Working memory in healthy aging and in Parkinson's disease: Evidence of interference effects

Elisa Di Rosa [1]*, Doris Pischedda [2,3], Paolo Cherubini [2,3], Daniela Mapelli [1,4,] Stefano Tamburin [5], Michele Burigo [6]

[1] Department of General Psychology, University of Padova, Padova, Italy
[2] Department of Psychology, University of Milano-Bicocca, Milan, Italy
[3] NeuroMI - Milan Center for Neuroscience, Milan, Italy
[4] Human Inspired Technologies Research Center, University of Padova, Padova, Italy
[5] Department of Neurosciences, Biomedicine and Movement Sciences, Neurology Section, University of Verona, Verona, Italy
[6] Cognitive Interaction Technology (CITEC), Bielefeld University, 33615, Bielefeld, Germany

***Corresponding Author:**

Elisa Di Rosa
Department of General Psychology,
University of Padova
Via Venezia, 835100 Padova, Italy
Phone: ++39 049 8276957
elisa.dirosa@unipd.it

Co-authors contacts:

Doris Pischedda: doris.pischedda@unimib.it
Paolo Cherubini: paolo.cherubini@unimib.it
Daniela Mapelli: daniela.mapelli@unipd.it
Stefano Tamburin: stefano.tamburin@univr.it
Michele Burigo: mburigo@cit-ec.uni-bielefeld.de

Running head: Task-rule inhibition in working memory

Acknowledgements: We wish to thank Dr. Alice Maier and Dr. Viola Bulgari for their help in data collection.

ABSTRACT

Focusing on relevant information while suppressing the irrelevant one are critical abilities for different cognitive processes. However, their functioning has been scarcely investigated in the working memory (WM) domain, in both healthy and pathological conditions. The present research aimed to study these abilities in aging and Parkinson's disease (PD), testing three groups of healthy participants (young, older and elderly) and one of PD patients, employing a new experimental paradigm. Results showed that the transient storing of irrelevant information in WM causes substantial interference effects, which were remarkable in elderly individuals on both response latency and accuracy. Interestingly, PD patients responded faster and were equally accurate compared to a matched control group. Taken together, findings confirm the existence of similar mechanisms for orienting attention inwards to WM contents or outwards to perceptual stimuli, and suggest the suitability of our task to assess WM functioning in both healthy aging and PD.

Keywords: Working memory; Inhibition; Interference; Aging; Parkinson's disease

1. Introduction

Working memory (WM), as proposed by Baddeley (Baddeley, 1986; Baddeley & Hitch, 1974), is a powerful explanatory concept that includes some of the fundamental properties of short term memory, such as the capacity to store information for a brief period of time, together with the presence of a superordinate control system that permits the use of stored information in the service of complex cognitive tasks. According to one of the most adopted models in cognitive psychology (Baddeley, 2000; 1986), WM is described as the implementation of storing and rehearsal processes of short term memory, that maintain information in an active state, and executive processes that enable cognitive operations to be done with the stored contents (Miyake & Shah, 1999).

Both cognitive and neurocognitive models agree in defining WM performance as the outcome of processing at multiple hierarchical levels, in which the analysis and the integration of low-level features interact with top-down processes to reconstruct a stable mental representation of previously experienced information (Ranganath, 2006). This tight integration of storage and processing by WM components provides functionality in higher cognitive domains, such as learning, problem-solving, decision making, and reasoning (Braver & West, 2008; Baddeley, 1992).

Even if the literature has mostly upheld this definition, the research on WM is still evolving, especially in terms of exploration of the mental operations that constitute the central executive (Smith & Jonides, 1999; Goldman-Rakic, 1995; Petrides, 1994). According to Reuter-Lorenz and Sylvester (2005) at least four key processes, or “top-down” operations, are crucial to the central executive: *executive attention*, which focuses resources on task relevant information, *inhibition*, which suppresses irrelevant information and resolves interference and conflict; *task management,* that is the ability to maintain a goal and to organize sub-goals and, finally, *set shifting,* which refers to the ability to change rule states and decision criteria.

Following the current challenges in defining the precise role of such operations in WM, the main goal of the present research is to investigate how *executive attention* and *inhibition* work in the management of information in WM. In order to do so, we designed a new experimental paradigm aimed at exploring the focusing of attention on the WM representations, that is on symbolic information. This new paradigm was designed assuming that during reasoning WM mechanisms manage symbolic information in the same way in which they manage the perceptual ones, according to previous evidence that showed as selective attention operates in a similar way to both perceptive and symbolic representations (Chun, 2011; Cherubini, Mazzocco, & Minelli, 2007; Cherubini, Burigo, & Bricolo, 2006).

More specifically, in this new paradigm two different condition-action rules were used as stimuli: we hypothesized that an irrelevant bi-conditional rule ("if X occurs, then do Y; otherwise do Z"), transiently encoded in WM can affect the use of a similar task-relevant rule that is also stored in WM. Specifically, we predicted that an incongruence between two task rules should cause significant interference costs, measured as delays in reaction times (RT) and decreased accuracy, while a congruency should result in facilitation benefits, that is faster RT and increased accuracy.

In fact, we assume that during reasoning WM mechanisms manage symbolic information in the same way in which they manage the perceptual information, basing on previous evidences that showed as selective attention operates in a similar way when focused on perceptive or symbolic representations (Chun, 2011; Cherubini et al., 2007, 2006). We used this new paradigm to investigate both normal age-related and pathological WM (dys)functions.

In this article, we first present an aging study, in which we examined age-related differences in executive attention and inhibition components of WM, across three age groups. We predicted that the degree of interference (both costs and benefits) measured by our experimental task should significantly increase with age. If this prediction is fulfilled, our results will increase the knowledge about how WM changes or decline in healthy aging. Indeed, the literature about aging and WM is mainly based on studies that employed span tasks (Sander, Lindenberger, & Werkle-Bergner, 2012; Bopp & Verhaeghen, 2007, 2005; Babcock & Salthouse, 1990). However, even if span tasks are both reliable and valid measures of WM capacity, they do not provide an evaluation of WM information-processing efficiency (Ma, Husain, & Bays, 2014), because they do not tap directly on the top-down, executive-control components of WM. In line with recent studies (for a review see Zokaei, Burnett Heyes, Gorgoraptis, Budhdeo, & Husain, 2015), we propose an alternative empirical approach that allows the estimation of executive attention and inhibition, two top-down components of WM that are of critical importance in information processing, which have not been yet sufficiently explored in aging (Braver & West, 2008).

