Vision Research 157 (2019) 252–263

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

When the whole is equal to the sum of its parts: A new approach to study face and body perception and representation

Baingio Pinna*, Katia Deiana

Dept. of Biomedical Sciences, University of Sassari, Italy

ARTICLE INFO

Keywords:
Visual perception
Face perception
Perceptual organization
Gestalt psychology
Canonical perspective

ABSTRACT

In this paper, a new approach and a novel method to study face perception is proposed and tested using several qualitative experiments. This method is based on three main tasks: a description task (subjects were asked to freely describe the target stimulus), a free pictorial task (free drawing/painting of what subjects were asked), and a pictorial reproduction task (making a copy of what subjects perceived). These tasks were carried out with children and adults and extended to conditions related to visual arts. The starting points of this work were the canonical perspective and the holistic processes involved in face perception. The aim of this work was to answer the two following basic questions: Are canonical perspective and holistic processes really effective for face perception? Is face perception other than the sum of its parts? The outcomes of the experiments clearly refuted the role of canonical perspective and weaken the holistic approach to face and body perception. The whole human body has been shown instead to appear as if built starting from every single component, therefore body and faces are like wholes, decomposable in a mosaic of juxtaposed independent components reduced to a reference image. In short, the whole is equal to the sum of its parts. Finally, our results also show evidence supporting the introduction of the notion of icon.

1. Introduction

The face is phenomenally the most important part of the human body. Located on the anterior (frontal, rostral) surface of the head, the face represents the region of senses and the center of expressions and emotions for most species and particularly for humans. It is commonly assumed that the face subsumes the set of features which best distinguish a person from others, thus it represents the keystone for human identity. By being the whole that embodies the entire identity, then face changes, local transformations, component alterations or damages, due to aging, accidents, the presence of scars or deformities, have important effects in the whole identity in such a way that they usually elicit personal and social consequences.

1.1. Canonical perspective of faces

There are several preliminary issues and problems useful to understand what is a face. They are related to the following questions. When we see or imagine a face what is its expected or preferred view? Which are the consequent facial main components? When we look at and think of another person what do we expect to see? What is the role of those expectations in processing sensory inputs and in representing what we perceive? When we look at a face or at a body, do we have a clear expectation of its basic structure in terms of each single element component and of their overall configuration? Can prior expectations at a perceptual and/or cognitive level trigger biases in the perception of facial and body attributes?

Palmer, Rosch, and Chase (1981) systematically studied perspective preferences in object categorization demonstrating what they called “canonical perspective”, i.e., the view that provides the most “diagnostic information” about an object. This is the information useful to best discriminate an object from others, on the basis of the views from which that object is most often perceived and on the basis of its familiarity. In a series of experiments, by using goodness rating, visual imagery, active selection and naming tasks, they demonstrated that participants consistently preferred the same viewpoints independently from the task. The canonical view is, in short, the most representative image of an object that first comes to mind when associating a name. This image elicits the most accurate and fastest recognition. In the case of a horse, Palmer et al. showed that the canonical perspective is the three-quarter view, that gives rise to the highest number of surfaces visible to the observer. In line with this theory, this is expected to occur also for a human head. The superior recognition effect imparted by the canonical perspective has been tested and confirmed in a large number

* Corresponding author at: Dept. of Biomedical Sciences, University of Sassari, V.le San Pietro 46/b, Sassari 07100, Italy.
E-mail address: baingio@uniss.it (B. Pinna).

https://doi.org/10.1016/j.visres.2018.03.011
Received 25 October 2017; Received in revised form 19 February 2018; Accepted 26 March 2018
Available online 20 July 2018

of studies (Bülthoff, Edelman, & Tarr, 1995; Cutzu & Edelman, 1994; Edelman & Bülthoff, 1992; Gomez, Shutter, & Rouder, 2008; Newell, Ernst, Tjan, & Bülthoff, 2001; Perrett & Harries, 1988; Perrett, Harries, & Looker, 1992).

It is worth to highlight for the next arguments and questions that canonical perspective refers to objects as wholes. However, ahead there is a composite visual object, whose components do not manifest the same phenomenal salience and importance. Ahead it includes a face, which, as previously stated, subsumes most of the useful information about a person identity. It also contains the back of the neck that is distinct from the face, i.e., it also includes the nape, its other side, which is much scarcely useful for a person recognition than it is the face. Moreover, these two main components can be seen from different angles. A face can be seen from an infinite number of angles and views in the 3D space. Fortunately, they are phenomenally reducible to the following: frontal, half-profile, full profile, back of the neck, from above and from below. A similar number of views can also be considered for the back of the neck. All these views are, of course, the result of a combination concerning the face and the back of the neck, except for the full frontal face and full nape, since they totally hide or occlude their opposite side.

The previous reduction, from infinite possibilities to only six views, is especially interesting. They emerge spontaneously as privileged and biased orientations already described by Gestalt psychologists by introducing the notion of Prägnanz (Köhler, 1930; Koffka, 1935; Metzger, 1953; Wertheimer, 1922). Furthermore, they are even more interesting if we consider that not all of them have the same phenomenal salience, but some appear to be far more singular and expected than others. It is an obvious thing to say that frontal views and full profiles are phenomenally far more important than the back of the neck. For example, when we think of a well-known person, we immediately imagine its frontal face. However, even when we see its back of the neck, we imagine his/her frontal face. It is common sense that, although we clearly remember a frontal face, the back of the neck is almost entirely unknown even if it is seen with about the same frequency of the frontal view. Full profiles are also more difficult to remember than frontal views (see Balas & Thomas, 2015; Royer et al., 2016).

These phenomenal remarks highlight some issues deduced from the theory of canonical perspective if we take it literally. Relevant questions are as follows: since each object is composite, is canonical perspective related to the whole or to every single component? If it is related to the whole object, according to which the canonical view is imparted and belong only to the whole object, as can be derived from the studies of canonical perspective, why do the single components should not play any role in creating a canonical view? Why should they be discharged? A further derived question is: how are single components combined when they manifest different weights as in the case of the back of the neck and the frontal face? Conversely, if a specific canonical view is processed for every single component, how are these views combined within a whole? What is the expected canonical perspective of the whole under these conditions?

1.2. Holistic processing of faces

Beyond the canonical perspective of the face, there are supplementary issues and problems useful to understand what is a face. Face perception is generally considered to be different from other visual objects since it is demonstrated to be more holistic (see Farah, Wilson, Drain, & Tanaka, 1998; Tanaka & Farah, 1993; Young, Hellawell, & Hay, 1987; Tsao & Livingstone, 2009): faces are wholes non-decomposable in a mosaic of independent components (eyes, mouth, nose, ears, hairs, etc.) and relations among them. It is commonly granted that face perception is not the mere sum of the perception of its parts and features, but a face is rather something meaningful in itself as a whole beyond its element components. A face is seen in its entirety, not by its individual parts. According to Gestalt psychologists, the face is one of the best examples of *Gestalten*, i.e., a whole that is not only more than the sum of the parts, but, at the same time, it is different since it determines the part properties and, in its turn it is determined by them (see also Wagemans et al., 2012). In this connection, Gestalt psychologists introduced the notion of systemic influence (Metzger, 1953), according to which the quality of a part depends upon the whole where this part is embedded and *vice versa* (Koffka, 1935; Köhler, 1930/1971). Briefly, the face perceived as a whole is something else than the sum of its parts and it is based on a part-whole systemic organization (Pomerantz & Kubovy, 1986; Biederman & Kalocsai, 1997; Palmer, 1999).

Some consequences of this approach are the following. A face is not perceived and built just by assembling its components. As a corollary, each face component, in its turn, can be considered as a whole, and as such it is not perceived just by assembling its further inner subcomponents (Lerner, Hendler, Ben-Bashat, Harel, & Malach, 2001; Schiltz & Rossion, 2006; Kubilius, Wagemans, & Op de Beeck, 2011).

