Emergence of Perceptual Gestalts in the Human Visual Cortex : The Case of the Configural-Superiority Effect

Jonas Kubilius, Johan Wagemans and Hans P. Op de Beeck

Psychological Science 2011 22: 1296 originally published online 20 September 2011
DOI: 10.1177/0956797611417000

Research Article

Psychological Science
22(10) 1296–1303

DOI: 10.1177/0956797611417000
http://pss.sagepub.com

Emergence of Perceptual Gestalts in the Human Visual Cortex: The Case of the Configural-Superiority Effect

Jonas Kubilius, Johan Wagemans, and Hans P. Op de Beeck
University of Leuven

Abstract

Many Gestalt phenomena have been described in terms of perception of a whole being not equal to the sum of its parts. It is unclear how these phenomena emerge in the brain. We used functional MRI to study the neural basis of the behavioral configural-superiority effect (i.e., visual search is more efficient when an odd element is part of a configuration than when it is presented by itself). We found that searching for the odd element in a display of four line segments (parts) was facilitated by adding two additional line segments to each of them (creating whole shapes). Functional MRI–based decoding of neural responses to the position of the odd element revealed a neural configural-superiority effect in shape-selective regions but not in low-level retinotopic areas, where decoding of parts was more pronounced. These results show how at least some Gestalt phenomena in vision emerge only at the higher stages of visual information processing and suggest that feed-forward processing might be sufficient to produce such phenomena.

Keywords

visual perception, neuroimaging

Received 10/7/10; Revision accepted 5/2/11

A hallmark of visual perception is that the percept of the whole is often qualitatively different from the sum of the parts. There is an impressive list of perceptual manifestations of superadditive Gestalt processes (for reviews, see Albright & Stoner, 2002; Spillmann, 2009; Wagemans, Wichmann, & Op de Beeck, 2005), such as configural superiority, symmetry, closure, and subjective contours. The configural-superiority effect, illustrated in Figure 1, is a particularly strong example of this superadditivity (Pomerantz, Sager, & Stoever, 1977). Despite the great appeal of these phenomena as demonstrations of the strength of perceptual organization, little is known about how they arise in the brain.

How Gestalts emerge was intensively debated even in the early days of Gestalt psychology (see Ash, 1995). The Graz school (e.g., von Ehrenfels, Meinong, Benussi) considered *Gestalten* as qualities that are dependent on objects but that have content beyond those objects (i.e., mental constructions with a specific phenomenological value added to the stimulus). In contrast, the Berlin school of Gestalt psychology (e.g., Wertheimer, Köhler, Koffka) referred to a different ontological idea, considering a Gestalt as a sui generis whole, not founded anymore on elementary objects. Wertheimer (1922, p. 55), for example, emphasized that Gestalts worked "von oben nach unten" ("from top to bottom")—the global configuration has a prevalence over the parts that compose it—and formulated this in clear opposition to Wundt's (1874) constructivist approach in which Gestalts work "von unten nach oben" ("from bottom to top")—the whole is a summing up of its parts.

This debate about Gestalt qualities and true Gestalts as ontological entities and about the interrelationships between wholes and parts in phenomenological experience was never settled because of the intrinsic limitations in ontology and phenomenology. The current state of science provides new conceptual distinctions and methodological tools that allow researchers to finally address these issues empirically. In fact, we identified two general hypotheses in the contemporary literature that are similar to the old divide but stated in explicit processing terms. These hypotheses relate explicitly to current neuroscientific knowledge and, in particular, to the fact that the visual system is comprised of multiple visual areas that are organized hierarchically (Felleman & Van Essen, 1991).

The first hypothesis is that global shape processing might be mainly a function of high-level visual areas (Biederman, 1987; Riesenhuber & Poggio, 1999). Accordingly, multiple

Corresponding Author:
Hans P. Op de Beeck, University of Leuven (K.U. Leuven), Tiensestraat 102 bus 3711, Leuven 3000, Belgium
E-mail: hans.opdebeeck@psy.kuleuven.be

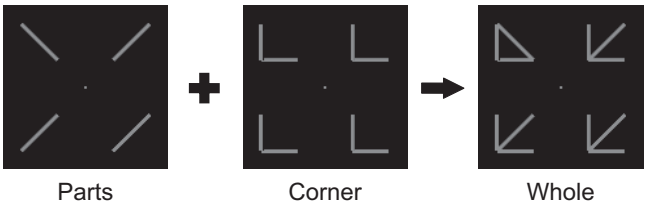

Fig. 1. Illustration of the configural-superiority effect (Pomerantz, Sager, & Stoever, 1977). Detecting the location of the odd line is both slower and less accurate when the display consists of four lines (illustration on the left) than when an irrelevant corner (i.e., two lines forming a 90° angle; illustration in the center) is added to each of the four lines (creating the "wholes" illustrated at the right). Even though the corner is irrelevant because it contains the exact same lines in each case, it has a strong beneficial effect on performance.

studies in monkey physiology, for example, have reported that removal of a part or a feature of a stimulus results in drastic decreases in neurons firing in the inferotemporal cortex (Desimone, Albright, Gross, & Bruce, 1984; Tanaka, Saito, Fukada, & Moriya, 1991). Similarly, in humans, selectivity for full shapes is mediated in the lateral occipital complex but not in retinotopic visual regions like primary visual cortex (Grill-Spector et al., 1998; Kourtzi & Kanwisher, 2001). According to this hypothesis, superadditive grouping effects should only emerge in mid- to high-level visual areas, but not in low-level visual areas.

