THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY, 2005, *5XA* (X), XXX–XXX

Age-related declines in context maintenance and semantic short-term memory

Henk J. Haarmann
University of Maryland, College Park, USA

Gemma E. Ashling
Goldsmith College, University of London, UK

Eddy J. Davelaar and Marius Usher
Birkbeck College, University of London, UK

This study reports age-related declines in context maintenance (Braver et al., 2001) and semantic short-term memory (STM) and evidence for a relation between the two. A group of younger and older adults completed a context maintenance task (AX-CPT), a semantically oriented STM task (conceptual span), a phonologically oriented STM task (digit span), and a meaning integration task (semantic anomaly judgement). In the AX-CPT task, a target response is required to the probe letter "X" but only when it is preceded by the letter "A" (the context). Either three (short interference) or six distractor letters (long interference) were presented between the cue and the probe. Results indicated an age-related deficit in context maintenance. Age-related declines were also observed for conceptual span and semantic anomaly judgement but not for digit span. Context maintenance was correlated with conceptual span and semantic anomaly judgement but not with digit span. These correlations were largely mediated by age differences, which also explained variance that was unique to (and not shared among) context maintenance, conceptual span, and semantic anomaly judgement.

It is well known that healthy ageing reduces information processing speed (Cerella, 1985; Myerson, Hale, Wagstaff, Poon, & Smith, 1990; Salthouse, 1996) and results in performance declines in various cognitive functions, including working memory (Morris, Craik, & Gick, 1990; Verhaeghen & Salthouse, 1997), inhibition (Hasher, Stoltzfus, Zacks, & Rypma, 1991; Hasher & Zachs, 1988; West & Alain, 2000; Zacks & Hasher, 1997), attention (Anderer, Pascual-Marqui, Semlitsch, & Saletu, 1998), and episodic memory (Craik, 1977; Moscovitch & Winocur, 1992). Furthermore, older adults show patterns of task performance that suggest

---

Correspondence should be addressed to Henk Haarmann at the Department of Hearing and Speech Sciences, University of Maryland, College Park, MD 20742, USA. Email: hhaarmann@hesp.umd.edu

EJD was supported by the Economic and Social Research Council (T0262701312).

http://www.tandf.co.uk/journals/pp/02724987.html DOI:10.1080/02724980443000214

age-related declines in various forms of attention, including selective attention (Brink & McDowd, 1999; Panek, Rush, & Slade, 1984; Spieler, Balota, & Faust, 1996; West & Baylis, 1998; West & Bell, 1997), divided and/or alternating attention (Brouwer, Waterink, Van Wolffelaar, & Rothengatter, 1991; Jensen & Goldstein, 1991; Korteling, 1993), and sustained attention (Filley & Cullum, 1994; Parasuraman & Nestor, 1991). Older adults take longer to respond on many tasks from very simple perceptual tasks to more complex cognitive tasks (Cerella, 1985; Myerson et al., 1990). Ageing also affects executive function, as indicated, for example, by age-related declines on neuropsychological tests for assessing prefrontal dysexecutive syndrome, such as the Wisconsin Card Sorting Test (Daigneault, Braun, & Whitaker, 1992), the fluency test (Keys & White, 2000), and the trail making test (Keys & White, 2000), even when speed of processing is controlled for. Finally, within the domain of episodic memory, age-related declines are especially apparent on tasks that involve free recall (Craik & Jennings, 1992), temporal order memory (Parkin, Walter, & Hunkin, 1995), source memory (Spencer & Raz, 1994), and release from proactive inhibition (Dobbs, Aubrey, & Rule, 1989).

One of the major challenges for a theory of cognitive ageing is to provide a unifying account for the age-related declines that are observed in the various cognitive domains, in a manner that is consistent with what is known about the neurobiological effects of ageing (Braver et al., 2001; Moscovitch & Winocur, 1992). Theories of ageing have tried to explain age-related performance declines across a variety of tasks in terms of a deficit in a single underlying cognitive function or a small set of cognitive functions, including processing speed (Myerson et al., 1990; Salthouse, 1996), working memory capacity (Craik, Morris, & Gick, 1990; Light & Anderson, 1985; Salthouse, 1992), inhibition (Hasher & Zacks, 1988; Zacks & Hasher, 1997), and attention (West & Bell, 1997). Recently, Braver, Barch, Cohen, and their colleagues have proposed a theory of cognitive ageing that has the potential to unify various previous theories (Braver et al., 2001; Braver & Barch, 2002). Their proposal is that cognitive ageing involves a working memory deficit in the ability to represent, maintain, and update information about task context. As a result, information about task context is used less effectively in governing the use of task-relevant information, explaining age-related problems on a variety of tasks in terms of a common mechanism. For example, in the Stroop task, the task instructions ("name ink colour") must be actively represented and maintained in working memory to bias attention allocation and response selection toward the ink colour dimension of a visually presented word. In the Stroop task, an age-related deficit in context representation and/or maintenance may appear as a problem in selective attention and/or a problem in inhibiting processing of the irrelevant task dimension (i.e., word reading; Braver et al., 2001). In dual-task paradigms, there are frequent shifts between tasks, requiring a continuous updating of task context. Thus, an age-related deficit in the representation, maintenance, and/or updating of task context may appear as a problem in dividing attention between two tasks. The same age-related deficit may interfere with performance on working memory tasks, which, due to their simultaneous storage and processing demands, involve a dual-task component and a need to frequently update the task context (Braver et al., 2001). Finally, age-related declines in episodic memory, such as, free recall, recall of temporal order, and source memory, may also be due to a problem in the representation, maintenance, and/or updating of context (Braver et al., 2001), since such tasks require integration of outputs from long-term memory with relevant contextual information or strategic cues (Moscovitch & Winocur, 1992;

Perfect, 2003). The theory that ageing impairs the processing of context in working memory is consistent with neurobiological theories of ageing according to which age-related cognitive declines are due to deficits in the executive function of the prefrontal cortex (PFC), resembling the deficits of neuropsychological patients with PFC damage (Moscovitch & Winocur, 1992; Perfect, 2003). Such deficits may be the expression of an underlying deficit in the processing of task context in a specific region within the PFC, the dorso-lateral PFC (DL-PFC; Braver & Barch, 2002; Braver et al., 2001; MacPherson, Phillips, & Della Sala, 2002). Indeed, neuroimaging data indicate that the DL-PFC supports the processing of task context in working memory and is activated differently during the maintenance of task context in old adults from the way it is in young adults (i.e., showing a decline instead of increase of activation across time; Barch, Braver, Racine, & Satpute, 2001).

One of the most direct pieces of evidence for an age-related deficit in the processing of context in working memory has been obtained with a modified version of the AX-CPT test (Braver et al., 2001). The AX-CPT test is an adaptation of the classic continuous performance test (CPT; Rosvold, Mirsky, Sarason, Bransome, & Beck, 1956), which has been especially designed to place a demand on the processing of task-relevant contextual information in working memory (Cohen, Barch, Carter, & Servan-Schreiber, 1999; Servan-Schreiber, Cohen, & Steingard, 1996). Participants are presented with a series of letters and have to give a target response to each occurrence of the letter "X" (i.e., the target) but only when it is preceded by the letter "A" (valid cue), which serves as the context. The experimental condition in which this happens is known as AX. In all other conditions, including BX (invalid cue followed by a target), AY (valid cue followed by a nontarget), and BY (invalid cue followed by nontarget), a nontarget response is required. The letter"B" is used to indicate any letter other than the cue "A", whereas the letter "Y" is used to indicate any letter other than the target "X". In the ageing study by Braver et al. (2001), AX trials occurred much more frequently than each of the other three trial types (e.g., 70% versus 10% each), creating a strong bias to respond incorrectly with a target response in the BX and AY condition and, therefore, increasing the demands on the use of context information (i.e., valid vs. invalid cue; Cohen et al., 1999). Poor context representation and/or maintenance should *increase* errors in the BX condition, since the noncue (B) is less likely to be in working memory when the target (X) is presented and since the target response is a prepotent response due to high proportion of AX trials. By contrast, poor context representation and/or maintenance should *decrease* errors in the AY condition, since the cue (A) is less likely to be in working memory and create a bias towards a target response when a nontarget (Y) is presented. These context maintenance problems should be exacerbated when the delay between the context (a cue or noncue) and the probe (a target or nontarget) is made longer and/or filled with intervening distractor letters (Braver & Barch, 2002; Braver et al., 2001). To test this latter prediction, Braver et al.'s (2001) study manipulated the properties of the cue–probe interval (4.9 s). In a baseline condition it was unfilled, whereas in an interference condition three distractor letters (requiring a nontarget response) were presented during the interval. A degraded condition, identical to the baseline condition except for visual degradation of the cue and probe letter, was included to control for task difficulty.

