Zsuzsanna Kondor

Representations and Cognitive Evolution: Towards an Anthropology of Pictorial Representation

Abstract

In this paper I will argue for the possibility of a kind of evolutionary anthropology of images. I will rely on the notion of cognitive evolution as Merlin Donald construed it, on the one hand, and on the enactive account of vision, on the other hand. The two approaches can be seen as divergent, since the Donaldian theory is based on the evolution of representational capabilities, while enactivism (both conservative and radical wings) challenges the necessity of mental representation, at least in certain cases. But, as I suggest, this divergence does not lead to a contradiction, rather both illuminate the importance of the extra-cranial context, and the importance of the representational techniques that make ideas, fears, etc. manifest.

I will begin by outlining the special character of the image construed as the picture of something; then I will review the necessary cognitive capabilities of perception and depiction, and finally account for the role of pictorial representation in cognitive evolution.
1. The riddle of images

According to a widely accepted conception, »an image cannot be seen as such without the paradoxical trick of consciousness, an ability to see something as ›there‘ and ›not there‹ at the same time« (MİTCHELL 1987: 17); or as Belting formulated the same idea in a rather enigmatic way, »what an image is: the presence of an absence« (BELTING 2005).

This paradox ascribed to images is based on the triad of the vehicle (the physical picture), the pictorial subject (the subject of pictorial representation), and the pictorial image (that represents its subject), and relates to the mostly implicit consideration of images as based on similarity or convention. First, I will attempt to show how the distinction of the elements of this triad relate to representationalism, and then focus on similarity.

Evan Thompson epitomizes the special character of picture-viewing with the help of the three distinct elements of a picture of something. As he put it,

»[t]he physical picture is present to perception, while the absent pictorial subject is re-presented, brought to presence by the pictorial image. In viewing and appreciating a picture, we are interested mainly in neither the physical picture nor the pictorial subject as such, but rather in the pictorial image that appears in the physical picture and represents the pictorial subject« (THOMPSON 2007: 288).

Thompson, with the help of the phenomenological distinction between the intentional act of presentation and re-presentation (2007: 25f.), characterizes the main differences between perception, imagining, and picture viewing. According to this distinction, perceptual experience is always presentational (the subject of the experience is present), while remembering and imagination are re-presentational, since their subject is not present phenomenally. That is, re-presentational experience, though related to presentational experiences, is determined by the absence of its subject. Either we try remembering a face (and/or scenery), or by imagining it we represent it, even if we rely on earlier presentational experiences.
Even though we can, with the help of the conceptual distinction between the intentional act of presentation and re-presentation, nicely define the distinctive features of the abovementioned activities, this does not help us in understanding how these engagements necessitate and facilitate each other. Moreover, the given description of picture-viewing clearly performs the dualism of form and content, and therefore suggests the possibility of their separation. As James, Wittgenstein, and Merleau-Ponty declared, we have no ground for distinguishing pain or a thought from its expression. Can we see a pictorial image without its vehicle? Consider what it is that we recall when we try to remember Vermeer's Girl with a Pearl Earring: the girl (the subject of the pictorial image), the pictorial image, or the painting in its corporeality? How is it possible to identify one's intentions without the expression of them with either movements, gestures, a change of facial expression, or with words? As Arnheim quite explicitly states, «neither the formal pattern, nor the subject matter is the final content of the work of art. Both are instruments of artistic form. They serve to give body to an invisible universal» (Arnheim 1974: 461). That is, in the case of works of art, both the subject of the pictorial image and the pictorial image itself serve the purpose of expressing an idea.

The dualism of form and content is, I believe, rooted in representationalism, since it suggests from the very beginning (at least from Descartes) that representations are the mediators between the physical world and mental activity. Similarly, in the course of the imagery debate from the 1970's, descriptionalism asserts that mental representations have a propositional structure, while pictorialism (the opposite pole) suggests mental representations are depictive. If we conceive mental representations either as being symbol manipulation in accordance with certain rules, or pictorial items in consciousness, the floor is open for a fallacy, namely believing that mental representations can be considered as either propositional or pictorial. In the case of pictorialism, as Thompson recalls Sartre's criticism, »an image is nothing other than a relation«. That is, »a mental image properly understood is not a content contained in

49 For further details, see Thompson 2007: 287-297.
consciousness; rather, it is a dynamic and relational act of consciousness« (THOMPSON 2007: 301).

