

Article

Not peer-reviewed version

Attention Dynamics in Spatial-Temporal Contexts

[Yuying Wang](#), [Xuemin Zhang](#)^{*}, [Eva Van den Bussche](#)

Posted Date: 26 January 2025

doi: [10.20944/preprints202501.1888.v1](https://doi.org/10.20944/preprints202501.1888.v1)

Keywords: Attention; psychological timeline; short-term memory; Gestalt perceptual organization principles



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Article

Attention Dynamics in Spatial-Temporal Contexts

Yuying Wang ^{1,2}, Xuemin Zhang ^{1,3,*} and Eva Van den Bussche ²

¹ Beijing Key Laboratory of Applied Experimental Psychology, Faculty of Psychology, National Demonstration Center for Experimental Psychology Education, Beijing Normal University, Beijing, China

² Brain and Cognition, Faculty of Psychology and Educational Sciences, KU Leuven, Leuven, Belgium

³ State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University

* Correspondence: xmzhang@bnu.edu.cn; Tel.: +86-10-58807499

Abstract: This study systematically explored the impact of the *spatial metaphor of time* on attentional bias using visual order stimuli, ranging from fully order to 50% disorder, highlighting the regulatory roles of short-term memory and Gestalt perceptual grouping. Three experiments were conducted. Experiment 1, comprising Experiments 1a and 1b, investigated how the *spatial metaphor of time* shaped attentional bias across five disorder levels. Experiment 1a examined horizontal orientations, while Experiment 1b focused on vertical orientations. Experiment 2 compared attentional biases toward stimuli with the same disorder level in both orientations. The key distinction between the stimuli was that one represented short-term memory, while the other did not involve memory. Experiment 3 extended Experiment 2. It compared attentional biases between fully ordered structures (intact Gestalt structures) as non-memory representations and partially disordered structures in short-term memory. The results revealed a significant preference for future-related information, particularly on the right in horizontal orientations. Short-term memory representations enhanced attentional attraction and triggered Inhibition of Return (IOR), while fully ordered structures attracted attention as effectively as partially disordered structures, thereby neutralizing attentional biases. Overall, this study contributes to a deeper understanding of the cognitive mechanisms underlying visual attention and the influence of temporal metaphors.

Keywords: attention; psychological timeline; short-term memory; gestalt perceptual organization principles

1. Introduction

The human brain integrates spatial and temporal dimensions through complex cognitive mechanisms, often associating specific spatial positions with temporal sequences. This process, known as serial order, involves arranging events or items in a sequence [1] and is fundamental to cognitive processing. It reflects the *spatial metaphor of time* [2], illustrating the brain's ability to synthesize sensory inputs into coherent temporal representations [3,4]. Studies have shown that individuals from left-to-right reading cultures, such as English and Spanish speakers, are more likely to perceive stimuli on the left as occurring earlier than those on the right [5,6]. While this metaphor is most evident in left-right orientations, it also extends to other dimensions, such as front-back and top-bottom, though findings on the vertical dimension remain inconclusive [2,7–10]. Neuroimaging research has revealed that maintaining order information in working memory engages specific cortical networks, particularly the prefrontal cortex and parietal lobes, highlighting the neural basis of these cognitive processes [11,12]. Recent work by Liu et al. [13] further demonstrated a gradual shift of attention from past to future locations, mapping temporal progression within the *spatial metaphor of time*. These findings underscore the importance of serial order processing in spatial dimensions, suggesting that the *spatial metaphor of time* may significantly influence attentional mechanisms.

Attention, as a sophisticated cognitive process, integrates and prioritizes information based on various factors. Among these, the basic features of objects, such as color and shape, play a foundational role in guiding attention [14–17]. Beyond these individual attributes, the structural organization facilitated by Gestalt perceptual grouping is equally critical. Gestalt principles enhance the efficiency of perceptual organization, enabling attention to optimize information processing and underscoring their importance in visual cognition [18–23]. Building on these foundational features, object-based attention further enhances visual search performance by significantly improving both efficiency and accuracy [24]. This advantage arises from the ability of unified entities to capture and maintain attention more effectively. Such entities possess greater salience and stability, making attentional focus more resistant to environmental disruptions [17,25–27]. Thus, to comprehensively understand how individuals process information across spatial locations, it is insufficient to focus solely on the spatial positioning of objects. Instead, it is crucial to further investigate the influence of the *spatial metaphor of time* on object-based attentional bias, as well as the role of Gestalt perceptual grouping in shaping attentional mechanisms.

Memory is intricately linked to attention [28–31]. The effective allocation of attention enhances the encoding and maintenance of spatial-temporal cues in working memory, facilitating tasks reliant on visual cognition and spatial processing [32–36]. Moreover, research has demonstrated that working memory can actively guide attention allocation, particularly when external information aligns with stored memory representations. In such cases, attention is directed preferentially toward these consistent elements [37–39]. Xie et al. [40] further revealed that when visual sensory memory representations are relevant to the current task, attention is preferentially allocated to those representations. Given that working memory and visual short-term memory share largely overlapping underlying capacity limitations [41,42], visual short-term memory might consistently affect attention in a manner comparable to both working memory and sensory memory. Building on these insights, it becomes essential to conduct experiments that integrate short-term memory to investigate how objects organized within memory representations differ in attracting attention compared to those not represented in memory.

In summary, the present study aimed to explore the impact of the *spatial metaphor of time* on attentional bias, while further investigating the roles of visual short-term memory and Gestalt perceptual grouping in this effect. Using a spatial cueing paradigm [24,43], we manipulated visual order by using images with varying levels of “disorder” as cues. These levels (0D, <1/4D, 1/4D, 1/2AD, 1/2ND) were determined by the proportion of disrupted perceptual structure within a four-quartered image, systematically varying visual complexity to assess its influence on attentional bias. This manipulation of visual order was employed across three experiments. Experiments 1a and 1b investigated attentional bias under varying levels of disorder and orientations, while Experiment 2 examined the effect of the short-term memory representation on subsequent attentional bias. Experiment 3 further tested whether short-term memory representations elicited similar effects on attention as fully ordered (intact Gestalt structures) structures.

The primary hypotheses of this study are as follows: First, we hypothesized that object-based attentional bias is shaped by spatial position, reflecting the dynamic nature of attentional shifts across the *spatial metaphor of time*. Specifically, individuals were expected to exhibit an attentional bias toward stimuli on the right in the horizontal orientation, consistent with the future-oriented spatial metaphor [5,6,13]. Additionally, in the vertical orientation, attention was hypothesized to bias downward, potentially reflecting a similar temporal association with future-oriented positions. Second, stimuli represented in short-term memory were predicted to attract more attention compared to non-represented structures. Finally, we hypothesized that competition for attentional capture might occur between stimuli with fully ordered structures and those represented in short-term memory, resulting in no significant bias toward either type.

2. Experiment 1

2.1. Method

2.1.1. Participants

The minimal sample size was determined by a priori power analysis using G*Power software, version 3.1.9.7 [44]. The power analysis indicated that at least 14 participants were required to obtain a statistical power of 0.80, assuming a Type I error probability of .05 and a medium effect size ($f = 0.25$) for the F test. In Experiment 1a, 20 college students aged 18 to 24 years (5 males, $M_{\text{age}} = 21.00$ years, $SD = 1.55$) participated, 19 right-handed and one left-handed. In Experiment 1b, a different group of 20 college students aged 19 to 24 years participated (4 males, $M_{\text{age}} = 21.45$ years, $SD = 1.60$), all right-handed. The Ethics Committee of the Psychology Faculty at Beijing Normal University approved this study (BNU202310050144). All participants signed the informed consent and received payment for their involvement.

2.1.2. Apparatus and Stimuli

The desktop computer, connected to a 27-inch color monitor with a refresh rate of 60 Hz and a resolution of 2560×1440 pixels ($57.08^\circ \times 32.10^\circ$), displayed the stimuli and collected the data. The experiment was administered using PsychoPy (v2021.2.3). Participants were permitted to adjust their viewing distance slightly, maintaining an approximate distance of 60 cm for optimal comfort.

The stimuli used in the experiments were OCTA-generated images, developed by Van Geert [45]. Forty images were used, divided into five sets of eight images each. These sets were categorized based on the proportion of disorder present in the images. The disorder levels were determined by dividing each image into four quarters and introducing disorder in specified portions: complete order (0D), less than one-quarter disorder ($<1/4D$), one-quarter disorder ($1/4D$), two adjacent quarters disorder ($1/2AD$), and two non-adjacent quarters disorder ($1/2ND$). The classifications and their corresponding disorder levels are illustrated in Figure 1. The five levels of disorder were carefully designed to systematically manipulate perceptual organization. All the elements in the image were controlled for luminance, size, shape, color to minimize any unintended spatial attentional biases. Furthermore, the disordered regions across all image types were balanced to ensure an even distribution across different quadrants and orientations. For instance, in $1/4D$ images, the disordered region was balanced across the four quadrants, while in $1/2AD$ images, the disordered parts were balanced to appear alternately on the left, right, top, and bottom halves of different images. This balanced design ensured that any observed effects could not be attributed to a fixed spatial preference for the disordered portion of the image.

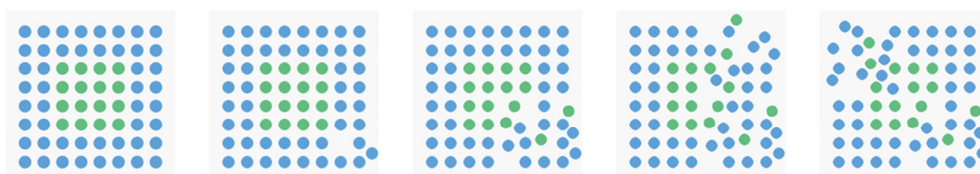


Figure 1. Examples of stimuli used in Experiment 1. From left to right, the images represent five disorder levels: complete order (0D), less than one-quarter disorder ($<1/4D$), one-quarter disorder ($1/4D$), two adjacent quarters disorder ($1/2AD$), and two non-adjacent quarters disorder ($1/2ND$).