The results of the Braver et al. (2001) study provided evidence for a context maintenance deficit in older adults (Barch et al., 2001). In the baseline condition, compared to young adults, older adults showed a trend towards more BX errors, fewer AY errors, reduced context sensitivity (as indicated by a d' measure using BX errors for false alarms), and slower responses on AX and BX (but not AY) trials. This pattern was significant and amplified in the interference condition, which placed a greater demand on context maintenance, strongly suggesting an age-related impairment in context maintenance. Both the interference and the degraded condition showed an increase in target errors relative to the baseline condition. However, unlike interference, perceptual degrading did not affect nontarget errors, suggesting that the interference effects on BX and AY errors were not caused by an increase in the general level of task difficulty (Braver et al., 2001) but by an impairment in context maintenance. This interpretation was confirmed by the simulation results of a neurocomputational model, which assumed that ageing impairs context maintenance and updating in the DL-PFC (Braver et al., 2001), consistent with neuroimaging results (Barch et al., 2001). When such a deficit was simulated, the model showed that increasing the cue–probe interval decreases performance on BX trials and increases performance on AY trials. The intact model, by contrast, showed no change in performance on BX trials and even a slight increase in AY errors with increased delay.

The proposal that ageing affects a working memory component that processes context information raises the issue of whether this component is related to short-term memory (STM) for information about the identity of previous lexical/semantic stimuli. This issue arises especially since STM may be viewed as the storage component of working memory (Cowan, 1999; Engle, Tuholski, Laughlin, & Conway, 1999; Haarmann, Davelaar, & Usher, 2003). One possibility is that there is no relation between working memory for context and short-term memory for identity information. As Braver et al. (2001) pointed out, this might explain why ageing is associated with declines in performance on complex working memory tasks, but with no or much smaller declines in standard STM tasks, such as digit span (Fisk & Warr, 1996; Humes, Nelson, Pisoni, & Lively, 1993). Whereas complex working memory tasks, such as reading span (Daneman & Carpenter, 1980) and operation span (Turner & Engle, 1989), place large demands on the representation, active maintenance, and updating of context, standard STM tasks, such as digit span, may incur those demands to a much lesser extent. For example, operation span requires solving a sequence of arithmetic problems, while maintaining digits (or words) presented in between problems for subsequent recall. Thus, operation (but not digit) span requires a continuous updating of the internal representation of the task instructions from a focus on arithmetic to one on digit recall. Nevertheless, another possibility is that there is a relation between working memory for context and STM for identity information. This is suggested by the need of context maintenance to allow semantic disambiguation in language processing. For example, in order to select the contextually appropriate meaning of an ambiguous word at the end of a sentence (e.g., whether the word *bank* refers to the edge of a river or a financial institution) one must be able to maintain the specific meanings of one or more prior words as part of its semantic context (e.g., "*Because the river was so beautiful, they decided to go for a walk along the bank*; Swinney, 1979). There is evidence that schizophrenia patients with positive thought disorder have difficulties maintaining context information in both a semantic disambiguation task (Bazin, Perruchet, Hardy-Bayle, & Feline, 2000; Cohen & Servan-Schreiber, 1992) and an AX-CPT task (Braver & Barch, 2002; Braver et al., 2001), suggesting that the mechanisms supporting STM for (semantic) identity information and context maintenance could be related.

If indeed such a relation exists, then the question arises of how to explain the finding of more severe age-related deficits on working memory tasks (e.g., reading span, operation span) than on standard span tasks, such as digit span (Fisk & Warr, 1996; Humes et al., 1993). The latter type of task may rely primarily on the classic phonological STM (Baddeley, 1986, 1992; Burgess & Hitch, 1999; Warrington & Shallice, 1969), which may not place a large demand on context maintenance, especially since it uses a specialized mechanism for retaining phonological codes (i.e., articulatory rehearsal of phonologically decaying codes). Recent neuropsychological findings by Martin and colleagues (Freedman & Martin, 2001; Martin & Romani, 1994; Martin, Shelton, & Yaffee, 1994; Romani & Martin, 1999) suggest that verbal STM consists not only of a phonological STM but also of a semantic STM, which maintains information about the identity of a small set of word meanings. They reported a double dissociation between two types of STM deficit. One patient, E.A, had a greater phonological than semantic STM impairment, and three patients, A.B., M.L., and G.R., had a greater semantic than phonological STM impairment (Freedman & Martin, 2001; Martin & Romani, 1994; Martin et al., 1994; Romani & Martin, 1999). Patient E.A. but not A.B. showed the patterns of effect typically associated with a phonological STM deficit including (a) impaired rhyme probe recognition, (b) a reversed modality effect (i.e. better performance in the visual than auditory modality), (c) lack of a phonological similarity effect in the visual modality, and (d) lack of a word length effect in both modalities. By contrast, patient A.B. but not E.A. showed a pattern of effects that indicated a deficit in semantic STM, including (a) severely impaired category probe recognition, (b) absence of a lexicality effect (i.e., words were not better recalled than nonwords), (c) problems with semantic attribute judgements (e.g., "Which is softer, cotton or sandpaper?"), and (d) problems with on-line semantic anomaly judgement, especially when the memory load is high (e.g., "*Jeeps*, men, and women were *walking* the streets"; Martin & Romani, 1994). Additional evidence for a semantic STM also comes from a recent cognitive study showing that lexical semantic effects are observed at recency in immediate free recall, even after contributions from long-term memory are factored out and also after contributions from phonological STM are minimized through articulatory suppression (Haarmann & Usher, 2001). This is also consistent with findings indicating the existence of semantic strategies that can eliminate word length and similarity effects in word span tasks even for subjects whose span is in the normal range (Logie, Della Sala, Laiacona, Chalmers, & Wynn, 1996).

Given such dissociable systems within verbal STM, it seems possible that the part of working memory that maintains information about task context is related to semantic but not phonological STM. This idea is consistent with the results of a recent correlation study with healthy young adults (Haarmann et al., 2003), showing that semantic but not phonological STM correlates well with performance on more complex cognitive tasks, including semantic anomaly judgement, text comprehension, and verbal problem solving. While phonological STM was assessed with a nonword span test requiring serial recall, semantic STM was assessed by means of the conceptual span test (Haarmann et al., 2003), which had been designed to maximize the contribution of semantic STM. In the *conceptual span* task, participants are presented with a memory list of nine randomly ordered words, consisting of three words in each of the three semantic categories. The memory list is followed by the name of one of the categories, and participants try to recall the words in the memory list belonging to that category. Since conceptual span uses a semantic cued recall procedure and does not require serial recall of the three words within a category, it is likely to primarily engage semantic STM and to involve only a minimal contribution of phonological STM.[1] Moreover, this span measure does not require alternation between memory encoding and production, as more complex span measures do. Nevertheless, the conceptual span test outperformed a complex span measure, the reading span (Daneman & Carpenter, 1980), in predicting semantic anomaly judgement and verbal problem solving (Haarmann et al., 2003), suggesting an important role for semantic STM in complex cognitive tasks.

Overview of the current study

The present study had three major aims, namely (a) to replicate and extend the finding of an age-related context maintenance deficit in the AX-CPT task, (b) to examine whether there is an age-related deficit in semantic (but not phonological STM), and (c) to test the hypothesis that there is a relation between the systems for context maintenance and semantic (but not phonological) STM. The semantic and the phonological STM abilities are tested with the conceptual span (described above) and the digit span, respectively. The digit span was used as it is one of the most standard tests (included in the Wechsler IQ test; test–retest reliability of .83; Wechsler, 1981) and is traditionally related to the phonological loop.

