old) words as well as unstudied (new) words. Accompanying each test stem was a response instruction (OLD or NEW), which indicated to participants whether they were to include or exclude previously presented words. For the Inclusion (OLD) condition, participants were to complete each stem with a previously presented word or, if they could not remember a studied word, to complete the stem with the first word that came to mind. The Inclusion condition is similar to a traditional cued recall task, but with instructions to guess when recollection fails. In the Exclusion (NEW) condition, the participants were to complete each stem with a word that had not been presented before; that is, they were to generate novel completions for the stems. As in the Inclusion condition, if the individual could not remember a previously presented word (and thus did not know what word to avoid), he or she was to complete the stem with the first word that came to mind. For both conditions the primary dependent measure was the probability of responding with previously presented words.

The rationale underlying these two conditions is that memory performance reflects the contributions of both consciously controlled recollection and more automatic, involuntary effects of memory. Automatic mnemonic processes are hypothesized to increase the probability of responding with old words in both the Inclusion and Exclusion conditions, whereas recollection is hypothesized to increase or decrease this probability depending on test instructions. By casting these conditions into equations (essentially, a theory of performance), one can combine performance in the two conditions and thereby derive separate estimates of the controlled and automatic processes. Formally, the probability of using old words in the Inclusion condition is equal to the probability of recollection (C) plus the probability that the old responses will automatically come to mind (A) when recollection is unsuccessful (1–C). Thus, $p(\text{old}|\text{Inclusion}) = C + A(1-C)$. For the Exclusion condition, in contrast, old words should only be used when they come to mind automatically (A) in the absence of recollection (1–C), because successful recollection would act to exclude previously presented words. Thus, $p(\text{old}|\text{Exclusion}) = A(1-C)$. Combining these equations, mnemonic control (recollection) is estimated as the difference between performance in the Inclusion and Exclusion conditions ($C = \text{Inclusion}-\text{Exclusion}$). Given an estimate of C, one can algebraically derive an estimate of automatic mnemonic processes [e.g., $A = \text{Exclusion} / (1-C)$].

The theory and assumptions underlying the process-dissociation procedure have been discussed extensively in the literature (see Jacoby, 1991; Jacoby et al., 1993; Jacoby, Begg, & Toth, in press; Reingold & Toth, 1966; Toth, Reingold, & Jacoby, 1994, 1995b); we refer the interested reader to these sources. We should note, however, the central assumption underlying the above equations; namely that controlled and automatic processes make independent contributions to performance. Support for this assumption comes from experiments showing invariance in one parameter (e.g., A) across levels of a variable assumed to selectively affect the processes indexed by the other parameter (C). To date, variables producing this pattern in stem- or fragment-cued recall include age (Jacoby, Jennings, & Hay, 1996), divided attention (Jacoby et al., 1993), and levels-of-processing (Toth et al., 1994). The converse pattern (invariance in C across change in A) has also been shown (Hay & Jacoby, 1996).

In the present study we used the equations described above to estimate controlled and automatic mnemonic processes for each research participant. To assess reliability of these estimates, each participant studied and was tested on two different sets of (intermixed) words. A novel aspect of the present study concerned the testing conditions under which mnemonic estimates were obtained. A potential problem with the inclusion/exclusion instructions is their complexity (Graf & Komatsu, 1994). To ensure participants' understanding of instructions, most previous process-dissociation studies have used highly trained examiners in a one-on-one testing session; however, this strategy is labor intensive and limits the conditions under which the procedure can be employed. As part of a larger effort to increase the generality of the procedure (see Hay & Jacoby, 1996; Jacoby, Jennings, & Hay, 1996), we automated all aspects of the experiment on computer (i.e., stimulus presentation, instructions, and response collection) and had the test administered by relatively naïve examiners.