2.1.3. Design and Procedure

Experiments 1a and 1b used a spatial cueing paradigm [43] and a $2 \times 2 \times 5$ within-subjects design. This design included two levels of cue validity (valid/invalid), two image positions (Experiment 1a: left/right, Experiment 1b: top/bottom), and five disorder levels (0D, $<1/4D$, $1/4D$, $1/2AD$, and $1/2ND$).

Figure 2 shows the experimental procedure for both Experiment 1a and Experiment 1b. In Experiment 1a, a fixation point was displayed at the center of the screen for 500-1000ms. Next, a

randomly selected image, either ordered or disordered and measuring $8.43^\circ \times 8.43^\circ$, was presented as the cue. The image appeared horizontally on the screen, positioned randomly to the left or right of the fixation point. Each image was displayed for either 100 ms or 200 ms, with both durations equally represented across 50% of the trials. After a 50ms interval with a fixation point displayed, a probe dot ($1.30^\circ \times 1.30^\circ$) appeared for 200ms at either the same location as the image (valid cue) or the opposite location (invalid cue). The proportion of valid to invalid cues was balanced at 50% across all trials. Participants were instructed to press the "F" key for the left probe dot and the "J" key for the right probe dot. Response times and accuracy were recorded, with each trial having a response deadline of 1500ms. Experiment 1b followed the same procedure as Experiment 1a, except that the images were presented vertically, either at the top or bottom of the screen, positioned 7.86° away from the fixation point. Participants pressed the "T" key for a probe dot presented at the top of the screen and the "N" key for a probe dot presented at the bottom of the screen. In both Experiment 1a and Experiment 1b, each participant completed 320 trials and was instructed to respond as quickly and accurately as possible.

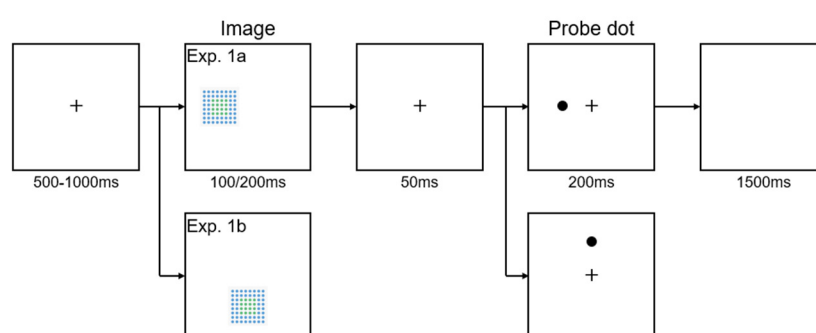


Figure 2. The procedure of Experiments 1a and 1b. In both experiments, participants responded to the position of probe dots. In Experiment 1a, they pressed "F" for dots on the left side of the screen (as shown in the image) and "J" for dots on the right. In Experiment 1b, they pressed "T" for dots at the top of the screen (as shown in the image) and "N" for dots at the bottom (as shown in the image). The procedure for Experiment 1a shows a valid cue trial; the procedure for Experiment 1b shows an invalid cue trial.

2.2. Results

The data were analyzed using IBM SPSS Statistics 26.0. Initially, mean reaction times (RTs, in milliseconds or ms) and mean error (in %) rates were calculated for each participant. Trials with RTs falling beyond ± 2.5 SDs from the participant's mean were excluded, which on average accounted for 3.6% of trials in Experiment 1a and 4.1% in Experiment 1b. Repeated measures ANOVA was conducted on the mean RTs and mean error rates. The within-subjects factors included Cue Validity (valid/invalid), Image Position (left/right for Experiment 1a; top/bottom for Experiment 1b), and Disorder Level (0D, <1/4D, 1/4D, 1/2AD, and 1/2ND).

Results for RTs in Experiment 1a: The interaction effect between Image Position and Disorder Level was significant. $F(4, 16) = 9.200, p < .001, \eta_p^2 = .697$. To further examine this interaction, we conducted post-hoc paired sample *t*-tests with Bonferroni correction. The analysis focused on the *discrepancy score of Image Position*, calculated as the difference in RTs between stimuli presented on the left and those on the right (RT for left images minus RT for right images). This score reflects the directionality of attentional bias: a positive value indicates faster responses to probes on the right, suggesting a rightward attentional bias, while a negative value suggests a leftward bias. Additionally, the magnitude of the discrepancy score represents the strength of the attentional bias, with larger values indicating stronger biases across conditions. In this experiment, post-hoc paired sample *t*-tests with Bonferroni correction revealed significant differences in the *discrepancy score of Image Position* across certain disorder levels. Specifically, the discrepancy score of 1/4D ($M = -5.21\text{ms}, SE = 5.72$) was significantly smaller than that of 1/2AD ($M = 12.70\text{ms}, SE = 4.36$), $t(1, 19) = -3.886, p = .001$; the

discrepancy score of 1/4D was significantly smaller than that of 1/2ND ($M = 15.00\text{ms}$, $SE = 4.75$), $t(1, 19) = -3.329$, $p < .001$. No significant differences were observed in the discrepancy score of Image Position between any other pairs of disorder levels (all $ps > .013$). The interaction is depicted in Figure 3. None of the other main effects or interactions reached significance (all $ps > .12$).

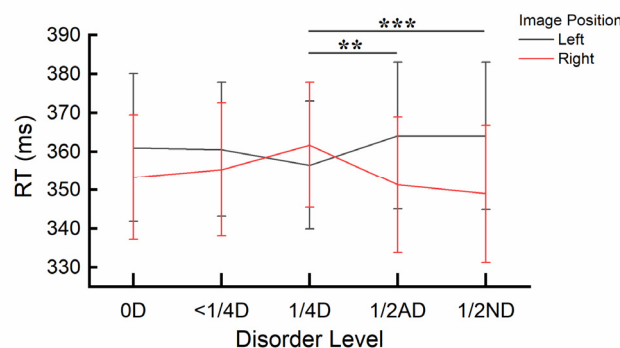


Figure 3. The interaction effect between Image Position and Disorder Level for RTs. Error bars represent the Standard Error of the Mean (SEM). ** $p < .01$, *** $p < .001$.

Results for error rates in Experiment 1a: The main effect of Image Position reached significance, $F(1, 19) = 6.181$, $p = .022$, $\eta_p^2 = .245$, with on average more errors when the image was displayed on the left of the screen ($M = 4.09\%$, $SE = 0.42\%$) compared to the right of the screen ($M = 2.97\%$, $SE = 0.42\%$). None of the other main effects or interactions reached significance (all $ps > .15$).

Results for RTs in Experiment 1b: A significant main effect of Image Position was observed, $F(1, 19) = 7.171$, $p = .015$, $\eta_p^2 = .274$, with RTs on average being faster when the image was displayed on the bottom of the screen ($M = 370.893$, $SE = 13.427$) compared to the top of the screen ($M = 382.622$, $SE = 13.143$). None of the other main effects or interactions reached significance (all $ps > .06$).

Results for error rates in Experiment 1b: No significant main effects or interactions were found, all $ps > .12$.

2.3. Discussion

In Experiments 1a and 1b, we used a spatial cueing paradigm to examine attentional biases and to implicitly activate associations between spatial positions and temporal concepts [46,47]. In other words, using this paradigm could reduce any direct, goal-directed focus on the stimuli or the position, as the task required responses to the probe dot independently of stimulus presentation. Therefore, differences in reaction times and accuracy across stimulus positions can be attributed to participants' ingrained spatial-temporal attentional biases. Experiment 1a revealed a higher error rate for probes following left-sided stimuli compared to right-sided stimuli. Analysis of the *discrepancy score of Image Position* further showed a rightward attentional bias under higher disorder levels (1/2AD and 1/2ND), which was consistent with the error rate results. This rightward bias aligns with the future-oriented spatial metaphor, as participants from left-to-right reading cultures tend to associate the right side with "future" and the left side with "past" [5,6,13]. Interestingly, a leftward bias was observed in the 1/4D condition, deviating from the expected pattern. This result suggested that intermediate disorder levels may impose unique visual or attentional demands, temporarily overriding typical spatial-temporal associations. However, as the underlying mechanisms remain unclear, further research is needed to explore how varying levels of visual complexity interact with attentional allocation and spatial-temporal associations. Experiment 1b showed faster responses to bottom-presented images compared to top-presented images, indicating a downward attentional bias. However, this effect was less pronounced than the horizontal orientation results, which may reflect a less stable mental timeline for vertical spatial representation. This finding is consistent with previous studies reporting varied vertical attentional biases [7,48–50]. Additionally, the downward

bias could be partially attributed to the visual system's natural preference for the lower visual field, which is evolutionarily associated with detecting potential threats in the environment [51–53].