In discussing their results, Braver et al. (2001) pointed out that “a strong claim of our theory is that inhibitory deficits in healthy older adults will be greatest under task conditions in which successful inhibition is dependent upon actively maintaining context information over a delay period. This claim remains to be tested” (p. 20). Although in Braver et al., the age effect on AY errors in the interference condition was as predicted, the effect of interference on the AY errors in the older group did not show the predicted improvement (relative to baseline). It is possible, however, that the comparison between the interference and the baseline conditions involves processes additional to context maintenance. Therefore, in the present study, we manipulated the difficulty of maintaining context information in the AX-CPT task, by contrasting a short and long interference condition during which three and six distractor letters were presented, respectively. To ensure encoding of these distractor letters and to prevent rehearsal of the cue, participants read the cue and distractor letters out loud. In addition, we intermixed trials belonging to the long and short conditions, rather than presenting them in separate blocks (as in Braver et al., 2001), to ensure that differences are not due to strategy variations between the two conditions (e.g., different encoding of context cue or increased attention) but can be attributed to the amount of time over which the context is maintained and exposed to interference. Finally, in order to obtain a more sensitive measure of performance, we used a speeded performance measure by imposing a 1-s response deadline. Notice that although older people are known to be sensitive in particular to time pressure, our main interest focuses on the predicted within-group effect of interference interval on performance. Accordingly, maintenance of the cue is more difficult in the long than in the short interference condition. The added maintenance difficulty reduces the ability of an invalid cue to inhibit a target response on BX trials. It furthermore reduces the ability of a valid cue to create a bias towards a target response on AX and AY trials. Hence, if older

---

[1] This is suggested by the absence of word length effects in semantic category cued recall (Haarmann et al., 2003).

adults have a deficit in maintaining context information (i.e., valid vs. invalid cue) in working memory, then they should produce more AX and BX errors but fewer AY errors in the long than in the short interference condition.

Using this rationale, we defined an index of context maintenance, based on the equation (AX short − AX long) + (BX short − BX long) + (AY long − AY short), where AX, BX, and AY refer to the error percentages on the corresponding trial types and short and long to the short and long interference condition. Semantic STM was assessed with the conceptual span test, while phonological STM was assessed with a digit span test. If the systems for context maintenance and semantic (but not phonological) STM are related, then there should be an age-related decline not only in context maintenance in the AX-CPT task but also in conceptual (but not digit) span. In addition, the index of context maintenance should correlate positively with conceptual span, whereas it should not correlate with digit span. We expected the correlation to be mediated largely by age-related effects, because ageing may affect overlapping or adjacent systems (in the prefrontal cortex) that support context maintenance and semantic STM. To assess the role of such a system for performance on a complex cognitive task that involves the on-line storage and integration of word meanings, a speeded anomaly judgement task requiring detection of semantic errors among the words in a sentence was included as well.

Method

Participants

Participants in the study were 43 young adults (age range 21–37 years) and 36 older adults (age range 66–85 years). The young adult group included 29 females and 14 males, while the older adult group included 23 females and 13 males. Young participants were recruited from Birkbeck College and from University College London, while older participants were recruited from a geriatric day centre (Edith Cavell House of Peterborough Hospital), from a residential home for the Elderly (Tanglewood), and from Birkbeck College in London. Informed consent was obtained in accordance with the institutional review board, and a small cash payment was given in return for participation. Inclusion criteria for all participants included (a) normal or corrected normal (20/30) vision and (b) at least 5 years of formal education. In addition, participants were excluded for (a) non-English native language, (b) lifetime history of psychiatric disorders or substance dependence, on the basis of *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.; American Psychiatric Association, 1994) criteria, (c) evidence of dementia (on the basis of using *DSM-IV* criteria), or (d) history or evidence of any cognitive and/or neurological disorder or head trauma or other sensory, motor, or medical problems that could affect cognition or performance. There were no known differences between the older adults recruited from the three different sites in terms of cognitive or physical background, other than that the older adults at the day centre and residential home (but not at Birkbeck College) were provided with meals and recreational activities in a protective setting. All participants were ambulatory, except for two older adults at the residential home who used walking aids. Focused contrasts indicated that the young and older adults did not differ on educational level. The demographic characteristics of both participant groups are shown in Table 1.

Tests

Each participant was tested individually and performed four tests, given in the same order—namely, digit span, conceptual span (Haarmann et al., 2003), AX-CPT, and anomaly judgement. A test

TABLE 1
Demographic characteristics

	Young adults			*Older adults*		
Characteristic	*M*	*SD*	*%*	*M*	*SD*	*%*
Age[a]	26.8	3.9		74.8	5.7	
Education[a]	14.84	1.6		14.36	2.0	
% female			67			64
% right-handed			91			88

[a]In years.

session lasted about 40 minutes. Presentation of all tests was visual and controlled by a laptop computer. In the digit and conceptual span tests, the experimenter recorded participants' spoken responses manually using a score sheet. In the AX-CPT and anomaly judgement tasks, the computer recorded both response choice and response time with 1-ms accuracy, using a button box with designated yes and no buttons.

Digit span. On each trial, participants silently read a random sequence of digits (1 to 9), presented at a computer-controlled rate of one digit/s. Immediately after the offset of the last digit, a question mark appeared, and participants attempted to recall aloud all digits in the set in their order of presentation. Testing started at a set size of two digits. There were two trials per set size. The set size was increased with one digit if there was at least one correct response (i.e., all words recalled in correct order). A participant's digit span was defined as the largest set size at which a correct response was obtained.

Conceptual span. On each trial, participants silently read a sequence of nine nouns (in small letters) followed by a category name (in capital letters), presented at a computer-controlled rate of 1 word/s. The nine words consisted of three groups of three nouns, with each group belonging to a different semantic category, and were presented in a random order. Participants were instructed to try to recall aloud the three nouns in the named category in any order (e.g., *lamp, pear, tiger, apple, grape, elephant, horse, fax, phone, FRUIT?* Answer: *pear grape apple*). Categories and words were sampled at random from six semantic categories with eight nouns each. There were 2 practice trials and 10 test trials. A participant's conceptual span was defined as the number of words recalled across the 10 test trials (the maximum possible score was 30).

AX-CPT. The letter sequences used for the AX-CPT task comprised a three-by-two design, crossing the factors trial type (AX, BX, or AY) and delay (short or long). Each letter sequence consisted of a cue letter (the context), several distractor letters, and a probe letter. The cue A and probe X refer to the actual letters A and X, while the cue B and probe Y refer to any other randomly selected letter in the alphabet (except "K", since it is visually similar to X and Y). In the short and long delay condition, three and six randomly selected distractor letters (different from the cue and the probe) intervened between the cue and probe letter, respectively. AX, BX, and AY sequences occurred with 70, 15, and 15%, respectively, half of them occurring in the short delay condition and the other half in the long delay condition. A total of 120 letter sequences (trials) were presented in random order across two blocks with a brief break in between. There were six practice trials. On each trial, a fixation cross appeared for 1,000 ms, followed by the presentation of the cue, the distractors, and the probe, one at a time at 550 ms per letter. All stimuli were shown centrally on a white background in 65-point uppercase Lucida font.

The cue letter and the distractor letters were shown in red, whereas the probe letter was shown in black. Participants were instructed to say out loud the cue letter and the distractor letters as they saw each of them, to ensure encoding and prevent rehearsal of the cue. They were instructed to otherwise ignore the distractor letters when monitoring for targets. Participants were further instructed to respond to the probe letter by pressing a yes button on AX trials and a no button otherwise, as quickly and accurately as possible. The assignment of buttons (left vs. right button) to answers ("yes" versus "no" answer) was counterbalanced across the participants in each age group. To prevent ceiling level performance, participants had to respond within 1,000 ms from the onset of the probe letter. Responses that were slower than this limit were recorded as incorrect. The computer recorded both response choice and response time with 1-ms accuracy. The response initiated the next trial.

