Acta Psychologica 145 (2014) 54–64

Contents lists available at ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/locate/actpsy

When the mind wanders: Age-related differences between young and older adults

Michela Zavagnin *, Erika Borella *, Rossana De Beni

Department of General Psychology, University of Padova, Italy

ARTICLE INFO

Article history:
Received 28 December 2012
Received in revised form 16 October 2013
Accepted 29 October 2013
Available online xxxx

PsycINFO classification:
2340

Keywords:
Mind wandering
Working memory
Inhibition
Processing speed
Aging

ABSTRACT

Interest in mind wandering (MW) has grown in recent years, but few studies have assessed this phenomenon in older adults. The aim of this study was to assess age-related differences between young, young–old and old–old adults in MW using two versions of the sustained attention to response task (SART), one perceptual and one semantic. Different indicators were examined (i.e., reported MW episodes and behavioral indices of MW such as response time latency and variability, incorrect response and omission errors). The relationship between MW, certain basic mechanisms of cognition (working memory, inhibition and processing speed), cognitive failures and intrusive thoughts in everyday life was also explored. Findings in both versions of the SART indicated that older adults reported a lower frequency of MW episodes than young adults, but some of the behavioral indices of MW (response time variability, incorrect response and omission errors) were higher in old–old adults. This seems to suggest that MW becomes less frequent with aging, but more pervasive and detrimental to performance. Our results also indicated that the role of age and cognitive mechanisms in explaining MW depends on the demands of the SART task considered.

1. Introduction

Mind wandering (MW) can be defined as a shift of attention from the environmental context to stimuli and mental representations associated with personal thoughts or ongoing activities, i.e., task-unrelated thoughts (TUTs) (Antrobus, Singer, & Greenberg, 1966; Giambra, 1995; Smallwood, Obonsawin, & Heim, 2003). Although this phenomenon is very common (Kane et al., 2007) and can be useful in some circumstances (Baird, Smallwood, & Schooler, 2011; Baird et al., 2012), it is often unintentional and leads to a less accurate information encoding with consequent cognitive failures and related psychological stress (Schupak & Rosenthal, 2009; Smallwood, Riby, Heim, & Davies, 2006). Hence the interest in MW in numerous studies on different age groups (healthy younger and older adults), as well as in specific populations, including adults with ADHD (Giambra, 1993) and people with depression (e.g., Smallwood, O'Connor, Sudberry, & Obonsawin, 2007).

It is now well documented that aging coincides with a decline in some basic cognitive mechanisms, such as working memory, inhibition and processing speed (e.g., Craik & Salthouse, 2008; Hasher & Zacks, 1988). This decline also explains age-related differences in many cognitive domains and everyday life abilities (e.g., Borella, Ghisletta, & de Ribaupierre, 2011).

Recent theories on MW claim that executive control lies behind the MW phenomenon (McVay & Kane, 2009). As older adults generally have more limited executive control resources than young adults, i.e., less efficient inhibitory mechanisms (attentional control), and a weak working memory performance (Hasher & Zacks, 1988), we might expect them to experience more MW, and report more TUTs than young adults. Instead, an apparently paradoxical decrease in TUT frequency with aging has been reported in various studies and using various models. In particular, Giambra (1973, 1993, 2000) used self-report questionnaires to investigate daydreaming, i.e., task-unrelated image and thought intrusions, finding that their frequency decreased with aging in both longitudinal (Giambra, 2000) and cross-sectional comparisons (Giambra, 1973, 1993). Giambra (1989) also obtained much the same results using more objective methods, which were not sensitive to participants' beliefs and therefore unaffected by any biases (such as those induced by questionnaires). Giambra used a vigilance task (in which participants responded to rare targets) in a retrospective correlational study (1989) that involved 5 experiments (conducted from 1977 to 1980). The results confirmed that TUTs did not increase with aging. In particular, four of the five studies on the frequency of TUTs reported that old–old adults (over 70 years of age) had fewer TUTs than young or middle-aged adults; and three of the five studies also identified a lower frequency of TUTs in young–old (60- to 70-year-olds) than in younger adults. A negative correlation between age and TUTs across the lifespan also emerged in vigilance and reading tasks of variable difficulty (Grodsky & Giambra, 1990–1991). Giambra attributed the results concerning the decrease in TUTs (incompatible

* Corresponding authors at: Department of General Psychology, Via Venezia, 8, 35131 Padova, Italy. Tel.:+39 049 8276622; fax: +39 049 8276600.
E-mail addresses: michela.zavagnin@email.it (M. Zavagnin), erika.borella@unipd.it (E. Borella).

0001-6918/$ – see front matter
http://dx.doi.org/10.1016/j.actpsy.2013.10.016

with the inhibitory hypothesis [Hasher & Zacks, 1988]) to various factors, such as age-related memory decline, or an age-dependent reduction in nonconscious information processing due to older people having less “unfinished business”, or fewer matters of concern, whereas young adults can devote their attention to MW as well because of their greater attentional capacity (Giambra, 1989).

Giambra's results (Giambra, 1989, 1993) might also be due, however, to the introspective procedure used – in some experiments participants were asked to report when their mind was wandering during a task (self-caught method) – without taking into account the age-related differences in people's ability to monitor their internal states, and thus report on their TUTs (Einstein & McDaniel, 1997). To test this hypothesis, Einstein and McDaniel (1997) used a more objective and reliable performance-based measure, assessing the frequency of MW on the basis of participants' performance when they were occasionally interrupted while recalling long lists of words: if a participant recalled significantly fewer words than the average, this was interpreted as being due to MW, which led to their failure to encode the stimuli. The results of this study showed that older adults did not differ significantly from young people in this performance-based measure, and the authors concluded that there was no evidence of age-related differences in MW. Parks, Klinger, and Perlmutter (1988–89) came to a similar conclusion when they used a thought sampling procedure during the performance of more or less difficult tasks. They found a higher frequency of evaluative thoughts focusing on steps towards a goal and of attention-control utterances in older than in younger adults, regardless of the task's difficulty. Finally, Jackson and Balota (2012) recently investigated MW in younger and older adults using more recently developed experimental models. The authors presented different versions of the SART (a go/no-go task in which participants had to inhibit a habitual response), in which they also considered response latency as a possible performance-based indicator of MW, and a more demanding reading comprehension task. They expected to find the same level of MW in young and older adults in the more demanding test (the reading comprehension). Instead, they found that older adults reported less MW than their younger counterparts in both the SARTs and the reading comprehension task. The manipulation of the tasks' presentation had a different influence on the frequency of MW episodes reported by younger and older adults, however. In the first two experiments, the young adults reported four times as many MW episodes as the older adults; in the others, the young adults reported only twice as many MW episodes as the older adults (particularly in the more demanding tasks, i.e., the reading comprehension and a slower, longer version of the SART). Older adults also showed a disproportionate post-error slowing in their completion of the SARTs, probably because they found it more difficult to re-engage in the task (possibly due to a greater concern about their performance). The authors suggested that this pattern of results might reflect a greater engagement of older adults in the task (supported by personality measures indicating greater conscientiousness in this age group than in the young adults) that could lead to less MW and more self-evaluation thoughts after errors and that would explain the age-related differences in post-error latency.

Taken together, the above studies indicate either a similar frequency of MW being reported by young and older adults or (in the majority of the studies) a decreasing frequency of MW with aging; regardless of older adults' decline in inhibitory mechanisms (or more generally in executive control), the ability to maintain task goals or context over time could make it difficult to control intrusive thoughts (e.g., Hasher & Zacks, 1988; Balota, Black, & Cheney, 1992).

It is noteworthy that it seems difficult to predict findings concerning age-related differences in MW (between young and older adults) on the strength of the two main hypotheses proposed in the literature (based mainly on evidence of this phenomenon in younger adults), i.e., the “decoupling” hypothesis (Smallwood & Schooler, 2006), and the “control failure × current concerns” hypothesis (McVay, Kane, & Kwapil, 2009). The decoupling hypothesis considers MW as a spontaneous process involving a state in which attention becomes coupled to an internal process and decoupled from the external information (e.g., Barron, Riby, Greer, & Smallwood, 2011; Smallwood, 2010); accordingly, the frequency of MW would be modulated by the amount of resources required by the task and the amount of resources that the person possesses (Smallwood & Schooler, 2006). Thus, older adults' more limited cognitive resources (Craik & Salthouse, 2008) would result in less frequent MW. The results that this hypothesis would lead us to expect were not seen in all of the above-cited studies (i.e., two of the five studies conducted by Einstein & McDaniel (1997), Giambra (1989) and Parks et al. (1988–89)). Nor are the findings reported by Jackson and Balota (2012) consistent with a similar decrease in the frequency of MW episodes that younger and older adults reported with increasingly difficult tasks, as predicted by this decoupling hypothesis.

