Consciousness = learning? The role of recurrent processing in perceptual learning
Meuwese, J.D.I.

Citation for published version (APA):
Meuwese, J. D. I. (2014). Consciousness = learning? The role of recurrent processing in perceptual learning

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Chapter 3. Does perceptual learning require consciousness or attention?

It has been proposed that visual attention and consciousness are separate and possibly even orthogonal processes. Attention and consciousness converge when conscious visual percepts are attended and hence become available for conscious report. In such a view, a lack of reportability can have two causes: the absence of attention or the absence of a conscious percept. This raises an important question in the field of perceptual learning. It is known that learning can occur in the absence of reportability, but it is unclear which of the two ingredients — consciousness or attention — is not necessary for learning. We presented textured figure-ground stimuli and manipulated reportability either by masking (which only interferes with consciousness) or with an inattention paradigm (which only interferes with attention). During the second session (24 hr later), learning was assessed neurally and behaviorally, via differences in figure-ground ERPs and via a detection task. Behavioral and neural learning effects were found for stimuli presented in the inattention paradigm and not for masked stimuli. Interestingly, the behavioral learning effect only became apparent when performance feedback was given on the task to measure learning, suggesting that the memory trace that is formed during inattention is latent until accessed. The results suggest that learning requires consciousness, and not attention, and further strengthen the idea that consciousness is separate from attention.

INTRODUCTION

Attention and awareness are intimately related processes, as we are generally conscious of stimuli that are in the focus of our attention. It has long been thought that these processes are identical or at least indistinguishable (Chun & Wolfe, 2001; De Brigard & Prinz, 2010; Merikle & Joordens, 1997; O’Regan & Noë, 2001; Posner, 1994; Prinz, 2010). However, in recent years, evidence is accumulating that attention and awareness are distinct processes that can be dissociated empirically and on a theoretical basis (Koch & Tsuchiya, 2007; Lamme, 2003; van Boxtel, Tsuchiya, & Koch, 2010a). Empirical evidence came from various experiments. An ERP component called visual awareness negativity, a negative difference wave around 200 msec after stimulus onset, correlates with visual awareness but is independent from top–down attention (Koivisto, Revonsuo, & Lehtonen, 2006; Railo, Koivisto, & Revonsuo, 2011). Distinct and independent neural correlates of visual awareness and spatial attention were also revealed by stimulus-induced oscillatory brain activity in the gamma range (Wyart & Tallon-Baudry, 2008). At the single-cell level, Supèr, Spekreijse, and Lamme (2001) showed that early neural responses in monkey primary visual cortex (V1) were unaffected by reportability of visual stimuli. A recent study independently manipulated attention and consciousness in a two-by-two design and found that V1 BOLD signals were selectively altered by the one (attention) and not the other (awareness) (Watanabe et al., 2011; see Koch & Tsuchiya, 2011, for a review). Also by means of a two-by-two design, van Boxtel, Tsuchiya, and Koch (2010b) showed that attention decreases afterimage duration, whereas consciousness prolongs the duration. This shows that the recently proposed orthogonality of attention and consciousness makes sense and can be tackled experimentally.

Separately manipulating attention and consciousness thus guides us toward the following two-by-two scheme where four types of neural processing can be distinguished: unattended and unconscious, attended yet unconscious, unattended yet conscious, and attended and conscious (see Figure 3.1). The latter condition is the “classical” case, where a conscious visual percept is attended, so that attention and awareness converge and the stimulus becomes available for conscious report. But as can be judged from Figure 3.1, a lack of reportability can have two causes: the absence of attentional access or the absence of a conscious percept (or, obviously, of both), with different neural consequences.

Disrupting consciousness can be done in a variety of experimental ways, the most well known of which is masking (Enns & Di Lollo, 2000; Fahrenfort et al., 2007). A properly masked stimulus remains fully invisible, even when it is attended. The lack of reportability of the stimulus in such a case is clearly because of the absence of consciousness, not because of a lack of attention. In our diagram (Figure 3.1), masked stimuli hence fall in the top right quadrant (“attended yet unconscious”). Taking away attention can be done by manipulations like change blindness or inattentional blindness. The result is attentional blindness, a lack of attentional access to low-level sensory signals (Kanai, Walsh, & Tseng, 2010; Mack & Rock, 1998; Rock, Linnert, Grant, & Mack,
Does perceptual learning require consciousness or attention?

These manipulations have been shown to affect the extent or depth of processing of stimuli, that is, whether stimuli evoke activity in higher, postperceptual levels or not (Rees, Russell, Frith, & Driver, 1999; Scholte et al., 2006; Vuilleumier et al., 2001). Perceptual processing typically remains largely intact. We put this kind of processing in the bottom left quadrant of the scheme of Figure 3.1, hence call it “unattended yet conscious.”

**Figure 3.1.** Four stages of neural processing emerge when manipulating either attention or consciousness or both. The manipulations that are being used in the current experiment are highlighted in white (top right and bottom left quadrant).

This is obviously a somewhat counterintuitive notion: it implies consciousness without cognitive access and reportability (Block, 2007; Lamme, 2003, 2006, 2010). Is this quadrant a realistic option? Do subjects in these conditions simply have no conscious sensation of the stimuli because of the removal of attention (i.e., absence of report implying absence of consciousness)? Or do they still have a conscious (or phenomenal) experience, which is however not available for report—the so-called “overflow argument” (Block, 2007, 2011), somewhat akin to the early versus late selection theories of attention, and related to the distinction between so called P- versus A-consciousness? Our stance is the latter, and we have provided arguments elsewhere (Lamme, 2003, 2006, 2010), but we acknowledge that the topic is subject to fierce debate (Cohen & Dennett, 2011; Kouider et al., 2010). We do not want to enter that complex discussion here but simply note that there are visual representations outside what is attended, that these are rich in content (Sligte, Vandenbroucke, Scholte, & Lamme, 2010; Vandenbroucke, Sligte, & Lamme, 2011), have perceptual characteristics (Moore & Egeth, 1997; Vandenbroucke, Fahrenfort, & Lamme, 2010; Vandenbroucke, Sligte, Fahrenfort, Ambroziak, & Lamme, 2012), and may deserve the status of being called “conscious.” At the least, there is agreement on the fact that these unattended representations are very different from truly unconscious representations, such as when stimuli are masked. We will call these representations “conscious” in this article, without implying a strong claim on our position in this debate. “Visible” or “potentially conscious” might have been equally proper denominations. Either way, we use this term for stimuli that fall into the bottom left quadrant of the scheme of Figure 3.1.
If we indeed acknowledge that attention and awareness are dissociable, the question arises whether this differential effect on stimulus processing has any long-term consequences. A typical long-term consequence of stimulus processing is perceptual learning. Experience and practice with sensory stimulation increases the ability to extract information from the environment (Gibson, 1969). Perceptual learning is stimulus-specific and seems to take place at early cortical stages where low-level stimulus attributes are processed (Ahissar & Hochstein, 1993). To date, the question whether attention or awareness is required for perceptual learning is still subject of debate. It is known that attended yet masked stimuli, although fully invisible, can cause short-lived priming effects. Lately, the boundaries have been pushed regarding the duration of these unconscious effects, but at most they seem to extend for about half an hour instead of eliciting any long-term learning (Capa, Cleeremans, Bustin, & Hansenne, 2011; see van Gaal, de Lange, & Cohen, 2012, for a review).