Anomaly judgement. This test came from a previous study (see Haarmann et al., 2003, Exp. 2, for details). This test consisted of 68 sentences. They were presented in a random order, one word at a time at a base rate of 450 ms (plus 30 ms for every additional letter in a word). Half of the sentences were semantically sensible, and half were semantically anomalous, including either an absurd (i.e., semantically anomalous) adjective–noun combination or an absurd noun–verb combination, but not both (e.g., "*The boys admired the* curly, *new* car *of the secretary in the office*", "*He* lifted *the bright* sun *outside the factory*, underlining shown for illustrative purposes only). Participants silently read each sentence word by word and had to press a no button as soon as the sentence stopped making sense and to press a yes button if at the end of the sentence it turned out the sentence was sensible. There was a 1.5-s response deadline to prevent ceiling level performance. Participants received feedback to indicate whether or not their response was correct and whether or not it occurred within the deadline. The computer recorded both response choice and response time with 1-ms accuracy.

Data analysis

Average error rates (misses plus false alarms) and RTs (for correct trials only) were calculated per subject per condition in the AX-CPT task. They were then entered as data points into mixed-factor analyses of variance (ANOVA) with age (young or old) as a between-subjects factor and delay (short or long) as a within-subjects factor. Error type (BX or AY) was an additional within-subjects factor in the ANOVAs of nontarget trials (Braver et al., 2001). Separate ANOVAs were carried out for target (i.e., AX) and nontarget trials (i.e., BX and AY trials), because of their different response requirements (target button vs. nontarget button press) and their different frequencies of occurrence (i. e., 70% for AX trials, 15% for each of the nontarget trials (Braver et al., 2001; Servan-Schreiber et al., 1996). Significant interactions were broken down into lower order interactions and/or simple main effects with further ANOVAs. ANOVAs were also calculated for each of the other measures (i.e., conceptual span, digit span, and anomaly judgement accuracy) separately, with age (young or old) as a between-subjects factor. For purposes of correlation and regression analyses, a single variable, "context maintenance", was calculated as an aggregate measure of the delay effect on the error rate across the three trial types, as follows: $(\text{AX short} - \text{AX long}) + (\text{BX short} - \text{BX long}) + (\text{AY long} - \text{AY short})$, where AX, BX, and AY denote error percentages on the corresponding trial types, and short and long refer to the short and long delay condition. A lower score on this variable indicates greater problems with context maintenance, since such problems are expected to increase AX and/or BX errors and to decrease AY errors in the long condition relative to the short condition (Braver et al., 2001). First-order and partial product moment correlations were calculated among age, context maintenance, conceptual span, digit span, and anomaly judgement. The partial correlation analyses were augmented with several stepwise hierarchical regressions. In anomaly judgement, older adults showed a much higher percentage of trials on which the 1,500-ms response deadline was surpassed than did young adults ($M = 59\%$ old adults vs. $M = 16\%$ young adults). We therefore calculated two measures of anomaly judgement—Anomaly Judgement 1 and Anomaly Judgement 2—which did and did not, respectively, take the deadline into account, in order to assess its effect on the results. Anomaly Judgement 1 took the deadline into account by scoring only responses prior to the deadline as correct (1) or incorrect (0), while scoring correct and incorrect responses past the deadline as if they were "don't know" responses (.5). Several participants showed a tendency to be over-rejecting in the anomaly judgement task. To correct for response bias, A' was used as an accuracy measure in this task. A' provides an unbiased estimate of the proportion correct in a two-alternative forced-choice procedure (Pollack & Norman, 1964) and was calculated as $A' = 0.5 + (y - x)(1 + y - x)/4y(1 - x)$, where x = false alarm rate, and y = hit rate (Grier, 1971; Linebarger, Schwartz, & Saffran, 1983).

Results

AX-CPT accuracy

Figure 1 shows the average performance accuracy in the AX-CPT task as a function of age, delay, and error type. The analysis of error rates on target trials (AX) revealed main effects of age, $F(1, 77) = 138.42, p < .001$, and delay, $F(1, 77) = 69.36, p < .001$, and an interaction of age and delay, $F(1, 77) = 25.72$, $p < .001$, due to an increase in AX errors with increased age and delay and due to a larger delay effect for older adults, $F(1, 35) = 70.21, p < .001$, than for young adults, $F(1, 42) = 6.91$, $p < .05$. However, when expressed in relative rather than absolute terms, young and older adults showed almost equal delay effects (50% and 45% more AX

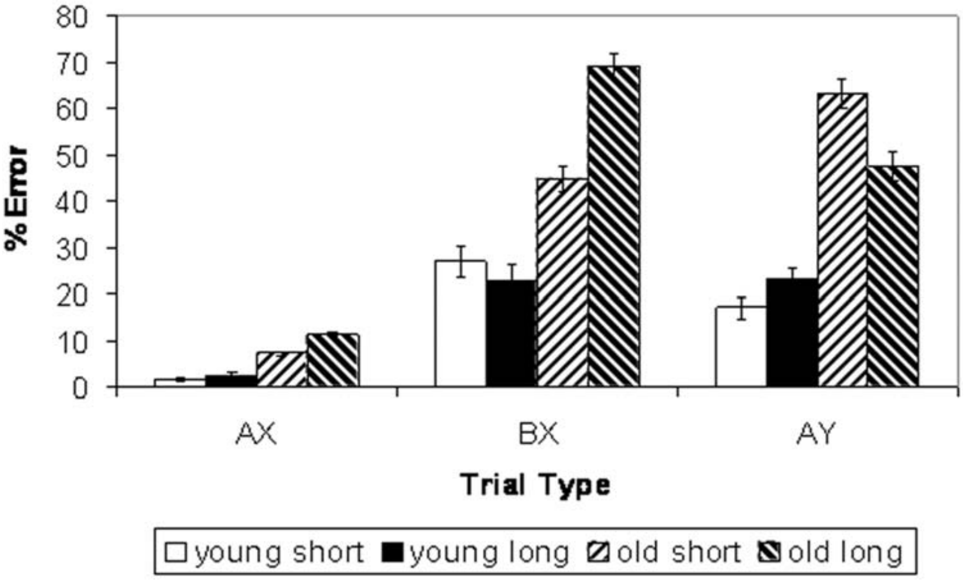

Figure 1. Percentage of error in the AX-CPT task as a function of trial type, age, and delay. Error bars depict standard errors of the means. Old adults made more AX, BX, and AY errors than did young adults. Old adults made more AX and BX errors but fewer AY errors in the long than in the short delay condition, suggesting a problem with context maintenance.

errors, respectively[2]). The analysis of error rates on nontarget trials (BX and AY) strongly suggested the presence of a context maintenance deficit in old but not young adults. There was an interaction of delay and error type, $F(1, 77) = 22.43$, $p < .001$, and an interaction of age, delay, and error type, $F(1, 77) = 64.37$, $p < .001$, due to different effects of age and delay on the BX versus AY errors. For BX errors, there were main effects of age, $F(1, 77) = 50.82$, $p < .001$, and delay, $F(1, 77) = 28.20$, $p < .001$, and an interaction of age and delay, $F(1, 77) = 55.93$, $p < .001$, due to the fact that older adults produced more BX errors than did young adults and due to the fact that older adults produced more BX errors in the long than in the short delay condition, $F(1, 35) = 76.18$, $p < .001$, whereas young adults did not. Young adults showed more BX errors in the short than in the long delay condition, but this effect fell short of significance, $F(1, 42) = 2.55$, $p < .12$. For AY errors, there were main effects of age, $F(1, 77) = 149.94$, $p < .001$, and delay, $F(1, 77) = 5.02$, $p < .05$, and an interaction of age and delay, $F(1, 77) = 26.58$, $p < .001$, due to the fact that older adults produced more AY errors than did young adults and due to opposite effects of delay in young versus older adults. Whereas AY errors increased with delay in young adults, $F(1, 42) = 4.90$, $p < .05$, they decreased with delay in old adults, $F(1, 35) = 23.68$, $p < .001$. The pattern of more BX errors and fewer AY errors in the long than in the short condition obtained for older adults suggests that they have a deficit in context maintenance. Such a deficit was also indicated by an age effect on d' context[3], $F(1, 77) = 73.54$, $p < .001$, which was larger in the long than in the short delay condition, $F(1, 77) = 42.52$, $p < .001$, replicating Braver et al. (2001).