Second, we asked whether this new task could potentially provide a more sensitive measure of information processing in WM, and whether it may be useful in the clinical setting. For this reason, we extended our investigation by testing a group of treated Parkinson's disease (PD) patients. In fact, despite several neuropsychological studies reported significant WM impairment in PD patients (Lewis, Slabosz, Robbins, Barker, & Owen, 2005; Owen, 2004; Fournet, Moreaud, Roulin, Naegele, & Pellat, 2000, 1996; Stebbins, Gabrieli, Masciari, Monti, & Goetz, 1999; Owen, Iddon, Hodges, Summers, & Robbins, 1997; Gabrieli, Singh, Stebbins, & Goetz, 1996; Cooper, Sagar, Jordan, Harvey, & Sullivan, 1991), to our knowledge there is no evidence on how PD patients perform in WM tasks that do not rely on span measures.

2. Materials and Methods

Participants. A total of 82 participants took part in this study: 20 young adults (18-40 years), 23 older adults (40-69 years), 19 elderly subjects (70-80 years), and 20 PD patients (56-77 years). Demographical characteristics of the four samples are presented in Table 1. PD patients were matched for age, gender, and education with the group of elderly participants (see Table 2). For every participant, inclusion criteria were a normal or corrected to normal vision, and a Mini Mental State Examination (MMSE; Folstein, Folstein, & McHugh, 1975) score over or equal to the cut-off value (score range 25-30 in each group).

Exclusion criteria were the presence of psychiatric disorders, and the use of any psychotropic drug excluding levodopa, dopamine agonists (DA) or monoamine-oxidase inhibitors (MAOI) in the PD group).

Healthy participants with history of any neurological disease were also excluded. PD patients fulfilled diagnostic criteria for PD according to the PD Society Brain Bank Criteria (Hughes, Daniel, Kilford, & Lees, 1992), and the mean disease duration was 5.5 years (SD = 3.4 years, range 1-14 years). Patients were tested under the effect of the pharmacological treatment (ON phase). Specifically, 10 patients were on levodopa, 2 on DA, 3 on a MAOI, and 5 on a combination of levodopa and DA. Information about levodopa equivalent daily dose (LEDD) was available only for 12 PD patients (mean 481 mg, SD 250, range 100 - 1000). Since they were asked to continue taking their medication at the required time on the day of the testing session, patients showed neither motor fluctuations nor dyskinesia and the mean ON score in the Unified Parkinson's Disease Rating Scale (UPDRS; Fahn & Elton, 1987) part III (motor sub-score) was 9.9 (SD = 4.5, range 3.5-20). All participants gave signed informed consent prior to inclusion in the study and after the protocol had been explained to them.

Table 1. Demographical characteristics of the four groups recruited. For age and education the mean values and standard deviations are reported.

	Sample size	Gender	Age (years)	Education (years)
Young adults	*N* = 20	7 M, 13 F	23.5 ± 2.3	15.9 ± 1.4
Older adults	*N* = 23	13 M, 10 F	61.7 ± 3.1	7.4 ± 3.4
Elderly adults	*N* = 19	11 M, 8 F	72.1 ± 2.3	6.1 ± 3.1
PD patients	*N* = 20	12 M, 8 F	69.5 ± 7.3	7.1 ± 3.4

Neuropsychological assessment. In order to exclude the presence of a cognitive decline and to specifically assess executive functions, both elderly participants and PD patients were invited to attend a neuropsychological assessment session before the execution of the experimental task.

The presence of general cognitive decline was excluded by the screening with MMSE, employed to assess the study eligibility. Subsequently, additional standardized tests were administered. The phonemic fluency test and the Frontal Assessment Battery (FAB; Appollonio et al., 2005; Dubois, Slachevsky, Litvan, & Pillon, 2000) were administered to assess executive functions, while Brown Peterson technique 10" and 30" (BPT; Peterson & Peterson, 1959; Brown, 1958) was selected from the battery Esame Neuropsicologico Breve-2 (Brief Neuropsychological Examination, version 2; ENB2; Mondini, Mapelli, Vestri, Arcara, & Bisiacchi, 2011) to assess temporary storage within WM (Wang, Ren, Li, & Schweizer, 2015). The performance of the two groups at every test was then compared in order to evaluate the presence of group differences.

As reported in Table 2, the two groups had comparable performance at the MMSE and at the phonemic fluency test. On the contrary, PD patients scored significantly lower than the elderly participants (i.e., a matched control group) on the FAB, BPT 10" and BPT 30" tests.

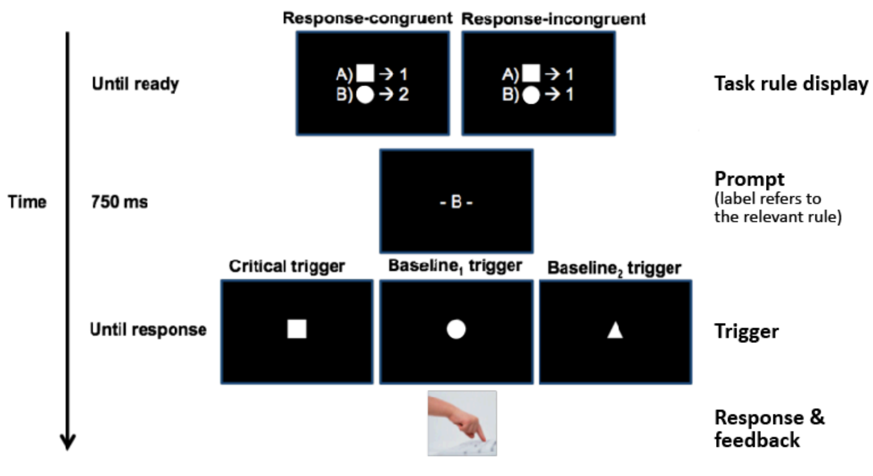
Table 2. Comparison between PD patients and the control group. Demographical data and mean scores at neuropsychological tests are reported.

	PD patients N=20	Healthy elderly N=19	*Statistics* (df)	*p*-value	*Cohen's d*
Age	69.5 ± 7.3	72.1 ± 2.3	$t_{(23)}$ = -1.49	.15	.49
Education	7.1 ± 3.4	6.1 ± 3.1	$t_{(37)}$ = 1.04	.30	.34
MMSE	27.5 ± 2.6	28.5 ± 1.1	$t_{(25)}$ = -1.76	.09	.54
FAB*	13.6 ± 2.9	15.5 ± 1.8	$t_{(31)}$ = -2.59	.014	.71
Phonemic fluency	9.0 ± 4.1	8.3 ± 2.4	$t_{(30)}$ = 0.86	.40	.25
BPT 10"*	5.1 ± 2.8	6.6 ± 1.8	$t_{(32)}$ = -2.14	.04	.73
BPT 30"*	4.5 ± 2.9	6.5 ± 1.9	$t_{(32)}$ = -2.49	.02	.84

Variations in the degrees of freedom (df) calculation are due to significant Levene test. MMSE: Mini Mental State Examination; FAB: frontal assessment battery; BPT: Brown Peterson technique. In the phonemic fluency, the score represents the mean number of words pronounced.