In addition, the process of perceptually building and representing a face is expected to start from the whole and then to proceed toward the inner elements (whole-part portrayal effect). Moreover, the final face portrait is also expected to be realistic, recognizable, predictable, related to prior knowledge. In other terms, face perception combines current sensory evidence with prior expectations (see Büthoff & Mallot, 1988; Landy, Maloney, Johnston, & Young, 1995; Rosas, Wichmann, & Wagemans, 2007) about the structure of the face. The Bayesian approach provides a formal description of the way in which this occurs (see Balas, 2012; Balas & Nelson, 2010; Kersten, Mamassian, & Yuille, 2004; Kersten & Yuille, 2003; Landy & Kojima, 2001; Maloney, 2002; Mamassian, Landy, & Maloney, 2002; Landy et al., 1995). Finally, all inner antinomies among components are assumed to be embodied, charged or cancelled within the resulting perception of the whole face (Schiltz & Rossion, 2006; see also Harris & Aguirre, 2008; Andrews, Davies-Thompson, Kingstone, & Young, 2010; Schiltz, Dricot, Goebel, & Rossion, 2010; Goffaux, Schiltz, Mur, & Goebel, 2012). This is, for example, the case of deformities, of different kinds of element changes or of alterations that become part of the face as a whole.

Given the systemic hypothesis in the part-whole organization, a unidirectional dominance of the whole upon the component attributes, i.e., the holistic unidirectional primacy, cannot be fully assumed. As a matter of fact, the configural dominance, studied in many different contexts beyond face perception, has been demonstrated only for some configurations but not for others (e.g., Pomerantz, 1981). Moreover, it has been found to depend on the task (Han, Humphreys, & Chen, 1999; Pomerantz & Pristach, 1989).

In spite of these results, as far as face perception is concerned, the holistic processing has been considered as distinct from non-face object processing (Kobatake & Tanaka, 1994). The composite effect has been demonstrated by showing that when subjects attempt to process only part of the face, they suffer interference from the other parts. This entails a lack of access to element components and a mandatory processing of the face as a whole (Diamond & Carey, 1986; Young et al., 1987). The composite effect has been confirmed also in other contexts (see Coltheart, Curtis, Atkins, & Heller, 1993; Anstis, 2005).

These hypotheses and results run against the well-known tool for the apprehension of suspects based on the process of facial composite production. Face composites are used by law enforcement agencies to identify people who commit crimes. To build a composite, witnesses or victims usually describe a suspect's face, selecting and choosing what appears to them as the most appropriate set of facial features. The resulting face is then obtained from a composite production and representation (see Frowd, Bruce, et al., 2004, Frowd, Carson, et al., 2004, Frowd, Hancock, & Carson, 2004; Koehn & Fisher, 1997).

Given these preliminary experimental and theoretical outcomes, the main purpose of this paper is to introduce a new approach and a novel method useful to reexamine and revise what we know about face perception. Moreover, the previous issues and questions are here reviewed,

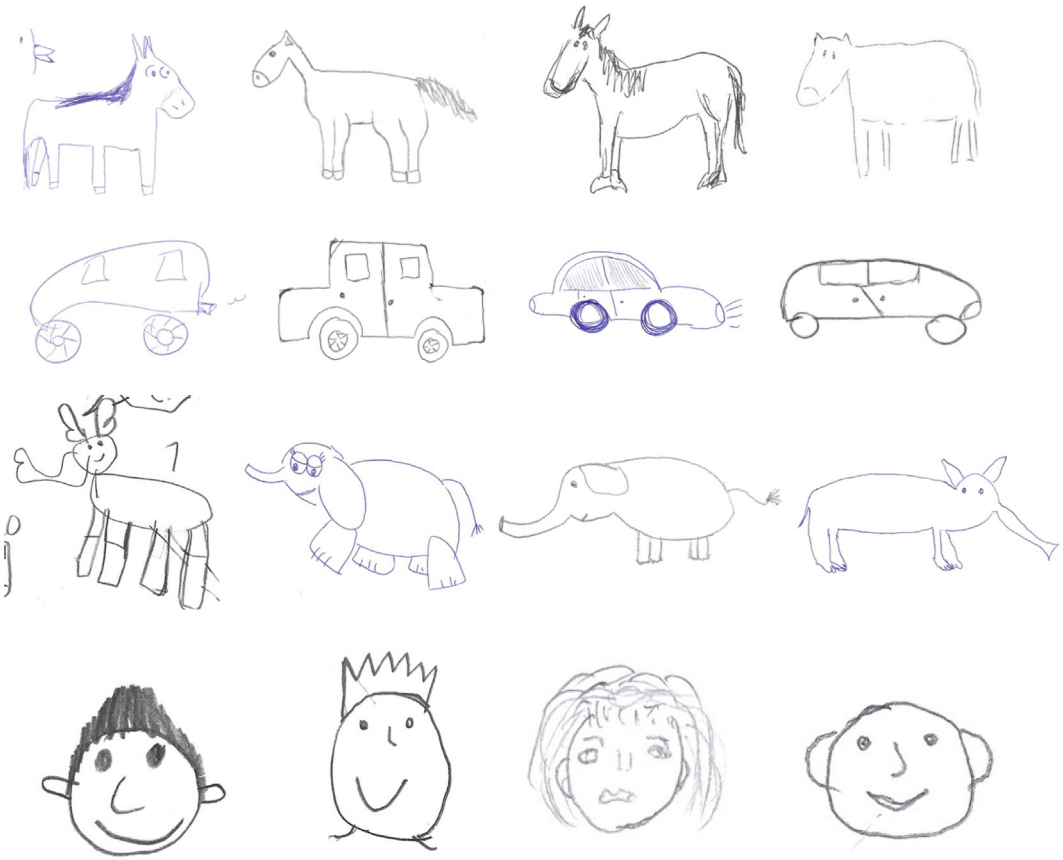

Fig. 1. Examples of the way children and adults freely depict horses, cars, elephants as if they were seen from the full left or right side. The human head was instead mainly depicted as if it was seen from the front.

answered and subsumed within the following two more general questions: Are canonical perspective and holistic processes really effective for face perception? Is face perception other than the sum of its parts?

2. General methods

2.1. Subjects

For each stimulus, described in the procedure and in the Results section, different groups of 50 children, ranging from 5 to 7 years old, and 50 undergraduate students of psychology, architecture, social service, design and linguistics, from 20 to 24 years old, participated in the experiments. Both children and undergraduate students were naive both to this topic and to our purposes. Within each group, subjects were about 50% male and 50% female and all had normal or corrected-to-normal vision.

The children were recruited from schools. To test the children, a special permission was obtained from the schools (headmasters and teachers) and the children's parents and under preliminary informed consent signed by all participants and by parents or legal guardian in compliance with the Helsinki declaration. Children were tested individually in a quiet room and always under the constant supervision of the experimenter, who remained neutral without making any suggestions or giving any sort of help which might have spoiled the genuine production of the subjects. The undergraduate students spontaneously participated in the experiments.

2.2. Stimuli

The stimuli were all the drawings, pictures and paintings as illustrated in the following sections, displayed on an Apple IPAD in an ambient illuminated by an Osram Daylight fluorescent light (250 lx, 5600 K). Stimuli were viewed binocularly in the fronto-parallel plane at a distance of 50 cm.

2.3. Procedure

The tasks of the subjects were to describe (description task), to draw/paint (free pictorial task) what they were asked and to reproduce by making a copy (pictorial reproduction task) what they saw on the IPAD screen. As a general example of the description task, subjects were simply asked: "could you please describe what you see?". During the free pictorial task, participants were, for instance, asked: "could you please draw a horse?". In the pictorial reproduction task, the question was "could you please make a copy of this (what was displayed on an Apple IPAD)?". The specific kind of task used for each experiment is described in the Results section. The descriptions reported in the next sections were those spontaneously communicated by more than 95% of the subjects.

