Looking and Seeing:

On the Abilities of Primates to Detect the Gaze and Attribute Visual Attention

Alex Leykin for "Animal Cognition"

Introduction

Vision is probably one of the most important functions of advanced organisms. When we look at objects, we not only detect their physical parameters but also are able to think about these objects and make complex conclusions. When we see another person looking somewhere, we are also able to understand that he or she focuses on some object and probably thinks about that. To what extent primates' vision is similar to ours? In other words, are they able to perceive others looking and understand what they are looking at?

This work attempts to address the argument about the existence of gaze detection in primates. I subject to critical analysis the most well known scientific papers in the field of primatology as well as provide parallels with the studies of human gaze. In order to understand the concept of gaze I will distinguish between gaze and visual attention and provide definitions of both concepts drawn from what we know about "gaze" as humans. Generally, I will try to answer and expand some of the questions posed by Keeley [1]. In particular I will address the question "Is it true that some animals follow gaze without attributing mental states to others?"

Neurophysiological background

A number of important anatomical factors can influence the way we think of gaze in primates, because often anatomical structure determines function and function is part of the issue we are trying to understand. The structure of primate vision as is must not be erroneously taken to be limited by the eye (or retina). It has been established by vision science [6] that primate vision is a complex system comprised of the structures present in the *eye* (retina, cones/rods, bipolar cells, ganglions), in the *midbrain* (optical fiber, LGN or lateral geniculate nucleus) and in the *cortex* (visual cortex with more than a dozen specialized areas, already singled out)

All primates possess a highly specialized and well evolved spatial vision and an ability to reconstruct 3-dimensional information (i.e. locations of objects in three dimensional worlds) from the visual input. The usage of both eyes produces the stereo-vision phenomenon, which enables calculating disparity maps. The visual parallax, i.e. the ability to assess the changes in the visual world while we yaw, pitch or roll of the head or entire body helps greatly to reason about the location of the objects in 3D. Finally, comparing the sizes of the objects in question to the objects of the known size, primates can approximate the distance to the target objects.

A highly developed *fovea* in combination with the frontal location of the eyes allows us as all the predators to exercise a "smooth pursuit". In other words, primates are able to concentrate their "gaze" (used in quotes to denote just the center of the visual focus) on a certain area (which will feed a high-resolution signal to the cortex) and receive a stable image of this area even when undergoing through spatial perturbations (e.g. while running)

Information about structure and basic functions of vision is essential to keep in mind while constructing an experiment related to the visual abilities. Moreover knowledge of the physical background can provide additional insights during interpretation of the experiment.

Definitions

For the purposes of this paper *gaze* will be defined as the act of fixating the eyes on the object. Note, however, that it is not required to look in the eyes (or "gaze at the gaze") to detect the gaze, because it can be noticed by other means, such as using the peripheral vision or inferring its location using other cues, e.g. head orientation. Notice the gaze is defined as a mechanistic operation, i.e. as nothing more then a physical concentration of one's eyes on some location in the 3D world. In other words this action can be described as "looking" as opposed to seeing.

Visual attention, on the other hand, is a psychological state when an object being looked at is "…taking possession of the mind, in clear and vivid form, of one out of what seem several simultaneously possible objects…" [6]. This psychological state we know as "seeing." A clear distinction between "looking" and "seeing" comes with the realization that however it is necessary to "look" in order to "see," it is still possible to look and not see. Given the definitions, one has to be clear about what is the primary object of the study: the gaze or the visual attention.

In an experimental setup there must be two agents: the subject and the "gazer", where the former one is detecting the gaze or acknowledging visual attention of the latter. I build my conclusions under the assumption that the abilities to look and see do not differ significantly depending on the object producing the gaze, so the gazer can be either a human being or another primate (typically conspecifics.) I will use the term *gaze detection* to refer to the act of mechanistic recognition by the subject of the direction of someone's gaze, based on the visual cues. Similarly, *visual attention detection* is the attribution of the concept of visual attention by the subject, which is a result of the performed gaze detection.

Experiments with primates

Povinelli's tests on chimpanzees gaze detection can be considered one of the most rigorous studies in the field. These tests will serve as a basis for my own argument for the low-level interpretation of the gaze. The experimenters in Povinelli's tests start from the initial observation and belief that the chimps seem to have a notion of "seeing" (perhaps due to antropomorphisation). When given empty buckets to play with, the lab animals put buckets over their head, run into things and then take buckets off. This behavior gave the authors the reason to believe that the chimps have the concept of seeing.