On the other hand, to our knowledge there is no evidence yet showing that perceptual learning does not occur for attended yet masked stimuli, as this has simply never been tested. For stimuli that are unmasked yet not attended however, perceptual processing is still intact and perceptual learning has been shown (Gutnisky et al., 2009; Jeneson, Kirwan, & Squire, 2010; Jessup & O’Doherty, 2009; Manns & Squire, 2001; Seitz et al., 2009; Seitz & Watanabe, 2003; Watanabe et al., 2001). Also, in studies of conditioning, that is, associating one stimulus with another stimulus, it is often found that this can occur without reportability of the stimulus contingencies (Clark, 1998; Clark, Manns, & Squire, 2001; Custers & Aarts, 2011; Knight, Nguyen, & Bandettini, 2003). Many studies have claimed that either attention or awareness is (not) needed for learning, however, while assuming that they are identical processes or at least without explicitly distinguishing the two (Ahissar & Hochstein, 1993, 1997; Gilbert, Sigman, & Crist, 2001; Jessup & O’Doherty, 2009; Shiu & Pashler, 1992). Therefore, it seems relevant to investigate systematically which of the two is really required for perceptual learning, using an experimental design where both mechanisms are independently manipulated.

Here, we tested which ingredient—consciousness or attention—is necessary for learning to detect a texture-defined stimulus. We chose texture-defined figure-ground stimuli, as consistent perceptual learning effects have been observed in the domain of texture segregation (Casco, Campana, Grieco, & Fuggetta, 2004; Censor, Bonneh, Arieli, & Sagi, 2009; Censor, Karni, & Sagi, 2006; Karni & Sagi, 1991; Ofen, Moran, & Sagi, 2007; Pourtois, Rausch, Vuilleumier, & Schwartz, 2008; Sagi & Tanne, 1994). When conscious and attended, figure-ground stimuli (or very related types) have been shown to induce strong perceptual learning effects at the earliest stages of visual processing, which are retinotopically specific (Karni & Sagi, 1991; Schoups, Vogels, & Orban, 1995; Schwartz, Maquet, & Frith, 2002). We independently manipulated consciousness and attention by exposing subjects to these stimuli rendered unreportable in two different ways: either by perceptual blindness (masking) or by attentional blindness (inattention paradigm) (Kanai et al., 2010). Respectively, stimuli were either masked yet attended or unmasked yet not attended. This yields stimuli unreportable in both conditions, yet for orthogonal and different reasons. This
enabled us to study visual field location specific learning in subjects that could not report about the stimuli that induced the learning (either because of the absence of attention or consciousness) and hence answer the question whether perceptual learning requires consciousness or attention.

METHODS

Subjects
Fifty-four subjects (all women) with no relevant psychiatric or neurological history participated in the experiment for course credit or financial compensation. We included only female subjects, as it is our personal experience that they are generally more reliable in attending (both) sessions, and more likely to follow instructions closely (especially in the Inattention condition). Fixating on the center is important to render subjects “inattentually blind”). Possibly our results do not generalize to male subjects, although so far the only known sex effects show more robust (task-irrelevant) perceptual learning effects in male compared with female subjects (Leclercq & Seitz, 2012a). Subjects were all right-handed and had normal or corrected-to-normal vision based on the Landolt eye chart. We obtained written informed consent from each subject before experimentation. The experiment was approved by the ethical committee of the Psychology Department of the University of Amsterdam.

Four subjects from the Inattention group were excluded because they reported to have seen the figure presented on Day 1 (as assessed by the 10AFC task, see below). Two subjects (one from each group) were excluded because of apparatus failure. After exclusion, 24 subjects were tested in both groups (Inattention group $M = 21.31$ years, $SD = 2.04$ years, Masked group $M = 21.62$ years, $SD = 2.49$ years).

Task Design

Stimuli and Setup
Stimuli were textured patterns consisting of a full field of patches of homogeneously oriented line elements, oriented at either 30°, 60°, 120°, or 150°. Target textures (presented at the onset of each white letter, see explanation of 2-back task below) formed either a figure or a no-figure. Texture defined figures contained a background with line elements oriented at either 30°, 60°, 120°, or 150°, and six squares (individual width 3.6°) consisting of line elements oriented orthogonal to the background, divided over two diagonally opposite quadrants (three squares in each quadrant; see Figure 3.2). Another six squares were present in the two remaining quadrants; however, their line elements had the same orientation as the background texture, so that they were not visible (however, because of a different jittering of the line elements, very vague and generally not perceived “boundaries” between square and background are discernible on close scrutiny; see “no-figure” panel in Figure 3.2). The no-figure texture also contained 12 squares (three in each quadrant); however, all squares had the same orientation as the background texture and were hence invisible (except that again, vague “boundaries” were created between square and background).
background, by the different jittering of the line elements). The reason for doing this is to make figure and no-figure stimuli fully comparable from the point of view of low-level features, such as line orientation and length (see also Zipser, Lamme, & Schiller, 1996). Namely, when figure trials of all orientations are collapsed, overall (and on each location) the number of line elements is equal for each orientation compared with when all no-figure trials are collapsed. The same is true for the number of boundaries; only the orthogonality of the line elements in the figure squares differs from the no-figures. This enabled us to calculate figure-no-figure difference waves in the EEG and hence isolate neural activity specifically related to figure-ground segregation (and not just boundary detection; for a similar procedure, see Caputo & Casco, 1999; Fahrenfort et al., 2007; Lamme, Van Dijk, & Spekreijse, 1992; Scholte et al., 2006).

Stimuli were created using Matlab (The MathWorks, Natick, MA, USA). Stimuli were presented using Presentation (Neurobehavioral Systems, Albany, CA, USA) and displayed on a 19-inch Iiyama VisionMaster Pro 450 with display settings of 1024×768, 32 bits and 60 Hz. Subjects were seated in a comfortable chair at fixed distance to the screen. An in-house manufactured chin rest positioned the subjects at equal height and fixed eye distance (75 cm) to the screen on both days.

**General Procedure**

The experiment consisted of two separate sessions on consecutive days; a learning phase (on Day 1) and a testing phase (24 hr later, on Day 2). Every subject was assigned to either the Inattention or the Masked group on the basis of a predetermined randomization schedule. On Day 1, both groups were exposed to a rapid serial visual presentation (RSVP) letter stream, presented against a background of textured figures. However, for the Inattention group reportability of the to-be-learned figure was manipulated by inattention, and for the Masked group reportability was manipulated by masking the figure. On Day 2, both groups performed the same (RSVP) task, while in the background figures were presented both at trained and novel positions to assess learning neurally. Finally, a staircased detection task with figures at trained and novel positions was administered to measure behavioral learning effects.