It seems also difficult to fully explain the above-mentioned results from the “control failure × current concerns” perspective (e.g., McVay & Kane, 2010), according to which TUTs are automatically generated in response to environmental cues, current concerns and personal goals. They enter the sphere of our awareness as a result of an attentional control failure, disrupting goal maintenance processes, so MW would presumably be more common in people with a poor working memory capacity (McVay & Kane, 2009), such as older adults, who may also have difficulty in modulating their MW in relation to the resources demanded by the task. But, according to these authors, it is also important to consider both the amount of thoughts automatically generated and the category of thoughts reportedly involved. In fact, older adults may have fewer current concerns and personal goals, and this would lead to less MW, as Giambra (1989) suggested, but may report more thoughts about their performance, i.e., concern-related thoughts (McVay, Meier, Touron, & Kane, 2013).

We could therefore conclude that it is still not clear how some of the cognitive mechanisms (i.e., working memory and inhibition) evoked by the two different hypotheses in the literature contribute to explaining MW in young and older adults. Likewise, it is therefore still hard to say how a task's complexity (in terms of the demands of the task) modulates MW. In fact, it has been well documented that the frequency of MW decreases for more complex tasks in younger adults (see Smallwood & Schooler, 2006), but this is not the case for older adults. Jackson and Balota (2012) tried to shed light on this aspect, and their experiments showed the different influence of task manipulation on MW frequency as a function of age. But their findings cannot be used to draw any final conclusions because they used a between-groups design and several variables (not only the difficulty of the tasks) were manipulated across the experiments, including the type of task, the rate of presentation of the stimuli, and the duration of the task.

Finally, it is important to make the point that an aspect of the issue has been overlooked, i.e., that the different MW patterns described in the literature may also depend on the age range of the older adults considered, given the difference in the extent of cognitive decline between the young-old (65–74 year-olds) and the old-old (75–85 year-olds), which is more accentuated in the latter group (e.g., Baltes, 1987; Borella, Carretti, & De Beni, 2008). The age brackets of older adults used in previous studies may have masked some of the differences in MW patterns with aging: Giambra divided his sample into two age groups (60–69, 70–89); Einstein and McDaniel (1997) studied a sample of young–old adults (60–76, $M = 65.6$), and Jackson and Balota (2012) considered elderly adults as a homogeneous group across experiments (their mean age ranged between 75.8 and 77.3 years).

Aiming to shed further light on this complex phenomenon, the present study explored the age-related differences in MW, comparing young, young-old and old–old adults, by: i) manipulating the demands of the tasks, presenting different versions of a SART; and ii) directly assessing the relationship between MW and certain cognitive mechanisms (working memory, inhibition and processing speed) using different versions of the SART. A multivariate design was adopted in which all the tasks considered were administered to all the participants.

The SART was chosen because it is the most commonly used paradigm for assessing MW. Concerning our first goal, to examine whether MW can be explained by the demands of a task (in terms of the cognitive resources involved), two versions of the SART were developed, one perceptual and one semantic. The semantic SART has the same structure as the perceptual one, but it involves participants simultaneously in two activities, i.e., encoding stimuli and detecting targets. Such a manipulation has been shown to prompt a different involvement of people's cognitive resources (see Smallwood & Schooler, 2006). In particular, less MW reportedly coincided with the more demanding SART in a sample of younger adults (Smallwood & Schooler, 2006). We newly explored whether the greater resources needed for the semantic SART than for the perceptual version would reduce the frequency of MW in older adults too, a hypothesis that would be consistent with the decoupling hypothesis (Smallwood & Schooler, 2006). To our knowledge at least, no studies conducted so far have examined this issue in both young and older adults using a multivariate design.

The frequency of MW was assessed using the probe-caught method. This is a widely-adopted method particularly useful in the case of older adults, who may be scarcely aware of their own mental processes because they do not have to constantly monitor their thoughts. The content of participants' MW episodes was also classified to obtain additional information on the source of these thoughts as a function of age and see whether there were any age-related differences in people's susceptibility to thoughts about personal goals and concerns, or to external cues, for instance. We thus distinguished between personal thoughts and daydreams (SITUTs), task-related interferences (TRIs) and external distractions (EDs) (see Stawarczyk, Majerus, Maj, Van Der Linden, & D'Argembeau, 2011). Using this classification, we assessed the generality vs. specificity of MW content. In particular, we expected to find: i) an age-related decrease in SITUTs, as predicted by Giambra (1989) and consistent with the two hypotheses in the literature on MW; and ii) an age-related increase in intrusive thoughts due to contextual cues (EDs and TRIs) in older adults as a result of inhibitory control failures and a greater focus on the smooth completion of the task (Jackson & Balota, 2012; Parks et al., 1988–89), which would be consistent with the ‘control failure × current concerns’ view of MW (e.g., McVay & Kane, 2010).

Given that older adults may have more difficulty in monitoring their thoughts, as Einstein and McDaniel (1997) suggested, the behavioral indicators of MW (i.e., response time latency and variability, incorrect responses and omission errors) usually assessed in younger populations (see Cheyne, Solman, Carriere, & Smildek, 2009) were also considered. Although these indices are sensitive to other factors (e.g., contextual distractions) as well as to MW, they can give us an important indication of the level of a participant's disengagement from a task. For instance, omission errors can be interpreted as being due to a deep level of MW because they have been found correlated with zoning-out episodes (i.e., MW without awareness) (Cheyne et al., 2009). If what Einstein and McDaniel said is true, we should see discrepancies between the reported MW episodes and the behavioral indicators: with increasing age, we might expect fewer reported MW episodes to coincide with a worse performance in the task and: a) more incorrect responses and omissions for non-target stimuli; b) pre-error speeding and post-error slowing (as found by Jackson & Balota, 2012); and c) a variance in the response times (RT CV). All these indicators would relate to more frequent, less conscious MW episodes.

Regarding the second goal of our study, we tried to shed more light on the relationship between the MW phenomenon, considered as a whole, and cognitive resources, using classical measures of working memory capacity (the Categorization Working Memory Span test, Borella et al., 2008), inhibitory efficacy (intrusion errors committed in the working memory task, Robert, Borella, Fagot, Lecerf, & de Ribaupierre, 2009), and processing speed (the pattern comparison test, Salthouse & Babcock, 1991). To our knowledge, no previous studies have investigated the joint role of these resources in explaining MW and individual and age-related differences in this phenomenon. According to the above-discussed hypotheses, we might expect either a decrease in MW with declining cognitive resources in aging, in agreement with the decoupling hypothesis, or else an increase in MW with decreasingly efficient inhibitory mechanisms and working memory (McVay & Kane, 2009). The present study should clarify this issue. Moreover, we expected cognitive resources to play a different part on the two versions of the SART considered, their relationship being stronger with the semantic SART, which demands more cognitive resources, than with the easier perceptual SART.

Finally, the relationship between MW and intrusive thoughts and cognitive failures due to inattention in everyday life was assessed with the aid of a new ad hoc questionnaire, which specifically indicated errors due to inattention (and not only to memory, as is typically the case in questionnaires on cognitive failure in everyday life). In particular, by means of this measure we wanted to see whether or not there was a sort of congruence between subjective (questionnaire) and more objective (performance-based, response latency, etc.) measures of MW in aging, as highlighted in younger adults after assessing self-reported everyday attention failures (Cheyne, Carriere, & Smildek, 2006), and mindfulness (Cheyne et al., 2006, 2009).

2. Method

2.1. Participants

Our sample consisted of twenty young adults (20–30 years old), 20 young–old adults (65–74 years old), and 19 old–old adults[1] (75–85 years old), who volunteered to take part in the study and were recruited by word of mouth. They were all native Italian speakers, and healthy. We ruled out anyone meeting any of the “exclusion criteria” proposed by Crook et al. (1986) — i.e. a history of head trauma, any neurological or psychiatric illness, a history of brain fever, dementia or any other state of altered consciousness, use of benzodiazepines in the previous 3 months, use of illicit drugs, any visual, auditory or motor impairments, any symptomatic cardiovascular conditions, breathing problems, or diseases capable of causing cognitive impairments.

