The quality of perception without attention

Vandenbroucke, A.R.E.

Seeing without knowing: neural signatures of perceptual inference in the absence of report.

CHAPTER 6

Based on: Vandenbroucke, A. R. E., Fahrenfort, J. J., Sligte, I. G., & Lamme, V. A. F. Seeing without knowing: neural signatures of perceptual inference in the absence of report. *In revision, Journal of Cognitive Neuroscience*.

Abstract

Every day, we experience a rich and complex visual world. Our brain constantly translates meaningless fragmented input into coherent objects and scenes. However, our attentional capabilities are limited and we can only report the few items that we happen to attend to. So what happens to items that are not cognitively accessed? Do these remain fragmentary and meaningless? Or are they processed up to a level where perceptual inferences take place about image composition? To investigate this, we recorded brain activity using functional Magnetic Resonance Imaging (fMRI) while participants viewed images containing a Kanizsa figure; an illusion in which an object is perceived by means of perceptual inference. Subjects were presented with the Kanizsa figure and three matched non-illusory control figures while they were engaged in an attentionally demanding distractor task. After the task, one group of subjects was unable to identify the Kanizsa figure in a forced choice decision task; hence they were 'inattentionally blind'. A second group had no trouble identifying the Kanizsa figure. Interestingly, the neural signature that was unique to the processing of the Kanizsa figure was present in both groups. Moreover, within-subject multi-voxel pattern analysis showed that the neural signature of unreported Kanizsa figures could be used to classify reported Kanizsa figures, and that this cross-report classification worked better for the Kanizsa condition than for all control conditions. Together, these results suggest that stimuli that are not cognitively accessed are processed up to levels of perceptual interpretation.

Introduction

Perception does not directly emerge from the physical stimulation of photoreceptor cells in the retina. Rather, the brain continuously interprets incoming information to make sense of it: through perceptual inference, visual input is translated from meaningless fragmented input into bound objects and scenes. For example, when we see a pen lying on top of a paper, we do not perceive the paper as having a pen-shaped hole in it. Instead, the paper is filled in underneath the pen and we perceive the paper as an uninterrupted rectangle. In this study, we investigated whether this type of inference depends on the ability to attend to and cognitively access visual percepts. When a part of the visual field is neither attended nor reported, does vision represent its constituent parts as consisting of bound and completed objects? Or do they remain fragmentary and meaningless? The answer to this question has important implications for understanding the nature of vision, and may ultimately change our view on conscious perception.

A prime example of perceptual inference is the Kanizsa illusion (Kanizsa, 1976), in which a set of inducers is aligned in such a way that observers perceive an occluding surface lying on top of black disks (Fig. 6.1A). This occluding object is defined by illusory contours and by the illusory contrast difference between surface and background. The illusory contours and illusory contrast difference do not emerge when the inducers are not properly aligned (Fig. 6.1B and D), or when the inducers are not likely to be completed as occluded objects (Fig. 6.1C). The formation of the illusion involves feedback from higher level visual areas such as the Lateral Occipital Complex (LOC) to lower visual areas V1/V2 (Halgren et al., 2003; Knebel & Murray, 2012; Lee, 2001; Maertens, Pollmann, Hanke, Mildner, & Möller, 2008; Seghier & Vuilleumier, 2006). Moreover, the perceptual nature of the Kanizsa figure has been shown to depend on activation in these regions in a reverse-hierarchical manner (Wokke et al., 2013). These studies suggest that the inference mechanisms at play in the Kanizsa illusion depend on interactions between functionally divergent visual areas.

A remarkable observation is that when the inducers of the figure are rendered invisible by continuous flash suppression, the illusion is not perceived, even when attended, showing that perception of the Kanizsa illusion requires conscious processing of the inducers (Harris et al., 2011). This is in contrast with the simultaneous brightness illusion - an illusion of a white disc seeming brighter when presented on a black background than on a grey background – that persists even when the background is not perceived due to flash suppression. Perceptual inference in the Kanizsa illusion thus occurs at a higher level of processing, and is susceptible to manipulations that selectively interfere with conscious perception. Studying to what extent the processes underlying the Kanizsa illusion require conscious access or reportability is therefore of direct relevance to the question whether access is necessary for the formation of a full perceptual representation.