**Day 1—Learning Phase**

**Grayscale task**

Within each group (Inattention and Masked), there were two conditions: A and B, referring to the position of the figures presented on Day 1 (see Figure 3.2). In each group, half of the subjects (12 subjects) were assigned to Condition A, and the other half were assigned to Condition B. Visual field biases (an asymmetry in representation and perception of/attention for one side of the visual field) may influence learning effects (Dickinson & Intraub, 2009). Therefore, before being assigned to a condition every subject was tested for visual field bias by performing a computerized version of the Grayscale task.
Does perceptual learning require consciousness or attention?

Figure 3.2. Schematic representation of the task designs. On Day 1 (learning phase), both groups perform a different task. The Inattention group performs an RSVP 2-back letter task, whereas a figure texture (layout A or B) is presented in the background together with every target letter. The Masked group performs a figure detection task, where they have to detect whether or not a masked figure texture is present in the background while the target letter is presented. Easily detectable unmasked luminance defined “catch” figure trials (layouts A and B) have been added to keep subjects motivated and to check whether the task instructions were being followed. On Day 2 (testing phase), both groups perform the same two tasks to measure learning effects neurally (vowel task) and behaviorally (staircase task). In the vowel task, subjects have to indicate whether the target letter is a vowel or not while both the trained and novel figure (layouts A and B) are...
presented in the background. In the staircase task, subjects have to detect whether or not the masked target is a figure. According to their performance, the SOA between target and mask varies, in a double staircase manner for both figures independently. Two versions of the staircase task are presented, one without and one with performance feedback.

(Nicholls, Bradshaw, & Mattingley, 1999; both the original version and a 90° rotated one, so biases could be calculated for both left vs. right and top vs. bottom of the visual field). On the basis of the outcome of this test, subjects were assigned to a condition such that visual field bias differences between conditions were minimized.

Inattention group (2-back task)
We used an inattention paradigm based on the paradigm used by Scholte et al. (Scholte et al., 2006). It consisted of a primary stream of stimuli of which the subjects were informed and a secondary stream of stimuli of which the subjects were not informed.

The primary stream consisted of a foveally presented RSVP of letters at a speed of 217 msec per letter (A, G, P, S, R, O, L, B, D, and T were used), which were presented on a red fixation point with a gray center of 1.1° in diameter. These letters were usually black, but once every six to eight letters, a white letter was presented. Subjects were instructed to indicate whether every white letter was the same as or different from the white letter presented two instances earlier (2-back task), by pressing one of two buttons of a response box. At the onset of every white letter, a short beep sounded to minimize the number of missed responses.

The secondary stream of stimuli consisted of textured patterns that were presented with the onset of each new letter for 50 msec (the remaining 167 msec an isoluminant gray dotted pattern was presented). Together with every black letter, a homogeneous pattern was presented. However, at the onset of each white letter a textured pattern was presented that either formed a figure (containing six squares, presented in 47% of the cases) or a no-figure (containing no squares, presented in 47% of the cases; for a detailed description of these textured patterns, see “Stimuli and Setup” above). The remaining 6% of the cases so-called “filler trials” were presented, with homogeneous textured patterns, which acted as a dummy to match the “catch trials” in the Masked group paradigm (see below for further explanation of “catch trials”).

For every trial, the background orientation of each textured pattern was chosen randomly from the four orientations (30°, 60°, 120°, or 150°; see “Stimuli and Setup” above), within the following constraints: subsequent textures always contained line elements with a different orientation to ensure that a new image was presented with the onset of each subsequent letter (black or white). The four background orientations were equally divided over figures and no-figures, and each orientation was equally divided over each white letter (A, G, P, S, R, O, L, B, D, and T). In total, 1,440 trials were presented (680 per figure/no-figure, 80 “filler trials”) in six blocks, with a total duration of approximately 45 min (including breaks). The task was preceded by a short practice version (12 trials,
with only no-figure backgrounds) to make subjects acquainted with the task. Buttons were counterbalanced across subjects.

Subjects were instructed to perform as good as possible on the primary letter stream task. They were not told about the potential presence of figure textures in the background, so as to induce inattentional blindness for these stimuli. We evaluated the presence or absence of inattentional blindness for the textured figures at the end of the session, by means of a 10 alternative forced-choice (10AFC) task in which subjects had to select the figure presented during the 2-back task from a set of nine other figure textures (see Figure 3.3). They were told that one of the figure textures had been presented regularly during the previous task, at the onset of the white letter, and were instructed to select it. The actual figures (both the stimuli from conditions A and B) and eight others consisting of circles and rectangles with similarly oriented line elements and textures of 0° and 90° were presented in random order (numbered 1 to 10). If subjects failed to select the correct figure, they were considered to be unable to report about the figure, that is, to have suffered from inattentional blindness during the exposure to the texture figures.

Figure 3.3. This set of figure textures was presented (in random order, numbered 1–10) during the 10AFC task, in which subjects from the Inattention group had to select the figure presented during the 2-back task from a set of nine other figure textures. If subjects failed to select the correct figure, they were considered to have suffered from inattentional blindness for the figure presented in the background on Day 1. Figures A and B are depicted in the top two pictures.
**Masked group (Detection task)**

Subjects in the Masked group were presented with a paradigm similar to the Inattention group; however, there were some crucial differences. Our aim was to manipulate reportability of the figure by masking instead of inattention. Therefore, all background textures were masked (the mask consisting of two superimposed homogeneous textures with orthogonally oriented line elements; see “mask” panel in Figure 3.2). Masking was achieved by having each figure/no-figure stimulus (50 msec), presented together with the white letters, being followed by an isoluminant gray screen (for 100 msec), followed by a mask with a duration of 67 msec. Instead of the 2-back letter task, subjects directed their attention toward the secondary stream of stimuli to perform a figure detection task. With every white letter (accompanied by a short beep), they had to indicate whether or not a figure was present in the background by pressing one of two buttons of a response box. Because of the mask, this task was practically undoable (but that was indeed our goal, unreportability of the figures by masking, even with attention). To keep our subjects motivated and to check whether the task instructions were being followed, “catch trials” were presented in 6% of the trials, which were much easier to detect. These “catch trials” contained unmasked and fairly visible luminance defined catch figures (both layouts A and B to prevent any learning effects of location to occur and to ensure fixation at the center of the screen) and catch no-figures. Note that in the Inattention group, to keep number of trials and task length equal, “filler trials” were presented instead of these “catch trials,” containing only catch no-figures (see Figure 3.2) while subjects were performing the 2-back task.

**Day 2—Testing Phase**

**Vowel task**

Subjects of both groups (Inattention and Masked) and assigned to each condition (A and B) performed the same task to measure neural activity evoked by both the trained and the novel figure (its spatial configuration in opposite corners of the trained figure). Subjects performed an RSVP letter task to ensure fixation in the center while being passively exposed to figures and no-figures in the background. The task is similar to the task performed by the Inattention Group on Day 1, except that here for every white letter (A, E, O, U, S, T, G, and B were used), subjects had to indicate whether it was a vowel or a consonant by pressing one of two buttons of a response box. The task is different from the 2-back task (Inattention group, Day 1) to ensure equal task difficulty for both groups. With every white letter (as on Day 1, accompanied by a short beep), either figure A, figure B, or no-figure was presented, which allowed us to make relevant subtractions in the EEG to assess neural learning effects.