The three age groups did not differ in terms of their years of formal education, or their scores in the *Wechsler Adult Intelligence Scale—Revised* (*WAIS—R*) vocabulary test (Wechsler, 1981); see Table 1.

2.2. Materials and methods

2.2.1. Perceptual and semantic SARTs

Both the SARTs contained 172 stimuli (144 non-target, 28 target) consisting of five strings of letters organized into blocks containing 5, 6 or 7 strings each, with 0, 1, or 2 targets. The presentation of the two blocks and the lists they contained were randomized.

Following Giambra's (1995) suggestions, and in the light of comments in Jackson and Balota (2012) on the need to present stimuli more slowly if a task is also intended for older people, each string was presented in the center of the screen for 2000 ms, with an interstimulus interval (ISI) of 2000 ms. The screen was black during the ISI. In the perceptual SART, there were five “x” target stimuli (XXXXX) and five “o” non-target stimuli (OOOOO). In the semantic SART, the targets were five-letter nouns identifying animals and the non-target stimuli were five-letter non-animal nouns (adapted from Smallwood et al., 2006). The experiment was conducted using the E-Prime software.

Participants sat about 50 cm away from the screen and, adopting the Stawarczyk et al. (2011) procedure, they were asked to press a green button when a non-target stimulus appeared and a red button when a target stimulus appeared (the target stimuli were evenly distributed throughout the task).

Probes were presented at the end of each block, and the time interval between one probe and the next ranged from 22 to 30 s (as suggested

[1] One old–old participant dropped out.

Table 1
Characteristics of the sample by age group and results of ANOVA.

	Young		Young–old		Old–old		ANOVA results			
	M	SD	M	SD	M	SD	*F*	df	η_p^2	*p*
Age	24.15	2.82	69.05	2.87	80.00	3.37				
Background										
Education	13.10	1.77	10.95	2.82	11.68	3.51	3.08	2,58	0.10	0.54
Vocabulary	50.85	10.66	47.70	9.44	45.74	10.20	1.27	2,58	0.04	0.29
Working memory										
CWMS	57.50	7.67	45.15	8.273	34.58	11.94	28.86	2,58	0.51	<.001
Inhibition										
Proportion of intrusion errors	0.06	0.06	0.11	0.08	0.18	0.13	9.22	2,58	0.25	<.001
Processing speed										
Pattern comparison (sec.)	115.38	30.61	172.79	42.13	190.41	46.87	19.56	2,58	0.41	<.001
Mind wandering										
MW Questionnaire	77.10	15.51	68.40	14.26	62.00	14.15	5.22	2,58	0.16	<.01

Note. CWMS: Categorization Working Memory Span test.

by Giambra (1989)). The probes consisted in the following prompt: "If your attention was completely on the task, press the RED button. If you had other thoughts, press the GREEN button and report the type of thoughts you had using the proposed classification".

Participants had to classify their thoughts in four categories (adapted from Stawarczyk et al., 2011) as follows: i) SITUTs: personal thoughts or daydreams unrelated to the task (e.g., about personal life or worries); ii) TRIs: thoughts related to the task (e.g., thoughts about the task's duration, the stimuli or the participant's performance); iii) EDs: thoughts due to exteroceptive or introceptive perceptions and sensations; and iv) "not known": participants were aware that their mind had wandered, but were unable to say what they had been thinking.

Several examples of MW episodes were given before presenting the perceptual and semantic SARTs. The proposed classification for any MW episodes was presented and participants were asked to classify some examples of MW episodes suggested by the experimenter: if they were able to do so easily, two practice trials (11 stimuli, and 2 probes) were run before the task started; if not, more examples were provided. The term "mind wandering" was avoided when explaining the task.

After completing both versions of the SART, participants were asked to indicate on a Likert scale how difficult it had been to classify their MW episodes (from 0 = very difficult to 6 = very easy), and how accurately they felt they had done so (0 = not at all, 6 = perfectly). A final debriefing was conducted to obtain participants' impressions, particularly focusing on any difficulties encountered.

The order of presentation of the perceptual and semantic SARTs was balanced across participants.

The dependent variables were: 1) frequency of MW episodes computed as a proportion of the total number of probes (the number of MW episodes out of the number of probes); 2) the proportion of each type of thoughts (SITUTs, TRIs, EDs and "not known") out of the total MW episodes (e.g., frequency of SITUTs out of overall MW episodes); and 3) the correct responses, or accuracy, calculating the d-prime index from the percentage of correct hits and correct rejections.

In addition, to better elucidate the level of disengagement from the task, we considered: 4) the proportion of errors, divided – as suggested by Cheyne et al. (2009) – into four categories, i.e., incorrect responses and omissions involving target and non-target stimuli. The proportion of each of these types of error out of the total errors was calculated (e.g., number of omissions in target items out of the total number of errors). We also calculated: 6) the response time latency (standardized RTs) and the response time variability in the four non-target items that were (off-task blocks) or were not (on-task blocks) preceded by a reported MW episode, as well as the pre- and post-error standardized RT latency and variability (see Supplementary data for the latter two indices). We chose to standardize the RTs, dividing the mean RTs by the standard deviation of the RTs for each individual, to take age-related differences in RT latency into account (older adults' RTs were slower than younger adults, due to the age-related decline in processing speed). To compute the variability of the RT latency, we used the response time coefficient of variability — RT CV (i.e., the mean RT divided by the standard deviation of the RT) (see Mrazek et al., 2011).

2.3. Categorization Working Memory Span test (CWMS, Borella et al., 2008)

This task is similar to the classic working memory tasks, the only difference being that it involves processing lists of words rather than sentences to limit the influence of semantic processing. Four sets were presented: each set included 18 lists of words grouped into series of 3 to 6 word lists. Each list contained 5 words of high-to-medium frequency. The lists could contain no, one, or two animal nouns, in any position, including the last (a typical list could be: year, mother, dog, word, night).

Participants heard the lists of words, which were read by the experimenter at a rate of 1 s per word, and were asked to tap their hand on the table whenever they heard an animal noun. The interval between two lists of words was 2 s (the presentation was paced by the experimenter). At the end of the series, participants had to recall the last word in each list in serial order. Two practice trials were given before the experiment started.

The total number of correctly recalled words was taken as the measure of their working memory capacity (maximum score of 72).

The proportion of intrusion errors (words presented in the task and recalled by the participant that were not the last words in each list) was also computed to measure the individual's ability to control the persistence of information in their working memory (see, for example, Borella, Carretti, Cornoldi, & De Beni, 2007; De Beni, Palladino, Pazzaglia, & Cornoldi, 1998), which represents a measure of inhibitory efficacy (e.g., Robert et al., 2009).

2.3.1. Pattern comparison task (adapted from Salthouse & Babcock, 1991)

This task consists of two columns of 60 line segments set out on two pages. Participants had to decide as quickly as possible whether two items were identical (writing *S*, for *Si* [*Yes*]) or not (writing *N*, for *No*). The experimenter used a stopwatch to record the time it took to complete each page. Three practice items were given before the experiment started.

The dependent variable was the total time taken to complete the answers for the two pages.

2.3.2. MW questionnaire (adapted from Borella et al., 2007)

This new questionnaire was developed specifically to assess attentional control failures (possibly reflecting MW episodes) that could occur in Italian people's everyday life (see Appendix 1). A pilot study was first conducted to collect descriptions of the most common situations in which people interviewed had persistent thoughts and the things that often happened when they were distracted. These descriptions were used to obtain 29 items for assessing the frequency of intrusive thoughts (e.g., "How often does it happen that you cannot prevent thoughts that are disturbing you and slowing down your performance in daily life?"), and cognitive failures due to inattention in everyday life (e.g., "How often does it happen that you do not remember what a person has just said to you because you have not been paying attention?"). Participants had to rate the frequency with which they made these mistakes or had these thoughts on a 5-point scale (from *Never* — 1, to *All the time* — 5); they scored from 29 to 145, a higher score indicating a higher frequency of intrusive thoughts. At the end of the questionnaire, there were three questions to assess participants' beliefs about this phenomenon (their answers were not considered in the total score).