Before starting the experiment, a paintbox of 36 different colors was provided by the experimenter to each subject as well as an eraser. The colored pastels could be freely used. Children were also free to change, adjust and erase mistakes or any other unwanted result. Subjects were tested individually. There were given no time limits to both descriptions, pictorial and reproduction tasks. While, the descriptions occurred spontaneously and fast, the duration of the pictorial and reproduction tasks changed according to the skills, the age and various other factors peculiar to each subject. The stimuli were shown continuously during the experiment.

Fig. 2. Examples of portraits made by artists showing the frequency dominance of the frontal view (a), followed by the profiles (b) and, finally, by the very rare painting of the back of the head (c). Picasso's portraits with juxtaposed mixed views are mostly perceived as full frontal face portraits (d).

2.3.1. Pros and cons of the direct and indirect representation of canonical perspectives and holistic views

One of the tasks used by Palmer et al. (1981) to study perspective effects in object categorization is very related to our method. They showed many pictures of the same object such as horses or cars and collected subjective ratings using a scale from 1 (very much like) to 7 (very unlike) about how much every single instance looked like the objects they depicted. Our methods are much more direct than the one used by Palmer et al. By asking to freely draw a horse or a car (free pictorial task), to reproduce (copy) what subject they were shown in a model (pictorial reproduction task) or to describe what they see and draw (description task), we should expect a more direct representation of the canonical target object. Another advantage in using these methods together, when they are applied to subjects of different ages, is the possibility to "observe" the potential development of the canonical object and its possible adaptations to the mental age and development stages of the subjects.

Despite these advantages, there are downsides, suggesting some basic limitations of our methods as tools to study canonical perspective. The most significant is the fact that with these direct techniques the main target is not necessarily what has been called "canonical perspective" but something similar/else that could be related to it but that is likely something different. While the subjective rating used by Palmer et al. plays with different perspectives and allows a rated choice among them, our methods do not enable any selection among alternatives but ask directly to depict or draw something that might precede the canonical perspective. This could be some kind of primitive imprinted image of the target object without any perspective or with a perspective not yet included in any 3D space. In spite of this difference, our tasks might reveal some strong relations between the two kinds of "canons". Moreover, our direct methods could be more suitable and helpful than the subjective ratings to understand what is a face and how a face is imprinted as a canonical image and consequently used in object and face categorization.

There is a further non-negligible issue that it is worthwhile considering. In drawing the image they have in mind, subjects are depicting and representing something that requires some painting ability and, thus, they likely experience technical drawing difficulties. Therefore, what emerges at the end is not the exact image of the canonical object but it is just its "imperfect" representation that could be easily contaminated by previous experiences of drawing the same or similar objects or just by what has been learned at school or through past experiences.

In addition, the main consequence of the drawing difficulties is related to the quality of the resulting image, which is expected to be the simplest image that could possibly be drawn, some kind of draft or basic schema of the "true" image which is in the subject's mind. However, in light of this, the same issue can be reconsidered as an advantage since this schema might represent a basic primitive structure, a true perceptual canon of the object similarly to the ancient Greek canon used in sculpture, for example by Polykleitos. The term "canon" comes from the Greek word κανών, meaning principle, standard, measure, law, or rule accepted as axiomatic and universally binding. A canon sets the standards of *symmetria* representing an ideal mathematical proportion among the parts of the human body suggesting as sculpture of the human figure, a dynamic symmetry or equilibrium between relaxed and tensed parts of the body and between the directions where each part moves. Something similar occurs in Egyptian canon of proportions. We suggest that the pictorial and representation outcomes of our tasks can be considered as drawings related to the standardization of proportions similar to these schemas or canons. Therefore they do not necessarily need to be the exact pictures of the image within the subjects' mind.

Last but not least, a final relevant argument is derived from the assumption that the outcomes expected with our methods reproduce some general and common canon of views and proportions among parts and whole. If this is true, then most of our previous questions and issues

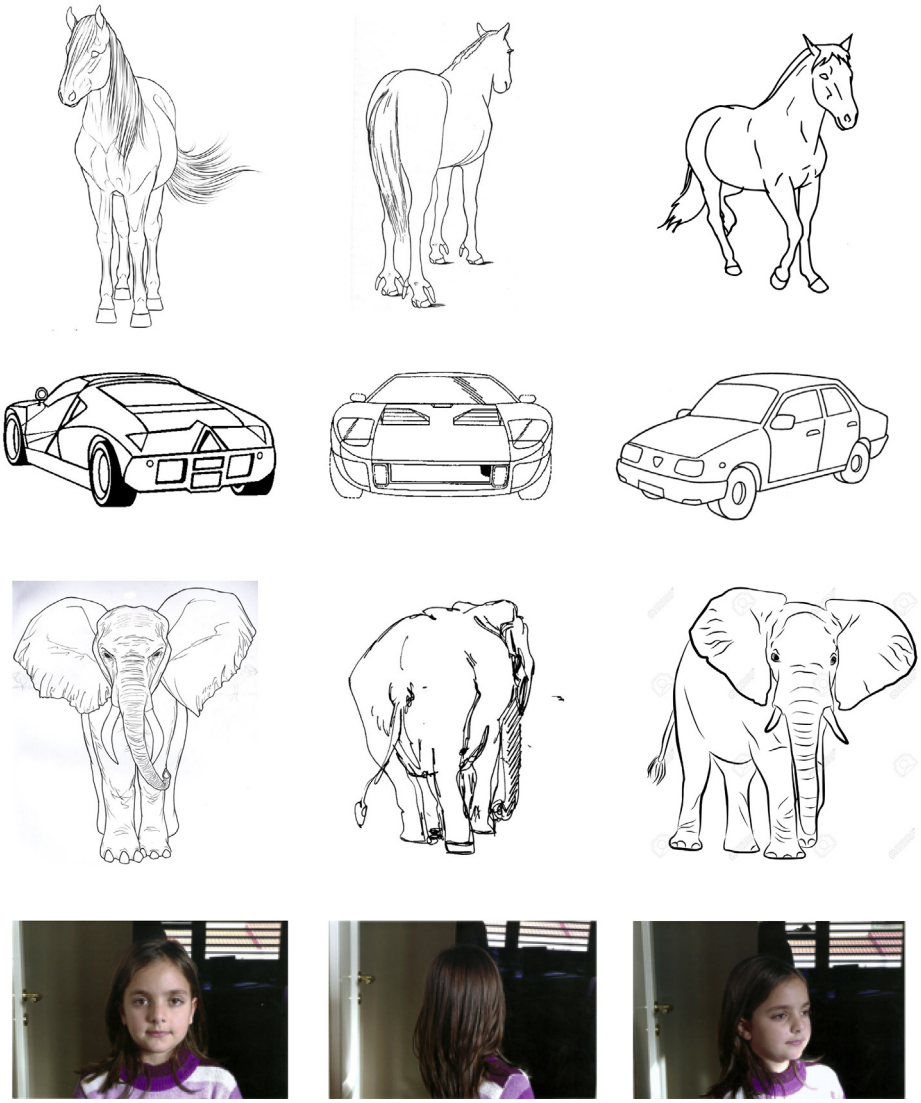

Fig. 3. Three views (front, back and three quarter) of each target used as stimuli in the pictorial reproduction and description tasks.

related to the holistic process of face perception might be answered.

All these issues and arguments will be deepened during the discussion of results. In any case, the new methods here used may cast fresh light within the studies on canonical perspective and, more importantly, on face perception.

