

**Short Report****Configural Gestalts remain nothing more than the sum of their parts in visual agnosia**

Lee H. de-Wit

Laboratory of Experimental Psychology, KU Leuven, Leuven, Belgium; e-mail: [Lee.deWit@ppw.kuleuven.be](mailto:Lee.deWit@ppw.kuleuven.be)

Jonas Kubilius

Laboratory of Biological Psychology, KU Leuven, Leuven, Belgium; e-mail: [Jonas.Kubilius@ppw.kuleuven.be](mailto:Jonas.Kubilius@ppw.kuleuven.be)

Hans P. Op de Beeck

Laboratory of Biological Psychology, KU Leuven, Leuven, Belgium; e-mail: [Hans.OpDeBeeck@ppw.kuleuven.be](mailto:Hans.OpDeBeeck@ppw.kuleuven.be)

Johan Wagemans

Laboratory of Experimental Psychology, KU Leuven, Leuven, Belgium; e-mail: [Johan.Wagemans@ppw.kuleuven.be](mailto:Johan.Wagemans@ppw.kuleuven.be)

Received 31 July 2013, in revised form 7 October 2013; published online 22 October 2013.

**Abstract.** We report converging evidence that higher stages of the visual system are critically required for the whole to become more than the sum of its parts by studying patient DF with visual agnosia using a configural superiority paradigm. We demonstrate a clear dissociation between this patient and normal controls such that she could more easily report information about parts, demonstrating a striking reversal of the normal configural superiority effect. Furthermore, by comparing DF's performance to earlier neuroimaging and novel modeling work, we found a compelling consistency between her performance and representations in the early visual areas, which are spared in this patient. The reversed pattern of performance in this patient highlights that in some cases visual Gestalts do not emerge early on without processing in higher visual areas. More broadly, this study demonstrates how neuropsychological patients can be used to unmask representations maintained at early stages of processing.

**Keywords:** visual agnosia, perceptual grouping, perceptual organization, configural superiority.

**1 Main text**

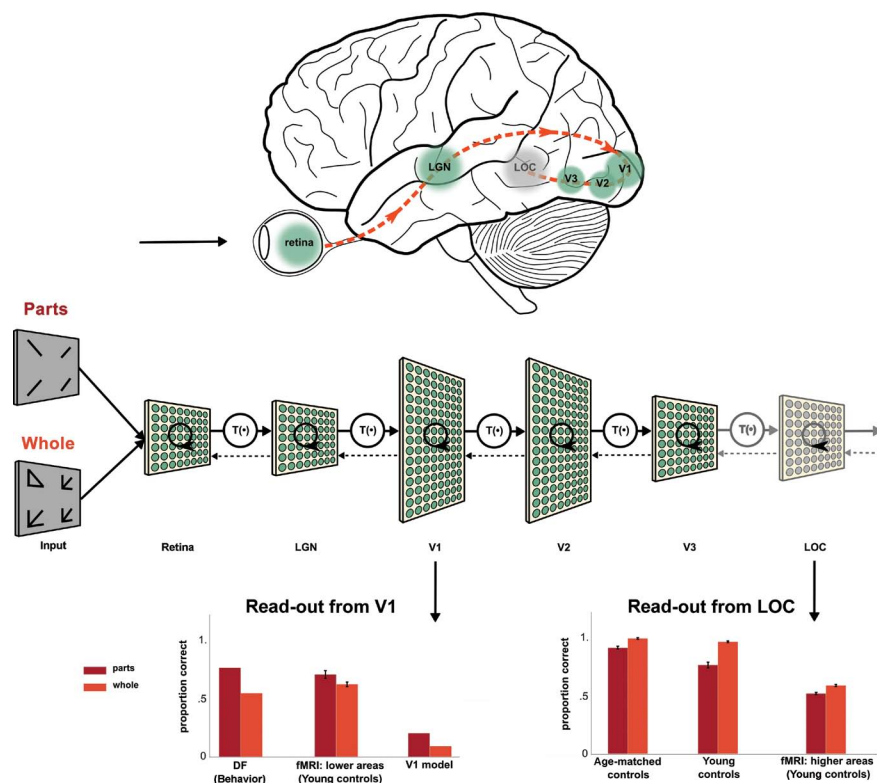
How do wholes become different from the sum of their parts? This classic Gestalt question is elegantly brought to experimental life in a simple configural superiority paradigm devised by James Pomerantz and colleagues (Pomerantz, Sager, & Stoever, 1977). When observers have to indicate which of four oriented lines has a different orientation (see Figure 1), Pomerantz et al. (1977) found that adding an identical, and thus objectively uninformative, corner to each line led to a pronounced increase in the perceptual salience of the resulting shape. The behavioral advantage for this configuration has been argued to reflect the role of higher visual areas in Gestalt formation (Kubilius, Wagemans, & Op de Beeck, 2011). More specifically, Kubilius et al. (2011) used multi-voxel pattern analysis (MVPA) techniques in a functional magnetic resonance imaging (fMRI) experiment to demonstrate that patterns of activation evoked by this paradigm revealed an advantage for the configural “whole” condition only in the higher visual areas lateral occipital (LO) and posterior fusiform (pFs). Early areas, like the primary visual cortex (V1), in fact showed the opposite, i.e., a better discrimination in the isolated “part” condition. This result potentially demonstrates that configural information exists neither in the stimulus nor in its representation in early visual areas—Gestalts only become different from the sum of their parts at higher stages of processing.