2.4. Procedure

All participants were tested individually in a quiet, well-lit room during a single session that lasted about 90 min. The tasks were presented in the following order: the health and demographics questionnaire, the vocabulary test, the first version of the SART (semantic or perceptual, balanced across participants), the processing speed task (pattern comparison), the working memory task, the second version of the SART, and the MW questionnaire.

3. Results

First, ANOVA were carried out on the MW questionnaire, the working memory task, the proportion of intrusion errors in the CWMS (inhibitory measure), and the processing speed, with group as the independent variable. The results are summarized in Table 1. Then the results of the SARTs were analyzed (see below).

3.1. MW questionnaire

The old–old reported significantly fewer intrusive thoughts and cognitive failures than the young adults ($p < .01$), while the young–old did not differ significantly from either of the other two groups.

3.2. Cognitive resources

The young adults recalled more words in the working memory task (CWMS) than the other two groups ($p < .001$), and the young–old recalled more words then the old–old ($p < .01$).

As concerns in inhibitory efficiency, the old–old again produced a higher proportion of intrusion errors than the young ($p < .001$) or young–old adults ($p < .05$), while the latter two groups did not differ significantly from one another.

Finally, in terms of processing speed, the young adults were faster than the young–old ($p < .001$) or old–old ($p < .001$), and the young–old were faster than the old–old ($p < .01$) (see Table 1).

As in the literature, older adults had a poor working memory performance and a lower processing speed and inhibitory efficacy than younger people, and the old–old performed worse than the young–old. These results confirm the age-related decline in cognitive resources and the more accentuated cognitive decline in the fourth age than in the third (e.g., Borella et al., 2007).

3.3. Perceptual and semantic SARTs

Descriptive statistics are presented in Table 2.

Repeated ANOVAs with a mixed design, using Group (young, young–old, old–old) as a between-subjects factor, and Condition

Table 2
Descriptive statistics (M and SD) for the measures of interest by SART version (perceptual vs. semantic) and Group (young, young–old and old–old).

		Young		Young–old		Old–old	
		M	SD	M	SD	M	SD
Perceptual SART							
MW episodes	Overall MW	0.59	0.25	0.38	0.32	0.16	0.15
	SITUTs	0.43	0.23	0.29	0.27	0.24	0.29
	TRIs	0.28	0.22	0.35	0.30	0.23	0.28
	EDs	0.27	0.17	0.24	0.29	0.24	0.28
	Not known	0.01	0.02	0.01	0.06	0.03	0.12
Accuracy	*d* prime index	4.56	0.24	4.53	0.27	4.30	0.53
Incorrect responses	Target item	0.30	0.13	0.35	0.13	0.36	0.13
	Non-target item	0.45	0.14	0.10	0.14	0.16	0.14
Omissions	Target item	0.15	0.11	0.10	0.11	0.37	0.11
	Non-target item	0.35	0.33	0.60	0.33	1.79	0.34
RTs	On-task	0.08	1.30	−0.14	1.90	−0.01	2.97
	Off-task	−0.25	0.21	−0.26	0.16	−0.23	0.12
RT CV	On-task	4.64	1.87	6.51	3.02	5.79	2.51
	Off-task	4.40	1.24	6.30	3.14	8.18	6.35
Semantic SART							
MW episodes	Overall MW	0.61	0.33	0.31	0.32	0.23	0.26
	SITUTs	0.40	0.24	0.27	0.26	0.19	0.29
	TRIs	0.30	0.22	0.37	0.33	0.56	0.34
	EDs	0.24	0.15	0.22	0.30	0.09	0.17
	Not known	0.01	0.02	0.04	0.11	0.01	0.02
Accuracy	d prime index	4.33	0.51	4.10	0.76	3.64	1.04
Incorrect responses	Target item	1.20	0.43	0.80	0.44	1.63	0.45
	Non-target item	0.20	0.27	0.49	0.27	1.12	0.27
Omissions	Target item	0.15	0.30	0.65	0.30	1.05	0.30
	Non-target item	0.95	1.16	1.45	1.16	4.21	1.20
RT latency	On-task	−0.15	0.20	−0.23	0.20	−0.18	0.12
	Off-task	0.04	0.24	−0.14	0.32	0.02	0.73
RT variability (RT CV)	On-task	4.47	1.23	6.36	3.06	6.05	1.97
	Off-task	4.64	1.54	8.26	5.41	8.89	6.95

Note: SITUTs = proportion of personal thoughts; TRIs = proportion of thoughts related to task; EDs = proportion of thoughts due to external stimuli; RT: response times; RT CV: response time coefficient of variability.

(perceptual vs. semantic SART) as a within-subject factor, were run on the following dependent variables, calculated for each participant: i) proportion of the frequency of MW episodes; ii) proportion of the frequency of MW episodes by type of thought (SITUTs, TRIs, EDs, "not known"; iii) accuracy (*d*-prime index); iv) incorrect responses involving target stimuli or non-target stimuli; and v) omissions involving target stimuli or non-target stimuli. We also analyzed response latency in blocks in which MW episodes were or were not reported.

The interactions were broken down using post-hoc pairwise comparisons with Bonferroni's correction at $p < .05$, adjusted for multiple comparisons. The results of the ANOVA are summarized in Table 3.

3.3.1. Frequency of MW episodes

On a descriptive level, the frequency of MW episodes was higher in young than in young–old or old–old adults in both SARTs but – unlike the other two age groups – the old–old adults reported a higher frequency of MW episodes in the semantic than in the perceptual SART (see Table 2).

ANOVA showed that, regardless of the SART Condition, young adults reported a higher frequency of MW episodes than the young–old ($p < .01$) or old–old ($p < .001$), while there was no difference between the latter two age groups (see Table 3).

3.3.2. Content of MW episodes[2]

The main effect of Group emerged in the proportion of SITUTs. The old–old reported a lower proportion of SITUTs than the young adults ($p < .05$), irrespective of the version of the SART, while the young–old adults did not differ significantly from the other two groups (see Table 2).

The main effect of Condition was significant for the proportion of TRIs: participants reported a greater proportion of TRIs in the semantic than in the perceptual SART. For TRIs, the significant interaction (Group × Condition) also revealed that old–old participants reported proportionally more TRIs than young adults in the semantic SART ($p < .05$). The old–old also reported more TRIs in the semantic than in the perceptual SART ($p < .01$). The young–old adults did not differ significantly from the other two groups (see Table 3).

No significant results were obtained for the other types of thoughts.

3.3.3. Accuracy

Results in the d-prime index – based on the percentage of correct hits and correct rejections – showed that the old–old were less accurate than the young ($p < .01$), while the young–old did not differ significantly from the other two age groups. On the whole, participants were also less accurate in the semantic than in the perceptual SART ($p < .001$) (see Table 2). The interaction was not significant.

3.3.4. Errors[3]

For incorrect responses involving both target and non-target stimuli, our results showed more incorrect responses in the semantic than in the perceptual SART ($p < .01$ and $p < .05$, respectively) (see Tables 2 and 3). The effect of age was not significant, and the interaction (Group × Condition) was only significant for non-target stimuli, showing that old–old adults produced more incorrect responses in the semantic than in the perceptual SART ($p < .01$).

For the omissions involving both target and non-target stimuli, the old–old tended to omit more responses than the young adults ($p < .05$ for both types of omission), while the young–old adults did not differ significantly from either of the other two age groups. More omissions were made, for both target and non-target stimuli, in the semantic than in the perceptual SART ($p < .05$ for both). No significant interactions emerged for either type of omission (see Table 3).

3.3.5. Response latency in blocks where MW episodes were or were not reported

Repeated ANOVA with a mixed design, with Group (young, young–old, old–old) as a between-subjects factor, and Condition (on-task vs. off-task) as a within-subject factor, were calculated on the means of the standardized RT latency and RT CV for the four non-target stimuli before probing. Descriptive statistics are given in Table 2 and the results of the ANOVA are summarized in Table 4.

Regarding RT latency, our results showed that it was only in the semantic SART that the RTs were longer before a reported MW episode (off-task trials) than in on-task trials ($p < .05$). No significant differences were seen in the perceptual condition.

On the other hand, the analysis of the RT CV pointed to a significant interaction in the perceptual SART, showing that the old–old participants' RTs in off-task blocks varied more than those of the young adults ($p < .05$). It was also only the old–old adults' RTs that varied more in off-task blocks than in on-task blocks ($p < .01$).

