



UNIVERSITY OF GOTHENBURG

*“Compared to the flowers
and the birds and the trees –
I am an ape man”
– Davies (1970)*

Allocentric Perception in Primates

and the Role of Vision in the Evolution of Thought

GUSTAF LINDBLAD

Master of Cognitive Science

Supervisor: Mikael Jensen

Report No. 2010:120

ISSN: 1651-4769

Abstract

This thesis explores the evidence of allocentric visual processing in primates with the purpose of illuminating a proposed co-evolution of more advanced visual perception and cognition in humans and closely related primates. The thesis is divided into three distinct parts, that in turn deal with three different aspects of this task. Part I reviews the last twenty years of research into the two visual systems hypothesis, proposed originally by Ungerleider and Mishkin in 1982, that separated the human visual system into two discrete subsystems named the dorsal and ventral streams. Different models for the functional subdivision of the human visual system are reviewed and discussed. The ventral stream is considered to handle allocentric visual processing, and is only found in primates, and it is argued that this is a prerequisite for visual object recognition and conscious sight, making this neurological pathway a stepping stone in the evolution of our cognitive systems. Part II considers the behavioural evidence of allocentric perception in different species of primates by linking such perception to different behavioural and cognitive abilities, and proposes a model to interpret the available data and guide further research into the visual capabilities of apes and monkeys. Part III discusses why allocentric perception is an important factor in the evolution of human cognition, and considers some more specific cognitive capabilities that could have evolved from the benefits of having allocentric visual processing.

Index

PRELIMINARIES

- Abstract
- Index
- Acknowledgements
- Introduction
- Thesis purpose and overview
- About this thesis
- A note on Swedish animal research laws

PART 1: TWO VISUAL SYSTEMS

- 1.1 Introduction
- 1.2 Neuroanatomy of the primate visual system
- 1.3 Two cortical visual systems: “What” and “Where”
- 1.4 Separate visual pathways for perception and action
 - 1.4.1. Vision for action
 - 1.4.2. Vision for perception
- 1.5 Allocentric versus egocentric frames of reference
- 1.6 Methods for studying the two streams hypothesis
 - 1.6.1. Patient studies
 - 1.6.2. Optical illusion
- 1.7 Other approaches to the two visual systems
 - 1.7.1. Semantic and pragmatic processing
 - 1.7.2. Three visual systems
- 1.8 Interaction between the dorsal and ventral streams
- 1.9 Lateralization of processing
- 1.10 Summary

PART 2: ALLOCENTRIC PERCEPTION IN PRIMATES

- 2.1 Introduction
- 2.2 Primate taxonomy
- 2.3 Primate cladistics
- 2.4 Evolution of the primate visual cortex
- 2.5 Allocentric perception in non-primates
- 2.6 Studying allocentric processing in primates
 - 2.6.1. Tool use
 - 2.6.2. Social and communicative skills
 - 2.6.3. Causal understanding
 - 2.6.4. Sorting objects on the basis of relative size
 - 2.6.5. Ability to recognize relational similarity
 - 2.6.6. Other skills and behaviours
 - 2.6.7. Skill grouping overview
- 2.7 Behavioural data
 - 2.7.1. Legend
 - 2.7.2. Tool use category
 - 2.7.3. Social category
 - 2.7.4. Reasoning category
- 2.8 On domains and modules

PART 3: VISION FOR COGNITION

- 3.1 Introduction
- 3.2 Solving the unsolvable problem
- 3.3 Vision versus olfaction

- 3.3.1 What is a primary sense
- 3.3.2 Multimodal representational capabilities
- 3.3.3 Speed, precision, distance and details
- 3.3.4 Functionality doesn't smell
- 3.3.5 Spatial relations
- 3.3.6 Movements and temporal relations
- 3.3.7 Vision versus the other senses
- 3.4 Specific cognitive abilities derived from vision
 - 3.4.1 Understanding causality
 - 3.4.2 Problem solving by visualizing
 - 3.4.3 Spatial thinking and arithmetics
 - 3.4.4 Sequential thinking and planning
 - 3.4.5 Social anticipation and reading thoughts
 - 3.4.6 Imitation
- 3.5 Vision driving cognitive evolution
 - 3.5.1 Vision in social interaction
 - 3.5.2 Representations and thought
- 3.6 Conclusion

REFERENCES

Acknowledgements

I would like to thank my wife Hanna Lindblad for the support that she has given me during the laborious process that finally resulted in this thesis. I would also like to thank my father Per Lindblad for reading and commenting on the manuscript at several points during the writing process, and my friend Markus Larsson who provided me with feedback on the nearly finished version. My thank also goes out to Josep Call and Ruth Millikan who provided me with literature and much needed encouragement and inspiration, as well as Mathias Osvath who provided much needed information on Swedish animal research laws. Also a big thank you to all my colleagues at SSKKII for encouragement and support, especially my supervisor Mikael Jensen and Alexander Almér for very important discussions of certain tricky bits.

Figure 1 released under the GNU Free Documentation License.

Figure 2 released to the public domain.

Introduction

It is no longer believed that the only function of the human visual system is to construct a mental visual model of the world, and to use this inner model to guide behaviour in the world. Rather, the visual system is divided into two parts, one that handles the control of motor behaviour in response to visual information, and the other that makes us aware of our environment and allows us to identify objects and make decisions about what actions to take. The evidence actually suggests, perhaps contra-intuitively, that the latter kind of function is not the most common function of animals visual systems – it seems that it might only be primates (and perhaps some other species) that are able to use visual information in this way. In most other seeing animals, visual information is used to control motor action without any conscious processing, the information is directly translated into motor commands. This means that other animals, even though they have sight, are not aware of any visual properties of the world, and rely on other kinds of sensory information for identifying objects and making decisions about actions. Humans on the other hand rely extensively on visual information, and so does other primates. The evolution of this function of the visual system seems to have happened in the primate lineage, and any such function in other animals is likely to have evolved separately.

There are a number of aspects of the primate visual system that sets it apart from other closely related species, and among them are more detailed central vision, extended colour vision and the ability to encode visual information in allocentric coordinates. These are properties that are essential for the visual system to be able to identify objects and create conscious perception of the world in the way that we experience it. Allocentric processing is needed to relate visual stimuli to each other, and it can be argued that object identification is not possible unless different parts of the object can be related to each other and the object can be related to its environment. I will argue that being able to process visual information in this way is a prerequisite for certain other cognitive abilities and skills that we find in primates but generally not in any other mammals.

Thesis purpose and overview

This thesis is aimed at describing why allocentric perception is such an important prerequisite for other cognitive abilities and linking it to such abilities. In order to do this it is necessary to dissect the visual system into functional parts to crystallize what allocentric processing is and what it is not. It is also of interest to see which primate species have what abilities, as this could potentially reveal something about the evolution of allocentric perception. Because of insufficient data about many primate species, the present thesis will not be able to make any such claims, but rather make suggestions for future research that could illuminate this question. I will present a model that can be used to guide research in order to make progress this field.

This thesis is divided into two main parts. The first part is a review of the last thirty years of research into the two visual systems of the primate, placing emphasis on models that make

a distinction between a system for unconscious visuomotor control and another for conscious visual perception. Different perspectives on functional distinctions are presented, with focus on the perceptual function and its ability to encode visual information in allocentric coordinates. The second part considers the evidence for allocentric perception in different species of primates and how it is linked to different cognitive abilities and behaviours. I also propose a model for further research on allocentric perception. The third part takes a more philosophical perspective on visual perception and discusses what I believe to be its important role in the evolution of more advanced cognitive abilities in primates, especially humans. I will also hint at some possible ways of doing empirical research to further investigate my ideas. The third part also serves the purpose of tying the thesis together.

About this thesis

Since this thesis is meant to be readable by all cognitive scientists, it has been necessary to explain some of the biological facts at a pretty basic level. The biologically inclined reader might be annoyed by the explanations and oversimplifications that are abound, but I hope that he or she can look past this and hopefully find something interesting and useful in the reasoning about cognition, which after all is the main area for this thesis. I am *not* a biologist, so there is the possibility that I might have misconstrued some of the finer points of this field. However, I am reasonably assured that any such error would not be in conflict with the main points of this work.

The decision to limit this investigation only to primate species has good reasons. While there are other species that might have allocentric perception, they are not closely related to primates, meaning that their visual systems most likely have separate evolutionary histories. Comparisons with closely related animals without allocentric perception would of course also have been interesting, but the research into the visual systems of most other species than primates is very sparse, not to say erratic. Allocentric perception is closely linked to the bifurcation of the cortical visual system that only seems to be present in primates. Whether or not it is present in other species is not completely settled (for a brief review of some conflicting research into the matter, see Ellard, 1998), but there are reasons to believe that this cortical organization might not be present or certainly not as developed in other species. Many species, for instance rodents that are among the closest relatives of primates, do not have a foveated retina (Ellard, 1998) and without a fovea the acuity of the available visual information is arguably not fine enough to support an equivalence of the ventral stream. One of the main functions of the ventral stream seems to be object identification, something which many primates, and especially humans, do by means of vision. For most other mammals the main way of identifying objects is by olfaction (primate, 2009, In Encyclopædia Britannica Online), which, while it does not rule out the presence of an equivalence of the ventral stream, would make it much less likely to be present.

Comparative studies of more closely related species is often easier to do and can supply more reliable data. So, when we are interested in the evolution of human traits, a good place to start is therefore in other primate species, and especially the great apes. This is however

not enough, since you really need knowledge about other less closely related species as well, in order to create a context for your comparisons to make sense. For instance, the notion that chimpanzees and humans are similar in many respects is only an interesting observation in the light of both species being less similar to other species, like rabbits or haddock (not to mention amoeba, trees or even gloves or newspaper subscriptions). There is also a risk of focusing too much attention to chimpanzees and other great apes when doing research into the evolution of human traits; *“It is equally important to look at distantly related species to gain insights about the way cognitive systems evolved when faced with particular selective pressures.”* (Call, 2004, p. 248) This is of course yet another reason to hope for more research into the cortical visual systems of a wider variety of species.

A note on Swedish animal research laws

Originally this thesis was supposed to be an empirical enquiry into the visual systems of primates with ethological studies of a number of different species (mainly great apes). This endeavour was regrettably hindered by Swedish law and bureaucratic inflexibility. Swedish law forbids any research that includes animals (excluding humans for some reason) without specific permits from the National Board of Agriculture. While it is easy to see the reason for this law – the protection and welfare of animals – and to sympathize with this cause, it is unfortunate that the law does not differentiate between invasive and non-invasive research. Even just looking at an animal and documenting its behaviour is punishable by up to two years in prison, which illustrates the lack of sense in this oversimplified legislation. In the case of great apes, this law is actually counterproductive since captive great apes are stimulated by and enjoy the interaction with researchers (Josep Call, Mathias Osvath, 2009, personal communications).

In order to protect and hopefully increase great ape populations, research into said species is of course of great value, so also in this sense is the Swedish legislation actually counterproductive. Strict control of invasive animal research is a good thing to limit animal suffering, but non-invasive observational studies should not be subject to the same control and unnecessary bureaucracy. It should also be noted that EU law demands that zoo's to partake in (non-invasive) animal research in order to be allowed to hold animals captive. The zoo representatives that I have been in contact with have very positive attitudes towards observational research on their animals and see no risk of animals being harmed in the process.

In the case of this paper, the necessary resources (time and money) were not available to get the required permissions. This means that the present study relies on observational data from various other researchers. I have also changed the emphasis from the original empirical framework towards a more philosophical angle on the evolution of the cortical visual system in primates and its role in the evolution of cognition as a whole, and sketching guidelines for future research. In retrospect it has become a completely different creation than what was initially intended. It might just be that the result is a better paper in the end, since the time available for observations would have been somewhat limited compared to what is preferable in behavioural studies. However it is regrettable that I was not able to get a deeper understanding

of the behaviour of various species of primates that is only available through thorough observation, as this could have had an enhancing influence on the present paper.

PART 1

Two visual systems

1.1 Introduction

There is a massive body of scientific evidence that the primate cortex processes visual information in two neurologically separated and distinct streams emanating from the primary visual cortex (V1). While the anatomical and neurological properties of the streams have been quite clearly established (Goodale & Milner, 2006), their functional characteristics have been a source of debate for the last three decades, a debate that is still active and fuelling research projects and articles. The most widely accepted model for the functional division was suggested by Goodale and Milner in 1992, and it has since been elaborated and well investigated. This model is the main foundation for this thesis and will receive the most attention, but other proposals will be examined as well.

1.2 Neuroanatomy of the primate visual system

The general structure of the neuroanatomy of the visual system is similar across all primates, including humans (Kaas, 2004). It begins, of course, with the retina of the eye, which in all primates has a fovea with a much denser concentration of photoreceptors, in particular photoreceptors that can differentiate colour information (cones). The remainder of the retina contains mainly photoreceptors that are only sensitive to light and not colour (rods). Visual information is passed from the eyes through the optic nerves. The nerves intersect at the optic chiasm, and the information is reorganized into left and right visual field instead of left and right eye. The information from the left visual field is then transmitted in the right optic tract to the right lateral geniculate nucleus (LGN) in the thalamus, and vice versa for the information from the right visual field. The LGN is the first major processing centre of visual information from the retinas, and transmits information to the primary visual cortex (V1) as well as other cortical areas (Gazzaniga et al., 2009).

From V1 information is transmitted through the ventral and the dorsal streams. The ventral stream is defined as beginning in V1 and ending in the inferior temporal cortex (IT), via V2, V3, and V4. The dorsal stream is defined as beginning in V1 and ending in the posterior parietal cortex, going through V2, V3, middle temporal area (MT) and the medial superior temporal area (MST) (Creem & Proffitt, 2001). There are however differences in the cortices of different primates; these definitions and nomenclature apply to human cortex. In the macaque, that was used for the original definitions, the dorsal stream ends in the inferior parietal area rather than the posterior, but it has been argued that these two areas in the monkey and human are homologue (Creem & Proffitt, 2001). Similar arguments can be made for other parts of the

human and monkey cortices, that while the exact location on the cortex may vary there is a large degree of homology (Van Essen, 2005). The dorsal and ventral streams of apes can be divided into a larger number of differentiated areas than than in the monkeys or other primates, but the general orientations and pathways are very similar (Kaas, 2004). The general description is that the dorsal stream extends from the striate cortex (V1) into the parietal cortex, and the ventral stream from V1 into the temporal cortex.

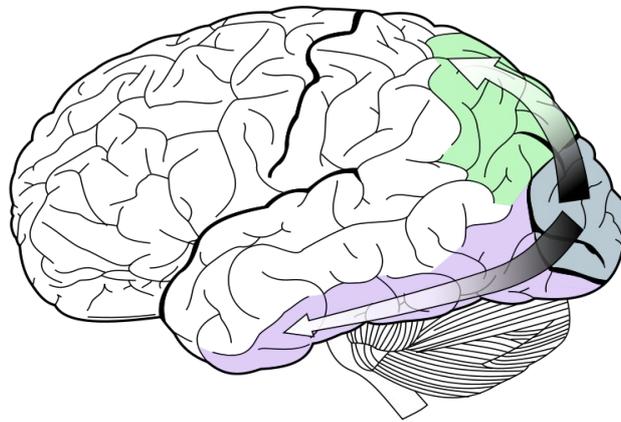


Figure 1: Illustration of the cortical streams of visual processing in the human brain, emanating from V1. The dorsal stream is signified by the green field (upper arrow) and the ventral stream by the purple field (lower arrow).

1.3 Two cortical visual systems: 'What' and 'Where'

The original proposal for the two visual systems was made by Leslie Ungerleider and Mortimer Mishkin (UM) in 1982 in a paper that has proven surprisingly hard to find, considering that it is seminal to this whole field and is referenced by almost every article on this matter since¹. There had been numerous studies putting forward ideas of partitioned visual systems in all kinds of vertebrate species, and a few suggesting a two-fold division according to Goodale and Milner (1992) and Jeannerod and Jacob (2005), but UM were the first to supply the anatomical and behavioural evidence for the two visual streams that has since been demonstrated to be present in most higher primates. The neurologically separated streams were named the 'dorsal' and the 'ventral' stream respectively², reflecting their locations in the cerebral cortex.

Much of their evidence came from lesion studies with macaques, combining different loci of the lesions with behavioural tasks requiring different visual information, namely an object identification task and a landmark (spatial) task. They found that a lesion in the dorsal stream led to impaired performance in the landmark task, but not in the object identification task, and vice versa for the ventral stream. These results led them to conclude that the dorsal stream was responsible for processing spatial visual information and the ventral stream for pro-

1 It is only available as a chapter in a book that is long since out of print, and no library in Sweden seems to have a copy.

