Cognition 256 (2025) 106041

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

Full Length Article

Hidden size: Size representations in implicitly coded objects

Elif Memis [a,*], Gizem Y. Yildiz [a], Gereon R. Fink [a,b], Ralph Weidner [a]

[a] Institute of Neuroscience and Medicine (INM-3), Forschungszentrum Jülich GmbH, Jülich, Germany
[b] Department of Neurology, University Hospital Cologne, Cologne University, Germany

ARTICLE INFO

Keywords:
Size illusion
Ebbinghaus illusion
Object-substitution masking (OSM)
Size averaging
Rescaled size perception in masked objects
Context integration

ABSTRACT

Its angular representation on the retina does not solely determine the perceived size of an object. Instead, contextual information is interpreted. We investigated the levels of processing at which this interpretation occurs. Combining three experimental paradigms, we explored whether masked and more implicitly coded objects are already size-rescaled. We induced object size rescaling using a modified variant of the Ebbinghaus illusion. In this variant, six dots altered the size of a central stimulus and served as inducers generating Object-Substitution Masking (OSM). Participants reported the average size of multiple circles using the size-averaging paradigm, allowing us to test the contribution of masked and non-masked central target circles. Our Ebbinghaus illusion variant altered perceived stimulus size and showed robust masking via OSM. Furthermore, size-averaging was sensitive enough to detect perceived size changes in the magnitude of the ones induced by the Ebbinghaus illusion. Finally, combining all three paradigms, we observed that masked and non-masked stimuli contributed to size averaging in a size-rescaled manner. In a control experiment testing the general effects of Ebbinghaus inducers, we observed a contrast-like effect on size averaging. Large inducers decreased the perceived average size, while small inducers increased it. In summary, our experiments indicate that context integration, induced by the Ebbinghaus illusion, alters size representations at an early stage. These modified size representations are independent of whether a target is recognisable. Moreover, perceived average size appears to be coded relative to surrounding perceptual groups.

1. Introduction

Perceiving our environment is an overwhelming computational challenge for our visual system, given that the processing capacities of our perceptual system are limited. At the same time, the amount of information in the outside world is infinite. One strategy the brain uses to cope with the vast amount of information entering our visual system via the retina is to process ensemble representations rather than individual items in the visual scene (Ariely, 2001; Chong & Treisman, 2005). The visual system forms ensemble representations by averaging certain visual features (i.e., size, orientation, speed) of items belonging to the same perceptual group (Alvarez & Oliva, 2008; Ariely, 2001; Watamaniuk & Duchon, 1992) and hence can efficiently code the average size of multiple objects. Interestingly, these representations only indirectly contain individual item features, so while observers can compare the average size of groups of items, they struggle to report visual features of individuated items in these groups (Ariely, 2001).

A recent study demonstrated that the perceived average size of a group of items changes based on size-distance rescaling mechanisms, indicating that the average estimates were computed after size-distance rescaling (Markov & Tiurina, 2021). Thus, the ensemble representation contributing to the formation of an average size of multiple objects is not merely a low-level coding of the angular size represented on the retina. Instead, it is a size value that resembles size as we perceive it. This is particularly intriguing in light of previous findings indicating that the brain estimates the average size of a group of items by considering the size of items that are not consciously perceived (Choo & Franconeri, 2010). These findings suggest that it should be possible to assess the size representations of masked and, hence, more implicitly coded objects and test whether context information modifies their size representations.

To answer the question of whether size averaging involves the retinal size of masked and, hence, more implicitly coded objects or if it is based on perceived size, i.e., involving size representations that are already size-rescaled, we combined object-substitution masking (OSM) (Enns &

* Corresponding author at: Cognitive Neuroscience, Institute of Neuroscience and Medicine (INM3) Forschungszentrum Jülich, Leo-Brand-Str. 5, 52425 Jülich, Germany, Department of Neurology, University Hospital Cologne, Cologne University, Kerpener Str. 62, 50937 Köln, Germany.

E-mail addresses: e.memis@fz-juelich.de (E. Memis), Gizem.Yildiz@psychiat.med.uni-giessen.de (G.Y. Yildiz), g.r.fink@fz-juelich.de (G.R. Fink), r.weidner@fz-juelich.de (R. Weidner).

https://doi.org/10.1016/j.cognition.2024.106041
Received 14 June 2024; Received in revised form 4 December 2024; Accepted 5 December 2024
Available online 12 December 2024

Di Lollo, 1997) with the Ebbinghaus illusion. The paradigms were linked so that the stimuli constituting the mask also served as inducers generating the Ebbinghaus illusion.

In OSM, an object is masked if the surrounding dots persist on the screen longer than the object itself. OSM is ideal for studying implicitly coded objects since it is assumed to alter only later representations while keeping early representations intact (Di Lollo, Enns, & Rensink, 2000). Changing the size of these dots allows the generation of Ebbinghaus-like inducer patterns. In the classic Ebbinghaus illusion, a target surrounded by small Ebbinghaus inducers appears larger than its physical size, while a target surrounded by large Ebbinghaus inducers appears smaller than its physical size (Ebbinghaus, 1902; Massaro & Anderson, 1971). To investigate early contextual modulations of size, it is important to use an illusion that operates at early levels of visual processing. It is known that the Ebbinghaus illusion is weaker when the inducer and target stimuli are presented to separate eyes, indicating the involvement of monocular neurons in early visual processing (Nakashima & Sugita, 2018; Song, Schwarzkopf, & Rees, 2011). This is also supported by fMRI studies, which indicate an important role for V1 in the Ebbinghaus illusion (Schwarzkopf & Rees, 2013; Schwarzkopf, Song, & Rees, 2011). The Ebbinghaus illusion is therefore particularly well suited as a means of altering early size representations. In this study, six surrounding dots were presented to operate as Ebbinghaus inducers and induce OSM. This way, we could manipulate both the perceived size of the target objects and their level of encoding.

To establish that each of these paradigms generates the expected effects, three behavioural experiments were conducted before the main experiment, testing whether the stimulus configurations used in our experiments a) allow to induce changes in perceived sizes using our Ebbinghaus stimulus configuration (Experiment 1), b) allow to generate a robust masking effect using OSM (Experiment 2) and c) allow to confirm that size averaging was sensitive enough to detect size changes in the magnitude of the ones induced by our Ebbinghaus stimulus configuration (Experiment 3). Having successfully tested these three phenomena in isolation, we combined the Ebbinghaus illusion, OSM and size averaging into one paradigm (Experiment 4 A-B) to investigate whether masked and, therefore, less recognisable objects are already size-rescaled. Participants were presented with a typical size-averaging paradigm, while six red inducers surrounded three out of eight green circles to induce both the Ebbinghaus illusion and OSM.

We hypothesised that the size of the inducers would generate an Ebbinghaus effect on the target objects, leading to a change in their perceived size, either an increase or a decrease. Additionally, because the target objects were included in the group of items used for the size averaging task, we expected that the altered size representations would also influence the perceived average size, as previously demonstrated by Markov and Tiurina (2021). The current study's experimental design also permits investigating the role of the recognisability of the target stimulus and its contribution to perceived size averaging. In particular, if context integration via the Ebbinghaus illusion alters size representations before an object is recognised, then the size of even masked objects would incorporate size-averaging mechanisms in a size-rescaled manner. If this is the case, participants would report a larger perceived average size in blocks with small inducers and a smaller perceived size in blocks with large inducers, regardless of whether the targets are masked. This pattern would be consistent with size-scaling effects that operate at an early level of processing, even before a stimulus is recognised.

Alternatively, if masked objects are not yet affected by context effects and, therefore, unchanged by the Ebbinghaus illusion, we would expect no difference between large and small inducer blocks.

2. Methods

2.1. Overview of experiments

To investigate whether context-induced size-rescaling affects representations of masked objects, we combined OSM with the Ebbinghaus illusion in a size-averaging paradigm. In a series of experiments, we tested each of these paradigms separately to confirm the efficiency of the respective experimental manipulation.

2.2. General setup

All experiments were presented via Visual Studio 1.68.1 using PsychoPy 2021.2.3 scripts on an AORUS F048U 47.53-in. monitor at a distance of 57 cm. The screen resolution was 1920 × 1080 pixels with a refresh rate of 120 Hz. The distance between participants and the monitor was maintained using a chin- and forehead rest. Each participant completed the experiment in a darkened room, and self-paced breaks were presented between the experimental blocks. All four experiments involve the presentation of a circular array (10 degrees of visual angle) of equally spaced eight green (76.71 cd/m^2) circles around a black (0.00 cd/m^2) fixation cross. Red inducers with a luminance of 18.27 cd/m^2 were used in all experiments, except for Experiment 3. The distance between the large inducers and the targets was 1.5 degrees of visual angle and the distance between the small inducers and the target was 1.2 degrees of visual angle. The distance between adjacent green circles was 3.8 degrees of visual angle. Additionally, the distance from the centre to the green circles is 5 degrees of visual angle, and the distance from the centre to the nearest inducer was 3.5 degrees of visual angle. The distances between the stimuli are comparable to those in similar studies (Choo & Franconeri, 2010; Jacoby, Kamke, & Mattingley, 2013). Experimental stimuli were presented over a grey (49.90 cd/m^2) background at 40 degrees of visual angle over a black screen. Pre-cues (used in Experiment 1) and post-cues (used in Experiment 2) shared the same luminance value as the green circles (76.71 cd/m^2).

Experiment 1 - Screening: Ebbinghaus illusion.

Experiment 1 had two primary objectives. First, we aimed to test whether our stimulus configurations reliably altered a stimulus's perceived size by surrounding it with either large or small inducers. Second, we aimed to identify participants exhibiting a strong illusion effect, who we recruited for the subsequent experiments.

2.3. Participants

Sixty healthy participants (32 females) attended Experiment 1. To identify participants with a strong illusion, the criterion for inclusion in subsequent experiments required participants to exhibit an illusion strength of at least 10 % (>10 %) in small (0.9 degrees of visual angle) target conditions of Experiment 1 (see Stimuli subsection for details). A 10 % illusion effect is comparable to typical values observed in the previous studies (Chen, Qiao, Wang, & Jiang, 2018; Chen, Wu, Yu, & Sperandio, 2024; Wu, Feng, Han, Chen, & Luo, 2023). This criterion applied to a group of twenty-nine participants ($M = 29.38$ years, $SD = 5.52$, 17 females) who participated in the remaining experiments on separate days. All participants had normal or corrected-to-normal vision. We obtained written informed consent before the experiment and paid all participants 10 euros per hour for their participation. The ethics committee of the German Society of Psychology approved the study. The sample size was determined based on effect sizes detected in previous studies (Choo & Franconeri, 2010) to ensure the current sample size would be sufficient to reveal significant differences between the experimental conditions (in a repeated-measures analysis of variance) with 95 % power and an alpha level of 0.05. Power estimates were computed using G*Power (Erdfelder, Faul, & Buchner, 1996).

2.4. Stimuli

One out of eight circles was randomly selected as a target circle, surrounded by either large (Fig. 1A) or small (Fig. 1B) circular inducers, generating an Ebbinghaus figure. We used the same stimulus configuration but without Ebbinghaus inducers as a control condition (Fig. 1C). Since there were no inducers in the control condition to locate the target circle, a pre-cue (Fig. 1, time window 2) was employed in all experimental conditions to indicate the target position. The target size was either 0.9 (small target) or 1.1 (large target) degrees of visual angle. Incorporating two different target sizes allowed us to test whether participants correctly performed the task.

