# In the here and now: Short term memory predictions are preserved in Alzheimer's disease

Julie Bertrand, Audrey Mazancieux, Chris. J.A. Moulin, Yannick Bejot, Olivier Rouaud, Céline Souchay

**To cite this version:**

Julie Bertrand, Audrey Mazancieux, Chris. J.A. Moulin, Yannick Bejot, Olivier Rouaud, et al.. In the here and now: Short term memory predictions are preserved in Alzheimer's disease. Cortex, 2019, 119, pp.158-164. 10.1016/j.cortex.2019.03.027 . hal-02366994

**HAL Id: hal-02366994**

**https://hal.univ-grenoble-alpes.fr/hal-02366994**

Submitted on 25 Oct 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Version of Record: https://www.sciencedirect.com/science/article/pii/S0010945219301595
Manuscript_d5cc53d6bae29ee4582decd50e215455

# In the here and now: Short term memory predictions are preserved in Alzheimer's disease

Julie Marilyne Bertrand[1,2*]

Audrey Mazancieux[1*]

Chris John Anthony Moulin[1]

Yannick Béjot[2]

Olivier Rouaud[3]

Céline Souchay[1]

[1] Laboratoire de Psychologie et Neurocognition, LPNC CNRS 5105, Université Grenoble Alpes, France

[2] EA4184, Medical School of Dijon, Université de Bourgogne, France

[3] Centre Leenaards de la mémoire, Lausanne, Switzerland

* Joint first authors.

**Word count**: 3639

**Corresponding author**

Céline Souchay

Laboratoire de Psychologie et Neurocognition, LPNC

Grenoble Alpes University, France

Email: celine.souchay@me.com

Tel: +33 6 72 23 12 97

© 2019 published by Elsevier. This manuscript is made available under the CC BY NC user license
https://creativecommons.org/licenses/by-nc/4.0/

## Abstract

According to neuropsychological models of anosognosia, there is a failure to transfer on-line awareness of dysfunction into a more generalised long term belief about memory function in Alzheimer's disease. This failure results in specific metamemory deficits for global predictions: patients overestimate their performance before the task but are able to monitor their memory performance after having experienced the task. However, after a delay, they are still not able to make accurate predictions. As previous work has mainly focused on long-term memory, the present study investigates this issue in short-term and working memory. Using both global and item-by-item metacognitive judgements in a digit span task, we showed that Alzheimer's disease patients are as accurate as older adults in monitoring their performance despite impaired memory. When they have the opportunity to test themselves, or when they have already performed the task, patients are able to use feedback to adjust their metacognitive judgements. Overall, these results show that even for a relatively complex task, patients with Alzheimer's disease are aware of their difficulties in the here-and-now.


**Keywords**: Metacognition, Short-term memory, Alzheimer's disease, Global predictions

**In the here and now: Short term memory predictions are preserved in Alzheimer's disease**

Metamemory, defined as the ability to monitor, regulate and predict one's memory performance (Flavell, Miller & Miller, 2002) has often been explored in Alzheimer's disease (e.g. Souchay, 2007; Sunderaraman & Cosentino, 2017). The literature has focused on patients' ability to predict their memory performance on an upcoming test. Such predictions can occur at an item level (i.e., how people expect to perform for each specific item) or a global level (i.e., participants' expectations for their overall test performance). At the item level, metacognitive tasks with episodic and semantic materials have showed a diversity of spared and impaired performance (e.g., Bäckman & Lipinska, 1993; Souchay, Isingrini, & Gil, 2002), supporting the idea that the metamemory impairment in Alzheimer's disease is a consequence of the memory deficit (for a review see Ernst, Moulin, Souchay, Mograbi, & Morris, 2016). Critically, experimental metamemory tasks are known to be related to clinical classifications of awareness (e.g. Consentino, et al., 2016).

Fewer studies have explored metamemory in Alzheimer's patients using global judgments but all have shown that patients overestimate their performance initially (Ansell & Bucks, 2006; Barrett, Eslinger, Ballentine, & Heilman, 2005; Moulin, Perfect, & Jones, 2000a; 2000b). These studies also compared predictions made before (prospective judgements) and after the task (retrospective judgements), finding that Alzheimer's patients revise their predictions to accurate levels after having experienced the task (Ansell & Bucks, 2006; Moulin et al, 2000a; 2000b; Stewart, McGeown, Shanks, & Venneri, 2010). Thus patients are poorly calibrated but they show intact awareness of their memory performance after having experienced the task. Interestingly, after a delay, people with Alzheimer's disease continue to overestimate their function even after having made accurate evaluations during a task (Silva, Pinho, Macedo, Souchay, & Moulin, 2017): there is a failure to transfer the online awareness of dysfunction into a more generalised long-term belief about memory function. Several models have been proposed to explain awareness deficits in Alzheimer's disease such as the Levels of Awareness framework (Clare, Marková, Roth, & Morris, 2011) or the Cognitive Awareness Model (CAM, Morris & Mograbi, 2013). Such models converge on a failure to consolidate their knowledge regarding their memory abilities over a long period; *mnemonic anosognosia*.

