EUROPEAN JOURNAL OF COGNITIVE PSYCHOLOGY
2009, 21 (2/3), 303–329

Psychology Press
Taylor & Francis Group

Memory accuracy in old age: Cognitive, metacognitive, and neurocognitive determinants

Ainat Pansky, Morris Goldsmith, and Asher Koriat
Department of Psychology, University of Haifa, Haifa, Israel

Shiri Pearlman-Avnion
Department of Education, Tel-Hai Academic College, Tel Hai, Israel

Age differences in memory accuracy were examined within a conceptual framework specifying the mediating role of metacognitive monitoring and control processes (Koriat & Goldsmith, 1996b). Replicating previous results, older adults showed poorer memory quantity and accuracy performance compared to young adults. Even when memory quantity performance was equated, by dividing the young adults' attention during encoding, the difference in memory accuracy was not eliminated. Examination of the underlying metacognitive processes revealed that the age-related reduction in memory accuracy stemmed partly from less effective memory monitoring, apparently the result of poorer encoding, and also from differences in two aspects of metacognitive control: (1) a more liberal report criterion—greater tendency to volunteer incorrect (and correct) answers, and (2) reduced control sensitivity—less reliance on subjective monitoring as a basis for responding. This latter control reduction was associated with lower neuropsychological measures of executive functioning, suggesting a decline in frontal-lobe efficiency.

Keywords: Memory accuracy; Metamemory; Ageing; Executive processes; Memory control.

Memory decline in old age is both ubiquitous and multifaceted. Much of the work on this topic has focused on the amount of information that can be recalled or recognised. In the last few years, however, accumulating evidence indicates an age-related decline in memory accuracy as well. Older adults have been found to be particularly vulnerable to memory slips and errors that can have serious consequences, such as taking the same medicine twice

Correspondence should be addressed to Ainat Pansky, Department of Psychology, University of Haifa, Haifa 31905, Israel. E-mail: pansky@research.haifa.ac.il

This research was supported by the Charles E. Smith Extended Grant on the Subject of Memory, awarded by the National Institute for Psychobiology in Israel to AK, MG, and AP.

http://www.psypress.com/ecp DOI: 10.1080/09541440802281183

(Koriat, Ben-Zur, & Sheffer, 1988) or being taken in by scams and frauds (Jacoby & Rhodes, 2006).

There are numerous references in the literature to possible links between the memory impairments associated with ageing and neuropsychological deficits in frontal-lobe functioning. Thus, the age-related pattern of memory errors that is characteristic of old age resembles the pattern found in patients suffering frontal lobe lesions (see Moscovitch & Winocur, 1992, 1995; Perfect, 1997, for reviews). This similarity suggests that one source of age-related deficits in memory accuracy may be a decline in frontal lobe efficiency. This idea is supported further by converging anatomical, physiological, and neurological evidence indicating pronounced frontal lobe deficits in old age (see Raz, 2000; Stuss, Craik, Sayer, Franchi, & Alexander, 1996; West, 1996; but see Greenwood, 2000).

Frontal-lobe functioning, in turn, has been broadly associated with so-called "executive functions". These functions are generally assumed to include the monitoring and control of one's own behaviour, the ability to suppress irrelevant information and to inhibit prepotent responses, the ability to shift between mental sets or tasks and to control attention, and the monitoring of working memory contents (see, e.g., Miyake et al., 2000; Shimamura, 2000). Executive functions are also assumed to include higher cognitive processes such as reasoning and planning (Waltz et al., 1999; West, 1996). Focusing specifically on memory processes, converging evidence suggests the frontal lobe is involved in strategic processes that support memory encoding, recovery, monitoring, and verification (see Moscovitch & Winocur, 2002).

Finally, some of the age-related impairments in memory functioning appear to be linked to executive functions. Thus, recent behavioural studies employing a wide range of memory paradigms have found superior memory performance for older adults achieving high scores on neuropsychological tests of executive functioning compared to older adults scoring poorly on such tests (e.g., Butler, McDaniel, Dornburg, Price, & Roediger, 2004; Chan & McDermott, 2007; Davidson & Glisky, 2002; Glisky, Polster, & Routhieaux, 1995; Glisky, Rubin, & Davidson, 2001; McDaniel, Glisky, Rubin, Guynn, & Routhieaux, 1999; Roediger & Geraci, 2007).

In this study, we aimed to shed further light on the links between memory accuracy deficits, neuropsychological deficits, and ageing, within a general theoretical model specifying the mediating role of metacognitive processes. That model was developed in previous work with young adults (Koriat & Goldsmith, 1996b) with the goal of clarifying the role of metacognitive processes in the strategic regulation of both memory quantity and memory accuracy performance.

Traditionally, memory research has focused on *memory quantity*—the amount of information that can be recalled or recognised after a retention interval. Hence, measures of memory performance have typically been calculated conditional on the input, by expressing the number of items recalled or recognised as the proportion or percentage of the total number of items presented. Such measures reflect the amount of presented or studied information that has been retained and is currently accessible. More recently, however increasing interest has been directed to *memory accuracy*—the extent to which the person's memory report accords with what has been presented. A very common and simple measure of memory accuracy is *output-bound proportion correct*: The number of correct items recalled or recognised is expressed as a proportion or percentage of the total number of items reported. Unlike the input-bound quantity measures, this type of accuracy measure is conditional on the output, reflecting the probability that a reported item is correct. Consider, for example, a participant who is presented with 30 words (items of information), and in a recall test reports 15 words (provides answers to 15 questions), 12 of which are correct and 3 are commission errors (wrong answers). Input-bound memory quantity performance in this case is 0.40 (12/30), that is, 40% of the input-studied items (questions) have been successfully recalled (answered). In contrast, output-bound memory accuracy is 0.80 (12/15). That is, 80% of the output-reported items (answers) are, in fact, correct. This latter measure uniquely reflects the *dependability* of the information that is reported—the degree to which each reported item can be trusted to be correct. Essentially, then, whereas the input-bound quantity measure holds the rememberer responsible for what he or she fails to report, the output-bound accuracy measure holds the person accountable only for what he or she does report.

Importantly, output-bound accuracy and input-bound quantity measures can be distinguished operationally only under conditions of *free report*, that is, only when the rememberer has the option of deciding which items of information to report and which to withhold. This condition is typical of most real-life memory situations. In contrast, under forced-report conditions, in which people are required to answer each and every question (as in standard forced-choice recognition tests), the input-bound (quantity) and output-bound (accuracy) memory measures are operationally equivalent. This is because the number of output items is the same as the number of input items (see Koriat & Goldsmith, 1994, 1996a).

The option of free report is essential when the focus is on output-bound memory accuracy. Just as an eyewitness cannot be expected to uphold the oath to tell "nothing but the truth" under forced-report conditions, neither does it make sense to hold participants accountable for the errors that they make under such conditions. Indeed, only under free-report conditions, when rememberers have the option to respond "don't know", can we assume that they are actually committed to the accuracy of their memory output.

THE STRATEGIC REGULATION OF MEMORY REPORTING: A METACOGNITIVE FRAMEWORK

The focus on memory accuracy in addition to memory quantity has led researchers to incorporate a wider range of phenomena and processes into the study of memory than has been done in traditional memory research (e.g., Koriat, Goldsmith, & Pansky, 2000). For example, a great deal of research in the area of metacognition has examined the processes by which people monitor the validity of their memories, as well as the accuracy of this monitoring (e.g., Koriat, 1993; Koriat, Lichtenstein, & Fischhoff, 1980; Schwartz, 1994). Other studies have examined the manner in which monitoring processes are actually used in controlling the process of remembering, and in regulating memory performance (e.g., Barnes, Nelson, Dunlosky, Mazzoni, & Narens, 1999; Goldsmith & Koriat, 1999, 2008; Higham, 2002, 2007; Koriat & Goldsmith, 1994, 1996b; Koriat, Goldsmith, & Halamish, 2008; Nelson & Narens, 1990). Most relevant to present purposes, the examination of monitoring and control processes operating during memory retrieval has provided important insights regarding memory deficits in old age (e.g., Henkel, Johnson, & de Leonardis, 1998; Jacoby, Bishara, Hessels, & Toth, 2005; Kelley & Sahakyan, 2003; Koutstaal, 2006; Rhodes & Kelley, 2005).

