Statistical Learning Facilitates Access to Awareness

Luzi Xu*, Chris L.E. Paffen, Stefan Van der Stigchel, and Surya Gayet

Experimental Psychology, Helmholtz Institute, Utrecht University,
Heidelberglaan 1, 3584 CS, Utrecht, The Netherland

Word count: introduction (616) + discussion (1005) + figure captions (320)

Abstract

Statistical learning (SL) allows us to quickly extract regularities from sensory inputs. Although many studies have established that SL serves a wide range of cognitive functions, it remains unknown whether SL impacts conscious access. We addressed this question, seeking converging evidence from multiple paradigms across four experiments (total N = 153): Two reaction-time based b-CFS experiments showed that objects at probable locations and with probable features are released from suppression faster than improbable objects. In a visual masking experiment, we observed higher sensitivity to probable (versus improbable) objects, independent of conscious access to the stimulus dimension carrying the regularities. Finally, a pre-registered accuracy-based b-CFS experiment showed higher localization accuracy for interocularly suppressed probable (versus improbable) objects given identical presentation durations, thereby excluding processing differences emerging *after* conscious access (e.g., criterion shifts). Together, these findings demonstrate that SL prioritizes conscious access of probable over improbable visual input.

Keywords: statistical learning, consciousness, visual awareness, breaking continuous suppression, interocular suppression, pre-registration

Statement of Relevance

Our visual environment provides us with a continuous stream of complex information, most of which is highly structured. Cars, for example, have prototypical locations (e.g., on the road rather than in the sky), and share prototypical visual characteristics (e.g., a horizontally elongated shape). Human beings are extremely proficient at extracting such structural regularities to facilitate a wide range of cognitive functions. The present study demonstrates for the first time that implicit learning processes like statistical learning can influence the selection of sensory input for conscious perception, prioritizing probable events over improbable events in conscious access. This finding indicates that our conscious perception can be shaped by regularities that we are not explicitly aware of. The direct effect of statistical learning on conscious access may explain how statistical learning serves a wide range of cognitive functions that benefit from or depend on consciousness, such as memory, learning, and decision making.

Introduction

The human brain can quickly and implicitly extract regularities of sensory inputs across time and space from the environment through statistical learning (SL). Over the past two decades, SL has become a major area in cognitive research, as indicated by its pervasive influence on a wide range of basic and higher-level cognitive processes (for a review, see Bogaerts et al., 2020; Frost et al., 2019; Sherman et al., 2020). To name a few, SL enables language acquisition (e.g., Saffran et al., 1996), perception of high-level perceptual units (e.g., scenes and events, Brady & Oliva, 2008; Turk-Browne, 2012), recognition and association of meaningful chunks for learning, memory (e.g., Brady et al., 2009), and social inference (e.g., Dotsch et al., 2017). Despite extensive research on the function of SL, no study has investigated whether SL can facilitate the entry of sensory inputs into consciousness. This question is particularly important, given that consciousness of sensory inputs is considered a prerequisite for many cognitive functions (e.g., Baars & Franklin, 2003; Meuwese et al., 2013; Rich & Mattingley, 2002; Sabary et al., 2020), including cognitive functions that are -themselves- regulated by SL (e.g., memory, language, and inference). To understand how SL functions, it is therefore essential to establish whether SL impacts conscious access of sensory inputs.

The hypothesis that SL affects conscious access is supported by the overlap between the behavioral correlates of SL and conscious access. Previous studies suggested that SL generates memory chunks (Orbán et al., 2008), induces implicit anticipations (Turk-Browne et al., 2010), and regulates the allocation of attention

(Geng & Behrmann, 2005; Wang & Theeuwes, 2018) and working memory resources (Brady et al., 2009; Umemoto et al., 2010). Given that these factors all influence the detection of interocularly suppressed stimuli (working memory, Gayet et al., 2013, 2016; anticipation, Denison et al., 2011, 2016; Pinto et al., 2015; attention, Thibault et al., 2016), one might hypothesize that SL directly affects conscious access. Alternatively, it is conceivable that SL does not alter conscious access, since similar implicit learning processes (e.g., perceptual learning) do not affect detection of stimuli under interocular suppression (Heyman & Moors, 2014; Paffen et al., 2018).

To investigate whether SL affects conscious access, we first used a reaction-time based breaking continuous flash suppression paradigm (b-CFS, Gayet et al., 2014; Stein et al., 2011) in Experiments 1 and 2 (Fig. 1). Here, the time it takes for observers to report some aspect of interocularly suppressed targets is related to the competitive strength of stimuli for entering visual awareness (e.g. higher contrast targets will break suppression faster than low contrast ones). In our experiments, we hypothesized that targets occurring with high probability would break suppression faster than low probability ones. To exclude the possibility that differential reaction times in b-CFS can reflect differences emerging *after* conscious access (e.g., decisional and post-perceptual processes, Stein et al., 2019; Stein & Peelen, 2021), we conducted two follow-up experiments. In Experiment 3, we used the detection-discrimination dissociation paradigm (hereafter “DDD”; Stein & Peelen, 2021) with which we tested whether localization sensitivity is higher for high versus low probability feature stimuli when participants cannot consciously perceive the feature dimension that carries the regularities (in our case the orientation of a triangle). This dissociation is similar to blindsight (Weiskrantz et al., 1974) in which patients localize stimuli more accurately than chance, even though they do not consciously perceive those target stimuli (e.g., chance-level in presence-absence judgements). In Experiment 4, we used an accuracy-based b-CFS paradigm (Litwin et al., 2023) in which the localization accuracy for interocularly suppressed high and low probability feature stimuli were compared. Together, the four experiments provide converging evidence that SL prioritizes conscious access of high versus low probable visual input.

Methods (Experiment 1)

In Experiment 1, we examined whether the presence of statistical regularities of target *locations* influences conscious access to targets. Specifically, if SL of spatial locations affects conscious access, we would observe faster reaction times (in an orthogonal discrimination task) for targets that appear at high- versus low-probability locations.

Participants. The predetermined sample size for this first experiment was loosely based on previous SL studies (Wang & Theeuwes, 2018). Twenty-five healthy participants, naive to the purpose of the experiment, participated in Experiment 1 after signing an informed consent form. One observer who did not follow instructions (i.e., the participant did not look through the binocular stereoscope) was excluded from all data analyses and was replaced. The eventual sample contained twenty-four participants (19 women and 5 men, mean age = 24.13, SD = 2.49) for data analysis.

All participants had normal or corrected-to-normal vision and had no color blindness. They received monetary compensation for their participation.

Apparatus and stimuli. The experiment was conducted on a 27-inch LCD monitor (2,560 × 1,440 pixels, 60-Hz refresh rate). The experiment took place in a darkened laboratory with all light sources turned off, except for the computer screen, which was positioned at an effective viewing distance (the distance the light travels from monitor to eye) of 57 cm. The presentation area on the screen comprised two parts (half-images presented on the left and right half of the monitor) which were viewed dichoptically through a stereoscope mounted on a chin rest. The stereoscope made it possible to independently stimulate the left and right eyes of participants, thus triggering interocular competition. To promote binocular fusion of the two competing images, each display (that contains a competing image) had a gray (16.2 cd/m², x = 0.283, y = 0.298) background (9.5°×7.2°) surrounded by a Brownian noise square frame with a thickness of 0.5°. The remaining part of the screen was set as a uniform black background luminance of 0.05 cd/m².

The target stimulus consisted of an upright or inverted triangle (1.4° in height and length), which was presented to a single eye. This target was distanced 2.4° from a central fixation dot (a 0.3° white dot with a black edge). The other eye was presented with high-contrast colored masks that were changed at a rate of 10 Hz. These masks were made up of randomly arranged circles (diameter 0.3 to 1.5°) of different colors (they differed both in hue and luminance). One hundred and twenty different CFS masks were generated before the experiment and appeared in a random order (without replacement) across different trials. The dynamic high-contrast CFS masks were presented to one eye to perceptually suppress the static stimulus shown to the other eye, thereby rendering the static stimulus initially unconscious (Tsuchiya & Koch, 2005). The experiment was programmed using Psychtoolbox (Brainard, 1997) in MATLAB (R2021a; The Mathworks, Natick, MA).

Procedure. Each trial started with a fixation dot that appeared at the center of the screen for 500 ms. After this, a dynamic CFS mask was presented to one eye of the participant. Meanwhile, the target was presented to the other eye of the participant with its intensity (i.e., opacity) ramping up from zero to full opacity within 1,000 ms (Fig. 1A). At the start of the trial, participants were unable to consciously perceive the triangle because it was interocularly suppressed by the CFS masks. Over time, the visibility of the dynamic CFS mask presented to one eye was gradually reduced with its transparency increasing gradually. The increasing intensity of the target and the decreasing intensity of the mask jointly caused the target to be eventually released from interocular suppression, thus allowing participants to report upon the target. The presentation of the CFS masks and the target ended when a participant responded or after ten seconds had passed. Participants were required to press one of two buttons to indicate the orientation of the triangle as quickly and accurately as possible (‘↑’ for upright triangles, ‘↓’ for inverted triangles). We manipulated how often a triangle was presented at different locations (left vs. right) but required participants to report triangle orientation (upright vs. inverted), see Fig. 1B. As such, the response mapping was orthogonal to the experimental manipulation, minimizing the influence of response bias.