In total 768 trials were presented (256 per figure A/figure B/no-figure) in four blocks, with a total duration of approximately 25 min (including breaks). The task was preceded by a short practice version (12 trials, with only no-figure backgrounds) to make subjects acquainted with the task. Buttons were counterbalanced across subjects.
Staircase task

Subjects from both groups (Inattention and Masked) and conditions (A and B) performed the same task to measure behavioral detectability of the trained and the novel figure (its configuration in opposite corners of the trained figure). Every trial a target, either a figure (A or B) or no-figure, was presented (50 msec, accompanied by a short beep), followed by an isoluminant screen (variable duration, starting at 300 msec) and a meta-contrast mask consisting of orthogonally oriented line elements (500 msec, see “staircase mask” panel in Figure 3.2). After presentation of the mask, the red fixation dot turned green to indicate that a response could be made. Subjects then pressed one of two buttons of a response box to indicate whether or not they saw a figure preceding the mask. The next trial started after a response was made, but when no response was given 1500 msec after mask presentation, a warning message appeared “No response! Respond when the fixation dot is green” (but in Dutch), which stayed on screen until a response was made.

The SOA between the target and the mask was initially set at 350 msec for both figures but varied in a double staircased manner (for each figure independently), according to the subject’s responses. For every two correct responses, the SOA was shortened by 17 msec, and for every incorrect response, it was lengthened by 17 msec, thereby varying the duration of the isoluminant screen in between target and mask. The SOA of the no-figures was paired with both figures (in a 50:50 manner), such that with every SOA there was an equal probability for a figure or no-figure as a target. After 60 trials for each figure/no-figure pair, this resulted in two separate SOA curves, one for the trained (figure layout A for subjects exposed to Condition A on Day 1 and figure layout B for subjects exposed to Condition B on Day 1) and one for the novel figure (figure layout A for subjects exposed to Condition B on Day 1 and figure layout B for subjects exposed to Condition A on Day 1).

Subjects performed this task twice in a row. The first time as described above, the second time performance feedback was given at every trial. After each response, a green or red bar appeared (for 500 msec), with the words “correct” or “incorrect,” respectively. In each group, only 18 of the 24 subjects performed these two tasks, as we started out with a more difficult version (target presented for 33 msec instead of 50 msec, no beep, no feedback) for the first six subjects in each group. This version of the task turned out to be incomprehensible (11 of 12 subjects finished with an SOA > 350 msec); therefore, we changed the task to the above described and excluded these 12 subjects from the analysis for these tasks. A further three subjects from the Inattention group were excluded from the analysis of the Staircase detection task without feedback and two subjects from the one with feedback, as they reported in the exit interview to have been fixating on one corner of the screen instead of on the center.

In total 120 trials were presented in one block, with a duration of approximately 6 min. The task was preceded by a short practice version (10 trials, with only no-figure backgrounds) to make subjects acquainted with the task. Buttons were counterbalanced across subjects.
Exit interview
At the end of the second session, subjects filled out an exit interview, with questions about the amount of sleep they had the night before and in between the two sessions, how many units of alcohol they drank, and what strategy they had used for the detection task (Day 2). Three subjects from the Inattention group were excluded from behavioral analysis for the no-feedback Staircase task and two subjects for the feedback Staircase task, as they reported to have been fixating on one corner of the screen instead of the center during these tasks.

Behavioral data analysis
Behavioral data were analyzed using SPSS 17.0 (IBM, Armonk, NY). For Day 1, bias scores on the Grayscale task were calculated as in Nicholls et al. (1999), for both left versus right (horizontal) and top versus bottom (vertical), which results in one overall score (horizontal x vertical). Bias score differences between conditions and groups were tested with independent t-tests. Performance on the 2-back task (Inattention group, Day 1) was calculated in percentage correct. For the detection task (Masked group, Day 1), performance (percentage correct and d0) was analyzed separately for the catch stimuli and the masked stimuli. It was tested with paired two-tailed t-tests whether these performances differed significantly from chance.

Equal performance on the Vowel task (Day 2) was tested with an independent t-test between both groups (percentage correct). Behavioral learning effects on Day 2 were established by comparing the results of the Staircase task for the trained versus the novel figure, for both the feedback and no feedback version. For each group, the trials were binned to five bins (12 trials each), and for every fifth bin (the end point of the SOA curves), a paired t-test was performed for the trained versus the novel figure. Learning effects were then compared between both groups, by means of an independent t-test for the fifth bin.

EEG measurements and analysis
EEG data were recorded using a BioSemi ActiveTwo 64-channel active EEG system (BioSemi, Amsterdam, the Netherlands) during the learning phase (Day 1) and the Vowel task (Day 2). Data were sampled at 1024 Hz and referenced to electrodes placed on both ear lobes. Sixty-four scalp electrodes were measured, as well as two electrodes for horizontal and two for vertical eye movements. Data were filtered using a high-pass filter of 0.5 Hz, a low-pass filter of 30 Hz, and a notch filter of 50 Hz before down sampling to 256 Hz. An independent component analysis (Hyvarinen, Karhunen, & Oja, 2001) was used to remove eye blink and eye movement artifacts. Another low-pass filter of 20 Hz was performed. Data were segmented into epochs starting at −217 msec before stimulus onset and ending at 500 msec after stimulus onset. Baseline correction was applied to the interval from −100 to 0 msec preceding stimulus onset. Automatic artifact rejection was applied by removing segments falling outside the −250 to 250 μV range or deviating more than two standard deviations from average. Spherical interpolation was used to create a signal for removed
channels. Data were normalized toward peak signal responses of the no-figure stimuli on a per subject basis.

To investigate the extent of neural processing of the figures on Day 1 (learning phase) in both groups, we averaged ERPs for figure and no-figure trials, and difference waves were obtained by subtracting the average of no-figure trials from the average of figure trials. The signal that remains then is selectively reflecting the processes underlying figure-ground segregation, whereas the signals related to low-level processes such as the detection of line segments and line segment discontinuities (the vague boundaries mentioned above) have been eliminated from the EEG signal (see Caputo & Casco, 1999; Fahrenfort et al., 2007; Lamme et al., 1992; Scholte et al., 2006). ROIs were determined based on the figure-ground difference wave of the Inattention group (because in the Masked group the figure-ground signal was greatly reduced by the mask). Selected electrodes were pooled, and a time window for each ROI was determined based on visual inspection (by two independent observers). This resulted in in the following ROIs: an early peri-occipital ROI (P5, P7, P07, P6, P8, P08; 100–200 msec) and a late peri-occipital ROI (O1, P07, P03, Oz, P0z, O2, P08, P04; 200–350 msec). For these ROIs and time windows, we performed paired two-tailed t-tests on the average (figure vs. no-figure) of each time window and ran an ANOVA for Timing (early, late) × Figure (figure, no-figure). We then tested at which time points the figure trials differed significantly from the no-figure trials for both groups. We conducted a paired two-tailed t-test at each time point for the selected time window (treating the average of each subject at that time point as an observation, thresholding with a false discovery rate at a p < .05 level), to see at which exact time points figure evoked activity differs from no-figure evoked activity.