The small and large inducers were always presented at 0.7 and 1.3 degrees of visual angle, respectively. The average size of the seven distractor circles was determined in accordance with the target size to keep the physical average size of the screen equal across different target sizes. In detail, the distractor average size was 0.9 degrees of visual angle for the small target and 1.1 degrees of visual angle for the large target. The average size of the distractors was calculated based on the normal cumulative distribution in all experiments. Specifically, the size of each circle was varied with a standard deviation of 0.15 degrees of visual angle around the mean (0.9 or 1.1 degrees of visual angle). This resulted in distinct sizes for each distractor circle in each trial while keeping the mean size of the seven distractors constant.

The Method of Constant Stimuli was used to detect the perceived size of the target circle. The comparison circle's size varied around the target stimulus's with 0.1 degrees of visual angle increments, resulting in two different comparison size lists. Ten different comparison sizes were used (half of them were smaller than the size of the target stimulus, and the other half was larger). Each comparison circle was presented ten times.

2.5. Procedure

Participants were presented with a size judgment task where each trial started with a 100 ms fixation in which only the fixation cross was presented over a light grey background, followed by a pre-cue for 1000 ms (Fig. 1). The pre-cue was used to indicate the position of the target circle. After a fixation presentation for 100 ms, green circles appeared around the fixation cross in a circular array. Participants saw red inducers around the pre-cued green circle in the experimental blocks with the small and large inducers. All stimuli disappeared simultaneously after 32 ms, and only the fixation cross was displayed on the screen for 320 ms. A green comparison circle was then displayed at the centre of the screen until participants pressed a button to indicate whether the comparison circle was larger or smaller than the target circle (indicated by a pre-cue). Each trial ended with a grey noise pattern for 500 ms. Participants completed 600 trials (2 target sizes × 3 inducer types × 10 comparison sizes × 10 repetitions), resulting in an experiment length of around 40 min.

Statistical Analyses.

All statistical analyses were performed using the JASP software package version 0.16.0.0 (University of Amsterdam, Amsterdam, Netherlands). We measured the perceived size of the target stimulus in Experiment 1. To assess how participants perceived the target stimulus's size, we generated psychometric curves for each participant across six experimental conditions. These curves were constructed by analysing response proportions at precise intervals of 0.1 degrees, showing a preference for the comparison circle being larger than the target circle. We employed the logistic function to quantify the probability (P) of perceiving the comparison circle as larger than the target circle. Subsequently, the Point of Subjective Equality (PSE) was calculated as $P =$ .5, representing the point where the comparison circle was perceived at an equivalent size to the target circle. Width of each psychometric curve was calculated as $P_{0.75}$ - $P_{0.5}$, representing the degree of uncertainty in the participants' responses for each experimental condition. Higher values of curve width indicate a greater perceptual uncertainty. We calculated goodness of fit measures when fitting psychometric curves to the data. The obtained curves demonstrated a strong fit in Experiment 1 (r ranged between 0.728 and 0.989).

The PSEs are anticipated to shift towards the right in the case of small inducer conditions, indicating that the target size is perceived as larger than its physical size. Conversely, we would anticipate an opposite shift towards the left in large inducer conditions, indicating that the target size is perceived as smaller than its physical size.

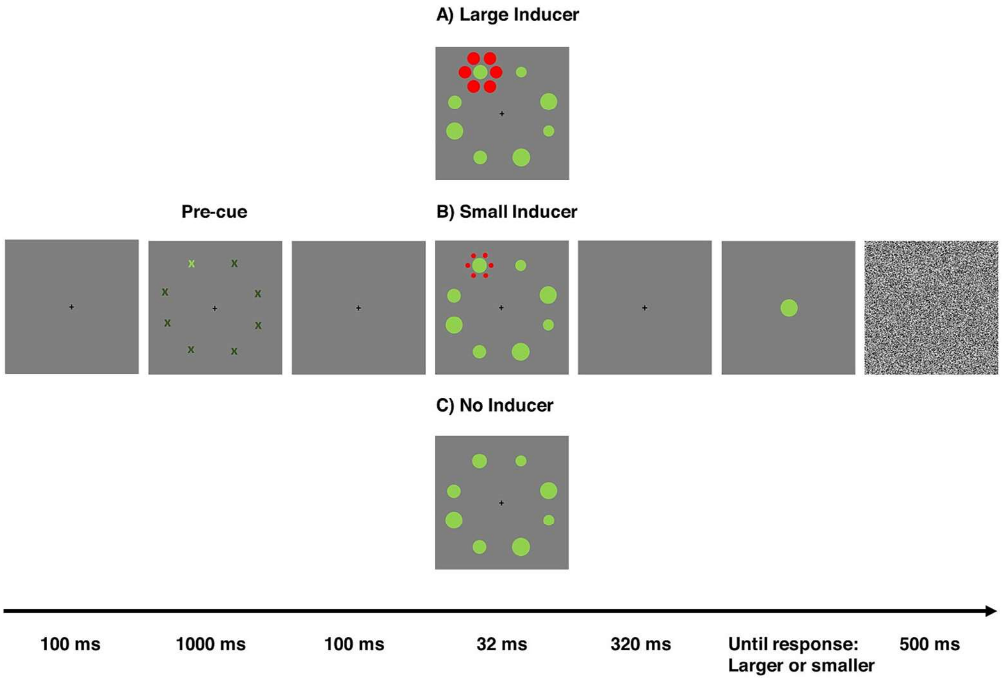

Fig. 1. Illustration of the task in Experiment 1. Participants were asked to determine whether the comparison circle (time window 6) was larger or smaller than the target circle (time window 4). The position of the target circle was indicated by a pre-cue (time window 2). The type of inducer was manipulated as either a large inducer (A) or a small inducer (B), while no inducers (C) were presented in the control condition.

In Experiment 1, the obtained PSEs and curve widths were entered into separate 2 × 3 ANOVAs with the factors target size (small, large) and inducer type (small inducer, large inducer, no inducer) to test whether the Ebbinghaus illusion altered the stimulus size.

3. Results

Fig. 2 displays the mean PSEs with standard errors for within-subject contrast for the small (Fig. 2A) and large (Fig. 2B) target sizes, along with different inducer-type conditions. ANOVA revealed significant main effects for both target size [small vs. large] (F (1, 28) = 111.81, $p <$ .001, $\eta_p^2 = 0.800$) and inducer type [small inducer vs. large inducer vs. no inducer] (F (2, 56) = 46.35, $p <$ .001, $\eta_p^2 = 0.623$). Moreover, the interaction between target size and inducer type was also significant (F (2, 56) = 5.86, p = .005, $\eta_p^2 = 0.173$).

Above all, participants estimated the perceived size of the target circle as larger in the small inducer conditions (M = 0.92, SE = 0.03) than in the large inducer conditions (M = 0.76, SE = 0.03) [planned t-test, t (28) = 8.40, *Bonferroni corrected* $p <$ .001, Cohen's d = 1.560], indicating a robust illusory size judgment (Fig. 2A-B). Furthermore, target size estimation was greater in the no inducer conditions (M = 0.93, SE = 0.02) compared to the large inducer conditions (M = 0.76, SE = 0.03) [planned t-test, t (28) = −8.96, *Bonferroni corrected* $p <$ .001, Cohen's d = −1.664]. However, the mean PSEs did not show a significant difference between the conditions with small inducers and those without inducers for any target size [planned t-test, t (28) = 0.59, *Bonferroni corrected* p = 1.870, Cohen's d = 0.109].

Consistent with our expectations, the averaged PSEs were higher in the large target condition (M = 0.93, SE = 0.03) than in the small target condition (M = 0.81, SE = 0.02). The inducer type and target size interaction suggests that the effect of inducer type on the mean PSEs varies based on target size. Overall, these findings indicate that inducer type significantly impacted participants' size judgment (illusion effect for small target: 19.12 %), showing that our variant of the Ebbinghaus illusion significantly altered the perceived size of our target stimulus.

Curve widths for the small and large targets were compared across inducer types (small inducer, large inducer, no inducer) by conducting a 2 × 3 ANOVA. ANOVA showed a significant main effect for inducer type [small inducer vs. large inducer vs. no inducer] (F (2, 55) = 3.256, p = .047, Greenhouse-Geisser corrected, $\eta_p^2 = 0.104$). Neither the main effect of target size (p = .137) nor the interaction between target size and inducer type was significant (p = .282). Specifically, the estimated curve widths were significantly greater in the small inducer conditions (M = 0.21, SE = 0.02) compared to the no inducer conditions (M = 0.17, SE = 0.01) [planned t-test, t (28) = −2.50, *Bonferroni corrected* p = .046, Cohen's d = −0.419]. However, the estimated curve widths did not reveal a significant difference between the large inducer conditions and the no inducer conditions (p = .286), and between the large inducer conditions and the small inducer conditions ($p \geq$ .999).

Experiment 2: Object substitution masking.

Since our primary goal was to investigate hidden size representations, we employed OSM to reduce target recognisability. Specifically, we tested whether the stimuli used as Ebbinghaus-like inducers in Experiment 1 could efficiently induce OSM by manipulating their presentation times. In a simultaneous viewing condition, the mask and the green circles disappeared simultaneously after 32 ms, aiming to preserve their accessibility levels to conscious visual awareness. In the delayed viewing condition, the mask persisted on the screen 320 ms longer than the green circles, inducing a masking effect.

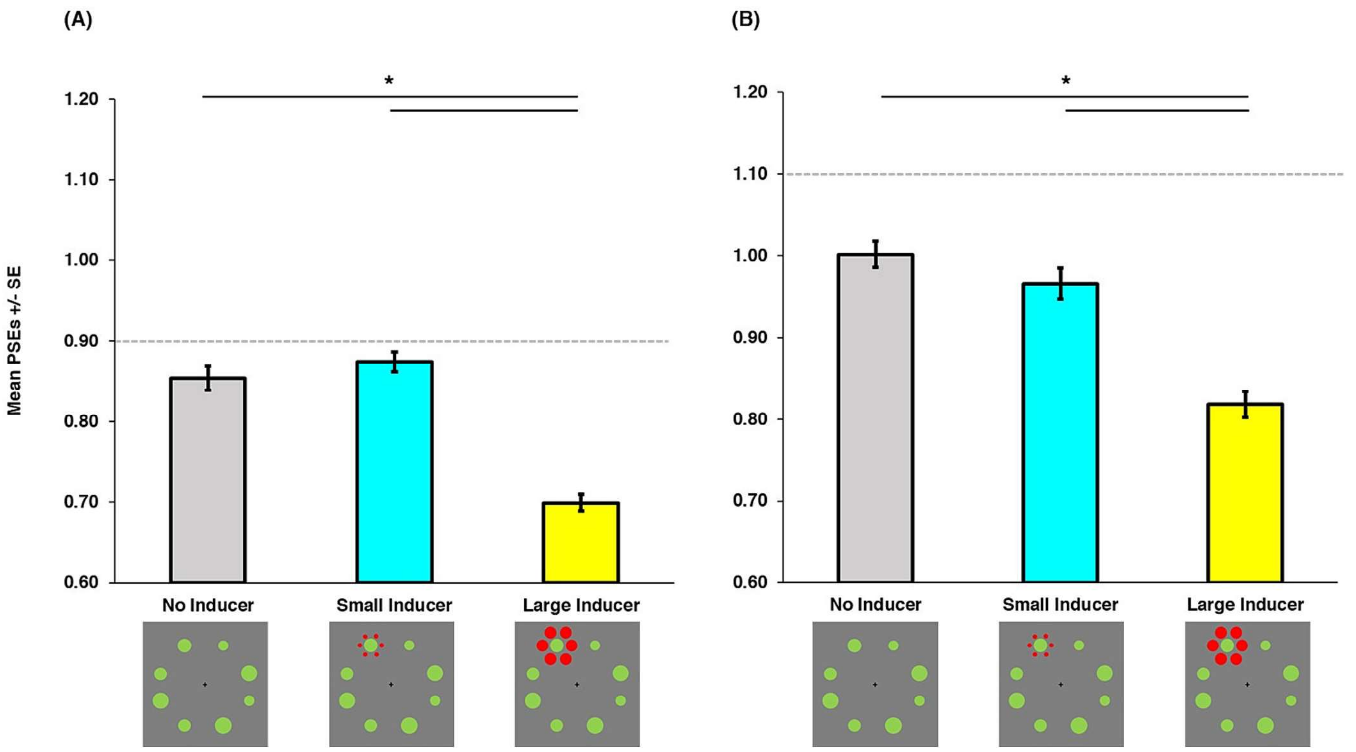

Fig. 2. Perceived size of the target stimulus for small (A) and large (B) target sizes in Experiment 1. Averaged PSEs across different inducer types were plotted. Grey bars indicate the no inducer condition, blue bars represent the small inducer condition and yellow bars represent the large inducer condition. Asterisks (*) represent significant differences at $p <$ .05. Error bars indicate the standard errors around the mean for within-subject contrasts (O'Brien & Cousineau, 2014). The horizontal dashed grey lines represent the physical size of the target stimulus. The figures shown below the x-axis are illustrations of the corresponding experimental conditions. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

4. Method

4.1. Participants

Experiment 2 featured a group of twenty-nine participants (M = 29.38 years, SD = 5.52, 17 females), all exhibiting an illusion effect of >10 % in Experiment 1.

