"What's HIDD'n in the HADD? – A cognitive conjuring trick?"[1]

Anders Lisdorf

Abstract

The consensus in the cognitive study of religion is that some sort of hyperactive agency detection in the human mind is responsible for the origin and spread of beliefs in superhuman agents such as gods, spirits and ancestors among human populations. While it is expressed differently in different authors, they all agree that some sort of hyperactive agency detection is a basic function of human cognition, which is what Justin Barrett has called the Hyperactive Agency Detection Device or HADD. But what is it? And isn't it a bit much to ask of one cognitive function to be the origin of religious belief? Problems arise when we begin to consider the neural basis: It is not there, or more precisely it doesn't work that way. Like the magician pulling rabbits from the hat this explanation may be a "self"-conjuring trick, only for us the hat is a HADD and the rabbits are superhuman agents (no reference to were-rabbits intended). This paper will try to point to a more parsimonious explanation.

Introduction

The basic line of argumentation for most research in cognition and culture is the following: Cultural representations are widespread in human cultures because they are transmitted often. What gets transmitted depends on how well it fits the human cognitive system. The human cognitive system is evolved through natural selection, so that adaptive cognitive functions have provided[2] This approach has been used to a large extent in the cognitive science of religion to explain why religion is widespread in human cultures (Atran 2003; Atran & Norenzayan 2004; Boyer 2003). A central feature of this endeavour has been to explain the ubiquity in human cultures of belief in superhuman agents (Atran & Norenzayan 2004; Barrett 2004; Boyer & Ramble 2001). This has in most cases been done with recourse to a stipulated hyper active agency detection function, what Justin Barrett has termed the Hyperactive Agency Detection Device or HADD. In this article I will consider whether the widespread belief in superhuman agents (gods, ghosts and ancestors) can be adequately explained by the HADD using the above mentioned approach.

[1] Revised and expanded version of a paper given at the "Origins of religion, cognition, and culture" conference in Aarhus, Denmark, 5-7 January 2006.

[2] The argument has been spelled out more in more detail by John Tooby, Leda Cosmides, Dan Sperber and others (Barkow et al 1995; Bloch & Sperber 2002; Sperber & Hirschfeld 2004; Sperber 1996)

Explaining cognitive functions by evolutionary psychology

The ultimate explanation of hyperactive agency detection is made with recourse to evolutionary psychology. Evolutionary psychology's strategy is to explain universal cognitive phenomena by evolved psychological functions. Let us consider this strategy as it is presented by Tooby and Cosmides. Any given cognitive function is stipulated to have had a selective advantage in the human ancestral environment thereby accounting for its present existence, as the slogan goes: "Our modern skulls house a stone age mind" The function may be triggered in other contexts than that for which it was selected. These are by-products of the original function[3].

But a couple of quite basic points need to be emphasised: cognitive functions are not inherited, only genes are, and cognitive functions do not have any selective advantages, only behaviour does. Second cognitive functions cannot be observed, only inferred from behaviour. Likewise only the scantiest evidence exists of any connection between a given gene and a cognitive function. Add to this the hypothetical function in a largely unknown environment (the stone age), and we end up with quite a mixture of "what if's", "maybes" and "probablies". That, of course, is the nature of science. But the nature of science is also to use all available evidence in the investigation, and not introduce more hypothesis than are needed.

For a cognitive function F to have been adaptive, it is necessary that a gene produces a neural structure S which reliably produces the cognitive function F. F leads to behaviour B, which has been adaptive in an ancestral environment. Empirically we only have access to B and S. B is observed in psychological experiments, and S can be observed through a number of techniques, such as various scanning techniques, lesion studies, single cell recordings, brain surgery and autopsies. Evolutionary psychology has almost exclusively focussed on B (eg. the primer). This is not a wise move in an environment with so many hypotheticals.

