Supporting Information

Schwiedrzik et al. 10.1073/pnas.1009147108

SI Materials and Methods

Participants. Twenty-two subjects (eight male, mean age 24 y, range 19–30 y) participated in the experiment after giving written informed consent. All but three subjects were right-handed as assessed with the Edinburgh inventory (1); all reported normal or corrected-to-normal vision and no history of neurological and/or psychiatric disease.

Stimuli. Stimuli were displayed on a CRT monitor (HP p1230, resolution 1024 × 768, visible screen size 30° × 22.9°) at a refresh rate of 100 Hz. Subjects viewed the screen from a distance of 75 cm. Background luminance was 3.18 cd/m^2. A square (0.35° × 0.35°) and a diamond (0.53° × 0.53°) were used as target stimuli. The outlines of the targets were 0.02° wide and had a luminance of 25.74 cd/m^2. The mask (25.74 cd/m^2) was star-shaped (Fig. 1*C*). Its inner edges were contiguous with the target stimuli from both sides without spatial overlap. In the main experiment, all stimuli were presented in the upper left quadrant at 4° eccentricity. For the transfer, stimuli were presented in the lower left quadrant, 6.6° from the trained position (center to center) at isoeccentricity (Fig. 1*B*). A fixation cross was always present at the center of the screen. Each trial started with a fixation period of 1,000–1,500 ms. Subsequently, the target was presented for 10 ms. The mask was presented for 50 ms at stimulus onset asynchronies (SOAs) ranging from 20 ms to 150 ms (20, 30, 40, 50, 70, 90, 110, 130, and 150 ms) for the threshold estimation or at an individually determined SOA during the training sessions (Fig. 1*A*).

Procedure. Subjects had to discriminate whether they saw a square or a diamond by pressing one of two buttons on a keyboard (counterbalanced within groups). Additionally, they had to rate the subjective visibility of the respective stimulus on the four-point Perceptual Awareness Scale (PAS; ref. 2) on a trial-by-trial basis by a button press. On this scale, 1 corresponds to "No experience," 2 to "Brief glimpse (a feeling that something has been shown)," 3 to "Almost clear experience (ambiguous experience of the stimulus)," and 4 to "Clear experience of the stimulus." Subjects were asked to maintain fixation on the center of the screen throughout the experimental sessions. The experiments were conducted in a darkened, sound-attenuating chamber. Constant head position was assured by the use of a chinrest with forehead support.

The experiment took place on 5 consecutive d. On the first day, we determined at which SOA a given subject performed the discrimination task at chance (20 ms: 7 subjects; 30 ms: 10 subjects; 40 ms: 4 subjects; 50 ms: 1 subject). This SOA was then used for the training. The first training session was conducted directly after the threshold measurement. On days 2–4, only training sessions were conducted. On the fifth day, the last training session took place. After this training session, we again assessed the masking threshold, followed by a threshold measurement at the transfer position.

For all threshold measurements, each target was presented 40 times at each SOA, yielding 80 trials per SOA and a total of 720 trials. After every 180 trials, we introduced a break of variable length. The occurrence of SOAs was randomized and counterbalanced over blocks. The sequence of target stimuli was fully randomized, and no feedback was given.

After the initial threshold measurement, subjects were randomly assigned to either a feedback or a no-feedback group. Subjects in the feedback group would receive blockwise percentage correct feedback during the training sessions. Subjects completed 600 trials per training session (a total of 3,000 trials). After every 100 trials, a break was introduced. The number of squares and diamonds was balanced per block. Subjects were paid €15 per hour. To assure constant motivation during the training sessions, subjects received a bonus of €2 if they improved by 10% from the previous training session or a fee of €2 if they did not improve.

Analysis. Squares were considered signal trials, and diamonds were considered noise trials. This setup yielded 40 signal and 40 noise trials per SOA for the threshold estimates and 300 signal and 300 noise trials per training sessions. For the calculation of sensitivity (d') and response bias (c), we used the log-linear correction to correct for extreme false alarm or hit-rate proportions (3). For subjective awareness, we calculated the mean PAS rating for correct and for incorrect responses, respectively. The mean PAS rating is suited to assess how subjective awareness changes gradually with learning. Furthermore, we split the PAS ratings into trials on which the target stimulus was minimally detected (PAS ≥ 2) and trials in which the subjects clearly saw the target stimulus (PAS = 4), and, thus, task-relevant information was subjectively available. For brevity, we refer to these splits as "subjective detection" and "subjective discrimination," respectively. Splitting the data this way allows us to investigate whether increases in subjective awareness are attributable to improved subjective detection and/or improved subjective discrimination. Furthermore, it allows for a more stringent comparison of the objective discrimination task with the subjective ratings because objective discrimination can be directly compared with subjective discrimination.

