

# Influence of goals on observation of actions: functional neuroimaging studies

Johannes Wolfram Robert Schultz

Wellcome Department of Imaging Neuroscience

University College London

Submitted for PhD in Neurological Studies

## **Abstract**

Mentalising or Theory-of-Mind (ToM) is defined as the attribution of mental states to other agents. While this capacity develops progressively in children, an important step is reached by passing the false-belief task, normally at about four years of age. Measures of brain activity during performance of a wide range of tasks requiring ToM have repeatedly demonstrated involvement of a particular set of brain regions. But how each of these regions contributes to this process is not yet clear. Based on previous data and a model of the cognitive components necessary for ToM, I performed three experiments using event-related functional magnetic resonance imaging in healthy volunteers to clarify the involvement of brain regions in important components of the ToM capacity. Two different cognitive processes were studied: 1) the identification of potential living entities in the environment and 2) the observation of human actions. In both types of processes, one variable appears to play an important role according to the literature: the presence of goals in the observed actions. This variable was therefore manipulated in all experiments. When healthy subjects watched two disks, moving in a seemingly animate way, interact with each other, activity in a region known to respond to biological motion (the posterior part of the superior temporal sulcus area, or pSTS) increased parametrically with the presence of a goal in the behaviour of the disks, as did attribution of animacy. In a second experiment using moving disks, the pSTS showed greater activation when a chasing disk appeared to attribute goals to the target rather than simply following it. The third experiment showed a role of the pSTS in the analysis of human movement kinematics during categorisation of actions depending on goal-directedness. The role of goals in the neural basis of mentalising is discussed.

## **Table of contents**

<b>Abstract</b>	<b>2</b>
<b>Table of contents</b>	<b>3</b>
<b>Acknowledgements</b>	<b>7</b>
<b>List of figures</b>	<b>8</b>
<b>List of tables</b>	<b>10</b>
<b>List of appendices</b>	<b>10</b>
<b>List of abbreviations</b>	<b>11</b>
<b>Part 1: Literature review</b>	<b>12</b>
<i>Chapter 1. General introduction</i>	<i>12</i>
1.1 Definitions of commonly used terms	13
Mental states	13
Intentions and intentionality	13
Goals and actions	15
Agents and animacy	17
1.1. A brief history of Theory-of-Mind	18
Does the chimpanzee have a Theory-of-Mind?	18
The False-belief task	19
A good control	21
1.2 Possible neural substrates for ToM	23
The medial prefrontal cortex	24
The temporal poles	26
The posterior part of the superior temporal sulcus	28
Other regions	28
<i>Chapter 2. Cognitive processes in ToM</i>	<i>30</i>
2.1. Representing mental states	30
Intentional relations	30
Simulation Theory and Theory Theory	31
But maybe we use both?	32
2.2. Recent cognitive models for ToM	34
Understanding others' actions with one's own action system	34
Internal models for prediction and control	35
An internal model-based system for understanding mental states	38

2.3. An extended model	41
Components	41
What it can do	42
The A-M module	43
Neural bases	44
<i>Chapter 3. Processes for mentalising studied in this thesis</i>	45
3.1.1 Animacy and intentionality in moving abstract shapes	45
Parameters for animacy and intentionality attribution	46
Animacy in single moving dots	47
Contingency	48
Identification of actions in terms of goals	49
Self-propelled movements and goal-directedness?	51
3.1.2 Biological bases for the perception of animate moving objects	52
Biological motion and the superior temporal sulcus	53
Attention effects	58
3.2.1 Action observation and imitation: the role of goals	58
Imitation and action observation in adults: the ideomotor principle	58
Goal-directed imitation in children	61
3.2.2 Biological bases for action observation	62
a) Studies in monkeys	63
b) Studies in humans	67
c) Anatomical regions of interest in this thesis	70
<i>Chapter 5. Conclusions</i>	72
<b>Part 2: Rationale for experiments</b>	<b>73</b>
<i>Chapter 1. General rationale</i>	73
<i>Chapter 2. Specific rationale for experiments involving observation of intentional objects.</i>	73
2.1. Experiment 1	74
2.2. Experiment 2	75
2.3. The role of attention	75
<i>Chapter 3. Specific rationale for action observation</i>	77
<b>Part 3. Materials and methods</b>	<b>79</b>
<i>Chapter 1. MRI and fMRI</i>	79
1.1. Basic MRI and fMRI physics	80
Spin	81
Equilibrium magnetisation	82
Radiofrequency magnetic fields	83
Relaxation	85

Image formation: frequency and phase encoding	88
Voxels	90
Image contrast	90
Ultrafast MRI sequences: Echo-Planar Imaging	91
1.2. FMRI and neural activity	93
BOLD contrast in fMRI	93
Physiology of BOLD	94
Correspondence between neural activity and the BOLD signal	95
<i>Chapter 2. FMRI analysis</i>	97
2.1. Introduction	97
2.2. Spatial preprocessing	99
Realignment	99
Slice-time correction	99
Normalisation	100
Spatial smoothing	101
2.3. Model-based analysis of fMRI data	102
Basic description of the SPM approach	102
Construction of the model	102
Eliminating temporal confounds	104
Fitting the model	104
t and F-statistics	106
2.4. Inference and the Theory of Gaussian Random Fields	107
i) Whole-brain inferences	107
ii) Anatomically constrained inferences	109
Inferences at the population level: Random Effects	110
Threshold used in this thesis	112
<b>Part 4. Experiments</b>	<b>113</b>
Overview	113
<i>Chapter 1. Experiment 1: Parametric study of animate motion</i>	<i>114</i>
Movement equation	114
Methods	121
Results	126
Discussion	140
Chapter 2. Experiment 2: Study of goal-attribution in a chasing paradigm	145
Methods	145
Results	149
Discussion	158
<i>Chapter 3. Experiment 3: action observation</i>	<i>164</i>

3.1. Introduction and overview	164
3.2. The fMRI experiment	166
Methods	166
Results	175
Discussion	189
<b>Part 5. General discussion</b>	<b>199</b>
1. <i>So what is this thesis all about?</i>	199
2. <i>Activation in pSTS during observation of abstract and human movements</i>	201
3. <i>Other functions of the superior temporal sulcus and gyrus</i>	204
4. <i>Potential confounding factors</i>	208
5. <i>How the 3 mentalising regions could interact</i>	210
6. <i>What comes next? Future directions</i>	212
7. <i>Conclusion</i>	214
<b>Part 6. Appendices</b>	<b>215</b>
<i>Appendix 1: Behavioural reports of animacy testing in Experiment 1.</i>	215
<i>Appendix 2: kinematic recordings of object manipulations used in Experiment 3</i>	217
<b>Part 7. References</b>	<b>220</b>

## **Acknowledgements**

My greatest thanks go to my principal supervisor Chris Frith, for his ideas, his calm, his humour and his constant support, and to my second supervisor Daniel Wolpert for his enthusiasm and technical prowess. Without these two persons, I would probably never have come to the Wellcome Department nor written a single line of this thesis. Further thanks are due to collaborators John O'Doherty, Hiroshi Imamizu, Karl Friston and Mitsuo Kawato. These people, and many others as well, have shown me how fascinating full-time research at a very high level can be, and have helped me in my attempts to enter this world. Additional thanks go to Katja Wiech and Frédérique de Vignemont for critical comments on this thesis. I would like to thank the Wellcome Trust for the Wellcome Prize Studentship that financed me during 3 years, the James McDonnell Foundation whose grant to my supervisors supplied additional funding for 9 months and enabled the collaboration with the ATR Lab in Japan, and the British Council whose Overseas Research Student Award paid my overseas graduate student fees. I would also like to thank the radiographers at the Functional Imaging Laboratory in London and at the ATR Brain Activity Imaging Center, and all the volunteers (from outside and inside the FIL!) who participated in my studies. Additional personal thanks go to my parents for their unconditional support, my friends at the FIL for help on various topics ranging from statistics to relationships and beer appreciation, and to the Queen's Larder for an unending supply of Guinness. Finally, thanks to my sister's cat Guinness (not the beer), whose sarcastic glances and relaxed attitude helped things to stay in perspective and my head to stay on my shoulders.

## **List of figures**

Figure Intro.1. The Sally-Ann task	20
Figure Intro.2. The Picture task.	22
Figure Intro.3. Three areas generally associated with mentalising.	24
Figure Intro.4. Gallese and Goldman's "retrodictive simulation" routine.	36
Figure Intro.5. MOSAIC (A) and HMOSAIC (B) models.	38
Figure Intro.6. Application of the MOSAIC model to action observation.	40
Figure Intro.7. A new cognitive model of mentalising.	43
Figure Intro.8. Action - Mental state (A-M) matching system.	43
Figure Intro.9. Johnson and colleagues' furry object.	49
Figure Intro.10. The classic point-light display of biological motion	54
Figure Intro.11. Activation in the superior temporal sulcus and gyrus during observation of biological motion.	57
Figure Intro.12. Imitation in young infants	62
Figure Intro.13. Visual and motor responses of a mirror neuron in area F5.	65
Figure Intro.14. Activity of a mirror neuron in F5 in response to action observation in full vision and in hidden conditions.	66
Figure Physics.1. Magnetic properties of hydrogen protons.	81
Figure Physics.2. Example of precession movement.	83
Figure Physics.3. RF excitation and the effect on proton orientation.	84
Figure Physics.4. T1 relaxation.	86
Figure Physics.5. T2 and T2* decay.	88
Figure Exp 1.1. Example of display.	113
Figure Exp 1.2. Results from the behavioural pilot study.	115
Figure Exp 1.3. Speed.	116

Figure Exp 1.4. Interactive and Control animations.	117
Figure Exp 1.5: Animacy ratings.	120
Figure Exp 1.6. Example of the 8 conditions.	122
Figure Exp 1.7. Interactivity and Speed ratings.	127
Figure Exp 1.8. Response times.	130
Figure Exp 1.9. Eye movements.	131
Figure Exp 1.10. Brain regions whose activation increased linearly with the amount of interactive motion.	135
Figure Exp 1.11. VOI anatomy and data.	137
Figure Exp 1.12. Parameter estimates and time-course data pSTS.	138
Figure Exp 1.13. Time-course of activation in other regions identified in the Interaction contrast (see Table Exp 1.4).	140
Figure Exp 2.1. Design and stimuli.	146
Figure Exp 2.2. Ratings in the Strategy and the Outcome task.	150
Figure Exp 2.3. Response times of the 14 participants.	152
Figure Exp 2.4. Brain areas responding more to Predict than Follow trials, irrespective of task.	154
Figure Exp 2.5. Brain regions: task effects.	156
Figure Exp 2.6. Activation time-courses in the superior temporal gyrus in both Strategy and Outcome tasks.	157
Figure Exp 2.7. Comparison between STS/STG activation in all 3 experiments.	163
Figure Exp 3.1. Experimental design and stimuli.	168
Figure Exp 3.2. Movement kinematics and animation.	170
Figure Exp 3.3. Behavioural results.	176
Figure Exp 3.4. Correlation between kinematic information and behaviour.	178

Figure Exp 3.5. Brain activation during all conditions with a moving actor vs. the resting actor image.	179
Figure Exp 3.6. Effects of object presence.	183
Figure Exp 3.7. Congruence effects.	186
Figure Exp 3.8. Correlation between stimuli, and pSTS activation.	189
Figure Discussion.1. Possible neural correlates of ToM cognitive components.	212

### **List of tables**

Exp 1.1 – Response times	129
Exp 1.2 – Eye movements	131
Exp 1.3 – fMRI 1 – Main effects	132
Exp 1.4 – fMRI 2 – Interaction	133
Exp 1.5 – fMRI 3 – Regions of interest	136
Exp 2.1 – fMRI data	153
Exp 3.1 – Number of joint measures with significant differences between Mimed and Actual movement kinematics	172
Exp 3.2 – fMRI 1 – commonly activated regions for all manipulation types	180
Exp 3.3 – fMRI 2 – differences between manipulation types	181
Exp 3.4 – fMRI 3 – object presence effects	184
Exp 3.5 – fMRI 4 – artificial vs. realistic conditions	186

### **List of appendices**

Appendix 1 – Behavioural reports from animacy rating task in Experiment 1	215
Appendix 2 – Joint variables time-series from Experiment 3	217

## **List of abbreviations**

ACC: Anterior Cingulate Cortex

ATR: Advanced Telecommunications Research laboratories

BOLD: Blood Oxygenation-Level Dependent signal

EEG: Electro-Encephalogram

FIL: Functional Imaging Laboratory

fMRI: functional Magnetic Resonance Imaging

HRF: Hemodynamic Response Function

MEG: Magneto-Encephalography

MIP: Maximum Intensity Projection

MPFC: medial Prefrontal Cortex

PET: Positron Emission Tomography

SPM: Statistical Parametric Mapping

STS, pSTS, STG and pSTG: (posterior) Superior Temporal Sulcus and Gyrus

TMPp: Temporal pole

TPJ: Temporo-Parietal Junction, an area including the pSTS, the inferior parietal lobule (angular and supramarginal gyri).

