

Conscious Vision Proceeds from Global to Local Content in Goal-Directed Tasks and Spontaneous Vision

 Florence Campana,^{1,2}  Ignacio Rebollo,¹  Anne Urai,^{1,3,4}  Valentin Wyart,¹ and  Catherine Tallon-Baudry¹

¹Laboratoire de Neurosciences Cognitives, INSERM U960, Ecole Normale Supérieure, Paris 75005, France, ²Department of Neuroscience, Georgetown University Medical Center Research, Washington, DC 20007, ³Department of Neurophysiology and Pathophysiology, University Medical Center Hamburg-Eppendorf, Hamburg 20246, Germany, and ⁴Department of Psychology, University of Amsterdam, Amsterdam 1018 WS, The Netherlands

The reverse hierarchy theory (Hochstein and Ahissar, 2002) makes strong, but so far untested, predictions on conscious vision. In this theory, local details encoded in lower-order visual areas are unconsciously processed before being automatically and rapidly combined into global information in higher-order visual areas, where conscious percepts emerge. Contingent on current goals, local details can afterward be consciously retrieved. This model therefore predicts that (1) global information is perceived faster than local details, (2) global information is computed regardless of task demands during early visual processing, and (3) spontaneous vision is dominated by global percepts. We designed novel textured stimuli that are, as opposed to the classic Navon's letters, truly hierarchical (i.e., where global information is solely defined by local information but where local and global orientations can still be manipulated separately). In line with the predictions, observers were systematically faster reporting global than local properties of those stimuli. Second, global information could be decoded from magneto-encephalographic data during early visual processing regardless of task demands. Last, spontaneous subjective reports were dominated by global information and the frequency and speed of spontaneous global perception correlated with the accuracy and speed in the global task. No such correlation was observed for local information. We therefore show that information at different levels of the visual hierarchy is not equally likely to become conscious; rather, conscious percepts emerge preferentially at a global level. We further show that spontaneous reports can be reliable and are tightly linked to objective performance at the global level.

Key words: behavior; consciousness; magnetoencephalography; perception; vision; visual system

Significance Statement

Is information encoded at different levels of the visual system (local details in low-level areas vs global shapes in high-level areas) equally likely to become conscious? We designed new hierarchical stimuli and provide the first empirical evidence based on behavioral and MEG data that global information encoded at high levels of the visual hierarchy dominates perception. This result held both in the presence and in the absence of task demands. The preferential emergence of percepts at high levels can account for two properties of conscious vision, namely, the dominance of global percepts and the feeling of visual richness reported independently of the perception of local details.

Introduction

The dominance of global percepts in vision has long been noted, from early Gestaltists (Koffka, 1935) and the “forest

before trees” (Navon, 1977), to more recent experimental evidence of gist extraction from natural scenes (VanRullen and Thorpe, 2001; Oliva, 2005; Fei-Fei et al., 2007) and perception of multiple objects as an ensemble (Parkes et al., 2001; Chong and Treisman, 2005; Alvarez and Oliva, 2008; Haberman and Whitney, 2011, 2012; Bronfman et al., 2014). In all those examples, visual percepts are dominated by global information, and local details can be difficult to retrieve.

The reverse hierarchy theory (Hochstein and Ahissar, 2002) is an influential functional model of the visual system that could account for the dominance of global information in conscious vision. In this model, during a first stage of unconscious hierarchical processing, local information encoded in low-level visual areas is automatically and rapidly combined in high-level visual areas into global information, such as overall shape, gist, or cat-

Received Sept. 30, 2015; revised Jan. 29, 2016; accepted Feb. 23, 2016.

Author contributions: F.C., V.W., and C.T.B. designed research; F.C., I.R., and A.E.U. performed research; F.C., I.R., and A.E.U. analyzed data; F.C., V.W., and C.T.B. wrote the paper.

This work was supported by The Agence Nationale de la Recherche grant ANR-BLAN-12-BSH2-0002-01, LABEX IEC-Institute for Cognition Studies grant ANR- ANR-10-LABX-0087 IEC, and IDEX PSL (Paris Sciences et Lettres) grant ANR-10-IDEX-0001-02 PSL. We thank Chloé Serryn for testing preliminary versions of these experiments; and Christophe Gitton and Antoine Ducorps for excellent technical help with MEG data acquisition.

The authors declare no competing financial interests.

Correspondence should be addressed to Dr. Florence Campana, Department of Neuroscience, Georgetown University Medical Center Research, 3970 Reservoir Road NW, Washington, DC 20007. E-mail: campana.florence@gmail.com.

DOI:10.1523/JNEUROSCI.3619-15.2016

Copyright © 2016 the authors 0270-6474/16/365200-14\$15.00/0

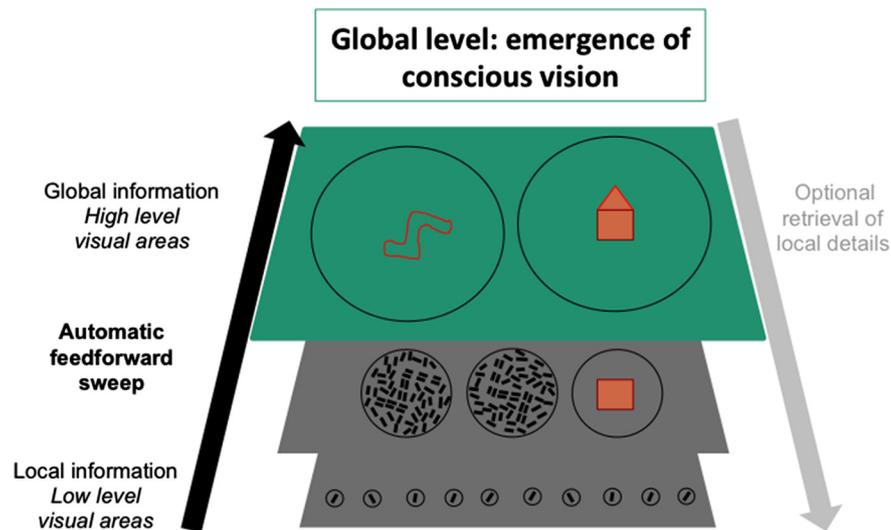


Figure 1. Schematic representation of the reverse hierarchy theory (Hochstein and Ahissar, 2002). Following a first phase of automatic unconscious feedforward sweep, conscious vision emerges in high-level visual areas with global information. Local details encoded in low-level visual areas can be consciously retrieved, but this would occur optionally if required by current goals. Recurrent activities are involved (Lamme and Roelfsema, 2000; Boehler et al., 2008) that are not represented here, for presentation purposes.

egory. This global information becomes conscious, whereas the local information only does so if required by current goals, during a later stage (Fig. 1). This optional second stage of conscious vision proceeds in the reverse direction along the visual hierarchy, from higher to lower visual areas. The model thus predicts that conscious vision has a default global content encoded in high-level visual areas and that this content can be optionally enriched with local information encoded in low-level visual areas (Campana and Tallon-Baudry, 2013).

Although it has been shown that attention operates from top to bottom along visual areas (Mehta et al., 2000; Buffalo et al., 2010), there is no corresponding experimental evidence that conscious vision emerges from top to bottom along these areas. Notably, the famous compound letters from Navon (1977, 1981) initially used in support of the theory cannot be used to probe the interaction between the visual hierarchy and conscious perception because they are not hierarchical: the global letter has a marked contour of luminance; therefore, it is possible to identify it without processing the local letters. Thus, to date, no experimental evidence supports the idea that global dominance in conscious percepts originates in the hierarchical organization of the visual system.

Here, we investigate whether conscious vision emerges preferentially at high levels of the visual hierarchy and is therefore dominated by global information or whether, alternatively, information at any level of the visual hierarchy is equally likely to become conscious (Kouider et al., 2010). We created a new kind of textures, which are, to our knowledge, the first existing truly hierarchical stimuli (Fig. 2A): in those textures, the global shape has no contour and is entirely defined by oriented lines. Therefore, it is necessary to process the orientation of the local lines to represent the orientation of the global shape.

In a first series of experiments, we investigated in goal-directed contexts whether conscious vision emerges at high levels of the visual hierarchy. Subjects were instructed to report either local (line orientation) or global (shape orientation) information, in separate blocks. If the hierarchical organization of the visual system constrains conscious perception to emerge at high

levels, global information should be reported faster than local information, global information is expected to bias performance when reporting local information, and global information should be computed regardless of task demands. We verified those predictions in behavioral and magnetoencephalographic data.

However, such paradigms do not inform us about more ecological situations in which we are not directly engaged in a goal-directed behavior. We therefore assessed whether, in unconstrained contexts without specific task demands, subjects preferentially report global information. Subjects were instructed to report the perceptual content they were experiencing spontaneously. This experiment reveals that spontaneous reports are dominated by global information and that those reports correlate with the reports of the global information in the goal-directed context, whereas no such correlation occurred for local information.

Materials and Methods

Stimuli

Stimulus structure. Stimuli were circular textures (diameter: 7.4° of visual angle) made of oriented lines (line length: 0.18° , distance between line centers: 0.2° , 36 lines on the radius) presented on a uniform gray background. Each texture consisted of lines of random orientation, except within a central rectangular area where lines were approximately aligned along a given orientation, called the local orientation. The central rectangular area in which lines were approximately collinear defined a rectangular shape whose orientation was called the global orientation (Fig. 2A). Thus, the local orientation and the global orientation are defined by the same physical information, namely, the ensemble of lines that are approximately aligned in the center of the stimuli.

Local and global orientations were chosen among 4 angles: -67.5° , -22.5° , 22.5° , 67.5° (with a vertical reference, negative angles on the left), and an angular distance between the local and the global orientation of either 45° or -45° (Fig. 2C). Therefore, subjects could not deduce one orientation from the other. Each of the four angles could be paired with two other angles, resulting in eight possible combinations of orientations.

Each line orientation was drawn from a von Mises distribution, a circular analog of the Gaussian distribution $f(x \setminus \mu, \kappa) = \frac{e^{\kappa \cos(x - \mu)}}{2\pi I_0(\kappa)}$ with I_0 the modified Bessel function of order 0, and μ and κ the two parameters characterizing the von Mises distribution. μ and κ are analogous, respectively, to the mean and the inverse of the variance in a normal distribution. If the coherence κ is large (e.g., Fig. 2A, left, $\kappa = 20$), the distribution is sharp and any line orientation will be close to μ . If κ is equal to zero, line orientation is random.