Fig. 1. (A) Trial overview of the b-CFS paradigm used in Experiments 1-2. (B) Targets could appear on the left or right of fixation, and could consist of upright or inverted triangles.

Targets (triangles) appeared on a specific side, either the left or right side, of the screen on 75% of all trials (‘the high-probability condition’), and appeared on the other side of the screen on the remaining 25% of all trials (‘the low-probability condition’). The location (left vs. right sides) being set as the high-probability location was counterbalanced between participants. In contrast, the orientation of triangles was counterbalanced (with random presentation orders) within participants across different trials. The distance from the target to the fixation dot was the same in all trials; that is, the target was positioned on the outline of an imaginary circle with a radius of 2.4°. To preserve a clear distinction between the left and right of fixation, targets could only appear in the left and right quadrants of the imaginary circle, on one of twenty equally interspersed locations. The exact position within a quadrant was drawn at random (thus not counterbalanced), so that targets had an equal probability of appearing either (slightly) above or below the horizontal midline.

After the response, participants kept viewing the target display with both eyes for 1,000 ms. Meanwhile, participants also received sound feedback (500 ms): a high-pitch (2,000 Hz), or a low-pitch (1,500 Hz) ‘beep’ sound for correct or incorrect responses. This phase was not only a feedback phase but also served as an important learning phase: given that it is unclear whether statistical regularities can be extracted from interocularly suppressed stimuli, we showed the targets without suppression at the end of each trial to ensure that participants could learn the statistical regularities from stimuli that were not interocularly suppressed.

The formal experiment comprised 8 practice trials and 6 blocks of 32 formal trials. Here, we manipulated the prevalence of the target locations, so that it was more likely to appear to the left of fixation (75% of trials) than to the right of fixation for half of the participants, and vice versa for the other half of the participants. We refer to these as ‘regularity blocks’. The color of the target was green (7.70 cd/m^2 , x = 0.288, y = 0.444 at full opacity) in the regularity blocks. To enhance the motivation of participants, participants received feedback about their average performance at the end of each block, showing their mean reaction times (RTs) and the mean accuracy of that block. After checking the performance, participants had the opportunity to take a self-initiated rest. We set up a mandatory rest every three blocks for all participants.

In addition, before the formal experiments, we also asked participants to finish 3 blocks of 32 trials where the statistical regularity was absent (‘non-regularity blocks’).

We used a red (6.12 cd/m^2 , x = 0.655, y = 0.332 at full opacity) color for the targets in the non-regularity blocks. This salient change of target color between the non-regularity blocks (of the pre-experiment) and regularity blocks (of the formal experiment) was aimed at minimizing the transfer of (the absence of) regularities from the non-regularity blocks to the regularity blocks. These non-regularity blocks were aimed at measuring within-participant differences in eye dominance and spatial biases. This pre-experiment data is not included in the formal data analysis and can be found in Supplementary Materials 2.

At the end of the formal experiment, we measured participants' subjective awareness of statistical regularities. They were asked to fill out a questionnaire after they completed the experiment. In this questionnaire, they were asked to guess the percentages of targets appearing on the left or right of fixation (the key manipulation dimension), the percentages of targets appearing at the upper or lower side of fixation (actual probability: 50% each), and the percentages of targets appearing in an upright or inverted orientation (actual probability: 50% each).

Results (Experiment 1)

The b-CFS task. Incorrect responses were excluded from all data analyses (2.41% of all trials). The accuracy of participants ranged from 89.58% to 100%, with an average accuracy of 97.59% (SD = 2.54). To test whether SL develops over time, we further compared the RTs to targets at high and low probability locations across time. To this aim, we divided the experiment into epochs of 64 trials (the conditions (high versus low probability) were fully counterbalanced within each of these individual epochs). We conducted a repeated measures ANOVA with the factors Probability (high versus low) and Epoch (1 to 3) to investigate how the influence of statistical regularities on response times would evolve over time.

Results show that the main effect of Probability was significant, ($F(1, 23) = 5.98$, $p = .023$, $\eta_p^2 = 0.21$). Specifically, RTs for the high-probability locations (3.66 s, SD = 1.81) were 0.47 seconds shorter than for the low-probability locations (4.13 s, SD = 1.20), see Fig. 2A. The shorter RTs for targets presented at high-probability locations indicates that visual input gains faster access to consciousness when appearing at a probable (rather than improbable) location. Besides, the main effect of Epoch was also significant ($F(2, 23) = 8.97$, $p = .001$, $\eta_p^2 = 0.28$), reflecting a general decrease in response times over the course of the experiment. Importantly, the absence of an interaction between Probability and Epoch, ($F(2, 23) = 0.40$, $p = .673$, $\eta_p^2 = 0.02$) suggests that the difference in RTs to targets appearing on high compared to low probability locations did not change over the course of the experiment. To test whether the SL effect has appeared at early stage of the experiment, we further conducted t-tests to compare RTs to targets on high and low probability locations in the first epoch. Results showed that RTs for the high-probability locations were already shorter than for the low-probability locations in Epoch 1 ($t(23) = 3.20$, $p = .004$, Cohen's $d = 0.27$, 95% CI = [0.20, 0.92]). Taken together, these data show that participants rapidly extracted the statistical regularities of the target location and that the effect of statistical regularities on conscious access did not change over time.

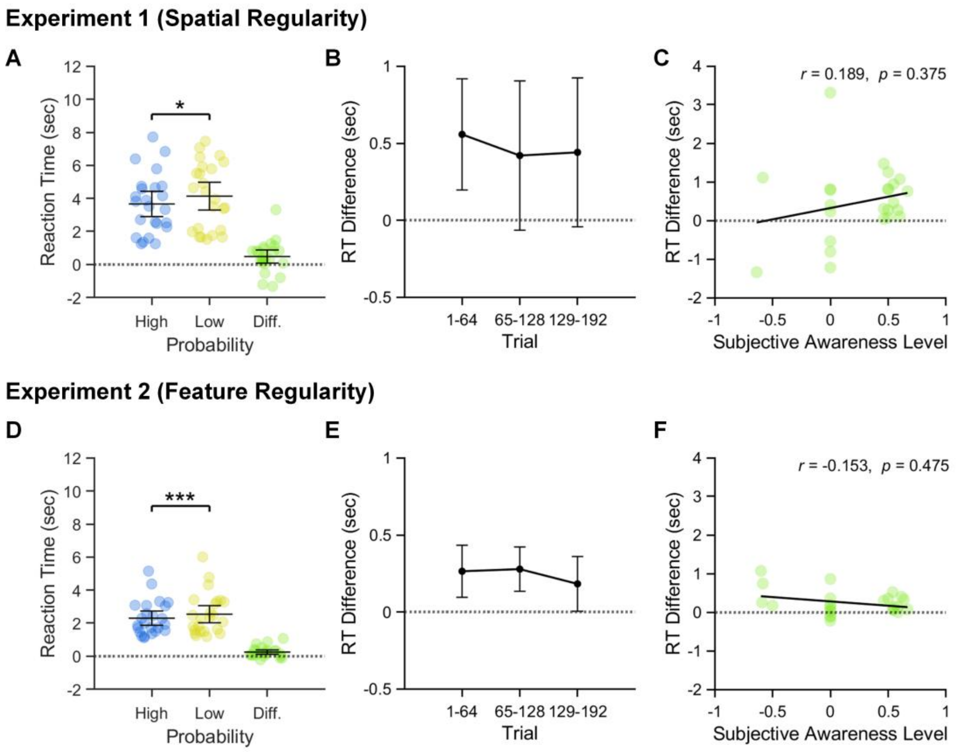

Fig. 2. The results of Experiments 1-2. (A) and (D) display mean response times (RTs) in the high-probability and low-probability conditions, with individual dots representing individual participants; (B) and (E) display RT differences between the high- and low-probability conditions, over time (on the x-axis). (C) and (F) show the correlation between participants' RT difference and their subjective awareness of the statistical regularities. Error bars represent the 95% confidence interval of the mean. Asterisks indicate significance (**p* < .05, ****p* < .001).

The subjective awareness ratings. We computed a subjective awareness rating, reflecting the extent to which participants were aware of the high versus low probability manipulation (see Supplementary Materials 1). A key aspect of this awareness metric is that we used individual participants' guesses to the other, non-manipulated, dimensions to scale the reported imbalance estimates to the manipulated dimension, thus accounting for individual differences in tendencies to report more (or less) extreme imbalance estimates. Basically, we divided the difference of the estimated percentage of the high- versus low-probability conditions (i.e., left/right in Experiment 1, up/down in Experiment 2) by the mean of the unsigned imbalance reported for the two non-manipulated dimensions. The sign of calculated subjective awareness ratings indicates whether the probability estimate is in line with the true probability of the stimulus occurrence (positive) or not (negative), while larger (absolute) values indicate a larger estimated imbalance in stimulus occurrence. Subjective awareness ratings were scaled, so that values above 1.0 indicate that participants provided more extreme probability estimations for the manipulated regularity dimension (i.e., left/right in Experiment 1, up/down in Experiment 2) compared with the two non-manipulated dimensions (e.g., up/down and upright/inverted in Experiment 1). Therefore, numbers equal to/below 1.0 can be interpreted as an indication that participants were unaware of the manipulated dimension (for details, see Supplementary Materials 1).