To directly evaluate the size of the learning effects in subjective detection and subjective discrimination, we calculated the respective gain at the trained and untrained locations, which takes into account the pretraining levels of subjective awareness: [(no. of trials with PAS ≥ 2 for posttraining or transfer)/(no. of trials with PAS ≥ 2 pretraining)] for subjective detection, and [(no. of trials with PAS = 4 for posttraining or transfer)/(no. of trials with PAS = 4 pretraining)] for subjective discrimination.

We also plotted the rate of incorrect trials against the rate of correct trials at three levels of visibility to obtain receiver operating characteristics (ROC) curves, which allows us to determine how well PAS ratings predict accuracy. Here, the hit rate refers to the percentage of correct trials with a high PAS rating, and the false-alarm rate refers to the percentage of incorrect trials with a high PAS rating. This is similar to the procedure known as Type II ROC analysis (4), where the relationship between confidence in one's response and the accuracy of that response is investigated by plotting the rate of correct responses with a high confidence rating (hits) against the rate of incorrect responses with a high confidence rating (false alarms). By varying the criterion at which a PAS rating is considered high (PAS = 1 vs. PAS = 2, 3, 4; PAS = 1, 2 vs. PAS = 3, 4; PAS = 1, 2, 3 vs. PAS = 4), we obtain three inflection points, to which we fitted ROC curves by using a proper binormal model (5, 6) in ROCkit (Kurt Rossmann Laboratories for Radiologic Image Research, http://xray.bsd.uchicago.edu/krl/). This model assumes likelihood ratio as the decision variable. From the ROC curves, we calculated the area under the curve (AUC), which indexes how well correct responses can be distinguished from incorrect responses based on PAS ratings. Cells for which the maximum-likelihood estimation in ROCkit did not converge were replaced by the average AUC of the respective group for further analyses.

In all repeated-measures ANOVA (rmANOVA) with more than one degree of freedom, we used the Greenhouse–Geisser correction and report adjusted degrees of freedom and P values. All P values for t tests are Bonferroni corrected for multiple comparisons, unless otherwise stated.

1. Oldfield RC (1971) The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia* 9:97–113.
2. Ramsøy TZ, Overgaard M (2004) Introspection and subliminal perception. *Phenom Cogn Sci* 3:1–23.
3. Hautus MJ (1995) Corrections for extreme proportions and their biasing effects on estimated values of d′. *Behav Res Methods Instrum Comput* 27:46–51.
4. Galvin SJ, Podd JV, Drga V, Whitmore J (2003) Type 2 tasks in the theory of signal detectability: Discrimination between correct and incorrect decisions. *Psychon Bull Rev* 10:843–876.
5. Metz CE, Pan X (1999) "Proper" binormal ROC curves: Theory and maximum-likelihood estimation. *J Math Psychol* 43:1–33.
6. Pesce LL, Metz CE (2007) Reliable and computationally efficient maximum-likelihood estimation of "proper" binormal ROC curves. *Acad Radiol* 14:814–829.

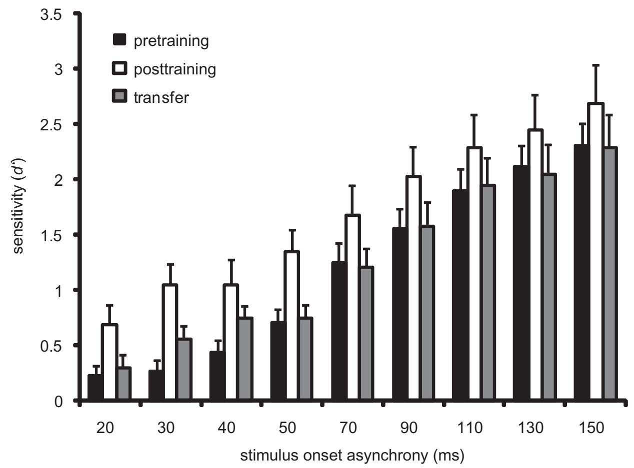

Fig. S1. Sensitivity (d') as a function of SOA before and after the training phase and at the untrained location. Sensitivity always increased linearly with SOA. Linear fits of the mean d' to the SOAs were highly significant at each threshold (all $R^2 > 0.9$, all $P < 0.01$). Error bars represent the SE of the estimated marginal mean.