To estimate controlled and automatic processes in the attentional domain, we used a spatial S-R compatibility task modeled after Toth et al. (1995a). In this task, left- (<) or light-pointing (>) arrows were presented in the middle, left, or right portion of a computer screen, and participants were instructed to make left- and right-handed responses that corresponded to the arrow's direction (e.g., a right-handed response to the presentation of '>'). Previous work has shown that, analogous to Stroop interference, participants are unable to avoid processing of the irrelevant dimension (the arrow's spatial location) as shown by faster reaction times (RTs) when the arrow's direction and spatial location are congruent (e.g., a right-pointing arrow on the right side of the screen) as compared with when they are incongruent (e.g., a right-pointing arrow on the left side of the screen).

Toth et al. (1995a) hypothesized that performance in this task could be described as reflecting two independent influences, controlled processing of stimulus form (i.e., arrow direction) and automatic processing of spatial location. Similar to the strategy used for cued-recall, they developed equations for performance on congruent and incongruent trials that were used to derive separate estimates of spatial (S) and form (F) processing. Specifically, they hypothesized that the probability of a correct response on congruent trials is equal to the probability that spatial processing directs responding, plus the probability of responding on the basis of form when spatial location does not direct responding; that is, p(correct|congruent) = S + F(1-S). Correct responding on incongruent trials, in contrast, should only occur when responding is based on form and not spatial location; thus p(correct|incongruent) = F(1-S). Subtracting the proportion correct on incongruent trials from that on congruent trials provides an estimate of spatial processing (S = congruent–incongruent). Given S, an estimate of form processing can be derived algebraically [F = incongruent/(1-S)].

As with cued-recall, the central assumption underlying the equations for this attention task is that spatial and form processes make independent contributions to performance. Toth et al. (1995a) found support for this assumption with the discovery that a manipulation of the proportion of congruent trials influenced spatial, but not form, processing estimates (see also Lindsay & Jacoby, 1994). Additional evidence for the validity of the estimates was provided by showing that form, but not spatial, estimates accurately predicted observed performance on arrows presented at fixation (neutral trials) where spatial processing is minimal.

Note that, unlike most attention tasks in which RT is the primary dependent measure, Toth et al. (1995a) used a deadline procedure to obtain proportional data appropriate for use with the process-dissociation equations. That is, they required participants to respond within 500 ms after stimulus presentation to allow assessment of the accuracy of performance. In the present study we did not use response deadlines because of the well-established relations between age and speed of processing (e.g., Salthouse, 1996). Instead, we used an RT version of the task and used post hoc deadlines to analyze the probability of a correct response at successive time intervals. This strategy allowed us to examine both RT and accuracy as a function of age and to reconstruct the time-course of both observed behavior (i.e., performance on congruent, incongruent, and neutral trials) and process estimates (S and F). Reliability of the process estimates was assessed by administering three separate blocks of 100 trials, each block containing all trial types.

In order to examine the age relations on the estimates of controlled and automatic processing in the context of age relations on other variables, all participants performed several additional cognitive tasks that were hypothesized to be related to age in varying degrees. Several computer-administered and paper-and-pencil speed tasks were used to assess highly age-sensitive perceptual speed abilities, and two fluency tasks were used to assess performance in situations where both speed and word knowledge might be relevant. A spatial test of line orientation was administered because it was assumed to reflect right hemisphere functioning, which has been hypothesized to decline with increasing age; the Trail Making Test was administered because it is sometimes postulated to be sensitive to frontal lobe functioning, which has also been hypothesized to decline with age. Finally, an episodic memory task was administered to provide an age-sensitive measure of verbal memory that could function as a criterion measure in some of the analyses.