To date, the bias in such research has overwhelmingly been towards examining the ability to monitor long term memory. Clearly, however, the global judgements literature reviewed above points to an ability to make accurate judgements in the short term which are not maintained in the long term. Here we sought to directly assess the ability to monitor short term memory, hitherto unexamined in Alzheimer's disease. Short-term memory underpins many activities of daily living, and evaluation this domain is critical. Our experimental design is based on Flavell's original design (Flavell, Friedrichs & Hoyt, 1970; see also Murphy, Sanders, Gabriesheski, & Schmitt, 1981) and adapted by Bertrand, Moulin, and Souchay (2017) in a recent study. A novelty of our protocol is that it allows the measure of both global predictions and item-by-item judgements. As it has been suggested by previous work in episodic memory, we hypothesised that Alzheimer's disease patients would be impaired on initial global predictions (made before the task). However, according to the idea that on-line metacognitive processes are intact, they should be preserved for item-by-item judgements and global postdiction because these are based on access to short term representations of task performance. Following standard practice in neuropsychological assessment, we examined both forward and backward span, although we made no specific predictions about differences between the two tasks, although the fact that backwards span is more demanding than forward span may be of interest (although this difference between the two tasks is far from clear-cut, e.g. Hester, Kinsella & Ong, 2004).

# Method

## Participants

Twenty-three older adults ($M_{age}$ = 73.09, $SD_{age}$ = 6.04; 17 females) and eighteen Alzheimer's patients ($M_{age}$ = 76.44, $SD_{age}$ = 5.89; 6 females) participated in the study. The healthy older adults were recruited from in the local community. Participants were defined as cognitively healthy if they had a mini-mental state exam (MMSE; Folstein, Folstein, & McHugh, 1975) score of 28-30.

All patients were recruited from the Memory Clinic at the Dijon university hospital. Diagnosis was determined by a group of neurologists at the memory clinic. Patients had a MMSE score ranging from 14 to 28 ($M$ = 21.67, $SD$ = 4.38). Participants were excluded if they had a history of clinical stroke, traumatic brain injury, alcohol or drug abuse or medical/psychiatric condition. The study was approved by the Institutional Review Board of the Dijon Hospital.

## Materials and procedure

For the two tasks (i.e., forward digit span and backward digit span), there were three phases (for a summary of the entire procedure see Figure 1). The first phase was a global prospective judgement task, where participants had to report how many digits they thought they would be able to remember (from 0 to 9). The second phase was an online task where participants gave item-by-item metacognitive judgements for the short-term memory task (either forward digit span or backward digit span). Here, there were two types of judgements: the prospective judgements (made before a trial) and the retrospective judgements (made after the trial). Item-by-item judgements were made after being presented a set of digits of a certain length: they were based on the participant's recent experience of the to-be-remembered digits. Participants were presented with first one series of digits and then asked whether they would be able to recall the number by giving a Yes/No answer. The number of digits increased sequentially (to a maximum of 9), until the participants said 'No'. For example, if after the presentation of 4 items a participant decided to say 'No', the item-by-item judgement stopped, with the participants therefore predicting having a span of 3. Note that in this paradigm, performance and predictions are not taken for the same trial. Rather, there is blocked presentation, such that a first set of digit spans are used to make the item-by-item prospective judgements, then there is the block of digit spans where recall is measured in the standard fashion, and finally a third block where the retrospective item-by-item predictions are made (see Figure 1). For the span tasks where recall was measured, the digit forward and backward span tasks from the Wechsler Adult Intelligence Scale-IV (WAIS-IV, Wechsler, 2011) were used. For the item-by-item judgements, lists of numbers matched in length with the span task were created and these differed for the two judgements. Lists were counterbalanced across judgements and participants. The third phase was a global retrospective judgement task. As in the first phase, participants had to say how many digits they thought they would be able to remember.