We then analyzed neural learning effects for Day 2 (testing phase). The same ROIs and time windows were used as for the Day 1 (learning phase) analysis, and ERP averages were calculated for trained figure trials and novel figure trials. Similar to the analysis for Day 1, using the same ROIs and time windows, we tested whether the average value of the trained figure (minus no-figure) differs from the novel figure (minus no-figure) for each time window using paired two-tailed t-tests and an ANOVA for Timing (early, late) × Figure (trained figure, novel figure). We then tested at which time points within the selected time window the trained figure differed significantly from the novel figure, within both the Inattention group and the Masked group (again, thresholding with a false discovery rate at a p < .05 level). Preprocessing was performed using Brain Vision Analyzer (Brain Products GmbH, Munich, Germany). For statistical analysis and visualization of the time courses, Matlab (The MathWorks, Natick, MA, USA) was used.
RESULTS

Learning Phase (Day 1)

Behavioral results: manipulation checks
To minimize visual field bias differences between conditions, for every subject a bias score was measured by performing a computerized version of the Grayscales task (Nicholls et al., 1999). On the basis of the outcome of this test, subjects were assigned to a condition such that visual field bias differences between conditions were minimized. The bias scores indeed did not differ significantly between conditions (Inattention group: \( t(1, 11) = -0.322, p = .75 \); Masked group: \( t(1, 11) = -0.822, p = .43 \) and groups \( t(1, 46) = 0.151, p = .88 \).

Performance measures on the tasks for Day 1 are presented in Figure 3.4. For the Inattention group, attention was successfully directed towards the primary stream of letter stimuli, as performance on the 2-back task was significantly better than chance (90.7%, \( t(1, 23) = 38.95, p < .00001 \)). Reporatability of the figure presented in the secondary stream of stimuli was measured with a 10AFC task. Four of 28 subjects (14%) from the Inattention group were excluded from further analysis because they chose the correct figure presented on Day 1 (see also “Subjects” section above). For the Masked group, detection of the catch trials was significantly better than chance (85.1%, \( t(1, 23) = 14.77, p < .00001 \)), indicating that the task instructions were being followed correctly. The other trials were successfully masked, as detection of the masked figures did not differ from chance significantly (50.5%, \( t(1, 23) = 1.60, p = .12 \); see Figure 3.4). However, the d’ value for detection of the masked figures, although being very close to zero (0.11), did differ significantly from chance (i.e., a d’ value of 0) \( (t(1, 23) = 2.48, p = .02) \), suggesting that the masked figures may have been just barely detectable for some subjects on some trials.

**Figure 3.4.** Behavioral performance on Day 1. For the Masked group, percentage correct does not differ significantly from chance for detection of masked figure textures (50.5% correct, \( t(1, 23) = 1.60, p = .12 \)), whereas unmasked “catch” figures are detected above chance level (85.1%, \( t(1, 23) = 14.77, p < .00001 \)), which indicates that subjects were performing the task correctly but just were not able to detect the masked figure textures. The Inattention group performs above chance level at the 2-back letter task (90.7%, \( t(1, 23) = 38.95, p < .00001 \)), indicating that top–down attention was directed at the letter task instead of the background figures.

EEG results: more late figure related processing in the Inattention group

As expected, because of the different task manipulations, on Day 1 figure textures were processed differently in both groups (see Figure 3.5). In the Inattention group, early (100–200 msec) and late (200–350 msec) figure-ground specific activity can be observed when the no-figure evoked signal is subtracted.
Does perceptual learning require consciousness or attention?

from the figure-evoked signal. On the basis of earlier work, we interpret the early figure-ground signal as feedforward sweep (FFS) and the late signal as recurrent processing (RP) activity (Fahrenfort, Scholte, & Lamme, 2008; Lamme & Roelfsema, 2000; Scholte, Jolij, Fahrenfort, & Lamme, 2008). For both time windows, figure evoked activity differs significantly from no-figure evoked activity (early ROI: $t_{(1, 23)}=-2.700, p = .013$; late ROI: $t_{(1, 23)}=-3.077, p = .005$). The ANOVA showed a main effect of Figure (vs. no-figure, $F_{(1, 23)} = 12.383, p = .002$, and for Timing, $F_{(1, 23)} = 30.779, p = .00001$, no interaction of Timing x Figure), indicating that figure evoked signal deviates from no-figure evoked signal at both time windows. When the difference between figure and no-figure activity is tested per time point, the following time periods differ significantly: 117–199 msec (early ROI) and 223–238 msec, 277–348 msec (late ROI). Thus, figure-ground segregation processes proceeded normally during inattention (see also Scholte et al., 2006).

For the Masked group, late (RP) figure-related activity is entirely blocked by the mask (late ROI: $t_{(1, 23)}=0.226, p = .82$; none of the time points differ significantly), whereas early (FFS) processing is still present, although marginally significant (early ROI: $t_{(1, 23)}=-1.935, p = .065$ (but note that this is tested using a two-sided paired t-test); time period 148–156 msec differs

![Figure 3.5.](image)

**Figure 3.5.** Neural processing of the figure and no-figure texture on Day 1 for both groups. Activation maps depict figure minus no-figure activity for the early (100–200 msec) and late (200–350 msec) time window. The difference wave of figure minus no-figure evoked signal shows early (FFS) and late (RP) figure-related activity in the Inattention group (early ROI: $t_{(1, 23)}=-2.700, p = .013$, late ROI: $t_{(1, 23)}=-3.077, p = .005$). When the difference between trained figure and novel figure activity is tested per time point, significant differences were found for the following time periods: 117–199 msec (early ROI) and 223–238 and 277–348 msec (late ROI). For the Masked group, late figure-related activity is entirely blocked by the mask ($t_{(1, 23)}=0.226, p = .82$), whereas early (FFS) processing is still present, although marginally significant (148–156 msec, early ROI: $t_{(1, 23)}=-1.935, p = .065$ (but note that this is tested using a two-sided paired t-test)). Please note that prestimulus activity is large for the nonsubtracted ERP signal (figure and no-figure) because of the RSVP design of our experiment, where the target stimulus is presented amidst a series of distractors in a rapid stream.
significantly). The ANOVA showed an interaction of Timing × Figure at trend level ($F_{(1, 23)} = 3.178, p = .088$; no main effects of Figure, $F_{(1, 23)} = 0.480, p = .50$, or Timing, $F_{(1, 23)} = 0.490, p = .49$), in line with our finding that figure evoked signal deviates from no-figure evoked signal at the early, but not the late time window. The masks successfully blocked RP of the figures, which is also reflected in the performance on the detection task (see above). This is in line with previous results (Fahrenfort et al., 2007).

**Testing Phase (Day 2)**

*Behavioral results: manipulation checks*

During the testing phase, both groups performed the same task (Vowel task), on which performance was equal for both groups (Inattention group: 95.4%, Masked group: 95.6%; $t_{(1, 46)} = 0.174, p = .863$).