However, this strict bottom-up view of visual processing is challenged on many fronts. Some researchers have proposed theoretical models to argue for top-down influences in object perception (Bar et al., 2006; Hochstein & Ahissar, 2002; Vecera & O'Reilly, 1998) and, specifically in the case of Gestalt phenomena, computational models have implicated feedback and lateral connections (Craft, Schutze, Niebur, & von der Heydt, 2007; Grossberg, Mingolla, & Ross, 1997; Kogo, Strecha, Van Gool, & Wagemans, 2010). In addition, a few empirical studies have indicated that some grouping effects emerge in low-level visual areas (Altmann, Bülthoff, & Kourtzi, 2003; Kourtzi, Tolias, Altmann, Augath, & Logothetis, 2003; Ostwald, Lam, Li, & Kourtzi, 2008), that in some task contexts even primary visual cortex has the capacity to contain high-level shape representations (Sigman & Gilbert, 2000; Williams et al., 2008), and that as a result of top-down processing, low-level visual areas no longer process the actual physical properties of stimuli but the subjectively experienced properties instead (e.g., Murray, Boyaci, & Kersten, 2006). Thus, according to a second hypothesis, effects such as configural superiority might involve low-level retinotopic areas as much as or even more than high-level shape-selective areas.

To differentiate between these two hypotheses, researchers need to measure the content of representations in the brain. Traditionally, noninvasive brain-imaging techniques were not able to do this because they could reveal only where representations reside in the brain but could not reveal their content. Recently, new analytical tools were developed to perform multivoxel pattern analysis (MVPA); these tools have the power to determine the properties of representations within specific brain areas (Haxby et al., 2001; Kamitani & Tong, 2005; Op de Beeck, Torfs, & Wagemans, 2008; Ostwald et al., 2008). In the study reported here, we used MVPA to examine whether the configural-superiority effect can be observed only in the higher-level visual regions implicated in object processing (Hypothesis 1), or whether its neural correlates can also be observed in low-level visual regions, thus implicating an interplay between low- and higher-level regions (Hypothesis 2).

Our results reveal that the representation of a whole emerges throughout the cortex; however, earlier regions contain more information about the parts, with an equal or better neural detection of parts than of wholes, and higher regions represent the whole shape, with better neural detection of the whole than of its constituent parts. We suggest that the configural-superiority effect provides indirect evidence for the feed-forward account of visual information processing.

Method

Participants and procedure

Nine affiliates of the University of Leuven (1 female, 8 male; ages 21–37) participated in the study. The first participant was excluded from further analyses because of multiple software failures during the scanning session (including many omitted trials and unreliable timing). All participants reported normal or corrected-to-normal vision and provided written informed consent. The experiment was approved by the ethical committee of the Faculty of Psychology and Educational Sciences and the committee for medical ethics at the University of Leuven.

The experiment was coded in Python 2.5 and presented using PsychoPy software (Peirce, 2007, 2009). Source code is available online at http://www.gestaltrevision.be/sources/confsup. The experiment consisted of 11 runs of 336 s each, in the following sequence: three experimental runs, one localizer run, one meridian-mapping run, three experimental runs, one localizer run, and two meridian-mapping runs. A high-resolution anatomical scan was conducted after all runs were complete.

Figure 2 shows the trial sequence for each type of run. Stimuli were projected on a screen, and participants were instructed to respond using a four-key button box. For all runs, the background color was gray.

Experimental runs

Stimuli and presentation. In each experimental trial, a display of four elements was presented. Three of these elements were identical, and one was different. In each of two conditions, there were four possible displays, in each of which the odd shape was located in a different quadrant. In the parts condition (equivalent to the no-context condition in Pomerantz et al., 1977), three quadrants contained straight lines oriented at a 45° angle and one quadrant contained a straight line oriented at a 135° angle. In the wholes condition (equivalent to