The results of Kubilius et al. (2011) are limited, however, first in the extent that all neuroimaging results are limited to an observation of correlation and, second, by an assumption often made when using MVPA, namely that the patterns of activation at the level of voxels (averaging over hundreds of thousands of neurons) can provide a direct measure of the representational content of a given area of the brain.

L. H. de-Wit and J. Kubilius contributed equally to this work.

Here we provide direct causal evidence that higher visual areas are required to construct the configural superiority Gestalt. In particular, we examined perceptual grouping in the visual form agnosia patient DF. This patient has been most famously studied in terms of a dissociation in her ability to perceive the world versus her ability to guide visual responses (Goodale, Milner, Jakobson, & Carey, 1991). In this study, we do not focus on this patient’s “vision for action” but rather on the fact that her lesion centers bilaterally on the LO area in the ventral stream (James, Culham, Humphrey, Milner, & Goodale, 2003). Earlier stages of her visual system, including the primary visual cortex, seem to be relatively spared (Bridge et al., 2013) and support functional behavior, including the guidance of action and basic orientation illusions such as the McCullough effect (Keith, Goodale, & Gernsey, 1991). This patient provides a perfect test case for the configural superiority effect: If higher stages of the ventral stream are required for the construction of such seemingly basic Gestalts, then one predicts that DF will not reveal the configural superiority advantage observed in normal participants.

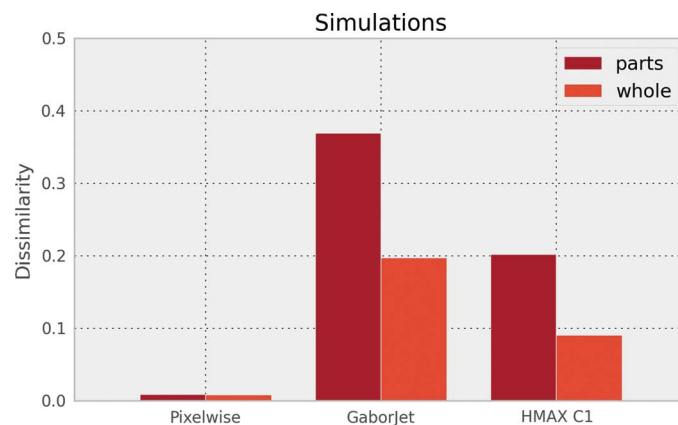
We tested DF with the same paradigm used by Kubilius et al. (2011) with two modifications that were expected to optimize the design for testing this patient and age-matched controls, as explained below. In each trial, a display of four elements was presented (Figure 1, left, see “parts” and “wholes”). Three of these elements were identical, and one was different. In the parts condition, the identical



**Figure 1.** The hypothesized read-out of parts and wholes in the visual hierarchy. *Top and middle rows:* schematic depiction of DF’s visual lesion along the ventral stream. Functional MRI scans show that DF lacks mid- and higher-level visual areas in the ventral visual pathway particularly LO (James et al., 2003), which forms part of the lateral occipital complex (LOC). *Bottom row:* representation of the assumed read-out of the content in early visual areas, based on fMRI data (from Kubilius et al., 2011), DF’s performance, and a V1 model (HMAX layer C1; Riesenhuber & Poggio, 1999). The performance of age-matched and young controls presumably reflects a read-out of representations available in LOC, given the consistency between the fMRI decoding in this area and the behavioral advantage in normal adults (figure adapted from DiCarlo & Cox, 2007; see Kubilius, 2013b).

elements were straight lines oriented at a 45° angle, and the odd element was a line oriented at a 135° angle. In the whole condition, additional lines forming a right angle (a corner) were added to each of the four lines, resulting in three arrows and a triangle. These elements were presented 7° away from the central fixation dot. DF was asked to choose which quadrant contained the odd element.