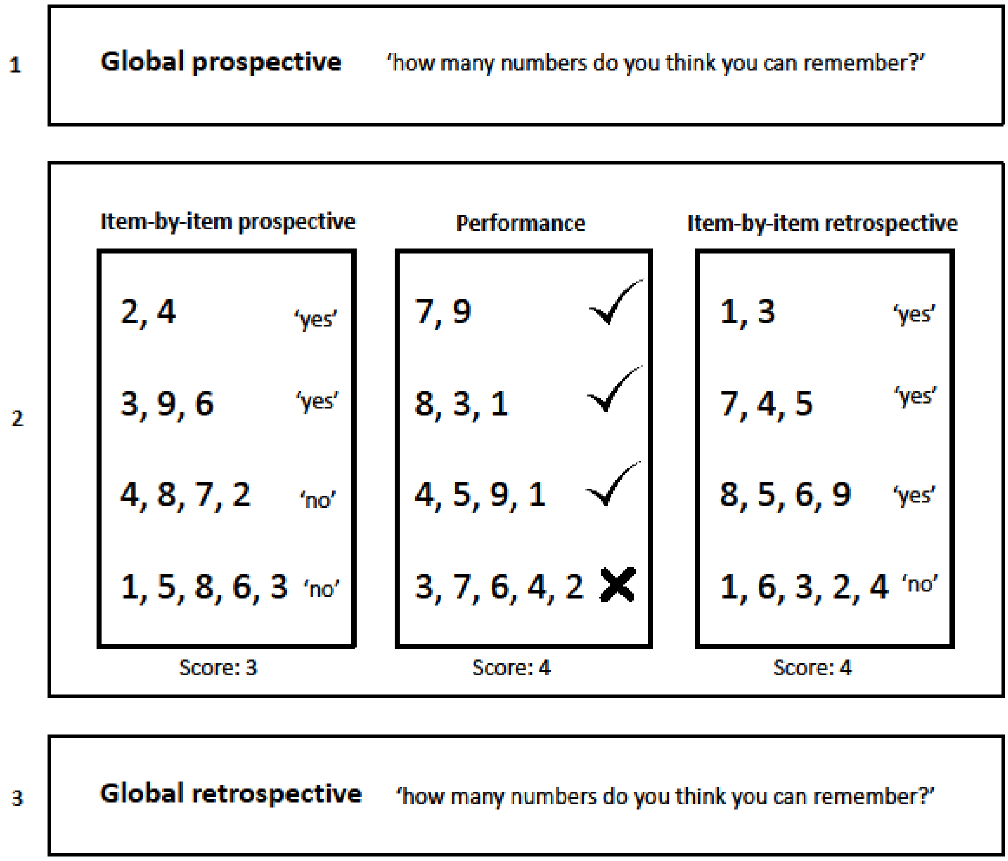


**Figure 1**. Summary of the procedure. The first phase is a global prospective judgement. The second includes prospective judgements, the actual task (either forward span or backward span) and retrospective judgements. The third phase is a global retrospective judgement.

## Data analyses

Analyses were conducted in R and included linear regressions (yielding *t* statistics which can be interpreted exactly as factorial ANOVAs) and Pearson’s correlations. We use a standard analysis protocol, starting by examining the mean recall (span performance) and the mean prediction values (prediction magnitude). Then, we focus on metacognitive accuracy in a standard fashion, considering the non-directional discrepancy between prediction and performance (e.g., Moulin et al, 2000a). This procedure allows the estimation of metacognitive accuracy independently from the bias (underestimation or overestimation of performance). A score of zero reflects perfect accuracy and the higher the score, the bigger the discrepancy between the metacognitive judgement and the performance (see Table 1).

[Following publication we will make the dataset and analysis script available on-line. Data and script are part of the submission.]

## Results

### Span performance

The mean span performance for each group is found in Table 1. We conducted linear regressions with group as a between-subjects factor and task as a within-subjects factor. As expected, we found a main effect of group, revealing that Alzheimer's patients have a lower performance than older adult controls, $t(39) = 3.28$, $p = .002$, $d = 1.03$. We also found a main effect of task. Performance on the forward span was higher than performance on the backward span, $t(39) = 6.20$, $p < .001$, $d_z = 0.97$, i.e. spans were significantly longer for forwards rather backwards recall. There was no interaction between the two factors, $t(39) = 1.36$, $p = .180$, $d = 0.43$.

### Magnitude of metacognitive judgements

**Forward span.** The mean values of predictions are given in Table 1. We conducted linear regressions with, group, prediction time (prospective vs retrospective), and judgement type (global vs item-by-item) as factors. Analyses revealed a significant effect of group, $t(39) = 2.61$, $p = .013$, $d = 0.82$, and a non-significant trend of judgement type, $t(39) = 1.82$, $p = .077$, $d_z = 0.28$. Patients made lower judgements overall than controls therefore predict having a lower span, which is appropriate given the differences in performance reported above. The analysis showed neither a significant effect of prediction time, $t(39) = 1.12$, $p = .269$, $d_z = 0.17$, and we found no significant interactions (all $t(39) < 1.43$, and all $p > .05$).

**Backward span.** The mean values of predictions are given in Table 1. As for forward span task, we conducted linear regressions with, group, prediction time (prospective vs retrospective), and judgement type (global vs item-by-item) as factors. Analyses revealed only that patients have a non-significant trend for lower judgements than older adults, $t(39) =$

1.85, $p$ = .072, $d$ = 0.58. The analysis again showed neither a significant effect of prediction time, $t(39)$ = 1.33, $p$ = .192, $d_z$ = 0.21, nor a significant effect of judgement type, $t(39)$ = 1.30, $p$ = .203, $d_z$ = 0.20. There were no significant interactions (all $t(39) < 1.29$, and all $p > .05$). Whereas we found significant differences in magnitude of predictions for forward spans, no such pattern was observed for the backwards span.

**Table 1**. Mean and standard errors for global judgements, item-by-item judgements, and performance for AD patients and older adults.

| | Forward span | | Backward span | |
|---|---|---|---|---|
| | Older adults | AD patients | Older adults | AD patients |
| Global prospective | 5.17 (1.03) | 4.78 (2.02) | 4.35 (0.78) | 3.89 (1.81) |
| Item-by-item prospective | 5.96 (1.26) | 4.67 (1.88) | 4.48 (0.95) | 3.83 (2.04) |
| Performance | 6.00 (1.09) | 4.67 (1.14) | 4.61 (1.31) | 3.78 (0.94) |
| Item-by-item retrospective | 6.26 (1.60) | 4.89 (2.03) | 4.96 (1.11) | 4.00 (1.50) |
| Global retrospective | 5.61 (1.26) | 4.61 (2.00) | 4.39 (0.84) | 3.83 (1.62) |

## Metacognitive accuracy

**Forward span.** We conducted linear regressions with, group, prediction time (prospective vs retrospective), and judgement type (global vs item-by-item) as factors. These analyses showed neither a significant effect of group, $t(39)$ = 1.09, $p$ = .284, $d$ = 0.34, nor a significant effect of prediction time, $t(39)$ = 1.76, $p$ = .086, $d_z$ = 0.27. We did however, find a main effect of judgement type, $t(39)$ = 2.18, $p$ = .036, $d_z$ = 0.34. Global judgements were less accurate that item-by-item judgments. There was also a significant interaction between judgement type and judgement time, $t(39)$ = 2.17, $p$ = .036, $d_z$ = 0.34. Retrospective judgements are more accurate for global predictions, $t(39)$ = 2.41, $p$ = .021, $d_z$ = 0.38, which is not the case for item-by-item judgements, $t(39)$ = -0.56, $p$ = .582, $d_z$ = -0.08. No other interaction was significant (all $t(39) < 1.59$, and all $p > .05$). As there was no group difference, these results show that patients are as accurate as controls at predicting their short-term memory performance (Figure 2A).