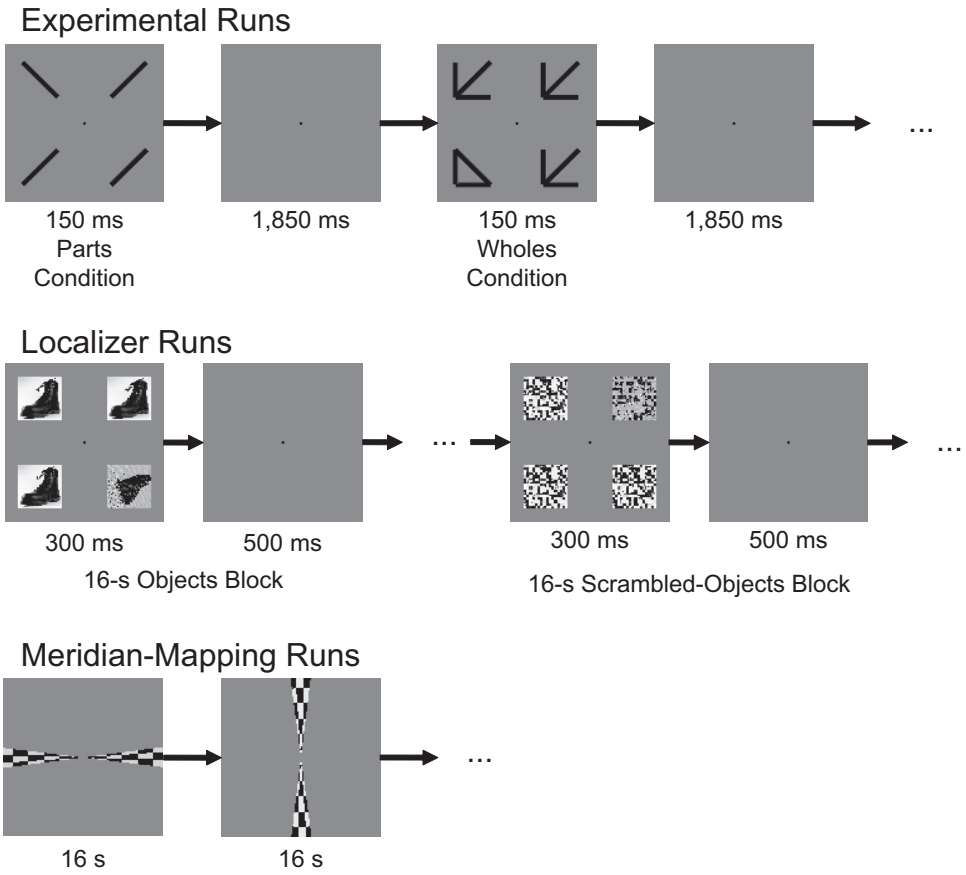

Fig. 2. Illustration of trial sequences for the three types of runs used in the study. In experimental runs (top row), a display containing three identical elements and one odd element was presented. In each trial of the parts condition, the elements consisted of straight lines; in each trial of the wholes condition, two additional lines were added to the single lines to form shapes. A single trial in each condition consisted of a 150-ms display followed by an 1,850-ms interstimulus interval. In the localizer runs (middle row), each trial consisted of a 300-ms presentation of four images (one of which did not match the others), followed by a 500-ms interstimulus interval. In the objects block, the images were intact, and in the scrambled-objects block, the images were scrambled. In meridian-mapping runs (bottom row), flickering wedges composed of checkerboard patterns were presented along either the horizontal or the vertical meridian. Wedges in each orientation were displayed for 16 s each. Note that in both the localizer and the meridian-mapping runs, the stimuli were presented in color.

the good-context condition in Pomerantz et al., 1977), two additional lines forming a right angle (a corner) were added to each of the four lines, so the total display contained three arrows and a triangle. Each element was presented approximately 7° away from a central fixation dot and subtended 4° of visual angle.

Each trial consisted of a 150-ms stimulus presentation (with a central fixation dot) followed by an 1,850-ms interstimulus interval (during which only the fixation dot was presented). The order of trials was optimized using code by Kao, Mandal, Lazar, and Stufken (2009). Each run started and finished with an 8-s fixation block, in which only the fixation dot was present. There were also 22 fixation trials, in which subjects were asked only to fixate on the central dot. These fixation trials were optimally interleaved in between the eight possible displays; the optimization calculations led to six out of eight displays in each run being presented 17 times, and the remaining two being presented 18 times.

Task. Participants were asked to "indicate which one of the four shapes is different" by pressing a key that corresponded to the position of each shape on the display.

Localizer runs

The localizer runs were designed to localize both shape-selective and retinotopic brain areas that were activated by the four stimuli locations in the experimental runs.

Stimuli and presentation. We localized shape-selective lateral occipital cortex (LOC) using a standard localizer scan (Grill-Spector et al., 1998). There were two conditions: intact objects and scrambled objects. The set of 20 intact objects was retrieved from imageafter.com and morguefile.com and consisted of images of human-made objects, foods, and plants. All images were full color and 256 × 256 pixels in size. Each image featured a prominent object presented on a simple (but not

uniform) background, which was included to ensure that images of both intact objects and scrambled objects subtended the same area. The set of 20 scrambled objects was created from the 20 images of intact objects by dividing each image into 256 tiles (16 × 16 pixels) and shuffling them randomly within the image.

Stimuli were presented approximately 7° away from a central fixation dot and subtended 5° of visual angle (to ensure we captured the whole region occupied by the stimuli in the experimental runs). Each trial consisted of a 300-ms stimulus presentation (with a fixation dot present) followed by a 500-ms interstimulus interval with only a fixation dot. There were eight blocks of 20 intact images each, eight blocks of 20 scrambled images each, and five fixation blocks. Blocks were counterbalanced and lasted for 16 s each.

Task. Participants were asked to press a key when they spotted an image repeated from the immediately preceding display (a one-back task). Images were repeated between two and four times per block.

Meridian-mapping runs

Stimuli and presentation. We used a standard procedure for the meridian-mapping runs (Tootell et al., 1995). Two wedges composed of a checkerboard pattern and subtending 15° of visual angle were presented. Wedges were oriented at the horizontal or vertical meridians of the display. Three different parameters of the internal checkerboard pattern of the wedges changed randomly eight times per second: the amount of angular cycles (the number of squares in each row of the wedge; within 20 and 40 cycles), the amount of phase cycles (the number of squares from the tip of a wedge to its base; within 4 and 18 cycles), and the color.

Each run started and finished with an 8-s fixation block. Flickering wedges of either the horizontal or vertical orientation were presented simultaneously for 16 s, followed by wedges of the other orientation. There were 10 blocks of each orientation per run, without fixation blocks in between.