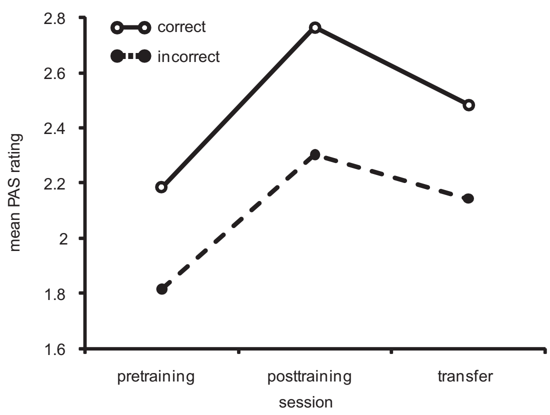

Fig. S2. Average subjective awareness for all nine SOAs in the threshold assessments before and after the training phase and at the untrained location. Empty cells were replaced by the average PAS rating from the respective group (feedback/no feedback).

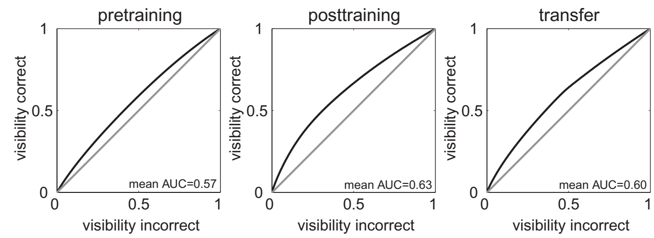

Fig. S3. Average ROC for the trained SOA for pretraining (*Left*) and posttraining (*Center*) and at the transfer location (*Right*). The ROC curves were obtained by fitting a line of 1,000 points to the three inflection points by means of maximum-likelihood estimation. The resulting AUC was significantly different from 0.5 (chance) in all three sessions (all $P < 0.01$, one-sided, Bonferroni corrected).

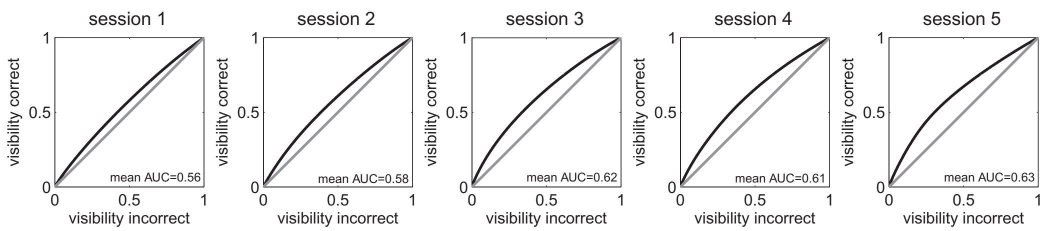

Fig. S4. Average ROC per training session. The AUC was always above chance (all $P < 0.01$, one-sided, Bonferroni corrected) and increased linearly with session [$F(1, 20) = 15.680$, $P < 0.01$, $\eta^2 = 0.439$].

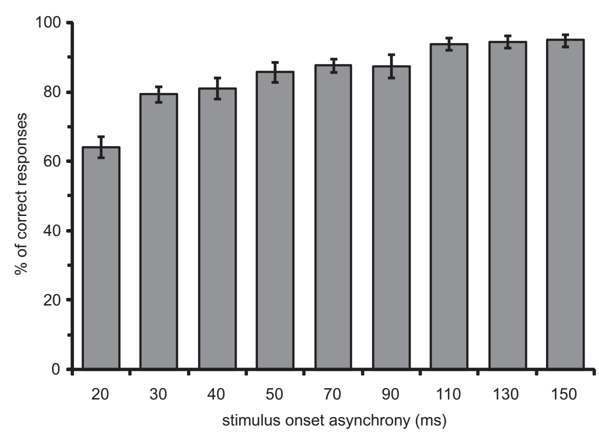

Fig. S5. Average percentage of correct PAS = 4 responses per SOA. Objective performance on clearly seen trials increases with SOA [main effect of SOA: $F(4.295, 85.891) = 21.567$, $P < 0.01$, $\eta^2 = 0.519$] and is at ceiling for the longest SOAs, which indicates that the relationship between objective performance and subjective awareness depends on the amount of available bottom-up information. Such behavior would be predicted by a model in which subjective awareness and objective performance rely on parallel channels with common input but independent source of noise: The more bottom-up information is available, the lower the influence of uncorrelated noise in the two channels. Suboptimal performance on clearly seen trials can also be attributed to motor errors or illusory percepts. Error bars represent the SE of the estimated marginal mean.