To build a stimulus, a local orientation (μ) was chosen among the four possible angles. Over a rectangular area made of 4×18 lines oriented at $\pm 45^\circ$ from μ , κ had a fixed value (called the coherence of the lines). The value of κ decreased progressively when moving away from the central rectangular area according to a sigmoidal function. The steepness of the sigmoidal function was such that κ decreased from its maximal value to zero over three lines (Fig. 2B). This defined a rectangular area in the middle of the stimuli (the global shape) made of similarly oriented lines and surrounded by lines randomly oriented, with a buffer zone of intermediate coherence values. During the experiment, low values of

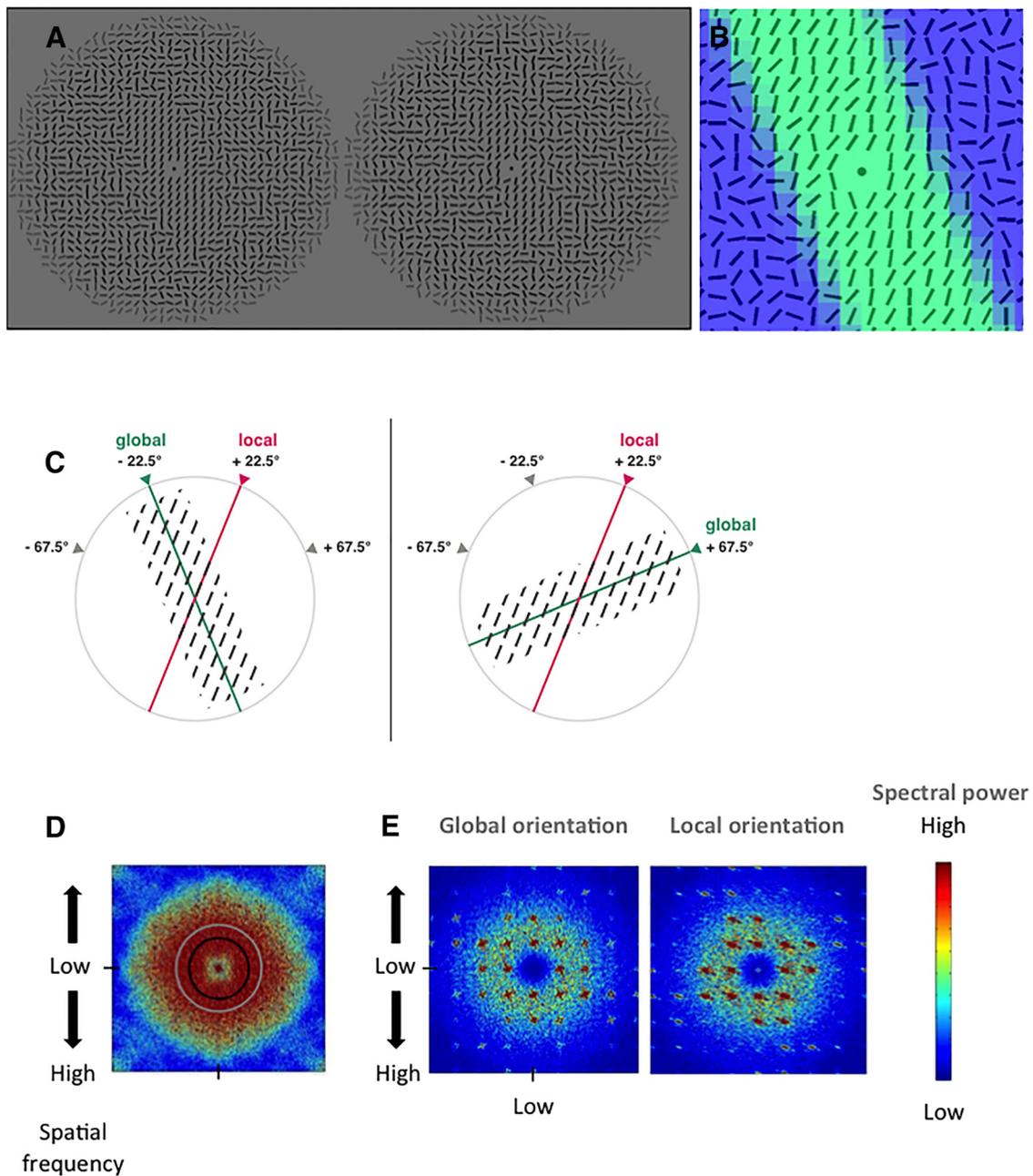


Figure 2. Stimuli. **A**, Two examples of the stimuli, at two coherence levels (left, $\kappa = 20$, high coherence level for presentation purpose; right, $\kappa = 4$, a coherence level representative of the stimuli used in the present study). The local and global orientations (22.5° , -22.5° respectively, vertical reference) are illustrated in **C**. **B**, Schematic representation of the coherence (κ) values within a stimulus. There is a smooth decrease of the coherence between the global oriented shape and the background: green represents maximal coherence ($\kappa = 5$); dark blue represents minimal coherence ($\kappa = 0$, random orientations). **C**, A local orientation was always combined with a global orientation at $\pm 45^\circ$ (modulo 180°) from the local orientation, and vice versa. In this example, a local orientation at 22.5° is combined with a global orientation of either -22.5° (left) or 67.5° (right). Orientations were 67.5° , 22.5° , -67.5° , and -22.5° , defining eight possible combinations of local and global orientations. **D**, 2D-Fourier spectrum of the stimuli. Low spatial frequencies are presented in the center. The frequency cutoff of the filters used in the control experiment on the role of spatial frequency (5.7 and 7.9 cpd) are indicated by a black and a light gray circle respectively. **E**, The 2D spectra of the global orientation (left) and of the local orientation (right), were derived from the subtraction of the Fourier spectrum of stimuli of different global orientations, and of different local orientations, respectively (for details, see Material and Methods). The spatial frequency domains supporting global and local orientation information were largely overlapping.

coherence were used (Fig. 2A, right, between 3 and 5) to enforce subjects to process many lines in the local task.

Stimulus properties

The stimuli were characterized by several unique properties. First, as opposed to many existing textured stimuli (Wolfson and Landy, 1995; Supér et al., 2001; Jolij et al., 2011; Poort et al., 2012), the global shape was exclusively defined by the distribution of a single local feature, line orientation. It had no luminance contour, no T-junction, or any abrupt transition between shape and background. The smooth transition was

implemented by a progressive sigmoidal decrease of coherence κ from shape to background, which is reminiscent of the “texture blur” proposed by Landy et al. (Landy and Bergen, 1991; Landy and Kojima, 2001). The absence of contour is critical to enforce a local-to-global hierarchical processing of the stimuli: to represent the global shape, it was mandatory to process the ensemble of homogeneously oriented lines and whose orientation is the local orientation. Second, the local information and the global information refer to the same basic feature, orientation, and the values this orientation can take, namely, 22.5° , 67.5° , -22.5° , -67.5°

degrees, are the same at the local and global level. Thus, the local information and the global information derive from the same physical information and only differ by their level of representation. To check that the global shape orientation and the local orientation solely differed by their level in the visual hierarchy while deriving from the same physical information, we compared the spatial frequency ranges in which global and local information was encoded. We averaged 24 stimuli of global orientation 67.5° and of local orientation -67.5° , and computed its 2D Fourier spectrum. We then created another mean image with a different global orientation (-22.5°) but with the same local orientation (-67.5°), and computed its 2D Fourier spectrum. To identify the spatial frequencies associated solely with global orientation (Fig. 2E, left), we then computed the difference between those two spectra. The same procedure was repeated to compute the difference between the 2D Fourier spectra of two mean images of different local orientations (67.5° and -22.5°) but of same global orientation (-67.5°) (Fig. 2E, right). The spectral domains associated with local and global orientations were highly similar, thus confirming that global and local orientations are encoded in similar spatial frequency ranges. To further confirm that low spatial frequencies were not involved in fast reports of the global information, we ran a control experiment with stimuli in which this information is filtered out.

Procedure

All participants were naive subjects with normal or corrected-to-normal vision. They gave informed written consent and were paid for their participation. Procedures were approved by the national ethics committee (Comité Consultatif de Protection des Personnes dans la Recherche Biomedicale).

Trials considered for analysis

On average, across all experiments, subjects failed to respond in less than 5% (maximum across subjects and experiments: 11%) of the total amount of trials. Only trials with a response were considered for analysis. The only trial with a reaction time inferior to 0.3 s (across all participants and experiments) was discarded.

Experiment 1: local and global tasks

Participants. Twenty-one right-handed subjects took part in the study (10 women, ranging from 19 to 26 years of age, mean: 20.8 years). In addition to these 21 subjects, 2 subjects were excluded during the training phase because they failed to report the orientation of the stimuli (performance lower than 30%, chance level at 25%).

Paradigm

Each trial began with a fixation point presented for 0.6–0.8 s and followed by a stimulus presented on an LCD screen (refresh rate 60 Hz, resolution 1680×1050 pixels, size of the screen, 43.6×27.1 cm) at a viewing distance of 70 cm. A new stimulus was created at each trial. The stimulus remained on the screen until subjects responded, or at most after a stimulus presentation time of 1.5 s. Subjects were engaged in a four-alternative forced-choice task on either the local or the global orientation in a blocked design. Subjects were told that, in the center of the stimulus, lines were homogeneous regarding their orientation and that this homogeneity defined a rectangular shape. In the global task, they were instructed to report the orientation of the shape made of this homogeneous ensemble of lines. In the local task, subjects were instructed to report the general direction along which they felt the lines were oriented. Also, to avoid artificially long reaction times, subjects were explicitly told that the local task was not requiring to compute the exact average orientation of the lines.

The nature of the block-local or global-and the four possible orientations were displayed on the screen at the beginning of each block. Subjects answered by pressing a response button with their right hand, from index to little finger, each button corresponding to an orientation. After a variable delay of 1.8–2.2 s with a gray screen, a new trial began. Subjects performed 11 local blocks and 11 global blocks in alternation (528 trials over the whole experiment, 24 trials per block). The nature of the first block was counterbalanced across subjects.

Before the experiment, subjects were shortly trained with stimuli at a coherence of 5. They first performed one local and one global block with an auditory feedback indicating whether their response was correct, followed by one local and one global block without feedback. To estimate a local coherence threshold ($\text{Coh}_{\text{local}}$) and a global coherence threshold ($\text{Coh}_{\text{global}}$), we engaged subjects in 2 one-up-one-down psychophysical staircases, for which they would perform at 80% accuracy. In practice, subjects were engaged in the blocked paradigm previously described, except that the coherence κ was varied depending on the correctness of the response. Threshold coherences were thus estimated separately for the local and the global task (initial threshold, 8).

Experiment 2: decoding of global orientation from MEG data

Participants. Sixteen right-handed subjects took part in the MEG experiment (8 women, ranging from 18 to 30 years of age, mean: 23.14 years). Two subjects were discarded from the analysis due to an insufficient number of correct trials per stimulus orientation (<40 correct trials for a given orientation in either the local or global task). The results are therefore based on data from 14 subjects.

Recordings

Continuous magnetoencephalographic signals were collected using a whole-head MEG system with 102 magnetometers and 204 planar gradiometers (Elekta Neuromag TRIUX, sampling rate of 1000 Hz; online 330 Hz low-pass filtered). A vertical electro-oculogram and electrocardiogram (bipolar EKG with electrodes on the right clavicle and left abdomen) were simultaneously recorded with the ground electrode on the left shoulder blade. Horizontal and vertical eye positions as well as pupil diameter were monitored using an eye-tracker (EyeLink 1000 system, SR Research).