We now outline the conceptual framework proposed by Koriat and Goldsmith (1996b), which guided the present investigation. Their model of the strategic regulation of memory reporting allows the examination of memory quantity and memory accuracy performance within a common framework, focusing on the manner in which performance is mediated by metacognitive processes that operate during memory reporting. As shown in Figure 1, in addition to an unspecified retrieval (or ecphory, reconstruction, etc.) mechanism, the model includes a *monitoring* mechanism that is used to subjectively assess the correctness of potential memory responses, and a *control* mechanism that determines whether or not to volunteer the best available candidate answer (for similar models, see Barnes et al., 1999; Higham, 2002). The control mechanism operates by setting a report criterion on the monitoring output: The answer is volunteered if its assessed probability of being correct passes the criterion, but is withheld otherwise. The criterion is set on the basis of implicit or explicit payoffs, that is, the perceived gain for providing correct information relative to the cost of providing incorrect information.

Based on this simple model, the overall quality of free-report memory performance (quantity and accuracy) may be seen to depend on four components: (1) *overall retention*—the amount of correct information (i.e., the number of correct candidate answers) that can be retrieved; (2) *monitoring effectiveness*—the extent to which the subjective confidence

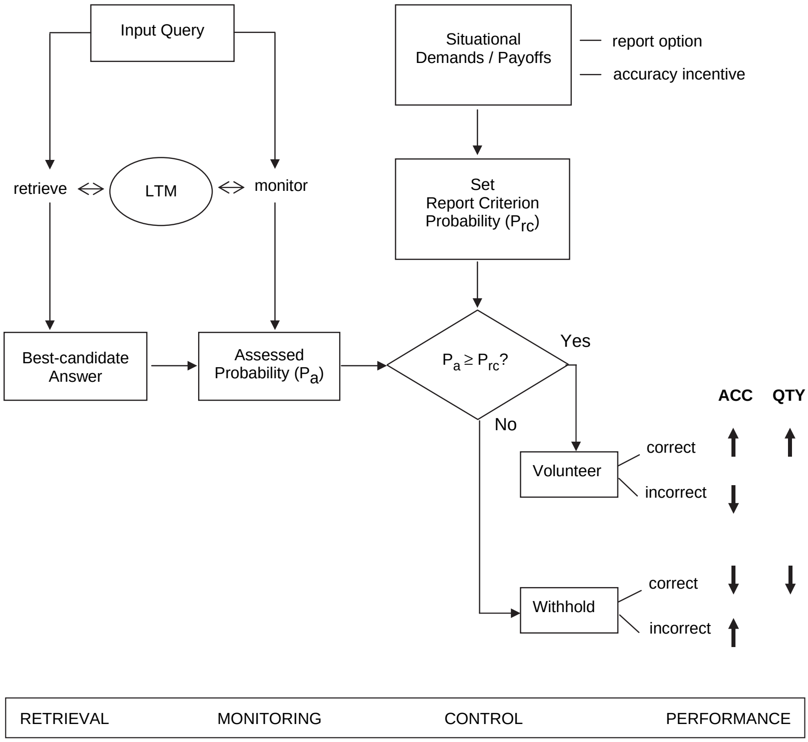

Figure 1. A schematic model of the strategic regulation of memory accuracy and memory quantity performance, utilising the option of free report. The upward and downward pointing arrows on the right of the figure signify positive and negative performance outcomes. LTM = long-term memory; ACC = accuracy; QTY = quantity; P_a = assessed probability; P_{rc} = response criterion probability (adapted from Koriat & Goldsmith, 1996b).

judgements successfully differentiate correct from incorrect candidate answers; (3) *control sensitivity*—the extent to which the volunteering or withholding of answers is in fact based on the monitoring output; (4) *report criterion setting*—the report criterion above which answers are volunteered, below which they are withheld.

Results from various studies have provided strong support for the model, revealing the manner in which monitoring and control processes mediate between memory retrieval on the one hand and memory performance on the other (Danion, Gokalsing, Robert, Massin-Krauss, & Bacon, 2001; Goldsmith, Koriat, & Pansky, 2005; Goldsmith, Koriat, & Weinberg-Eliezer, 2002; Higham, 2002; Higham & Tam, 2005; Kelley & Sahakyan, 2003; Koren et al., 2004; Koriat & Goldsmith, 1996b; Payne, Lambert, & Jacoby, 2002; Rhodes & Kelley, 2005; for a review see Goldsmith & Koriat, 2008).

Most important for the present study is the methodology that was developed to assess the contribution of each component to free-report memory performance (Koriat & Goldsmith, 1996b). This methodology, which was adapted here to the study of age-related deficits, will be described briefly. It is based on a two-phase, free-forced paradigm. Participants are exposed to a memory input stimulus and later given either a recall or recognition test using the following procedure:

1. *Forced-report phase*: A cue is presented for each studied item, and participants are required to provide an answer even if they have to guess. In addition, the participant also indicates his/her confidence regarding the correctness of each answer.
2. *Free-report phase*: The participant is presented with the same cues again (either the cues alone or the cue-answer pairs from the previous phase) under free-report instructions: He/she must indicate whether he/she would like to "volunteer" the answer or "withhold" it. Monetary or point incentives for accurate volunteering are offered, for example, a 1-point bonus for each correct volunteered answer but a 2-point penalty for each incorrect volunteered answer (withheld answers receive no bonus but also no penalty).

The measures that can be derived are as follows:

1. *Free-report accuracy* and *quantity* performance (output-bound and input-bound proportions correct, respectively), based only on the Phase-2 answers. These can be treated as measure of "actual" memory performance.
2. *Forced-report quantity correct*, based only on Phase-1 answers. This can be used as an estimate of how much information is accessible in memory ("retention" or "ecphory").
3. *Monitoring effectiveness*, based on the relationship between confidence judgements and the forced-report answers (Phase 1). Within-subject gamma correlations (Nelson, 1984) are generally used to measure *monitoring resolution*—the extent to which the confidence judgements distinguish between correct and incorrect answers. Calibration bias scores, reflecting over/underconfidence (i.e., the extent to which the confidence judgements are higher or lower than the actual proportions correct; see Lichtenstein, Fischhoff, & Phillips, 1982) may sometimes also be calculated.
4. *Control sensitivity*, the extent to which the free-report volunteering and withholding decisions in Phase 2 are based on the monitoring output (confidence ratings in Phase 1). This can be indexed by within-

participant gamma correlations between confidence (Phase 1) and volunteering-withholding decisions (Phase 2).

5. *Control policy (report criterion)*, the minimum level of confidence that is required by the participant before he/she is willing to volunteer an answer. This can be estimated by finding a cutoff on the confidence ratings that best separates the items that were volunteered and those that were withheld in Phase 2. Alternatively, Phase-2 volunteering or withholding ratios can be compared, while partialling out differences in forced-report quantity correct and/or differences in mean confidence ratings (cf. B''_D in Higham, 2002).

Neuropsychological evidence suggests that frontal lobes are at least partially involved in metamemory judgements (e.g., Janowsky, Shimamura, & Squire, 1989; Vilkki, Servo, & Surma-aho, 1998; see also Fernandez-Duque, Baird, & Posner, 2000). Assuming that old age is associated with a decline in frontal lobe functioning, then we may expect age-related declines in metamemory processes, but the evidence so far has been mixed and inconclusive (see Hertzog & Dunlosky, 2004; Hertzog & Hultsch, 2000; Souchay &Isingrini, 2004; Souchay, Isingrini, & Espagnet, 2000).

Recently, Kelley and colleagues (Kelley & Sahakyan, 2003; Rhodes & Kelley, 2005) utilised Koriat and Goldsmith's (1996b) model and methodology in conjunction with Kato's (1985) associative interference paradigm to compare the strategic regulatory processes of younger and older adults. Kelley and Sahakyan (2003, Exp. 1) found that for control word-pairs (not expected to elicit associative interference), forced-report performance was superior for younger than for older participants. However, the older participants were successful in taking advantage of the option to withhold answers to narrow the gap between their level of free report accuracy performance and that of the younger participants. In contrast, for "deceptive" word-pairs (in which the retrieval cues evoke a highly accessible associate that competes with the target, thereby presenting a difficult challenge to memory monitoring), the age difference in accuracy performance became, if anything, somewhat larger under free-report than under forced report. This interactive pattern was accounted for in terms of age-related differences in monitoring effectiveness. Indeed, although both older and younger participants were highly overconfident in the correctness of their responses to the deceptive items, the degree of overconfidence was more pronounced for the older participants. Furthermore, the older participants exhibited lower levels of monitoring resolution for both deceptive and control items.

Additional experiments suggested that the impaired monitoring of the older participants derived from impoverished encoding: When the encoding of the younger participants was disrupted by having them study the word list under divided attention, they exhibited a pattern of performance that was very similar to that of the older participants in terms of both memory accuracy and memory monitoring. Kelley and Sahakyan (2003) suggested that the poorer memory monitoring of older adults derives primarily from their greater reliance on the familiarity of candidate responses rather than on recollection of details of the study experience (Jacoby, 1999; Jacoby, Debner, & Hay, 2001), which in turn stems, at least in part, from poor encoding. A similar conclusion was reached by Rhodes and Kelley (2005), who used the same approach to investigate age differences in memory performance, but now tying these to neuropsychological measures of executive functioning. In their study, path analyses supported a model in which ageing impairs executive functioning, which in turn impairs retention (forced-report performance—a product of both encoding quality and retrieval), which in turn impairs free-report memory accuracy, both directly, and also indirectly by way of impaired monitoring. No age-related differences in control processes were found in either of these studies.

