Experimental Brain Research (2019) 237:2939–2956
https://doi.org/10.1007/s00221-019-05623-x

RESEARCH ARTICLE

Working memory in action: inspecting the systematic and unsystematic errors of spatial memory across saccades

Adam Frost[1] · George Tomou[1,2,3] · Harsh Parikh[1] · Jagjot Kaur[1] · Marija Zivcevska[1] · Matthias Niemeier[1,3]

Received: 5 June 2019 / Accepted: 6 August 2019 / Published online: 10 September 2019

Abstract
Our ability to interact with the world depends on memory buffers that flexibly store and process information for short periods of time. Current working memory research, however, mainly uses tasks that avoid eye movements, whereas in daily life we need to remember information across saccades. Because saccades disrupt perception and attention, the brain might use special transsaccadic memory systems. Therefore, to compare working memory systems between and across saccades, the current study devised transsaccadic memory tasks that evaluated the influence of memory load on several kinds of systematic and unsystematic spatial errors, and tested whether these measures predicted performance in more established working memory paradigms. Experiment 1 used a line intersection task that had people integrate lines shown before and after saccades, and it administered a 2-back task. Experiments 2 and 3 asked people to point at one of several locations within a memory array flashed before an eye movement, and we tested change detection and 2-back performance. We found that unsystematic transsaccadic errors increased with memory load and were correlated with 2-back performance. Systematic errors produced similar results, although effects varied as a function of the geometric layout of the memory arrays. Surprisingly, transsaccadic errors did not predict change detection performance despite the latter being a widely accepted measure of working memory capacity. Our results suggest that working memory systems between and across saccades share, in part, similar neural resources. Nevertheless, our data highlight the importance of investigating working memory across saccades.

Keywords Transsaccadic integration · Transsaccadic memory · Working memory · Change detection · *n*-Back · Remapping

Introduction

We make three or four saccadic eye movements per second to take highly resolved foveal snapshots of the world, and we use transsaccadic integration (Niemeier et al. 2003, 2007; Deubel et al. 1998) and remapping (Henriques et al. 1998; Medendorp et al. 2003) to combine information across saccades. Yet, the foveal information quickly fades and cannot be joined into rich, 'photographically' detailed images of our visual surround—contrary to what subjective impressions might suggest. For example, we are surprisingly poor at spotting dramatic changes to scenes if the changes coincide with saccades or other brief visual interruptions and transients (e.g. Grimes 1996; Phillips 1974; Simons and Levin 1998; Simons and Rensink 2005). That is, we retain only a small fraction of the available visual information across eye movements.

Our limited ability to retain transsaccadic information resembles visual working memory, with its ability to store information for about four items of a given set over a span of seconds (e.g. Luck and Vogel 1997). Trying to remember more information places increased demands on our working memory's limited resources such that whole items are dropped from memory (Luck and Vogel 1997), or their representation becomes gradually less precise (Bays et al. 2009).

Transsaccadic visual working memory is now known to have limitations similar to those observed in visual working memory (e.g. Irwin 1991; Prime et al. 2007). Thus, it seems reasonable to assume that transsaccadic memory is the same as visual working memory.

✉ Adam Frost
adam.frost@mail.utoronto.ca

✉ Matthias Niemeier
niemeier@utsc.utoronto.ca

1 University of Toronto Scarborough, Toronto, Canada

2 York University, Toronto, Canada

3 Centre for Vision Research, Toronto, Canada

However, the similarities could be coincidental. Transsaccadic memory (TSM) might use neural resources that are rather separate from those of visual working memory—especially because TSM faces challenges that are rather different from those memory functions that are more commonly probed in visual working memory paradigms.

One obvious difference is that in usual visual working memory tests, participants keep their eyes steady at a fixation point, whereas transsaccadic tasks require the eyes to move. So, the intervening planning and execution of eye movements could affect stored information. When lateral eye movements are made, information in spatial memory is translated according to its retinotopic coordinates, and items crossing the visual midline are transferred across cortical hemispheres (Merriam et al. 2003, 2007). Vuilleumier et al. (2007) reported that patients suffering from left hemispatial neglect showed deficits in memory performance when items were remapped into or within the impaired left hemifield, even when the items were remapped back into the right hemifield before recall. Transcranial magnetic stimulation of posterior parietal cortex (Prime et al. 2008) or frontal eye fields (Prime et al. 2010) is disproportionately disruptive to visual working memory when a saccade is made before recall, likely reflecting that functions related to spatial remapping are disrupted. Further, remapping in general can incur significant costs to visual working memory performance relative to a no-saccade baseline (Vasquez and Danckert 2008; Brink et al. 2019).

The saccade-related drop in recall could reflect drops in accuracy or precision. In other words, they could reflect systematic or unsystematic memory errors resulting from representational noise that accumulates as memory buffers compete with attentional and oculomotor processes, or errors could come with an imperfect remapping process. Further, saccades could interfere more with greater memory loads.

In the present study, we addressed two central questions: first, how is the integrity of spatial information held in TSM affected when larger compared to smaller set sizes are consuming more or less memory resources? That is, do participants commit more unsystematic errors in their responses, or are there systematic effects that change the topography of TSM representations? Second, to what degree does TSM share underpinnings with visual working memory as captured in classic working memory paradigms?

Experiment 1

Participants

Twenty-six healthy undergraduate students (19 females with a median age of 18 years) gave their informed and written consent to participate in the experiment. All participants had normal or corrected-to-normal vision. All procedures were approved by the Human Participants Review Sub-Committee of the University of Toronto and were performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki.

Apparatus

Participants performed the experiments in a dark room with their heads stabilized by a chin rest. Stimuli were presented on an LCD monitor centered at eye level that subtended 57.5° of visual angle (115 cm wide at an eye–screen distance of 107 cm) placed eye level. Stimuli were generated and presented using Matlab (MathWorks) and the Psychophysics and EyeLink Toolbox extensions (Brainard 1997; Pelli 1997; our experimental code for this and the other two experiments is available upon request). One task in this experiment required gaze tracking, and for it participants were fitted with a head-mounted eye tracker which sampled the position of the left eye at 500 Hz (EyeLink II, SR Research, Ottawa).

Procedure

Transsaccadic memory task

The TSM task constitutes an extension of a previous transsaccadic integration task (Prime et al. 2006). Trials of the task began with fixation on a point (0.15° across) located at the vertical screen centre and 10° to the left or right of screen centre. Participants fixated this point and then clicked a computer mouse. 900–1100 ms later, a memory display of either one or two coloured lines (any pair of red, green or blue, 1 pixel across) was flashed on screen for 400 ms. The lines always continued through to the edges of the display. They were positioned and angled so that they intersected both the fixation point and one of nine points in a virtual three-by-three square grid (10° wide) centred horizontally between the initial fixation point and the saccade target, and vertically 8° above the fixation point (see Fig. 1). Immediately after the offset of the memory display, a mask showed 40 randomly distributed and oriented lines of the same widths and colours as those in the memory array for 200 ms. Five hundred milliseconds after the offset of the mask, a new fixation point appeared 20° to the opposite side of the screen relative to the first fixation point and participants were asked to move their eyes to it. After 900–1100 ms, a new angled line appeared that matched the colour of one of the previously presented lines. Crucially, the postsaccadic line was placed so that it crossed the postsaccadic fixation point as well as the same intersection point in the virtual three-by-three square grid that had been crossed earlier by the initial memory line of the same colour. Four hundred milliseconds later, a

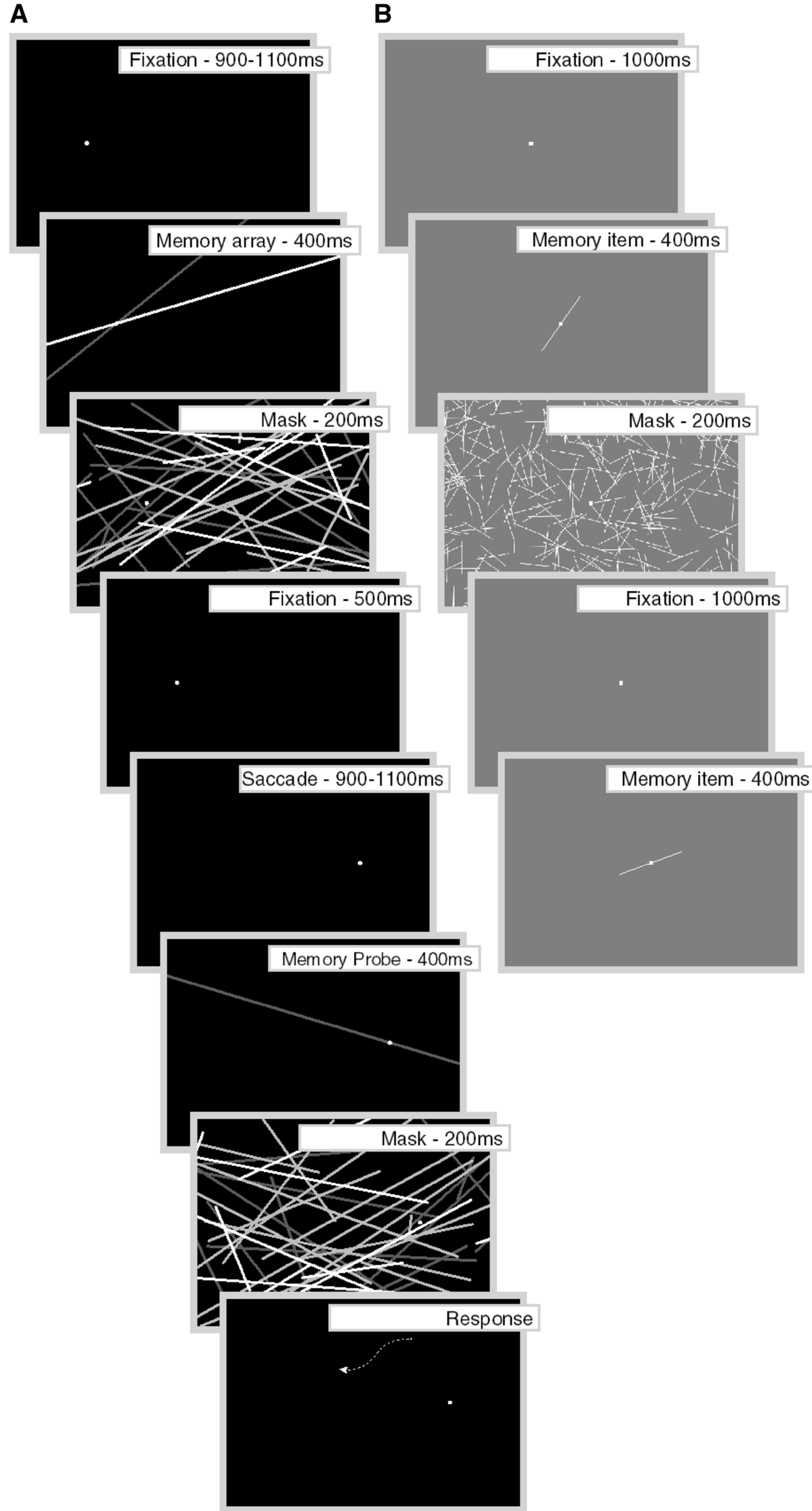

Fig. 1 Experimental paradigm. **a** TSM task. **b** 2-Back line orientation task

second mask (same parameters as the first mask) followed, after which a cursor was placed at the fixation point. Participants used the cursor to indicate the position where the recall line would have intersected the colour-matched memory line, had they appeared on the screen at the same time. A mouse click at the estimated intersection location concluded the trial. Blocks comprised 126 trials, and the task comprised three blocks for a total of 378.