Paradigm

Stimuli and tasks were the same as in Experiment 1, except that the coherence κ was fixed across subjects and equal to 4. Stimuli were back-projected on a translucent screen at a viewing distance of 80 cm. After a variable delay of 1.5–2 s during which subjects were instructed to blink, a new trial began. Subjects performed 5 local blocks and 5 global blocks in alternation (72 trials per block, 720 trials over the whole experiment). The nature of the first block (local or global) was counterbalanced across participants.

MEG data preprocessing

Data were preprocessed with Maxfilter to remove extracranial noise and realign data between runs. Trials contaminated by artifacts (blinks, saccades $>1^\circ$ visual angle, as well as visually identified artifacts in MEG data) and occurring between -0.4 s to 0.8 s around stimulus onset were discarded from further analysis; 6% of trials were removed on average across subjects. The number of remaining trials did not differ between local and global blocks (paired t test: $T_{(13)} = 0.74$, $p > 0.4$).

Global orientation decoding

Decoding analysis was performed using the linear Support Vector Machine implemented in the Fieldtrip toolbox (Oostenveld et al., 2011). Data were band-passed filtered (fourth-order Butterworth filter) between 0.5 and 25 Hz and downsampled to 100 Hz. For each subject, the multivariate pattern analysis was conducted on single-trial data from all sensors (correct trials only, -0.2 to 0.85 s around stimulus onset) smoothed over time windows of 30 ms shifted by steps of 10 ms.

To test whether information about the global orientation was present in the MEG signal, two linear classifiers were trained in the global blocks, where global orientation was task-relevant, and two classifiers were trained in the local blocks, where global orientation was task-irrelevant. A first classifier was trained to distinguish between a global orientation at -67.5° (associated local orientation, -22.5° or 67.5°) and a global orientation at 22.5° (associated local orientation, -22.5° or 67.5°). The signals to classify were thus obtained in response to stimuli that had the same local orientation but which differed in their global orientation. A second classifier was trained to distinguish between a global orientation at 67.5° (associated local orientation, 22.5° or -67.5°) and a global orientation at -22.5° (associated local orientation, 22.5° or -67.5°).

tation, 22.5° or –67.5°). We then computed for each subject, at each time point, the mean accuracy of the two classifiers, separately in the local task and in the global task.

The classification procedure ran as follows: (1) The signal at each time point, for each subject, was z-score normalized across sensors and a linear Support Vector Machine was used as a classifier. (2) For each subject, for each classifier, a sixfold cross-validation leave-one-out procedure was used. The data were split in 6 folds; the classifier was trained on 5 of them and tested on the remaining one. This procedure was repeated 6 times, with the data used for training and testing drawn randomly. (3) For each subject, a single measure of accuracy was obtained by averaging the classification performance across the two classifiers. We then estimated whether those accuracies differed from chance at the group level. Chance level was estimated by randomly assigning task labels to the MEG data, computing decoding accuracy, repeating 100 times the random assignment of task labels and computing the mean decoding accuracy thus obtained. Significant differences between decoding accuracy in original data and chance level were then detected across subjects using paired *t* tests and a clustering procedure (Maris and Oostenveld, 2007) as implemented in the FieldTrip toolbox. This procedure identifies time windows of significant differences while correcting for multiple comparisons over time. Significant differences in decoding accuracy between local and global blocks were obtained through the same clustering procedure.

Local orientation decoding

The same procedure was applied on local information, to test whether information about the local orientation was present in the MEG signal, in local and global blocks.

Experiment 3: spontaneous and instructed tasks

Participants. Nineteen right-handed subjects took part in the study (11 women, ranging from 20 to 26 years of age, mean: 22.5 years).

Paradigm

The material (the screen parameters, the viewing distance, the stimuli, the visuomotor associations between orientations and response buttons) was the same as in the first experiment, except that the coherence was now fixed across subjects (values of 3, 4, and 5 randomly intermixed within each block). Subjects had 2 s to answer after stimulus onset.

They were engaged in two successive paradigms: the spontaneous condition, detailed below, followed by the instructed global/local tasks as in previous experiments. In the spontaneous condition, subjects were instructed to report the orientation they spontaneously perceived when viewing the stimulus. Subjects were told that there was no good/bad answer and that only their impression mattered. Therefore, they had the possibility to report having no impression of orientation by pressing a specific response button with their right thumb. To assess spontaneous vision in the absence of prior knowledge and goals, we did not inform subjects on the local and global distinction. We took great care not to bias subjects toward one type of orientation, local or global. For this purpose, in the instructions, the four orientations were represented as colored oriented lines of a size intermediate between that of the global shape and of local lines. Subjects were first trained, in the presence of the experimenter, by reporting the orientation of 12 congruent stimuli (i.e., stimuli with an identical global and local orientation). Subjects were then engaged in the spontaneous experiment: 12 consecutive congruent stimuli were followed by 432 trials split into 18 blocks of 24 trials, with a random alternation of 216 congruent stimuli and 216 incongruent stimuli. The analysis was restricted to incongruent trials to determine whether subjects spontaneously reported more the local or the global orientation.

After the spontaneous condition, subjects completed the local and global tasks described in the first experiment, with the same stimulus duration and coherence values as in the spontaneous task. All stimuli were incongruent. An initial training session consisted of two local and two global blocks with an auditory feedback at a coherence of 5. Subjects were then engaged in the tasks for 288 trials (144 trials in the local task, 144 in the global task, split in 6 local blocks and 6 global blocks of 24 trials, in alternation).

Control experiment with filtered stimuli

Participants. Eighteen right-handed subjects took part in the study (13 women, ranging from 20 to 26 years of age, mean: 23.7 years). One subject was discarded from the analysis because his individual accuracy, both in the local and in the global task, differed from group accuracy by more than 2 SDs. Analyses are therefore presented on 17 subjects.

Stimuli

Stimuli were the same as in previous experiments, except that a fourth-order Butterworth high-pass filter was applied on some stimuli. The coherence κ was fixed across subjects and equal to 4. We used two levels of filtering (Fig. 2D): one-third of the stimuli were filtered in a medium spatial frequency range (F1, frequency cutoff: 5.7 cpd), one-third of the stimuli were filtered at higher spatial frequencies (F2, frequency cutoff: 7.9 cpd), and the remaining stimuli were unfiltered.

Paradigm

The experiment consisted of 432 stimuli split in 9 local blocks and 9 local blocks of 24 trials (72 F1-filtered and 72 F2-filtered stimuli, 72 unfiltered stimuli, randomly intermixed within each block). Subjects were instructed to report either the local or the global orientation in alternating blocks.

Results

Experiment 1: local and global tasks

In the first experiment, 21 subjects reported either the global or the local orientations, in a block design (24 trials per block). We predicted that, if the visual hierarchy constitutes a structural constraint upon conscious vision, global orientations encoded at high levels of the visual hierarchy should be reported faster than local orientations. Furthermore, if global orientations are perceived first, they might bias reports of the local orientation.

Shorter reaction times in the global task

Reaction times in the global task were significantly shorter than in the local task (Fig. 3A), both in correct and incorrect trials (repeated-measures, main effect of task: $F_{(1,20)} = 15.1, p < 0.001$; correct trials: global: 0.91 ± 0.01 s, local: 0.96 ± 0.15 s, paired *t* test: $T_{(20)} = -3.3, p < 0.01$; incorrect trials: global: 1.01 ± 0.02 s, local: 1.063 ± 0.019 s, paired *t* test: $T_{(20)} = -3.1, p < 0.01$).

The shortening of reaction times could not be accounted for by a speed-accuracy trade-off. First, as a consequence of the staircase procedure, subjects performed equally well in both tasks at the group level (mean accuracy: global task: $83.4 \pm 1.38\%$, local task: $82.5 \pm 1.27\%$, paired *t* test: $T_{(20)} = 0.45, p > 0.6$). Coherence levels were similar in the two tasks (global task: 3.3 ± 0.23 , local task: 3.7 ± 0.29 , paired *t* test: $T_{(20)} = 1.6, p > 0.1$). We further tested whether shorter reaction times in the global task could be due to subtle between-subject differences in accuracy between the global and local tasks, especially due to some speed-accuracy trade-off. As shown in Figure 3B, we computed the correlation between the local-global difference in accuracy and the local-global difference in reaction times across subjects. A speed-accuracy trade-off would correspond to a positive correlation (i.e., subjects faster in the global task would be less accurate in this task), whereas we observed a negative correlation (linear regression between accuracy, local – global, and reaction times in correct trials, local – global, model validity: $F_{(1,20)} = 12, p < 0.01$, strength of the correlation: $r = -0.6, T_{(20)} = -3.5, p < 0.01$). In other words, subjects who were faster in the global task than in the local task were also more accurate in the global task. To further probe whether the reaction time advantage in the global task could be due to a speed-accuracy trade-off, we computed the intercept of the regression that estimates the difference in reaction times an ideal observer with the exactly same accuracy in the global and local tasks would have. Such an ideal observer would be 0.049 s faster in the global task ($T_{(20)} = 4.41, p < 0.001$).