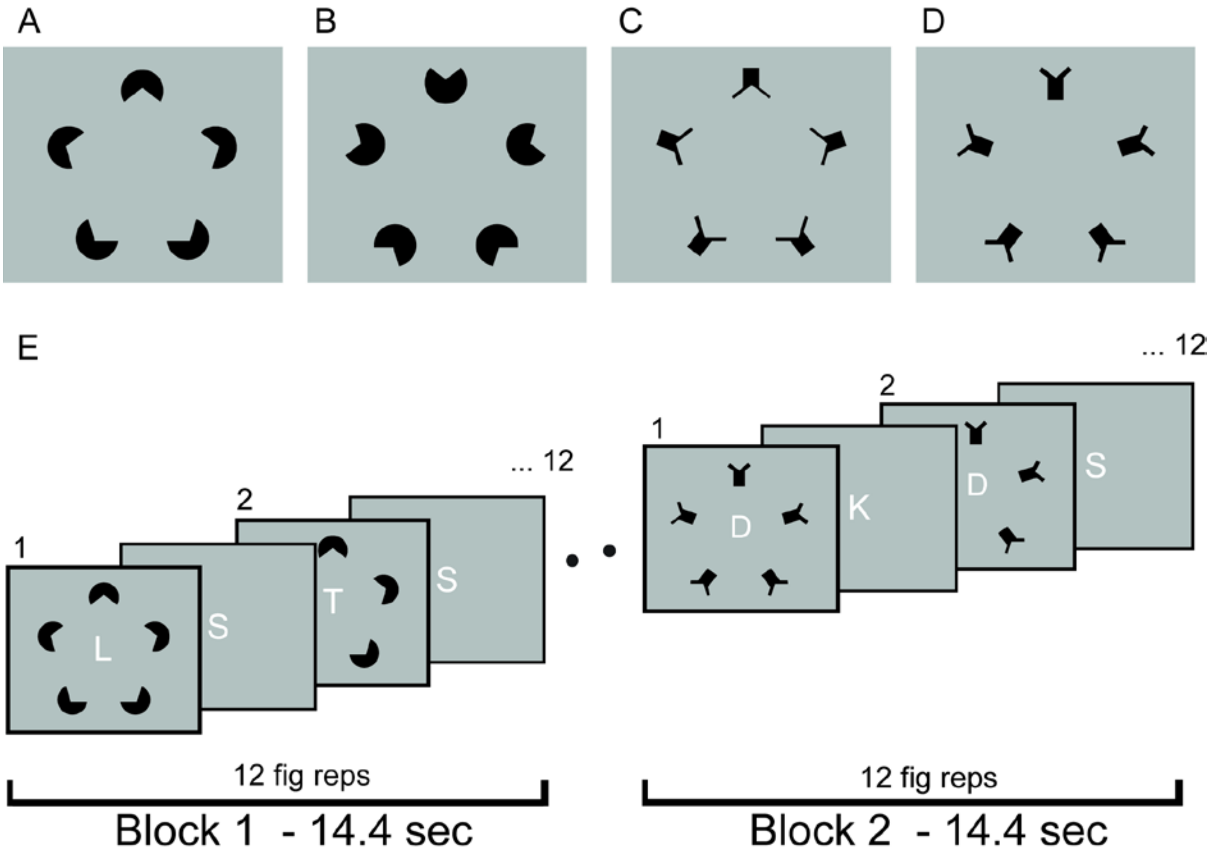

Figure 6.1. Stimuli and task design. *The four stimuli used in the experiment were a (A) Kanizsa figure, (B) Kanizsa control figure, (C) Line figure (same spatial layout and center of gravity as Kanizsa figure), (D) Line control figure. (E) Subjects performed a 2-back task in which they detected the repetition of a letter that was presented 2 serial positions before. Letters were presented for 600 ms each and every 1200 ms, one of the four figures was presented around the letters for 400 ms. Each figure was presented 12 times in a row, resulting in a blocked design with blocks lasting 14.4 seconds.*

To investigate whether the neural correlates of perceiving a Kanizsa figure are present when subjects do not cognitively access the figure, we combined an *inattentional blindness* paradigm (Rees et al., 1999; Rensink et al., 1997; Scholte et al., 2006) with fMRI measurements. Inattentional blindness occurs when a subject is attentionally engaged in another task, rendering a non-target stimulus unnoticed and unreported even when explicitly asked about it. When subjects are informed about the presence of a certain stimulus during the task, they have no trouble seeing the stimulus, even when engaged in the distractor task. This suggests that their inability to report these stimuli occurs because they simply did not access them, not because of perceptual load (Lavie, 2005; Yi, Woodman, Widders, Marois, & Chun, 2004). The paradigm of inattentional blindness formalizes the common intuition that many stimuli in plain sight remain unnoticed and are therefore never accessed, even though they are potentially accessible.

In an fMRI scanner, subjects performed an attentionally demanding 2-back letter task, while the Kanizsa and three control figures (Fig. 6.1A-D) were presented surrounding the letters (Fig. 6.1E). Subjects were instructed that black 'distractor' stimuli would be flashed around the letters, but that they should focus on the letter task to maximize their score on the 2-back task. After three runs of the task, subjects were unexpectedly asked whether they had seen which figure was presented surrounding the letters. The subjects that were unable to select the correct option were labeled as Inattentionally Blind (IB). If subjects selected the Kanizsa figure, they were labeled as Not Inattentionally Blind (NIB). We employed univariate and multivariate analysis techniques to compare the Kanizsa illusion to the three control figures. This allowed us to determine a neural signature that is unique to the illusion. The presence of a unique signature for the Kanizsa illusion for both IB and NIB subjects would indicate that access is not required to process the illusion. If however, only NIB subjects show a unique signature of the illusion, this would suggest that processing the illusion requires access mechanisms.

Methods

fMRI acquisition

Forty-two students (three male) from the University of Amsterdam participated in the experiment for course credit or a monetary reward. The experiment was approved by the Local Ethics Committee and subjects gave their written informed consent. All subjects had normal or corrected-to-normal vision and were screened on the possibility of metal in their bodies and other risk factors precluding participation in MRI studies. Scanning was performed on a 3T Philips Achieva MRI scanner at the Spinoza Center in Amsterdam. A high-resolution T1-weighted anatomical image (TR, 8.21 ms; TE, 3.81 ms; FOV, 256 × 256 × 160) was recorded for each subject. Functional MRI was recorded using a gradient-echo, echo-planar pulse sequence (TR, 2000 ms; TE, 27.63 ms; FA, 90°; 27 slices with interleaved acquisition; voxel size, 2 x 2 x 3 mm; 96 × 96 matrix; FOV, 89 x 89 x 192) centered around the calcarine sulcus. Stimuli were back-projected on a 61 x 36 cm LCD screen using Presentation software (Neurobehavioral Systems, Inc., Albany, CA) and viewed through a mirror attached to the head coil.

Stimuli

To isolate the neural signature associated with the Kanizsa figure, the figure was compared to three control figures (Fig. 6.1B-D). In addition to the classical control figure in which the inducers are rotated (Fig. 6.1B; Mendola, Dale, Fischl, Liu, & Tootell, 1999), we used two additional figures (Fig. 6.1C-D) to control for potential confounds. The Kanizsa illusion has both cognitive and perceptual attributes when compared to its traditional control figure. One can cognitively infer a pentagon from the layout of the Kanizsa figure by connecting the lines between the 'pacmen' inducers (similar to for example knowing that a car has moved because the second time you see it, it is in a different location versus actually perceiving the movement of the car; (Pylyshyn, 1999). Importantly, the illusion also has perceptual attributes: the surface of the pentagon seems to be 'real', as the surface of the pentagon is perceived as slightly brighter than the grey background and as lying on top of full disks instead of pacmen. Completion of the pacmen to full disks and the increased brightness are perceptually inferred (Kanizsa, 1976; Nakayama & Shimojo, 1992; Nakayama, He, & Shimojo, 1995). Using the traditional control figure in which the inducers are rotated outwards, does not allow one to tease apart the perceptual aspects of the illusion (completion of the disks, illusory contours, and the brightness illusion) from its cognitive aspects (presence or absence of a pentagon). Therefore, we devised two additional control figures that isolate the cognitive attributes of the illusion (Figure 6.1C-D) using 'hats' instead of pacman inducers. In these control figures a pentagon shape is either present or absent, yet only by cognitive inference and without the perceptual inferences that characterize the Kanizsa illusion.

The additional control figures also allowed us to circumvent other confounds in the traditional control condition. Particularly, it could be argued that the Kanizsa with rotated inducers is not a perfect control: inwardly rotated inducers create an image with a different low spatial frequency content compared to an image with outwardly rotated inducers (there is a larger 'white gap' in the former; see (Davis & Driver, 1998). If this is an important factor driving the neural signals we record, a difference found between the Kanizsa figure and its outward rotated counterpart should also be found between the line figure and its outward rotated control that contain a similar spatial frequency difference. Together, the three controls (Fig. 6.1B-D) allowed us to assess the unique influence of the illusory nature of the Kanizsa (Fig. 6.1A) on cortical processing when compared to more cognitive or low-level influences on stimulus processing.