Data from all blocks were screened for eye movement errors (eye movements during stimulus presentation, refixations further than 2° away from the target and/or trials with blinks—an average of 79%, SD 16%, of trials per participant were included in the final analysis). The remaining data were used to calculate systematic and unsystematic transsaccadic errors as reflected in the mouse clicks. To this end, we collected the horizontal and vertical coordinates of all clicks for each working memory load level and saccade direction separately. As shown in Fig. 2, this created nine clusters of mouse clicks, each pertaining to one of the nine line intersection locations. From the clusters we first removed outliers. To identify outliers, for each participant we calculated a 'global' standard deviation of the distance of each of their responses from its respective intersection location. Then, each response that deviated by more than 1.96 'global' standard deviations from its respective cluster was labelled as an outlier and removed from subsequent analyses (0.49% of responses on average). Note that it is possible that some of the outliers reflected working memory errors. For example, participants might have forgotten the true intersection location and simply guessed. Other outliers might have been due to confusion errors where participants aimed for an intersection location other than the correct one. However, a third kind of outliers was probably due to simple motor errors where participants pressed the mouse key before moving the mouse to the intended location. With the current paradigm, we were unable to tease apart memory and motor errors.

Next, we recalculated the horizontal and vertical standard deviations as measures of unsystematic error. To calculate systematic errors we recalculated the average coordinates of the nine clusters of mouse clicks and submitted the data to a Procrustes analysis to obtain quantitative measures of all linear transformations required to overlay the cluster locations with the three-by-three grid of the original line intersections. That is, we obtained measures of how much the nine clusters were systematically shifted horizontally and vertically, how much they were rotated, and how much they differed in scale relative to the original grid. As a fifth output, Procrustes analysis gave a measure of nonlinear distortion.

2-Back line orientation task

Participants' working memory ability was assessed using a 2-back task. Here, participants were presented with a series of white lines (3° in length, 0.06° wide) on a grey background, and they compared the angle line on screen with the angle of the line presented prior to the previously presented one (see Fig. 1). Participants fixated a central fixation point (0.2° across) for the entirety of the task, and the stimulus lines were each presented, extending from the fixation point outwards, for 400 ms. The angles of the lines were one of four possibilities, either 55°, 90°, 125°, or 160°. The probability of a match trial was 1/6. The angled line was followed immediately by a mask comprising 1000 lines placed at random across the screen, which was displayed for 200 ms, and followed by a blank screen and central fixation point. One second later, the next trial began. If the participant believed that the present line was a match, they responded with a key press on a computer keyboard. Blocks comprised 62 trials, and participants completed three blocks. Hit rates and false alarm rates were used to calculate d' values as measures of sensitivity free of any response bias.

Results and discussion

Group averages of systematic and unsystematic errors are visualized in Fig. 2 for each memory load condition and

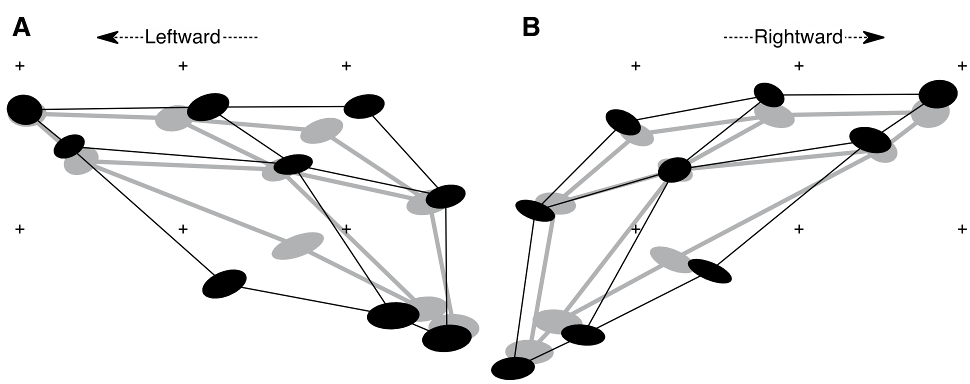

Fig. 2 Systematic and unsystematic errors. **a** Errors after leftward saccades. **b** Errors after rightward saccades. As a visual analogue for standard error, ellipses were fitted to group-level responses (mouse clicks) for each of the nine possible intersection locations, with radii shrunk by a factor of ten for graphical clarity. Ellipse centres represent the mean response for a given intersection location. Black: low memory load condition, grey: high memory load condition. Data are superimposed onto the true intersection locations (3×3 grid of plus signs)

saccade direction separately. The figure shows that participant responses were distorted in a centrifugal manner away from the saccade target. Also, responses in the low load condition covered a larger area than those of the high load condition. Finally, close inspection suggested that ellipses in the high load condition were slightly larger than in the low load condition. To quantify systematic errors we converted data, we flipped the leftward saccade data horizontally to register coordinates relative to the saccade, and submitted them to Procrustes analysis (see "Procedure"). Unsystematic errors were calculated using standard deviations (Fig. 3).

To test which of these measures of TSM performance reflect a working memory-like sensitivity to load, we conducted a series of seven 2-way repeated measures ANOVAs with memory load (one or two lines) as a first factor. As a second factor we included saccade direction (left vs. right) (Table 1). We found that an increase in memory load resulted in decreased performance with respect to unsystematic error (both horizontal and vertical), Procrustes *D* and scaling ($ps < 0.005$).Although the centrifugal distortions resulted in large values of horizontal and vertical offset opposite to the saccade as well as large Procrustes rotation (clockwise and counterclockwise after rightward and leftward saccades, respectively), these measures were not significantly influenced by load ($ps > 0.1$). Also, the tests showed no significant effect of saccade direction on any of the dependent variables, in terms of its main effects ($ps > 0.05$) or its interactions with load ($ps > 0.05$).

Subsequently, to test whether the TSM task captures performance features comparable to those reflected in classic

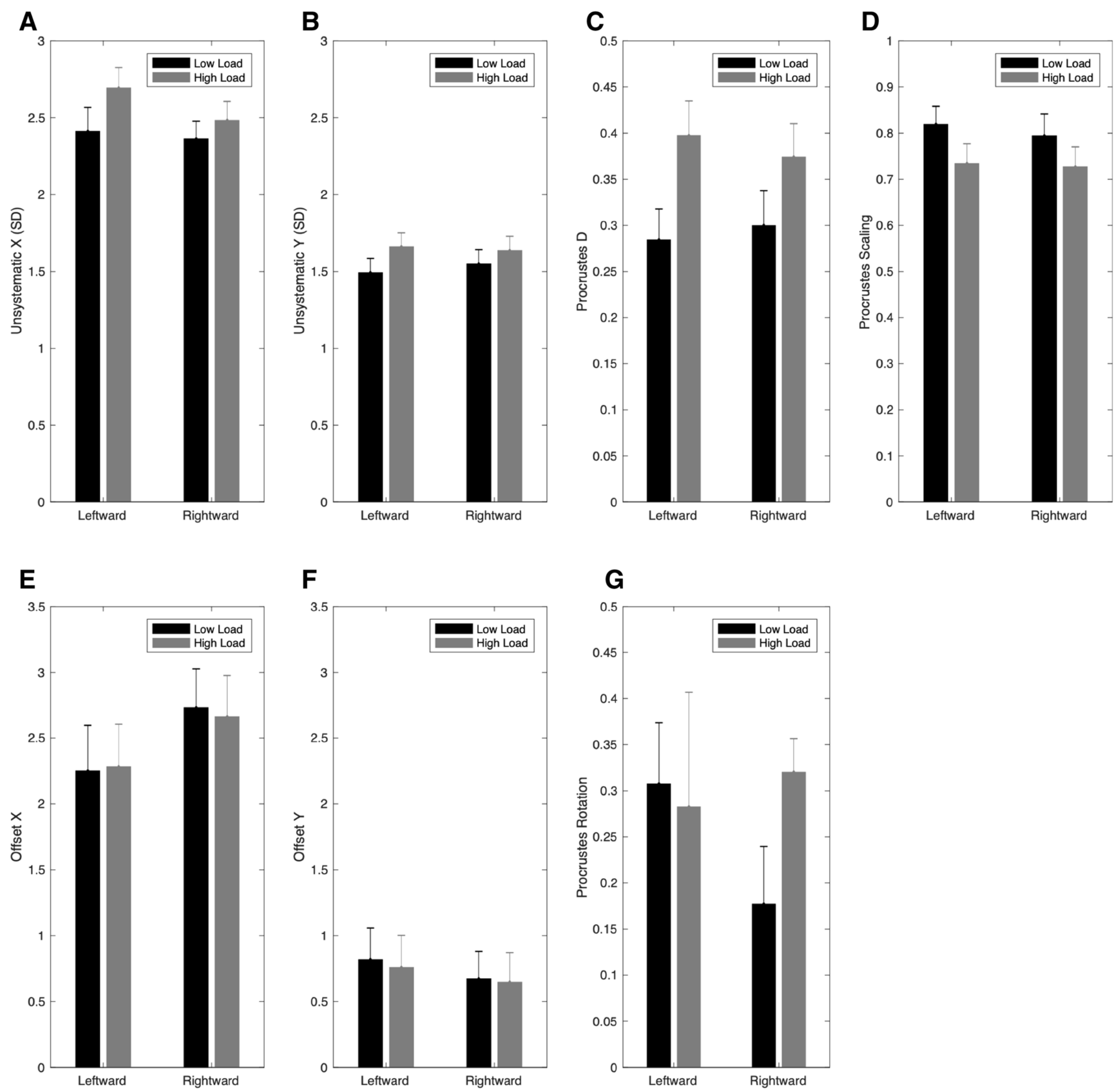

Fig. 3 Results of Experiment 1: TSM performance measures for the two saccade directions and memory loads separately. **a**, **b** Horizontal and vertical unsystematic error. **c**–**g** Measures of systematic error