RT: Response Time

WDIN: Wellcome Department of Imaging Neuroscience

## **Part 1: Literature review**

### **Chapter 1. General introduction**

This thesis will deal with a basic process that enables human beings to interact socially with each other. The first part of the introduction sets the background of the thesis, defines keywords and processes that will be used and referred to frequently, shows connections between them and mentions previous neuroscientific work that led to the association of specific brain structures to some of the processes.

Human beings and other animals live in environments that change frequently, and these changes can be good or bad for them. Also, to survive, humans as well as animals need to interact with each other. To understand the changes that happen in the world, predict future events and influence them, people need to understand underlying principles that cause these events. A classic philosophical view is that humans explain events in the world using either mentalistic or physicalist explanations. Mentalistic explanations are used for animate things, physicalist explanations for inanimate things. Therefore, an essential step for this process is the discrimination between animate and inanimate things; this will be discussed further and become an important part of this thesis. Mentalising, the process of applying mentalistic explanations by attributing mental states to entities in the world, is also known as "Theory-of-Mind" (ToM) (Premack & Woodruff, 1978). This is a useful heuristic for dealing with animate agents (i.e. entities that "do things", meaning that they behave in a goal-directed manner, see below), for predicting and influencing their behaviour if it is not compatible with one's own desires, beliefs, wishes, thoughts or intentions. Mentalising is the key background concept that will be used in this thesis, and experiments will address basic processes thought to be necessary for it.

## 1.1 Definitions of commonly used terms

In this thesis, mental states, goals, actions, intentions and intentionality, agency and animacy will be central concepts, so I would like to define these terms at the beginning in order not to confuse readers.

### Mental states

Mental states in the sense used in this thesis, such as remembering, believing, desiring, hoping, knowing, intending, feeling, experiencing are mental representations of a state of the world; this state could be present, past or future, real or hypothetical. For example, the sentence “Carolina wants to eat chocolate” describes the mental state of Carolina, which is the desire to eat chocolate. Mental states are useful for understanding and predicting the behaviour of others: to continue the previous example, if I knew that Carolina wants to eat chocolate, I will not be very surprised if I see her reaching for a chocolate bar. It is therefore worthwhile to be able to identify mental states of other organisms. But being *mental* states, they are solely present in the mind of the organism that has the mental state. However, as the organism’s behaviour can be related to its mental states, an observer can use the behaviour to identify potential mental states explaining that behaviour. This is done by mentalising.

### Intentions and intentionality

Intentions and intentionality are classic concepts in philosophy, and as I will use them repeatedly, I will briefly review the way they are used usually and then define how I will use them. They are classically defined in three ways:

Intentionality. This refers to the “aboutness” of mental states: mental states have an intrinsic relationship to things in the world, and they would not exist if not

completed by something other than themselves (this was defined by Brentano; see also Searle, 1983). For example, Carolina's desire to eat chocolate relates to Carolina and chocolate, which are an objective organism and item in the world, eating chocolate is an action which also exists in the world, and Carolina eating chocolate is (possibly) a future state of the world. Mental states such as believing, desiring, and others listed above all relate to states of the world; in that, they are said to be intentional. But this is not the way in which I will use the word "intentional", I will use it as relating to intentions. See below.

Intentions in actions. These are specific intentional states of mind that, unlike beliefs, judgments, hopes, desires or fears, play a clear role in the etiology of actions. While all intentions are intentional, not all intentional states are intentions. For example, "Carolina intends to eat chocolate" is quite likely to play a role in making her reach for a chocolate bar, whereas "Carolina believes that there is chocolate in the cupboard" does not, by itself, explain her act of taking the chocolate bar. When referring to intentions in this thesis, I will be referring to intentions in actions (Note: I will not make the difference between 'intention in action' and 'intention to act', as defined in Searle, 1983).

Communicative intention. A speech act has two types of intention: the informative intention (the desire to inform the listener of something), and the communicative intention, which is the intention to "make mutually manifest to audience and speaker the informative intention of the speaker" (Sperber & Wilson, 1986). The communicative intention only exists when dealing with other humans. Although this type of intention is certainly relevant for Theory-of-Mind, this will be only interesting at higher levels of ToM than those examined in this thesis. I will therefore not use this definition.

So I will use intention to mean intention in an action, and intentional as the adjective for intention, meaning something that has intentions. Intentional actions are therefore opposed to accidental actions, which are defined by not being directed by an intention.

### Goals and actions

Goals in actions are defined differently by different authors, as discussed in a recent study by Koski and colleagues (Koski et al., 2002). The goal-directed theory of imitation (Bekkering, Wohlschlaeger, & Gattis, 2002) defines goals as physical objects that can be targets for reaching and grasping movements, but also as a representation of the goal in neural codes, in a “functional mechanism necessary to initiate an imitative action” (Koski et al., 2002). Tomasello separates action goals from the means to achieve them (Tomasello, 1999). Travis defines a goal as a “mental state representing a desired state of affairs in the world” (Travis, 1997). In Meltzoff and Moore’s active intermodal mapping theory, an infant’s goal is to match the relations between their own body parts with those of the observed model, again a functional definition (Meltzoff & Moore, 1997). Dickinson and Balleine consider as goal-directed an action mediated by 1) instrumental knowledge of the causal relationship between the action and the outcome or goal, and 2) the current goal or incentive value of the outcome (Dickinson & Balleine, 2000). In summary, a goal can be defined as an object, the outcome of an action or a representation (mental or neuronal) of either the object or the desired end-state of an action. In this thesis, I will use the word “goal” only as the physical object toward which an action is directed, not a mental representation, by contrast to an intention. Therefore a goal-directed action is akin to an object-directed action, and I will sometimes use the second terminology to

make sure that what I mean is clear. However, some authors I cite will use a different definition, which I hope will appear clearly enough.

Depending on the definition of goal used, some actions might also be considered non-goal-directed. When someone mimes or pretends to perform an action, her or his movements are not directed towards a present object, and are thought to be based on a stored representation of the target object rather than perceptual input (Goodale, Jakobson, & Keillor, 1994). The following data suggest that mimed movements are controlled by different processes from object-directed actions. Mimed movements can have different kinematics from object-directed or actual movements (Goodale, Jakobson, & Keillor, 1994) and are much less affected by visual illusions (Westwood, Chapman, & Roy, 2000) than are actual movements (Aglioti, DeSouza, & Goodale, 1995; Haffenden & Goodale, 1998; Ellis, Flanagan, & Lederman, 1999; Flanagan & Beltzner, 2000). Together with the case of a patient with ventral visual stream lesions (James, Culham, Humphrey, Milner, & Goodale, 2003) who is unable to perform mimed movements based on perceptual cues (Goodale, Jakobson, & Keillor, 1994), this suggests that neural structures underlying control of mimed actions could be located in the ventral rather than the dorsal visual stream (Milner & Goodale, 1995; Westwood, Chapman, & Roy, 2000). Thus, the fact that actual and mimed movements differ based on their object-directedness, and maybe even goal-directedness, might determine their control by different visual streams. As will be mentioned below, observation of non-object-directed actions does not activate neurons responding both during execution and observation of human actions (the so-called “Mirror Neurons”, see below). These data will serve as the basis for an imaging experiment of action observation, Experiment 3 in this thesis. Also, I will use the terminology “non-object-

directed” as being similar to “non-goal-directed”, as for goal-directed and object-directed actions.

While most goal-directed actions imply that the agent has an intention, this does not need to be the case: an accidental action can have a goal, but does not have an intention. The opposite is true as well: an intentional action can be non-goal-directed: if one considers that a mimed movement is not a goal-directed action, it can nevertheless be executed on the basis of an intention.

### Agents and animacy

I will use the word agent as referring to an entity or an organism that “does something”, i.e. performs an action. Agents can perform goal-directed or (more rarely) non-goal-directed actions, which might be directed by intentions and other mental states, or not. Sometimes, we would explain actions of agents by mentalistic terms including intentions, even if we know that the agent has no mental states (ex: a computer “wants to connect to the web”).

Animacy I will use to mean “being alive”. A process I will study in this thesis is the identification of animate entities in the world. Animate things often perform actions, and are therefore often agents, and as will be discussed in more detail, performing a goal-directed action is a good cue for attributing animacy to an entity.

## 1.1. A brief history of Theory-of-Mind

### Does the chimpanzee have a Theory-of-Mind?

Premack and Woodruff published an article in 1978 discussing the existence of Theory-of-Mind in chimpanzees, which sparked off discussions about Theory-of-Mind in other fields as well (Premack & Woodruff, 1978). The authors describe an interaction with a chimpanzee called Sarah, in which Sarah watched a person trying to solve practical problems, such as trying to reach bananas hanging from the ceiling. Sarah had to choose the most plausible continuation of the person's action among various options, and chose a continuation showing the person stacking boxes to reach for the bananas. The authors argued that Sarah understood what the person was trying to do and used this information to predict the next actions of the person. They suggested that Sarah was able to understand a person's actions based on their desires and goals, an essential component for Theory-of-Mind.

But another explanation for Sarah's behaviour can be put forward. She might just have shown what *she* would do in the person's situation (stack boxes to reach the banana). She might just have projected her behaviour onto the situation of the person on the screen, and might therefore just have shown what she would do and not what she thought the person wanted to do. Her behaviour therefore does not need to involve understanding or manipulating mental states, but just understanding of the physical situation a person is in. In a comment on Premack and Woodruff's article, Bennett, Dennett and Harman suggested that a convincing demonstration for mentalising in animals or humans<sup>1</sup> would be to show that one can understand that someone has a

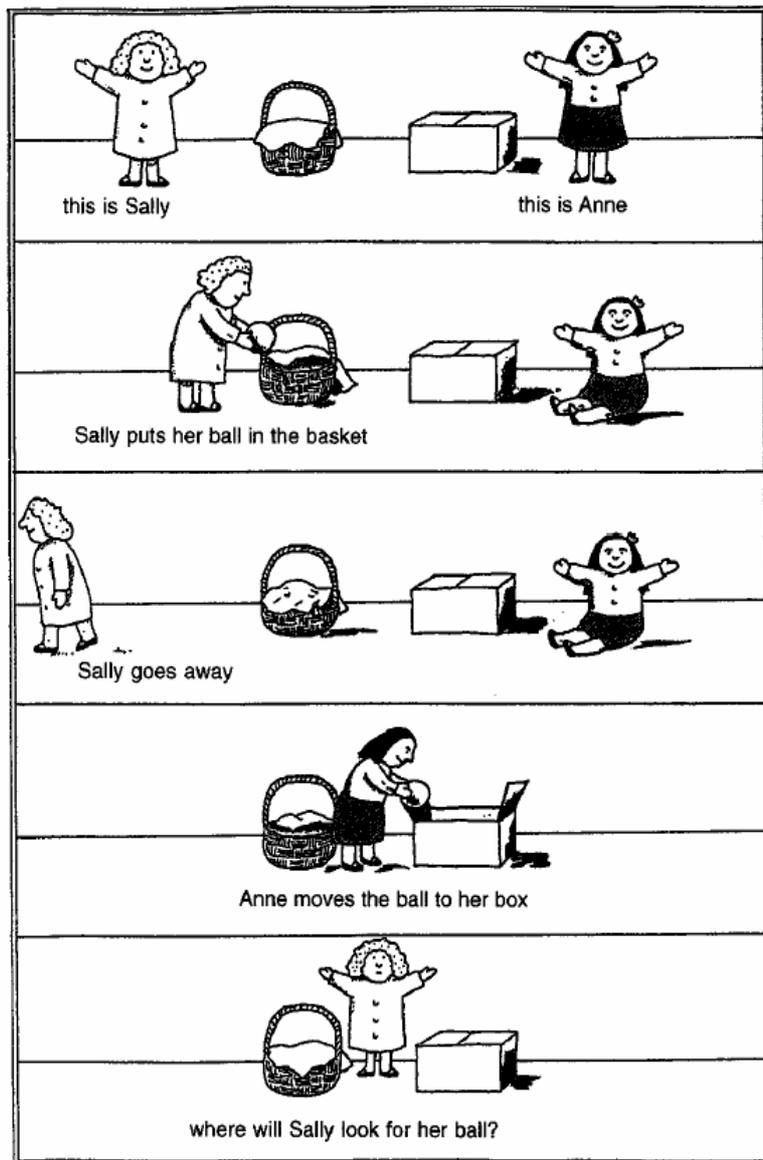
---

<sup>1</sup> NOTE: As the question of whether or not only humans are capable of mentalising is not directly relevant to the present thesis, I will only discuss behavioural data from studies with humans from now on, and only refer to animal work regarding single-cell neurophysiological recording data.

false belief about the world, and deduct his actions from it (Bennett, 1978; Dennett, 1978; Harman, 1978).