So, before we start to look for an extra hypothesis concerning the environment of evolutionary adaptation for any uniquely human cognitive function F we should expect to find:

1) A neural structure, S, reliably producing the stipulated cognitive function, F, under certain input conditions causing a range of behaviours B1..Bn.
2) S should be unique to humans
3) All instances of B1..Bn should be caused by S through F

[3] For a short and accessible introduction presenting these thoughts confer the primer on the web-site for the Center for Evolutionary Psychology (http://www.psych.ucsb.edu/research/cep/primer.html (9 January 2006))

The first point is the causal nexus explaining the existence of any given evolved cognitive function. Usually evolutionary psychology is manipulating B in psychological experiments (eg. the Wason selection task of Tooby & Cosmides); a difference in cognitive function is inferred from differences in behaviour. Tooby & Cosmides found that subjects performed differently in the "Wason selection task", when the problem was framed as a social problem (Cosmides & Tooby 2005: 597).

The second point naturally only goes for behaviours specific to humans. Behaviours we share with animals would not include this clause.

The third point is trickier. It is actually a clause of parsimony. It is possible that two apparently identical behaviours have different neural substructures. This is the case for example of ritualised behaviour associated with OCD and normal routine behaviour (Boyer & Lienard 2006: Section 6.4). They are then analogical features, and need two different explanations. The EP account has to make sure that the features are homological, or must assume that they are.

What does the HADD do? examples:

While a number of scholars have used some sort of hyperactive agency detection I will focus on only two examples namely Atran & Norenzayan and Justin L.Barrett. Let me just shortly mention those left out: Stewart Guthrie's account is the first, but he does not think of agency detection as a cognitive function, but merely as a perceptual function (Guthrie 1993; Guthrie 2004; Guthrie 1980; Guthrie 2001). Pascal Boyer does not explain the function with recourse to agency detection per se, but instead with reference to an inbuilt search for relevance and agentive concepts can produce more inferences than non-agentive. The explanation is thus made more with recourse to a cost/benefit algorithm than to hyperactivity (Boyer 2001; Boyer 1996; Boyer & Ramble 2001). I do believe that these arguments suffer from other set backs, but space does not allow me to treat them thoroughly.

Let us presently consider how hyperactive agency detection works according to Justin Barrett:

Justin Barrett

According to Justin L. Barrett : "Part of the reason people believe in gods, ghosts and, goblins also comes from the way in which our minds, particularly our agency detection device (ADD) functions. Our ADD suffers from some hyperactivity, making it prone to find agents around us,

including supernatural ones, given fairly modest evidence of their presence. This tendency encourages the generation and spread of god concepts" (Barrett 2004: 31). This is what he terms the Hyperactive Agency Detection Device or HADD (32). Examples of when the HADD is active include: Hearing a bump in the night (31), perception of geometric objects moving by themselves (32), perception of a wispy form (33), a computer malfunctioning (33), perception of crop circles (37). To this we might add all representations of "gods, ghosts and goblins". I do want to add that Barrett does not say that the HADD is the whole story, he actually has several chapters explaining other important factors, but he does say that it is the central function.

Atran and Norenzayan

According to Scott Atran and Ara Norenzayan: "Widespread counterfactual and counterintuitive beliefs in supernatural agents can be explained by the fact that they "(..)trigger our naturally selected agency-detection system, which is trip-wired to respond to fragmentary information, inciting perception of figures lurking in shadows and emotions of dread and awe" (Atran & Norenzayan 2004: 714). While they call it the agency-detection system, it is the same basic explanation as Barrett gives for the HADD. Examples of when this system is active are: Dots moving on a screen, voices in the wind, faces in the clouds, any complex design, Mother Theresa in a cinnamon bun, the virgin of Guadelupe in a melted ice cream on the pavement, the Virgin in windows, curtains and television afterimages, the face of the Evil-One in the smoke of the World Trade Center (719), sudden movement of an object stirred by the wind, representing ghosts or gods (720).

Atran and Norenzayan do have some other nuances than do Barrett, and their explanation also incorporates emotions, but these observations will suffice for the present purposes.