THE CURRENT STUDY

The current study also examined age-related differences in metacognitive and neurocognitive functioning that contribute to accurate memory performance. However, in contrast to Kelley and Sahakyan (2003) and Rhodes and Kelley (2005), we were primarily interested in potential age-related differences in metacognitive control. This focus was motivated by several results indicating systematic population differences in measures of control (see Goldsmith & Koriat, 2008; and see later). Consequently, rather than use a list-learning task with deceptive pairs, which placed a large burden on monitoring, we used a more naturalistic memory task, testing the memory for a simple slide show, without making any special attempt to mislead or hamper memory monitoring.

We followed the recommendation of previous researchers to facilitate the interpretation of age-related metamemory comparisons by equating the baseline quantity performance of the older and younger participants (see Perfect & Stollery, 1993). Because of the well-documented age-related decline in episodic memory quantity (Craik & Anderson, 1999), we included in our study a group of young participants who were presented with the stimulus materials under divided attention at study (e.g., Jennings & Jacoby, 1993; Kelley & Sahakyan, 2003). Thus, the experiment included three groups: young full-attention control, young divided-attention, and older adults. All participants were presented with a computerised slide show and accompanying narration describing an event in the life of a family. The young divided-attention pariticipants performed a concurrent auditory task while watching the show. Thirty questions about the show were then presented in multiple-choice recognition format. Report option was manipulated as follows: In the first phase, the participants were required to answer all 30 questions, and to indicate their confidence in the correctness of each answer. In the second phase, the participants were asked to take the same test again under free-report instructions, deciding which answers to volunteer and which to withhold, with a monetary incentive for accurate reporting. Note that, in contrast to Kelley and Sahakyan (2003) and Rhodes and Kelley (2005), who implemented the two phases on an item-by-item basis, we followed Koriat and Goldsmith's (1996b) procedure in which the confidence judgements and volunteering decisions are elicited in two separate phases of the experiment. Although the item-by-item method has its advantages (see Goldsmith & Koriat, 2008), we thought that the two-phase method might reduce possible demand characteristics, by which the expectation that one's volunteering decisions should be based on one's confidence judgements might be conveyed implicitly in the procedure itself, making it more difficult to find differences in control sensitivity.

In order to assess the role of executive functioning as a potential mediator of the age-related decline in memory accuracy, the older participants were administered two tests that are commonly used in neuropsychological research to evaluate executive functioning (see, e.g., Rhodes, 2004; Salthouse, Atkinson, & Berish, 2003).[1] The older adults were sampled across a wide range of ages, with the goal of achieving adequate variation on measures of executive functioning, metacognitive functioning, and free-report performance within this age group.

METHOD

Participants

The experiment included three groups of participants: (1) 20 older adults (aged 62 to 87, mean = 73, $SD = 7.11$) (2) 20 young adults (aged 17 to 26, mean = 23), assigned to the full-attention group, and (3) 20 young adults assigned to the divided-attention group (aged 20 to 28, mean = 22). These groups will be designated as older, Young-FA, and Young-DA, respectively. All the participants were healthy and had normal hearing and normal or corrected vision.

[1] We did not administer these neuropsychological tests to the young adults because, based on previous research (see Rhodes, 2004), we did not expect to find much meaningful variance for this population.

Materials

The participants were presented with a computerised slide show lasting about 4 min. It consisted of 27 colour photographs presented on a high-resolution (super VGA) screen controlled by an IBM-PC compatible computer. The pictures were presented at a rate of 8 s per picture, with about 0.5 s between pictures. Each picture was accompanied by a recorded narration spoken in Hebrew by a professional radio broadcaster, which was produced through speakers attached to the computer.

The show ("On the Way to the Picnic") depicts a staged incident in the life of a family preparing to go out on a picnic, in which the family cat climbed up an electricity pole and had to be enticed by various means to come down. The story begins with an introduction to each family member, which is followed by a sequence of events culminating in the successful rescue of the cat. For the Young-DA group, a pseudorandom sequence of 300 Hz and 600 Hz tones was presented concurrently with the slide show. The participants were asked to count how many tones of each type (low frequency, high frequency) were sounded.

A 30-item memory questionnaire concerning the slide show was developed in a five-alternative recognition format. The questions related both to aspects that were central to the story and to peripheral details (e.g., Why did the cat . . .? What was the colour of the mother's dress?). Most of the answers consisted of a single word or name. In the forced-report phase, each question was accompanied by a 5-point "thermometer" scale, ranging from "very cold" to "very hot", on which the participants were asked to indicate their confidence that their chosen answer was correct. Higher confidence was indicated by "hotter" ratings. We chose to use this rating scale, rather than subjective probability assessments of correctness used Koriat and Goldsmith's (1996b) study with young adults, because we wanted to obtain a "gut feeling" rather than an analytic judgement, and because we suspected that some of the older adults might encounter difficulties in assigning such probabilities. This decision precluded some of the standard measures associated with the Koriat and Goldsmith framework (e.g. calibration bias), and required a few modifications in the data analyses which will be described in the Results section.[2]

Two tests of executive function were used. The first test was a Hebrew version of the Verbal Fluency (FAS) test, including three speeded tasks (60 s each) requiring the listing of as many words as possible: (1) in the ANIMAL category, (2) in either the fruit or vegetable category, and (3) starting with the

[2] In light of the findings of two studies that were published after the current study was conducted (Kelley & Sahakyan, 2003; Rhodes & Kelley, 2005), perhaps we were overcautious regarding these considerations.

Hebrew letters SHIN or MEM. The total number of unique words produced served as the dependent measure. The second test was the Wisconsin Card Sorting Task (WCST; Hart, Kwentus, Wade, & Taylor, 1988). The two most commonly used scores—number of categories achieved and number of perseverative errors committed—served as dependent measures.

Procedure

After viewing the slide show, the participants' memory was tested in two phases. The first was a forced-report five-alternative multiple choice recognition test in which the participants were asked to answer each item (guessing if necessary) and to indicate their level of confidence in the correctness of their answer on the 5-point "thermometer" scale.

In the second, free-report phase, participants went over the answers they had provided in the previous phase, now without the associated confidence judgements (which were cut away), and were given the option to either volunteer or withhold each of the answers. This phase included a payoff schedule according to which participants were told that they would gain 10 bonus points for each correct answer volunteered, but would lose 10 points for each incorrect answer volunteered. However, they need not volunteer all the answers; for those answers withheld they would not gain any points but neither would they lose any points. The participants were told that when the study was over, they would receive a monetary prize proportional to the number of points each of them earned.

Finally, the older participants (only) were tested on the FAS and WCST.

RESULTS

Based on Koriat and Goldsmith's (1996b) model, we compared the three groups of participants on free-report memory performance, both accuracy and quantity, and on four component measures: (1) memory retention (the amount of remembered information), (2) monitoring effectiveness (the ability to discriminate between correct and incorrect answers), (3) control sensitivity (the extent to which the decision to volunteer a particular answer is based on the monitoring output), and (4) free-report volunteering rate (after equating mean confidence) as an indirect measure of control policy. Post hoc paired comparisons were performed using Scheffe's test (.05 criterion value).

As expected, memory retention, measured in terms of forced-report memory performance (quantity or accuracy), was superior for the Young-FA participants (.84) than for both the older participants (.69) and the Young-DA participants (.68), $F(2, 57) = 10.32$, $MSE = 139.43$, $p < .001$, $\eta^2 = .27$.

The latter two groups did not differ, indicating our success in equating baseline quantity performance.