Results showed that although participants performed above chance in reporting which was the high-probability location (as indicated by above zero awareness rating, $t(23) = 3.43$, $p = .002$, Cohen's $d = 0.70$, 95% CI = [0.10, 0.41]), the reported imbalance between high and low probability conditions was not higher for the manipulated stimulus dimension (75%–25%) than for the two non-manipulated stimulus dimensions (50%–50%), as indicated by below 1.0 awareness rating (mean:

0.26, SD = 0.37; $t(23) = 9.85$, $p < .001$, Cohen's $d = 2.01$, 95% CI = [0.59, 0.90]). That is, although participants tended to estimate the high-probability location as more probable (55.83, SD = 10.60) than the low-probability location (44.17, SD = 10.60), this bias fell within the range of biases reported in the absence of any probability manipulation. This can be interpreted as an indication that participants were unaware of the statistical regularities.

A rank-based Spearman's correlation test ($p < .05$ in the Shapiro-Wilk Normality Test) showed that there was no significant correlation between the calculated subjective awareness score and the difference in RTs between high and low probability locations in the b-CFS task, $r = 0.19$, $p = 0.38$, Fig. 2C. These results indicate that higher levels of awareness about the regularities during the experiment were not accompanied by faster responses to high-probability targets (relative to low-probability targets) in the b-CFS task.

Inter-trial priming effects. In this experiment, high-probability (location) trials occurred more often than low-probability trials due to the manipulation of regularity. This led to a higher incidence of consecutive high-probability trials and consequently to a lower incidence of consecutive low-probability trials. Therefore, it is possible that the targets of the high probability condition were prioritized because they occurred more often in consecutive trials (i.e., the inter-trial priming effect) instead of being a probable target. To test this possibility, we divided the trials into 'repeat trials' in which the target was at the same location as in the previous trial, and 'change trials' in which the target was at a different location. A two-tailed paired-sample t-test was conducted to examine the effect of Inter-trial Continuity (change vs. repeat) on SL effect (i.e., the difference between RTs of high probability and low probability trials). Results showed that there was no difference between SL effect of change trials and repeat trials, $t(23) = 0.15$, $p = .879$, Cohen's $d = 0.03$, 95% CI = [–0.39, 0.45]), showing that the faster responses to high-probability targets were not caused by an inter-trial priming effect. Therefore, the differential RTs between high- and low-probability locations is unlikely to have been caused by inter-trial priming.

Methods (Experiment 2)

Experiment 2 was aimed at extending the findings of Experiment 1, by testing whether the presence of statistical regularities of target *features* (instead of target *locations*) also accelerates conscious access of targets comprising high-probability features. If SL affects unconscious processing at the level of features, we should observe faster reaction times to targets comprising high-probability (vs. low-probability) features.

The methods were generally identical to those of Experiment 1 except for the following changes. First, a new group of twenty-four healthy participants (20 women and 4 men, mean age = 25.08, SD = 3.17) were recruited for the experiment. The sample size was set to match that of Experiment 1. Second, in the regularity blocks, we manipulated the probability of targets being upright or inverted, instead of manipulating the probability of target locations. Participants were asked to determine target locations (left or right) as quickly as possible (instead of reporting the target orientation, see Fig. 1).

Results (Experiment 2)

The b-CFS task. Incorrect responses were excluded from all data analyses (1.54% of all trials). The accuracy of participants ranged from 94.79% to 100%, with an average accuracy of 98.46% (SD = 1.43).

As in Experiment 1, we also conducted a repeated measures ANOVA with the factors Probability (high versus low) and Epoch (1 to 3) in Experiment 2. Results show that the main effect of Probability was, again, significant ($F(1, 23) = 14.79$, $p < .001$, $\eta_p^2 = 0.39$). Specifically, RTs for the high-probability features (2.30 s, SD = 1.03) were 0.24 seconds shorter than for the low-probability features (2.54 s, SD = 1.23), see Fig. 2D. The shorter RTs for targets presented with high-probability features indicate that visual input gains faster access to consciousness when appearing with a probable (rather than improbable) feature. Apart from this, the main effect of Epoch was also significant ($F(2, 23) = 9.10$, $p = .001$, $\eta_p^2 = 0.45$), which reflects a general decrease in response times over the course of the experiment.

The interaction between Probability and Epoch was not significant, $F(2, 23) = 0.40$, $p = .673$, $\eta_p^2 = 0.02$, suggesting that the difference in RTs to targets comprising high compared to low probability features did not change over the course of the experiment. As in Experiment 1, we further conducted t-tests to compare RTs to targets on high and low probability features in the first epoch, and showed that RTs for the high-probability features were already shorter than for the low-probability features in Epoch 1 ($t(23) = 3.24$, $p = .004$, Cohen's $d = 0.66$, 95% CI = [0.10, 0.43]). These data show that the statistical regularities of the target feature were extracted rapidly and affected unconscious processing accordingly.

The subjective awareness ratings. Results showed that participants performed at chance in reporting which was the high-probability location, as indicated by the insignificant difference between awareness ratings and zero, $t(23) = 1.88$, $p = .072$, Cohen's $d = 0.39$, 95% CI = [–0.02, 0.35]). Moreover, participants' subjective awareness scores (0.16, SD = 0.43) were significantly lower than 1.0 at the group level ($t(23) = 9.58$, $p < .001$, Cohen's $d = 1.96$, 95% CI = [0.66, 1.02]), indicating that the tendency to correctly estimate the high-probability feature as more probable (56.46, SD = 12.64) than the low-probability feature (43.96, SD = 13.19) fell within the range of biases reported in the absence of any probability manipulation. That is, most participants were unaware of the statistical regularities. The rank-based Spearman's correlation test ($p < .05$ in the Shapiro-Wilk Normality Test) showed that there was no significant correlation between the calculated values of subjective awareness and the differential RTs in the b-CFS task, $r = –0.15$, $p = .48$ (Fig. 2F). This shows that the effects of SL on b-CFS localization times did not depend on participants' explicit knowledge about these statistical regularities.

Inter-trial priming effects. To exclude the possible account of inter-trial priming effects, a two-tailed paired-sample t-test was conducted to examine the effect of Inter-trial Continuity (change vs. repeat) on SL effect (i.e., the difference between RTs of high probability and low probability trials). Results showed that there was no difference between SL effect (i.e., the difference between RTs of high probability and low probability trials) of change trials and repeat trials, $t(23) = 0.66$, $p = .519$, Cohen's $d = 0.13$, 95% CI = [–0.20, 0.38]), showing that the faster responses to high-probability targets were not caused by an inter-trial priming effect. Therefore, the differential RTs between high- and low-probability features is unlikely to have been caused by inter-trial priming.

Methods (Experiment 3)

In Experiments 1-2, we found that initially suppressed stimuli were responded to faster when they were positioned at a high-probability location or consisted of a high-probability feature. Even though this result seems, at first glance, to indicate that conscious access is sped up for high-probable stimulus information, we cannot know for certain that these effects originate prior to conscious access of the stimulus. Specifically, it is possible that the differential RTs that we found in Experiments 1 and 2 were due to conscious perceptual processes (e.g., high-probability stimuli are processed more rapidly *after* they break suppression, thus generating faster responses) or decisional factors (i.e., participants might instill a more liberal response tendency for high-probability stimuli, yielding faster responses at the cost of reduced accuracy; see Gayet et al., 2014; Stein & Peelen, 2021).

To test whether statistical regularities can exert an influence on the processing of stimuli that are not (yet) consciously perceived, we used a detection-discrimination dissociation paradigm (hereafter: DDD; Stein & Peelen, 2021). This recently developed paradigm offers the possibility to observe [A] a difference in detection/localization sensitivity between two stimulus conditions (e.g., higher sensitivity for localizing upright compared to inverted faces), while [B] participants are unable to discriminate between these two stimulus conditions (e.g., chance level performance in distinguishing between upright and inverted faces). The latter null effect [B] demonstrates that participants have no conscious access to the stimulus dimension that governs the performance difference [A], and thus that effect [A] could not have been caused by conscious processes (see Schmidt & Vorberg, 2006).

In our implementation of the DDD paradigm, we asked participants to perform a non-speeded two-alternative-forced-choice (2-AFC) localization (left vs. right) task as well as a non-speeded 2-AFC discrimination (upright vs. inverted) task on every trial. We manipulated the statistical regularity of the 'to-be-discriminated' dimension (upright vs. inverted triangles). The signal detection theory index *d'* was used to indicate the localization and discrimination sensitivities. We aimed at establishing a condition in which participants were unconscious of whether the target was upright or inverted (the dimension that carried statistical regularities), and subsequently test the effects of these regularities using a localization task (without regularities). If SL indeed affects the localization sensitivity of targets unconsciously, participants would have different localization sensitivities for high- vs. low-probability targets (e.g., upright versus inverted triangles), even when being unconscious of the identity of the target feature (i.e., when they are at chance at discriminating between upright and inverted triangles). Adversely, if SL cannot influence the detectability of objects at an unconscious level, we would observe no difference between high- and low-probability feature conditions whenever participants are unable to discriminate between upright and inverted triangles.

As stated above, we first needed to establish a condition where participants performed at chance level on the orientation discrimination task ($d' = 0$: participants have no conscious access to the manipulated feature dimension), and above chance performance on the localization task across both high and low probability conditions ($d' > 0$; participants had access to the target location). Having met these requirements, we then compared the localization sensitivity between targets with high vs. low-probability features.