Unlike in Kubilius et al. (2011), the stimuli were present on the screen until the patient made a response, rather than being briefly flashed. Second, rather than having to press buttons to identify the

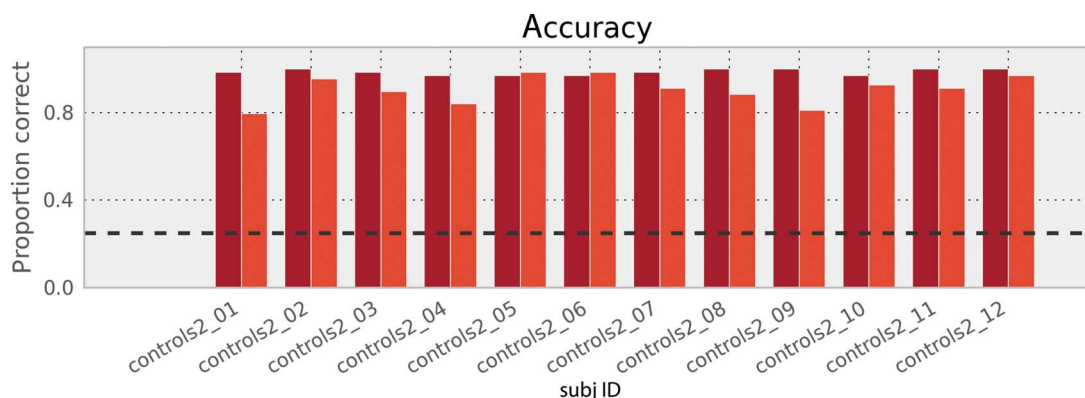


**Figure 2.** Discriminability between stimuli in simple V1 models: Pixelwise, Gabor Jet, and HMAX C1. GaborJet and HMAX layer C1 outputs resembled both DF’s performance and discriminability of stimuli in the primary visual cortex (V1) of young participants tested in Kubilius et al. (2011). Pixelwise model exhibited only a minute advantage of parts versus wholes ( $8.87 \times 10^{-3}$  vs.  $8.54 \times 10^{-3}$ ) due to a large number of identical gray background pixels.

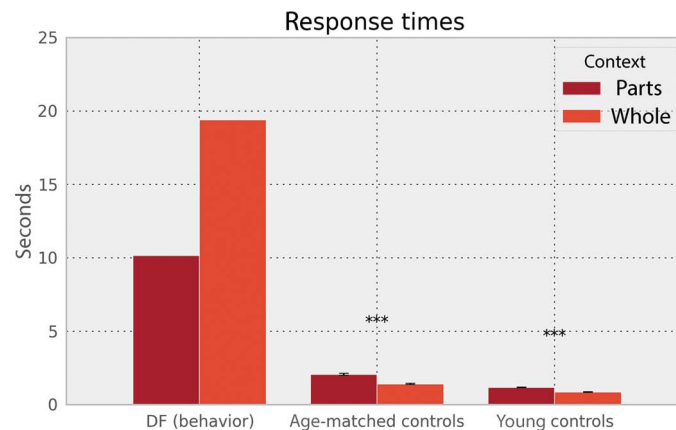
location of the “odd-one-out,” the patient responded by pointing to the option she wished to select. Note that this was not intended to make the task a “vision for action” task; rather, DF was using a simple egocentric action to symbolically provide her response. Her responses were coded by the experimenter. Twelve age-matched controls (six males, six females; ages 51–61 years) were tested using the same procedure.

The results revealed a striking dissociation between DF and the control group, both in terms of accuracy and correct reaction time (see Figure 1). Using a statistic developed by Crawford, Garthwaite, and Porter (2010) to compare the difference score of a single patient to a small control sample, we can confirm that DF’s results clearly dissociate from the normal population (two-tailed:  $p < .001$  for accuracy and  $p < .00001$  for reaction time). While the control participants exhibited a robust configural superiority advantage ( $t(11) = 4.07, p = .002$ ; see Figure 3 for individual performance), DF showed the opposite: she performed better in the isolated line or part condition (two-proportion  $z$ -test,  $z = 3.79, p < .0001$ ). DF’s performance is in fact strikingly consistent with the discriminability of patterns of activation measured with fMRI in the early visual cortex of healthy young participants (Kubilius et al., 2011). DF’s reaction times also revealed the same reversal, with much faster performance in the isolated line condition, though overall DF’s reaction times were much slower than healthy controls (Figure 4).

