

**Remembering spatial words: Sensorimotor simulation affects verbal recognition memory**

Journal:	<i>Quarterly Journal of Experimental Psychology</i>
Manuscript ID	QJE-STD-21-056.R2
Manuscript Type:	Standard Article
Date Submitted by the Author:	10-Oct-2021
Complete List of Authors:	Kumcu, Alper; Hacettepe University, Department of Translation and Interpreting Thompson, Robin; University of Birmingham, School of Psychology
Keywords:	sensorimotor simulation, conceptual cueing, spatial interference, compatibility effect, recognition memory, grounded-embodied cognition

SCHOLARONE™  
Manuscripts

**Remembering Spatial Words: Sensorimotor Simulation Affects Verbal Recognition Memory**Alper Kumcu<sup>1</sup> and Robin Thompson<sup>2</sup><sup>1</sup>Department of Translation and Interpreting, Hacettepe University<sup>2</sup>School of Psychology, University of Birmingham**Author Note**

Experiment 1 reported in this study is partly based on the PhD thesis “Looking for Language in Space: Spatial Simulations in Memory for Language” authored by the first author under the supervision of the second author and submitted to the University of Birmingham. We have no conflicts of interest to disclose. Supplemental online materials: <https://osf.io/4yjdz> .

Correspondence concerning this article should be addressed to Alper Kumcu, Department of Translation and Interpreting, Hacettepe University, Turkey, 06800.

Email: [alperkumcu@hacettepe.edu.tr](mailto:alperkumcu@hacettepe.edu.tr), ORCID: <https://orcid.org/0000-0003-0844-3562>

Word Count: 12,349 (including abstract)

Abstract Word Count: 248

**Abstract**

Previous evidence shows that words with implicit spatial meaning or metaphorical spatial associations are perceptually simulated and can guide attention to associated locations (e.g., *bird* – upward location). In turn, simulated representations interfere with visual perception at an associated location. The present study investigates the effect of spatial associations on short-term verbal recognition memory to disambiguate between modal and amodal accounts of spatial interference effects across two experiments. Participants in both experiments encoded words presented in congruent and incongruent locations. Congruent and incongruent locations were based on an independent norming task. In Experiment 1, an auditorily presented word probed participants' memory as they were visually cued to either the original location of the probe word or a diagonal location at retrieval. In Experiment 2, there was no cue at retrieval but a neutral encoding condition in which words normed to central locations were shown. Results show that spatial associations affected memory performance although spatial information was neither relevant nor necessary for successful retrieval: Words in Experiment 1 were retrieved more accurately when there was a visual cue in the congruent location at retrieval but only if they were encoded in a non-canonical position. A visual cue in the congruent location slowed down memory performance when retrieving highly imageable words. With no cue at retrieval (Experiment 2), participants were better at remembering spatially congruent words as opposed to neutral words. Results provide evidence in support of sensorimotor simulation in verbal memory and a perceptual competition account of spatial interference effect.

*Keywords:* sensorimotor simulation, conceptual cueing, spatial interference, compatibility effect, recognition memory, grounded-embodied cognition

### Remembering Spatial Words: Sensorimotor Simulation Affects Verbal Recognition Memory

In contrast to propositional and computational theories (Chomsky, 1957; Fodor, 1975; Pylyshyn, 1986) that propose abstract, amodal, and arbitrary symbols; grounded-embodied cognition theories assume that there are no strict delimitations between perception and cognition and that mental representations are sensorimotor in nature (Barsalou, 1999, 2008; Paivio, 1986; Pulvermüller, 1999). A grounded-embodied treatment of language suggests that upon reading or hearing a word, traces of previous perceptual experiences (e.g., visual, auditory, olfactory, gustatory or tactile) related to the meaning of the word are automatically reactivated and give rise to a *sensorimotor simulation* of its referent (Bergen, 2016; Zwaan & Madden, 2005; Zwaan & Radvansky, 1998). For example, reading a spatial word, that is a word with a spatial association such as *bird* which elicits a mental image, can give rise to a perceptual experience of an upward location based on the typical location of birds in the physical world.

Several studies have shown the involvement of sensorimotor simulations in language comprehension in perceptual tasks (Bergen et al., 2007; Estes et al., 2008, 2015; Ostarek & Vigliocco, 2017; Richardson et al., 2003; Stanfield & Zwaan, 2001; Zwaan et al., 2002; Zwaan & Yaxley, 2003a). For instance, in Zwaan and Yaxley (2003a), participants identified whether spatial words such as *root* and *branch* are semantically related. Results showed faster and more accurate reactions when words were presented in their canonical locations (e.g., *root* was located under *branch*) compared to a non-canonical arrangement (e.g., *branch* under *root*) (see also Šetić & Domijan, 2007). In a follow-up study, Zwaan and Yaxley (2003a) replicated the processing advantage with spatially congruent presentation but only when words were presented to the left visual field that corresponds to the right hemisphere, which is responsible for storing and processing visuospatial relations. However, the interaction between match effect and visual field was not replicated when the response side (i.e., right-handed vs. left-handed responses) was counterbalanced between participants (Berndt et al., 2019).

1  
2  
3 Emotionally charged words with positive or negative valence such as *hero* or *liar* similarly evoke  
4  
5 sensorimotor simulations based on the metaphorical relationship between valence and vertical space  
6  
7 (i.e., GOOD IS UP” and “BAD IS DOWN”) (Lakoff & Johnson, 1980). These sensorimotor simulations are  
8  
9 reflected in faster reaction times (RTs) in semantic judgment tasks when there is an alignment between  
10  
11 the word’s metaphorical location and its actual position on the screen even though neither location is  
12  
13 relevant for a successful semantic judgement. For example, Meyer and Robinson (2004) demonstrated  
14  
15 that positive words are evaluated faster (i.e., “rate *hero* on a scale from 1 – negative to 5 – positive”) if  
16  
17 they are presented at the top of the screen and negative words (e.g., *liar*) are evaluated faster if they  
18  
19 are presented at the bottom (see also Chasteen et al., 2010; Giessner & Schubert, 2007; Leitan et al.,  
20  
21 2015; Meyer & Robinson, 2004; Schubert, 2005; Taylor et al., 2015; Zanolie et al., 2012). Words with  
22  
23 spatial associations can also guide visual attention to an associated location (i.e., a *conceptual cueing*  
24  
25 *effect*). For instance, words such as *bird* or *hero* can orient attention upward, whereas words such as  
26  
27 *root* or *liar* can orient attention downward (Amer et al., 2017; Dudschig et al., 2013; Estes et al., 2015;  
28  
29 Gozli et al., 2016). Such attentional bias has been found to interfere with visual perception at an  
30  
31 associated location and thus; counterintuitively, hinders identification of visual targets. To illustrate, *bird*  
32  
33 elicits slower identification of a visual target presented at the top of a display than at the bottom of the  
34  
35 display. Several studies have demonstrated this phenomenon, known as a *spatial interference effect* or a  
36  
37 *mismatch advantage* (Bergen et al., 2007; Estes et al., 2008; Gozli et al., 2013; Richardson et al., 2003;  
38  
39 Verges & Duffy, 2009).

40  
41  
42 Although a spatial interference effect has been reliably reported (see Estes & Barsalou, 2018 for  
43  
44 a comprehensive meta-analysis), there are different theories posited to explain the phenomenon. Below  
45  
46 we discuss both the grounded-embodied point of view and the memory retrieval point of view. From a  
47  
48 grounded-embodied point of view, a spatial interference effect has been attributed to perceptual  
49  
50 competition between the spatial representation of a word triggered by sensorimotor simulation and an  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 actual visual cue. The result of this competition is slowed perception in the overlapping location (see  
4  
5 Quadflieg et al., 2011). Alternative accounts are mostly based on *spatial indexing* or *spatial coding*  
6  
7 (Pylyshyn, 1989) and assume that word locations on the computer screen are encoded in an amodal  
8  
9 fashion (i.e., without the perceptual simulation) along with the words themselves (see Spivey et al.,  
10  
11 2004 for a review). In support of spatial coding account against modal simulation in semantic  
12  
13 judgement, Estes et al. (2015) found that words with strong (e.g., *sky*) and weak (e.g., *flourish*) visual  
14  
15 strength and imageability elicit a similarly strong interference effect. In memory retrieval tasks, spatial  
16  
17 indexing is assumed to account for a plethora of studies showing a *compatibility effect* (see Wynn et al.,  
18  
19 2019 for a review). That is, a match between an initial location where information is presented and a  
20  
21 later location where participants look to retrieve information improves recognition. For example,  
22  
23 participants in Scholz et al. (2016) listened to four factual statements presented with a loudspeaker  
24  
25 image in one of the four spatial areas on a two by two grid. Another statement probed participants'  
26  
27 recognition memory and during retrieval, with gaze behaviour manipulated by a visual cue appearing  
28  
29 either in a congruent location or a diagonal location with the original location of the loudspeaker image  
30  
31 associated with the probed information. Results showed that participants who were visually cued to the  
32  
33 congruent location were more successful in retrieving the information from memory despite the spatial  
34  
35 information being completely irrelevant for successful completion of the task (see also Johansson &  
36  
37 Johansson, 2014). These findings suggest that coding locations of information leads to a compatibility  
38  
39 effect and thus, a type of *match advantage* in memory, in contrast to spatial interference due to  
40  
41 conceptual cueing as described above. Hence, the alternative account from the memory point of view  
42  
43 predicts that target objects are coded not only for their location but also for their match with the  
44  
45 location of the cue word: if the physical location matches, then memory is improved and if it does not,  
46  
47 then memory is deteriorated.  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 The objective of the present study is to investigate the effect of spatial congruency of verbal  
4 stimuli on short-term verbal recognition memory as well as to compare evidence for either modal (i.e.,  
5 sensorimotor simulation in the present study) or amodal (i.e., spatial indexing in the present study)  
6 theories of language representation and spatial interference. In particular, we aim to test whether  
7 implicit spatial information is perceptually (thus, modally) simulated and/or whether word locations are  
8 amodally indexed during short-term verbal recognition memory. Participants in Experiment 1 saw four  
9 words with implicit spatial meaning or metaphorical spatial associations presented simultaneously on a  
10 two-by-two grid at the encoding phase and they were auditorily probed with one of the four encoded  
11 words. During retrieval, a visual cue appeared on the grid either in the same location (i.e., congruent  
12 retrieval) or in a diagonal location (i.e., incongruent retrieval) as to the location of the probe word  
13 during encoding. Participants in Experiment 2 did not see a cue at retrieval but studied spatially neutral  
14 words in addition to upward and downward words.

15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30 Our predictions for Experiment 1 were straightforward: If sensorimotor simulation is indeed an  
31 intrinsic part of language representation with consequences in verbal recognition memory, then a word  
32 with a spatial association would be expected to activate the visuospatial experience of its referent and in  
33 turn, orient attention to the associated location (i.e., a modal explanation). In light of previous evidence  
34 showing spatial interference in semantic tasks, we predict that retrieval performance will be poorer in  
35 the case of a match between the location of the probe word and a visual cue during retrieval (i.e.,  
36 congruent retrieval) compared to an incongruent retrieval condition, where the simulated location is  
37 empty at the time of retrieval. However, if spatial indexing is reflective of the retrieval process (i.e., an  
38 amodal explanation), then memory performance will be better when there is a matching visual cue at  
39 the retrieval screen as per the evidence demonstrating compatibility effect in memory. Adjudicating  
40 modal and amodal explanations is timely and critical considering the conflicting evidence and non-  
41 replications (e.g., Berndt et al., 2019) discussed above.

## Experiment 1

### Method

#### Participants

Experiment 1 was carried out with 48 students at the University of Birmingham (five males;  $M_{\text{age}} = 20.13$ ,  $SD = 3.27$ , range: 18 – 41). One participant's accuracy (46%) was identified as an outlier based on the interquartile range and thus replaced. All participants were monolingual native speakers of British English (speaking/learning only English from birth and currently using English as their primary language) as determined with the Language History Questionnaire (version 2.0; Li et al., 2013). Participants reported normal or corrected-to-normal vision, no speech or hearing difficulties and no history of any neurological disorder. They received either £8 or course credit for participation. All participants were fully informed about the details of the experimental procedure and gave written consent. Post-experiment debriefing revealed that all participants were naïve to the purpose of the experiment. Ethical review and approval for data collection were conducted and granted by the University of Birmingham Ethics Committee.