Table 1 Two-way repeated measures ANOVAs for the measures of transsaccadic error

Error	Load			Saccade			Load × saccade		
	F(1, 25)	p	η_p^2	F(1, 25)	p	η_p^2	F(1,25)	p	η_p^2
Unsystematic X	10.20*	0.004	0.29	4.17	0.052	0.14	1.47	0.237	0.06
Unsystematic Y	9.98*	0.004	0.29	0.22	0.642	0.01	1.73	0.200	0.07
D	31.90***	< 0.001	0.56	0.06	0.807	0.00	3.13	0.089	0.11
Scale	19.60***	< 0.001	0.44	0.65	0.427	0.03	0.31	0.584	0.01
Offset X	0.03	0.853	0.00	1.83	0.188	0.07	0.58	0.452	0.02
Offset Y	0.57	0.457	0.02	1.73	0.200	0.07	0.22	0.647	0.01
Rotation	0.94	0.341	0.04	4.09	0.054	0.14	0.54	0.470	0.02

Unsystematic error comprises the mean horizontal and vertical standard deviation. Procrustes D, scaling, offset and rotation are all derived from Procrustes analysis (without reflection). Greenhouse–Geisser corrections produced identical values for all ANOVAS

Asterisks denote significance after serial Bonferroni corrections for multiple comparisons. *p <0.05, **p <0.01, ***p <0.001

working memory tasks without eye movements, we conducted a multiple regression analysis with stepwise inclusion of the TSM error variables (averaged across both saccade directions and loads) to predict participants' performance in the 2-back task (average $d' = 1.13$, SD 0.81). The resultant model included only horizontal unsystematic error as a significant predictor for performance in the 2-back task (greater unsystematic error associated with poorer 2-back performance, $\beta = -0.58$, $F(1, 24) = 12.11$, $p = 0.002$), with an R^2 of 0.335. In addition, we calculated individual correlations that revealed that the TSM error measures that were significantly affected by load also significantly predicted performance on the 2-back task (two-tailed bootstrapped 95% confidence intervals did not include zero) in the transsaccadic task (see Table 2, Fig. 4).

In sum, Experiment 1 indicates that greater memory load increased unsystematic errors as well as contracted and nonlinearly distorted topography in TSM. Further, the same measures, except for vertical unsystematic errors, correlated with the 2-back test, suggesting that TSM relies on similar cognitive resources as spatial working memory. However, these results come with several limitations. First, the TSM task as used here requires transsaccadic *integration* of one presaccadic and one postsaccadic line, and, so, it probably measures more than purely transsaccadic memory functions. Second, the influence of transsaccadic memory (and integration) could have been lessened if participants had pursued a certain "line search strategy" to respond; they could have used the postsaccadic fixation point as a reference and then travelled with the mouse cursor along the remembered postsaccadic line, more or less so depending on how steep the presaccadic line had been. Indeed, we noticed that, except for the top row of responses, systematic errors tended to be biased approximately along an imaginary line extending from the postsaccadic fixation point beyond the correct intersection location (Prime et al. 2006 for a similar observation). In other words, it is difficult to judge the extent to which participants memorized categorical information rather than spatial information across saccades. Furthermore, the "line searches" likely skewed systematic errors. Third, the 2-back task, although commonly used as a measure

Table 2 Correlations between 2-back performance and measures of transsaccadic error

Error	r	p	CI (95%)		CI corrected	
			Lower	Upper	Lower	Upper
Unsystematic X	−0.58[a]	0.002	−0.84	−0.23	−0.89	−0.07
Unsystematic Y	−0.39	0.051	−0.66	−0.10	−0.73	0.02
D	−0.51[a]	0.008	−0.73	−0.24	−0.80	−0.11
Scaling	0.48[a]	0.014	0.20	0.70	0.08	0.76
Offset X	0.07	0.719	−0.26	0.38		
Offset Y	0.04	0.828	−0.29	0.40		
Rotation	−0.32	0.112	−0.71	0.05		

Median r and 95% confidence intervals derived from bootstrapped regression analysis (100,000 resampling iterations)

[a]Zero falls outside of the Bonferroni-corrected 95% confidence interval

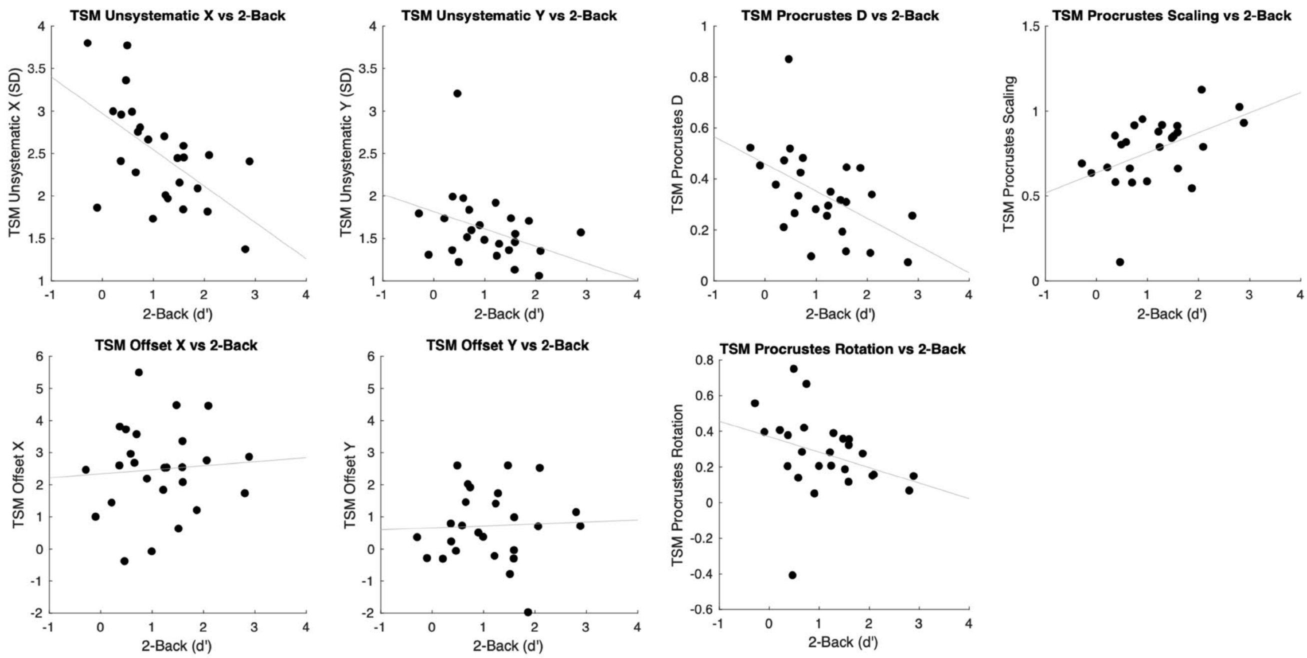

Fig. 4 Results of Experiment 1: correlations between 2-back performance and measures of transsaccadic error

of working memory, has been found to correlate poorly with some other measures of working memory (Kane et al. 2007; Jaeggi et al. 2010) and appears to represent a less common facet of working memory-related executive function. Therefore, we devised a second experiment in which we simplified the TSM task and also replaced the 2-back task with a change detection paradigm that is well accepted as a measure of working memory capacity (Vogel et al. 2005).

Experiment 2

Participants

Twenty-two healthy undergraduate students (13 females with a median age of 19 years) gave their written and informed consent to participate in the experiment. All participants had normal or corrected-to-normal vision. All procedures were approved by the Human Participants Review Sub-Committee of the University of Toronto and were performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki.

Apparatus

Participants were tested using the same setup described in Experiment 1.

Procedure

Transsaccadic memory task

This task (Fig. 5) was structurally similar to that in Experiment 1 (Fig. 1) with fixation point and saccade target set apart by 20° horizontally and the same virtual three-by-three grid of possible memory locations positioned between and 8° above them. However, all stimuli were shifted down by 5° to leave more space for responses in the upper part of the screen, and we jittered positions horizontally so that from trial to trial all stimuli might appear 5.75° further to the left or right, or anything in between. Also, gaze position was processed online. During a trial, if the participant's gaze deviated by 2° from the initial fixation point or 3° from the target fixation point after saccade, the participant was alerted with a beep tone and blue flash, and the trial was recycled. Crucially, to avoid any "line searches" the presaccadic memory arrays showed one or three squares (red, green and blue, 0.6° across) that marked positions randomly selected from the three-by-three grid. After the saccade, one of the squares would reappear at a random position within the three-by-three grid, and participants used the computer mouse to try to restore its correct location. In addition, we used masks that contained 1500 randomly distributed squares of the same dimensions and colours as those in the memory array. Blocks comprised 112 trials, and there were three blocks in total. We extracted the same systematic and unsystematic errors as dependent variables as in Experiment 1.