3. Results

3.1. Exp. 1. How horses, cars, elephants and faces are freely drawn

The first results here reported are those achieved with the free pictorial task of a horse, a car, an elephant and a human head. In line with the canonical perspective, these stimuli are expected to be painted in the three-quarter view. Some most representative examples of the outcomes are reported in Fig. 1.

In the case of children, horses (mean 93%, SD 5%, from now on 93-5), cars (95-6) and elephants (94-4) were significantly painted as if they were seen from the full left or right side. Although the three-quarter views were rarely painted, they were more frequent than the frontal and back sides with the frontal side more frequent than the back side that was never used. Adults drew these stimuli similarly but with a lower but still significant percentage: horses (76-6), cars (80-6) and elephants (72-7). It is worthwhile noting that by searching with Google and counting the number of different views, for example, of images of drawn horses, a similar order and frequency can be found. If the free painting task reveals the basic primitive structure of an object, this is not surprising but, all the more reason, it is fully expected and supporting our results.

The human head was mainly depicted as if it was seen from the front, i.e., by showing the full face (children 99-1; adults 90-4). The three-quarter views were much less frequently painted, the profiles came soon afterward and not any back side was depicted. Again, by browsing in the web portraits made in the history of art, it is very easy to find full frontal faces, very few profiles and very rare back sides (see Fig. 2a–c). It is interesting in the case of Picasso's portraits that the mixture and juxtaposition of different views within the same face are unnoticed but it appears as a full frontal face portrait (see Fig. 2d). According to these results, we suggest using the internet as a precious source of data already collected to study phenomena like face perception related to basic primitive canons, biases and shape prototypes as suggested by Rosch et al. (1976).

These outcomes weaken the three-quarter view predicted by the canonical perspective theory and suggest that only one of the sides is considered and elected as the most representative for the target object, i.e., its best example and the reference instance of all the possible ones. In other terms, one side only behaves like a reference image and, as

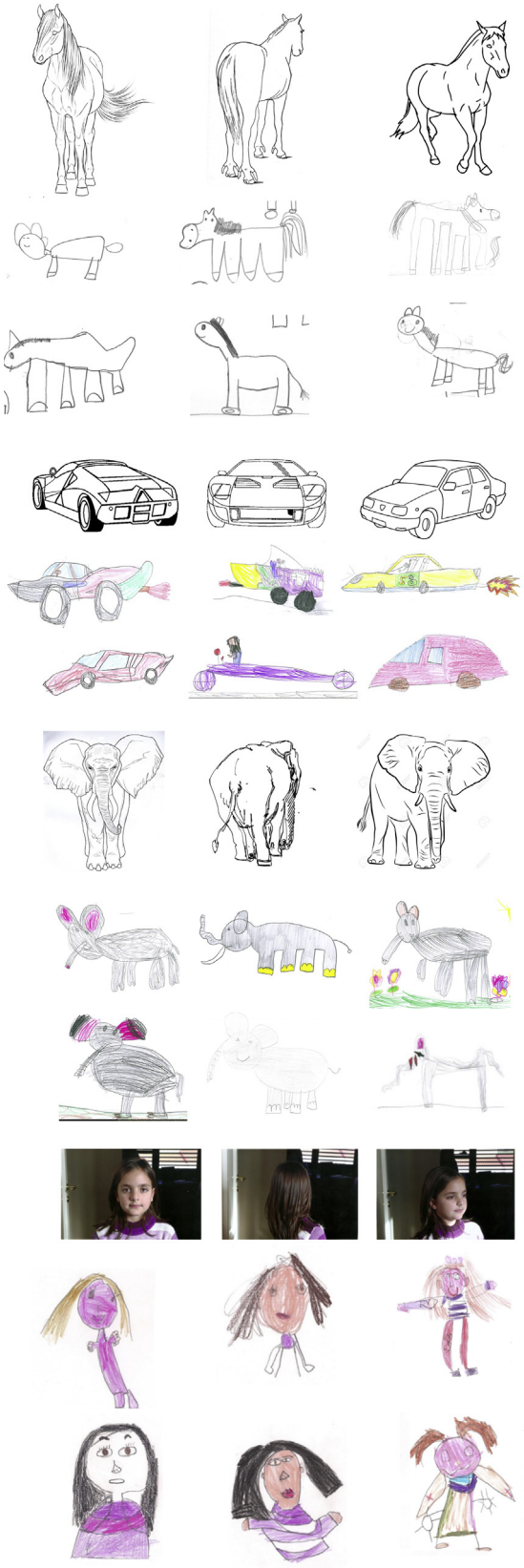

Fig. 4. Examples of children pictorial reproductions placed adjacent to the target stimuli. Horses, cars, elephants were copied as if they were seen from the full left or right side, while the human head was copied as if it was seen from the front. The heads are turned to show the frontal face also children perceived only the back of the neck.

such, it is designed to put together and subsume as a representative unity all the other possible sides, views and instances of the object.

If this is true, a way to test and challenge the strength of this kind of unitary organization is to adopt the pictorial reproduction task and by means of it to replicate the previous experiment.

3.2. Exp. 2. How horses, cars, elephants and faces are freely described and copied

The basic argument is the following: if the representative side behaves like a canon or like a basic idea of that specific object then it can also play a role even when the task is that to make a copy of the target object. Moreover, as in the previous task, through the pictorial representation, we expect a variation in strength by changing the age.

To explore more deeply the effectiveness of this task and of the reference image, three variations related to three views for each target were presented to three different groups of subjects. The three views are illustrated in Fig. 3. The pictorial reproduction task was supported by the description of the perceived model to be copied. The free description was introduced to be sure that each subject has recognized what was represented in the stimulus and to see if the description contains information about the illustrated view. For example, a description stating “an elephant seen from the back” is different from “an elephant”. This difference is important if it is in correlation with the result of the reproduction task. In the first case, the expected outcome is an elephant seen from the back, while in the second just an elephant as illustrated in Fig. 1. In other words, the name of the object tout court indicates the predominance of the free representation, while the description, enriched by the orientation of the object, is expected to be reproduced accordingly.

The results supported the free descriptions in line with the free reproductions. Children drew the horses mostly like they did with the free pictorial task although the mean percentages are reduced in the case of the models seen from the front and from the back (horse-front 85/5, horse-back, 83/6, car-front 82/6, car-back, 83/5 and elephant 83-5). The results of the adults were more consistent with the models (about 5-4 for all the targets). Some of the most representative and surprising outcomes made by children are illustrated in Fig. 4 next to each target stimulus.

The case of the faces reproduced by children is consistent with the results of the other objects although they might appear much more unexpected and astonishing especially in the case of the back of the head. The heads are turned to show the frontal face also in the extreme condition where they perceived only the back of the neck. These results demonstrate a compelling tendency to be in front, a tendency that should not be limited to this specific conditions made by children. As a matter of fact, also Fig. 2 supports this tendency. However, more convincing data are clearly illustrated in Byzantine, Orthodox and Ethiopian icons as shown in Fig. 5. Here, all the characters of the paintings, bas-reliefs, and mosaics show their frontal face even when it elicits a holistic unreal effect with superimposed halos or with strange scenes. This is the case of Christ's crucifixion (Fig. 5a), where although all the characters around Christ crucified appear loving and suffering for him, their face is not raised toward Christ but toward the observer, as if they were part of a selfie as it happens nowadays. Although the holistic effect appears unreal, it goes most of the time unnoticed. Only if someone makes us see it, we can become aware of it. This can be accounted for as being part of the natural tendency to be in front. All the examples shown in Fig. 5 reveal this strong tendency that explains also the contemporary biases and fashions to take pictures and make selfies.