What does the HADD do? a benchmark:

Since belief in superhuman agents by these accounts qualify as uniquely human, all three points mentioned above must apply to the HADD. Religion is considered a by-product of human cognitive functions. According to Barrett and Atran and Norenzayan the proper function of the HADD was to detect and evade predators, the by-product is a susceptibility to infer superhuman beings. Let me

highlight four representative examples of situations in which the HADD is stipulated to be activated as a sort of short hand benchmark[4]:

> B1 the sudden movement of a twig in the forest leads the inference of ghosts, which can be seen in subsequent caution, sacrifice or stories relating the event.
>
> B2 the perception of crop-circles leads to the inference of aliens, which can be seen in stories about the aliens communicating their far advanced technologies to us.
>
> B3 prayer activating the inference of a god listening, which can be seen from the communicative behaviour in a context where there is no one to communicate with.
>
> B4 Reading the book of Genesis in the Bible, which leads to talk about the design of nature.

These four examples are, or lead to, behaviours involving superhuman agents. So we have four distinct behaviours B1-B4, which are the explananda. Either the HADD is active in B1-4 or it is not the most parsimonious explanation.

Now let me jump to the first point, the neural structure. We should expect to find a neural structure which reliably produces hyperactive agency detection, and, which is specific to humans.

Neural basis - a review of possible candidates for the HADD

Superior temporal sulcus

The classic example of hyperactive agency detection, mentioned by Barrett and Atran and Norenzayan, is Michotte's experiments (Michotte 1963). Michotte had subjects watch a screen on which dots were moving. When they were moving freely, apparently self-propelled, and reacted to other dots, subjects were inclined to infer mental states such as chasing, helping etc. This obviously suggests some sort of hyper active agency detection.

A good portion of neurological research has been dedicated to this function (Blakemore et al 2001; Decety & Grezes 1999; Frith & Frith 2001; Grezes et al 2001). It is usually referred to as detection of biological movement, and the neural structure responsible for this is the superior

[4] I will term them B1-B4 because the empirical measure is some sort of behaviour. In some of the examples the behaviour is a product of something perceptual (B1, B2 and B4) while B3 does not include anything perceptual. B1 includes movement while this is not the case in the others. These examples have been chosen to indicate the breadth of the claims about the conditions under which the HADD functions.

temporal sulcus, or STS. While most studies have focussed on its activation by visual stimuli, a recent fMRI study has showed that part of the STS was also activated by auditory stimuli (Bidet-Caulet et al 2005). There are lateral specialisations in the STS (Blakemore et al 2001), and also some substructures sub-serving specific sub-functions (Bidet-Caulet et al 2005; Goodale & Milner 1992; Grezes et al 2001).

Indeed this neural structure would account for for B1, where something is moving, but not B2-4, where nothing is moving. Thus it cannot be the S explaining the HADD.

There is also another problem, namely, that it is not unique to humans.

Mirror neurons

Another possibility is that the, so called, mirror neuron circuit is responsible for the hyperactive agency detection. Mirror neurons were found in the ventral pre-motor cortex in macaque monkeys (Gallese et al 1996). They have the special property that they are activated both by the perception of a movement and the execution of a movement. This has been seen as the basis of mentalizing or empathy (Gallese 2001)[5]. Comparable findings have been reported in homologous areas in humans. These areas were found to be active during perception of a movement, execution of the movement, but also in the imagination of the movement (Jeannerod 1999) (Jeannerod 1999). Thus mirror neurons could enable us to simulate a non-present agent. This would account for B3- 4, where some sort of simulation of an agent takes place, but it wouldn't explain B1. Mirror neurons are not either very hyperactive, and they seem to function at a very low and boring level of movements. Another problem is, as mentioned, they are found also in monkeys, even macaques. So, the mirror neuron circuit cannot either be the S of the HADD

A combination

We could then stipulate that the STS and PMv together formed a distributed neurological instantiation of the HADD. There are two objections to this. The mirror neuron circuit is not hyperactive at all; at least not in the sense stipulated by Barrett[6]. It is also a pretty pedestrian movement centred system that does not care much about the mental states, or long term goals of the perceived agent. Second this system is not in any obvious way unique to humans.