### The False-belief task

Following these deliberations, Wimmer and Perner created the first false-belief task in 1983, now a classic in the study of theory-of-mind and mentalising (Wimmer & Perner, 1983). In their task (also known as the Sally-Anne task, see Figure Intro.1), the experimenter tells a child a story with two characters in a kitchen and some chocolate. While his mum is watching, Maxi, the first character, puts the chocolate in one of the cupboards and leaves the scene. While Maxi is away, his mum takes the chocolate from one cupboard and puts it in another cupboard. Then Maxi comes back and the experimenter asks the observing child: "Where will Maxi look for the chocolate?" To pass the test, the child must say that Maxi will look inside the empty cupboard where the chocolate was but no longer is. This implies that the child understood that because Maxi was not present when his mum switched the chocolate to the other cupboard, it cannot know where the chocolate really is and will make a mistake. Children usually pass this test around the age of 4. Younger children, who do not understand that Maxi needs to witness his mum's actions to know what happens to the chocolate, say that Maxi will look for the chocolate in the second cupboard, because that is where the chocolate really is, and as the child knows where the chocolate is, so does Maxi. These findings have since been refined and replicated many times: Wellman, Cross and Watson performed a meta-analysis of 178 ToM studies in 2001 and found confirming evidence of robust changes with age during preschool years (Wellman, Cross, & Watson, 2001).



**Figure Intro.1. The Sally-Anne task.**

Originally by Wimmer and Perner (1983). Drawing by Axel Scheffler, in Frith, 2001.

This task has since been tested in children with various developmental disorders, and an interesting finding has emerged: children with autism with a verbal mental age of 4 fail the test, but children with Down syndrome with a verbal mental age of 4 perform as well as healthy 4-year olds (Baron-Cohen, Leslie, & Frith, 1985; Frith, 2001). This and other findings led to the hypothesis that children with autism suffer

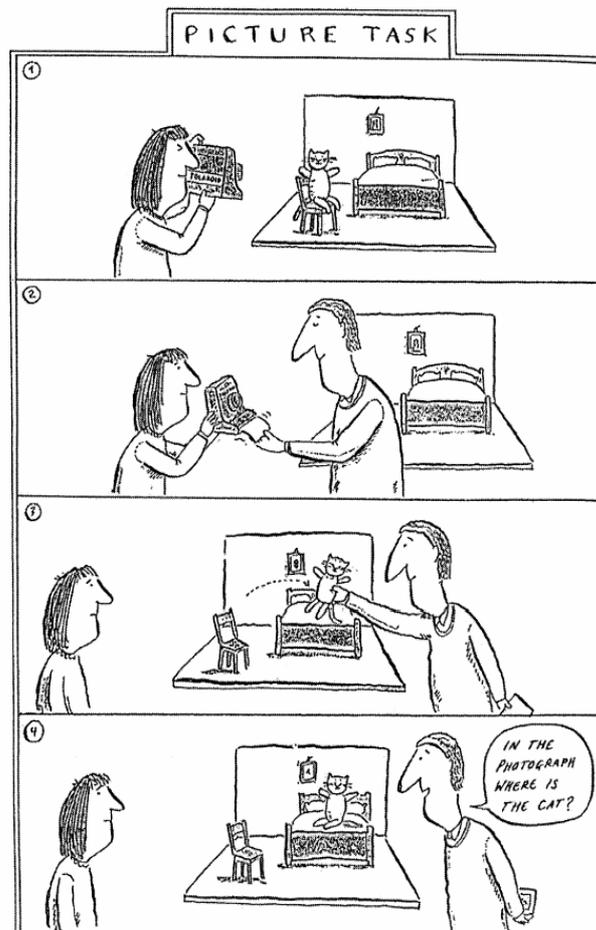
from "mind-blindness", a specific deficit in "reading other people's minds" (Baron-Cohen, 1995; for a recent review, see: Frith, 2001).

Gopnik and Astington developed a variant of the false-belief task to test whether the age of false belief understanding is similar when attributed to the self as to others: the Smarties test (Gopnik & Astington, 1988). In their test, a child is shown a tube of candies with "Smarties" written on it. The experimenter asks the child what he thinks is in it. The child, of course, answers: "Smarties!" The experimenter then opens the tube and shows the child that there are no Smarties in the tube, but pencils instead. Now the child is asked the decisive question: "When you saw the box first, what did you think was in it?" To pass the test, the child has to remember that at the time, it didn't know that there were pencils in the tube, but thought it was Smarties. Confirming previous findings, this usually happens at age 4-5, whereas 3-year-olds fail the test and say "Pencils".

#### A good control

To refine findings of a deficit in the Sally-Anne task and show that it is really due to a deficit in attribution of mental states and not of understanding of complex situations, Leslie and Thaiss compared the False Photograph Task with the Sally-Anne task in autistic and healthy children (Leslie & Thaiss, 1992). In this test (Figure Intro.2), the child is shown a teddy bear sitting on a chair. The experimenter takes a Polaroid photograph of the teddy on the chair in front of the child, and puts the photograph in his pocket. He then moves the teddy bear onto a bed beside the chair. Then he asks the child: "On the photograph, is the teddy on the chair or the bed?" To pass the test, the child has to answer "On the chair", as the photograph will not change even if the reality it used to represent changes. Autistic children that fail the Sally-

Anne Task pass the False Photograph Task, whereas normally developing children succeed in both tasks, even though the False Photograph Task appears more difficult than the Sally-Anne Task. This suggests that children with autism have a specific problem with an aspect of the Sally-Anne Task, most probably with the attribution of mental states (Frith, 2001).



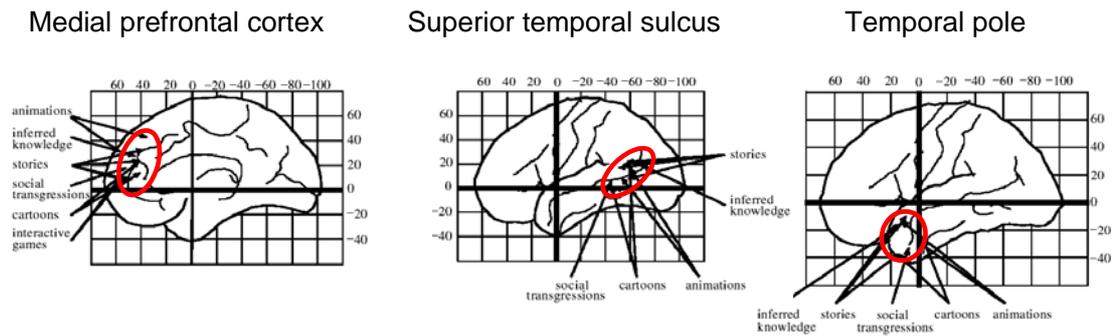
**Figure Intro.2. The Picture task.**

By Leslie and Thaiss (1992), also known as the (False) Photograph task. From Frith, 2001.

## 1.2 Possible neural substrates for ToM

The first two studies attempting to identify neural structures activated during attribution of mental states were performed by Fletcher and colleagues (Fletcher et al., 1995) and Goel and colleagues (Goel, Grafman, & Hallett, 1995) almost 10 years ago. Fletcher and colleagues asked volunteers to explain the behaviour of people using pretence and deception in one series of short stories, and presented them with stories in which the mental states of the characters did not play a role as a control. Activity in medial prefrontal cortex, posterior cingulate cortex and right temporo-parietal junction was increased during mentalising. Goel and colleagues asked healthy volunteers to judge whether someone like Christopher Columbus, living in the 15th century, would have known the use of a series of objects. Medial prefrontal cortex and left temporo-parietal junction were more activated during this task than during memory retrieval and simple inferencing.

Since these original studies, a number of fMRI and PET studies have followed, using verbal and non-verbal, on- and off-line tasks with a variety of media (pictures, stories, cartoons, animations, games...). In reviews of such studies, Chris and Uta Frith and colleagues (Frith & Frith, 1999; Gallagher & Frith, 2003; Frith & Frith, 2003) showed that three regions appear to show increased activation during mentalising (Figure Intro.3): the posterior superior temporal gyrus or temporo-parietal junction, the temporal poles and the medial prefrontal cortex.



**Figure Intro.3. Three areas generally associated with mentalising.**

FMRI and PET studies using diverse mentalising tasks, from the meta-analysis by Frith and Frith, 2003, activate the medial prefrontal cortex, the superior temporal sulcus and the temporal pole. Displayed are data from 10 studies, with tasks including inferred knowledge (Goel, Grafman, & Hallett, 1995), social transgressions (Berthoz, Armony, Blair, & Dolan, 2002), cartoons (Brunet, Sarfati, Hardy-Bayle, & Decety, 2000), interactive games (Gallagher, Jack, Roepstorff, & Frith, 2002), animations (Castelli, Happe, Frith, & Frith, 2000; Schultz et al., 2003) and stories (Fletcher et al., 1995; Gallagher et al., 2000; Vogeley et al., 2001; Ferstl & von Cramon, 2002).

### The medial prefrontal cortex

It is possible that the medial prefrontal cortex (mPFC, also referred to as anterior paracingulate cortex) is the region most specifically associated with mentalising, and that the pSTS and temporal poles reflect activities that aid mentalising and from which mentalising possibly developed (Gallagher & Frith, 2003). In their latest review (Frith & Frith, 2003), Uta and Chris Frith discuss the currently available information on the mPFC. This region has direct connections to TMPp and pSTS (Bachevalier, Meunier, Lu, & Ungerleider, 1997). It is the most anterior part of the paracingulate (BA 32), partly overlapping but mostly anterior to the anterior rostral cingulate area (RCZa) as defined by Picard and Strick in their review of premotor areas (Picard & Strick, 2001), and overlapping with the (functionally defined) “emotional” part of the ACC (Bush, Luu, & Posner, 2000). The paracingulate cortex is often considered part of ACC (BA 24, 25 and 33), but the mPFC has been described cytoarchitectonically as a cingulo-frontal transition area, different from the ACC proper (Devinsky,

Morrell, & Vogt, 1995). Frith and Frith also mention a specific particularity of the ACC which might correlate with the mentalising aptitude: the presence of spindle cells, a type of neurons found only in apes and hominids (Nimchinsky et al., 1999), appearing at the age of 4 months in humans (Allman, Hakeem, Erwin, Nimchinsky, & Hof, 2001). If the ACC has undergone recent evolutionary changes and is involved in mentalising, it could explain why only humans can perform mentalising fully, while apes are at best limited to representing very simple psychological states of others, such as seeing (Tomasello, Call, & Hare, 2003; Povinelli & Vonk, 2003).

