Journal of Biomed Research

Mersad et al., J Biomed Res. 2022;3(1):8-10.

Commentary

Blending into the crowd: electrophysiological evidence of gestalt perception of a human dyad: extended discussion and theoretical viewpoint

Karima Mersad[1*], Philippe Oliviero[2]

[1]Université Paris Cité, Vision Action Cognition, F-92100 Boulogne-Billancourt, France

[2]Université Paris Cité, Institut Universitaire de Technologie, 45 Rue des Saints-Pères, 75006 Paris, France

*Author for correspondence: Email: karima.mersad@parisdescartes.fr

Received date: May 02, 2022
Accepted date: May 19, 2022

Citation: Mersad K, Oliviero P. Blending into the crowd: electrophysiological evidence of gestalt perception of a human dyad: extended discussion and theoretical viewpoint. J Biomed Res. 2022;3(1):8-10.

Abstract

In this commentary, we provide further discussion and interpretation of a recent article entitled "Blending into the Crowd: Electrophysiological Evidence of Gestalt Perception of a Human Dyad", published one year ago by the first author of the present commentary. Firstly, drawing a parallel between the experiment described in the above article and another closely comparable experimental study, we propose that the neural integration process evidenced when seeing two human shapes close in space is a marker of the categorization of a stimulus as a group of humans (two here) represented as an entity per se. We also highlight that the original article provides a new kind of primacy of global visual processing over local elements. Lastly, we suggest that holistic perceptual processing of a dyad and more generally of a group, might guide individual's actions in response to intentions and behaviors at the group level.

The Frequency Tagging paradigm, where different visual stimuli are presented at different frequencies, has primarily been used to investigate low-level visual processing. More recently, studies focused on higher-level processes, notably, hierarchical visual perceptual organization of complex objects and scenes. In these paradigms, the response at a stimulation frequency constitutes a marker of the stimulus-specific neural response while objective evidence of a perceptual integration of stimuli in a holistic representation, is provided by the response at some intermodulation (IM) components, when such components exist.

Communication

Social Groups as a Mental Category

Mersad and Caristan [1] used nested complex stimuli: a dyad composed of two close but separated busts, each of them being itself a complex stimulus formed with parts (head and shoulders). In this case, the evidenced perceptual integration reflects indubitably high-level visual computations. However, what is exactly the nature of the integration process? A parallel with Boremanse et al. [2] study where the authors reported neural integration in a closely similar paradigm but with different stimuli, might afford insight into this question. In their study, these authors could objectively dissociate the neural response of each half of a face, reflected in the input frequencies, from the response to the face as a whole, which appeared as a combination of these input frequencies. This latter response provided a signature of non-linear integration of face halves into a perceptual unit, comparable to the neural convergence observed for the two silhouettes stimuli in Mersad and Caristan [1]. However, the nature of the stimuli constituting the parts is a critical difference between the two studies. Indeed, the half of a face do not belong to the realm of natural visual entities: we often perceive faces in profile, but we rarely see only the half of a face. Thus, perceptual integration of face halves into a coherent unit, observed in Boremanse et al. [2] study, is a form of (perceptual) reconstruction of an object, unnaturally divided in two parts presented at different temporal frequencies to be dissociated.

Moreover, the face halves were contiguous which could have induced low-level neural integration as the small receptive fields spanning the border of the two halves could be responsible for generating responses at IM components.

Unlike the half of a face, a human silhouette is an independent visual entity that is part of the ecological world. In Mersad and Caristan [1], two silhouettes were presented close to each other in space but with a clear gap between them. Here, response at IM frequency signals the existence of a neural network processing both of these physically distant visual items, joining them into a coherent percept [3].

Critically, the IM response was significantly reduced when human silhouette stimuli were replaced with another category of objects, two silhouettes of chairs, and when the dyad was presented upside down, an inversion effect observed in visual perceptual categorization of objects [4].

Our proposal is that the IM component reflecting perceptual integration evidenced in Mersad and Caristan [1] experiment is a neural marker of mental categorization of a visual stimulus as a *group of humans* (two here) represented as an entity *per se.*

Perception of Nested Structures: Global or Local Level?

An enduring question in psychology that can be tackled by the Frequency Tagging approach is the relations between whole (global) and part (local) levels of processing, an issue discussed in Mersad and Caristan [1] study. The authors noted that while inversion effect was observable at the global dyad level, it was absent from the local silhouette level, and this, in spite of abundant data in the literature showing that body pictures and silhouettes are subjected to inversion effect [5].

The dominance of global visual percept over local information has been documented from early Gestalt theories [6] and Navon's 'Forest before trees' experiment, to ensemble coding studies showing that a summary representation is favored over the representation of local details [7]. These studies showed that visual percepts are dominated by global information using paradigms such as shape detection and identification [8], processing time of gist extraction from visual scenes [9] or statistical summary ensemble computations [7]. Some studies highlighted the role of structural or configural information, for example, Haberman et al., [7] found a decline in the ability to compute the average emotional expression of a set of faces when the faces where presented upside down.

However, Mersad and Caristan [1] could experimentally dissociate the neural entrained response to the whole-dyad silhouette from the response to the individual silhouettes forming the dyad, while presenting these stimuli either in (typical) upright or (altered) inverted position (Figure 1). Thus, the authors could compare the processing of the *structural information* between the two levels (local and global) of the visual perceptual hierarchy. They showed that configural information at the global level dominates perception, and this, specifically when the nested stimuli represented human silhouettes. Hence, their results offer a new kind of demonstration of global primacy.