[5] For recent criticism of this see (Jacob & Jeannerod 2005)

[6] Although some sort of hyperactivity has been detected: activity in left PMv was detected in subjects viewing tools, and implied movement (Chao & Martin 2000), but these are more commonly interpreted as being simulations of the affordances of the object at hand. Strictly speaking these are not mirror neurons but so called "canocical neurons". This would not explain B2 and B3

Summary

We have not been able to find an S which reliably produces hyperactive agency detection leading to behaviours B1-4.

While there is ample evidence of hyperactive agency detection, the HADD is not the most parsimonious explanation of why the belief in superhuman agents and their associated behaviours is widespread among humans. So in short the HADD as explaining belief in superhuman agents is an illusion. We may be conjuring ourselves to think that we have explained what we wanted to. Like the magician pulling rabbits from the hat, the HADD does not contain all the rabbits.

Another trick?

Any skilled entertainer knows you must always have another trick ready if the first one doesn't work. So I would like to direct your attention to the title, in which I have introduced a clue. So my trick for know will be a guessing game.

A possible way out of the problem would be to focus on the central function before we eagerly make up some evolutionary scenario.

What we really want to know is why humans have a tendency to interact with physical entities as if they were intentional. This central function is most precisely put in the by Daniel Dennett, who calls this "taking the intentional stance": "first you decide to treat the object whose behaviour is to be predicted as a retional agent; then you figure out what beliefs the agent ought to have, given its place in the world and its purpose. Then yuou figure our what desites it ought to have on the same considerations and finally you predict this rational agent would act to further its goals in the light of its beliefs" (Dennett 1987, 17). This indeed seems to be a hyperactive function, that is, a function with many false positives, as Dennet already pointed out in the seventies.

We should then instead talk of a hyperactive intentionality detection device or a HIDD. Let us see if this move will fare better than the HADD.

Neural basis of the HIDD

Anterior paracingulate cortex (PCC)

In a series of fMRI studies Gallagher and Frith have tried to find out what happened when people took the intentional stance. Previous studies had shown a discrete system involved in ToM tasks, but

all of them were off-line, that is reflective like narrative comprehension. To isolate the intentional stance in on-line processing, they asked subjects to play a computerised version of the game stone, paper scissors. One group was told that they were playing against an experimenter, another that they were playing against a computerised using a predetermined rule-based strategy, while a third was told they were playing against a random sequence. The first group were thus prompted to take the intentional stance toward the game, reflecting quite accurately what takes place in most religion. What they didn't know was that all three groups were in fact playing against a random sequence. So the only difference between the groups was the stance taken towards the task. The results showed just one area, which had also been activated in all other mentalizing studies: the anterior paracingulate gyrus (PCC) (Gallagher et al 2002).

Now let us consider whether the stipulated neural structure could cause all instances of the behaviours we wished to explain:

> B1 the sudden movement of a twig in the forest leads the inference of ghosts, which can be seen in subsequent caution, sacrifice or stories relating the event.

This would amount to biological movement, since this is characterized by being movement by and of itself. Castelli and colleagues used animations of two characters, a blue triangle and a red triangle, to investigate the neural activation in ToM tasks (Castelli et al 2000). In some animations the triangles were interacting intentionally like dancing together, chasing eachother or imitating each other, while in others they were bouncing off the walls like billiard balls (316). This showed strong activation, among other places, in the PCC (319)[7].

> B2 the perception of crop-circles leads to the inference of aliens, which can be seen in stories about the aliens communicating their far advanced technologies to us.

Crop circles are traces of agency, or have evidence of design, as Barrett calls it. That is they seem non-random. One study by Gallagher and colleagues might show something similar (Gallagher et al 2000). They used cartoons without any text. This is perceptually similar in that you have straight lines and order. They showed that these cartoons also activated the PCC, but other cartoons did not (19). So obviously we have another variable here, which needs further investigation.

[7] The study mentions the medial prefrontal cortex, but the PCC is a sub-part of that area.