Regarding mental images, many claim that these mental phenomena are pictorial in consciousness as well; but it is rather the »illusion of immanence« as Sartre coined this kind of fallacy (THOMPSON 2007: 300f.). On the basis of these phenomenological considerations, and taking into account motor skills, the enactive approach\textsuperscript{51} suggests bodily skills and perception are inseparable, and thus provides an alternative for representationalism (be it either descriptive or pictorial), a promising way to bridge the gap between the world and consciousness.

I suggest considering resemblance and convention as being conflicting characteristics of pictorial representation (as compared to linguistic symbols), which is possible when we ignore the expressive medium and opens the floor to the abovementioned fallacy. Most pictorialists suggest linguistic symbols are conventional, while pictures are based on similarity. But defining what similarity consists of seems to be a hard nut to crack. Kristóf Nyíri, on the basis of Gombrich’s considerations, suggests »[r]esemblance is selective equivalence« (Nyíri 2011: 17). This apparently apt formulation leads us to a great deal of difficulty. Granted, in the case of geometry this definition works very well because we can work with well-defined axioms and definitions and even among theoretical entities. Two triangles are similar if their corresponding angles are equal, and their corresponding sides are proportional. (Of course, one of the conditions automatically implies the other.) But in everyday life, we do not exclusively face clear-cut cases. We can recognize emoticons and pictograms on the basis of similarity; we can see similarity between a dog and its keeper, or two persons, between animate and inanimate beings, etc. Since the notion of equivalence is much more demanding, numerous ques-

\textsuperscript{51} Enactivism has far-reaching roots in the history of philosophy (however, some of them are almost non-referred), closely relates to cognitive psychology, and interestingly, combines in a promising way some suggestions of so-called continental philosophy with the conceptual framework of the philosophy of mind. In 1991 F. J. Valera, E. Thompson and E. Rosch proposed »as a name the term enactive to emphasize the growing conviction that cognition is not the representation of a pregiven world by a pregiven mind but is rather the enactment of a world and a mind on the basis of a history of a variety of actions that a being in the world performs« (Valera et. al. 1991: 9).
tions arise: Can we give a list of equal items in each case? Or examined from another angle: can equal features provide sufficient ground for similarity? Is there a definite number of identical/equivalent features required to see two things as being similar? The list can continue, but my main concern is whether we can define equal features in the course of recognizing similarity when this process relates to presentational, representational, and motor skills. That is, I would say that since the recognition of resemblances is based not exclusively on visual capabilities, but is rather also related to motor activity and other modalities, then resemblance construed as selective equivalence may lead us astray. How can we talk about equivalence when we speak of different kinds of related capabilities? If we accept that visual perception and motor engagement are inseparable, analogy seems to be a more appropriate candidate. When we apprehend the ambient world, we rely on earlier experiences, and while utilizing them we act upon this world with the help of different motor skills.

As Harald Høffding suggested, Kant’s essential invention was the introduction of analogy, *i.e.* »likeness of the relations of different phenomena, not likeness of single qualities« (HØFFDING 1905: 200). Analogy opens up a new horizon in theorizing. Although we do not gain much regarding what similarity consists of, with the introduction of analogy, as opposed to the requirement of equivalence, the difficulty raised by the demand of relating different kinds is eliminated. (I have in mind the antinomic situation when we need to talk about different things in terms of identical features.) How can a physical and a mental image have identical features when they are different in kind? I believe analogy and the exclusion of the »illusion of immanence« provide an appropriate ground for the understanding of how images work regarding both perception and depiction. As Kant introduced analogy and underscored that a schema is not identical with an image, it became possible to anchor mental activities in the physical, flesh and blood world.
2. Visual perception and pictorial representation

Considered through the lens of Gestalt psychology, picture-viewing and pictorial depiction are based on the same ground, though the capability of depiction requires some additional knowledge, viz. the knowledge of the underlying structure of the perceived effect.