**Backward span.** As for forward span task, we conducted linear regressions with, group, prediction time (prospective vs retrospective), and judgement type (global vs item-by-item) as factors. The analysis showed neither a significant effect of group, $t(39) = 1.47$, $p = .149$, $d = 0.46$, nor a significant effect of judgement time, $t(39) = 0.51$, $p = .614$, $d_z = 0.08$, nor a significant effect of judgement type, $t(39) = 1.45$, $p = .156$, $d_z = 0.23$. No interaction was significant (all $t(39) < 1.23$, and all $p > .05$). For the backward span task, these results show that patients are as accurate as controls at predicting their performance (Figure 2B).

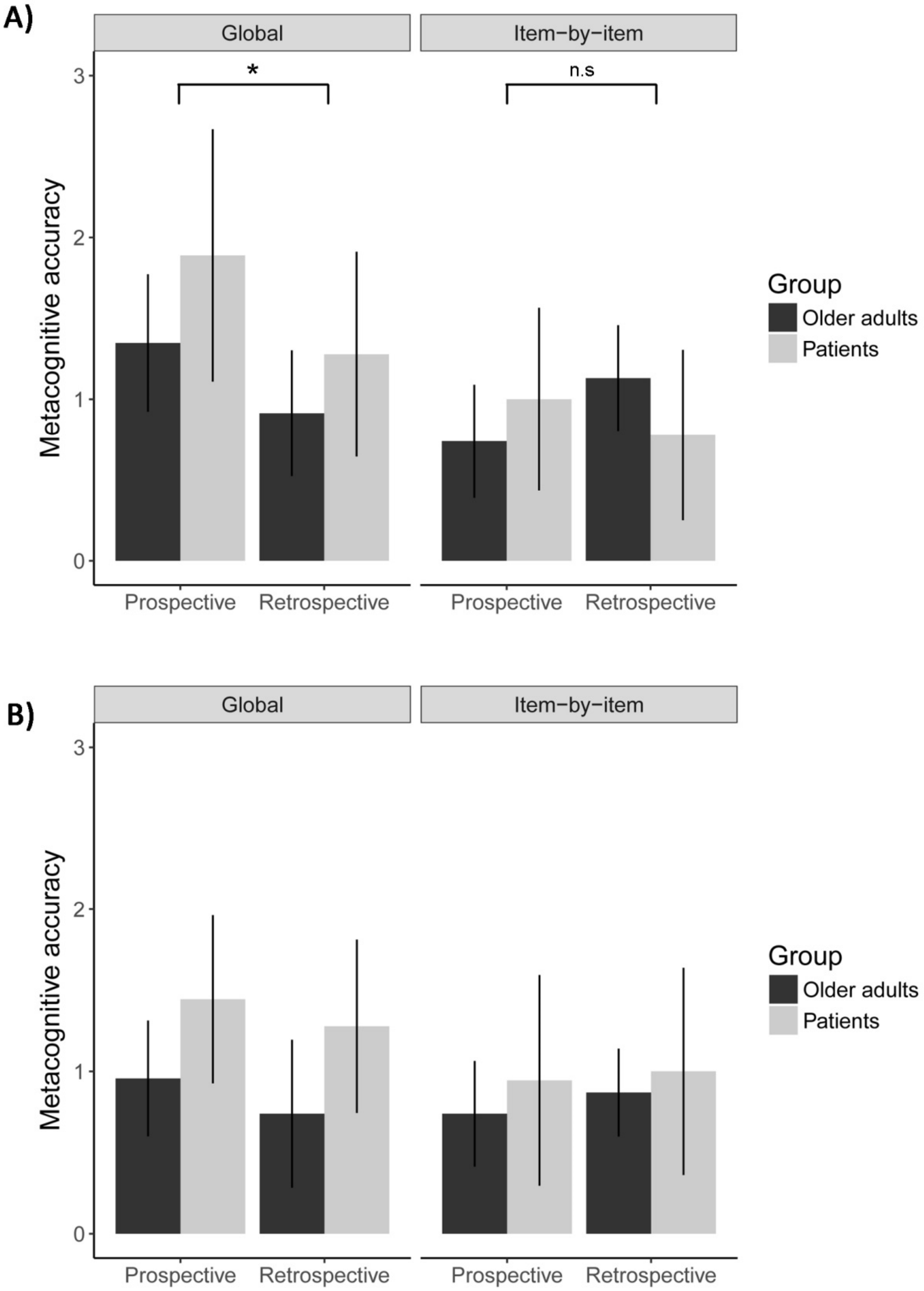


**Figure 2**. (A) Mean and confidence intervals for metacognitive accuracy according to judgement type, judgement time, and group for the forward span task. (B) Mean and confidence intervals for metacognitive accuracy according to judgement type, judgement time, and group for the backward span task.

## Correlational analyses

In order to examine the accuracy at the group level, we analyzed the correlations between the metacognitive judgements and digit span tasks. In these analyses, individuals' predictions are correlated with individuals' performance, such that as a group, we can see if those people with lower predictions actually have a worse performance (see Connor, Dunlosky & Hertzog, 1997). Table 2 shows that this relationship is always positive except for the global prospective judgements in both tasks where correlations are not significant. The same analysis can be carried out for each group individually, to compare these correlations for AD patients and older adults. Table 3 and Table 4 show that this relationship is overall positive between judgement and performance (although not always significant) except for the global prospective judgements in the forward span where correlations coefficients are near 0. Moreover, there was no difference in the magnitude of correlation across AD patients and older adults (all $z$ value < $|1.96|$).

