Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Words move the eyes: Spatial cueing without targets

Samuel Shaki

Samuel_shaki@hotmail.com

Ariel University

Bar Assia

Ariel University

Oria Pitem

Ariel University

Martin H. Fischer

University of Potsdam

Article

Keywords: attentional shift, semantic cueing effect, eye movements

Posted Date: November 4th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11307251/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Semantic information can influence spatial attention, but it remains unclear whether spatially meaningful words can directly orient overt attention without a subsequent target or response requirement. The present study investigated whether explicit and implicit spatial words induce spontaneous gaze shifts before target presentation and whether these shifts reflect bottom-up word meaning or top-down task instructions. Across two experiments, 99 native Hebrew-speaking adults completed a go/no-go spatial cueing task while eye movements were recorded. Experiment 1 used explicitly spatial words (e.g., “up,” “down”), whereas Experiment 2 used words implicitly associated with spatial locations (e.g., “sky,” “carpet”). In go trials, both word types produced spatial congruency effects in reaction times and progressively larger gaze shifts toward semantically associated locations with increasing cue-target intervals. Critically, no-go trials demonstrated meaning-driven gaze shifts even without required responses or targets, particularly for explicitly spatial cues. Overall, the findings indicate that semantic processing alone can automatically bias overt visual attention, with linguistic meaning exerting a predominantly bottom-up influence on gaze allocation.

INTRODUCTION

Understanding the relationship between the comprehension of symbols and our resulting behavior has implications for a wide range of cognitive tasks, from language processing to interface design. Consequently, the specific example of how linguistic meaning influences spatial attention allocation has been a subject of long-standing debate in cognitive psychology (e.g., [1, 2]). Words that *explicitly* denote space (e.g., "UP", "DOWN") reliably direct attention in a direction that is congruent with their meaning. In contrast, words that only *implicitly* connote spatial regions by referring to objects that occupy typical spatial locations (e.g., "ROOF", "CARPET") present a more complex semantic cueing condition that has yielded inconsistent results (review in [3]).

Researchers interested in understanding this semantic cueing mechanism have heavily relied on Posner’s [4] attentional cueing paradigm. In this paradigm, shifts of attention are inferred from manual reaction times (RTs) to peripherally presented targets that follow explicitly or implicitly spatial cues, such as directed arrows or spatial prepositions. The delay between cue and target is manipulated to investigate the temporal dynamics of the semantic cueing mechanism. The semantic cueing effect can be computed as the RT difference between spatially incongruent conditions (*invalid* trials: target appears at a location opposite from the location implied by the meaning of the cue) and spatially congruent conditions (*valid* trials: target appears at a location implied by the meaning of the cue). The resulting difference value is typically positive and interpreted as additional time needed to re-allocate one’s attentional resource across space during invalid trials. This rationale applies equally to *overt* attention allocation (performed by moving the eyes or other body parts in the expected direction of the target) and *covert* attention allocation (when there are no overt body movements).

Initially, the semantic cueing effect was thought to depend on cognitive processing demands of both the linguistic cue and the subsequent target ([5]; see review by [6]). More recently, Shaki and Fischer [7, 8] deconstructed the cue-target relationship by adopting a Go/No-go task. They instructed participants to respond to targets only if the preceding cues were real English words (vs. nonwords), or only if they had an upward (vs. downward) spatial connotation. Comparing performance across instruction conditions and stimulus onset asynchronies (SOAs) between cue and target, Shaki and Fischer [7, 8] demonstrated that semantic congruity effects from implicitly spatial words only emerged when the go rule required participants to process *both* semantic attributes of the cue and specific features (such as location or identity) of the target. If participants could fully ignore either the cue meaning or the target attributes, semantic cueing effects vanished, regardless of SOA. This suggested that the mechanism responsible for semantic cueing of attention required the combined, multi-stage processing of both the linguistic cue and the visual target.

However, this suggestion was recently challenged by Shaki et al. [9], who investigated whether deep (semantic) processing of the *cue alone* might suffice to trigger spatial shifts of attention. In a simple target detection task, the authors varied the level of semantic processing depth of implicitly spatial lexical cues across three task instructions - lexical decision, non-spatial categorization, and spatial categorization. Shaki et al. [9] demonstrated that differential target processing was not mandatory: When participants processed the cue words deeply enough to extract their spatial connotations, robust semantic cueing effects emerged in manual detection speed. Crucially, the authors paired these indirect chronometric measurements of covert attention allocation with eye-tracking to also capture overt attention deployment directly. They discovered that both spatial and non-spatial categorization of implicitly spatial lexical cues triggered spontaneous vertical gaze shifts toward the implied spatial locations before detection targets even appeared.

Shaki et al.'s [9] recent observations provided direct evidence that cognitive activation of word meaning *alone* modulates semantic cueing effect. Despite this insight into minimally necessary cognitive processes for semantic cueing of attention, a fundamental question raised in the Discussion of Shaki et al.'s [9] report still remained unanswered: Is mere exposure to spatially meaningful semantic cues alone sufficient to move the eyes, or is a subsequent spatial task or spatial target presentation necessary to generate these overt attentional shifts?

To answer this question, we examined separately participants' performance in go-trials and no-go trials. First, we inspected go-trials to see how the novel *overt* attention signature in this task [9] aligns with our earlier findings [7, 8] on explicit vs implicit cueing. Secondly and crucially, we focused on no-go trials where participants do not have to respond. This informs us about the importance of the response for semantic cueing effects and answers our question.

Moreover, we aimed to discover whether participants' eye gaze reflects the task instruction (top-down) or the current cue's meaning (bottom up). Note that no-go trials are diagnostic to understand these ingredients of the congruency effect: In go-trials both the rule and the cue convey the same direction,

resulting in a compound effect on target detection. In contrast, no-go trials dissociate the two directions. This reveals the origin of the overt attentional component underlying the congruency effect.

To summarize, by monitoring spontaneous eye movements following spatial lexical cueing in go and no-go trials, we examined whether gaze habitually maps meaning onto space purely driven by internal semantic processing. This approach determined whether linguistic concepts possess the power to orient visual attention in space, even when there is no target to be detected.

EXPERIMENT 1: EXPLICIT SPATIAL WORDS

The experimental paradigm and parts of the methodological descriptions, including the apparatus, eye-tracking procedures, and data preprocessing, were adapted from our previous study [9].

Participants

Sample-size planning was informed by the a-priori power analysis using MorePower 6.0 [10], as reported in our previous study [9], which targeted Cue × Target and Task × Cue × Target interactions in a repeated-measures ANOVA. In the present study, corresponding effects were examined through Cue Direction × Target Location interactions within each experiment and a between-experiment comparison of congruency effects by Word Type. Based on partial eta squared ηp^2 = .096 reported in Shaki and Fischer [7] for the three-way interaction, α = .05, and a desired power 1 − β = .80, the results indicated that at least 48 participants would be required to reliably detect the expected effects. Additional evidence comes from another study [8], which reported a Task × Congruency interaction (ηp^2 = .184) across 91 participants in two experiments and a main effect of Congruency (ηp^2 = .301) among the 46 participants analyzed in Experiment 2. Because eye-tracking studies often involve data loss due to blinks or calibration difficulties, we recruited slightly more participants than the minimum suggested.