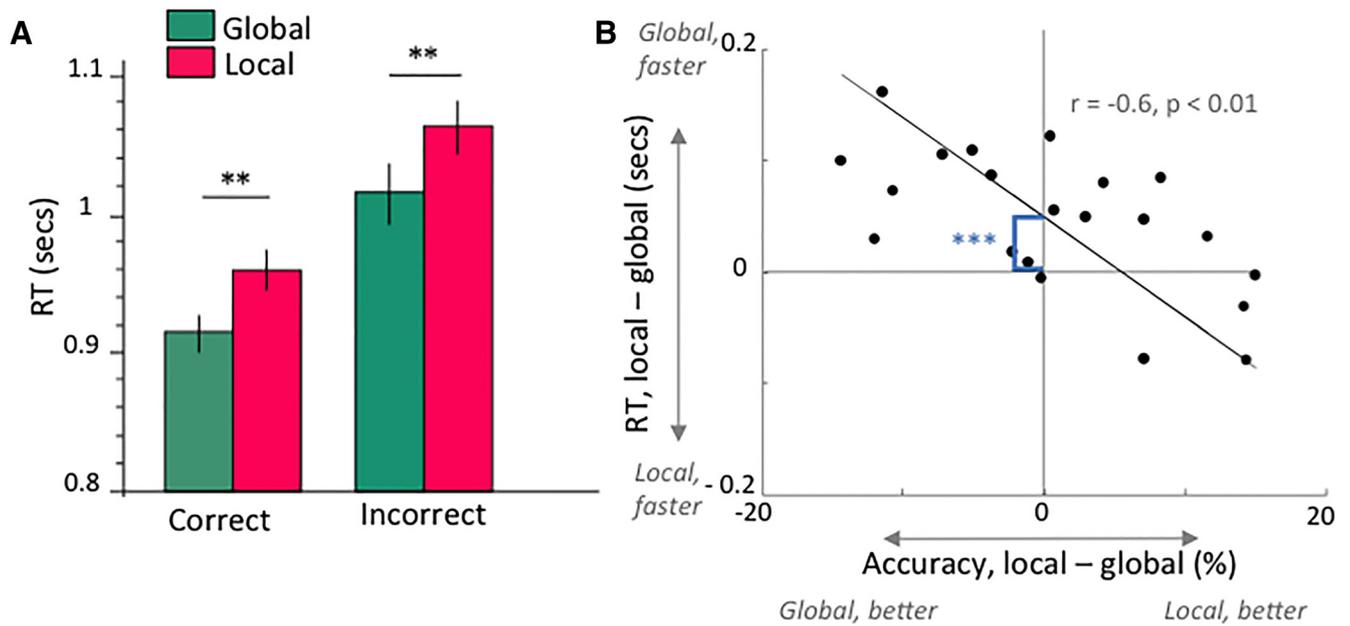


Figure 3. Experiment 1. Subjects are faster to report the global orientation. **A**, Reaction times in the global and local tasks, for both correct (left) and incorrect (right) trials. Reaction times in the global task were shorter. **B**, Difference in reaction times between the local and global tasks as a function of the difference in accuracy. The two variables are significantly linearly correlated (strength of the correlation: $r = -0.6, p < 0.01$) with a negative slope, showing that subjects who were faster at the global task were also more accurate. The significantly positive intercept indicates that an ideal subject with strictly identical accuracies at the local and global tasks would be 0.049 s faster in the global task. $**p < 0.01$. $***p < 0.001$.

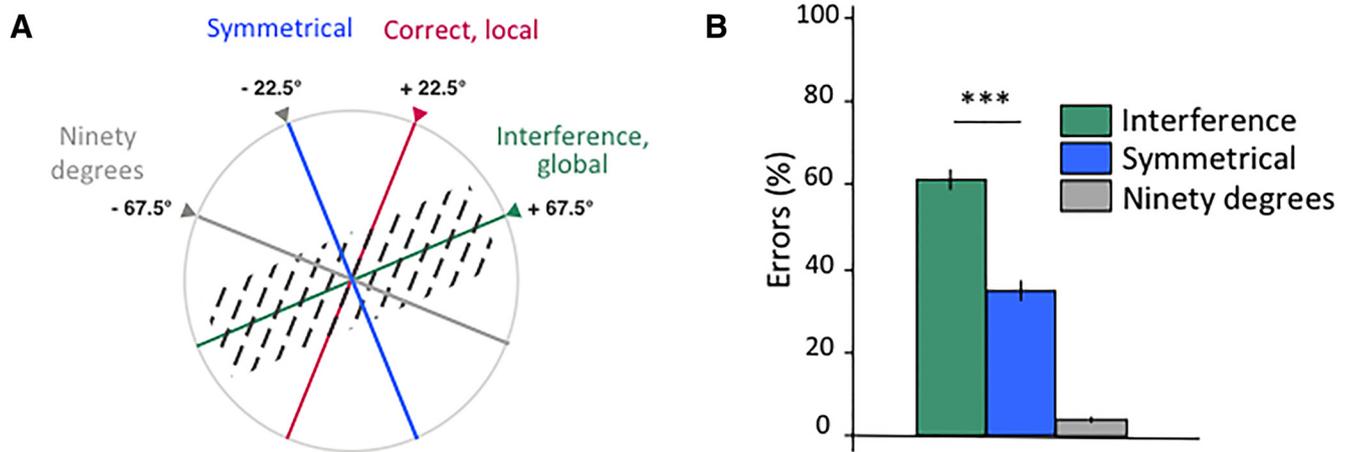


Figure 4. Experiment 1. The task-irrelevant global orientation impacts responses in the local task. **A**, For any given stimulus in the local task (here an example where the stimulus has a local orientation at 22.5°), subjects could make three different types of errors: (1) an interference error, corresponding to the task-irrelevant global orientation; (2) a symmetrical error, corresponding to the orientation symmetrical to the interference error with respect to the correct orientation (both these errors are at $\pm 45^\circ$ from the correct orientation); and (3) a ninety degrees error, at 90° from the correct orientation. **B**, Rate of interference errors, symmetrical errors, and ninety errors in the local task computed within error trials. Interference errors were more numerous and reflect a bias toward global orientation in the local task. $***p < 0.001$.

To conclude, accuracy does interact with reaction times, but accuracy differences cannot entirely account for shorter reaction times in the global task.

Global orientations influence reports of the local orientations

If the global orientation is perceived before the local orientation, subjects might tend to report the global orientation while performing the local task. To assess a bias toward global orientations, we analyzed the pattern of errors. Figure 4A describes the answers a subject can make in the local task: the correct local orientation, or different types of errors: (1) the global orientation that is always at $\pm 45^\circ$ from the local one, resulting in what we called an interference error; (2) the orientation symmetrical to the interference error with respect to the correct orientation, also at $\pm 45^\circ$

of the correct answer, resulting in what we called a symmetrical error; and (3) the remaining orientation, at 90° of the correct answer, resulting in what we called a 90° error.

If the global orientation influences perceptual local reports in the local task, subjects should be biased toward the global orientation. In other words, in the local task, subjects should commit more interference errors (i.e., global reports) than symmetrical errors, although these errors are similar in terms of their angular distance from the correct orientation. Subjects made significantly more interference errors than symmetrical errors (Fig. 4B): within errors in the local task, interference errors represented $61.57 \pm 2.3\%$ and symmetrical errors only $34.68 \pm 2.4\%$ (paired t test: $T_{(20)} = 5.78, p < 0.001$). The few remaining errors were 90°

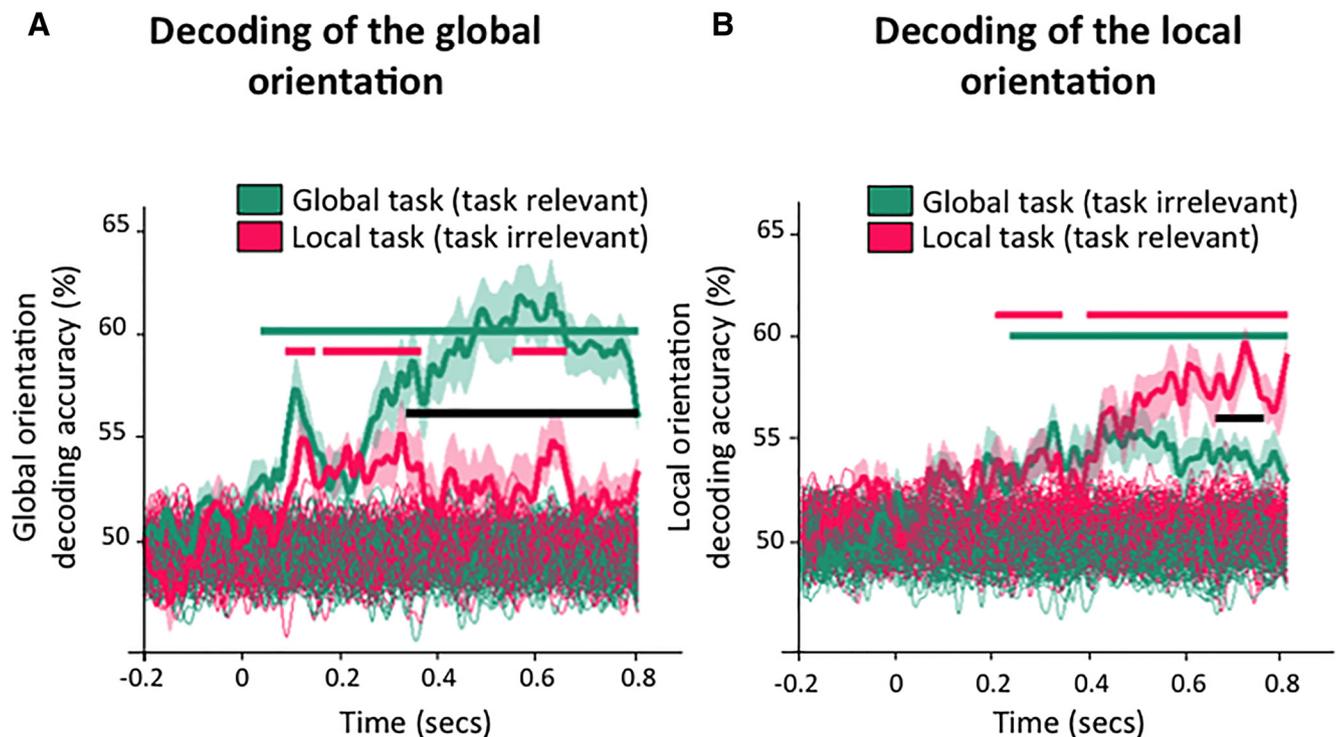


Figure 5. Experiment 2. Time course of the accuracy of the decoding of the global orientation and of the local orientation from MEG data. **A**, Decoding accuracy (thick lines) of the global orientation, averaged across subjects, in the global task where global orientation is task-relevant (green) and in the local task (pink) where global orientation is task-irrelevant. The global orientation could be decoded in both tasks. **B**, Decoding accuracy of the local orientation, averaged across subjects, in the local task where local orientation is task-relevant (pink) and in the global task (green) where local orientation is task-irrelevant. The local orientation could be decoded in both tasks. **A, B**, Thick lines indicate the decoding accuracy from the MEG data. Thin dotted lines indicate decoding accuracy > 100 shuffled datasets and represent an estimate of chance level. Green and pink horizontal bars represent time windows with a better-than-chance decoding of global orientation in the global and local task, respectively (Clustering test, Monte-Carlo $p < 0.01$ one-tailed, corrected for multiple comparison). Black bar represents significant between-tasks differences.

errors: they represented $3.75 \pm 0.74\%$ of the errors. Thus, the global orientation in the stimuli influences the reports in the local blocks.

To rule out the possibility that subjects reported the global orientation in the local task because they failed to switch to this task, we checked that interference errors occurred at a similar rate in the beginning and end of each local block (local blocks, first half of trials, interference errors: $62.9 \pm 2.7\%$, symmetrical errors: $33.7 \pm 2.7\%$, paired t test: $T_{(20)} = 5.54$, $p < 0.001$; local blocks, second half of trials, interference errors: $59.9 \pm 4\%$, symmetrical errors: $35.8 \pm 3.9\%$, $T_{(20)} = 3$, $p < 0.01$; ANOVA with factors error type and block half, no effect of block half nor interaction, both $F_{(1,20)} < 0.6$, both $p > 0.4$).

Interim conclusion

The first two predictions are met. First, subjects report global orientations faster. Second, the global orientation interferes with the local task, although this information is task-irrelevant.