As shown in Figure 2, the option of free report allowed all three groups of participants to enhance the accuracy of their reports considerably. However, a one-way analysis of variance (ANOVA) indicated that there were still substantial differences between the groups in free-report memory accuracy, $F(2, 57) = 5.72$, $MSE = 108.51$, $p < .01$, $\eta^2 = .17$. The older adults showed inferior accuracy performance (.87) than both the Young-DA (.94) and the Young-FA adults (.98). In comparing memory accuracy under free and

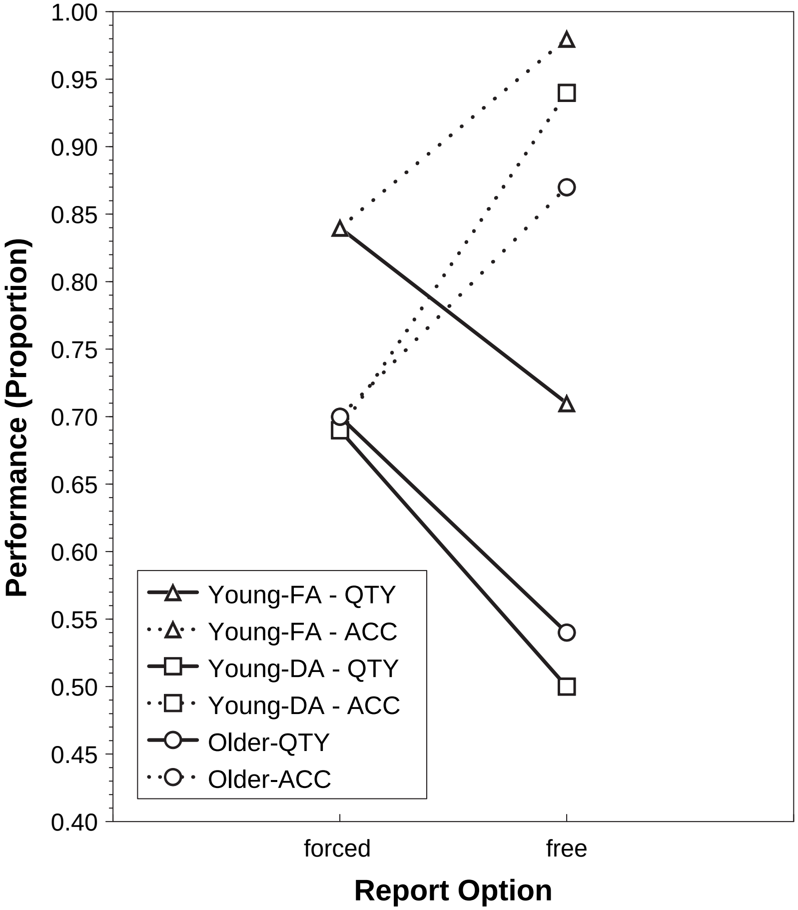

Figure 2. Quantity (QTY) and accuracy (ACC) performance, under free versus forced memory reporting, for the three groups of participants: (1) young adults in the full attention condition (Young-FA), (2) young adults in the divided attention condition (Young-DA), and (3) older adults (Older).

forced instructions, it appears that although the older participants' memory performance was practically identical to that of the Young-DA group under forced-report conditions, the older group did not gain as much as the Young-DA group did from the opportunity to screen out wrong answers, with free-forced accuracy improvements of .17 and .25, respectively, $F(1, 38) = 4.25$, $MSE = 147.03$, $p < .05$, $\eta^2 = .10$. Note that by exercising the option to withhold answers, the Young-DA group was able to close the gap with the Young-FA group, achieving comparable free-report accuracy. The increase in memory accuracy due to the exercise of free-report option, however, came at the cost of a reduction in memory quantity performance. This cost was more pronounced for the Young-DA group (.19) than for the older group (.13), $F(1, 38) = 6.48$, $MSE = 58.70$, $p < .05$, $\eta^2 = .15$. Thus, it appears that the two groups chose different strategies in trading quantity for accuracy, with the older group putting more emphasis on quantity and the Young-DA group putting more emphasis on accuracy. In fact, given the particular incentive matrix used in the current experiment, these two different tradeoffs yielded equivalent payoffs: The average number of points earned in the free-report stage did not significantly differ for the older group (143) and for the Young-DA group (138.5), but both groups earned significantly less bonus points than the Young-FA group (209), $F(2, 57) = 10.88$, $MSE = 2864.12$, $p < .001$, $\eta^2 = .28$.

What are the mechanisms underlying age-related changes in memory performance? Our paradigm allowed us to trace the observed age differences in free-report accuracy and quantity performance to differences in metacognitive monitoring and control. Regarding monitoring, one can use the option of free report to enhance accuracy only to the extent that one is able to discriminate correct from incorrect answers. *Monitoring effectiveeness* (resolution) was evaluated in terms of the within-subject Goodman-Kruskal gamma correlation between the assessed confidence and the correctness of each answer (Nelson, 1984). Gamma averaged .81, .83, and .92, for the older group, the Young-DA group, and the Young-FA group, respectively, $F(2, 57) = 3.20$, $MSE = 0.02$, $p < .05$, $\eta^2 = .10$. Thus, both the older adults and the Young-DA group exhibited poorer monitoring than the Young-FA group. However, the monitoring resolution of the older adults was comparable to that of the Young-DA group. This pattern of results indicates, in line with Rhodes and Kelley's (2005) results, that age-related differences in monitoring resolution are eliminated when the quality of encoding is equated. Hence, age-related monitoring differences may in fact stem from differences in the efficiency of encoding processes. The results also suggest, however, that part of the age-related reduction in memory accuracy, observed in comparing the older and Young-DA groups, stems from sources other than differences in monitoring resolution and/or encoding quality.

As mentioned earlier, the use of the confidence rating scale, rather than probability assessments of correctness, precluded the calculation of calibration bias scores (but see comparable measure in the later path analysis). Nonetheless, comparison of the mean confidence rating of the older adults (3.9) to that of the Young-DA group (3.7) does not reveal a significant difference between the two groups, $F(1, 38) = 2.07$, $MSE = 0.30$, *ns*, $\eta^2 = .05$. Given the comparable forced-report quantity performance of the two groups (reported previously), there is no evidence of an age-related difference in over- or underconfidence (the interaction between age group and measure, forced-report quantity vs. confidence, was not significant, $F < 1$).

As noted earlier, effective monitoring should enhance memory accuracy only to the extent that the control decision (i.e., whether or not to volunteer each answer) is sensitive to the monitoring output. To assess *control sensitivity*, we calculated within-subject gamma correlations between subjective confidence and the decision to volunteer an answer. These correlations were found to differ significantly between the groups. $F(2, 57) = 4.20$, $MSE = 0.01$, $p < .05$, $\eta^2 = .13$. The relationship between confidence and volunteering was less strong for the older participants (mean gamma = .90) than for both the Young-FA participants (mean gamma = .99) and the Young-DA participants (mean gamma = .96).

In addition, the older participants chose to volunteer significantly more answers (mean = 19.7) than did the Young-DA participants (mean = 15.9), $F(1, 57) = 6.73$, $MSE = 22.02$, $p < .05$, $\eta^2 = .15$. Given that the two groups exhibited comparable memory quantity performance and provided equally high confidence ratings in the forced-report phase, this pattern indicates that the older adults employed a more liberal *control policy* than their younger counterparts. This difference was also implied in the different accuracy–quantity tradeoff patterns observed earlier in comparing the different accuracy gains and quantity costs of free versus forced reporting exhibited by the two groups.

We turn next to examination of the results of the neurocognitive measures collected for the older group. In order to assess the potential role of executive functioning in the age-related memory and metamemory deficits, we calculated the correlations across participants within the older age group, between the memory measures, metacognitive measures, the neurocognitive indices obtained from WCST and the verbal fluency test, and age (in months).

As shown in Table 1, age differences within the older age group were found to be negatively correlated with the positive indices of executive functioning (verbal fluency, categories achieved in WCST) and positively correlated with the negative indices (total errors and preservation errors in WCST). Thus, the older the participant, the worse his or her performance on executive tests. In order to avoid the confound of neuropsychological

TABLE 1
Correlations between memory/metamemory indices and measures of executive functioning of the older participants

	Age	*Memory retention*	*Memory accuracy*	*Monitoring resolution*	*Control sensitivity*
FAS	−.68*	.58*	.62*	.43	.59*
WCSTc	−.53*	.24	.57*	.39	.42
WCSTt	.54*	−.30	−.58*	−.38	−.53*
WCSTp	.53*	−.46*	−.70*	−.44*	−.74*
Age	—	−.57*	−.63*	−.45*	−.42

*$p < .05$. FAS = Word fluency; WCSTc = WCST categories achieved; WCSTt = WCST total errors; WCSTp = WCST perseverative errors.

functioning with age, we examined the relationships between the executive measures and the cognitive measures using second-order partial correlations to partial out the age factor. These data, presented in Table 2, show that the executive measure of perseverative errors on WCST is strongly and negatively correlated with both memory accuracy ($r = -.56$) and control sensitivity ($r = -.67$), and that control sensitivity is correlated with the FAS index as well ($r = .46$).