The four figures (1 Kanizsa, 3 control) consisted of five black inducers placed on a grey background (Fig. 6.1A-D). The Kanizsa figure (Fig. 6.1A, total span 12.6˚x 12.7˚) and its control

(Fig. 6.1B, total span 13.1˚ x 13.2˚) consisted of pacman-like inducers; five black circles (diameter 2.6˚) with a gap taken out (width 2.1˚, angle 78˚). The Line figure (Fig. 6.1C, total span 12.6˚ x 12.5˚) and its control (Fig. 6.1D, total span 13.4˚ x 13.2˚) consisted of hat-like inducers (2.3˚ x 2.4˚): line elements with a rectangle on top that had the same gap angle. In the Kanizsa figure the inducers were aligned in such a way that a pentagon could be inferred lying on top of black discs. The support ratio (the length of the real contours relative to the illusory contour) of the Kanizsa figure is strongly related to the perception of the illusion. In this study, we used a support ratio of 0.42, which is sufficient to produce the illusion (Seghier & Vuilleumier, 2006; Wokke et al., 2013). In the classical Kanizsa control condition the inducers were rotated 180˚ around their center of gravity - that was calculated by taking the pixel for which the amount of black pixels surrounding it was equal above and below, left and right. The Line figure, which constituted the second control figure, resembled the Kanizsa figure in its spatial layout and its inducers had the same center of gravity as the Kanizsa inducers. In this figure a pentagon could be inferred, however, the illusion of contours and a contrast difference between surface and background was not present. The Line figure was compared to a third control figure, which contained the same elements as the Line figure, but now rotated 180˚ around their center of gravity, just as the Kanizsa control figure.

Behavioral task

Subjects performed a 2-back task on letters that were presented in a Rapid Serial Visual Presentation (RSVP). They were instructed to press a button when the same letter was presented as two serial positions before. Letters (0.5˚) were presented in the center of the screen for 600 ms each. In every sequence of 8 letters a repetition occurred (jittered between location 3 and 8) and a total of 78 sequences was presented. At the end of each run, subjects received feedback about the percentage correctly detected targets.

Procedure

Subjects were informed that they were participating in a study on the ability to perform a memory task while visually distracted. Before the start of the MRI session, they practiced the behavioral task for blocks of 2 minutes until they reached a performance of at least 80%. Then, they performed a block of around 6 minutes, the same length as in the MRI scanner. During this block, distracter stimuli were presented surrounding the letters. These distracter stimuli were similar to those used in the actual experiment, consisting of black stimuli (rectangles and half circles) at the same position and changing every 14.4 seconds. However, they did not form a

Kanizsa figure and were intended to get subjects used to the flashing stimuli while focusing on the letter task.

Each functional run started with 10 seconds fixation. Subsequently, two letter sequences (9.6 seconds) were presented. After two sequences, the first surrounding figure was presented for four sequences (18.8 seconds). It was expected that the first presentation of a figure would elicit a heightened activity due to its sudden onset regardless of figure type. We therefore presented a figure that was different from the four experimental figures at the start of each run (consisting of five half circles presented at the same position as the inducers of the four experimental figures). This way, none of the experimental figures had the advantage of being the first figure that was presented. After the presentation of the start figure, the four experimental figures were presented surrounding the letters in blocks of 14.4 seconds (Figure 6.1E). In each block, one of the four figures was flashed around the letters with a duration of 400 ms and an ISI of 800 ms, making a total of 12 presentations per block. There was no rest between blocks and the same figure was never repeated. The blocks were counterbalanced such that each figure was followed by one of the other figures for an equal amount of times. In each run, 6 blocks of each stimulus were presented, resulting in a total of 24 blocks per run. Each run ended with a 16 seconds rest period, making the total runtime 400 seconds (200 volumes).

After three runs in the scanner, subjects were presented with a surprise question. In this question, subjects were informed that the black stimuli surrounding the letters had formed figures, and that they should choose the figure they thought had been presented during the three runs. After reading this question, they received 8 options (Figure 6.2) of which only one contained the illusory Kanizsa figure that was presented during the experimental runs. They were asked to choose one of these options, even if they had to guess. We embedded two other Kanizsa type figures (Fig. 6.2B and C) to prevent subjects guessing the Kanizsa figure because it was the figure that popped out compared to the other figures. All subjects that selected the correct figure were categorized as Not Inattentionally Blind (NIB), as they might have either explicit or implicit knowledge or familiarity with the figure, even if they felt they were guessing. All subjects that selected an incorrect figure were labeled Inattentionally Blind (IB). After subjects answered the surprise question, the correct answer was not given to them. Instead, they were asked to perform the exact same run again and while performing the letter task, to try to detect which out of the 8 options was shown. The control run was identical to the experimental runs. After the control run, subjects were asked again to identify the correct figure from the same 8 options, after which the correct answer was revealed. Only participants that had answered the second question correctly were included in further analyses, ensuring their ability to perform the letter task and detect the Kanizsa figure at the same time.

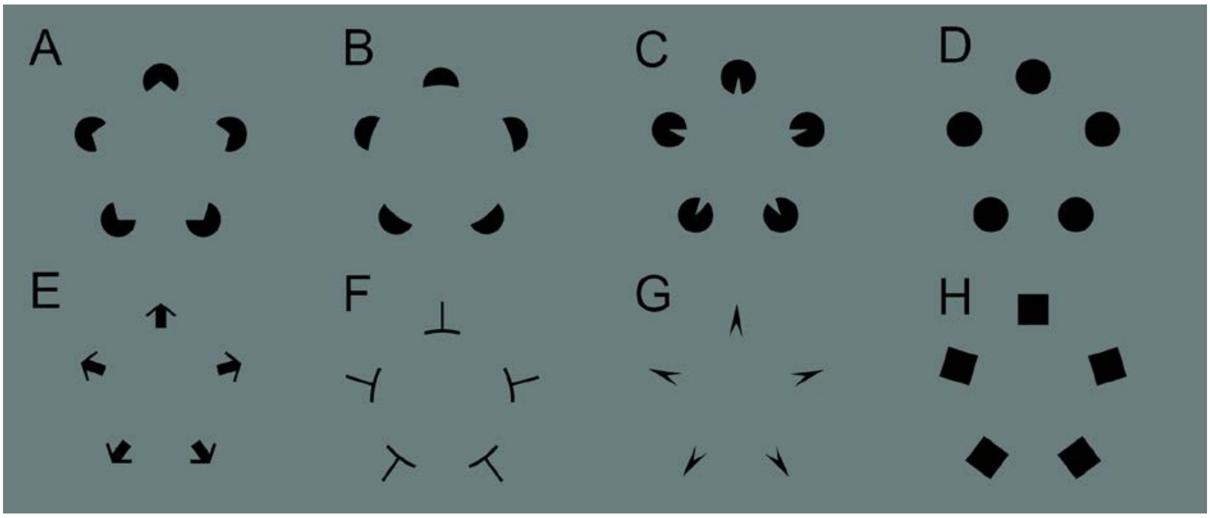

Figure 6.2. Multiple choice options. *The 8 options presented in the Multiple Choice question after Experimental runs 1-3 and the control run. For the Not Inattentionally Blind group, option A was chosen by all subjects after run 1-3 and the control run. For the Inattentionally Blind group, option A was never chosen after run 1-3. Instead, option B was chosen 5 times, option C was chosen 2 times, option F was chosen 2 times, option H was chosen 3 times and option D, E and G were never chosen. After the control run, all subjects chose option A.*