The results for the semantic SART only showed a significant main effect of age group: the old–old adults' RTs varied more than those of the young adults ($p < .01$), while the young–old adults did not differ significantly from either of the other two age groups.

The pre-error and post-error latency and variability data are given in the Supplementary data section.

3.4. Correlation analyses

Correlation analyses (see Table 5) were performed considering age, proportion of MW episodes and accuracy (*d*-prime index), RT latency and RT CV in on-task blocks, and off-task block in the perceptual and semantic SARTs, working memory (CWMS word recall), inhibition (intrusion errors in the CWMS), score in the MW questionnaire, and processing speed (pattern comparison).

Significant negative correlations were found between age and both the proportion of MW episodes and the variability of RTs in both SART conditions. Age also correlated with cognitive resources. These findings confirmed the results of the ANOVA.

As expected, our findings also indicated a strong correlation between the frequency of MW episodes in the perceptual and semantic SARTs, while no significant correlations came to light between the accuracy measures in the two versions of the SART. MW in the perceptual SART correlated negatively with RT variability in off-task blocks, but this correlation was no longer significant after controlling for age. More variable RTs in the on-task blocks in the perceptual SART correlated with a greater accuracy, but this effect was also lost after controlling for age. So it seems that there was an increasing variability in RTs with increasing age on the one hand, while on the other there were no significant correlations between RT variability and the frequency of MW episodes reported by participants after controlling for age.

The relationships between the frequency of MW episodes while performing the SARTs and accuracy, cognitive resources and self-perceived MW were assessed in more detail by means of regression analyses, as outlined below.

3.5. Regression analyses

The role of age, cognitive resources (working memory, inhibition, processing speed), and cognitive failures, due to inattention in everyday life, on the frequency of MW episodes (overall MW) experienced by participants during the perceptual and semantic SARTs was assessed using hierarchic regression analyses. Two models, differing in terms of order entry, were used to assess the role of age versus cognitive resources and cognitive failures due to inattention in everyday life — MW Questionnaire. In Model 1, age was entered in Step 1, while

[2] In the questionnaire administered at the end of each version of the SART, participants reported having no difficulty and being highly accurate in classifying their thoughts. The mean values assigned by participants were always higher than 5 for all age groups, which did not differ significantly from one another.

[3] Nineteen and ten participants made no errors in the perceptual and semantic SARTs, respectively.

Table 3
Mixed-design 3 × 2 ANOVA results for the measures of interest, with Group (young, young–old, old–old) as a between-subjects factor and Condition (perceptual vs. semantic SART) as repeated measures.

				F	df	η_p^2	*P*
MW episodes	Overall MW	Between subjects	Group (G)	13.00	2,56	0.15	<.001
		Within subjects	Condition (C)	0.08	1,56	0.00	0.78
			G × S	2.06	2,56	0.07	0.14
	SITUTs	Between subjects	Group (G)	4.56	2,56	0.14	<.05
		Within subjects	Condition (C)	0.66	1,56	0.01	0.42
			G × S	0.06	2,56	0.00	0.94
	TRIs	Between subjects	Group (G)	0.73	2,56	0.03	0.49
		Within subjects	Condition (C)	8.73	1,56	0.14	<.01
			G × S	6.39	2,56	0.19	<.01
	EDs	Between subjects	Group (G)	1.25	2, 56	0.04	0.30
		Within subjects	Condition (C)	3.06	1,56	0.05	0.09
			G × S	0.97	2,56	0.03	0.39
	Not known	Between subjects	Group (G)	0.48	2,56	0.02	0.66
		Within subjects	Condition (C)	0.89	1,56	0.00	0.89
			G × S	2.09	2,56	0.07	0.13
Accuracy	*d* prime index	Between subjects	Group (G)	6.04	2,56	0.18	<.01
		Within subjects	Condition (C)	15.29	1,56	0.21	<.001
			G × S	1.23	2,56	0.04	0.30
Incorrect response	Target item	Between subjects	Group (G)	0.77	2,56	0.03	0.47
		Within subjects	Condition (C)	12.26	1,56	0.18	<.01
			G × S	0.89	2,56	0.03	0.42
	Non-target item	Between subjects	Group (G)	1.51	2,56	0.10	0.23
		Within subjects	Condition (C)	5.32	1,56	0.09	<.05
			G × S	4.61	2,56	0.14	<.05
Omission	Target item	Between subjects	Group (G)	2.97	2,56	0.10	<.05
		Within subjects	Condition (C)	5.32	1,56	0.09	<.05
			G × S	1.38	2,56	0.05	0.26
	Non-target item	Between subjects	Group (G)	3.46	2,56	0.11	<.05
		Within subjects	Condition (C)	4.28	1,56	0.07	<.05
			G × S	0.82	2,56	0.03	0.45

Note: SITUTs = Personal thoughts; TRIs = Thoughts related to task; EDs = Thoughts due to external stimuli.

words recalled in the CWMS, the proportion of intrusion errors in the CWMS, pattern comparison time (sec.), and MW Questionnaire score were entered in Step 2. For Model 2, the order of entry of the various variables was reversed, with CWMS, intrusion errors, pattern comparison time and MW Questionnaire in Step 1 and age in Step 2.

Concerning the frequency of MW episodes reported during the perceptual SART, Model 1 showed that age explained a significant part of the variance ($R^2 = .54$, $p < .001$), while the MW Questionnaire (Step 2) accounted for 16% of the variance; the CWMS ($\beta = .50$, $p < .05$), and the MW Questionnaire ($\beta = .35$, $p < .01$) were the only salient predictors. On the other hand, when age was entered in the last step (Model 2) only cognitive resources and cognitive failures due to inattention in everyday life explained a significant part of the variance ($R^2 = .44$, $p < .001$), and age was no longer a significant predictor. Here again, CWMS ($\beta = .50$, $p < .05$) and the MW Questionnaire score ($\beta = .35$, $p < .01$) made a unique contribution to the explained variance.

When we analyzed the predictors of the frequency of MW episodes during the semantic SART, only age ($\beta = -.49$, $p < .001$) explained a significant part of the variance ($R^2 = .24$, $p < .001$) in Model 1, whereas cognitive resources and MW Questionnaire, but not age, explained a significant amount of variance in Model 2 ($R^2 = .29$, $p < .001$). The MW Questionnaire made a unique contribution to the variance explained ($\beta = .35$, $p < .01$).

Table 4
Mixed-design 3 × 2 ANOVA results for standardized response time (RT) latency and response time coefficient of variability (RT CV) in semantic and perceptual SART, with Group (young, young–old, old–old) as a between-subjects factor and Condition (on-task vs. off-task blocks) as repeated measures.

			F	df	η_p^2	*p*
Perceptual SART						
RT latency	Between subjects Within subjects	Group (G)	0.06	2,49	0.00	0.95
		Condition (C)	0.55	1,49	0.01	0.46
		G × S	0.05	2,49	0.00	0.96
RT CV latency	Between subjects Within subjects	Group (G)	2.97	2,49	0.11	0.06
		Condition (C)	1.47	1,49	0.03	0.23
		G × S	4.14	2,49	0.15	<.05
Semantic SART						
RT latency	Between subjects Within subjects	Group (G)	1.33	2,46	0.06	0.28
		Condition (C)	4.67	1,46	0.09	<.05
		G × S	0.22	2,46	0.01	0.80
RT CV latency	Between subjects Within subjects	Group (G)	4.29	2,46	0.16	<.05
		Condition (C)	3.46	1,46	0.07	0.07
		G × S	0.63	2,46	0.03	0.51

Note: RT: response times; RT CV: response time coefficient of variability.