AX-CPT response times

Table 2 lists the average RTs and standard errors in the AX-CPT task as a function of age, delay, and error type. The analysis of RTs on target trials (AX) revealed a main effect of age, $F(1, 77) = 274$, $p < .001$, due to an age-related slowing in response times. The analysis of RTs on nontarget trials (BX and AY) revealed a main effect of age, $F(1, 77) = 275$, $p < .001$, an interaction between error type and age, $\mathrm{F}(1, 77) = 5.25$, $p < .05$, and an interaction among error type, age, and delay, $F(1, 77) = 5.58$, $p < .05$, reflecting the following pattern of results. For RTs on AY trials, there was a main effect of age, $F(1, 77) = 265$, $p < .001$, due to slower responses for the older group. For RTs on BX trials, there was not only a main effect of age, $F(1, 77) = 238$, $p < .001$, but also an interaction of age and delay, $F(1, 77) = 4.60$, $p < .01$, due to a delay effect in older, $F(1, 35) = 6.33$, $p < .05$, but not young adults, $F < 1$. Older adults responded more slowly on BX trials in the long than in the short delay condition.

Age effects on other tasks

Table 3 shows descriptive statistics for context maintenance in the AX-CPT task, conceptual span, digit span, and anomaly judgement accuracy for young and old adults. Ageing effects were found for all task measures except digit span. Compared to young adults, older adults showed

---

[2] The relative percentage change in error rate was calculated as (long − short)/((long + short)/2), where short and long designate error rates in the short and long delay conditions, respectively.

[3] The d' context provides a focused and unbiased measure of context sensitivity, which is based on AX hits and BX false alarms and corrects for perfect hit rates (1.0) and false alarms (0.0) with a correction factor (i.e., hits = $2^{-(1/N)}$ and false alarms = $1 - 2^{-(1/N)}$, where N equals the number of target and nontarget trials, respectively; Braver et al., 2001).

TABLE 2
AX-CPT response times[a]

		Young		Old	
Trial type	*Delay*	*M*	*SE*	*M*	*SE*
AX	Short	277	17	609	17
	Long	277	12	623	16
BX	Short	295	15	611	16
	Long	282	15	636	19
AY	Short	309	14	629	18
	Long	312	14	605	14

[a] In ms.

TABLE 3
Descriptive statistics of all task measures

	Age group					
	Young[a]			Old[b]		
Measure	*M*	*SD*	*Range*	*M*	*SD*	*Range*
Context maintenance[c]	9.3	25.8	−46.8–55.6	−44.2	30.3	−93.7–26.2
Conceptual span	20.02	2.99	14–25	11.22	2.19	7–17
Digit span	4.4	1.29	2–7	4.14	1.15	2–6
Anomaly Judgement 1[d]	.91	.09	.62–1	.69	.16	.38–.96
Anomaly Judgement 2[e]	.95	.06	.71–1	.81	0.13	.45–.98

[a] $N = 43$. [b] $N = 36$. [c] Context maintenance = (AX short − AX long) + (BX short − BX long) + (AY long − AY short), where AX, BX, and AY denote the error percentages in the corresponding trial types, and short and long refer to the length of the delay in the AX-CPT task. [d] A prime measure, which scored only responses within the deadline as correct (0) or incorrect (1) and treated responses beyond it as "don't know" answers (.5). [e] A prime measure, which ignored whether or not the responses were made within the deadline.

poorer context maintenance, $F(1, 77) = 71.99$, $p < .001$, conceptual span, $F(1, 77) = 214.36$, $p < .001$, and anomaly judgement, $F(1, 77) = 65.34$, $p < .001$ (Anomaly Judgement 1), $F(1, 77) = 39.81$, $p < .001$ (Anomaly Judgement 2). However, older adults did not differ from young adults on digit span, $F < 1$. In anomaly judgement, ageing had a larger negative impact when the response deadline was imposed on scoring (see Table 3), and ageing had a negative impact not only on performance accuracy, but also on RT (for correct trials only), $F(1, 77) = 55.02$, $p < .001$, excluding a speed–accuracy trade-off explanation of the effect.

Correlation and regression analyses

Table 4 lists the size and *p* values of the product moment correlations among context maintenance, conceptual span, digit span, anomaly judgement, and years of age. Significant correlations of a moderate to large size were found among context maintenance, conceptual span, anomaly judgement, and years of age, whereas digit span did not correlate with any of

TABLE 4
Correlations among task measures and age

			Anomaly Judgement[b]		
Measure	*Conceptual span*	*Digit span*	*1*	*2*	*Age*
Context maintenance	.60**	.07	.42**	.39**	−.67**
Partial corr.[a]	.03, age				−.38**, conceptual span
Conceptual span	–	−.06	.61**	.58**	−.87**
Partial corr.[a]					−.79**, context maintenance
Digit span	–	–	.03	−.02	−.08
Anomaly Judgement 1	–	–	–	.86**	−.67**
Anomaly Judgement 2	–	–	–	–	−.59**

[a] Partial correlations among context maintenance, conceptual span, and age, and the variable that is controlled for.
[b] See the data analysis section or Table 3 for an explanation of the difference between Anomaly Judgements 1 and 2.
$*p < .05$, $**p < .01$.

these three measures. The correlation between conceptual span and anomaly judgement ($r = .58$) was of similar magnitude ($z = .58$, $p = .56$, $N_1 = 79$, $N_2 = 64$, two-tailed)[4] as the one obtained in a previous study with young adults ($r = .51$; Haarmann et al., 2003). The correlations involving the Anomaly Judgements 1 and 2 measures were of similar magnitude, with one exception. Age correlated better, $t(76) = 1.77$, $p = .04$, one-tailed, with Anomaly Judgement 1 than with Anomaly Judgement 2. This difference in the size of the correlation with age was to be expected given that older adults had greater difficulty in responding within the deadline (41% of all trials) than young adults (84% of all trials) due to their age-related slowing and given that the Anomaly Judgement 1 but not Anomaly Judgement 2 measure took the deadline into account.

Table 4 also lists the partial correlations (a) between context maintenance and conceptual span, controlling for age, (b) between age and context maintenance, controlling for conceptual span, and (c) between age and conceptual span, controlling for context maintenance. The correlation between context maintenance and conceptual span was abolished when age was controlled for, consistent with the notion of age-related effects upon a system that supports common aspects of context maintenance and semantic STM. However, there appeared to be additional age-related effects upon context maintenance above and beyond those shared with semantic STM, suggested by the finding that the correlation between age and context maintenance was reduced but not abolished when conceptual span was controlled

---

[4] The test for the significance of the difference between independent correlations (i.e., comparing correlations obtained from two samples) was computed with the indepcor.exe program, accompanying an article by Crawford, Mychalkiw, Johnson, and Moore (1996). This program implements Howell's (1997) procedures for such a test, following which both correlations are first converted to Fisher's z', and the difference between them is divided by the standard error of the difference to yield a normal curve deviate (z).

for. To further test this, context maintenance was regressed onto conceptual span (entered first) and age (entered second). Conceptual span accounted for 27% of the variance in context maintenance, $R^2 = 27$, $F(1, 77) = 28.52$, $p < .001$, which was entirely shared with age. Age added another 21.5% of unique variance, R^2 *change* $= 21.5$, $F(1, 76) = 26.99$, $p < .001$. Together, conceptual span and age accounted for 48.5% of the variance in context maintenance, $R^2 = 48.5$, $F(2, 76) = 35.85$, $p < .001$. The correlation between age and conceptual span was somewhat reduced but remained highly significant, and age still explained 62% of the variance on conceptual span when context maintenance was controlled for, consistent with the existence of age-related effects that are unique to semantic STM.