With the aim of gaining a clearer picture of the direct and indirect contributions of age and executive functioning to free-report memory accuracy, and how these are mediated by memory and metamemory components, a series of path analyses were conducted (see Figure 3).[3] Based on Koriat and Goldsmith's (1996b) framework, each path model assumed that free-report accuracy was affected directly by forced-report quantity correct (the amount of correct information that is accessible in memory), monitoring resolution (within-participant gamma correlation between confidence rating and actual correctness of the answer), mean confidence, answer withholding rate (percentage of answers withheld under free report), and control sensitivity (within-participant gamma correlation between confidence rating and the decision to volunteer/withhold the answer). In addition, monitoring resolution was assumed to depend in part on the amount and quality of accessible information in memory (see Kelley & Sahakyan, 2003; Koriat, 1993, 1995; Rhodes & Kelley, 2005). Hence, age or

[3] Because of the small sample sizes, these path analyses were conducted using least-squares multiple regression rather than structural equation modelling (Kline, 1998). Power analyses based on Green (1991), with power set at .80, indicated adequate sample size to detect partial correlations ≥.43 in Models A and B and ≥.57 in Model C. Thus, the nonsignificant path coefficients in these models should be interpreted with caution, as they may be concealing low-to-moderate sized effects. (Note that the path coefficients are standardised beta coefficients, which are generally lower than the corresponding partial correlations.)

TABLE 2
Partial correlations between memory/metamemory indices and measures of executive functioning of the older participants

	Memory retention	*Memory accuracy*	*Monitoring resolution*	*Control sensitivity*
FAS	.32	.34	.19	.46*
WCSTc	−.09	.37	.19	.25
WCSTt	.00	−.38	−.18	−.39
WCSTp	.23	−.56*	−.26	−.67*

*p <.05. FAS = Word fluency; WCSTc = WCST categories achieved; WCSTt = WCST total errors; WCSTp = WCST perseverative errors.

executive functioning might affect monitoring resolution directly or indirectly via an effect on forced-report quantity correct. Of course, mean confidence was assumed to be strongly determined by forced-report quantity correct. Hence, any additional effect of age or executive functioning on mean confidence would indicate an effect that could not be explained by differences in the actual correctness of the answers, and would therefore reflect either differences in the use of the confidence scale (i.e., in assigning numbers to subjective confidence levels) or differences in monitoring calibration bias (over- or underconfidence). Similarly, the answer withholding rate was assumed to be strongly determined by mean confidence. Hence, any additional effect of age or executive functioning on this measure would indicate a difference in the withholding rate that could not be explained by a difference in mean confidence alone, and would therefore reflect a more conservative (positive effect) or more liberal (negative effect) control policy.

The first two path models (Figure 3, Models A and B) allowed for the possibility that each of the included cognitive or metacognitive variables might mediate the effects of age group, in comparing the older participants to the Young-DA group (Model A), or in comparing the older participants to the Young-FA group (Model B). The third model (Model C), focusing on the older age group alone, examined whether some of the effects of age within the older age group might be mediated by differences in executive functioning, which in turn might be mediated by any or all of the cognitive or metacognitive factors. Possible direct effects of age group (Models A and B), chronological age (Model C), and executive functioning (Model C) on free-report accuracy were also examined.

Beginning with the analysis comparing the older age group to the Young-DA group (Figure 3, Model A), we see that the cognitive and metacognitive factors together accounted for 87% of the variance in free-report memory accuracy, with no remaining direct effect of age group. In fact, the overall effect of age group on free-report accuracy ($r = -.28$) was mediated entirely

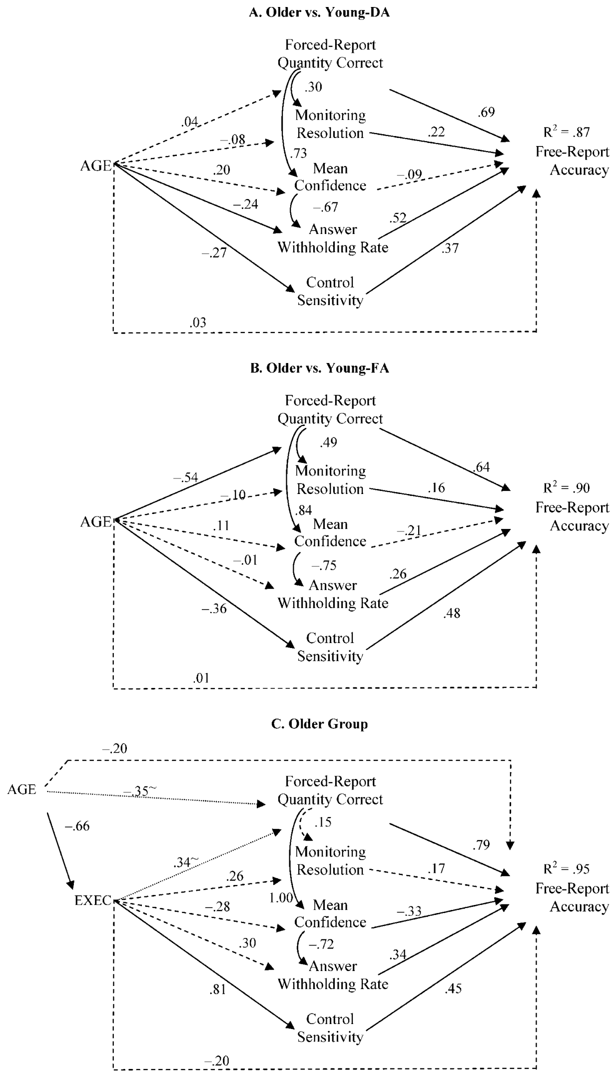

Figure 3. Path models examining the effects of age (Model A: Older vs. Young-DA participants; Model B: Older vs. Young-FA participants; Model C: chronological age within the older group) and executive functioning (EXEC; Model C only) on free-report accuracy performance, and the mediation of these effects by memory and metamemory components. Coefficients for each path represent standardised β weights. Statistically significant coefficients/paths are designated by solid arrows, and nonsignificant coefficients/paths are indicated by dashed arrows. The two coefficients/paths in Model C marked by a " ∼" and dotted arrows are nonsignificant individually, but together comprise a significant regression model. Also in Model C, for the sake of clarity, nonsignificant paths between age and several of the cognitive and metacognitive variables have been omitted. See text for further details.

by the two metacognitive control variables: withholding rate and control sensitivity, with the indirect effects of age group via these two variables calculated as $-.12$ and $-.10$, respectively. Thus, the older participants tended to be more liberal in their reporting, and to base their reporting less strongly on their subjective confidence than did the Young-DA participants, with negative consequences for free-report accuracy stemming from both differences. Other variables affecting free-report accuracy that were not affected by age group in this analysis were forced-report quantity correct and monitoring resolution. The lack of age effect on forced-report quantity correct reflects the success of the divided-attention manipulation in equating forced-report performance (presumably by equating the quality of memory encoding) between the two age groups. Although not related to age group in this analysis, the unique contribution of monitoring resolution beyond that of forced-report quantity correct is worth noting, because it reinforces the general theoretical claim that effective memory monitoring is not just a matter of "good memory" (for a similar result, see Rhodes & Kelley, 2005).

Turning to the analysis comparing the older age group to the Young-FA group (Figure 3, Model B), the pattern of results is somewhat more complex. Once again, the cognitive and metacognitive factors together accounted for a very large proportion of the variance in free-report memory accuracy (90%), with no remaining direct effect of age group. Moreover, once again part of the overall effect of age group on free-report accuracy ($r = -.42$) was mediated by control sensitivity (indirect effect $= -.17$). In this comparison, however, there was no direct age effect on the answer withholding rate, and hence no indication of an age effect on the control policy. Instead, several indirect effects of age on free-report accuracy were mediated by forced-report quantity correct, with negative age effects stemming from the decreased accessibility of correct information per se (indirect effect $= -.35$) and from an ensuing decrease in monitoring effectiveness (indirect effect $= -.04$). Interestingly, these negative age effects were offset to a small degree by a positive indirect effect, in which lower accessibility of correct information leads to lower mean confidence, leading to more withholding of answers, leading to higher free-report accuracy (indirect effect $= .09$).

In the preceding two analyses, we examined the effects of age group on the various cognitive and metacognitive variables, and how these might mediate age-related effects on free-report accuracy. In this final analysis (Figure 3, Model C), we examined the extent to which age-related effects within the older group may be mediated by differences in executive functioning. To index executive functioning, we used a composite score based on the mean of the FAS score and WCST perseverance error score, after first standardising these scores and then reversing the sign of WCST perseverance error score (so that high values of the composite score would reflect good executive functioning). Unlike in the preceding analyses, chronological age (in months) was used to tap the potential age-related effects within the older group.