4.2. Stimuli

Eight green circles were presented around the fixation cross, and two circles, randomly chosen (one from the right visual field, the other one from the left), were surrounded by either large (Fig. 3A) or small (Fig. 3B) inducers. This diagonal arrangement of inducers was intended to ensure a diffuse distribution of attention during the entire task. One of the green circles surrounded by red inducers was randomly defined as the target stimulus, while the other was assigned as a nontarget stimulus. Target and nontarget stimuli were presented at either 0.8 or 1.1 degrees of visual angle. The target and nontarget sizes were independent from each other. Trials where the target and nontarget stimuli were the same size were categorized as congruent, while those with different-sized target and nontarget stimuli were classified as incongruent. To prevent unequal distributions, we ensured that target and nontarget sizes had the same size in half of the trials (i.e., congruent trials: both target and nontarget at 0.8 or 1.1 degrees of visual angle), while in the remaining trials, they were different (i.e., incongruent trials: target stimulus at 0.8 degrees of visual angle and nontarget stimulus at 1.1 degrees of visual angle and vice versa). The participant was informed which of the stimuli marked by inducers was task-relevant using a post-cue. The average size of the six distractor circles was always 0.95 degrees of visual angle, and the average size was calculated based on the normal cumulative distribution (M = 0.95, SD = 0.15).

4.3. Procedure

We manipulated the viewing type (simultaneous, delayed) to test the efficiency of OSM, using a paradigm similar to that of Choo and Franconeri (2010). Participants were presented with a size-discrimination task where they were supposed to identify the size of the target. Given that there were only two size categories, we asked them to report

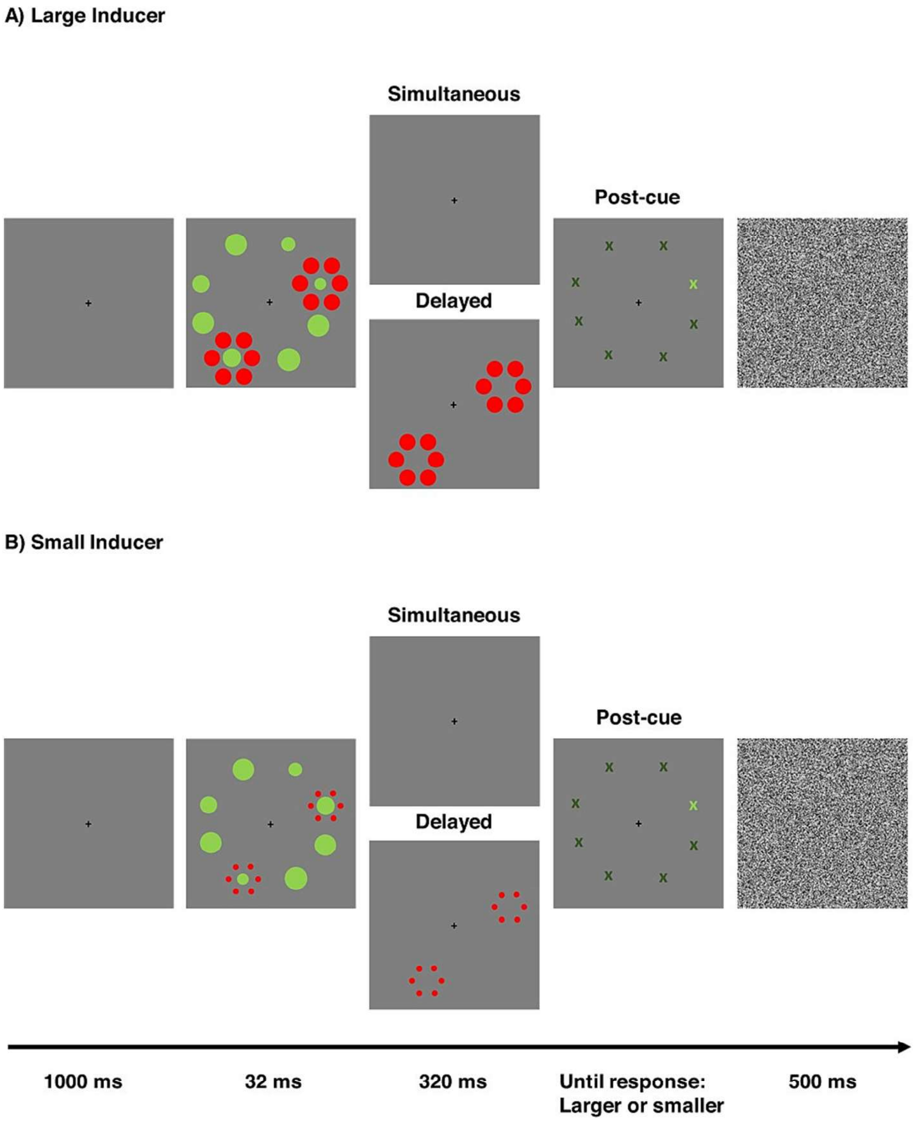

Fig. 3. Illustration of the task in Experiment 2. Participants were asked to indicate whether the target circle, marked by a post-cue, was large or small. Two out of eight circles were surrounded by either large (A) or small (B) inducers. The post-cue indicated which of the two circles was the target. In the *simultaneous* viewing condition, all stimuli disappeared at the same time. In contrast, in the *delayed* viewing condition, the mask remained visible for an additional 320 ms after the test circles had disappeared.

whether the target indicated by a post-cue was a large or a small target. Each trial started with a 1000 ms fixation period (Fig. 3). Following this, the distractors, target, non-target, and inducers were presented for 32 ms. Afterward, either the fixation cross alone (Fig. 3A-B, Simultaneous) or both the fixation cross and inducers (Fig. 3A-B, Delayed) remained on the screen for 320 ms. A post-cue then appeared, staying visible until participants pressed a button to indicate whether the target circle marked by the post-cue was large or small. At the end of each trial, a grey noise pattern was displayed for 500 ms. Each participant completed 640 trials (2 inducer types × 2 viewing conditions × 2 target sizes × 2 nontarget sizes × 40 repetitions). This experiment took around 40 min.

4.4. Statistical analyses

The percentage of correct responses was used as an outcome variable in Experiment 2. We performed a 2 × 2 × 2 × 2 ANOVA to test the effectiveness of OSM in our experimental paradigm. This analysis involved examining the impact of viewing condition (simultaneous, delayed), inducer type (small, large), target size (small, large) and congruency (congruent, incongruent). While the latter was irrelevant to our hypothesis, it was added to explain variance induced by target congruency.

5. Results

Fig. 4 shows the percentages of correct responses with standard errors for within-subject contrast for the simultaneous and delayed viewing conditions under different inducer-type conditions. An ANOVA revealed the main effects of viewing condition [simultaneous vs. delayed] ($F(1, 28) = 117.46, p < .001, \eta_p^2 = 0.808$), inducer type [small vs. large] ($F(1, 28) = 9.76, p = .004, \eta_p^2 = 0.258$) and congruency [congruent vs. incongruent] ($F(1, 28) = 10.86, p = .003, \eta_p^2 = 0.280$). However, the main effect of target size did not reach significance ($p = .313$). Additionally, significant interactions were observed between target size and inducer type ($F(1, 28) = 17.40, p < .001, \eta_p^2 = 0.383$), as well as between inducer type and congruency ($F(1, 28) = 7.18, p = .012, \eta_p^2 = 0.204$).

A robust masking effect was revealed by a sharp decrease in the percentage of correct responses in the delayed viewing condition ($M = 0.60$, $SE = 0.02$) compared to those in the simultaneous viewing condition ($M = 0.75$, $SE = 0.02$). Interestingly, performance demonstrated superiority in the conditions with large inducers ($M = 0.69$, $SE = 0.02$) over the conditions with small inducers ($M = 0.66$, $SE = 0.02$), indicating a stronger masking effect induced by the small inducer. Additionally, the interaction between target size and inducer type revealed that the impact of inducer type was present only when the target size was small. Specifically, small targets in the large inducer conditions (M = 0.70, SE = 0.03) were detected more accurately than in the small inducer conditions (M = 0.62, SE = 0.03). However, such an effect was not observed when the target size was large.

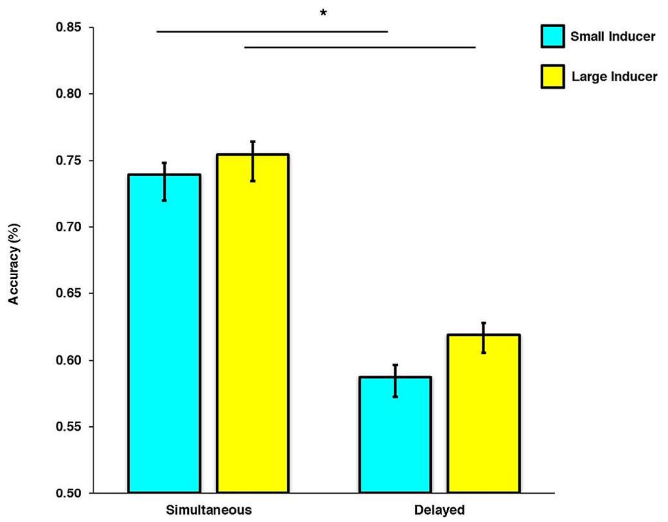

Fig. 4. The percentages of correct responses in Experiment 2. The percentage of correct responses was plotted against inducer type and viewing condition. Blue bars represent the small inducers, and yellow bars represent the large inducers. Asterisks (*) represent significant differences at $p < .05$. Error bars indicate the standard errors around the mean for within-subject contrasts (O'Brien & Cousineau, 2014). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Incongruent trials ($M = 0.70$, $SE = 0.02$), where the target and nontarget had different sizes, generated more accurate responses than congruent trials ($M = 0.65$, $SE = 0.02$). However, inducer type and congruency interaction revealed that the size of the inducer produced a difference only for the congruent trials. Specifically, the percentage of correct responses was greater in the large inducer conditions (M = 0.67, SE = 0.03) than in the small inducer conditions (M = 0.63, SE = 0.03) only if the target and nontarget were the same size. Overall, these findings indicate that OSM efficiently reduced the reportability of the masked circle size.

Experiment 3: Size averaging.

Experiment 3 tested whether the size averaging task can effectively capture the illusory size changes induced by the Ebbinghaus illusion. Employing two distinct target sizes (small and large) that reflected the induced perceived size changes observed in Experiment 1, we explored whether our size averaging task can detect such alterations in target size.

6. Method

6.1. Participants

The same group of twenty-nine participants ($M = 29.38$ years, $SD = 5.52$, 17 females) attended Experiment 3.