Task. These were passive-viewing runs, so participants had no task to perform.

Functional MRI scanning parameters

Functional MRI (fMRI) data were obtained using a 3-T Philips Intera scanner with an eight-channel SENSE head coil using an echo-planar imaging sequence. We recorded 38 slices from the first 2 participants and 37 slices from the remaining 6 participants. Slices were oriented downward for full inferotemporal cortex coverage and covered almost the entire brain (voxel size = 2.75 × 2.75 × 2.75 mm, interslice distance = 0.2 mm, acquisition matrix = 80 × 80). Each run consisted of 168 measurements; the interval between measurements (repetition time) was set to 2,000 ms with an echo time of 30 ms. The T1-weighted anatomical scan had 0.85- × 0.98-mm in-plane resolution, 1.37 mm between the slices (acquisition matrix = 256 × 256), a 9.6-ms repetition time, a 4.6-ms echo time, 182 coronal slices, and a duration of 383 s.

Imaging data analysis

Data from the fMRI scans were analyzed using Statistical Parametric Mapping software (Version 8; Wellcome Trust Centre for Neuroimaging, London, England). We performed standard data preprocessing, including spatial smoothing of the functional images with a 5.5-mm full-width half-maximum Gaussian kernel, followed by a separate statistical analysis for each run type (experimental, localizer, meridian mapping). (For further details, see Imaging Data Analysis in the Supplemental Material available online.)

Definition of regions of interest. Regions of interest (ROIs) were defined in caret5 software (Van Essen et al., 2001) on a flattened image of the brain taken separately from each participant (Fig. 3). First, the borders between three regions—V1 and V2, V2 and V3, and V3 and higher regions—were defined by observing where activations for the horizontal wedge were greater than activations for the vertical wedge in the meridian-mapping runs. Using this border information, ROIs were selected based on the localizer runs. To identify ROIs in regions V1, V2, and V3, we identified areas in which activations for all stimuli were greater than activations during fixation blocks. For the shape-selective lateral occipital (LO) and posterior fusiform (pFs) cortex, we identified brain regions in which activations for intact objects were greater than activations for scrambled objects (Grill-Spector et al., 1998). LO was defined as a lateral shape-selective region, and pFs cortex was chosen on the ventral surface. Talairach coordinates of all ROIs are provided in Table S1 in the Supplemental Material. Note that data from the experimental runs were not used to define the ROIs.

Multivoxel pattern analysis. We used a PyMVPA package (Hanke et al., 2009) to extract fMRI responses for each voxel individually and the Linear Nu Support Vector Machine provided in PyMVPA (using the Library for Support Vector Machines, or LIBSVM; Chang & Lin, 2001) to perform classification. We followed a 10-fold cross-validated training procedure, repeated 100 times, each time with a different random sampling and splitting of the time points. Performance is reported as the proportion of correct identification of the test data labels.

Results

Behavioral results

Given the results of Pomerantz et al. (1977), we expected a robust configural-superiority effect in the experimental runs, with participants reporting the odd elements both faster and more accurately in the wholes condition (with three lines per quadrant forming a triangle or an arrow) than in the parts

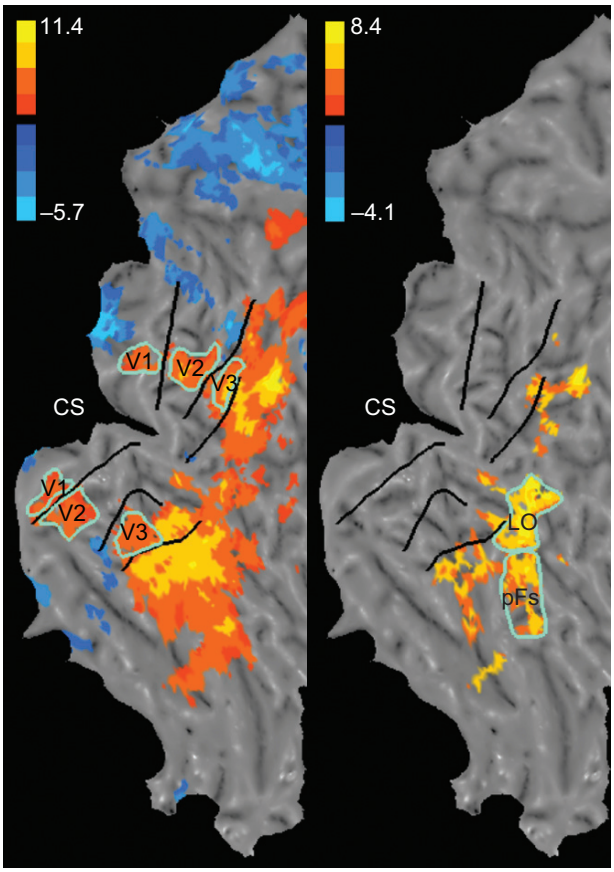

Fig. 3. Illustration of typical regions of interest used for support-vector-machine (SVM) classification. In these flattened images of the right hemisphere of Participant 7, black lines indicate borders between regions, as identified using meridian maps. Cyan outlines mark the regions of interest in which *t* tests (localizer task) indicated that activations during viewing of both intact and scrambled objects were greater than activations during fixation blocks (V1, V2, V3; left panel) or in which activations during viewing of intact objects were greater than activations during viewing of scrambled objects (lateral occipital, LO; posterior fusiform cortex, pFs; right panel). CS = calcarine sulcus.