Table 5
Correlations between measures of interest.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1. Age	–																
2. Perceptual SART (MW)	−.54**		.78**	−.19	.06	.00	−.24	−.19	−.16	.04	−.24	.01	−.36*	.41**	.17	−.02	−.11
3. Semantic SART (MW)	−.49**	.73**		−.18	−.18	−.14	.06	−.31	−.14	.08	−.15	−.03	−.38	.31*	.10	.03	−.12
4. Perceptual SART (accuracy)	−.25	−.05	.01		.02	−.23	.24	.14	.06	.35	.37	.40	.36	−.09	.28	.30	.02
5. Semantic SART (accuracy)	−.28	.15	−.10	.13		.02	−.14	.05	.06	.11	−.04	.19	.27	−.03	.37*	−.08	−.18
6. Perceptual SART (RTs on-task)	−.03	.03	−.18	−.07	.11		−.31	.13	−.07	−.34*	−.26	−.42**	−.23	−.24	−.12	.01	.01
7. Perceptual SART (RTs off-task)	−.10	−.19	−.02	.12	.07	.07		.16	.30	.06	.33*	.31	.22	−.01	−.04	.08	−.05
8. Semantic SART (RTs on-task)	−.10	−.14	−.19	.12	.08	.07	.17		.06	.19	.20	−.02	.32	−.23	.22	−.16	−.13
9. Semantic SART (RTs off-task)	−.05	−.08	−.11	.03	−.02	−.18	−.18	−.01		.10	.25	.16	.38	−.07	.04	−.01	.13
10. Perceptual SART (RT CV on-task)	.31*	−.26	−.17	.30*	.05	−.29*	.19	.00	.14		.61**	.22	.36*	−.08	.33*	−.05	−.34*
11. Perceptual SART (RTCV off-task)	.36**	−.42**	−.43**	.08	.14	−.14	.21	.22	.23	.41**		.31	.44**	−.09	.20	−.09	−.37*
12. Semantic SART (RTCV on-task)	.33*	−.28*	−.16	.23	.07	−.20	.06	.22	.05	.43**	.33*		.36*	.07	.24	−.14	−.01
13. Semantic SART (RTCV off-task) self-report of MW	.31*	−.31*	−.23	.20	−.01	−.13	.25	.10	.13	.39**	.54**	.66**		.02	.36*	−.30	−.06
14. MW Questionnaire working memory	−.40**	.48**	.42**	.12	.13	−.17	−.08	−.13	−.06	.01	−.05	−.11	−.06		−.17	.23	−.16
15. CWMS (words recalled) inhibition	−.68**	.52**	.36**	.29*	−.48**	.08	−.10	.12	.04	−.08	.02	−.13	−.03	.22		−.74**	−.12
16. Proportion of intrusions processing speed	.47**	−.30*	−.20	−.29*	−.20	−.17	.14	−.09	−.03	−.01	−.06	−.17	.08	−.03	−.74**		−.10
17. Pattern comparison (sec.)	.64**	−.41**	−.36**	−.10	−.35**	−.04	.01	−.10	.08	.18	.20	.20	.05	−.55**	.29*	−.36**	

Note: Raw correlations are presented below the diagonal; correlations above the diagonal are controlled for age.
TUTs: Task-unrelated thoughts; RT: response times; RT CV: response time coefficient of variability; CWMS: Categorization Working Memory Span test; MW: mind wandering episodes.
* $p < .05$.
** $p < .01$.

It is worth noting that similar results were obtained, in terms of significant predictors, when the proportion of SITUTs (which is a stricter measure of MW) was considered as a dependent variable[4].

Regression analyses were also run to estimate the percentage of variance in accuracy (*d*-prime) in the perceptual and semantic SARTs, considering age, CWMS, proportion of intrusion errors, score in the MW Questionnaire, processing speed, and reported frequency of MW during the two SARTs, and entering the variables as explained above for the previous analysis. None of the predictors accounted for a significant part of the variance in accuracy in the perceptual SART. In contrast, age explained a significant part of the variance ($R^2 = .07$, $p < .05$) for the semantic SART in Model 1, and cognitive resources and MW (the MW Questionnaire and the proportion of MW episodes) added 29% of the variance. The CWMS score ($\beta = .79$, $p < .01$) and the frequency of MW episodes in the semantic SART ($\beta = -.33$, $p < .01$) thus made a unique contribution to explaining the variance. But when cognitive aspects were entered in the first step, the contribution of age was no longer significant, and cognitive resources and MW accounted for a larger portion of the variance ($R^2 = .36$, $p < .001$). In particular, the CWMS ($\beta = .79$, $p < .001$), and the frequency of reported MW episodes ($\beta = -.40$, $p < .01$) made a unique contribution to explaining the variance in accuracy.

Overall, these results suggest, on the one hand, that age explains a large part of the variance in MW reported by participants in both versions of the SART, while working memory and the MW Questionnaire only contributed significantly in the perceptual task. On the other hand, working memory explained the majority of the variance in accuracy (while the contribution of age was limited), but only in the more demanding version of the SART.

4. Discussion and conclusion

MW is a phenomenon that has recently received considerable attention, but few studies have assessed whether and how it changes with age, also in relation to the demands of a given task. To the best of our knowledge, none of the studies conducted to date have considered the relationship between MW and cognitive resources (working memory, inhibition, processing speed) in aging. We therefore newly investigated the age-related differences in MW between young, young–old and old–old individuals using a multivariate design to analyze: i) whether the frequency of MW was modulated by the demands of the task in terms of the processing resources required (using both the probe-caught method and performance-based measures), and for this purpose we considered the amount and type of reported MW episodes, also distinguishing between different types of thoughts; and ii) the relationship between MW and mechanisms evoked to explain MW, such as working memory, inhibition, and processing speed, as well as the relationship between MW and the self-perceived frequency of intrusive thoughts and cognitive failures due to inattention in daily life, assessed by means of a new MW Questionnaire.

Concerning the first goal, young–old and old–old adults reported fewer MW episodes than young adults in both versions (perceptual and semantic) of the SART administered — a finding consistent with other studies (e.g., Jackson & Balota, 2012) and with the decoupling

[4] When the proportion of SITUTs among all the probes was considered as a dependent variable in the perceptual SART, Model 1 showed that only age ($\beta = -.41$, $p < .01$) explained a significant part of the variance ($R^2 = .17$, $p < .05$); cognitive resources and the score in the MW questionnaire (Step 2) did not contribute to explaining the variance. In contrast, when age was entered in the last step (Model 2), only CWMS ($\beta = .57$, $p < .05$) and the score in the MW questionnaire ($\beta = .30$, $p < .05$) made a unique contribution to the explained variance ($R^2 = .24$, $p < .01$), and age (Step 2) was no longer a significant predictor. For the SITUTs in the semantic SART, only age ($\beta = -.52$, $p < .001$), not cognitive resources or the score in the MW questionnaire, explained a significant part of the variance ($R^2 = .27$, $p < .001$) in Model 1. In Model 2, cognitive resources and the score in the MW questionnaire explained 31% of the variance, and age (Step 2) an additional 4%. The score in the MW questionnaire ($\beta = -.37$, $p < .05$) and age ($\beta = -.37$, $p < .05$) made a unique contribution to the explained variance.

hypothesis. Some noteworthy findings emerged, however, when we analyzed the types of thought reported when MW episodes occurred. Our results showed an age-related decline in the proportion of personal thoughts (SITUT) regardless of the type of SART, with the old–old reporting fewer SITUTs then younger adults. This may be due to age-related changes in emotional goals and regulation (e.g., Carstensen & Charles, 1998), meaning that the old–old may have fewer personal thoughts because they are less anxious about everyday life issues or, as suggested by Giambra (1989), they could have less "unfinished business". The old–old also reported a higher proportion of task-related thoughts (TRI) than younger adults. As Jackson and Balota (2012) suggested that older adults (and the old–old in particular in the present study) may approach such tasks differently from younger people, and be more inclined to engage in the tasks, for instance. It could also be that the cue provided by the words in the semantic SART facilitated MW by creating more interferences than perceptual stimuli (Carretti, Mammarella, & Borella, 2011). This was mentioned by participants afterwards and might also explain why old–old adults reported having more TRIs in the semantic than in the perceptual SART. McVay et al. (2013) recently found a similar difference between the TRIs and the SITUTs reported by older adults. The authors interpreted this pattern of findings in the light of the 'control failure × current concerns' hypothesis, judging it to be due to reduced control processes interacting with the generation of thoughts about current concerns. Older adults may have fewer worries and thoughts about their day-to-day life, but more off-task thoughts about their performance in the laboratory, which might go unsuppressed due to executive deficits. It is particularly worth noting, regarding both SITUTs and TRIs, that we found differences between young and old–old adults, while the young–old's performance seemed to come in between the other two age groups. The present results thus suggest that it is not older adults in general, but only old–old adults who differ in their sensitivity to these types of thoughts.