It can be seen that in this analysis, the entire path model accounted for fully 95% of the variance in free-report accuracy. Notably, although both executive functioning and age yielded large overall effects on free-report accuracy ($r = .73$ and $-.63$, respectively), once the cognitive and metacognitive factors were included in the model, the residual direct effect of executive functioning ($-.20$) was no longer significant, and once executive functioning was included in the model, the residual direct effect of age ($-.20$) was also no longer significant. In fact, the effects of executive functioning on memory accuracy in this analysis were mediated entirely by the included cognitive and metacognitive variables, whereas the effects of age were mediated mostly, if not entirely, by executive functioning. Examining first the effects of executive functioning, the most notable result is that control sensitivity was again found to mediate age-related differences in memory accuracy, now being shown to mediate the effects of executive functioning (indirect effect = .36). In contrast, after partialling out the relevant variables, the unique effects of executive functioning on forced-report quantity correct, monitoring resolution, mean confidence (calibration), and answer withholding rate (control policy), were not significant, though this may be due in part to the small sample size ($N = 20$). Presumably, despite the nonsignificant path coefficient (.34), some of the effect of executive functioning on memory accuracy is mediated by forced-report quantity correct. This is because the executive-functioning and age variables together accounted for a significant and moderately large proportion of the variance in forced-report quantity correct ($R^2 = .40$). Yet, with both variables included in the model, neither the executive-functioning nor the age coefficient individually reached statistical significance.

Turning to the age variable, notwithstanding the problem just mentioned in interpreting the nonsignificant direct effect of age on forced-report quantity correct (versus the indirect effect of age on this variable via executive functioning), it is fair to say that much of the age effect on free-report memory accuracy observed in this analysis is mediated by the executive functioning composite score. This includes the indirect effect mediated by executive functioning and control sensitivity as well as the possible indirect effects mediated by executive functioning and forced-report quantity performance.

Taken together, the results of all three path analyses converge on the conclusion that diminished control sensitivity is an important mediator of the observed old-age decline in free-report memory accuracy, and this in turn stems from an old-age decline in executive functioning. This is in addition to inconsistent evidence of effects stemming from the use of a more liberal control policy (found in Model A but not in Models B and C) and to the effects due to an age-related decline in memory per se.

DISCUSSION

The present findings replicate those of numerous studies, demonstrating an age-related decline in the amount of retrieved information (for reviews, see Anderson & Craik, 2000; Balota, Dolan, & Duchek, 2000). More illuminating are the findings regarding memory accuracy, particularly when memory quantity was equated across the two age groups by dividing attention at study for the younger group. On the whole, these findings reveal a link between decreased memory accuracy, poorer executive functioning, and ageing within a general theoretical framework specifying the mediating role of metacognitive processes. Several differences were found between older and younger adults, both in memory accuracy and in its metacognitive determinants. First, the older participants exhibited lower free-report accuracy than did either group of younger participants. Second, the monitoring resolution of the older adults was lower than that of the Young-FA adults, but not lower than that of the Young-DA adults. This pattern suggests that the impaired monitoring (and ensuing consequences for memory accuracy) may be due to deficient memory encoding by older adults and Young-DA adults alike (see also Kelley & Sahakyan, 2003; Rhodes & Kelley, 2005).

Third, the higher volunteering rate of the older participants compared to the Young-DA participants, given the comparable level of retention and confidence of the two groups, suggests that they may have set a more lax criterion for volunteering information. This, together with the finding of a smaller reduction in memory quantity performance stemming from the exercise of free-report option by the older participants compared to the Young-DA participants, could be taken to reflect a general difference between older and younger adults in control policy. This possibility is consistent with previous findings demonstrating that older adults are less likely than younger adults to withhold answers when given the option of free report, even though this leads them to higher rates of false memory (see Jacoby et al., 2005; Kelley & Sahakyan, 2003). One reason why older adults might be reluctant to utilise the option of free report is the desire to "save face" by providing a reasonable amount of information despite their poor memory retention (cf. the "minimum informativeness" criterion recently proposed by Ackerman & Goldsmith, in press). This reluctance might be reinforced by metaknowledge that their memory quantity performance is generally poor (Hertzog & Hultsch, 2000) and by low self expectations with regard to their memory performance (Kausler, 1991). In other words, older adults may tend to adopt a liberal report criterion in an attempt to compensate for relatively low memory quantity, even though this comes at a cost in memory accuracy. In contrast, young adults would be more likely to attribute low memory quantity performance to external, situational factors. Indeed, the Young-DA participants in the present experiment could easily blame their relatively low memory quantity performance on the difficulty of attending to the presented information with an interfering tone-counting task.

Similar patterns of relatively low forced-report quantity performance accompanied by a relatively liberal control policy have been found in other special populations. For example, young children are particularly reluctant to withhold answers in response to memory questions and show a greater tendency than adults to provide wrong answers even when they are reminded that they have the option to say "don't know" (e.g., Cassel, Roebers, & Bjorklund, 1996). Thus, Koriat, Goldsmith, Schneider, and Nakash-Dura (2001) found a developmental trend, in which younger children exhibited both inferior memory quantity performance and a more liberal control policy than older children, leading to lower free-report accuracy as well.

The final age-related difference in metacognitive functioning observed in the current study is the finding of somewhat diminished control sensitivity in the older participants, indicating a reduced reliance on subjective confidence in deciding which answers to volunteer and which to withhold. As brought out in the path analyses, this difference in control sensitivity was responsible for some of the observed age-related accuracy reduction, and was quite robust, mediating age-related accuracy differences in all three analyses. This finding is particularly interesting in light of the strong correlation found between control sensitivity and measures of executive functioning within the older group. Taken together, the results suggest that the age-related reduction in memory accuracy results not only from poor memory retention and a consequential decline in memory monitoring, but also from a decline in metacognitive control (reduced sensitivity and perhaps a more lax control policy as well), which seems to be related to a decline in executive functioning, one of the important functions of the frontal lobe (Moscovitch & Winocur, 1995).

The concept of control sensitivity is an important feature of Koriat and Goldsmith's (1996b) theoretical framework, distinguishing it from similar frameworks such as signal-detection theory (see Goldsmith & Koriat, 2003, 2008). However, there are methodological (and hence interpretational) issues concerning the assessment of control sensitivity that must be considered. On the one hand, a procedure in which participants make their control decision immediately after providing a confidence judgement, on an item-by-item basis, might be suspected of demand characteristics, creating an artificially high correlation between the subjective measure and actual behaviour. On the other hand, eliciting the control decisions in a second phase, after the confidence ratings for all the items have been provided (and subsequently removed), as we did in the present study, also poses potential problems in interpreting the results.

One such problem (raised by an anonymous reviewer), is that despite the temporal separation, some participants may nonetheless remember the confidence ratings that they provided in the first phase when making their control decision in the second phase, which may in turn guide that decision. If so, age-related differences in control sensitivity might actually reflect poorer memory for the first-phase confidence judgements on the part of the older participants compared to the young participants. Another possibility is that participants' confidence in their answers might change somewhat from one phase to the next, and that the older participants' confidence may be less stable than that of the younger participants. Regarding the memory issue, we believe that given the nearly perfect correlations between confidence and volunteering for the young participant groups in this study (gamma averaging .97), as well as in other studies (e.g., Koriat & Goldsmith, 1996b), and the large number of confidence ratings that would need to be remembered (30 in this study, up to 90 in previous studies), it is doubtful that memory for one's previously provided confidence ratings could be contributing much variance. With regard to the stability issue, we note that the difference in control sensitivity was found between the older participants and the Young-DA participants (as well as the Young-FA participants), who were equated on memory retention. This reduces the possibility that confidence in one group would be less stable than in the other, but we acknowledge that this issue must remain open until it is examined more directly in future research.

To the extent that the observed age-difference in control sensitivity is real, it joins other recent evidence suggesting that control sensitivity is an important factor to consider in explaining memory performance in special populations. For example, in a series of studies, Koren and colleagues (Koren et al., 2004, 2005; Koren, Seidman, Goldsmith, & Harvey, 2006) adapted Koriat and Goldsmith's (1996b) framework and methodology to examine the performance of first-episode schizophrenic patients on a metacognitive free-report version of the WCST. In that adaptation, patients rated their confidence in each sort, and decided whether or not they wanted that sort to count towards their performance payoff. Several of the metacognitive measures from the adapted task were found to correlate more strongly with clinical measures relevant to real-world functioning ("insight into illness" and "competence to consent to treatment") than did traditional neuropsychological measures. Most prominent was control sensitivity, which was more highly correlated with the clinical measures of insight than any of the standard neuropsychological measures that were examined (Koren et al., 2004).