2 Standard anatomical terms, dorsal meaning towards the top of the brain and ventral towards the bottom. From Latin dorsum = back, venter = abdomen.

cessing information about object properties, and called them the 'space-channel' and 'object-channel' respectively (Jeannerod & Jacob, 2005). In most of the literature they are called the 'where' and the 'what' streams of visual processing though.

This discovery led to an upsurge of scientific interest in the visual cortex, and much research into the two visual systems of different primates including humans ensued. The existence of two neurologically distinct streams through the cortex was confirmed through several different methods, and for the first ten years the theory of a functional division into 'what' and 'where' was more or less retained by the scientific community and gathered more evidence from the research (Goodale & Milner, 1992).

1.4 Separate visual pathways for perception and action

In 1992 Melvyn Goodale and David Milner (GM) published an article, followed by a book in 1995, proposing another theory that proved to be even more influential on the subsequent research of the two visual streams. Their proposal was that the two streams served two completely different purposes for the organism. UM had assumed that the overall function of the two systems was perceptual awareness and that the two streams subserved two functions of this larger goal. GM proposed that this function was performed by the ventral stream alone, and that the function of the dorsal stream was of a completely different nature, namely that of visuomotor transformation. This concept can be described as a direct link between vision and motor control, the visual information is put directly to use in controlling bodily movements without any prior conscious awareness of it (Goodale & Milner, 1992). In their words, the dorsal stream handles vision for action, while the ventral stream handles vision for perception. The systems have also been named the 'how' and 'what' streams of visual perception.³

While the 'where' and 'what' model proposes differences in the informational input to the systems, the action-perception model is focused on the output of the systems and the functional use of that output (i.e. motor control and perceptual awareness). *“Both streams process information about the structure of objects and about their spatial locations, and both are subject to the modulatory influences of attention.”* (Milner & Goodale, 2008, p. 774) Basically both systems have access to the same information, but it is processed for different purposes and therefore producing very different output. Even though both systems have potential access to the same visual information, that doesn't necessarily mean that both systems always or ever use the same information for their respective tasks.

3 The basic division of vision into action and perception can be experienced by doing a small experiment. Without looking, reach out to an object within your visual periphery and grasp that object. You will most likely succeed in grasping it in an effective and precise way, even though you were not aware of the object's exact position or identity. The interesting thing about peripheral vision is that while you can discern that there are distinct objects present, you will not be able to identify them, or most of the time even tell what colour they are. But reaching out and grasping any such object is most of the time no problem at all, provided that the object does not have a form that is difficult to perform a grasping action on.

1.4.1 Vision for action

“[...]there seems [to be] no reason to suppose that representational system tailored to safely guide an animal's continuous motions through its immediate environment would also be suitable for encoding and amplifying its theoretical knowledge.” (Millikan, 2004, chapter 18)

This paper is mainly concerned with the ventral stream, but for the sake of clarity there is reason to give a little more elaborated account of the dorsal stream and its role in vision. In part this is also of importance for understanding what the ventral stream does and does not do, especially in respect to allocentric and egocentric frames of reference.

One important aspect of vision-for-action is that it is intrinsically unconscious throughout the whole process chain including the actual modulation of motor action. While the decision to perform a certain motor action can be consciously deliberate, the adaptation of that motor action using available visual stimuli is an unconscious and automatic process. Another important point is that the dorsal stream has no working memory (e.g. Goodale & Haffenden, 1998; Goodale & Milner, 2006), all processing is done 'on-line' with the information that is being fed into the system at that very moment from the retinas via V1, but also directly through LGN in the thalamus (Ro, 2008). There is continuous informational input into the system and the system only has access to the current information, nothing is retained in memory for later use. This also means that there could be a potential difference in the control of action when this control is mediated by memory of visual aspects rather than direct visual input, since the memory of visual aspects will necessarily have resulted from processing in the ventral stream and not the dorsal stream (see section about visual illusions).

Jeannerod and Jacob (2005, p. 302) gave this interpretation of visuomotor transformation in GM's model: *“The visuomotor transformation is the automatic conversion of visual information into hand commands for reaching and grasping objects.”* This account seems to be lacking in scope, since it is likely that processing in the dorsal stream can support more adverse motor control than just reaching for or grasping objects with ones hands. Firstly, many primates are quite adept at 'handling' objects with their feet (and so are football players!), and it makes no sense that such behaviour should be subserved by another visual system than that which subserves control of the hands. Secondly, the dorsal stream is probably an evolutionarily older system than the ventral stream (see the next section), in fact it is likely that some equivalence of it is present in most seeing vertebrates, and most vertebrates don't have hands. That the dorsal stream should have evolved in primates into a specific system for controlling object manipulation by hand seems unlikely; it is more probable that it subserves all motor control that is dependent on visual information for successful execution, which would include transportation (e.g. walking or climbing) and obstacle avoidance, head and eye movements and possibly also behaviour pertaining to predators and prey (fight and flight behaviour). It is not clear to me whether or not this limited interpretation of visuomotor transformation is what GM had in mind, but their experimental designs and arguments does suggest that they might be content with such a narrow scope. Goodale and Humphrey (1998) does initially describe visuomotor transformation as being object-directed action, which does

imply a narrow interpretation, but goes on to extend the account so that it includes actions that could be described as object-related rather than object-directed.

The ordinary housefly is a good example of an organism with a very simple visual system (but of course no cortical visual system, since flies don't have cortices) with only rudimentary visuomotor transformation. Basically, the fly measures the speed by which an object in its visual field is growing and this information is directly implemented into one of two behaviours: flight or no flight/landing. If it is flying and an object is growing more and more rapidly in its visual field it will elicit a landing behaviour, if it is stationary and an object is growing in its visual field it will take flight. (Marr, 1982) There is probably a bit more to it (like flying towards more light areas and so forth), but this is the basic plan for how the fly's visual system works. Other behaviours are mediated by chemical perception.

1.4.2 Vision for perception

Vision for perception is the kind of vision that we experience and readily think of when considering vision, and might not need any more explanation. However there are a couple of feats that must be accomplished by the visual system in order for visual perception to be of any use to the organism (actually there are lots of them, but I will only review some that are relevant to this thesis).

First of all it must organize the informational input. What reaches the eye is an array of light of different wavelengths at different points. It is up to the visual system to analyse this mass of almost intelligible information and parse it into meaningful components that it would make sense to identify. In the identification of objects, the system must match the visual object to conceptual information already present in memory. This whole process is extremely complex, considering that the visual system must be able to recognize an object based on virtually an infinite number of different retinal projections. An object can be perceived from this angle or that, in daylight or in lamplight, from close or afar, through a keyhole, on a photograph, half hidden behind uncle Tom, covered in soot and so on (examples inspired by Milikan). The retinal projection from an object in the world will probably never be exactly the same at two different points in time, and still the visual system can recognize the object. It must also be able to track the object as being the same object, even though the projection from the object changes due to saccades, head movements, body movements, object movements, light condition changes etc. The object can also change its actual shape – while walking the human body changes its shape constantly. Even though this is the case we don't perceive any change in the identity of objects – our visual systems achieves object constancy. I will argue that in order to do this, the visual system needs to code the visual information in scene- and/or object-relative coordinates, what is called allocentric coordinates. I will return to this shortly.

In contrast to the visuomotor system, the perception system also needs to remember parts of the scene that is not presently in focus. The eyes can only focus on one object or even on one part of an object at a time, and the construction of the whole scene or object is a patchwork of different focus points over time, with input and guidance from memories and expectations. While the visuomotor system works on-line with low resolution information from the

whole retina, the perception system uses mainly the high resolution colour information from the fovea. This is why it is easy to direct actions at objects in the visual periphery, but much harder to identify them.

1.5 Allocentric versus egocentric frames of reference

The notion that visual information can be processed according to different frames of reference is the most central to this thesis. There is a good deal of evidence suggesting that the cortical visual streams can be classified according to what frame of reference is utilized. Two basic frames of reference are discussed in the literature: allocentric and egocentric.

In an egocentric frame of reference the visual information is coded in relation to some point on the perceiver. It could be relative the perceiver's body as a whole, the head, the eye, the retina, the hand or any other part of the body that is relevant for the processing of the visual information. In an allocentric frame of reference, the information is processed in relation to the object itself, or some point of the object (an object-centred frame of reference), or in relation to some other position in the objects environment (an environment-centred frame of reference).

For visuomotor transformation the processing must be done within an egocentric frame of reference (e.g. Creem & Proffitt, 2001; Goodale & Milner, 1992; Jeannerod & Jacob, 2005), since it is the objects relation to the agent that is critical for a successful motor action. The objects relation to other objects or its environment is normally not important for acting upon the object, so an allocentric frame of reference would be of no use to visuomotor transformation. Jeannerod and Jacob (2005) argue that it would be an impossible amount of work for the visuomotor system to code all retinal information in allocentric coordinates, as it would yield so many coordinates and reference points that the number of calculations would expand in an explosive fashion. This is assuming that the visuomotor system actually processes all retinal information indiscriminately and that there is no selection or focusing in the dorsal stream, which is what at least Jeannerod et al. believe.

In conscious perception, and specifically in object recognition, an allocentric perspective is needed. To be able to select and recognize an object from different angles, distances and other conditions from the visual information available, the visual system needs to encode the relative information of different objects.

“Since perceptual recognition of an object must be achieved from many different spatial perspectives on many different occasions, it requires encoding of visual information about an object's enduring properties. In other words, perceptual recognition of an object demands that visual information about a perceived object matches conceptual information and knowledge about it stored in long-term memory.” (Jeannerod & Jacob, 2005, p. 303)

And since it is unreasonable to suppose that we store information about what an object looks like under all conceivable conditions (how could we ever learn?), this means that the perception system must encode information in an allocentric frame of reference to enable pro-

cessing of object constancy. The best way to explain is through some examples.

The colour of a surface is often an important quality to be able to identify what that surface is, but the colour changes with the lighting conditions. It is often said that a piece of coal in the sunlight reflects more light than a ball of snow in the shadow⁴, still we perceive the first as black and the latter as white. This visual feat is possible because of the relation between the object and its surroundings. Without an allocentric frame of reference this relation would be lost. Similarly, size estimation is aided by the relation between object and surroundings, the allocentric information helping to decide whether you are seeing a horse far away or a really realistic toy right in front of you.

Another good example is that of face recognition. Almost all faces consist of the same features with very little variation in absolute terms (nobody has a giant eye in the middle of the chin, or two noses etc.). Without being able to process the proportions and relations of those features, face recognition would be a very hard task indeed. It has been argued that the human brain (and perhaps also other primates brains) have a special centre or module for face recognition. The question is not settled, but if there is such a module it would most likely be part of the ventral stream and it would certainly process the visual information in allocentric coordinates.

The allocentric perspective can also be argued to be necessary for object recognition because it allows the processing of different parts of an object relative to each other, which of course would be crucial to recognizing that object. This hypothesis rests on the assumption that the object recognition system does not represent objects in a holistic manner, but rather the different features of objects and their relations to each other, which is what the research suggests (Tanifuji et al., 2005).

According to the GM model, this would place allocentric processing in the ventral stream and egocentric processing in the dorsal stream. There is at least one study that claims that this could actually be the main basic division of function between the two streams. Schenk (2006) found that a patient with severe damage to the ventral stream showed impaired performance on tasks requiring allocentric processing, whereas her performance on tasks requiring egocentric processing was normal. An interesting note is that the patient (known as DF in the literature) was actually the same person that GM did much of their research on. So there is good reason to suppose that the ventral stream codes visual information in an allocentric frame of reference, and the dorsal stream in an egocentric frame of reference. This would also be compatible with Jeannerod's model, however Creem and Proffitt (2001) indicates that while there surely is allocentric processing in the ventral stream, this could also be the case in the inferior dorsal stream. However, they present no direct evidence that this is the case, only that it could be. Zaehle et al. (2007) provide evidence that the processing of allocentric and egocentric spatial information is done in at least partly separated neural networks. Summing up, the evidence seems to suggest that the dorsal stream processes information in an egocentric frame of reference and the ventral stream in an allocentric frame of reference, but it is possible that there is a little of both going in the dorsal stream.

4 I am not sure whether this is true though. But it seems plausible, and while it might be false it still illustrates the point.

1.6 Methods for studying the two streams hypothesis

Research of the two visual streams has been done by a lot of different means. Up until circa the 1990's lesion studies with monkeys (or other animals) was quite common, and the first evidence for the hypothesis came from Ungerleider and Mishkins lesion studies with macaques in 1982. While this kind of research provides very good control for the scientist and therefore relatively reliable and clear results, it is a hurtful and indecent way of conducting research and the hurt done to the animals can in my opinion never be justified by any results. Invasive animal research should never be performed unless there is absolutely no other way getting important results that benefits humans or other animals in a significant way. This line of research does not fit either bill, since there are both other methods and the results are not vital for our quality of life.

Today there are brain scanning techniques available that weren't feasible thirty years ago, and while it is very hard to use most of them on any other primates than humans it can to a certain degree provide the same information as invasive techniques. Brain scanning techniques are also more economic, faster and can supply more detailed data. Another technological advancement that has not yet been used in this field (to the best of my knowledge) is eye-tracking, through which it should be possible to study eye movements during different tasks. It would be interesting to see how the patterns of eye movement differ when identifying objects from the patterns when performing object related actions, as this could possibly tell us important things about what information is needed, in what order and how fast it can be processed.

1.6.1 Patient studies

The evidence for Goodale and Milners model came first and foremost from double dissociations in patients with brain lesions affecting either stream. Patients with lesions in the ventral stream suffer from what is called visual form agnosia (or just visual agnosia), and are impaired in recognition and identification tasks, but can interact motorically in a normal way with the objects which they cannot identify. Conversely patients with lesions in the dorsal stream are impaired in their physical interaction with objects which they have no problem recognizing and identifying, a condition called optic ataxia (e.g. Goodale & Milner, 1992; Jeannerod & Jacob, 2005; Milner & Goodale, 2008). Visual agnosia is not to be confused with associative agnosia, in which the patient is able to consciously see the object but still unable to identify what the object is, even though they can describe features of the object accurately. In visual agnosia the patients vision of the object is impaired, but the patient is still able to identify objects from descriptions or through the other senses.

Patients that are suitable subject for researching such double dissociations are very rare, as pure visual agnosia requires the patient to have damage in both left and right ventral streams but no damage to the dorsal stream, and the opposite for optic ataxia. One such patient that has been well described in the literature, is known as DF. She suffers from visual agnosia fol-

lowing carbon monoxide poisoning, that damaged both ventral streams but left the rest of her cortex fairly intact. DF is described as completely unable to recognize shape, size and orientation of objects and not surprisingly she is also unable to identify objects visually. If she is allowed to touch the objects with her hands she can identify them and describe their features as well as anyone could by means of tactile interaction, showing that it cannot be associative agnosia. When DF is instructed to reach out and grasp objects, her movements are described as strikingly accurate, in terms of both arm movements and grip scaling, showing that her visuomotor transformation is intact (e.g. Goodale & Milner, 1992).

One of the problems with using patients with brain lesions for research is that the damages are rarely localized to only the areas that are of interest to the researchers. There is often more diffuse and collateral damage, and damage in one area can very well impair the functioning of another area because of interconnection. Experiments with non-impaired subjects is therefore of great importance to corroborate the evidence from patient studies.

1.6.2 Optical illusions

It has been posited that experiments with optical illusions can show the separation of the visual system in intact brains. Certain optical illusions deceive the visual system because of the relations between the different parts of the illusions, and this should mean that if the action-perception model is correct such visual illusions should not affect visuomotor transformation in the dorsal stream, as this is done in egocentric coordinates. Allocentric processing on the other hand should be tricked by this kind of illusions.

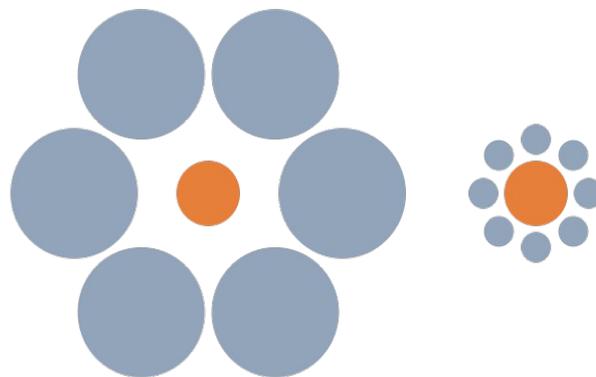


Figure 2: The Ebbinghaus illusion.

Experiments have been done with for instance the Ebbinghaus illusion where two identical circles seem to be of different sizes because of their surroundings (see Figure 2). According to the action-perception model this illusion should trick conscious perception of the circles, but not actions aimed at the same circles. If the subject is instructed to 'grasp' two circles of sizes that match the perceived sizes of the circles in the illusion, approaching grip size will reflect the difference, but if the subject is instructed to 'grasp' the centre circles in the illusion, approaching grip size should not vary according to the illusion effect. As a contrast, if the subject is instructed to estimate an appropriate grip size without directing an action at the actual

circles this grip size should reflect the effect of the illusion. Most experiments have confirmed the hypothesis, but there are some that have contended that these results are not conclusive as the outer circles could act like obstacles affecting grip size in grasping actions (see for instance Milner & Goodale, 2008, for a review of such experiments).