3.3. Exp. 3. On human body representation: from canonical perspectives to icons

Besides these supporting effect of the previous icons, they can be even more useful for our purposes since they can let us understand and

Fig. 5. Examples of the tendency to be in front in Byzantine, Orthodox and Ethiopian icons as if the represented subjects were part of a contemporary selfie.

define more deeply the kind of basic reference schema children are dealing with when they perform the pictorial reproduction task. Although related to, this schema cannot be fully assimilated to the canonical perspective as studied by Palmer et al. (1981). Our results and the icons of Fig. 5 demonstrate something else that precedes the choice of a canonical perspective, something to which all the possible instances of a single object and inside it, of every single component are reduced. This schema is an image that is derived from its instances but not reduced to any of them, not to any specific perspective. Moreover, it may lead to an unreal or wrong representation that goes unnoticed and tends to appear normal and natural.

If these differences are correct it is epistemologically necessary to distinguish and separate the two issues by introducing a new term for the phenomenon emerging in this work.

The term “symbol”, from Greek “σύμβολον” derived from “συμβάλλω”, might be useful for our purposes, meaning “putting together, bringing together and matching”. However, it suffers the fact that in semiotics it is related to the connection between signifier and signified that is completely arbitrary and must be culturally learned.

A better candidate is the term “icon”, from Greek “εικών -όνος”, as described by the online etymology dictionary and meaning “image, figure, picture”, also “statue”, from Late Latin icon “likeness, image, portrait; image in a mirror; a semblance, phantom image”; in philosophy, “an image in the mind”, related to eikenai “be like, look like”. In semiotics the icon bears a physical resemblance to what is being represented, thus manifesting a real connection between the signifier and the signified. A further pro in favor of the term Icon is related to the use of them in Eastern Orthodox Church, especially associated with portraits and considered as visual 2D representations of Christ not very lifelike or literal.

The choice of the term “icon”, with the meaning previously described, suggests several challenging hypotheses that are relevant for our purposes. If icons are not directly connected with views, indeed beyond canonical perspectives, then an object and its element

Fig. 6. Free drawings related to different parts of human body.

Fig. 7. Examples of a posteriori assembled male and female bodies made by using the parts of the bodies drawn separately.

Fig. 8. Examples of the whole human body, drawn freely by children and adults, that appears as the juxtaposition of the free drawing of every single component (cf. these results with those of Fig. 7).

components have their own iconic model to refer during the pictorial reproduction task. This entails that the wholeness of an object is the result of some kind of juxtaposition or summation of the icons of each component. In other terms, by taking an object clearly made of several parts, like a human body, the main point emerging from the results of the previous experiments is that, under the pictorial reproduction task, icons behave like tiles in an overall mosaic staying together as juxtaposed elements to compose the whole object.

To test this deduction, a first group of children and adults were asked to paint/draw the following parts of the human body: nose, eye, ears, hair, shoulder, hands, legs, feet. The task of the second group was instead to paint/draw a human body.

The main hypothesis was that the overall body built starting from the icons of the single parts, i.e., putting them together, would have been at the end very close or the same as the icon of the whole body created independently from the single parts.

Fig. 6 shows the most frequent outcomes for every single component of the human body as drawn by children and adults. The results reveal again that children and adults draw spontaneously similar icons (nose 75-5, eye 100-0, ears 98-2, hair 85-5, shoulder 99-2, hands 88-6, legs 89-4, feet1 55-5, feet2 45-5). Furthermore, adults drawings are more related to specific perspectives and they are also more realistic than those drawn by children, which are on the contrary, more abstract and beyond any spatial orientation.

The resulting overall male and female bodies created a posteriori just by putting together or assembling the most frequent icons of the body components taken separately is illustrated in Fig. 7.

In Fig. 8, some results of the whole human body drawn freely by the second group of children and adults are illustrated. The above considerations made by children and adults occurred also under this condition (children 92-5, adults 85-6). By comparing the assembled icons of Fig. 7 with the holistic ones drawn by participants of Fig. 8, strong qualitative similarities emerge. Under our qualitative conditions, the holistic approach to face and body perception is clearly weakened. Since the whole human body appears as built starting from the icons belonging to every single component, our data suggest that body and faces are like wholes decomposable in a mosaic of independent components in the form of juxtaposed elementary icons. Although it is

Fig. 9. The five basic ballet positions of the feet are very similar to the outcomes of the free paint/draw task of only parts of the human body (legs and feet in this case).

commonly granted that a face and a human body are not the mere sum of the perception of their parts and features, but rather something meaningful in themselves as a whole beyond their element components, the answer to the question "Are they other than the sum of their parts?" is, under our conditions, negative (see also Gold, Mundy, & Tjan, 2012, 2014; Shen & Palmeri, 2015).

Further interesting properties suggested by these results are related to the unnatural, unreal, not familiar and impossible human body posture. Although these kinds of results can be considered as childish stuff, even if also adult subjects reported similar drawings, indeed they are not really childish. As a matter of fact, in the classical ballet, there are five basic positions in which both feet and arms are required in certain positions (see Fig. 9), what needs to be stated here is that at least two of these positions are likely related to the icons drawn by our subjects when their task was to freely draw legs and feet. The basic first position in the ballet is the position with the feet standing with their heels together and the toes going outwards. The second position is similar to the first, but the feet point in opposite directions, with heels, spaced more or less twelve inches apart.

The question is now: are the icons painted by our subjects influenced by prior knowledge of the ballet positions or are these positions a direct consequence of the icons of the feet as drawn by our subjects? The right answer should be the second if the reader did not connect the body icon to the ballet positions before reading it in the previous words. This argument suggests that the canons of beauty and the aesthetic appreciation, for example in art, are likely related to the icons as here qualitatively described and, more formally, defined. Moreover, what appears to be childish is a basic root that continues in adult life, as demonstrated with the ballet positions. Finally, prior expectations and knowledge might depend on the building process ruled by icons and not *vice versa*. This possibility weakens the assumption that face and body perception combine current sensory evidence with prior expectations as suggested in a Bayesian approach.

3.4. Exp. 4. Frontal and profile face representation: when and why of Egyptian art

The introduction of the notion of "icon" has been helpful in suggesting new conditions but it can also be helpful to solve a strong contradiction and counterexample in the public eye that some reader might have noticed. Since we previously demonstrated the strong role of the frontal face as a reference icon for the pictorial reproduction task, the Egyptian representation canon seems apparently to contradict our findings.

First of all, each character is depicted in a view that art historians called twisted perspective, which is considered as a valid explanation for the way the head was almost always shown in profile, with a full view of the eyebrow and eye on it. The shoulders were mostly shown frontally, but the waist, buttocks, and limbs in profile. Legs drawn in profile in order to show the knees. Nipples on male figures and breast on females were drawn in profile, while collars, necklaces, pectorals and clothing in frontal view. The navel was full view. The two feet were mostly depicted identically (see Fig. 10). The twisted perspective could be considered as an ingenious way to project a three-dimensional world on a two-dimensional surface without creating the illusion of the third dimension by means of well-known depth cues.

The first immediate remark to this kind of explanation is the fact that the twisted perspective does not introduce any clear depth effect, but all the painted stories are seen as taking place on a flat frontal two-dimensional space. This argument weakens and disassembles the explanation of the multiple views belonging to each body part. However, if the icon model here proposed can explain these multiple views and the fact that, as previously shown, the human figure is formed from a composite built up from its individual icons, the very critical and intriguing matter for our model is the head depicted in profile.