A total of 57 native Hebrew-reading Israeli adults were recruited from the student population of Ariel, Israel. There were 11 males and 46 females with ages ranging from 19 to 40 years (M = 23.4 years). Exclusion criteria were: (a) reported history of learning disabilities, brain injuries, memory decrease, visual impairments or prescription glasses; (b) consumption of drugs on the day of the experiment; and (c) wearing eye make-up, such as mascara and eye-liner, to ensure good quality eye tracking. All participants received course credit in exchange for their participation.

Stimuli

The stimuli set was informed by previous work and consisted of four Hebrew words denoting upper locations (meaning: up, upper, high, and above) and four Hebrew words denoting lower locations (meaning: down, lower, low, and below). All Hebrew words were 3–5 letters long and shown in black Arial font with 35-point size on white background. The lexical frequencies of words denoting upper locations (M = 35.94, SD = 2.86) and words denoting lower locations (M = 32.99, SD = 4.61) were similar [11]. Cues

were displayed at the centre of the screen, while the target (an asterisk "*" in 20-point Times New Roman font) appeared 8º vertically above or below the centre of the display (cf. [7, 8]).

Apparatus

The eye-tracking apparatus and display configuration followed the protocol previously described in [9]. The experiment was programmed using the EyeLink Experiment Builder software package (SR Research Ltd., version 2.4.77). Eye tracking was performed binocularly using the EyeLink Portable Duo eye-tracker (SR Research Ltd., Ontario, Canada) at a sampling rate of 500 Hz. The signals of horizontal and vertical gaze positions from both eyes were averaged across all selected non-blink samples in a current bin. The stimuli were displayed on an AOC 2470W 23-inch display monitor (1920 x 1080 pixels resolution and 60 Hz refresh rate). The presentation of task instructions, stimuli, event timing and response recording was controlled by Experiment-Builder software. Responses were made by pressing the space bar of a QWERTY keyboard centred in front of the screen with the dominant hand. All other keyboard keys were covered.

Design

Each participant worked on 160 trials in total, divided into two separate counterbalanced blocks 80 trials with different response rules. These rules were (translated verbatim):

- respond to a target only when the preceding word denoted an upper location
- respond to a target only when the preceding word denoted a lower location

In all other cases, participants were instructed to withhold their response. Each block began with eight practice trials, during which feedback was provided by the experimenter before the experimental phase started. Next there were 72 experimental trials: these consisted of 48 go trials (66.7%) and 24 no-go trials. The go trials reflected the full crossing of four task-relevant cue words, three SOAs (400, 600 and 800 ms between cue onset and target onset), and two target locations (above and below), yielding 24 unique combinations, each presented twice. The 24 no-go trials comprised two types: (1) a task-relevant cue word that was not followed by a target (8 catch trials); and (2) a task-irrelevant cue word followed by a target appearing at either the upper or lower location (16 lure trials). All trial types were randomly presented within each block.

Procedure

The general trial sequence and eye-tracker calibration procedure were adapted from our previous study [9]. This study was approved by the Ethics Committee of Ariel University (reference number AU-SOC-SS-20190204-1). The procedure complied with guidelines and regulations based on the Declaration of Helsinki. Informed consent was obtained from all participants prior to their involvement in the study. Personal data were collected only for administrative purposes (awarding course credit) and kept confidential. Published data are fully anonymized, ensuring no personal information was disclosed.

The experiment was conducted in the Cognition Lab at Ariel University in a room with controlled lighting (but no natural light). The distance from the participant's eye to the display screen was set to 78 cm using a chin rest to ensure maximum accuracy. Before the task began, each participant was instructed to gaze toward five points on the screen for eye tracker calibration. After completing the calibration, a one-point drift correction was performed after the task instructions (including the response rule) had appeared on the screen but before data collection started. Participants were instructed to quickly and accurately make a go/no-go decision in each trial according to the response rule.

Each trial was initiated by a central fixation dot presented for 250 ms. Then, the cue word was presented for 250 ms and disappeared. Finally, if it was not a catch trial, a target was presented after one of three different randomly chosen SOAs (400, 600 or 800 ms) and remained visible until the participant's response or until 2000 ms had elapsed (in no-go trials) (see Fig. 1 for a schematic illustration of the trial sequence). Reaction time (RT) was defined as the time from target onset until the participant's response on the space bar. No feedback was given.

Figure 1

A schematic illustration of the trial sequence. *Note*. The target (asterisk) could appear either above or below fixation.

Analysis

Eye-movement data were preprocessed using an adapted version of the five-step procedure previously reported in [9], with adjustments to account for the different SOAs. A total of 57 participants were tested but three were excluded from the analyses due to a technical error that prevented their experimental data from being saved properly. Practice trials were not analysed. Next, we trimmed both manual and ocular data to ensure high-quality results.

Considering first the manual responses, 146 no-go trials (1.9%) contained commission errors and 236 trials (3%) reflected omitted responses in go trials. Five participants with more than 20% errors were excluded. Error trials of the remaining 49 participants ranged only from 0% to 9.4%. Finally, RTs for go trials deviating by more than 2.5 standard deviations from each individual's mean were excluded (138 trials, 2.9%). After this data trimming, the remaining manual responses were analysed with a repeated measures ANOVA that evaluated the effects of 2 Cue Directions (down, up), 2 Target Locations (down, up) and 3 SOAs (400, 600 and 800 ms) on correct RTs. Results appear in Fig. 2 (left upper panel).

Considering now the eye movement data, these were stored as .EDF files in the EyeLink Data Viewer software (Version 4.3.210; [12]) and later exported to Excel. We wrote a custom code in MATLAB R2024b [13] to pre-process the eye position recordings in five steps: First, trials with more than 25% of recordings missing were excluded from the analysis (123 trials). Second, recordings for which manual RTs fell outside of the mean and 2.5 standard deviations from each individual's mean were excluded (138 trials, as above). Third, no trials were excluded due to gaze positions falling outside the predefined screen

boundaries. Fourth, we created 50-millisecond bins to evaluate eye positions during the inter-stimulus interval when the screen was blank (for 150 ms in the 400 ms SOA; for 350 ms in the 600 SOA; or for 550 ms in the 800 ms SOA). We converted eye position data into gaze shifts, corresponding to the difference in average gaze position between subsequent bins (cf. [9]). For each trial, the sum of gaze shifts along the Y-axis (in pixels) was calculated from the cue word disappearance to the appearance of the target, thus reflecting the vertical eye position change over 3, 7 or 11 averages, respectively (for the three SOA conditions). For example, if the average vertical fixation coordinates from two adjacent bins were 10 and 50, respectively, the resulting gaze shift would be + 40, indicating an upward displacement of eye position at the beginning of the blank screen interval. Positive values corresponded to upward shifts, and negative values corresponded to downward shifts of eye position. It is important to note that we did not analyse eye movements along the X-axis because both cues and targets appeared horizontally centred on the screen. In order to calculate gaze shifts, we need to compare successive eye positions to the position of the eye at the beginning of the blank screen. Therefore, we excluded trials with missing eye position data from the first bin (35 trials or 0.7%). Fifth and last, we linearly interpolated the remaining data to fill in the missing data caused by blinks or random data loss (1 trial or 0.02%). Following this preprocessing, one participant was excluded because more than 25% of their trials had been removed due to missing eye-tracking recordings. This left a final sample of 48 participants, with both RT and gaze-shift analyses conducted on the same set of participants.