B3 prayer activating the inference of a god listening, which can be seen from the communicative behaviour in a context where there is no one to communicate with.

Prayer is the same as interacting with non-present agents. This was shown in the already mentioned study of Gallagher et al, where a game of rock, paper scissors was played with a non-present agent. Something similar was also shown by a study of McCabe and colleagues, where subjects were told they were a game playing against human (McCabe et al 2001). This study also showed increased activity in activation of the PCC[8].

B4 Reading the book of Genesis in the Bible, which leads to talk about the design of nature.

Reading the bible like any other book is a matter of story comprehension. A couple of studies have shown that the PCC is activated in such tasks (Fletcher et al 1995; Gallagher et al 2000). There might of course be differences in neural activation in reading the bible, but what concerns us here is that reading the bible will activate the PCC.

So in conclusion we can say that the stipulated neural structure responsible for our cognitive function, the HIDD, is indeed activated in the different behaviours we wished to explain.

The key here is that there are different ways of activating this function. Earlier it has been assumed more or less that agency detection was stimuli driven, but the above mentioned indicates that it can just as well be context driven.

But how far is this uniquely human? The anterior cingulate cortex, of which the PCC is considered to be part, is an ancient structure belonging to the limbic lobe. This would immediately indicate that the neural area is not even specific to primates, but an unusual type of neuron is found here. It is found only in higher primates, not in monkeys. There is evidence that in humans this structure has undergone further evolution from the primates (Gallagher & Frith 2003: 80). This is consistent with a uniquely human structure. Second, this type of neuron is not present at birth in human children. It doesn't appear until around 4 months of age, which would give some clues to the delayed development of mentalizing capabilities in infants (cf. Meltzoff & Prinz 2003).

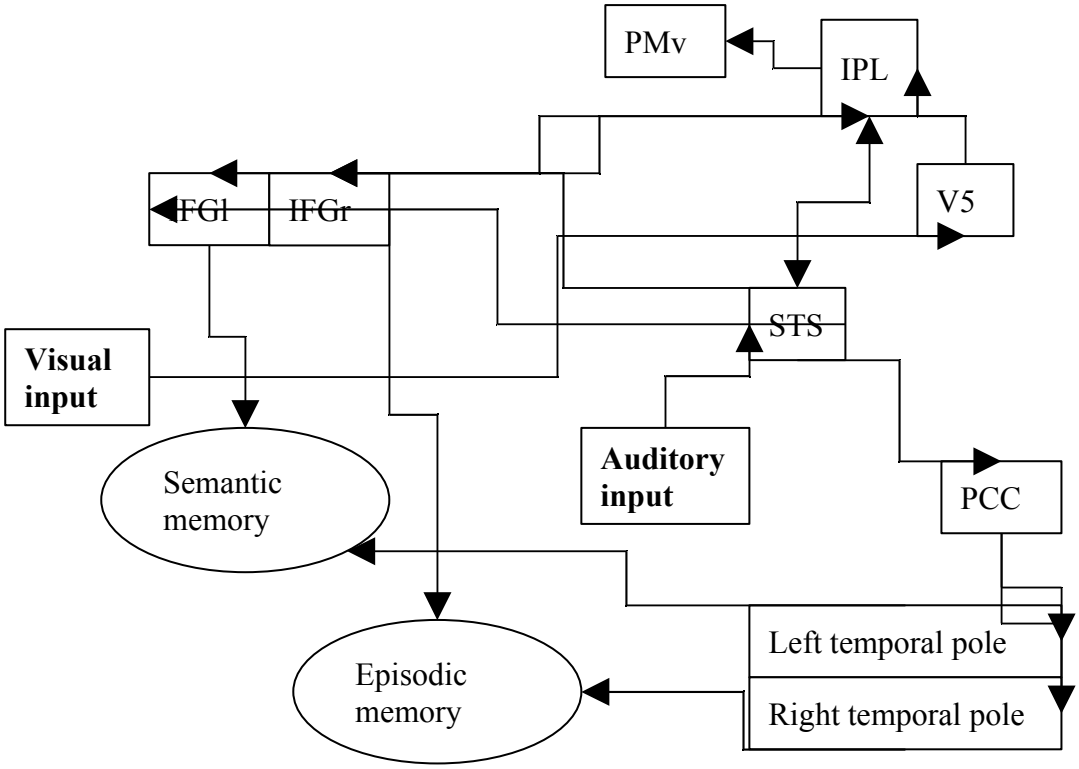
[8] Also here the designation is medial prefrontal cortex (cf. previous note)