Age-related effects upon anomaly judgement may be associated with common effects on semantic STM and context maintenance, with additional unique effects of context maintenance, and with still further unique effects of age. To test this, anomaly judgement (Anomaly Judgement 2) was regressed onto conceptual span (entered first), context maintenance (entered second), and age (entered third). Conceptual span accounted for 17% of the variance in anomaly judgement, $R^2 = 17$, $F(1, 77) = 17.09$, $p < .001$, which was entirely shared with both context maintenance and age. While context maintenance did not add any further significant variance over and beyond conceptual span, R^2 *change* $= 0.02$, $F(1, 76) = 2.09$, $p = .15$, age added another 17% of variance, R^2 *change* $= 17$, $F(1, 75) = 20.26$, $p < .001$. Together, conceptual span, context maintenance, and age accounted for 34% of the variance in anomaly judgement, $R^2 = 34$, $F(3, 75) = 14.87$, $p < .001$. An identical regression analysis that used the Anomaly Judgement 1 instead of Anomaly Judgement 2 measure yielded analogous results, supporting the same conclusions. These results are consistent with the notion that anomaly judgement is affected by a factor that accounts for shared variance among ageing, semantic STM, and context maintenance and by one or more additional age-related factors.

Discussion

The main focus of the AX-CPT analysis is the error rates, which already include a temporal deadline (the response-time analysis mirrors the error rates, despite the response deadline that may diminish the range of the effects). The different error patterns of young and older adults on the AX-CPT task provide further support for the proposal of an age-related deficit in the maintenance of context information in working memory (Braver et al., 2001). On nontarget trials, young adults produced an equal amount of BX errors in the short and long interference condition and more AY errors in the long than in the short interference condition. This error pattern indicates that the added delay and interference in the long interference condition did not affect the ability of young adults to maintain the context letter (i.e., the cue). Moreover, the longer maintenance of the cue may have increased the a priori bias towards a target response on AY trials, explaining why AY errors increased with delay in young adults. By contrast, older adults produced more BX errors and fewer AY errors in the long than in the short interference condition. This error pattern indicates that the added delay and/or interference had a negative effect on the ability of older adults to maintain the context letter in working memory. On BX trials, a loss of context information across the delay makes it more difficult to use the invalid cue to inhibit the prepotent target response, explaining the increase in BX errors with delay. On AY trials, a loss of context information across the delay reduces the ability of the valid cue to create an inadvertent bias towards the prepotent target response, explaining the decrease in AY errors with delay.

These error patterns are consistent with the theory of Braver et al. (2001) for an age-related context maintenance deficit. As expected (since responses were scored under a deadline), older people made more errors than did the younger ones in all conditions. As predicted, however, the increase in interference led to an increase in AY errors rate for the young subjects but to a decrease for the older ones. This data pattern was not obtained in Braver et al., which focused on between-group effects. A number of procedural differences can explain the subtle variation in the results. First, while in the Braver et al. study, the difficulty in context maintenance was manipulated by adding interference within an unfilled cue–probe delay interval, in our study we lengthened the duration of the filled cue–probe delay interval. Second, while in Braver et al.'s study, the trials in the easy and difficult context maintenance condition were presented in separate blocks, we intermixed the trials, reducing the likelihood that effects of task condition reflect strategy variations (e.g., difference in the encoding of the context cue or increased attention). Third, we used a shorter response deadline, and we did not provide response feedback. This is likely to explain why older adults in our study had a larger overall error rate. Despite these subtle differences, we believe that our results provide complementary data supporting the context theory of ageing, by demonstrating within-group effects of interference predicted by this theory.

A prominent theory of cognitive ageing is that it involves a global decline in processing speed (Myerson et al., 1990; Salthouse, 1996). While this theory explains an important aspect of the reaction time data in the AX-CPT task (i.e., the overall slower response times in older adults), it is insufficient to explain all age effects on reaction time in the AX-CPT task, in particular on AY trials (Braver et al., 2001). The processing speed theory of ageing predicts that the slowest condition in young adults will show the largest age effect on reaction time (Brinley, 1965; Cerella, 1985; Myerson et al., 1990) and that reaction times in all conditions correlate positively with age. However, Braver et al. (2001) found that AY trials, while being the slowest condition in young adults, did not show the largest age effect on reaction time, and that AY reaction times correlated negatively with age, after factoring out a global decline in processing speed. Furthermore, the processing speed theory of ageing predicts that factors affecting processing difficulty, such as interference, should have effects that are in the same direction in young and older adults. However, this was not the case for AY trials, where young adults had more difficulty in the long than in the short interference condition, whereas the reverse was true for older adults.

In addition to the evidence for an age-related decline in context maintenance, the span and anomaly judgement data suggest that there may be an age-related decline in semantic STM. The results for the two span tasks revealed negative effects of ageing on conceptual span but not on digit span. Conceptual span relies to a large extent on the contribution of semantic STM (Haarmann et al., 2003). In contrast, digit span may primarily rely on phonological STM, due to the relatively shallow meaning of digits and due to the fact that serial recall of digits relies on their sequential rehearsal in the phonological loop (Baddeley, 1986). This is consistent with neuropsychological data showing that patients with phonological STM deficits are impaired in this test (e.g., patient E.A. above or K.F. reported by Warrington & Shallice, 1969). Moreover, the digit span is less likely to be contaminated by semantic/visual strategies, which affect word span tests (Logie et al., 1996).[5] Despite the procedural differences (cued vs. serial recall), the span results may therefore suggest that there is an age-related deficit in semantic STM but not in phonological STM. Such a selective deficit is consistent with the semantic anomaly judgement findings. Semantic anomaly judgement is dependent on semantic STM for the active maintenance of word meanings in order to support their on-line integration and to allow detection of semantic errors (Haarmann et al., 2003; Hanten & Martin, 2000; Martin & Romani, 1994). In line with this task analysis, we found a positive correlation between anomaly judgement and conceptual (but not digit) span and age-related performance declines on both tasks.

Our claim that the age-related decline in conceptual span reflects a semantic STM impairment can only be accepted tentatively at the present time. The present study used a nonclustered version of conceptual span, where words from different semantic categories are randomly mixed in a memory list and thus need to be clustered into their respective category to support category-cued recall. As a result, the nonclustered version of the conceptual span may measure not only semantic STM, but also clustering ability, which may engage cognitive functions other than semantic STM. It would, therefore, be of much relevance to determine whether there is an age-related effect on a clustered version of conceptual span, where words are grouped by their category, and clustering ability is less required. We predict this to be the case, however, on the basis of a correlational study where we find that both the clustered and the nonclustered versions of the conceptual span show high correlations with reading comprehension and problem solving in young adults (Haarmann et al., 2003). Finally, it is possible that (although constructed as a storage-only test) the conceptual span may still capture individual differences in the ability to control attention (Kane, Bleckley, Conway, & Engle, 2001; Kane & Engle, 2000). It is currently an open question as to what extent the construct of semantic STM and the ability to control attention rely on a common underlying cognitive mechanism.

In addition to investigating whether there is evidence for age-related deficits in context maintenance and semantic STM, this study sought to investigate whether there is a relation between these two cognitive functions. Consistent with the existence of such a relation, there was a positive correlation of context maintenance with conceptual span and with anomaly judgement, while neither of these measures correlated with digit span. From a theoretical perspective, two hypotheses regarding the relation between cognitive decline in context maintenance and in semantic STM can be contrasted. First, it is possible that in contextual tasks such as the AX-CPT, task-relevant representations (e.g., A preceding an X) are created on-line (and from scratch) for the task and are utilized to bias information processing. The brain system involved in this function is, according to this hypothesis, distinct from the system involved in active maintenance of semantic information. The two systems may, however, be located within adjacent areas of the frontal cortex (Braver & Bongiolatti, 2002) and may both therefore to some degree be affected by ageing, explaining the source of the age-mediated correlation between context maintenance (in AX-CPT) and semantic STM (in conceptual span). The second hypothesis is that context representations are not built for each task from scratch but are rather piggybacking on top of system engaged in the maintenance of lexical-semantic information (Gabrielli, Poldrack, & Desmond, 1998). Further (behavioural, neuropsychological, and imaging) studies, which explicitly contrast delayed semantic judgements and contextual word meaning disambiguation, are necessary to determine whether

---

[5] Most of the subjects reported using a covert rehearsal strategy in the digit span test and only a very small minority of subjects (< 5%) reported using a visual strategy.

and how semantic STM and context maintenance are related. **Q2**

While ageing may affect overlapping or separate mechanisms for context maintenance and semantic STM, our data suggest that ageing has additional effects as well. The partial correlations and regression analyses indicated that ageing accounts for variance in context maintenance, conceptual span, and anomaly judgement beyond the age-related variance that is shared by those measures. This may indicate that ageing affects not only a common mechanism that contributes to all three measures but also additional, task-specific mechanisms. Based on the present data, we can only speculate as to what those additional ageing effects may be. A deficit in the ability to update context information in working memory may interfere more with its activation in the long than in the short condition (due to the presence of more distractor letters in the long condition) and thus account for additional ageing effects on the context maintenance measure. A deficit in the ability to cluster word meanings belonging to a cued category may account for additional ageing effects on conceptual span, while a deficit in the ability to integrate word meanings into a sentence representation may account for additional ageing effects on anomaly judgement.