Experimental Task. Some previous studies on young participants showed that selection and inhibition mechanisms operate in a similar way, disregarding whether attention is oriented outwards to perceptually available, external stimuli, or inwards to the contents of WM (Kiyonaga & Egner, 2014; Chun, 2011; Cherubini et al., 2007, 2006). Here we capitalized on those previous findings, by building a task where the irrelevant distractors in WM could be either congruent or incongruent with the appropriate trial response. As occurs in visual attention tasks, if a response-congruent distractor was not effectively inhibited – and thus interfered with processing – a benefit in performance is expected, resulting in decreased RT and reduced error rates. By contrast, if a response-incongruent distractor was not effectively inhibited and interfered with processing, a cost in performance is expected, resulting in increased RTs and increased error rates (as a result of "capture errors"; Reason, 1990; Rasmussen, 1982). However, it must be noticed that in the present study – differently from more classical empirical paradigms, such as the Stroop task (Stroop, 1935) or the flankers task (Eriksen & Eriksen, 1974) – both the relevant information and the irrelevant distractor were not stimuli concurrently presented in the visual field of the participants, but they were task rules retained in WM.

At the start of each trial, two condition-action task rules were displayed at the center of a computer screen. Each rule was to be interpreted as a double-conditional ("if *condition* then perform *action 1*; else, perform *action 2*") instruction: it stated what to do if a condition occurred, and what to do if that condition did not occur. For example, the rule coded "□ → 1" means: "if a square is displayed, press key 1; otherwise, press key 2". The two rules were labeled by the letters A and B. The condition (leftmost part) of each rule was one out of three possible geometrical figures (circle, square, or triangle). The action (rightmost part) of each rule specified one of two response keys, labeled "1" and "2". Participants were asked to memorize both rules for the incoming trial. The duration of the rule presentation was self-regulated by the participant (Figure 1). After the participant declared that they had memorized the two rules, the rules were removed from the screen, and the prompt letter A or B was immediately displayed, for 750 ms. The prompt indicated which one of the two rules was relevant in the current trial, and was to be applied to the incoming trigger stimulus. Participants were instructed that the unprompted rule was irrelevant, and was to be ignored. Immediately after the prompt, a trigger stimulus was displayed, consisting in one of three geometrical figures (circle, square, or triangle). The trigger remained on the screen until the participant responded. Participants had to apply the prompted rule to the trigger figure in order to determine the appropriate response key. Participants responded by using the left or the right button of an analog computer mouse, labeled with the number 1 and 2 in a counterbalanced way across participants. Their correct responses were followed by an acute tone feedback, whereas their incorrect responses were followed by a grave tone. A new trial followed 350 ms after the feedback. The trigger stimuli were classified as critical, $baseline_1$, or $baseline_2$. Critical and $baseline_2$ triggers mismatched the condition of the prompted rule, indicating that the correct response was the other key than that mentioned in the rule action. However, different from baseline$_2$, critical triggers *always* matched the condition of the irrelevant rule. Thus, they might activate it (Reverberi, Pischedda, Burigo, & Cherubini, 2012), causing interference with the computation of the response. In that case, if the action of the irrelevant rule was different from that of the prompted rule, the irrelevant rule suggested the *appropriate* response (response-congruent irrelevant rules). Otherwise, the irrelevant rule suggested an *inappropriate* response (response-incongruent irrelevant rules). Baseline$_1$ triggers directly matched the condition of the prompted rule. Thus, they provide an estimation of the time required for inferring the response after a positive match on that rule. Baseline$_2$ triggers did not match the condition of any rule. Thus, they provide an estimation of the time required for inferring the response after a negative match on the relevant rule, in absence of interference by the irrelevant one. Because the baseline$_2$ triggers also mismatched the condition of the irrelevant rule, it was unlikely that the irrelevant rule was activated; even if it was, half of the times – that is in the response-incongruent condition – it suggested the appropriate response, and half of the times – that is, in the response-congruent condition – it suggested the inappropriate response. Therefore, all residual interference by the irrelevant rule was canceled out when all the baseline$_2$ trials were averaged. Six experimental conditions originated from a fully orthogonal 2x3 within-participant factorial design (see Figure 1).



The two experimental factors were the response congruency of the irrelevant rule with respect to critical triggers (response-congruent vs. response-incongruent) and the type of triggers (critical vs. baseline_1 vs. baseline_2 triggers). As a result of all combinations of the 24 possible pairs of rules (12 response-congruent and 12 response-incongruent; pairs displaying two rules with the same condition were omitted), by two prompts (A or B), by three trigger figures (triangle, square, or circle), 144 trials were presented in the experimental session, 24 for each cell of the experimental design. Notice that each possible coupling of two rules and one trigger was only presented *once* during the experiment, making the engaging of long-term learning of each specific trial's rules both unlikely and, also, useless for performance.

The experiment was run with the E-Prime 2 software (Psychology Software Tools, Pittsburgh, PA) installed on a personal computer equipped with a 17" monitor. The authors will make available the code of the experimental task on request, for non-commercial research or clinical-screening purposes.

The experimental session was structured as follows: after participants read self-paced instructions displayed on a computer, two blocks of practice trials were administered. In the first practice block, participants were trained to interpret correctly each task rule. Only one rule at a time was displayed followed by a trigger figure (no prompt was presented). An accuracy feedback was given after each response. The practice block ended as soon as the participant provided five correct responses across the last six trials. In the second practice section, detailed instructions about the two-rule structure were given, explaining how to interpret the prompt that followed the rule display. Participants were explicitly told that only the prompted rule should govern responses, while the uncued rule was irrelevant and had to be ignored. After five practice trials, randomly selected from the 144 experimental trials, the experimental phase began.

Predictions. Our first prediction was that significant response costs should emerge in response-incongruent conditions, namely in conditions where the irrelevant distracting rule suggested the wrong response. At the same time, response benefits were expected in response-congruent conditions, that is when the distracting rule suggested the correct response. If these predictions were confirmed, our second prediction would be to find an age-related increase in the size of these interference effects. Finally, given the exploratory nature of our clinical investigation, clear predictions about PD patients' performance were not defined a priori.