Functionally also, cognitive processes associated with the anterior paracingulate cortex hint at a connection to mentalising. This area is activated by attention to diverse sorts of events and sensations, such as emotion, pain, tickling and irrelevant thoughts, and could therefore store 2<sup>nd</sup>-order representations (i.e. decoupled from the "physical world") of these sensations and events, used for attention and report. Representing mental states of the self or others appears to be a similar type of process and could therefore be supported by the same neural structure. Studies of autobiographical memory also yielded activation in the mPFC, and they might implicate a common component to mentalising by calling upon representations of the self. A simpler explanation would be that mentalising tasks might be forms of complex problem-solving. While there are connections between mentalising and executive functions in the development of cognitive functions, they are still not fully understood (Perner & Lang, 1999), and imaging studies of executive functions suggest that different regions are involved in these processes (Frith & Frith, 2003).

Neuropsychological studies in patients with frontal lesions showed that some of these patients have deficits in Theory-of-Mind tasks. One study of patients with frontal variant frontotemporal dementia showed that these patients have deficits in

first-order and second-order false belief tasks, faux pas detection and a mentalising task based on pictures of the eye region of the face (Baron-Cohen, Jolliffe, Mortimore, & Robertson, 1997), while having no problem in control tasks testing memory and general comprehension. Alzheimer patients only had deficits in the second-order false belief task (Gregory et al., 2002). Interestingly, damage to ventromedial frontal cortex was associated with deficit in the ToM tasks, while deficits in executive function did not correlate with deficits in ToM. In another study comparing patients with unilateral left and right frontal lesions to matched controls, patients of both groups were found to have deficits in first- and second-order ToM tasks (Rowe, Bullock, Polkey, & Morris, 2001). No effect of laterality or lesion size was found. Comparing patients with orbitofrontal and dorsolateral prefrontal lesions showed that only the former performed like individuals with Asperger's syndrome, with deficits in recognizing social faux pas; the latter had problems only when task demands on working memory were high (Stone, Baron-Cohen, & Knight, 1998). In another study, patients with medial frontal lesions showed impaired detection of deception; the authors suggest that this deficit may depend on connections between the medial frontal lobe and the amygdala (Stuss, Gallup, & Alexander, 2001). But in a very recent study, a patient with symmetric medial prefrontal lesions showed no impairment on a range of ToM tasks (Bird, Castelli, Malik, Frith, & Husain, 2004).

### The temporal poles

In their review mentioned above, Frith and Frith (Frith & Frith, 2003) also discuss explanations for the association between mentalising and the temporal poles, particularly in the left hemisphere. The authors mention that the anterior temporal lobe has been considered a potential convergence point for all sensory modalities

(Moran, Mufson, & Mesulam, 1985). Reviewing neuroimaging data available on the area, they mention that activation increases in the temporal poles, particularly in the left hemisphere, are found in language paradigms, such as the comparison of sentences to word strings or unrelated sentence strings, or the comparison between more coherent vs. less coherent narratives, or in semantic decision tasks. Another process associated with this brain area is retrieval of autobiographic memory, retrieval of emotional context in single-word recognition, and the recognition of familiar faces, scenes and voices. Episodic memory in itself might be useful for mentalising: remembering past interactions with a person might help us to recall what we said to them at the last encounter, what their likely attitude towards us could be, or which mental states were associated with a particular behaviour they exhibited (Gallagher & Frith, 2003).

Frith and Frith (Frith & Frith, 2003) suggest that an overall function of the temporal poles might be involved in generating a wider semantic and emotional context for material being processed. A part of this wider semantic context that could be useful for mentalising are the so-called scripts: the habitual sequences of events and activities that take place in a given setting and time. One popular example is the restaurant script: we choose a restaurant, then expect what we will find on the menu, then that we will order, then taste the wine, then receive and enjoy the food, then pay the bill. Such scripts could be useful for understanding mental states of other people by matching their behaviour with possible scripts for their situation, and noticing the deviations. These scripts are gradually lost in patients with semantic dementia, who show atrophy in the anterior temporal lobes, particularly in the left hemisphere.

### The posterior part of the superior temporal sulcus

A recent study suggests that the pSTS is not only involved in mentalising tasks but is also the region that shows the most consistent activation during ToM stories, and no activation during control stories about a false photograph, mechanical inference, human actions and other controls (Saxe & Kanwisher, 2003). The authors of this study therefore conclude that it is the pSTS that is the most important region for mentalising. Arguing that pSTS is necessary for mentalising, a recent neuropsychological study showed that three patients with lesions of the left temporo-parietal junction including the posterior STS all had deficits in a false-belief task, while not all showed deficits in story-based and video-based control tasks involving memory, counterfactual reasoning, reality checking, or response inhibition (Samson, Apperly, Chiavarino, & Humphreys, 2004). Other processes associated with the posterior part of the STS will be discussed in greater detail later on.

### Other regions

Other researchers suggest that there might not be a specific neural circuit for mentalising, but that this ability could instead rely on neural structures associated with other cognitive processes (Siegal & Varley, 2002). Candidate processes and brain structures are executive functioning in the frontal lobes, visuo-spatial processing (particularly identification of animate entities in higher-order visual cortex, see Frith & Frith, 1999; Frith, 2001), language abilities (particularly grammatical abilities) in the left hemisphere or emotional processing in the amygdala. Although language is helpful in solving ToM tasks, disorders of language or grammar only (either acquired from a brain lesion or present during development) do not eliminate mentalising abilities (Siegal & Varley, 2002). Patients with frontal lesions (specially right) have

difficulties with deception tasks, but deficits in tests of executive functioning do not correlate with deficits in ToM in patients with prefrontal brain lesions (Rowe, Bullock, Polkey, & Morris, 2001). Brain regions associated with social cognition, often termed the "social brain" (orbitofrontal cortex, the amygdala and the superior temporal gyrus, see Brothers, 1990; Adolphs, 1999; Adolphs, 2003) are certainly important and probably necessary for the emergence of ToM, but these regions are probably not sufficient for this cognitive process once it is developed. Some researchers argue that the "social brain" is the core component for ToM, but others argue that it is rather the above-mentioned triad, identified in functional imaging studies, that is the network underlying ToM. Also important for the development of ToM are conversational skills and access to interactions with other human beings, with which a child gets exposed to mental states of other people.

## **Chapter 2. Cognitive processes in ToM**

In this section I will describe what could happen during mentalising, then review some classic and more recent cognitive models attempting to explain the mentalising process.

### **2.1. Representing mental states**

#### Intentional relations

One theory relevant to mentalising that has been proposed is Barresi and Moore's framework of social understanding based on representations of intentional relations (Barresi & Moore, 1996). They define an intentional relation as a person's activity or state (emotional or other) related to an object or state of the world<sup>2</sup> (example: John likes Mary, the tiger wants to eat the antelope, Maiko drinks tea). Both self and others have intentional relations between them and objects or states of the world. The core of their framework is a schema that generates representations of intentional relations that apply to both the self and the other. For example, if Maiko and I are talking about a cup of tea in front of us, we both create a representation of each of our relation to the cup of tea. To create such a representation, the schema needs to have access to information about the intentional relations of both the self and the other (this information could be visual, auditory or of some other modality). While such information about the self is directly accessible (I know what the cup of tea looks like to me), information about the intentional relations of the other would be acquired by emotional empathy, joint attention (adequate for the case of the cup of tea), and particularly goal-directed imitation. Also, attention must be sufficient for processing

---

<sup>2</sup> Their use of "intentional" is much like Brentano's definition of it, and not always related to intentions. They therefore use intentional differently than I do in the rest of this thesis.

both types of information at the same time. Once this common representation is created, it can be applied to both self and other and enables social understanding.

This framework emphasizes the link between perception of goals and understanding of intentional relations (culminating in mentalising), and in this, is relatively close to the relationship between actions, goals and intentions which will also be defended in this thesis.

### Simulation Theory and Theory Theory

In philosophy of mind, two classic accounts of mentalising have been proposed: 'simulation theory' and 'theory theory'.

Simulation theory explains that people use their own mental mechanisms to understand and predict the behaviour of others, by generating ourselves actions and processes similar to the other (Goldman, 1993). Observers would represent other peoples' mental states by adopting their perspective or the context of their actions, by "putting themselves in the other person's shoes". Simulation theory would therefore rely on mechanisms that can be triggered both by the observed person performing an action and the observer performing an action himself. Such a common representation for both perception and execution of an action will be described below, and a possible neurophysiological basis for such mechanisms might have been found with the "mirror neurons" also described below<sup>3</sup> (Gallese & Goldman, 1998).

In contrast, 'theory theory' proposes that people acquire and use a commonsense knowledge of minds and mental states, and from there develop a set of causal and explanatory laws, akin to physical laws used to explain the behaviour of matter

---

<sup>3</sup> Although, as noticed by Pierre Jacob (in a presentation of work in progress), mirror neurons code for actions, there is no evidence that they also code for the mental states behind the action. Therefore, although they might be necessary for many applications of mental state simulation, they are certainly not sufficient for it in that they would be similarly activated by a given action performed with different intentions.

(Gopnik, 1993). Therefore, to understand the mental states of another, I would enter into my database the observed behaviour and on its basis infer which mental state could explain the behaviour. To develop this theory, one would act like a child scientist, performing experiments with oneself and others to understand the mind.

But maybe we use both?

The explanation of what happens during mentalising that I will try to defend is as follows. To attribute mental states to another agent, the observer relies mainly on the behaviour of this agent. Using his own knowledge of mental states associated with the observed behaviour and the information he has about the other agent (the context of the person's action), he will then try to select the mental state that is most likely to be present in the mind of the other agent. To test and refine his selection, the observer can make predictions about the behaviour of the other agent based on the mental state he has chosen and compare them to the real behaviour, then modify his choice of a mental state. In this, he will act like Gopnik's child scientist (see above), and compile a database of mental states associated with different actions and contexts. I would not like to exclude however the possibility that the observer's action system is directly activated by the actions of another, which would call up mental states associated with these actions when performed himself, which would be useful to compare the action-mental state associations of the other and the self (For example: I cry when I am sad, but she cries when she is happy. Same action, but different mental state. Good to know when I observe her the next time!). Also, in different situations simulation of the other's action might help (when one can put oneself in the other's shoes) but in others it will be necessary to theorise (how can I understand a mother's actions when

she has lost a child, when I have never experienced this?). I believe that I am therefore following completely neither Theory Theory nor Simulation Theory.

In my view, the basic processes for the accomplishment of mentalising are: 1) the identification of potential agents or detection of agency, 2) observation and recognition / understanding of their actions, 3) knowledge of potential mental states that the other agent could have, 4) the correct association between the mental state and the behaviour and 5) the evaluation and testing of the mental state by comparison between expected and actual behaviour of the other agent. Different models of how these processes interact will be discussed below.

## **2.2. Recent cognitive models for ToM**

Recently, more detailed models based on control systems used by engineers and on neuroscientific data have been proposed.

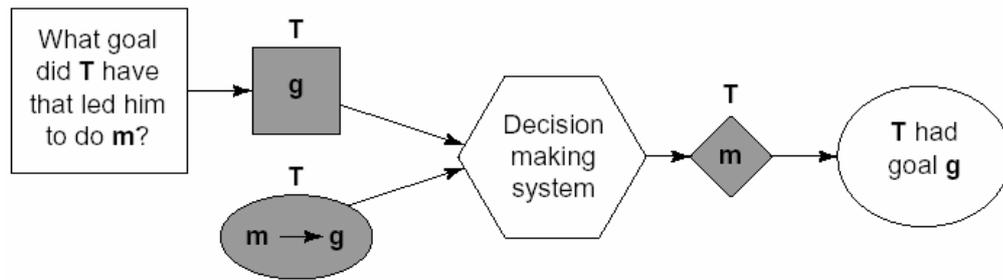
### Understanding others' actions with one's own action system

Based on previous psychophysical and neuroimaging work, Blakemore and Decety (Blakemore & Decety, 2001) have suggested that humans automatically infer intentions from observed actions of other people and other types of biological motion. Inferring intentions of others by observing their actions could be a basic form of theory of mind and the basis of higher levels of understanding of others' minds. They propose that such a mechanism for intention inference might be based on the system labelling the consequences of one's own actions from one's own intentions. This system could be implemented as a forward model predicting the sensory consequences from a given intention, on the basis of a store of sensory predictions associated with actions of the self. Understanding others' intentions could be based on simulating (covertly imitating) the observed action and estimating the intentions of the actor on the basis of one's own intentions. The authors propose the following sequence of events: "the observed sensory consequences (of another person's actions) would be mapped onto stored sensory predictions (of the sensory consequences of one's own actions). These stored representations could then be used to estimate the motor commands and intentions that would normally precede such an action. This could be achieved by automatically and unconsciously simulating the observed action and estimating what our intentions would be if we produced the same action within the same context" (Blakemore & Decety, 2001). In their article, Blakemore and Decety suggest that the forward model could use efference copy signals created in

parallel with the motor commands of an action to predict the sensory consequences from that action (psychological evidence is provided by the example of tickling: Weiskrantz, Elliott, & Darlington, 1971; Blakemore, Frith, & Wolpert, 1999), but they do not explain how this mechanism could be reversed to retrieve the intentions associated with the mapped sensory consequences of the observed actor's actions.