*EEG results: early and late learning effects only for Inattention group*

Neural learning effects were found only in the Inattention group (see Figure 3.6). For both early and late ROIs, Trained Figure minus no-figure evoked activity

![Figure 3.6](image-url)

*Figure 3.6.* Neural processing of the trained figure and novel figure texture on Day 2 for both groups. Activation maps depict trained figure minus novel figure activity for the early (100–200 msec) and late (200–350 msec) time window. When comparing trained figure and novel figure evoked activity on Day 2, learning effects were found only in the Inattention group. For both early and late ROIs, trained figure evoked activity differs significantly from novel figure evoked activity (early ROI: $t_{(1, 23)} = -2.213, p = .037$, late ROI: $t_{(1, 23)} = -3.137, p = .005$). When the difference between trained figure and novel figure activity is tested per time point, significant differences were found for the following time periods: 156–172 msec (early ROI) and 273–328 msec (late ROI). No learning seems to have taken place in the Masked group, as no differences were found between trained figure and novel figure evoked activity, for both ROIs (early ROI: $t_{(1, 23)} = -0.566, p = .58$, late ROI: $t_{(1, 23)} = -0.593, p = .56$). Also none of the time points differ significantly between trained figure and novel figure ERPs. Please note that prestimulus activity is large for the nonsubtracted ERP signal (trained figure and novel figure) because of the RSVP design of our experiment, where the target stimulus is presented amidst a series of distractors in a rapid stream.
differs significantly from novel figure minus no-figure evoked activity (early ROI: \( t_{(1,23)} = -2.213, p = .037 \); late ROI: \( t_{(1,23)} = -3.137, p = .005 \)). The ANOVA showed a main effect of Trained Figure (vs. Novel Figure, \( F_{(1, 23)} = 10.415, p = .004 \), and no main effect of Timing, \( F_{(1, 23)} = 0.693, p = .41 \), nor an interaction of Timing \( \times \) Figure), indicating that trained figure evoked signal deviates from novel figure evoked signal at both time windows. When the difference between trained figure and novel figure activity is tested per time point, the following time periods differ significantly: 156–172 msec (early ROI) and 273–328 msec (late ROI).

For the Masked group, no differences were found between trained figure minus no-figure and novel figure minus no-figure evoked activity, for both ROIs (early ROI: \( t_{(1,23)} = -0.566, p = .58 \); late ROI: \( t_{(1,23)} = -0.593, p = .56 \)). The ANOVA showed no main effect of Trained Figure (vs. Novel Figure, \( F_{(1, 23)} = 0.469, p = .50 \), neither for Timing, \( F_{(1, 23)} = 0.111, p = .74 \), nor an interaction of Timing \( \times \) Figure), indicating that Trained Figure evoked signal does not deviate from novel figure evoked signal at both time windows. Also none of the time points differ significantly between trained figure and novel figure ERPs. So no neural learning seems to have taken place in this group.

**Behavioral results: learning effect only for Inattention group with performance feedback**

In Figure 3.7, behavioral performance on the Staircase task is depicted. In each group, only 18 of the 24 subjects performed these two tasks, as we started out with a more difficult version for the first six subjects in each group, which turned out to be incomprehensible. Therefore, we changed the task and excluded these 12 subjects from the analysis for these tasks (see “Staircase task” in “Methods” section above). From these 18 subjects in the Inattention group, three subjects were excluded from behavioral analysis for the no-feedback Staircase task and two subjects for the feedback Staircase task, as they reported to have been fixating on one corner of the screen instead of the center during these tasks. No further exclusions were required in the Masked group (see “Exit Interview” in “Methods” section above). There was a learning effect in the Inattention group, but not for the Masked group as measured by the Staircase task with feedback. In the Inattention group SOA’s for the trained figure were significantly shorter than for the novel figure (\( t_{(1, 15)} = -2.147, p = .049 \)). For the Masked group, there was a trend going in the opposite direction: shorter SOA’s for the novel figure as opposed to the trained one (\( t_{(1, 17)} = 1.891, p = .076 \)). Overall, SOA’s were lower in the Masked group compared with the Inattention group. Although this difference was not significant (\( t_{(1, 32)} = 1.082, p = .29 \)), it might be because of the fact that subjects in the Masked group have more prior experience with figure detection tasks, as they had to detect the masked figure on Day 1.

Remarkably, these effects were only present for the second version of the Staircase task, during which performance feedback was given at every trial. For the first (no feedback) version, no learning effects were found in both groups (Inattention group: \( t_{(1, 14)} = 0.233, p = .819 \), Masked group: \( t_{(1, 17)} = 0.010, p = .992 \).
Figure 3.7. Behavioral performance on the staircase task on Day 2. The version without performance feedback (on the left) does not reveal any learning effects in both groups (Inattention group: $t(1, 14) = 0.233, p = .819$, Masked group: $t(1, 17) = 0.010, p = .992$). Once performance feedback is provided on a trial-by-trial basis, a learning effect becomes apparent only for the Inattention group (Inattention group: $t(1, 15) = -2.147, p = .049$; Masked group: $t(1, 17) = 1.891, p = .076$). Note that for the Masked group, this is a trend going in the opposite direction: shorter SOA’s for the novel figure as opposed to the trained one, whereas in the Inattention group SOA’s for the trained figure are significantly shorter than for the novel figure.

**DISCUSSION**

*Inattentinal learning*

We found perceptual learning during inattention, but not for masked stimuli to which attention is directed. Neurally, learning effects were found for both early and late ROIs in the Inattention group, whereas the Masked group did not show any neural learning effects. This learning effect for the Inattention group was apparent behaviorally, measured with a staircased detection task, but only when performance feedback was provided.

These results suggest that different kinds of unreportability, perceptual or attentional blindness, can lead to different neural and behavioral outcomes. Attentional blindness (inattention) to stimuli does not fully preclude that the unreported stimuli induce learning, whereas perceptual blindness (masking) does. The fact that we find a learning effect for the Inattention group, but not for the Masked group, indicates that attention is not required for perceptual learning, whereas consciousness is. This further strengthens the idea that consciousness is separate from attention.

Did we successfully divert attention away from the background figures in the Inattention group and direct attention toward the masked figures in the Masked group? Although not widely used, the Inattentional Blindness paradigm—originally devised by Mack and Rock (1998) and popularized by the famous Gorilla-suit movie (Simons & Chabris, 1999)—is in fact very effective when it comes to shutting down both bottom–up and top–down attention. Key to the Inattentional Blindness paradigm is that subjects do not know in advance about the possible presentation of figures. This prevents any (hidden) top–down motivation for attending these figures. Also, bottom–up attention is precluded, as
is evidenced by the absence of knowledge of the figures after the task: typically, when bottom-up attention is grabbed by the figure, subjects notice the object and are readily able to report it (as happened in a subset of our subjects that were hence excluded from further analysis). This is also shown by the classical Mack and Rock experiments, showing that bottom-up attention grabbers such as faces typically do “break through” and people readily remember having seen these (Mack & Rock, 1998). This strongly argues that subjects who do not report the figures indeed devoted neither top-down nor bottom-up attention to these figures. We therefore think that our Inattention condition sufficiently exhausted all attentional resources. As for the Masked condition, the question is whether attending to the “catch” stimuli implies that the masked figures were attended as well. We can only show successful detection of the “catch” stimuli that are identical to the masked stimuli in terms of spatial layout. The situation is similar to paradigms in which masked and hence unconscious primes are presented together with visible stimuli (Dehaene et al., 2001). Such invisible primes at the same location as visible stimuli are known to affect subsequent processing, whereas unconscious primes that are not attended do not (Naccache et al., 2002). This shows that attention to visible stimuli at a particular location automatically transfers to other—invisible—stimuli at that same location in some tasks, suggesting that our masked figures were indeed attended.