Data processing. Raw RTs and accuracy ratings (AC) were processed for each participant to derive two normalized scores that estimated the degree of interference caused by the irrelevant rule (Figure 2). The rationale of the normalized scores is as follows: correct responses to both critical and $baseline_2$ stimuli followed from inferences triggered by *negative* matches on the condition of the prompted rule, whereas correct responses to $baseline_1$ stimuli followed from *positive* matches on the condition of the prompted rule (see Figure 1). It is known that inferences from negative matches are more difficult – that is they take longer and cause more errors – than inferences from positive matches on the condition of a rule (Evans & Handley, 1999; Klauer & Oberauer, 1995; Wason, 1959). Thus, the relative cost of the negative versus positive inferences was computed by dividing, respectively, RTs to critical stimuli and RTs to $baseline_2$ stimuli by the corresponding RTs to $baseline_1$ (formulas 1 and 2 in Figure 2). Theoretically these two ratios are both > 1, and they measure the baseline cost of an inference triggered by a negative match on the condition of a rule.
However, normalized critical RTs are also affected by *interference* by the critical rule, whereas normalized $baseline_2$ RTs are not. Thus, the difference between the two normalized measures conveys a pure estimation of the interference effect on correct-response latencies, ΔRT (formulas 5 and 7 in Figure 2). ΔRT expresses the relative acceleration/delay (thus, it is not measured in milliseconds, but as a percentage of decrease/increase of the negative inference delay) of the response to critical triggers (subject to interference by the irrelevant rule) with respect to responses to the $baseline_2$ triggers (that are not subject to interference by the irrelevant rule). ΔRT is negative when interference by the irrelevant rule causes a response benefit (acceleration) to critical triggers, positive when interference causes a response cost (delay) to critical triggers, and zero if interference

by the irrelevant rule has no effect on the responses (i.e., the irrelevant rule does not interfere with performance at all).

Following the same logic, we devised normalized scores and a single measure of interference, ΔAC, for accuracy rates in each condition (formulas 3, 4, 6 and 8 in Figure 2). ΔAC expresses the relative increase/decrease of accurate responses to critical triggers (that are prone to interference by the irrelevant rule) with respect to $baseline_2$ triggers (that are not subject to interference). ΔAC is positive when capture by the irrelevant rule causes performance benefits, negative when capture by the irrelevant rule causes performance costs, and zero if the irrelevant rule does not capture performance at all.

It is worth noticing that the normalization procedure and the Δ measures have two further advantages. First, Δ measures remove any intrinsic interindividual and intergroup variability between absolute response speed and baseline accuracy of different participants. By using those measures, the trivial and known facts that older participants and patients are, on average, slower and less accurate than younger healthy participants is canceled out, and thus they can bear no effects whatsoever on the analyses of the degree of interference. Second, Δ measures – being computed by using condition-specific baselines – remove the non-interesting effects caused by the different WM load imposed by the response-congruent conditions (generated by pairs of rules with different actions) and response-incongruent conditions (generated by pairs of rules with the same actions, and thus with a lighter load on WM; see Figure 1).

Δ measures, which are expected to be positive in one condition and negative in the other condition, cancel each other out when congruent and incongruent conditions are averaged. Therefore, they are not useful for between-groups comparisons, and for correlations between the total amount of interference and other continuous variables, such as age or neuropsychological test scores. We thus defined Σ measures (formulas 9 and 10 in Figure 2), which are estimates of the total size of interference effects, summing up the absolute values of costs and benefits.

NORMALIZATION PROCEDURE

1. ***Normalized RT critical:*** *Mean RT critical / Mean RT baseline$_1$*
2. ***Normalized RT baseline$_2$:*** *Mean RT baseline$_2$/ Mean RT baseline$_1$*

3. ***Normalized accuracy critical:*** *Accuracy critical / Accuracy baseline$_1$*
4. ***Normalized accuracy baseline$_2$:*** *Accuracy baseline$_2$/ Accuracy baseline$_1$*

BENEFIT

5. ***Δ RT*** *= (Normalized RT congruent critical) – (Normalized RT congruent baseline$_2$)*
6. ***Δ Accuracy*** *= (Normalized accuracy congruent critical) – (Normalized accuracy congruent baseline$_2$)*

COST

7. ***Δ RT***= *(Normalized RT incongruent critical) – (Normalized RT incongruent baseline$_2$)*
8. ***Δ Accuracy*** *= (Normalized accuracy incongruent critical) – (Normalized accuracy incongruent baseline$_2$)*

ABSOLUTE VALUES (Σ) OF THE INTERFERENCE EFFECTS

9. **ΣRT** = ΔRT (response-incongruent condition) - ΔRT (response-congruent condition)
10. **ΣAC** = ΔAC (response-congruent condition) – ΔAC (response-incongruent condition)

Data analysis. Performance at the experimental task was evaluated considering ΔRT and ΔAC, obtained by the normalization of the raw RT and accuracy data, as dependent variables (see Table 1A in the supplementary material).

Planned analyses on the performance of healthy participants comprised therefore two 2x3 linear mixed models, factoring as a within-subject factor the *congruency* of the irrelevant rule (response-congruent vs. response-incongruent), and as a between-subjects factor the *group* (young adults, older adults, and elderly). Given the difference in the educational level between the three groups (see Table 1), the years of education were inserted in the models as a covariate.

For the analysis of PD patients' performance, a 2x2 ANOVA was conducted, factoring the *congruency* as within-subject factor, and the *group* (PD patients vs. healthy elderly participants) as between-subjects factor. In this case no covariates were considered, because the two groups were matched for both age and years of education (see Table 2).

In all analyses, if a main effect of the *congruency* factor or a two-way interaction between *congruency* and *group* were found, the total interference (ΣRT and ΣAC measures, see Table 2A of the supplementary materials) was compared in the different groups by means of independent samples *t*-tests. The Bonferroni correction was applied to all comparisons.

3. Results

Healthy participants

Results of the linear mixed model on ΔRT[1] showed that the factors *congruency* and *group* did not have significant main effects [*congruency*: $F(1,57) = .011$, $p = .92$; *group*: $F(1,57) = .28$, $p = .75$]. However, a significant two-way *congruency* x *group* interaction was found [$F(2,57) = 4.33$, $p = .018$, $np^2 = .132$; see Figure 3(a)]. Post hoc analyses revealed that the differences between congruent and incongruent conditions were significant only in older adults ($p < .05$) and elderly subjects ($p < .001$), but they were not significant in young adults. Furthermore, results of the post-hoc *t*-tests on ΣRT[2] values indicated that: a) young adults interference effects were not significantly different from those of older adults [$t(41) = .50$, $p = .62$]; b) young adults interference effects were significantly different from those of elderly participants [$t(36) = -2.1$, $p < .05$]; c) older adults interference effects were significantly different from those of elderly participants [$t(39) = -2.3$, $p < .05$; see Figure 3(a)]. Finally, neither a main effect nor an interaction with the covariate *years of education* was found ($p = .28$).