Participants. A new group of twenty-four participants (20 women and 4 men, mean age = 24.25, SD = 2.25) were recruited for an experiment, based on the sample size of Experiments 1 and 2. In a subsequent, preregistered replication experiment (http://256.so/i1n), another group of twenty-five participants (15 women and 10 men, mean age = 25.52, SD = 2.79) were recruited based on the effect size of the exploratory experiment. Because of a complication in the data analysis proposed in the preregistration (i.e., the lack of consideration of the possible influence of unbalanced trial numbers on the comparison of signal detection parameter *d'*; for details, see Supplementary Materials 3), we combined the data from both groups together for an exploratory analysis outlined below.

Apparatus and stimuli. The experiment was conducted in the same lab environment as Experiments 1-2, but on another 27-inch LCD monitor (2,560 × 1,440 pixels, 120-Hz refresh rate), positioned at an effective viewing distance of 57 cm. Here, participants directly viewed the monitor without the stereoscope used in the previous experiments. All stimuli were presented on a presentation area consisting of a gray background (9.5° × 7.2°; 16.2 cd/m², x = 0.283, y = 0.298) surrounded by a Brownian (1/f2) noise frame with a thickness of 0.5°, and a fixation dot in the center. The remaining part of the screen was set as a uniform black background.

The target stimulus consisted of a green triangle (pointing upwards or downwards; 1.4° × 1.4°) on top of a green circle (diameter of 2.6°), and appeared either to the left or right side of fixation at a fixed eccentricity (2.4° from fixation). The exact hue and luminance of the triangle (mean across participants: 7.74 cd/m², x = 0.304, y = 0.509) and the circle (mean across participants: 12.4 cd/m², x = 0.304, y = 0.521) were determined for each participants individually, using an adaptive staircases procedure (see below) during a separate experimental session preceding the main experiment session. We used the same masks as in Experiments 1 and 2.

Procedure. The regularity manipulation was identical to that of Experiment 2 in which we also manipulated feature regularity. For any given participant, the triangle pointed towards one direction (up or down) in 75% of all trials and pointed towards the opposite direction (down or up) in the remaining 25% of all trials. Which one was selected as the high-probability feature was counterbalanced across participants. Within both the low and high-probability trials, each combination of presentation time (8.3, 16.7, and 33.3 + 8.3 ms) and target location (left or right of fixation) occurred equally often. The longer presentation times (16.7 and 33.3 + 8.3 ms) were included for the purpose of keeping participants motivated; we feared that participants would become unmotivated when only engaging in 'unaware' trials. We were primarily concerned with the effects at the shortest presentation time (8.3 ms) in which participants were most likely to be unaware of the target features. These conditions were intermixed, and the trial order was randomized for each participant.

At the beginning of each trial, a black fixation dot appeared at the center of the presentation area for 1000 ms (see Fig. 3A). Next, a central white fixation dot (with a black edge) appeared for 500 ms to indicate the start of the stimulus presentation. A mask appeared for 100 ms before the target presentation, serving as a forward mask. Next, the target (a triangle) appeared on the screen for a short duration (8.3ms, 16.7ms, or 33.3 ms followed by an 8.3 ms blank). This was followed by three consecutive backward masks that were presented for 100 ms each. After the masks, participants were asked to perform two tasks. In the 2-AFC localization task, participants pressed the left or right arrow to indicate the location of the target. In the 2-AFC discrimination task, participants pressed up or down arrows to indicate the orientation of the target (a triangle pointing up or down). Notably, these two tasks were non-speeded, and participants were instructed to respond as accurately as possible without any time pressure. The order of the two questions was fixed within participants but counterbalanced across all participants.

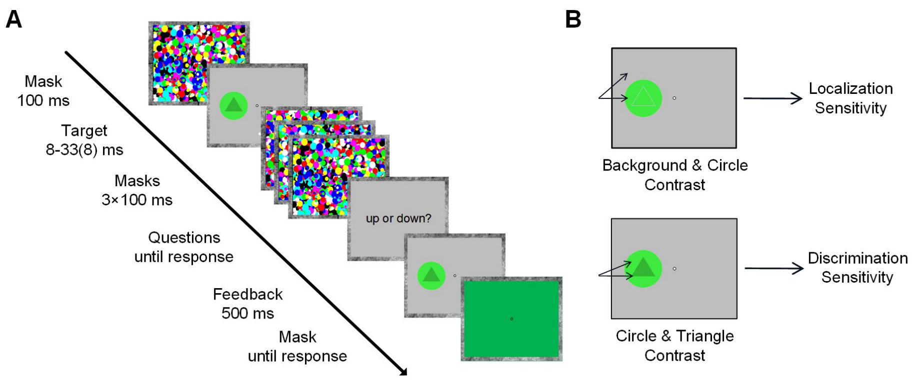

Fig. 3. (A) Trial outline of the detection-discrimination dissociation paradigm used in Experiments 3. (B) Schematic representation of the stimulus properties that were varied using an adaptive staircase approach, to obtain the desired performances in the localization and discrimination tasks.

Before the main experiment, we conducted a pre-experiment that was aimed at determining RGB values for the circle and triangle constituents of the target stimulus (Fig. 3B) leading to desired performance level on the localization and discrimination tasks, for each individual participant. To this end, we used Accelerated Stochastic Approximation (Kesten, 1958), a non-parametric adaptive procedure that rapidly converges to any accuracy level. Participants performed the same tasks (localization and discrimination tasks) as in the formal experiment. The localization task threshold was titrated by varying the intensity (i.e., opacity) of the circle relative to the background, and the discrimination task threshold was titrated by varying the intensity (i.e., opacity) of the triangle relative to the circle. The background color was fixed for all participants. Participants first performed in a staircase-procedure adjusting the contrast between the circle and background (more specifically, changing the RGB of the circle) to obtain above-chance localization performance (aiming to converge at 75% correct). After this, and in order to find chance-level discrimination performance, they performed in another staircase in which we used the RGB of the circle color obtained in this first staircase, and adjusted the contrast between the circle and triangle (more specifically, changing the RGB of the triangle).

We aimed to find chance level (50% accuracy) feature discrimination performance in the 8.3 ms condition. It is important to consider that such a staircase procedure, by definition, cannot converge to a performance level just below the threshold of visibility (i.e., a stimulus that is just below the threshold of visibility, and a stimulus that is not even presented would both yield a performance level of 50%). Therefore, we used a slightly longer presentation time (16.7 ms) with a slightly higher desired accuracy (60% correct) in the staircase. Initial piloting suggested that a reduction in presentation time of 8.3ms would indeed lead to a decrease in accuracy of about 10%.

The pre-experiment, aimed at determining the RGB values for the circle and the triangle of the target stimuli consisted of two steps. In the first step, we aimed to adjust the RGB values of the circle (thus changing the contrast with the background) to get above-chance sensitivity in the localization task. Specifically, we aimed at 75%

accuracy, by staircasing the RGB values of the circle relative to the gray background. We simultaneously ran two independent staircases for this first step, with initial 'contrast' values of 0.1 and 0.9, where a value of 0 indicates that the circle had the same color as the gray background, and a value of 1 indicates a green circle of maximum luminance (RGB balance [1 0 0]). The ASA algorithm converges to a desired accuracy level, by (A) adaptively adjusting the stimulus 'contrast' depending on response accuracy, and (B) by gradually decreasing the size of the contrast change (or step size). The initial step size was set to 0.8, and decreased over time, and as a function of the number of reversals. In the second step, we aimed to adjust the RGB values of the triangle (thus changing the contrast with the circle) to obtain near chance-level sensitivity for discriminating between upright and downward pointing triangles. The circle color was obtained from the previous step (described above), and participants' discrimination sensitivity was manipulated by adjusting the RGB values of the triangle (relative to the circle). We set the desired accuracy level at 60%, and the starting values at 0.1 and 0.9. The initial step size was set to 0.8.

Participants finished the main experiment on another day than the pre-experiment. In the pre-experiment, each set of staircases consisted of 128 trials (64 trials for each starting value; 0.1 and 0.9). In the main experiment, we set different numbers of trials for the two groups of participants that we recruited. For the purpose of exploring the stability of the results over time, the first group of participants performed 384 experimental trials. Then, with the original purpose of replicating the effects of the first epoch that we found in the first group of participants (see http://256.so/i1n), the second group of participants only performed 96 experimental trials (i.e., the length of one epoch). Except for the actual length of the experiment, the experimental procedures were identical for both groups.

Data Analysis. The (first) 96 trials of the two groups of participants were pooled together for exploratory analyses. Because we are only interested in interpreting localization performance when orientation discrimination sensitivity *d'* is at chance level, we only analyzed the data in the shortest presentation time condition, in which the unconsciousness of feature discrimination dimension was most likely to be established.

All responses from the localization/discrimination tasks were transformed to the signal-detection theory sensitivity index *d'*. For the discrimination measure, the correct responses for high-probability features (e.g., 'upright' responses for 'upright' triangles) were coded as hits in the high-probability feature trials, while the same responses (e.g., 'upright' responses for 'inverted' triangles) were coded as false alarms in the low-probability trials. In order to test whether participants were unconscious of the triangle features (which is the prerequisite for comparing the localization performance), we conducted a two-tailed, one-sample t-test to compare the discrimination sensitivity *d'* with 0 (discrimination sensitivity $d' > 0$ indicates awareness of the triangle features).