**Table 2**. Bonferroni corrected correlations between metacognitive judgements and performance for both forward digit span and backward digit span. As there are 4 correlations per tasks the significance threshold is equal to 0.05/4 = 0.013.

| | Forward span | Backward span |
|---|---|---|
| Global prospective | $r = 0.06$, $p = .727$ | $r = 0.30$, $p = .057$ |
| Item-by-item prospective | **$r = 0.66$, $p < .001$** | **$r = 0.45$, $p = .003$** |
| Item-by-item retrospective | **$r = 0.72$, $p < .001$** | **$r = 0.49$, $p = .001$** |
| Global retrospective | **$r = 0.47$, $p = .002$** | **$r = 0.43$, $p = .005$** |

**Table 3**. Bonferroni corrected correlations between metacognitive judgements and performance for the forward digit span. As there are 4 correlations for each group the significance threshold is equal to 0.05/4 = 0.013.

| | Forward span | | |
|---|---|---|---|
| | AD patients | Older adults | *z* value |
| Global prospective | *r* = -0.03, *p* = .914 | *r* = 0.04, p = .854 | -0.21 |
| Item-by-item prospective | ***r* = 0.60, *p* = .009** | ***r* = 0.56, *p* = .005** | 0.18 |
| Item-by-item retrospective | *r* = 0.47, *p* = .049 | *r* = 0.20, *p* = .362 | 0.93 |
| Global retrospective | ***r* = 0.77, *p* < .001** | ***r* = 0.55, *p* = .007** | 1.21 |

**Table 4**. Bonferroni corrected correlations between metacognitive judgements and performance for the backward digit span. As there are 4 correlations for each group the significance threshold is equal to 0.05/4 = 0.013.

| | Backward span | | |
|---|---|---|---|
| | AD patients | Older adults | *z* value |
| Global prospective | *r* = 0.26, *p* = .297 | *r* = 0.37, *p* = .087 | -0.37 |
| Item-by-item prospective | *r* = 0.28, *p* = .256 | ***r* = 0.60, *p* = .003** | -1.22 |
| Item-by-item retrospective | *r* = 0.56, *p* = .015 | *r* = 0.36, *p* = .098 | 0.77 |
| Global retrospective | *r* = 0.17, *p* = .508 | ***r* = 0.65, *p* < .001** | -1.82 |

## Discussion

Previous studies of metacognition in Alzheimer's disease have focused on long term memory. Here we investigated the awareness of short term memory and working memory. We replicated the documented deficits in both digit span forward and digit span backward in Alzheimer's disease (e.g. Morris & Baddeley, 1988). In addition, we showed that people with Alzheimer's disease are as accurate as controls at assessing this function, despite the deficit in performance. To consider the importance of this finding for our understanding of metacognition and anosognosia in Alzheimer's disease, we must consider how participants are able to make accurate judgements on these tasks.

Regarding item-by-item predictions, we propose that when given the digits to memorise, even in the 'dry-run' prediction phase, participants test themselves. People with

Alzheimer's disease run through the digits presented, as do controls, and have access to whether or not they will be able to complete the task. Because there are no dual demands of performance and prediction, participants are able to directly report this information: in the here-and-now they can accurately gauge whether they can retain (or retain and manipulate, in the case of digits backward) the information. For retrospective judgements, where there is a preservation in Alzheimer's disease in long term memory (e.g., Gallo, Cramer, Wong, & Bennett, 2012; Moulin, James, Perfect, & Jones, 2003), patients are able to use correctly the feedback arising from this self-test to make accurate predictions.

Turning to global predictions, we found that for the forward span the first prediction is less accurate than the retrospective one and item-by-item judgements (for both patients and older adults). This effect is also typically observed in long term memory tasks for Alzheimer patients (e.g., Silva et al., 2017). Moreover, correlational analyses at the group level bring additional evidence to this. It has been shown in both Alzheimer's disease (Silva et al., 2017) and with older adults (Connor, Dunlosky, & Hertzog, 1997) very low correlations between initial global predictions and performance. This was not the case for later retrospective predictions. Thus, when they can experience the task, both older adults and Alzheimer's disease patients update their knowledge about the task and make accurate judgements. Although results are clear for the forward span task (i.e., significant difference and very low correlation, $r = 0.06$), this is less the case for the backward digit span. We did not find a significant difference in terms of accuracy for this task, and correlational analyses showed no significant relationship between global prospective judgement and performance. However, this low correlation was not different from the other three (Fisher's z, all $p > .05$). The backwards digit span data do not therefore follow exactly the pattern of overestimation and initial inaccuracy found in the Alzheimer's group on previous long term memory tasks and here in our own experiment. Critically, we find no evidence for any group differences in accuracy or magnitude of predictions on this task either. We might hypothesise in general that people anticipate the backwards digit span is a difficult task.

It is important to discuss the large variability for accuracy in patients. It is possible that more patients overestimate their performance (see supplementary results). We counted the number of participants who overestimated and indeed found that patients overestimate more than controls for the prospective global prediction. For the forward span, 33% of patients overestimate their performance compared to 17% for controls. The same is observed for backward span, 56% of patients overestimate their performance compared to 26% for controls.