For No-go trials, in which no manual response was required, only gaze shift data were analysed. No-go trials presented cue words whose explicit meaning mismatched the meaning of the response rule of the current block. Thus, when participants were instructed to respond to targets following words denoting upper locations, gaze shifts were analysed following words denoting lower locations, and vice versa in the other block. Across the two blocks, 16 no-go trials followed words denoting upper locations and 16 followed words denoting lower locations. Eye movements were analysed during the inter-stimulus interval, using the same pre-processing procedure and gaze shift calculation described above. Of these trials, 165 (6.7%) were excluded based on the eye-tracking data-quality criteria described above. After this data trimming to ensure high-quality results, the remaining trials were analysed with a repeated measures ANOVA that evaluated the effects of 2 Response (go, no-go), 2 Cue Directions (down, up) and 3 SOAs (400, 600 and 800 ms) on vertical gaze shifts. Given that all eye data were collected prior to target onsets (which would trivially capture overt attention), no meaningful congruency computation was possible. Results appear in Fig. 2 (right upper panel). All results were statistically evaluated with SPSS version 29 and degrees of freedom were adjusted with the Greenhouse-Geisser correction when distributional assumptions were violated. All raw data, the analysis script, and result tables are available at https://osf.io/v3mbw.

Results and Discussion

Reaction Times

In the 2 Explicit Cue Directions (down, up) x 2 Target Locations (down, up) x 3 SOAs (400, 600 and 800 ms) ANOVA on correct RTs, there was a significant main effect of Explicit Cue Direction, $F(1, 47) = 13.69$, $p < .001$, $\eta p^2 = .226$, with faster average RTs following upward cue words (453 ms) than following downward cue words (479 ms). The main effect of Target Location was also significant, $F(1, 47) = 11.53$, $p = .001$, $\eta p^2 = .197$, with faster average RTs for upward targets (459 ms) than for downward targets (473 ms). The main effect of SOA was also significant, $F(2, 94) = 15.46$, $p < .001$, $\eta p^2 = .248$. Average RTs for the 400-, 600-, and 800-ms SOAs were 476 ms, 450 ms, and 472 ms, respectively, indicating fastest responses following intermediate SOAs.

More importantly, there was a significant interaction between Explicit Cue Direction and Target Location, $F(1, 47) = 30.23$, $p < .001$, $\eta p^2 = .391$. When cues denoted downward locations, average RTs were similar for downward and upward targets (476 ms and 482 ms, respectively), $t(47) = -0.93$, $p = .358$, $d = -0.13$. In contrast, when cues denoted upward locations, responses were faster to upward targets than to downward targets (437 ms and 470 ms, respectively), $t(47) = 6.48$, $p < .001$, $d = 0.94$.

The interaction between Explicit Cue Direction and SOA, $F(2, 94) = 0.40$, $p = .672$, $\eta p^2 = .008$, and the interaction between Target Location and SOA, $F(2, 94) = 2.20$, $p = .117$, $\eta p^2 = .045$, were non-significant. The three-way interaction between Explicit Cue Direction, Target Location, and SOA was also non-significant, Greenhouse–Geisser-corrected $F(1.77, 83.07) = 0.57$, $p = .547$, $\eta p^2 = .012$.

Gaze shifts

Consider first the subset of trials for which manual RTs were reported above (go trials). We excluded from this analysis those trials where no target appeared (catch trials) as well as the lure trials, although they unfolded identically to the go trials. This ensured direct comparability of manual and ocular responses (see Fig. 2, top left and middle panels) and also means that all data in this analysis reflect congruency between response rule (top-down) and word meaning (bottom-up).

In the 2 Explicit Cue Directions (down, up) x 3 SOAs (400, 600 and 800 ms) ANOVA on vertical gaze shifts, there was a significant main effect of Explicit Cue Direction, $F(1, 47) = 24.49$, $p < .001$, $\eta p^2 = .343$, with downward cues inducing gaze shifts of 2.76 pixels and upward cues inducing gaze shifts of 22.94 pixels. In addition, the main effect of SOA was significant, $F(2, 94) = 9.80$, $p < .001$, $\eta p^2 = .172$, with gaze shifts increasing across SOAs: 0.97 pixels at the 400-ms SOA, 15.50 pixels at the 600-ms SOA, and 22.08 pixels at the 800-ms SOA.

Moreover, SOA interacted reliably with Explicit Cue Direction, $F(2, 94) = 17.72$, $p < .001$, $\eta p^2 = .274$. A polynomial contrast testing the linear trend revealed a significant Explicit Cue Direction × SOA interaction, $F(1, 47) = 20.64$, $p < .001$, $\eta p^2 = .305$. The estimated linear increase was substantially steeper for explicitly upward cue words, $b = 19.40$ pixels per 200-ms increase in SOA, than for explicitly downward cue words, $b = 1.71$ pixels per 200-ms increase. This pattern indicates that the interaction was driven by an increasing divergence between explicit cue directions at longer SOAs (see Fig. 2, middle upper panel).

Follow-up directional paired-samples *t*-tests for these go trials were conducted to titrate the cognitive mechanism of interest in accordance with our previous expectations [9]. They showed a significant difference between upward and downward implied cue directions already at 400 ms, $t(47) = -1.91$, $p = .031$. Similarly, gaze shifts were significantly larger for upward than downward implied cue directions at both 600 ms, $t(47) = -4.78$, $p < .001$, and 800 ms, $t(47) = -4.62$, $p < .001$.

Consider now trials for which the cue word did not match the go instruction (no-go trials). Here, participants knew they would not have to respond, even if a target appeared. All data in this analysis come from trials where response rule (top-down) and word meaning (bottom-up) mismatched. Looking at these data answers our second research question, whether the eyes are controlled by the currently instructed go rule (top-down) or by the current word's meaning (bottom-up). In the 2 Explicit Cue Directions (down, up) x 3 SOAs (400, 600 and 800 ms) ANOVA on vertical gaze shifts, there was a marginally significant main effect of Explicit Cue Direction, $F(1, 47) = 3.56$, $p = .065$, $\eta p^2 = .070$, with downward cues inducing gaze shifts of 5.25 pixels and upward cues inducing gaze shifts of 14.14 pixels. In addition, the main effect of SOA was significant, $F(2, 94) = 14.07$, $p < .001$, $\eta p^2 = .230$, with gaze shifts increasing across SOAs: 0.55 pixels at the 400-ms SOA, 9.22 pixels at the 600-ms SOA, and 19.31 pixels at the 800-ms SOA. SOA did not interact reliably with Explicit Cue Direction, $F(2, 94) = 2.60$, $p = .080$, $\eta p^2 = .052$.