For the localization task, 'right' responses were coded as hits in 'right' trials and as false alarms in 'left' trials. Hit and false alarm rates of 0 or 1 were converted to

1/(2N) and 1−1/(2N), respectively; N refers to the number of trials on which the rates are based (Macmillan & Creelman, 2004). The hit and false alarm rates were eventually transformed to z-scores to calculate the final *d’* values. In order to test whether high-probability feature leads to higher localization sensitivity *d’*, we compare the localization *d’* of the high-probability (feature) trials with that of the low-probability (feature) trials. In the originally planned (and preregistered) data analysis, we conducted a straightforward, two-tailed paired-sample t-test to compare the localization sensitivity *d’* of the high-probability (feature) trials and that of the low-probability (feature) trials. However, after running simulations, we found that when trial numbers are low (as is the case in the low-probability condition), SDT sensitivity *d’* will tend to go toward zero. Therefore, the lower localization *d’* in the low compared to the high probability condition could be explained by the lower number of trials in the low-probability condition. To exclude this possible account, we used a bootstrapping method to equate the trial numbers for the two experimental conditions. Specifically, in each iteration of the bootstrapping procedure, we sampled the same number of trials from the high-probability trials (8 out of 24) and the low-probability trials (all 8 trials, within a given timing condition). This led to a total of 245,025 iterations, with 8 high-probability trials and 8 low-probability trials (reflecting the total number of possible ways to draw 8 trials from 24, while keeping the number of left and right target locations balanced to four on each side). In each iteration, a difference between the localization *d’* of the 8 sampled high-probability trials and the 8 low-probability trials was computed for each individual participant, and averaged across all participants. This resulted in the exhaustive set of 245,025 group level differences between high and low-probability localization *d*', based on equal trial numbers in both conditions. To test for significance, we computed the fraction of iterations in which high probability trials yielded a larger localization *d′* than low probability trials out of the full set of 245,025 group level difference scores (akin to a bootstrap test, but based on the exhaustive set of comparisons rather than on random permutations of the data). If a positive value was observed on more than 95% of iterations, this was regarded as evidence for higher *d'* in high-probability trials, given an alpha level of 0.05 (Fig. 4B). Note that this approach entails a directional test, which follows from the strong prediction that localization *d′* is larger for high-probability trials compared with low-probability trials.

Results (Experiment 3)

Before we balanced the numbers of high- and low-probability trials, we first conducted the planned paired-samples t-test on the pooled data of Experiment 3 (see Supplementary Materials 3 for the original separate analyses of Experiment 3A and 3B). At the shortest presentation time (8.3 ms), the orientation discrimination *d'* (0.18, SD = 0.68) was not significantly higher than zero, $t(48) = 1.82$, $p = .075$, Cohen's $d = 0.26$, 95% CI = [-0.02, 0.37]. This result provides evidence (albeit weak) that participants did not have conscious access to the critical feature dimension at the group level in the shortest presentation duration condition. In line with our hypothesis, however, we found significantly higher localization sensitivity *d'* for high-probability features (1.43, SD = 0.61) than low-probability features (0.95, SD = 86), *t*(48) = 3.69, *p* < .001, Cohen's *d* = 0.53, 95% CI = [0.22, 0.73]). The lower localization *d*' in the low-probability condition could have been caused by the smaller number of trials in the low-probability compared to the high-probability condition. To account for this, we next conducted a bootstrap test in which we equated the numbers of trials in the high- and low-probability (feature) conditions. As shown in Fig. 4B, results of the bootstrap test show that localization sensitivity *d'* was higher in the high-probability (1.16, SD = 0.42) compared to the low-probability condition (0.95, SD = 0.86), *p* = .027 (245,025 bootstrap samples).

A rank-based Spearman's correlation test showed that there was no significant correlation between the discrimination *d'* and the difference in localization *d'* between high and low probability features across participants, *r* = -0.02, *p* = 0.871 (Fig. 4C). These results suggest that the difference in localization performance were not related to the discrimination performance. Therefore, the effect of statistical regularities on localization performance was independent of conscious access to the feature carrying the regularity.

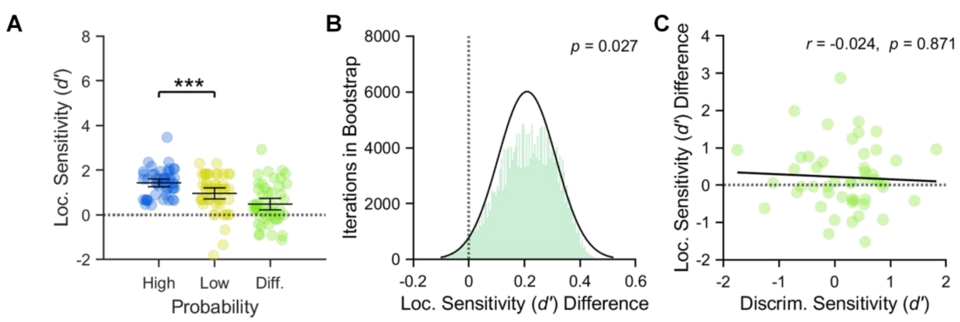

Fig. 4. The results of Experiments 3 (the shortest presentation time of the first 96 trials). (A) depicts the localization sensitivity *d'* for high and low-probability features (before bootstrapping) and the difference between the two conditions, with individual dots representing individual participants. (B) shows the distribution of all group-level differences in localization *d*' between equal numbers of high and low-probability trials Positive values on the x axis represent a higher localization *d'* for high-probability compared to low-probability features. (C) shows the correlation between (bootstrapped) discrimination sensitivity *d'* and the average localization *d'* difference between high and low-probability trials for each participant. Error bars represent the 95% confidence interval of the mean. Asterisks indicate significance (***$p < .001$).

Taken together, we did not observe compelling evidence for zero visibility of the critical feature at the group level. However, we did observe that higher localization sensitivity for high-probability features was not correlated with the level of visibility of feature, which indicates that the higher localization sensitivity for high-probability feature does not depend on conscious access to the feature dimension that carried the regularity. This implies that, the better localization sensitivity for high- compared to low-probability features also applied to participants with chance-level feature discrimination sensitivity. These results provide evidence that statistical regularities can affect the processing of visual input that is not available to consciousness. This conclusion, however, is based on a data analysis that was not originally planned, and therefore exploratory. To confirm these findings, we conducted Experiment 4.

Methods (Experiment 4)

In Experiments 3, we found that participants had a higher localization sensitivity for high- compared to low-probability features and that this difference did not depend on their awareness of the dimension that carried the regularities (i.e., feature discrimination). Although the results were in line with our original hypotheses, the evidence was based on exploratory analyses, which we regard as insufficient evidence to bolster our claims. Therefore, a pre-registered replication experiment was required to confirm the exploratory results of Experiment 3. For this purpose, we turned to a simpler paradigm, which is similarly capable of isolating differences in conscious access from post-perceptual processes. We decided against conducting another detection-discrimination dissociation experiment, in consideration of the following factors: First, a substantial number of trials is required for signal detection measures (such as d') to be reliable, but we cannot trivially increase the number of low-probability trials. This is because the DDD paradigm requires finding a very specific stimulus presentation intensity that - on the one hand - yields zero sensitivity for discriminating between the two target features, and - on the other hand - does not completely abolish stimulus processing. Such a setting is not only difficult to approximate empirically, but is nearly impossible to preserve over the course of experimental trials due to training effects (i.e., the increase of trials leads to a higher discrimination sensitivity). Second, conducting a properly powered replication of Experiment 3 would cost an unrealistic amount of financial resources, as it requires many participants (110 for 80% power in a one-tailed test).

Akin to Experiment 3, the goal of Experiment 4 was to test whether statistical learning affects conscious access, or whether it only affects processes arising after stimulus detection. To this aim, we used an accuracy-based variant of the b-CFS paradigm (Litwin et al., 2023) in which we compared the localization accuracy for high-probability versus low-probability features while excluding processing (e.g., criterion) differences emerging *after* conscious access. The advantage of this method, compared with the DDD paradigm used in Experiment 3, is that the interpretability of the data does not hinge on establishing chance level performance on a secondary (i.e., discrimination) task. The methods and hypothesis of this experiment were preregistered becore data collection (http://256.so/fopr).

In this paradigm, participants performed a non-speeded two-alternative forced-choice (2AFC) localization task during viewing of a (b-)CFS presentation (i.e., CFS masks to one eye and a target to the other eye). The duration of the CFS presentation was pre-determined before every trial and kept identical between conditions of interest (e.g., high- and low-probability conditions in our case). Because responses in this paradigm are non-speeded, participants' responses on the forced-choice task reflect how much information they obtained about a stimulus within a given presentation duration. If for a specific presentation duration (e.g., yielding ~80% localization accuracy on average across high and low-probability conditions), participants have more information about (e.g., the identity or location of) a stimulus in condition A compared to condition B, we can establish that conscious access of the stimulus (location/identity) was faster in condition A than in condition B.

This precludes any effect of decisional biases and post-detection effects, as it exhaustively measures the amount of information available to the participant within a specific timeframe (Litwin et al., 2023). As in Experiment 3, we manipulated the regularity of target features (i.e., upright vs. inverted triangles) and hypothesized that targets with high-probability (vs. low-probability) features break the suppression of CFS masks earlier and thus can be localized better in the non-speeded forced-choice task.