Moreover, the main effect of cue validity was not significant in either Experiment 1a or 1b, which might suggest a potential inhibition of return (IOR). IOR is a cognitive process where attention is less likely to return to a previously attended location, leading to slower responses or more errors for validly cued positions [54]. Typically, if attention were consistently maintained at the cued location, RTs for validly cued targets would be significantly faster, and error rates significantly lower, than for invalidly cued ones. The absence of such significant differences in the present experiments suggested that attention may have been suppressed at previously attended locations, although the extent of this suppression appears to be mild. One potential explanation for this result is that the stimuli used in these experiments possess fewer complexity factors and adhere to an ordered structure reminiscent of Gestalt perception, facilitating rapid processing within short stimulus durations. An ERP study has demonstrated that Gestalt organization leads to swift processing occurring approximately 100ms after stimulus onset [55], providing evidence for this explanation. Such rapid processing may reduce the time window during which attentional orienting effects, such as cue validity, can be robustly observed, thereby contributing to the lack of a significant cue validity effect.

3. Experiment 2

3.1. Method

3.1.1. Participants

The minimal sample size was determined using the same method as in Experiments 1a and 1b. The power analysis indicated that at least 24 participants were required to obtain a statistical power of 0.80, assuming a Type I error probability of .05 and a medium effect size ($f = 0.25$) for the F test. 35 college students aged 18 to 29 years (5 males, $M_{age} = 21.69$ years, $SD = 2.23$) participated, 34 right-handed and one left-handed.

3.1.2. Apparatus and Stimuli

The apparatus and stimuli were identical to those used in Experiments 1a and 1b.

3.1.3. Design and Procedure

The design of Experiment 2 was a $2 \times 2 \times 4$ within-subjects design, incorporating two levels of cue validity (valid/invalid), two levels of image similarity (same/different), and four image positions (left/right/top/bottom). Experiment 2 also used a spatial cueing paradigm, but, unlike Experiments 1a and 1b, a central image was introduced before the cue image. This central image could be either identical to or different from the cue image, as shown in [Figure 4](#).

Experiment 2 consisted of two parts with different orientations: horizontal (left/right) and vertical (top/bottom). Each participant first completed the horizontal part, then the vertical part, with 320 trials per part, totaling 640 trials per participant. Each trial began with a fixation point at the center of the screen, displayed for 500-1000ms. Following the fixation, a central image appeared at the screen's center for 500ms. This central image was randomly chosen from five types (0D, <1/4D, 1/4D, 1/2AD, and 1/2ND, see Experiment 1). Next, a cue image was presented for either 100ms or 200ms. This cue image could appear on the left or right side (horizontal orientation) or on the top or bottom side (vertical orientation) of the screen. The cue image could be the same as (50% of the trials) or different from (50% of the trials) the central image but would always have the same disorder level as the central image. Thus, the two levels of image similarity (same/different) represent whether the cue image was completely identical to the central image or not. Trials were presented randomly. After a brief interval of 50ms again displaying the fixation cross, a probe dot appeared for 200ms. Participants responded to the probe's location using the same keys as in Experiment 1a and 1b: "F" for left, "J" for right, "T" for top, and "N" for bottom. By structuring the experiment into two distinct

parts and ensuring each participant completed both, the design aimed to comprehensively assess the effects of different image positions and cue validity on RTs and error rates. This approach allowed for a robust analysis of spatial attention across both horizontal and vertical orientations.

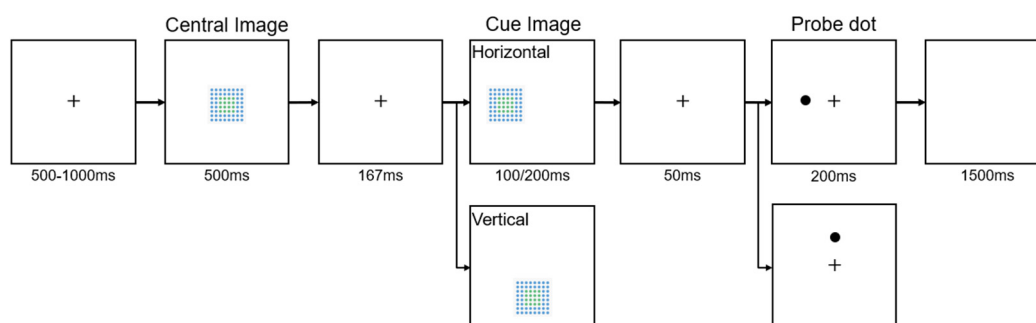


Figure 4. The procedure of Experiment 2. In this experiment, participants responded to the position of probe dots. In Experiment 1a, they pressed “F” for dots on the left side of the screen (as shown in the image) and “J” for dots on the right. In Experiment 1b, they pressed “T” for dots at the top of the screen (as shown in the image) and “N” for dots at the bottom. The procedure for the horizontal orientation shows a valid cue trial; the procedure for the vertical orientation shows an invalid cue trial.

3.2. Results

The exclusion criteria for Experiment 2 were consistent with those used in Experiments 1a and 1b, where trials deviating more than ± 2.5 SDs from the individual participant's mean were excluded, accounting for an average of 4.0% of total trials. Repeated measures ANOVA was performed on the RTs and error rates, with Cue Validity (valid/invalid), Image Similarity (same/different), and Image Position (left/right/top/bottom) as within-subjects factors.

Results for RTs: The main effect of Cue Validity was significant, $F(1, 34) = 59.136$, $p < .001$, $\eta_p^2 = .635$. RTs were significantly slower for valid cue images ($M = 363.99\text{ms}$, $SE = 8.05$) compared to invalid ones ($M = 341.63\text{ms}$, $SE = 8.83$). This finding suggested the occurrence of inhibition of Return (IOR). The main effect of Image Similarity was significant, $F(1, 34) = 10.086$, $p = .003$, $\eta_p^2 = .229$, with RTs being faster when the cue image matched the central image ($M = 351.90\text{ms}$, $SE = 8.61$) than when they differed ($M = 353.72\text{ms}$, $SE = 8.63$). The main effect of Image Position was also significant, $F(3, 32) = 13.028$, $p < .001$, $\eta_p^2 = .550$. Post-hoc comparisons using Bonferroni correction indicated that RTs were significantly faster when the image was presented on the left ($M = 342.89\text{ms}$, $SE = 9.555$) or right ($M = 336.50\text{ms}$, $SE = 8.98$) compared to the top ($M = 367.83\text{ms}$, $SE = 8.74$) and bottom ($M = 364.03\text{ms}$, $SE = 8.73$), $ps < .001$. There were no significant differences in RT between any other pair of Image Positions, $ps > .16$. Combining the data for the left and right positions into a horizontal orientation and the data for the top and bottom positions into a vertical orientation revealed a significant difference in RTs between these orientations, $t(1, 34) = -6.360$, $p < .001$, with RT being significantly faster in the horizontal orientation ($M = 339.69\text{ms}$, $SE = 9.17$) compared to the vertical orientation ($M = 365.93\text{ms}$, $SE = 8.54$).

A significant interaction effect was found between Image Similarity and Image Position, $F(3, 32) = 3.617$, $p = .024$, $\eta_p^2 = .253$. Post-hoc paired sample t -tests using Bonferroni correction indicated a significant difference in RT for the *discrepancy score of Image Similarity*. This score reflects the difference between trials where the central image and the cue image were either the same or different (calculated as different minus same), highlighting individual variations in responses to these conditions. For RTs, a positive discrepancy score indicates faster responses to the same image, while a negative score indicates slower responses. Similarly, for error rates, a positive score reflects a lower error rate for the same image, whereas a negative score reflects a higher error rate. In this experiment, the post-hoc analyses confirmed significant RT differences in the discrepancy score of Image Similarity across specific Image Positions. Specifically, the discrepancy score was significantly smaller when the image was presented on the left ($M = -1.68\text{ms}$, $SE = 1.32$) compared to when it was presented on the right (M

= 3.13ms, $SE = 1.10$), $t(1, 34) = -2.924$, $p = .006$. No significant differences in Image Similarity were observed between any other pairs of Image Positions (all $ps > .012$). The interaction is also depicted in Figure 5. None of the other interactions reached significance (all $ps > .11$).

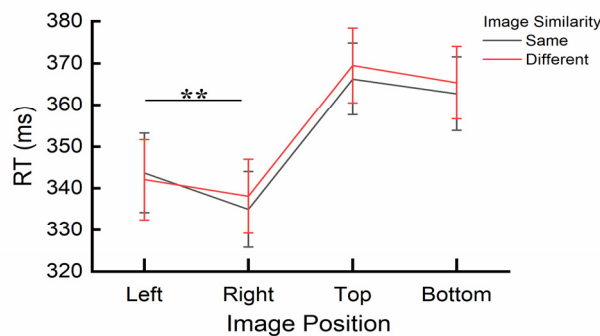


Figure 5. The interaction effect between Image Similarity and Image Position for RTs. Error bars represent SEM. $**p < .01$.

Results for error rates: The main effect of Cue Validity was significant, $F(1, 34) = 10.807$, $p = .002$, $\eta_p^2 = .241$, with higher error rates when the cue image was valid ($M = 4.51\%$, $SE = 0.39\%$) compared to when it was invalid ($M = 3.33\%$, $SE = 0.24\%$), indicating IOR. The main effect of the Image Position was also significant, $F(3, 32) = 5.080$, $p = .005$, $\eta_p^2 = .323$. Post-hoc comparisons using Bonferroni correction indicated that error rates were significantly lower when the image was presented on the left ($M = 3.54\%$, $SE = 0.34\%$) than on the top ($M = 4.70\%$, $SE = 0.44\%$), $p = .009$; similarly, error rates were lower when the image was presented on the right ($M = 3.30\%$, $SE = 0.34\%$) than on the top ($p = .014$) or the bottom ($M = 4.14\%$, $SE = 0.31\%$, $p = .047$). There were no significant differences in error rates between any other pairs of Image Positions, $ps > .29$. Combining the data for the left and right positions into a horizontal orientation and the data for the top and bottom positions into a vertical orientation revealed a significant difference in error rates between these orientations, $t(1, 34) = -2.768$, $p = .009$, with error rates being significantly lower in the horizontal orientation ($M = 3.73\%$, $SE = 0.307\%$) compared to the vertical orientation ($M = 4.42\%$, $SE = 0.333\%$).