6.2. Stimuli

Eight heterogeneously sized green circles were presented around the fixation cross. Unbeknown to the participants, three green circles were assigned as the target stimuli (e.g., Fig. 6). For each participant, these three targets were presented with a size matching the PSE values estimated in Experiment 1. Namely, all three targets were either presented at the stimulus's perceived size surrounded by large inducers (small target size) or presented at the stimulus's perceived size surrounded by small inducers (large target size) in Experiment 1. The average size of the five green distractor circles was always 0.9 degrees of visual angle, and the average size was calculated based on the normal cumulative distribution ($M = 0.90$, $SD = 0.15$). Constant stimuli were used to detect the perceived average size of the green circles. The comparison circle's size varied around the average screen size with 0.1 degrees of visual angle increments, resulting in two different comparison size lists. There were 10 different comparison sizes and each comparison circle was presented 10 times.

6.3. Procedure

Each trial started with a 500 ms fixation (Fig. 5). Then, eight green circles were presented for 32 ms, followed by a 320 ms fixation. In the response window, a comparison circle was displayed until participants pressed a button to indicate whether the comparison circle was larger or smaller than the average size of all eight circles. Each trial ended with a grey noise pattern for 500 ms. Participants performed 200 trials (2 target sizes × 10 comparison sizes × 10 repetitions). This experiment took around 15 min.

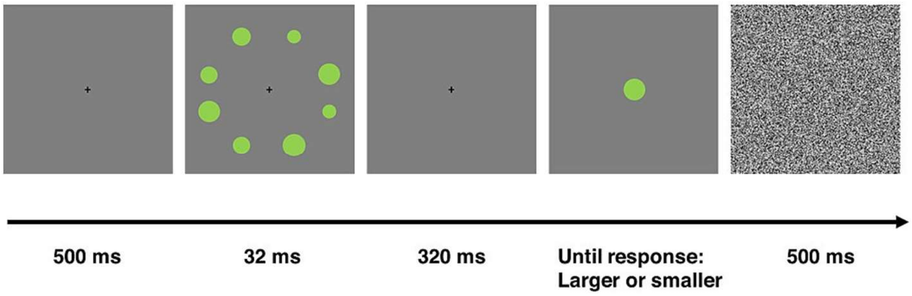

Fig. 5. Illustration of the task in Experiment 3. Participants were asked to indicate whether the comparison circle (time window 4) was larger or smaller than the average size of all eight circles (time window 2).

6.4. Statistical analyses

To assess how participants perceived the average size of the stimuli, we generated psychometric curves for each participant across experimental conditions. The PSE represented the point where the comparison circle was perceived at an equivalent size to the stimuli's average size. Curve widths reflected the level of uncertainty in the participants' responses, with wider curves indicating greater perceptual uncertainty. We calculated goodness of fit measures when fitting psychometric curves to the data. The obtained curves demonstrated a strong fit in Experiment 3 (*r* ranged between 0.833 and 0.995). The perceived average size of all green circles was used as an outcome variable. Given our directional hypothesis, we conducted a one-tailed paired sample *t*-test to assess how changes in target size (small and large) affected the perceived average size.

7. Results

Fig. 6 depicts the mean PSEs with standard errors for within-subject contrast for the small and large average sizes. The results indicated a significant difference in the mean PSE values for small and large targets (t (28) = 1.94, p = .032, Cohen's d = 0.359), while no significant difference was found between the curve widths (t (28) = 1.41, p = .170, Cohen's d = 0.262). Specifically, participants estimated the average size as significantly smaller in the small target condition (M = 0.80, SE = 0.02) than in the large target condition (M = 0.83, SE = 0.03), showing that the size averaging paradigm was sensitive enough to detect size changes induced by the Ebbinghaus illusion (Experiment 1).

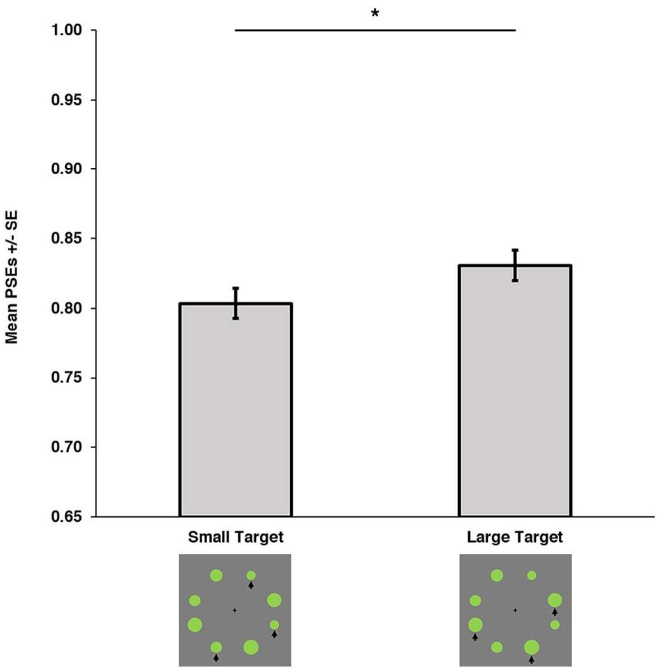

Fig. 6. Perceived average size across different target sizes in Experiment 3. Averaged PSEs were plotted for small and large target size. Asterisks (*) represent significant differences at $p < .05$. Error bars indicate the standard errors around the mean for within-subject contrasts (O'Brien & Cousineau, 2014). Black arrows (↑) illustrate an example of randomly assigned three target positions across experimental conditions. Please note that these arrows were not part of the original display.

Essentially, these findings suggest that the presence of small targets (the mean PSE of the large inducer per participant detected in Experiment 1) could reduce the perceived size of the entire stimulus, while the presence of large targets (the mean PSE of the small inducer per participant detected in Experiment 1) could increase it.

Experiment 4 A: Size rescaling and masking.

The results of Experiments 1–3 indicated that our stimulus configurations were suitable to generate a robust Ebbinghaus illusion and an efficient OSM. Likewise, the size averaging paradigm was sensitive enough to reveal size changes with a magnitude as induced by our variant of the Ebbinghaus illusion. In Experiment 4 A, we combined these three paradigms to reveal the effect of context integration on hidden size averaging.

8. Method

8.1. Participants

One participant from the group of twenty-nine participants couldn't take part in Experiment 4 A-B. So, the remaining twenty-eight participants (M = 29.32 years, SD = 5.62, 17 females) took part in Experiments 4 A-B. We calculated goodness of fit measures when fitting psychometric curves to the data (see below in the Statistical Analysis Subsection). Participants with goodness of fit values smaller than r = 0.63 (corresponding to a $p > .05$) were removed from the sample. In the final sample, there were 23 participants (M = 30 years, SD = 5.83, 13 females).

8.2. Stimuli

Three out of eight circles were randomly chosen as target circles and surrounded by small or large inducers that allowed them to act as Ebbinghaus inducers and masking stimuli inducing OSM (Fig. 7A-B). All three targets were always presented at 0.9 degrees of visual angle. Accordingly, every change in the perceived average size could be attributed to an illusion effect induced by the Ebbinghaus illusion. The physical average size of the stimulus display (excluding inducers) was either 0.9 or 1.1 degrees of visual angle (e.g., Fig. 8, horizontal dashed grey lines). This was achieved by setting the average size of the five distractors to 0.9 or 1.2 degrees of visual angle. The average size of the five distractors was calculated based on the normal cumulative distribution, and the size of each circle was varied with a standard deviation

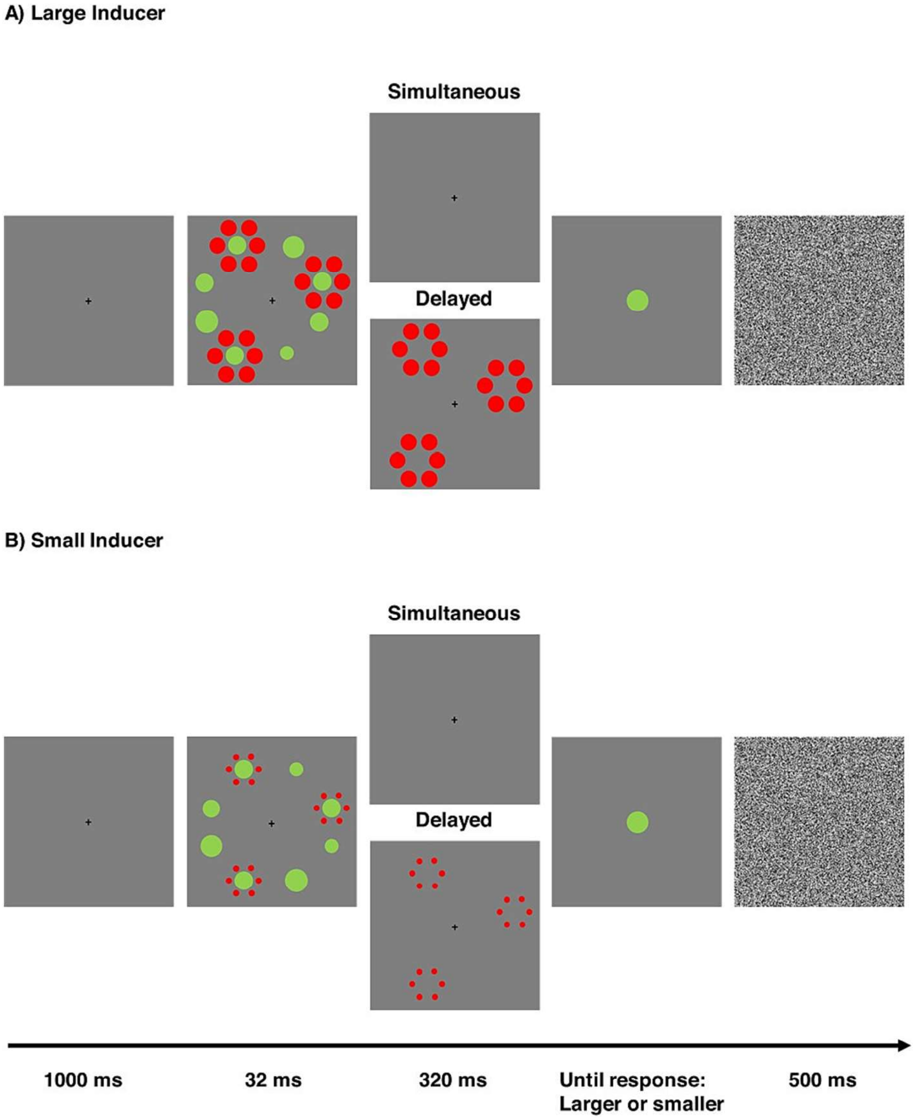

Fig. 7. Illustration of the task in Experiment 4A. Participants were asked to indicate whether the comparison circle (time window 4) was larger or smaller than the average size of all green circles (time window 2). Three out of eight circles were surrounded by either large (A) or small (B) inducers. In the *simultaneous* viewing condition, all stimuli disappeared at the same time. In contrast, in the *delayed* viewing condition, the mask remained visible for an additional 320 ms after the test circles had disappeared. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

of 0.15 degrees of visual angle around the mean (0.9 or 1.2 degrees of visual angle). The target size was kept constant, and only the distractor average size was manipulated. This way, we ensured that participants averaged all green circles on the screen without strategically selecting a subset. Specifically, if participants performed the task correctly, then their reported perceived average size is supposed to be larger in conditions with a large distractor average than those with a small distractor average.