Experiment 2: decoding of global orientation from MEG data

In the second experiment, MEG signals were recorded while subjects reported either the global or local orientations in alternation, in a blocked design similar to Experiment 1. The major difference was that coherence level was fixed at 4 for all subjects and tasks. According to the reverse hierarchy theory, global information is computed during the automatic feedforward sweep (i.e., during a first stage of hierarchical processing), independently of task demands. We therefore tested whether global information could be decoded from MEG data in the global as well

as in the local task. Correct trials only were considered for analysis.

Behavior

Subjects were both faster and more accurate in the global than in the local task (reaction times: global, 0.83 ± 0.16 s, local, 0.94 ± 0.18 s, paired t test: $T_{(13)} = 2.8$, $p < 0.05$; accuracy: global, $93.5 \pm 5\%$, local $81.8 \pm 9\%$, paired t test: $T_{(13)} = 6.16$, $p < 0.001$). As in Experiment 1, the global orientation influenced reports of the local orientation: in the local task, subjects made significantly more interference errors (i.e., reports of the global orientation in the local blocks: $69 \pm 2.99\%$) than symmetrical errors (symmetrical errors: $29.9 \pm 3.08\%$, paired t test: $T_{(13)} = 6.45$, $p < 0.001$).

Information about global orientation is present in early brain signals in both the local and global tasks

Decoding accuracy of global information presented an initial peak ~ 0.130 s in both the local and global tasks (Fig. 5A), followed by a rebound in decoding accuracy (starting at 0.3 s after stimulus) much more sustained in the global task. This qualitative description was backed up by the statistical analysis. Across participants, global information could be decoded significantly better than chance in the global task in a sustained manner, from 0.055 to 0.825 s (Monte-Carlo $p < 0.01$, corrected for multiple comparison over time). Crucially, global information could be decoded significantly better than chance in the local task as well (Fig. 5A) in three time windows: 0.105–0.165, 0.185–0.385, and 0.575–0.685 s. Decoding accuracy of global orientation differed between tasks only in the late part of MEG signals, from 0.355 s to 0.825. Thus, global information could be decoded early from

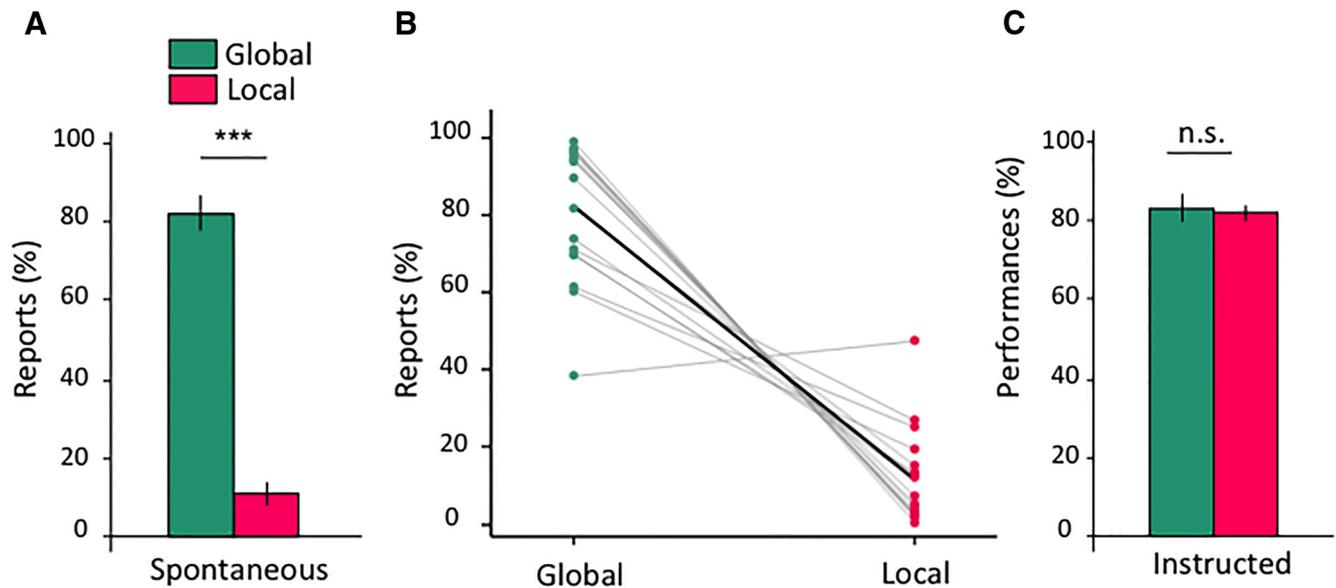


Figure 6. Experiment 3. Reports in the spontaneous condition and in the instructed global and local tasks. **A**, Report rate of the global and the local orientation in the spontaneous condition. Subjects massively reported the global orientation. **B**, Individual data in the spontaneous condition. All subjects but one reported the global orientation more often. **C**, Accuracy in the global and local tasks for the subjects previously engaged in the spontaneous condition. With instructions, subjects could report the local and the global orientation equally well. *** $p < 0.001$. n.s., Not significant.

MEG signals. This was the case even when subjects were engaged in the local task that does not require the use of the global information. Overall, decoding accuracy of the local information (Fig. 5B) was rather low, possibly because of the noise in line orientations as well as because of the larger variability in the anatomy of early visual areas across healthy individuals relatively to higher visual areas (Dougherty et al., 2003) encoding the global information.

Interim conclusion

Decoding results indicate that the global information was computed early, with a peak of decoding accuracy ~ 0.130 s and that this early computation of the global information took place even when global information was not task-relevant: global information seems therefore to be computed automatically, in accordance with the predictions of the reverse hierarchy theory.

Experiment 3: spontaneous and instructed tasks

The third experiment aimed at testing the hypothesis that the automatic processing of global information comes along with the perception of the global information, constituting thereby the default content of perception in the absence of specific task demands.

Nineteen subjects were asked to report the orientation they perceived spontaneously. They were presented with stimuli identical to those of previous experiments (incongruent stimuli, with different local and global orientations) as well as congruent stimuli (identical global and local orientations), and had to either select one of four orientations or use a fifth key to indicate an absence of clear percept of orientation. Subjects were told that there was no good answer and that only their impression mattered. The training procedure and instructions were formulated to avoid biasing subjects toward either local or global orientations, as detailed in Material and Methods. After this spontaneous condition, the local and the global information was described to the subjects, who were then engaged in the local and the global tasks in alternating blocks.

Spontaneous reports are reliable

Because, in the spontaneous condition, subjects were informed that there was no correct answer, they might have answered randomly. We checked that this was not the case over congruent trials, which were not used in the remaining analysis: if subjects reported the content of their perception, their reports should depend on stimulus properties, and the congruent orientation should be reported more often than the three other orientation possibilities. This was found in all subjects (average reports of the dominant orientation among the (19–1) subjects: $77 \pm 4.71\%$) but one, who was excluded from further analysis (binomial test with chance at 25%, all $p < 0.001$ except one subject with $p > 0.9$). Subjects reported no percept of orientation in $3.3 \pm 1.2\%$ of the answered trials. Further results are within trials in which an orientation was reported.

Subjects spontaneously report the global orientation

We then assessed over the incongruent trials whether subjects spontaneously report the global orientation rather than the local orientation (Fig. 6A). Subjects reported the global orientation significantly more often than the local orientation (global: $82.3 \pm 4\%$, local: $11.5 \pm 2.9\%$, $T_{(17)} = 10.3$, $p < 0.001$). The dominance of global percepts in spontaneous reports was present for all subjects but one (Fig. 6B). Reaction times were again shorter for global reports than for local reports (global: 1.036 ± 0.007 s, local: 1.24 ± 0.01 s; paired t test: $T_{(17)} = 4.95$, $p < 0.001$).

Instructed global/local tasks

At the predefined levels of coherence used in the spontaneous condition, subjects may have been able to perceive the global orientation and unable to perceive the local orientation, and would therefore have rarely reported the local orientation. To rule out this possibility, the same subjects were then described the stimuli and instructed to report either the local or the global orientation of incongruent stimuli as in previous experiments, but using the same coherence levels as in the spontaneous condition. Subjects were equally accurate in the local and global tasks

Table 1. Experiment 3: similar rate of reports and reaction times for the global information in the spontaneous condition and in the global instructed task^a

	Coherence 3		Coherence 4		Coherence 5	
	Spontaneous	Global	Spontaneous	Global	Spontaneous	Global
Reports of the global orientation (%)	79.7 ± 4.6	79.2 ± 3.9	82.4 ± 3.9	84.7 ± 3.7	84.4 ± 3.9	86 ± 3.6
T ₍₁₇₎ p value	−0.2 > 0.8		0.7 > 0.5		0.4 > 0.7	
Reaction time (s)	1.084 ± 0.033	1.072 ± 0.041	1.032 ± 0.035	1.011 ± 0.1046	0.999 ± 0.033	0.966 ± 0.044
T ₍₁₇₎ p value	−0.5 > 0.6		−1.1 > 0.3		−1.4 > 0.2	

^aData are mean ± SEM. Reports of the global information in the spontaneous condition and in the global instructed task at each coherence value (Experiment 3), and corresponding reaction times. Comparison between the spontaneous task and the global task (paired *t* test). At each coherence value, the rate of reports of the global information in the spontaneous condition and in the global task is similar. This is also the case for reaction times.

(Fig. 6C) (global: 83.3 ± 3.6%, local: 81.8 ± 2.35%, paired *t* test: T₍₁₇₎ = −0.67, *p* > 0.5). This shows that subjects were able to report the local orientation as well as the global orientation if explicitly instructed to do so, but they did not report it spontaneously.

The reports of global information are correlated between the spontaneous condition and the global task

Rate of reports of global orientation were similar in the spontaneous condition and in the global instructed task (rate of reports of global orientation, spontaneous condition: 82.3 ± 4%, global task: 83.3 ± 3.6%, paired *t* test: T₍₁₇₎ = −0.3, *p* > 0.7; for report rates by coherence value, see Table 1) and significantly correlated within subjects (Pearson correlation, $r_{\text{global}} = 0.72$, Bonferroni corrected *p* < 0.001) (Fig. 7A). By contrast, the rate of reports of the local orientation was not correlated between the spontaneous condition and the local instructed task (Pearson correlation, $r_{\text{local}} = -0.2$, Bonferroni corrected *p* > 0.9) (Fig. 7C).

Comparison of correlation coefficients

To conclude safely about a difference regarding the correlation for global information versus local information, we tested for statistical differences between r_{global} and r_{local} . We first created a distribution of differences between r_{global} and r_{local} under the null hypothesis by creating 1000 surrogate datasets in which the labels global and local were randomly switched. In practice, a surrogate dataset was created by randomly assigning, for each subject, the labels “Local” and “Global” to the two pairs of values (Accuracy in global task, Global reports in spontaneous condition) and (Accuracy in local task, Local reports in spontaneous task), for each subject. We then computed the difference between the correlation coefficients r_{global} and r_{local} on each surrogate dataset. The Monte-Carlo *p* value corresponds to the fraction of surrogate $r_{\text{global}} - r_{\text{local}}$ differences exceeding the original difference. This procedure shows that the correlation coefficient r_{global} linking global reports in the spontaneous condition and in the global task was significantly larger than r_{local} linking local reports in the spontaneous condition and in the local task (Monte-Carlo *p* < 0.01).