A significant interaction effect was also found between Cue Validity and Image Position, $F(3, 32) = 3.414$, $p = .029$, $\eta_p^2 = .242$. Post-hoc paired sample t -tests using Bonferroni correction indicated significant differences in error rates for IOR. Specifically, the IOR was significantly larger when the cue image was presented on the left ($M = 2.43\%$, $SE = .628\%$) compared to the top ($M = 0.32\%$, $SE = .602\%$), $t(1, 34) = 2.799$, $p = .008$. No other significant differences in IOR were observed between any other pairs of Image Positions (all $ps > .018$). The interaction is also depicted in Figure 6. Furthermore, combining the data of IOR for the left and right positions into a horizontal orientation and the data for the top and bottom positions into a vertical orientation revealed a significant difference IOR between them, with IOR being significantly larger in the horizontal orientation ($M = 1.98\%$, $SE = .414\%$) compared to the vertical orientation ($M = .38\%$, $SE = .477\%$), $t(1, 34) = 3.024$, $p = .005$.

None of the other main effects or interactions reached significance (all $ps > .12$).

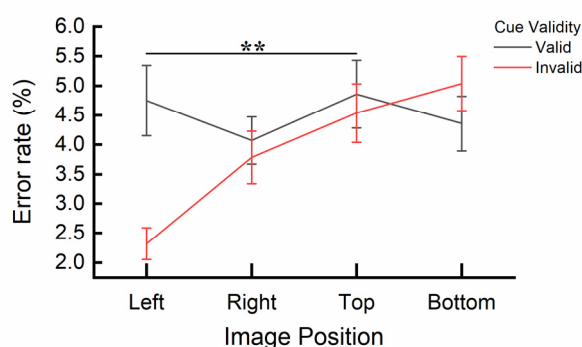


Figure 6. The interaction effect between Cue Validity and Image Position for error rates. Error bars represent SEM. $**p < .01$, $***p < .001$.

2.4. Discussion

Consistent with our expectations, participants responded faster when the central image matched the cue image, indicating that short-term memory enhances the processing of matching stimuli. Notably, the significant inhibition of return (IOR) was observed in Experiment 2, with participants responding more slowly and less accurately to valid cues. This suggested that memory representations enhanced attentional engagement as well as amplified IOR, even though only half of the trials used the same central and cue images. In contrast, the lack of memory representations in Experiments 1a and 1b likely reduced attentional engagement and IOR, resulting in no significant differences between valid and invalid cues. These findings align with previous research showing that memory representations enhance attentional engagement [37–39], while further IOR inhibits the return of attention to previously attended locations.

As for the orientation, the result demonstrated faster responses and less error rates in the horizontal orientation than vertical orientation. This finding suggested proficiency in processing horizontal information, likely due to common horizontal eye movements and attentional shifts in everyday activities such as reading and scanning environments [56,57]. Moreover, the IOR was more pronounced in the horizontal orientation than in the vertical, which might be due to greater susceptibility to habituation in the horizontal direction [58].

The interaction between Image Similarity and Image Position was found in results of RTs. When cue and central images were the same and positioned on the left side, responses were slower. Furthermore, the result of error rates highlighted greater IOR effects on the left, underscoring possibly stronger inhibitory control in this position under specific conditions. This finding implied that locations corresponding to previously information are more likely to inhibit attention and responses to the information stored in memory. This observation supports Liu et al.'s finding [13] from another perspective, suggesting that attention can be influenced not only by the temporal properties of the positions but also by the temporal information stored in the objects at those locations.

4. Experiment 3

4.1. Method

4.1.1. Participants

The minimal sample size was determined using the same method as in the previous experiments. The power analysis indicated that at least 24 participants were required to obtain a statistical power of 0.80, assuming a Type I error probability of .05 and a medium effect size ($f = 0.25$) for the F test. 33 college students aged 18 to 26 years (6 males, $M_{age} = 21.22$ years, $SD = 2.61$) participated, 31 right-handed and two left-handed.

4.1.2. Apparatus, Design, Stimuli and Procedure

The apparatus in this experiment was identical to those used in Experiments 1 and 2. The design of Experiment 3 was similar to that of Experiment 2, featuring a $2 \times 2 \times 4$ within-subjects design: 2 levels of cue validity (valid/invalid), 2 levels of image similarity (same/different), and 4 image positions (left/right/top/bottom). The stimuli and procedure for Experiment 3 mirrored that of Experiment 2, with the following key differences: when the cue image differed from the central image, it was not an image with the same disorder level as the central image, but a completely ordered image. Additionally, the central image no longer included a completely ordered image, but only partially disordered images at four different levels (<1/4D, <1/4D, 1/2AD, 1/2ND). Participants first

completed the horizontal orientation of the experiment, followed by the vertical orientation, with 256 trials in each part, for a total of 512 trials per participant.

4.2. Results

The exclusion criteria for Experiment 2 were aligned with those of the previous experiments, excluding trials that accounted for an average of 4.6% of the total trials. Repeated measures ANOVA was conducted on the RTs and error rates, with Cue Validity (valid/invalid), Image Similarity (same/different), and Image Position (left/right/top/bottom) as within-subjects factors.

Results for RTs: Cue Validity had a significant main effect, $F(1, 32) = 75.495$, $p < .001$, $\eta_p^2 = .702$. The RTs were significantly slower for valid cue images ($M = 357.49\text{ms}$, $SE = 7.03$) compared to invalid ones ($M = 334.40\text{ms}$, $SE = 7.92$), indicating IOR. None of the other main effects or interactions reached significance (all $ps > .08$). Furthermore, combining the data for the left and right positions into a horizontal orientation and the data for the top and bottom positions into a vertical orientation revealed a significant difference in RTs between these orientations, $t(1, 32) = -2.366$, $p = .024$, with RTs being significantly faster in the horizontal orientation ($M = 339.32\text{ms}$, $SE = 8.74$) compared to the vertical orientation ($M = 352.57\text{ms}$, $SE = 6.93$).

Results for error rates: The main effect of Cue Validity was significant, $F(1, 32) = 13.093$, $p = .001$, $\eta_p^2 = .290$, with error rates significantly higher when the cue image was valid ($M = 5.48\%$, $SE = 0.44\%$) compared to when it was invalid ($M = 3.89\%$, $SE = 0.33\%$), indicating IOR. A significant main effect of Image Position was also found, $F(3, 30) = 6.652$, $p = .001$, $\eta_p^2 = .399$. Post-hoc comparisons using Bonferroni correction indicated that error rates were significantly smaller when the image was presented on the left ($M = 4.21\%$, $SE = 0.37\%$) compared to the top ($M = 5.82\%$, $SE = 0.43\%$), $p = .002$; and when presented on the right ($M = 4.07\%$, $SE = 0.43\%$) compared to the top, $p = .004$. Combining the data for the left and right positions into a horizontal orientation and the data for the top and bottom positions into a vertical orientation revealed a significant difference in error rates between these orientations, $t(1, 32) = -3.922$, $p < .001$, with error rates being significantly smaller in the horizontal orientation ($M = 4.14\%$, $SE = 0.340\%$) compared to the vertical orientation ($M = 5.23\%$, $SE = 0.352\%$).

A significant interaction effect between Image Similarity and Image Position was also observed, $F(3, 30) = 3.672$, $p = .023$, $\eta_p^2 = .269$. Post-hoc paired sample t -tests using Bonferroni correction indicated that the *discrepancy score of Image Similarity* (different versus same) was significantly smaller when the image was presented on the left ($M = -1.51\text{ms}$, $SE = 0.53$) compared to when it was presented on the bottom ($M = 0.76\text{ms}$, $SE = 0.58$), $t(1, 32) = -3.000$, $p = .005$. No significant differences in Image Similarity were observed between any other pairs of Image Positions (all $ps > .011$). The interaction is also depicted in [Figure 7](#). None of the other interactions reached significance (all $ps > .48$).

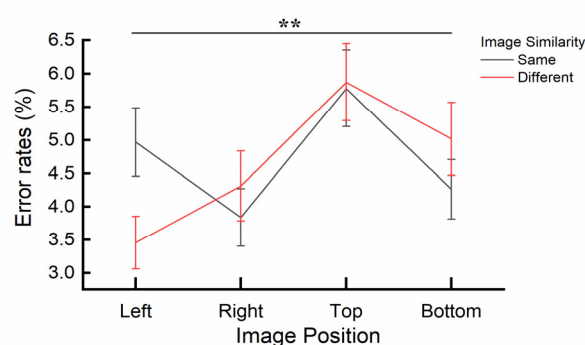


Figure 7. The interaction effect between Image Similarity and Image Position for error rates. Error bars represent SEM. ** $p < .01$.

4.3. Discussion

In line with the findings of Experiment 2, Experiment 3 demonstrated IOR in both RTs and error rates. Responses were also faster and more accurate in the horizontal orientation compared to the vertical orientation. Moreover, the error rate results revealed a significant interaction effect between Image Similarity and Image Position, similar to the RT results of Experiment 2. When images on the left matched the central image, error rates were higher compared to non-matching images. This pattern was different from the results for images on the right, top, and bottom, with a significant difference observed between the left and bottom positions.

Interestingly, unlike Experiment 2, this experiment did not reveal a significant main effect of image similarity in either RTs or error rates. This result suggested that there was no significant difference in responses when the central and cue images were the same or different. This finding aligned with the possibility that both the short-term memory representations of the central image and fully ordered structures, which conform to Gestalt principles, may equally capture attention. As a result, when both are present simultaneously, any potential attention bias might be balanced out.