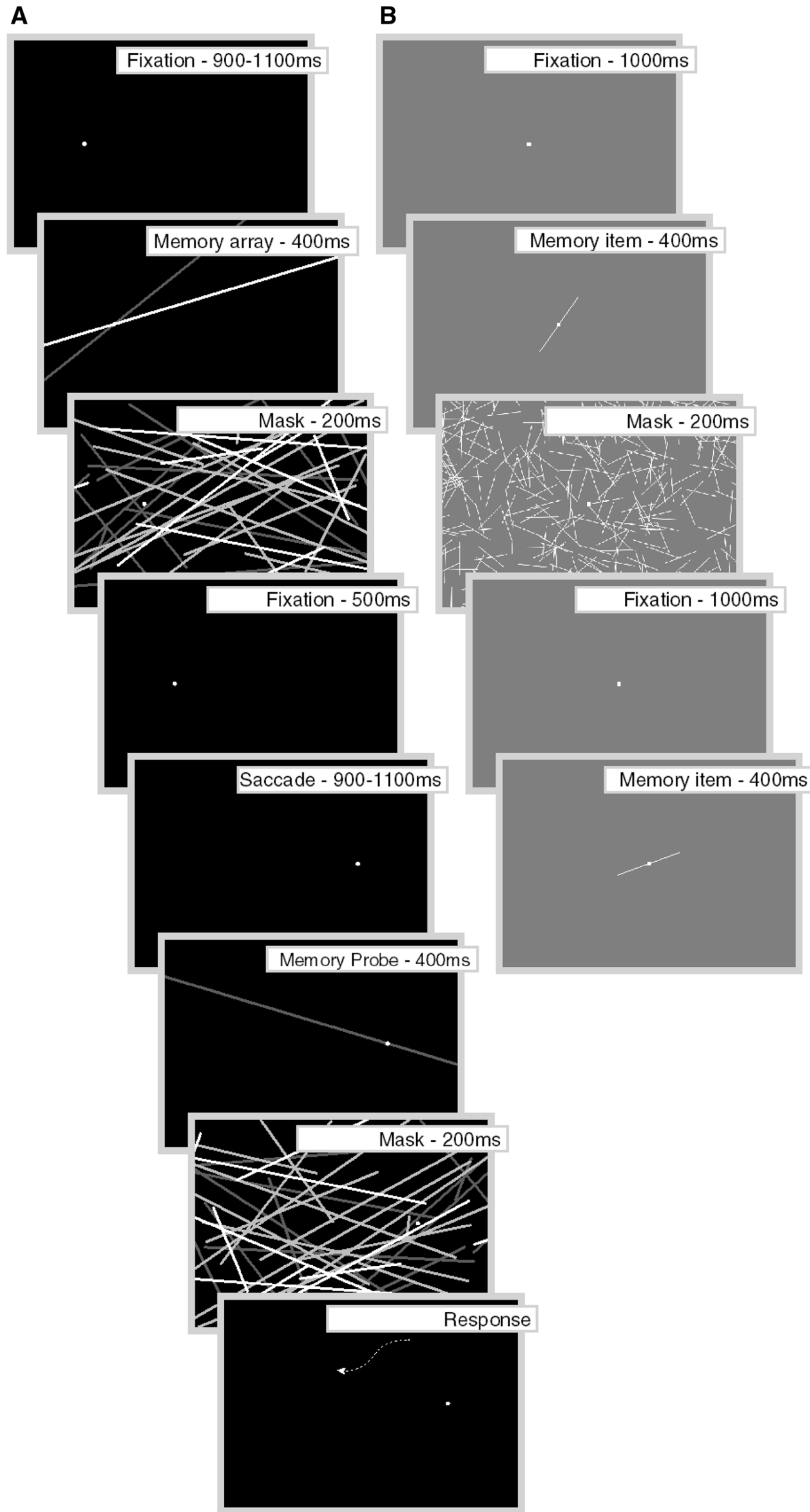

Fig. 5 Experimental paradigm. **a** Transsaccadic memory task. **b** Change detection task

Change detection task

We adopted the paradigm from Vogel and colleagues (Vogel et al. 2005; Fig. 5). Participants fixated centrally on a white fixation point (0.2° across) on a grey background, and a set of two, three or four uniquely coloured (red, orange, green, blue, purple) squares (0.65° in width) was presented either to the left or to the right of fixation. The squares were assigned random locations within a rectangular area (4° wide, 7.3° high), centred 5° away from fixation on the horizontal axis, with a minimum spacing of 1.7° between squares. The memory array remained on screen for 100 ms, followed by a blank display (save for the fixation point) for 900 ms, after which the squares reappeared at their original spatial locations. However, on 50% of trials, one square's colour was changed. Participant used the "d" and "s" key on a computer keyboard to indicate whether they believed the colours to be different or the same, respectively. The recall array remained on screen for 1.5 s or until a response was given, and then the next trial began after a 1.5 s inter-trial interval. Participants completed two blocks of 150 trials (one trial was lost due to a programming glitch). Based on the data we calculated d' values (data from set sizes 3 and 4 merged-together), just like for the 2-back task in Experiment 1.

Results and discussion

Group averages of systematic and unsystematic errors for each memory load condition and saccade direction separately are visualized in Fig. 6. This shows that participant responses were much less distorted than in Experiment 1, consistent with our intention to dissuade participants from using a line search strategy during their responses. Low load data formed a slightly skewed parallelogram. Low load data formed a smaller trapezoid that was slightly rotated clockwise after leftward saccades and counterclockwise for rightward saccades. Also, ellipses were noticeably larger in the high load condition than in the low load condition (Fig. 6).

Just like in Experiment 1, we converted the data relative to saccade direction and tested the influence of memory load and saccade direction on the systematic and unsystematic errors in a series of seven 2-way repeated measures ANOVAs (Fig. 7, Table 3). Again, we found that higher load resulted in greater unsystematic error. Some measures of systematic error were affected as well: scaling again declined with higher load ($p < 0.05$). Rotation now showed a significant load-dependent effect as well ($p < 0.001$). As with Experiment 1, there was no significant effect of load on systematic offset (horizontal and vertical) ($ps > 0.5$), but unlike before, Procrustes D now also showed no effect of load. Also consistent with Experiment 1, the tests produced no significant effect of saccade direction on any of the dependent variables, in terms of its main effects ($ps > 0.05$) or its interactions with load ($ps > 0.05$).

Next, we conducted a multiple regression analysis with stepwise inclusion of the TSM error variables (averaged across both saccade directions and loads) to predict participant performance in the change detection task (mean $d' = 2.72$, SD 0.71). However, none of the TSM error measures were significant predictors. Of the individual correlations (Table 4), only vertical translation showed a trend but failed to reach significance when corrected for multiple comparisons. In addition, it is of note that the bootstrapped confidence intervals were often very large, with a span width of r values exceeding 1 in four out of the seven regressions. This could suggest that a small number of influential data points concealed true relationships present within the majority of the data. To rule out this possibility, we computed robust regressions (Maronna et al. 2006), which reweight data points such that individual cases do not have disproportionate influence over the regression model. Nevertheless, the resulting correlations were equivalent to or smaller than

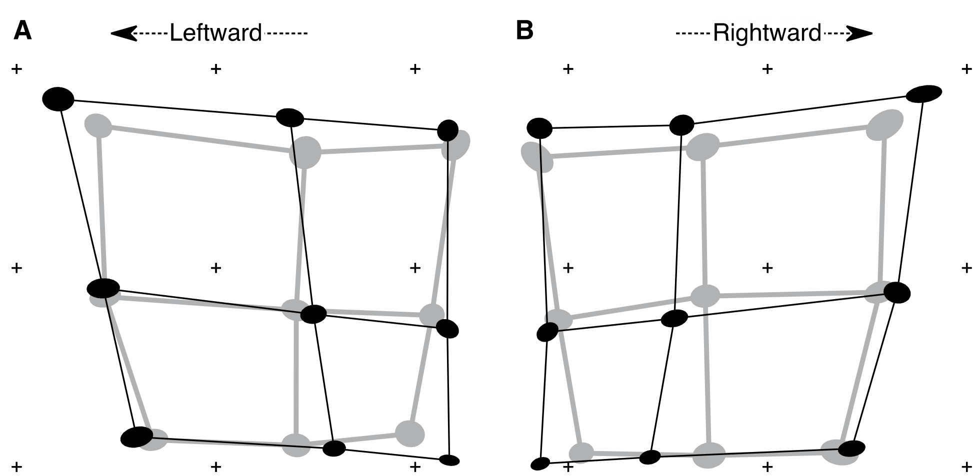

Fig. 6 Systematic and unsystematic errors. **a** Errors after leftward saccades. **b** Errors after rightward saccades. As a visual analogue for standard error, ellipses were fitted to group-level responses (mouse clicks) for each of the nine possible memory locations, with radii shrunk by a factor of ten for graphical clarity. Ellipse centres represent the mean response for a given intersection location. Black: low memory load condition, grey: high memory load condition. Data are superimposed onto the true intersection locations (3×3 grid of plus signs)

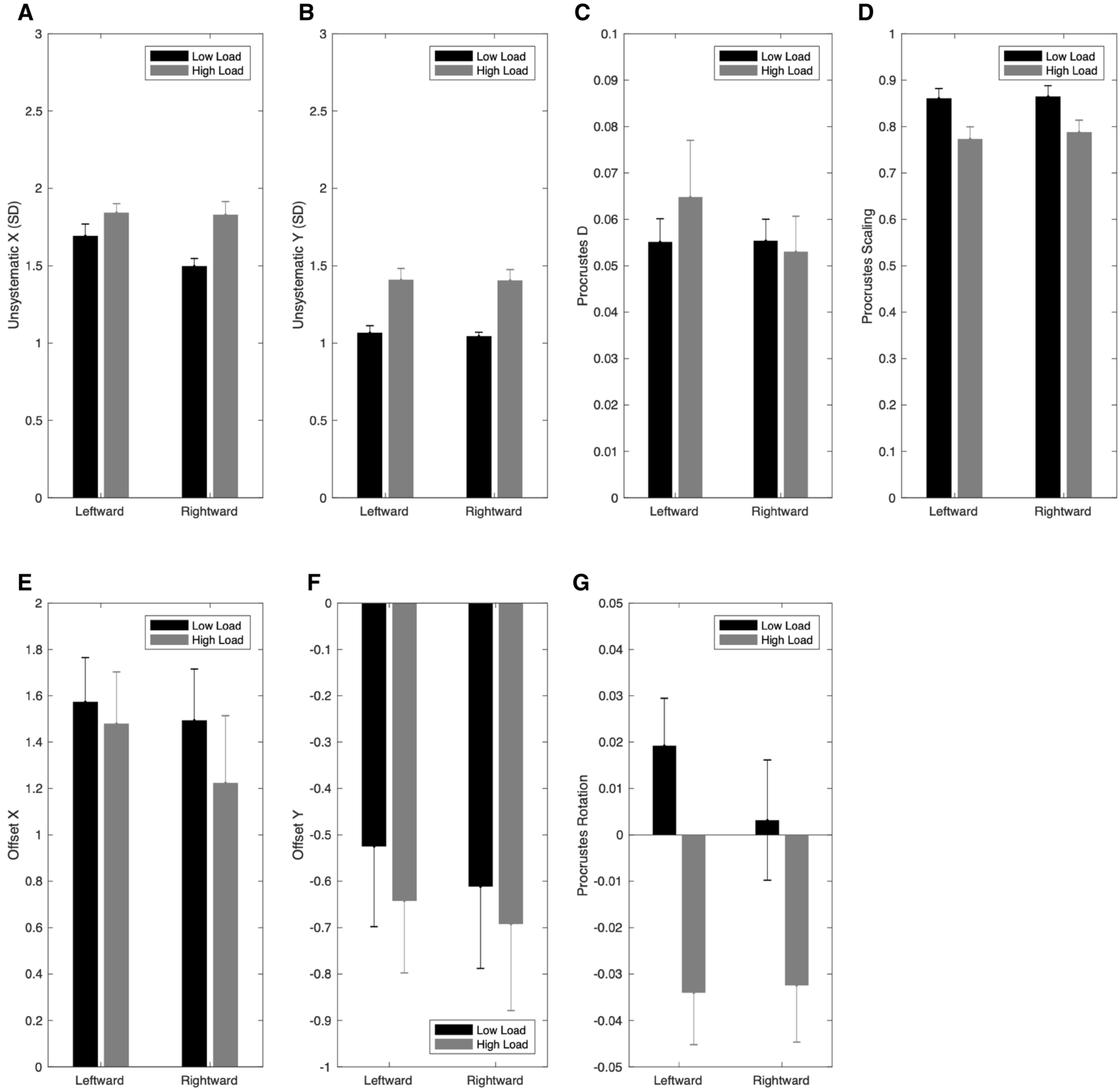

Fig. 7 Results of Experiment 2: TSM performance measures. Bar colour indicates load level (black for low load, grey for high), and grouping indicates the saccade direction. Vertical axis labels indicate the variable displayed. **a**, **b** Horizontal and vertical unsystematic error. **c**–**g** Measures of systematic error

those that resulted from the initial regression tests. Finally, we found that converting change detection performance into measures of Kowan's k (mean $k = 2.75$ SD 0.47) produced no correlations either.