Conclusion

This investigation has also shown that supplying an evolutionary explanation about how a stipulated cognitive function would have been adaptive is sometimes detrimental to understanding the phenomenon at hand. Because of the eagerness and easiness with which an adaptive function was supplied, reflection on the phenomenon seized. Further reflection and research into philosophy and neuroscience would have revealed the insufficiency of the argument, and led to what I have presented here.

The proposed HIDD, or just taking the intentional stance, fits all the criteria stipulated in the beginning: it has a well defined neural structure (PCC), which reliably produces the cognitive function (intentionality detection); it is uniquely human, which can be seen neurologically from recent mutation after the phylogenetic departure from the great apes; all instances of the behaviours B1-4 was found to be consistent with activation of the neural structure. The proposed explanation of the widespread belief in superhuman agents, dare I say the origin of religion, is thus more parsimoniously explained by a HIDD than a HADD.

But this is not the end point of the investigation. It is merely the starting point. The next problem is the conditions of activation of the HIDD. We need to understand the input conditions. In order to do this we need to understand the neural connections. I have here compiled a figure, which only highlights some of the most important areas and their most important interconnections at a coarse level.



So to avoid complacency the hyperactive agency detection function calls for increased research to understand the neural, cognitive and behavioural basis of the widespread representations of gods, ghosts and ancestors. I hope only to have shown that the functions needed to explain this is much more complicated, and we need to integrate all the knowledge we have from a wide variety of disciplines to begin to solve the riddle.

References

Atran, S. 2003: "A cognitive theory of culture: An evolutionistic alternative to sociobiology and collective selection", *Homme*: 107-143

Atran, S. and Norenzayan, A. 2004: "Religion's evolutionary landscape: Counterintuition, commitment, compassion, communion", *Behavioral and Brain Sciences*, nr. 27: 713-+

Barkow, J, Cosmides, L, and Tooby, J 1995: *The Adapted Mind*, Oxford University Press, New York,

Barrett, Justin L. 2004: *Why Would Anyone Believe in God?*, Altamira, Walnut Creek,

Bidet-Caulet, A., Voisin, J., Bertrand, O., and Fonlupt, P. 15-10-2005: "Listening to a walking human activates the temporal biological motion area", *Neuroimage*, nr. 28: 132-139

Blakemore, S. J., Fonlupt, P., Pachot-Clouard, M., Darmon, C., Boyer, P., Meltzoff, A. N., Segebarth, C., and Decety, J. 4-12-2001: "How the brain perceives causality: an event-related fMRI study", *Neuroreport*, nr. 12: 3741-3746

Bloch, M. and Sperber, D. 2002: "Kinship and evolved psychological dispositions - The mother's brother controversy reconsidered", *Current Anthropology*, nr. 43: 723-748

Boyer, P. 2003: "Religious thought and behaviour as by-products of brain function", *Trends in Cognitive Sciences*, nr. 7: 119-124

Boyer, P. and Lienard, P 2006: "Why Ritualized Behavior? Precaution Systems and Action Parsing in Developmental, Pathological and Cultural Rituals", *Behavioral and Brain Sciences*, nr. Forthcoming

Boyer, P. and Ramble, C. 2001: "Cognitive templates for religious concepts: cross-cultural evidence for recall of counter-intuitive representations", *Cognitive Science*, nr. 25: 535-564

Castelli, F., Happe, F., Frith, U., and Frith, C. 2000: "Movement and mind: A functional imaging study of perception and interpretation of complex intentional movement patterns", *Neuroimage*, nr. 12: 314-325