condition (with only one line per quadrant). Replicating Pomerantz et al.'s findings, our results showed a pronounced configural-superiority effect (Fig. 4a), with the proportion of correct identifications of the location of the odd shape greater in the wholes condition than in the parts condition (difference = ~.3), paired-samples $t(7) = 4.23$, $p = .004$, two-tailed, and response times approximately 300 ms faster in the parts condition (1,172 ms) than in the wholes condition (860 ms), $t(7) = 20.6$, $p < .001$. Moreover, every participant exhibited this effect, and it remained robust even after many hundreds of trials—last scanning run: accuracy, $t(7) = 3.09$, $p = .018$; response time, $t(7) = 18.7$, $p < .001$ (see Fig. S1 in the Supplemental Material for accuracy and response time in each experimental run).

fMRI results

None of the ROIs defined for each participant exhibited differences in the mean percentage signal change across the parts

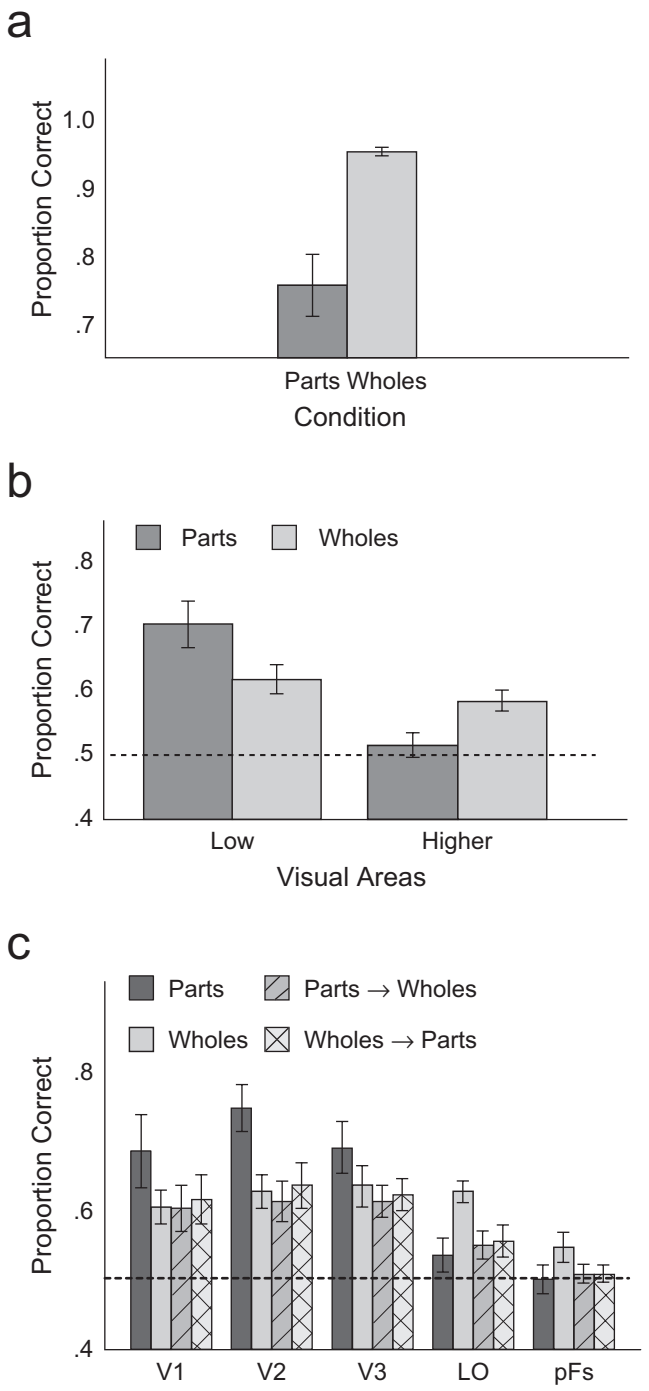

Fig. 4. Behavioral and support-vector-machine (SVM) classification performance. The top graph (a) shows subjects' accuracy in the experimental runs as a function of condition. Using data from a random subset of voxels from each region of interest (ROI), we trained and tested the SVM; the SVM's accuracy during testing is presented in the middle and bottom graphs. The graph in (b) shows results averaged across low-level visual areas (V1, V2, and V3) and across higher-level shape-selective areas (lateral occipital cortex, LO, and posterior fusiform cortex, pFs) when the SVM was trained and tested on the same condition (e.g., trained on parts and tested on parts). The graph in (c) shows results separately for each ROI, both when the SVM was trained and tested on the same condition and when it was trained on one condition and tested on the other (e.g., trained on parts and tested on wholes). Dashed lines represent chance performance (.5). Error bars show ±1 *SEM* ($N = 8$).

and wholes conditions—$t(7)$s ≤ 1.95, ps ≥ .092 (see Fig. S2 in the Supplemental Material).

We therefore performed an MVPA because it is more sensitive than a univariate analysis, in which the response of the voxels in an ROI is averaged together (Haxby et al., 2001; Kamitani & Tong, 2005; Op de Beeck, Brants, Baeck, & Wagemans, 2010). Within each ROI, the fMRI response in each voxel to individual trials was extracted. These patterns of responses were used by a linear support vector machine (SVM) to evaluate differences between representations contained within a particular ROI. In particular, the SVM was trained (learning a linear decision rule) with a part of the fMRI responses to two displays containing the odd element at a different location in the experimental runs (the training set), for example, when the odd element was in the top left quadrant and in the top right quadrant. The two odd-element locations were always picked either both from the parts condition or both from the wholes condition.