Method

Participants

Of the 124 adults examined, complete data were obtained from 115 between 18 and 78 years of age. Participants were recruited from appeals to groups and acquaintances. The primary criteria for inclusion in the study were that the individual had to be in reasonably good health, had to have completed at least 11 years of education, and was not currently a student. Descriptive characteristics of the sample are summarized in Table 1. [Nine participants were excluded because of incomplete data. Two participants in the stem completion task skipped 25% or more of the trials, apparently because they kept the ENTER key depressed during the presentation of the test items. The data from another individual were lost because the computer was inadvertently turned off in the middle of the stem completion task. Six participants in the arrow task used a reversed stimulus-response mapping (i.e., they consistently made left-handed responses to right-pointing arrows, and vice versa), and hence their data were deleted from the analyses. Three other participants had missing data but were kept in the overall sample. One had missing data on the digit symbol RT test, and that value was replaced with an estimate predicted from the individual's age and score on the digit-digit RT test. One individual each had missing data on the CVLT recognition measure and the Judgment of Line Orientation Test, and in both cases the mean of the age group was used as the estimate for the missing value.]

Procedure

To maximize convenience for the participants and to facilitate recruiting, most of the testing was conducted in the

Table 1. Descriptive Characteristics of Research Participants and Summary Measures of Performance