The present findings suggest several theoretically motivated directions for future research. To obtain further evidence for a relation between context maintenance and semantic STM, one could use a semantic disambiguation task and vary the interference (i.e., number of intervening words) between the context word (e.g., river) and the ambiguous target word (e.g., bank). A similar interference manipulation may be also applied in a delayed semantic judgement task. The hypothesis that there is overlap in the mechanisms for context maintenance and semantic STM predicts a strong association between the effect of interference in these tasks. Further research should also contrast indices of semantic and phonological STM that are more similar in (number of) items and response requirements. This could be done, for example, by contrasting cued recall with a semantic category versus a rhyme cue.

In conclusion, the results of this study suggest that the age-related deficit in the processing of information about task context (Braver et al., 2001) is in part due to a problem in the maintenance of that information in working memory and that there is an age-related decline in semantic but not phonological STM. The results of this study further suggest that there might be a relation between context maintenance and semantic STM, such that context maintenance depends on semantic STM. Additional studies are needed to further evaluate this possibility and differentiate it more conclusively from a model in which context maintenance and semantic STM are supported by separate cognitive mechanisms. These studies should utilize behavioural experiments, complex correlational analysis, and functional neuroimaging, as well as comparison of young and older adults. Such a use of convergent methodology will help characterize the interplay between context maintenance and semantic STM in controlling cognition and action in young and older adults.

REFERENCES

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: American Psychiatric Association.

Anderer, P., Pascual-Marqui, R. D., Semlitsch, H. V., & Saletu, B. (1998). Differential effects of normal aging on sources of standard N1, target N1 and target P300 auditory event-related brain potentials revealed by low resolution electromagnetic tomography (LORETA). *Electroencephalography and Clinical Neurophysiology*, *108*, 160–174.

Baddeley, A. D. (1986). *Working memory* (Oxford Psychology series No. 11). Oxford, UK: Clarendon Press.

Baddeley, A. D. (1992). Is working memory working? The fifteenth Bartlett lecture. *Quarterly Journal of Experimental Psychology, 44A,* 1–31. Q3

Barch, D. M., Braver, T. S., Racine, C. A., & Satpute, A. B. (2001). Cognitive control deficits in healthy aging: Neuroimaging investigations. *Neuroimage*, *13*, S1025. Q4

Bazin, N., Perruchet, P., Hardy-Bayle, M. C., & Feline, A. (2000). Context-dependent information processing in patients with schizophrenia. *Schizophrenia Research*, *45*, 93–101.

Braver, T. S., & Barch, D. M. (2002). A theory of cognitive control, aging cognition, and neuromodulation. *Neuroscience and Biobehavioral Reviews*, *26*(7), 809–817.

Braver, T. S., Barch, D. M., Keys, B. A., Carter, C. S., Cohen, J. D., Kaye, J. A., et al. (2001). Context processing in older adults: Evidence for a theory relating cognitive control to neurobiology in healthy aging. *Journal of Experimental Psychology: General*, *130*, 746–763.

Braver, T. S., & Bongiolatti, S. R. (2002). The role of frontopolar cortex in subgoal processing during working memory. *Neuroimage*, *15*, 523–536.

Brink, J. M., & McDowd, J. M. (1999). Aging and selective attention: An issue of complexity or multiple mechanisms. *Journal of Gerontology: Psychological Sciences*, *54*, 30–33.

Brinley, J. F. (1965). Cognitive sets, speed and accuracy of performance in the elderly. In A. T. Welford & J. E. Birrent (Eds.), *Behavior, aging, and the nervous system* (pp. 114–149). Springfield, IL: Charles C. Thomas.

Brouwer, H., Waterink, W., Van Wolffelaar, P. C., & Rothengatter, T. (1991). Divided attention in experienced young and older drivers: Lane tracking and visual analysis in a dynamic driving simulator. *Human Factors*, *33*, 573–582.

Burgess, N., & Hitch, G. J. (1999). Memory for serial order: A network model of the phonological loop and its timing. *Psychological Review*, *106*, 551–581.

Cerella, J. (1985). Information processing rates in the elderly. *Psychological Bulletin*, *98*, 67–83.

Cohen, J. D., Barch, D. M., Carter, C., & Servan-Schreiber, D. (1999). Context-processing deficits in schizophrenia: Converging evidence from three theoretically motivated cognitive tasks. *Journal of Abnormal Psychology*, *108*, 120–133.

Cohen, J. D., & Servan-Schreiber, D. (1992). Context, cortex, and dopamine: A connectionist approach to behavior and biology in schizophrenia. *Psychological Review*, *99*, 45–77.

Cowan, N. (1999). An Embedded-Processes Model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 62–101). Cambridge, UK: Cambridge University Press.

Craik, F. I. M. (1977). Age differences in human memory. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (pp. 384–420). New York: Van Nostrand Reinhold.

Craik, F. I. M., & Jennings, J. M. (1992). Human memory. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 51–110). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Craik, F. I. M., Morris, R. G., & Gick, M. (1990). Adult age differences in working memory. In G. Vallar & T. Shallice (Eds.), *Neuropsychological impairments of short-term memory* (pp. 247–267). Cambridge, UK: Cambridge University Press.

Crawford, J. R., Mychalkiw, B., Johnson, D. A., & Moore, J. W. (1996). WAIS-R short-forms: Criterion validity in healthy and clinical samples. *British Journal of Clinical Psychology*, *35*, 638–640.

Crosson, B., Rao, S. M., Woodley, S. J., Rosen, A. C., Bobholz, J. A., Mayer, A., et al. (1999). Mapping of semantic, phonological, and orthographic verbal working memory in normal adults with functional magnetic resonance imaging. *Neuropsychology*, *13*, 171–187. Q5

Daigneault, S., Braun, C. M. J., & Whitaker, H. A. (1992). Early effects of normal aging on perseverative and nonperseverative prefrontal measures. *Developmental Neuropsychology*, *8*, 99–114.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*, 450–466.

Dobbs, A. R., Aubrey, J. B., & Rule, B. G. (1989). Age-associated release from proactive inhibition: A review. *Canadian Psychologist*, *30*, 331–344.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory and general fluid intelligence: A latent variable approach. *Journal of Experimental Psychology: General*, *128*, 309–331.

Filley, C. M., & Cullum, C. M. (1994). Attention and vigilance functions in normal aging. *Applied Neuropsychology*, *1*, 29–32.

Fisk, J. E., & Warr, P. (1996). Age and working memory: The role of perceptual speed, the central executive, and the phonological loop. *Psychology and Aging*, *11*, 316–323.

Freedman, M. L., & Martin, R. C. (2001). Dissociable components of short-term memory and their relation to long-term learning. *Cognitive Neuropsychology*, *18*, 193–226.

Gabrieli, J. D. E., Poldrack, R. A., & Desmond, J. E. (1998). The role of left prefrontal cortex in language and memory. *Proceedings of the National Academy of Sciences*, *95*, 906–913. **Q6**

Gray, J. R., Chabris, C. F., & Braver, T. S. (2003). Neural mechanisms of general fluid intelligence. *Nature Neuroscience*, *6*, 316–322. **Q7**

Grier, J. B. (1971). Nonparametric indexes for sensitivity and bias: Computing formulas. *Psychological Bulletin*, *75*, 424–429.