#### Internal models for prediction and control

In a recent article, Gallese and Goldman (Gallese & Goldman, 1998) detail the "simulation routine" that represents the mechanism of Simulation Theory (Figure Intro.4). Based on Goldman's model (Goldman, 1989), they propose that while an observer is watching another person, a "pretend belief and desire generator" would generate potential mental states that the observed person might have, on the basis of all information available about the observed person, such as previous actions or contextual information. This mental state would be fed into a "decision-making system", which would predict the appropriate behaviour of the observed person corresponding to the mental state. The decision-making system used for this purpose would be the same as used by the observer himself to perform everyday actions. This predicted behaviour can then be used to anticipate the next actions of the other and interact appropriately with him.



**Figure Intro.4. Gallese and Goldman’s “retrodictive simulation” routine.**

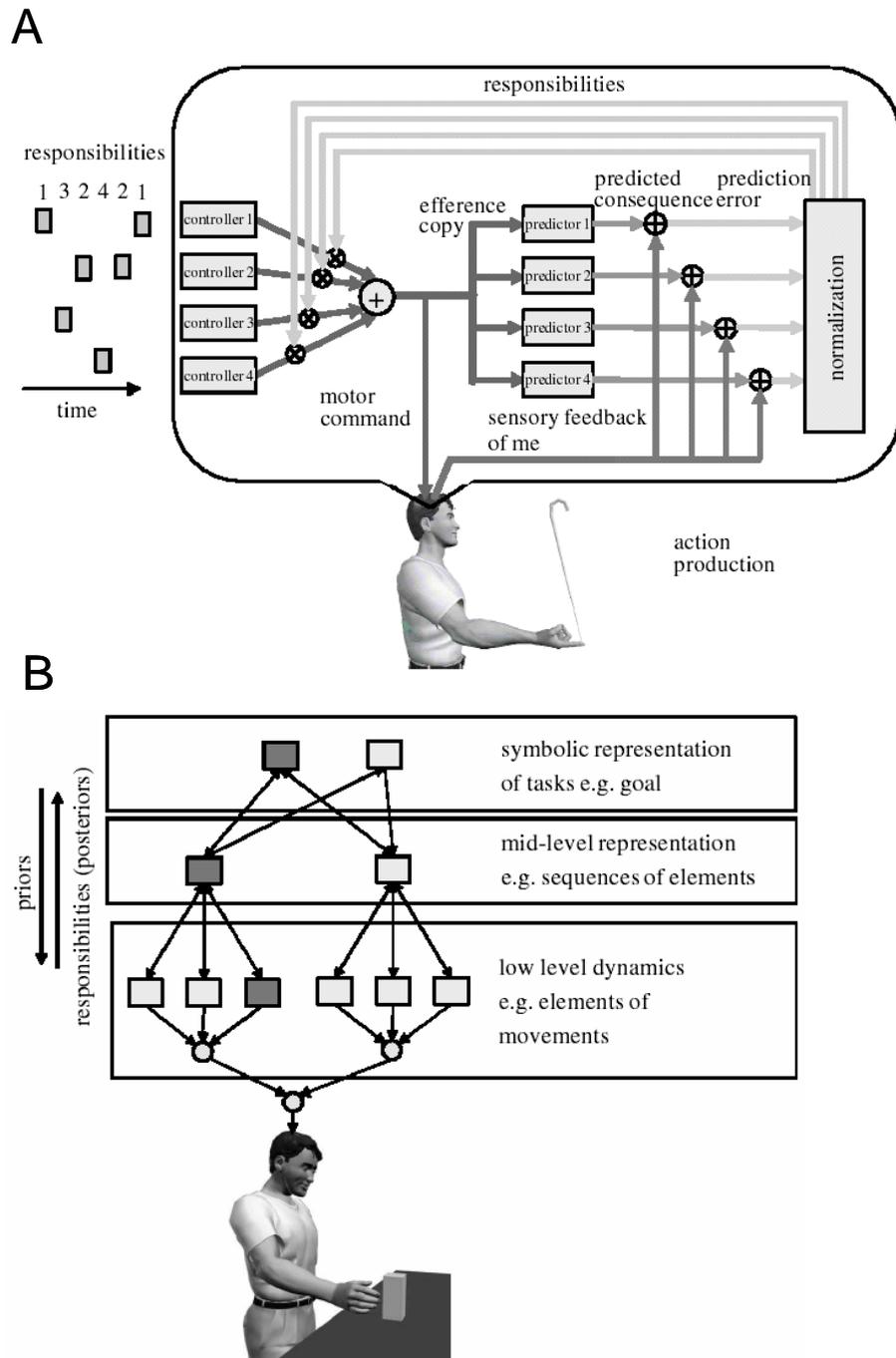
The authors describe it as follows: “After observing the target agent (**T**) perform action **m**, the attributor uses simulation to test whether goal **g** would have fitted with the choice of **m**. Goal **g** is re-created and fed into his decision-making system, which does output **m**.” From Gallese and Goldman (1998). Which goal (**g**) is introduced into the model depends on information about the observed person and contextual information.

The sort of system that might create these predicted behaviours resembles a type of control systems used in engineering involving internal models. Such models have been used to explain human sensorimotor learning (for a review, see Wolpert, Ghahramani, & Flanagan, 2001) and are thought to be implemented in the cerebellum (Shidara, Kawano, Gomi, & Kawato, 1993). There are two types of internal models with complementary roles: forward models which act as predictors, and inverse models which represent controllers. A forward model works in the following way: based on the current state of a given system and the action the system is going to perform, the forward model will predict the next state of the system. An inverse model calculates the action required for a particular goal (or desired state) given the current state of the system. A useful control mechanism could include an inverse model to select the best action for a given goal combined with a forward model to predict the effects of the action on the state of the system (Wolpert, Ghahramani, & Flanagan, 2001). The difference between the predicted and the desired state of the system can be used to select the best action, before it is even initiated. Once the best action has been chosen and started, the difference between the predicted and the

actual state of the system would be fed back into the inverse model, modifying or fine-tuning the chosen action to minimise the difference and finally achieve the goal.

To make such a control system more flexible for various sensorimotor contexts, Wolpert and Kawato have proposed to use multiple parallel systems: the modular selection and identification for control model, or MOSAIC model (Wolpert & Kawato, 1998; Haruno, Wolpert, & Kawato, 2001, see Figure Intro.11 A). In this system, prior information about the context is given by sensory information such as visual input. Given the state of the system and the goal to be achieved, the parallel inverse models corresponding best to the context then calculate actions corresponding to the specified goal. The predicted effects of each action on the state of the system are compared with the desired state of the system. The inverse model that proposed the action which results in the closest state to the desired state receives the greatest confidence rating, and the action is executed. The actual state of the system is then compared to the desired state. When the inverse model is performing well and has selected the right context and action, there is no difference between the desired and the actual state of the system, the goal is reached.

A recent development of this architecture was to propose a hierarchical version of MOSAIC: the HMOSAIC (Wolpert, Doya, & Kawato, 2003, see Figure Intro.11 B). This structure is composed of higher-level MOSAICs that control lower-level ones. The higher level systems deal with action goals and intentions, intermediary ones with action sequences, and the lowest with the actions themselves. Different pathways between the higher and the lower models allow a flexible use of actions to achieve a given goal. Such a system might allow control of multiple objects, and a very flexible adaptation to a number of situations.



**Figure Intro.5. MOSAIC (A) and HMOSAIC (B) models.**

From Wolpert, Doya and Kawato (2003).

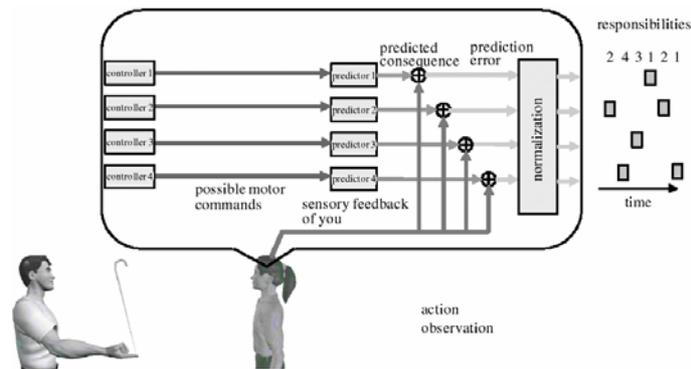
An internal model-based system for understanding mental states

Instead of controlling the effects of actions on physical states of the world, such a system could also be used to control interactions between people. Blakemore and Decety propose that such a control system composed of a forward model coupled to an inverse model could be used to understand other people's mental states (Blakemore

& Decety, 2001). These authors build on the idea of a common representation for perceived and executed actions as postulated in the Ideomotor theory and embodied in the mirror-neuron system. They propose that such a representation would link an action with its sensory consequences such as lifting a glass and the proprioceptive feedback and visual aspect of the action. This representation is itself associated with mental states, such as the desire to quench a thirst. When observing a person lifting a glass, the representation of the sensory consequences of this action would be activated, and used to estimate what the motor commands for this action might have been, and the mental states associated with the action would be retrieved (probably by using an inverse model, although this is not specified by Blakemore and Decety). The system would then attribute this mental state to the observed person: it would suppose that the person is trying to quench his thirst. To test whether this is correct, the mental state would be fed to a forward model, which would calculate the expected next actions of the observed person. The actual actions of the person and the predicted actions would then be compared, if there is no difference, the system would assume to have successfully attributed a mental state to the observed person.

Wolpert, Doya and Kawato (Wolpert, Doya, & Kawato, 2003) have proposed that such a system could be based on a Hierarchical MOSAIC model (Figure Intro.6). As described above, hierarchic MOSAIC models are made of multiple MOSAICs coding different levels of the action, from the actions themselves to the goals and intentions behind them. As it is assumed that two humans have a roughly similar HMOSAIC, observation of another person's actions could activate the observer's own model. An observed movement with a clear goal which is represented in the HMOSAIC of the observer would activate all levels of the observer's system. An action without a clear goal would only activate the lower levels. Activation of the

highest levels would represent understanding of the action up to the level of the goal and intention behind it. As described above, activation of lower levels by higher levels can happen through multiple pathways. This would account for the following observed effect in imitation. Imitating a meaningful action can be performed through different pathways, and is known to be driven by the action's goal (see Section 3.2.1 below). Therefore the imitation could be achieved by making different movements than the original observed movements (Bekkering, Wohlschlaeger, & Gattis, 2002; Gergely, Bekkering, & Kiraly, 2002). But observation of meaningless actions, not activating the higher levels of the observer's HMOSAIC, can only be imitated by performing exactly the same actions as the observed person.



**Figure Intro.6. Application of the MOSAIC model to action observation.**

From Wolpert, Doya and Kawato (2003).

Relating to social interactions, the same idea could apply to all forms of communication. The more similar two people's HMOSAICs are, the easier communication between them would be, as similar representations could be found for both the observer and the observed person.

Another advantage of MOSAIC models is that this system can be used to simulate interactions with other people. As people have similar MOSAICs, a person could test out a particular action on his prediction system to estimate the effect on the

other person. If differences between ones' own MOSAIC and the other person's are known, a separate system for that person might be built, based on one's own system and the known difference. This would refine the simulation and tailor it to a particular person, thus increasing the likelihood of making realistic predictions. Such a system, representing other people's mental states linked to their possible actions, might be the basis for mentalising (Wolpert, Doya, & Kawato, 2003).