The targets surrounded by the three inducers were never adjacent and always separated by at least one distractor circle. Furthermore, targets were present on every trial in both hemifields, one in the left and two in the right hemifield or vice versa. Constant stimuli were used to detect the perceived average size of the green circles. The size of the comparison circle varied around the physical average size of the entire stimulus display with 0.1 degrees of visual angle increments. Accordingly, different comparison size lists were used for the two average size conditions. In each block, there were 10 different comparison sizes (half of them were smaller than the average size of all green circles, and the other half was larger), and each comparison circle was presented 10 times.

8.3. Procedure

The procedure of Experiment 4 A was identical to Experiment 2, except for the following differences. First, in Experiment 4 A, three targets were surrounded by inducers, compared to just one target in Experiment 2. Second, participants were asked to decide whether the average size of all green circles was smaller or larger than that of the comparison circle. Finally, Experiment 4 A included not only simultaneous, but also delayed viewing condition.

Each trial started with a 1000 ms fixation, followed by a test screen consisting of green circles and inducers for 32 ms (Fig. 7). In the simultaneous viewing condition, all stimuli disappeared at the same time, and only the fixation cross appeared on the screen for 320 msec (Fig. 7A-B, Simultaneous). In the delayed viewing condition, the target circles disappeared, but the inducers persisted on the screen for 320 ms (Fig. 7A-B, Delayed). A comparison circle was presented at the centre of the screen until participants pressed a button to indicate whether the comparison circle was larger or smaller than the average size of all green circles. A grey noise pattern was presented for 500 ms to indicate the end of a trial. Participants completed 800 trials (2 inducer types × 2 viewing

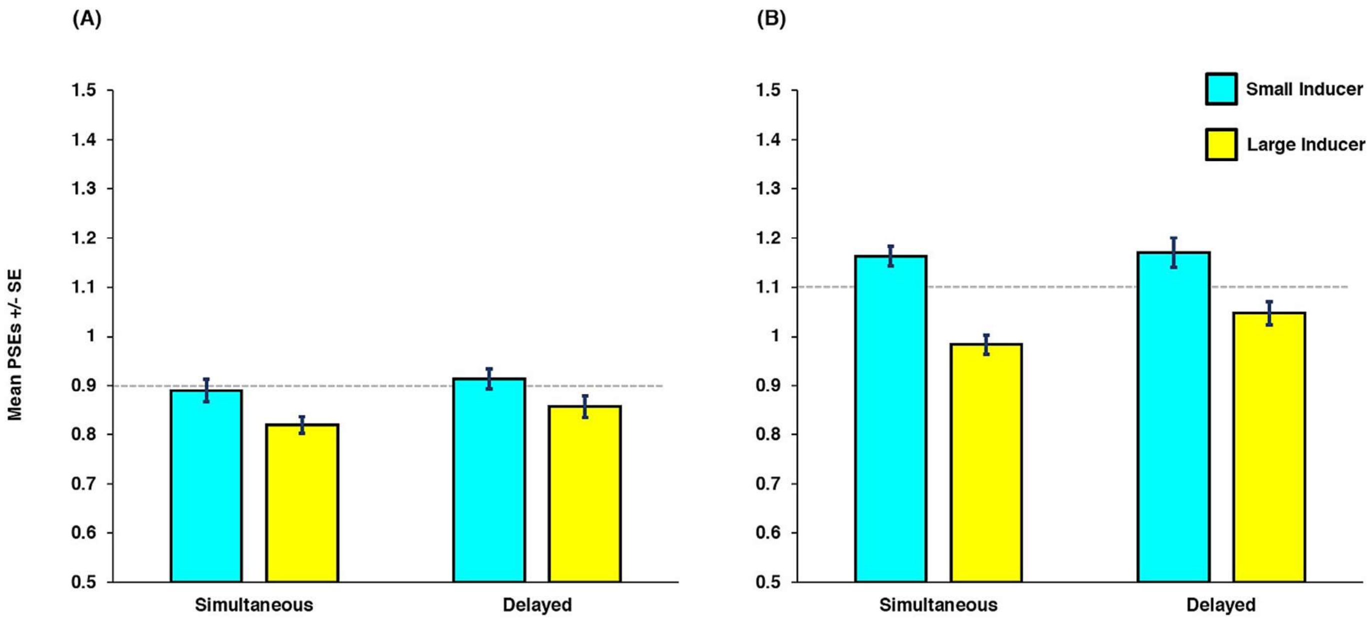

Fig. 8. Perceived average size for small (A) and large (B) distractor averages in Experiment 4A. Averaged PSEs were plotted against the inducer type and viewing condition. Blue bars represent the small inducers, while yellow bars indicate the large inducers. Error bars indicate the standard errors around the mean for within-subject contrasts (O'Brien & Cousineau, 2014). The horizontal dashed grey lines represent the physical average size of the stimulus display. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

conditions × 2 distractor average sizes × 10 comparison sizes × 10 repetitions). This experiment took around 50 min.

9. Eye-tracking data acquisition

We collected eye movement data using Eyelink 1000 (SR Research, Mississauga, Ontario, Canada) to verify that participants maintained their gaze on the fixation cross throughout Experiments 4 A-B. Eye movement data were recorded from the right eye at a sampling rate of 500 Hz. Before the experiment, participants performed a five-point calibration and validation procedure. A three-times bigger circular area (1.2 degrees of visual angle) around the fixation cross was defined as a region of interest (ROI), and we then calculated the time participants spent within this ROI throughout the experiment. Eye movement data were recorded throughout the entire experiment, but only critical periods were further analysed (Fig. 7, time windows 1–2-3). The average coordinates of the fixation cross for each trial were employed as a drift check for that specific trial. Preprocessing of the eye movement data was performed using RStudio Version 4.2.0 (RStudio Team, 2015).

Statistical analyses.

As in Experiment 3, we assessed how participants perceived the average size of the stimuli by fitting psychometric curves to each participant's data for each experimental condition. The PSE represented the point where the comparison circle was perceived at an equivalent size to the stimuli's average size. Curve widths reflected the level of uncertainty in the participants' responses for each experimental condition, with wider curves indicating greater perceptual uncertainty. We calculated goodness of fit measures when fitting psychometric curves to the data. The obtained curves demonstrated a strong fit in Experiment 4 A (r ranged between 0.743 and 0.997). The perceived average size of all green circles was used as an outcome variable. A 2 × 2 × 2 ANOVA was conducted to detect the illusory effects on hidden size averaging, with the factors viewing condition (simultaneous, delayed), inducer type (small, large) and distractor average size (small, large). The estimated curve widths were analysed in a separate 2 × 2 × 2 ANOVA with viewing condition (simultaneous, delayed), inducer type (small, large), and distractor average size (small, large) as within-subjects factors.

10. Results

10.1. Eye movement data

Due to technical problems, we could not use two participants' eye movement data. Following the exclusion of five outliers based on r values (see above in the Participants Subsection), the analysis of eye movement data involved the remaining twenty-one participants. A 2 × 2 x 2 ANOVA was employed to examine the percentage of time that participants maintained fixation on the ROI in each experimental condition, with the factors of inducer type (small, large), viewing condition (simultaneous, delayed) and distractor average size (small, large). During the experiment's critical periods, participants spent an average of 95.45 % of their time within the ROI. ANOVA did not reveal significant main effects of viewing condition [simultaneous vs. delayed] (F (1,20) = 0.77, p = .391, $\eta_p^2 = 0.037$), inducer type [large vs. small] (F (1, 20) = 2.44, p = .134, $\eta_p^2 = 0.109$) and distractor average size [small vs. large] (F (1, 20) = 0.44, p = .513, $\eta_p^2 = 0.022$). Besides, none of the interactions reached significance (all $ps \geq 0.052$), indicating that participants maintained their gaze similarly across experimental conditions.

10.2. Analysis of the perceived average size

Fig. 8 represents the mean PSEs with standard errors for within-subject contrast for the viewing condition and inducer type, in the (A) small and (B) large distractor averages. ANOVA revealed significant main effects of viewing condition [simultaneous vs. delayed] (F (1,22) = 8.50, p = .008, $\eta_p^2 = 0.279$), inducer type [large vs. small] (F (1, 22) = 51.92, $p < .001$, $\eta_p^2 = 0.702$) and distractor average size [small vs. large] (F (1, 22) = 49.70, $p < .001$, $\eta_p^2 = 0.693$). Furthermore, an interaction between inducer type and distractor average was observed (F (1, 22) = 19.02, $p < .001$, $\eta_p^2 = 0.464$).

Most importantly, participants estimated the average size as larger in the small inducer conditions (M = 1.03, SE = 0.03) than in the large inducer conditions (M = 0.93, SE = 0.02), indicating the effect of contextual information. Given that no significant interaction was observed between inducer type and viewing condition (F (1, 22) = 1.50,

$p = .234$, $\eta_p^2 = 0.064$), the data suggest that the effect of contextual information did not differ between masked and non-masked objects (Fig. 8A-B). Moreover, the mean PSEs were overestimated in the delayed viewing condition ($M = 1.00$, $SE = 0.02$) compared to the simultaneous viewing condition ($M = 0.96$, $SE = 0.03$) (Fig. 8A-B).

As expected, we observed a notable difference between the small and large distractor averages. Specifically, the averaged PSEs in the large distractor average condition ($M = 1.09$, $SE = 0.03$) were greater than the ones in the small distractor average condition ($M = 0.87$, $SE = 0.03$) (Fig. 8A-B).

10.3. Curve widths

The estimated curve widths were analysed by conducting a 2 × 2 × 2 ANOVA with the factors viewing condition (simultaneous, delayed), inducer type (small, large) and distractor average size (small, large). ANOVA revealed neither main effects (all $ps \geq 0.188$) nor interactions (all $ps \geq 0.077$).

Experiment 4B: Size rescaling or size contrast?.

The results from Experiment 4 A could be attributed to a context-based modulation of the target items. However, one could argue that the inducers generated a more general effect on elements present in the display other than the target stimuli. The goal of Experiment 4B was to investigate whether the observed effect in Experiment 4 A could be attributed to the size rescaling of masked or non-masked target objects or whether it signifies a more general influence of Ebbinghaus inducers. Specifically, we explored whether the apparent size of one set (e.g., Ebbinghaus inducers) would influence the perceived average size of the other set (e.g., distractor circles), indicating a size-contrast effect which may represent a broader impact of inducers on other elements on the screen. To explore this, we conducted a control experiment identical to Experiment 4 A, with the only distinction being the absence of targets within the inducers. If the results from Experiment 4 A were independent of target-related size-rescaling and instead stemmed from the more general effects of inducers, then we would anticipate identical results in Experiment 4B. Specifically, the obtained results would not differ based on the presence or absence of targets. However, if the size rescaling induced by Ebbinghaus inducers altered the target size representation, we would expect to observe an additional modulation by the Ebbinghaus inducers in Experiment 4 A compared to Experiment 4B.

11. Method

11.1. Participants

The same twenty-eight participants who attended Experiment 4 A took part in Experiment 4B. We calculated goodness of fit measures when fitting psychometric curves to the data (see below in the Statistical Analysis Subsection). Participants with goodness of fit values smaller than $r = 0.63$ (corresponding to a $p > .05$) were removed from the sample. In the final sample, there were 23 participants (M = 30 years, SD = 5.83, 13 females).

11.1.1. Stimuli

All stimuli used in Experiment 4B were the same as in Experiment 4 A. However, the only difference was the absence of target circles within the inducers. Hence, the average size of the overall screen was either 0.9 or 1.2 degrees of visual angle corresponding to the distractor average size since there were no more target circles to influence the overall screen average in the conditions with a large distractor average (e.g., Fig. 9, horizontal dashed grey lines).

11.1.2. Procedure

The procedure was identical to the one in Experiment 4 A.

11.1.3. Eye-tracking data acquisition

We applied the same protocol to collect eye movement data.