In sum, once again the TSM task showed load effects on unsystematic errors and on scaling. Unlike in Experiment 1, here we also observed a load effect on rotation but none on Procrustes D, likely due to differences in the format of the two transsaccadic tests. Further, the correlational analyses of the current experiment suggest that TSM bears little functional relevance to working memory as measured in the change detection task. This is surprising given that change detection is a well-researched and robust measure of visual working memory (REFs). However, in this experiment we also used a novel TSM task, so as to confirm that the new TSM task truly shows no correlation with change detection and, further, to test whether the new task correlates with the 2-back task, in Experiment 3 we

Table 3 Two-way repeated measures ANOVAs for the measures of transsaccadic error

Error	Load			Saccade			Load × saccade		
	$F(1, 21)$	p	η_p^2	$F(1, 21)$	p	η_p^2	$F(1, 21)$	p	η_p^2
Unsystematic *X*	21.02***	< 0.001	0.50	5.91*	0.024	0.22	2.97	0.099	0.12
Unsystematic *Y*	41.90***	< 0.001	0.67	0.24	0.627	0.01	0.07	0.793	0.00
D	0.14	0.716	0.01	2.24	0.150	0.10	1.59	0.221	0.07
Scale	8.77*	0.007	0.30	0.81	0.380	0.04	0.55	0.467	0.03
Offset *X*	4.87	0.039	0.19	0.48	0.496	0.02	1.00	0.330	0.05
Offset *Y*	1.15	0.296	0.05	0.75	0.397	0.03	0.21	0.650	0.01
Rotation	45.78***	< 0.001	0.69	0.29	0.596	0.01	0.46	0.505	0.02

Unsystematic error comprises the mean horizontal and vertical standard deviation. Procrustes *D*, scaling, offset and rotation are all derived from Procrustes analysis (without reflection). Greenhouse–Geisser corrections produced identical values for all ANOVAS

Asterisks denote significance after serial Bonferroni corrections for multiple comparisons. * $p<0.05$, ** $p<0.01$, *** $p<0.001$

Table 4 Correlations between change detection performance and measures of transsaccadic error

Error	*r*	*p*	CI (95%)		*r* robust	*p* robust
			Lower	Upper		
Unsystematic *X*	− 0.23	0.297	− 0.71	0.51	0.14	0.648
Unsystematic *Y*	− 0.19	0.389	− 0.66	0.45	0.09	0.886
D	− 0.23	0.297	− 0.66	0.52	0.25	0.244
Scaling	0.14	0.522	− 0.40	0.62	0.16	0.597
Offset *X*	0.08	0.717	− 0.33	0.45	0.07	0.730
Offset *Y*	− 0.36	0.104	− 0.68	− 0.01	− 0.37	0.070
Rotation	− 0.22	0.325	− 0.47	0.03	− 0.22	0.109

95% confidence intervals derived from bootstrapped regression analysis (100,000 resampling iterations). Robust r and p were computed using robust linear regression analysis (Maronna et al. 2006)

tested both 2-back and change detection in conjunction with TSM (Fig. 8).

Experiment 3

Participants

Twenty-four healthy undergraduate students (15 females with a median age of 21 years) gave their informed and written consent to participate in the experiment. All participants had normal or corrected-to-normal vision. All procedures were approved by the Human Participants Review Sub-Committee of the University of Toronto and were performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki.

Apparatus

Participants were tested using the setup as described in Experiment 1.

Procedure

Transsaccadic memory task

The TSM task for this experiment (Fig. 9) was quite similar to that of Experiment 2 (Fig. 5). However, we re-designed its display to probe memory in both the upper and lower visual field. Moreover, with the new design participants could no longer anticipate the direction of the saccade and perhaps bias their attention accordingly. To remove all directional cues, we always showed the fixation point in the screen centre (trial by trial with ± 6.9° horizontal jitter) and the saccade target 15° to the left or right of it with equal probability. Also, we abandoned the virtual three-by-three grid of possible positions. Instead, we used a virtual octagon (10° diameter) that was always centred on the fixation point. That is, the (one or three) squares of the memory array would appear on positions randomly sampled from the octagon's eight vertices.

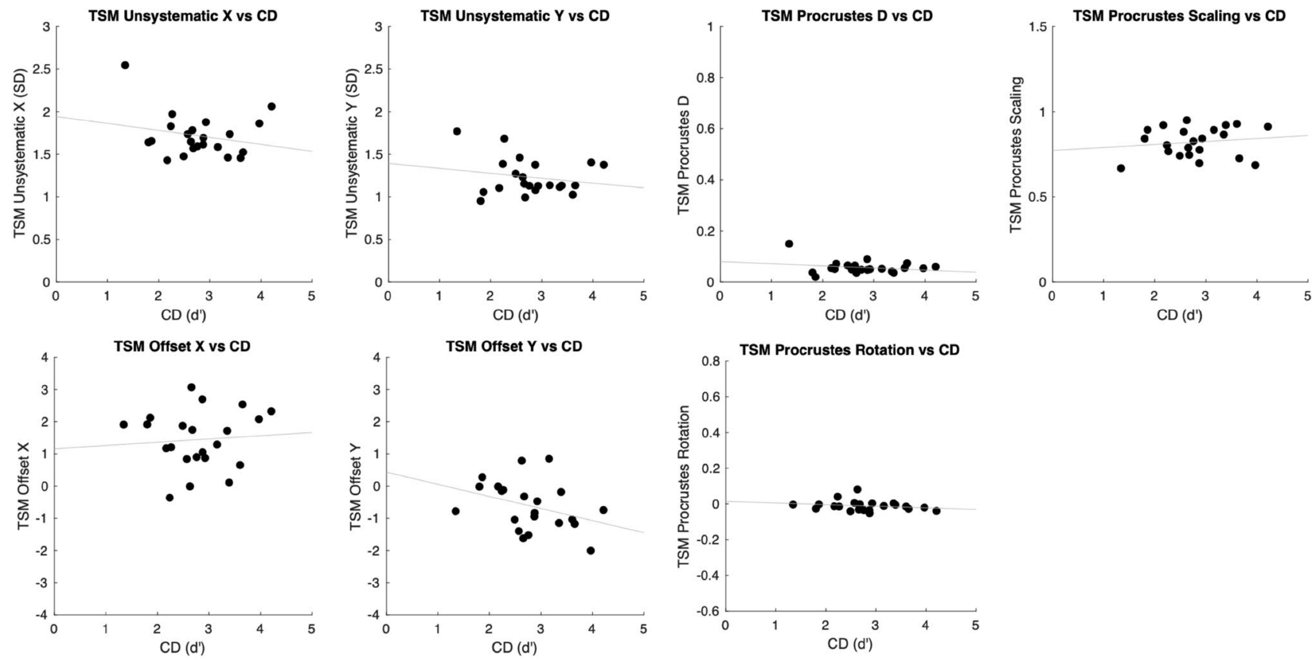

Fig. 8 Results of Experiment 2: correlations between CD performance and measures of transsaccadic error

Change detection task

We used the same task as in Experiment 2, except we only tested set sizes 3 and 4 and so there were 200 trials in total (one trial was lost due to a programming glitch).

2-Back task

Participants' working memory was also assessed using a 2-back task. To correspond with the change from line stimuli to squares in the TSM task, here participants fixated a central fixation point (0.2° across; white on a middle-grey background) for the entirety of the task, and then monitored a series of white squares (0.6° in width) that appeared, one at a time, for 400 ms at one of eight evenly spaced locations of a circular array (12° in diameter), followed by a 200-ms mask showing 1000 random squares, and then by a blank screen and central fixation point. 1 s later, the next trial began. Every time a square appeared at the same position as two trials earlier, participants were asked to respond with a key press on a computer keyboard. There were 186 trials tested in three blocks.

Results

Group averages of systematic and unsystematic errors are visualized in Fig. 10 for each memory load condition and saccade direction separately. To confirm and extend the results of Experiments 1 and 2, we again converted the data relative to saccade direction and then tested systematic and unsystematic errors for load-dependent and saccade direction effects in a series of seven 2-way repeated measures ANOVAs. Once again, unsystematic errors increased significantly with load (corrected ps < 0.001) (Fig. 11, Table 5).

In terms of systematic errors, the F tests for scaling and horizontal offset yielded load effects with p values of $p = 0.035$ or smaller, though these failed to reach significance after correcting for multiple comparisons. Procrustes D showed a trend (uncorrected $p = 0.027$) towards greater distortions with smaller loads that contrasted with Experiment 1 and that reflected that high load responses formed less regular octagons with shorter oblique edges compared to horizontal and vertical ones. This effect was not significant after serial Bonferroni correction though. As with the two previous experiments, there was no significant effect of load on vertical offset, nor was there any effect of load on Procrustes rotation. Consistent with both previous experiments, the tests showed no significant effect of saccade direction on any of the dependent variables, in terms of its main effects or its interactions with load (Fig. 11, Table 5).