Results of the linear mixed model on ΔAC showed that the main effect of *group* and the two-way interaction *congruency* x *group* were not significant [respectively: $F(2,58) = .21$, $p = .81$; $F(2,58) = 1.87$, $p = .16$]. However, a significant main effect of the *congruency* factor was found [$F(1,58) = 9.26$, $p = .004$; $np^2 = .138$; mean ΔAC: congruent condition = .082; incongruent condition = -.155],

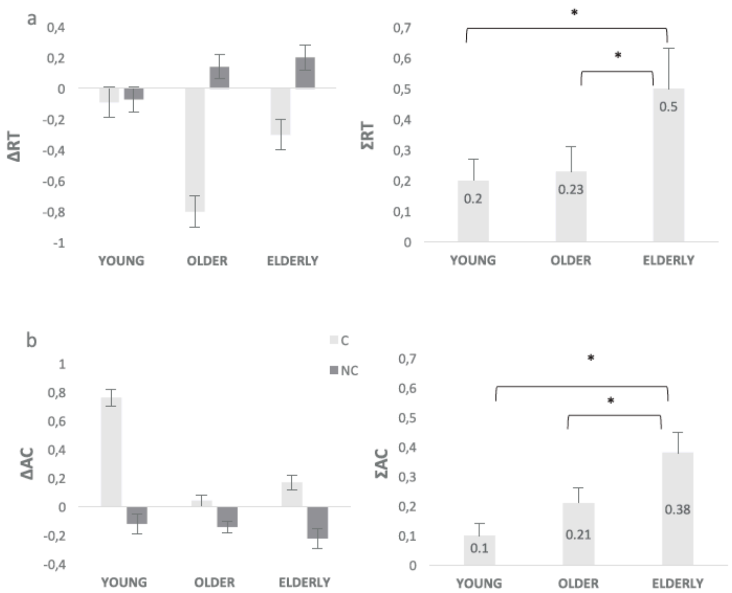
---

[1] One elderly participant had a ΔRT of -5.53 in the congruent condition, that is a value that is more than 10 standard deviations (SD) away from the mean of his group (mean ΔRT = - 3; SD = .55) and was therefore excluded as an

[2] Data about ΣRT and ΣAC are presented in Table 2A (see supplementary materials).

indicating overall the presence of a response benefit in the congruent condition and a response cost in the incongruent condition. A set of *t*-tests on ΣAC values were then conducted, and indicated that: a) young adults interference effects were not significantly different from those of old adults [$t(37) = -1.58$, $p = .06$]; b) young adults interference effects were significantly different from those of elderly participants [$t(36) = -3.53$, $p < .001$]; c) old adults interference effects were significantly different from those of elderly participants [$t(40) = -2.0$, $p < .05$; see Figure 3(b)].

Finally, as for the analysis on RTs, neither a main effect nor an interaction with the covariate *years of education* was found ($p = .51$).



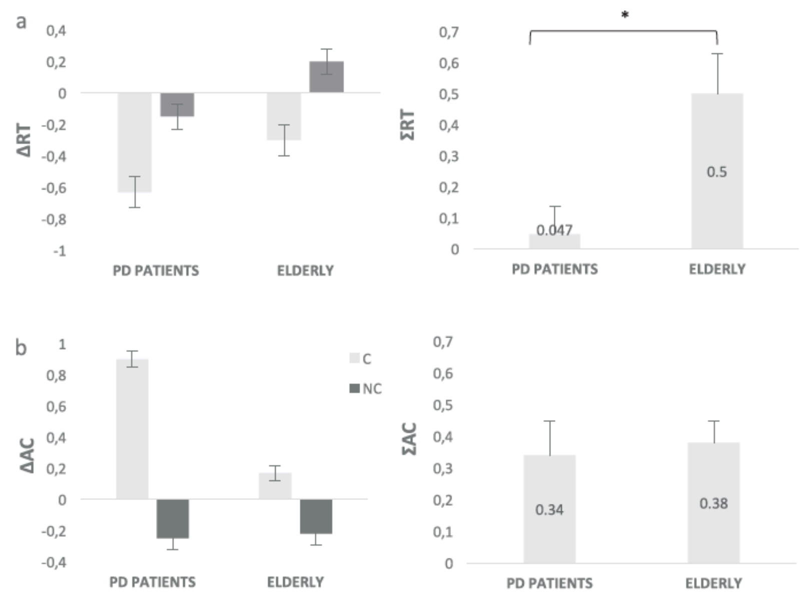
PD patients

Results of the ANOVA on ΔRT[3] showed a significant main effect of *congruency* [$F(1,36) = 10.75$, $p = .002$, $np^2 = .23$; mean ΔRT: congruent condition = -.18; incongruent condition = .09], confirming the presence of response benefits in the congruent condition and response costs in the incongruent condition. Furthermore, while there was no effect of *group* [$F(1,36) = .013$, $p = .91$], a significant two-way interaction *congruency* x *group* was found [$F(1,36) = 7.36$, $p = .01$, $np^2 = .17$].

[3] The analysis was run on the ΔRT values of 38 participants, since one healthy elderly participant was removed as an outlier (see footnote 1).

A post hoc analysis revealed that the difference between congruent and incongruent conditions was significant only in elderly subjects ($p < .001$), and not in PD patients ($p > .60$). Results of the *t*-tests on ΣRT values confirmed the previous result, indicating that the interference effect was significantly higher in elderly subjects with respect to PD patients [$t(36) = 2.71$, $p < .05$; see Figure 4(a)].

Finally, the ANOVA on ΔAC showed a significant main effect of *congruency* [$F(1,37) = 30.45$, $p = .000$, $np^2 = .451$; mean ΔAC: congruent condition = .13; incongruent condition = -.24], while both the effect of *group* and the two-way interaction *congruency* x *group* were not significant [*group*: $F(1,37) = 1.36$, $p = .25$, *congruency* x *group:* $F(1,37) = 0.1$, $p = .76$]. Results of the *t*-tests on ΣAC indicated that both groups showed comparable interference effect [$t(37) = .31$, $p > .50$; see Figure 4(b)].



4. Discussion

The main goal of the present study was to investigate how the cognitive system manages symbolic information temporarily stored in WM. We showed that rules temporarily encoded in working memory can interfere with performance. Although they are tagged as irrelevant, these rules influence behavior, by either causing performance costs – when they activate a response that is incongruent with the appropriate one – or performance benefits – when they accidentally converge on the appropriate response. By using a new experimental paradigm with pre-planned normalization of responses according to two different baselines, we devised a pure measure of interference between two task rules that allowed us to test WM efficiency in selecting the relevant information and at the same time in inhibiting the irrelevant one, that is to investigate the top-down WM components defined as *executive attention* and *inhibition* (Reuter-Lorenz & Sylvester, 2005). To our knowledge, these two top-down WM components have been scarcely investigated in a lifespan perspective, since most studies about age-related changes in WM relied on span measures. For this reason, we decided to employ this new experimental paradigm to investigate age-related (dys)function in the management of symbolic information in WM, with the aim to increase the knowledge about how WM changes or decline in healthy aging.

