Mechanisms underlying corruption of working memory in Parkinson's disease

Sean James Fallon¹,² | Chevonne van Rhee¹ | Annika Kienast¹ | Masud Husain¹,³
Sanjay G. Manohar¹,³

¹Department of Experimental Psychology, University of Oxford, Oxford, UK
²School of Psychology, University of Plymouth, Plymouth, UK
³Nuffield Department of Clinical Neurosciences, John Radcliffe Hospital, Oxford, UK

Correspondence
Sean James Fallon, Department of Experimental Psychology, University of Oxford, Oxford, UK.
Email: sean.fallon@plymouth.ac.uk and sj.fallon@bristol.ac.uk

Abstract
Working memory (WM) impairments are reported to occur in patients with Parkinson's disease (PD). However, the mechanisms are unclear. Here, we investigate several putative factors that might drive poor performance, by examining the precision of recall, the order in which items are recalled and whether memories are corrupted by random guessing (attentional lapses). We used two separate tasks that examined the quality of WM recall under different loads and retention periods, as well as a traditional digit span test. Firstly, on a simple measure of WM recall, where patients were asked to reproduce the orientation of a centrally presented arrow, overall recall was not significantly impaired. However, there was some evidence for increased guessing (attentional lapses). On a new analogue version of the Corsi-span task, where participants had to reproduce on a touchscreen the exact spatial pattern of presented stimuli in the order and locations in which they appeared, there was a reduction in the precision of spatial WM at higher loads. This deficit was due to misremembering item order. At the highest load, there was reduced recall precision, whereas increased guessing was only observed at intermediate set sizes. Finally, PD patients had impaired backward, but not forward, digit spans. Overall, these results reveal the task- and load-dependent nature of WM deficits in PD. On simple low-load tasks, attentional lapses predominate, whereas at higher loads, in the spatial domain, the corruption of mnemonic information—both order item and precision—emerge as the main driver of impairment.

KEYWORDS
attention, misbinding, Parkinson's disease, spatial span, working memory
INTRODUCTION

Working memory (WM) refers to our ability to maintain and manipulate information over a short time period in order to perform goal-directed tasks (Baddeley, 2012). As such, ensuring WM functions effectively is essential for our daily lives. Reflecting this, measures of WM performance have been found to be related to real-world measures of success, such as academic achievements (Gathercole et al., 2003). WM functioning is also a key index of health and is impaired in brain disorders (Baddeley et al., 1991; Christopher & MacDonald, 2005). Though, traditionally conceived as a movement disorder, it has now been established that cognitive deficits are also a core feature of Parkinson's disease (PD; Aarsland et al., 2017; Lawson et al., 2021), with deficits in WM forming part a key component of the executive dysfunction prominent in this group (Morris et al., 1988).

Deficits in short-term recall in PD are widely considered to result from the dopaminergic degeneration that characterises the disorder (Sawamoto et al., 2008). Despite this, PD patients taking dopamine-enhancing medication have still been reported to display poor WM performance (Owen et al., 1992). These residual WM deficits may instead result from non-dopaminergic pathology to noradrenergic or cholinergic systems (Kehagia et al., 2010) or possibly even dopaminergic overdosing (Fallon et al., 2015; Fallon, Gowell, et al., 2019). However, the generation of pharmacological strategies to overcome this problem has been hampered by a lack of a full characterisation of medicated patients' WM deficits. Specifically, the circumstances under which impairments appear and the recall parameters that are affected have not been fully delineated.

Spatial WM has been argued to be particularly vulnerable in PD (Owen et al., 1997; Postle et al., 1997). This issue has most extensively been examined using Corsi-like block-tapping tasks (Milner, 1971) in which individuals encode the location and order of a series of spatial locations (physical blocks or highlighted blocks on a screen), which they are then asked to subsequently reproduce. Although some authors have reported impairments (Fournet et al., 2000; Kemps et al., 2005; Stoffers et al., 2003), preserved performance has also been observed in PD (Owen et al., 1993), with deficits sometimes being more prominent for verbal than spatial material (Gruszka et al., 2016). A subsequent study isolated some of these inconsistencies as being due to variations in the structure (or ‘chunkability’) of the memoranda, with patients performing as well as controls when recalling easy-to-chunk memoranda, but worse when this information was not easy to parse (Fallon, Bor, et al., 2017). However, that study was unable to specify why PD patients found the unstructured material difficult to remember, that is define which parameters of patients’ recall are deficient. This is important as the previous study cannot, even though an impairment was found, rule out that this was due to impaired memory or impaired performance, for example attentional lapses or a response bias (e.g. shifting all responses left or right). In this context, a response bias means a tendency for there to be a systematic bias in reproducing spatial information such that the whole constellation of remembered spatial locations are reported as being to the left or right of the actual location.