**Task-irrelevant versus task-relevant learning**

Several studies are in line with the idea that attention is not necessary for learning. A relevant paradigm is so-called “task-irrelevant learning” (Seitz & Watanabe, 2003; Seitz & Watanabe, 2009; Watanabe et al., 2001; for a review, see Choi & Watanabe, 2009). It is claimed that internal rewards gate this type of learning, as learning only occurs when the task-irrelevant stimulus is temporally paired with a task-relevant stimulus, which triggers an internal reward (release of neuromodulatory factors that boost sensory signals) resulting in reinforcement learning (Seitz et al., 2009; Seitz & Watanabe, 2003, 2005, 2009; Seitz, Náñez, Holloway, Koyama, & Watanabe, 2005). This explanation could possibly account for our results, as in our study internal rewards might have been present (triggered by the “task-relevant” letters, presented simultaneously with the “task-irrelevant” background figures). On the one hand, it must be noted that internal rewards paired to the target stimuli might have been less present in the Masked group compared with the Inattention group, as task performance for masked figure detection was at chance level (as opposed to a high performance on the 2-back letter task in the Inattention group), which would be in line with our results. On the other hand, Seitz, Náñez, Holloway, Tsushima, and Watanabe (2006) show that correct performance may not always generate internal reinforcement signals, and Seitz and Watanabe (2008) found that learning was greatest in the most difficult task condition where performance was lowest. They suggest that reinforcement learning may not depend upon correct responses, but rather on target-uncertainty (the more difficult target recognition is, the greater the reinforcement signal upon successful recognition). In our case, there would be more target-uncertainty in the Masked group than in the Inattention group.
(although there are fewer instances of successful recognition), and this should then have resulted in stronger learning for masked than for unattended stimuli—that is, the opposite of our results. Also, learning has been shown to occur just because of the temporal pairing of a target and stimulus, such that task-driven and stimulus-driven signals coincide, irrespective of the task-relevancy of the stimulus (Seitz & Watanabe, 2005, 2008). This temporal pairing is equal for both groups in our experiment. Either way, to figure out the exact relation of our results to other established learning paradigms would require more study.

Task requirements for the Inattention and Masked group during the learning phase were asymmetrical in terms of task-relevancy, which is potentially an essential difference. For the Inattention group the to-be-learned stimulus was task-irrelevant, whereas for the Masked group it was task-relevant. The fact that the Staircase task with performance feedback revealed a behavioral trend in the opposite direction for the Masked group (lower SOA’s for the novel figure compared with the trained figure) might indicate that in this group inhibition of learning has taken place. This might have been caused by the fact that in the Masked group attention was directed toward the figure stimuli during the learning phase. However, attention is thought to suppress learning only when stimuli are distracting from the main task (or task-irrelevant; Choi & Watanabe, 2009; Roelfsema, van Ooyen, & Watanabe, 2010; Tsushima & Watanabe, 2009; Tsushima, Seitz, & Watanabe, 2008). In our experiment, the masked figures are task-relevant; therefore, attention should, in general, enhance learning of these figures (Ahissar & Hochstein, 1993; Gilbert et al., 2001; Jessup & O’Doherty, 2009; Shiu & Pashler, 1992). However, attention may not be a sufficient requirement for learning to occur (Ahissar & Hochstein, 1997) and there are multiple examples where explicitly attending to a task-relevant stimulus leads to decreased learning compared with implicit exposure. This seems to be the case in particular for complex stimuli that cannot be described with simple rules, such as in a recognition task with kaleidoscopic visual stimuli (Voss, Baym, & Paller, 2008; Voss & Paller, 2009), rule discovery in an artificial grammar task (A S Reber, 1976), and complex auditory stimuli (Seitz et al., 2010; Vlahou, Protopapas, & Seitz, 2012; Wade & Holt, 2005; see Reber, 1989, for a review).

Neural changes associated with task-irrelevant learning are largely unclear (Sasaki, Nanez, & Watanabe, 2009), and one could argue that this type of learning (and thus the learning reported in the current study) may not be wholly perceptual, but associated more with decision-making regions instead. In general, some have suggested that visual training does not improve how sensory information is represented in the brain, but rather how this sensory representation is interpreted by higher areas that guide behavior (Dosher & Lu, 1998; Law & Gold, 2008). However, as the task-irrelevant stimulus is actually dissociated from the decision process concerning the task, we believe that task-irrelevant perceptual learning really has a low-level basis (Seitz & Watanabe, 2009). Indeed, here we report neural learning effects recorded by (peri-)occipital electrodes. Further study (using a technique with higher spatial
resolves) would be required to investigate the possibility that neural changes took place in (connections with) higher brain areas.

**Location specificity of learning**

Location-specific learning is often considered a hallmark of perceptual as opposed to other types of learning (Karni & Sagi, 1991; Mednick et al., 2002; Ofen et al., 2007; Pourtois et al., 2008; Sagı & Tanne, 1994; Schouppe et al., 1995; Schwartz et al., 2002; Sowden, Rose, & Davies, 2002). Recently, however, location specificity of perceptual learning has become subject of debate (Harris, Gliksberg, & Sagi, 2012; Le Dantec & Seitz, 2012; Zhang, Klein, Levi, & Yu, 2011). If perceptual learning in our experiment would be location nonspecific, learning of one figure would transfer to the orthogonal (novel) figure, which would make the comparison of trained figure versus novel figure processing meaningless. Perhaps there was location nonspecific learning, possibly even in the masked condition, that we did not pick up. On the one hand, Harris et al. (2012) state that location-specific learning is a consequence of low-level sensory adaptation. They show that, as soon as trials are interleaved with dummy trials suppressing local adaptation, transfer of learning to other locations takes place. This happens for dummy trials containing textures oriented 45° away from the target textures, whereas dummy trials with an orientation 90° orthogonal to the targets have no such effect. In our study, presentation of textured target stimuli was interrupted during learning by “catch” and “filler” trials. These temporal interruptions caused by “catch” and “filler” trials were equal for both groups, and to induce learning at the level of figure-ground segregation that generalizes across orientations (instead of merely inducing learning at the level of orientation discrimination), we made sure that all stimuli (targets as well as “catch” and “filler” textures) consisted of line elements with the same set of different orientations (30°, 60°, 120°, and 150°). However, luminance properties of these interruptions differed (luminance defined “catch” figures (both configurations A and B) in the Masked group, no-figure “filler” trials in the Inattention group). This may have caused location-invariant learning in the Masked group, which would be an alternative explanation for our finding that no learning effects were found in this group. On the other hand, according to Zhang et al. (2011), only location nonspecific additional tasks (so called “double training”) can cause transfer to untrained locations. As our additional task (detection of “catch” trials) is location specific, according to Zhang et al. (2011) no transfer could have taken place. Transfer induced by double training has further been challenged by Le Dantec and Seitz (2012), who found that transfer during double training is incomplete. It is clear that transfer is controlled by many different factors and that further research is needed to know whether transfer could have occurred in our experimental setup.