Similarly, reaction times to report global orientation in the spontaneous condition and in the global task were similar (reaction times to report the global orientation, spontaneous condition: 0.1037 ± 0.044 s, global task: 0.1015 ± 0.047 s, paired *t* test: T₍₁₇₎ = 1.02, *p* > 0.3; for reaction times by coherence value, see Table 1) and significantly correlated (Pearson correlation, $r_{\text{RTglobal}} = 0.89$, Bonferroni corrected *p* < 0.001) (Fig. 7B). Reaction times to report the local orientation in the spontaneous condition and in the local instructed task were not significantly correlated (Pearson correlation, $r_{\text{RTlocal}} = 0.45$, Bonferroni corrected *p* > 0.25) (Fig. 7D). The correlation coefficient r_{RTglobal} linking global reaction times in the spontaneous condition and in the global task was significantly larger than r_{RTlocal} linking local

reaction times in the spontaneous condition and in the local task (Monte-Carlo *p* < 0.01, using the procedure described above to compare the correlation coefficients).

Those results confirm that subjects did not respond randomly in the spontaneous condition and show that subjects spontaneously performed the global task. Importantly, the absence of correlation in the case of the local information shows that the correlation found for the global information cannot be trivially explained by idiosyncratic differences, some subjects being fast and good in any task and other subjects being always slower and less accurate.

Replication of results from Experiment 1

We also replicated the findings of Experiment 1 in the instructed task across the three levels of coherence (Table 2), with faster decisions in the global than in the local task and a bias toward global orientations in the local task.

Interim conclusion

The results of this experiment validate the third prediction we made, namely, that subjects spontaneously report global information, although they can report equally well global and local information when instructed to do so. Furthermore, reports were correlated within subjects between the spontaneous condition and the global instructed task. Together, these results suggest that conscious vision has a preferential global content, encoded in high levels of the visual hierarchy. We ran a number of control analyses and a control experiment. The results are described below.

Control analysis and experiment

We consistently observed longer reaction times when subjects reported the local orientation, which we attribute to a delayed access to local information. We have ruled out that speed-accuracy trade-off differences between local and global tasks could account for reaction time differences. In this section, we consider alternative interpretations.

Could longer reaction times in the local task have strategic grounds?

Subjects might have performed the local task by actively looking for the global shape and then averaging the orientation of the lines within the global shape. This would have resulted in longer reaction times in the local task. Two results speak against this possibility: in two instances where it was very unlikely that subjects could perceive the global shape, their reaction times were still longer to report the local orientation. First, in Experiment 1, 8 subjects had a coherence threshold lower in the local task relatively to the global task. In other words, those subjects performed the local task at a coherence level that was too low to properly see the global shape. In these subjects, reaction times were nevertheless longer in the local task (reaction times, local: 0.957 ± 0.176 s,

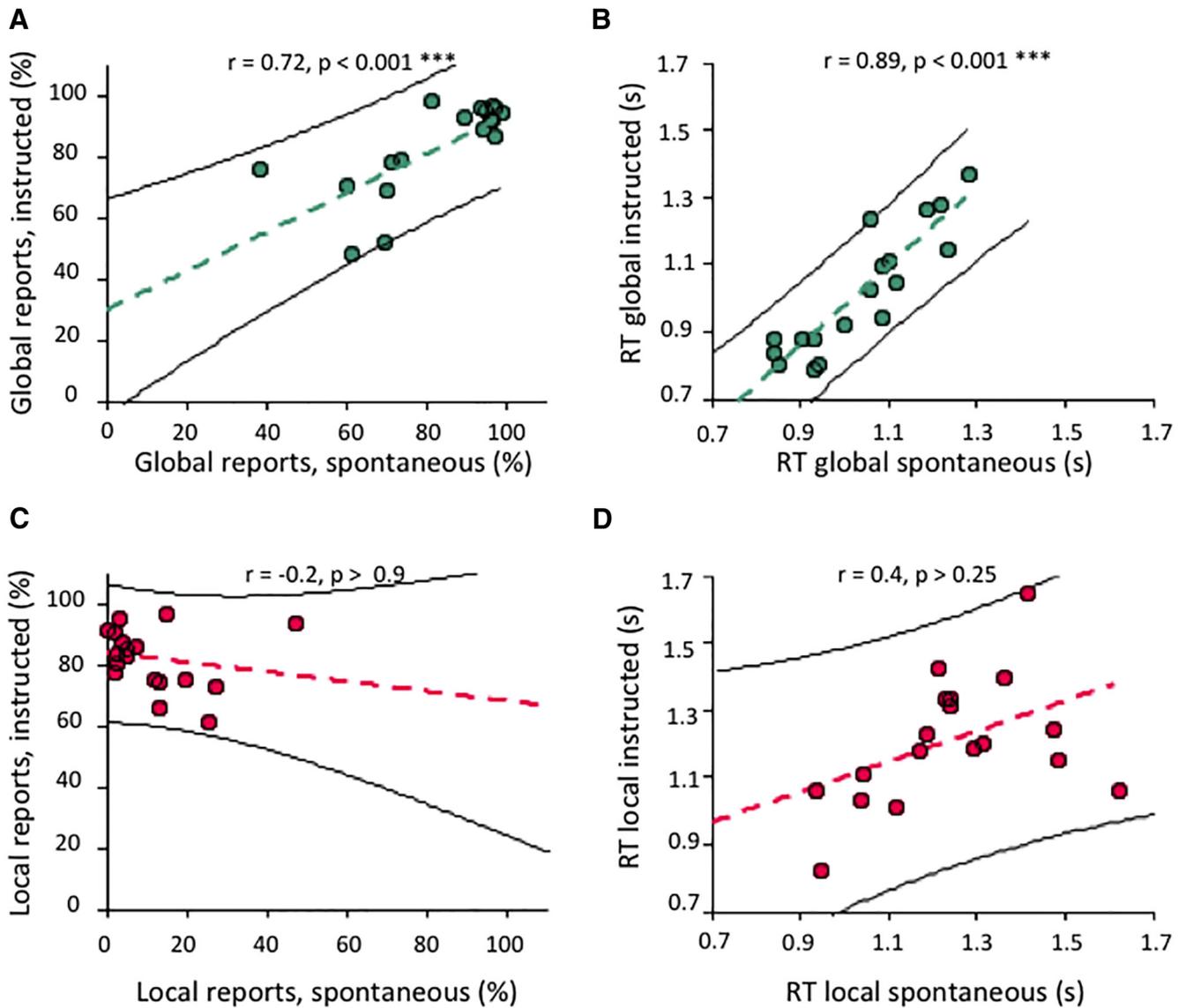


Figure 7. Experiment 3. Performance in the spontaneous condition and in the global task are correlated. **A**, Significant Pearson correlation between the accuracy in the global task and the percentage of global reports in the spontaneous condition, for each subject. **B**, Significant Pearson correlation between reaction times in the global task and when reporting the global orientation in the spontaneous condition, for each subject. **C**, Accuracy in the local task as a function of the percentage of local reports in the spontaneous condition, for each subject. No significant Pearson correlation was found. **D**, Reaction times in the local task as a function of the reaction times when reporting the local orientation in the spontaneous condition, for each subject. No significant Pearson correlation was found. *** $p < 0.001$ (Bonferroni corrected).

Table 2. Experiment 3: replication of the results of Experiment 1^a

	Coherence 3		Coherence 4		Coherence 5	
	Global	Local	Global	Local	Global	Local
Reaction times (s)	1.072 ± 0.041	1.223 ± 0.042	1.011 ± 0.046	1.215 ± 0.046	0.966 ± 0.044	1.186 ± 0.045
$T_{(17)} p$ value	−3.6 < 0.01		−4.5 < 0.001		−5.3 < 0.001	
Intercept > 0	3.6 < 0.01		4.6 < 0.01		7 < 0.001	
Errors (local task) (%)	66.3 ± 3 (interference)	21.5 ± 4.5 (symmetrical)	70 ± 6.1 (interference)	34.5 ± 7.2 (symmetrical)	71.3 ± 6.2 (interference)	15.7 ± 3.5 (symmetrical)
$T_{(17)} p$ value	8.4 < 0.001		2.9 < 0.01		8.3 < 0.001	

^aData are mean ± SEM. Results from the local and global tasks at each coherence value (Experiment 3), replicating the findings of Experiment 1. Reaction times are shorter in the global task. An ideal observer without any imbalance in accuracy between tasks would have shorter reaction times in the global task. There are more interference than symmetrical errors in the local task.

global: 0.903 ± 0.179 s, two-tailed paired t test: $T_{(7)} = -3.38, p = 0.012$). Second, in the spontaneous condition in Experiment 3, subjects were not informed about stimuli structure and on how the global shape was defined. However, their reaction times were

still longer when reporting the local orientation. We can therefore safely conclude that the longer reaction times in the local task are not based on a strategy requiring the perception of the global shape.

Table 3. Control experiment: effect of filtering on reaction times and accuracy^a

Condition	Reaction times (s)			Accuracy (%)		
	UF	F1	F2	UF	F1	F2
Global	0.898 ± 0.035	0.901 ± 0.038	0.913 ± 0.039	95 ± 1	95 ± 0.01	96 ± 0.1
Main effect of filtering	$F_{(1,312,16)} = 0.95, p > 0.3$			$F_{(1,858,16)} = 2.6, p > 0.09$		
Local	1065 ± 27	1005 ± 34	995 ± 31	93 ± 2	95	94 ± 1
Main effect of filtering	$F_{(1,680,16)} = 20.5, p < 0.001$			$F_{(1,369,16)} = 1.4, p > 0.2$		
$T_{(16)}$ p value	8.22, $p < 0.001$	4.44, $p < 0.001$	3, $p < 0.01$	-0.8, $p > 0.3$	0.7, $p > 0.4$	1.3, $p > 0.18$

^aData are mean ± SEM. Reaction times and accuracy in the global and local tasks (Control experiment), for unfiltered stimuli (UF) and at 2 different levels of high-pass filtering (F1 and F2). The main effect of the ANOVA in both tasks, and t test (global vs local task) over reaction times and accuracy are reported. There is no effect of filtering in the global task on either reaction times or accuracy. Reaction times are shorter in the global task whatever the level of filtering. The Greenhouse-Geisser correction for the degrees of freedom was applied when appropriate.

Are the stimuli truly hierarchical?