	Age Group			
	18–39	40–59	60–78	Age Correlation
N	40	38	37	—
Age	29.0 (4.8)	49.1 (5.1)	69.2 (5.1)	—
% Female	57.5	50.0	51.4	—
Education	15.5 (1.7)	15.2 (2.5)	15.3 (2.6)	−.01
Health (1 = excellent, 5 = poor)				
Satisfaction	2.4 (0.5)	2.2 (0.8)	2.0 (0.8)	−.19
Rating	2.2 (0.8)	2.3 (0.9)	2.1 (1.0)	−.06
Limitations	1.4 (0.6)	1.5 (0.8)	1.8 (1.0)	.18
Cardiovascular surgery (%)	0	0	13.5 (34.7)	.26*
Blood pressure medications (%)	5.0 (22.1)	15.8 (37.0)	43.2 (50.2)	.41*
Head injury (%)	5.0 (22.1)	2.6 (16.2)	2.7 (16.4)	−.07
Neurological treatment (%)	2.5 (15.8)	5.3 (22.6)	8.1 (27.7)	.11
Vocabulary (no. correct out of 10)				
Synonym	5.5 (2.9)	6.3 (2.7)	8.1 (2.2)	.40*
Antonym	4.9 (2.9)	5.2 (3.1)	6.6 (3.2)	.24*
CVLT (no. of items correct)				
List A				
Trial 1	7.9 (1.9)	6.6 (1.6)	6.2 (2.1)	−.35*
Trial 2	10.5 (2.4)	9.6 (2.3)	8.2 (2.6)	−.38*
Trial 3	11.3 (2.6)	10.4 (2.5)	9.4 (2.7)	−.34*
Trial 4	12.5 (2.4)	11.3 (2.3)	10.1 (2.7)	−.40*
Trial 5	12.8 (2.7)	11.7 (2.3)	10.7 (2.8)	−.40*
List B	7.0 (2.4)	6.5 (1.9)	6.4 (1.8)	−.17
Immediate free recall	11.9 (2.8)	10.6 (2.8)	8.4 (3.3)	−.45*
Immediate cued recall	12.5 (2.4)	11.8 (2.5)	9.7 (3.0)	−.41*
Delayed free recall	12.2 (2.8)	10.8 (2.8)	9.0 (3.2)	−.44*
Delayed cued recall	12.6 (2.6)	12.0 (2.5)	9.6 (2.8)	−.43*
Recognition (no. correct)	14.9 (1.2)	14.0 (1.7)	13.2 (2.1)	−.41*
Arrow				
Neutral RT	434 (64)	499 (61)	600 (142)	.59*
Congruent RT	434 (79)	495 (64)	565 (123)	.52*
Incongruent RT	471 (74)	551 (74)	661 (147)	.62*
Form estimate @ 500 ms	.688 (.259)	.403 (.269)	.183 (.236)	−.64*
Space estimate @ 500 ms	.136 (.122)	.221 (.120)	.245 (.143)	.33*
Stem completion				
p(old|inclusion)	.543 (.094)	.480 (.096)	.464 (.113)	−.30*
p(old|exclusion)	.234 (.181)	.236 (.153)	.255 (.113)	−.04
p(pass|old-incl)	.088 (.093)	.085 (.089)	.108 (.114)	.07
p(pass|old-excl)	.109 (.092)	.120 (.124)	.114 (.138)	.02
p(pass|new-incl)	.221 (.192)	.209 (.174)	.232 (.226)	.03
p(pass|new-excl)	.098 (.099)	.101 (.124)	.103 (.159)	.01
Controlled estimate	.309 (.194)	.245 (.188)	.239 (.184)	−.14
Automatic estimate	.299 (.183)	.284 (.150)	.277 (.123)	−.07
Reaction time (ms)				
DDRT	665 (101)	744 (95)	920 (225)	.58*
DSRT	1428 (186)	1693 (280)	2076 (456)	.65*
Paper-and-pencil speed (no. items/sec)				
SD Copy 1	1.69 (0.35)	1.59 (0.32)	1.31 (0.31)	−.47*
SD Copy 2	1.70 (0.31)	1.62 (0.29)	1.37 (0.26)	−.45*
XO comparison 1	1.18 (0.24)	1.03 (0.31)	0.87 (0.16)	−.49*
XO comparison 2	1.22 (0.24)	1.09 (0.32)	0.90 (0.20)	−.46*
Letter comparison 1	0.38 (0.08)	0.33 (0.09)	0.27 (0.09)	−.52*
Letter comparison 2	0.33 (0.08)	0.31 (0.08)	0.23 (0.12)	−.41*
Pattern comparison 1	0.64 (0.14)	0.56 (0.11)	0.46 (0.10)	−.51*
Pattern comparison 2	0.58 (0.12)	0.54 (0.09)	0.44 (0.10)	−.47*
Fluency (no. items in 60 sec)				
VF-C	16.2 (4.7)	14.9 (4.9)	14.8 (5.2)	−.15
VF-F	14.9 (4.2)	13.7 (4.6)	14.6 (4.7)	−.07
VF-L	15.0 (4.1)	13.4 (3.3)	13.5 (4.9)	−.15
VF-animals	20.6 (5.1)	18.8 (5.7)	17.1 (4.6)	−.26*
VF-furniture	12.9 (2.8)	12.3 (3.4)	11.8 (3.3)	−.14
VF-vegetables	13.8 (3.3)	14.0 (3.9)	14.0 (3.3)	.03
Judgment of Line Orientation	12.7 (1.7)	12.0 (2.2)	12.1 (2.3)	−.13
Trail Making (no. seconds)				
Trail A	21.0 (4.6)	26.2 (7.9)	32.9 (10.5)	.53*
Trail B	53.6 (20.3)	66.7 (21.3)	87.2 (36.6)	.46*

Notes: CVLT = California Verbal Learning Test; RT = reaction time; DDRT = digit-digit reaction time; DSRT = digit-symbol reaction time; VF = verbal fluency.
*$p < .05$.

participants' homes. The testing session was preceded by a short description of the tasks, and written informed consent was then obtained. The session lasted between 1.5 and 2 hours, depending on the individual's pace and desire for breaks. All participants received the tests in the following order: background questionnaire, SD copy, XO comparison, letter comparison, pattern comparison, synonym and antonym vocabulary, digit-digit RT, digit-symbol RT, arrow RT, stem completion, CVLT, letter and category verbal fluency, Judgment of Line Orientation, delayed portion CVLT, and Trail Making Test Parts A and B. The background questionnaire contained a variety of questions about health status and education (Table 1).