We also verified the regression findings from the previous two experiments by testing TSM in conjunction with both 2-back and change detection performance within the same participants. To do so, we first conducted multiple regression analyses with stepwise inclusion of the TSM error variables (averaged across both saccade directions and loads) as predictors of participants' performance in the 2-back task and subsequently the change detection task. As was the

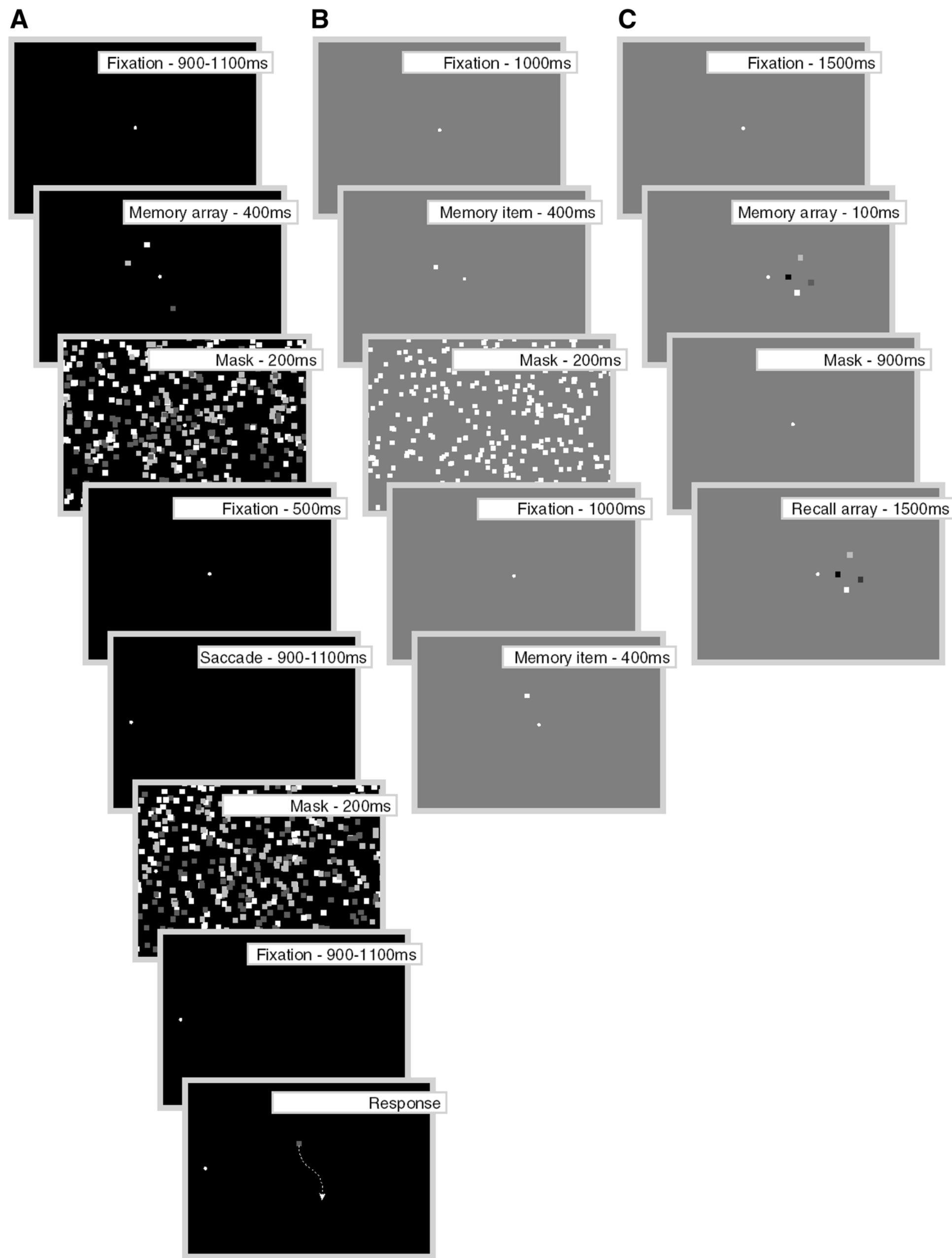

Fig. 9 Experimental paradigm. **a** Transsaccadic memory task. **b** 2-Back task. **c** Change detection task

case in Experiment 1, 2-back performance (mean $d' = 1.16$, SD = 0.96) was predicted by error in the TSM task, with a model comprising horizontal unsystematic error ($R^2 = 0.414$) as before, and horizontal translation as well ($R^2 = 0.123$) for a whole-model R^2 of 0.613 (horizontal systematic error $\beta = -0.71$, horizontal translation $\beta = 0.45$). As with Experiment 2, none of the TSM error variables significantly predicted performance on the change detection task (mean $d' = 2.40$, *SD* 0.92). Therefore, once again, it appears that the relationship between TSM and 2-back is stronger than the relationship between TSM and change detection.

Next, we proceeded to examine each of the TSM error variables individually for relationships with performance on the two traditional working memory tasks (Table 6). We found that 2-back performance was significantly predicted by TSM unsystematic error (both horizontal and vertical), and there was a trend towards a negative correlation with Procrustes *D* that, although failing to reach significance after correcting for multiple comparisons, was consistent with our finding in Experiment 1. The correlation reflected that participants with high *D* values produced octagons that were more irregular than participants with low *D* values. As for change detection, we found no significant predictors amongst our TSM error variables, which is consistent with our results from Experiment 2. In addition, there was no significant correlation between

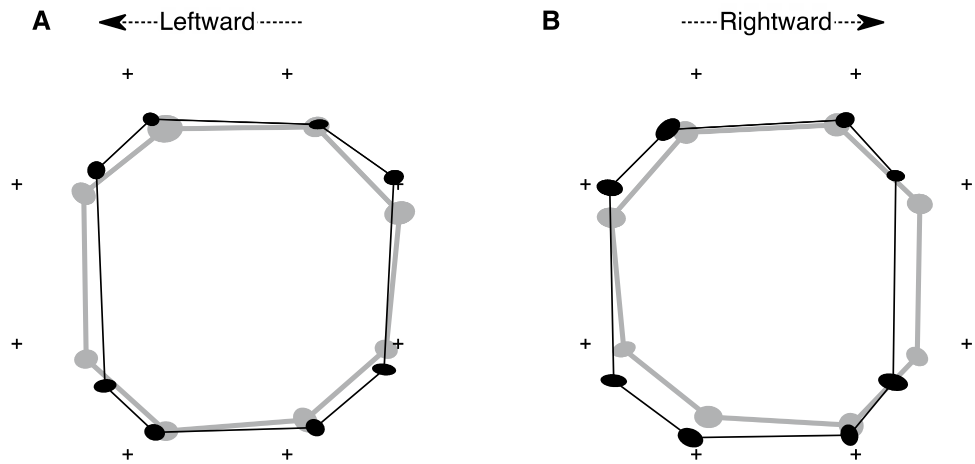

Fig. 10 Systematic and unsystematic errors. **a** Errors after leftward saccades. **b** Errors after rightward saccades. As a visual analogue for standard error, ellipses were fitted to group-level responses (mouse clicks) for each of the eight possible memory locations, with radii were shrunk by a factor of ten for graphical clarity. Ellipse centres represent the mean response for a given intersection location. Black: low memory load condition, grey: high memory load condition. Data are superimposed onto the true intersection locations (crosses)

2-back and change detection performance ($r=0.13$, 95% CI [-0.28, 0.50]) (Figs. 12, 13).

In sum, the results of this third experiment largely replicate the findings from Experiments 1 and 2, and intriguingly suggest a dissociation between 2-back and TSM on the one hand—and change detection performance on the other hand. We found that greater Procrustes *D* values predicted poorer 2-back performance, whereas higher load did not increase *D* values. Thus, in part between-subject and within-subject effects were governed by different mechanisms. The latter probably included perceptual effects. For example, participants might have misperceived memory squares as closer to an oblique position, but less so when multiple squares appeared.

General discussion

The current study asked how transsaccadic memory (TSM) relates to other forms of working memory. This is of significance given the fact that saccadic eye movements are extremely common in daily life while potentially having a substantial impact on memory functions. Yet, they are systematically eliminated from standard working memory paradigms. In three experiments, we found that unsystematic TSM errors were sensitive to working memory load. Load also exacerbated certain systematic types of TSM errors, although these errors varied, probably depending on the different spatial layouts of our tests and on people's response strategies. Crucially, we observed that those TSM errors that were sensitive to load also correlated with the 2-back task. However, we found no such correlations with the change detection task (CD). Our results show that TSM shares mechanisms at least with certain other types of working memory and that the representations of spatial information maintained across eye movements are impacted by representational noise as well as by imperfect remapping across eye movements.

The fact that TSM correlates with 2-back but not CD is surprising. Our TSM and CD paradigms were quite similar at the time of encoding, and both were rather different from the serialized presentation schedule of the 2-back task. It is very likely then that this dissociation can be attributed to differences between CD and TSM that occur after encoding, namely: remapping as one form of information manipulation, interference, and/or the degree to which gestalt-like or allocentric information is available at recall.

Remapping is clearly a potential error source for TSM in contrast to CD. That is, in the TSM the remembered stimulus location must be translated to its pre-saccadic position in eye-centred coordinates to accurately indicate its position on the display. Essentially, these translations constitute a form of manipulation of working memory content (e.g., Baddeley 2003) and, so, there is a possible connection to the 2-back task where manipulation of memory content is required to compare the present stimulus to the stimulus from two trials earlier. It is therefore possible that working memory manipulation is one source of errors, arguably systematic errors, that 2-back and TSM have in common, but not the CD task.