11.1.4. Statistical analyses

The obtained curves demonstrated a strong fit in Experiment 4B (r ranged between 0.772 and 0.995). The perceived average size of all green circles was used as the dependent variable. ANOVA with the mean PSE values was performed with the factors viewing condition (simultaneous, delayed), inducer type (small, large) and distractor average size (small, large). After that, we applied a normalisation procedure to facilitate a comparable assessment of the PSE values across different

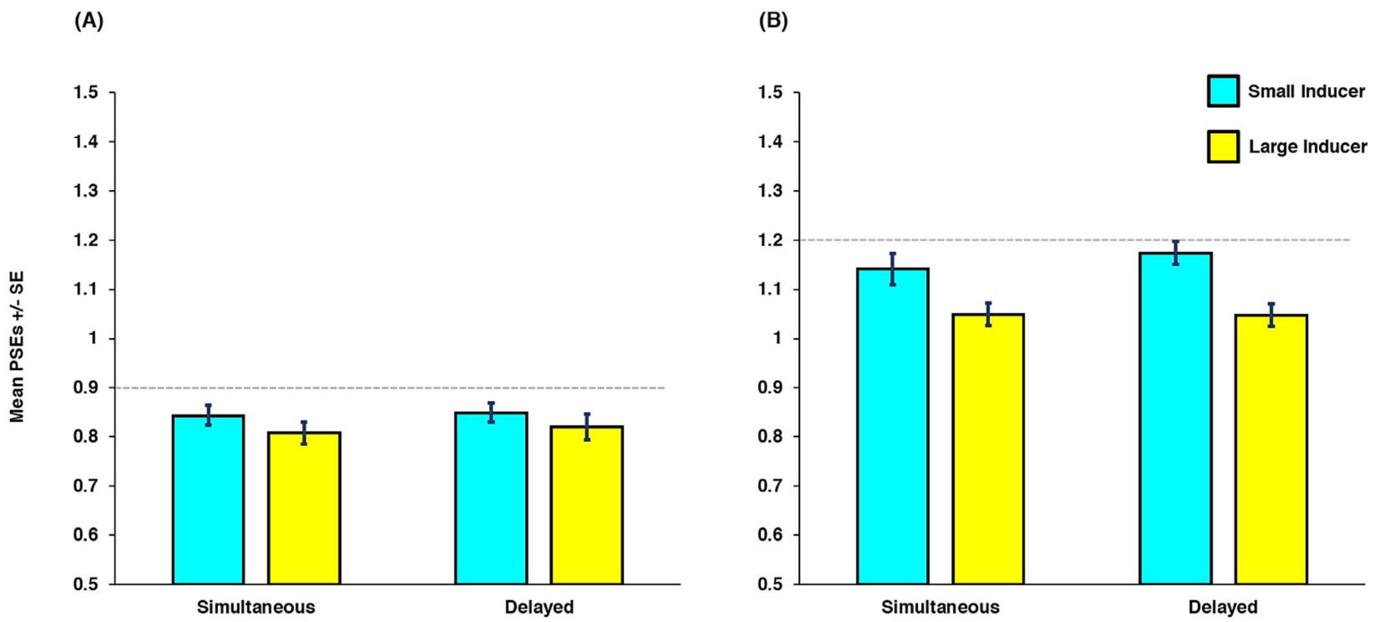

Fig. 9. Perceived average size for small (A) and large (B) distractor averages in Experiment 4B. Averaged PSEs were plotted against the inducer type and viewing condition. Blue bars represent the small inducers, while yellow bars indicate the large inducers. Error bars indicate the standard errors around the mean for within-subject contrasts (O`Brien & Cousineau, 2014). The horizontal dashed grey lines represent the physical average size of the stimulus display. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

distractor averages. This normalisation was necessary due to the differences in the physical average size of the green circles in the large distractor average condition between Experiment 4 A and Experiment 4B. In Experiment 4 A, we employed a constant target size (0.9 degrees of visual angle) and manipulated the distractor average size (0.9 and 1.2 degrees of visual angle), whereas no targets were presented in Experiment 4B. This resulted in different average sizes for the large distractor conditions. Consequently, in Experiment 4 A, the physical average size of the green circles was 1.1 degrees of visual angle, calculated as follows: $((5 \times 1.2) + (3 \times 0.9)) / 8 = 1.1$. In contrast, the physical average size of the green circles was 1.2 degrees of visual angle in Experiment 4B, calculated as follows: $(5 \times 1.2) / 5 = 1.2$. To normalise the raw PSE values, we calculated each PSE value as a percentage relative to the particular physical average size. This procedure accounts for variations in the distractor's average size. For example, for a given PSE value (e.g., 1.1) and a screen average (e.g., 0.9), the normalised value was computed using the formula: (PSE × 100) / Screen Average. Then, we subtracted target-absent PSE percentages (Experiment 4B) from target-present PSE percentages (Experiment 4 A). This subtraction enables us to observe the effect of the target's presence on the PSE values.

12. Results

12.1.1. Eye movement data

This analysis included the same twenty-one participants as those in Experiment 4 A. ANOVA was conducted to assess the percentage of time participants spent on ROI in different conditions, with the factors of inducer type (small, large), viewing condition (simultaneous, delayed) and distractor average size (small, large). On average, participants spent 95.79 % of their time within the ROI. ANOVA did not show significant main effects of viewing condition [simultaneous vs. delayed] (F (1, 20) = 2.13, p = .160, $\eta_p^2 = 0.096$), inducer type [large vs. small] (F (1, 20) = 2.07, p = .165, $\eta_p^2 = 0.094$) and distractor average size [small vs. large] (F (1, 20) = 0.83, p = .373, $\eta_p^2 = 0.040$). Similar to Experiment 4 A, no significant interactions were observed (all $ps \geq 0.329$). Accordingly, fixation was comparable across the various experimental conditions.

Analysis of the perceived average size.

Fig. 9 represents the mean PSEs with standard errors for within-subject contrast for the viewing condition and inducer type in the small (A) and large (B) distractor averages. ANOVA revealed significant main effects of both inducer type [small vs. large] (F (1, 22) = 24.49, $p <$ .001, $\eta_p^2 = 0.527$) and distractor average size [small vs. large] (F (1, 22) = 60.00, $p <$ .001, $\eta_p^2 = 0.732$). However, the main effect of viewing condition did not reach significance. Moreover, an interaction between inducer type and distractor average was detected (F (1, 22) = 21.71, $p <$ .001, $\eta_p^2 = 0.497$).

As in Experiment 4 A, where the targets were present, the average size was estimated as larger in the small inducer conditions (M = 1.00, SE = 0.03) compared to the large inducer conditions (M = 0.93, SE = 0.03) (Fig. 9A-B). Also, the averaged PSEs in the large distractor average condition (M = 1.10, SE = 0.03) were greater than in the small distractor average condition (M = 0.83, SE = 0.03) (Fig. 9A-B). Additionally, the inducer type and distractor average interaction showed that the difference between the small and large inducers was significant for the large distractor average. This pattern is consistent with the notion of a more general effect of the Ebbinghaus inducers.

However, we explicitly compared the results of the two experiments in one analysis to test for additional modulations induced by Ebbinghaus inducers in Experiment 4 A.

Exploring on top effects: The inducer effect on the target.

Fig. 10 represents the normalised PSEs with standard errors for within-subject contrast for the viewing condition and inducer type. Experiment 4B was conducted to estimate the general effects of the Ebbinghaus stimuli on the results of Experiment 4 A. While both experiments involved the effects of inducers on distractor stimuli, only Experiment 4 A involved additional effects of inducers on the masked and non-masked target items. Subtracting the target-absent PSE values (Experiment 4B) from the target-present PSE values (Experiment 4 A), we could reveal target-related modulations. Therefore, we directly compared Experiments 4 A and 4B. After normalising the raw PSE values by calculating the percentage of each PSE value considering the physical average size of the screen, we subtracted the target-absent PSE values (Experiment 4B) from the target-present PSE values (Experiment 4 A). The resulting values were included in the analysis. A 2 × 2 × 2 ANOVA was conducted with the factors viewing condition (simultaneous, delayed), inducer type (small, large) and distractor average size (small, large). ANOVA revealed significant main effects of inducer type [small vs. large] (F (1, 22) = 6.52, p = .018, $\eta_p^2 = 0.229$) and distractor average size [small vs. large] (F (1, 22) = 4.52, p = .045, $\eta_p^2 = 0.170$). However, neither the main effect of viewing condition (p = .191) nor the interactions (all $ps \geq 0.082$) reached statistical significance.

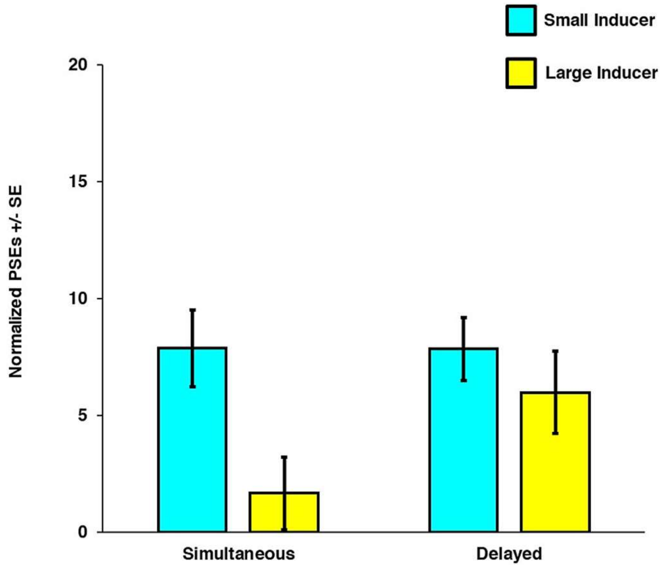

Fig. 10. Difference between Experiments 4A and 4B. Normalised PSEs were plotted against inducer type and viewing condition. Blue bars represent the small inducers, while yellow bars indicate the large inducers. Error bars indicate the standard errors around the mean for within-subject contrasts (O'Brien & Cousineau, 2014). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Most importantly, participants estimated the average size as larger in the small inducer conditions (M = 7.86, SE = 1.10) relative to that in the large inducer conditions (M = 3.82, SE = 1.28), indicating evidence of target-based size modulations (Fig. 10). Since, the interaction between the factors inducer type and viewing condition was not significant, we did not have evidence that this effect was different for masked and non-masked targets (F (1, 22) = 2.01, p = .170, $\eta_p^2 = 0.084$). Furthermore, as expected, the average size was estimated larger in the large distractor average conditions (M = 7.26, SE = 1.27) than in the small distractor average conditions (M = 4.42, SE = 0.94).

13. Curve widths

The estimated curve widths were analysed by conducting a 2 × 2 x 2 ANOVA with the factors viewing condition (simultaneous, delayed), inducer type (small, large) and distractor average size (small, large). As in Experiment 4 A, ANOVA revealed neither main effects (all $ps \geq 0.184$) nor interactions (all $ps \geq 0.533$).

14. Discussion

Through behavioural experiments, we investigated how contextual information influences the perceived size of masked and non-masked objects. A combined paradigm allowed the induction of an illusion via the Ebbinghaus inducers. The Ebbinghaus inducers also efficiently functioned as a mask, inducing object substitution masking. Efficient masking decreased target recognisability and hence rendered target item encoding more implicit. Finally, size averaging allowed us to indirectly assess the size representations of the target stimuli so that it was possible to infer the effect of size rescaling of both masked and non-masked visual stimuli in the same way. The results of Experiment 4 A showed that large Ebbinghaus inducers decreased the reported average size, whereas small inducers led to an increase. This effect was independent of the target items being masked.

Moreover, in a control experiment (Experiment 4B), we tested for a more general inducer effect on the display configuration by removing all target items while keeping all other display features identical to Experiment 4 A. Experiment 4B indeed showed a general effect of inducers on distractor items. We explicitly tested for specific target effects by directly comparing the results of Experiments 4 A and 4B. This analysis showed that over and above the general effects found in Experiment 4B, there was a target-related modulation of size by Ebbinghaus inducers, which was not significantly different for masked and non-masked objects.