Finally, we also asked whether this new task could potentially provide a sensitive measure of top-down WM components that can aid researchers working with PD patients, a clinical population in which WM impairments have been frequently reported but, to our knowledge, never investigated using a measure that specifically explores WM top-down components (see Uitvlugt, Pleskac, & Ravizza, 2015 for a recent review; see also Kehagia, Barker, & Robbins, 2010). To reach these goals, the new experimental task was administered to three groups of healthy people, namely young adults (18-40 years), older adults (40-69 years), and elderly subjects (70-80 years), and, additionally, to a group of PD patients, matched for age, gender, and education with the group of elderly subjects. Results confirmed all our predictions. First, in all analyses the response-congruency/incongruence of the irrelevant rule had a significant effect on the response, causing significant benefit and cost effects respectively. This finding was robust and confirmed the suitability of the experimental design for measuring interference effects in WM; furthermore, this result also supports the claim that the orienting of attention has similar behavioral effects when attention is oriented inward to the contents of WM as well as when attention is oriented outward to information that is perceptually available (Kiyonaga & Egner, 2014; Chun, 2011; Cherubini et al., 2007, 2006).

Second, in both analyses on RT and accuracies, we found a robust age-related increase of the interference effects. Specifically, our data showed that elderly subjects (over 70 years) are significantly less able to inhibit irrelevant rules and to focus on relevant ones than under 70 years' participants. The fact that the significant increase of interference effects emerges especially in elderly subjects and not also in older adults, suggests a speculative conclusion: while a relatively stable and non-compromised functioning is preserved until late mid-age, an increasingly steep decay in top-down WM components emerges only in advanced age (De Carli, 2003; Graham et al., 1997). Further studies are needed for establishing the exact shape of the function linking age and interference effects in WM.

Finally, in the present study we tested a group of PD patients under dopaminergic treatment, considering the group of elderly subjects as a matched control condition. Both groups were evaluated with standardized neuropsychological tests to exclude general cognitive decline and to compare their performance in standard executive function tests, and in a WM test that assesses the storage capacity in condition of interference. From this preliminary neuropsychological assessment, it emerged that PD patients, even in absence of generalized cognitive decline, had executive function deficits and manifested worse ability to remember verbal material in conditions of interference. Surprisingly, results from the application of our experimental task showed a complex pattern of performance. Despite the outcome of the neuropsychological assessment, PD patients did not present any significant interference effect in their response latencies, showing better performance than elderly participants. At the same time, interference effects measured by accuracy levels were not significantly different between PD patients and healthy elderly participants, which were matched for age, gender, and education. Hence, our PD patients were comparable with their control group when the performance was evaluated in terms of accuracy, while they showed a better performance when evaluated by their RT. This lack of interference effect on RT first highlights the contrast between the performance at standard neuropsychological tests that assess WM and executive functions and the performance in our experimental paradigm.

Despite we did not have precise predictions about the performance of PD patients, we propose some possible interpretations for this peculiar result. First, it is possible that a standard WM task, such as the BTP, evaluates only the storage of information in WM, even if requiring coping with elements of interference, while our task better assesses top-down aspects of WM functioning. It is thus possible that the same PD patients who showed a deficit in information storage might not have any significant impairment in the top-down components of WM with respect to a matched control group of elderly subjects. Therefore, we may speculate that PD itself might have a differential effect on WM and executive functions, as tested by standard techniques and our experimental task.

However, given the significant interference effects on accuracy rates of PD patients, a second interpretation seems to be more likely. In our study, patients were tested under the effects of dopaminergic treatment ("ON" phase); therefore, it could be possible that the absence of interference effects measured with latency was caused by an enhancement/restoration of motor function due to the medication itself, that is the dopaminergic treatment.

It is tempting to speculate that motor changes to dopaminergic drugs might have a larger influence in our task than standard executive functions test. However, this is only a hypothesis, even if in line with the literature about the differential effects of dopaminergic treatment in WM performance (for a review, see Moustafa, Sherman, & Frank, 2008; see also Lewis et al., 2005; Costa et al., 2003; Marini, Ramat, Ginestroni, & Paganini, 2003; Owen et al., 1995; Lange et al., 1992; Cooper et al., 1991; Poewe, Berger, Benke, & Schelosky, 1991). Actually, even if statistically significant differences were found in the executive function neuropsychological tests, it is unclear whether they were clinically significant or translated into true executive dysfunction, as the mismatch between executive tests and performance in everyday life has been demonstrated in PD patients (Koerts et al., 2011). Nevertheless, because in the present study it was not possible to test PD patients in OFF state, and we did not have information about the level of LEDD and Hoehn-Yahr score

(Goetz et al., 2004; Hoehn & Yahr, 1967) for all them, any relation between PD patients' performance and the dopaminergic treatments refers to a speculative hypothesis, even if in line with the literature about the differential effects of dopaminergic treatment in WM performance (for a review see Moustafa et al., 2008; see also Lewis et al., 2005; Costa et al., 2003; Marini et al., 2003; Owen et al., 1995; Lange et al., 1992; Cooper et al., 1991; Poewe et al., 1991). Future studies testing PD patients in "OFF" phase could be able to confirm this hypothesis by overcoming the limits of the present research.

In conclusion, to our opinion the present study represents an important contribution to the study of cognition, aging, and neuropsychology. First, we added some information about the functioning of top-down WM components in healthy conditions. When individuals must focus on one relevant rule held in WM while preventing distractions from an irrelevant one, they do so by deploying executive attention to the relevant rule, while inhibiting the irrelevant one. Our results show that these operations have a "price" for the cognitive system, in strict analogy to attentional costs and benefits observed for the orienting of visual attention to visual stimuli in the Stroop's (1935) and flankers' (Eriksen & Eriksen, 1974) paradigms. A failure in inhibiting the irrelevant information temporarily stored in WM may indeed cause interference effects, which can result in either benefits or costs. This finding supports the view that the basic neural mechanisms of attentional orienting are similar when orienting attention inwards to the contents of WM, or outwards to perceptually available stimuli (Kiyonaga & Egner, 2014; Chun, 2011; Cherubini et al., 2007, 2006). Second, showing for the first time the existence of these interference effects in WM, here we also show that their magnitude grows as people get older. This indicates that the ability to inhibit irrelevant, distracting contents of WM significantly declines with age. To our opinion this result is important for a better understanding of the age-related cognitive decline in activities such as reasoning, problem solving, planning, and decision-making. Finally, since a key variable of this new WM task is represented by the induction of the so-called "capture errors" (Reason, 1990; Rasmussen, 1982), our paradigm may represent not only a valid measure to highlight the age-related decline in top-down WM components, but also to evaluate WM functioning in PD patients. However, only future studies could confirm this last conclusion, by testing PD patients in OFF condition and comparing this new paradigm with different traditional WM measures.

References

Appollonio, I., Leone, M., Isella, V., Piamarta, F., Consoli, T., Villa, M. L., ... & Nichelli, P. (2005). The frontal assessment battery (FAB): Normative values in an Italian population sample. *Neurological Sciences*, *26*(2), 108-116.