After training the SVM, we evaluated how well it was able to detect differences between the two displays by providing it with the remaining fMRI responses and counting how many times the SVM classified those samples correctly as being from one or another display. Classification performance (proportion of correct classifications) reflects the amount of information contained in that ROI. For example, in those ROIs where parts are represented accurately, the classifier should be able to tell the difference among patterns of fMRI responses to the parts conditions better than patterns of fMRI responses to the wholes conditions. We performed such pair-wise classification among all four possible displays in each condition and averaged across all these pair-wise combinations. Therefore, the chance level was 50%. Effects were averaged across hemispheres and participants, given that the difference between the decoding of parts and wholes was not modulated by hemisphere nor was it related to the behavioral performance of individual subjects (see Figs. S3 and S4 in the Supplemental Material).

First, we compared classification performance of the parts and the wholes conditions in low-level versus higher-level visual areas (Fig. 4b). We observed pronounced differences between these areas: The pattern of activity in low-level regions was more informative about parts than about wholes—V1, V2, and V3 averaged together: $t(7) = 3.30$, $p = .013$; however, in higher regions, classification was more accurate for whole shapes—LO and pFs cortex averaged together: $t(7) = 4.05$, $p = .005$. In fact, we could no longer decode parts more often than expected by chance in these higher regions, one-sample $t(7) = 0.81$, $p = .44$, two-tailed.

Further inspection across the five ROIs separately (Fig. 4c) confirmed the observed differences between low-level retinotopic regions and shape-selective regions. In low-level regions V1, V2, and V3, there was a lack of preference for whole shapes, with parts surpassing whole shapes in V2, $t(7) = 7.19$, $p < .001$. Nonetheless, both parts and whole shapes could be decoded more often than expected by chance: $t(7)$s ≤ 3.46, ps ≥ .011. However, in higher-level visual regions LO and pFs cortex, the pattern of results was reversed, with the classification of whole shapes surpassing the classification of parts—LO: $t(7) = 3.60$, $p = .009$; pFs cortex: $t(7) = 3.56$, $p = .009$. Moreover, parts could no longer be decoded reliably more often than expected by chance in these regions—LO: $t(7) = 1.31$, $p = .23$; pFs cortex: $t(7) = 0.068$, $p = .95$. (Note that it is not meaningful to compare the absolute classification performance across ROIs because they differ in a number of properties, such as the functional contrasts used to select the voxels and the size of the original ROI from which these voxels were randomly taken to equate ROI size.)

Higher-level visual areas appear to represent wholes better than parts. Are parts still represented in those wholes? To investigate this question, we next trained our SVM on wholes and tested its performance on parts. If any part information was present in the representation of a whole, we should have been able to decode the part information. Indeed, in the more-posterior-region LO, we were able to decode parts marginally more often than expected by chance, $t(7) = 2.33$, $p = .053$ (Fig. 4c), but this performance was not significantly better than when we trained and tested on parts, $t(7) = 0.79$, $p = .46$. Furthermore, in the more anterior-region pFs cortex, no information appeared to be transferred from wholes to parts—difference from chance level: $t(7) = 0.48$, $p = .65$; comparison with training and testing on parts: $t(7) = 0.39$, $p = .71$. Although the lack of a significant difference from chance might be an issue of statistical power, these results nonetheless suggest that in higher-level visual areas, the representation of wholes might be qualitatively different from the representation of parts.

Discussion

In the study reported here, we investigated how parts are combined into whole shapes in the visual cortex. We used the configural-superiority effect (Pomerantz et al., 1977) to obtain information about the neural underpinning of this process. We demonstrated that whole shapes can be decoded more accurately than their component parts in object-selective regions but not in retinotopic areas, and that whole shapes cannot be decoded successfully from parts information alone in these regions. This finding provides support for the hypothesis regarding the encoding of Gestalts in the brain—that superadditive global shapes emerge at higher-level visual areas—and provides evidence against the hypothesis that lower-level visual areas are able to encode Gestalts as a result of feedback from higher areas. Taken together, the results of our study suggest that, at least in some cases, Gestalt phenomena can emerge gradually in the visual system, with higher regions playing a crucial role in their formation.

More specifically, we demonstrated that the behavioral effect is reflected in the fMRI pattern of responses in higher visual areas, where whole-shape decoding outperformed part decoding, but not in the pattern of responses in retinotopic regions (Fig. 4). Furthermore, we trained pattern classifiers to the whole shape and attempted to decode parts, testing whether higher performance on whole shapes in these regions would lead to an

enhanced decoding of parts, a finding that would suggest at least a partial representation of the parts within the whole shape. However, we observed no such improvement in these higher regions. In fact, the most anterior shape-selective region (pFs cortex) failed to show any transfer of such information, a finding that suggests a complete absence of part information in response patterns at this level of processing.