#### Stimuli

The experiment consisted of 192 trials in total. Trials were evenly divided into two groups ( $n = 96$ ) as experimental trials (positive probe) and fillers. Probe words in the experimental trials were among the four study words in the encoding phase, whereas a different, not seen, word was probed in fillers. Thus, 384 ( $96 \times 4$ ) words were used in positive probe trials and 480 ( $96 \times 5$ ) words were used in fillers. In total, there were 864 ( $384 + 480$ ) unique words across 192 trials. All words were nouns (see supplemental online materials for the list of words used in the experiments). 384 words in the experimental trials have implicit spatial meaning or metaphorical spatial associations. The spatial associations of experimental words were verified with a separate spatial association norming study (see



1  
2  
3 below). Experimental words were also a subset of the extension of Paivio et al. (1968) norms (Clark &  
4  
5 Paivio, 2004).

6  
7  
8 Words in the experimental trials were arranged into sets with four words. Words in filler trials  
9  
10 were sets of five words, including four for the presentation and one unseen word for the following  
11  
12 probe. Words within both experimental and filler sets were matched on imageability, concreteness and  
13  
14 word length (all  $SDs < 2.00$ ). Words were further controlled so that no word started with the same letter  
15  
16 or was semantically related with any other word in the trial set. Monosyllabic, disyllabic and trisyllabic  
17  
18 words were evenly distributed making 1/3 of each type (e.g., [3, 3, 3, 3], [1, 2, 1, 2] or [3, 2, 3, 2, 3] etc.).  
19  
20 Welch's t-tests revealed no significant difference between the probe and distractor words in  
21  
22 imageability, concreteness, word length or number of syllables (all  $ps > .05$ ). Thus, any word among the  
23  
24 four words in each trial set was as likely to be remembered as any other word. We further formed 192  
25  
26 unique mathematical equations (e.g.,  $[2 \times 3] - [2 + 3] = 1$ ) to present as memory interference between  
27  
28 encoding and retrieval phases (see Conway & Engle, 1996 for a similar design). Half of the equations  
29  
30 were correct. Incorrect equations were further divided into two equal groups: The results were either  
31  
32 plus or minus one from the correct result.  
33  
34  
35

### 36 37 **Spatial Association Norming**

38  
39 Before running Experiment 1, norming was conducted with a different group of 30 monolingual  
40  
41 native speakers of British English (five males;  $M_{age} = 19$ ,  $SD = 0.69$ , range: 18 – 20) who rated the spatial  
42  
43 associations of 1482 English nouns horizontal and vertical (x, y) coordinate space in a lab-based  
44  
45 experiment (Brysbaert et al., 2014 for a similar design). Handedness was tested with the Edinburgh  
46  
47 Handedness Inventory (Oldfield, 1971). Results showed that there were 16 right-handed, one left-  
48  
49 handed and 13 ambidextrous participants. The mean laterality quotient was 65 on a scale of -100  
50  
51 (absolute left-handedness) to +100 (absolute right-handedness) showing that right-handedness was  
52  
53 dominant among participants in particular including participants who can use their both hands, which is  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 a frequently seen case considering the prevalence of right-handedness (Corballis, 2003). The coordinate  
4 space was a blank, central square of 480 x 480 in pixels ( $11^\circ \times 11^\circ$  of visual angle) on the computer  
5 screen. Horizontal and vertical axes of the coordinate were 960 pixels long ( $\pm 480$  pixels in both  
6 directions from the origin) giving 921,600 different clickable points. The material set was randomly  
7 distributed over 5 lists of 290 words. The order of words in the lists was fully randomised. Participants  
8 were presented a centrally placed single word (Courier, font size = 52) in capital letters. They were  
9 instructed to read the word carefully and visualise its meaning before moving onto the rating phase.  
10 There was no time limit for the study phase and it ended by hitting the space key. Participants were  
11 presented a rating screen during the rating phase. They were expected to identify the word's location by  
12 left clicking any point within the square. When a participant clicked a location, the experimental  
13 software automatically recorded the click location both horizontally and vertically in the coordinate  
14 space.

15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30 Control words ( $n = 32$ ) were included to test inter-rater reliability. These words spanned the full  
31 range of locations and were randomly distributed into each list. Analyses indicated excellent inter-rater  
32 reliability based on intraclass correlation coefficients (ICC) for horizontal (ICC = 0.95, CI [95%] = 0.92 –  
33 0.97) and vertical positions (ICC = 0.98, CI [95%] = 0.97 – 0.99) (see Cicchetti, 1994; Hallgren, 2012 for  
34 cut-offs). The mean standard error in ratings was 3.02 for the horizontal plane and 5.37 for the vertical  
35 plane, indicating a high degree of consistency between participants in location ratings. We also  
36 computed rater correlation coefficients between each participant's average rating for a given word and  
37 the mean rating for the word in question within each presentation list. The mean coefficient of the  
38 correlation sets was 0.76 ( $SD = 0.07$ , range: 0.59 – 0.84) for the horizontal position and 0.65 ( $SD = 0.12$ ,  
39 range: 0.36 – 0.82) for the vertical position. Words that overlapped words with a data set from Estes et  
40 al. (2015) ( $n = 95$ ) showed a strong, positive correlation in vertical locations,  $r(92) = .92$ ,  $p < .001$ .

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

Variances (horizontal position = 13,128.40; vertical position = 41,536.82) clearly showed that ratings on the vertical axis were more spread out than the ratings on the horizontal axis. In other words, participants tended to use the whole range of the vertical scale but rated the words more conservatively on the horizontal scale. A heat map showing the distribution of the words across the coordinate system indicated that the majority of the words were rated on a diagonal axis from bottom left to top right (see Figure 1). This was also evident from the strong, positive correlation between horizontal and vertical positions;  $r(1437) = .44, p < .001$ .

[Figure 1 near here]

384 words in the experimental trials were selected from the normed and labelled words to be used as the experimental words in Experiment 1. An equal number of words ( $n = 24$ ) in the experimental set were associated with top, top right, bottom and bottom left locations in space. Horizontal positions were not considered as a variable due to the dominance of vertical space over horizontal space (Franklin & Tversky, 1990; Marmolejo-Ramos et al., 2013). Words were grouped as upward (top and top right) and downward words (bottom and bottom left) because the majority of the words were normed in a diagonal axis from bottom left to top right and there were very few words located in top left and bottom right positions. All participants saw an equal number of upward and downward words in total ( $n = 48$ ). That said, the horizontal position was taken into consideration when placing the words in the grid: While words associated with the upward and rightward position (e.g., *friend*) were placed in the top right quadrant, words associated with upward position only (e.g., *heaven*) were placed in the top left quadrant. Similarly, words associated with the downward and leftward position (e.g., *jail*) were placed in the bottom left quadrant and words associated with the downward position only (e.g., *mule*) were presented in the bottom right quadrant. Descriptive and spatial statistics of words used as a probe are

presented in Table 1. In addition to horizontal and vertical positions, we also calculated two spatial measures: spatiality and spatial variability. Spatiality shows the distance of a word from the centre. Hence, it is the extent to which a word is spatial as opposed to being spatially neutral. The spatiality value is the sum of the horizontal and vertical locations ratings (negative values were transformed into positive), normalised into a range from 0 to 1. Spatial variability was computed to investigate the words that participants rated on vastly different locations of the coordinate system. The spatial variability value is the average of the standard deviations between ratings among participants within each list, normalised into a range from 0 to 1.

[Table 1 near here]

### Apparatus

Stimuli were presented on a TFT LCD 24-inch widescreen monitor operating at 144 Hz with a resolution of 1920 x 1080 pixels (508 mm x 285.75 mm). The experiment was programmed in and run on OpenSesame (version 3.1.7; Mathôt et al., 2012). Probe words that were presented auditorily in the retrieval phase were produced by a native female speaker of British English in a sound-attenuated room and recorded using Audacity (version 2.1.10; <https://www.audacityteam.org>). Participants responded (yes/no they had seen the word) by pressing one of two keys on a standard keyboard.

Data were analysed and visualised in R programming language and environment (R Core Team, 2020). Mixed-effects models were constructed with *lme4* package (Bates et al., 2015). Significance values of the coefficients in models were computed based on the t-distribution using the Satterthwaite approximation with *lmerTest* package (Kuznetsova et al., 2016). Bayes factors were computed with *BayesFactor* package (Morey & Rouder, 2015).

### Procedure

The task was a yes/no verbal recognition memory test. As spelt out in detail below, each trial was composed of five consecutive phases: (1) fixation, (2) encoding, (3) interference, (4) retrieval and (5) feedback (see Figure 2). The task was to decide as quickly and accurately as possible whether an auditorily presented word had appeared previously or not. As soon as participants made a yes/no judgement (by hitting one of the response keys) the trial ended, and a new trial began. (1) Fixation: A fixation cross appeared at the centre of the screen for 500 ms. (2) Encoding: Participants were presented four words on a two by two grid for 1800 ms (see Richardson & Spivey, 2000 for a similar, “Hollywood Squares” design). The duration of the encoding phase was based on previous studies in which a compatibility effect was observed with a similar experimental design (Kumcu & Thompson, 2016, 2020). Encoding location was a between-subjects variable to minimise the learning effects between conditions. Half of the participants ( $n = 26$ , randomly chosen) saw the probe word in a spatially congruent location and the other half in an incongruent location. (3) Interference: Participants were presented mathematical equations and asked to identify whether the equation was correct or not within 10,000 ms (or they timed out). (4) Retrieval: The probe word was auditorily presented as participants looked at the screen with a blank grid. Participants were tasked to make a yes/no judgement to determine whether they had seen the probe word among the four words shown in the encoding phase within 3,000 ms (or they timed out). During the retrieval phase, a square appeared either in the same (congruent retrieval) or in the diagonal quadrant (incongruent retrieval) as the original location of the probe word in the encoding phase. Auditory words and visual probes were presented at the same time. Participants were instructed to look at the square as they gave their responses. Retrieval manipulation was a within-subjects variable. All participants were asked to retrieve the probe word in congruent and incongruent retrieval conditions. Participants were asked to respond as quickly and as accurately as possible. Therefore, they were instructed to keep their index fingers

1  
2  
3 positioned on the response keys at all times. The key press was randomised such that half of the  
4  
5 participants pressed the 'J' key for yes and the 'F' key for no, while the other half pressed 'F' for yes and  
6  
7 'J' for no. (5) Feedback: Participants were shown their accuracy (correct or incorrect) and response time  
8  
9 (RT) for 500 ms after each trial. Total accuracy and average RT were shown at the end of the  
10  
11 experiment. For a detailed schema of the experiment, see Figure 2.  
12  
13

14  
15 The order of trials and equations were fully randomised independently of each other. The  
16  
17 experiment was divided into four equal blocks with 48 trials in each block and a short break between the  
18  
19 blocks. A typical session lasted approximately 45 minutes. Overall accuracy in interference equations  
20  
21 and the recognition memory test for words (including fillers) were 79% and 81% respectively, suggesting  
22  
23 that participants attended to the task with high concentration.  
24

25  
26 Following the experiment, a computerised version of the Corsi block-tapping task (Corsi, 1972)  
27  
28 was used to measure visuospatial short-term memory and a forward digit span test was used to  
29  
30 measure short-term memory (using Psychology Experiment Building Language [PEBL], version 0.13, test  
31  
32 battery version 0.7, <http://pebl.org>) (Mueller & Piper, 2014). Participants were instructed to reproduce  
33  
34 lit-up blocks in the same order in Corsi block-tapping task. Corsi block span was reported as the  
35  
36 dependent measure, which was calculated by taking the minimum list length, adding the total number  
37  
38 correct, and dividing by the number of lists at each length. Participants were asked to repeat the  
39  
40 visual/auditory presentation of number strings by typing in the forward digit span test. The reported  
41  
42 dependent measure was digit span, which is the sum of the longest string of digits repeated without  
43  
44 error over two trials. Participants were also asked to report their level of daytime sleepiness on the  
45  
46 Epworth Sleepiness Scale.  
47  
48  
49  
50  
51

52  
53 *[Figure 2 near here]*  
54  
55  
56  
57  
58  
59  
60

## Results

### Data Preparation

#### Homogeneity

Variability between the participants was considered as a random effect in the mixed-effects models and thus, was controlled in the analysis. Along with that, participants who were tested on congruent and incongruent encoding conditions were compared on individual variability measures (i.e., Corsi Block-tapping test, forward digit span, Epworth Sleepiness Scale) to verify the homogeneity between the groups. There was not a significant difference between participants tested on congruent and incongruent encoding conditions in any of the measures (see Table 2).