Region of Interest localization

For each subject, V1, V2, V3, V3AB and V4 (Fig. 6.3) were localized using a polar angle mapping, an eccentricity mapping and a study specific localizer. For polar angle mapping, a checkerboard (red-green, flickering at 8 Hz) wedge rotated around fixation (complete revolution in 30s; 8 repetitions) and for eccentricity mapping, a checkerboard ring (red-green, flickering at 8 Hz) expanded from center to periphery (complete revolution in 30s; 8 repetitions). During these two runs, subjects fixated at the center while detecting blue squares presented in the red-green checkerboard stimuli. The TR of these two runs was set to 2500 ms as the phase of the wedge and expanding ring was set at 2500 ms (6 phases resulted in one cycle of 15 seconds). In addition to the retinotopic mapping, a study specific localizer was used in which a black circle (diameter 2.6˚; flickering at 2 Hz in 16 second blocks) presented in the center of the screen was alternated with five black circles presented at the inducer positions (total figure 12.7˚ x 12.7˚, flickering at 2 Hz in 16 second blocks). These two conditions were separated by a 16 second rest period and were repeated 5 times. Throughout the run, subjects maintained fixation and performed a fixation task in which they had to detect a rotation of the fixation cross. Data of these runs were projected onto an inflated surface reconstruction and V1, V2, V3, V3AB and V4 were defined for each subject.

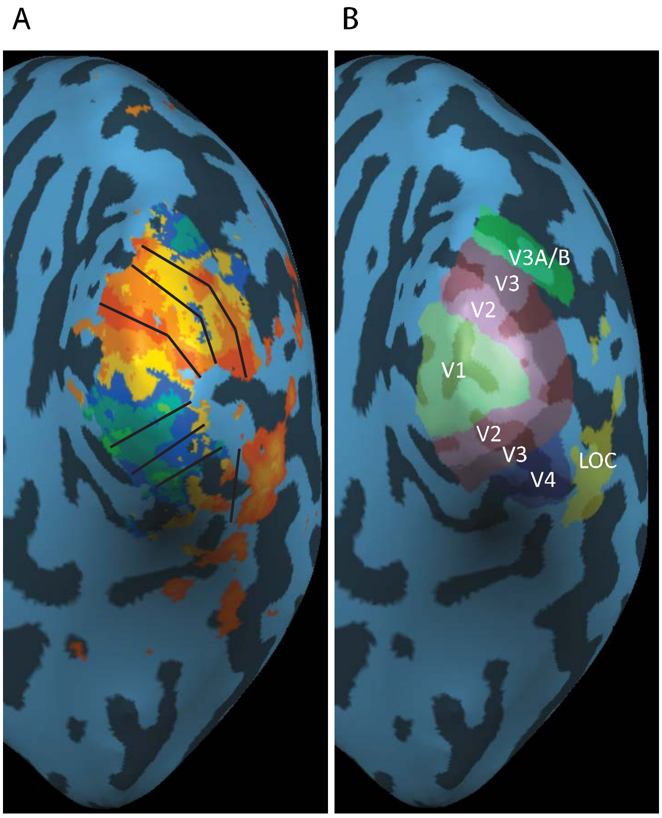

Figure 6.3. Regions of Interest. *Each individuals' cortex was inflated and a retinotopic map (A, depicted here on a representative subject), eccentricity map, study specific localizer and a mapping for Lateral Occipital Complex (also shown here) were projected onto the inflated surface. On the basis of these mappings, six Regions of Interest (ROIs: V1, V2, V3, V3A/B, V4, LOC) were identified for each participant (B, depicted here on the representative subject in A).*

In addition to these lower visual areas, LOC was localized using a mapper in which blocks of houses, faces, objects (chairs, scissors, bottles) and phase scrambled versions of these objects were presented (Scholte, Jolij, Fahrenfort, & Lamme, 2008). Each block lasted 16 seconds (8 presentations of 1000 ms per block) with a rest of 12 seconds between each block. Each stimulus category was repeated 4 times. Activity for the contrast between objects and scrambled pictures was mapped on an inflated surface reconstruction and LOC was defined for each subject. All localizer runs were performed during the same session as the experimental runs.

Univariate fMRI analysis

Data were analyzed using Brainvoyager 2.1 (Brain Innovation, Maastricht, The Netherlands; Goebel, Esposito, & Formisano, 2006) and Matlab 2010 (MathWorks Inc). Functional scans were slice-time corrected, motion corrected, spatially smoothed with a Gaussian of 2 mm FWHM and high-pass filtered using a GLM with Fourier basis set (3 cycles). All functional scans were aligned to the first functional scan, which was coregistered to the T1-weighted anatomical image. Structural images were transformed to Talairach space using an ACPC transform (Talairach & Tournoux, 1988).

A General Linear Model with 5 predictors (4 experimental figures and start figure) was defined for each subject. A whole brain analysis (correcting for multiple comparisons using a False Discovery Rate (FDR) of 0.05) was performed for the two groups separately, combining the three experimental runs (z-transformed) for each subject. For the ROI analyses, the GLM was modeled in each subject and ROI separately for the three experimental runs combined. To test the effect of Figure, ROI and Group, a 4 (Figure: Kanizsa, Kanizsa Control, Lines, Lines Control) x 6 (ROI: V1, V2, V3, V3AB, V4, LOC) x 2 (Group: IB vs NIB) mixed repeated measures ANOVA on the beta-values was performed.

Multivariate fMRI analysis

The same preprocessing steps as described for the univariate analyses were performed. For each block separately, the response for each voxel in each ROI was calculated. This was done by first z-transforming the whole timeseries and then averaging over 6 volumes following the first stimulus presentation in a block, with a 2 volume delay to account for the haemodynamic lag. The response for each block of the three experimental runs was fed into a training algorithm implemented in the Princeton MVPA Toolbox (http://code.google.com/p/princeton-mvpa-toolbox) using the backpropagation algorithm of the Netlab Neural Network Toolbox (http://www1.aston.ac.uk/eas/research/groups/ncrg/resources/netlab/). This yielded a specific voxel pattern for each figure in each ROI (all voxels in each ROI were used). First, the four figures were classified within experimental runs by training on 2 experimental runs (while subjects were IB) and testing the remaining experimental run in a leave-1-out procedure. Then, the patterns of the blocks in the control run (while subjects no longer suffered IB) were classified based on the three experimental runs together. This yielded a classification score (percentage correct) per subject per ROI within experimental runs and for the control run. Chance performance was calculated for each individual and each training set by shuffling the test labels and calculating the baseline performance on each figure for each specific training set. On average, chance performance was 25 % correct. To investigate whether classification of the Kanizsa figure was better than classification of the three control figures, and whether there were any differences between the two groups, a 4 (Figure: Kanizsa, Kanizsa Control, Lines, Lines Control) x 6 (ROI: V1, V2, V3, V3AB, V4, LOC) x 2 (Group: IB vs. NIB) mixed repeated measures ANOVA was performed.