These findings in pFs cortex are very consistent with subjective Gestalt experience. In terms of phenomenology, a typical property of the formation of a Gestalt is that any notion of the real attributes of the parts gets lost once the Gestalt is formed, as is particularly evident in the case of embedded figures. We have demonstrated here that in the case of the configural-superiority effect, this phenomenology is only reflected in the neural representations of higher-level visual regions and not in retinotopic areas. Although low-level visual areas contain information about parts, they do not seem to contribute to the Gestalt experience, in contrast to reports on the neural basis of some other Gestalt phenomena (e.g., Murray et al., 2006). In fact, retinotopic areas V1, V2, and V3 exhibited an opposite pattern, such that decoding of parts was actually more pronounced when the context of the full shape was not present. This observed disadvantage of the whole shapes in the lower regions could, for example, result from the processing of extra parts (the corner), which were the same in all conditions and thus provided only noise. It is interesting that despite a clear disadvantage for whole shapes in these lower regions, the configural-superiority effect was nonetheless present at the higher level in the visual system's hierarchy as well as in the phenomenological experience.

Such double dissociation between lower and higher regions suggests a feed-forward emergence of the configural-superiority effect. If low-level regions are employed by higher regions via feedback in order to facilitate whole-shape processing, then we would have expected to see also an advantage for whole-shape decoding in these low-level regions. Nevertheless, the poor temporal and spatial resolution of our method did not allow us to directly observe feed-forward, lateral interactive, and feedback effects and their typical timing characteristics, so it is not possible to rule out that feedback would play a role in the observed findings in a way that would not lead to a neural configural superiority in low-level regions.

In this context, it is interesting to note that previous studies recording the brain's electrical activity have already suggested that some perceptual-grouping phenomena could reflect bottom-up processing (Nikolaev, Gepshtein, Kubovy, & van Leeuwen, 2008; Tanskanen, Saarinen, Parkkonen, & Hari, 2008). For instance, using high-density event-related potential (ERP) measurements, Nikolaev et al. (2008) showed that the sensitivity to different groupings of dots in multistable dot lattices was correlated positively with the amplitude of the earliest ERP peak (C1, about 60 ms after stimulus onset)—a correlation believed to reflect spontaneous feed-forward processes—and negatively with the amplitude of the next ERP peak (P1, about 110 ms after stimulus onset)—a correlation believed to reflect lateral and feedback interactions, which are under more attentive control than spontaneous feed-forward processes. Lamme and his colleagues developed a rich theory about the interplay between feed-forward and recurrent processing, and the role of this interplay in the emergence of visual awareness, on the basis of their neurophysiological work on the time course of selectivity in V1 (for a review, see Lamme & Roelfsema, 2000).

Taken together, our results provide evidence that the configural-superiority effect gradually emerges throughout the feed-forward cortical shape-processing hierarchy. In addition, we propose that our findings provide a new perspective on the rather conflicting literature on the neural basis of Gestalts. Specifically, findings from studies such as ours enable the development of a taxonomy of Gestalt phenomena with two broad categories: bottom-up Gestalts and top-down Gestalts. Our findings seem to indicate that the configural-superiority effect is a bottom-up Gestalt process.

Acknowledgments

We thank Lee de-Wit, Wouter Braet, M. Dorothee Augustin, two anonymous reviewers, and Shaun P. Vecera for their excellent comments on the manuscript. J. K. is a Research Assistant of the Research Foundation – Flanders (FWO).

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Funding

This work was supported by a Methusalem grant (METH/08/02) from the Flemish Government, by the Fund for Scientific Research – Flanders (Grant KAN 1.5.022.08), and by the Human Frontier Science Program (Grant CDA-0040/2008).

Supplemental Material

Additional supporting information may be found at http://pss.sagepub.com/content/by/supplemental-data

References

Albright, T. D., & Stoner, G. R. (2002). Contextual influences on visual processing. *Annual Review of Neuroscience*, *25*, 339–379.

Altmann, C. F., Bülthoff, H. H., & Kourtzi, Z. (2003). Perceptual organization of local elements into global shapes in the human visual cortex. *Current Biology*, *13*, 342–349.

Ash, M. G. (1995). *Gestalt psychology in German culture, 1890–1967: Holism and the quest for objectivity*. Cambridge, England: Cambridge University Press.

Bar, M., Kassam, K. S., Ghuman, A. S., Boshyan, J., Schmid, A. M., Dale, A. M., . . . Halgren, E. (2006). Top-down facilitation of visual recognition. *Proceedings of the National Academy of Sciences, USA*, *103*, 449–454.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, *94*, 115–147.

Chang, C. C., & Lin, C. J. (2001). LIBSVM: A library for support vector machines [Computer software]. Retrieved from http://www.csie.ntu.edu.tw/~cjlin/libsvm

Craft, E., Schutze, H., Niebur, E., & von der Heydt, R. (2007). A neural model of figure-ground organization. *Journal of Neurophysiology*, *97*, 4310–4326.

Desimone, R., Albright, T., Gross, C., & Bruce, C. (1984). Stimulus-selective properties of inferior temporal neurons in the macaque. *Journal of Neuroscience*, *4*, 2051–2062.

Felleman, D. J., & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*, *1*, 1–47.

Grill-Spector, K., Kushnir, T., Hendler, T., Edelman, S., Itzchak, Y., & Malach, R. (1998). A sequence of object-processing stages revealed by fMRI in the human occipital lobe. *Human Brain Mapping*, *6*, 316–328.

Grossberg, S., Mingolla, E., & Ross, W. D. (1997). Visual brain and visual perception: How does the cortex do perceptual grouping? *Trends in Neurosciences*, *20*, 106–111.