Their first experiment proceeded as follows. The experimenter looks at the object hidden behind the glass half of which is transparent but the other half is made opaque. As the animals did not seem to trace the human gaze past the obstacle, the initial intuition was that they "understand that the person is looking at something that they cannot see" [2]. Moreover the subjects were typically looking behind the glass, following the experimenter, which also speaks in favor of the high-level explanation. But will this high-level explanation hold during further interpretation? Taking the physiological model of primates' vision into account this behavior could be explained simply by the visual ability to map distances and reconstruct the 3D scene, followed by the occlusion map, which speaks in favor of the low-level explanation.

The authors have tried to confirm their findings by making the animals address their natural begging gesture to one of the two people in the room. One person's gaze was obstructed by bucket, screen, blindfold or hand or simply turned away. All but one setting have made the subjects guess randomly, confirming the functional, i.e. low-level explanation. However, almost all chimpanzees guessed right when the person was facing away.

The hierarchical model of chimpanzees' gaze detection arises at this point consisting of four levels of recognition/detection*: body, head, face and eyes*. The subjects give a strong preference to the features at the body/torso level. Given those are the same, head orientation serves as an additional cue. The lowest weight is given to the smallest feature geometrically the eyes. "Over the shoulder", screens and "attended vs. distracted" test proves the hierarchical model. Open eyes vs. closed eyes produce the same result, although subjects can learn a bit from experience, still the face was a more salient feature than the eyes.

The study concludes that the chimpanzees naturally do not have the "visual attention" or "seeing concept", but they have constructed a set of skills to imitate it during Povinelli's experiments. Gaze following is their skill, but they lack the conceptual understanding of seeing! It means that they don't know "how gaze is related to subjective states of attention" [2].

Training with rewards improved performance, but did not generalize to the blindfolds: covering the eyes vs. covering the mouth. Does it mean that submerging chimps into an extensive social context with humans develops human-like gaze recognition abilities in chimps? Was the learning experience for infants and chimps comparable in its extensiveness and multitude of "seeing vs. non-seeing tasks?" "Unlike the apes, children were correct in most or all of the conditions from their very first trial period" [Ibid.] A year after, and then 3 years after the subjects forgot almost all their skills and had to be presented 4 dozen screen tests and open/closed eyes test to develop the skill again.

Earlier assumption about the presence of "seeing" concept in apes is reconciled with the general conclusion as follows. The chimps putting buckets over their heads have merely found a fun "roller-coaster" like game to enjoy and "they place a bucket over their heads because it produces an interesting, pleasurable experience". This does not necessarily imply that they have discovered the concept of seeing by obstructing their vision. Another, even higher level explanation is that "it is also possible that chimpanzees have a better understanding of their own mental states than the mental states of others, a possibility addressed by Povinelli and Prince (1998)".

Some researchers argue that "Povinelli's chimps" differ from chimps in the wild in the fact that they lack a social experience in the natural setting to gain the skills required [9]. Although such a claim stands in the case of tool use, which occurs in chimpanzees' natural life cycle, one can argue that with respect to "seeing experiments", Povinelli's chimps even had an advantage, because from their infant they

were exposed to the significant interactions with human beings, where gaze plays an important role.

Santos et al. [8] show on a similar set of experiments that rhesus monkeys have a more reliable eye-detection in favor of head detection and body tracking. The researchers construct a competitive setup, where the animals are intended to steal, not beg, which poses a more stimulating task under the evolution-based motivational premises. Indeed, this experiment, as the authors admit, may better approximate the normal social conditions in which rhesus monkeys have been observed to use their visual perception. Two baskets with food are placed before each subject with each near a human. One of the human has his/her vision obstructed in on of the ways used in [2]. The monkeys consistently preferred the basket in front of the "non-looking" individual.

Across the six experiments the authors have established that rhesus monkeys can correctly utilize the information about what competitor can see to retrieve a piece of food – thus demonstrating exceptional gaze detection capabilities. Let us analyze their following statement in terms of Morgan's Canon: "…it is difficult to interpret these results in terms of a simple mechanism for responding to the gaze of another individual without representing that individual's perceptions." In my opinion this is a case where the higher level faculty is attributed to the animal, without carefully considering the lower one. One might hypothesize that mere tracing of the direction of gaze is enough to establish the best food target, without reasoning about what the human can actually "see". This objection is later confronted in the study by arguing that in at least two experimental setups gaze direction could not have been inferred from the stance or head orientation. But that argument misses the core of the objection: a quite impressive ability to tell where one is "looking" by using solely the eye information does not necessarily prove the understanding that one is "seeing".