The interpretation of Experiments 4 A's and B's findings critically depends on the effectiveness of each experimental manipulation combined in our paradigm. Therefore, in addition to the main experiment, we verified the effectiveness of each paradigm by conducting Experiments 1–3.

Experiment 1 demonstrated the efficiency of our experimental setup in altering the perceived size of objects. The results showed a substantial perceived size difference between small and large inducers compared to the control condition without inducers, indicating that our variant of the Ebbinghaus illusion generated a robust size illusion effect in Experiment 1. However, the no-inducer condition differs from the two illusion conditions in terms of the number of stimuli, which could explain why no significant increase in perceived average size was detected in the small inducer condition compared to the no-inducer condition. A variant with inducer stimuli of the same size as the target might have provided a more controlled baseline for comparison. Please note, that the no-inducer condition was not used in the subsequent experiments.

Furthermore, it was crucial to ensure that our masking procedure efficiently altered target recognisability to infer the effects of size-rescaling of masked vs. non-masked stimuli. We used OSM following the experimental setup previously introduced by Choo and Franconeri (2010). The results from Experiment 2 revealed that our stimulus configuration reliably induced OSM, as evident from a robust performance decline in the delayed viewing condition. In principle, the simultaneous presentation of targets and inducers may have decreased target recognisability, hence generating a masking effect itself. However, results from Experiment 2 show that accuracy was significantly reduced in the delayed viewing condition as compared to the simultaneous viewing condition hence indicating that our experimental set-up successfully altered target recognisability. We used the stimulus parameters proven efficient in Experiment 2 to ensure robust masking in Experiment 4 A.

Lastly, we tested whether the size averaging paradigm was sensitive enough to detect perceived size changes in the magnitude of the ones induced by the Ebbinghaus illusion in Experiment 3. We employed the individual PSE values from Experiment 1 as small and large target sizes to achieve this. Specifically, we adjusted the physical sizes of the target items to align with the perceived sizes altered by the Ebbinghaus illusion in Experiment 1. We employed three target circles in a single display to enhance the impact of target size. Experiment 3 confirmed that both small and large targets significantly influenced the perceived display's average size, leading to a decrease and an increase in the perceived average size, respectively.

After successfully confirming the efficiency of each paradigm in Experiments 1–3, the objective of Experiment 4 A was to integrate these three paradigms to reveal the effect of context integration on hidden size averaging. The findings of Experiment 4 A demonstrated a modulation in the perceived average size, regardless of the object being masked or not. No notable distinctions were evident between the simultaneous (non-masked) and delayed (masked) viewing conditions. These findings replicate those reported by Choo and Franconeri (2010), emphasising that even when OSM highly impairs an object's visibility, its contribution to the average size remains intact and is comparable to that of an unmasked object. The results of Choo and Franconeri (2010) have been criticised in a study by Jacoby et al. (2013). The main criticism raised in this study was directed towards the mask positions, particularly the possibility of adjacent masks, which could potentially introduce an attentional effect. Besides, Choo and Franconeri (2010) found that the average influence of the two target circles was weaker in masked conditions compared to non-masked conditions, showing a bias towards smaller responses. Following the critique by Jacoby et al. (2013), this result could be interpreted as evidence that OSM disrupts size averaging. We addressed these concerns by implementing three masks to ensure a diffuse distribution of attention mandatory for OSM (Enns & Di Lollo, 1997). Further, we included two distinct distractor averages in Experiment 4 A, which allowed testing participants' task performance. Contrary to Jacoby et al. (2013), which demonstrated that OSM disrupts single object representation and impairs averaging performance, the findings of Experiment 4 A strongly support the notion that even masked objects contribute to size averaging. The results of Experiment 4 A align with a prior study showing that ensemble summary statistics vary depending on contextual information, and the perceived average range is formed after size-rescaling mechanisms (Markov & Tiurina, 2021). Notably, our results take these findings one step further, showing that even the size of masked objects is first rescaled and then incorporated into the averaged size. Interestingly, even though size information was implicitly coded and not explicitly reportable in the delayed viewing conditions, a bottom-up signal still revealed its influence on perceptual decisions. This is consistent with a previous study showing that neural activity patterns in occluded areas of early visual cortex provide key insights into the categorization and specific details of nearby images (Morgan, Petro, & Muckli, 2019). Despite V1 and V2 receiving only feedback from higher cortical areas and lacking a direct feedforward signal, information can still be extracted from these occluded areas. This suggests that category information can be effectively anticipated by the internal models in the early visual cortex.

Previously, Im and Chong (2009) combined size averaging and the Ebbinghaus illusion to test whether the reported average size reflects a stimulus's perceived or physical size. In contrast to our study, their stimuli were clearly visible and unmasked. Small and large Ebbinghaus inducers induced, similar to our study, contextual information. Following a reasoning similar to ours, the authors argued that if the perceived size of target objects contributes to the reported average size, there would be either an overestimation or an underestimation of the average size. Instead, if the physical size contributes to the reported average size, no difference would emerge between small and large inducer blocks. Their results showed that the perceived size of the objects contributed to the reported average size. Thus, their data align with our findings, indicating that each target size is rescaled through contextual information and then contributes to the average size, consistent with Tiurina and Utochkin (2019). These findings collectively demonstrate that contextual information influences the objects before they contribute to ensemble summary statistics.

In light of the current findings, it is reasonable to argue that size averaging comes into play just after the size representations are rescaled through contextual information. This hypothesis is supported by the finding that conscious perception of the surrounding inducers or target item is not essential for the Ebbinghaus illusion, indicating a subconscious contextual modulation (Chen et al., 2018). In Chen et al. (2018), the target items and the surrounding inducers were rendered invisible by continuous flash suppression and backward masking. Significant illusory size judgments were observed in both masking procedures, demonstrating that the illusory effect induced by the Ebbinghaus illusion persisted even when the target or surrounding inducers were invisible. In contrast, similar findings could not be detected for the Ponzo illusion, which relies on feedback mechanisms from higher-level visual areas (Schwarzkopf et al., 2011). Chen et al.'s findings indicate that subconscious contextual modulations already occur in early visual processing, consistent with the results from a functional imaging study indicating that neural activity in V1 represents the perceived size of the objects (Sperandio, Chouinard, & Goodale, 2012).

Even though the Ebbinghaus illusion is widely recognised for its association with activations in early visual processing regions, particularly V1, recent studies suggest that this illusion may involve brain regions beyond V1 (Chen et al., 2024; Wu et al., 2023). Wu et al. (2023) reported an important role of the parietal cortex for the Ebbinghaus illusion. Namely, the connection from the right V1 to the superior parietal lobule (SPL) was found to predict the strength of the Ebbinghaus illusion. A stronger Ebbinghaus illusion after inhibitory rTMS to the right SPL might be explained by altered inhibitory feedforward connectivities from V1 to SPL. This effect aligns with the observed negative correlation between the strength of this connection and the magnitude of the Ebbinghaus illusion. Furthermore, Chen et al. (2024) demonstrated that the Ebbinghaus illusion is influenced by feedback projections from higher to lower visual areas, highlighting the role of top-down signals in modulating this illusion. The influence of feedback mechanisms on size rescaling was further demonstrated by Zeng, Fink, and Weidner (2020). Their study, which applied TMS stimulation, revealed that there is a top-down influence on context integration, indicating that feedback mechanism from LOC modulates the altered perceived size representations in the early visual cortex.

Despite object recognisability being highly impaired by OSM in Experiment 4 A, the perceived average size was larger in the delayed viewing than the simultaneous viewing condition. A similar tendency to overestimate the average size in the delayed viewing condition was observed in previous studies (Choo & Franconeri, 2010; Jacoby et al., 2013). Nevertheless, this effect was not evident when target items were removed in Experiment 4B. Two perspectives may account for this overestimation in the delayed condition compared to the simultaneous condition in Experiment 4 A. Initially, visual information processing differs between the simultaneous and delayed viewing conditions in Experiment 4 A. In the simultaneous condition, where targets are presented within the inducers, the perceptual process benefits from faster processing due to the immediate availability of visual signals. However, in the delayed condition, the absence of visual signals due to OSM necessitates a more prolonged accumulation of evidence, resulting in extended processing time. This prolonged processing time aligns with previous studies demonstrating that stimuli presented or processed longer tend to be perceived as larger (Rammsayer & Verner, 2014; Thomas & Cantor, 1976).

Secondly, neuronal activity initially progresses from lower-level visual areas to higher-order visual areas in visual processing, referred to as feedforward processing (Boehler, Schoenfeld, Heinze, & Hopf, 2008; Krasich et al., 2022). Once this initial feedforward activity reaches a specific visual area, the activation is sent back to lower-level visual areas through feedback connections (recurrent processing). These feedback connections are believed to convey predictions about incoming information, while any error in these predictions is transferred via feedforward connections, aiding to refine subsequent processing. OSM is thought to interfere with recurrent processing while preserving the initial feedforward signal (Boehler et al., 2008; Di Lollo et al., 2000; Enns & Di Lollo, 1997; Harris, Ku, & Woldorff, 2013). Specifically, the longer presentation of mask in the delayed viewing condition results in additional feedforward processing of the mask-alone representation. This additional feedforward processing of the prolonged mask-alone presentation disrupts the recurrent processing of the stimulus-mask representation. Consequently, context-induced perceptual biases may have a stronger impact and, may enhance the impact of the Ebbinghaus illusion in the delayed viewing condition. Integrating these perspectives, we suggest that the observed overestimation in the delayed viewing condition in Experiment 4 A arises from the prolonged processing time required for accumulating evidence and the interruption of recurrent processes. These explanations could also clarify why we did not observe overestimation in Experiment 4B, where the absence of targets removed potential processing differences between the simultaneous and delayed viewing conditions.

Interestingly, we observed a modulation of the perceived average size in the control experiment (Experiment 4B). This finding cannot be attributed to rescaling the target size, as demonstrated in Experiment 4 A, nor to the influence of inducers contributing to average size. If Ebbinghaus inducers were incorporated into the items to be averaged, an opposing effect would be anticipated. Specifically, the presence of large inducers would lead to an increase in the reported average, while small inducers would result in a decrease in the reported average. However, we observed the exact opposite pattern. The data suggest that the inducers directly act on the perceived average size in more general terms. The fact that large inducers decreased while small inducers increased perceived average size suggests a size contrast effect, which may arise on multiple levels of processing. It is possible that the size of the items to be judged is already biased at the initial perceptual level, where all stimuli are presented simultaneously. This interpretation is supported by psychophysical (Nakashima & Sugita, 2018; Song et al., 2011) and neuroimaging (Schwarzkopf et al., 2011; Schwarzkopf & Rees, 2013) studies suggesting that the Ebbinghaus illusion occurs on early levels within the visual system. Alternatively, the decision might be influenced at a subsequent stage, where the contrast effect becomes more pronounced, leading participants to judge the distractors relative to the inducers in Experiment 4B. This latter view is consistent with early accounts of the Ebbinghaus illusion such as the one suggested by Massaro and Anderson (1971). Based on this account, the bias observed in Experiment 4B might reflect a size-contrast effect occurring at a post-perceptual stage. In particular, the inducers may have served as standards or references when the distractor circles are judged (Massaro and Anderson, 1971).