4.4. Combined Results of Experiment 2 and Experiment 3

As Experiment 2's central image included five levels of disorder, whereas Experiment 3's central image included only four levels of disorder (excluding the fully ordered images), we first excluded the trials where the central image was 0D for each participant in Experiment 2 to ensure that both experiments had the same experimental design and number of trials (512 trials). Repeated measures ANOVA was performed on the RTs and error rates, with Cue Validity (valid/invalid), Image Similarity (same/different), and Image Position (left/right/top/bottom) as within-subjects factors, and Group (Experiment 2/Experiment 3) as between-subjects factor.

Results for RTs: The main effect of Cue Validity was significant, $F(1, 66) = 128.593, p < .001, \eta_p^2 = .661$, with RTs significantly slower for valid cue images ($M = 360.64\text{ms}, SE = 5.59$) compared to invalid ones ($M = 337.86\text{ms}, SE = 5.95$), indicating IOR. The main effect of Image Similarity was significant, $F(1, 66) = 6.391, p = .014, \eta_p^2 = .088$, with RTs being faster when the cue image matched the central image ($M = 348.65\text{ms}, SE = 5.68$) than when they differed ($M = 349.85\text{ms}, SE = 5.70$). A significant main effect of Image Position was also found, $F(3, 64) = 11.313, p < .001, \eta_p^2 = .347$. Post-hoc comparisons using Bonferroni correction indicated that RTs were significantly faster when the image was presented on the left ($M = 342.25\text{ms}, SE = 6.65$) or right ($M = 336.43\text{ms}, SE = 6.21$) compared to the top ($M = 360.65\text{ms}, SE = 5.70$) or bottom ($M = 357.67\text{ms}, SE = 5.60$), $ps < .001$. There were no significant differences in RTs between any other pair of image positions, $ps > .08$. Moreover, the main effect of Group was significant, $F(1, 66) = 3773.619, p < .001, \eta_p^2 = .983$. RTs in Experiment 2 ($M = 352.56\text{ms}, SE = 7.92$) were on average slower than those in Experiment 3 ($M = 345.94\text{ms}, SE = 8.16$). Combining the data for the left and right positions into a horizontal orientation and the data for the top and bottom positions into a vertical orientation revealed a significant difference in RTs between these orientations, $t(1, 67) = -5.725, p < .001$, with RTs being significantly faster in the horizontal orientation ($M = 339.34\text{ms}, SE = 6.28$) compared to the vertical orientation ($M = 359.36\text{ms}, SE = 5.54$).

A significant interaction effect was found between Cue Validity and Image Position, $F(3, 64) = 5.542, p = .006, \eta_p^2 = .176$. Post-hoc paired sample t -tests using Bonferroni correction indicated significant differences in RTs for IOR depending on the Image Position. Specifically, the IOR was significantly larger when the cue image was presented on the bottom ($M = 26.17\text{ms}, SE = 2.688$) compared to the top ($M = 19.58\text{ms}, SE = 2.376$), $t(1, 67) = 3.209, p = .002$. No significant IOR differences were observed between other pairs of Image Positions (all $ps > .015$). The interaction is also depicted in [Figure 8](#).

None of the other interactions reached significance (all $ps > .14$).

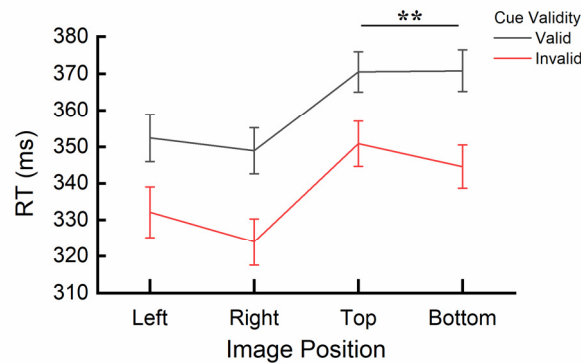


Figure 8. The interaction effect between Cue Validity and Image Position for error rates. Error bars represent SEM. ** $p < .01$.

Results for error rates: Cue validity had a significant main effect, $F(1, 66) = 26.400$, $p < .001$, $\eta_p^2 = .286$, with error rates significantly higher when the cue image was valid ($M = 4.96\%$, $SE = 0.28\%$) compared to when it was invalid ($M = 3.57\%$, $SE = 0.21\%$), indicating IOR. A significant main effect of Image Position was also found, $F(3, 64) = 11.528$, $p < .001$, $\eta_p^2 = .351$. Post-hoc comparisons using Bonferroni correction indicated that error rates were significantly smaller when the image was presented on the left ($M = 3.85\%$, $SE = 0.25\%$) compared to the top ($M = 5.27\%$, $SE = 0.31\%$), $p < .001$; and when presented on the right ($M = 3.54\%$, $SE = 0.24\%$) compared to both the top ($p < .001$) and the bottom ($M = 4.40\%$, $SE = 0.26\%$, $p = .025$). Furthermore, the main effect of Group was significant, $F(1, 66) = 4.285$, $p = .042$, $\eta_p^2 = .061$. The error rates in Experiment 2 ($M = 3.84\%$, $SE = 0.29\%$) were on average smaller than those in Experiment 3 ($M = 4.69\%$, $SE = 0.29\%$).

A significant three-way interaction effect was found between Image Similarity, Image Position and Group, $F(3, 64) = 5.219$, $p = .003$, $\eta_p^2 = .197$. Post-hoc independent sample t -tests using Bonferroni correction indicated significant differences between the two groups in error rates for the *discrepancy score of Image Similarity* between specific Image Positions. Specifically, when the cue image was presented on the left, the discrepancy score in Experiment 2 ($M = .27\%$, $SE = .564\%$) was significantly greater than that in Experiment 3 ($M = -1.52\%$, $SE = .533\%$), $t(1, 66) = .927$, $p = .008$. No significant differences were observed between any other pairs of discrepancy scores (all $ps > .26$). The interaction is also depicted in Figure 9.

None of the other main effects or interactions reached significance (all $ps > .17$).

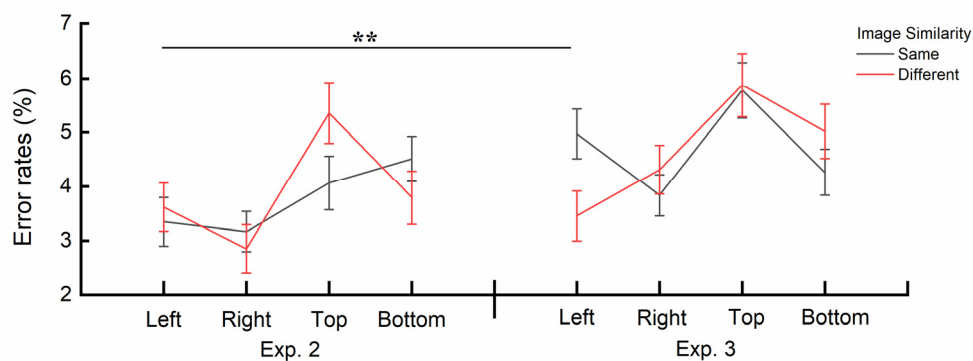


Figure 9. The three-way interaction effect among Image Similarity, Image Position and Group. Error bars represent SEM. ** $p < .01$.

4.5. Discussion

The results combining Experiments 2 and 3 were consistent with those of the two separate experiments. Firstly, IOR was found in both RTs and error rates. Furthermore, the RT results indicated that participants responded faster to horizontal orientations than to vertical orientations,

and faster when the cue image matched the central image compared to when they were different. Additionally, a significant group difference was found in both RTs and error rates, with Experiment 2 showing slower response and higher accuracy than Experiment 3. The faster response in Experiment 3 might be due to the use of fully ordered images as the different cue images, leading to faster information processing. The accompanying lower accuracy could be due to the faster responses (i.e., speed-accuracy-tradeoff [59,60]). As for the IOR, the RT results showed that IOR was significantly greater when the cue image was presented on the bottom compared to the top. This result indicated a significant difference in IOR between the two vertical positions, with greater IOR observed for stimuli presented at the bottom. This effect, as found in Experiment 1b, is likely due to the attentional bias induced by the lower visual field [51–53], as what we found in Experiment 1b.

Notably, the analysis of error rates revealed a three-way interaction effect among Image Similarity, Image Position and Group. Specifically, when presented on the left, the *discrepancy score of Image Similarity* showed significant differences between the two experiments. In Experiment 2, the error rate was slightly higher for cue images different from the central image compared to when the two images were the same, while in Experiment 3, it was the opposite (see [Figure 9](#)). This difference may be attributed to the fact that, in Experiment 3, the cue images different from the central images were complete Gestalt structures. This finding suggests that when cue images are processed more efficiently (e.g., fully ordered images that are easier to process and require less cognitive effort [61]), structures represented in short-term memory are more strongly suppressed at locations metaphorically associated with past temporal meanings. As shown in [Figure 9](#), error rates under the same conditions for the left and top positions in Experiment 3 were consistently higher than those in Experiment 2, with the increase on the left side being particularly pronounced.

5. General Discussion

The present study reveals how the *spatial metaphor of time* shapes attentional bias, as well as how visual short-term memory representations and Gestalt perceptual grouping modulate this effect. Our investigation focused on three main questions: how attentional bias shifts with spatial orientation associated with the temporal concept, how short-term memory representations influence this bias, and whether short-term memory representations affect attention similarly to fully ordered structures aligned with Gestalt principles. These findings shed new light on the dynamic roles of spatial cues and memory in guiding attention.