To solve this contradiction, we can go back to the previous remark according to which the objects, Egyptians wished to depict, were placed over a flat drawing surface. Moreover, another basic phenomenal

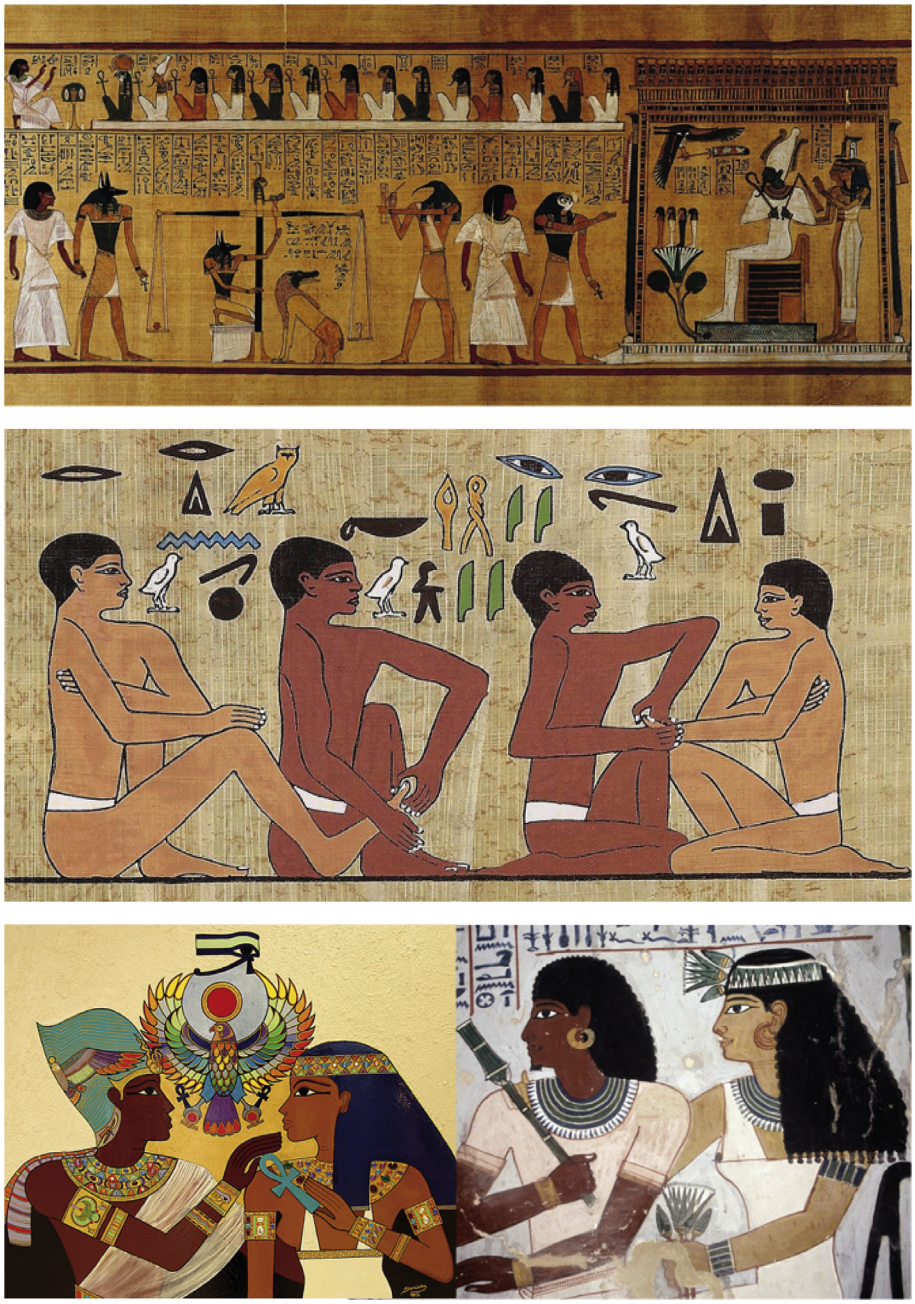

Fig. 10. Examples of Egyptian art where each character is depicted in a view that art historians called twisted perspective.

observation is the fact that all the characters depicted appear to be part of a story that is happening along the flat storytelling plane. In short, they are doing something. This suggests that the human bodies previously painted by our subjects were related to a static icon, this could be the reason why the full frontal face appears in contradiction with the Egyptian profiles. Therefore, the main hypothesis to be tested is how our subjects would have drawn the human bodies under dynamic conditions, i.e., when they are doing something along the flat horizontal plane.

In this experiment, subjects were asked to draw/paint a man or a woman moving left or right perpendicularly to subjects' viewpoints. The results, illustrated in Fig. 11, support the previous hypothesis. More in details, the profiles are now represented as expected even though the composite assemblage previously described remains with each component painted according to its own icon.

It is worthwhile highlighting that also in Egyptian Art what can be considered mistakes goes unnoticed, totally invisible or not disturbing at all. The main reason is again explained by means of the icons, considered as reference images and portraits. As such, they are not perspective mistakes but something similar to a caricature, which, although wrong and deformed, is perceived as righter and more straight than the original target to which the icon is referred. In Fig. 12, unnoticed mistakes are illustrated: a Picasso painting revealing the eyes painted frontally and a Liv Strömquist's illustration showing Adam with the face turned in the opposite direction of the arms, legs and feet but turned in the same direction of the chest and with the sight directed behind the face.

Again, these pictures and findings support the notion of icon as a basic reference element useful for object categorization and representation.

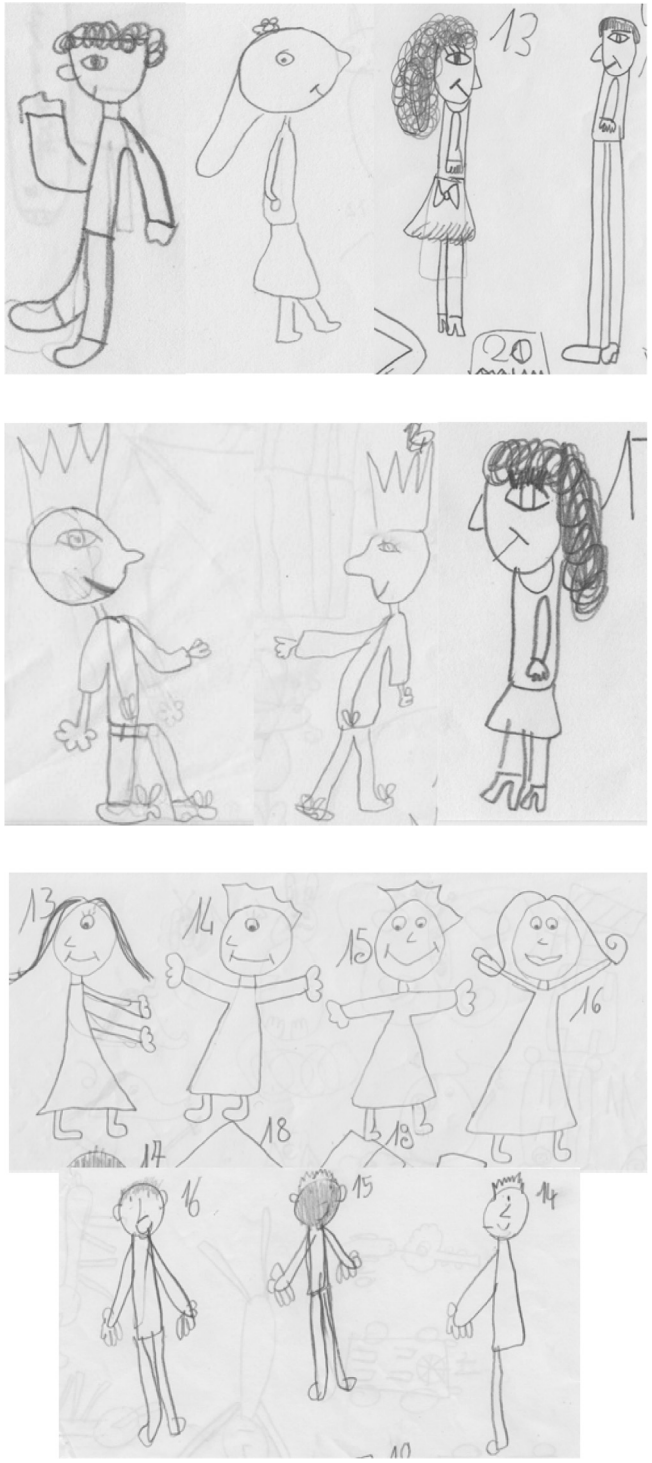

Fig. 11. Free pictorial drawings of human body moving left or right perpendicularly to subject's viewpoint. The emergence of the profiles under these dynamic conditions is similar to the way ancient Egyptians used to paint.