As a (perhaps related) commonality, TSM and 2-back, in contrast to the current CD task, include more distracting events that could interfere with working memory. Being able to suppress interference is a part of what is measured with visuospatial working memory tasks (Vogel et al. 2005; Oberauer and Lin 2017; Sun et al. 2017). For example, in both the TSM and the 2-back task, visual stimuli appear and must be attended in between encoding

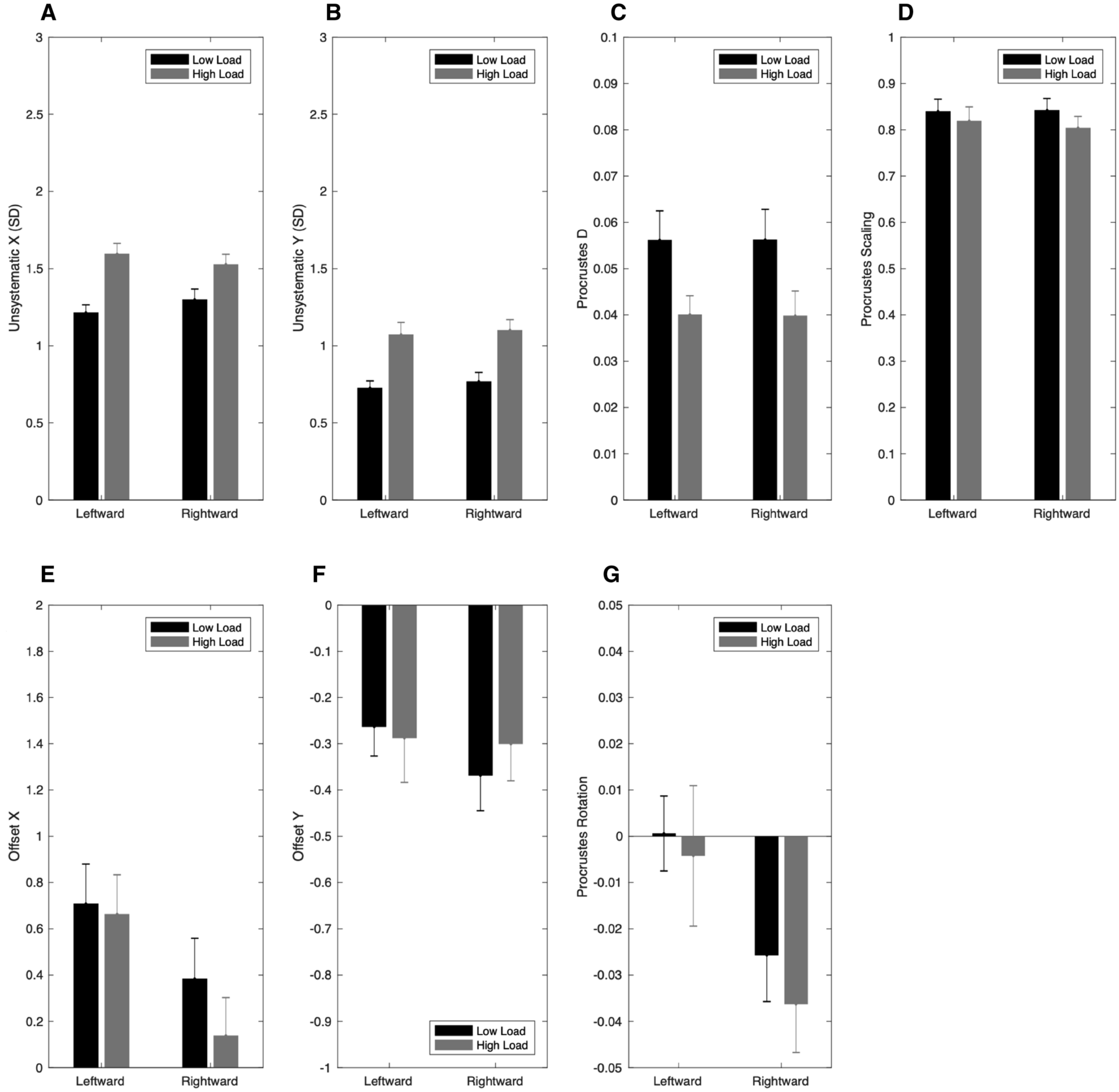

Fig. 11 Results of Experiment 3: TSM performance measure means. Bar colour indicates load level (black for low load, grey for high), and grouping indicates the saccade direction. Vertical axis labels indicate the variable displayed. **a**, **b** Horizontal and vertical unsystematic error. **c**–**g** Measures of systematic error

and recall. However, other interferences are quite different from one another: TSM requires intervening saccades while 2-back affords back and forth of suppressing and re-activating working memory content, which (if relevant) would be difficult to reconcile with the observed correlations. Finally, the masks in TSM and 2-back task represent a common visual form of distraction, but then it remains unclear why other visual commonalities between TSM and CD task did not produce correlations.

Another possibility is that the CD task differs from the other two because it provides allocentric information during recall. That is, when the original memory array reappears, participants only have to decide whether within the gestalt formed by the squares one has changed its colour (participants could also compare lists of colour names, but that is unlikely given the time pressure of the task). In contrast, with the TSM and the 2-back task only a single item is presented at the time of recall and, thus, little

Table 5 Two-way repeated measures ANOVAs for the measures of transsaccadic error

Error	Load			Saccade			Load × saccade		
	$F(1, 23)$	p	η_p^2	$F(1, 23)$	p	η_p^2	$F(1, 23)$	p	η_p^2
Unsystematic *X*	58.92***	0.000	0.72	0.06	0.816	0.00	7.99	0.010	0.26
Unsystematic *Y*	34.32***	0.000	0.60	1.03	0.320	0.04	0.05	0.826	0.00
D	5.56	0.027	0.20	0.00	0.979	0.00	0.00	0.962	0.00
Scale	5.01	0.035	0.18	0.38	0.544	0.02	1.11	0.302	0.05
Offset *X*	5.78	0.025	0.20	4.43	0.046	0.16	3.43	0.077	0.13
Offset *Y*	0.25	0.619	0.01	1.84	0.188	0.07	3.29	0.083	0.13
Rotation	1.05	0.316	0.04	4.88	0.037	0.18	0.04	0.852	0.00

Unsystematic error comprises the mean horizontal and vertical standard deviation. Procrustes *D*, scaling, offset and rotation are all derived from Procrustes analysis (without reflection). Greenhouse–Geisser corrections produced identical values for all ANOVAS

Asterisks denote significance after serial Bonferroni corrections for multiple comparisons. $*p<0.05$, $**p<0.01$, $***p<0.001$

Table 6 Correlations between 2-back performance and measures of transsaccadic error

TSM error	r	p	CI (95%)		Corrected CI	
			Lower	Upper	Lower	Upper
2-Back						
Unsystematic *X*	−0.64[a]	0.001	−0.83	−0.40	−0.87	−0.28
Unsystematic *Y*	−0.51[a]	0.011	−0.74	−0.25	−0.80	−0.15
D	−0.57	0.004	−0.83	−0.13	−0.88	0.06
Scaling	0.28	0.185	−0.15	0.67		
Offset *X*	0.35	0.096	−0.04	0.63		
Offset *Y*	0.24	0.261	−0.11	0.61		
Rotation	−0.17	0.433	−0.48	0.20		
CD						
Unsystematic *X*	−0.03	0.893	−0.33	0.24		
Unsystematic *Y*	−0.21	0.315	−0.60	0.17		
D	−0.28	0.188	−0.61	0.05		
Scaling	0.31	0.145	0.01	0.56	−0.13	0.64
Offset *X*	0.17	0.417	−0.40	0.56		
Offset *Y*	0.40	0.056	−0.04	0.69		
Rotation	−0.24	0.251	−0.56	0.12		

95% confidence intervals derived from bootstrapped regression analysis (100,000 resampling iterations)

[a]Zero falls outside of the serial Bonferroni-corrected confidence interval

allocentric information remains. Allocentric recall seems to recruit different neural mechanisms than egocentric recall (e.g. Feigenbaum and Morris 2004). However, this degree of dissociation based only upon the presence or absence of allocentric information would, to our knowledge, be a newly identified phenomenon.

Whichever of these explanations will hold true in the future, with regard to TSM itself, across the present three experiments we found that information is preserved across eye movements in terms of their spatial positions and topography, though this information was sensitive to working memory load and degraded by systematic and unsystematic errors. Contrary to the findings of Vasquez and Danckert (2008), we did not observe any saccade direction-specific working memory performance costs, consistent with a recent report (Brink et al. 2019).

In conclusion, our current study shows that several measures of TSM are sensitive to load and predict performance at least in one other working memory task. Remarkably, we found that TSM does not correlate with the CD task as employed here, although both tasks provide quite similar information at the time of encoding and working memory capacity correlates with many other cognitive functions (Fukuda et al. 2010; Gold et al. 2010; Johnson et al. 2013; Luck and Vogel 2013; Unsworth et al. 2014). In contrast, the 2-back task and TSM have more in common than either

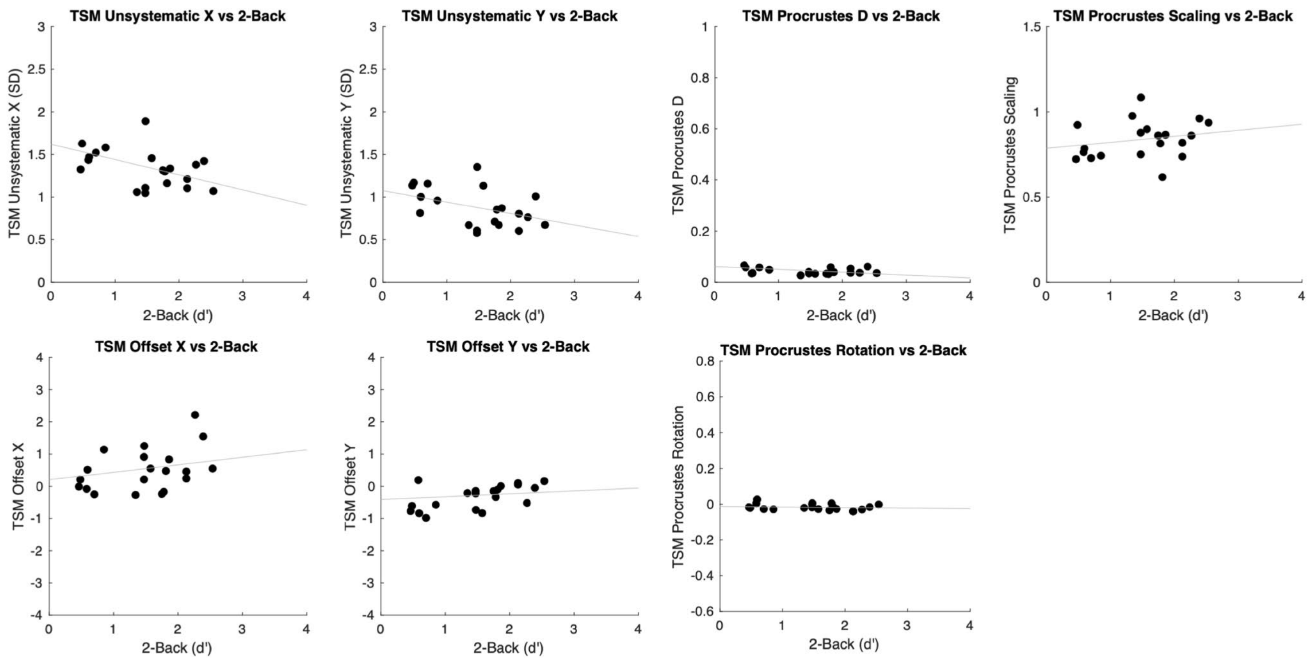

Fig. 12 Results of Experiment 3: correlations between 2-back performance and measures of transsaccadic error