Traditional means of probing the contents of memory usually consist of participants making binary judgement about whether a probe item was previously presented, for example delayed match-to-sample or discrete locations in the Corsi-like block-tapping tasks. These methods may misrepresent mnemonic abilities or, at the very least, are a relatively crude measure of memory fidelity. Recent methodological developments in measuring WM have fundamentally challenged traditional conceptions of the processing governing short-term recall (Ester et al., 2012; Ma et al., 2014; Oberauer & Hein, 2012). The development of delayed report methods, in which participants are required to reproduce the exact features of memoranda such as a line’s orientation, has revealed that recall can be impaired for several reasons: the precision of feature memory can be reduced (noisier representation), participants might guess randomly (attentional lapses), or the features of different objects might become confused with each other—sometimes termed misbinding (Bays et al., 2009).

Studies that have used these methodologies and computational techniques have revealed that they are more sensitive in detecting deficits in PD (Zokaei, Burnett Heyes, et al., 2014). Moreover, different neurocognitive mechanisms appear to be responsible for impaired precision, guessing and misbinding. On
a spatial WM task, increased rates of misbinding have been linked to demonstrable, or suspected, damage to medial temporal lobe regions. (Liang et al., 2016; Pertzov et al., 2013; Zokaei et al., 2019) In contrast, in PD—unlike Alzheimer’s disease (Zokaei et al., 2020)—there is evidence for increased guess responses (Zokaei et al., 2020). Increased guess responses have not been found in and can be exacerbated by disruptions to the dopaminergic system (Fallon, Mattiesing, et al., 2017; Fallon, Muhammed, et al., 2019). As well as the implication that separate neural systems underlie these dissociations, cognitive variables such as delay (between encoding and retrieval), set size, filtering and ignoring irrelevant information have also been found to have distinct effects on precision, guessing and misbinding (Fallon et al., 2018; Pertzov et al., 2012; Tabi et al., 2021). Further, the results of a previous study show that precision working metrics can be more sensitive than simple span measures in PD (Zokaei, Burnett Heyes, et al., 2014). However, despite the utility these measures can have in fractionating the mnemonic impairments different clinical cohorts display, a spatial span task that allows these measures to be extracted has not been developed. Thus, there is a pressing need for a new paradigm that provides a continuous response space, that is where responses are not confined to discrete spatial locations. Using a more continuous response space in span tasks [as in previous spatial WM studies (Pertzov et al., 2012)] may enable measures of precision, misbinding and guesses to be extracted.

A common problem is that measures of spatial WM and memory for orientation have rarely been performed in the same patients, alongside traditional measures such as digit span. This has prevented researchers from determining whether patient have domain-dependent deficits, and whether consistent deficits on these sub-measures of WM are found across domains. Here, we try to overcome this limitation by using a new analogue spatial span task where we can experimentally manipulate the number of items encoded and also track the precision of recall for individual items. In this new version of the task, a series of spatial locations are presented on a touchscreen. After encoding, the participant has to reproduce the sequence by touching the locations on screen. This makes it possible to measure the precision (Euclidean distance) between the encoded item and the probe and also examine how the sequence of spatial locations is systematically corrupted during recall, that is are the items reported in the wrong order. In this study we use this task to clarify how robust spatial WM in PD is under highly mnemonic loads and isolate the components of recall that are affected. To buttress these findings and enable generalisations about non-spatial visual WM in patients we also administered a simple short-term memory task that assessed the memory for orientations of single items in patients.

METHOD

Participants

The Parkinson's disease (PD) group was recruited from a database of PD patients. Only non-colour-blind participants with normal or corrected-to-normal vision were eligible to take part. All participants provided written informed consent prior to participation. Approval for this study was given by the local ethics committee. Table 1 provides a summary of participant demographics and performance on neuropsychological tasks. Non-parametric analyses did not reveal significant differences in age ($p = .98$) or ACE scores ($p = .181$). Gender balance tended to differ, but this was no significant ($p = .055$).