*[Table 2 near here]*

#### Measures

Accuracy and RT were used as measures of memory performance. Accuracy was the proportion of experimental trials to which the participants correctly responded yes. RT was the time in milliseconds between the onset of auditory presentation of the probe word and correct keyboard response. Responses identified as outliers that were 2 *SD* away (678.5 ms) from the mean (1079.41 ms) were removed ( $n = 179$  out of 3,721 observations, 4.81%) from the RT dataset. The average accuracy in the memory retrieval task (excluding the fillers) was 80.75% ( $SD = 0.39$ ) and the average RT (excluding fillers and incorrect responses) was 1029.02 ms ( $SD = 253.17$ ). Overall, 4,608 observations for accuracy (excluding fillers) and 3,542 observations for RTs (excluding fillers and incorrect responses) collected from 48 participants were analysed.

### Mixed-effects modelling

Data were analysed using linear and binomial logit mixed-effects modelling. There are several advantages of mixed-effects models with crossed random effects over traditional by-participant ( $F1$ ) by-item ( $F2$ ) ANOVAs (Baayen et al., 2008). The use of mixed-effects models is particularly safer, more powerful and more informative in psycholinguistics (see Hutchinson et al., 2013) because language materials such as words used as test items in these studies show substantial variation even if they are controlled on several properties. Further, it is misleading to generalise the findings beyond the language materials specifically used as a sample in a given study (Clark, 1973). In the current study, both participants and items were treated as random effects to explain by-participant and by-items variation and to neutralise potentially relevant covariates such as learning, fatigue or order effects. Thus, we captured the fact that some participants were more accurate and faster than others and some items were easier to remember than others, which is not possible with a traditional ANOVA.

Visual inspections of residual plots did not reveal any obvious deviations from homoscedasticity or linearity. Linear models were fit for continuous target variables (RT). Binomial models were fit for categorical target variables (accuracy) and with *bobyqa* optimiser to prevent nonconvergence.

We started fitting models by building the random effects structure and followed a maximal approach. That is, random effects were included as both random intercepts and correlated random slopes (random variations) as long as they converged and were justified by the data (Barr et al., 2013). Random intercepts and slopes were included even if they did not improve the model fit to control for possible dependence due to repeated measures or order effects. Random slope models allowed the explanatory variable to have a different effect for each group. Thus, we also captured that the fact that the effect of encoding or retrieval condition on memory performance may vary by the lexical or semantic properties of words such as word length.



1  
2  
3 Random effects structure was simplified step by step as per the magnitude of the contribution  
4 of a random effect to the explanation of the variation in the data. That is, the random effect with the  
5 weakest contribution was dropped first and if necessary, the structure was further reduced accordingly  
6  
7  
8  
9  
10 (Winter, 2013). Results are reported in terms of the unstandardised regression coefficient,  $B$ .

### 11 **Factor analysis**

12 Diagnostic tests indicated collinearity between word variables (i.e., imageability, concreteness,  
13 word length, number of syllables, horizontal position and vertical position) (MeanVIF = 2.86, RangeVIF =  
14 2.06 – 3.46; kappa = 2400.33). To address collinearity, we performed exploratory factor analysis using a  
15 principal component analysis extraction method with an orthogonal (varimax) rotation method. Kaiser's  
16 criterion and the scree test criterion showed the presence of three factors in our data. Three factors  
17 were interpreted as follows based on the loadings (see Table 3): (1) Imagery: Imageability and  
18 concreteness, (2) Length: word length and syllable length and (3) Word Position: Horizontal and vertical  
19 position. The three-factor solution explained 89.38% of the variance in the data. VIF of the factors were  
20 below one and thus, below our threshold of two. Regression scores calculated for three factors were  
21 employed in the subsequent linear mixed-effects multiple regression models.  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38

39 *[Table 3 near here]*  
40  
41  
42

### 43 **Memory Performance**

44 Mixed-effects models with accuracy and RT as target variables were fit. Fixed effects were  
45 encoding condition with two levels (i.e., spatially congruent encoding and spatially incongruent  
46 encoding) and retrieval condition with two levels (i.e., spatially congruent retrieval and spatially  
47 incongruent retrieval) and their interaction (see supplemental online materials for analysis codes).  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

All factors (i.e., imagery, word length and word position) were added as random slopes into participants and items for the general accuracy model. Summary of the general accuracy model with encoding and retrieval conditions as fixed effects indicated that overall, participants retrieved the probe word more accurately when they were cued to the congruent position (i.e., where the probe word appeared before) ( $M = 0.81$ ,  $SD = 0.39$ ) during retrieval;  $B = 0.25$ ,  $z = 2.14$ ,  $p = .03$  compared to the incongruent position ( $M = 0.80$ ,  $SD = 0.40$ ). For the general RT model, imagery and word length factors were added as random slopes into participants and imagery factor was added as random slopes into items as the maximal model did not converge. Summary of the general RT model with encoding and retrieval conditions as fixed effects indicated that overall, words encoded in congruent locations (e.g., *bird* in the top left quadrant) were retrieved slower ( $M = 1075.26$  ms,  $SD = 252.59$ ) than words encoded in incongruent locations (e.g., *bird* in the bottom right quadrant) ( $M = 985.63$  ms,  $SD = 245.99$ );  $B = 103.49$ ,  $t = 2.95$ ,  $p = .005$  regardless of the cue position at retrieval. Importantly, there was an interaction effect between retrieval condition and imagery factor, showing that highly imageable words slowed down retrieval when there was a visual cue in the same quadrant as to the initial location of the probe word,  $B = 17.32$ ,  $t = 2.39$ ,  $p = .02$ .

Likelihood tests showed that encoding condition;  $\chi^2(1) = 0.49$ ,  $p = .48$  or retrieval condition;  $\chi^2(1) = 1.23$ ,  $p = .27$  did not improve the null model fit for accuracy. Encoding condition improved the null model fit for RT;  $\chi^2(1) = 6.30$ ,  $p = .01$ ,  $BF = 2.71$  but not retrieval condition;  $\chi^2(1) = 0.75$ ,  $p = .39$ .

No interaction was found between imagery factor and encoding for accuracy or RT and also, imagery factor and retrieval for accuracy (the smallest  $p = .08$ ). There was no correlation between the performance in the mathematical equation and the performance in the recognition task (the smallest  $p = .08$ ).

The general accuracy model revealed an interaction effect between encoding and retrieval and indicated worse accuracy with congruent encoding and congruent retrieval;  $B = -0.30$ ,  $t = -1.93$ ,  $p = .05$ .

1  
2  
3 A condition variable with four levels involving all four encoding and retrieval combinations was  
4  
5 formulated to investigate the interaction of encoding and retrieval conditions on memory performance  
6  
7 more precisely (see Table 4). Descriptive measures of memory performance across conditions were  
8  
9 presented in Table 4 in the order of accuracy from the most accurate to the least accurate.  
10  
11  
12  
13

14 [Table 4 near here]  
15  
16  
17  
18

19 Smaller models were fit only with congruently and incongruently encoded words. All factors  
20 (i.e., imagery, word length and word position) were added as random slopes into participants and items  
21 for the accuracy model. Retrieval condition within incongruent encoding condition improved the null  
22 model fit for accuracy;  $\chi^2(1) = 4.59, p = .03, BF = 0.39$ . Summary of the accuracy model demonstrated  
23 that words were retrieved more accurately when they were encoded in an incongruent position and  
24 (e.g., *bird* in the bottom right quadrant) and a visual cue appeared in the same position (e.g., top right  
25 quadrant) during retrieval;  $B = 0.23, z = 2.11, p = .03$ . Retrieval condition within congruent;  $\chi^2(1) = 0.13, p$   
26 = .72 or incongruent encoding;  $\chi^2(1) = 2.42, p = .12$  did not improve model fit for null RT model.  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38

39 [Figure 3 near here]  
40  
41  
42  
43

## 44 Discussion

45 Results of Experiment 1 demonstrated that spatial manipulation at encoding and retrieval  
46 influenced memory performance. Incongruent encoding, with a congruent retrieval cue, was the most  
47 accurate combination. We argue that this is because reading words with literal or metaphorical spatial  
48 associations simulated the suggested locations. The performance was better when the simulated  
49 location did not overlap with the visual cue in an incongruent encoding (e.g., *bird* in the bottom right  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 quadrant) followed by a congruent retrieval (e.g., visual cue in the bottom right quadrant. Participants  
4  
5 retrieved the words encoded in incongruent locations faster compared to congruent locations. That  
6  
7 said, we also found an interaction effect between imageability of the words and cue position on  
8  
9 retrieval times: Highly imageable words (e.g., *mountain*, *volcano* or *university* etc.) were retrieved more  
10  
11 slowly than less imageable words (e.g., *outcome*, *magnitude* or *truth* etc.) when there was a visual cue in  
12  
13 the same quadrant as the location of the probe word regardless of the spatial arrangement at encoding.  
14  
15 Such a finding suggests that highly imageable words with spatial associations elicit stronger simulations  
16  
17 of the associated locations and, in turn, lead to a stronger interference effect when matched with visual  
18  
19 information in the same location. In light of our findings and following the converging evidence  
20  
21 supporting perceptual competition (e.g., Verges & Duffy, 2009), we argue that sensorimotor simulation  
22  
23 dictated memory performance in Experiment 1. We attribute the spatial interference effect to  
24  
25 perceptual competition between the sensorimotor representation of words and visual cues. If an  
26  
27 amodal account, such as spatial indexing, were reflective of the retrieval process, memory performance  
28  
29 would have been better whenever there was a match between the location of the probe word and the  
30  
31 visual cue (i.e., congruent retrieval) as shown in previous studies (Johansson & Johansson, 2014; Scholz  
32  
33 et al., 2016). However, a match between the simulated location and the visual probe resulted in worse  
34  
35 performance.  
36  
37  
38  
39  
40

## 41 Experiment 2

42  
43 Experiment 1 examined spatial manipulation both at encoding and retrieval stages with the  
44  
45 encoding manipulation as a between-participants variable. Hence, we conducted a second experiment  
46  
47 in an attempt to replicate the findings with a larger sample and a simpler, more powerful design. To be  
48  
49 more precise, we tested the effect of spatial congruency on memory performance with a repeated  
50  
51 measures within-participants design. In other words, all participants saw the words in congruent and  
52  
53 incongruent positions at encoding. The retrieval manipulation was removed altogether from Experiment  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 2 to isolate the effect. Further, a control encoding condition with spatially neutral words normed to  
4  
5 central locations was introduced. In Experiment 1, words encoded in incongruent positions were  
6  
7 remembered better as compared to words encoded in congruent position. However, it was not clear  
8  
9 whether the mismatch advantage stemmed from an actual advantage for mismatched information or a  
10  
11 disadvantage for the matched information as there was no control condition with neutral words. The  
12  
13 neutral condition in Experiment 2 was added to disambiguate direction of the effect more precisely. If  
14  
15 spatial information was a part of verbal memory, then memory performance would be different across  
16  
17 encoding conditions, which would support results from Experiment 1. As there was no cue in this  
18  
19 experiment, we predicted that the performance would be better in the congruent encoding condition  
20  
21 compared to incongruent and/or neutral conditions.  
22  
23  
24

## 25 Method

### 26 Participants

27  
28  
29  
30 Experiment 2 was carried out with 55 monolingual native speakers of English recruited via  
31  
32 Prolific (<https://www.prolific.co>), an online participant recruitment system. The participant pool was  
33  
34 filtered to recruit individuals with normal or corrected-to-normal vision, no speech or hearing difficulties  
35  
36 and no history of any neurological disorder. They received £10 for participation. All participants were  
37  
38 fully informed about the details of the experimental procedure and gave consent. Ethical review and  
39  
40 approval for data collection were conducted and granted by the University of Birmingham Ethics  
41  
42 Committee. One participant's accuracy (48%) was identified as an outlier based on interquartile range  
43  
44 and thus removed from the dataset. The analysis reported below was conducted with 54 participants  
45  
46 (28 males;  $M_{\text{age}} = 21.13$ ,  $SD = 2.28$ , range: 18 – 25).  
47  
48  
49

### 50 Stimuli

51  
52 As in Experiment 1, Experiment 2 consisted of 192 trials in total. The same materials as in  
53  
54 Experiment 1 were used except for spatially neutral words ( $n = 32$ ) (e.g., *board* etc.) associated with the  
55  
56  
57  
58  
59  
60