### 2.3. An extended model

#### Components

The internal-model based system presented above is very appealing, and potential neural substrates, at least for motor applications of it, might already have been found. I propose to extend the model presented above, to include additional functions I think are necessary for the attribution and manipulation of mental states of other people (Figure Intro.7). Inputs for the model are the visual aspect of the agent and a representation of the desired mental state of the agent. The components of the model are:

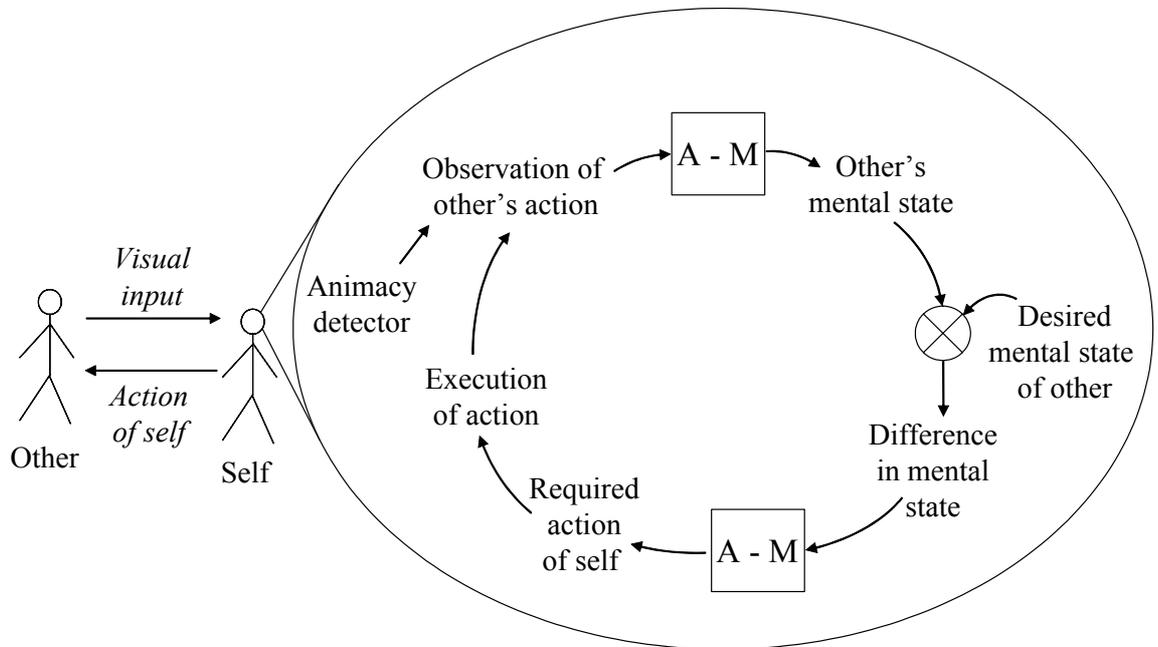
- an agent detector module
- an action observation system
- an action - mental state matching system (Figure Intro.8)
- a comparator between the desired and the actual mental state of the agent
- an action execution system (gives orders to motor system or language system)

The system aims to minimise the differences between the desired and actual mental state of the other, and could very well rely on internal models. Mental states can be desires, intentions, beliefs or any of the others detailed in section 1.1. Actions

can be all kinds of body movements, whether goal-directed or not, including language.

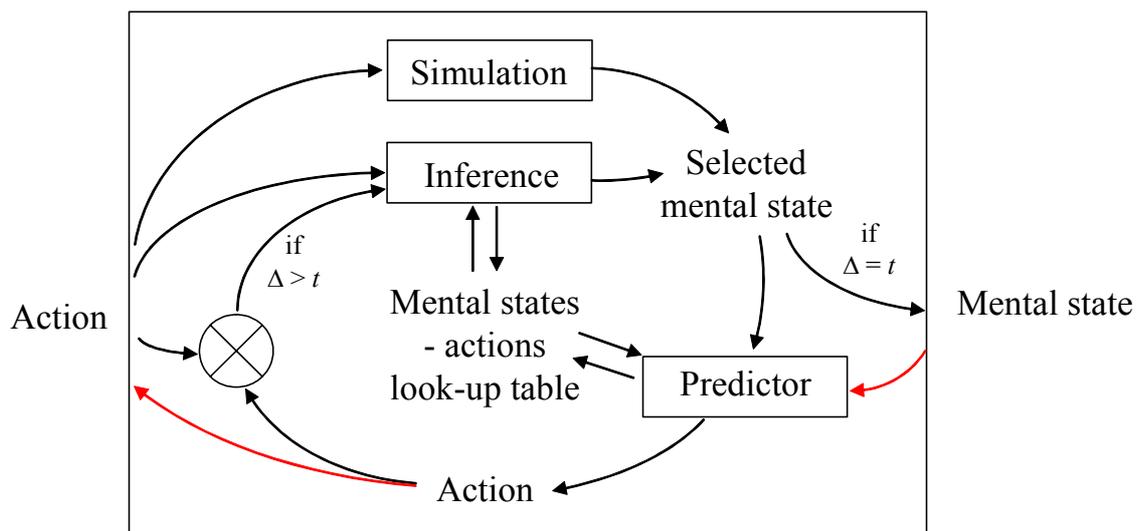
### What it can do

I propose that this model can account for multiple situations involving mentalising, including understanding and manipulating mental states of others. It is triggered whenever the agent detector identifies an object in the world whose characteristics (movement or aspect) make it appear animate. The movements of this object (let's call it an entity) are then observed, and fed into the "A-M" (Action-Mental state matching) module, which identifies a potential mental state responsible for the animate entity's action. The module is detailed below (Figure Intro.8). What happens then is dependent on the identity of the entity, and the plans of the observer for it. If the mental state should not be changed the loop stops here. If the mental state should be changed (i.e. the entity believes something erroneous, for example a dog believing that the mailman is an enemy), then the difference in mental state is fed into the A-M module in the reverse way, and this determines an action that the observer can execute to change the entity's mental state. The new behaviour of the entity is then again fed into the A-M module for a second round, and so on until the mental state is satisfactory. This type of loop might also be used to explain teaching: if a child believes the world is flat (which we can identify during a conversation, which is also a type of action that could be covered by the model), this false belief can be changed by explaining to her the scientific bases of this fact. The way she responds can then be used to ascertain that she has understood. Of course, teaching new actions can be explained in a similar way.



**Figure Intro.7. A new cognitive model of mentalising.**

The A-M module is detailed in Figure Intro.8. See text for explanations.



**Figure Intro.8. Action - Mental state (A-M) matching system.**

Blown up from the model presented in Figure Intro.7. When used to match a mental state to an action, only the black arrows are used. When used to match an action to a mental state, the red arrows are used in addition.  $t$  represents the threshold above which the mental state is accepted. See text for details.

### The A-M module

The A-M module is the heart of the model. It matches mental states and actions, in both directions, and is composed of 1) a simulation mechanism in which the action execution system of the observer is activated almost as if the observer performed the

same action as the observed entity, 2) an inference system that springs into action when the observed action cannot be simulated, it uses commonsense theory and 3) a mental states-actions look-up table acquired through experience, 4) a predictor that will put out an action on the basis of a mental state and 5) a comparator that will find differences between the action corresponding to the mental state and the real action. The predictor could very well be a forward model, and the inference system an inverse model in the sense described above. The system will cycle until the difference between the predicted and the real action has reached a threshold value ( $t$  in Figure Intro.8). One could propose that during learning, all components are improved (maybe except the simulation component), and particularly the catalogue is increased.

### Neural bases

Potential neural structures for many of these components have already been identified. Detection of an agent could be performed by the superior temporal sulcus (STS) and the fusiform gyrus, analysis of its actions could be performed by the STS and a parieto-premotor network, representation of the mental states of others and the self and their comparison could all be performed by the medial prefrontal cortex or the STS. Some of the neural structures in which the A-M system could be encoded have also probably been identified. The action execution system in the model represents the motor system very generally, including language production. These associations will be discussed further in the General Discussion, in light of the results of the experiments of the thesis.

### **Chapter 3. Processes for mentalising studied in this thesis**

In the experiments of this thesis I will mostly deal with the identification of potential agents and with the observation of the actions of other agents. I will try to identify some key variables of animate motion and human action observation, and based on these, try to identify which brain regions deal with these variables.

#### **3.1.1 Animacy and intentionality in moving abstract shapes**

The identification of objects with intentions appears intuitively to be very important for adequate interaction with other living organisms in the environment, and indeed, Premack and Premack (Premack & Premack, 1995) have proposed a theory of human social competence based on the identification of intentional objects. While these authors suggest that humans characterise moving objects as animate and intentional when their movement in space appears self-propelled, which is considered an important cue by many researchers as we will see, other factors such as the aspect of the entity (its morphological characteristics) play an important role as well (Gelman, Durgin, & Kaufman, 1995; Opfer, 2002; Johnson, 2003). As object morphology was not manipulated in the experiments of this thesis, I will only come back to it briefly in the General Discussion.

A systematic study of the necessary movement characteristics for the attribution of animacy is based on the interesting observation that simple disks moving in particular ways can appear animate and even intentional to observing children and adults. This phenomenon has been studied since the early 1900s, and two classic examples of such work are Michotte's book "The perception of Causality" (Michotte, 1946) and an article by Heider and Simmel (Heider & Simmel, 1944). In their study, Heider and Simmel showed healthy volunteers objects (a circle, a large and a small

triangle) that appeared to interact with each other in complex ways. Observers consistently described these objects as chasing or escaping each other, and other intentional terms, and observers even attributed personality traits and emotions to the objects.

As yet, it is not entirely clear which characteristics of an object's movements induce attribution of animacy in observing children or adults. Some characteristics that are necessary for a moving object to appear animate have begun to emerge from behavioural studies in children and adults. I will now review some of these behavioural studies and their approaches.

#### Parameters for animacy and intentionality attribution

A number of studies following from these early results were aimed at determining which parameters of the objects' motion were responsible for inducing an attribution of animacy or intentionality. Bassili (Bassili, 1976) discovered that temporal contingency between the changes in direction of two circles on a dark background made the objects appear to be interacting with each other, whereas the impression of intentionality was correlated with spatial contingencies. But when the experimenter asked his subjects to rate the moving objects for animacy, ratings varied widely, suggesting that different participants used different cues to attribute animacy to the objects. Dittrich and Lea (Dittrich & Lea, 1994) tested subjects with animations consisting of many moving distractor objects (shown as letters on a screen) together with a "target" object (also shown as a letter), which either appeared to be stalking one of the distractors or following it to prevent it from getting lost. The experimenters varied the number of distractors, their movement characteristics and the directness of the targets' behaviour, and subjects had to identify the target object, rate the

purposefulness of its movements, the degree of interactivity with the other objects and the impression of animacy it conveyed. Dittrich and Lea concluded that the impression of animacy depended on both the impression of intentionality of the target (which depended on the spatiotemporal kinematics of its trajectory) and on the impression of interactivity between the target object and its goal (which depended on the relationship between the movements of the target and the goal, that is their relative spatiotemporal kinematics).

#### Animacy in single moving dots

Single moving objects can also appear animate. A common hypothesis, originally proposed by Stewart (unpublished, but cited in: Opfer, 2002), is that this impression arises if the movements of the objects do not respect "Newtonian physics". By this is meant that an object appears to receive new energy, visible as accelerations, stops or sharp turns. This makes the object appear self-propelled. Stewart's results showed that three types of the many different movements she tested produced attributions of animacy: starts from rest, sharp turns to avoid a collision and direct movements towards a goal. Gelman and colleagues (Gelman, Durgin, & Kaufman, 1995) replicated these findings, and showed that animate interpretations appeared when the movements of an object were contingent upon an obstacle or a goal present in their environment. They concluded that attribution of animacy depended on early-developing knowledge of causal principles.