As Thompson suggests, in the case of picture-viewing we focus on the pictorial image. However, with a work of art or a piece of scientific visualisation, this focus is inseparable from the pictorial subject, i.e. the idea/intention of the artist or the mutation of the investigated organ. On the one hand, I believe the elements of the triad on the basis of the importance of the context; on the other hand, because pictorial representations are objectifications of certain thought processes, and accordingly, their appearance heavily depends on the scope and limits of their vehicle, the elements of the abovementioned triad can hardly be isolated even within the framework of theorizing. But instead of getting ahead of myself, first I will attempt to clarify the peculiarities of the apprehension of a picture of something.

Pictures are representations, or more precisely objectifications, as they are accessible to others. Since they represent something, they can be considered as symbols. The working definition of a symbol is very akin to the riddle of images. In accordance with the general view, the psychologist Judy S. DeLoache suggests »a symbol is something that someone intends to represent something other than itself« (DELOACHE 2004: 66). Many things can be a symbol, as symbolic competence evolves from »general symbolic ability«, and relates to intentionality. Young children, when focusing on a symbol itself, may encounter difficulties: it turns out that the more appealing and conspicuous an object is, the more difficult it is to consider as other than itself, and having physical contact with an object also increases the difficulty of using it as a symbol. Taking pictures as symbols into account, although pictures have a unique characteristic, viz. they are unavoidably perspectival\(^2\), experiments provide evidence that pers-

---

\(^2\) Pictures as being unavoidably perspectival should be considered in a restricted sense, viz. with the exception of graphs, diagrams and the like. For more details see Roberto Casati and Valeria Giardino, Public Representations and Indeterminacies of Perspectival Content. (Manuscript. Forthcoming in Z. Kondor (Ed.): Enacting Images. Representation Revisited. Cologne [von Halem Verlag] 2012.)
pective does not play a role in the case of younger children. In one experiment, pictures of tulips taken from above and others from directly in front of the flowers were shown to children. The photos did not seem to differ considerably for children under three years of age; it appears that 3-year-old children do not consider the difference of perspective when the depicted objects are identical (DeLoache 2004: 69).

This suggests that motor skills and earlier experiences help us recognize the referent of a symbol with no regard to the perspective from which the pictorial symbol is depicted. This option creates a strange situation: according to anatomical evidence there are two visual pathways that are usually supposed to be attached to different tasks, and the motor component seems to be deeply integrated into the act of visual perception, particularly in case of picture viewing. According to the representational theory of the visual mind, one pathway is devoted to visual percepts in the service of further cognitive processing, and the other is in the service of action (Jacob et.al. 2003). Though the transit between the two pathways is beyond question, the functional differences seem to provide ground for researchers and philosophers to separate the motor component from visual perception. But enactivism underscores the cooperation between the two capabilities. Valera and his co-authors suggest »sensory and motor processes, perception and action, are fundamentally inseparable in lived cognition. Indeed, the two are not contingently linked in individuals; they have also evolved together« (Valera et.al. 173). Noë goes a bit further when he suggests that »[t]he basis of perception ... is implicit practical knowledge of the ways movement gives rise to changes in stimulation« (Noë 2004: 8). The anatomic distinction of dorsal (in the service of action) and ventral (dedicated to further processing of what is seen, perceptual representation, identification, etc.) paths is not to be considered either as evidence for or against the enactive approach, since enactivism proposes that any kind of perceptual representation is determined by the applied sensorimotor skills. Noë highlights the importance of motor skills even in the case of observational concepts. Experiencing something as cubical is possible because you possess some practical knowledge according to which you experience a cubic thing »as something whose appearance would vary in precise ways as you move in relation to it, or as it moves in
relation to you. You have an implicit practical mastery of these patterns of change« (Noë 2004:117). Similarly, as the DeLaoche experiments show, young children recognize highly realistic two-dimensional depictions of three-dimensional objects with the help of their body: they try exploring them with their hands, using them as real objects. This kind of manual exploration starts to decrease after 18 months of age (DeLoache 2004: 68). That is, the necessary condition of picture-viewing, viz. the recognition of a picture as the picture of something, is a capability gained thanks to bodily experiences by motor engagements in ontogeny.

The relation between perception and motor skills is fairly obvious in case of depiction. But this relation reaches beyond the requirements of the fine tuning of movements in the course of creating drawings, paintings, sculptures and the like; perception and depiction are rooted in embodied cognition. As Merleau-Ponty and current conceptual metaphor theory underscore, our concepts are based on our primordial bodily experiences.