Results

Behavioral performance

Out of 42 subjects, 5 subjects did not select the correct answer after the control run and these were excluded from all analyses. Twelve subjects did not select the correct answer after the three experimental runs, but did select the correct answer after the control run (IB group). From the subjects that were correct on both questions, twelve subjects (NIB group) were matched to the IB group on age (IB: 20.3 yrs, NIB: 20.2 yrs), gender (all female) and overall performance on the letter task (IB: 88%, NIB: 83%, $F(1,22) = 1.345$, $p = .259$). There was a main effect of task, showing that n-back performance during the experimental runs was higher than performance during the control run (87 % vs 83 %; $F(1,22) = 6.703$, $p = .017$), but this effect was the same for both groups ($F(1,22) = .016$, $p = .900$).

Univariate fMRI analysis

We examined visual areas that are known to be involved in generating the Kanizsa illusion (for an overview, see Seghier & Vuilleumier, 2006). We defined Regions of Interests (ROIs) for V1, V2, V3, V3AB, V4 and LOC based on retinotopic and object-specific localizers (see Fig. 6.3 and Methods). For each of these regions, a GLM was fitted to determine the activity corresponding to each condition (Fig. 6.4). We compared the mean activity for each figure and the activity for the two groups by performing a 4 (Figure: Kanizsa, Kanizsa Control, Lines, Lines Control) x 6 (ROI: V1, V2, V3, V3AB, V4, LOC) x 2 (Group: NIB vs. IB) mixed repeated measures ANOVA. A significant effect of figure was found, $F(3,66) = 21.5$, $p < .001$, showing that the Kanizsa illusion resulted in stronger activations of both lower- and higher-tier visual areas compared to the three control figures, post-hoc t-tests: all $p < .001$, Bonferonni corrected. Moreover, the amount of activity for the Line figure and its control did not significantly differ, post-hoc t-test: $p > .999$, Bonferonni corrected, suggesting that these two figures could not be dissociated based on activity in these visual areas. This shows that even though cognitive inference is possible for the Line figure (i.e. a pentagon can be inferred because logically, it is the only possible configuration), there is no specific visual neural signature that accompanies the figure. This confirms that there is no perceptual inference for the Line figure and the heightened activity associated with the Kanizsa illusion was due to the perceptual characteristics of the illusion, and not due to other visual properties such as layout, spatial frequency, collinear contours, or the 'cognitive binding' of the inducers into a single object.

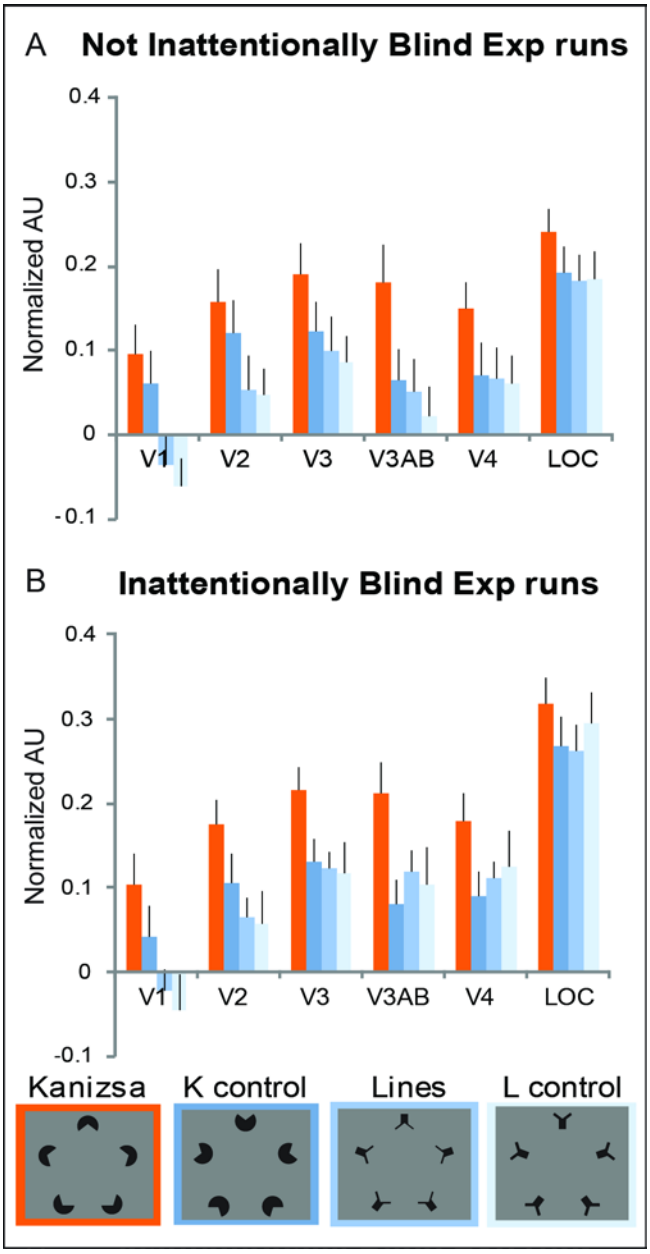

Figure 6.4. Univariate results. *Activity (in normalized arbitrary units) associated with the four figures in each ROI for the Not Inattentionally Blind - NIB - group (A) and the Inattentionally Blind – IB - group (B). In both groups, the Kanizsa figure elicited more activity across lower and higher visual areas. Error bars denote standard error.*

The Kanizsa illusion elicited heightened activity across all of visual cortex, regardless of whether subjects reported the figure. Moreover, the lack of an interaction effect showed that the pattern of activation for the four figures was the same for the NIB and the IB group, $F(3,66) = .64$, $p = .59$. This suggests that the Kanizsa illusion is processed even when the percept is not accessed or reported. There were 7 subjects in the IB group that chose one of the two Kanizsa-type foils in the experimental question (Fig. 6.2). Although these foils have a different perceptual appearance than the target Kanizsa, it could be that these subjects were able to report that they saw a Kanizsa-type figure, but not able to report the details of the figure they saw. To investigate whether the results of the IB group were determined by these subjects, we analyzed the 5 subjects that chose option D-H to see whether the pattern of results was the same. Although statistical testing with 5 subjects yields too little power, the pattern for these 5 subjects was the same as the pattern for the whole group (Fig. 6.5).

In addition to a main effect of Figure, there was a main effect of ROI, $F(5,110) = 32.2$, $P < .001$, Greenhouse-Geisser corrected, and an interaction between Figure and ROI, $F(15,330) = 23.8$, $p < .001$, Greenhouse-Geisser corrected. The main effect of ROI was driven by the fact that overall activity in V1 was lower and activity in LOC was higher than the other ROIs. The interaction effect revealed that the difference between the three figures was largest in V3AB and smaller in LOC. The involvement of V3AB in addition to LOC has been found in previous literature as well (Mendola et al., 1999). However, no interaction effects with Group were found, all $p > .25$, Greenhouse-Geisser corrected, suggesting that the pattern of activity across ROIs was similar for the IB and NIB group.