This result is consistent with the hypothesis that early visual cortex is not sufficient to produce compelling Gestalt representations. It might appear counterintuitive that DF should not only reveal no configural advantage, but also perform better in the isolated line condition than in the configural condition. To address this issue further, we compared DF’s performance to different models of V1 that apply Gabor filters at various orientations, spatial frequencies, and scales (Figure 1; Riesenhuber & Poggio, 1999; Xu, Yue, Lescroart, Biederman, & Kim, 2009; see Figure 2 for comparisons to other



**Figure 3.** Accuracies for all age-matched control participants. Dashed line indicates chance level (25%).



**Figure 4.** Response times (plotted for correct responses only) for patient DF, age-matched controls, and young controls revealed a robust interaction (two-tailed  $p < .0001$  using Crawford et al., 2010). Moreover, there was a highly statistically significant difference between the parts and whole conditions in each group (DF: generalized linear model using binomial distribution  $t(275) = 5.75$ ,  $p < .0001$ ; age-matched controls: two-tailed related samples  $t(11) = 9.64$ ,  $p < .0001$ ; young controls: two-tailed related samples  $t(7) = 20.62$ ,  $p < .0001$ ).

models). Just like DF, these models discriminated better between parts than whole shapes, indicating that at an early stage of processing, the additional corner units simply add distracting noise or clutter that requires further processing to be usefully organized as a genuine Gestalt.

Together, our results provide converging neuropsychological and computational evidence that higher visual areas are instrumental in the emergence of these configural Gestalts. Thus, while the primary visual cortex might show some sensitivity to certain grouping cues (Wannig, Stanisor, & Roelfsema, 2011), our study highlights the critical importance of higher-level vision in organizing visual input such that the whole becomes quantifiably different from its parts. Viewed as a visual search task, our results with the “configural superiority effect” also suggest that the representational differences that are most “salient” for the visual system are not computed in earlier areas, providing an important challenge to models that assume that salience is computed by V1 (Li, 1999, 2009). More broadly, our results highlight how neuropsychological patients can be used to test computational models by “unmasking” the representations at earlier stages of processing (Mannan, Kennard, & Husain, 2009; Ossandón et al., 2012).

## 2 Methods

### 2.1 Participants

Patient DF, aged 59, participated in the study. Twelve age-matched participants (six females, six males; ages 51–61 years) participated in the study. The experiment was approved by the ethical committee of the Faculty of Psychology and Educational Sciences at KU Leuven.

### 2.2 Software

The experiment was coded, presented, and analyzed in Python 2.7 (with an exception of comparison of patient DF to control participants, which was analyzed using custom software developed by Crawford et al., 2010, accessible at <http://homepages.abdn.ac.uk/j.crawford/pages/dept/SingleCaseMethodology.htm>) using *psychopy\_ext* extension (Kubilius, 2013a) for *PsychoPy* (Peirce, 2009, 2009), *pandas*, and *matplotlib*. Full source code and all collected data are available online at <https://bitbucket.org/qbilius/df>.

### 2.3 Model simulations

We used three simple models of V1:

- (i) *Pixelwise*, where raw pixel values are used for comparing stimuli;
- (ii) *GaborJet* (Xu et al., 2009), where a given image is decomposed using eight orientations Gabor filters of five spatial scales at 100 image locations, resulting in a 4,000-dimensional output vector;
- (iii) *layer C1 of the original HMAX model*, presumably similar to V1 (Riesenhuber & Poggio, 1999) where four orientations and 12 spatial scales of Gabor filters are applied at each image location (layer S1) and pooled over nearby locations and sizes; each model was provided with  $256 \times 256$

px images of each configuration observed by participants. Discriminability was computed using the dissimilarity measure as proposed by Xu et al. (2009). Briefly, this measure reflects an absolute value of a sine of an angle between the (vectorized) outputs of a model. If the two outputs are very similar, the angle is small, resulting in a low discriminability. Conversely, if the two outputs are dissimilar, the angle is close to 90°, resulting in a high discriminability value of nearly 1.