Polynomial contrasts did revealed no significant linear Explicit Cue Direction × SOA interaction, $F(1, 47) = 3.03$, $p = .088$, $\eta p^2 = .061$. Numerically, the linear increase was steeper for explicitly upward cue words, $b = 14.38$ pixels per 200-ms increase in SOA, than for explicitly downward cue words, $b = 4.38$ pixels per 200-ms increase, suggesting a tendency toward greater divergence between explicit cue directions at longer SOAs (see Fig. 2, right upper panel).

Despite the lack of reliable main effect, follow-up directional paired-samples *t*-tests for these no-go trials were again conducted to titrate the cognitive mechanism of interest in accordance with our previous expectations [9]. They showed no difference between upward and downward cue directions at 400 ms, $t(47) = 0.20$, $p = .420$. In contrast, gaze shifts were larger for upward than downward cue directions at both 600 ms, $t(47) = -1.84$, $p = .036$, and 800 ms, $t(47) = -1.74$, $p = .045$ (see Fig. 2, right upper panel).

These results, obtained with explicitly spatial words, align with the findings of Shaki et al. [9]: While they used implicitly spatial cue words, we used explicitly spatial cue words and replicated two main findings: We observed a spatial congruity effect in manual RTs and a general upward drift of eye gaze that is stronger in upward-associated compared to downward-associated cueing condition. Moreover, we extended the time course of these effects by examining both earlier and later delays between cue and target onsets, showing that the effect becomes stronger with longer SOAs. Finally, the novel observation obtained from gaze positions in no-go trials will be discussed in the General Discussion in light of complementary results with implicitly spatial words, reported next.

Figure 2: Results from both experiments. Upper panels: explicitly spatial cue words (Experiment 1); lower panel: implicitly spatial cue words (Experiment 2). Left side panels: manual reaction times in

milliseconds (ms). Middle and right side panels: Ocular gaze shifts in pixels, separately shown for go trials (middle panels) and no-go trials (right side panels). Error bars reflect 1 SEM (barely visible at 400 ms SOA). Statistical significance indicated by $p < .05$ (*), $p < .01$ (**), and $p < .001$ (***).

EXPERIMENT 2: IMPLICIT SPATIAL WORDS

Participants

Based on the above rationale, a total of 56 native Hebrew-reading Israeli adults were recruited from the student population of Ariel, Israel. There were 8 males and 48 females with ages ranging from 19 to 35 years (M = 22.9 years). Exclusion criteria were identical to Experiment 1.

Stimuli

The stimuli set now consisted of four Hebrew words merely connotating upper locations (meaning: sky, airplane, ceiling, tower) and four Hebrew words merely connotating lower locations (meaning: roots, carpet, floor, basement). All Hebrew words were 4–6 letters long and shown as before. The lexical frequencies for the two sets of cue words were tested using the M1 corpus by the National Institute for Testing and Evaluation [11]. The lexical frequencies of words implying upper locations (M = 31.97, SD = 2.75) and words implying lower locations (M = 29.60, SD = 2.97) were similar.

Apparatus, Design and Procedure

These were identical to Experiment 1.

Analysis

A total of 56 participants were tested but one did not complete all the experimental blocks and therefore was excluded from analyses. Practice trials were not analysed. Considering manual responses, 140 no-go trials (1.7%) contained commission errors and 158 trials (2%) reflected omitted responses in go trials. Two participants with more than 20% errors were excluded. Error trials of the remaining 53 participants ranged from 0% to 13.5%. Finally, RTs for go trials outside of the mean and 2.5 standard deviations from each individual's mean were excluded (139 trials, 2.7%). Remaining trials were analysed as before, and results appear in Fig. 2 (left lower panel).

Eye-movement data were again preprocessed in five steps, following the procedure described in Experiment 1 and adapted from our previous study [9]: First, trials with more than 25% of recordings missing were excluded from the analysis (180 trials). Second, recordings for which manual RTs fell outside of the mean and 2.5 standard deviations from each individual's mean were excluded (139 trials, as above). Third, we excluded recordings in which eye position of either eye fell outside the screen (4 trials). Fourth, we created 50-millisecond bins to evaluate eye positions during the inter-stimulus interval between cue offset and target onset when the screen was blank (for either 150 ms in the 400 ms SOA; or 350 ms in the 600 ms SOA; or 550 ms in the 800 ms SOA). We converted eye position data into gaze

shifts as before and excluded trials with missing eye position data from the first bin (103 trials or 2%). Fifth and last, we linearly interpolated the remaining data to fill in the missing data caused by blinks or random data loss (4 trials or 0.08%). Following this preprocessing, two more participants were excluded because more than 25% of their trials had been removed due to missing eye-tracking recordings. This left a final sample of 51 participants, with both RT and gaze-shift analyses conducted on the same set of participants.

No-go trials were analyzed using the same procedure as in Experiment 1, with only gaze shift data included because no manual response was required. These 32 trials per participant comprised cue words whose implied spatial direction mismatched the response rule of the current block; across the two blocks, 16 trials implied upper locations and 16 implied lower locations. Eye movements during the blank screen inter-stimulus interval were pre-processed and quantified as described above, and 167 trials (6.6%) were excluded based on the eye-tracking data-quality criteria described above.

Remaining trials were analysed with a repeated measures ANOVA that evaluated the effects of 2 Response (go, no-go), 2 Implicit Cue Directions (down, up) and 3 SOAs (400, 600 and 800 ms) on gaze shifts. Given that all eye data were collected prior to target onsets (which would trivially capture overt attention), no meaningful congruency computation was possible. Results appear in Fig. 2 (right lower panel). All results were statistically evaluated with SPSS version 29 and degrees of freedom were adjusted with the Greenhouse-Geisser correction when distributional assumptions were violated. All raw data, the analysis script, and result tables are available at https://osf.io/v3mbw.

Results and Discussion

Reaction Times

In the 2 Implicit Cue Directions (down, up) x 2 Target Locations (down, up) x 3 SOAs (400, 600 and 800 ms) ANOVA there was a main effect of Implicit Cue Direction, $F(1, 50) = 7.13$, $p = .010$, $\eta p^2 = .125$. Average RTs following downward- and upward-implying cues were 493 ms and 517 ms, respectively. Thus, contrary to Experiment 1, we obtained here faster responses following downward-implying cues.

The main effect of Target Location was non-significant, $F(1, 50) = 0.82$, $p = .369$, $\eta p^2 = .016$. The main effect of SOA was significant, $F(2, 100) = 17.02$, $p < .001$, $\eta p^2 = .254$. Average RTs for the 400-, 600-, and 800-ms SOAs were 524 ms, 491 ms, and 500 ms, respectively. Thus, responses were again fastest at the intermediate SOA.

More importantly, there was a significant interaction between Implicit Cue Direction and Target Location, $F(1, 50) = 39.77$, $p < .001$, $\eta p^2 = .443$. When cue words implied a downward direction, average RTs were significantly faster for downward than for upward targets (482 ms and 505 ms, respectively), $t(50) = -3.61$, $p < .001$, $d = -0.51$. In contrast, when cue words implied an upward direction, average RTs were faster for upward than for downward targets (501 ms and 533 ms, respectively), $t(50) = 4.83$, $p < .001$, d

= 0.68. In contrast to Experiment 1, this pattern reflects a bilateral semantic congruity effect. All remaining interactions were non-significant, all *p*-values ≥ .485.