> »[I]f the words ›enclose‹ and ›between‹ have a meaning for us, it is because they derive it from our experience as embodied subjects. In space itself independent of the presence of a psycho-physical subject, there is no direction, no inside and outside. A space is ›enclosed‹ between the sides of a cube as we are enclosed between the walls of our room« (Merleau-Ponty 1962: 236).

The Gestalt laws of perception53 and the idea of embodiment provided the ground for revealing the relation between perception and depiction. And while both depiction and perception

---

53 Regarding visual perception, Gestalt psychology recognized some key characteristics of visual perception. There are different groupings of these laws/principles depending on nuances and time period. Attempting to recapitulate the main characteristics, emergence is one of the main principles. Emergence suggest that a thing in our environment is perceived as a whole, all at once, i.e. the recognition of a dog comes not by identifying its parts and then by the composition of these parts, but rather via the recognition of certain patterns. The generative aspect of perception is reification. This means that the experienced percept provides more spatial information then the sensory stimulus. That is, we recognize certain shapes although they have no contours like in the case of a Kanizsa triangle. Multistability of perception refers to a tendency in ambiguous perceptual situations. As in the case of Rubin’s heads/vase illusion (1919) the figure (the substantial appearance of an object) and the ground (the general environment of the object) oscillate in accordance with the interpretation. Thanks to invariance, it is possible to recognize simple geometrical objects independently from the perspective we see them in and elastic deformations. According to the law of Prägnanz, we tend to arrange our experience in a manner that is simple, symmetric, regular and orderly. There are six principles of Gestalt perception: proximity, similarity, common fate, good continuation, closure, and symmetry. The principle of proximity/contiguity state that
are rooted in the same bodily experiences, depiction necessitates the knowledge of the representational means.

Arnheim came close to the enactive approach when he considered the perception of shapes as »an eminently active occupation«. As he writes, »in looking at an object, we reach out for it. With an invisible finger we move through the space around us, go out to the distant places things are found, touch them, catch them, scan their surfaces, trace their borders, explore their texture«.\(^{54}\) (Arnheim 1974: 43).

He also suggests considering vision as »a creative activity of the human mind«. It is because »vision deals with the raw material of experience by creating a corresponding pattern of general forms, which are applicable not only to the individual case at hand but to an indeterminate number of other similar cases as well« (Arnheim 1974: 46). In visual perception, all our earlier visual experiences play a part. Because we form perceptual concepts while perceiving the ambient world, we are able to grasp some »outstanding features of objects« and at the same time notice slight changes. Arnheim argues on the basis of the holistic nature of perception. As his example suggests

»The minute modifications of muscle tension or skin color that make a face tired or alarmed are quickly observed. The viewer may, however, be unable to pinpoint what has caused the change in overall appearance because the telltale signs fit smoothly into an integrated framework« (Arnheim 1974: 44).

54 Cf. »My theme ... is the touch-like character of vision itself, by which I mean precisely the fact that vision too is intrinsically active« (Noë 2004: 97).
That is, we perceive general patterns deeply integrated into the given context, therefore even slight changes can modify the overall structure, and therefore they are easily noticeable.

Arnheim distinguishes perceptual and representational concepts. The latter are necessary when in a particular medium the relevant features of a model are to be represented.

»[Image-making of any kind requires the use of representational concepts. Representational concepts furnish the equivalent, in a particular medium, of the visual concepts one wishes to depict, and they find their external manifestation in the work of the pencil, the brush, the chisel.

The formation of representational concepts, more than anything else, distinguishes the artist from the nonartist« (ARNHEIM 1974: 169).