Design and procedure

Prior to the main experimental paradigms, the Addenbrookes Cognitive Examination (ACE; Bak & Mioshi, 2007) and Digit Span (DS; e.g. Blackburn & Benton, 1957) were administered to the HC and PD groups. Participants were seated in front of a touchscreen monitor (Dell Inpsiron All-In-One) with a resolution of $1920 \times 1080$ pixels with a viewing distance of approximately 42 cm between the participant and the screen. All computerised tasks were carried out in MATLAB Psychtoolbox version 3.0 (Kleiner
et al., 2007) All patients were tested ‘ON’ their usual dopaminergic medication and took these medications according to their usual schedule.

Standard digit span

Participants were asked to perform the standard digit span task. In this task a series of numbers are read out aloud to the participant. Two manipulations were included in the present study. Participants either had to repeat the digits in the order presented (forward) or repeat the digits back in the reverse order to that presented (backward). Participants were given a score of one for each sequence correctly recalled and zero if the sequence was incorrectly recalled. If recall was unsuccessful on two consecutive trials for each set size, the test was stopped.

Simple one-item delayed reproduction task

Simple short-term recall was assessed in a task that has been used previously (Fallon, Muhammed, et al., 2019). Briefly, participants had to remember the orientation of a single, centrally presented arrow after a variable delay period (1000 ms or 2000 ms; Figure 1a). The fidelity of recall was assessed by asking participants to match the orientation of a probe arrow to the orientation of the previously presented arrow (e.g. by rotating the probe arrow clockwise (‘A’ key) or anti-clockwise (‘Z’ key). When participants were happy with the rotated orientation, they had to press the ‘Space’ bar. There 48 trials for each delay duration (96 in total).

Precision of recall was measured using assess the standard deviation of responses from the target (1/circular standard deviation). This was calculated according to the JV10_error function (Bays et al., 2009). We also examined whether there were biases in recall (clockwise or counterclockwise).

To complement this model-free metric we also subjected the data to computational modelling using Memfit (Suchow et al., 2013). We applied a very basic model (Zhang & Luck, 2008). This model sees WM recall as comprising two components, represented in the following equation:

$$p(\hat{\theta}) = (1 - \gamma)\phi_\kappa(\hat{\theta} - \theta) + \gamma \frac{1}{2\pi}$$

1. Variability in recall (Standard deviation).
2. Probability of guessing ($\gamma$).

where $\hat{\theta}$ is the response angle participants provide, $p(\hat{\theta})$ is the probability of the given response, $\phi_\kappa$ is a von Mises probability density function centered on zero with concentration reflecting the standard deviation of
responses, $\theta$ is the probed target angle (memoranda). The probability of guess responses are given by the $\gamma$ parameter and the probability of responding to the target orientation is given by $1 - \gamma$. These parameters were obtained using maximum-likelihood-derived parameters. In simple terms, this model tries to estimate the proportion of the response distribution that is driven by changes in precision (standard deviation around the target orientation) and guesses (responses that are uniformly distributed from 0 to 2$\pi$). An additional wrapper was included to measure response biases (clockwise or counterclockwise shifts in responses). Maximum likelihood estimates were obtained separately by running the model on each trial type (1000 ms or 2000 ms) and subject.

### Precision spatial span task

Traditional span paradigms such as the Corsi blocks task require participants to remember discrete spatial locations and then assess recall by asking participants to reproduce these locations by tapping on discrete spatial locations, in the order in which they appeared. Thus, recall is dichotomised into successful or failed. In this new paradigm, we develop a procedure for measuring the precision of spatial memory across different set sizes.

Participants had to remember the locations and order in which a series of dots that appeared on the screen (Figure 1). They then had to reproduce this spatial sequence, touching the empty screen on the locations on the screen where the items appeared (in the correct order). Set size was experimentally manipulated (1–6 locations). Participants performed 30 trials (five for each set size). The order of the trials and the spatial locations of the trials was the same for each participant.

Prior to the study commencing, the spatial locations used for each set size were generated randomly under the proviso that there was a minimum distance between each spatial location within a sequence (three times the visual size of each item roughly subtended $\sim$13° visual angle). Randomly generated spatial locations were visually examined, curated and checked for their distribution around the screen. These spatial locations were kept constant across all participants (see Figure S1).