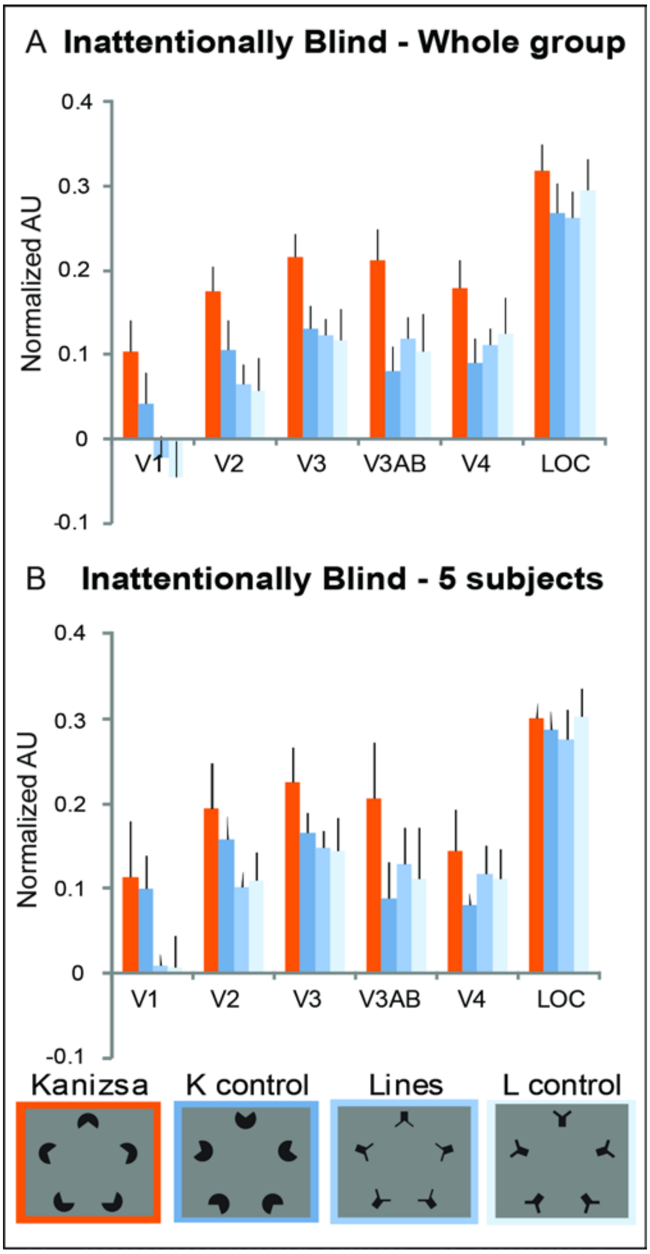

Figure 6.5. Univariate results for the Inattentionally Blind (IB) group. Results for the whole IB group (A) and the 5 subjects that chose a non-Kanizsa-type figure in the experimental question (B, See Figure 6.2). *The activity pattern for the whole group and these 5 subjects is similar, showing that the results are not dominated by the subjects that selected a Kanizsa-type foil.*

To further investigate whether there were any differences between the NIB and the IB group apart from the ROIs specified, a multi-subject whole brain analysis was performed (note that the scans did not cover the front of the brain). Figure 6.6 shows the multi-subject whole brain analyses for the two groups separately (FDR = .05), which show that mainly the lower visual areas are involved, similarly for the NIB and IB group. There were no regions that were significantly more activated for the NIB group for the Kanizsa figures versus the other figures (FDR = .05).

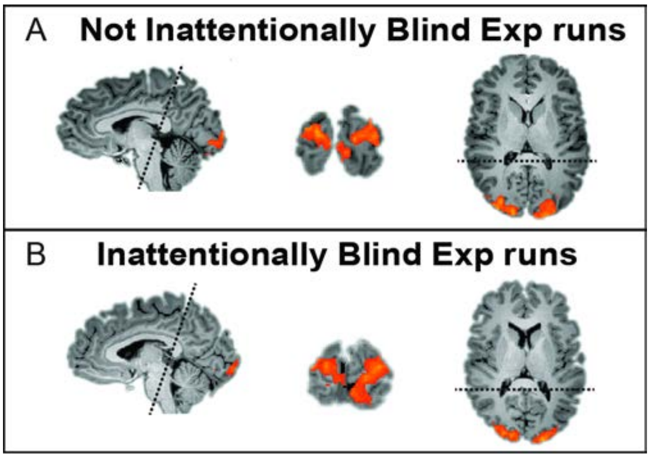

Figure 6.6. Whole brain analysis. *A whole brain analysis was performed for the Not Inattentionally Blind - NIB - (C) and Inattentionally Blind- IB – group (D), in which the Kanizsa figure was contrasted with the three control figures (Kanizsa>Controls). Activity in lower and higher visual areas and does not differ between the two groups. Note that the scans did not cover the whole brain but only recorded visual cortex and extended into a part of the temporal and parietal cortex. Dotted lines indicate upper and lower limits of epi acquisition.*

Multivariate pattern analysis

We hypothesized that because only the Kanizsa figure elicited a perceptually integrated percept, its multi-voxel pattern should be more consistent than the voxel patterns underlying the three control figures and therefore result in higher decoding performance. To test this hypothesis, we used ROI voxel patterns for the four stimulus figures within each subject to classify each of the figures during the experimental runs. We confirmed that classification performance was highest for the Kanizsa condition compared to the control conditions, $F(3,66) = 9.5$, $p < .001$, and there was no interaction between performance for the NIB and IB group, $F(3,66) = .7$, $p = .52$. This shows that the integrated perceptual experience resulting from a Kanizsa figure is reflected in a more consistent multi-voxel representation, regardless of whether subjects were able to report the figure. Next, we wanted to examine whether this more consistent multi-voxel representation for Kanizsa figures persists across runs in which the figure was reported compared to the runs in which it was not reported. To do so, we tested whether the patterns elicited during the experimental condition could be used to classify the patterns elicited in the control condition. We trained a neural pattern classifier on the three experimental