1  
2  
3 central location in the norming study. Neutral words were shown in the neutral encoding condition.  
4  
5 Based on results from Experiment 1 and in Kumcu and Thompson (2020), the lexicosemantic variables  
6  
7 that were used to control the words were reduced: Only imageability and word length in letters were  
8  
9 selected from the imagery and length factors of Experiment 1. There was a high correlation between  
10  
11 spatiality and emotionality in the norming study;  $r(1437) = .37, p < .001$  showing a difference between  
12  
13 spatial and neutral words in emotionality. Words that were located in the centre were less emotional.  
14  
15 Emotionality was thus added to control the words to make sure that any difference in memory between  
16  
17 spatial and neutral words stems from the spatial congruency manipulation rather than emotionality. As  
18  
19 expected, spatial and neutral words in Experiment 2 were different in emotionality;  $t(94) = 4.87, p < .001$   
20  
21 but there was no such difference between up and down words;  $t(62) = 1.50, p = 0.14$ . One-way ANOVAs  
22  
23 showed that there was no difference in imageability;  $p = 0.49$  or word length;  $p = 0.3$  between the up,  
24  
25 down or neutral words. There was no difference between spatial (up and down words combined) and  
26  
27 neutral words in imageability;  $p = 0.76$  or word length;  $p = 1$  either. There were also reliable differences  
28  
29 between the up, down and neutral words in horizontal;  $F(2,93) = 71.09, p < .001$  and vertical positions;  
30  
31  $F(2,93) = 661.09, p < .001$  as per the spatial norming (see Table 5).  
32  
33  
34  
35

### 36 Procedure

37  
38  
39 The same procedure as in Experiment 1 was used. The exceptions were the neutral encoding  
40  
41 condition and the exclusion of spatial cues shown at retrieval. Participants were now asked to retrieve  
42  
43 the probe word as quickly and as accurately as possible without any looking instructions. The main  
44  
45 manipulation of Experiment 2 was the spatial match between the spatial association of the probe word  
46  
47 and its position on the grid during encoding. Neutral probe words were counterbalanced across  
48  
49 participants to be shown in each one of the four quadrants of the grid. The spatial match condition,  
50  
51 location of the distractors, location of the neutral probe words, response hand-sides (old/new for the  
52  
53 memory task and correct/incorrect for the equation task) were counterbalanced across participants. As  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 a result, there were eight versions of the experiments. There was no difference in accuracy or RTs  
4  
5 between the test versions (all  $ps > .05$ ). The experiment was conducted online via Gorilla  
6  
7 (<https://gorilla.sc>), an online experiment development and hosting service.  
8  
9

10  
11  
12 [Table 5 near here]  
13  
14  
15

## 16 Results

### 17 Measures

18  
19  
20  
21 Outliers at the higher and lower extremes ( $n = 94$  out of 4,069 observations, 2.31%) that were  
22  
23 identified based on the interquartile range were removed from the RT dataset. The average accuracy in  
24  
25 the memory retrieval task (excluding fillers) was 80% ( $SD = 0.4$ ) and the average RT (excluding fillers and  
26  
27 incorrect responses) was 1310.40 ms ( $SD = 459.97$ ). Overall, 5,063 observations for accuracy (excluding  
28  
29 fillers) and 3,975 observations for RTs (excluding fillers and incorrect responses) collected from 54  
30  
31 participants were analysed.  
32  
33

### 34 Memory Performance

35  
36  
37 Word type with two levels (spatial vs. neutral) significantly improved both the null accuracy  
38  
39 model;  $\chi^2(1) = 5.27, p = .02, BF = 0.40$  and the null RT model;  $\chi^2(1) = 5.62, p = .02, BF = 0.63$ . Participants  
40  
41 retrieved the spatial words less accurately ( $M = 0.80, SD = 0.4$ );  $B = 0.19, z = 2.20, p = .03$  but faster ( $M =$   
42  
43 1301.07 ms,  $SD = 457.31$ ) than the spatially neutral words ( $M = 0.82, SD = 0.38$  and  $M = 1328.91$  ms,  $SD$   
44  
45 = 464.82);  $B = 26.26, t = 2.37, p = .02$  regardless of the spatial match.  
46  
47

48  
49 Spatial match condition with three levels (congruent, incongruent and neutral) improved the  
50  
51 null accuracy model;  $\chi^2(2) = 6.53, p = .04, BF = 0.06$  (see Figure 1 and Table 2). Participants retrieved the  
52  
53 spatial words in incongruent positions (e.g., *victory* in downward position) ( $M = 0.79, SD = 0.41$ ) less  
54  
55 accurately than the spatially neutral words associated with the central position ( $M = 0.82, SD = 0.38$ );  $B =$   
56  
57  
58  
59  
60

0.24,  $z = 2.57$ ,  $p = .01$ . There was no such difference in accuracy between the retrieval of spatial words in congruent position (e.g., *victory* in upward position) ( $M = 0.80$ ,  $SD = 0.4$ ) and neutral words;  $B = 0.14$ ,  $z = 1.46$ ,  $p = .14$ . A smaller model fitted with spatial words only indicated that there was no difference in accuracy between the retrieval of spatial words in congruent and incongruent positions either;  $\chi^2(2) = 1.29$ ,  $p = .26$ . Spatial match condition improved the null RT model significantly;  $\chi^2(2) = 6.12$ ,  $p = .05$ ,  $BF = 0.06$ . Participants retrieved the words in congruent positions (e.g., *victory* in upward position) ( $M = 1290.57$  ms,  $SD = 461.14$ ) faster than the spatially neutral words associated with the central position ( $M = 1328.91$  ms,  $SD = 464.82$ );  $B = 30.77$ ,  $t = 2.41$ ,  $p = .02$  (see Table 6 and Figure 4). There was no such difference between the retrieval of spatial words in incongruent positions (e.g., *victory* in downward position) ( $M = 1311.75$  ms,  $SD = 453.31$ ) and neutral words;  $B = 21.49$ ,  $t = 1.68$ ,  $p = .09$ . A smaller model fitted with spatial words only indicated that there was no difference in accuracy between the retrieval of spatial words in congruent and incongruent positions either;  $\chi^2(1) = 0.36$ ,  $p = .55$ .

[Table 6 near here]

[Figure 4 near here]

Spatial word type with two levels (up words vs. down words) did not improve the null accuracy model fitted with only spatial words;  $\chi^2(1) = 0.07$ ,  $p = .8$ . Participants retrieved the up ( $M = 0.79$ ,  $SD = 0.41$ ) and down words ( $M = 0.80$ ,  $SD = 0.4$ ) equally accurately. That said, spatial word type improved the null RT model;  $\chi^2(1) = 5.85$ ,  $p = .02$ ,  $BF = 1.19$ . Overall, participants retrieved the up words ( $M = 1315.57$  ms,  $SD = 470.81$ ) slower than down words ( $M = 1286.69$  ms,  $SD = 443.24$ );  $B = 31.24$ ,  $t = 2.43$ ,  $p = .02$ . Smaller models with only up and down words indicated that the words associated with downward location were retrieved slower when they were presented in incongruent positions ( $M = 1322.17$  ms,  $SD = 452.10$ ) compared to congruent positions ( $M = 1251.68$  ms,  $SD = 431.82$ );  $B = 50.48$ ,  $t = 2.86$ ,  $p = .004$ .



1  
2  
3 There was no such difference between the retrieval of up words in congruent ( $M = 1329.63$  ms,  $SD =$   
4  
5 486.02) and incongruent positions ( $M = 1301.19$  ms,  $SD = 454.64$ );  $\chi^2(1) = 2.77$ ,  $p = .1$ .

6  
7 Results showed that imageability (but not word length;  $p = .42$  or emotionality;  $p = .32$ )  
8  
9 improved the accuracy in general;  $\chi^2(1) = 4.33$ ,  $p = .04$ ,  $BF = 0.42$ . Overall, participants retrieved the high  
10  
11 imageable words more accurately than the low imageable words;  $B = 0.08$ ,  $z = 2.13$ ,  $p = .03$ . Interaction  
12  
13 models showed that imageability improved the accuracy for neutral words to the most extent;  $B = 0.10$ ,  
14  
15  $z = 2.60$ ,  $p = .009$  and also spatially congruent words with a smaller magnitude;  $B = 0.08$ ,  $z = 2.07$ ,  $p = .04$ .  
16  
17 However, imageability did not improve the accuracy of spatial words presented in incongruent  
18  
19 positions;  $B = 0.06$ ,  $z = 1.55$ ,  $p = .12$ . Results showed that word length (but not imageability;  $p = .52$  or  
20  
21 emotionality;  $p = .96$ ) affected the RTs in general;  $\chi^2(1) = 4.32$ ,  $p = .04$ ,  $BF = 0.36$ . Overall, participants  
22  
23 retrieved the longer words more slowly than the shorter words;  $B = 7.62$ ,  $t = 2.09$ ,  $p = .04$ . The effect of  
24  
25 word length was evident in the retrieval of spatially neutral words with the largest magnitude;  $B = 10.73$ ,  
26  
27  $t = 2.82$ ,  $p = .005$  and also in the retrieval of spatial words in incongruent positions;  $B = 7.26$ ,  $t = 1.94$ ,  $p =$   
28  
29  $.05$ . Word length did not affect the retrieval of spatial words in congruent positions;  $B = 6.07$ ,  $t = 1.63$ ,  $p$   
30  
31  $= .1$ .

### 32 33 34 35 36 37 Discussion

38  
39 Results from Experiment 2 demonstrate that spatial manipulation at encoding and spatial  
40  
41 meaning of words manipulate memory performance. Participants retrieved spatial words less accurately  
42  
43 but faster than the spatially neutral words even though lexicosemantic differences between spatial and  
44  
45 neutral words were strictly controlled both at the stimuli development and analysis stages. Congruency  
46  
47 of spatial information affected retrieval performance as well. Participants retrieved spatial words in  
48  
49 incongruent positions less accurately than neutral words. Likewise, they remembered spatial words in  
50  
51 congruent positions faster than neutral words. Results from Experiment 2 support and extend results  
52  
53 from Experiment 1, although the direction of the effect is different as discussed below.  
54  
55  
56  
57  
58  
59  
60

### General Discussion

The present study investigated the effect of spatial associations on short-term verbal recognition memory and disambiguated between modal and amodal accounts of spatial interference effects in two experiments. In particular, Experiment 1 explored the role of sensorimotor simulation vs. spatial indexing in verbal recognition short-term memory. Previous studies have yielded mixed results as to how words with intrinsic or metaphorical spatial associations are processed: While there is evidence that visually-loaded words guide attention to the associated location which in turn slows down the identification of visual targets in the same location (e.g., Verges & Duffy, 2009), there is also evidence that spatial compatibility between a word and a visual cue improves recognition memory (Wynn et al., 2019). Against this background, we investigated modal (sensorimotor simulation) and amodal (spatial indexing) theories of language representation and spatial interference effects in a short-term verbal recognition memory task in Experiment 1. Words encoded in a spatially incongruent position from their spatial association (e.g., *bird* in the bottom right quadrant or *mule* in the top left quadrant) were retrieved more accurately when there was a visual cue in the same quadrant during retrieval (e.g., bottom right and top left, respectively) compared to a spatially incongruent retrieval cue. Words encoded in a spatially incongruent position (e.g., *bird* in the bottom right quadrant or *mule* in the top left quadrant) were retrieved faster than words encoded in a spatially congruent position (e.g., *bird* in the top left quadrant or *mule* in the bottom right quadrant) regardless of the cue position during retrieval. As a follow-up study, Experiment 2 investigated the effect of encoding manipulation on memory performance with a control condition, a larger sample and more powerful within-participants design. Here, participants retrieved the spatially-loaded words faster but less accurately than spatially neutral words even when lexical and semantic differences between spatial and neutral words were controlled. As to the congruency between the location of the probe word on the screen and the location

1  
2  
3 implied by the word, the results indicated that participants retrieved incongruent words less accurately  
4  
5 and congruent words faster than the control words, suggesting a match advantage.  
6

7  
8 The results from both experiments overall suggest that participants tapped into the spatial  
9  
10 associations of words during their retrieval such that spatial meanings and physical locations of words  
11  
12 on the computer screen impacted how well they were remembered. In this regard, results seem to  
13  
14 support the claims of embodied semantics that perceptual relations are encoded in the language  
15  
16 (Louwerse, 2008) with consequences for verbal memory. Crucially, these consequences occurred even  
17  
18 though spatial information was neither relevant nor necessary for successful retrieval. In other words,  
19  
20 participants were only instructed to retrieve the words as quickly and as accurately as possible in both  
21  
22 experiments. They were not asked to remember where they saw the words, visualise them or judge  
23  
24 their spatial iconicity. This occurred even though there was a difference between the conditions  
25  
26 manipulated based on the location of the words and their referents. If spatial associations were not  
27  
28 simulated in the current study, there would have been no systematic differences between the  
29  
30 conditions. Further evidence in this direction is the effect of word imageability on perceptual simulation:  
31  
32 In Experiment 1, highly imageable words were retrieved more slowly when there was a visual cue in the  
33  
34 same quadrant as the location of the probe word (i.e., congruent retrieval) regardless of the spatial  
35  
36 arrangement at encoding. The fact that imageability, which is a semantic variable, magnified the  
37  
38 interference effect is important. This suggests that participants did not merely encode the word  
39  
40 locations, as the spatial indexing theories assume but accessed the spatial meanings of words.  
41  
42  
43  
44