Gaze shifts

Consider first the subset of trials for which manual RTs were reported above (go trials). We again excluded the few trials where no target appeared (catch trials), although they unfolded identically to go trials. This ensured comparability of manual and ocular responses (see Fig. 2, bottom left and middle panels) and means that all data in this analysis reflect congruency between response rule (top-down) and word meaning (bottom-up).

In the 2 Implicit Cue Directions (down, up) x 3 SOAs (400, 600 and 800 ms) ANOVA on vertical gaze shifts, there was a main effect of Implicit Cue Direction, $F(1, 50) = 12.42$, $p < .001$, $\eta p^2 = .199$, with downward cues inducing gaze shifts of 3.60 pixels and upward cues inducing gaze shifts of 13.10 pixels. In addition, the main effect of SOA was significant, $F(2, 100) = 11.24$, $p < .001$, $\eta p^2 = .184$, with gaze shifts increasing across SOAs: 0.14 pixels at the 400-ms SOA, 7.60 pixels at the 600-ms SOA, and 17.31 pixels at the 800-ms SOA. Moreover, SOA interacted with Implicit Cue Direction, $F(2, 100) = 9.01$, $p < .001$, $\eta p^2 = .153$.

A polynomial contrast testing the linear trend revealed an Implicit Cue Direction × SOA interaction, $F(1, 50) = 10.60$, $p = .002$, $\eta p^2 = .175$. The estimated linear increase was substantially steeper for implicitly upward cue words, $b = 14.26$ pixels per 200-ms increase in SOA, than for implicitly downward cue words, $b = 3.19$ pixels per 200-ms increase. This pattern indicates that the interaction was driven by increasing divergence between implicit cue directions at longer SOAs (see Fig. 2, lower middle panel).

Follow-up directional paired-samples *t*-tests for these go trials were again conducted to titrate the cognitive mechanism of interest in accordance with our previous expectations [9]. They showed no difference between upward and downward implied cue directions at 400 ms, $t(50) = -0.88$, $p = .193$. However, gaze shifts were larger for upward than downward implied cue directions at both 600 ms, $t(50) = -3.24$, $p < .001$, and 800 ms, $t(50) = -3.24$, $p < .001$.

Consider now no-go trials in which participants knew they would not have to respond, even if a target appeared. All data in this analysis reflect a mismatch between response rule (top-down) and word meaning (bottom-up). Again, looking at these data answers our second question, whether the eyes are controlled by the currently instructed go rule (top-down) or by the current word's meaning (bottom-up). In the 2 Implicit Cue Directions (down, up) x 3 SOAs (400, 600 and 800 ms) ANOVA on vertical gaze shifts, there was no effect of Implicit Cue Direction, $F(1, 49) = 0.92$, $p = .342$, $\eta p^2 = .018$, with downward cues inducing gaze shifts of 4.97 pixels and upward cues inducing gaze shifts of 9.09 pixels. The main effect of SOA was significant, $F(2, 98) = 6.91$, $p = .002$, $\eta p^2 = .124$, with gaze shifts increasing across SOAs: -0.10 pixels at the 400-ms SOA, 6.18 pixels at the 600-ms SOA, and 15.01 pixels at the 800-ms SOA. SOA did not interact with Implicit Cue Direction, $F(2, 98) = 0.30$, $p = .741$, $\eta p^2 = .006$.

Polynomial contrasts did not reveal a linear trend for the Implicit Cue Direction × SOA interaction, $F(1, 49) = 0.24$, $p = .624$, $\eta p^2 = .005$. The estimated linear increase was somewhat steeper for implicitly upward cue words, $b = 8.84$ pixels per 200-ms increase in SOA, than for implicitly downward cue words, $b = 6.27$ pixels per 200-ms increase, although the two implicit cue directions showed similar linear increases across SOAs (see Fig. 2, lower right panel).

Despite the lack of reliable main effect, follow-up directional paired-samples *t*-tests for these no-go trials were again conducted to titrate the cognitive mechanism of interest in accordance with our expectations [9]. They showed a significant difference between implicitly upward and downward cue word directions already at 400 ms, $t(50) = -1.85$, $p = .035$. Gaze shifts were also larger for implicitly upward than downward cue word directions at 600 ms, $t(50) = -1.85$, $p = .035$, but not at 800 ms, $t(50) = -0.71$, $p = .239$. This indicates a dilution of semantic effects in no-go trials over time.

These results, obtained with implicitly spatial words, closely replicate the findings of Shaki et al. [9]: We again observed a spatial congruity effect in manual RTs and a general upward drift of eye gaze that is stronger in upward-associated compared to downward-associated cueing conditions. As was the case with explicit cue words, the gaze shift effect becomes stronger with longer SOAs. Finally, before we address our question about top-down vs. bottom-up control over these gaze shifts in the General Discussion by focusing on no-go trials, we compare results between experiments.

COMPARING EXPLICIT AND IMPLICIT SPATIAL WORDS

As might be expected for the cognitively more demanding extraction of implicit meaning, RTs were 39 ms slower for implicitly compared to explicitly spatial cue words (505 ms vs 466 ms), $t(97) = 1.73$, $p = .044$, $d = 0.35$. It is unclear whether this reflects different lexical access processes required by different cue words used, or perhaps individual differences between the participant samples, or additional spatial mapping steps required with implicitly compared to explicitly spatial words. Given that this latter possibility is of interest for our goal to characterize the cognitive mechanism involved in semantic cueing, we conducted a between-experiment analysis of the congruency effects, thereby removing absolute RT differences. Specifically, RTs in congruent cue-target conditions (up-up and down-down) were subtracted from RTs in incongruent cue-target conditions (up-down, down-up) to compute congruency effects for each participant in each condition. A mixed-factors ANOVA evaluated effects of Word Type (explicit, implicit) and SOA (400, 600, 800 ms) on these congruency effects and found a main effect of Word Type, $F(1, 97) = 68.88$, $p < .001$, $\eta p^2 = .415$, with 16 ms smaller congruency benefit following explicitly compared to implicitly spatial cue words (39 and 55 ms, respectively). There was no main effect of SOA, $F(2, 194) < 1$, and no interaction, $F(2, 194) < 1$. Thus, the RT congruency effect remained stable over time.

A corresponding congruency analysis cannot be applied to gaze data because gaze was measured before target onset, and therefore before cue-target congruency relationships were defined. However, gaze-shifts allow us to test whether upward versus downward cue meaning produced a systematic linear

change in eye positions across SOAs, and whether this temporal pattern differed between explicitly and implicitly spatial words.