Overall, patients and controls have the same judgement accuracy distributions (see supplementary results). This result has implications for metacognition and anosognosia more generally in Alzheimer's disease. Despite having a deficit in short-term and working memory, the patients with Alzheimer's are nonetheless able to reliably report their difficulties with this task: the magnitude of judgements is different from controls (although a trend for the backward span). This is in direct contrast with tasks which require memory retrieval. On (long term) episodic memory feeling of knowing tasks (e.g. Souchay et al., 2002), patients are unable to reliably gauge whether a previously studied word is available or not when tested by recognition. This is proposed to be due to the impoverished information available to the person with Alzheimer's disease: they cannot evaluate their memory accurately, because they cannot retrieve from memory enough pertinent information on which to base their judgement. In contrast, even for a relatively complex task, such as reversing and repeating a series of digits as tested here, in the here-and-now patients with Alzheimer's disease are aware of their difficulties.

Taken together, these results support the idea of a preservation of online monitoring in Alzheimer's disease. When they can test themselves or when they have already performed the task, both older adults and patients are able to use the performance feedback to adjust their metacognitive judgements. Naturally, this has major clinical implications. On-line, whilst struggling with a task, a patient with Alzheimer’s disease will be aware of their difficulties, even if when asked later they are not aware of how difficult the task was, or indeed, when encountering the same task again, they will not beforehand know how difficult they will find it. It would be important to replicate the likely deficit for patients in initial global prediction and to add a second trial after a delay. If this impairment is also found for a second trial, this would be in line with the failure to transfer information from online evaluations into long-term representations (Morris & Mograbi, 2013). Interestingly, Stewart et al. (2010) show that whilst global judgements may be accurate for long term memory tasks, the accuracy that is acquired is forgotten as soon as one hour later. We add another task to the literature for which people with Alzheimer’s can accurately gauge their performance. The impact of this work is that people with Alzheimer’s are able to reflect upon their performance in a task which is critical for daily function: short term memory.

Anosognosia, however, is likely to remain a multifaceted construct, with varying causes and manifestations. Whilst it is clear memory mechanisms are pertinent to tasks which involve memory, different domains should be compared (Chapman, Colvin, Vuorre, et al., 2018) and the involvement of other process such as executive function (Scherling, Wilkins, Zakrezewski, et al., 2017) perspective taking (Serino & Riva, 2017) , and emotion (need to be examined in detail (Bertrand, Dourado, Laks, et al., 2016).

**Acknowledgments**

This work was supported by a research studentship given to Julie Bertrand, provided by the Region of Bourgogne (FABER), “Conscience et Mémoire dans la Maladie d’Alzheimer” awarded to Céline Souchay and Chris Moulin. Chris Moulin is supported by the Institut Universitaire de France.

## Conflict of interest

The authors have no conflict of interest to report.

## References

Ansell, E. L., & Bucks, R. S. (2006). Mnemonic anosognosia in Alzheimer's disease: A test of Agnew and Morris (1998). *Neuropsychologia*, *44*(7), 1095-1102.

Bäckman, L., & Lipinska, B. (1993). Monitoring of general knowledge: Evidence for preservation in early Alzheimer's disease. *Neuropsychologia*, *31*(4), 335-345.

Barrett, A. M., Eslinger, P. J., Ballentine, N. H., & Heilman, K. M. (2005). Unawareness of cognitive deficit (cognitive anosognosia) in probable AD and control subjects. *Neurology*, *64*(4), 693-699.

Bertrand, E., Dourado, M. C., Laks, J., Morris, R. G., Landeira-Fernandez, J., & Mograbi, D. C. (2016). Mood-congruent recollection and anosognosia in Alzheimer's disease. *Cortex*, *84*, 55-62.

Bertrand, J. M., Moulin, C. J. A., & Souchay, C. (2017). Short-term memory predictions across the lifespan: monitoring span before and after conducting a task. *Memory*, *25*(5), 607-618.

Chapman, S., Colvin, L. E., Vuorre, M., Cocchini, G., Metcalfe, J., Huey, E. D., & Cosentino, S. (2018). Cross domain self-monitoring in anosognosia for memory loss in Alzheimer's disease. *Cortex*, *101*, 221-233.

Clare, L., Marková, I. S., Roth, I., & Morris, R. G. (2011). Awareness in Alzheimer's disease and associated dementias: theoretical framework and clinical implications. *Aging & mental health*, *15*(8), 936-944.

Cosentino, S., Zhu, C., Bertrand, E., Metcalfe, J., Janicki, S., & Cines, S. (2016). Examination of the metacognitive errors that contribute to anosognosia in Alzheimer's disease. *Cortex*, *84*, 101-110.

Ernst, A., Moulin, C. J., Souchay, C., Mograbi, D. C., & Morris, R. (2016). Anosognosia and Metacognition in Alzheimer's disease: Insights from Experimental Psychology. *The Oxford handbook of metamemory, 1*, 456-79.

Flavell, D. M., Jamshidi, Y., Hawe, E., Pineda Torra, I., Taskinen, M. R., Frick, M. H., ... & Miller, G. (2002). Peroxisome proliferator-activated receptor α gene variants influence progression of coronary atherosclerosis and risk of coronary artery disease. *Circulation*, *105*(12), 1440-1445.

Flavell, J. H., Friedrichs, A. G., & Hoyt, J. D. (1970). Developmental changes in memorization processes. *Cognitive psychology*, *1*(4), 324-340.

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.

Gallo, D. A., Cramer, S. J., Wong, J. T., & Bennett, D. A. (2012). Alzheimer's disease can spare local metacognition despite global anosognosia: Revisiting the confidence–accuracy relationship in episodic memory. *Neuropsychologia*, *50*(9), 2356-2364.