The first finding supported the *spatial metaphor of time* [5,6,62,63], demonstrating an inherent future-oriented bias in attentional processing. Specifically, we observed a marked rightward attentional bias in the horizontal orientation in Experiment 1a. The spatial cueing paradigm employed in this study effectively activated implicit spatial-temporal metaphors by relying solely on participants' responses to probe dot at different positions, without inducing additional cognitive operations related to stimuli or spatial factors. This approach allowed us to detect implicit spatial-temporal associations through differences in reaction times and error rates. While previous research has focused on memory-driven, past-oriented attention [28–31], our findings using the spatial cueing paradigm provided stronger evidence for a natural future-oriented spatial bias. However, hemispheric processing biases may also contribute to these findings. Geffen et al. [64] found that stimuli processed in the right visual field (RVF) are often associated with faster responses due to the left hemisphere's specialization for verbal processing. These interhemispheric effects, including transmission delays across the corpus callosum, could partially explain the observed rightward bias in Experiment 1a. Nevertheless, the systematic manipulation of disorder levels in our study revealed shifts in attentional biases that cannot be fully accounted for by hemispheric asymmetries alone. For example, in the 1/4D condition, we observed a pattern opposite to the rightward attentional bias found under higher disorder levels (1/2AD and 1/2ND), indicating that verbal processing does not play a key role in shaping attentional biases. In Experiment 1b, the downward bias was less pronounced but, together with the findings from Experiment 1a, provided further evidence that spatial-temporal associations play a more dominant role in shaping attentional biases. This

preference for future-related positions [8,48] highlights the consistent influence of spatial-temporal metaphors across different orientations. Additionally, the findings from Experiments 1a and 1b, together with recent work by Liu et al. [13], could deepen our understanding of the *spatial metaphor of time*. These findings suggest that these metaphors are not merely symbolic mappings but functionally embedded biases within our perceptual systems. In other words, the future-oriented bias observed here is not just an abstract concept (such as “left for past, right for future”) but an automatic attentional shift deeply rooted in perceptual processing. This directional bias could help us respond more quickly and effectively to future-related information, reflecting an adaptive mechanism that facilitates anticipatory processing and enhances our interactions with the environment.

When introducing short-term memory representations in Experiment 2, we found that participants responded faster and more accurately when the central image matched the cue image. This finding aligned with our expectation, suggesting that visual short-term memory elicited attentional bias and is consistent with previous research on the influence of working and sensory memory on attention [37–40]. Both RTs and error rates showed significant inhibition of return (IOR), which was absent in Experiment 1. This result suggested that short-term memory representations may play a key role in suppressing previously processed information and triggering IOR. Although Reppa et al. [24] observed that attentional bias through object-based cueing can improve processing speed and accuracy by enhancing target detection and identification, our findings demonstrate that IOR may counteract this benefit by limiting reorienting, especially when short-term memory cues are involved. This effect was particularly strong when the central and cue images were identical, increasing suppression at the cue image location. This finding provided insight into how short-term memory representations may selectively facilitate attentional engagement, intensifying suppression where these cues were presented. Additionally, this selective direction of attentional resources was also evident in the prioritization of certain spatial orientations. The stronger IOR observed in horizontal orientations, compared to vertical ones, likely reflects the impact of habitual horizontal task processing, which has been shown to enhance processing efficiency and attentional stability along this orientation [65,66].

Experiment 3 reinforced the findings from Experiment 2, showing similar IOR patterns in both RTs and error rates, as well as faster and more accurate responses in horizontal orientations compared to vertical ones. However, unlike in Experiment 2, there were no significant differences in RTs or error rates between the two levels of Image Similarity (where cue and central images were either identical or different). This contrast likely arose from the variations in the ‘different cue image’ across experiments: in Experiment 3, the different cue image was a fully ordered structure, while in Experiment 2, it shared a similar level of disorder with the central image. The absence of the significant finding in Experiment 3 suggested that fully ordered and partially disordered structures held in short-term memory capture attention at a similar level. This outcome underscored the stability of fully ordered Gestalt structures in capturing attention, suggesting that their inherent perceptual organization can draw attentional resources comparably to short-term memory representations. As shown in studies on perceptual grouping, this process is largely pre-attentive, occurring even without spatial attention [67,68]. Specifically, the Gestalt principle of proximity allows for cohesive groupings to form automatically, enhancing the perceptual stability of ordered stimuli. Moreover, a recent study found that neural representations of grouping progress across early visual areas, with V3 exhibiting stronger representational strength for grouping than V1 and V2. This grouping effect was further amplified when attention was directed toward the grouped stimuli, indicating that attentional engagement can provide top-down modulation in V3, boosting the effects of grouping [69].

The combined analysis of Experiments 2 and 3 not only confirmed significant IOR but also the faster RTs when participants responded to cue images that matched the central image. This result suggested that representations maintained in short-term memory are processed more efficiently than other types of structures (including both completely ordered and partially disordered structures). Error rate outcomes revealed significant differences in IOR across positions, with the cue images displayed at the top showing significantly less IOR than those at the bottom. This result may be

influenced by the vertical presentation. In this setup, the bottom position likely draws more attention and triggers IOR as it is associated with future concepts. Moreover, the bottom position is situated in the lower visual field. Compared to the top, this placement results in higher processing efficiency [48,51–53], which may also contribute to a more pronounced IOR at the bottom than at the top position. In addition, there were noticeable differences in RTs and error rates between the groups. Responses in Experiment 3 were faster than in Experiment 2, yet the error rates were also higher. This discrepancy could be attributed to the speed-accuracy trade-off, where faster processing of cue images in Experiment 3 may have accelerated responses to the probe dot but reduced accuracy [59,60,70,71].

Notably, both Experiments 2 and 3 observed unique characteristics when cue images were presented on the left compared to other positions. In Experiment 2, RTs were slower when the central and cue images were identical on the left side, as opposed to the other three positions, particularly the right side. The error rate results showed a greater IOR on the left side than the other three positions, and significantly more so than at the top. In Experiment 3, error rates were higher when images on the left matched the central image compared to non-matching images. This pattern was not consistent with observations for images on the right, top, and bottom, and a significant difference was noted between the left and bottom positions. Additionally, a significant three-way interaction effect among image similarity, image position, and group was found. In Experiment 2, the error rate was higher for cue images different from the central image when presented on the left, while in Experiment 3, the opposite was true. These findings suggested a more pronounced IOR and suppression of structures represented in short-term memory when cue images are displayed on the left, particularly in Experiment 3 when the cue image differed from the central image and was completely ordered. This phenomenon is likely related to the psychological timeline, as demonstrated in Experiment 1, where individuals preferred future-related right space positions. Furthermore, the representation in short-term memory associated with the past and present on the left side might lead to compounded suppression effects, resulting in slower response or higher error rates. The differences in error rates on the left side between Experiments 2 and 3 further suggest that when the overall efficiency of processing cue images is high (i.e., due to the presentation of fully ordered images, which are easier to process and require less cognitive effort), there is stronger suppression of structures represented in short-term memory. This finding may be influenced by the predominance of right-handed participants in our study. Previous research has shown that right-handed individuals often perform significantly worse than chance on directional recall tasks, whereas left-handed individuals do not exhibit the same level of impairment [72]. Given that recall tasks involve retrieving past-related information, this may point to a potential suppression of past-related information among right-handed individuals. Therefore, it is essential to explore how handedness affects cognitive processes related to temporal information. Future research could explore whether different temporal contexts (e.g., past versus future) elicit distinct effects based on handedness, offering valuable insights into how individuals process information.

Taken together, this study highlights the dynamic interplay between spatial-temporal metaphors, short-term memory representations, and Gestalt principles in shaping attentional biases. The findings provide compelling evidence for a future-oriented attentional bias, demonstrate the role of short-term memory in modulating attention, and reveal the influence of Gestalt perceptual grouping on attentional allocation. However, several limitations of the study warrant consideration. The stimuli employed in our experiments consisted of simple geometric shapes, devoid of complex attributes or deeper semantic associations. Additionally, the experimental paradigm required participants to engage in relatively straightforward tasks. These design choices, while ensuring experimental control, may have constrained the ability to fully capture the potential effects of disorder on attention. Although Experiment 1a revealed that varying levels of disorder exert distinct influences on attentional biases, other underlying phenomena may not have been adequately captured or fully explained. Moreover, the generalizability of these findings remains limited and warrants further investigation. Future research should address these limitations by employing more

complex and meaningful stimuli and refining experimental paradigms to better capture the nuanced effects of order/disorder on cognitive mechanisms such as attention and memory. Such efforts would not only advance our understanding of how the brain interacts with the environment but also offer significant practical applications. For instance, insights from this research could inform the design of user interfaces in high-stakes domains, such as air traffic control or autonomous vehicles, by aligning designs with human attentional preferences. Additionally, understanding the neural mechanisms underlying attentional allocation could lead to the development of interventions or training programs aimed at optimizing cognitive performance in demanding environments, ultimately improving safety and efficiency in real-world applications.

6. Conclusions

The present study reveals how spatial-temporal metaphors shape attentional biases and how visual short-term memory representations and Gestalt perceptual grouping modulate this effect. Specifically, it demonstrated a future-oriented preference across spatial orientations, with a distinct rightward bias on the horizontal orientation and a downward tendency on the vertical orientation. Short-term memory was shown to not only elicit attentional bias but also suppress previously attended information, resulting in significant Inhibition of Return (IOR). Furthermore, Gestalt principles were found to play a fundamental role in visual attention, balancing the attentional capture driven by memory representations. These findings deepen our understanding of how spatial-temporal cues, memory, and perceptual organization interact to influence attention and provide practical insights for optimizing interface design and decision-making systems in safety-critical contexts.