To address this question, we conducted a between-experiment linear trend analysis of gaze-shift data, restricted to go-trials. This was motivated by the preceding analyses of no-go trials, which did not reveal a linear Cue Direction × SOA trend with either explicit or implicit cue words. For go-trials, these new polynomial contrasts revealed no linear Word Type × Cue Direction × SOA interaction, $F(1, 97) = 1.66$, $p = .201$, $\eta p^2 = .017$, indicating similar linear trends across SOAs for all spatially associated cue words. However, the corresponding quadratic interaction was significant, $F(1, 97) = 5.11$, $p = .026$, $\eta p^2 = .050$, reflecting earlier eye elevation for explicitly compared to implicitly up-associated cue words, resulting in a larger difference at 600 ms than at 800 ms for upward-associated cue words (see Fig. 3). In contrast, the time course for explicitly and implicitly down-associated cue words was similar.

GENERAL DISCUSSION

The present study aimed to characterize the cognitive mechanism involved in semantic cueing, i.e., how visual attention is spatially directed as a result of meaning extraction from linguistic symbols. Our main question was whether mere exposure to spatially meaningful semantic cues alone moves the reader's eyes, or whether a subsequent spatial target search is necessary. After inspecting go-trials to see whether our task [9] aligns with earlier findings [7, 8], we focused on no-go trials where cue word meaning was in conflict with the go rule. This allowed us to discover whether participants' eye gaze reflects the current task instruction (top-down) or the current cue word's meaning (bottom up).

Go Trials

The go trials replicated earlier results: both explicit and implicit spatial cue words produced results that were aligned with our earlier findings [9]: We observed spatial congruency effects in RT and a general upward trend in gaze positions that was larger following upward-associated cue words.

Both experiments revealed fastest RTs at intermediate SOAs, consistent with standard fore-period effects: Participants prepared their responses by adjusting their temporal target expectation to the middle interval [14].

Surprisingly, the congruency benefit in RT was significantly smaller following explicitly compared to implicitly spatial cue words. This somewhat counterintuitive result is explained by the fact that overall RTs were 39 ms slower with implicitly compared to explicitly spatial cue words. The additional time is presumably needed to extract the "hidden" spatial meaning. Consequently, spatial codes have more time to become effective, leading to larger spatial congruency effects [15]. A similar finding appeared in our earlier work [7] (see Fig. 1, semantic condition).

For gaze shifts in go trials, where response rule and word meaning were aligned, cueing was extended across a larger time course than previously [9] and became stronger with time. We found no linear

interaction between Word Type, Cue Direction, and SOA, suggesting similar linear development of the gaze-shift effect for both word types. More work is needed to formally compare gaze-shift effects for explicitly and implicitly spatial words.

No-go Trials

Gaze shifts in no-go trials were studied to answer two main questions: First, is mere exposure to spatially meaningful semantic cues alone sufficient to move the eyes, or is a subsequent spatial task or spatial target presentation necessary to generate overt attentional shifts? And secondly, if the first answer is positive, does participants' eye gaze reflect the task instruction (blocked and top-down) or the current cue's meaning (randomized and bottom-up)?

Regarding the first question, whether spatially meaningful semantic cues alone change overt attention deployment, the answer is yes: Meaning-driven gaze shifts occurred even without the need to respond (see Fig. 2, right side panels). This finding of semantically induced spatial bias extends previous work, where spatial associations were computed in the context of response selection [7–9, 16]. It also aligns with neuroscientific evidence for semantic processing without overt responding, such as the N400 effect in language comprehension ([17]; see also [18]). The fact that this result holds for both explicitly and implicitly spatial words testifies to the power of semantic cueing [2, 3].

Regarding the second question, whether eye gaze in no-go trials is controlled bottom-up or top-down, our answer is bottom-up: We found stronger upward gaze shifts following words with up-related meaning (under the down rule) than with down-related meaning (under the up rule). We do not know whether this effect occurs because bottom-up semantic computation is inherently more salient, or because non-alternating top-down semantics decay in intensity over time. Future studies with trial-based rule randomization can address this question.

Finally, we observed an earlier shift of gaze position following upward associated cue words with explicit compared to implicit spatially associated cue words (see Fig. 3). This can be understood as a combination of faster semantic activation for explicitly spatial cue words, together with the visual system's inherent setting to advantage the upper visual field [19]: Across a wide range of tasks, upward eye movements are initiated more rapidly than downward eye movements (see also [20]).

Limitation and Outlook

Although we learned more about minimal requirements for semantic cueing of attention, our participants were still actively engaged in spatial detection. This may have created a spatial or motoric mind set which contributed to the observed results. Future work should measure semantically induced attention shifts in tasks without spatial or motoric processing requirements (e.g., passive brain-computer interface tasks [21]).

Eye movement recording is widely used to study cognitive processes during reading, assuming that gaze reflects current lexical processing [22–24]. However, our findings suggest that simply viewing spatially

associated words can shift gaze in the lexically implied direction. This semantic influence on eye behavior warrants further investigation into its dimensional generalizability and its relative strength compared with reading-task effects in models of eye movement control

Declarations

Competing Interests

The authors declare no competing interests.

Author Contribution

S.S. and M.F. conceptualized the study and methodology, performed the statistical analysis, and wrote the main manuscript text. O.P. and B.A. designed and conducted the experiments, contributed to the Methods and Results sections, and prepared the figures. All authors provided revisions and approved the final manuscript.

Acknowledgement

This work was supported by funding from ISF 1398/24 to SS and from DFG 1915/12-1 “Boundary conditions of conceptual spaces” to MHF.

Data Availability

All data obtained in this study are available on the Open Science Framework (https://osf.io/v3mbw).

References

1. Amer, T., Gozli, D. G. & Pratt, J. Biasing spatial attention with semantic information: An event coding approach. *Psychol. Res.* **82**, 840–858. https://doi.org/10.1007/s00426-017-0867-5 (2018).
2. Logan, G. D. Linguistic and conceptual control of visual spatial attention. *Cogn. Psychol.* **28**, 103–174. https://doi.org/10.1006/cogp.1995.1004 (1995).
3. Petrova, A. et al. Spatial congruency effects exist, just not for words: Looking into Estes, Verges, and Barsalou *Psychol. Sci.* 29, 1195–1199. (2008). https://doi.org/10.1177/0956797617728127 (2018).
4. Posner, M. I. Orienting of attention. *Q. J. Exp. Psychol.* **32**, 3–25. https://doi.org/10.1080/00335558008248231 (1980).
5. Gibson, B. S. & Sztybel, P. The spatial semantics of symbolic attention control. *Curr. Dir. Psychol. Sci.* **23**, 271–276. https://doi.org/10.1177/0963721414536728 (2014).