Blythe and colleagues (Blythe, Todd, & Miller, 1999) argued that only a small set of motion cues are necessary to attribute animacy to a moving object and even to identify what intention motivated its movements. They asked 10 pairs of participants to move 2 abstract "bugs" on a computer screen in order to simulate 6 different types

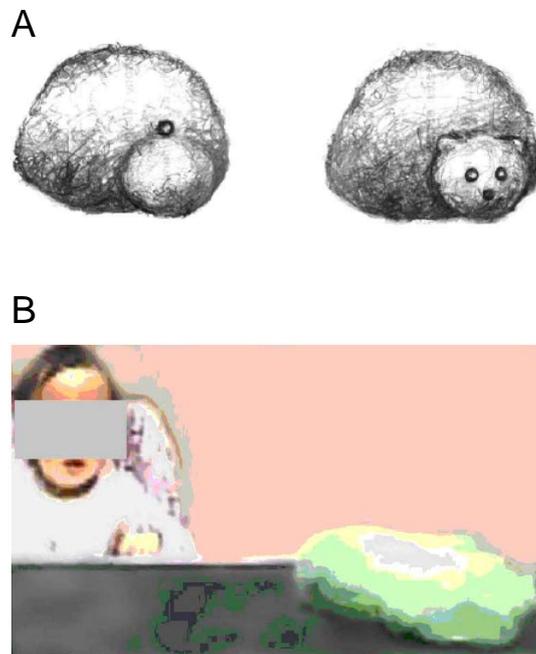
of dyadic interactions (courting, being courted, play, fighting, pursuit, evasion) between them. 10 other participants could correctly identify 49 percent of the 300 recorded dyadic interactions. They then trained a neural network on the same examples, using as inputs for the network 7 motion cues per agent [relative distance between the agents, relative angle between one agent's heading and the other's position, relative heading (angle between their headings), absolute velocity, relative velocity, absolute vorticity (change in heading with respect to background) and relative vorticity]. The network performed even better than human observers in categorising the original 300 examples (82 percent correct) and 300 new, unknown examples (67 percent correct). The authors then showed how 3 other types of algorithm also performed better than the humans on the task, including one algorithm that performed almost as well as the neural network by using on average only 3.6 motion cues out of the 7 available. They suggest that people and other animals might use simple cues such as the 7 they have identified to categorise animate-looking movements.

Scholl, Tremoulet and Feldman (Tremoulet & Feldman, 2000; Scholl & Tremoulet, 2000) have investigated the attribution of animacy to a single object moving on a uniform background, and have shown that changes in direction and changes in speed can induce this percept. Opfer (Opfer, 2002) has shown that goal-directedness in the motion of an object is probably the most important cue for the attribution of animacy to single moving objects.

### Contingency

Despite Bassili's inconclusive results about the correlation between temporal or spatial contingencies with animacy (see above), Johnson showed that attribution of

animacy to an oval, furry object emitting beeping sounds (Figure Intro.9) depended on the contingency between the movements of the object and of another agent, such as a person (Johnson, 2003). Adult observers interpret the behaviour of a contingently moving object in mentalistic terms, and infants treat such objects as if they could perceive, attend, communicate and display goal-directed behaviour. They will follow their attentional orientation and seem to use the object's behaviour to determine its perceptual / attentional orientation and object-directed goals.



**Figure Intro.9. Johnson and colleagues' furry object.**

(A) It embodies many of the proposed cues for mentalistic agents, without being person-like: it can beep and flash a light in response to the experimenter's actions, and one version has facial features. (B) Both the facial features and the behaviour were varied, and the infants followed the orientation by shifting their attention in the same direction as the agent more often than in another direction in conditions where the object had either facial features, displayed contingent behaviour, or both. From Johnson, 2003.

### Identification of actions in terms of goals

At present, the most interesting candidate parameter for the attribution of animacy to moving objects seems to be goal-directed behaviour. Csibra argues that

children understand goal-directed actions and identify goals without needing to identify intentions behind the action (Csibra, 2003). In his article, Csibra defines a goal-directed as an action that is “about” the end-state of that action. Based on his and colleagues’ experiments on observation of moving objects by infants (Gergely, Nadasdy, Csibra, & Biro, 1995; Gergely & Csibra, 2003; Csibra, 2003) and studies of observation of human actions, also in infants (Woodward, 1998; Woodward & Sommerville, 2000), he concludes that infants adopt a teleological stance during the first year of life, before they attribute mental states to goal-directed actions, which happens in the second year (Gergely & Csibra, 1997). He defines the attribution of a teleological stance as "seeking to construe an event in terms of its goals", which is different from the intentional stance where there is attribution of mental states, and implies relating three aspects of a goal-directed movement to each other: behaviour, physical context and end state of the action. Csibra argues that goal-directed movements are automatically recognised by infants (and adults) on the basis of the motion of the agent. This occurs only if the movement is instrumental in obtaining the goal, whether the observed movements are performed by moving disks or humans grasping things. In order to explain the results of the experiment by Meltzoff in which infants imitated a failed action by performing it until the intended end-state (Meltzoff, 1995: only 18-month old children did this, 12-month old only imitated the failure), he argues that both 12-month old and 18-month old infants construed the observed failed attempt as goal-directed, but only 18-month olds are capable of the counterfactual reasoning necessary for ignoring the observed end of the action and replacing it with the inferred goal (Csibra, 2003).

### Self-propelled movements and goal-directedness?

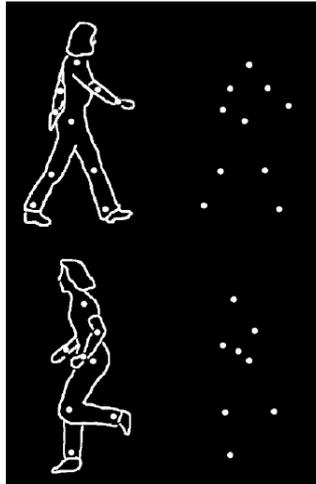
Self-propelled movements and goal-directed motion might be related to each other. Premack (Premack, 1990) proposes that infants will treat as intentional any self-propelled object, and Gergely and Csibra (Gergely & Csibra, 2003) show that infants attribute goals more often to self-propelled objects than objects that are launched by other objects. But even when its movements show no signs of self-propelled motion an object can appear goal-directed (Csibra, Gergely, Biro, Koos, & Brockbank, 1999), which indicates that there are probably other cues for goal-directedness, such as adjustments of the object's behaviour depending on the relevant aspects of the environment.

### 3.1.2 Biological bases for the perception of animate moving objects

A number of studies in humans show that the cortex surrounding the posterior part of the superior temporal sulcus or pSTS in both hemispheres (also known as temporo-parietal junction in the human, or TPJ) is the most consistently activated brain region during the observation of interactions between simple moving objects that appear causal or intentional (Castelli, Happe, Frith, & Frith, 2000; Blakemore et al., 2001; Schultz et al., 2003; Blakemore et al., 2003). Castelli and colleagues showed that abstract geometrical shapes moving in a goal-directed way induce activation increases in the superior temporal sulcus, and this activation increases further when the agent appears to act intentionally (Castelli, Happe, Frith, & Frith, 2000). The following other areas were also activated in these studies. Comparing movements appearing causal with movements appearing non-causal activated bilateral V5/MT and left intraparietal sulcus or angular gyrus in addition to the pSTS (Blakemore et al., 2001). Contingent movements between animate-looking objects activated the superior parietal cortex more than non-contingent movements in the same objects (Blakemore et al., 2003), and only attention to contingency induced activation of the pSTS in this experiment. Basal temporal, occipital cortex and medial prefrontal cortex in addition to the pSTS showed activation correlating with intentionality scores during observation of Heider and Simmel-like animations (Castelli, Happe, Frith, & Frith, 2000). A similar sort of animation showed activation in the fusiform gyrus, the temporal poles, the medial prefrontal cortex and the inferior frontal gyrus in addition to the pSTS (Schultz et al., 2003).

### Biological motion and the superior temporal sulcus

Different sorts of evidence suggest that biological motion is probably a particular type of motion. Johansson, in a now classic study, presented participants with films of humans dressed in black and white markers attached to their joints performing simple actions (Johansson, 1973, Figure Intro.10). While only the moving markers could be seen in the films (this type of display is now commonly known as a point-light display), participants could reliably identify walking or running movements. Additional psychophysical studies suggest that detection of biological motion is probably effected by very flexible mechanisms under the influence of learning, different from those involved in the detection of other forms of complex motion (Neri, Morrone, & Burr, 1998; Giese & Poggio, 2003). Biological motion could obey certain specific laws that could determine its specificity, such as the " $2/3^{\text{rds}}$  power law" that describes how the speed of hand movement changes with the radius of a curve drawn by a person writing (Lacquaniti, Terzuolo, & Viviani, 1983). While the  $2/3^{\text{rds}}$  power law describes characteristics of fine hand movements, a similar law could describe the characteristics of other types of biological movements. Several motion-blind patients were able to discriminate biological motion stimuli (Vaina, Lemay, Bienfang, Choi, & Nakayama, 1990; McLeod, Dittrich, Driver, Perrett, & Zihl, 1996), whereas some subjects with relatively normal motion perception were unable to discriminate biological motion stimuli (Schenk & Zihl, 1997). These results suggest a dissociation between the perception of normal visual motion and biological motion.



**Figure Intro.10. The classic point-light display of biological motion**

By Johansson (1973). From Puce and Perrett, 2003.

Also, experiments showed that execution of an action is affected by the observation of a human arm performing a similar but incongruent action (e.g. grasping a small object is performed less well during observation of a person grasping a big object, while observing a person grasping a small object does not interfere with the movement: Castiello, 2003). However, observation of a robotic arm (roughly resembling a human arm, but without its biological movement characteristics) performing the similar but incongruent action does not affect the execution of an action (Castiello, 2003; Kilner, Paulignan, & Blakemore, 2003). As Castiello and colleagues and Kilner and colleagues concluded, observation of biological movements appears to interfere with motor execution, whereas similar movements without the characteristics of biological motion do not. Confirming this idea is the finding that the human premotor cortex is only activated when a grasping movement directed towards an object is performed by a human rather than a robotic arm (Tai, Scherfler, Brooks, Sawamoto, & Castiello, 2004).

The posterior part of the superior temporal sulcus in humans (pSTS), also known as temporo-parietal junction, is a cortical region located at the posterior and superior end of the temporal lobe that has received much interest in recent years. Anatomically, the pSTS region receives inputs from both the dorsal and ventral visual streams, and is connected to regions thought to be involved in social and emotional processing (Brothers, 1990; Adolphs, 2003). The part of monkey STS called STP (the Superior Temporal Polysensory area) receives input from area MST (Medial Superior Temporal area) in the dorsal visual stream and from the anterior inferotemporal area in the ventral visual stream (Boussaoud, Ungerleider, & Desimone, 1990; Felleman & Van Essen, 1991; reviewed in Puce & Perrett, 2003). The cortex in the STS also has connections to the amygdala (Aggleton, Burton, & Passingham, 1980) and the orbitofrontal cortex (Barbas, 1988).

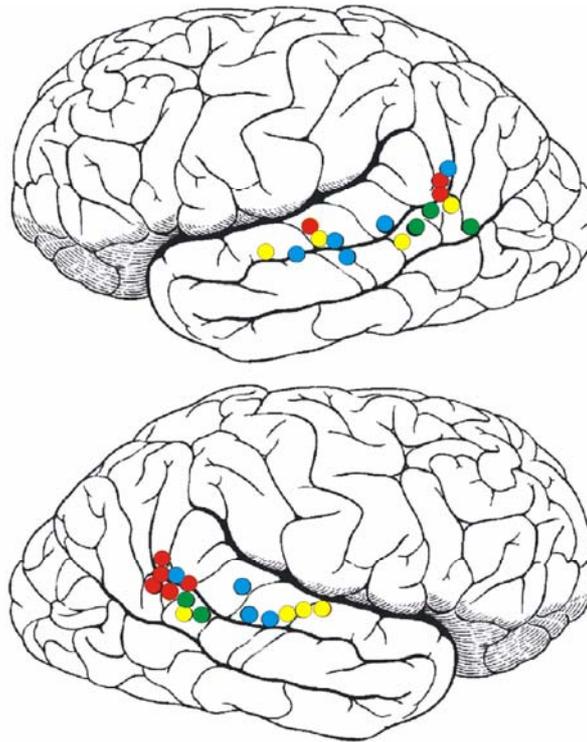
In humans, there appears to be a difference in specialisation between the anterior and the posterior part of the STS and STG. The *posterior* part of the human STS responds to the same stimuli that induce increases in firing in cells of the *anterior* part of the monkey STS (STSa), which led some researchers to argue that the monkey's anterior STS corresponds functionally to the posterior part of the human STS (Karnath, 2001; Jellema & Perrett, 2003a). The *anterior* parts of the STS and STG appear to be involved in auditory processing, including pitch, objects, and voices (Belin, Zatorre, Lafaille, Ahad, & Pike, 2000; Warren, Uppenkamp, Patterson, & Griffiths, 2003; von Kriegstein, Eger, Kleinschmidt, & Giraud, 2003; Zatorre, Bouffard, & Belin, 2004). Therefore, only data from the posterior part of the human STS will be discussed in this thesis.