In a similar vein, Danion et al. (2001) used Koriat and Goldsmith's (1996b) metacognitive free-forced paradigm to compare schizophrenic patients with healthy controls on semantic (general-knowledge) memory tasks and found significantly lower control sensitivity for the schizophrenic patients (.83) compared to the healthy controls (.94). Interestingly, a similar reduction in control sensitivity was found with healthy participants who were under the influence of lorezapam (Massin-Krauss, Bacon, & Danion, 2002). Taken together, these lines of research point to an impaired relationship between subjective experience and behaviour associated with mental illness and the effects of psychoactive drugs that is similar to the one we found for normal older adults.

Diminished control sensitivity is reminiscent of what Stuss and Benson (1987) have called a dissociation between self-consciousness and self-knowledge, which is typical of patients with prefrontal damage (West, 1996). For example, this is the type of behaviour that frontal patients tend to exhibit when making perseverative errors on the WCST. The WCST requires the person to sort a series of cards based on a continually changing rule. When a category shift occurs, patients frequently continue to sort the cards to the now incorrect category, even as they express with each sort that this response is inappropriate (Duncan, 1995; Heaton, 1981). This suggests that the individual is aware that he or she is committing errors in behaviour, but is unable to modify that behaviour based on this knowledge. Thus, diminished control sensitivity could perhaps derive from a general age-related difficulty in inhibiting prepotent responses. This capacity has been attributed to the frontal lobe (see, e.g., Miyake et al., 2000; Rabbitt, Lowe, & Shilling, 2001), and has been suggested to be impaired among older adults (see, e.g., Zacks, Hasher, & Li, 2000), and perhaps among older adults with reduced executive functioning in particular (Davidson & Glisky, 2002).

Older adults' diminished control sensitivity may also stem in part from low memory self-efficacy (Hertzog & Hultsch, 2000), mentioned earlier in connection with the possibility of a general age-related relaxation of control policy. If older adults are distrustful of their memory monitoring as well as of their memory per se, this might make them less likely to rely on their monitoring output to guide their memory reporting.

Hence, an optimistic implication of our results is that part of the reduction in memory accuracy in old age may be avoidable: Training older adults to rely more on their monitoring output, and perhaps to be more conservative in their control policies, could help improve their free-report accuracy.

REFERENCES

Ackerman, R., & Goldsmith, M. (in press). Control over grain size in memory reporting—with and without satisficing knowledge. *Journal of Experimental Psychology: Learning, Memory and Cognition*.

Anderson, N. D., & Craik, F. I. M. (2000). Memory in the aging brain. In F. I. M. Craik & E. Tulving (Eds.), *The Oxford handbook of memory* (pp. 411–425). New York: Oxford University Press.

Balota, D. A., Dolan, P. O., & Duchek, J. M. (2000). Memory changes in healthy older adults. In F. I. M. Craik & E. Tulving (Eds.), *The Oxford handbook of memory* (pp. 395–409). New York: Oxford University Press.

Barnes, A. E., Nelson, T. O., Dunlosky, J., Mazzoni, G., & Narens, L. (1999). An integrative system of metamemory components involved in retrieval. In D. Gopher & A. Koriat (Eds.), *Attention and performance XVII: Cognitive regulation of performance: Interaction of theory and application* (pp. 287–313). Cambridge, MA: MIT Press.

Butler, K. M., McDaniel, M. A., Dornburg, C. C., Price, A. L., & Roediger, H. L. (2004). Age differences in veridical and false recall are not inevitable: The role of frontal lobe function. *Psychonomic Bulletin and Review*, *11*, 921–925.

Cassel, W. S., Roebers, C. E. M., & Bjorklund, D. F. (1996). Developmental patterns of eyewitness responses to repeated and increasingly suggestive questions. *Journal of Experimental Child Psychology*, *61*, 116–133.

Chan, J. C. K., & McDermott, K. B. (2007). The effects of frontal lobe functioning and age on veridical and false recall. *Psychonomic Bulletin and Review*, *14*, 606–611.

Craik, F. I. M., & Anderson, N. D. (1999). Applying cognitive research to problems of aging. In D. Gopher & A. Koriat (Eds.), *Attention and performance XVII: Cognitive regulation of performance: Interaction of theory and application* (pp. 583–615). Cambridge, MA: MIT Press.

Danion, J. M., Gokalsing, E., Robert, P., Massin-Krauss, M., & Bacon, E. (2001). Defective relationship between subjective experience and behavior in schizophrenia. *American Journal of Psychiatry*, *158*, 2064–2066.

Davidson, P. S. R., & Glisky, E. L. (2002). Neuropsychological correlates of recollection and familiarity in normal aging. *Cognitive, Affective and Behavioral Neuroscience*, *2*, 174–186.

Duncan, J. (1995). Attention, intelligence, and the frontal lobes. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 721–733). Cambridge, MA: MIT Press.

Fernandez-Duque, D., Baird, J. A., & Posner, M. I. (2000). Executive attention and metacognitive regulation. *Consciousness and Cognition*, *9*, 288–307.

Glisky, E. L., Polster, M. R., & Routhieaux, B. C. (1995). Double dissociation between item and source memory. *Neuropsychology*, *9*, 229–235.

Glisky, E. L., Rubin, S. R., & Davidson, P. S. R. (2001). Source memory in older adults: An encoding or retrieval problem? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 1131–1146.

Goldsmith, M., & Koriat, A. (1999). The strategic regulation of memory reporting: Mechanisms and performance consequences. In D. Gopher & A. Koriat (Eds.), *Attention and performance XVII: Cognitive regulation of performance: Interaction of theory and application* (pp. 373–400). Cambridge, MA: MIT Press.

Goldsmith, M., & Koriat, A. (2003). Dolphins on the witness stand? The comparative psychology of strategic memory monitoring and control. *Behavioral and Brain Sciences*, *26*, 345–346.

Goldsmith, M., & Koriat, A. (2008). The strategic regulation of memory accuracy and informativeness. In A. Benjamin & B. Ross (Eds.), *The psychology of learning and motivation* (Vol. 48, pp. 1–60). London: Academic Press.

Goldsmith, M., Koriat, A., & Pansky, A. (2005). Strategic regulation of grain size in memory reporting over time. *Journal of Memory and Language*, *52*, 505–525.

Goldsmith, M., Koriat, A., & Weinberg-Eliezer, A. (2002). Strategic regulation of grain size memory reporting. *Journal of Experimental Psychology: General*, *131*, 73–95.

Green, S. B. (1991). How many subjects does it take to do a regression analysis? *Multivariate Behavioral Research*, *26*, 499–510.

Greenwood, P. M. (2000). The frontal aging hypothesis evaluated. *Journal of the International Neuropsychological Society*, *6*, 705–726.

Hart, R. P., Kwentus, J. A., Wade, J. B., & Taylor, J. R. (1988). Modified Wisconsin Sorting Test in elderly normal, depressed and demented patients. *The Clinical Neuropsychologist*, *2*, 49–56.

Heaton, W. C. (1981). *Wisconsin Card Sorting Test*. Odessa, TX: Psychological Assessment Resources.

Henkel, L. A., Johnson, M. K., & de Leonardis, D. M. (1998). Aging and source monitoring: Cognitive processes and neuropsychological correlates. *Journal of Experimental Psychology: General*, *127*, 251–268.

Hertzog, C., & Dunlosky, J. (2004). Aging, metacognition, and cognitive control. In B. H. Ross (Ed.), *The psychology of learning and motivation: Advances in research and theory* (pp. 215–247). San Diego, CA: Elsevier.

Hertzog, C., & Hultsch, D. F. (2000). Metacognition in adulthood and old age. In F. I. M. Craik (Ed.), *The handbook of aging and cognition* (2nd ed., pp. 417–466). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Higham, P. A. (2002). Strong cues are not necessarily weak: Thomson and Tulving (1970) and the encoding specificity principle revisited. *Memory and Cognition*, *30*, 67–80.

Higham, P. A. (2007). No special K! A signal detection framework for the strategic regulation of memory accuracy. *Journal of Experimental Psychology: General*, *136*, 1–22.

Higham, P. A., & Tam, H. (2005). Generation failure: Estimating metacognition in cued recall. *Journal of Memory and Language*, *52*, 595–617.

Jacoby, L. L. (1999). Deceiving the elderly: Effects of accessibility bias in cued-recall performance. *Cognitive Neuropsychology*, *16*, 417–436.

Jacoby, L. L., Bishara, A. J., Hessels, S., & Toth, J. P. (2005). Aging, subjective experience, and cognitive control: Dramatic false remembering by older adults. *Journal of Experimental Psychology: General*, *134*, 131–148.

Jacoby, L. L., Debner, J. A., & Hay, J. F. (2001). Proactive interference, accessibility bias, and process dissociations: Valid subject reports of memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 686–700.

Jacoby, L. L., & Rhodes, M. G. (2006). False remembering in the aged. *Current Directions in Psychological Science*, *15*, 49–53.