Chao, L. L. and Martin, A. 2000: "Representation of manipulable man-made objects in the dorsal stream", *Neuroimage*, nr. 12: 478-484

Cosmides, L and Tooby, J 2005: "Neurocognitive Adaptations for Social Exchange", i Buss, D. M., *Evolutionary Pshycology Handbook*, Wiley, NY: 584-627

Decety, J. and Grezes, J. 1999: "Neural mechanisms subserving the perception of human actions", *Trends in Cognitive Sciences*, nr. 3: 172-178

Fletcher, P. C., Happe, F., Frith, U., Baker, S. C., Dolan, R. J., Frackowiak, R. S. J., and Frith, C. D. 1995: "Other Minds in the Brain - A Functional Imaging Study of Theory of Mind in Story Comprehension", *Cognition*, nr. 57: 109-128

Frith, U. and Frith, C. 2001: "The biological basis of social interaction", *Current Directions in Psychological Science*, nr. 10: 151-155

Gallagher, H. L. and Frith, C. D. 2003: "Functional imaging of 'theory of mind'", *Trends in Cognitive Sciences*, nr. 7: 77-83

Gallagher, H. L., Happe, F., Brunswick, N., Fletcher, P. C., Frith, U., and Frith, C. D. 2000: "Reading the mind in cartoons and stories: an fMRI study of 'theory of mind' in verbal and nonverbal tasks", *Neuropsychologia*, nr. 38: 11-21

Gallagher, H. L., Jack, A. I., Roepstorff, A., and Frith, C. D. 2002: "Imaging the intentional stance in a competitive game", *Neuroimage*, nr. 16: 814-821

Gallese, V. 2001: "The 'shared manifold' hypothesis - From mirror neurons to empathy", *Journal of Consciousness Studies*, nr. 8: 33-50

Gallese, V., Fadiga, L., Fogassi, L., and Rizzolatti, G. 1996: "Action recognition in the premotor cortex", *Brain*, nr. 119: 593-609

Goodale, M. A. and Milner, A. D. 1992: "Separate Visual Pathways for Perception and Action", *Trends in Neurosciences*, nr. 15: 20-25

Grezes, J., Fonlupt, P., Bertenthal, B., on-Martin, C., Segebarth, C., and Decety, J. 2001: "Does perception of biological motion rely on specific brain regions?", *Neuroimage*, nr. 13: 775-785

Guthrie 1993: *Faces in the clouds*, Oxford University Press, New York,

Guthrie, S. 1980: "A Cognitive Theory of Religion", *Current Anthropology*, nr. 21: 181-203

Guthrie, S. 2001: "Rethinking animism", *Journal of the Royal Anthropological Institute*, nr. 7: 156-157

Guthrie, S. 2004: "Folk Religion in a Rural Japanese Hamlet: A Fieldwork Perspective", *Temenos*, nr. 39-40: 101-123

Jacob, P. and Jeannerod, M. 2005: "The motor theory of social cognition: a critique", *Trends in Cognitive Sciences*, nr. 9: 21-25

Jeannerod, M. 1999: "To act or not to act: Perspectives on the representation of actions", *Quarterly Journal of Experimental Psychology Section A-Human Experimental Psychology*, nr. 52: 1-29

McCabe, K., Houser, D., Ryan, L., Smith, V., and Trouard, T. 25-9-2001: "A functional imaging study of cooperation in two-person reciprocal exchange", *Proceedings of the National Academy of Sciences of the United States of America*, nr. 98: 11832-11835

Meltzoff and Prinz 2003: *The Imitative Mind*, Cambridge University Press, Cambridge, U.K,

Michotte, A. E. 1963: *The perception of causality*, Basic Books,

Sperber, D. and Hirschfeld, L. A. 2004: "The cognitive foundations of cultural stability and diversity", *Trends in Cognitive Sciences*, nr. 8: 40-46

Sperber, Dan 1996: *Explaining culture*
a naturalistic approach, Blackwell, Oxford, UK,