These age-related differences may, nonetheless, be due to a lesser propensity with aging to report some types of MW episode, either because they may be seen as a normal product of a person's mental activity, or because older people would have greater difficulty in monitoring them (Einstein & McDaniel, 1997) while simultaneously completing the task in hand. To take these aspects into account, our participants' performance indices were also analyzed, revealing some notable differences vis-à-vis the self-reported measures of MW considered in the present study.

One aspect concerned the task's manipulation: as expected, our results indicated that the semantic SART was more difficult than the perceptual version, since we observed a lower accuracy in the former, with more omissions and incorrect responses, particularly in the old–old adults. When we considered the age-related differences, our results showed that only the old–old were less accurate in both versions of the SART than the younger groups, while the young–olds' performance (like their reported MW episodes) came in between the other two age groups. In particular, the old–old adults produced more omissions for non-target stimuli in both SARTs, suggesting that their MW went deeper and they were less aware of it. This lesser awareness could explain the lower frequency of reported MW and also its greater influence on the old–old adults' performance. According to Cheyne et al. (2009), omissions represent a failure to prepare a response due to exogenous events and endogenous rhythmic attention swings, and may become an important opportunity for reactive MW. Such errors could indicate a MW state that is less likely to be conscious (Smallwood, McSpadden, & Schooler, 2007), and that can interfere with the primary task (Cheyne et al., 2009).

Another significant behavioral issue regards the RTs, and it proved particularly interesting to compare RT latency and variability in blocks of stimuli that were or were not reportedly preceded by participants' MW episodes. Our participants proved slower to respond when their minds wandered than when they were focused on the task, but only when performing the semantic SART, not the perceptual task. Performance in the semantic task thus seemed to be more sensitive to MW episodes, suggesting that the semantic task was more difficult than the perceptual SART because of the nature of the stimuli. When age differences were considered, we found that only the old–old adults' RTs varied more when their minds wandered than when they were focused on the task, and this applied to both versions of the task. This can be seen as an indication of a transient disengagement of their attention from the task in hand (Cheyne et al., 2009), when their attention was decoupled between internal and external information, a state in which errors in performance become more likely, as suggested by Smallwood, McSpadden, et al. (2007), Smallwood, O'Connor, et al. (2007). When we considered the pre- and post-error latency and variability (see Supplementary data) instead, we found that only young adults, not older adults (as might be expected), were faster before an error than before a correct response in the semantic SART, and their RT variability was greater before errors than before correct responses, in both the perceptual and the semantic SARTs. These results are at odds with previous findings (Cheyne et al., 2009; Jackson & Balota, 2012; Smallwood et al., 2004), however, and they should be considered with caution because of the limited data available for our analysis.

Finally, regression analyses enabled us to clarify the role of basic cognitive mechanisms in MW: the frequency of MW episodes was explained by age, cognitive failures and intrusive thoughts in everyday life and individual differences in cognitive resources, and in working memory in particular. The same predictors emerged when we considered a strict measure of MW (i.e., the proportion of SITUTs out of the number of probes). Therefore, people who reported more frequent cognitive failures and intrusive thoughts in everyday life, and who had a higher working memory performance, reported more MW episodes but, oddly enough, only in the less demanding (perceptual) task. They probably found it easy to complete the task and this enabled their minds to wander. The findings for the perceptual SART would be consistent with the decoupling hypothesis, but it would be hard to explain why working memory capacity does not predict the frequency of MW in the semantic SART too (Levinson, Smallwood, & Davidson, 2012). This could have to do with the greater influence of individual differences in the more demanding task, as predicted by the 'control failure × current concerns' view (Kane & McVay, 2012), but it is important to note that inhibitory efficacy did not seem to play a part in explaining the reported frequency of MW episodes, as predicted in younger adults (Kane & McVay, 2012). Only intrusion errors were considered, however, a different involvement of inhibitory efficacy in MW might emerge using other inhibitory measures. We might also attribute the differences between the two SARTs to the fact that our older people may have been less able to monitor their MW when they had to cope with a more demanding task. Finally, our finding may be due simply to variables (e.g., the type of stimuli used) relating to the type of task: these aspects need to be better investigated in future studies.

A different pattern emerged for accuracy: regression analyses showed that cognitive resources (and especially working memory capacity) and age (to a marginal degree) explained the variance in the accuracy of response in the semantic SARTs. As expected, the role of working memory was stronger in the more demanding task. The frequency of MW episodes also influenced accuracy in this more difficult condition. Regression analyses also showed a close relationship between cognitive failures due to inattention in everyday life (MW Questionnaire) and the frequency of MW episodes. The particular interest of these findings lies in that they demonstrate the importance not only of an age-related decline in cognitive mechanisms, and working memory capacity in particular, but also of individual differences in cognitive resources and non-cognitive mechanisms (e.g., emotions, personal goals, interest in the tasks) and/or difficulties in monitoring MW, in explaining why older adults are less likely to report MW. It is also worth noting that this result is consistent with some studies on younger adults that highlighted the role of mood (e.g., Smallwood, McSpadden, et al., 2007; Smallwood, O'Connor, et al., 2007), personal goals (e.g. Stawarczyk et al., 2011), and beliefs about personal ability (Mrazek et al., 2011) in MW.

Our results thus suggest an important role for working memory in explaining both the frequency of MW episodes and performance in the SART, but this seems to depend on the difficulty of the task. This interesting relationship between working memory and the resources demanded by a task, and especially the insignificant role of working memory in the more demanding (semantic) SART warrants further, more detailed studies to confirm our present findings. Future studies on this phenomena should also include non-cognitive factors to assess the relationships emerging from the present results more thoroughly. In contrast, our MW Questionnaire had a significant role in both the perceptive and the semantic SART, showing a close relationship between the self-reported measurement of intrusive thoughts and cognitive failures in everyday life (MW Questionnaire) and MW as measured by experimental tasks (SARTs) in our older people. These results extend to older adults the findings already reported in younger populations (Cheyne et al., 2006, 2009).

In conclusion, this study found that older adults reported fewer MW episodes than younger adults, and the content of these MW episodes changed in relation to the age group considered: the old–old reported proportionally fewer personal thoughts and more task-related thoughts than younger adults. Meanwhile, the age-related lower accuracy in performing the tasks, and some behavioral indicators of disengagement from the task (particularly in old–old adults) – though not necessarily an indication of MW episodes – seemed to suggest that MW interfered with older adults' performance. Younger adults had more such episodes, but appeared to be more aware of them and largely able to control their mind wandering, so their performance was unaffected. Although older adults reported experiencing fewer MW episodes, they may have been less able to refocus on the task in hand when it became necessary to do so, so their MW had more evident detrimental effects on their performance. In other words, there is probably a lower production of MW episodes in older people, particularly in very old age, accompanied by a weaker ability to suppress these thoughts. Given the discrepancy between performance-based measures and the frequency of MW reported by our older adults, future studies should use more objective procedures (e.g., psychophysiological indices) to identify any differences between the MW patterns of younger and older adults that probe-based methods are unable to capture (e.g., longer episodes of disengagement, difficulty in refocusing on the task after a MW episode has occurred) and, more importantly, the old–old adults' responses to the probe, which may be affected by their monitoring difficulties.

Future studies could also focus on finding any differences in the MW patterns of older versus younger adults (distinguishing between young–old and old–old adults as in the present study), and other cognitive and non-cognitive mechanisms that might explain the greater influence of MW on older people's performance. If this stronger influence of intrusive thoughts in advanced old age is confirmed, it would be useful to develop methods for containing aging adults' MW.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.actpsy.2013.10.016.

References

Antrobus, J. S., Singer, J. L., & Greenberg, J. (1966). Studies in the stream of consciousness: Experimental suppression of spontaneous cognitive process. *Perceptual and Motor Skills*, *23*, 399–417.

Baird, B., Smallwood, J., Mrazek, M.D., Kam, J., Franklin, M. S., & Schooler, J. W. (2012). *Mind-wandering facilitates creative incubation.*

Baird, B., Smallwood, J., & Schooler, J. W. (2011). Back to the future: Autobiographical planning and functionality of mind wandering. *Consciousness and Cognition*, *20*, 1604–1611.

Balota, D. A., Black, S. R., & Cheney, M. (1992). Automatic and attentional priming in young and older adults: Reevaluation of the two process model. *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 485–502.