Janowsky, J. S., Shimamura, A. P., & Squire, L. R. (1989). Source memory impairment in patients with frontal lobe lesions. *Neuropsychologia*, *27*, 1043–1056.

Jennings, J. M., & Jacoby, L. L. (1993). Automatic versus intentional uses of memory: Aging, attention, and control. *Psychology and Aging*, *8*, 283–293.

Kato, T. (1985). Semantic-memory sources of episodic retrieval failure. *Memory and Cognition*, *13*, 442–452.

Kausler, D. H. (1991). *Experimental psychology, cognition, and human aging* (2nd ed.). New York: Springer-Verlag.

Kelley, C. M., & Sahakyan, L. (2003). Memory, monitoring, and control in the attainment of memory accuracy. *Journal of Memory and Language*, *48*, 704–721.

Kline, R. B. (1998). *Principles and practice of structural equation modeling*. New York: Guilford Press.

Koren, D., Poyurovsky, M., Seidman, L. J., Goldsmith, M., Wenger, S., & Klein, E. M. (2005). The neuropsychological basis of competence to consent in first-episode schizophrenia: A pilot metacognitive study. *Biological Psychiatry*, *57*, 609–616.

Koren, D., Seidman, L. J., Goldsmith, M., & Harvey, P. D. (2006). Real-world cognitive—and metacognitive—dysfunction in schizophrenia: A new approach for measuring (and remediating) more "right stuff". *Schizophrenia Bulletin*, *32*, 310–326.

Koren, D., Seidman, L. J., Poyurovsky, M., Goldsmith, M., Viksman, P., Zichel, S., & Klein, E. (2004). The neuropsychological basis of insight in first-episode schizophrenia: A pilot metacognitive study. *Schizophrenia Research*, *70*, 195–202.

Koriat, A. (1993). How do we know that we know? The accessibility model of the feeling of knowing. *Psychological Review*, *100*, 609–639.

Koriat, A. (1995). Dissociating knowing and the feeling of knowing: Further evidence for the accessibility model. *Journal of Experimental Psychology: General*, *124*, 311–333.

Koriat, A., Ben-Zur, H., & Sheffer, D. (1988). Telling the same story twice: Output monitoring and age. *Journal of Memory and Language*, *27*, 23–39.

Koriat, A., & Goldsmith, M. (1994). Memory in naturalistic and laboratory contexts: Distinguishing the accuracy-oriented and quantity-oriented approaches to memory assessment. *Journal of Experimental Psychology: General*, *123*, 297–315.

Koriat, A., & Goldsmith, M. (1996a). Memory metaphors and the real-life/laboratory controversy: Correspondence versus storehouse conceptions of memory. *Behavioral and Brain Sciences*, *19*, 167–228.

Koriat, A., & Goldsmith, M. (1996b). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review*, *103*, 490–517.

Koriat, A., Goldsmith, M., & Halamish, V. (2008). Control processes in voluntary remembering. In I. H. L. Roediger (Ed.), *Cognitive psychology of memory* (Vol. 2, pp. 307–324). Oxford, UK: Elsevier.

Koriat, A., Goldsmith, M., & Pansky, A. (2000). Toward a psychology of memory accuracy. *Annual Review of Psychology*, *51*, 481–537.

Koriat, A., Goldsmith, M., Schneider, W., & Nakash-Dura, M. (2001). The credibility of children's testimony: Can children control the accuracy of their memory reports? *Journal of Experimental Child Psychology*, *79*, 405–437.

Koriat, A., Lichtenstein, S., & Fischhoff, B. (1980). Reasons for confidence. *Journal of Experimental Psychology: Human Learning and Memory*, *6*, 107–118.

Koutstaal, W. (2006). Flexible remembering. *Psychonomic Bulletin and Review*, *13*, 84–91.

Lichtenstein, S., Fischhoff, B., & Phillips, L. D. (1982). Calibration of probabilities: The state of the art to 1980. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 306–334). New York: Cambridge University Press.

Massin-Krauss, M., Bacon, E., & Danion, J. M. (2002). Effects of benzodiazepine lorazepam on monitoring and control processes in semantic memory. *Consciousness and Cognition*, *11*, 123–137.

McDaniel, M. A., Glisky, E. L., Rubin, S. R., Guynn, M. J., & Routhieaux, B. C. (1999). Prospective memory: A neuropsychological study. *Neuropsychology*, *13*, 103–110.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*, 49–100.

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

Moscovitch, M., & Winocur, G. (1995). Frontal lobes, memory, and aging. In J. Grafman, K. J. Holyoak, & F. Boller (Eds.), *Structure and functions of the human prefrontal cortex* (pp. 119–150). New York: New York Academy of Sciences.

Moscovitch, M., & Winocur, G. (2002). The frontal cortex and working with memory. In D. T. Stuss & R. T. Knight (Eds.), *Principles of frontal lobe function* (pp. 188–209). New York: Oxford University Press.

Nelson, T. O. (1984). A comparison of current measures of the accuracy of feeling-of-knowing predictions. *Psychological Bulletin*, *95*, 109–133.

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 26, pp. 125–123). San Diego, CA: Academic Press.

Payne, B. K., Lambert, A. J., & Jacoby, L. L. (2002). Best laid plans: Effects of goals on accessibility bias and cognitive control in race-based misperceptions of weapons. *Journal of Experimental Social Psychology*, *38*, 384–396.

Perfect, T. (1997). Memory aging as frontal lobe dysfunction. In M. A. Conway (Ed.), *Cognitive models of memory: Studies in cognition* (pp. 315–339). Cambridge, MA: MIT Press.

Perfect, T. J., & Stollery, B. (1993). Memory and metamemory performance in older adults: One deficit or two? *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *46A*, 119–135.

Rabbitt, P. M. A., Lowe, C., & Shilling, V. (2001). Frontal tests and models for cognitive ageing. *European Journal of Cognitive Psychology*, *13*, 5–28.

Raz, N. (2000). Aging of the brain and its impact on cognitive performance: Integration of structural and functional findings. In F. I. M. Craik & T. Salthouse (Eds.), *The handbook of aging and cognition* (2nd ed., pp. 1–90). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Rhodes, M. G. (2004). Age-related differences in performance on the Wisconsin Card Sorting Test: A meta-analytic review. *Psychology and Aging*, *19*, 482–494.

Rhodes, M. G., & Kelley, C. M. (2005). Executive processes, memory accuracy, and memory monitoring: An aging and individual difference analysis. *Journal of Memory and Language*, *52*, 578–594.

Roediger, H. L., & Geraci, L. (2007). Aging and the misinformation effect: A neuropsychological analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*, 321–334.

Salthouse, T. A., Atkinson, T. M., & Berish, D. E. (2003). Executive functioning as a potential mediator of age-related cognitive decline in normal adults. *Journal of Experimental Psychology: General*, *132*, 566–594.

Schwartz, B. L. (1994). Sources of information in metamemory: Judgments of learning and feelings of knowing. *Psychonomic Bulletin and Review*, *1*, 357–375.

Shimamura, A. P. (2000). Toward a cognitive neuroscience of metacognition. *Conciousness and Cognition*, *9*, 313–323.

Souchay, C., & Isingrini, M. (2004). Age related differences in metacognitive control: Role of executive functioning. *Brain and Cognition*, *56*, 89–99.

Souchay, C., Isingrini, M., & Espagnet, L. (2000). Aging, episodic memory feeling-of-knowing, and frontal functioning. *Neuropsychology*, *14*, 299–309.

Stuss, D. T., & Benson, D. F. (1987). The frontal lobes and control of cognition and memory. In E. Perecman (Ed.), *The frontal lobes revisited* (pp. 141–158). New York: IRBN Press.

Stuss, D. T., Craik, F. I. M., Sayer, L., Franchi, D., & Alexander, M. P. (1996). Comparison of older people and patients with frontal lesions: Evidence from word list learning. *Psychology and Aging*, *11*, 387–395.

Vilkki, J., Servo, A., & Surma-aho, O. (1998). Word list learning and prediction of recall after frontal lobe lesions. *Neuropsychology*, *12*, 268–277.

Waltz, J. A., Knowlton, B. J., Holyoak, K. J., Boone, K. B., Mishkin, F. S., de Menezes Santos, M., et al. (1999). A system for relational reasoning in human prefrontal cortex. *Psychological Science*, *10*, 119–125.

West, R. L. (1996). An application of prefrontal cortex function theory to cognitive aging. *Psychological Bulletin*, *120*, 272–292.

Zacks, R. T., Hasher, L., & Li, K. Z. H. (2000). Human memory. In F. I. M. Craik (Ed.), *The handbook of aging and cognition* (pp. 293–357). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.