The main index of gross performance on this task were the average spatial error for each set size. This was calculated as the pairwise Euclidean distance between the respective target and response locations according to the following equation:

$$\text{Average Error} = \frac{\sum_{i=1}^{N} \sqrt{(\text{response}X_i - \text{target}X_i)^2 + (\text{response}Y_i - \text{target}Y_i)^2}}{N}$$

where $X$ is the horizontal axis position, $Y$ is the vertical axis position, and $N$ is the set size.

Average error indicates overall performance, but it does not indicate whether the ordinal positions of items in memory has been retained. Examining the serial reproduction errors participants make during recall may provide a more sensitive or complementary measure of memory. Accordingly, we measured the propensity to confuse the order in which items were recalled. Algorithmically, this was done by matching each response location to the nearest target location. We can then assign a number based on the ordinal position to each of the response locations corresponding to the order in the encoded memoranda. For example, if in the six-location condition, the 3rd and 6th locations were swapped, the response sequence would be 1 2 6 4 5 3. The Chebyshev distance (maximal displacement) can then be used to provide a simple metric to quantify the extent to which the sequence was displaced, that is (6–3) = 3.

Given that we can gauge some indication of serial reproduction errors, we can also try and examine the precision of spatial memory as if these errors never took place, that is we can virtually ‘repair’ the responses by matching each response location to the target locations irrespective of the order in which they were reproduced and then recompute the Euclidean distance as in the earlier equation. This gives us a measure (upper bound) of the precision of memory independent of misbinding errors that is similar to the precision values obtained for the orientation data.
We can also construct a metric to examine how much participants’ responses differed from pure chance, that is giving an index of how much trial-specific information was encoded. To do this, for each trial we generated 10,000 random location sets (10,000 × N, where N is the set size). We then computed the average Euclidean error (nearest neighbour) for each of these random sequences just as we did for the participants’ responses. This gives us a distribution of error distances (chance distribution), which can be used to assess how likely participants’ responses were to come from a pure guessing distribution, or not. To assign a probability of chance responding to each trial for each participant, we can rank participants actual repaired error distance for a trial by comparing it with the error distribution.

Analysis

R (3.6.3)/RStudio (1.2.5003) was used to analyse the data. The data were analysed using non-parametric methods. Specifically, we used the Aligned rank transform (Wobbrock et al., 2011) to perform non-parametric ANOVAs. Where applicable alternative non-parametric analyses (lmPerm) are indicated (Wheeler & Torchiano, 2016).

RESULTS

PD patients have reduced backward digit spans

First, we examined whether patients and controls differed in terms of their performance on forward and backward digit spans. There was no significant difference between patients [11.6 (2.3)] and controls [11.9 (1.8)] for forward digit span scores (U = 147, p = .628). However, backward span score was significantly lower in patients [6.8 (2.1)] compared to controls (8.7 (2.2), U = 260, p = .020).

Enhanced guess rates and delay-dependent response biases

Short-term memory performance in patients and controls was assessed by asking them to reproduce the orientation of a presented arrow after 1000 ms or 2000 ms (Figure 1a). First, we examined the precision of recall (1/circular standard deviation) according to disease (control, PD) with delay (1000, 2000 ms). There was a trend for PD patients to have poor precision compared to controls (F(1,36) = 3.35, p = .07; Figure 2a). However, there was no significant main effect of delay (1000 ms vs. 2000 ms) on recall (F(1,36) = 1.84, p = .18) and no significant interaction between disease and delay (F < 1).

We next examined bias (clockwise, counterclockwise) in responding (Figure 2b). Patients were not found to have a significant overall response bias compared to controls (main effect of disease; F(1,36) = 2.11, p = .15). Bias did not significantly vary according to delay (F < 1). However, there was evidence for a significant interaction between disease and delay (F(1,36) = 6.45, p = .015). This was due to there being no significant difference between patients and controls for the short (1000 ms) duration (W = 168, p = .73), but patients were found to have a counterclockwise bias compared to controls for the 2000 ms period (W = 100, p = .018).

Thus, overall, these results suggest that PD patients have a delay-dependent bias in the recall of orientations, but that the precision with which this information is recalled is similar between patients and controls.