## 2.4 Additional details

For details about the experimental procedure and young participants (ages 21–37 years) whose behavioral and fMRI data we used in this study, please refer to an earlier report by Kubilius et al. (2011).

**Acknowledgments.** This work was supported by a Methusalem grant (METH/08/02) awarded to Johan Wagemans from the Flemish Government. L.H.D. is a postdoctoral fellow, and J. K. is a research assistant of the Research Foundation-Flanders (FWO).

## References

- Bridge, H., Thomas, O. M., Minini, L., Cavina-Pratesi, C., Milner, A. D., & Parker, A. J. (2013). Structural and functional changes across the visual cortex of a patient with visual form agnosia. *Journal of Neuroscience*, *33*(31), 12779–12791. doi:10.1523/JNEUROSCI.4853-12.2013
- Crawford, J. R., Garthwaite, P. H., & Porter, S. (2010). Point and interval estimates of effect sizes for the case-controls design in neuropsychology: Rationale, methods, implementations, and proposed reporting standards. *Cognitive Neuropsychology*, *27*, 245–260. doi:10.1080/02643294.2010.513967
- DiCarlo, J. J., & Cox, D. D. (2007). Untangling invariant object recognition. *Trends in Cognitive Sciences*, *11*(8), 333–341. doi:10.1016/j.tics.2007.06.010
- Goodale, M. A., Milner, A. D., Jakobson, L. S., & Carey, D. P. (1991). A neurological dissociation between perceiving objects and grasping them. *Nature*, *349*(6305), 154–156. doi:10.1038/349154a0
- James, T. W., Culham, J., Humphrey, G. K., Milner, A. D., & Goodale, M. A. (2003). Ventral occipital lesions impair object recognition but not object-directed grasping: An fMRI study. *Brain*, *126*(11), 2463–2475. doi:10.1093/brain/awg248
- Keith, G. H., Goodale, M. A., & Gurnsey, R. (1991). Orientation discrimination in a visual form agnostic: Evidence from the McCollough effect. *Psychological Science*, *2*(5), 331–335. doi:10.1111/j.1467-9280.1991.tb00161.x
- Kubilius, J. (2013a). *psychopy\_ext*. Retrieved from [https://github.com/qbilius/psychopy\\_ext](https://github.com/qbilius/psychopy_ext)
- Kubilius, J. (2013b). *Ventral visual stream*. Retrieved September 23, 2013, from <http://dx.doi.org/10.6084/m9.figshare.106794>
- Kubilius, J., Wagemans, J., & Op de Beeck, H. P. (2011). Emergence of perceptual Gestalts in the human visual cortex: The case of the configural-superiority effect. *Psychological Science*, *22*(10), 1296–1303. doi:10.1177/0956797611417000
- Li, Z. (1999). Contextual influences in V1 as a basis for pop out and asymmetry in visual search. *Proceedings of the National Academy of Sciences*, *96*(18), 10530–10535. doi:10.1073/pnas.96.18.10530
- Li, Z. (2002). A saliency map in primary visual cortex. *Trends in Cognitive Sciences*, *6*(1), 9–16.
- Mannan, S. K., Kennard, C., & Husain, M. (2009). The role of visual salience in directing eye movements in visual object agnosia. *Current Biology*, *19*(6), R247–R248. doi:10.1016/j.cub.2009.02.020
- Ossandón, J. P., Onat, S., Cazzoli, D., Nyffeler, T., Müri, R., & König, P. (2012). Unmasking the contribution of low-level features to the guidance of attention. *Neuropsychologia*, *50*(14), 3478–3487. doi:10.1016/j.neuropsychologia.2012.09.043
- Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1–2), 8–13. doi:10.1016/j.jneumeth.2006.11.017
- Peirce, J. W. (2009). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics*, *2*, 10. doi:10.3389/neuro.11.010.2008
- 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*(3), 422–435. doi:10.1037/0096-1523.3.3.422
- Riesenhuber, M., & Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nature Neuroscience*, *2*, 1019–1025. doi:10.1038/14819
- Wannig, A., Stanisor, L., & Roelfsema, P. R. (2011). Automatic spread of attentional response modulation along Gestalt criteria in primary visual cortex. *Nature Neuroscience*, *14*(10), 1243–1244. doi:10.1038/nn.2910
- Xu, X., Yue, X., Lescroart, M. D., Biederman, I., & Kim, J. G. (2009). Adaptation in the fusiform face area (FFA): Image or person? *Vision Research*, *49*(23), 2800–2807. doi:10.1016/j.visres.2009.08.021