6. Estes, Z. et al. A comprehensive meta-analysis of spatial interference from linguistic cues: Beyond Petrova. *Psychol. Sci.* 29, 1558–1564. (2018). https://doi.org/10.1177/0956797618794131 (2018).
7. Shaki, S. & Fischer, M. H. How does language affect spatial attention? Deconstructing the prime-target relationship. *Mem. Cognit*. **51**, 1115–1124. https://doi.org/10.3758/s13421-022-01390-3 (2023).
8. Shaki, S. & Fischer, M. H. Language directs spatial attention differently in explicit and implicit tasks. *PLoS One*. **18**, e0291518. https://doi.org/10.1371/journal.pone.0291518 (2023).
9. Shaki, S., Pitem, O. & Fischer, M. H. Lexical priming of space depends on how deeply you think about it. *Sci. Rep.* **15**, 38410. https://doi.org/10.1038/s41598-025-22265-y (2025).
10. Campbell, J. I. D. & Thompson, V. A. MorePower 6.0 for ANOVA with relational confidence intervals and Bayesian analysis. *Behav. Res. Methods*. **44**, 1255–1265. https://doi.org/10.3758/s13428-012-0186-0 (2012).
11. NITE NLP Tools. (Version 15.06.23) Corpus. Jerusalem: National Institute for Testing and Evaluation. https://hlp.nite.org.il/WebCorpora.aspx
12. SR Research Ltd. EyeLink Data Viewer (Version 4.3.210). SR Research Ltd. (2023). https://www.sr-research.com/data-viewer/
13. The MathWorks. Inc. MATLAB (Version 24.2, R (2024b). https://www.mathworks.com/products/matlab.html (2024).
14. Luce, R. D. *Response Times: Their Role in Inferring Elementary Mental Organization* (Oxford University Press, 1991). https://doi.org/10.1093/acprof:oso/9780195070019.001.0001
15. Gevers, W., Verguts, T., Reynvoet, B., Caessens, B. & Fias, W. Numbers and space: A computational model of the SNARC effect. *J. Exp. Psychol. Hum. Percept. Perform.* **32**, 32–44. https://doi.org/10.1037/0096-1523.32.1.32 (2006).
16. Zwaan, R. A. & Yaxley, R. H. Spatial iconicity affects semantic relatedness judgments. *Psychon Bull. Rev.* **10**, 954–958. https://doi.org/10.3758/BF03196557 (2003).
17. Kutas, M. & Hillyard, S. A. Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science* **207**, 203–205. https://doi.org/10.1126/science.7350657 (1980).
18. Miklashevsky, A., Lindemann, O. & Fischer, M. H. The force of numbers: Investigating manual signatures of embodied number processing. *Front. Hum. Neurosci.* **14**, 590508. https://doi.org/10.3389/fnhum.2020.590508 (2021).
19. Previc, F. H. Functional specialization in the lower and upper visual fields in humans: Its ecological origins and neurophysiological implications. *Behav. Brain Sci.* **13**, 519–542. https://doi.org/10.1017/S0140525X00080018 (1990).
20. Fischer, M. H., Deubel, H., Wohlschläger, A. & Schneider, W. X. Visuomotor mental rotation of saccade direction. *Exp. Brain Res.* **127**, 224–232. https://doi.org/10.1007/s002210050792 (1999).
21. Metzger, S. L. et al. A high-performance neuroprosthesis for speech decoding and avatar control. *Nature* **620**, 1037–1046. https://doi.org/10.1038/s41586-023-06443-4 (2023).

22. Rayner, K. Eye movements in reading and information processing: 20 years of research. *Psychol. Bull.* **124**, 372–422. https://doi.org/10.1037/0033-2909.124.3.372 (1998).
23. Bai, Y., Jin, X., Zhao, S. & Oulasvirta, A. Hierarchical resource rationality explains human reading behaviour. *Nat. Hum. Behav.* https://doi.org/10.1038/s41562-026-02534-0 (2026).
24. Rayner, K. & Fischer, M. H. Mindless reading revisited: Eye movements during reading and scanning are different. *Percept. Psychophys*. **58**, 734–747. https://doi.org/10.3758/BF03213106 (1996).

Figures

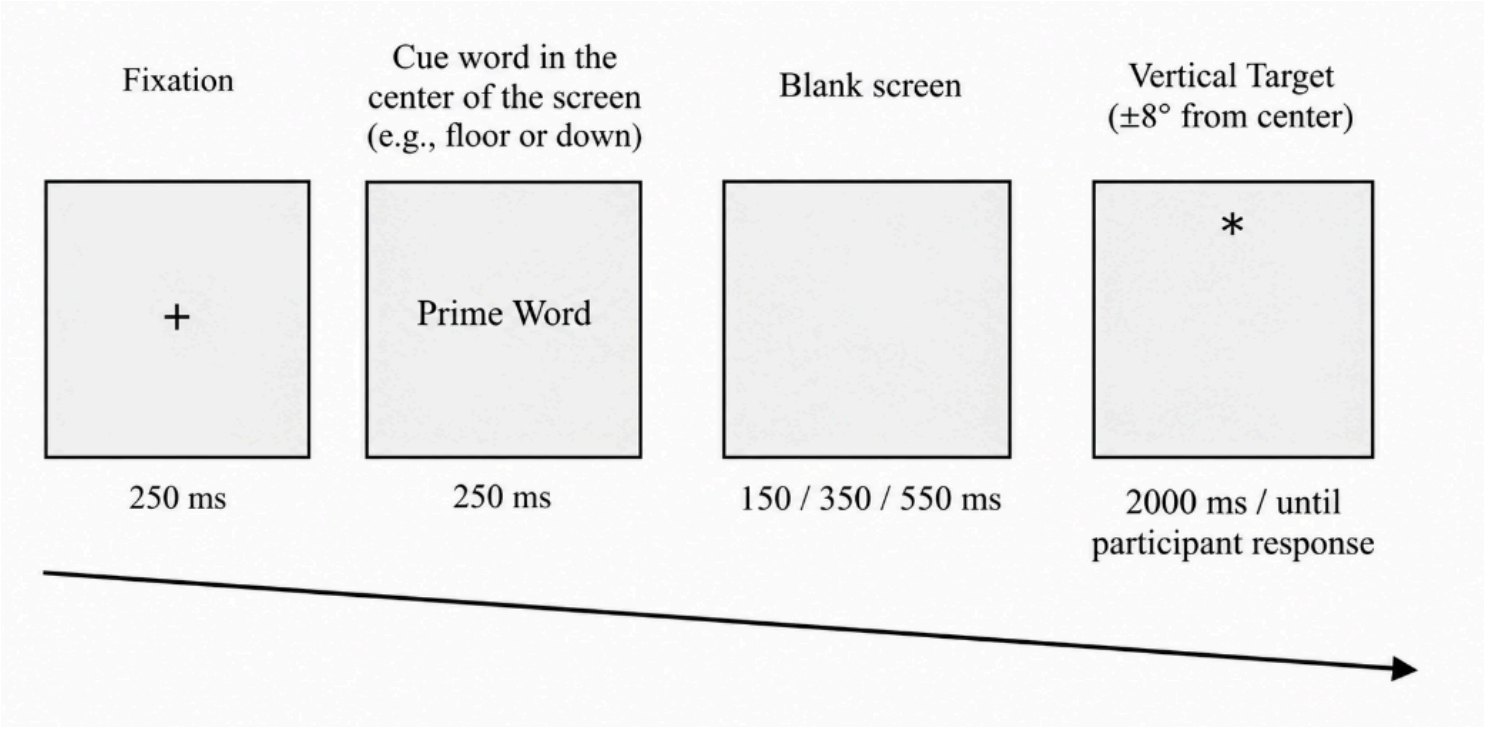


Figure 1

A schematic illustration of the trial sequence. *Note*. The target (asterisk) could appear either above or below fixation.