Hare et al [4] evaluated the possibility of the theory of seeing in chimpanzees as well. In their study a competitive environment was created, where dominant and subordinate conspecifics had to compete for two items of food: one of which was usually in-visible to the dominant (occluded by the barrier). Experiments have confirmed the screen results obtained by Povinelli: chimps can discern the spatial location of the barrier in relation to the conspecifics' gaze. Speaking in favor of primates' imaginative abilities is the fact that the subordinates were able to "predict what the dominant would see" even before the dominant was actually present in the cage! The authors admit that the test quite fits the "cognitively weakest hypothesis", that the observed behavior is based on "learned behavioral contingencies, accompanied

by no understanding of the other's visual experience at all". Finally, a mixed conclusion is presented that chimpanzees are natural-born gaze followers, and in addition they reflect what they have seen through their individual experience, this way coming to learn important things about the visual access of conspecifics.

Constructing an experiment to reveal "seeing"

How an experiment to reveal the presence of "visual attention" concept in primates can be created? A study aiming to operationalize the concept of visual attention as defined above can be constructed. This assumes a series of obvious manipulations that primates can perform instead of using language (as the humans might do) to show not only that they acknowledge the fact that the experimenter is looking at something, but also realize that the human experimenter can see. One experimental setup I can envision is as follows.

The experimenter is presented with a number of distinct objects (bananas, apples, a pineapple, etc.), such that would be easy to identify not only for the human, but for the primate under the study as well. The experimenter's gaze is concentrated on one of these objects, in other words, as it would appear to any human observer (and according to my definitions), her mind is possessed by the object. The subject of such a study would be a primate, who is presented with the exact same set of objects. The research technique than becomes to determine in a number of pass/fail trials if primate can pick the object that coincides with the one experimenter's attention is directed at.

The subject might be subsequently rewarded when selecting the correct object from her own set. The research hypothesis here is that due to the presence of two distinct yet similar sets of objects simple tracing of the direction of the gaze will be not enough to establish the correct object, without reasoning about what the human can actually "see". In order to succeed the subject will have to form a mental model of the object and link it to the "looker's" attention (otherwise the subject would be picking object from the experimenter's set). I suppose that reinforcing the correct choice with food rewards will not diminish the value of the experiment. Why?

If only one set of objects were present, then the subjects could have simply inferred a conditioned rule that selecting the objects which the humans look at is good. In the proposed setting, however, this rule would fails and the only successful strategy, it seems would be forming a higher level concept of what is selected. Special interest is presented by the scenarios where the experimenter's face, head, or both are obstructed – to replicate the setup of Povinelli's experiments.

A similar experiment could be conducted on humans. The details are quite trivial because humans can use language to describe the choice.

Conclusions and more questions

The work reviewed in this paper did not convincingly demonstrate that the primates have visual attention to understand what others look at and see. At most, a smooth gradation of skills in different species in terms of gaze detection was established. In order to demonstrate that primates have the "theory of seeing" one has to consider, perhaps, other experimentation techniques to reveal more complex concepts that can be constructed based on seeing, e.g. guilt. Because of the absence of natural-language in primates one has to think in terms of an experiment which is aiming to reveal the conceptualization of objects seen by humans or conspecifics, operating though on an identical, yet separate set of objects.

More general questions also arise. Most importantly, considering that chimps are cognitive beings, who constantly find themselves in novel social situations and adapt to them, why they evolve so little during their life spans if compared to humans? Why their collective or societal organization remains on the relatively low (again comparing to humans) level?

References

1. Brian L. Keeley "Eye gaze information processing theory: A case study in primate cognitive neuroethology."
2. Povinelli, D.J. 2000. "Folk Physics for Apes: The Chimpanzee's Theory of How The World Works" New York/Oxford: Oxford University Press
3. Tomasello, M., Hare, B., & Agnetta, B. (1999) Chimpanzees follow gaze geometrically. *Animal Behaviour*, **58**, 769-77
4. Hare, B., Call, J., Agnetta, B., and Tomasello, M. 2000. Chimpanzees know what conspecifics do and do not see. *Animal Behaviour*, **59,** 771-785.
5. Hare, B., Call, J., and Tomasello, M. 2001. Do chimpanzees know what conspecifics know? *Animal Behaviour*, **61,** 139-151.
6. Rodieck, R.W., 1998, "The First Steps in Seeing", Sunderland MA: Sinauer
7. James, W, "The principles of psychology", New York: Holt, 1890, pp. 403-404.
8. Flombaum JI, Santos LR "Rhesus Monkeys Attribute Perception to Others" *Current Biology*, Vol. 15 (8 March 2005), pp. 447-452.
9. Allen, C. "A skeptic's progress." *Biology & Philosophy* (2002), Vol. 17: 695-702.