45  
46 Metaphorical spatial referents of abstract words such as *liar* had similar effects on how they  
47  
48 were remembered. As the large body of literature on the embodiment of abstractness has indicated  
49  
50 (e.g., Barsalou, 2003), abstract words are also grounded in the sensorimotor system through conceptual  
51  
52 metaphors (e.g., "GOOD IS UP" and "BAD IS DOWN") (Lakoff & Johnson, 1980). We argue that  
53  
54 participants in the current study accessed the spatial part of their meaning which then cued them to  
55  
56  
57  
58  
59  
60

1  
2  
3 certain locations (Experiment 1) or facilitated/hindered their retrieval according to their position on the  
4  
5 screen (Experiment 2). This occurred even though abstract words intuitively have weaker spatial  
6  
7 associations compared to highly imageable, concrete words. Estes et al. (2015) cast doubt on the  
8  
9 perceptual competition account on the grounds that, in their study, words with strong and weak visual  
10  
11 strength and imageability elicited equivalent spatial interference. Our results are at odds with these  
12  
13 results. The results provide evidence that comprehenders perceptually simulate the intrinsic spatial  
14  
15 associations of words. Thus, our data support perceptual, grounded-embodied theories of language  
16  
17 representation (Barsalou, 1999, 2008; Paivio, 1986; Pulvermüller, 1999). That said, we do not argue that  
18  
19 spatial indexing and a compatibility effect should be discarded altogether given the abundant evidence  
20  
21 (Wynn et al., 2019). The argument here is that when analysed in the same experiment (Experiment 1),  
22  
23 the perceptual simulation evoked by presented words is so salient that any compatibility effect due to  
24  
25 spatial indexing is overshadowed. Thus, amodal explanations, such as spatial indexing, are not sufficient  
26  
27 to explain findings in studies showing spatial iconicity.  
28  
29  
30  
31

32 The effect of the sensorimotor simulation in the experiments was systematically in opposite  
33  
34 directions: In Experiment 1, words encoded in incongruent locations were retrieved more accurately  
35  
36 when there was a spatially matching cue. In Experiment 2, however, words encoded in incongruent  
37  
38 locations were retrieved less accurately compared to the control condition. Likewise, words encoded in  
39  
40 congruent locations were retrieved slower in Experiment 1 but they were retrieved faster in Experiment  
41  
42 2 compared to the control condition. These results are in line with the studies investigating the spatial  
43  
44 words in other semantic tasks and reporting match/mismatch advantage under different conditions. As  
45  
46 briefly discussed in the Introduction, tasks in which participants are asked to judge semantic relations  
47  
48 between spatial words result in a match advantage (e.g., Zwaan & Yaxley, 2003a), that is, spatial  
49  
50 congruency typically results in improved accuracy and/or response speed. On the other hand, a  
51  
52 mismatch advantage has been found in tasks in which participants acted on a visual target after reading  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 a spatial word (e.g., Bergen et al., 2007). We argue that this is also what we observed in both  
4  
5 experiments.  
6

7  
8 To be more precise, in Experiment 1 there was a mismatch advantage due to the spatial cueing  
9  
10 effect triggered by reading spatial words and competition between the simulated locations and the  
11  
12 visual cue. To illustrate, reading *bird* gave rise to a simulation of an upward location and seeing a visual  
13  
14 cue in the same position deteriorated retrieval accuracy due to a perceptual competition between the  
15  
16 spatial representation of the simulated location and the cognitive cost of perceiving actual visual  
17  
18 information in the same location. In this regard, results confirm previous studies reporting spatial  
19  
20 interference effect or mismatch advantage due to conceptual cueing effect (Bergen et al., 2007; Estes et  
21  
22 al., 2008; Gozli et al., 2013; Richardson et al., 2003; Verges & Duffy, 2009) and extend the mismatch  
23  
24 advantage to short-term verbal recognition memory, a relatively understudied task in the literature.  
25  
26 However, participants were not faced with such a perceptual competition because there was no visual  
27  
28 cue in Experiment 2 at retrieval. In turn, memory performance was better whenever the physical and  
29  
30 the simulated locations matched at encoding.  
31  
32  
33

34  
35 An issue to consider is why spatial manipulations in the current study affected accuracy and RTs  
36  
37 differently under different conditions. To be more precise, words encoded in congruent locations  
38  
39 slowed down their retrieval in Experiment 1 and words encoded in incongruent locations (with a  
40  
41 congruent cue) were retrieved more accurately. Likewise, results in Experiment 2 indicate that there  
42  
43 was a disadvantage for the mismatched words in accuracy and an advantage for the matched words in  
44  
45 response speed. Further, spatial words were remembered less accurately but faster than neutral words.  
46  
47 Taken together, it seems as the spatial incongruency affected accuracy rates while spatial congruency  
48  
49 affected retrieval times. Some other studies investigating sensorimotor simulation with spatial probes  
50  
51 noted similar discrepancies between accuracy rates and RTs. For example, top targets elicited faster but  
52  
53 less accurate responses than bottom targets in Estes et al. (2015). They attributed the difference to a  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 speed-accuracy tradeoff. However, Pearson correlation of the response times and accuracy scores in the  
4  
5 present study was weak in Experiment 1 (-0.24) and negligible in Experiment 2 (-0.06), ruling out a  
6  
7 speed-accuracy tradeoff. In any case, our results suggest that sensorimotor simulation affects accuracy  
8  
9 and response speed differently in memory, and further research disambiguating the effects of  
10  
11 sensorimotor simulation on memory is necessary.  
12  
13

14           Our results are crucial both for memory and language. Spatial and other sensory attributes of  
15  
16 words are often disregarded when developing verbal material for memory experiments. The current  
17  
18 study demonstrates that these variables are as important as more traditional variables such as  
19  
20 imageability and frequency and might affect how well individuals remember. As to language, we extend  
21  
22 the growing literature of embodiment of language processing. Given the experimental design and the  
23  
24 variety of verbal stimuli in the present study, our results have further implications in favour of modal  
25  
26 explanations over amodal and symbolic accounts. For example, Louwerse (2008) claimed that spatial  
27  
28 iconicity results can be explained by the frequency of linguistic relations between spatial pairs. In  
29  
30 consequence, iconic word pairs (*branch over root*) are processed faster than non-iconic counterparts  
31  
32 (*root over branch*) because language comprehenders are exposed to iconic linguistic symbols more  
33  
34 often (e.g., *branch* more frequently precedes *root* in everyday language use). Similarly, Zwaan and  
35  
36 Yaxley (2003a) speculated that words such as *attic* and *basement* could have links to the concepts of top  
37  
38 and bottom respectively given that attics are typically found in the top parts of the houses. Accordingly,  
39  
40 they argued that when the position of the word *attic* is amodally encoded on the computer screen, it is  
41  
42 tagged with the concept of top. Thus, a conflict between the tag assigned to the word and the contents  
43  
44 of its meaning yields a slower response. Our data speak against these amodal, symbolic accounts of  
45  
46 spatial iconicity. Participants in our study were not asked to judge semantic relations of pairs but to  
47  
48 retrieve a single word among three others, which were controlled based on several predictors of verbal  
49  
50 memory. Given that conformity of a single word location with its meaning influenced how well the word  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 was retrieved from memory, it is unlikely that linguistic relations played a role in our results. Even if  
4  
5 amodal theories could explain the findings for concrete words with literal spatial associations such as  
6  
7 *attic* or *basement*, they fail to explain why retrieval of words with metaphorical spatial meanings such as  
8  
9 *hope* or *chaos* was modulated according to their spatial arrangement. The fact that abstract words  
10  
11 triggered the simulation of metaphorical locations with consequences in memory in the present study  
12  
13 sheds further light on the grounding of abstract words, which is a long-standing challenge for grounded-  
14  
15 embodied cognition (Dove, 2016). Along with that, future studies in non-Western languages with a right-  
16  
17 to-left reading direction are particularly important to generalise the results.  
18  
19

20  
21 To conclude, our results provide experimental support for converging evidence in favour of a  
22  
23 simulation-based account of language representation against amodal, symbolic processes. Thus, in  
24  
25 addition to the growing evidence, we argue that experiential and/or metaphorical spatial associations of  
26  
27 words are perceptually reactivated in a task that is seemingly unrelated to the spatial information.  
28  
29 Moreover, the mediating role of word imageability on the spatial interference effect in our study  
30  
31 endorses a perceptual competition account to explain the spatial interference effect. The unique  
32  
33 contribution of the present study is extending these results to short-term recognition memory involving  
34  
35 language with further implications in support of a simulation account of language based on grounded-  
36  
37 embodied cognition.  
38  
39  
40  
41  
42

#### 43 **Disclosure statement**

44  
45 No potential conflict of interest was reported by the authors.  
46  
47  
48  
49

#### 50 **Data Accessibility Statement**

51  
52 The data and materials from the present experiment are publicly available at the Open Science  
53  
54 Framework website: <https://osf.io/4yjdz> .  
55  
56  
57  
58  
59  
60

## References

- Amer, T., Gozli, D. G., & Pratt, J. (2017). Biasing spatial attention with semantic information: an event coding approach. *Psychological Research*, 1–19. <https://doi.org/10.1007/s00426-017-0867-5>
- Amorim, M., & Pinheiro, A. P. (2019). Is the sunny side up and the dark side down? Effects of stimulus type and valence on a spatial detection task. *Cognition and Emotion*, 33(2), 346–360. <https://doi.org/10.1080/02699931.2018.1452718>
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(4), 390–412. <https://doi.org/10.1016/j.jml.2007.12.005>
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255–278. <https://doi.org/10.1016/j.jml.2012.11.001>
- Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, 22(4), 577–609; discussion 610–660. <https://doi.org/10.1017/S0140525X99252144>
- Barsalou, L. W. (2003). Abstraction in perceptual symbol systems. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 358(1435), 1177–1187. <https://doi.org/10.1098/rstb.2003.1319>
- Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, 59, 617–645. <https://doi.org/10.1146/annurev.psych.59.103006.093639>
- Bates, D. M., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Bergen, B. K. (2016). Embodiment, simulation and meaning. In N. Riemer (Ed.), *The Routledge handbook of semantics* (pp. 142–156). Routledge. <https://doi.org/10.4324/9781315685533.ch8>
- Bergen, B. K., Lindsay, S., Matlock, T., & Narayanan, S. (2007). Spatial and linguistic aspects of visual



1  
2  
3           imagery in sentence comprehension. *Cognitive Science*, 31(5), 733–764.

4  
5           <https://doi.org/10.1080/03640210701530748>

6  
7  
8           Berndt, E., Dudschig, C., Miller, J., & Kaup, B. (2019). A replication attempt of hemispheric differences in  
9           semantic-relatedness judgments (Zwaan & Yaxley, 2003). *Acta Psychologica*, 198, Article 102871.

10  
11           <https://doi.org/10.1016/j.actpsy.2019.102871>

12  
13  
14           Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally  
15           known English word lemmas. *Behavior Research Methods*, 46(3), 904–911.

16  
17           <https://doi.org/10.3758/s13428-013-0403-5>

18  
19           Chasteen, A. L., Burdzy, D. C., & Pratt, J. (2010). Thinking of God moves attention. *Neuropsychologia*,  
20           48(2), 627–630. <https://doi.org/10.1016/j.neuropsychologia.2009.09.029>

21  
22           Chomsky, N. (1957). *Syntactic structures*. Mouton de Gruyter.

23  
24           Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized  
25           assessment instruments in psychology. *Psychological Assessment*, 6(4), 284–290.

26  
27           <https://doi.org/10.1037/1040-3590.6.4.284>

28  
29           Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological  
30           research. *Journal of Verbal Learning and Verbal Behavior*, 12(4), 335–359.

31  
32           [https://doi.org/10.1016/S0022-5371\(73\)80014-3](https://doi.org/10.1016/S0022-5371(73)80014-3)

33  
34           Clark, J. M., & Paivio, A. (2004). Extensions of the Paivio, Yuille, and Madigan (1968) norms. *Behavior*  
35           *Research Methods, Instruments, and Computers*, 36(3), 371–383.