As we can see, representational concepts are rooted in perceptive and motor skills and, importantly, in the recognition of the underlying structure of the perceivable effect on the one hand, and the knowledge of the possibilities and limits of the given expressive medium on the other hand. Young children are capable of recognizing the difference between vertical/horizontal and oblique lines, but only later are able to actually produce the latter. Children’s drawings reveal much about the relation of perceptive, depictive and motor skills. Though a straight line is visually the simplest line, the first scribbles are whirl-like »circular path of lines«. Arnheim suggests two reasons for this interesting phenomenon. On the one hand, »the child’s arm re-enacted the motion as a gesture during the drawing« (1974: 172), and on the other hand, according to the law of simplicity, a circle is the »simplest visual pattern« as it »does not single out any one direction« (1974: 175) and moreover, certain anatomic features subserve curved motion. Experiments suggest the young children favour circles over other shapes (ARNHEIM 1974:175), and early childhood drawings tend to show humans as round

55 Arnheim mentions the »lever construction of the human limbs« and the structure of the shoulder joint. The movement by the elbow, the wrist, and the fingers also facilitate rotation. (Arnheim 974:174f.)
tadpole-like shapes. Of course, these scribbles are not circles in a strict sense, but the result of the child’s more or less uncontrolled movements is most similar to circles.

Recent research in the field of developmental psychology has provided evidence that the intention of communication plays an important role in drawings in ontogeny. »The first scribbles of a child are not intended as representation. They are a form of the enjoyable motor activity in which the child exercises his limbs, with the added pleasure of having visible traces« (ARNHEIM 1974:171). With practice, the schemas become more detailed, and details gain importance when it turns out the drawing can be considered as a representation of something/someone. Four-year-old children do not take into account details that can help when it is important to distinguish a balloon and a lollipop. But

»[t]hree- and four-year-olds produced better drawings when they understood that their drawings would be used to communicate to an adult which object to use in a game than when they were just asked to draw the objects. When informed that their drawings had not communicated clearly enough, children of this age subsequently improved them. Thus, having a communicative intent enhanced the children’s appreciation of the symbolic function of their drawings« (DELOACHE 2004: 67f.).

The idea of embodiment provides a frame of reference in which perception and depiction are inseparably attached to motor skills; and depictive capability relates to social commitments and necessitates material engagement as it enables expression within a given expressive medium. The necessity of considering the other as an intentional agent, the intention of communication, i.e. beyond the joyful activity of drawing, being aimed at sharing some ideas is a necessary condition of pictorial representation in ontogeny.

3. Representations and material engagement

The view that sociability is a fundamental characteristic of human nature is widely accepted in the fields of sociology, linguistics, cognitive evolution theories, developmental psychology,
etc. In his social brain hypothesis, anthropologist and evolutionary psychologists Robin Dunbar suggests that the size of the neocortex and that of the group correlate.

»The social brain hypothesis implies that constraints on group size arise from the information-processing capacity of the primate brain, and that the neocortex plays a major role in this. ...The constraint on group size could be a result of the ability to recognize and interpret visual signals for identifying either individuals or their behavior; limitations on memory for faces; the ability to remember who has a relationship with whom (e.g., all dyadic relationships within the group as a whole); the ability to manipulate information about a set of relationships; and the capacity to process emotional information, particularly with respect to recognizing and acting on cues to other animals’ emotional states« (DUNBAR 1998: 184).

That is, taking part in a group’s life requires more than just the capability of remembering, it requires additional capacities of information processing.

Merlin Donald, in his cognitive evolutionary theory, relies on Dunbar’s hypothesis (expressed in earlier writings by him) and goes a bit further when he underscores that new institutions/inventions can extend our brain capacity. He suggests »that the evolution of cognitive structure at the modular level might have continued well beyond the point at which physical evolution had stopped«. Cultural patterns have a far-reaching effect on cognition, even »in terms of its fundamental neurological organization« (DONALD 1991: 14). He differentiates three main transitions in phylogeny: a shift from episodic to mimetic, from mimetic to mythical, and finally from mythical to theoretical culture. Each of these changes is a consequence of the emergence of a new kind of representational skill and there is a concomitant increasing load on biological memory, with the exception of the last change.56

56 »Human memory had, from its inception, expanded the range of primate memory. The earliest form of hominid culture, mimetic culture, depended on an expansion in the self-representational systems of the brain and created the initial base for semantic memory storage, which consisted initially of representational action scenarios reflected in mime, gesture, craft, and skill. With the evolution of speech and narrative ability, there were even greater increases in the load on biological memory, adding not only the storage networks for phonologi-
Mimetic representation as a supplement to episodic memory yielded the ground for creating a community with special habits and organization, conferring a sense of identity on the members of community. The emergence of linguistic representation, (beside episodic memory and mimetic skills), made the evolution of a framework possible in which mythical constructions could emerge as structuring and orienting potentials within the communities. The transition from mythical to theoretical culture was thanks to a new invention, viz. the external symbolic storage system. While earlier changes inserted a huge load on memory capacity, this latter one rather exempted biological memory. This apparently technological invention established considerable alterations in cognitive habits.