Participants. Based on the effect size (Cohen's $d = 0.34$) of a previous study that used the bias-free b-CFS paradigm to measure conscious access (Litwin et al., 2023), a sample of 55 participants was needed for an experimental power of 80% with an alpha level of 0.05 for a planned one-tailed paired-samples t-test (power calculation performed in G*Power). We opted to preregister a one-tailed test because we have clear predictions on the directionality of the effect following Experiments 1-3. In order to counterbalance the between-subject condition (i.e., upright or inverted triangles are the high or low-probability feature), we recruited one more participant than was specified in the pre-registration (i.e., fifty-six participants). A new group of sixty-one participants were recruited. For each participant, we simultaneously ran two independent staircases (for stimuli appearing left and right of fixation; see Data Analysis section below) ; participants were excluded according to their performance in either staircase. We excluded five participants whose average accuracy (in both staircases) was lower than 65% or higher than 95% from data analysis (see preregistered analysis plan at http://256.so/fopr). This resulted in the planned sample size of fifty-six included participants (49 women and 7 men, mean age = 24.57, SD = 2.82).

Apparatus and stimuli. The apparatus and stimuli used in Experiment 4 were the same as in Experiments 1 and 2.

Procedure. For any given participant, either the upward pointing triangle or the downward pointing triangle was selected as the high-probability feature (e.g., the triangle pointed upwards in 75% of all trials), while the other one was the low-probability feature (e.g., the triangle pointed downwards in the remaining 25% of all trials). Which one was selected as the high-probability feature was counterbalanced across participants. Within both the low and high-probability conditions, different target location conditions (left or right of fixation) occurred equally often.

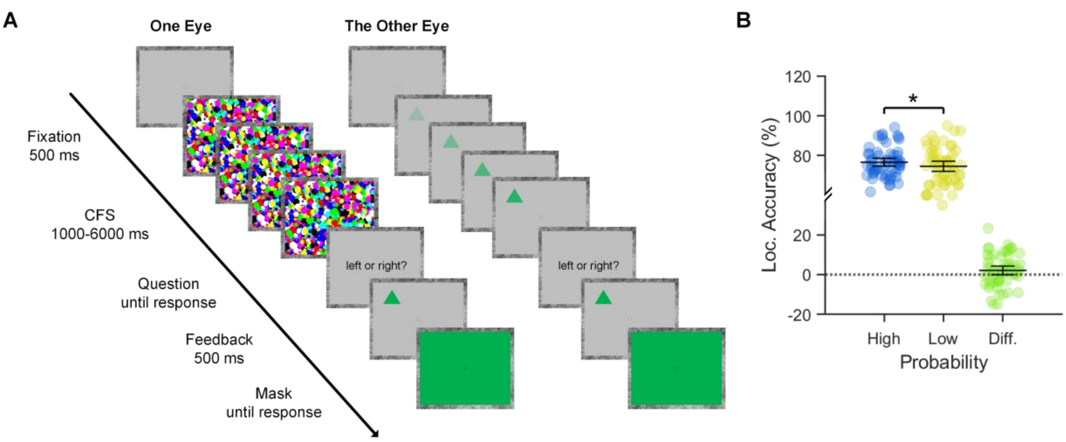

Fig. 5. (A) Trial outline of the accuracy-based breaking continuous flash suppression (b-CFS) paradigm used in Experiments 4. (B) The localization accuracy in the high-probability and low-probability conditions, with

individual dots representing individual participants. Error bars represent the 95% confidence interval of the mean. Asterisks indicate significance (*$p < .05$).

At the beginning of each trial, a central white fixation dot (with black edge) appeared at the center of the presentation area for 500 ms. After this, a dynamic CFS mask (consisting of a number of so-called Mondrian images, randomly chosen from 120 generated images, and replaced at 10 Hz without repetition) randomly appeared to the dominant eye of participants. Between 300 and 600 ms after the onset of the dynamic CFS mask (to the dominant eye), the target (a triangle) was presented to the non-dominant eye of the participant (either to the left or right side of fixation, with equal probability), and remained on the screen for the duration that was pre-determined by the staircase procedure (ranging between 1 and 6 seconds). During the presentation of the target, the intensity (i.e., opacity) of the target linearly ramped up from zero to the eventual opacity (30%, 50% or 60% of the original stimulus opacity of Experiments 1 and 2; depending on the performance in practice session, see below) within 2 seconds, regardless of the determined target presentation duration.

After the presentation of target and mask stimuli, a message 'presented left or right?' appeared on the screen, requiring participants to press one of two arrow keys ('←' for left, '→' for right) to indicate on which side of fixation the target was presented (i.e., a two-alternative forced-choice localization task). Participants were instructed to respond as accurately as possible, without any time pressure. After the response, the target stimulus remained present for 500 ms for both eyes, which ensured that participants got the opportunity to learn that one triangle orientation was more prevalent than the other. At the same time, they received auditory feedback - a high-pitch (2,000 Hz), or a low-pitch (1,500 Hz) ‘beep’ sound, indicating a correct or incorrect response respectively. At the end of the trial, the presentation area was filled with the same (green) color that was used for the target triangle to minimize after-images at the target location before onset of the next trial. After pressing the space bar, the next trial began.

Participants completed 24 trials for determining eye dominance and 32 practice trials in the pre-experiment, and then completed 5 blocks of 32 trials in the formal experiment. At the end of the experiment, we measured participants’ awareness of statistical regularities in a questionnaire as in Experiments 1 and 2.

To avoid ceiling or floor effects (or have more trials available for data analysis), it is necessary to keep the overall localization performance at an consistent level across different visual fields and different phases of the experiment. As mentioned above, the presentation duration (ranging from 1 to 6 seconds) of the target on a given trial was determined by an ongoing adaptive staircase procedure. Specifically, a 2-down/1-up adaptive staircase procedure decreased target presentation duration after two consecutive correct responses and increased target presentation duration after an incorrect response. The 2-down/1-up adaptive staircase method allows the algorithm to reliably converge on the individual presentation time thresholds that yields a localization accuracy of 80.35% (García-Pérez, 2001). Because suppression durations are known to substantially differ between nasal and temporal visual hemifields (Sahakian et al., 2022), we ran two interleaved staircases for different presentation positions (left vs. right side of the central fixation) respectively. The localization performance of the high- and low-probability trials were thus compared within staircases first, and averaged afterwards. Presentation times were increased or decreased in a stepwise manner, following a logarithmic scale. For practice trials, the range of presentation durations one 1-6 seconds was divided into only ten steps, with presentation durations in each step being 1.195 times longer or shorter compared to the adjacent steps. This allowed the algorithm to quickly (but coarsely) converge to an appropriate performance level for each participant. The initial duration of target presentation for the practice trials was set at an intermediate level (about 2.914 sec). In experimental trials, the starting value of the staircase was the final value obtained in the practice trials. Here the range of 1-6 seconds was divided into 30 logarithmically spaced steps (yielding a factor of 1.061), thus allowing to more precisely adjust the presentation duration to keep performance stable throughout the experiment.

To account for extreme individual differences in localization performance, we also adjusted the target opacity during the practice session. The default (full) opacity of the target stimulus was 50% of the original stimulus opacity (of Experiments 1 and 2). The stimulus opacity would be increased to 60% (or decreased to 30%) of the original opacity if the target presentation duration (i.e., stimulus intensity) of a single staircase in practice trials went above 4.161 sec (or went below 1.707 sec). Target opacity was only adjusted during the practice session, and the opacity for formal experimental trials depended on the final opacity values obtained in the practice trials.

Data analysis. According to the preregistered analysis plan (http://256.so/fopr), we used a one-tailed paired-samples t-test to compare the localization accuracy of the high-probability condition to that of the low-probability condition. To avoid ceiling or floor effects, we excluded the data from an entire staircase (i.e., left or right target location) if the average accuracy of that staircase exceeded the pre-defined ceiling or floor (i.e., accuracy $< 65\%$ or $> 95\%$). According to this preregistered data exclusion principle, five participants were completely excluded from data analysis, and for twenty participants, the data from one staircase was excluded. For each of these twenty participants, the exclusion of one staircase may have resulted in decreased precision of estimation for the average localization accuracy However, there were still considerable number of trials (80 trials) after data exclusion.

Results (Experiment 4)

The b-CFS task. Results show that localization accuracy was higher for high-probability targets (0.77, SD = 0.07) compared to low-probability targets (0.75, SD = 0.10), $t(55) = 1.88$, $p = .033$, Cohen's $d = 0.25$, 95% CI = [0.00, ∞]. The higher localization accuracy for high-probability (versus low-probability) features was in line with our hypothesis in the preregistration, providing evidence that conscious access is enhanced for targets with high-probability (versus low-probability) features. Importantly, since we used an accuracy-based measurement, the differential localization performances cannot be accounted for by decisional biases and post-detection effects. To further assert robustness of the results, we plotted the effect size as a function of included participant for Experiments 1-4 (see Supplementary Materials 4), and showed that a wide range of predetermined sample sizes resulted in the same conclusions across different paradigms.

Inter-trial priming effects. As in Experiments 1 and 2, an exploratory two-tailed paired-sample t-test was conducted to examine the effect of Inter-trial Continuity (change vs. repeat) on SL effect (i.e., the difference between accuracy of high probability and low probability trials). Results showed that there was no difference between SL effect of change trials and repeat trials, $t(55) = 1.47$, $p = .146$, Cohen's $d = 0.20$, 95% CI = [–0.02, 0.12]), showing that there was no inter-trial priming effect. Therefore, the differential RTs between high- and low-probability locations is unlikely to have been caused by inter-trial priming.