Also predicted from the model is that if a delay is introduced between presentation of the stimulus and the time when the subject is required to act on the stimulus, visual memory of the figure can interfere with the direct processing in the dorsal stream, thus making the action subject to the illusory effect. If the subject is presented with the illusion and the illusion is then removed, requiring the subject to *pretend* to grasp the circle, this action will be guided by visual memory alone and is therefore affected by the illusion.

Several other visual illusions have been used in similar experiments (for instance the Müller-Lyer and Ponzo illusions), and most experiments seem to support the hypothesis. However not all optical illusions should show this pattern, as some illusions arise prior to the stream separation of the visual cortex and should therefore affect processing in both streams (e.g. Milner & Dyde, 2003).

1.7 Other approaches to the two visual systems

The action-perception model had a profound effect on the research field of human (and primate) vision, however the great impact of the 'what' and 'where' model can still be discerned. Most scientists have accepted the main premise of the action-perception model – that vision serves two purposes in primates and that this is reflected in the neuroanatomical organization of the cortical visual system. There are still those who maintain the UM model, as well as some completely different suggestions, for instance Rossetti and colleagues (in Milner & Goodale, 2008) that argue against both models, stating that the evidence cannot support either and propose that the functional division might actually be for processing peripheral and central vision. During the last ten years, there have been research that indicates a possible union of the UM and GM models, with both 'where' and 'how' in the dorsal stream, as well as suggestions of further subdivision of the visual system into three or even more streams. I will review two of the more interesting proposals.

1.7.1 Semantic and pragmatic processing

Jeannerod et al. (e.g. 2005), have proposed a model that is quite similar to the action-perception model, but argue that the GM model seriously underestimates the complexity of visual cognition, and that merely matching perception and action onto the two visual streams is not enough to model the cortical processing of visual information. Their suggestion is a model with emphasis on interaction between the two streams, and that the streams can be further subdivided into parts with different functions. However, they are in agreement with GM about the two main functions of the visual system in primates using their own terminology: 'semantic' processing for object recognition, and 'pragmatic' processing that uses visual information for programming motor function. The purposes of the two kinds of processing are

largely similar to the action-perception model, but they are distributed over both streams.

To me, it seems that Jeannerod et al. and GM are not really talking about the same thing when they speak about action and perception and visual systems. For GM, action and perception are the terms used to describe the functions of the neurological streams in visual cortex, whereas Jeannerod et al. are interested in knowing the full picture of how vision for action and vision for perception are actually achieved. There is not necessarily a complete correlation between the two models. The systems that Jeannerod is interested in are more functional, encompassing different parts of the cortex and possibly a longer chain of cortical interaction. What GM mean by visual systems are the actual distinct neurological streams. As a result of this it might seem that their respective models are more incompatible than might actually be the case. However there are significant differences about the actual localization of function as well, as Jeannerod et al. much like UM places spatial processing in the dorsal stream, whereas GM argue that spatial processing is performed in both streams, but for different purposes and therefore in different ways.

An intriguing idea in Jeannerod et al.'s model is that human visual cognition makes use of several different kinds of visual representations and that there needs to be separate visual systems to accommodate each kind of representation. They propose that these visual systems are distributed over the visual streams, with substantial interaction and also lateralisation of function. However, they explicitly state that at least two of these kinds of representations are unique to humans (Jeannerod & Jacob, 2005) which would also make the corresponding visual systems unique to humans. I have not been able to find any evidence supporting this claim, and at present this makes their model unfit to serve as the basis for this thesis even though the more elaborate subdivision of the visual system is an attractive idea to serve as a basis for comparative studies.

1.7.2 Three visual systems

It is certainly implied in the Jeannerod model, but it is spelled out in a proposal made by Creem and Proffitt (2001), that there might in fact not be just two cortical visual systems, but actually three⁵. The proposal can be regarded as an integration of the UM and GM model, and asserts that the dorsal stream might in fact be two streams encompassing respectively 'where' and 'how'. Jeannerod's model also places spatial processing in the dorsal stream, but neither Jeannerod nor Creem and Proffitt declare definitively whether this means that there is no spatial processing at all in the ventral stream, as is the case in the UM model. Creem and Proffitt suggest that the dorsal stream can be divided into a superior (upper) and inferior (lower) part, with the superior stream subserving visuomotor transformation ('how') and the inferior handling spatial processing ('where'). They also present some evidence that points to a lateralisa-

5 Actually Jeannerod and Jacob (2005) suggests a functional division of even more areas in the dorsal stream, but are unclear to whether these parts are to be considered as modules in their own right or as components of a consolidated dorsal stream. However, their main idea is that the visual system has to principal functions and that these functions can be divided into a number of subfunctions that are implemented by a number of areas located in both streams.

tion of spatial processing to the right hemisphere, a lateralisation also suggested by Jeannerod and Jacob (2005).

1.8 Interaction between the dorsal and ventral streams

The divide of vision into the two cortical streams is a neuroanatomical fact. Regarding functionality there is good evidence for the basic division into action and perception, even though, as Jeannerod and Jacob stress (2005), this is likely to be an oversimplification. In normal situations there is of course no real need to separate the two functions, as we use both simultaneously and in an intertwined fashion; There is considerable crosstalk between the two streams, and other cortical areas get input from both. Much of the evidence for the division comes from cases where one stream has been compromised, and the subject is left with only one kind of visual processing, resulting in optic ataxia or visual agnosia depending on which stream it is. In non-impaired vision, both streams are likely to be used in all kinds of situations, though not to the same degree (Goodale & Humphrey, 1998).

Perception provides both foundation for choice of actions as well as information that is relevant for the completion of actions, and is highly relevant for all action. Likewise, vision-for-action is likely to affect the perceptual system as our actions in the world is likely to affect our perception of the world (see for instance Milner, 1998). Jeannerod and Jacob (2005, p. 311) goes so far as to say: *“Not only is what one can do shaped by what one perceives, but also conversely what one can do shapes what one can perceive.”* The word 'shape' is of course quite vague, but from this quote you can get the impression that what one can see influences what one can do in an equal amount compared to the opposite. This seems rather questionable to me, as perceiving will surely be more important for acting than the other way around. If I were to become paralysed, surely this would not influence my ability to perceive the world in a comparable amount to how much my ability to act in the world would be impaired by the loss of conscious perception. On the other hand, in an evolutionary perspective, what the organisms have been able to do have probably had an effect on the evolution of the visual system, as the visual system ultimately is in the service of behaviour whether it be directly through visuomotor transformation or by means of conscious perception.

1.9 Lateralisation of processing

I will digress for a moment to consider an interesting issue that I have found no real discussion of in the literature, though it intrigues me. It is commonly called the 'two cortical streams of visual processing', but this is really a somewhat one-sided description, as the brain has two hemispheres, each with its own dorsal and ventral streams. The lateralisation of these streams is just as much a neuroanatomical division as the separation of the dorsal and the ventral in one hemisphere. There is considerable crosstalk between the ventral and the dorsal, but less is known about the communication between the left and right streams. There is bound to be communication via the corpus callosum, but I've found very little information on the nature of

this communication. Studies with split brain patients could perhaps illuminate this matter.

Creem and Proffitt (2001) and Jeannerod and Jacob (2005) have written briefly about lateralisation in the dorsal stream, where they notice a tendency that spatial processing might be localized in the right inferior parietal cortex. Jeannerod and Jacob also have some additional ideas about lateralisation in the dorsal stream, but nothing is said about the ventral stream. This is actually quite puzzling.

Lateralisation in visuomotor transformation makes sense, as motor control of the left side of the body is separated from control of the right side. It would also make sense to code the left versus the right visual field separately. However it is not so clear that this would be the case for visual perception in the ventral stream. There is no need to separate the right from left visual field when identifying objects, and the conscious experience of seeing is unified. So how is this unification achieved and what does the left versus the right ventral stream do? Both ventral streams seem to get input originating in both visual fields, so the question is what difference there is in the processing of such information. Perhaps it is a segregation of left and right side of the perceiver, with the left side processing information about objects to the right of the perceiver and vice versa, even though information for both fields is available. But this really begs the questions why then there is no segregation of informational input, and where and how objects that are right in front of the perceiver are processed.

It could be that the left and right ventral streams process information about one side of objects each regardless of the availability of information. But that leaves the questions of why there should exist such a division and how and where the unification is achieved. Object identification has no obvious benefit from left-right segregation, and the fact that mainly foveal information is used would further make such a segregation unlikely. There are cases of unilateral neglect where the patient is oblivious of one side of objects (typically the left), which possibly could support such a hypothesis, but the leading theory about neglect states that it generally is caused by lesions in the parietal or frontal cortex, and that it is a deficit in attention, rather than in perceptual processing (Gazzaniga et al., 2009). This would mean that the problem is not in the ventral stream itself, but in what information reaches consciousness. There is evidence that even though the patient does not become aware of information in the neglected side, the information is processed and can influence judgement unbeknownst to the patient.

Another alternative is to suppose that the left and the right ventral streams process for different aspects of visual processing. The left hemisphere of the brain is generally associated with semantics (language) and details, while the right side is more important for patterns and understanding the gist of things. If this (rather simplified) account of the division of propensities for different tasks is applied to the visual system as well, there are some quite interesting and testable hypotheses that could be made. Alas, it is not the aim of this thesis, but the question is an interesting one and merits further investigation.

1.10 Summary

While most of the different models that have been proposed within the research field of the

two visual systems framework are focused on defining more or less exclusive functions for the two visual streams, two research teams stand out. On the one hand we have Jeannerod et al. that believe that there are two or more visual systems in the primate brain, but that these systems are not strictly related to the neuroanatomical structure of the cortex, but are derivatives of the overall functioning of the whole visual cortex. On the other hand Creem and Proffitt propose three main visual systems in what could be described as a compromise between the two most influential two visual systems models, UM and GM. The following table is a schematic summary of the different models that have been reviewed above.

	Ventral	Dorsal	
Ungerleider & Mishkin	what	where	
Goodale & Milner	what	how	
Rossetti et al.	central	peripheral	
Schenk	allocentric	egocentric	
Jeannerod et al.	semantic	pragmatic	
Creem & Proffitt	what	where	how

Table 1: Overview of the different models reviewed in this thesis.

There is one thing that most researchers today agree on, and that is that primate vision serves two main functions for the organism: vision for conscious perception (vision for perception in GM's model or semantic processing in Jeannerods model), and vision for regulation of motor behaviour (vision for action in GM's model, and pragmatic processing in Jeannerods model). Vision of the latter kind is also agreed to be a function that most organisms with vision have⁶, whereas vision for perception is held to be a less common function. The neurological evidence suggests that if it is present in animals other than primates, it has evolved separately.

There is much evidence that conscious perception is associated with the ventral stream, and likewise is visuomotor transformation linked to the dorsal stream. And while it is reasonable to assume that Jeannerod and Jacob are correct when they say that visual cognition is far more complex than this, there stills seems to be a great deal of merit to this simple divide. It might perhaps be that visual perception is achieved in the ventral stream, but that processing in the dorsal stream can add to this function, like my coffee maker helps in heating up my house even though I would not take this to be its function (and I would certainly not go out and buy

⁶ There might be organisms that have vision solely for other purposes, for instance for synchronizing circadian rhythm to daylight cycles.

forty coffee makers to get through the winter). One thing seems certain and that is that there seems to be no chance of conscious perception without a functional ventral stream (Milner & Goodale, 2008). Using visual information for guiding motor behaviour is of course possible without a functional dorsal stream, but it becomes slow, inefficient and prone to error as it must pass through the ventral stream first and demands conscious attention to be successful.

The integrative use of processing in both streams is bound to have synergistic effects on visual cognition as a whole. There is known to be considerable crosstalk between the two streams and there is no reason to suppose that the rest of the associative cortex cannot use informational output from both systems. Because of this, I see no reason that not both GM and Jeannerod can be right about the essential principles of their models. For the purposes of the following ethological analysis I will adhere to GM's basic model, even though I believe it is far from settled that it is correct. At present it is the model with the most convincing scientific support, and it is also more parsimonious than other proposed models.

So even though it is still unresolved exactly how and where allocentric perception is achieved, at present the evidence favours the ventral stream.

PART 2

Allocentric perception in primates

2.1 Introduction

There have been speculations that processing of visual information in allocentric coordinates might not be possible for all primates, and it is a valid question since there does not seem to exist any integrative study that supports or invalidates such a hypothesis. Three main lines of research are integral to answer this question: research on the evolution of primate cognition, research on the neuroanatomy of primate visual systems and experimental and ethological studies of primate behaviour, all of which have some degree of overlap. I will argue that allocentric perception need not be an all or nothing question – few abilities are. There might be differences in degrees or it is possible that different kinds or areas of allocentric processing can be identified, and that not all the various species have developed all kinds. For instance there could be different levels of how fast or how elaborate such processing can be done, or allocentric processing of objects (or specific objects like faces) but not for whole scenes etc. The present thesis cannot answer any such questions with any empirical data, but will give suggestions for how to think about these matters and possible future research strategies.

Allocentric processing has probably evolved in the primate lineage and it should therefore be possible to study this evolution to some extent. The evolution of the trait would most likely be closely tied to any discernible degrees or areas of such processing. Analysis of skills and behaviours of different species of primates could thus give an account of the evolutionary history of the perceptual system, according to how the primate species are historically related, something which is called a cladistic analysis. As the behavioural data that I have available are not enough to support any specific claims about such evolution I will mainly present a model for how this research could progress, as well as some speculation based on the data that is available.

2.2 Primate taxonomy

In order to investigate the evolution of a cognitive trait (or any trait for that matter) you need to know the evolutionary history and kinship of the species with that trait. For the purposes of this thesis it is necessary to be familiar with the primate lineage. While this is perhaps common knowledge to a certain extent, there might be a need to review this taxonomy.

Classification of animals into taxa is not a simple thing, since there are many things to account for and you could have different reasons and uses for the classifications. Creating a taxonomy that reliably reflects actual historical kinship requires that all available knowledge and methods be used. Genetics, zoology, osteology, geology, ethology and chemistry are the most important research fields that contribute to this enormous task. As could be expected

there are is no classification that is universally agreed upon by all or even a majority of all scientists, but much of the basic structure is pretty solid. Regarding the order of primates there are naturally some different suggestions, but I will stick to the most common classification and not really go into the details since the differences are small and not of any major importance to my work (see for instance Dawkins (2004), Begun (2004), Ross & Martin (2007)).

The biological order of Primates is divided into two suborders, Strepsirrhini (e.g. lemurs) and Haplorrhini (or Anthropoidea), and the later group is further divided into Tarsiiformes and Simiiformes. The group Simiiformes contains monkeys and apes and is the order with which this thesis is concerned. Research on primate cognition has mainly been done within this group, and even though much of our knowledge from this research might be extendible to the rest of the primate order, one should only do so with caution. If one is picky, many instances of the word *primate* in this paper should read *simiiformes* or *simians*, but it seems that a common practice is to refer to this group simply as primates. I too will follow this convention (since *primate* is a much more familiar and graceful sounding word than *simiiformes*) except in cases where it is important to separate the simians from other primates. The simians are sometimes referred to as “higher primates”, and even though I find this term a bit misleading (altitude has got very little to do with these matters), I find it quite useful to give the text more flow.

Regarding the neglected primate species (tarsiers, lemurs etc.), the applicable research is extremely scarce. These animals are in many ways substantially different from the higher primates and even in the absence of much research it is probably safe to assume that much of their perceptual and cognitive systems lack the complexity that is distinctive of the higher primates (Kaas, 2005).

The infraorder of simians is further divided into Platyrrhini, that consists of New World monkeys, and Catarrhini encompassing Old World monkeys (Cercopithecoidea) and apes (Hominoidea). Apes can be further sub categorized into gibbons that are also known as lesser apes (family Hylobatidae), and great apes (family Hominidae) which includes humans. This nomenclature is not universally agreed upon, some would for instance not place humans in the same family as other great apes, but rather have a separate family just for our species. I will follow the former classification as it makes more sense from a cladistic perspective. For instance chimpanzees and humans are more closely related than chimpanzees and gorillas in terms of when the last common ancestors lived (by a million years or so), and this makes the latter kind of classification less tenable (Dawkins, 2004). The family of great apes consist of humans (*Homo*), chimpanzees (*Pan*), gorillas (*Gorilla*) and orang-utans (*Pongo*). Except for *Homo* that only consist of one extant species, multiple species are recognized in each of these genera today.