Precision (1/SD) can be an amalgam of different components. To further dissect participants’ performance on this task we ran a computational model to extract different aspects of performance (precision, guessing parameter and bias). For guessing, a non-parametric linear model with disease and delay suggested that there was some evidence for an effect of disease (F(1,36) = 4.08, p = .050), with PD patients showing higher levels of chance responding compared to controls (Figure 3a). There was no effect of delay or interaction between disease and delay (Fs < 1). For the modelled precision (Figure 3b), there was no signif-
significant effect of delay ($F(1,36) = 2.10, p = .15$) or interaction between delay and disease ($F$'s $< 1$), consistent with non-modelled (raw) results.

Finally, we examined the modelled response bias (clockwise, counterclockwise). This analysis also agreed with the non-modelled results. There were no significant main effects of disease or delay ($F$'s $< 1$), but there was some evidence for a significant interaction between disease and delay ($F(1,36) = 5.53, p = .02$). Again, this was due to there being no significant difference between patients and controls for short delays ($W = 189, p = .80$), but there was a non-significant trend for a difference at longer delays ($W = 246, p = .054$).

In summary, both the analogue analysis of the raw error and modelling point to altered recall in PD patient even for short (<2000 ms) durations. Modelling suggests that the differences are most likely due to guessing, which was found to be significantly higher in patients compared to controls. Analogue measures of response bias—and to a lesser extent modelled measures of response bias—suggest that patients have a significant bias in recall that is apparent only at longer (2000 ms) delays.
Healthy Controls
Parkinson’s disease

(a) Probability of Guessing

(b) Standard deviation (radians)

(c) Bias (radians)
Impaired precision and increased serial order errors at higher set sizes

Next, we examined spatial WM on the analogue version of the Corsi task. One participant in the healthy control group was excluded from the analysis due to high imprecision on 1-location trials. For mean precision across set sizes (Figure 4a), there was a main effect of set size, with error (from true location) progressively increasing across the higher set sizes ($F(5,180) = 221.12, p < .00001$). There was a non-significant trend for a main effect of disease ($F(1,36) = 3.01, p = .09$) and a significant interaction between disease and set size ($F(5,180) = 2.35, p = .04$). This was due to PD patients tending to have higher errors for larger set sizes (e.g. for 6-item condition, $W' = 115, r_{biserial} = .36, p = .058$) but not for lower set sizes (e.g. one-item condition, $W' = 209, r_{biserial} = .16, p = .40$). Thus, PD patients showed evidence for relatively impaired performance at larger set sizes.

We then examined three components of error that can contribute to poor performance (repaired error distance, serial reproduction errors and guessing). For the virtually ‘repaired’ sequences error (see “Methods” section) the average error was also found to increase with set size ($F(5,180) = 287, p < .00001$; Figure 4b). PD patients showed significantly higher repaired error distances than controls ($F(1,36) = 5.43, p = .025$). However, the presence of this impairment was found to vary significantly with set size (disease × set size interact; $F(5,180) = 3.44, p = .005$). These match the overall error results, suggesting that the disease effect is not ‘repairable’ by ignoring item bindings. Pairwise comparisons revealed that PD patients were impaired only at the highest set size ($W' = 107, r_{biserial} = .40, p = .03$). There was no significant difference between patients and controls for the 1-item condition ($W' = 209, r_{biserial} = .16, p = .40$).

With regard to serial misplacements (Chebyshev distance for set size greater than 1; Figure 4c), these significantly increased, unsurprisingly, with increasing set size ($F(4,144) = 214, p < .00001$). PD patients made significantly more errors in reproducing the correct sequence of responses ($F(1,37) = 10.76, p = .002$). However, the deficit they displayed varied significantly with set size ($F(4,144) = 5.34, p = .0004$). Pairwise comparisons revealed that it was only at the highest set size (six items) that patients were significantly impaired compared to controls ($W' = 97.5, p = .016, r_{biserial} = .45$). None of the differences for the other set sizes were significant (all $p's > .21$). Thus, this analysis shows that PD patients systematically misreported the order in which items were presented for the highest set size, an effect that was not significant for lower set sizes.