Haarmann, H. J., Davelaar, E. J., & Usher, M. (2003). Individual differences in semantic short-term memory capacity and reading comprehension. *Journal of Memory and Language*, *48*, 320–345.

Haarmann, H. J., & Usher, M. (2001). Maintenance of semantic information in capacity-limited item short-term memory. *Psychonomic Bulletin & Review*, *8*(3), 568–578.

Hanten, G., & Martin, R. C. (2000). Contributions of phonological and semantic short-term memory to sentence processing: Evidence from two cases of closed head injury in children. *Journal of Memory and Language*, *43*, 335–361.

Hasher, L., Stoltzfus, E. R., Zacks, R. T., & Rypma, B. (1991). Age and inhibition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *17*, 163–169.

Hasher, L., & Zachs, R. T. (1988). Working memory, comprehension and aging: A review and a new view. In G. H. Bower (Ed.), *The psychology of learning and motivation XXII* (pp. 193–225). New York: Academic Press.

Howell, D. C. (1997). *Statistical methods for psychology* (4th ed.). Belmont CA: Duxbury Press.

Humes, L. E., Nelson, K. J., Pisoni, D. B., & Lively, S. E. (1993). Effects of age on serial recall of natural and synthetic speech. *Journal of Speech and Hearing Research*, *36*, 634–639.

Jensen, G. D., & Goldstein, L. (1991). A microcomputerized task assessment of cognitive change in normal elderly and young adults. *Experimental Aging Research*, *17*, 119–121.

Kane, M. J., Bleckley, M. K., Conway, A. R. A., & Engle, R. W. (2001). A controlled-attention view of working-memory capacity. *Journal of Experimental Psychology: General*, *130*, 169–183.

Kane, M. J., & Engle, R. W. (2000). Working-memory capacity, proactive interference, and divided attention: Limits on long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *26*, 336–358.

Keys, B. A., & White, D. A. (2000). Exploring the relationship between age, executive abilities, and psychomotor speed. *Journal of the International Neuropsychological Society*, *6*, 76–82.

Korteling, J. E. (1993). Effects of age and task similarity on dual-task performance. *Human Factors*, *35*, 99–113.

Light, L. L., & Anderson, P. A. (1985). Working-memory capacity, age, and memory for discourse. *Journals of Gerontology*, *40*, 737–747.

Linebarger, M. C., Schwartz, M., & Saffran, E. M. (1983). Sensitivity to grammatical structure in so-called agrammatic aphasics. *Cognition*, *13*, 361–392.

Logie, R. H., Della Sala, S., Laiacona, M., Chalmers, P., & Wynn, V. (1996). Group aggregates and individual reliability: The case of verbal short-term memory. *Memory and Cognition*, *24*, 305–321.

MacPherson, S. E., Phillips, L. H., & Della Sala, S. (2002). Age, executive function, and social decision making: A dorsolateral pre-frontal theory of cognitive aging. *Psychology and Aging*, *17*, 598–609.

Martin, R. C., & Romani, C. (1994). Verbal working memory and sentence comprehension: A multiple-components view. *Neuropsychology*, *9*(4), 506–523.

Martin, R. C., Shelton, J. R., & Yaffee, L. S. (1994). Language processing and working memory: Neuropsychological evidence for separate phonological and semantic capacities. *Journal of Memory and Language*, *33*, 83–111.

McDowd, J., & Oseas-Kreger, D. (1991). Aging, inhibitory processes, and negative priming. *Journals of Gerontology*, *46*, 340–345. **Q8**

Morris, R. G., Craik, F. I. M., & Gick, M. L. (1990). Age differences in working memory tasks: The role of secondary memory and the central executive system. *The Quarterly Journal of Experimental Psychology*, *42*, 67–86.

Moscovitch, M., & Winocur, G. (1992). The neuropsychology of memory and aging. In F. I. M. Craik & T. A. Salthouse (Eds.), *Handbook of aging: Cognition* (pp. 315–372). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Myerson, J., Hale, S., Wagstaff, D., Poon, L. W., & Smith, G. A. (1990). The information-loss model: A mathematical theory of age-related cognitive slowing. *Psychological Review*, *97*, 475–487.

Panek, P. E., Rush, M. C., & Slade, L. A. (1984). Locus of the age–Stroop interference relationship. *Journal of Genetic Psychology*, *145*, 209–216.

Parasuraman, R., & Nestor, P. G. (1991). Attention and driving skills in aging and Alzheimer's disease. *Human Factors*, *33*, 539–557.

Parkin, A. J., Walter, B. M., & Hunkin, N. M. (1995). Relationships between normal aging, frontal lobe function, and memory for temporal and spatial information. *Neuropsychology*, *9*, 304–312.

Perfect, T. (2003). Memory aging as frontal lobe dysfunction. In M. A. Conway (Ed.), *Cognitive models of memory* (pp. 315–339). Cambridge, MA: MIT Press.

Pollack, L., & Norman, D. A. (1964). A non-parametric analysis of recognition experiments. *Psychonomical Science*, *1*, 125–126.

Romani, C., & Martin, R. (1999). A deficit in the short-term retention of lexical-semantic information: Forgetting words but remembering a story. *Journal of Experimental Psychology: General*, *128*, 56–77.

Rosvold, H. E., Mirsky, A. F., Sarason, I., Bransome, E. D., & Beck, L. H. (1956). A continuous performance test of brain damage. *Journal of Consulting Psychology*, *20*, 343–350.

Salthouse, T. A. (1992). Working-memory mediation of adult age differences in integrative reasoning. *Memory and Cognition*, *20*, 413–423.

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, *103*, 403–428.

Servan-Schreiber, D., Cohen, J. D., & Steingard, S. (1996). Schizophrenic deficits in the processing of context. A test of a theoretical model. *Archives of General Psychiatry*, *53*, 1105–1112.

Spencer, W. D., & Raz, N. (1994). Remembering facts, source, and context: Can frontal dysfunction explain adult age differences? *Psychology and Aging*, *9*, 149–159.

Spieler, D. H., Balota, D. A., & Faust, M. E. (1996). Stroop performance in healthy younger and older adults and in individuals with dementia of the Alzheimer's type. *Journal of Experimental Psychology: Human perception and performance*, *22*, 461–479.

Swinney, D. A. (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Leaning and Verbal Behavior*, *18*, 645–659.

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, *28*, 127–154.

Verhaeghen, P., & Salthouse, T. A. (1997). Meta-analyses of age-cognition relations in adulthood: Estimates of linear and nonlinear age effects and structural models. *Psychological Bulletin*, *122*, 231–249.

Warrington, E. K., & Shallice, T. (1969). The selective impairment of auditory verbal short-term memory. *Brain*, *4*, 885–896.

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale–Revised*. New York: Psychological Corporation.

West, R., & Alain, C. (2000). Age-related decline in inhibitory control contributes to the increased Stroop effect observed in older adults. *Psychophysiology*, *37*, 179–189.

West, R., & Baylis, G. C. (1998). Effects of increased response dominance and contextual disintegration on the Stroop interference effect in older adults. *Psychology and Aging*, *13*, 206–217.

West, R., & Bell, M. A. (1997). Stroop color–word interference and electroencephalogram activation: Evidence for age-related decline in the anterior attentional system. *Neuropsychology*, *11*, 421–427.

Zacks, R., & Hasher, L. (1997). Cognitive gerontology and attentional inhibition: A reply to Burke and McDowd. *Journals of Gerontology: Series B. Psychological Sciences and Social Sciences*, *52*, 274–283.

Original manuscript received 5 August 2003
Accepted revision received 4 February 2004

PQJA
Queries
Haarmann, Ashling, Davelaar, & Usher

Q1 Just first author as corresponding author ok?
Q2 Gabrielli et al: Gabrieli in Refs.
Q3 Baddeley, 1986: Oxford ok as added?
Q4 Is S1025 the page number?
Q5 Crosson et al., 1999: text ref.?
Q6 Gabrieli or Gabrielli, Poldrack, & Desmond?
Q7 Gray, Chabris, & Braver, 2003: text ref.?
Q8 McDowd & Oseas-Kreger, 1991: text ref.?