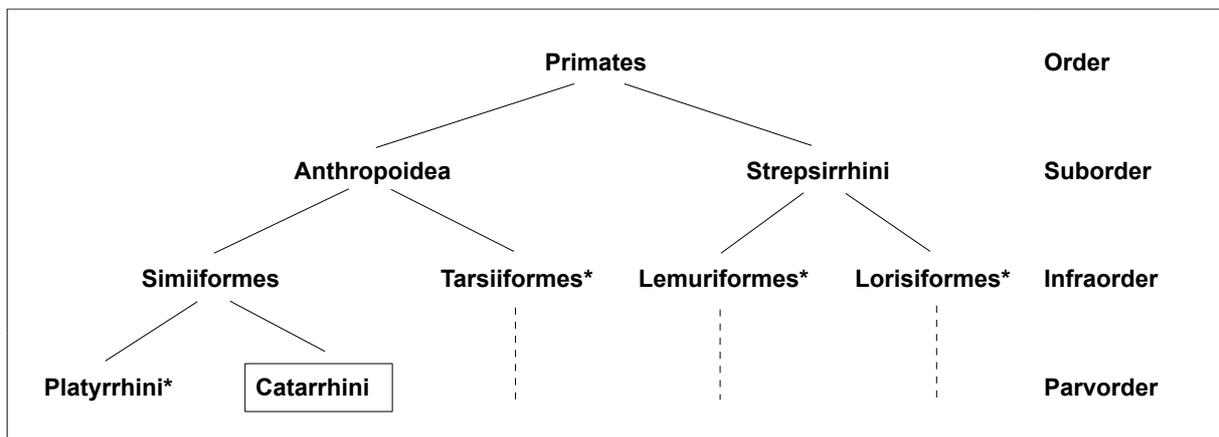


Figure 3: The orders of primates. Created from Dawkins (2004), *primate* (2009) In *Encyclopædia Britannica Online*, Ross & Martin (2007). *Further subdivision omitted.

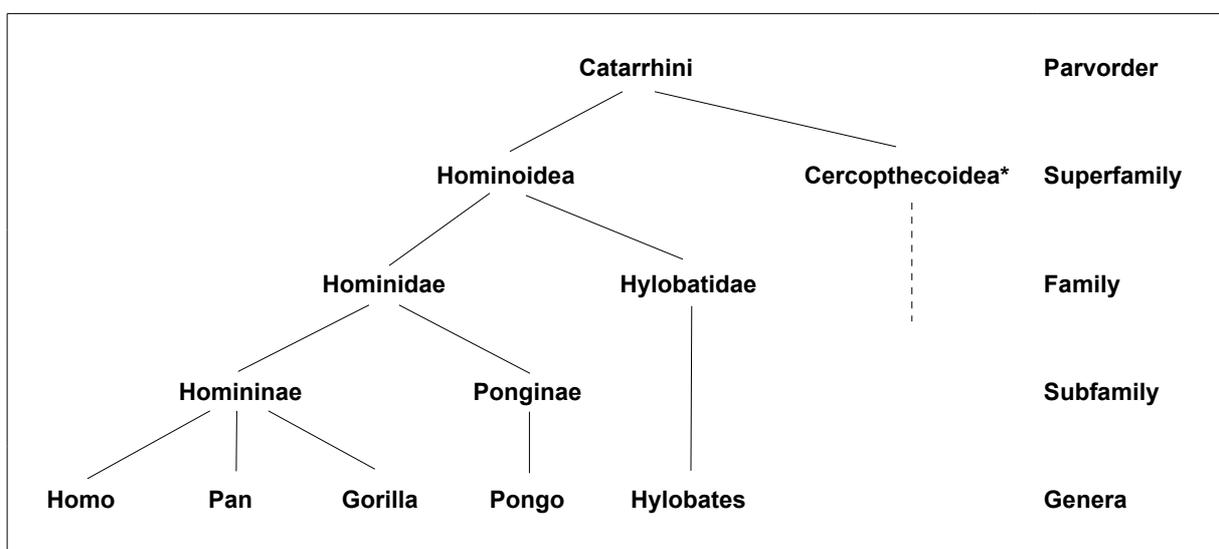


Figure 4: The families of Catarrhini. Created from Dawkins (2004), *primate* (2009) In *Encyclopædia Britannica Online*, Ross & Martin (2007). *Further subdivision into different genera of Old World monkeys omitted.

2.3 Primate cladistics

The research on the cognitive skills of primates is not evenly distributed among the different species. After humans, the most studied species are the chimpanzees. There is significantly less research on gorillas and orang-utans, and gibbons have been the focus of very little research thus far. There are more than ten different species of gibbons and over a hundred species of Old World monkeys as well as about a hundred species of New World monkeys. Among monkeys there is a significant amount of research on both Old World and New World monkeys, but most of it is restricted to only a few species. For a research model it is not feasible to do research on all these species, and there needs to be a grouping of species to make it practical and useful, therefore a select few species have been treated as representatives of their families. For the model I will propose I will adhere to this tradition, and have made a grouping according to the following table (table 2).

Group	Species
Humans	Homo sapiens
Chimpanzees	Pan troglodytes Pan paniscus
Gorillas	Gorilla gorilla Gorilla beringei
Orang-utans	Pongo pygmaeus Pongo abelii
Gibbons	All species of Hylobatidae
Old World monkeys	All species of Cercopithecidae
New World monkeys	All species of Platyrrhini

Table 2: Functional groupings that will be used in the research model.

This grouping is the cladistic grouping that emerges if the human ancestry is traced backwards along a line of ancestors. Every group of species shares a (in principle) single common ancestor with humans at some point in history. A group with a common ancestor is called a clade. This means that the genera *Homo* and *Pan* constitutes a common clade as well as being two separate clades, and the whole order of primates is also a clade. Tracing back from humans through the primate lineage and plotting the different branches of the extant clades would produce these groupings. There is fairly reliable information of when the last common ancestors (LCA) of every branch lived, and this information can be used to get an idea of when traits are likely to have evolved. The following schematic shows estimations of when the LCA of humans and the other primate clades considered in this thesis lived.

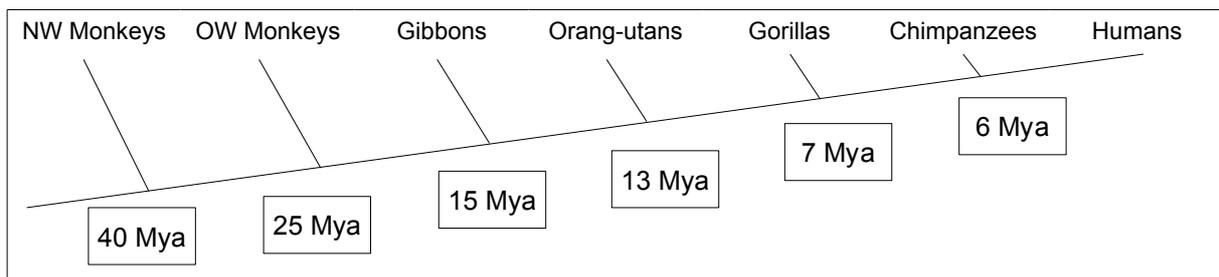


Figure 3: Timeline of when the last common ancestors of the different primate groups lived. The numbers are approximations. Mya = Million years ago. The figure is not to scale, and distances should not be understood as representing anything. Created from Glazko & Nei (2003).

The main idea for appraising the evolution of cognitive or other traits is that if it is present in all (or most) species within a clade, then it is most likely that their common ancestor also had that trait. If the trait is only present in a minority of the species within the clade, then it is

likely that their LCA did not have the trait. Knowing when the LCA's of the different clades lived makes it possible to estimate the time period during which a certain trait evolved or disappeared.

I will illustrate with some different examples. Within the primate clade, only humans exhibit spoken language so it is therefore likely that this trait has evolved after the last LCA of chimpanzees and humans, that is less than six million years ago (and probably more recently). Or an example related to vision: Trichromatism is present in virtually all Old World monkeys and apes, but not in New World monkeys or other primates, which makes it most likely that this trait evolved somewhere between the LCA of New World monkeys and the LCA of all other anthropoidea, meaning that it evolved between 40 and 25 million years ago. Almost all monkeys have tails, and even though no apes have them this makes it likely that the common ancestor of all anthropoidea had tails since most other primates also have tails, placing the loss of the tail somewhere between the LCA of Old World monkeys and the LCA of apes.

2.4 Evolution of the primate visual cortex

The primate visual cortex consists of a large number of areas that are unique to primates, and the human cortex seems to be the one with the most number of differentiated areas. This is not very surprising considering the fact that the human brain is larger relative to all other primate species with more room for such differentiation. Areas V1 and V2 are present in most mammals, but so far there is no evidence of the existence of area V3 or any of the visual areas designated with numbers in humans in any other mammal than primates (Kaas, 2005). Some of the more caudal visual areas, like V3, are likely to have evolved with the first primates as they are present in all studied primate species. In early primates there was an expansion of the visual cortex, especially in the temporal lobe, and in the anthropoids there was an even greater expansion with addition of numerous visual areas (Kaas, 2005). It is believed that early primates were nocturnal, which makes it likely that these animals, like their closest ancestors, used olfaction as their primary sensory system. The switch to vision as primary sensory system is likely to have co-occurred with or happened after primates became diurnal, but it makes sense to speculate that some development of the visual system preceded and perhaps even was a factor in initiating the switch from night to day living.

New World monkeys stand out from the other primates considered in this thesis. In many respects, their sensory faculties are less like the rest of the anthropoidea, but more like those of other mammals. For example, New World monkeys are dichromates and use olfaction as the primary sensory system for a number of functions (including social functions), which is typical for most other mammal species (primate, 2009, in Encyclopædia Britannica Online). The higher primates are trichromates and use olfaction to a much lower degree. Trichromatism allows for more extensive colour vision, which in turn makes the visual system more sensitive to finer details. Not much is known about the differences in visual cortex organization between New World and Old World monkeys, but *“a number of differences seem likely, given their long, separate evolutionary history”* (Kaas, 2005, p. 277).

“[T]he origins of vision may be related more to its contribution to the control of action than to its role in conscious perception, a function which appears to be a relative newcomer on the evolutionary scene” (Goodale & Humphrey, 1998, p. 183).

There is a great deal of evidence that visuomotor transformation is something that is present in all vertebrates with vision. This makes it likely that the dorsal stream that performs this function in primates is either an old evolutionary structure or a development from such a structure (Goodale & Milner, 1992). The primitive vertebrates in which vision evolved had no cortex, so visuomotor transformation must have been implemented by more primitive brain circuits. In extant species without cortex but with visual guidance of action, this is achieved in the mesencephalon and other parts of the brain stem. Such visual systems are assumed to consist of specialized modules that control specific motor outputs in response to innately specified stimuli (Goodale & Humphrey, 1998; Sowards & Sowards, 2002). Subcortical structures still play a role in vision in mammals, including humans, but the emergence of a visual cortex has allowed for more complex visual processing and mediation of a wider behavioural repertoire.

The ventral stream can only be found in primates, which would make it a much later evolutionary development. Primates seem to be the only animals that can encode visual information in allocentric coordinates (Zaehle et al., 2006), which would make primates the only animals capable of true object identification through vision.⁷ Many animals can of course react selectively to objects giving the impression of being able to identify objects by means of vision, but there are good reasons to believe that this is not by means of conscious visual perception of the objects (Sowards & Sowards, 2002). Goodale and Humphrey (2008) sees the emergence of the new organization and features in the primate visual system as a result of pressure from more complex cognitive and social behaviours, but give no indication as to why they believe this was the direction of causation. I would like to emphasize that it could just as easily be the other way around, that the more capable visual system allows for more complex cognitive and social behaviour, or even more probably that more complex behaviour and more capable vision evolved together, providing pressure and possibilities in both directions. In the last part of this thesis I will discuss this topic at more length.

There seems to be some kind of scaling in how complex and differentiated the visual cortex is, from certain New World monkeys that has the least convoluted cortices of the studied species, to the apes that have the most complex with the apex in human cortex (Kaas, 2005). This might not be all too surprising, considering that the same rough pattern can be found for overall relative brain size, and a larger brain will have more room for additional differentiation. Kaas (2005) argues that the brain of certain New World monkeys (e.g. the galago) is likely to be relatively similar to the brains of our ancestral primitive primates. Likewise, it is believed that the shared ancestor of humans and chimpanzees was much more like chimps than humans, a reasonable assumption considering that chimps are more similar to other apes than humans are (at least in terms of cognitive abilities). Since there seems to be an increase

⁷ Though there are some other species that are good candidates, see next section (2.5).

in complexity of the visual cortex that roughly follows the lineage of the primates, this could tell us something about the evolution of vision in primates. Humans obviously have more advanced cognitive capabilities than other primates, and it is reasonable to assume that species that are more closely related to humans will also have more advanced cognition. This is also, in general, what the research shows. But there is cause for caution in these matters, as there are no hierarchies linking species to each other but only common ancestors. Every lineage of species are subject to their own selectional pressures and the evolution of that lineage will reflect this, not the selectional pressures of any other lineage no matter how closely related they are.

2.5 Allocentric perception in non-primates

The literature boldly claims that only primates have allocentric perception, but there is reason to pause and think a little about this claim. The literature also claims that in order to identify objects through vision, the animal needs to process the visual information in allocentric coordinates. This would mean that only primates are capable of identifying objects through vision. There is evidence of other animals than primates doing just that, as certain species of birds as well as elephants and dolphins seem to be able to recognize themselves in a mirror (e.g. Marino, 2004). Since the mirror image cannot supply any other kinds of information than visual, it seems that at least one of these statements must be wrong. Certain birds, elephants and dolphins engage in other activities that arguably would imply that they might both be able to identify objects through vision as well as processing information in allocentric coordinates (e.g. Clayton & Emery, 2005; Bates et al., 2008).

The coding of sensory information in allocentric coordinates can only be done with certain kinds of sensory information, such information that reflects the spatial configuration of the world. This directly disqualifies such senses as taste and touch, since taste only gives information about that part of the world which is in the mouth and touch only about what is in contact with the body. In the case of hearing and smell there is a certain amount of spatial information that can be perceived, but arguably not enough to be processed in allocentric coordinates. Animals with a keen sense of smell are sensitive enough to locate the general direction from which a smell comes, but because of the general properties of air the smell is diffused in a way that prohibits any very precise such judgements of location to be made. The spatial relation between two sources of smell could not be precisely judged from even a short distance. Using ordinary hearing to judge the spatial properties of the world would of course be an impossible task as most things in the world does not make any distinctive sounds, a fact that also applies to smell (a hole in the ground probably smells the same as the ground).

Constructing a detailed spatial representation of the environment with preserved relations between constituents through hearing, smell, touch or taste seems to be a very hard task to accomplish. Adding the identification of those constituents from afar makes it basically impossible. Creating such a mental representation of the world is in the case of humans done mainly through vision and memory, with relevant modification and input from the other senses. There is only one known sensory system other than vision that could in principle

allow for spatial perception of the kind necessary for constructing such a representation and that is echolocation. As echolocation also reveals the spatial configuration of the world with its different surfaces, it could in principle provide sensory information that could be processed in an allocentric frame of reference. Though judging from the behavioural repertoire of most echolocating animals, I suspect that it is predominantly used in a fashion similar to visuomotor transformation. A possible exception could be the dolphin.

In the case of certain species of birds (mainly corvids) and dolphins, these animals show signs of cognitive capabilities that are similar to those of primates, and are rare or non-existent among other animals. Corvids are known to use tools, and both dolphins and certain birds seem to be able to discriminate among objects and solve problems with the aid of vision. Assuming that this is so, there are two possibilities: Either allocentric perception is not needed for this kind of vision, or other animals than primates process sensory information in allocentric coordinates. I would assume the latter. There are good arguments for why allocentric perception is needed for object identification, but the evidence that only primates have this capability is lacking. The only available evidence is that the ventral stream in which allocentric processing is achieved has evolved only in the primate lineage. But this is by no means conclusive, as allocentric processing as easily could be achieved by different and independent neural developments in other lineages. It is true however, that this capacity seems to have evolved separately in the primate from any other such visual system in any other group of animals. Visual perception has been studied quite extensively in several species of primates, but the research of the visual systems in other animals is lacking. Further research will probably show that visual systems similar in function to the ventral stream have evolved in other lineages as well.