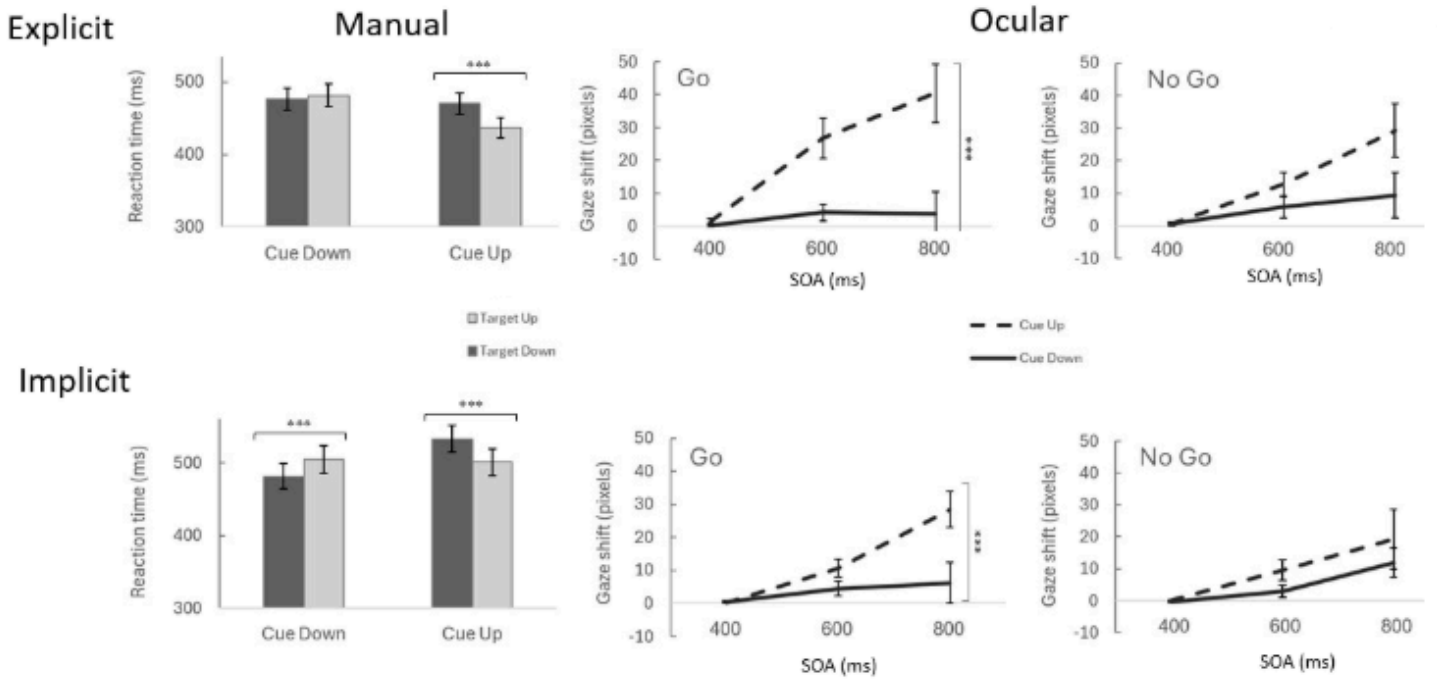


Figure 2

Results from both experiments. Upper panels: explicitly spatial cue words (Experiment 1); lower panel: implicitly spatial cue words (Experiment 2). Left side panels: manual reaction times in milliseconds (ms). Middle and right side panels: Ocular gaze shifts in pixels, separately shown for go trials (middle panels) and no-go trials (right side panels). Error bars reflect 1 SEM (barely visible at 400 ms SOA). Statistical significance indicated by $p < .05$ (*), $p < .01$ (**), and $p < .001$ (***).

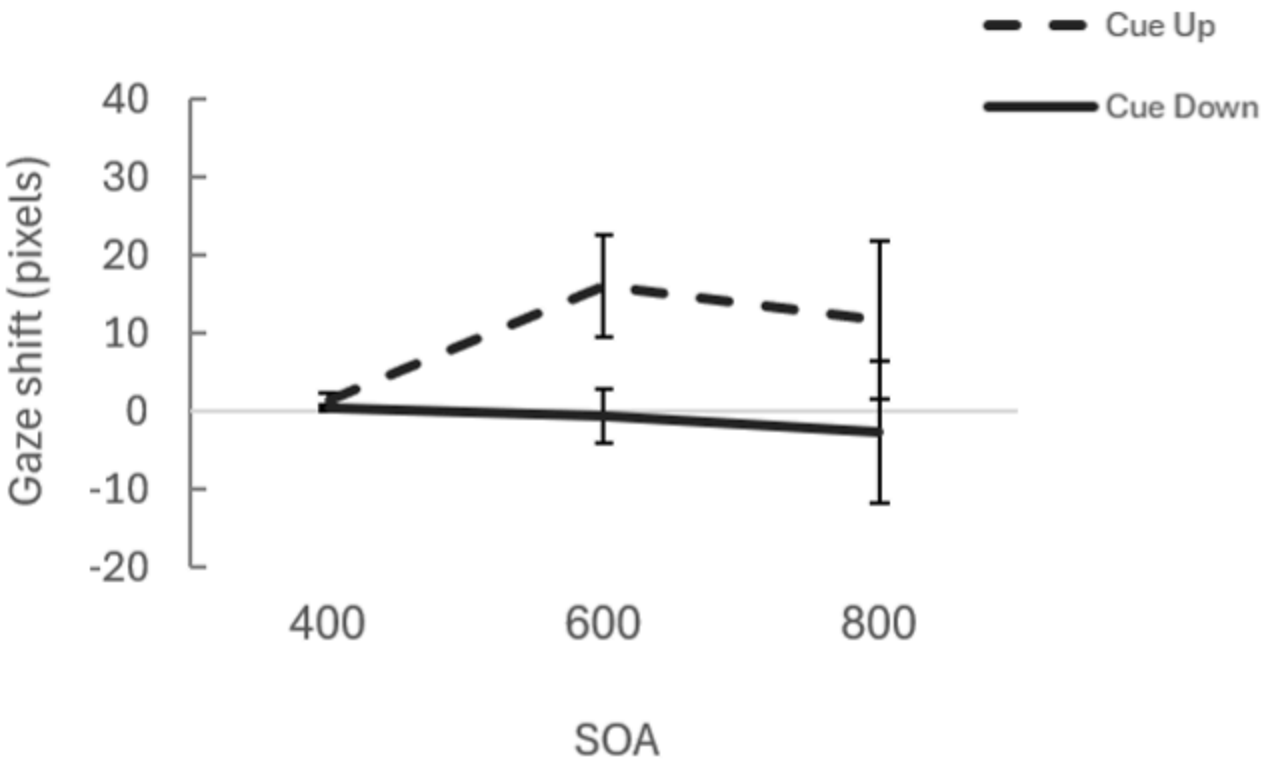


Figure 3

Gaze shift dynamics. Values represent the progressive location changes in eye position following explicit and implicit cue words (Explicit - Implicit). Error bars indicate ±1 SEM. Upward associated cue words are shown by the dashed line, downward associated cue words are shown by the solid line.