Babcock, R. L., & Salthouse, T. A. (1990). Effects of increased processing demands on age differences in working memory. *Psychology and Aging*, *5*(3), 421-428.

Baddeley, A. D. (1986). *Working memory*. Oxford: Oxford University Press.

Baddeley, A. D. (1992). Working memory. *Science*, *255*(5044), 556-559.

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in cognitive sciences*, *4*(11), 417-423.

Baddeley, A. D., & Hitch, G. (1974). Working memory. The psychology of learning and motivation, *8*, 47-89.

Bopp, K. L., & Verhaeghen, P. (2005). Aging and verbal memory span: A meta-analysis. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *60*(5), P223-P233.

Bopp, K. L., & Verhaeghen, P. (2007). Age-related differences in control processes in verbal and visuospatial working memory: Storage, transformation, supervision, and coordination. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *62*(5), P239-P246.

Braver, T. & West, R. (2008). Working memory, executive control, and aging. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (3rd edition, pp. 311-372). New York: Psychology Press.

Brown, J. (1958). Some tests of the decay theory of immediate memory. *Quarterly Journal of Experimental Psychology*, *10*(1), 12-21.

Cherubini, P., Burigo, M., & Bricolo, E. (2006). Inference-driven attention in symbolic and perceptual tasks: Biases toward expected and unexpected inputs. *Quarterly journal of* *experimental psychology*, *59*(3), 597-624.

Cherubini, P., Mazzocco, A., & Minelli, S. (2007). Facilitation and inhibition caused by the orienting of attention in propositional reasoning tasks. *Quarterly Journal of Experimental Psychology*, *60*(11), 1496-1523.

Chun, M. M. (2011). Visual working memory as visual attention sustained internally over time. *Neuropsychologia*, *49*(6), 1407-1409.

Cooper, J. A., Sagar, H. J., Jordan, N., Harvey, N. S., & Sullivan, E. V. (1991). Cognitive impairment in early, untreated Parkinson's disease and its relationship to motor disability. *Brain*, *114*(5), 2095-2122.

Costa, A., Peppe, A., Dell'Agnello, G., Carlesimo, G. A., Murri, L., Bonuccelli, U., & Caltagirone, C. (2003). Dopaminergic modulation of visual-spatial working memory in Parkinson's disease. *Dementia and geriatric cognitive disorders*, *15*(2), 55-66.

De Carli, C. (2003). Mild cognitive impairment: prevalence, prognosis, aetiology, and treatment. *The Lancet Neurology*, *2*(1), 15-21.

Dubois, B., Slachevsky, A., Litvan, I., & Pillon, B. (2000). The FAB. A frontal assessment battery at bedside. *Neurology*, *55*(11), 1621-1626.

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & psychophysics*, *16*(1), 143-149.

Evans, J. S. B., & Handley, S. J. (1999). The role of negation in conditional inference. *The Quarterly Journal of Experimental Psychology: Section A*, *52*(3), 739-769.

Fahn, S., Elton, R. L., & UPDRS Program Members (1987). Unified Parkinson's disease rating scale. In S. Fahn, C. D. Marsden, M. Goldstein, & D. B. Calne (Eds.), *Recent developments in Parkinson's disease*, Vol. 2 (pp. 153-163, 293-304). Florham Park, NJ: Macmillan Healthcare Information.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician. *Journal of psychiatric* *research*, *12*(3), 189-198.

Fournet, N., Moreaud, O., Roulin, J. L., Naegele, B., & Pellat, J. (1996). Working memory in medicated patients with Parkinson's disease: the central executive seems to work. *Journal of Neurology, Neurosurgery & Psychiatry*, *60*(3), 313-317.

Fournet, N., Moreaud, O., Roulin, J. L., Naegele, B., & Pellat, J. (2000). Working memory functioning in medicated Parkinson's disease patients and the effect of withdrawal of dopaminergic medication. *Neuropsychology*, *14*(2), 247-253.

Gabrieli, J. D., Singh, J., Stebbins, G. T., & Goetz, C. G. (1996). Reduced working memory span in Parkinson's disease: Evidence for the role of frontostriatal system in working and strategic memory. *Neuropsychology*, *10*(3), 322.

Goetz, C. G., Poewe, W., Rascol, O., Sampaio, C., Stebbins, G. T., Counsell, C., ... & Yahr, M. D. (2004). Movement Disorder Society Task Force report on the Hoehn and Yahr staging scale: status and recommendations the Movement Disorder Society Task Force on rating scales for Parkinson's disease. Movement disorders, 19(9), 1020-1028.

Goldman-Rakic, P. S. (1995). Architecture of the prefrontal cortex and the central executive. *Annals of the New York Academy of Sciences*, *769*(1), 71-84.

Graham, J. E., Rockwood, K., Beattie, B. L., Eastwood, R., Gauthier, S., Tuokko, H., & McDowell, I. (1997). Prevalence and severity of cognitive impairment with and without dementia in an elderly population. *The Lancet*, *349*(9068), 1793-1796.

Hoehn, M. M., & Yahr M. D. (1967). Parkinsonism: onset, progression and mortality. *Neurology*, 17, 427-442.

Hughes, A. J., Daniel, S. E., Kilford, L., & Lees, A. J. (1992). Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. *Journal of Neurology, Neurosurgery & Psychiatry*, *55*(3), 181-184.

Kehagia, A. A., Barker, R. A., & Robbins, T. W. (2010). Neuropsychological and clinical heterogeneity of cognitive impairment and dementia in patients with Parkinson's disease. *The*

Lancet Neurology, *9*(12), 1200-1213.

Kiyonaga, A., & Egner, T. (2014). The Working Memory Stroop Effect When Internal Representations Clash With External Stimuli. *Psychological science*, *25*(8), 1619-1629.

Klauer, K. C., & Oberauer, K. (1995). Testing the mental model theory of propositional reasoning. *The Quarterly Journal of Experimental Psychology*, *48*(3), 671-687.

Koerts, J., Tucha, L., Leenders, K. L., van Beilen, M., Brouwer, W. H., & Tucha, O. (2011). Subjective and objective assessment of executive functions in Parkinson's disease. Journal of the neurological sciences, *310*(1), 172-175.

Lange, K. W., Robbins, T. W., Marsden, C. D., James, M., Owen, A. M., & Paul, G. M. (1992). L-dopa withdrawal in Parkinson's disease selectively impairs cognitive performance in tests sensitive to frontal lobe dysfunction. *Psychopharmacology*, *107*(2-3), 394-404.

Lewis, S. J., Slabosz, A., Robbins, T. W., Barker, R. A., & Owen, A. M. (2005). Dopaminergic basis for deficits in working memory but not attentional set-shifting in Parkinson's disease. *Neuropsychologia*, *43*(6), 823-832.