2.6 Studying allocentric perception in primates

As has been mentioned, allocentric perception is not unique to man. This should be quite clear considering the fact that other primates than humans use vision for object identification, which calls for coding of visual information in allocentric metrics. Haun et al. (2006) showed in a series of experiments that four species of apes (humans, chimpanzees, gorillas and orangutans) prefer an allocentric frame of reference to an egocentric when solving a spatial problem relying on vision. This can also be established through studying behaviours that are reliant on allocentric visual information. But as I've already stated, it would probably be an oversimplification to regard this question from a standpoint that posits allocentric processing to “be or not to be”. It would be interesting to find out if there could be levels or degrees of allocentric processing, or if it could be area specific. A relevant question is to enquire into to what extent it is integrated with other cognitive capacities. The coding of visual information in allocentric metrics is of course of no use if the organism cannot put that information to use. Turned on its head, there is probably no more allocentric processing than there needs to be in order for the animal to perform the actions that it performs. For reasons of developmental and energy consumption economics, a system will not evolve to be more “expensive” than it needs to be. The larger brain area that is needed to accommodate more extensive allocentric processing leads to

a larger brain which typically requires a longer ontogenic developmental time, and consumes more energy to function. It should therefore be possible to say something about the probable evolution of allocentric processing from studying the distribution of skills in different primate species. This is the foundational idea that the following proposition rests on.

So what behaviours are relevant to study when trying to ascertain the presence of allocentric perception, and what behaviours can be said to pertain to what levels or areas of allocentric perception? To start with the second part of the question, levels can only be identified through the behaviours that they mediate. The conscious experience of different levels of allocentric perception is, as is all other conscious experience, private and inaccessible to others than the experiencer. Different levels of such visual processing could then be studied by means of brain scanning, but would still only make sense in relation to the different behaviours that are hypothesised to accompany these levels. This is of course a somewhat circular chain of inference, but so is basically all inferences at some point.

To a certain extent the behaviours that are of interest are behaviours that are, if not unique to, then at least characteristic of primates but non-existent or scarce in other mammals. If the behaviour can be argued to be conditional on allocentric perception, the behaviour should not be present in animals lacking this ability. However, this is a statement that needs some qualifying, as behaviours that seem similar need not be so. Different animals have different phenotypes and different ecological conditions, and so two behaviours that seem similar can serve very different purposes, and two behaviours that seem dissimilar might in fact serve the same purpose (Papini, 2008). To understand the behaviours of a certain species, you must be familiar with that species' biology and ecological niche.

Of course not all behaviours that are typical of, or unique to primates are served by allocentric perception. Many primates are adept at grasping objects with their hands and feet, something that few other animals do, but such a skill would be served mainly by the dorsal stream and in an egocentric frame of reference as we have seen. The behaviours of interest for studying allocentric perception should in some way or another require the perception of relations between objects (object is to be understood in a wide sense: an apple, a tool, a conspecific, a tree or the ground) and/or events (any kind of phenomenon that can be perceived visually). In the next section I will present an inventory of behavioural kinds that I will argue fulfil these criteria.

There is no scientific evidence for any such levels of allocentric perception at present, it is just a suggestion for a model that could help explain the apparent expansion and diversification of the visual cortex, and especially the ventral stream, in the primate lineage. As there is much evidence that humans and primate species more closely related to humans have a larger visual cortex with a more complex and compartmentalized organization, this begs the question what it is used for. It could of course just be as simple as that it provides more acute vision with finer detail, but the explorations of the visual system suggests otherwise. My suggestion is that the larger visual cortex is directly involved in subserving more advanced cognitive processes that uses visual information, with different parts of the visual system evolved to process information for different tasks, perhaps in a modular fashion.

Melvyn Goodale (1998) made a good point when he said that you cannot really tell where

vision ends and motor control begins in the dorsal stream – there are no boundaries between a visual cortex and a motor cortex. I would like to make a similar point: Where does vision end and cognition begin? The ventral stream processes visual information and turns it into conscious perception and makes it available for other cognitive manipulations, such as memory, problem solving, reasoning etc. But where is the boundary where it ceases to be visual processing and begins to be other cognitive processing? My answer is that there is no such point, as it all is part of the same overall system, evolved to solve a number of problems in order to facilitate survival and procreation. When we think about objects or space there is bound to be activation in the visual cortex, and this is of course part of thinking about such things.

The proposed method is to use behavioural data from primates and map it to certain cognitive skills judged to be interconnected with allocentric visual perception. My theory is that primate species that are more closely related to humans and have larger and more differentiated visual cortices also have a larger repertoire of cognitive skills of this kind. There should also be a roughly systematic increase in the distribution of skills, such that a skill shown by a less closely related species should also in principle be present in more closely related species. Obviously this might not be the case for a number of reasons that would not contradict my claim, but there should be some kind of discernible trend. Further, should this trend prove to be both clear and systematic (which I hope it will be), then there is good reason to be excited about it.

This presented scheme is an interim working model. To a large part it is somewhat speculative, as is necessary for now as there is not enough research to support it. I am not an expert on primates or primate cognition or behaviour, the arrangement reflects my present knowledge of these matters, and it is therefore feasible that there exists scientific evidence contradicting certain parts of it. Such evidence should be accumulated and incorporated into the model, extending and attuning it. All my suggestions might be proven wrong eventually, but it is a starting point. Perhaps it will not be fruitful to talk about 'levels' of allocentric processing, but rather modes or varieties of visual processing that also happen to be allocentric in nature. Perhaps all these skills rely on the same basic visual capacities. But at present I see this way of framing the problem as a useful and promising model. I will also suggest some directions for future research that could help develop and corroborate the model, as well as research within the model with the purpose of extending the knowledge of the evolution of allocentric perception in primates.

2.6.1 Tool use

Many species of apes manufacture and use tools, mainly for food retrieval purposes like cracking open nuts, fishing for termites and such. Tool use has also been documented in many species of monkeys. The evidence comes from both ethological studies of wild animals and experimental studies with animals in captivity. There is of course no such thing as a specific 'tool use behaviour', different tool use behaviours most certainly rely in part on different cognitive abilities. The etiquette 'tool use' is a collecting term for a wide variety of behaviours that have in common that some kind of object is used to achieve the goals of the agent.

There are several important aspects of primate tool use behaviour that make it a good candidate for studying allocentric processing. A pivotal part of using tools is selecting what objects would be appropriate tools, and this requires visual processing of relational properties of the object itself and in relation other objects that the tool might be used with. For instance a stick for termite fishing has to have a certain length, be slender and straight enough etc. More than that, it might also have to be stripped of twigs and leaves, requiring the primate to actually manufacture the appropriate tool. The use in itself of such a stick does not require allocentric perception, as it can be viewed as an elongation of the arm in itself, and thus reliant on egocentric coordinates. However, choosing and manufacturing it cannot be done within an egocentric frame of reference. Modification by removing parts from an object or putting two or more parts together requires perception of the relations between the different parts. Some research indicates that certain species of primates might be able to choose appropriate tools that don't need modification, but are not able to modify an object into an appropriate tool (e.g. Call, 2000). Tool manufacture can be of two basic kinds, either removing parts of an object to create a tool or assembling two or more parts to create a tool. A general tendency seems to be that most primates are good at destroying things, but not that good at creating new things. In other words, tool manufacture by subtractive methods is relatively prevalent, while additive manufacture (combining objects) is rarer (Josep Call, personal communication).

Cracking nuts with hammer and anvil is a much more complicated procedure, requiring the relations between as much as four objects to be 'just so'. The nut is placed on an anvil, often a larger rock, and is pounded with another rock. Sometimes the anvil is supported by another piece of rock or wood to make it more steady and level. This means that for successful assembly and use of this tool set up, allocentric coordinates are essential. The different parts of the arrangement could be regarded as individual tools, but I would argue that it is the combination of objects that constitutes what I call a compound tool. It could be argued that this is a case of additive tool manufacture, but I would contend that it is not. Adding a handle to the hammer by attaching a stick to a rock would be additive manufacture, or joining two short sticks to make a long stick. Using objects together as in the hammer and anvil case is similar, but not the same.

This seems to indicate that there are four aspects of tool use in primates that could vary between species and that could be indicative of allocentric capacities: tool selection, subtractive tool manufacture, additive tool manufacture and compound tool use.

2.6.2 Social and communicative skills

Many species of animals have some means of communicating with conspecifics: through sound, smell, visual displays and other behavioural patterns. Much of this communication seem to have a very strong innate component, like in the case of glands that excrete communicative smells. Such communication might require the receiver to learn what different smells mean, but the arsenal of available messages is predetermined by the glandular system of that species. Many vocalizing behaviours also seem to be innate, as are visual displays for aggression, submission or presenting for procreation. In the case of primates and gesture communic-

ation on the other hand, there is a much weaker innate component, as such communication varies considerably between different groups of the same species.

Gestural communication has been widely studied among chimpanzees, and different populations can use the same gestures to communicate different things, different gestures to communicate the same things or any combination. Different populations also vary as to how much they use gestural communication, and for what purposes (Blake, 2004). This means that though there is likely to be some innate propensity for using communicative gestures, the gestures and their meaning are learned and culturally transferred. Of course gestural communication is closely tied to visual perception, but most gestures in themselves might not need any allocentric processing to be perceived, as it is not necessary for the animal to consciously identify the gesture, just react to it as a signal stimulus.

There is evidence that great apes adapt their gestures in such a way that the conspecific with which they want to communicate is likely to see the gesture (Povinelli et al., 2003). This includes making sure to be in the line of sight of the conspecific as well as paying attention to whether or not that individual is looking. There are two important aspects for the investigation here: Firstly, the primate must be able to know what the conspecific *can* see from where that individual is situated – taking that individual's perspective. Secondly, it is a question of being able to see and understand whether that individual is paying attention. These aspects have been used in experimental studies trying to ascertain whether the chimpanzee has a theory of mind or not (Call & Tomasello, 2008), raising the question about how vision and allocentric processing might be important for such a capacity. There is much research indicating that visual information about bodily postures and movements affect how humans understand and appraise others in very profound ways (e.g. Oberzaucher & Grammer, in press). Thus our visual systems clearly play an important part in our social interaction, and it might in fact determine some of the foundations for such interaction.

A kind of gesture that is of particular importance in connection with allocentric perception is pointing. This kind of gesture concerns a relation between the sender and an object, and the meaning of the gesture is inherent in that relation, i.e. it is not specified through convention. Pointing and understanding of pointing gestures have been observed in chimpanzees (Call, 2009, personal communication). Arguably, this kind of gesture is more relying on allocentric perception of the perceiver than any other kind of gesture. Similar to pointing is gaze following, something that several primate species are clearly capable of (Tomasello et al., 1998) While a pointing gesture is arguably an intentional communicative behaviour, looking at something could potentially give the same information to the perceiver, albeit unintentionally.⁸ Gaze following is very similar to what I described earlier about perspective taking and knowing when a conspecific is looking. But the purposes are different, in the former case it is necessary to know that the conspecific can perceive my communication, in the latter case it is used to understand something of communicative value from the conspecific. However for the

8 This is in itself possibly communicating something different. Apart from communication the focus of attention, the pointing gesture also communicates the sender's intention for the receiver to become aware of whatever he's pointing to, while looking does not communicate any such information. While such information can be perceived and be useful to humans, it is unclear whether this is so for other primates.

purposes of my model I will treat both as the same skill, as it is likely that the visual processing for these two skills is the same as it involves processing of identical kinds of visual information, i.e. head and eye positions of the conspecific.

While pointing and looking are clearly different from the sender's perspective, at first glance it might look like it could be the same skill that is needed to perceive such signals. However there are two differences that make it likely that they are actually two different skills. Pointing is quite a clear signal that is highly visible. Even without paying much attention to it, such a gesture would still be noticeable. On the other hand the direction of someone's gaze is not very pronounced, and requires more selective attention to be noticed. This could point towards gaze following being a more difficult task than understanding a pointing gesture. On the other hand, while pointing is a gesture that is quite rare among primates, and virtually non-existent among other mammals, looking at objects of interest is not. This information is readily available from most animals that have vision, as the direction of the head and eyes will always coincide with the direction of whatever that animal is looking at, and this means that there is a much greater chance of a skill developing for reading such signals from both an evolutionary perspective as well as from an individual learning perspective. Understanding of pointing gestures and gaze following will therefore be treated as two different skills in my model. Understanding of what another animal can see is the third skill in this category. It should also be noted that it need not be a conspecific that the subject can read this signals from, it could also be a member of another species, for instance a human as would be necessary in an experimental setting.

2.6.3 Causal understanding

It has been shown in experiments that some primates might be capable of understanding physical causality (e.g. Hanus & Call, 2008; Seed & Call, 2009). There are numerous experiments testing primates in different problem solving tasks that require the animal to use untrained causal clues to deduct the location of hidden food, or to reach food items from a distance.

To understand that one event causes another event you need to understand the relation between those two events. In the case of physical causality this is basically that one object causes an effect onto another object (billiard ball hitting another billiard ball causing the latter to move), and for this you need full allocentric processing of the individual objects and their relation. When causing things in the world yourself it is sufficient with an egocentric perspective, for instance when pushing something causing it to move. However, if the effect that you are after is caused indirectly by your actions, this would call for an allocentric perspective. For instance the effect of shaking a branch is that the branch moves, but doing that to cause fruit to fall off requires a perception of the relation between the branch and the fruit, hence an allocentric perspective.

I would strongly suggest that causal understanding is highly unlikely to occur without allocentric visual processing, that understanding the causality of the world begins in the visual system. Immanuel Kant proposed an idea similar to that, when he in his *Critik der reinen Vernunft* (1781) claimed that accounts of time, space and also causality are aspects that our cog-

nitive systems necessarily shapes our perception of the world in. To explain the primates conception of causal events “...we need more than Humean causality, we might need at least some sprinkles of Kantian causality” (Josep Call, in conversation). Kant's most basic idea was that because the perceptual and cognitive systems shape the way we perceive the world we cannot really know what the world is like in itself (beyond our perception of it) and that there might not really be any such thing as causality in the world itself. But the idea that perception of time, space and causality is entirely shaped by the perceptual and cognitive systems is a valid point, whether or not you want to make the claim that this means that we cannot really know anything about the world-in-itself or not. It is likely that the different cognitive and perceptual systems of different species make these species experience time, space and causality in different ways. The main point here is that to have an account of causality, the animal needs allocentric perception. I will discuss this further in the last part of this essay.

2.6.4 Sorting objects on the basis of relative size

Experimental studies have shown that chimpanzees can be taught to sort objects of different size in order relative to their sizes (Call, 2000), which means that they are able to compare objects relative to each other. To the best of my knowledge, sorting of objects on basis of other relations than size have not been tested. It would be interesting to see if apes that can sort objects in order based on relative size could also sort objects based on other visual properties or non-visual properties, like for instance weight. Sorting objects in order from smallest to largest might be a cognitively more challenging task than for instance sorting objects into groups according to similarity. It might be the case that certain animals can do the latter but not the former. If so, this category should be divided into two.

2.6.5 Ability to recognize relational similarity

In an experiment Daniel Haun and Josep Call (2008) tested four primate species on their ability to infer relation between two sets of objects based on their relative locations in their respective groups. For instance the leftmost object in the first group is related to the leftmost object in the second groups etc. To solve this task the animal must be able to perceive relations between relations, firstly relating the objects within one group to each other and then relating the objects of the two groups to each other. Two of the species were successful in the task.

2.6.6 Other skills and behaviours

There are bound to be other kinds of skills that rely on allocentric perception that I will not consider. A very straight forward skill of this type would be the ability to identify different objects by means of vision. While it is assumed that all or most primates do this, and most other mammals don't, there could be differences in how well different species can discrimin-

ate different types of objects or between different objects with similar appearances, for instance objects consisting of the same parts but with different relations between the parts. I have not yet had the time to look into any eventual research in this direction. There could also be behaviours related to orienting in an environment or other problem solving situations that could be illuminating for the purpose of my study. The skills that I've settled on investigating offer experimental data that is useful for this thesis, and are easily identifiable as relating to allocentric perception.

2.6.7 Skill grouping overview

The skills that have been presented can be organized into three categories. If some species have positive results in one category this is an indication that allocentric processing could be domain specific, or there could show other significant patterns of skill distribution.

Tool use	Social	Reasoning
Selection	Pointing	Causal understanding
Subtractive manufacture	Gaze following	Relational sorting
Additive manufacture	Perspective understanding	Recognizing similarity
Combinatory tool use		

Table 3: Skill grouping chart.

2.7 Behavioural data

All the skills that I have described are confirmed through experimental research to be at least partly displayed by at least one primate species other than humans. For most of them, however, there is very little data available. Also, there seems to be a bias towards positive results. There are very few studies that have been published that present evidence of primate species lacking this ability or that, and that is of course to be expected. Much research is aimed at getting positive results and proving what different species *can* do, studying what animals *can't* do is just not as exciting. For my purposes though, all results are results, and since negative results are quite rare they would stand out as particularly interesting. The problem with negative results is that you never can be sure whether they actually mean anything. If an animal has been able to perform an action, that means that it is possible. If it does not perform the action, that doesn't really mean anything more than just that. Designing experiments aimed at getting reliable negative results is indeed a very hard task. Comparative studies are complicated since experimental set ups does not necessarily translate well between different species. And using different set ups for different species to accommodate for this will naturally be even more complicated from a comparative standpoint. This problem is unavoidable, and the only way to deal with it is to always be very careful when interpreting results.