Biases in responding (undershooting or overshooting in the vertical or horizontal plane) may also contribute to produce higher errors distances in patients. To examine this, we investigated whether the participants responses for the whole response sequence (crucially after correcting for misbinding or repairing the sequence) displayed a bias to the left or right. Biases in the horizontal plane were examined in a non-parametric ANOVA with set size (1,2,3,4,5,6) as a within-subject factor and disease group (PD or control) as a between-subject factor. Whether responses veered to the left or right was significantly affected by set size ($F(5,180) = 10.81, p < .00001$), with rightward responses most occurring at set size 4, in contrast to the leftward responses at higher set sizes. There was no significant effect of group ($F < 1$), and a non-significant trend towards a significant group by set size interaction ($F(5,180) = 1.99, p = .08$).

Similarly, biases in the responding in the vertical plane (after correcting for misbinding using the repaired sequences), were examined in a non-parametric ANOVA with set size (1,2,3,4,5,6) as a within-subject factor and disease group (control or PD) as a between-subject factor. Biases in the up/down direction did significantly differ according to set size ($F(5,180) = 12.24, p < .0001$), mainly due to biases for responses to move vertically down at higher set sizes (Figure S1). However, there was no significant effect of group ($F < 1$) or significant interaction between group and set size ($F(5,180) = 1.13, p = .34$). In summary, there was no evidence for systematic responses biases between patients and controls in this task.
Finally, we examined whether participants were responding randomly by using the relative average Euclidean distance between actual responses to target locations compared to the Euclidean distance between randomly chosen spatial locations and target locations. A non-parametric linear model was used to examine performance on this metric according to disease and set size. There was evidence for an interaction between disease and set size ($F(4,144) = 3.02, p = .019$). This was due to PD patients tending to show greater evidence random responding for set size 3 than other set sizes whereas the opposite tended to happen for other set sizes. Cumulatively, these supplementary analyses revealed that PD patients’ impaired spatial WM at higher set sizes was not due to impaired guesses or response biases, but there was evidence for a set-size-dependent increase in misbinding.

**DISCUSSION**

This study sought to define the circumstances in which medicated PD patients display impaired WM. The results reveal that patients’ deficits vary according to task and set size. When reproducing remembered orientations, patients did not have significantly reduced precision, but, importantly, did show evidence for increased guessing, irrespective of the short delays (1000 ms or 2000 ms) used in this task (Figures 2 and 3). Moreover, patients showed higher (counterclockwise) response biases for longer durations. Whilst there was some evidence that increased guessing also occurred in the analogue Corsi-like task, this was delay-dependent and occurred only for intermediate (three item) loads. Like the one-item orientation task, precision was not robustly impaired in PD patients on the analogue Corsi task at low loads. However, it was at the highest set size (six items; Figure 4).

Utilising the ability to decompose the sources of this error enabled by our design, revealed that recall was corrupted at higher loads: recall of the order of items in memory became corrupted. Notably, though, even after correcting for these errors, patients still showed reduced precision at the highest set size. In line with this, patients were impaired on the more demanding backward variant of the traditional digit span task. These findings suggest that WM impairments in PD patients are highly nuanced and appear in a load- and task-specific manner. However, there was also evidence for cross-cutting deficits. Patients showed a tendency for enhanced guess responses across both experiments, suggesting that attentional lapses can contribute to impaired WM performance in PD patients across domains.

**PD patients showed impaired backward digit span performance**

Meta-analyses have consistently shown that PD patients are impaired in visual and visuospatial WM tasks, with substantially larger effects sizes observed on paradigms containing spatial material or tasks that require more executive control (e.g. manipulation rather than just maintenance; Ramos & Machado, 2021; Siegert et al., 2008). The observations from the present study are consistent with these observations. PD patients were found to be impaired only on the backward and not forward digit span. There is a large corpus of evidence that suggests that maintaining information in WM has different neural substrates compared to the active manipulation of items (Masse et al., 2019). Most directly for the current issues, PD patients and controls have been found to show increased activity in the dorsolateral, ventrolateral prefrontal cortex and striatum during manipulation compared to the maintenance of information (Lewis et al., 2003).