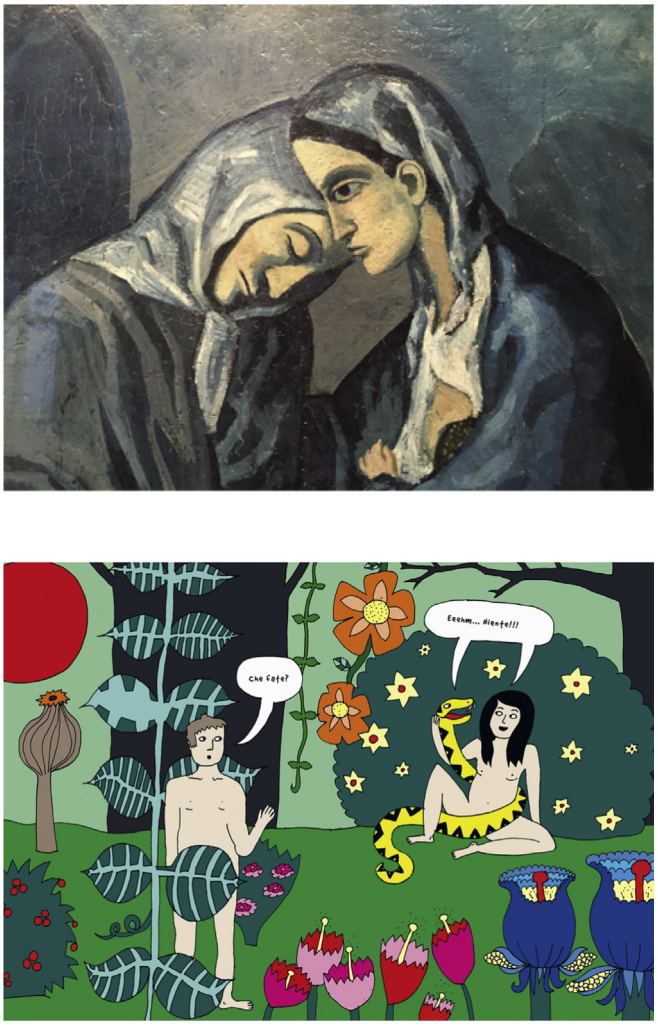

Fig. 12. Picasso's painting and Liv Strömquist's illustration showing significant "mistakes" are unnoticed, totally invisible or not disturbing at all.

4. Conclusions

In this paper, a new approach and a novel method to study face perception have been introduced and tested through several qualitative experiments. This method was based on three main tasks: description task (the subjects were asked to freely describe the target stimulus), free pictorial task (free drawing/painting of what subjects were asked); pictorial reproduction task (making a copy of what subject perceived). These tasks were used with children and adults and extended to conditions related to visual arts.

The starting points of this work were the canonical perspective first studied by Palmer et al. (1981) and the holistic processes involved in face perception. The main purposes were to answer the two following basic questions: Are canonical perspective and holistic processes really effective for face perception? Is face perception other than the sum of its parts?

The results of the experiments weaken both the role of canonical perspective and of the holistic approach to face and body perception. The whole human body has been shown instead to appear as built starting from every single component, therefore body and faces are like wholes decomposable in a mosaic of juxtaposed independent components reduced to reference images. In short, the answer to the question "Are faces and bodies other than the sum of their parts?" is, under our conditions, not positive (see also Gold et al., 2012, 2014; Shen & Palmeri, 2015).

In contrast with the canonical perspective, our results prove and endorse the introduction of the notion of icon, that is a primitive non-literal image of an object used for categorization and as a reference tool to reproduce pictorially any object. The icon subsumes all the instances of an object and elicits its connection with other objects by means of possible juxtapositions and linear summations. More in detail, the icon reduces the number of possible instances of an object to only one not necessarily present among the true instances. The process of *Reductio ad Unum* using the icon elicits an effective and efficient use of the objects

and a minimum information load. Moreover, by means of the linear dynamics here demonstrated and discussed, it becomes very easy to combine icons to create iconic wholes that are combined in their turn within other icons and so on. The evolutionary advantage of creating internal icons is based on the need to develop and elicit a quick and low-effort means by which understand, recognize and categorize the external world with all its objects.

Using the notion of icon, a high number of effects were explained including the apparent contradiction shown by the face profiles of Egyptian art.

Although more quantitative data are necessary for a deep math analysis of the notions here developed, these results show a new way useful to further explore and deepen the understanding of the laws aimed to build icons and to create new icons starting from more elementary ones. Future studies with the same methods are in progress addressed to children and adults belonging to different cultures. Still open questions related to the notion of icon are as follows: how are icons created and for what main purpose beyond the reproduction and the pictorial tasks? Do they change across time? How do they change during psychological development? How are they related to the canonical perspectives?

Acknowledgement

This work is supported by Fondazione Banco di Sardegna (finanziato a valere sulle risorse del bando Fondazione di Sardegna – Annualità 2015).

References

Andrews, T. J., Davies-Thompson, J., Kingstone, A., & Young, A. W. (2010). Internal and external features of the face are represented holistically in faces elective regions of visual cortex. *Journal of Neuroscience, 30*, 3544–3552.

Anstis, S. (2005). Last but not least. *Perception, 34*, 237–240.

Balas, B., & Thomas, L. E. (2015). Competition makes observers remember faces as more aggressive. *Journal of Experimental Psychology: General, 144*, 711–716.

Balas, B. (2012). Bayesian face recognition and perceptual narrowing in face-space. *Developmental Science,* 578–588. https://doi.org/10.1111/j. 1467-7687.2012.01154.x.

Balas, B., & Nelson, C. A. (2010). The role of face shape and pigmentation in other-race face perception: An electrophysiological study. *Neuropsychologia, 2010*(48), 498–506.

Biederman, I., & Kalocsai, P. (1997). Neurocomputational bases of object and face recognition. *Philosophical Transactions of the Royal Society B: Biological Sciences, 352*, 1203–1219.

Bülthoff, H. H., Edelman, S., & Tarr, M. J. (1995). How are three-dimensional objects represented in the brain? *Cerebral Cortex, 5*, 247–260.

Bülthoff, H. H., & Mallot, H. A. (1988). Integration of depth modules: Stereo and shading. *Journal of the Optical Society of America A, Optics and Image Science, 5*, 1749–1758.

Coltheart, M., Curtis, B., Atkins, P., & Heller, M. (1993). Models of reading aloud: Dual-route and parallel distributed-processing approaches. *Psychological Review, 100*, 589–608.

Cutzu, F., & Edelman, S. (1994). Canonical views in object representation and recognition. *Vision Research, 34*, 3037–3056.

Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General, 115*, 107–117.

Edelman, S., & Bülthoff, H. H. (1992). Orientation dependence in the recognition of familiar and novel views of three-dimensional objects. *Vision Research, 32*, 2385–2400.

Farah, M. J., Wilson, K. D., Drain, M., & Tanaka, J. N. (1998). What is "special" about face perception? *Psychological Review, 105*(3), 482–498.

Freiwald, W. A., Tsao, D. Y., & Livingston, M. (2009). A face feature space in the macaque temporal lobe. *Nature Neuroscience, 12*, 1187–1196.