Hanke, M., Halchenko, Y., Sederberg, P., Hanson, S., Haxby, J., & Pollmann, S. (2009). PyMVPA: A Python toolbox for multivariate pattern analysis of fMRI data. *Neuroinformatics*, *7*, 37–53.

Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*, *293*, 2425–2430.

Hochstein, S., & Ahissar, M. (2002). View from the top: Hierarchies and reverse hierarchies in the visual system. *Neuron*, *36*, 791–804.

Kamitani, Y., & Tong, F. (2005). Decoding the visual and subjective contents of the human brain. *Nature Neuroscience*, *8*, 679–685.

Kao, M., Mandal, A., Lazar, N., & Stufken, J. (2009). Multi-objective optimal experimental designs for event-related fMRI studies. *NeuroImage*, *44*, 849–856.

Kogo, N., Strecha, C., Van Gool, L., & Wagemans, J. (2010). Surface construction by a 2-D differentiation-integration process: A neurocomputational model for perceived border ownership, depth, and lightness in Kanizsa figures. *Psychological Review*, *117*, 406–439.

Kourtzi, Z., & Kanwisher, N. (2001). Representation of perceived object shape by the human lateral occipital complex. *Science*, *293*, 1506–1509.

Kourtzi, Z., Tolias, A. S., Altmann, C. F., Augath, M., & Logothetis, N. K. (2003). Integration of local features into global shapes: Monkey and human fMRI studies. *Neuron*, *37*, 333–346.

Lamme, V. A. F., & Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neurosciences*, *23*, 571–579.

Murray, S. O., Boyaci, H., & Kersten, D. (2006). The representation of perceived angular size in human primary visual cortex. *Nature Neuroscience*, *9*, 429–434.

Nikolaev, A. R., Gepshtein, S., Kubovy, M., & van Leeuwen, C. (2008). Dissociation of early evoked cortical activity in perceptual grouping. *Experimental Brain Research*, *186*, 107–122.

Op de Beeck, H. P., Brants, M., Baeck, A., & Wagemans, J. (2010). Distributed subordinate specificity for bodies, faces, and buildings in human ventral visual cortex. *NeuroImage*, *49*, 3414–3425.

Op de Beeck, H. P., Torfs, K., & Wagemans, J. (2008). Perceived shape similarity among unfamiliar objects and the organization of the human object vision pathway. *Journal of Neuroscience*, *28*, 10111–10123.

Ostwald, D., Lam, J. M., Li, S., & Kourtzi, Z. (2008). Neural coding of global form in the human visual cortex. *Journal of Neurophysiology*, *99*, 2456–2469.

Peirce, J. W. (2007). PsychoPy: Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*, 8–13.

Peirce, J. W. (2009). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics*, *2*, 10. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2636899/

Pomerantz, J. R., Sager, L. C., & Stoever, R. J. (1977). Perception of wholes and of their component parts: Some configural superiority effects. *Journal of Experimental Psychology: Human Perception and Performance*, *3*, 422–435.

Riesenhuber, M., & Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nature Neuroscience*, *2*, 1019–1025.

Sigman, M., & Gilbert, C. D. (2000). Learning to find a shape. *Nature Neuroscience*, *3*, 264–269.

Spillmann, L. (2009). Phenomenology and neurophysiological correlations: Two approaches to perception research. *Vision Research*, *49*, 1507–1521.

Tanaka, K., Saito, H., Fukada, Y., & Moriya, M. (1991). Coding visual images of objects in the inferotemporal cortex of the macaque monkey. *Journal of Neurophysiology*, *66*, 170–189.

Tanskanen, T., Saarinen, J., Parkkonen, L., & Hari, R. (2008). From local to global: Cortical dynamics of contour integration. *Journal of Vision*, *8*(7), Article 15. Retrieved from http://www.journalofvision.org/content/8/7/15

Tootell, R., Reppas, J., Kwong, K., Malach, R., Born, R., Brady, T., . . . Belliveau, J. W. (1995). Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. *Journal of Neuroscience*, *15*, 3215–3230.

Van Essen, D. C., Drury, H. A., Dickson, J., Harwell, J., Hanlon, D., & Anderson, C. H. (2001). An integrated software suite for surface-based analyses of cerebral cortex. *Journal of the American Medical Informatics Association*, *8*, 443–459.

Vecera, S. P., & O'Reilly, R. C. (1998). Figure-ground organization and object recognition processes: An interactive account. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 441–462.

Wagemans, J., Wichmann, F. A., & Op de Beeck, H. (2005). Visual perception I: Basic principles. In K. Lamberts & R. L. Goldstone (Eds.), *Handbook of cognition* (pp. 3–47). London, England: Sage.

Wertheimer, M. (1922). Untersuchungen zur Lehre von der Gestalt, I: Prinzipielle bemerkungen [Investigations of the principles of Gestalt, I: Fundamental remarks]. *Psychologische Forschung*, *1*, 47–58.

Williams, M. A., Baker, C. I., Op de Beeck, H. P., Mok Shim, W., Dang, S., Triantafyllou, C., & Kanwisher, N. G. (2008). Feedback of visual object information to foveal retinotopic cortex. *Nature Neuroscience*, *11*, 1439–1445.

Wundt, W. M. (1874). *Grundzüge der physiologischen psychologie* [Principles of physiological psychology]. Leipzig, Germany: W. Engelman.