Perception for Action: When Facing Human Pluralities

Gibson's ecological theory holds that perception is a guide for adaptive behavior [10]. Indeed, following Gibson's Perceiving is for doing principle, perception, while offering affordances, is closely linked to action.

Do distinctive mechanisms for perception of human groups constrain individual's attitudes and actions?

To our knowledge, no answer has been proposed to this question in the literature. However, we can distinguish two different types of processes that has been reported for visual perception of human aggregates: statistical summary principles described in ensemble coding (e.g. Habermann and Whitney, 2012 [7]) and holistic processing highlighted in Mersad and Caristan [1], although in the

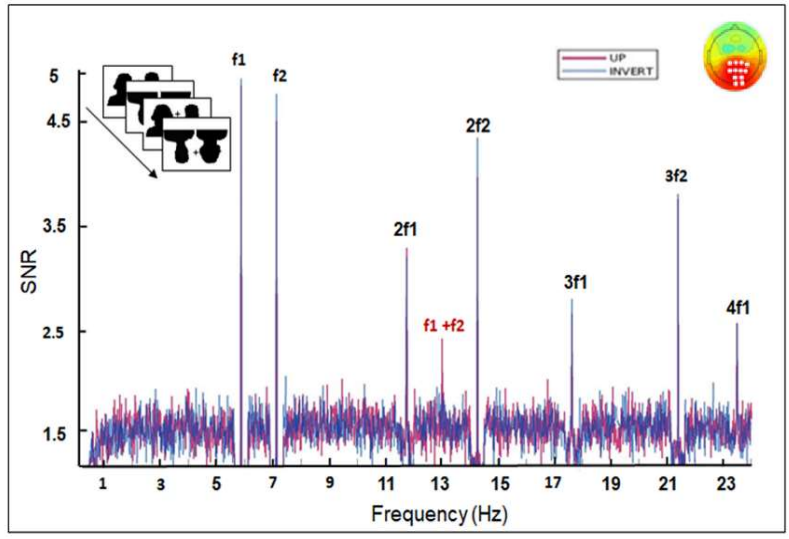

Figure 1: Signal to Noise Ratio (SNR) of EEG signals (0–24 Hz frequency range) averaged across participants and cortical region of analysis (white dots on topographical map). Responses at stimulation frequencies and harmonics identify the visual processing at the dyad elements-level. A clearly distinct response, at the second order sum IM component (f1+f2), probes visual processing at the global dyad shape-level. The latter response is significantly higher when silhouettes were presented Upright than Inverted. Reproduced from Mersad and Caristan [1].

latter case the plurality was consisting only of two humans.

A tentative framework encompassing these two mechanisms is that distinctive perception of human groups is useful for having access to a picture of the *majority trend* by generating a compressed representation of human ensembles, while disregarding precise information about individuals. Note that the two mechanisms cited above are likely to be complementary: while summary statistics allows gathering average information with better speed, holistic processing captures properties at the global level that are not reducible to individual levels.

Interestingly, the latter Gestalt perceptual mechanism that captures emergent properties resulting from the interaction of the parts is not without pointing out multi-agent models describing unpredictable emergent behaviors that appears when collectives of agents are in interaction [11].

Our proposal is that holistic perceptual mechanism of human groups might intervenes in the identification and analysis of emergent behavior characterizing human pluralities. Indeed, these are critical abilities for a human being to be able to read intentions at the group level and to tune behavior and action consequently.

Future investigations should shed light on the nature of the link between group perceptual mechanisms and the adaptive action they trigger in the individual, extending the framework to variable sizes of human groups.

References

1. Mersad K, Caristan C. Blending into the Crowd: Electrophysiological Evidence of Gestalt Perception of a Human Dyad. Neuropsychologia. 2021 Sep 17;160:107967.
2. Boremanse A, Norcia AM, Rossion B. An objective signature for visual binding of face parts in the human brain. Journal of Vision. 2013 Sep 2;13(11):6.
3. Norcia AM, Appelbaum LG, Ales JM, Cottereau BR, Rossion B. The steady-state visual evoked potential in vision research: A review. Journal of Vision. 2015 May 1;15(6):4.
4. Rossion B, Boremanse A. Robust sensitivity to facial identity in the right human occipito-temporal cortex as revealed by steady-state visual-evoked potentials. Journal of Vision. 2011 Feb 1;11(2):16.
5. Reed CL, Stone VE, Grubb JD, McGoldrick JE. Turning configural processing upside down: part and whole body postures. Journal of Experimental Psychology: Human Perception and Performance. 2006 Feb;32(1):73.
6. Koffka K. Principles of Gestalt psychology. Routledge; 2013 Oct 8.
7. Haberman J, Whitney D. Ensemble perception: Summarizing the scene and broadening the limits of visual processing. From perception to consciousness: Searching with Anne Treisman. 2012:339-49.
8. Navon D. Forest before trees: the precedence effect of global features in visual perception. Cognitive Psychol. 1977;7:476-84.
9. VanRullen R, Thorpe SJ. The time course of visual processing: from early perception to decision-making. Journal of Cognitive Neuroscience. 2001 May 15;13(4):454-61.
10. Gibson JJ. The ecological approach to visual perception: classic edition. Psychology Press; 2014 Nov 20.
11. Schelling TC. Micromotives and macrobehavior. WW Norton & Company; 2006 Oct 17.