Author Contributions: Conceptualization, Y.W.; methodology, YY.W. and E. V. d. Bussche.; formal analysis, YY.W.; investigation, YY.W.; writing—original draft preparation, YY.W.; writing—review and editing, All authors; supervision, E. V. d. Bussche and XM.Z; funding acquisition, XM.Z.. All authors have read and agreed to the published version of the manuscript.

Funding: The present study was supported by STI 2030—the Major Projects(2021ZD0200500) & STI2030-Major Projects+2021ZD0204300. The Key Program of National Natural Science Foundation of China (61632014).

Institutional Review Board Statement: The Institutional Review Board and Ethics Committee of Human Participant Protection, Faculty of Psychology at Beijing Normal University, approved the study (IRB number: BNU202310050144).

Informed Consent Statement: All participants gave their consent to participate in the study and were informed that their responses would be anonymous.

Data Availability Statement: The stimuli and the experimental data, including both raw and cleaned data are available on OSF (doi: 10.17605/OSF.IO/BDZ5K).

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Hurlstone, M.J.; Hitch, G.J.; Baddeley, A.D. Memory for Serial Order across Domains: An Overview of the Literature and Directions for Future Research. *Psychological Bulletin* **2014**, *140*, 339–373, doi:10.1037/a0034221.
2. Clark, H.H. SPACE, TIME, SEMANTICS, AND THE CHILD1. In *Cognitive Development and Acquisition of Language*; Moore, T.E., Ed.; Academic Press: San Diego, 1973; pp. 27–63 ISBN 978-0-12-505850-6.
3. Oliveri, M.; Koch, G.; Caltagirone, C. Spatial–Temporal Interactions in the Human Brain. *Exp Brain Res* **2009**, *195*, 489–497, doi:10.1007/s00221-009-1834-1.
4. Zampini, M.; Shore, D.I.; Spence, C. Audiovisual Temporal Order Judgments. *Exp Brain Res* **2003**, *152*, 198–210, doi:10.1007/s00221-003-1536-z.

5. Pan, L.; Huang, X. The Influence of Spatial Location on Temporal Order Perception. *Curr Psychol* **2024**, *43*, 2052–2061, doi:10.1007/s12144-022-04225-1.
6. Scozia, G.; Pinto, M.; Pellegrino, M.; Lozito, S.; Pia, L.; Lasaponara, S.; Doricchi, F. How Time Gets Spatial: Factors Determining the Stability and Instability of the Mental Time Line. *Atten Percept Psychophys* **2023**, *85*, 2321–2336, doi:10.3758/s13414-023-02746-w.
7. Beracci, A.; Fabbri, M. Vertical Mental Timeline Is Not Influenced by VisuoSpatial Processing. *Brain Sciences* **2024**, *14*, 184, doi:10.3390/brainsci14020184.
8. He, J.; Bi, C.; Jiang, H.; Meng, J. The Variability of Mental Timeline in Vertical Dimension. *Front. Psychol.* **2021**, *12*, doi:10.3389/fpsyg.2021.782975.
9. Leone, M.J.; Salles, A.; Pulver, A.; Golombek, D.A.; Sigman, M. Time Drawings: Spatial Representation of Temporal Concepts. *Consciousness and Cognition* **2018**, *59*, 10–25, doi:10.1016/j.concog.2018.01.005.
10. Shinohara, K. Conceptual Mappings from Spatial Motion to Time: Analysis of English and Japanese. In Proceedings of the Computation for Metaphors, Analogy, and Agents; Nehaniv, C.L., Ed.; Springer: Berlin, Heidelberg, 1999; pp. 230–241.
11. Marshuetz, C.; Smith, E.E.; Jonides, J.; DeGutis, J.; Chenevert, T.L. Order Information in Working Memory: fMRI Evidence for Parietal and Prefrontal Mechanisms. *Journal of Cognitive Neuroscience* **2000**, *12*, 130–144, doi:10.1162/08989290051137459.
12. Wise, T.B.; Barack, D.L.; Templer, V.L. Geometrical Representation of Serial Order in Working Memory. *Learn Behav* **2022**, *50*, 443–444, doi:10.3758/s13420-022-00541-5.
13. Liu, B.; Alexopoulou, Z.-S.; van Ede, F. Jointly Looking to the Past and the Future in Visual Working Memory. *eLife* **2024**, *12*, RP90874, doi:10.7554/eLife.90874.
14. Chen, Y.; Wang, Y.; Guo, S.; Zhang, X.; Yan, B. The Causal Future: The Influence of Shape Features Caused by External Transformation on Visual Attention. *Journal of Vision* **2021**, *21*, 17, doi:10.1167/jov.21.11.17.
15. Frey, H.-P.; Honey, C.; König, P. What's Color Got to Do with It? The Influence of Color on Visual Attention in Different Categories. *Journal of Vision* **2008**, *8*, 6, doi:10.1167/8.14.6.
16. Glavan, J.J.; Haggitt, J.M.; Houpt, J.W. Temporal Organization of Color and Shape Processing during Visual Search. *Atten Percept Psychophys* **2020**, *82*, 426–456, doi:10.3758/s13414-019-01863-9.
17. Zhou, Y.; Wu, F.; Wan, X.; Shen, M.; Gao, Z. Does the Presence of More Features in a Bound Representation in Working Memory Require Extra Object-Based Attention? *Mem Cognit* **2021**, *49*, 1583–1599, doi:10.3758/s13421-021-01183-0.
18. Ben-Av, M.B.; Sagi, D.; Braun, J. Visual Attention and Perceptual Grouping. *Perception & Psychophysics* **1992**, *52*, 277–294, doi:10.3758/BF03209145.
19. Freeman, E.; Sagi, D.; Driver, J. Lateral Interactions between Targets and Flankers in Low-Level Vision Depend on Attention to the Flankers. *Nat Neurosci* **2001**, *4*, 1032–1036, doi:10.1038/nn728.
20. Glicksohn, A.; Cohen, A. The Role of Gestalt Grouping Principles in Visual Statistical Learning. *Atten Percept Psychophys* **2011**, *73*, 708–713, doi:10.3758/s13414-010-0084-4.
21. Kimchi, R.; Yeshurun, Y.; Cohen-Savransky, A. Automatic, Stimulus-Driven Attentional Capture by Objecthood. *Psychonomic Bulletin & Review* **2007**, *14*, 166–172, doi:10.3758/BF03194045.
22. Kimchi, R.; Yeshurun, Y.; Spehar, B.; Pirkner, Y. Perceptual Organization, Visual Attention, and Objecthood. *Vision Research* **2016**, *126*, 34–51, doi:10.1016/j.visres.2015.07.008.

23. White, B. Attention, Gestalt Principles, and the Determinacy of Perceptual Content. *Erkenn* **2022**, *87*, 1133–1151, doi:10.1007/s10670-020-00234-3.
24. Reppa, I.; Schmidt, W.C.; Leek, E.C. Successes and Failures in Producing Attentional Object-Based Cueing Effects. *Atten Percept Psychophys* **2012**, *74*, 43–69, doi:10.3758/s13414-011-0211-x.
25. Cavanagh, P.; Caplovitz, G.P.; Lytchenko, T.K.; Maechler, M.R.; Tse, P.U.; Sheinberg, D.L. The Architecture of Object-Based Attention. *Psychon Bull Rev* **2023**, *30*, 1643–1667, doi:10.3758/s13423-023-02281-7.
26. Chen, Z. Object-Based Attention: A Tutorial Review. *Atten Percept Psychophys* **2012**, *74*, 784–802, doi:10.3758/s13414-012-0322-z.
27. Li, A.; Wolfe, J.M.; Chen, Z. Implicitly and Explicitly Encoded Features Can Guide Attention in Free Viewing. *Journal of Vision* **2020**, *20*, 8, doi:10.1167/jov.20.6.8.
28. Awh, E.; Vogel, E.K.; Oh, S.-H. Interactions between Attention and Working Memory. *Neuroscience* **2006**, *139*, 201–208, doi:10.1016/j.neuroscience.2005.08.023.
29. Downing, P.E. Interactions between Visual Working Memory and Selective Attention. *Psychol Sci* **2000**, *11*, 467–473, doi:10.1111/1467-9280.00290.
30. Kruijne, W.; Meeter, M. Implicit Short- and Long-Term Memory Direct Our Gaze in Visual Search. *Atten Percept Psychophys* **2016**, *78*, 761–773, doi:10.3758/s13414-015-1021-3.
31. Oberauer, K. Working Memory and Attention - A Conceptual Analysis and Review. *J Cogn* **2019**, *2*, 36, doi:10.5334/joc.58.
32. Delogu, F.; Nijboer, T.C.W.; Postma, A. Encoding Location and Serial Order in Auditory Working Memory: Evidence for Separable Processes. *Cogn Process* **2012**, *13*, 267–276, doi:10.1007/s10339-012-0442-3.
33. Galeano-Keiner, E.M.; Pakzad, S.; Brod, G.; Bunge, S.A. Examining the Role of Attentional Allocation in Working Memory Precision with Pupillometry in Children and Adults. *Journal of Experimental Child Psychology* **2023**, *231*, 105655, doi:10.1016/j.jecp.2023.105655.
34. Kar, B.R.; Kenderla, P.K. Working Memory and Executive Attention: Insights from Developmental Studies and Implications for Learning and Education. *J Indian Inst Sci* **2017**, *97*, 497–510, doi:10.1007/s41745-017-0044-2.
35. Cowan, N.; Bao, C.; Bishop-Chrzanowski, B.M.; Costa, A.N.; Greene, N.R.; Guitard, D.; Li, C.; Musich, M.L.; Ünal, Z.E. The Relation Between Attention and Memory. *Annual Review of Psychology* **2024**, *75*, 183–214, doi:10.1146/annurev-psych-040723-012736.
36. Rasoulzadeh, V.; Sahan, M.I.; van Dijck, J.-P.; Abrahamse, E.; Marzecova, A.; Verguts, T.; Fias, W. Spatial Attention in Serial Order Working Memory: An EEG Study. *Cereb Cortex* **2021**, *31*, 2482–2493, doi:10.1093/cercor/bhaa368.
37. Carlisle, N.B.; Kristjánsson, Á. How Visual Working Memory Contents Influence Priming of Visual Attention. *Psychological Research* **2018**, *82*, 833–839, doi:10.1007/s00426-017-0866-6.
38. Foerster, R.M.; Schneider, W.X. Task-Irrelevant Features in Visual Working Memory Influence Covert Attention: Evidence from a Partial Report Task. *Vision* **2019**, *3*, 42, doi:10.3390/vision3030042.
39. Trentin, C.; Slagter, H.A.; Olivers, C.N.L. Visual Working Memory Representations Bias Attention More When They Are the Target of an Action Plan. *Cognition* **2023**, *230*, 105274, doi:10.1016/j.cognition.2022.105274.
40. Xie, T.; Nan, W.; Fu, S. Attention Can Operate on Object Representations in Visual Sensory Memory. *Atten Percept Psychophys* **2021**, *83*, 3069–3085, doi:10.3758/s13414-021-02323-z.