We designed our stimuli to impose a hierarchical dependence between levels: the global shape is a second-order property of local orientation, without contours. Creating truly hierarchical stimuli is a difficult task since, usually, global and local information not only differ by their level in the visual hierarchy but also by their spatial frequency spectrum, global information spanning lower spatial frequency ranges (Shulman et al., 1986). This imbalance is involved in the global precedence: typically, in Navon's stimuli, the global advantage is partly due to the fast processing of low spatial frequency information specific to the global letters (Badcock et al., 1990; Hughes et al., 1990; Kimchi, 1992; Lamb and Yund, 1993, 1996; Han et al., 2003; Flevaris et al., 2014). To assess in isolation the role of the visual hierarchy in the global precedence, it is essential to create stimuli in which the local and the global information is carried by similar spatial frequencies.

As shown in Figure 2E, the 2D Fourier spectrum associated with local and global orientations appears highly similar. We computed differences between local and global orientations by splitting the spectrum into three equally sized frequency domains (low, middle, and high) and found that, in all three frequency domains, local and global orientation information differed by <0.3% of total spatial power. Differences were thus extremely small, but nevertheless systematic (two-tailed paired t tests, all uncorrected $p < 0.002$). In the low spatial frequency range, there was actually slightly less global than local information, suggesting that faster reports of global orientation were not based on a preferential and faster processing of low spatial frequencies. To further confirm that low spatial frequencies were not involved in fast reports of the global information, we ran a control experiment with stimuli in which this information was filtered out. In this experiment, subjects were presented with unfiltered stimuli (UF) similar to those used in Experiments 1–3, and with two stimuli where an increasing amount of low spatial frequencies were removed (F1 and F2) (Fig. 2D). The crucial question is whether reaction times in the global task are lengthened by low spatial frequency removal. A one-way ANOVA with filtering as factor reveals that it is not the case ($F_{(1,16)} = 0.95, p = 0.3$; Table 3). Besides, paired t tests show that reaction times were significantly shorter in the global task than in the local task for the three types of stimuli (Table 3). Unexpectedly, reaction times were affected by filtering in the local task: removal of low frequencies speeded up reaction times in the local task (Table 3). This effect was present at both F1 and F2 compared with UF (paired t tests, UF vs F1: $T_{(16)} = 4.64, p < 0.001$; UF vs F2, $T_{(16)} = 5.55, p < 0.001$). However, reaction times in F1 and F2 did not differ ($T_{(16)} = 1.1, p > 0.2$). Together, the results from this control experiment show that a differential use of spatial frequencies in the local and global tasks may have contributed to increase the reaction time advantage in the global task, but that this reaction time advantage persists when low spatial frequencies are removed.

We also predicted that, if the local orientation is inevitably processed for the global orientation to be computed, as expected from the hierarchical structure of the stimuli, local orientation might influence reports in the global task. Subjects indeed committed more interference than symmetrical errors in the global task (Experiment 1: interference errors, $65.67 \pm 2.3\%$, symmetrical errors, $25.81 \pm 1.54\%$, two-tailed paired t test: $T_{(20)} = 9.5, p < 0.001$; Experiment 3: interference errors, $68.71 \pm 4.3\%$, symmetrical errors: $23.99 \pm 4.3\%$, two-tailed paired t test: $T_{(17)} = 8.04, p < 0.001$). This further control analysis confirms the hierarchical structure of our stimuli, compared with Navon's letters, for instance, where "global information interferes with the local response but not vice versa" (Lamb and Robertson, 1989).

Finally, we predicted that, if the global information is solely defined by the local information, increasing noise on line orientation (i.e., decreasing coherence κ) should decrease accuracy in both the local and the global tasks, and to the same extent. This prediction was met: in Experiment 3, noise level affected similarly the performance in both tasks. An ANOVA on performance with factors task and coherence κ showed a main effect of coherence ($F_{(1,17)} = 14.87, p < 0.005$), but no task \times coherence interaction ($F_{(1,17)} = 0.37, p > 0.6$) and no main effect of task ($F_{(1,17)} = 0.52, p > 0.4$).

Discussion

In this work, we have assessed whether high-level global information in the visual hierarchy dominates perception as proposed by the reverse hierarchy theory (Hochstein and Ahissar, 2002; Campana and Tallon-Baudry, 2013) or whether, alternatively, information at any level of the visual hierarchy is equally likely to be accessed and reported (Kouider et al., 2010). Our results directly support the former hypothesis. First, we show that, although local information had to be processed for global information to be computed, the global information was reported faster. This reaction time advantage could not be explained by subjects' strategy or by the use of low-frequency information. Second, global information computation appeared to be a mandatory processing step, even when it was not task-relevant: global information biased reports of local information, and the global information could be decoded from MEG signals in the first 0.15 s of visual processing with the same accuracy independently of the task in which subjects were engaged. Third, our conclusions apply also to spontaneous reports: subjects who were not informed on stimulus structure spontaneously reported the global orientation, although, once informed about stimulus structure, they could report the global and the local orientations equally well if instructed to do so. Crucially, the frequency and speed of the spontaneous reports were correlated with the frequency and speed of the reports in the global task, but not in the local task. Thus, spontaneous vision seems akin to perception voluntarily directed

at a global level. Together, these findings indicate that reported percepts correspond preferentially to integrated, global information at high levels of the visual hierarchy (Hochstein and Ahissar, 2002; Campana and Tallon-Baudry, 2013), both in the presence and in the absence of task demands.

Perception is constrained by the visual hierarchy

A direct link between the dominance of global percepts and the hierarchical organization of the visual system has been missing so far, presumably partly due to a lack of appropriate stimuli. Here, we were able to investigate such a link through the use of truly hierarchical stimuli, in which the local information and the global information emerge from the encoding, at different levels of the visual hierarchy, of the orientation of an ensemble of local lines. The local information and the global information do not differ in any other respects. To achieve this, we created a global shape exclusively defined by its local components because it has no contour.

The absence of a contour is critical because it avoids an imbalance in the spatial frequency content of local and global information, which might contribute to the faster perception of global information. In Navon's letters, typically, the fast identification of global letters involves the fast processing of low spatial frequency information: once these letters are high-pass filtered, the fast identification of the global letter is altered or even lost (Shulman and Wilson, 1987; Badcock et al., 1990; Hughes et al., 1990; LaGasse, 1993; Lamb and Yund, 1993). Similarly, in natural scenes, the global dominance (VanRullen and Thorpe, 2001; Oliva, 2005; Fei-Fei et al., 2007) and the later reports of detailed information (Macé et al., 2009) involve some fast processing of low spatial frequencies (Schyns and Oliva, 1994; Bullier, 2001; Bar, 2003; Peyrin et al., 2010; Peyrin and Musel, 2012). Therefore, these stimuli cannot be used to probe in isolation the role of the visual hierarchy in the dominance of global percepts.

The representation of the spatial frequency spectrum of the global and of the local orientations confirms that, in our stimuli, this information taps into similar spatial frequency ranges. In addition, we showed in a control experiment that the faster reaction times when reporting the global orientation were not based on a faster processing of low spatial frequency information: reaction times when reporting the global orientation were not affected by the removal of low spatial frequency information. Thus, low spatial frequency does not account for the dominance of global information observed here. There is no doubt that with more ecological stimuli, such as visual scenes, a faster processing of low spatial frequency will reinforce the tendency to perceive global information. However, while the faster processing of low spatial frequencies is known to be susceptible to task manipulation (Oliva and Schyns, 1997; Schyns and Oliva, 1997, 1999; Morrison and Schyns, 2001), in the present study, the global dominance is task-independent. Noteworthy, the bias toward global reports we consistently found in Experiments 1 and 3 cannot be trivially explained by a difference in the detectability of local versus global information: objective accuracies were matched in those experiments. Thus, the reports, either spontaneous or task-driven, appear constrained by the visual hierarchy.

Our MEG data are also consistent with the predictions of the reverse hierarchy theory (Hochstein and Ahissar, 2002; Campana and Tallon-Baudry, 2013). These data show that global information is processed within the first 0.15 s of visual processing, even when not task-relevant. The latency of the peak of decoding accuracy (0.13 s) is compatible with the

notion that such global integration takes place at the end of the feedforward sweep, around 100–150 ms (Lamme and Roelfsema, 2000). However, we do not know whether subjects become conscious at this moment or whether additional recurrent processing is necessary for consciousness to emerge (Koivisto, 2012; Supér et al., 2001). Similarly, we do not know which mechanisms are involved in the late access to local information. As proposed in the reverse hierarchy theory (Hochstein and Ahissar, 2002), top-down processing might be important, but lateral interactions between neurons in low-level visual areas might also be involved (Roelfsema, 2006).

Spontaneous reports can be reliably measured and should be compared with performance on global information

To mimic ecological situations in which one is not engaged in a visual goal-directed behavior (e.g., mind-wandering), we asked subjects to report what they were spontaneously seeing when viewing our stimuli, in the absence of instruction regarding a specific level of information. In the literature on conscious vision, spontaneous reports and, more generally, the reliability of measures relying on introspection have been heavily questioned (Dehaene et al., 2006; Cohen and Dennett, 2011) due to massive discrepancies between spontaneous reports and performance in instructed tasks, in seminal paradigms, such as the change blindness paradigm (Simons and Levin, 1997; O'Regan and Noë, 2001) and the partial report paradigm (Sperling, 1960).

Here, we show that spontaneous reports have a preferential global content, thereby implying that those reports need to be compared with performance about global information. In the aforementioned paradigms, performance was probed on local elements only; this could explain why the poorness of performance did not match the visual richness described in the spontaneous reports (Campana and Tallon-Baudry, 2013; Haberman et al., 2015).

This proposal is strengthened by the fact that the positive correlation we found here between spontaneous reports and performance holds for global information only and not for local information. Our results resolve an inconsistency in the literature and suggest that the feeling of visual richness inherent to conscious vision might arise from its default high-level global, integrated content (shapes, categories) (Campana and Tallon-Baudry, 2013).

Together, the correlation we found between reports in the spontaneous and in the instructed conditions, as well the consistency of the spontaneous reports between subjects (all subjects but one majoritarily reported the global information), indicate that spontaneous reports can be reliable and should not be discredited. Nevertheless, it is undeniable that truly spontaneous reports, where participants freely describe verbally what they see (e.g., Fei-Fei et al., 2007), make it difficult to extract quantitative information. The paradigm we developed to probe spontaneous vision avoids this caveat by engaging subjects in a very controlled task. By providing subjects with questions about the orientations, we limited the parameter space of spontaneous vision while allowing subjects to skip a trial if their percept was not adequately described by an orientation. Thus, whereas introspection and objective measures (e.g., discrimination tasks) are usually considered in opposition, as if there were a choice to make between them, it appears fruitful to combine them to probe conscious vision in a controlled fashion, especially its spontaneous content.

In conclusion, through a new kind of stimuli, we provide the first direct evidence that conscious vision is constrained by

the visual hierarchy, both in the presence and in the absence of task demands. This result is in agreement with the reverse hierarchy theory (Hochstein and Ahissar, 2002) but contradicts the idea that information at all levels of the visual hierarchy is equally likely to be accessed (Kouider et al., 2010). Whereas most of the studies rely on goal-directed tasks, here we show that spontaneous vision is also accessible to empirical investigation, and we propose a concrete methodology to probe it reliably. Spontaneous reports appear tightly linked to objective performance at the global level. This offers a simple explanation to the apparent mismatch between spontaneous reports and performance in the change blindness and Sperling paradigm, where objective performance is probed at a local level.