For reasons of space I cannot review the methods used in the experiments from which I have collected the data that I will present, but is of course readily available in the literature.

This is only intended to be a brief overview of some of what has been established about the distribution of these skills in the different primate groups. I do not claim that it is in any way a fair representation of the state of the research in this field, it is only a presentation of some data points to give an example of how this plotting is intended to work. My main effort in writing this thesis has gone in to the theoretical side in reviewing the two visual streams research and constructing the research platform. There is obviously a lot of very relevant research that I have not yet had the time to familiarize myself with.

Humans are just included for reference. A similar charting could also be constructed for use in a human developmental perspective, which would indeed be an interesting complementary study.

The fundamental idea and design for this research model is owes a great debt to the work of Josep Call, especially Call (2000).

2.7.1 Legend

In the following charts these codes apply (adapted from Call, 2000):

- ++ convincing evidence of skill
- + some evidence of skill
- +/- conflicting evidence
- negative evidence of skill

If the field is empty, it means that there is no data available. Note that a single plus sign does not denote that there is evidence of “weaker” skill than two plus signs, just that there is less evidence confirming the existence of the skill.

2.7.2 Tool use category

	Selection	Subtractive manufacture	Additive manufacture	Combinatory tool use
Humans	++	++	++	++
Chimpanzees	++	++	+/-	++
Gorillas				
Orang-utans	++	++		
Gibbons	+			
OW Monkeys				
NW Monkeys	++	+	-	

Data from Byrne (2004), Call (2000), Cunningham et al. (2009), Nakamichi (2004).

2.7.3 Social category

	Pointing	Gaze following	Perspective understanding
Humans	++	++	++
Chimpanzees		++	++
Gorillas			
Orang-utans			
Gibbons			
OW Monkeys		++	
NW Monkeys			

Data from Call & Tomasello (2008), Tomasello et al. (1998),

2.7.4 Reasoning category

	Causal understanding	Relational sorting	Recognizing similarity
Humans	++	++	++
Chimpanzees	++	++	++
Gorillas	++		-
Orang-utans	++		-
Gibbons			
OW Monkeys			
NW Monkeys			

Data from Call (2000), Hanus & Call (2008), Haun & Call (2009).

2.8 On categories and modules

The three categories I have constructed here could be collapsed into two to fit with the hypothesis about the evolution of the human brain that identifies three basic modules of human (and to a certain extent primate) cognition: folk psychology, folk biology and folk physics. What I have called the social category has a clear fit with the folk psychology module, and my other two categories could be linked to folk physics. I would hypothesize that allocentric vision has been a necessary prerequisite for the cognition described by all three modules, but there is not any readily available observational or experimental data that illuminates the role of vision for folk biology.

While it could have been desirable to make my model streamlined with this highly interesting modules approach, and collapsing my tool and reasoning categories into a physics category, I have chosen not to do so. This decision is driven by the literature and data presented. There is a strong tradition of studying tool use as a specific behavioural domain in evolution-

ary studies and primate ethology, and there good reason to adhere to this tradition to make this model fit better with the data. Also, it is not immediately clear that all the skills I've presented would fit into the “folk” modules (for example the skill of recognizing similarity might be linked to folk physics, but it seems to be a bit of a stretch) so easily. This could be a shortcoming of my model, showing that perhaps some other categorization could be more productive, but at the time being I think my model shows coherence with regards to available data, and I will leave it at that.

PART 3

Vision for cognition

3.1 Introduction

This last part of my thesis is written in a slightly philosophical and highly speculative manner. I do not intend for all the arguments and ideas that I put forward here to be regarded as finalized scientific claims; I will present no or very little empirical substantiation to back up any of them because I have no such data (yet!). Instead this part is intended to give an idea on a more general plane what it is that I am aiming for, and fuel discussion and hopefully generate new ideas and research. I also intend to discuss and clarify some of the ideas that is put forward in the previous parts.

In this part I will argue that more advanced vision has both been an important prerequisite for the development of more advanced cognition in primates, and also a driving force in said evolution both for general cognition and certain specific abilities. This thesis is mainly focused on the first, more general way that vision, and especially allocentric vision, has been important for evolution of cognition by giving primates richer representational capabilities. However, I will sketch a few different specialized areas where it is feasible that allocentric vision has contributed in a more direct way to specific cognitive abilities. I might be wrong about some, most or all details, but I think that the general idea is strengthened by such examples even if they turn out to be wrong, because at least it is capable of generating falsifiable hypotheses.

3.2 Solving the unsolvable problem

How shape constancy is achieved by the visual system, the capacity to recognize the same shape as the same under a wide range of conditions, is a problem of nearly unimaginable complexity on which psychologists of perception are still hard at work. (Millikan, 2009, p. 176)

I have discussed the complexity of the problem of vision-for-perception in section 1.4.2. The visual system of humans obviously solves this problem, and I believe that there are two principle features that makes it possible: processing power and specialization. I will discuss only the first of these, as any description of the specialized modules of the humans visual system would be way to lengthy for this thesis.

The human brain is very large compared to other animals both in absolute measures and relative measures. Larger animals normally have larger brains, but this does not in itself make them “smarter”. The larger brain is a bi-product of having a larger body to control and receive sensory input from. What does this tell us in the case of vision? The tactile senses scales up

with the size of the body, but is this also true when it comes to vision, hearing, smell and taste? This does not seem to be the case from a phenomenological standpoint. To be sure, having a larger tongue with more gustatory sensors means that there is more information to be processed, but not in the same way that a larger body means more tactile sensory input and motor output. The function of the gustatory sensory system is to detect different tastes, and in a way it is not that important whether the tongue is large or small for this purpose, as long as it has enough sensors to reliably detect the different tastes that it needs to differentiate between. That our tongue is larger than that of rat nevertheless is probably just a consequence of us having a larger body, not because we have or need a more developed sense of taste. Just because an animal has a larger body it does not mean that this animal needs more gustatory information. The same holds for smell and hearing as well, basically.

Vision is special in this regard. Being a large animal means that you need to see large things, because these are the things that you generally interact with. This means that your field of vision needs to be bigger, which means larger eyes, more photoreceptors and more visual information to process. But it does not scale in the way that the tactile senses scale, as the surface area of a body grows exponentially with the enlargement of the body, but the visual field does not. And the amount of different smells and sounds that the animal needs to smell and hear do not scale at all, at least not as a direct factor of body size. This means that while we expect the brain to have a certain size relation to the body that it resides in, the relation between areas processing visual information and the general size of the brain is expected to be inverted; that is, when the brain is enlarged due to increased body size, the relative area of the brain that processes visual information is expected to get smaller. The same relation should also hold for hearing, smell and taste.

This is not the case with the human and primate brain. The amount of brain tissue that is involved in processing visual information is larger than expected compared to other mammals (Kaas, 2005), even in relation to the massively enlarged brain that humans have ($EQ \approx 6^9$). This is a clear indication of the idea that the human (and to a large degree primate) visual system can do things that other mammal's visual systems cannot. Philosophers, neurologists, engineers and cognitive scientists are in agreement that vision-for-perception (as defined by Goodale and Milner (e.g. 1992)) is an almost impossibly complex problem to solve (e.g. Pinker, 1997; Jeannerod & Jacob, 2005; Millikan, 2009). I am fairly certain that the vast increase in processing power that the enlarged human and primate brains affords, is essential in the evolutionary solution to this problem, and this is probably the best indication that I can present to defend the (perhaps controversial) idea that most animals don't use vision in this way at all. Their small brains just can't deliver that processing power. The following is a summary of the main points presented in this thesis that support the idea that only primates have vision-for-perception, that is, conscious visual representations of the world.

- Primates have a fovea with densely packed photoreceptors, other animals have not. This gives primates central vision with much more detail than what is pos-

9 EQ = Encephalization Quotient, a relative measure of brain size, where the size of the body is taken into account. An EQ of 6 means that the human brain is six times larger, in relative terms, than the average brain of a mammal (Geary, 2009).

sible without a fovea.¹⁰

- Humans and our closest primate relatives have trichromatic colour vision, most other animals have not; many do not have any colour vision at all. Being able to process the different wavelengths of light in this way yields more information about that which the light is reflected off of, which is likely to be of great help in identifying objects quickly and correctly.
- Humans and primates have a much larger portion of the brain devoted to processing visual information than other animals, both in absolute and relative terms.
- The primate and especially human brain have a much more extensive differentiation of neurologically distinct areas involved in visual processing compared to other animals. In the human brain there are about forty distinctly different areas, whereas many other animals only have a few. The problem of achieving vision-for-perception is tremendously hard, and its solution probably demands this great differentiation together with the larger brain affording more processing power.
- Behavioural evidence suggest that many other mammals rely much more on olfaction for identification purposes, not least in identifying conspecifics and food. (Humans don't smell each others butts to learn about each other.) Vision is thought to aid in locomotive behaviour and eliciting conditioned or genetically programmed responses, without much room for adaptive modification of behaviour and no higher cognitive processes dealing with that visual information.

3.3 Vision versus olfaction

There are two fundamental ways in which I believe that the switch from olfaction to vision as primary sense for world interaction in primates have influenced and been an important factor in the evolution of primate cognition. The first is a more general way, in which more advanced vision can support more detailed representations of the world than any other sense can provide, which allows for more specialized and diverse behaviour, which in turn demands more advanced cognition. The second way is that allocentric vision allows for “allocentric cognition”, specialized ways of representing and processing information about the world.

To avoid any unnecessary misunderstanding, it is worth pointing out that it is not necessary for an individual to actually be able to see for it to have the cognitive abilities that I consider to be evolutionary reliant on vision. Blind people are expected to have exactly the same cognitive abilities as seeing people, as long as their blindness is not due to deficits in critical brain regions. In other word, my claim is only that vision has been important in the evolutionary

¹⁰ It could also be noted that the design of the vertebrate eye is vastly superior to other designs common among for instance insects. It has been calculated that if humans had compound eyes like flies, we would need eyes with a diameter of 27 metres to see as accurately as the vertebrate design allows us to do (Dawkins, 1996).

history of cognition, not in the developmental history of singular individuals. It is for example likely that our ability to mentally represent space and spatial qualities is closely tied to the visual cortex in the brain, and it would be the same brain region that is used by both blind and seeing persons. It has evolved through and for processing mainly visual input, but once it is there it can be used without having access to any visual input at all.

The ultimate goal for the sensory systems of an organism is to create representations of that organism's environment, in order to adapt behaviour in a way that maximizes reproductive success. By representing its environment, an agent can adapt its behaviour in a purposeful way and not rely on trial and error behaviour which is likely to sooner and later get it killed. You can only walk off a cliff once, so once you've tried that you don't get the chance to try a different approach. If you on the other represent the cliff you can choose the appropriate behaviour to steer clear of such detrimental actions. Such representations can be conscious, as in the case of visual representations originating the ventral stream¹¹, or unconscious, like the visuomotor transformations of the dorsal stream.

Representations of the world can originate from any sensory system of an organism, and can be of varying complexity. There are simple anaerobic bacteria called magnetosomes that through a magnetic sensory system represent the direction of less oxygen in their environment, which is all the information the bacteria needs to survive (Dretske, 1986). Their sensory systems consist of a single molecule that orients itself towards magnetic north, thereby representing that crucial piece of information. Humans on the other hand have access to a plethora of incredibly complex representations of very varying kinds that often can supply us with the information that we need to survive, and very much more. Other animals have other representational capabilities, that have been shaped by evolution according to their needs.

One of the main points of this thesis is that primates seem to have evolved a unique representational capacity that allows us to construct a conscious visual representation of the world. This representation have some unique properties that include identification of objects, spatial structure and preserved relations of content. I argue that this representational capacity is a cornerstone in the evolution of more advanced cognition, as well as certain specific cognitive abilities that are closely tied to vision and allocentric processing.

Most other mammals use olfaction as their primary sensory system, primarily meaning that they use smell to identify objects in the world. This does not mean that these animals do not have any conscious visual perception of the world at, but it is very possible that this is in fact the case. It would certainly be a completely different kind of visual perception that does not include identification and recognition of objects. The absence of allocentric visual processing has the consequence that all visual stimuli are only processed in relation to the agent, and probably only used in visuomotor transformation that is unconscious.

Of course, many animals don't need to identify objects in the world at all. Frogs use vision to catch prey, but not through identifying objects in the world as prey, just by reacting to black dots of the right size moving across their retinas in the right way. As such phenomena ordinarily happens to coincide with flies and other things that frogs eat being present, such a simple

¹¹ N.b. that not all representations originating the ventral stream have to be conscious, just that they *could* be, as opposed to those originating in the dorsal stream.

behavioural mechanism is sufficient for a frog to survive in most settings (barring in a laboratory with devious scientists projecting black dots onto frogs retinas). Thus the frog has no need for a more sophisticated visual system, so it has never evolved.

Perception through vision and through olfaction are very different processes. Olfaction works through a number of different chemical receptors that activate when they encounter the type of molecule that they are built to detect. Molecular structure does not vary with lighting conditions, distance or any of the conditions that affect the retinal projections from an object, it is fixed and will activate the system exactly the same way every time (as long as the system is intact). Identification through this kind of chemical reception is therefore (comparatively) simple and reliable. But it is also inflexible. The olfactory system will only react to the kind of molecules that it has receptors for. Suppose an organism with olfaction as its primary perceptual channel cannot identify the materials uranium and ytterbium (not unlikely seeing that they are virtually non-existent in nature and so has not had any influence over the evolution of the organisms perceptual system) as it does not have any chemical receptors for such molecules. It would be impossible for this organism to make any distinction between the two, or even to re-identify the a lump of uranium as the same lump of uranium on two occasions. These substances would literally be blanks in the organisms perceptual world, the same way that we can't through any of our senses perceive ultra-violet light, magnetic force or dark matter. We know (or suspect) that these things exist because we can see their effects or have mathematical proof that they should exist, but we will never be able to have a perceptual experience of them. It is reasonable to suspect that the olfactory system of any organism has chemical receptors for molecules that have been around during its evolution, and in particular for molecules that have carried information that has been of value to the organisms reproductive success. The visual system has an advantage in this respect, because it is able to register the light reflected of any surface whether or not that kind of surface has figured in the evolutionary history of the visual system. I would not be able to identify a lump of uranium if I found one, but I could probably learn. At the very least I would be able to perceive such a lump as a lump of something, and I would be able to re-identify it and to recognize it as being different from other things that I have not seen before.

3.3.1 What is a primary sense?

This might be a good place to be more specific about what I mean by the term “primary sense”. It is not a widely used concept, although it appears from time to time in the literature. Most organisms have some way of getting information about their surroundings, as well as their own body (like balance, proprioception or hunger). The organism can have any amount of sensors, from only one to a plethora, functioning by the same principle (that is, reacting to the same kind of stimuli), and such a perceptual system is in common english called “a sense”. Most, if not all, vertebrae have more than one sensory system reacting to external stimuli, and these systems are more often than not used for different purposes, or at least in different contexts. Most sighted animals use vision as the primary source of information to

guide locomotive behaviour, while using olfaction for identifying objects in the world, such as food, conspecifics etc. Humans and other primates use vision for both purposes. This is not to say that we can't use other senses for the same purposes. We also use touch to aid locomotion, especially when it is dark, and some blind people allegedly can use their hearing to get spatial information that can aid their moving around. We also use our sense of smell to identify food, especially to assess whether or not something has gone bad. However in most cases we use vision to identify objects. In order to identify an object as my bed, I can't smell, taste or listen to it. I might touch it and through a process of touching it in different places come to the conclusion that it is a bed, perhaps mine, but through vision I can see that it is in the blink of an eye. Considering the ease with which we humans, and allegedly other primates, can use vision to identify objects in the world it is no wonder that we are so reliant on it for so many of our sensory needs.

What I mean by primary sense or primary sensory system is the sensory system through which the organism identify objects in the world. There must of course not be any such thing as a primary sense for a certain organism, it is completely conceivable that there are organisms that use two (or more) senses in an equivalent way to get such information (however see the next section on multimodal representational capabilities). It is also possible that the notion of a primary sense breaks down for all organisms except certain vertebrae, due to them not identifying objects at all, or just reacting in a genetically programmed way to predetermined stimuli without ever identifying anything. However, in the case of humans it makes perfect sense to regard vision in this way, and especially in comparison to animals that are reliant on olfaction for object identification purposes.

3.3.2 Multimodal representational capabilities

In humans, and possibly other primates too, there is a large degree of integration of the different senses. We can experience the same thing or phenomenon in the world through more than one sense modality at the same time, and thus construct multimodal representations of things in the world. Not all animals have this ability, as described by Gärdenfors (2000) an adder, for instance, uses its different sense modalities for quite different purposes at the same time without (it seems) integrating them into a multimodal representation. This is to the disadvantage of the snake in certain circumstances, but apparently not often enough for the snake to have evolved a multimodal representational ability.