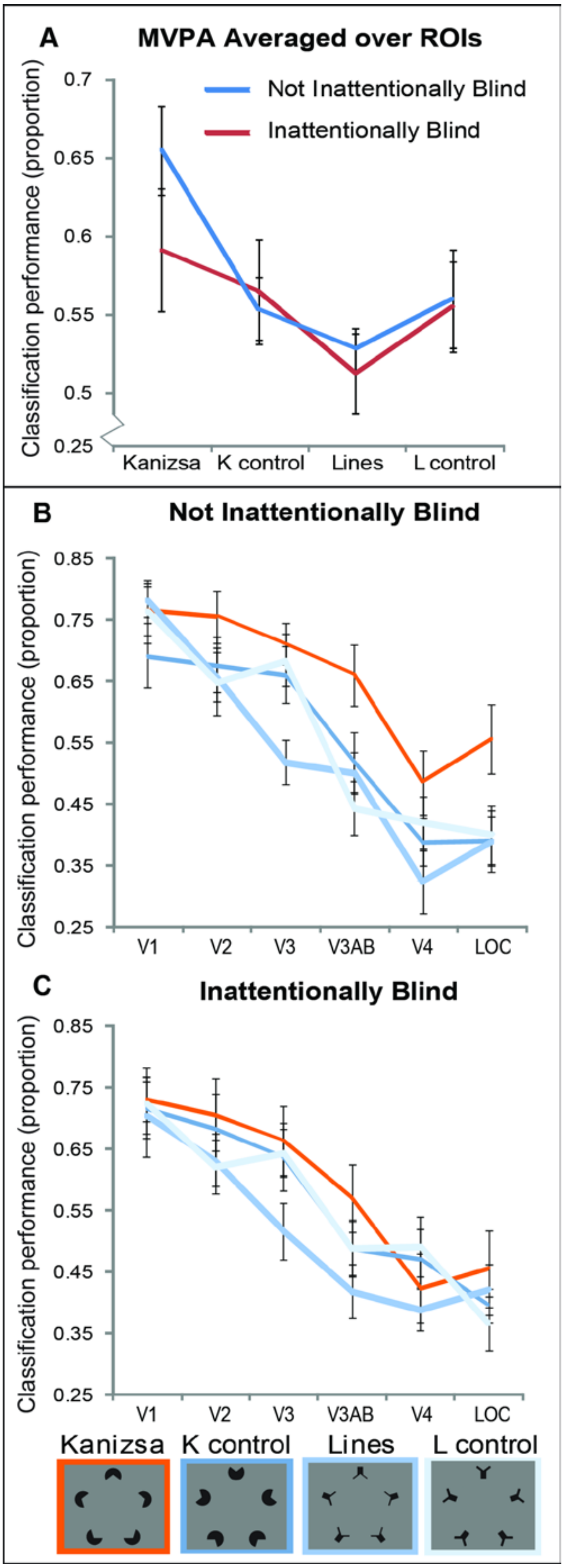
A
MVPA Averaged over ROIs
Classification performance (proportion)
0.7
0.65
0.6
0.55
0.5
0.25
Not Inattentionally Blind
Inattentionally Blind
Kanizsa
K control
Lines
L control
B
Not Inattentionally Blind
Classification performance (proportion)
0.85
0.75
0.65
0.55
0.45
0.35
0.25
V1
V2
V3
V3AB
V4
LOC
C
Inattentionally Blind
Classification performance (proportion)
0.85
0.75
0.65
0.55
0.45
0.35
0.25
V1
V2
V3
V3AB
V4
LOC
Kanizsa
K control
Lines
L control

Figure 6.7. Multivariate results. *Classification performance in percentage correct averaged over all ROIs for the Not Inattentionally Blind (NIB) group and the Inattentionally Blind (IB) group (A) and for each separate ROI (B and C). The classification for Kanizsa figures is better than classification for the three control figures. Average chance performance was 25%, corrected for each figure separately (see Methods). Error bars denote standard error.*

runs and classified the patterns from the control run. If the patterns underlying the Kanizsa illusion remained the same and reportability has no influence on its neural representation, classification between the experimental and control runs is predicted to be better for the Kanizsa figures than for the control figures. If, however, the underlying neural pattern changed due to access to the figure, classification should not work between the experimental and the control runs, resulting in similar or even lower classification performance for the Kanizsa figures than for the control figures.

Pattern classification was obtained for each ROI separately. Classification performance for all ROIS is shown in Figure 6.7 (average chance performance = 25%, see Methods). We were able to determine which figure was presented during the control run based on the patterns resulting from the experimental run, as all figures could be classified well above chance. Importantly, classification worked best for Kanizsa figure, $F(3,66) = 8.1$, $p < .001$; post-hoc t-tests compared to controls, all $p < .059$, Bonferonni corrected. This shows that the multi-voxel pattern underlying the Kanizsa figure is more consistent even across the reported and unreported conditions. There was no interaction between figure type and group, confirming that classification performance for the NIB and the IB group was the same, $F(3,66) = 1.2$, $p = .33$. To test whether the lack of an interaction effect may have been due to a lack of power, we maximized power by averaging classification performance for the three control figures, and tested against the Kanizsa figure. There was still no interaction between Figure and Group $F(1,22) = 3.3$, $p = .09$, although it did show a trend towards significance. Although we cannot exclude the possibility that more subjects would bring out a significant interaction, the current dataset suggests that classification performance is best for the Kanizsa figure, irrespective of whether it was reported or not.

There was a main effect of ROI, $F(5,110) = 43.5$, $p < .001$, driven by overall better classification in areas V1, V2 and V3, and an interaction between ROI and figure, $F(15,330) = 2.7$, $p = .001$, driven by a larger difference in classification between figures for V2, V3AB and LOC , but again, no interaction with Group, all $p > .75$. Taken together, both the univariate and multivariate results suggest that the ability to access or report the Kanizsa figure does not influence the neural signature associated with the formation of the illusory percept.

Discussion

In this study, we examined the influence of access on the perceptual processing of the Kanizsa illusion using an inattentional blindness paradigm (Rees et al., 1999; Scholte et al., 2006; Simons & Chabris, 1999) in combination with fMRI. Although a large group of subjects was not able to select the Kanizsa figure after the experimental runs, these subjects still displayed a unique neural pattern associated with processing a Kanizsa figure. The illusory figure elicited heightened activity in the areas that are critical for the perception of the illusion (Seghier & Vuilleumier, 2006). Also, multivariate pattern analysis showed that the neural signature of the Kanizsa during the unreported state could be used to classify its neural signature during the reported state, and classification performance was better for the Kanizsa figure than for any of the non-illusory conditions. Together these results suggest that access is not necessary for the type of perceptual inference that underlies the Kanizsa illusion to take place in visual cortex.

Mechanisms underlying Kanizsa processing

The Kanizsa illusion is a prime example of perceptual inference, a process that is linked to conscious processing (Fahrenfort et al., 2007; Harris et al., 2011; Wokke et al., 2013). It has been shown that the Kanizsa illusion is not perceived when its inducers are masked (Harris et al., 2011), suggesting that conscious processing of the inducers is necessary to perceive the figure. However, other studies have shown that the Kanizsa illusion survives crowding (Lau & Cheung, 2012) and breaks through interocular suppression more easily (Wang et al., 2012), suggesting that the Kanizsa illusion can be processed unconsciously. Although these findings seem to contradict each other, the perception of the Kanizsa illusion may depend on multiple mechanisms.