36  
37           <https://doi.org/10.3758/BF03195584>

38  
39           Conway, A. R., & Engle, R. W. (1996). Individual differences in working memory capacity: more evidence  
40           for a general capacity theory. *Memory*, 4(6), 577–590. <https://doi.org/10.1080/741940997>

41  
42           Corballis, M. C. (2003). From mouth to hand: Gesture, speech, and the evolution of right-handedness.  
43           *Behavioral and Brain Sciences*, 26(2), 199–208. <https://doi.org/10.1017/S0140525X03000062>

1  
2  
3 Corsi, P. (1972). *Human memory and the medial temporal lobe region of the brain* [Unpublished doctoral  
4 dissertation]. McGill University.

5  
6  
7 Dove, G. (2016). Three symbol ungrounding problems: Abstract concepts and the future of embodied  
8 cognition. *Psychonomic Bulletin & Review*, 23(4), 1109–1121. [https://doi.org/10.3758/s13423-015-](https://doi.org/10.3758/s13423-015-0825-4)  
9  
10  
11  
12 0825-4

13  
14 Dudschig, C., Souman, J., Lachmair, M., de la Vega, I., & Kaup, B. (2013). Reading “sun” and looking up:  
15 The influence of language on saccadic eye movements in the vertical dimension. *PLoS ONE*, 8(2),  
16  
17 e56872. <https://doi.org/10.1371/journal.pone.0056872>

18  
19  
20  
21 Estes, Z., & Barsalou, L. W. (2018). A comprehensive meta-analysis of spatial interference from linguistic  
22 cues: Beyond Petrova et al. (2018). *Psychological Science*, 29(9), 1558–1564.  
23  
24  
25  
26 <https://doi.org/10.1177/0956797618794131>

27  
28 Estes, Z., Verges, M., & Adelman, J. S. (2015). Words, objects, and locations: Perceptual matching  
29 explains spatial interference and facilitation. *Journal of Memory and Language*, 84, 167–189.  
30  
31  
32  
33 <https://doi.org/10.1016/j.jml.2015.06.002>

34  
35 Estes, Z., Verges, M., & Barsalou, L. W. (2008). Head up, foot down: Object words orient attention to the  
36 objects' typical location. *Psychological Science*, 19(2), 93–97. [https://doi.org/10.1111/j.1467-](https://doi.org/10.1111/j.1467-9280.2008.02051.x)  
37  
38  
39 9280.2008.02051.x

40  
41 Fodor, J. A. (1975). *The language of thought*. Harvard University Press.

42  
43 Franklin, N., & Tversky, B. (1990). Searching imagined environments. *Journal of Experimental*  
44  
45  
46  
47 *Psychology: General*, 119(1), 63–76. <https://doi.org/10.1037/0096-3445.119.1.63>

48  
49 Giessner, S. R., & Schubert, T. W. (2007). High in the hierarchy: How vertical location and judgments of  
50 leaders' power are interrelated. *Organizational Behavior and Human Decision Processes*, 104(1),  
51  
52  
53 30–44. <https://doi.org/10.1016/j.obhdp.2006.10.001>

54  
55 Gorman, A. M. (1961). Recognition memory for nouns as a function of abstractness and frequency.

- 1  
2  
3 *Journal of Experimental Psychology*, 61(1), 23–29. <https://doi.org/10.1037/h0040561>
- 4  
5  
6 Gozli, D. G., Chasteen, A. L., & Pratt, J. (2013). The cost and benefit of implicit spatial cues for visual  
7  
8 attention. *Journal of Experimental Psychology: General*, 142(4), 1028–1046.  
9  
10 <https://doi.org/10.1037/a0030362>
- 11  
12 Gozli, D. G., Pratt, J., Martin, K. Z., & Chasteen, A. L. (2016). Implied spatial meaning and visuospatial  
13  
14 bias: Conceptual processing influences processing of visual targets and distractors. *PLoS ONE*,  
15  
16 11(3), e0150928. <https://doi.org/10.1371/journal.pone.0150928>
- 17  
18  
19 Hallgren, K. A. (2012). Computing Inter-rater reliability for observational data: An overview and tutorial.  
20  
21 *Tutorials in Quantitative Methods for Psychology*, 8(1), 23–34.  
22  
23 <https://doi.org/10.20982/tqmp.08.1.p023>
- 24  
25  
26 Hunt, R. R., & Worthen, J. B. (Eds.). (2006). *Distinctiveness and memory*. Oxford University Press.
- 27  
28  
29 Hutchinson, S., Wei, L., & Louwerse, M. M. (2013). Avoiding the language-as-a-fixed-effect fallacy: How  
30  
31 to estimate outcomes of linear mixed models. In P. Bello, M. Guarini, M. McShane, & B. Scassellati  
32  
33 (Eds.), *Proceedings of the 36th Annual Conference of the Cognitive Science Society* (pp. 2393–2398).  
34  
35 Cognitive Science Society.
- 36  
37  
38 Johansson, R., & Johansson, M. (2014). Look here, eye movements play a functional role in memory  
39  
40 retrieval. *Psychological Science*, 25(1), 236–242. <https://doi.org/10.1177/0956797613498260>
- 41  
42  
43 Johns, M. W. (1991). A new method for measuring daytime sleepiness: The Epworth sleepiness scale.  
44  
45 *Sleep*, 14(6), 540–545. <https://doi.org/10.1016/j.sleep.2007.08.004>
- 46  
47  
48 Kessels, R. P. C., Zandvoort, M. J. E. Van, Postma, A., Kappelle, L. J., & Haan, E. H. F. De. (2000). The Corsi  
49  
50 block-tapping task: Standardization and normative data. *Applied Neuropsychology*, 7(4), 252–258.  
51  
52 [https://doi.org/10.1207/S15324826AN0704\\_8](https://doi.org/10.1207/S15324826AN0704_8)
- 53  
54  
55 Kumcu, A., & Thompson, R. L. (2016). Spatial interference and individual differences in looking at  
56  
57 nothing for verbal memory. In A. Papafragou, D. Grodner, D. Mirman, & J. C. Trueswell (Eds.),  
58  
59  
60

1  
2  
3 *Proceedings of the 38th Annual Conference of the Cognitive Science Society* (pp. 2387–2392).

4  
5 Cognitive Science Society.

6  
7 Kumcu, A., & Thompson, R. L. (2020). Less imageable words lead to more looks to blank locations during  
8  
9 memory retrieval. *Psychological Research*, *84*, 667–684. [https://doi.org/10.1007/s00426-018-1084-](https://doi.org/10.1007/s00426-018-1084-6)  
10  
11  
12 6

13  
14 Kuznetsova, A., Brockhoff, P., & Christensen, R. (2016). lmerTest: Tests in linear mixed effects models. R  
15  
16 package version 3.0.0. <https://cran.r-project.org/package=lmerTest>.

17  
18 Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. University of Chicago Press.

19  
20 Leitan, N. D., Williams, B., & Murray, G. (2015). Look up for healing: Embodiment of the heal concept in  
21  
22 looking upward. *PLoS ONE*, *10*(7), e0132427. <https://doi.org/10.1371/journal.pone.0132427>

23  
24 Lezak, M. D., Howieson, D. B., Bigler, E. D., & Tranel, D. (2012). *Neuropsychological assessment* (5th ed.).  
25  
26 Oxford University Press.

27  
28 Li, P., Zhang, F., Tsai, E., & Puls, B. (2013). Language history questionnaire (LHQ 2.0): A new dynamic  
29  
30 web-based research tool. *Bilingualism: Language and Cognition*, *17*(3), 673–680.  
31  
32 <https://doi.org/10.1017/S1366728913000606>

33  
34 Louwerse, M. M. (2008). Embodied relations are encoded in language. *Psychonomic Bulletin & Review*,  
35  
36 *15*(4), 838–844. <https://doi.org/10.3758/PBR.15.4.838>

37  
38 Marmolejo-Ramos, F., Elosúa, M. R., Yamada, Y., Hamm, N. F., & Noguchi, K. (2013). Appraisal of space  
39  
40 words and allocation of emotion words in bodily space. *PLoS ONE*, *8*(12), 18–20.  
41  
42 <https://doi.org/10.1371/journal.pone.0081688>

43  
44 Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment  
45  
46 builder for the social sciences. *Behavior Research Methods*, *44*(2), 314–324.  
47  
48 <https://doi.org/10.3758/s13428-011-0168-7>

49  
50 Meyer, B. P., & Robinson, M. D. (2004). Why the sunny side is up: Associations between affect and  
51  
52

1  
2  
3 vertical position. *Psychological Science*, 15(4), 243–247. <https://doi.org/10.1111/j.0956->

4  
5 7976.2004.00659.x

6  
7 Montoro, P. R., Contreras, M. J., Elosúa, M. R., & Marmolejo-Ramos, F. (2015). Cross-modal

8  
9 metaphorical mapping of spoken emotion words onto vertical space. *Frontiers in Psychology*,

10  
11 6(August), 1–13. <https://doi.org/10.3389/fpsyg.2015.01205>

12  
13 Morey, R. D., & Rouder, J. N. (2015). BayesFactor: Computation of Bayes factors for common designs. R

14  
15 package version 0.9.10-1. <http://cran.r-project.org/package=BayesFactor>

16  
17 Mueller, S. T., & Piper, B. J. (2014). The Psychology Experiment Building Language (PEBL) and PEBL Test

18  
19 Battery. *Journal of Neuroscience Methods*, 222, 250–259.

20  
21 <https://doi.org/10.1016/j.jneumeth.2013.10.024>

22  
23 Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh Inventory.

24  
25 *Neuropsychologia*, 9, 97–113. [https://doi.org/10.1016/0028-3932\(71\)90067-4](https://doi.org/10.1016/0028-3932(71)90067-4)

26  
27 Ostarek, M., & Vigliocco, G. (2017). Reading sky and seeing a cloud: On the relevance of events for

28  
29 perceptual simulation. *Journal of Experimental Psychology: Learning Memory and Cognition*, 43(4),

30  
31 579–590. <https://doi.org/10.1037/xlm0000318>

32  
33 Paivio, A. (1986). *Mental representation: A dual coding approach*. Oxford University Press.

34  
35 Paivio, A., Yuille, J. C., & Madigan, S. A. (1968). Concreteness, imagery, and meaningfulness values for

36  
37 925 nouns. *Journal of Experimental Psychology*, 76(1, Pt.2). 1-25.

38  
39 <https://doi.org/10.1037/h0025327>

40  
41 Pulvermüller, F. (1999). Words in the brain's language. *Behavioral and Brain Sciences*, 22(2), 253–279;

42  
43 discussion 280-336. <https://doi.org/10.1017/S0140525X9900182X>

44  
45 Pylyshyn, Z. W. (1986). *Meaning and cognitive structure: Issues in the computational theory of mind*.

46  
47 Ablex.

48  
49 Pylyshyn, Z. W. (1989). The role of location indexes in spatial perception: A sketch of the FINST spatial-

- 1  
2  
3 index model. *Cognition*, 32(1), 65–97. [https://doi.org/10.1016/0010-0277\(89\)90014-0](https://doi.org/10.1016/0010-0277(89)90014-0)
- 4  
5 Quadflieg, S., Etzel, J. A., Gazzola, V., Keysers, C., Schubert, T. W., Waiter, G. D., & Macrae, C. N. (2011).  
6  
7 Puddles, parties, and professors: linking word categorization to neural patterns of visuospatial  
8  
9 coding. *Journal of Cognitive Neuroscience*, 23(10), 2636–2649.  
10  
11  
12 <https://doi.org/10.1162/jocn.2011.21628>
- 13  
14 R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for  
15  
16 Statistical Computing. <https://www.r-project.org/>.
- 17  
18 Richardson, D. C., & Spivey, M. J. (2000). Representation, space and Hollywood Squares: Looking at  
19  
20 things that aren't there anymore. *Cognition*, 76(3), 269–295. <https://doi.org/10.1016/S0010->  
21  
22 [0277\(00\)00084-6](https://doi.org/10.1016/S0010-0277(00)00084-6)
- 23  
24 Richardson, D. C., Spivey, M. J., Barsalou, L. W., & McRae, K. (2003). Spatial representations activated  
25  
26 during real-time comprehension of verbs. *Cognitive Science*, 27(5), 767–780.  
27  
28  
29 [https://doi.org/10.1016/S0364-0213\(03\)00064-8](https://doi.org/10.1016/S0364-0213(03)00064-8)
- 30  
31 Scholz, A., Mehlhorn, K., & Krems, J. F. (2016). Listen up, eye movements play a role in verbal memory  
32  
33 retrieval. *Psychological Research*, 80(1), 149–158. <https://doi.org/10.1007/s00426-014-0639-4>
- 34  
35 Schubert, T. W. (2005). Your highness: Vertical positions as perceptual symbols of power. *Journal of*  
36  
37 *Personality and Social Psychology*, 89(1), 1–21. <https://doi.org/10.1037/0022-3514.89.1.1>
- 38  
39 Šetić, M., & Domijan, D. (2007). The influence of vertical spatial orientation on property verification.  
40  
41  
42  
43 *Language and Cognitive Processes*, 22(2), 297–312. <https://doi.org/10.1080/01690960600732430>
- 44  
45 Spivey, M. J., Richardson, D. C., & Fitneva, S. A. (2004). Thinking outside the brain: Spatial indices to  
46  
47 visual and linguistic information. In J. Henderson & F. Ferreira (Eds.), *Interfacing language, vision,*  
48  
49 *and action: Eye movements and the visual world* (pp. 161–190). Psychology Press.  
50  
51  
52 <https://doi.org/10.4324/9780203488430>
- 53  
54 Stanfield, R. A., & Zwaan, R. A. (2001). The effect of implied orientation derived from verbal context on  
55  
56  
57  
58  
59  
60