The question then arises, whether the effects observed in Experiment 4 A can be fully explained by the same mechanism, namely by a contrast effect of the inducers on the visible items on the screen. We addressed this question by comparing the data from Experiments 4 A and 4B. We subtracted the size average judgments in the control Experiment 4B with its respective judgments in Experiment 4 A to eliminate any size contrast effect by the inducers on non-target stimuli. The rationale behind this was that if the changes in size-averaging were solely due to a general size-contrast effect, then the two experiments should show no differences. On the other hand, any observed difference could be attributed to a rescaling effect of the target. What we found were on-top effects in Experiment 4 A over Experiment 4B, indicating target-related effects. In particular, we found a significant impact of inducer size, with larger reported size averaging values for small inducers and smaller size averaging values for large inducers, indicating target-size-rescaling. Importantly, we did not find any difference between viewing conditions. Hence, target-size-rescaling was not found to be different between explicitly and implicitly coded objects.

Two key considerations could be raised when interpreting our findings. Initially, one might question whether our results have been influenced by post-perceptual decision biases (Firestone & Scholl, 2016). While such concerns are typical in psychophysics, we implemented several precautions to minimise the impact of these biases. Participants were instructed to report their responses without using mental strategies and were unaware of the experimental manipulations (inducer types, viewing conditions, and distractor sizes), making it unlikely that differences between conditions were due to general decision biases. The randomisation of block order further reduced potential biases from prior knowledge, and participants' lack of familiarity with psychophysical methods (e.g., the Method of Constant Stimuli) decreased the likelihood of systematic bias. Finally, we used two distractor averages to evaluate task performance. The target size was consistently set at 0.9 degrees of visual angle, while the distractor average sizes were manipulated to be either small (0.9) or large (1.2) degrees of visual angle. This manipulation altered the overall average size of objects on the screen. Our findings demonstrated a clear upward shift in perceived average size when the distractor average size increased, which would not be expected to observe when the effects were purely due to decisional biases.

Subsequently, as our paradigm involves several circles, it could potentially induce crowding effects. However, the stimuli that are located close enough to generate crowding effects are separated by colour. It is known, that perceptual grouping is one of the key factors for crowding. Stimuli from different sets of colours are unlikely to generate crowding as demonstrated by Kennedy and Whitaker (2010). In their study, they manipulated chromaticity channels (red-green, blue-yellow) and demonstrated that crowding occurs only when the target and flankers shared the same colour. Furthermore, most crowding paradigms typically involve longer stimulus durations (≥150 ms) that allow a proper allocation of attention (Freeman & Pelli, 2007; Greenwood & Parsons, 2020; Kennedy & Whitaker, 2010; Kooi, Toet, Tripathy, & Levi, 1994; Levi, Klein, & Hariharan, 2002; Li, Joo, Yeatman, & Reinecke, 2020), whereas our stimuli were presented for a notably brief duration of 32 ms, making it unlikely for a crowding mechanism to occur. Finally, Parkes, Lund, Angelucci, Solomon, and Morgan (2001) demonstrated that crowding did not prevent participants from obtaining summary statistics, such as the average orientation of a cluster of items.

15. Conclusion

In summary, we identified two critical aspects related to size-rescaling in the context of size-averaging. First, we found that the Ebbinghaus inducers play a pivotal role in altering the perceived average size of distractor stimuli. When large inducers were present, the perceived average size decreased, while smaller ones had the opposite effect. The results indicate that the average size of different groups of objects are not independent, but mutually alter their perceived size in a contrast-like fashion. These findings align with earlier theories regarding the Ebbinghaus illusion (Massaro & Anderson, 1971), suggesting that the context circles within the illusion serve as standards for size judgment.

Second, our results revealed that the inducers consistently influenced their size representation even when stimuli were masked. This observation indicates that the size-rescaling effects operate at the early stages of visual processing, preceding explicit stimulus encoding.

CRediT authorship contribution statement

Elif Memis: Visualization, Validation, Software, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization, Writing – original draft. **Gizem Y. Yildiz:** Visualization, Validation, Supervision, Software, Project administration, Methodology, Formal analysis, Conceptualization, Writing – review & editing. **Gereon R. Fink:** Supervision, Resources, Methodology, Funding acquisition, Conceptualization, Writing – review & editing. **Ralph Weidner:** Validation, Supervision, Project administration, Methodology, Conceptualization, Writing – review & editing.

Declaration of competing interest

The authors declare no competing interests.

Data availability

The data is made available via the internal Jülich platform Jülich DATA (https://data.fz-juelich.de/) and via Re3data (https://www.re3data.org/).

Acknowledgement

Open access publication funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 491111487. GRF gratefully acknowledges support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) (Project-ID 431549029—SFB 1451). The funders have/had no role in the decision to publish or preparation of the manuscript. We are grateful to our colleagues from the Institute of Neuroscience and Medicine for many valuable discussions.

Appendix A. Supplementary Data

The links to the raw behavioural data and psychometric curves for each experiment can be found in Supplementary Material A. Supplementary data to this article can be found online at [https://doi.org/10.1016/j.cognition.2024.106041].

References

Alvarez, G. A., & Oliva, A. (2008). The representation of simple ensemble visual features outside the focus of attention. *Psychological Science, 19*(4), 392–398.

Ariely, D. (2001). Seeing sets: Representation by statistical properties. *Psychological Science, 12*(2), 157–162.

Boehler, C. N., Schoenfeld, M. A., Heinze, H. J., & Hopf, J. M. (2008). Rapid recurrent processing gates awareness in primary visual cortex. *Proceedings of the National Academy of Sciences, 105*(25), 8742–8747.

Chen, L., Qiao, C., Wang, Y., & Jiang, Y. (2018). Subconscious processing reveals dissociable contextual modulations of visual size perception. *Cognition, 180*, 259–267.

Chen, L., Wu, B., Yu, H., & Sperandio, I. (2024). Network dynamics underlying alterations in apparent object size. *Brain Communications, 6*(1). fcae006.

Chong, S. C., & Treisman, A. (2005). Statistical processing: Computing the average size in perceptual groups. *Vision Research, 45*(7), 891–900.

Choo, H., & Franconeri, S. L. (2010). Objects with reduced visibility still contribute to size averaging. *Attention, Perception, & Psychophysics, 72*(1), 86–99.

Di Lollo, V., Enns, J. T., & Rensink, R. A. (2000). Competition for consciousness among visual events: The psychophysics of reentrant visual processes. *Journal of Experimental Psychology: General, 129*(4), 481.

Ebbinghaus, H. (1902). *Grundzuge der Psychologie* (Vol. 1., Band 2. Theil.). Leipzig: Viet & Co..

Enns, J. T., & Di Lollo, V. (1997). Object substitution: A new form of masking in unattended visual locations. *Psychological Science, 8*, 135–139.

Erdfelder, E., Faul, F., & Buchner, A. (1996). GPOWER: A general power analysis program. *Behavior Research Methods, Instruments, & Computers, 28*(1), 1–11.

Firestone, C., & Scholl, B. J. (2016). Cognition does not affect perception: Evaluating the evidence for "top-down" effects. *Behavioral and brain sciences, 39*, e229.

Freeman, J., & Pelli, D. G. (2007). An escape from crowding. *Journal of Vision, 7*(2), 22.

Greenwood, J. A., & Parsons, M. J. (2020). Dissociable effects of visual crowding on the perception of color and motion. *Proceedings of the National Academy of Sciences, 117*(14), 8196–8202.

Harris, J. A., Ku, S., & Woldorff, M. G. (2013). Neural processing stages during object-substitution masking and their relationship to perceptual awareness. *Neuropsychologia, 51*(10), 1907–1917.

Im, H. Y., & Chong, S. C. (2009). Computation of mean size is based on perceived size. *Attention, Perception, & Psychophysics, 71*(2), 375–384.

Jacoby, O., Kamke, M. R., & Mattingley, J. B. (2013). Is the whole really more than the sum of its parts? Estimates of average size and orientation are susceptible to object substitution masking. *Journal of Experimental Psychology: Human Perception and Performance, 39*(1), 233.

Kennedy, G. J., & Whitaker, D. (2010). The chromatic selectivity of visual crowding. *Journal of Vision, 10*(6), 15.

Kooi, F. L., Toet, A., Tripathy, S. P., & Levi, D. M. (1994). The effect of similarity and duration on spatial interaction in peripheral vision. *Spatial Vision, 8*(2), 255–280.

Krasich, K., Simmons, C., O'Neill, K., Giattino, C. M., De Brigard, F., Sinnott-Armstrong, W., & Woldorff, M. G. (2022). Prestimulus oscillatory brain activity interacts with evoked recurrent processing to facilitate conscious visual perception. *Scientific Reports, 12*(1), 22126.

Levi, D. M., Klein, S. A., & Hariharan, S. (2002). Suppressive and facilitatory spatial interactions in foveal vision: Foveal crowding is simple contrast masking. *Journal of Vision, 2*(2), 2.

Li, Q., Joo, S. J., Yeatman, J. D., & Reinecke, K. (2020). Controlling for participants' viewing distance in large-scale, psychophysical online experiments using a virtual chinrest. *Scientific Reports, 10*(1), 904.

Markov, Y. A., & Tiurina, N. A. (2021). Size-distance rescaling in the ensemble representation of range: Study with binocular and monocular cues. *Acta Psychologica, 213*, Article 103238.

Massaro, D. W., & Anderson, N. H. (1971). Judgmental model of the Ebbinghaus illusion. *Journal of Experimental Psychology, 89*(1), 147.

Morgan, A. T., Petro, L. S., & Muckli, L. (2019). Scene representations conveyed by cortical feedback to early visual cortex can be described by line drawings. *Journal of Neuroscience, 39*(47), 9410–9423.

Nakashima, Y., & Sugita, Y. (2018). Size-contrast illusion induced by unconscious context. *Journal of Vision, 18*(3), 16.

O'Brien, F., & Cousineau, D. (2014). Representing error bars in within-subject designs in typical software packages. *The Quantitative Methods for Psychology, 10*(1), 56–67.

Parkes, L., Lund, J., Angelucci, A., Solomon, J. A., & Morgan, M. (2001). Compulsory averaging of crowded orientation signals in human vision. *Nature Neuroscience, 4*(7), 739–744.

Rammsayer, T. H., & Verner, M. (2014). The effect of nontemporal stimulus size on perceived duration as assessed by the method of reproduction. *Journal of Vision, 14*(5), 17.

RStudio Team. (2015). *RStudio: Integrated development for R*. Boston, MA: RStudio, Inc.

Schwarzkopf, D. S., & Rees, G. (2013). Subjective size perception depends on central visual cortical magnification in human V1. *PLoS One, 8*(3), Article e60550.

Schwarzkopf, D. S., Song, C., & Rees, G. (2011). The surface area of human V1 predicts the subjective experience of object size. *Nature Neuroscience, 14*(1), 28–30.

Song, C., Schwarzkopf, D. S., & Rees, G. (2011). Interocular induction of illusory size perception. *BMC Neuroscience, 12*, 1–9.

Sperandio, I., Chouinard, P. A., & Goodale, M. A. (2012). Retinotopic activity in V1 reflects the perceived and not the retinal size of an afterimage. *Nature Neuroscience, 15*(4), 540–542.

Thomas, E. A., & Cantor, N. E. (1976). Simultaneous time and size perception. *Perception & Psychophysics, 19*(4), 353–360.

Tiurina, N. A., & Utochkin, I. S. (2019). Ensemble perception in depth: Correct size-distance rescaling of multiple objects before averaging. *Journal of Experimental Psychology: General, 148*(4), 728.

Watamaniuk, S. N., & Duchon, A. (1992). The human visual system averages speed information. *Vision Research, 32*(5), 931–941.

Wu, B., Feng, B., Han, X., Chen, L., & Luo, W. (2023). Intrinsic excitability of human right parietal cortex shapes the experienced visual size illusions. *Cerebral Cortex, 33*(10), 6345–6353.

Zeng, H., Fink, G. R., & Weidner, R. (2020). Visual size processing in early visual cortex follows lateral occipital cortex involvement. *Journal of Neuroscience, 40*(22), 4410–4417.