Hester, R. L., Kinsella, G. J., & Ong, B. E. N. (2004). Effect of age on forward and backward span tasks. *Journal of the International Neuropsychological Society*, *10*(4), 475-481.

Morris, R. G., & Baddeley, A. D. (1988). Primary and working memory functioning in Alzheimer-type dementia. *Journal of clinical and experimental neuropsychology*, *10*(2), 279-296.

Morris, R. G., & Mograbi, D. C. (2013). Anosognosia, autobiographical memory and self knowledge in Alzheimer's disease. *Cortex*, *49*(6), 1553-1565.

Moulin, C. J., James, N., Perfect, T. J., & Jones, R. W. (2003). Knowing what you cannot recognise: Further evidence for intact metacognition in Alzheimer's disease. *Aging, Neuropsychology, and Cognition*, *10*(1), 74-82.

Moulin, C. J., Perfect, T. J., & Jones, R. W. (2000a). Global predictions of memory in Alzheimer's disease: Evidence for preserved metamemory monitoring. *Aging, Neuropsychology, and Cognition*, *7*(4), 230-244.

Moulin, C. J., Perfect, T. J., & Jones, R. W. (2000b). Evidence for intact memory monitoring in Alzheimer's disease: Metamemory sensitivity at encoding. *Neuropsychologia*, *38*(9), 1242-1250.

Moulin, C. J. A., Perfect, T. J., & Jones, R. W. (2000). The effects of repetition on allocation of study time and judgements of learning in Alzheimer's disease. *Neuropsychologia*, *38*(6), 748-756.

Murphy, M., Sanders, R., Gabriesheski, A., & Schmitt, F. (1981). Metamemory in the aged. *Journal of Gerontology*, *36*(2), 185–193.

Scherling, C. S., Wilkins, S. E., Zakrezewski, J., Kramer, J. H., Miller, B. L., Weiner, M. W., & Rosen, H. J. (2016). Decreased self-appraisal accuracy on cognitive tests of executive functioning is a predictor of decline in mild cognitive impairment. *Frontiers in aging neuroscience*, *8*, 120.

Serino, S., & Riva, G. (2017). The proactive self in space: How egocentric and allocentric spatial impairments contribute to anosognosia in Alzheimer's Disease. *Journal of Alzheimer's Disease*, *55*(3), 881-892.

Silva, A. R., Pinho, M. S., Macedo, L., Souchay, C., & Moulin, C. (2017). Mnemonic anosognosia in Alzheimer's disease is caused by a failure to transfer online evaluations of performance: Evidence from memory training programs. *Journal of clinical and experimental neuropsychology*, *39*(5), 419-433.

Stewart, G., McGeown, W. J., Shanks, M. F., & Venneri, A. (2010). Anosognosia for memory impairment in Alzheimer's disease. *Acta Neuropsychiatrica*, *22*(4), 180-187

Souchay, C. (2007). Metamemory in Alzheimer's disease. *Cortex*, *43*(7), 987-1003.

Souchay, C., Isingrini, M., & Gil, R. (2002). Alzheimer's disease and feeling-of-knowing in episodic memory. *Neuropsychologia*, *40*(13), 2386-2396.

Souchay, C., Isingrini, M., & Gil, R. (2002). Alzheimer's disease and feeling-of-knowing in episodic memory. *Neuropsychologia*, *40*(13), 2386-2396.

Sunderaraman, P., & Cosentino, S. (2017). Integrating the constructs of anosognosia and metacognition: a review of recent findings in dementia. *Current neurology and neuroscience reports*, *17*(3), 17-27.

## Supplementary results

To explore more deeply the accuracy of metacognitive judgements in both patients and older adults we calculated the signed difference between judgements and performance. For all judgements we then plotted the number of patients and older adults according to the signed accuracy score (Figure 3). We computed Chi² to compare distributions of accuracy in both group. These analyses revealed no difference for the forward span task, $X^2 = 56.83$, $p = .179$, and the backward span task, $X^2 = 41.25$, $p = .251$.

Using the signed difference between judgements and performance for global prospective judgement only, we classified both patients and older adults in 3 categories: accurate (when accuracy was equal to 0), underestimators (when accuracy was negative), and overestimators (when accuracy was positive). Table 2 and 3 show the frequency of each category for patients and controls.

**Table 2**. Frequency of overestimators, accurate estimators, and underestimators for patients and older adults for the forward span task.

| | Underestimators | Accurate | Overestimators |
|---|---|---|---|
| Older adults | 0.61 | 0.22 | 0.17 |
| Patients | 0.44 | 0.22 | 0.33 |

**Table 3**. Frequency of overestimators, accurate estimators, and underestimators for patients and older adults for the backward span task.

| | Underestimators | Accurate | Overestimators |
|---|---|---|---|
| Older adults | 0.43 | 0.30 | 0.26 |
| Patients | 0.27 | 0.17 | 0.59 |