A recent study added to our understanding of the precise physiological mechanisms responsible for manipulation deficits by demonstrating that PD patients showed impaired activation in the subthalamic nucleus (Ye et al., 2021). PD patients not taking their l-dopa medication also display relative difficulties with the manipulation of items of items in WM (Lewis et al., 2005). This would suggest that dopaminergic abnormalities in frontostriatal regions could be driving this impairment, possibly due to altered
dynamics between the direct and the indirect corticostriatal pathways that have been proposed to gate the entry and exit of information in WM (Wiecki & Frank, 2010). Contrary to this, however, a recent study (Grogan et al., 2018) found that withdrawal from dopaminergic medication did not affect patients' backward span performance and that, irrespective of medication, deficits were still observed in PD. This could suggest that pathological processes that are not amenable to dopaminergic medication, perhaps involving the subthalamic nucleus (Ye et al., 2021), could be driving this deficit and should be a target for future treatment strategies.

PD patients show increased error and misbinding during spatial WM

Although results from the forward digit span task suggest that PD patients can show intact WM when no manipulation is required, our novel analogue version of the Corsi spatial span task revealed impaired performance. This suggests that simple recall—at least in the spatial domain—can be compromised in patients if the task is sensitive enough to reveal these deficits. A strength of the task is that it allowed us to explore several putative sources of error. These analyses revealed disturbed performance across multiple parameters in PD patients but also highly nuanced fluctuations consistent with these parameters being sensitive to damage to different neural systems and cognitive variables (Tabi et al., 2021; Zokaei et al., 2020). Earlier studies have found that PD patients can show impaired performance on a computerised, touchscreen spatial span task (Fallon, Bor, et al., 2017). However, that study was unable to specify the underlying neurocognitive mechanisms responsible for this: for example, did patients simply just give up at higher loads and make random responses or did the items in WM just become less precisely represented? Thus, prior to the present study, it was unclear why PD patients show impairments on spatial span tasks. Clinically, this is important if we use data from such tasks to drive specific pharmacological treatments in different patient groups.

Patients showed reduced overall precision for higher set sizes, suggesting that as the number of to-be-recalled items increased the resolution of each of these items decreased (Figure 4a). This deficit was found to occur even after correcting for possible errors in serial reproduction (Figure 4b). Thus, this provides strong evidence that PD patients have a reduced capacity to maintain information over a short period of time. The neural substrates for this deficit remain to be fully articulated. The dopaminergic receptor agonist cabergoline affects the precision of WM (Fallon et al., 2016). However, withdrawing PD patients from their medication does not appear to affect their mnemonic precision (Fallon, Gowell, et al., 2019; Fallon, Mattiesing, et al., 2017). The power of alpha (10 Hz) oscillations, which are thought to index functional inhibition (Jensen & Mazaheri, 2010), is associated with the precision of memory (Myers et al., 2014). Thus, there may be a role for the functional inhibition cortical areas in PD patients that is responsible for driving their impaired recall. Subsequent studies might profitably examine the role of dopamine and oscillatory neural components in WM in PD patients.

Previous investigations that have examined spatial WM in PD found no evidence that patients confuse the spatial location of different items retained in memory (Zokaei et al., 2020). However, these studies have generally only examined this tendency at relatively low set sizes (three items) and, crucially, have not examined the sequential reproduction of items. Here, in the present investigation, PD patients did show evidence for enhanced misbinding (higher Chebyshev distance), which become prominent at higher mnemonic loads. Several lines of evidence implicate hippocampal damage as having a key role in underlying these deficits. For example, lesions to the hippocampus are associated with an increased likelihood of making these errors (Pertzov et al., 2013; Zokaei et al., 2019). However, it is unclear whether the present findings are due to hippocampal disruption in PD patients. Although it is relatively spared in PD relative to other neurodegenerative conditions, there are reports of hippocampal atrophy in PD patients (Camicioli et al., 2003). Moreover, there is growing evidence that dopaminergic disruption in the hippocampus can drive cognitive deficits in PD patients (Calabresi et al., 2013). Thus, by taxing patients' recall at higher load we might have been able to uncover the cognitive sequelae of hippocampal dysfunction in PD patients. However, this is speculative and further work is needed to verify this.
PD patients may show enhanced guessing

Guessing responses occur when there is no, or minimal, evidence for an association between the response item in an analogue report task and the encoded features. As such, they can also be conceptualised as an attentional lapse resulting from an absence of mnemonic representations to guide responding. We evaluated the tendency to make guessing responses in two very different tests. In the simple one-item delayed reproduction task, PD patients showed increased probability of guessing irrespective of delay. This makes it less likely that disease impairs the stability of maintenance processes directly, indicating instead a tendency for more attentional lapses. In contrast, the tendency to guess or make random responses was found to be load-dependent on the analogue spatial span task, with patients showing higher guess rates for intermediate set sizes.