Multimodal representation in humans is often centred around the visual modus. There are few circumstances where two or more senses are combined into a multimodal representation without vision being one of the senses. Provided you are not blind, how many things in the world do you know about only by a combination of sound and smell? Touch and taste? Any other such combination? I cannot think of any for myself. This should not be surprising at all, it is just a natural consequence of vision being our primary sense. I have no empirical data to justify the following claim, but it is a distinct possibility that the evolution of allocentric vision in primates was a prerequisite for humans to develop multimodal representational ability. Allocentric processing allows the visual system to relate parts of the visual field to each other,

and it is not much of a stretch to speculate that this ability might have been a stepping stone in developing the ability to relate different sense data to each other.

3.3.3 Speed, precision, distance and details

Visual perception has a quite few overarching attributes that makes it a more powerful sensory system than olfaction. I will describe a few that are suitable for philosophical pondering.

The visual system can be updated with information about the world at the speed of light, the olfactory system with the speed of wind or the speed of the animal depending on the situation. Speed might not always be crucial when perceiving the world, but it could be in certain situations, and it provides a general benefit.

The visual system provides very precise information about the location of things, whereas olfaction detects the smell of a more loosely defined location and or direction. By seeing a tree I know exactly where that tree is, both in relation to me as well as in relation to other visually salient aspects of the world. An animal relying on olfaction for this information would certainly have access to more information than a human trying to locate something by smell, but because particles move erratically compared to photons, it could never come close in precision even in theory.

Vision allows us to get information about parts of the world that are quite distant to our own position. Olfaction can provide information from some way away, but hardly from kilometres away, not to mention the impossibility of smelling the moon or the sun. With distance the quality of the information that reaches the sensory system is degraded, reflected light becomes fainter with distance as it spreads out over ever larger volumes. But compared to the diffusion of smells over distance light is tremendously accurate. I find it hard to imagine that even the most elaborate of olfactory systems can distinguish two closely positioned targets from even a hundred metres away, something that is easily achieved in vision. Vision also makes it possible to distinguish details of that which is perceived, once again because of the physical properties of light; more precisely that light travels in straight paths, whereas odours (molecules) travel in more or less stochastic ways through the air blending and diffusing even over very short distances. Different parts of a uniform object will reflect light towards a specific point (e.g. a retina) in different ways as a result of distance and angle, making it possible to calculate the distance and angle of the object. However the smell of a surface at a 90 degree angle will not differ from the smell of the same object at a 45 degree angle, making it impossible to judge the angle by olfaction, whereas this is readily available information for visual processing. Having conscious access to such information is a necessity for perceiving shapes, at least in three dimensions and probably in two as well.

3.3.4 Functionality doesn't smell

That is not true, of course. Functionality does smell in many cases, from perfumes and other man made objects to pheromones and other signalling systems in biological organisms. But

much functionality does not smell. You cannot tell the straight twig that is suitable for termite fishing from the crooked one, or the sharp edged rock suitable for cutting from the blunt one, using your nose. Using objects as tools very often requires us to see them to determine their usefulness. A plate and a cup made from the same chunk of clay will at average smell the same, but only one of them will hold my coffee in a useful way. Tool use and manufacture would be incredibly hard without vision, all but impossible through olfaction alone, and more to the point I would dare assert that it would probably never have evolved without vision at all. Again, this is not to say that blind people cannot use tools, of course they can, but this is because they are the same species that have evolved this ability intertwined with the evolution of vision.

3.3.5 Spatial relations

One of the main feats of allocentric vision is that it can process and maintain spatial relations. Olfaction can possibly process and maintain some spatial relations, but as I sketched in a previous paragraph it has some severely limiting disadvantages compared to vision. Being able to process perceptual information that way means that you should be able to think about spatial relations too. Not being able to process such information however does not directly imply that you cannot think in such terms, but it makes it rather unlikely. It is hard to see what information you would use for thinking in terms of relations if such information is not available, but even more to the point it is hard to see how and why such a thinking capacity could evolve without allocentric perceptual input. Even if one does not agree with Hobbes that there is nothing in the mind that has not first been in the senses (except for the mind itself), it is hard to imagine what our thought would contain if we had no sensory input.

From an evolutionary point of view it makes much more sense to say that the mind is there to process information from the senses, than saying that the senses are there to give the mind something to process. Thinking begins with sensory perception, what else would there be to think about? I would say that the same line of arguing holds for the evolution of thought also, the evolution of perception comes before the evolution of thought. It is therefore a good starting point to assume that perception of spatial relations is a prerequisite for thinking about spatial relations. Thinking about spatial relations is in turn important for a lot of problem solving, including judging size or distance differences or determining what fits together. Relying on olfaction to solve such problems could mean being stuck with a round peg in a square hole without ever realizing the problem. Such problems could of course be solved with the aid of tactile perception or echolocation, but I will get back to such sensory systems in a bit to discuss their drawbacks.

3.3.6 Movements and temporal relations

Appreciating movements is a special case of processing spatial relations, involving temporal relations. Basically the same line of argumentation as for perception of spatial relations is applicable, since perceiving movements in essence is tracking the change of spatial relations

over time. However, this capacity seems to have its own locus of processing the brain (Gazzaniga et al., 2009), making it worth mentioning in its own right. It might also be possible to link perception of movements to experiencing time, as conscious awareness of time probably would be radically different if you were not able to track movements. Animals other than humans probably have different time experiences, some might not have any conception of time passing at all, and it is more than likely that this is linked to the way that the animals perceive the world through their sensory systems. Experiencing movements and temporal relations through vision could therefore have shaped our time conception, laying the foundation for sequential thinking and planning. Experiencing spatial and temporal relations is arguably also the foundation for experiencing causality.

3.3.7 Vision versus the other senses

We should not be surprised that very few (if any) animals use ordinary hearing as a means of identifying objects in the world. Most objects make very little sound at all, and much that makes some sound only does this from time to time and often for specific purposes. Imagine trying to listen to a mushroom to conclude whether it is poisonous or not. Thus, hearing is not a likely candidate to ever be used as primary sensory system.

Gustation is similar to olfaction in many ways, they both react to chemical stimulation and are bound to use similar neural mechanisms. So why then is gustation rarely (or never) used as a means of identifying objects in the world? The answer is obvious of course. While olfaction can detect chemicals over a distance, gustation requires the organism to take the object into its mouth or at least lick it. This is both cumbersome and possibly dangerous compared to olfaction. Should the object prove to be poisonous (or an enemy), it is too late. By smell an animal can steer towards food or other necessities, while the animal reliant on taste would wander aimlessly tasting the world hoping that something will be edible soon.

Some of the representations of the world that we can get through vision is also possible to get by means of tactile perception, such as certain spatial relations and shapes. Using our hands it is possible to avoid putting a round peg in a square hole, to judge relative size of two objects or to determine which of two stones is the sharpest. This is however possible only when such objects are directly available, trying to find these objects in the world probably entails both very lengthy and possibly dangerous searching by literally fumbling in the dark. Tactile sensory systems utilizing whiskers should have some advantage for certain tasks, but are still limited to a very narrow receptive field, and are probably only useful for sensorymotor transformation.

In very many respects echolocation is similar to vision; it provides a three dimensional spatial array of the world with textural information about surfaces. Echolocation sensory systems can in many respects allow for allocentric processing, but I have found no literature that proclaim the existence of any such systems. It is known to be used by bats, dolphins and whales, and while dolphins are regarded as among the most cognitively advanced animals, it is unclear whether they use echolocation for anything else than basic sensorymotor transform-

ation. The fact that they also have a visual sensory system ought to make the matter of connecting any cognitive abilities to a sensory system more complicated. Echolocation has several disadvantages compared to vision as a sensory system though. The speed of sound is considerably slower than that of light and the energy of sound waves also dissipates more rapidly over longer distances, meaning that the sensory radius will be smaller. Also, since sound is a true wave form (periodical displacement of matter) the spatial resolution of sensory arrays is bound to be less fine grained compared to vision (since light propagates directly through air), meaning that it is probably much more difficult to discriminate between small details in the sensory field. A completely different disadvantage of echolocation is that it requires the organism to constantly make sounds to receive information about the world, which requires energy and also signals your position to potential predators.

Compared to echolocation, vision is “free” in the sense that it uses the energy of the sun directly as information source, however this is a disadvantage of vision compared to all other senses as none of them are dependent on lighting conditions. No organisms that I know of (except humans) provide their own light source to see during night or underground.

3.4 Specific cognitive abilities derived from vision

I will present some specific abilities that I believe could have evolved by utilizing and developing processing capabilities of the visual system, or at the very least would not have evolved the way that they have in humans and those primates that possess them were it not for vision.

3.4.1 Understanding causality

In the research model that I presented in the previous part, I summarized some experimental and ethological evidence for causal understanding in primates. Said presentation does not proclaim to show that allocentric perception is necessary for causal understanding, but here I will argue that such a case could be made.

According to David Hume (1739) we can never perceive causation, all we ever experience is one event following after another. This is in a sense trivially true, as causation is not a really a visible or perceivable phenomenon. But, as Hume also points out, we are naturally compelled to interpret the world as containing a natural phenomenon such as causality. Kant (1781) would express this idea more directly by making the distinct point that experiencing causality in the world is inescapable, because causality is projected onto the world from our cognitive apparatus (my wording). If this is indeed true, and I believe that it is, then there is no real difference between perceiving a certain projection of light onto our retinas as an object and perceiving a temporal sequence of events as causally related. Our perceptual and cognitive systems create both experiences in a compelling way. Whether there is actual causation in the world or actual objects in the world is not knowable, we only know that which our perceptual and cognitive systems construe out of the sense data that they receive¹². In this sense there

¹² However it is reasonable to believe that our perceptions actually tell us something real about the world, otherwise this information would probably not be useful for guiding successful behaviour for surviving and

is no difference between causation and other phenomena. We understand the world in terms of causal relations, and we act in accordance with this understanding, in much the same way that we act in accordance with our visual perception of objects (how else could we ever enjoy, or even stand, watching a game of football?).

Animals lacking allocentric vision might indeed perceive the world in the way that Hume described, as one event following another with no causal relation what so ever. But being able to experience visually one event following after another in a predictable way lays the foundation for a perceptual system that picks up on this regularity. And once such a perceptual system is in place it can also be utilized for representing such causal relations even when they are not happening right in front of you, and ipso facto you have a mental account of causality.

3.4.2 Problem solving by visualizing

Many cognitive problems can be solved by visualizing the problem and solution. Problem solving by visualizing involves creating a compound image like representation with relations between different parts. Examples of such visualizations of cognitive tasks include cognitive maps, mental rotation and comparing visual memories of objects. These are all classic examples of cognitive tasks that we solve by utilizing visual thinking, and are often put forward as arguments for the thesis that we think with the aid of images. Thinking in images, which is hard to deny that we seem to be doing, is clearly related to vision and in these cases of solving certain problems it involves representing, maintaining and processing the relations within and between these images. In this case I believe that it is patently clear that allocentric processing is necessary, and that the evolution of such skills has sprung out of allocentric vision could hardly be considered a controversial idea.

3.4.3 Spatial thinking and arithmetics

Spatial extension and counting goes together, we generally think about numbers as a two dimensional extension and we plot things in space. You could of course count a number of auditory signals, or a number of tactile stimulations, but it is hard to think of any naturally occurring situations where this could have been useful for any organism. Counting occurrences of a phenomenon in time is just not as much use as being able to count spatially occurring phenomena¹³, such as the number of rivals in a certain direction or fruits on a tree, etc. Auditory and tactile sensory systems allow the organism to count chronological occurrences of phenomena, while vision allows it to count spatially occurring phenomena. I don't even know how any counting could be achieved through olfaction or gustation, counting taste or smell seems similar to counting water or air – it just isn't countable.

Counting things in the visual field is in a way relating these things to each other, at least to

propagating genes. I.e. the fact that our perceptual and cognitive systems have evolved through natural selection makes it likely that they convey real information about the world, at least in some respects.

13 Being able to count measures of time is a wholly different thing than being able to count sequential phenomena, and is conceivably useful to many organisms, however I do not know if any animals can do it.

avoid counting the same thing more than once. In an ordered universe where everything was placed in neat lines or other patterns, this would not be a problem, but in the less ordered world where things are more scattered, being able to relate different visual objects to each other seems to be at the very least helpful when counting. A related ability is to be able to assess which of two amounts is larger, which maybe could be considered a kind of proto-counting ability. In this situation there is no doubt that being able to relate two parts of the visual field to each other is necessary.

Humans can of course do a lot more mathematical operations than relating amounts and counting objects (what different primates can do in these fields is a little less well known). But I find it entirely possible that both arithmetics as well as geometrics as mental operations have a foundation in the visual system. We tend to rely on visual representations both when learning to count (counting on our fingers or using abaci) but also later in life for both everyday calculations as well as more advanced mathematics (graphs, geometrical figures, etc).

Just to speculate a little about why mathematical abilities has arisen among primates, it could very well be the case that since most primates live in socially organized groups and eat a lot of things like fruit and nuts that are collected, it could maybe be a matter of distributing foodstuffs among group members according to hierarchical other social structures. Many other mammals eat leaves or grass that is not easily countable, or share a single prey.

3.4.4 Sequential thinking and planning

True planning behaviour is almost unknown among animals, except for humans that sometimes seem to do nothing but plan ahead (e.g. Gärdenfors, 2000). Recently there have been reports of planning among chimpanzees (Osvath, 2009), and evidence for planning behaviour among other apes is not unlikely to be presented within a few years. Planning requires thinking in steps: first this, then that, and sometimes if... then. Sequential thinking of this kind can be argued to demand an understanding of temporal relations, which I have already discussed. It is possibly also related to causal understanding.

Thinking in sequences requires us to be able to relate two or more instances of thought to each other, a kind of allocentric cognition or processing. There could of course be allocentric processing without allocentric perception, but the question arises what use allocentric processing would be without that kind of sensory input. There could be some use to be sure (anything could be useful for something), but could it provide a benefit that is large enough for it to have evolved in any animal? And if so, would it provide enough benefit to be sustained through further evolution, considering that it demands an expensive¹⁴ neural system to be achieved. If a part of an organism has a higher cost to develop and maintain than the benefits it bestows to that organism, evolution will favour those that do not have that part resulting in its eventual elimination in the highly unlikely event that it is ever developed.

¹⁴ Expensive in three senses: Firstly that it costs the organism to develop the system during ontogeny, secondly that it is expensive in the sense that it demands energy to be sustained and used, and thirdly that it is a complex system that has a high risk of failure due to injury or disease.

3.4.5 Social anticipation and reading thoughts

Most primate species live in groups, some – like us humans – in very large groups compared to other mammals. Group dynamics and social interaction is an important part of primates lives; understanding and negotiating social situations can be fundamentally important for survival and gene propagation (e.g. de Waal, 1982; McGrew, 2004), an outcast animal will have a hard time staying alive and probably even harder time procreating. Many things in an animal's environment will behave in a highly predictable way: all inanimate objects are just that, vegetation is highly sedentary and so forth. For many species of animals conspecifics will behave in very predictable ways too, but this is not necessarily the case for primates, and certainly not for humans. Even though human behaviour can seem very predictable to us at times, it is far less predictable than the behaviour of virtually all other things that we come in contact with. From the diverse repertoire of behaviour that has been demonstrated by apes and monkeys alike, this also seems to be the case in the social contexts of our primate cousins.

Being able to predict conspecific behaviour becomes a very important part of life, to understand and negotiate social interaction. So how can you predict that which is unpredictable? Humans do it by assessing each others cognitive and emotional states. We are very adept at reading each others mental states and attitudes through facial expressions and bodily postures, and there is good reason to believe that this could also be the case for other primates. Other animals also react to such visual displays, dogs for instance are intimidated or agitated by the showing of teeth. However it is likely that allocentric processing of the subtle variations in conspecifics facial expressions and bodily postures allows for a much more flexible and diverse understanding of social situations. After all, the variations that we are able to pick up on can be very minute, and still have a profound influence on our assessment of the person or situation. Millions of people have for centuries been fascinated by the ever so slight smile of Mona Lisa by da Vinci. Dogs can be trained to react to even more subliminal variations (mostly behavioural patterns), as in the case of seizure response dogs, but I feel fairly certain that the signals that they pick up on only elicit their conditions response, not an understanding of their ward's mental state.