Speed tests. — The letter comparison and pattern comparison tests were identical to those used in earlier studies (e.g., Salthouse, 1996). These paper-and-pencil tests consist of pairs of 3, 6, or 9 letters (letter comparison) or line segments (pattern comparison), with one-half of the pairs differing in one element. The task for the participant was to decide whether the two members of the pairs were the same or different, and to write the letter S (for same) or D (for different) between the members of the pairs as quickly as possible. Two separate administrations of each test were presented with 30 seconds allowed in each administration. For ease of comparison with other variables, the scores were converted to the number of items completed per second.

Two simpler paper-and-pencil tests were designed to assess presumed components of the comparison tests. In the XO comparison test the items consisted of pairs of letters that were either both X, both O, or an X and an O. As in the other comparison tests, the participant was instructed to classify the pairs as same or different as quickly as possible by writing the letter S or the letter D between the two members of the pair. This test was assumed to require a very elementary type of comparison, of a single element, relative to the letter comparison and pattern comparison tests, which contained multiple elements. The SD copy test consisted of columns of blank lines adjacent to columns of the letters S and D intermixed. The task for the participant was to copy the indicated letter as rapidly as possible on the blank line. This test was assumed to involve some of the same sensory and motor requirements of the comparison tasks but without any requirement for comparison. Each of these tests was administered twice with 30 seconds allowed for each administration. The scores in each test were the number of items completed per second.

The digit-digit and digit-symbol RT tasks have been used in numerous recent studies (e.g., Salthouse, 1996) and consist of a choice RT response to a pair of items presented in the middle of the computer screen. In the digit-digit task the items were digits, and in the digit-symbol task the top item was a digit and the bottom item was a symbol. In both versions a code table was presented at the top of the screen, but in the digit-digit version it merely consisted of identical digits, whereas in the digit-symbol version it contained pairs of digits and symbols. The task for the participant was to decide, as rapidly as possible, whether the pair of items presented was the same or different, either according to physical identity (digit digit) or according to correspondence in the code table (digit symbol). Same responses were indicated by pressing the / key, different responses by pressing the Z key. A practice block of 18 trials preceded the test blocks of 27 trials each in both the digit-digit and digit-symbol versions of the task. Average accuracy was greater than 95% in both of the tasks and was not significantly correlated with RT; performance in each task is therefore summarized in terms of the median RT in milliseconds (ms).

Vocabulary tests. — The vocabulary tests were identical to those used in several earlier studies (e.g., Salthouse, 1993), and consisted of 10 five-alternative multiple-choice synonym items and 10 five-alternative multiple-choice antonym items. Five minutes were allowed for the completion of the two parts of this test. The scores were the numbers of words answered correctly in each part of the test.

Arrow Task. — Stimuli in the arrow task consisted of right- (>) and left- (<) pointing arrows presented along the medial-horizontal axis of the computer screen in one of three locations: left, right, or center. Trial-type (congruent, incongruent, neutral) was defined by the relationship between the location and direction of each arrow: For congruent trials, the direction and location of the arrow denoted the same response (e.g., a left arrow on the left side of screen); for incongruent trials, direction and location denoted opposite responses (e.g., a left arrow on the right side of screen); and for neutral trials, the arrow was presented in the center of the screen and had no apparent spatial displacement (relative to the fixation point). In each of three blocks participants responded to 100 consecutive trials consisting of 40 congruent, 40 incongruent, and 20 neutral trials. Trial types occurred randomly within each set of 10 trials so that, although trial type could not be predicted from trial to trial, the proportion of trial types was consistent throughout testing.

Participants were instructed to sit a comfortable distance from the computer screen. At a distance of 50 cm, each arrow subtended a visual angle of approximately 1.1° square and, for congruent and incongruent trials, appeared approximately 8° to the left and right of fixation. Prior to testing, participants read instructions and were shown visual examples of the arrows and trial types. Instructions stressed both speed and accuracy. The test began immediately after a practice block of 50 trials. Each trial began with a fixation point (asterisk) presented in the center of the screen for 250 ms. Immediately upon its removal, an arrow was presented for approximately 90 ms. Participants responded by pressing the Z key for left-pointing arrows and the / key for right-pointing arrows. Response timing began at arrow onset and continued until a response was made. After a response was detected, the screen was cleared and remained blank for 1 second, at which time the fixation point for the next trial was presented.