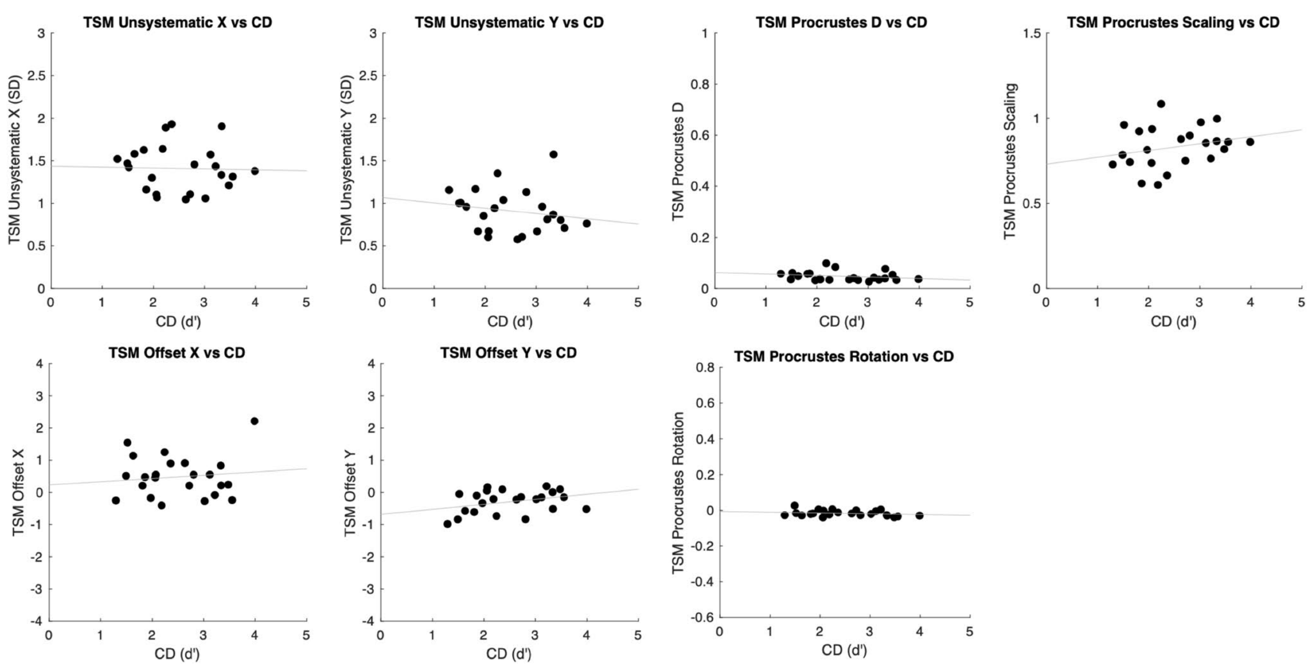

Fig. 13 Results of Experiment 3: correlations between CD performance and measures of transsaccadic error

one does with CD, even though stimulus presentation in the two tasks is rather different and 2-back is sometimes considered as a specialized and non-general measure of working memory compared to other tasks that preclude any eye movements (Kane et al. 2007; Jaeggi et al. 2010). Our study therefore demonstrates the significance of testing working memory functions across saccades.

Acknowledgements The authors acknowledge the funding from the Natural Sciences and Engineering Research Council of Canada.

References

Baddeley A (2003) Working memory: looking back and looking forward. Nat Rev Neurosci 4:829–839. https://doi.org/10.1038/nrn1201

Bays PM, Catalao RFG, Husain M (2009) The precision of visual working memory is set by allocation of a shared resource. J Vis 9:1–11. https://doi.org/10.1167/9.10.7

Brainard D (1997) The psychophysics toolbox. Spat Vis 10:433–436. https://doi.org/10.1163/156856897X00357

Brink AFT, Nijboer TCW, Fabius JH, Van Der Stigchel S (2019) No direction specific costs in trans-saccadic memory. Neuropsychologia 125:23–29. https://doi.org/10.1016/j.neuropsychologia.2019.01.014

Deubel H, Bridgeman B, Schneider WX (1998) Immediate post-saccadic information mediates space constancy. Vis Res 38:3147–3159. https://doi.org/10.1016/S0042-6989(98)00048-0

Feigenbaum JD, Morris RG (2004) Allocentric versus egocentric spatial memory after unilateral temporal lobectomy in humans. Neuropsychology 18:462–472. https://doi.org/10.1037/0894-4105.18.3.462

Fukuda K, Vogel E, Mayr U, Awh E (2010) Quantity, not quality: the relationship between fluid intelligence and working memory capacity. Psychon Bull Rev 17:673–679. https://doi.org/10.3758/17.5.673

Gold JM, Hahn B, Zhang WW et al (2010) Reduced capacity but spared precision and maintenance of working memory representations in schizophrenia. Arch Gen Psychiatry 67:570. https://doi.org/10.1001/archgenpsychiatry.2010.65

Grimes J (1996) On the failure to detect changes in scenes across saccades. In: Akins KA (ed) Vancouver studies in cognitive science, Vol 5. Perception. Oxford University Press, New York, NY, US, pp 89–110

Henriques DY, Klier EM, Smith MA et al (1998) Gaze-centered remapping of remembered visual space in an open-loop pointing task. J Neurosci 18:1583–1594

Irwin DE (1991) Information integration across saccadic eye movements. Cogn Psychol 23:420–456. https://doi.org/10.1167/14.1.28

Jaeggi SM, Buschkuehl M, Perrig WJ, Meier B (2010) The concurrent validity of the *N*-back task as a working memory measure. Memory 18:394–412. https://doi.org/10.1080/09658211003702171

Johnson MK, McMahon RP, Robinson BM et al (2013) The relationship between working memory capacity and broad measures of cognitive ability in healthy adults and people with schizophrenia. Neuropsychology 27:220–229. https://doi.org/10.1037/a0032060

Kane MJ, Conway ARA, Miura TK, Colflesh GJH (2007) Working memory, attention control, and the *N*-back task: a question of construct validity. J Exp Psychol Learn Mem Cogn 33:615–622. https://doi.org/10.1037/0278-7393.33.3.615

Luck SJ, Vogel EK (1997) The capacity of visual working memory for features and conjunctions. Nature 390:279–281. https://doi.org/10.1038/36846

Luck SJ, Vogel EK (2013) Visual working memory capacity: from psychophysics and neurobiology to individual differences. Trends Cogn Sci 17:391–400. https://doi.org/10.1016/j.tics.2013.06.006

Maronna RA, Martin RD, Yohai VJ (2006) Robust statistics: theory and methods Chichester. England, Wiley

Medendorp WP, Tweed DB, Crawford JD (2003) Motion parallax is computed in the updating of human spatial memory. J Neurosci 23:8135–8142

Merriam EP, Genovese CR, Colby CL (2003) Spatial updating in human parietal cortex. Neuron 39:361–373. https://doi.org/10.1016/S0896-6273(03)00393-3

Merriam EP, Genovese CR, Colby CL (2007) Remapping in human visual cortex. J Neurophysiol 97:1738–1755. https://doi.org/10.1152/jn.00189.2006

Niemeier M, Crawford JD, Tweed D (2003) Optimal transsaccadic integration explains distorted spatial perception. Nature 422:76–80. https://doi.org/10.1038/nature01439

Niemeier M, Crawford JD, Tweed D (2007) Optimal inference explains dimension-specific contractions of spatial perception. Exp Brain Res 179:313–323. https://doi.org/10.1007/s00221-006-0788-9

Oberauer K, Lin HY (2017) An interference model of visual working memory. Psychol Rev 124:21–59. https://doi.org/10.1037/rev0000044

Pelli D (1997) The VideoToolbox software for visual psychophysics: transforming numbers into movies. Spat Vis 10:437–442. https://doi.org/10.1163/156856897X00366

Phillips WA (1974) On the distinction between sensory storage and short-term visual memory. Percept Psychophys 16:283–290. https://doi.org/10.3758/BF03203943

Prime SL, Niemeier M, Crawford JD (2006) Transsaccadic integration of visual features in a line intersection task. Exp Brain Res 169:532–548. https://doi.org/10.1007/s00221-005-0164-1

Prime SL, Tsotsos L, Keith GP, Crawford JD (2007) Visual memory capacity in transsaccadic integration. Exp Brain Res 180:609–628. https://doi.org/10.1007/s00221-007-0885-4

Prime SL, Vesia M, Crawford JD (2008) Transcranial magnetic stimulation over posterior parietal cortex disrupts transsaccadic memory of multiple objects. J Neurosci 28:6938–6949. https://doi.org/10.1523/jneurosci.0542-08.2008

Prime SL, Vesia M, Crawford JD (2010) TMS over human frontal eye fields disrupts trans-saccadic memory of multiple objects. Cereb Cortex 20:759–772. https://doi.org/10.1093/cercor/bhp148

Simons DJ, Levin DT (1998) Failure to detect changes to people during a real-world interaction. Psychon Bull Rev 5:644–649. https://doi.org/10.3758/bf03208840

Simons DJ, Rensink RA (2005) Change blindness: past, present, and future. Trends Cogn Sci 9:16–20. https://doi.org/10.1016/j.tics.2004.11.006

Sun SZ, Fidalgo C, Barense MD et al (2017) Erasing and blurring memories: the differential impact of interference on separate aspects of forgetting. J Exp Psychol Gen 146:1606–1630. https://doi.org/10.1037/xge0000359

Unsworth N, Fukuda K, Awh E, Vogel EK (2014) Working memory and fluid intelligence: capacity, attention control, and secondary memory retrieval. Cogn Psychol 71:1–26. https://doi.org/10.1016/j.cogpsych.2014.01.003

Vasquez B, Danckert J (2008) Direction specific costs to spatial working memory from saccadic and spatial remapping. Neuropsychologia 46:2344–2354. https://doi.org/10.1016/j.neuropsychologia.2008.03.006

Vogel EK, McCollough AW, Machizawa MG (2005) Neural measures reveal individual differences in controlling access to working memory. Nature 438:500–503. https://doi.org/10.1038/nature04171

Vuilleumier P, Sergent C, Schwartz S et al (2007) Impaired perceptual memory of locations across gaze-shifts in patients with unilateral spatial neglect. Neurology 19:1388–1406. https://doi.org/10.1162/jocn.2007.19.8.1388