General Discussion

The present study examined the effects of statistical learning (SL) on conscious access using three different paradigms and provided converging evidence that SL prioritizes conscious access for probable items over improbable items. In two b-CFS experiments, targets broke through interocular suppression faster when they appeared at probable locations or contained probable features, providing preliminary evidence for the influence of SL on conscious access. In the third (DDD) experiment, we observed that the perceptual advantage for probable (versus improbable) feature items was not correlated to the conscious access of the feature dimension that carried the regularity. In the last (accuracy-based b-CFS) experiment, we excluded potential contributions of decisional and post-perceptual factors, and again showed higher localization performance for probable (versus improbable) features.

Our study goes beyond existing work in showing that statistical learning, as an implicit learning process, alters the priority of visual input for conscious access.. Consistent with previous studies suggesting that SL operates implicitly and consumes few cognitive resources (including conscious resources; Turk-Browne et al., 2005), most of our participants had no explicit knowledge of these regularities, while they nonetheless differentially prioritized the conscious access. This suggests that the selection of information for conscious access (what we become aware of), is itself governed by unconscious processes. This makes sense, considering that conscious resources are scarce (e.g., Dehaene et al., 2001) and important for human cognition (e.g., Baars et al., 2005). In contrast to our findings, previous studies did not find an effect of implicit learning on the detection of interocularly suppressed stimuli (e.g., Paffen et al., 2018). This may indicate that only certain types of implicit learning processes can influence conscious access.

The effect of SL on conscious access provides a new explanation for how SL influences a range of cognitive functions (Bogaerts et al., 2020; Frost et al., 2019): SL may enhance the perception of probable information by making it more consciously accessible. For example, the enhanced allocation of attentional resources (e.g., Geng & Behrmann, 2005; Hoffmann & Kunde, 1999; Miller, 1988) for probable (versus improbable) stimuli might be partially attributed to their faster entry into conscious awareness. Furthermore, the effect of SL on conscious access may be crucial in many high-level cognitive functions that depend on conscious resources (e.g., memory, language, and inference; Baars & Franklin, 2003; Rich & Mattingley, 2002; Sabary et al., 2020). SL might, for example, enhance the encoding of probable stimuli into working memory systems (e.g., Umemoto et al., 2010) which operate largely on conscious resources (Giattino et al., 2018; Baars & Franklin, 2003).

We do not claim that SL *always* causes probable items to enter consciousness faster: probable items might be prioritized or de-prioritized, depending on task goals. In support of this, Denison et al. (2016) showed that statistically unlikely images can be prioritized over statistically likely images when the former are more informative. Moreover, Wang and Theeuwes (2018) found that probable singleton distractors capture *less* attention than improbable ones. These effects can come about by probable distractors being de-prioritized for conscious access, or, alternatively, by prioritizing probable distractors for conscious access, after which they are (consciously) disengaged from faster. In sum, it remains open for investigation whether or not high-probability items are always prioritized for conscious access.

It also remains unknown by what mechanism highly probable events are prioritized by the visual system. One possibility is that SL evokes preparatory responses in anticipation of upcoming visual events. Anticipation is indeed a typical consequence of SL (Turk-Browne et al., 2010), and has been shown to modulate the detection of interocularly suppressed stimuli (Denison et al., 2011, 2016; Pinto et al., 2015). Predicting the upcoming image from a sequence of images, for instance, can facilitate detection of expected images during interocular suppression (Denison et al., 2011), and the anticipation of visual stimuli can evoke feature-specific activity patterns in early visual cortex, resembling visually evoked responses (Kok et al., 2012, 2014, 2017; Gayet & Peelen, 2022). Thus, SL may pre-activate stimulus-specific representations in primary visual cortex, thereby lowering the effective threshold for probable stimuli to breach conscious access. Another possibility is that the influence of SL on conscious access is mediated by attention. Since unconscious information has been shown to guide the allocation of spatial attention (Jiang et al., 2007; Hsieh et al., 2011), probable items might gain faster conscious access because more (spatial or feature-based) attention is directed towards them (e.g., Sun et al., 2016). Note however, that it also been shown that sustained attention to a target stimulus does not always alleviate suppression (e.g., Gayet et al., 2020).

The SL effects shown here might be reminiscent of the preferential conscious access of familiar over unfamiliar stimuli (Gobbini et al., 2013; Jiang et al., 2007; Ramon & Gobbini, 2018; Stein et al., 2012). Familiarity and SL are distinct, however, as they differ in a number of key aspects. First, SL refers specifically to the ability to (often implicit and rapid, see Turk-Browne et al., 2005) extract visual patterns from varying sensory inputs. In contrast, familiarity (of human faces, language et cetera) as assessed in previous b-CFS studies (Gobbini et al., 2013; Jiang et al., 2007; Ramon & Gobbini, 2018; Stein et al., 2012) was caused by explicit cognitive processes, and was likely due to long-lasting learning processes. Notably, short-term extraction of statistical regularities (e.g., SL in the current task setting) has been shown to influence perception differently from life-long learned familiarity, even within the same experimental context (e.g., Dogge et al., 2019; Aldegheri et al., 2023). Furthermore, not all learning processes necessarily affect conscious access (e.g., Heyman & Moors, 2014; Paffen et al., 2018), although learning processes (in theory) could lead to increases in familiarity.

In conclusion, we show that the visual system rapidly and implicitly extracts statistical regularities from streams of sensory input to promote the selection of information for conscious processing. Given that conscious resources are scarce, and that conscious access is a prerequisites for a myriad of cognitive functions, our findings provide a mechanism for how statistical learning underlies a broad range of cognitive functions.

References

Aldegheri, G., Gayet, S., & Peelen, M. V. (2023). Scene context automatically drives predictions of object transformations. *Cognition*, *238*, 105521.

https://doi.org/10.1016/j.cognition.2023.105521

Baars, B. J. (2005). Global workspace theory of consciousness: toward a cognitive neuroscience of human experience. *Progress in Brain Research*, *150*, 45–53. https://doi.org/10.1016/S0079-6123(05)50004-9

Baars, B. J., & Franklin, S. (2003). How conscious experience and working memory interact. *Trends in Cognitive Sciences*, *7*(4), 166–172. https://doi.org/10.1016/S1364-6613(03)00056-1

Bogaerts, L., Frost, R., & Christiansen, M. H. (2020). Integrating statistical learning into cognitive science. *Journal of Memory and Language*, *115*, 104167. https://doi.org/10.1016/j.jml.2020.104167

Brady, T. F., & Oliva, A. (2008). Statistical learning using real-world scenes: Extracting categorical regularities without conscious intent. *Psychological Science*, *19*(7), 678–685. https://doi.org/10.1111/j.1467-9280.2008.02142.x

Brady, T. F., Konkle, T., & Alvarez, G. A. (2009). Compression in visual working memory: Using statistical regularities to form more efficient memory representations. *Journal of Experimental Psychology: General*, *138*(4), 487–502. https://doi.org/10.1037/a0016797

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*(4), 433–436. https://doi.org/10.1163/156856897X00357

Dehaene, S., Naccache, L., Cohen, L., Bihan, D. L., Mangin, J. F., Poline, J. B., & Rivière, D. (2001). Cerebral mechanisms of word masking and unconscious repetition priming. *Nature Neuroscience*, *4*(7), 752–758.

https://doi.org/10.1038/89551

Denison, R. N., Piazza, E. A., & Silver, M. A. (2011). Predictive context influences perceptual selection during binocular rivalry. *Frontiers in Human Neuroscience*, *5*, 13757. https://doi.org/10.3389/fnhum.2011.00166

Denison, R. N., Sheynin, J., & Silver, M. A. (2016). Perceptual suppression of predicted natural images. *Journal of Vision*, *16*(13), 6. https://doi.org/10.1167/16.13.6

Dogge, M., Custers, R., Gayet, S., Hoijtink, H., & Aarts, H. (2019). Perception of action-outcomes is shaped by life-long and contextual expectations. *Scientific Reports*, *9*(1), 5225. https://doi.org/10.1038/s41598-019-41090-8

Dotsch, R., Hassin, R. R., & Todorov, A. (2017). Statistical learning shapes face evaluation. *Nature Human Behaviour*, *1*(1), 1–6. https://doi.org/10.1038/s41562-016-0001

Frost, R., Armstrong, B. C., & Christiansen, M. H. (2019). Statistical learning research: A critical review and possible new directions. *Psychological Bulletin*, *145*(12), 1128–1153. https://doi.org/10.1037/bul0000210

García-Pérez, M. A. (2001). Yes-no staircases with fixed step sizes: psychometric properties and optimal setup. *Optometry and Vision Science*, *78*(1), 56–64. https://doi.org/10.1097/00006324-200101010-00015

Gayet, S., & Peelen, M. V. (2022). Preparatory attention incorporates contextual expectations. *Current Biology*, *32*(3), 687–692. https://doi.org/10.1016/j.cub.2021.11.062

Gayet, S., Douw, I., Van der Burg, V., Van der Stigchel, S., & Paffen, C. L. (2020). Hide and seek: Directing top-down attention is not sufficient for accelerating conscious access. *Cortex*, *122*, 235–252. https://doi.org/10.1016/j.cortex.2018.08.027

Gayet, S., Paffen, C. L. E., & van der Stigchel, S. (2013). Information matching the content of visual working memory is prioritized for conscious access. *Psychological Science*, *24*(12), 2472–2480. https://doi.org/10.1177/0956797613495882