Baltes, P. B. (1987). Theoretical propositions of life-span developmental psychology: On the dynamics between growth and decline. *Developmental Psychology*, *23*, 611–626.

Barron, E., Riby, L. M., Greer, J., & Smallwood, J. (2011). Absorbed in thought: The effect of mind wandering on the processing of relevant and irrelevant events. *Psychological Science*, *22*, 596–601.

Borella, E., Carretti, B., Cornoldi, C., & De Beni, R. (2007). Working memory, control of interference and everyday experience of thought interference: When age makes the difference. *Aging Clinical and Experimental Research*, *19*, 1–7.

Borella, E., Carretti, B., & De Beni, R. (2008). Working memory and inhibition across the adult life-span. *Acta Psychologica*, *128*, 33–44.

Borella, E., Ghisletta, P., & de Ribaupierre, A. (2011). Age differences in text processing: The role of working memory, inhibition and processing speed. *Journal of Gerontology: Psychological Sciences*, *66*, 311–320.

Carretti, B., Mammarella, I. C., & Borella, E. (2011). Age differences in proactive interference in verbal and visuo-spatial working memory. *Journal of Cognitive Psychology*, *1–13*.

Carstensen, L. L., & Charles, S. T. (1998). Emotion in the second half of life. *Current Directions in Psychological Science*, *5*, 144–149.

Cheyne, J. A., Carriere, J. S. A., & Smildek, D. (2006). Absent-mindedness: Lapses in conscious awareness and everyday cognitive failures. *Consciousness and Cognition*, *15*, 578–592.

Cheyne, J. A., Solman, G. J. F., Carriere, J. S. A., & Smildek, D. (2009). Anatomy of an error: A bidirectional state of task engagement/disengagement and attention-related errors. *Cognition*, *111*, 98–113.

Craik, F. I. M., & Salthouse, T. A. (2008). *Handbook of aging and cognition* (3rd ed.) Mahwah, NJ: Lawrence Erlbaum.

Crook, T., Bartus, R. T., Ferris, S. H., Whitehouse, P., Cohen, G. D., & Gershon, S. (1986). Age-associated memory impairment: Proposed diagnostic criteria and measures of clinical change. Report of a National Institute of Mental Health work group. *Developmental Neuropsychology*, *2*, 261–276.

De Beni, R., Palladino, P., Pazzaglia, F., & Cornoldi, C. (1998). Increases in intrusion errors and working memory deficit of poor comprehenders. *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology*, *51a*, 305–320.

Einstein, G., & McDaniel, M. (1997). Aging and mind wandering: Reduced inhibition in older adults? *Experimental Aging Research*, *23*, 343–354.

Giambra (1973). Daydreaming in males from seventeen to seventy-seven: A preliminary report. *Proceedings of the Annual Convention of the American Psychological Association*, 769–770.

Giambra, L. M. (1989). Task-unrelated thought frequency as a function of age: A laboratory study. *Psychology and Aging*, *4*, 136–143.

Giambra, L. M. (1993). The influence of aging on spontaneous shifts of attention from external stimuli to the contents of consciousness. *Experimental Gerontology*, *28*, 485–492.

Giambra, L. M. (1995). A laboratory-based method for investigating influences on switching attention to task-unrelated imagery and thought. *Consciousness and Cognition*, *4*, 1–21.

Giambra, L. M. (2000). Daydreaming characteristics across the life-span: Age differences and seven to twenty year longitudinal changes. In R. Kunzendorf, & B. Wallace (Eds.), *Individual differences in conscious experience*. Amsterdam: John Benjamins.

Grodsky, & Giambra (1990–1991). The consistency across vigilance and reading tasks of individual differences in the occurrence of task-unrelated and task-related images and thoughts. *Imagination, Cognition and Personality*, *10*, 39–52.

Hasher, L., & Zacks, 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* (pp. 193–225). New York: Academic Press.

Jackson, J.D., & Balota, D. A. (2012). Mind-wandering in younger and older adults: Converging evidence from the sustained attention to response task and reading for comprehension. *Psychology and Aging*, *27*, 106–119.

Kane, J. K., Brown, L. K., McVay, J. C., Silvia, P. J., Myin-Geremeys, I., & Kwapil, R. T. (2007). For whom the mind wanders, and when: An experience-sampling study of working memory and executive control in daily life. *Psychological Science*, *18*, 614–621.

Kane, M. J., & McVay, J. C. (2012). What mind wandering reveals about executive-control abilities and failures. *Current Directions in Psychological Science*, http://dx.doi.org/10.1177/0963721412454875.

Levinson, D. B., Smallwood, J., & Davidson, R. J. (2012). The persistence of thought: Evidence for a role of working memory in the maintenance of task-unrelated thinking. *Psychological Science*, *23*, 375–380.

McVay, J., & Kane, M. J. (2009). Conducting the train of thought: Working memory capacity, goal neglect, and mind wandering in executive-control task. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *31*, 196–204.

McVay, J. C., & Kane, M. J. (2010). Does mind wandering reflect executive function or executive failure? Comment on Smallwood and Schooler (2005) and Watkins (2008). *Psychological Bulletin*, *136*, 188–197.

McVay, J., Kane, M. J., & Kwapil, R. T. (2009). Tracking the train of thought from the laboratory into everyday life: An experience-sampling study of mind wandering across controlled and ecological context. *Psychonomic Bulletin & Review*, *16*, 857–863.

McVay, J. C., Meier, M. E., Touron, D. R., & Kane, M. J. (2013). Aging ebbs the flow of thought: Adult age differences in mind wandering, executive control, and self-evaluation. *Acta Psychologica*, *143*, 136–147.

Mrazek, M.D., Chin, J. M., Schmader, T., Hartson, K. A., Smallwood, J., & Schooler, J. W. (2011). Threatened to distraction: Mind wandering as a consequence of stereotype threat. *Journal of Experimental Psychology*, *47*, 1243–1248.

Parks, C. W., Klinger, E., & Perlmutter, M. (1988–89). Dimensions of thought as a function of age, gender and task difficulty. *Imagination, Cognition and Personality*, *8*, 49–62.

Robert, C., Borella, E., Fagot, D., Lecerf, T., & de Ribaupierre, A. (2009). Working memory and inhibitory control across the life span: Intrusion errors in the reading span task. *Memory & Cognition*, *37*, 336–345.

Salthouse, T. A., & Babcock, R. L. (1991). Decomposing adult age differences in working memory. *Developmental Psychology*, *27*, 763–776.
Schupak, C., & Rosenthal, J. (2009). Excessive daydreaming: A case history and discussion of mind wandering and high fantasy proneness. *Consciousness and Cognition*, *18*, 290–292.
Smallwood, J. (2010). Why the global availability of mind wandering necessitates resource competition: Reply to McVay and Kane. *Psychological Bulletin*, *136*, 202–207.
Smallwood, J., Davies, J. B., Heim, D., Finnegan, F., Sudberry, M. V., O'Connor, R. C., et al. (2004). Subjective experience and the attentional lapse. Task engagement and disengagement during sustained attention. *Consciousness and Cognition*, *4*, 657–690.
Smallwood, J., McSpadden, M., & Schooler, J. W. (2007). The lights are on but no one's home: Meta-awareness and the decoupling of attention when the mind wanders. *Psychonomic Bulletin and Review*, *14*, 527–533.
Smallwood, J., O'Connor, R. C., Sudberry, M. V., & Obonsawin, M. C. (2007). Mind-wandering and dysphoria. *Cognition and Emotion*, *21*, 816–842.
Smallwood, J., Obonsawin, M. C., & Heim, S. D. (2003). Task-unrelated thought: The role of distributed processing. *Consciousness and Cognition*, *12*, 169–189.
Smallwood, J., Riby, L., Heim, D., & Davies, J.D. (2006). Encoding during the attentional lapse: Accuracy of encoding during the semantic SART. *Consciousness and Cognition*, *15*, 218–231.
Smallwood, J., & Schooler, J. W. (2006). The restless mind. *Psychological Bulletin*, *6*, 946–958.
Stawarczyk, D., Majerus, S., Maj, M., Van Der Linden, M., & D'Argembeau, A. (2011). Mind-wandering: Phenomenology and function as assessed with a novel experience sampling method. *Acta Psychologica*, *136*, 370–381.
Wechsler, D. (1981). *Wechsler Adult Intelligence Scale — Revised manual*. New York: The Psychological Corporation.