References

- Alvarez GA, Oliva A (2008) The representation of simple ensemble visual features outside the focus of attention. *Psychol Sci* 19:392–398. [CrossRef Medline](#)
- Badcock JC, Whitworth FA, Badcock DR, Lovegrove WJ (1990) Low-frequency filtering and the processing of local-global stimuli. *Perception* 19:617–629. [CrossRef Medline](#)
- Bar M (2003) A cortical mechanism for triggering top-down facilitation in visual object recognition. *J Cogn Neurosci* 15:600–609. [CrossRef Medline](#)
- Boehler CN, Schoenfeld MA, Heinze HJ, Hopf JM (2008) Rapid recurrent processing gates awareness in primary visual cortex. *Proc Natl Acad Sci U S A* 105:8742–8747. [CrossRef Medline](#)
- Bronfman ZZ, Brezis N, Jacobson H, Usher M (2014) We see more than we can report: “cost free” color phenomenality outside focal attention. *Psychol Sci* 25:1394–1403. [CrossRef Medline](#)
- Buffalo EA, Fries P, Landman R, Liang H, Desimone R (2010) A backward progression of attentional effects in the ventral stream. *Proc Natl Acad Sci U S A* 107:361–365. [CrossRef Medline](#)
- Bullier J (2001) Integrated model of visual processing. *Brain Res Rev* 36:96–107. [CrossRef Medline](#)
- Campana F, Tallon-Baudry C (2013) Anchoring visual subjective experience in a neural model: the coarse vividness hypothesis. *Neuropsychologia* 51:1050–1060. [CrossRef Medline](#)
- Chong SC, Treisman A (2005) Statistical processing: computing the average size in perceptual groups. *Vision Res* 45:891–900. [CrossRef Medline](#)
- Cohen MA, Dennett DC (2011) Consciousness cannot be separated from function. *Trends Cogn Sci* 15:358–364. [CrossRef Medline](#)
- Dehaene S, Changeux JP, Naccache L, Sackur J, Sergent C (2006) Conscious, preconscious, and subliminal processing: a testable taxonomy. *Trends Cogn Sci* 10:204–211. [CrossRef Medline](#)
- Dougherty RF, Koch VM, Brewer AA, Fischer B, Modersitzki J, Wandell BA (2003) Visual field representations and locations of visual areas V1/2/3 in human visual cortex. *J Vis* 3:586–598. [CrossRef Medline](#)
- Fei-Fei L, Iyer A, Koch C, Perona P (2007) What do we perceive in a glance of a real-world scene? *J Vis* 7:10. [CrossRef Medline](#)
- Flevaris AV, Martínez A, Hillyard SA (2014) Attending to global versus local stimulus features modulates neural processing of low versus high spatial frequencies: an analysis with event-related brain potentials. *Front Psychol* 5:1–11. [CrossRef Medline](#)
- Haberman J, Brady TF, Alvarez GA (2015) Individual differences in ensemble perception reveal multiple, independent levels of ensemble representation. *J Exp Psychol* 144:432. [CrossRef Medline](#)
- Haberman J, Whitney D (2011) Efficient summary statistical representation when change localization fails. *Psychonom Bull Rev* 18:855–859. [CrossRef Medline](#)
- Haberman J, Whitney D (2012) Ensemble perception: summarizing the scene and broadening the limits of visual processing. In: *From perception to consciousness: searching with Anne Treisman* (Wolfe J, Robertson L, eds). Oxford: Oxford UP.
- Han S, Yund EW, Woods DL (2003) An ERP study of the global precedence effect: the role of spatial frequency. *Clin Neurophysiol* 114:1850–1865. [CrossRef Medline](#)
- Hochstein S, Ahissar M (2002) View from the top: hierarchies and reverse hierarchies in the visual system. *Neuron* 36:791–804. [CrossRef Medline](#)
- Hughes HC, Fendrich R, Reuter-Lorenz PA (1990) Global versus local processing in the absence of low spatial frequencies. *J Cogn Neurosci* 2:272–282. [CrossRef Medline](#)
- Jolij J, Scholte HS, van Gaal S, Hodgson TL, Lamme VA (2011) Act quickly, decide later: long-latency visual processing underlies perceptual decisions but not reflexive behavior. *J Cogn Neurosci* 23:3734–3745. [CrossRef Medline](#)
- Kimchi R (1992) Primacy of wholistic processing and global/local paradigm: a critical review. *Psychol Bull* 112:24–38. [CrossRef Medline](#)
- Koffka K (1935) *Principles of Gestalt psychology*. New York: Harcourt Brace.
- Koivisto M (2012) Is reentry critical for visual awareness of object presence? *Vision Res* 63:43–49. [CrossRef Medline](#)
- Kouider S, de Gardelle V, Sackur J, Dupoux E (2010) How rich is consciousness? The partial awareness hypothesis. *Trends Cogn Sci* 14:301–307. [CrossRef Medline](#)
- LaGasse LL (1993) Effects of good form and spatial frequency on global precedence. *Percept Psychophys* 53:89–105. [CrossRef Medline](#)
- Lamb MR, Robertson LC (1989) Do response time advantage and interference reflect the order of processing of global-and local-level information? *Percept Psychophys* 46:254–258. [CrossRef Medline](#)
- Lamb MR, Yund EW (1993) The role of spatial frequency in the processing of hierarchically organized stimuli. *Percept Psychophys* 54:773–784. [CrossRef Medline](#)
- Lamb MR, Yund EW (1996) Spatial frequency and attention: effects of level-, target-, and location-repetition on the processing of global and local forms. *Percept Psychophys* 58:363–373. [CrossRef Medline](#)
- Lamme VA, Roelfsema PR (2000) The distinct modes of vision offered by feedforward and recurrent processing. *Trends Neurosci* 23:571–579. [CrossRef Medline](#)
- Landy MS, Bergen JR (1991) Texture segregation and orientation gradient. *Vision Res* 31:679–691. [CrossRef Medline](#)
- Landy MS, Kojima H (2001) Ideal cue combination for localizing texture-defined edges. *J Opt Soc Am A Opt Image Sci Vis* 18:2307–2320. [CrossRef Medline](#)
- Macé M, Joubert OR, Nespoulous JL, Fabre-Thorpe M (2009) The time course of visual categorizations: you spot the animal faster than the bird. *PLoS One* 4:e5927. [CrossRef Medline](#)
- Maris E, Oostenveld R (2007) Nonparametric statistical testing of EEG- and MEG-data. *J Neurosci Methods* 164:177–190. [CrossRef Medline](#)
- Mehta AD, Ulbert I, Schroeder CE (2000) Intermodal selective attention in monkeys: I. Distribution and timing of effects across visual areas. *Cereb Cortex* 10:343–358. [CrossRef Medline](#)
- Morrison DJ, Schyns PG (2001) Usage of spatial scales for the categorization of faces, objects, and scenes. *Psychonom Bull Rev* 8:454–469. [CrossRef Medline](#)
- Navon D (1977) Forest before trees: the precedence of global features in visual perception. *Cogn Psychol* 9:353–383. [CrossRef](#)
- Navon D (1981) The forest revisited: more on global precedence. *Psychol Res* 43:1–32. [CrossRef](#)
- Oliva A (2005) Gist of the scene. *Neurobiol Attention* 696:251–258.
- Oliva A, Schyns PG (1997) Coarse blobs or fine edges? Evidence that information diagnosticity changes the perception of complex visual stimuli. *Cogn Psychol* 34:72–107. [CrossRef Medline](#)
- Oostenveld R, Fries P, Maris E, Schoffelen JM (2010) FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput Intell Neurosci* 2011:156869. [CrossRef Medline](#)
- O’Regan JK, Noë A (2001) A sensorimotor account of vision and visual consciousness. *Behav Brain Sci* 24:939–973. [CrossRef](#)
- Parkes L, Lund J, Angelucci A, Solomon JA, Morgan M (2001) Compulsory averaging of crowded orientation signals in human vision. *Nat Neurosci* 4:739–744. [CrossRef Medline](#)
- Peyrin C, Michel CM, Schwartz S, Thut G, Seghier M, Landis T, Marendaz C, Vuilleumier P (2010) The neural substrates and timing of top-down processes during coarse-to-fine categorization of visual scenes: a combined fMRI and ERP study. *J Cogn Neurosci* 22:2768–2780. [CrossRef Medline](#)
- Peyrin C, Musel B (2012) On the specific role of the occipital cortex in scene perception. In: *Visual cortex: current status and perspectives* (Molotchnikoff S, Rouat J, eds), pp 61–82. Rijeka: Intech.
- Poort J, Raudies F, Wannig A, Lamme VA, Neumann H, Roelfsema PR

- (2012) The role of attention in figure-ground segregation in areas V1 and V4 of the visual cortex. *Neuron* 75:143–156. [CrossRef Medline](#)
- Roelfsema PR (2006) Cortical algorithms for perceptual grouping. *Annu Rev Neurosci* 29:203–227. [CrossRef Medline](#)
- Schyns PG, Oliva A (1997) Flexible, diagnosticity-driven, rather than fixed, perceptually determined scale selection in scene and face recognition. *Perception* 26:1027–1038. [CrossRef Medline](#)
- Schyns PG, Oliva A (1994) From blobs to boundary edges: evidence for time- and spatial-scale-dependent scene recognition. *Psychol Sci* 5:195–200. [CrossRef](#)
- Schyns PG, Oliva A (1999) Dr. Angry and Mr. Smile: when categorization flexibly modifies the perception of faces in rapid visual presentations. *Cognition* 69:243–265. [CrossRef Medline](#)
- Shulman GL, Sullivan MA, Gish K, Sakoda WJ (1986) The role of spatial-frequency channels in the perception of local and global structure. *Perception* 15:259–273. [CrossRef Medline](#)
- Shulman GL, Wilson J (1987) Spatial frequency and selective attention to local and global information. *Perception* 16:89–101. [CrossRef Medline](#)
- Simons DJ, Levin DT (1997) Change blindness. *Trends Cogn Sci* 1:261–267. [CrossRef Medline](#)
- Sperling G (1960) The information available in brief visual presentations. *Psychol Monogr* 74:1–29. [CrossRef](#)
- Supér H, Spekreijse H, Lamme VA (2001) Two distinct modes of sensory processing observed in monkey primary visual cortex (V1). *Nature* 4:304–310. [CrossRef Medline](#)
- VanRullen R, Thorpe SJ (2001) The time course of visual processing: from early perception to decision-making. *J Cogn Neurosci* 13:454–461. [CrossRef Medline](#)
- Wolfson SS, Landy MS (1995) Discrimination of orientation-defined texture edges. *Vision Res* 35:2863–2877. [CrossRef Medline](#)