Stem completion task. — Critical stimuli in the stem-completion task consisted of 120 five-letter words, mostly nouns, ranging in frequency from 1 to 200 (Kucera & Francis, 1967). The words were divided into three sets (1–3), and each set was further divided into two subsets (A and B) of 20 words each (mean frequency of the six subsets from 28.8 to 31.9). Two of the sets (1 and 2) were used as studied words. Stems corresponding to all of the critical words were presented at test under either inclusion (subsets 1A, 2A, and 3A) or exclusion (1B, 2B, and 3B) instructions. As is common in individual difference research, all participants received the same items for a given condition to avoid confounding experimental treatment with the characteristics of participants. However, the presentation order of items at study, and within conditions at test, was randomized for each participant. In addition to the 120 critical words, there were 16 buffer words (8 primacy, 8 recency), 2 of which were used as examples in the test instructions (see below). Word stems were created by replacing the last two letters in each word with two underscores. The word stems were unique within the set of critical words, and each had at least two completions.

All aspects of the task were presented on computers and were self-paced. Study instructions informed participants that a list of words would be presented that they were to rate for pleasantness on a scale of 1 to 5 (1 = most unpleasant and 5 = most pleasant). They were told that their ratings should be based on the meaning of the words, but that we were interested in their first impression, so they should make their ratings quickly. No mention was made of the subsequent memory test. Study words remained on the screen until the participant responded by pressing one of the number keys (from 1 to 5) on the keyboard. Once an appropriate key was pressed, the next word was automatically presented after a blank-screen delay of 1 second.

Test instructions were presented immediately following study. Participants were informed that their memory would be tested for words presented in the pleasantness-rating task using word-stems ["the first three letters of five-letter words followed by two dashes (e.g., pea_ _)"]. They were also told that, in addition to word stems, the message OLD or NEW would appear above each stem, and that they were to respond differently depending on this message:

> If the message is OLD your job is to complete the stem with a word presented in the pleasantness task; that is, try to complete the stem with an old word. If the message is NEW your job is to complete the stem with a word that was NOT presented in the pleasantness task; that is, try to come up with a new word. Regardless of the message (OLD or NEW), if you cannot remember a word from the pleasantness task that fits into the stem, then just complete the stem with the first five-letter word that comes to mind. All of the stems can be completed with more than one word so try to come up with a completion for all of them. However, do not use plurals or proper nouns. Also, if the message tells you to give a NEW word, but all you can think of is ONE completion that you are SURE is old, then it is okay to pass that stem by entering xx.

Participants made their responses by typing in two letters, which appeared on the screen under the stem's two underscores, and then pressing the ENTER key.

Following test instructions, the computer led participants through four practice test trials. The first two used stems corresponding to the last two buffer words in the study list; the last two practice trials used new stems. The first and third practice trials were Inclusion (OLD) trials, the second and fourth were Exclusion (NEW) trials. For each practice trial, the computer represented the appropriate instructions and provided feedback concerning the participant's responses. For example, if a participant responded with a plural word, the computer informed him or her of the error and presented the test stem again until an appropriate response was provided. Following practice, the participant was given the option of beginning the actual test or repeating the practice phase. No feedback was provided during the test proper.

Test stems were presented in the center of the screen, in large lower-case letters, and in white against a black background. The Inclusion (OLD) and Exclusion (NEW) cues were bright green and bright red, respectively. They appeared 500 ms prior to the presentation of the test stem and were positioned in the center of the screen just above the stem. Stems and response cues remained on the screen until participants responded; if 15 seconds elapsed without a response, the computer beeped and the message "Please enter a response" appeared below the stem. Once a response was entered, the response cue for the next trial (OLD or NEW) appeared immediately.