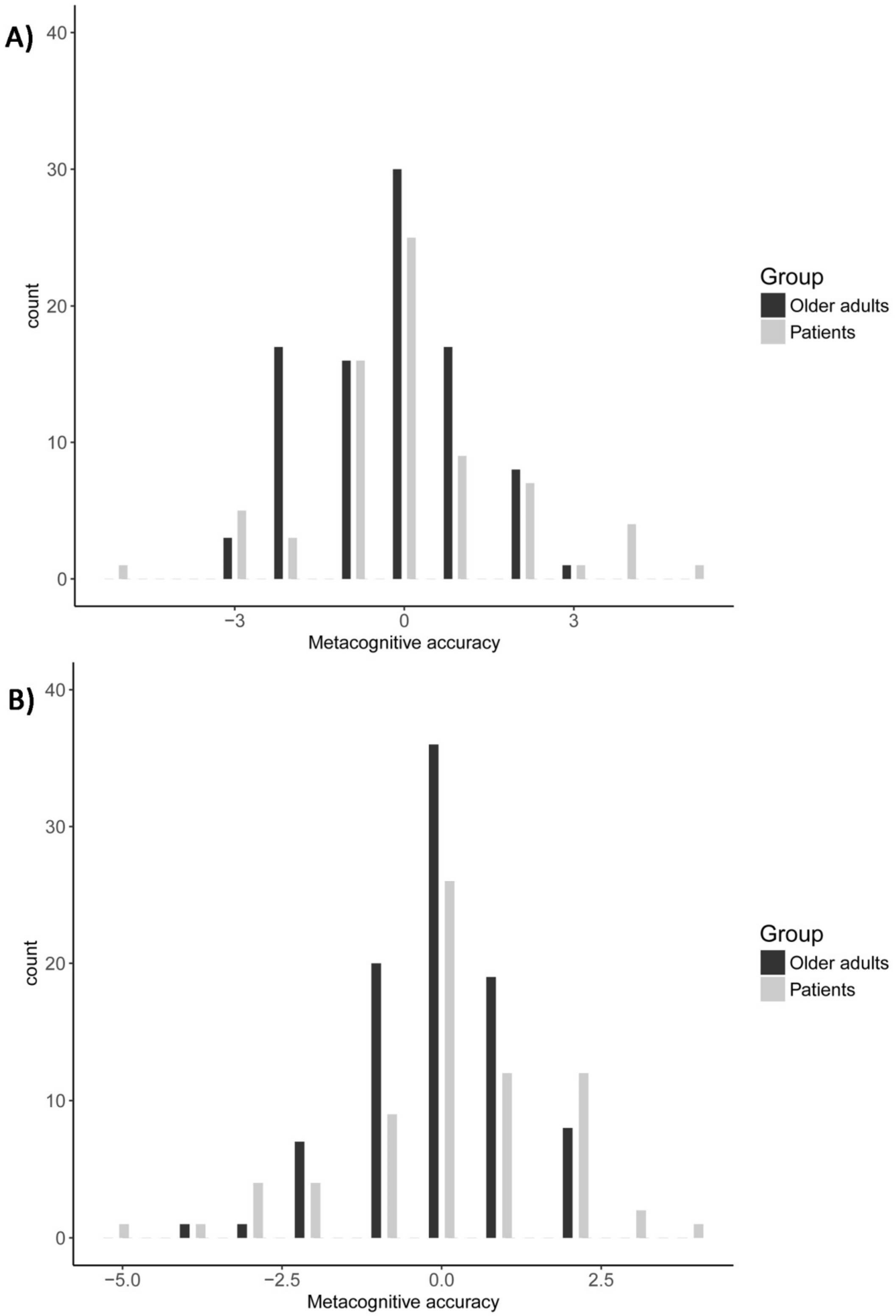


**Figure 3**. (A) Distribution of metacognitive accuracy for all judgements for patients and older adults for the forward span task. (B) Distribution of metacognitive accuracy for all judgements for patients and older adults for the backward span task.

Finally, we focus here on analyses regarding accuracy of judgement we outliers exclusion. We chose to report them here as the excluded participants are two patients (i.e., one for forward span and one for backward span). We understand that excluded patients might be tricky in a sense that they might have a non-normal behaviour by definition. These outliers were excluded because they had a gap in Cook's distance more that three and a Studentized Deleted Residuals < |4|.

**Forward span.** Analyses showed neither a significant effect of group, $t(38) = 0.98$, $p = .335$, $d = 0.31$. We did however, find a non-significant trend main effect of judgement type, $t(38) = 1.96$, $p = .058$, $d_z = 0.31$, and a main effect of judgements time, $t(38) = 2.62$, $p = .012$, $d_z = 0.41$. Prospective judgements were less accurate than retrospective judgements. There was also a significant interaction between judgement type and judgement time, $t(38) = 3.42$, $p = .002$, $d_z = 0.54$. Retrospective judgements are more accurate for global predictions, $t(38) = 4.16$, $p < .001$, $d_z = 0.66$, which is not the case for item-by-item judgements, $t(38) = -0.69$, $p = .494$, $d_z = -0.11$. We also find a significant interaction between judgement time and group, $t(38) = -2.42$, $p = .020$, $d = -0.77$. However, when we observed simple main effects, we just find a trend for a less accurate prospective judgements for patients, $t(38) = -1.88$, $p = .068$, $d = 0.60$, and no difference for retrospective judgements, $t(38) = 0.41$, $p = .688$, $d = 0.13$. No other interaction was significant (all $t(38) < 1.01$, and all $p > .05$).

**Backward span.** As for forward span task, we conducted linear regressions with, group, prediction time (prospective vs retrospective), and judgement type (global vs item-by-item) as factors. The analysis showed neither a significant effect of group, $t(38) = 1.47$, $p = .149$, $d = 0.47$, $d_z = 0.01$, nor a significant effect of judgement time, $t(38) = 0.09$, $p = .929$, nor a significant effect of judgement type, $t(38) = 1.14$, $p = .261$, $d_z = 0.18$. No interaction was significant (all $t(38) < 0.84$, and all $p > .05$).