**The role of recurrent processing in learning**

No learning occurred when the task-irrelevant figures were masked yet still attended, whereas learning did occur for visible ("conscious") yet unattended stimuli. As masking blocks RP (late, figure-ground segregation related activity,
see Bridgeman, 1980; Di Lollo et al., 2000; Fahrenfort et al., 2007; Lamme, Zipser, & Spekreijse, 2002; Ro, Breitmeyer, Burton, Singhal, & Lane, 2003), while inattention does not, RP seems to be a requirement for learning.

What are the differences between feedforward (early activity, propagating from lower to higher areas) and RP that could explain the differential involvement in perceptual learning? A potentially important difference is that feedforward propagation is more strongly supported by AMPA receptors, whereas the prolonged recurrent activation of cells will more strongly trigger the activation of N-methyl-D-aspartate (NMDA)-type receptors, following Hebbian rules, and thus mediate synaptic plasticity. This has been shown in monkeys (Self et al., 2012), which confirmed predictions from modeling studies (Dehaene, Sergent, & Changeux, 2003; Lumer, Edelman, & Tononi, 1997). The efficacy of NMDA receptors is enhanced if there is a concurrent bottom–up depolarization by sensory inputs, which gives top–down projections a mainly modulatory role (Dehaene et al., 2003). Thus, recurrent connections might gate perceptual learning by activating NMDA receptors (Berardi, Pizzorusso, Ratto, & Maffei, 2003; Dinse, Ragert, Pleger, Schwenkreis, & Tegenthoff, 2003; Fino, Deniau, & Venance, 2009; Mori & Mishina, 1995; Roelfsema et al., 2010; Sheng & Kim, 2002).

Several other studies suggest a role for recurrent activity in synaptic plasticity as well: (Perceptual) learning involves establishing a task-specific activation of an appropriate subset of neurons by feedback connections (Crist, Li, & Gilbert, 2001; Gilbert et al., 2001; Gilbert & Sigman, 2007). Moreover, recurrent connections from higher to lower areas seem to underlie persistent use-dependent functional plasticity, whereas feedforward connections are thought to lose much of their plasticity as development proceeds (Gilbert, 1996; Singer, 1995). Also, evidence from lesion studies in rats indicates that feedback from other cortical regions is necessary for the consolidation of learning (Cauller, 1995; Diamond, Petersen, & Harris, 1999). Schäfer, Vasilaki, and Senn (2007) propose a model, consistent with biological data, showing that perceptual learning can modulate top–down input to V1 while feedforward (and lateral) pathways remain unaffected.

Our claim that RP is needed for perceptual learning may be inconsistent with findings by Seitz et al. (2009), who show learning of a stimulus rendered unreportable by continuous flash suppression. Although the level of (recurrent) processing for the suppressed stimulus during continuous flash suppression is still unclear (Jiang & He, 2006; Yang & Blake, 2012), this would imply that for these low-level orientation gratings learning could take place at the level of feedforward processing. As orientation selectivity is thought to arise from feedforward processing (Hubel & Wiesel, 1968), this actually makes sense. In contrast, figure-ground segregation requires RP (Fahrenfort et al., 2008; Lamme & Roelfsema, 2000; Scholte et al., 2008), which might explain the difference in learning requirements for orientation discrimination (Seitz’ stimuli) versus figure detection (our stimuli).

Instead of explaining the lack of learning effects in the Masked group in terms of (disruption of) RP, these results could also be explained in terms of erasure of iconic memory in this condition. Iconic memory relies on the
persistence of neural signals in the visual system, which in turn depends on recurrent signals (Sligte, Scholte, & Lamme, 2008; Sligte et al., 2010). It is known that masking erases RP (Boehler, Schoenfeld, Heinze, & Hopf, 2008; Fahrenfort et al., 2007; Lamme et al., 2002) and, therefore, erases iconic memory as well.

**Effects of performance feedback**

A surprising outcome of this experiment was the finding that behavioral learning effects were only present when performance feedback was given on the Staircase task (although a neural learning effect was found before any behavioral tasks were conducted). This suggests that some sort of latent learning occurred, which only became apparent after feedback was provided. This finding stresses that the absence of performance enhancement after a learning task does not necessarily mean that the stimulus has not been learned, it just means that learning is not expressed in that particular task (see, e.g., Frensch, Lin, & Buchner, 1998; Frensch, Wenke, & Rünger, 1999). As we did not include a condition in which subjects did not receive feedback on the second run of the Staircase task, nor a condition in which subjects were not exposed to any figure stimuli on Day 1, it must be noted that we do not know for sure whether the observed behavioral effects are a direct result of the performance feedback. A more conservative interpretation would be that performance feedback led to a better understanding of the task itself, which enabled us to measure behavioral output more reliably, as subjects were now executing the task correctly. This explanation fits with the “Eureka effect,” where subjects suddenly learn to perform a difficult detection task after being presented with a single long exposure stimulus (Ahissar & Hochstein, 1997, 2004) or an easy trial (Liu, Lu, & Dosher, 2010). Perhaps the performance feedback acted as top-down guidance of perception, facilitating neuronal pattern recognition.

Since our behavioral learning effect seems to be latent, it would be interesting to remeasure neural responses after the behavioral learning effect was revealed by the Staircase task with performance feedback, to see whether behavioral manifestation of learning had any effect on its neural manifestation. We will do so in our next study, and we expect to find that the behavioral expression of latent learning effects will enhance neural learning effects.

**Conclusions**

In summary, our results are in line with studies that show location specific perceptual learning effects for figure-ground stimuli (Casco et al., 2004; Censor et al., 2006; Censor & Sagi, 2009; Karni & Sagi, 1991; Ofen et al., 2007; Pourtois et al., 2008; Sagi & Tanne, 1994; Schwartz et al., 2002) and studies that show “task-irrelevant” perceptual learning (Seitz et al., 2009; Seitz & Watanabe, 2003; Watanabe et al., 2001; for a review, see Choi & Watanabe, 2009). However, as we only found perceptual learning during inattention, but not for masked stimuli to which attention was directed, we also show that different kinds of unreportability (perceptual or attentional blindness) can lead to different neural and behavioral outcomes. These results should however be interpreted with caution, as the two groups differ in terms of task-relevancy of the trained
stimulus. Remarkably, behavioral learning effects (for the Inattention group) were only present when performance feedback was given on the Staircase task. This potentially suggests that some sort of latent learning trace has been formed that is only accessed after feedback is provided. A further study of this phenomenon of “latent learning” seems necessary.