To process a Kanizsa figure, its inducers should be grouped and processed as one object. This process might be driven by grouping mechanisms that depend on fast, feedforward activity and could be performed unconsciously (Roelfsema, 2006). Then, the details of the figure – the specific illusory shape that is seen – are filled in by feedback mechanisms. In a recent TMS experiment, it was shown that the critical time window for V1/V2 in which discrimination of Kanizsa figures was affected occurred after the critical time window in which the LOC (Wokke et al., 2013) - an area that sits higher up the visual hierarchy and is involved in object detection (Malach, Levy, & Hasson, 2002) - was involved. Moreover, this effect occurred only when the support ratio of the Kanizsa inducers was large enough to clearly cause an illusory percept. Critically, for all the support ratios that were used, also for those not evoking an illusory percept, the inducers could be grouped. This suggests that V1 is causally involved in the shape formation of the illusion and not in the initial grouping of elements. These findings match with the behavioral findings of Wang et al. (2012) on the one hand and Harris et al. (2011) on the other hand. The Kanizsa figure may break through interocular suppression easier than a control figure: if the grouping of elements occurs unconcsciously, a Kanizsa figure will be seen more easily than a control figure that cannot be grouped. In the study in which the Kanizsa inducers were masked, however, the critical manipulation was for subjects to perceive which direction the illusory triangle was facing, and thus the shape of the figure should be processed. If shape processing depends on feedback interactions, masking should indeed inhibit the formation of the shape (Fahrenfort et al., 2007; Harris et al., 2011). Together, these results suggest that perceiving the Kanizsa illusion depends on unconscious grouping mechanisms and conscious figure formation, which are supported by feedforward and feedback mechanisms respectively.

Although in the current study, we were not able to directly test the involvement of feedforward and feedback mechanisms due to the temporal resolution of the fMRI signal, we found strong modulation in V1 and V2 even though the receptive field sizes in V1 and V2 are an order of 6-12 times too small to encapsulate the entire Kanizsa figure (Smith, Singh, Williams, & Greenlee, 2001). This suggests that feedback from higher areas modulated activity in lower visual areas, thereby suggesting that processes underlying shape formation had occurred and thus the figure was fully processed (Harris et al., 2011; Wokke et al., 2013).

Importantly, the neural correlates underlying Kanizsa figure processing were related to the illusory nature of the percept and not to physical or cognitive stimulus attributes. By using two additional control figures (Fig. 6.1C and D), we were able to dissociate perceptual inference from cognitive inference; even though the Line figure had the same inducer layout and a pentagon could be cognitively inferred, the perception of illusory contours and a contrast difference between figure and background were absent. The absence of this illusory percept was reflected in the neural signature such that the Line figure (Fig. 6.1C) showed the same modulation as the Line control figure (Fig. 6.1D), both in the accessible and unreported states. Moreover, the spatial frequency difference for the Line figure and its control was comparable to the spatial frequency difference for the Kanizsa figure and its control. The similarity between the neural signature for the Line figure and its control confirm that there was no perceptual characteristic added to this figure and that the neural modulation observed for the Kanizsa figure was not due to differences in spatial frequency or collinearity, but due to the process of perceptual inference.

Inattentional blindness and the Kanizsa figure

To create inattentional blindness, we specifically manipulated cognitive load and not perceptual load (Yi et al., 2004). We chose to manipulate cognitive load to test the hypothesis that cognitive access is necessary for perceptual inference. IB participants could not identify the presented figure when they were uninformed about the configuration of the distracting stimuli, while when they were informed about the configuration, they were able to select the correct figure even though they maintained the same performance on the n-back task. This warrants the conclusion that the figure was potentially accessible, yet not accessed during its presentation. It may well be that would we have manipulated perceptual instead of cognitive load, results would have been different. Outcomes of neuronal modeling support the prediction that during inattentional blindness incoming sensory information might be processed, but blocked from access, because the network is engaged in processing distracting information (Dehaene & Changeux, 2005). In that sense, the paradigm of inattentional blindness formalizes the common intuition that many stimuli in plain sight remain unnoticed and are therefore never accessed, even though they are potentially accessible, and forms the most rigorous test of the fate of unaccessed visual stimuli.

In the present study, we found no evidence for a difference in processing of the Kanizsa figure between the IB and the NIB group, and thus no evidence for an influence of cognitive access. Possibly, whether the percept could be reported or not depended on the neural pattern in frontal or fronto-parietal areas (not included in the current EPI sequence; Carmel, Lavie, & Rees, 2006; Lumer & Rees, 1999, although see Thakral, 2011). However, the neural patterns that were unique to the Kanizsa illusion were present in both lower and higher-level visual areas for the IB group. Therefore, even if the difference in reportability depended on activity in more frontal areas, this did not change the neural patterns associated with the perceptual characteristics of illusory contours and an illusory contrast difference as isolated in this study and other studies investigating Kanizsa processing (Seghier & Vuilleumier, 2006). This suggests that access to a stimulus does not alter the neural, and hence perceptual characteristics of that stimulus. It merely makes the representation globally available for cognitive manipulation and report. Possibly, subjects in the NIB group had some left-over attention that they allocated to the distractor stimuli, thereby gaining access to the Kanizsa figure: attention might have amplified the neural signature, making the neural pattern better distinguishable. This might explain why, even though there was no interaction between groups, the classification scores for the Kanizsa figure in the NIB group trended to be higher than those in the IB group. At the same time, this does not weaken or disqualify our main finding that the processes of perceptual inference associated with the Kanizsa illusion are fully present without access (or attention).

A potential alternative explanation for the subjects' behavior is that the inability to select the correct figure was not a result of inattentional blindness, but of inattentional amnesia (Wolfe, 1999). It could be the case that subjects were able to access the figure at the moment of presentation, but a memory failure prevented them from selecting the correct figure when asked about it. However, in comparison to studies where the target figure was presented just once (Simons & Chabris, 1999; Thakral, 2011), in our study the figure was presented 18 times, each for a period of 14.4 seconds. The question about these figures was then asked within 1 minute after the last stimulus presentation. This makes it improbable that the failure to select the correct figure was due to simple memory failure. Moreover, the same subjects were perfectly able to select the figure when their task instruction was to pay attention to the figures; thus, it is not the case that these subjects simply had bad visual memory. The main point, however, is that in both frameworks – inattentional blindness or amnesia - the subject was not able to report or recognize a recently presented figure even when forced to make a decision. One of the functions of cognitive access is storage in working memory, and one could say that access has failed when subjects cannot report about objects that were presented multiple times only a few seconds ago. At the same time, the neural signature coding for the perceptual state accompanying the figure was clearly present. This confirms the importance of using brain measurements when investigating the nature of visual representations instead of relying on behavioral measures of report only (Kanai & Tsuchiya, 2012; Lamme, 2010).

Conclusions

The present study is extends previous work showing that several perceptual processes such as figure ground segregation (Pitts et al., 2012; Scholte et al., 2006), feature grouping (Moore & Egeth, 1997; Pitts et al., 2012) visual context effects (Lathrop, Bridgeman, & Tseng, 2011) occur during inattentional blindness. In this study we show that the neural Kanizsa signature does not subside when subjects are not able to report about their percept. Importantly, the Kanizsa figure is accompanied by a unique signature that is absent when controlling for confounding factors such as collinearity, spatial frequency and cognitive inference. Including these controls confirms that the signature is a correlate of the illusory percept itself, and not of something else. This implies that this type of perceptual inference occurs in the absence of access to the percept, potentially putting non-accessed states in the realm of conscious rather than unconscious processing (Lamme, 2010).

Acknowledgements

This work was made possible by an advanced ERC grant to VAFL. We thank Ned Block for his helpful comments on this manuscript.