- 1  
2  
3 picture recognition. *Psychological Science*, 12(2), 153–156. <https://doi.org/10.1111/1467->  
4  
5 9280.00326  
6  
7 Taylor, J. E. T., Lam, T. K., Chasteen, A. L., & Pratt, J. (2015). Bow your head in shame, or, hold your head  
8  
9 up with pride: Semantic processing of self-esteem concepts orients attention vertically. *PLoS ONE*,  
10  
11 10(9), e0137704. <https://doi.org/10.1371/journal.pone.0137704>  
12  
13  
14 van Ede, F., de Lange, F. P., & Maris, E. (2012). Attentional cues affect accuracy and reaction time via  
15  
16 different cognitive and neural processes. *Journal of Neuroscience*, 32(30), 10408–10412.  
17  
18 <https://doi.org/10.1523/JNEUROSCI.1337-12.2012>  
19  
20 Verges, M., & Duffy, S. (2009). Spatial representations elicit dual-coding effects in mental imagery.  
21  
22 *Cognitive Science*, 33(6), 1157–1172. <https://doi.org/10.1111/j.1551-6709.2009.01038.x>  
23  
24 Winter, B. (2013). *Linear models and linear mixed effects models in R with linguistic applications*.  
25  
26 <http://arxiv.org/abs/1308.5499>  
27  
28 Wynn, J. S., Shen, K., & Ryan, J. D. (2019). Eye movements actively reinstate spatiotemporal mnemonic  
29  
30 content. *Vision*, 3(2), 21. <https://doi.org/10.3390/vision3020021>  
31  
32  
33 Zanolie, K., Dantzig, S. van, Boot, I., Wijnen, J., Schubert, T. W., Giessner, S. R., & Pecher, D. (2012).  
34  
35 Mighty metaphors: Behavioral and ERP evidence that power shifts attention on a vertical  
36  
37 dimension. *Brain and Cognition*, 78(1), 50–58. <https://doi.org/10.1016/j.bandc.2011.10.006>  
38  
39  
40 Zwaan, R. A., & Madden, C. J. (2005). Embodied sentence comprehension. In D. Pecher & R. A. Zwaan  
41  
42 (Eds.), *Grounding cognition: The role of perception and action in memory, language, and thinking*  
43  
44 (pp. 224–245). <https://doi.org/10.1017/CBO9780511499968.010>  
45  
46  
47 Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory.  
48  
49 *Psychological Bulletin*, 123(2), 162–185. <https://doi.org/10.1037/0033-2909.123.2.162>  
50  
51  
52 Zwaan, R. A., Stanfield, R., & Yaxley, R. (2002). Language comprehenders mentally represent the shapes  
53  
54 of objects. *Psychological Science*, 13(2), 168–171. <https://doi.org/10.1111/1467-9280.00430>  
55  
56  
57  
58  
59  
60

1  
2  
3 Zwaan, R. A., & Yaxley, R. H. (2003a). Hemispheric differences in semantic-relatedness judgments.  
4

5 *Cognition*, 87, B79–B86. [https://doi.org/10.1016/S0010-0277\(02\)00235-4](https://doi.org/10.1016/S0010-0277(02)00235-4)  
6

7 Zwaan, R. A., & Yaxley, R. H. (2003b). Spatial iconicity affects semantic relatedness judgments.  
8

9 *Psychonomic Bulletin & Review*, 10(4), 954–958. <https://doi.org/10.3758/BF03196557>  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

Peer Review Version



### Figure Captions

**Figure 1.** The Heat Map Shows the Distribution of Word Location Ratings over Positions on the Horizontal and Vertical Axis

Note. Lighter colours indicate lower frequency and darker colours indicate higher frequency. Note that the blank area outside the pattern does not refer to an absence of ratings, but rating frequency under a certain threshold (see supplemental online materials for an interactive scatterplot showing all word locations on the horizontal and vertical axes).

**Figure 2.** A Schematic Illustration of the Temporal Order of Events in an Example Trial Showing Two Different Encoding and Two Different Retrieval Conditions in Experiment 1

Note. A schematic illustration of the temporal order of events in an example trial showing two different encoding and two different retrieval conditions in Experiment 1. In this example, the relevant quadrant is the top left location, where the probe word (i.e., BIRD) appears. Feedback on this diagram (Correct 1200 ms) is given as an example.

**Figure 3.** Accuracy and RTs Between Four Combinations of Encoding and Retrieval Conditions in Experiment 1

Note. Violin plots show the distribution of accuracy and RTs per condition. The box plots within the violins show the whole range (vertical line), the mean (black square), and the median (horizontal line). Grey dots represent data points for individual participants. (CE + CR: Congruent Encoding and Congruent Retrieval, CE + IR: Congruent Encoding and Incongruent Retrieval, IE + CR: Incongruent Encoding and Congruent Retrieval, IE + IR: Incongruent Encoding and Incongruent Retrieval). \*  $p \leq .05$  \*\*  $p \leq .01$ .

1  
2  
3 **Figure 4.** Accuracy and RTs across Word Types and Encoding Conditions in Experiment 2

4  
5 Note. Violin plots show the distribution of accuracy and RTs per word type and condition. The box plots  
6  
7 within the violins show the whole range (vertical line), the mean (black square), and the median  
8  
9 (horizontal line). Grey dots represent data points for individual participants. \*  $p \leq .05$

10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

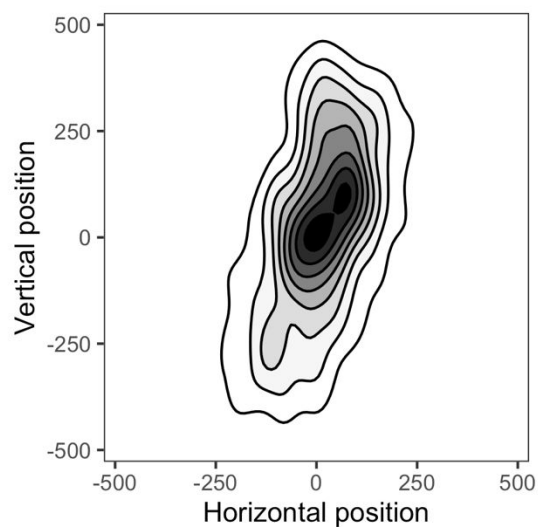
Peer Review Version

## REMEMBERING SPATIAL WORDS

1

**Figure 1**

*The Heat Map Shows the Distribution of Word Location Ratings over Positions on the Horizontal and Vertical Axis*



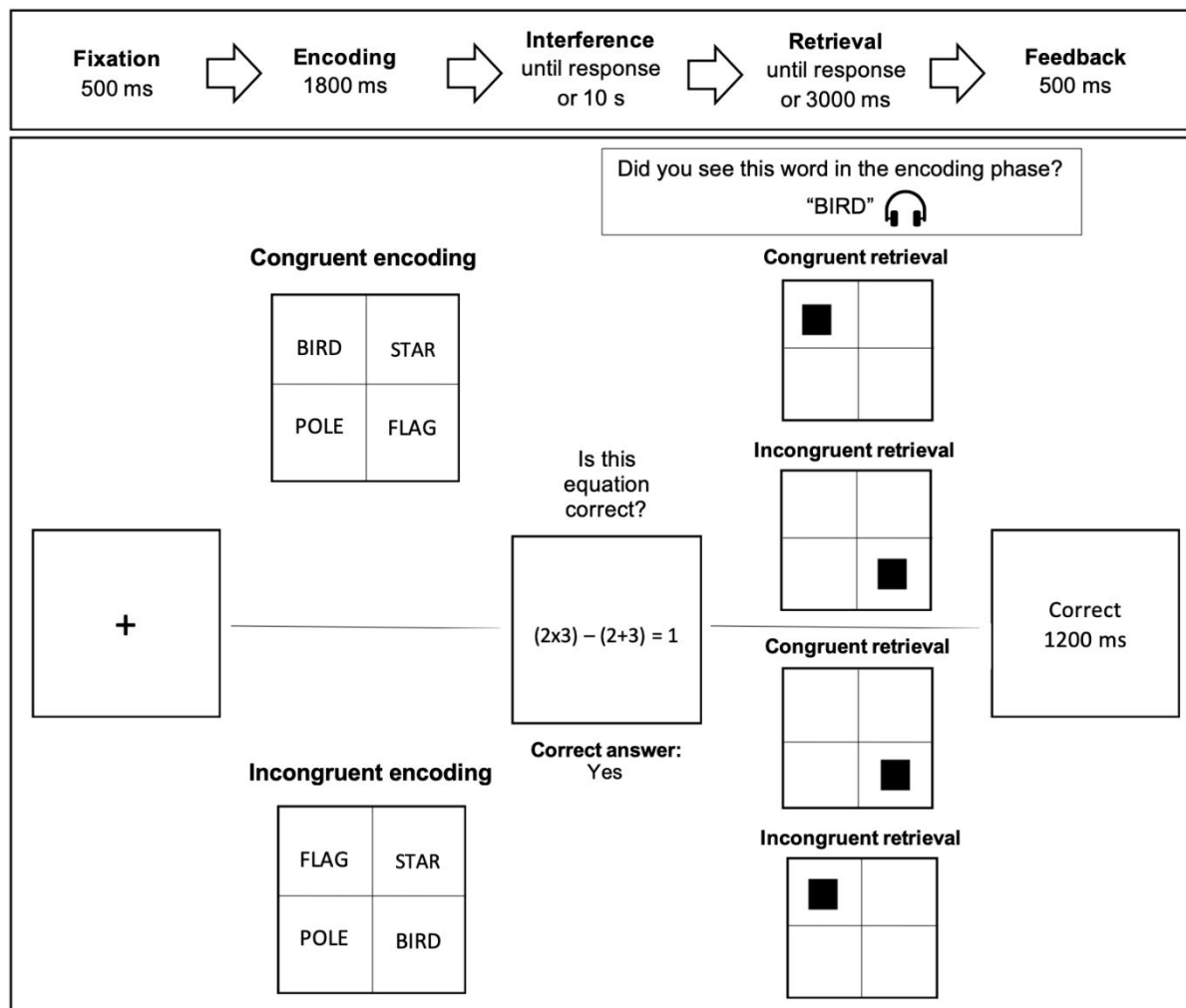
*Note.* Lighter colours indicate lower frequency and darker colours indicate higher frequency. Note that the blank area outside the pattern does not refer to an absence of ratings, but rating frequency under a certain threshold (see supplemental online materials for an interactive scatterplot showing all word locations on the horizontal and vertical axes).