»[T]heoretic culture broke with the metaphoric style of meaning in oral-mythic culture. Where narrative and myth attribute significances, theory is not concerned with significance in the same sense at all. Rather than modeling events by infusing them with meaning and linking them by analogy, theory dissects, analyzes, states laws and formulas, establishes principles and taxonomies, and determines procedures for the verification and analysis of information. It depends for its advanced development on specialized memory devices, languages, and grammars« (DONALD 1991: 274f.).

According to the Donaldian chronology, pictorial representations emerged within the framework of oral-mythical culture and gain importance in Donald’s theory because of their mediating role in the development of external storage; images are considered as the first steps towards external storage. As he put it »the first pictorial images themselves were also external representations. They existed outside of the individual, rather than in visual memory. Therefore a technological bridge was under construction that would eventually connect the biological individual with an external memory architecture« (DONALD 1991: 284).

Donald considered cognitive evolution as being an encapsulating development of representational skills. The refinement of representational skills provides the ground for the evolution of...
increasingly complex communities, and this entails further representational capabilities. Accordingly, Donald suggests that while the invention of pictorial representation presupposes earlier cognitive skills, »pictorial images, on one level, must possess episodic reference«, and »mimetic manufacturing skill and linguistic conceptual skill« (DONALD 1991: 283), pictorial images also paved the way for the creation of an external and lasting form of representation. Therefore, these first steps opened up the horizon of cognitive development. An encapsulating evolution of representational skills towards external storage implies, as Donald suggests, that »we are permanently wedded to our great invention, in a cognitive symbiosis unique in nature« (1991: 356). Thus, human minds can be considered as being hybridizations.

At this point, I will outline the view of the enactive approach with regard to the role of images in phylogeny. Though this view calls into question the necessity of linguistic conceptual skills, I will underscore the importance of external representations in both cases.

The alternative view suggests emphasizing the importance of material engagement as compared with Donald’s account of the evolution of the human mind. It highlights the importance of the interaction between the cognizing agent and its environment. Under the heading of neuroarchaeology, we can find recent efforts towards converging the results of neuroscience and the findings of archaeology as it relates to material culture. As Lambros Malafouris put it,

»neuroarchaeology is strictly an interactionist approach, aiming primarily to understand the bidirectional links between brains, minds and culture. Brain activity is a crucial component of the human mind but so is also material culture. The human mind as an interactive, embodied, and distributed autonetic system is neither restricted to nor can it be identified with the temporal and spatial boundaries of the brain within«. (2010: 64)

Against the background of the notions of embodiment, scaffolding, and material engagement, pictorial images enter a new light. Within the framework of embodied cognition, the importance of the ambient world and cooperation with it radically increases. The responsive relation
suggests a symbiosis between the biological brain, the body, and its environment. According to the extended mind theory and also Donald himself, the ambient world provides a scaffold, or viewed the other way around, in order to accommodate to their conditions, humans (and other animals) rely on affordances, i.e. the easing elements of the environment. This kind of accommodation drives changes in the organism as well. Accordingly, the capability of pictorial representation as a kind of material engagement can be considered as a new way of accessing the world. As Malafouris suggests,

»[t]hrough the process of ›imaging‹ the underlying mechanisms of human perception are being transformed to an object for perception and contemplation. Those invisible mechanisms become now perceivable visual patterns arrayed and combined in real time and space. In this sense, the image offers a new mode of epistemic access to the world of visual experience« (2007: 299).

The main difference between the Donaldian approach and Malafouris’ suggestion lies in the relation between pictorial skills and conceptual capacity. Malafouris

»believe[s] that to say that a painting from Chauvet ›resembles‹, or ›looks like‹, a rhinoceros, a horse or a bull, does not necessarily imply that these paintings also ›represent‹ those animals in a concept-mediated referential manner. Iconicity does not in itself imply the existence of consciously manipulable content-bearing tokens, though it certainly sets up ... a visually stimulating channel of influence for their creation« (MALAFOURIS 2007: 293).