Other tests. — The CVLT (Delis, Kramer, Kaplan, & Ober, 1987) was administered according to the instructions in the manual. There are two lists of 16 words each in this test, with four words from each list in each of four categories (i.e., spices, fruits, clothing, tools). The procedure involves the presentation of five immediate free recall trials with List A, one immediate free recall trial with List B, recall of List A, and cued (by category) recall of List A. After a 20-minute delay occupied by other activities, there was a delayed free recall test of List A, a delayed cued recall of List A, and a delayed recognition test for items from List A.

Two different types of verbal fluency tests were administered. In the letter fluency tests the participants were allowed 1 minute each to say as many words as possible that began with the letters C, F, and L, with the constraint that none of the words should be proper nouns. In the category fluency test they were allowed 1 minute each to say as many words as possible that were members of the animals, furniture, or vegetables categories. The score in each test was the number of different words produced that were in the appropriate category.

The Judgment of Line Orientation Test (Benton, Hamsher, Varney, & Spreen, 1983) was administered according to the published instructions except that only the odd-numbered items were presented. This test consists of a display of line segments in different orientations with the participant instructed to select the lines of matching orientations from a semicircle of numbered lines. The score is the number of test lines matched correctly. This abbreviated form has been demonstrated to have acceptable reliability and validity (Woodard et al., 1997).

The Trail Making Test (Reitan, 1992) was also administered according to the published instructions. This test consists of two versions, A and B, and in both versions the test form consists of a haphazardly arranged set of 25 circles on a piece of paper that are to be connected as rapidly as possible. In the A version of the test the circles contain the numbers from 1 to 25, and the circles are to be connected in numerical sequence, but in the B version the circles contain alternating letters and numbers; thus, the participant has to connect the 1 circle with the A circle, the A circle to the 2 circle, the 2 circle to the B circle, etc. Because errors were very infrequent, the measure of performance in these tests is the time in seconds to complete the sequence.

Results

CVLT analyses. — There are many possible measures that can be derived in the CVLT, but most have similar relations to age (Table 1) and are not independent of one another. A principal components analysis was therefore conducted to identify a parsimonious set of variables for later analyses. Two components were identified with eigenvalues greater than 1, the first accounting for 68.3% of the variance, and the second accounting for 9.8% of the variance. All of the variables except List B recall had moderate to high loadings (between .62 and .92) on the first component, and List B recall was the only variable with a high loading (>.5) on the second component. Further analyses revealed that the sum of the scores across the first five trials in List A was correlated .96 with the first principal component score and only .15 with the second component score, whereas List B recall was correlated .74 with the second component score and only .49 with the first component score. These two variables were therefore used to summarize CVLT performance in all subsequent analyses.

Arrow analyses. — Two separate sets of measures were obtained in the arrow task. One set consisted of the median reaction times (RT) on neutral trials (arrows in the middle of the screen), congruent trials (e.g., left-pointing arrow on the left side of the screen), and incongruent trials (e.g., left-pointing arrow on the right side of the screen). Means of these medians for each age group are presented in Table 1, where it can be seen that RTs in each condition increased with age. Hierarchical regression analyses were conducted to determine whether there was a significant influence of age on congruent or incongruent RT after control of neutral-trial RT. In neither case was the residual age-related variance significantly different from zero (i.e., incongruent, $\Delta R^2 = .005$; congruent, $\Delta R^2 = .002$), indicating that there was no relation between age and either measures of interference (i.e., incongruent relative to neutral) or facilitation (i.e., congruent relative to neutral) after overall speed was taken into consideration.