## REMEMBERING SPATIAL WORDS

1

**Figure 2**

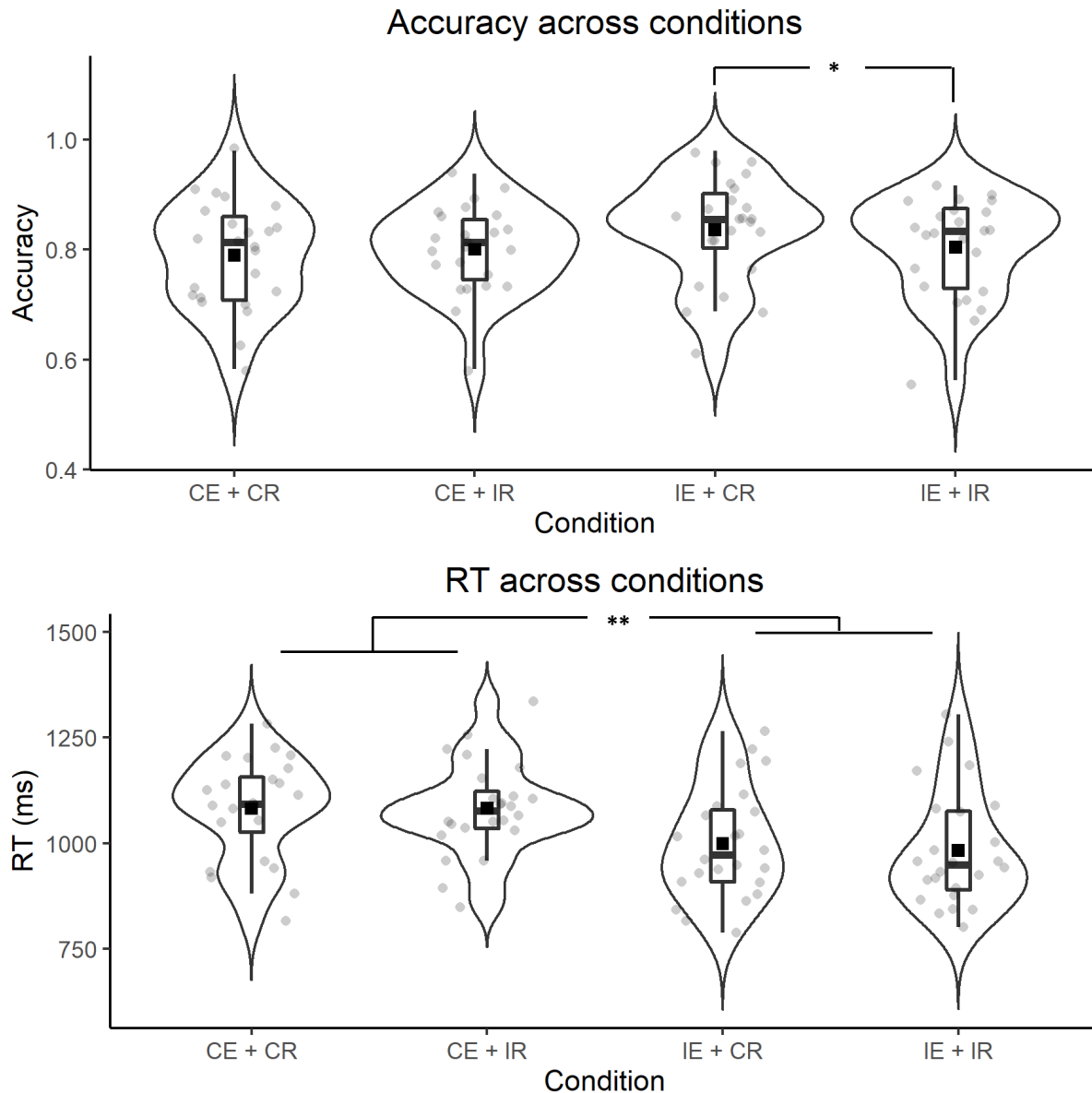
A Schematic Illustration of the Temporal Order of Events in an Example Trial Showing Two Different Encoding and Two Different Retrieval Conditions in Experiment 1



Note. A schematic illustration of the temporal order of events in an example trial showing two different encoding and two different retrieval conditions in Experiment 1. In this example, the relevant quadrant is the top left location, where the probe word (i.e., BIRD) appears. Feedback on this diagram (Correct 1200 ms) is given as an example.

**Figure 3**

*Accuracy and RTs Between Four Combinations of Encoding and Retrieval Conditions in Experiment 1*



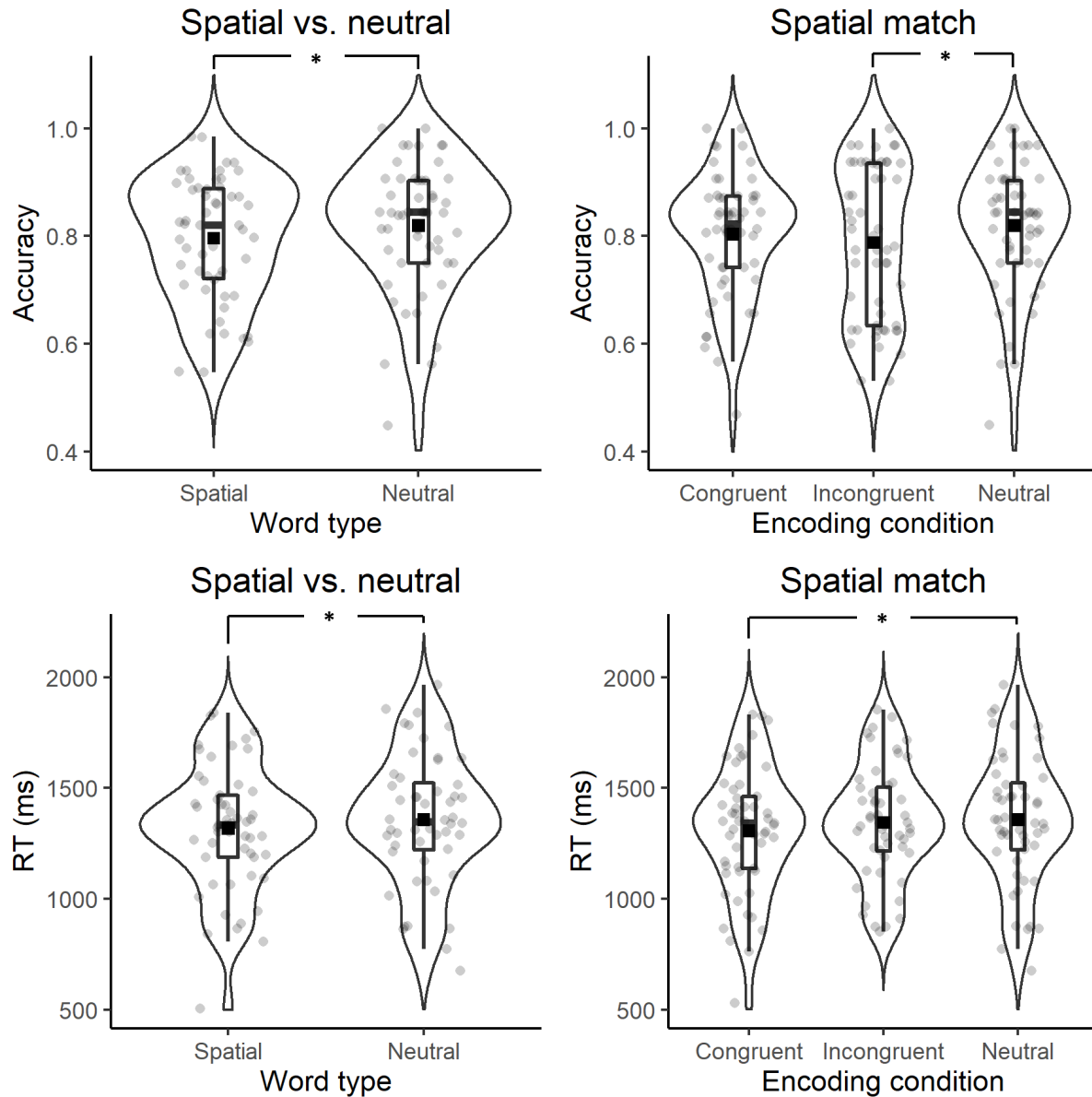
*Note.* Violin plots show the distribution of accuracy and RTs per condition. The box plots within the violins show the whole range (vertical line), the mean (black square), and the median (horizontal line). Grey dots represent data points for individual participants. (CE + CR: Congruent Encoding and Congruent Retrieval, CE + IR: Congruent Encoding and Incongruent Retrieval, IE + CR: Incongruent Encoding and Congruent Retrieval, IE + IR: Incongruent Encoding and Incongruent Retrieval). \*  $p \leq .05$  \*\*  $p \leq .01$ .

## REMEMBERING SPATIAL WORDS

1

Figure 4

Accuracy and RTs across Word Types and Encoding Conditions in Experiment 2



Note. Violin plots show the distribution of accuracy and RTs per word type and condition. The box plots within the violins show the whole range (vertical line), the mean (black square), and the median (horizontal line). Grey dots represent data points for individual participants. \*  $p \leq .05$

## REMEMBERING SPATIAL WORDS

1

**Table 1***Descriptive Measures (Mean and SD) of the Probe Words in Positive Trials of Experiment 1*

Variable	Upward words	Downward words	Total
Number	48	48	96
Imageability	4.84 (1.41)	5.02 (1.28)	4.93 (1.34)
Concreteness	4.19 (1.91)	4.67 (2)	4.43 (1.96)
Length in letters	7.02 (1.83)	6.5 (1.75)	6.76 (1.80)
Syllable length	2.33 (1)	2.12 (0.91)	2.23 (0.96)
Horizontal position	127.8 (114.85)	-118.88 (143.5)	4.46 (179.13)
Vertical position	308.6 (73.62)	-301.2 (98.22)	3.7 (318.43)
Spatiality	0.58 (0.13)	0.61 (0.15)	0.59 (0.14)
Spatial variability	0.33 (0.11)	0.35 (0.1)	0.34 (0.11)

## REMEMBERING SPATIAL WORDS

1

**Table 2***Results of the Individual Variability Measures (Mean and SD) Among All Participants and Across Encoding**Groups*

Measures	All	Congruent encoding	Incongruent encoding	<i>p</i>
Corsi block span	58.94 (24.25)	53.75 (25.01)	64.12 (22.81)	.140
Digit span	9.15 (2.41)	9.62 (2.58)	8.67 (2.16)	.170
Sleepiness	0.84 (0.55)	0.85 (0.51)	0.83 (0.6)	.897

*Note.* The average score for healthy adults in the Corsi block-tapping task is 55.7 (Kessels et al., 2000)

and 8 – 10 in forward digit span test (Lezak et al., 2012). The normal range of daytime sleepiness is 0 –

10 (Johns, 1991).



## REMEMBERING SPATIAL WORDS

1

**Table 3***Factor-Loadings of the Predictors*

Predictors	Factor			
	Imagery	Length	Position	h2
Imageability	<b>0.93</b>	-0.21	-0.02	0.92
Concreteness	<b>0.95</b>	-0.15	-0.06	0.92
Length in letters	-0.19	<b>0.93</b>	0.11	0.91
Number of syllables	-0.17	<b>0.94</b>	0.05	0.91
Horizontal position	-0.02	0.09	<b>0.92</b>	0.85
Vertical position	-0.05	0.06	<b>0.92</b>	0.86
Factor statistics				
Eigenvalue	2.62	1.60	1.13	7.99
Variance (%)	30.55	30.28	28.54	89.93

## REMEMBERING SPATIAL WORDS

1

**Table 4***Descriptive Measures (Mean and SD) of Memory Performance in Experiment 1*

Condition	Accuracy (%)	RT (ms)
Incongruent encoding + congruent retrieval	84 (0.37)	993.53 (252.77)
Incongruent encoding + incongruent retrieval	80 (0.40)	977.34 (238.54)
Congruent encoding + incongruent retrieval	80 (0.40)	1077.92 (254.31)
Congruent encoding + congruent retrieval	79 (0.41)	1072.54 (250.95)

## REMEMBERING SPATIAL WORDS

1

**Table 5***Descriptive Measures (Mean and SD) of the Probe Words in Positive Trials of Experiment 2*

Variable	Spatial words	Upward words	Downward words	Neutral words
Number	64	32	32	32
Imageability	4.90 (1.27)	4.73 (1.36)	5.08 (1.16)	4.98 (1.06)
Word length	6.72 (1.54)	7.00 (1.68)	6.44 (1.34)	6.72 (1.25)
Emotionality	4.12 (1.2)	4.34 (1.18)	3.90 (1.2)	2.94 (0.93)
Horizontal position	-4.93 (170.40)	131.60 (95.85)	-141.46 (106.48)	24.64 (71.02)
Vertical position	-2.81 (315.07)	303.14 (70.11)	-308.76 (59.86)	1.89 (71.26)
Spatiality	0.58 (0.09)	0.56 (0.15)	0.60 (0.16)	0.15 (0.09)
Spatial variability	0.34 (0.11)	0.32 (0.1)	0.35 (0.11)	0.39 (0.14)

## REMEMBERING SPATIAL WORDS

1

**Table 6***Descriptive Measures (Mean and SD) of Memory Performance in Experiment 2*

Spatial match condition	Accuracy (%)	RT (ms)
Spatial words	80 (0.40)	1301.07 (457.31)
Congruent encoding	80 (0.40)	1290.57 (461.14)
Incongruent encoding	79 (0.41)	1311.75 (453.31)
Neutral words	82 (0.38)	1328.91 (464.82)

Peer Review Version