The observation that PD patients show enhanced guessing accords well with other data. Across a variety of WM tasks incorporating memory for orientations (Zokaei, McNeill, et al., 2014) or locations (Zokaei et al., 2020), individuals with PD patients display enhanced guessing responses. There is strong evidence for suspecting that dopaminergic degeneration is responsible for driving an increase in enhanced guessing. Attenuating dopaminergic signal in healthy adults with haloperidol has produces a general increase in guessing response, irrespective of whether information had to be just maintained in WM or whether distractors needs to be filtered out (Fallon, Muhammed, et al., 2019). Whereas, in PD patients, withdrawal of dopamine has been associated with an increased tendency to make guess responses when confronted with distracting information (Fallon, Mattiesing, et al., 2017). Thus, there is evidence that the dopaminergic control of guessing can be both a generalised and task-specific process.

Although, the exact mechanisms through which dopamine could produce this effect are unknown, one hypothesis is that this occurs due to dopamine's ability to modulate endogenous rather than exogenous cognitive control (Fallon, Muhammed, et al., 2019). Support for this hypothesis comes from the fact that haloperidol can increase the connectivity between the dopaminergic midbrain regions and the default mode network (Cole et al., 2012), which is canonically associated with engaging in off-task processes (Raichle et al., 2001). Reinstatement of normal default mode deactivation has been associated with levodopa administration in PD patients (Delaveau et al., 2010). Thus, there is not a simple model through which increased guess responses could be observed in PD patients. However, future studies need to examine whether cognitive fluctuations in PD patients during cognitive tasks can occur either due to unstable dopaminergic dynamics or through non-dopaminergic pathology, for example acetylcholine (Wesnes et al., 2005). There also needs to be further work to understand how attentional lapses or guess responses vary according to load.

Response biases may influence recall error

A vexed question in research on cognition in PD patients, and other movement disorders, is the extent to which observed impairments are due to deficits or systematic biases in responding. Particularly in circumstances where a response or series of responses is required, patients may over or undershoot the desired location or response. These types of responses can be hard to control and eliminate, even in ideal experimental settings. Indeed, in the context of mentally rotating hand positions, PD patients have been found to show asymmetries in their proficiency with clockwise or counterclockwise rotations (Helmich et al., 2007). Scant attention has been paid to whether such biases exist in tests of WM in PD patients (though see Bek et al., 2022). Here, we were able to monitor, quantify and account for the variety of ways in which these errors could impinge on our performance metrics and ascertain whether there were group differences. Although ostensibly the analogue Corsi-like task contains the most opportunities for impaired motor control to affect performance, there was no evidence for systematic biases in patients responded in the horizontal or vertical plane. In contrast, response biases were seen in the simple one-item memory task. Crucially, however, this bias was delay-dependent. This illustrates that it was not the motoric requirements of the task, but the cognitive requirements induced by the delay that were responsible for generating this difference.
In summary, these data confirm the existence of impaired WM performance in PD patients and the need for further pharmacological augmentation strategies to combat these deficits. Moreover, they suggest that short-term recall deficits in PD patients are heterogenous in origin and the cause of these deficits (parameters responses) are different.

AUTHOR CONTRIBUTIONS
Sean James Fallon: Conceptualization; formal analysis; methodology; resources; writing – original draft.
Chevonne van Rhee: Data curation; validation.
Annika Kienast: Data curation; project administration.
Sanjay G. Manohar: Software; writing – review and editing.
Masud Husain: Project administration; supervision; writing – review and editing.

ACKNOWLEDGEMENTS
This research was supported by a Wellcome Trust Principal Research Fellowship to MH and the Oxford NIHR Biomedical Research Centre.

CONFLICT OF INTEREST
All authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT
Our ethics approval does not permit public archiving of the data supporting this study. Those wishing to access to this data should contact the lead author, SJF. Access can be obtained and granted to named individuals. To obtain the data, investigators must complete a formal data sharing agreement, including conditions for secure storage of sensitive data.

ORCID
Sean James Fallon https://orcid.org/0000-0002-1538-724X

REFERENCES


**SUPPORTING INFORMATION**

Additional supporting information can be found online in the Supporting Information section at the end of this article.