Gayet, S., van der Stigchel, S., & Paffen, C. L. E. (2014). Breaking continuous flash suppression: Competing for consciousness on the pre-semantic battlefield. *Frontiers in Psychology*, *5*, 460. https://doi.org/10.3389/fpsyg.2014.00460

Gayet, S., van Maanen, L., Heilbron, M., Paffen, C. L. E., & Van der Stigchel, S. (2016). Visual input that matches the content of visual working memory requires less (not faster) evidence sampling to reach conscious access. *Journal of Vision*, *16*(11), 26. https://doi.org/10.1167/16.11.26

Geng, J. J., & Behrmann, M. (2005). Spatial probability as an attentional cue in visual search. *Perception and Psychophysics*, *67*(7), 1252–1268. https://doi.org/10.3758/BF03193557

Giattino, C. M., Alam, Z. M., & Woldorff, M. G. (2018). Neural processes underlying the orienting of attention without awareness. *Cortex*, *102*, 14–25. https://doi.org/10.1016/j.cortex.2017.07.010

Gobbini, M. I., Gors, J. D., Halchenko, Y. O., Rogers, C., Guntupalli, J. S., Hughes, H., & Cipolli, C. (2013). Prioritized detection of personally familiar faces. *PloS ONE*, *8*(6), e66620. https://doi.org/10.1080/13506285.2017.1405134

Heyman, T., & Moors, P. (2014). Frequent words do not break continuous flash suppression differently from infrequent or nonexistent words: Implications for semantic processing of words in the absence of awareness. *PLoS ONE*, *9*(8), e104719. https://doi.org/10.1371/journal.pone.0104719

Hoffmann, J., & Kunde, W. (1999). Location-specific target expectancies in visual search. *Journal of Experimental Psychology: Human Perception and Performance, 25*(4), 1127–1141. https://doi.org/10.1037/0096-1523.25.4.1127

Hsieh, P. J., Colas, J. T., & Kanwisher, N. (2011). Unconscious pop-out: attentional capture by unseen feature singletons only when top-down attention is available. *Psychological Science*, *22*(9), 1220–1226. https://doi.org/10.1177/095679761141930

Jiang, Y., Costello, P., & He, S. (2007). Processing of invisible stimuli: Advantage of upright faces and recognizable words in overcoming interocular suppression. *Psychological Science*, *18*(4), 349–355. https://doi.org/10.1111/j.1467-9280.2007.01902.x

Kesten, H. (1958). Accelerated stochastic approximation. *The Annals of Mathematical Statistics*, *29*(1), 41–59. https://doi.org/10.1214/aoms/1177706705

Kok, P., Failing, M. F., & de Lange, F. P. (2014). Prior expectations evoke stimulus templates in the primary visual cortex. *Journal of Cognitive Neuroscience*, *26*(7), 1546–1554. https://doi.org/10.1162/jocn_a_00562

Kok, P., Jehee, J. F., & De Lange, F. P. (2012). Less is more: expectation sharpens representations in the primary visual cortex. *Neuron*, *75*(2), 265–270. https://doi.org/10.1016/j.neuron.2012.04.034

Kok, P., Mostert, P., & De Lange, F. P. (2017). Prior expectations induce prestimulus sensory templates. *Proceedings of the National Academy of Sciences*, *114*(39), 10473–10478. https://doi.org/10.1073/pnas.1705652114

Litwin, P., Motyka, P., & Gayet, S. (2023). Physiological arousal underlies preferential access to visual awareness of fear-conditioned (and possibly disgust-conditioned) stimuli. *Emotion*. Advance online publication. https://doi.org/10.1037/emo0001296

Macmillan, N. A., & Creelman, C. D. (2004). Detection theory: A user's guide: 2nd edition. In *Psychology Press*. https://doi.org/10.4324/9781410611147

Meuwese, J. D., Post, R. A., Scholte, H. S., & Lamme, V. A. (2013). Does perceptual learning require consciousness or attention? *Journal of Cognitive Neuroscience*, *25*(10), 1579–1596. https://doi.org/10.1162/jocn_a_00424

Miller, J. (1988). Components of the location probability effect in visual search tasks. *Journal of Experimental Psychology: Human Perception and Performance, 14*(3), 453–471. https://doi.org/10.1037/0096-1523.14.3.453

Orbán, G., Fiser, J., Aslin, R. N., & Lengyel, M. (2008). Bayesian learning of visual chunks by human observers. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(7), 2745–2750. https://doi.org/10.1073/pnas.0708424105

Paffen, C. L. E., Gayet, S., Heilbron, M., & van der Stigchel, S. (2018). Attention-based perceptual learning does not affect access to awareness. *Journal of Vision*, *18*(3), 7. https://doi.org/10.1167/18.3.7

Pinto, Y., van Gaal, S., de Lange, F. P., Lamme, V. A. F., & Seth, A. K. (2015). Expectations accelerate entry of visual stimuli into awareness. *Journal of Vision*, *15*(8), 13. https://doi.org/10.1167/15.8.13

Ramon, M., & Gobbini, M. I. (2018). Familiarity matters: A review on prioritized processing of personally familiar faces. *Visual Cognition*, *26*(3), 179–195. https://doi.org/10.1080/13506285.2017.1405134

Rich, A. N., & Mattingley, J. B. (2002). Anomalous perception in synaesthesia: a cognitive neuroscience perspective. *Nature Reviews Neuroscience*, *3*(1), 43–52. https://doi.org/10.1038/nrn702

Sabary, S., Devyatko, D., & Kimchi, R. (2020). The role of visual awareness in processing of global structure: Evidence from the perceptual organization of hierarchical patterns. *Cognition*, *205*, 104442. https://doi.org/10.1016/j.cognition.2020.104442

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926–1928. https://doi.org/10.1126/science.274.5294.1926

Sahakian, A., Paffen, C. L. E., Van der Stigchel, S., & Gayet, S. (2022). A nasal visual field advantage in interocular competition. *Scientific Reports*, *12*(1), 4616. https://doi.org/10.1038/s41598-022-08473-w

Schmidt, T., & Vorberg, D. (2006). Criteria for unconscious cognition: Three types of dissociation. *Perception & Psychophysics*, *68*(3), 489–504. https://doi.org/10.3758/BF03193692

Sherman, B. E., Graves, K. N., & Turk-Browne, N. B. (2020). The prevalence and importance of statistical learning in human cognition and behavior. *Current Opinion in Behavioral Sciences*, 32, 15–20. https://doi.org/10.1016/j.cobeha.2020.01.015

Stein, T. (2019). The breaking continuous flash suppression paradigm: Review, evaluation, and outlook. *Transitions between consciousness and unconsciousness*, 1–38. https://doi.org/10.4324/9780429469688-1

Stein, T., & Peelen, M. v. (2021). Dissociating conscious and unconscious influences on visual detection effects. *Nature Human Behaviour*, *5*(5), 612–624. https://doi.org/10.1038/s41562-020-01004-5

Stein, T., Hebart, M. N., & Sterzer, P. (2011). Breaking continuous flash suppression: A new measure of unconscious processing during interocular suppression? *Frontiers in Human Neuroscience*, *5*, 167. https://doi.org/10.3389/fnhum.2011.00167

Stein, T., Sterzer, P., & Peelen, M. V. (2012). Privileged detection of conspecifics: Evidence from inversion effects during continuous flash suppression. *Cognition*, *125*(1), 64–79. https://doi.org/10.1016/j.cognition.2012.06.005

Sun, S. Z., Cant, J. S., & Ferber, S. (2016). A global attentional scope setting prioritizes faces for conscious detection. *Journal of Vision*, *16*(6), 9. https://doi.org/10.1167/16.6.9

Thibault L, van den Berg R, Cavanagh P, Sergent C (2016) Retrospective attention gates discrete conscious access to past sensory stimuli. *PLoS ONE*, *11*(2), e0148504. https://doi.org/10.1371/journal.pone.0148504

Tsuchiya, N., & Koch, C. (2005). Continuous flash suppression reduces negative afterimages. *Nature Neuroscience*, *8*(8), 1096–1101. https://doi.org/10.1038/nn1500

Turk-Browne, N. B. (2012). Statistical learning and its consequences. *The influence of attention, learning, and motivation on visual search*, 117–146.

Turk-Browne, N. B., Jungé, J. A., & Scholl, B. J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology: General*, *134*(4), 552–564. https://doi.org/10.1037/0096-3445.134.4.552

Turk-Browne, N. B., Scholl, B. J., Johnson, M. K., & Chun, M. M. (2010). Implicit perceptual anticipation triggered by statistical learning. *Journal of Neuroscience*, *30*(33), 11177–11187. https://doi.org/10.1523/JNEUROSCI.0858-10.2010

Umemoto, A., Scolari, M., Vogel, E. K., & Awh, E. (2010). Statistical Learning Induces Discrete Shifts in the Allocation of Working Memory Resources. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(6), 1419–1429. https://doi.org/10.1037/a0019324

Wang, B., & Theeuwes, J. (2018). Statistical regularities modulate attentional capture. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(1), 13–17. https://doi.org/10.1037/xhp0000472

Weiskrantz, L., Warrington, E. K., Sanders, M. D., & Marshall, J. (1974). Visual capacity in the hemianopic field following a restricted occipital ablation. *Brain*, *97*(1), 709–728. https://doi.org/10.1093/brain/97.1.709