41. Aben, B.; Stapert, S.; Blokland, A. About the Distinction between Working Memory and Short-Term Memory. *Front Psychol* **2012**, *3*, 301, doi:10.3389/fpsyg.2012.00301.
42. Colom, R.; Shih, P.C.; Flores-Mendoza, C.; Quiroga, M.Á. The Real Relationship between Short-Term Memory and Working Memory. *Memory* **2006**, *14*, 804–813, doi:10.1080/09658210600680020.
43. Posner, M.I.; Snyder, C.R.; Davidson, B.J. Attention and the Detection of Signals. *Journal of Experimental Psychology: General* **1980**, *109*, 160–174, doi:10.1037/0096-3445.109.2.160.
44. Faul, F.; Erdfelder, E.; Lang, A.-G.; Buchner, A. G*Power 3: A Flexible Statistical Power Analysis Program for the Social, Behavioral, and Biomedical Sciences. *Behavior Research Methods* **2007**, *39*, 175–191, doi:10.3758/BF03193146.
45. Van Geert, E.; Bossens, C.; Wagemans, J. The Order & Complexity Toolbox for Aesthetics (OCTA): A Systematic Approach to Study the Relations between Order, Complexity, and Aesthetic Appreciation. *Behav Res* **2022**, doi:10.3758/s13428-022-01900-w.
46. Tse, C.-S.; Altarriba, J. Recognizing the Directionality of an Arrow Affects Subsequent Judgments of a Temporal Statement: The Role of Directionality in Spatial Metaphors. *Psychol Rec* **2012**, *62*, 497–506, doi:10.1007/BF03395816.
47. Weger, U.W.; Pratt, J. Time Flies like an Arrow: Space-Time Compatibility Effects Suggest the Use of a Mental Timeline. *Psychonomic Bulletin & Review* **2008**, *15*, 426–430, doi:10.3758/PBR.15.2.426.
48. Dalmaso, M.; Schnapper, Y.; Vicovaro, M. When Time Stands Upright: STEARC Effects along the Vertical Axis. *Psychological Research* **2023**, *87*, 894–918, doi:10.1007/s00426-022-01693-9.
49. Hartmann, M.; Martarelli, C.S.; Mast, F.W.; Stocker, K. Eye Movements during Mental Time Travel Follow a Diagonal Line. *Conscious Cogn* **2014**, *30*, 201–209, doi:10.1016/j.concog.2014.09.007.
50. Stocker, K.; Hartmann, M.; Martarelli, C.S.; Mast, F.W. Eye Movements Reveal Mental Looking Through Time. *Cogn Sci* **2016**, *40*, 1648–1670, doi:10.1111/cogs.12301.
51. Danckert, J.; Goodale, M.A. Superior Performance for Visually Guided Pointing in the Lower Visual Field. *Exp Brain Res* **2001**, *137*, 303–308, doi:10.1007/s002210000653.
52. Gottwald, V.M.; Lawrence, G.P.; Hayes, A.E.; Khan, M.A. Representational Momentum Reveals Visual Anticipation Differences in the Upper and Lower Visual Fields. *Exp Brain Res* **2015**, *233*, 2249–2256, doi:10.1007/s00221-015-4294-9.
53. Soret, R.; Prea, N.; Peysakhovich, V. Exploring the Impact of Body Position on Attentional Orienting. *Information* **2024**, *15*, 111, doi:10.3390/info15020111.
54. Posner, M.I.; Rafal, R.D.; Choate, L.S.; Vaughan, J. Inhibition of Return: Neural Basis and Function. *Cognitive Neuropsychology* **1985**, *2*, 211–228, doi:10.1080/02643298508252866.
55. Han, S.; Ding, Y.; Song, Y. Neural Mechanisms of Perceptual Grouping in Humans as Revealed by High Density Event Related Potentials. *Neuroscience Letters* **2002**, *319*, 29–32, doi:10.1016/S0304-3940(01)02502-2.
56. Yu, D.; Park, H.; Gerold, D.; Legge, G.E. Comparing Reading Speed for Horizontal and Vertical English Text. *Journal of Vision* **2010**, *10*, 21, doi:10.1167/10.2.21.
57. Hilchey, M.D.; Klein, R.M.; Satel, J.; Wang, Z. Oculomotor Inhibition of Return: How Soon Is It “Recoded” into Spatiotopic Coordinates? *Atten Percept Psychophys* **2012**, *74*, 1145–1153, doi:10.3758/s13414-012-0312-1.

58. Taylor, T.L.; Klein, R.M. On the Causes and Effects of Inhibition of Return. *Psychonomic Bulletin & Review* **1998**, *5*, 625–643, doi:10.3758/BF03208839.
59. Cohen, R.A. Processing Speed and Attentional Resources. In *The Neuropsychology of Attention*; Cohen, R.A., Ed.; Springer US: Boston, MA, 2014; pp. 751–762 ISBN 978-0-387-72639-7.
60. Maolin, Y. Speed-Accuracy Tradeoff. In *The ECPH Encyclopedia of Psychology*; Springer Nature: Singapore, 2024; pp. 1–1 ISBN 978-981-9960-00-2.
61. Matthew, J.S.; Michael, A.N. Gestalt and Feature-Intensive Processing: Toward a Unified Model of Human Information Processing. *Curr Psychol* **2002**, *21*, 68–84, doi:10.1007/BF02903160.
62. Bonato, M.; Zorzi, M.; Umiltà, C. When Time Is Space: Evidence for a Mental Time Line. *Neuroscience & Biobehavioral Reviews* **2012**, *36*, 2257–2273, doi:10.1016/j.neubiorev.2012.08.007.
63. von Sobbe, L.; Scheifele, E.; Maienborn, C.; Ulrich, R. The Space–Time Congruency Effect: A Meta-Analysis. *Cognitive Science* **2019**, *43*, e12709, doi:10.1111/cogs.12709.
64. Geffen, G.; Bradshaw, J.L.; Wallace, G. Interhemispheric Effects on Reaction Time to Verbal and Nonverbal Visual Stimuli. *Journal of Experimental Psychology* **1971**, *87*, 415–422, doi:10.1037/h0030525.
65. Miwa, K.; Dijkstra, T. Lexical Processes in the Recognition of Japanese Horizontal and Vertical Compounds. *Read Writ* **2017**, *30*, 791–812, doi:10.1007/s11145-016-9700-6.
66. Rakover, S.S. Explaining the Face-Inversion Effect: The Face–Scheme Incompatibility (FSI) Model. *Psychon Bull Rev* **2013**, *20*, 665–692, doi:10.3758/s13423-013-0388-1.
67. Kubovy, M.; Holcombe, A.O.; Wagemans, J. On the Lawfulness of Grouping by Proximity. *Cognitive Psychology* **1998**, *35*, 71–98, doi:10.1006/cogp.1997.0673.
68. Sasaki, Y. Processing Local Signals into Global Patterns. *Current Opinion in Neurobiology* **2007**, *17*, 132–139, doi:10.1016/j.conb.2007.03.003.
69. Wu, H.; Zuo, Z.; Yuan, Z.; Zhou, T.; Zhuo, Y.; Zheng, N.; Chen, B. Neural Representation of Gestalt Grouping and Attention Effect in Human Visual Cortex. *Journal of Neuroscience Methods* **2023**, *399*, 109980, doi:10.1016/j.jneumeth.2023.109980.
70. Heitz, R.P. The Speed-Accuracy Tradeoff: History, Physiology, Methodology, and Behavior. *Front. Neurosci.* **2014**, *8*, doi:10.3389/fnins.2014.00150.
71. Wickelgren, W.A. Speed-Accuracy Tradeoff and Information Processing Dynamics. *Acta Psychologica* **1977**, *41*, 67–85, doi:10.1016/0001-6918(77)90012-9.
72. Jones, G.V.; Martin, M. Handedness Dependency in Recall from Everyday Memory. *British Journal of Psychology* **1997**, *88*, 609–619, doi:10.1111/j.2044-8295.1997.tb02660.x.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.