Frowd, C., Bruce, V., Ness, H., Thomson-Bogner, C., Peterson, J., Mcintyre, A., & Hancock, P. (2004). Parallel approaches to composite production. *Psychology Crime and Law*.

Frowd, C., Carson, D., Ness, H., McQuiston, D., Richardson, J., Baldwin, H., et al. (2004). Contemporary composite techniques: The impact of a forensically-relevant target delay. *Legal and Criminological Psychology*.

Frowd, C., Hancock, P., Carson, D., & EvoFIT. (2004d). A holistic, evolutionary facial imaging technique for creating composites. *ACM Transactions on Applied Perception (TAP), 1*, 1–21.

Goffaux, V., Schiltz, C., Mur, M., & Goebel, R. (2012). Local discriminability determines the strength of holistic processing for faces in the fusiform face area. *Frontiers in Psychology, 3*, 604.

Gold, M. J., Mundy, J. P., & Tjan, S. B. (2012). The perception of a face is no more than the sum of its parts. *Psychological Science*. https://doi.org/10.1177/0956797611427407.

Gold, J. M., Barker, J. D., Barr, S., Bittner, J. L., Bratch, A., Bromfield, W. D., Goode, R. A., Jones, M., Lee, D., & Srinath, A. (2014). The perception of a familiar face is no more than the sum of its parts. *Psychonomic Bulletin & Review, 21*(6), 1465–1472. https://doi.org/10.3758/s13423-014-0632-3.

Gomez, P., Shutter, J., & Rouder, J. N. (2008). Memory for objects in canonical and non canonical viewpoints. *Psychonomic Bulletin & Review*. https://doi.org/10.3758/PBR.15.5.940.

Han, S., Humphreys, G. W., & Chen, L. (1999). Parallel and competitive processes in hierarchical analysis: Perceptual grouping and encoding of 100 years of visual gestalts II 1243 closure. *Journal of Experimental Psychology: Human Perception and Performance, 25*, 1411–1432. https://doi.org/10.1037/0096-1523.25.5.1411.

Harris, A., & Aguirre, G. K. (2008). The representation of parts and wholes in face-selective cortex. *Nature Neuroscience, 20*, 863–878.

Kersten, D., Mamassian, P., & Yuille, A. (2004). Object perception as Bayesian inference. *Annual Review of Psychology, 55*, 271–304.

Kersten, D., & Yuille, A. (2003). Bayesian models of object perception. *Current Opinion in Neurobiology, 13*, 150–158.

Kobatake, E., & Tanaka, E. (1994). Neuronal selectivities to complex object features in the ventral visual pathway of the macaque cerebral cortex. *Journal of Neurophysiology, 71*, 856–867.

Koehn, C., & Fisher, R. (1997). Constructing facial composites with the Mac-a-Mug Pro system. *Psychology, Crime and Law, 3*, 215–224.

Köhler, W. (1971). Human perception. In M. Henle (Ed.). *The selected papers of Wolfgang Köhler* (pp. 142–167). New York, NY: Liveright (Original work published 1930).

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*, 1296–1303. https://doi.org/10.1177/0956797611417000.

Landy, M. S., & Kojima, H. (2001). Ideal cue combination for localizing texture-defined edges. *Journal of the Optical Society of America A: Optics, Image Science, and Vision. 18*, 2307–2320.

Landy, M. S., Maloney, L. T., Johnston, E. B., & Young, M. (1995). Measurement and modeling of depth cue combination: In defense of weak fusion. *Vision Research, 35*, 389–412.

Lerner, Y., Hendler, T., Ben-Bashat, D., Harel, M., & Malach, R. (2001). A hierarchical axis of object processing stages in the human visual cortex. *Cerebral Cortex, 11*, 287–297.

Maloney, L. T. (2002). Statistical decision theory and biological vision. In D. Heyer, & R. Mausfield (Eds.). *Perception and the physical world: Psychological and philosophical issues in perception* (pp. 145–189). New York: Wiley.

Mamassian, P., Landy, M., & Maloney, L. T. (2002). Bayesian modelling of visual perception. In R. P. N. Rao, B. A. Olshausen, & M. S. Lewicki (Eds.). *Probabilistic models of the brain: Perception and neural function* (pp. 13–36). Cambridge, MA: MIT Press.

Metzger, W. (1953). *Gesetze des Sehens [Laws of seeing]* (2nd ed.). Frankfurt am Main, Germany: Kramer.

Newell, F. N., Ernst, M. O., Tjan, B. S., & Bülthoff, H. H. (2001). Viewpoint dependence in visual and haptic object recognition. *Psychological Science, 12*(1), 37–42.

Koffka, K. (1935). *Principles of gestalt psychology*. London: Lund Humphries.

Palmer, S. E. (1999). *Vision science: Photons to phenomenology*. Cambridge, Massachusetts, London, England: The MIT press.

Palmer, S. E., Rosch, E., & Chase, P. (1981). Canonical perspective the perception of objects. In J. Long, & A. Baddeley (Eds.). *Attention and performance IX* (pp. 135–151). Hillsdale, NJ: Erlbaum.

Perrett, D. I., & Harries, M. H. (1988). Characteristic views and the visual inspection of simple faceted and smooth objects: "tetrahedra and potatoes". *Perception, 17*, 703–720.

Perrett, D. I., Harries, M. H., & Looker, S. (1992). Use of preferential inspection to define the viewing sphere and characteristic views of an arbitrary machined tool part. *Perception, 21*, 497–515.

Pomerantz, J. R., & Kubovy, M. (1986). Theoretical approaches to perceptual organization: Simplicity and likelihood principles. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.). *Handbook of perception and human performance* (pp. 36-1–36-46). New York, NY: Wiley.

Pomerantz, J. R. (1981). Perceptual organization in information processing. In J. R. Pomerantz, & M. Kubovy (Eds.). *Perceptual organization* (pp. 141–180). Hillsdale, NJ: Erlbaum.

Pomerantz, J. R., & Pristach, E. A. (1989). Emergent features, attention, and perceptual glue in visual form perception. *Journal of Experimental Psychology: Human Perception and Performance, 15*, 635–649. https://doi.org/10.1037/0096-1523.15.4.635.

Rosas, P., Wichmann, F. A., & Wagemans, J. (2007). Texture and object motion in slant discrimination: Failure of reliability-based weighting of cues may be evidence for strong fusion. *Journal of Vision, 7*(6), 3. https://doi.org/10.1167/7.6.3.

Royer, J., Blais, C., Barnabé-Lortie, V., Carré, M., Leslerc, J., & Fiset, D. (2016). Efficient visual information for unfamiliar face matching despite viewpoint variations: It's not in the eyes!. *Vision Research*. https://doi.org/10.1016/j. visres.2016.04.004.

Shen, J., & Palmeri, T. J. (2015). The perception of a face can be greater than the sum of its parts. *Psychonomic Bulletin & Review, 22*(3), 710–716. https://doi.org/10.3758/s13423-014-0726-y.

Schiltz, C., & Rossion, B. (2006). Faces are represented holistically in the human occipito-temporal cortex. *NeuroImage, 32*, 1385–1394.

Schiltz, C., Dricot, L., Goebel, R., & Rossion, B. (2010). Holistic perception of individual faces in the right middle fusiform gyrus as evidenced by the composite face illusion. *Journal of Vision, 10*(2), 25. https://doi.org/10.1167/10.2.25.

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *The Quarterly Journal of Experimental Psychology A, 46*, 225–245.

Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configurational information in face perception. *Perception, 16*, 747–759.

Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., et al. (2012). A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure-ground organization. *Psychological Bulletin, 138*, 1172–1217.

Wertheimer, M. (1922). Untersuchungen zur Lehre von der Gestalt I. *Psychologische Forschung, 1*, 47–58.