Unlike Malafouris, Donald seems to believe that there is an underlying »linguistic conceptual skill« (DONALD 1991: 283).

This difference57, though not negligible, does not play an important role for our current purposes. Under both frameworks, it is possible to develop a theory of images as they inherently

57 I believe this difference is rooted in the broader context of the two investigations: Donald aimed to create continuity in cognitive evolution focusing on representational skills. Malafouris takes into account cave paintings from a different perspective, viz. from an enactive approach. But both scholars agree regarding the long-lasting
belong to the cognitive capacity of human beings. Beyond the striking similarities between the prerequisites of linguistic and pictorial capabilities, scaffolding and new expressive forms entail considering images as an important anthropological issue. Nevertheless, from an enactive point of view, images definitely play a crucial role in phylogeny.

»The Palaeolithic image-maker constructs an external scaffold that affords the world to be seen and experienced in ways that the physiology of the naked eye by itself does not allow. This scaffolding also enables a new direct understanding of the human perceptual system and thus offers to the Palaeolithic individual the opportunity to become in some sense, maybe for the first time, the engineer of his or her own perception. The image, as it is also the case with language, enabled humans to think about thinking« (Malafouris 2007: 300).

4. Conclusion

Being able to create a picture as a picture of something and using linguistic symbols require the same conditions: a social brain as a prerequisite, and all this entails: a community, representational skills, and the integration of new representational patterns into everyday practice. Representations as means of communication are bounded to habits, institutions (in a community) and/or some kind of material support. Casting a glance at ontogeny, we can see that symbol use, picture viewing, linguistic and depictive skills are the result of being a member of a group (family, nursery, etc.) and thus becoming involved in its communicational habits.

In ontogeny, the manual exploration of realistic images is a prerequisite of picture-viewing in a broad sense. The capability of pictorial depiction needs an intention of communication and

---

58 Michael Tomasello (1999) suggests cultural transmission is a decisive element in the evolution of the cognitive setting of the modern human mind. Accordingly, intersubjectivity (as a precondition), communication, and therefore linguistic capability play a crucial role. Regarding linguistic skill, he underscores the importance of the possibility of external expression as it facilitates cognition. (Tomasello 1999: 195) The perspectival character of language (namely, we express ourselves in a way that best fits our purposes) and the representational concept (as Arnheim suggested) similarly presuppose expertise in the given expressive medium.
perfection in the expressive medium. Accordingly, image production presupposes sociability. Since many scholars share the view that the social brain is the answer to the puzzling question of how human cognition could reach its recent capacity in a relatively short time period in the course of biological evolution, it can be considered as being particularly human. Having a social brain necessitates communication. Communication presupposes the capability of objectification, i.e. sharing an experience/idea is possible only if it is accessible for others. Accordingly, representation plays a crucial role. Since representation is only possible as presentation, the expressive medium is decisive.

Representationalism as a highly resistant view in philosophy entered the scene of philosophical debate with the Cartesian dualism of the extended body and the immaterial mind quite explicitly. Under this dualist framework, we do not have a satisfying answer to the question of how it is possible to make a physical percept accessible for mental processing. The enactive approach seems to pave the way for an answer beyond the horizon of representationalism. It regards perception as an active engagement, as opposed to a passive acceptance of external stimuli for further cognitive processing that results in an internal representation.

The inclusion of bodily skills, and hence the importance of the responsive relation between the cognizing subject and her environment, shed a new light on the act of picture viewing, pictorial depiction and, the role of pictorial representation in phylogeny. Similarly, recent psychological experiments illuminate the special role of pictures in ontogeny: we learn the difference between a picture and its subject with the help of bodily exploration, and we gradually learn how we can use an expressive medium through the joyful activity of leaving traces, and with finer coordination, how we can use this medium to express something and develop representational concepts. All these result in a permanent refinement of cognitive capacity. Images as external representations enter a new light: as Donald suggests, they pave the way for creating an extra-cranial storage system, and according to Malafouris, like language, they open the floor for reflective thinking within the framework of material engagement.
References


JAMES, W.: What is emotion? In: Mind, Vol. 09, Issue 34, 1884, 188-205