The second set of measures from the arrow task was estimates of the probabilities of spatial- and form-based responding. The initial step in these analyses consisted of partitioning the data according to response time, and then computing the cumulative probability of a response in the neutral, congruent, and incongruent conditions for each 100-ms interval from 200 to 1000 ms. The results of these computations are displayed in Figure 1. Notice that, as expected, the probabilities increase with greater time, and

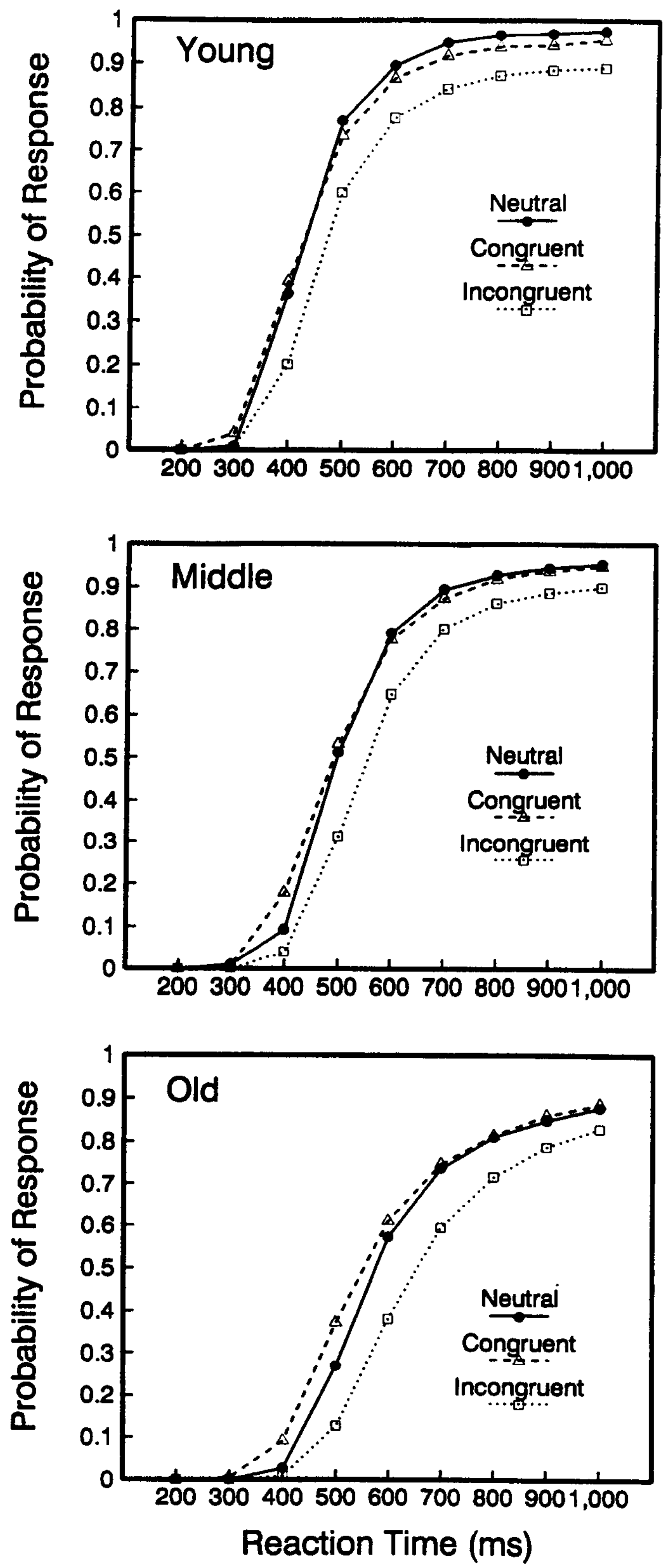

Figure 1. Probability of a response in the neutral, congruent, and incongruent conditions in the arrow task as a function of age group (young = 18 to 39, middle = 40 to 59, and old = 60 to 78) and post hoc response deadline.