Ma, W. J., Husain, M., & Bays, P. M. (2014). Changing concepts of working memory. *Nature neuroscience, 17*(3), 347-356.

Marini, P., Ramat, S., Ginestroni, A., & Paganini, M. (2003). Deficit of short-term memory in newly diagnosed untreated parkinsonian patients: reversal after L-dopa therapy. *Neurological Sciences*, *24*(3), 184-185.

Miyake, A., & Shah, P. (1999). *Models of working memory: Mechanisms of active maintenance and executive control*. Cambridge University Press.

Mondini, S., Mapelli, D., Vestri, A., Arcara, G., & Bisiacchi, P. S. (2011). *Esame neuropsicologico breve 2*. Milano: Raffaello Cortina.

Moustafa, A. A., Sherman, S. J., & Frank, M. J. (2008). A dopaminergic basis for working memory, learning and attentional shifting in Parkinsonism. *Neuropsychologia*, *46*(13), 3144-3156.

Owen, A. M. (2004). Cognitive dysfunction in Parkinson’s disease: the role of frontostriatal circuitry. *The Neuroscientist*, *10*(6), 525-537.

Owen, A. M., Iddon, J. L., Hodges, J. R., Summers, B. A., & Robbins, T. W. (1997). Spatial and non-spatial working memory at different stages of Parkinson's disease. *Neuropsychologia*, *35*(4), 519-532.

Owen, A. M., Sahakian, B. J., Hodges, J. R., Summers, B. A., Polkey, C. E., & Robbins, T. W. (1995). Dopamine-dependent frontostriatal planning deficits in early Parkinson's disease. *Neuropsychology*, *9*(1), 126.

Peterson, L., & Peterson, M. J. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology*, *58*(3), 193.

Petrides, M. (1994). Frontal lobes and working memory: Evidence from investigations of the effects of cortical excisions in nonhuman primates. In F. Boller & J. Grafman (Eds.), *Handbook of Neuropsychology* (Vol. 9, pp. 59-82). Amsterdam: Elsevier.

Poewe, W., Berger, W., Benke, T. H., & Schelosky, L. (1991). High-speed memory scanning in Parkmson's disease: Adverse effects of levodopa. *Annals of neurology*, *29*(6), 670-673.

Ranganath, C. (2006). Working memory for visual objects: complementary roles of inferior temporal, medial temporal, and prefrontal cortex. *Neuroscience*, *139*(1), 277-289.

Rasmussen, J. (1982). Human errors. A taxonomy for describing human malfunction in industrial installations. *Journal of occupational accidents*, *4*(2), 311-333.

Reason, J. (1990). *Human error*. Cambridge, UK: Cambridge University Press.

Reuter-Lorenz, P. A. & Sylvester, C. C. (2005). The cognitive Neuroscience of Working Memory and Aging. In: R. Cabeza, L. Nyberg & D. Park (Eds.), *Cognitive neuroscience of aging: linking cognitive and cerebral aging* (pp. 186-217). New York: Oxford University Press.

Reverberi, C., Pischedda, D., Burigo, M., & Cherubini, P. (2012). Deduction without awareness. *Acta psychologica*, *139*(1), 244-253.

Sander, M. C., Lindenberger, U., & Werkle-Bergner, M. (2012). Lifespan age differences in working memory: A two-component framework. *Neuroscience & Biobehavioral Reviews*, *36*(9), 2007-2033.

Smith, E. E., & Jonides, J. (1999). Storage and executive processes in the frontal lobes. *Science*, *283*(5408), 1657-1661.

Stebbins, G. T., Gabrieli, J. D., Masciari, F., Monti, L., & Goetz, C. G. (1999). Delayed recognition memory in Parkinsons disease: a role for working memory? Neuropsychologia, *37*(4), 503-510.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of experimental psychology*, *18*(6), 643-662.

Uitvlugt, M. G., Pleskac, T. J., & Ravizza, S. M. (2015). The nature of working memory gating in Parkinson's disease: A multi-domain signal detection examination. *Cognitive, Affective, & Behavioral* Neuroscience, *16*(2), 289–301.

Wang, T., Ren, X., Li, X., & Schweizer, K. (2015). The modeling of temporary storage and its effect on fluid intelligence: Evidence from both Brown–Peterson and complex span tasks. *Intelligence*, *49*, 84-93.

Wason, P. C. (1959). The processing of positive and negative information. *Quarterly Journal of Experimental Psychology*, *11*(2), 92-107.

Zokaei, N., Burnett Heyes, S., Gorgoraptis, N., Budhdeo, S., & Husain, M. (2015). Working memory recall precision is a more sensitive index than span. *Journal of neuropsychology*, *9*(2), 319-329.

Figure captions

Figure 1. Example of a trial, experimental design, and procedure.

Notes: *First factor of the design, Congruent vs. incongruent conditions*. The top-left display shows a typical congruent condition: the relevant rule, signaled by the prompt display, is rule B (circle -> 2); the bottom-left critical trigger "square" should then drive response "1", which actually coincides with the response dictated by the irrelevant A rule (square ->1) – in the event it captured performance. Symmetrically, the top-right display shows a typical incongruent condition: the relevant rule B (circle ->1) matched with the critical trigger "square" drives response "2", but if the irrelevant rule A (square ->1) captured performance, it would dictate the incorrect response "1". *Second factor of the design, Type of trigger*. Critical triggers (bottom-left display) always mismatch the condition of the relevant rule, and match the condition of the irrelevant one (thus, the irrelevant rule can interfere with performance); also $baseline_2$ triggers (bottom-right display) mismatch the condition of the relevant rule, but they mismatch the condition of the irrelevant rule as well (thus, the irrelevant rule should not interfere with performance); $baseline_1$ triggers (bottom-center display) are always direct matches of the condition of the relevant rule (and mismatches of the irrelevant one); they serve to measure the relative delay of an inference triggered by a negative match on a rule condition (as occurs for critical triggers and $baseline_2$ triggers), with respect to an inference triggered by a positive match on a rule condition. The combination of the congruency factor and of the three types of triggers generated the six-cells fully-orthogonal experimental design.

Figure 2. Normalization procedure for response latencies and accuracy rates.

Figure 3. Interference effects in healthy participants. Panel a) interference effects as measured on response latencies (RT); panel b) interference effects as measured on the rates of accurate responses (AC).

Notes: c = congruent; nc = incongruent; *significant difference = $p < .05$; error bars represent standard errors of the mean (SEM).

Figure 4. Interference effects in Parkinson's disease patients and healthy elderly participants. Panel a) interference effects as measured on response latencies (RT); panel b) interference effects as measured on the rates of accurate responses (AC).

Notes: c = congruent; nc = incongruent; *significant difference = $p < .05$; error bars represent standard errors of the mean (SEM).