By scanning another person's facial expression and bodily posture we can learn a great deal about that person's emotions, attitude and disposition, giving us a much better basis for prediction of that person's subsequent behaviour; to a certain degree one could be excused for saying that it is a kind of thought reading. And it is not hard to see why a propensity for thought reading would evolve. Of course there are other ways that this could be done, and indeed is done. We get very similar information from the tone of someone's voice and other prosodic features. Similarly animals that have a keen sense of smell are likely to be able to assess other animals internal states by variations in odour of that animal's bodily secretions. My point is not that allocentric vision gives primates some unique benefit in predicting conspecific behaviour, but that it affords a new way of social prediction that just might be more flexible and capable than other systems.

3.4.6 Imitation

One ability that is thought to be very important for social interaction and learning (Bjorklund and Bering, 2003), as well as the evolution of language (Arbib, 2009) is the ability to mimic. Most primates do not have the ability to imitate, and apes only have it to a weak degree (Arbib, 2009), while humans can get pretty good at imitating quite early in life.¹⁵ Nonverbal imitation is of course heavily reliant on vision; successfully imitating nonverbal behaviour on the basis of olfaction, gustation or hearing is probably impossible, with a few exceptions. Perceiving behaviour through vision does in most cases demand allocentric perception (relating different parts of the observed body to each other and the environment, etc.) and it is therefore quite possible that imitation would not have been possible without this skill. In turn this would mean that allocentric vision was a prerequisite in the evolution of language and different kinds of social cognition. I will not elaborate this thought further here, I merely include it as a possible selling argument for future investigation.

3.5 Vision driving cognitive evolution

Human cognition allows us to do so much more than any other species. I have tried to show several different ways through which our visual sensory system has been a pivotal factor in the evolutionary shaping of our cognition. There are two more general factors that I believe are more important than anything else in driving cognitive evolution, and that is social interaction and representational capabilities. I will end this thesis by discussing how the advancement of our visual system has been important in a more overarching way for making us able to think the way that we do.

Before I do that, however, I would like to emphasize that while I believe that visual allocentric perception is a key component in human cognitive evolution, that does not mean that the other sensory systems have been unimportant. Our auditory system is likely to have been a very driving force, especially in the last million years of our evolution, and especially with regards to social interaction and representational capabilities. The development of more advanced communicative skills has perhaps been the most important factor for our overall cognitive capacities, and this development is of course closely linked with our auditory system. For the most part, this development does not seem to be shared with our primate relatives and researchers believe that the evolution of human language happened between 1,2 million years and a hundred thousand years ago. In the scope of shared evolution with other primates it therefore seems safe to insist that vision has been the most important sensory system for cognitive evolution.

¹⁵ It is likely that imitation is a very important part of human development, making this a likely candidate for a biologically defined ability. Imitation seems to be something that comes very natural to children, even to the degree that they seem find great joy in imitating others.

3.5.1 Vision in social interaction

Allocentric visual perception has the advantage over other kinds of perception that it allows the agent to take in complex social situations, involving a multitude of conspecifics, their interaction and their spatial relation to each other. Even at a glance, a fairly rich overview of a social situation is readily available. More developed allocentric perception also allows the agent to get an understanding both of what conspecifics are looking at, but also what they can see from their perspective. Who is looking at who? Who is looking at me? Where could I go so that X cannot see me? Allocentric visual perception can answer these sorts of questions for the agent, which of course would help with planning behaviour in complex social settings, and also allows for certain kinds of deception.

It has been shown that animals that display higher levels of social interaction tend to have a higher EQ compared to similar animals with less social interaction, and the fairly obvious conclusion from this has been that social interaction drives the evolution of more complex cognition and therefore larger brains. I see no need to challenge this assumption as it seems both parsimonious in regards to the data, as well as plausible from a common sense view. But I also find it unlikely that this is a one-way street, but rather that more developed cognition allows for more social interaction as well, even when said cognition has not evolved to solve problems of social interaction. This is of course a pretty obvious statement to most researchers in this field, but I believe that it serves as a good parallel to what I believe is the case with the development of human visual perception.

At any rate, the switch from olfaction to vision as primary sense for recognizing conspecifics does allow for more complex interactive behaviour in social settings to evolve. It is a possibility that this switch happened due to increasing need for better information about social interaction, but I find it hard to believe that this could be the only factor, although it is very possible that it played an important part.

3.5.2 Representations and thought

The switch from olfaction to vision as the primary sensory system in primates made possible a much more flexible way of creating representations of the world. Representations are the cornerstones of concepts, and concepts form the basis for thinking. All concepts have their basis in perception in some way or another. This idea is heralded by Jesse Prinz (2004) and I agree. Concepts are our way of organizing our perceptions and thoughts of the world. A being without any perceptual systems would have no need for concepts. This is not the place to argue for this account, but those that do not agree on this premise will probably still agree that many of our concepts have a perceptual basis. For Prinz it is of importance that concepts cannot be innate, but must be learned by the cognitive system that uses them. This is where I and Prinz part ways in our views. Even if all concepts must have a perceptual basis, I see no reason to postulate that they must have a perceptual basis in the ontogeny of the organism that uses them. As long as they have a perceptual basis in the phylogeny of that organism, that is enough for me to satisfy the criteria of being perceptually based. Our tendency to acquire con-

cepts in a certain way is reliant on our perceptual systems, and in particular our visual system, and our perceptual systems have been constructed through the process of evolution to accommodate creating representations and concepts in a way that has been historically useful for us as organisms.

I would perhaps dare go so far as to say that the ability to use concepts is the same ability as being able to think. Using concepts can of course be a lot of different things. It could be argued that simply having a concept is to be able to think of that concept. That is something that I think that many animals with sensory systems can do. However, some animals might be able to manipulate concepts in different ways which is a more developed form of thinking. Concepts are of course essential for language.

The development of the representational powers of the primate visual system has allowed us to form and use more powerful and flexible concepts. But it has also laid the foundation for more powerful and flexible manipulation of our concepts. Visual perception allows us to understand the world in a way that is just not possible through our other senses.

Visual perception and cognition have most certainly evolved together in primates, the one would not be useful without the other. As the senses provide the information that is needed for cognition, the form of the sensory information is crucial to what cognition is possible. I have already stated that I don't believe to be possible to draw a definitive line between vision and cognition, it is an integrated system and in a way thinking starts in the eyes, at least from an evolutionary perspective.

3.6 Conclusion

The research into the two streams of the human visual system has shown good evidence for a functional division, but it is far from resolved exactly what is the best way to describe that division. However, it seems clear that the dorsal stream handles most or all visuomotor transformation in an egocentric frame of reference, and that the ventral stream is more involved in conscious perception with the aid of an allocentric frame of reference, and that the latter is a neurological development unique to primates. This also suggests that conscious visual experience might be an uncommon thing in the biological realm, perhaps even unique to primates.

It seems to be proven beyond doubt that many primates other than humans are able to process visual information in an allocentric frame of reference, though behavioural data is scarce for many species. The model presented in this thesis suggests that the evidence can be divided into three categories that encompass different kinds of behaviour, namely tool use, social interaction and reasoning, and propose guidelines for future research of the allocentric capabilities of different primate species.

The evolutionary switch from olfaction to vision as a primary sensory system in primates has served to enrich our inner worlds with a more organized representational capability, as allocentric vision preserves relational information about the world that is lost through other senses. This is likely to have been the prerequisite for the evolution of more advanced cognition in general as well as more specialized cognitive abilities, such as spatial thinking, planning behaviours and many skills for social interaction.

References

- Arbib, M. A. (2009) Invention and community in the emergence of language: Insights from new sign languages. In: *Foundations in Evolutionary Cognitive Neuroscience*, (Platek, S. M., & Shackelford, T. K. eds.) Cambridge: Cambridge University Press.
- Bates, L. A., Poole, J. H., & Byrne, R. W. (2008) Elephant cognition. *Current Biology*, 18, 544-546.
- Begun, D. R. (2004) Enhanced cognitive capacity as a contingent fact of hominid phylogeny. In: *The Evolution of Thought: Evolutionary Origins of Great Ape Intelligence*, (Russon, A. E., & Begun, D. R. eds.) Cambridge: Cambridge University Press.
- Bjorklund, D. F., Bering, J. M. (2003) A note on the development of deferred imitation in enculturated juvenile chimpanzees (*Pan troglodytes*). *Developmental Review*, 23, 389-412.
- Blake, J. (2004) Gestural communication in the great apes. In: *The Evolution of Thought: Evolutionary Origins of Great Ape Intelligence* (Russon, A., & Begun, D. eds.) Cambridge: Cambridge University Press.
- Byrne, R. W. (2004) The manual skills and cognition that lie behind hominid tool use. In: *The Evolution of Thought: Evolutionary Origins of Great Ape Intelligence* (Russon, A., & Begun, D. eds.) Cambridge: Cambridge University Press.
- Call, J. (2000) Representing space and objects in monkeys and apes. *Cognitive Science*, 24, 397-422.
- Call, J. (2004) Is there only one way to become sapiens? (2004) *Trends in Cognitive Sciences*, 8, 247-249.
- Call, J. & Tomasello, M. (2008) Does the chimpanzee have a theory of mind? 30 years later. *Trends in Cognitive Sciences*, 12, 187-192.
- Clayton, N., & Emery, N. (2005) Corvid cognition. *Current Biology*, 15, 80-81.
- Creem, S. H., & Proffitt, D. R. (2001) Defining the cortical visual systems: "What", "Where", and "How". *Acta Psychologica*, 107, 43-68.
- Cunningham, C. L., Anderson, J. R., & Mootnick, A. R. (2006) Object manipulation to obtain a food reward in hoolock gibbons, *Bunopithecus hoolock*. *Animal Behaviour*, 71, 621-629.
- Davies, R. (1970) *Apeman. Lola versus Powerman and the Moneygoround, Part One*. London: RCA Records.
- Dawkins, R. (1996) *Climbing Mount Improbable*. London: Penguin Books.
- Dawkins, R. (2004) *The Ancestors Tale: A Pilgrimage to the Dawn of Life*. London: Weidenfeld & Nicolson.
- Dretske, F. (1986) Misrepresentation. In: *Belief* (Bogdan, R., ed.). Oxford: Oxford University Press.
- Ellard, C. G. (1998) Comparative perspectives on multiple cortical visual systems. *Neuroscience & Behavioral Reviews*, 22, 173-180.
- Gazzaniga, M., Ivry, R., & Mangun, G. (2009) *Cognitive Neuroscience: The Biology of the Mind – Third Edition*. New York: W. W. Norton & Company, Inc.
- Geary, D. C. (2009) The evolution of general fluid intelligence. In: *Foundations in Evolutionary Cognitive Neuroscience*, (Platek, S. M., & Shackelford, T. K. eds.) Cambridge: Cambridge University Press.
- Glazko, G. V., & Nei, M. (2003) Estimation of Divergence Times for Major Lineages of Primate Species. *Molecular Biology and Evolution*, 20, 424-434.
- Goodale, M. A. (1998) Visuomotor control: Where does vision end and action begin? *Current Biology*, 8, 489-491.
- Goodale, M. A., & Haffenden, A. (1998) Frames of Reference for Perception and Action in the Human Visual System. *Neuroscience & Biobehavioral Reviews*, 22, 161-172.
- Goodale, M. A., & Humphrey, G. K. (1998) The objects of action and perception. *Cognition*, 67, 181-207.
- Goodale, M. A., & Milner, A. D. (1992) Separate visual pathways for perception and action. *Trends in Neuroscience*, 15, 20-25.
- Goodale, M., & Milner, D. (2006) One brain – two visual systems. *The Psychologist*, 19, 660-663.
- Gärdenfors, P. (2000) *Hur Homo blev sapiens: om tänkandets evolution*. Nora: Bokförlaget Nya Doxa.
- Hanus, D., & Call, J. (2008) Chimpanzees infer the location of a reward on the basis of the effect of its weight. *Current Biology*, 18, 370-372.

- Haun, D. B. M., & Call, J. (2009) Great apes' capacities to recognize relational similarity. *Cognition*, *110*, 147-159.
- Haun, D. B. M., Rapold, C. J., Call, J., Janzen, G., & Levinson, S. C. (2006) Cognitive cladistics and cultural override in Hominid spatial cognition. *PNAS*, *103*, 17568-17573.
- Hume, D. (1739) *A Treatise of Human Nature Book I: Of the Understanding*. London. (Swedish translation: (2002) *Avhandling om den mänskliga naturen bok I: Om förståndet*, Thales.)
- Jeannerod, M., & Jacob, P. (2005) Visual cognition: a new look at the two-visual systems model. *Neuropsychologia*, *43*, 301-312.
- Kaas, J. H. (2005) The Evolution of Visual Cortex in Primates. In: *The Primate Visual System: A Comparative Approach*, (Kremers, J., ed), 267-283.
- Kant, I. (1781) *Kritik der reinen Vernunft*. Hartknoch Riga (Swedish translation: (2004) *Kritik av det rena förnuftet*, Thales.)
- Marino, L. (2004) Dolphin cognition. *Current Biology*, *14*, 910-911.
- Marr, D. (1982) *Vision*. San Francisco: W. H. Freeman and Company.
- McGrew, W. (2004) *The Cultured Chimpanzee: Reflections on Cultural Primatology*. Cambridge: Cambridge University Press.
- Millikan, R. G. (2004) *Varieties of Meaning*. Cambridge: MIT Press.
- Millkan, R. G. (2009) Embedded Rationality. In: *Cambridge Handbook of Situated Cognition*, (Aydede, M., Robbins, P. eds.) Cambridge: Cambridge University Press.
- Milner, A. D. (1998) Streams and consciousness: visual awareness and the brain. *Trends in Cognitive Science*, *2*, 25-30.
- Milner, A. D., & Dyde, R. (2003) Why do some perceptual illusions affect visually guided action, when others don't?. *Trends in Cognitive Science*, *7*, 10-11.
- Milner, A. D., & Goodale, M. A. (2008) Two visual systems re-viewed. *Neuropsychologia*, *46*, 774-785.
- Nakamichi, M. (2004) Tool-use and tool-making by captive, group-living orangutans (*Pongo pygmaeus abelii*) at an artificial termite mound. *Behavioural Processes*, *65*, 87-93.
- Oberzaucher, E., & Grammer, K. (2008) Everything is movement: on the nature of embodied communication. In: *Embodied communication*. (Wachsmuth, I. & Knoblich, G. eds.) Oxford University Press.
- Osvath, M. (2009) Spontaneous planning for future stone throwing by a male chimpanzee. *Current Biology*, *19*, 190-191.
- Papini, M. R. (2008) *Comparative Psychology: Evolution and Development of Behavior (Second edition)*. New York: Psychology Press.
- Pinker, S. (1997) *How the Mind Works*. London: Penguin Books.
- Povinelli, D. J., Theall, L. A., Reaux, J. E., & Dunphy-Lelii, S. (2003) Chimpanzees spontaneously alter the location of their gestures to match the attentional orientation of others. *Animal Behaviour*, *66*, 71-79.
- primate. (2009) In *Encyclopædia Britannica*. Retrieved September 16, 2009, from Encyclopædia Britannica Online.
- Prinz, J. J. (2004) *Furnishing the Mind: Concepts and their Perceptual Basis*. Cambridge: MIT Press.
- Ro, T. (2008) Unconscious vision in action. *Neuropsychologia*, *46*, 379-383.
- Ross, C. F., & Martin, R. D. (2007) The Role of Vision in the Origin and Evolution of Primates.
- Schenk, T. (2006) An allocentric rather than perceptual deficit in patient D.F. *Nature Neuroscience*, *9*, 1369-1370.
- Seed, A., & Call, J. (2009) *Causal Knowledge for Events and Objects in Animals*. In: Watanabe, S., Blaisdell, A. P., Huber, L., & Young, A. (Eds), *Rational animals, irrational humans*. Tokyo: Keio University.
- Sewards, T. V., & Sewards, M. A. (2002) Innate visual object recognition in vertebrates: some proposed pathways and mechanisms. *Comparative Biochemistry and Physiology*, *132*, 861-891.
- Tanifuji, M., Tsunoda, K., & Yamane, Y. (2005) Representation of Object Images by Combinations of Visual Features in the Macaque Inferotemporal Cortex. In: *From Monkey Brain to Human Brain*, (Dahaena, S., Duhamel, J., Hauser, M. D., & Rizzolatti, G. eds.) Cambridge: MIT Press.
- Tomasello, M., Call, J., & Hare, B. (1998) Five primate species follow the visual gaze of conspecifics. *Animal*

Behaviour, 55, 1063-1069.

Van Essen, D. C. (2005) Comparisons of Macaque and Human Cortical Organization. In: *From Monkey Brain to Human Brain*, (Dahaena, S., Duhamel, J., Hauser, M. D., & Rizzolatti, G. eds.) Cambridge: MIT Press.

Zahle, T., Jordan, K., Wüstenberg, T., Baudewig, J., Dechent, P., & Mast, F. W. (2006) The neural basis of the egocentric and allocentric spatial frame of reference. *Brain Research*, 1137, 92-103.