In addition to the association with mentalising mentioned above, neuroimaging and neurophysiological data suggest that this part of the cortex can also be activated

by biological motion. The first studies using Johansson's classic point-light displays (Figure Intro.10) unexpectedly revealed activation in the pSTS (Howard et al., 1996; Bonda, Petrides, Ostry, & Evans, 1996), at a time when this brain region was thought to be participating in speech processing.

Further studies confirmed the pSTS activations: Grossman and colleagues showed point-light displays of various human movements which all activated the pSTS (particularly on the right side), whereas control moving dot displays including scrambled biological motion only activated area V5/MT/MST and the lateral occipital complex (Grossman et al., 2000). Even watching upside-down displays and imagining seeing biological motion activated the pSTS, albeit to a lesser degree (Grossman et al., 2000). Although some studies showed no STS activation during observation of biological motion (Vaina, Solomon, Chowdhury, Sinha, & Belliveau, 2001), a more recent study comparing observation of a walking human, a robot walking similarly, meaningful (a moving grandfather clock) and meaningless mechanical motion (coherently moving assembly of tubes and blocks) showed that observation of biological motion, independent of the aspect of the moving entity, yielded higher pSTS activation compared to the controls (Pelphrey et al., 2003). Comparing point-light displays and videos of tools moving in their characteristic fashion and human whole-body motion, Beauchamp and colleagues showed that the superior temporal sulcus responded strongly only to human motion, and slightly more to videos of human motion than point-light displays of it, suggesting that the STS integrates form, color and motion of biological entities (Beauchamp, Lee, Haxby, & Martin, 2003). These results confirm a previous study by Beauchamp and colleagues, which showed that STS responds to articulated human motion more than to unarticulated human motion, tool motion or static humans or tools (Beauchamp, Lee, Haxby, & Martin,

2002). A recent article by Puce and Perrett reviewed many of these studies and confirmed the involvement of the superior temporal sulcus in processing of biological motion (Puce & Perrett, 2003, Figure Intro.11).



**Figure Intro.11. Activation in the superior temporal sulcus and gyrus during observation of biological motion.**

Top and bottom panels represent left and right hemisphere, respectively. Data are from studies of observation of body movements (red), hand movements (blue), mouth movements (yellow), and eye movements (green). As can be seen, reported activation locations spread all along the temporal sulcus, with a concentration in the posterior part, particularly in the right hemisphere. From Puce and Perrett, 2003.

The posterior STS appears also to be involved in other tasks of social relevance and intentionality, particularly involving faces. This will be discussed further below.

### Attention effects

Previous studies indicate that recognition of biological motion can be influenced by attention (Thornton, Rensink, & Shiffrar, 2002). Other studies indicate that activation in the superior temporal gyrus area increases when subjects attend to more "socially relevant" dimensions of a visual display, such as emotion (Narumoto, Okada, Sadato, Fukui, & Yonekura, 2001), trustworthiness (Winston, Strange, O'Doherty, & Dolan, 2002) and contingency between the movements of two objects (Blakemore et al., 2003).

### 3.2.1 Action observation and imitation: the role of goals

Imitation and observation of human actions and their influence on action execution have been studied for a very long time, both in animals and humans. As this literature is very extensive, I will constrain my discussion by drawing on conclusions from the studies mentioned in the sections above. The importance of goals for the identification of potential living entities was discussed in section 3.1.1, and Barresi and Moore's proposal that the imitation of goals is necessary for the development of an intentional schema was mentioned in section 2.1. Also, as will be detailed later, neural structures involved in action observation and execution seem to be especially sensitive to the presence of goals. I will therefore discuss aspects of action observation by focusing mainly on imitation and execution of goal-directed actions.

### Imitation and action observation in adults: the ideomotor principle

An influential model for the approach of action observation and imitation is the ideomotor principle, especially as advocated by Prinz and colleagues (Prinz, 2002). This theory is based on ground laid by Lotze (Lotze, 1852) and James (James, 1890).

These authors argue that voluntary action needs 1) an idea of what is willed or intended and 2) a lack or removal of conflicting ideas. If regularities exist between actions and perceivable events, 1) one can expect certain events given certain actions and 2) one can select and initiate a certain act to achieve certain effects. The selection of an action to produce certain effects is the basis for the ideomotor principle. As Prinz recalls, the connection between action execution and action observation was first suggested by Greenwald's "ideomotor compatibility" (Greenwald, 1970). This refers to motor learning observed in situations where there are common features between a stimulus and the outcome of an action produced in response to this stimulus. The ideomotor principle builds on ideomotor compatibility. It postulates the following: the representation of an event that follows from a given action can evoke the same action again.

Guided by this principle, Prinz and colleagues performed a series of experiments on action reproduction, selection, initiation and production. In one series of experiments, subjects were asked to reproduce the movements of abstract objects moving on a screen while watching irrelevant moving distractors (Kerzel, Hecht, & Kim, 1999). In these experiments, the movement produced by the subjects was influenced by the distractors, in ways depending on the relationship between the target and distractor objects and the time constraints for reproduction. This led the authors to suggest that movement perception and production draw on common representational resources.

In his article, Prinz (Prinz, 2002) discusses another series of experiments, in which it was demonstrated that observing actions can influence action selection, effector selection and even action initiation. Prinz describes a first (unpublished) experiment by Stuermer and colleagues, in which the experimenters asked volunteers

to perform one of two simple hand movements, while watching similar movements of a hand displayed in one of two colours. The volunteers had to select their movement based on the hand colour. The experimenters found a compatibility effect such that watching movements congruent with the subject's response movement speeded their response. When the experimenters presented only the final frames of the gestures (still in either of two colours), they observed the same effect. In another experiment, Brass and colleagues (Brass, Bekkering, & Prinz, 2001) asked subjects to lift either their index or their middle finger in response to images of similar finger tappings, presented together with symbolic cues on the fingers. Subjects had to select the finger to lift depending either on the image of the finger or the symbolic cue. When the finger to be lifted was cued by a stimulus finger doing the same movement, response times were shorter than when cued by the symbol on a stationary finger. Also, when the moving finger shown was irrelevant to movement selection, response times were affected by the congruency between the movement associated with the symbolic cue and the movement shown. Further experiments confirmed these findings, suggesting again that observed movements influence movement execution. Brass and colleagues also showed such compatibility effects even when the subject's response was fixed and only its timing was dependent on the stimulus presented.

In the same review article, Prinz (Prinz, 2002) also discusses unpublished work by Knuf and colleagues who asked subjects to manipulate a joystick in order to set a ball on a screen moving towards a target. While the subjects could influence conditions and the ball's movement at the start of the trial, after a certain point subjects could not influence the movement of the ball anymore. Nevertheless, they often continued to act on the joystick even after this point, as if this could influence the outcome of the ball's movement. This type of behaviour is known as ideomotor

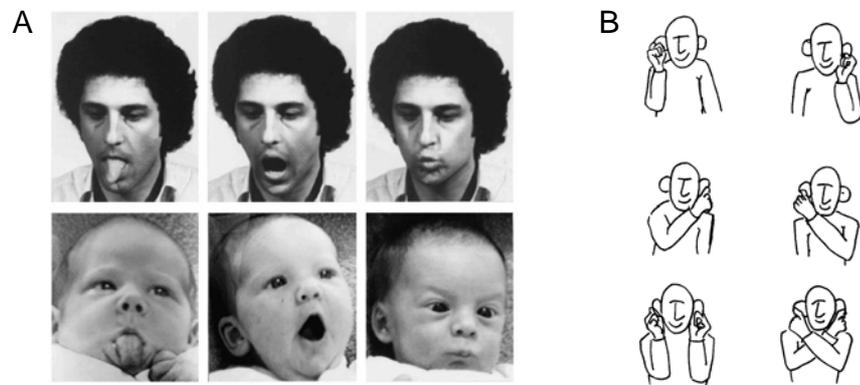
action, and Knuf and colleagues showed that it was influenced by the intentions of the subject rather than by what he perceived. The experimenters concluded that production of actions can be induced by the observation of actions or action effects, and that the observed actions are represented not only in physical terms but also terms of the underlying goals.

Reviewing all these experiments, Prinz concluded that there are common representational resources for perception and action, and that action imitation is therefore a natural by-product of perception. Most importantly, he suggests that action goals play a “dominant role in the representational structures mediating between perception and action”.

#### Goal-directed imitation in children

Since Meltzoff and Moore’s studies showing that newborn infants can imitate gestures suggested that imitation was an innate or early developing ability (Meltzoff & Moore, 1977, Figure Intro.12 A), imitation in children has been extensively studied. A number of studies have again revealed the important role played by goals. When asked to imitate a model who is performing an action that fails to reach its intended goal, 18-month-old infants perform the complete intended action leading to the goal state (Meltzoff, 1995). Bekkering, Wohlschlaeger and Gattis (Bekkering, Wohlschlaeger, & Gattis, 2002) asked pre-school children to imitate the actions of an experimenter sitting in front of them and reaching for one of his ears. In many cases, the children reached for the correct ear but used the hand ipsilateral to the ear. This did not happen when the experimenter always reached for the same ear or when his movements were directed at space instead of a physical object. So when there is more than one physical goal available, infants copy the goal of the action, but if there is

only one or no physical goal and the movement of the experimenter varies, they will imitate the movement itself. This speaks for a hierarchy of importance during imitation: most important is the goal, then the movements. This was nicely shown in the following study: if infants are asked to imitate an action performed in an unnecessarily complex way, they will often perform a different, more efficient movement than the experimenter, leading to the same goal, therefore imitating the goal but not the movement (Gergely, Bekkering, & Kiraly, 2002).



**Figure Intro.12. Imitation in young infants**

**A** From Meltzoff and Moore, 1977. **B** Experimenter's actions used by Bekkering and colleagues that children imitate an actions' goal rather than the exact body movements (when the experimenter reached for his opposite ear, children used reached their ipsilateral hand instead of the contralateral hand for their corresponding ear in 40% of trials using).

### 3.2.2 Biological bases for action observation

In addition to the above-mentioned connection between the posterior STS and processing of biological motion such as whole moving bodies or animate-looking abstract shapes, studies using more detailed human or animal actions have uncovered interesting neural structures, both in monkeys and humans.

### a) Studies in monkeys

#### *Execution of actions and observation of objects related to the action*

Neurophysiological studies in monkeys have revealed neurons in area F5 of the premotor cortex responding during execution of goal-directed actions (Rizzolatti et al., 1988; Rizzolatti, Fogassi, & Gallese, 2001). Some of these neurons respond during the presentation of objects that can be manipulated as well as during the manipulation; they are called “canonical neurons” (Rizzolatti et al., 1988; Rizzolatti, Fadiga, Gallese, & Fogassi, 1996; Gallese, Fadiga, Fogassi, & Rizzolatti, 1996). Neurons in the anterior intraparietal area (AIP) are also involved in the visual control of hand action and manipulation of objects (Sakata, Taira, Murata, & Mine, 1995; Sakata, Taira, Kusunoki, Murata, & Tanaka, 1997; Murata, Gallese, Luppino, Kaseda, & Sakata, 2000). The AIP has close links with Area F5 via reciprocal, monosynaptic connections (Godschalk, Lemon, Kuypers, & Runday, 1984; Cavada & Goldman-Rakic, 1989; Kurata, 1991; Luppino, Murata, Govoni, & Matelli, 1999). Because of their common properties and the connections between them, AIP and F5 are thought to transform the properties of objects (also known as affordances) into the appropriate hand actions: AIP would identify the object’s affordances, and F5 would use this information to select appropriate actions (Murata et al., 1997; Fagg & Arbib, 1998).

#### *Observation and execution of actions: the "mirror neurons"*

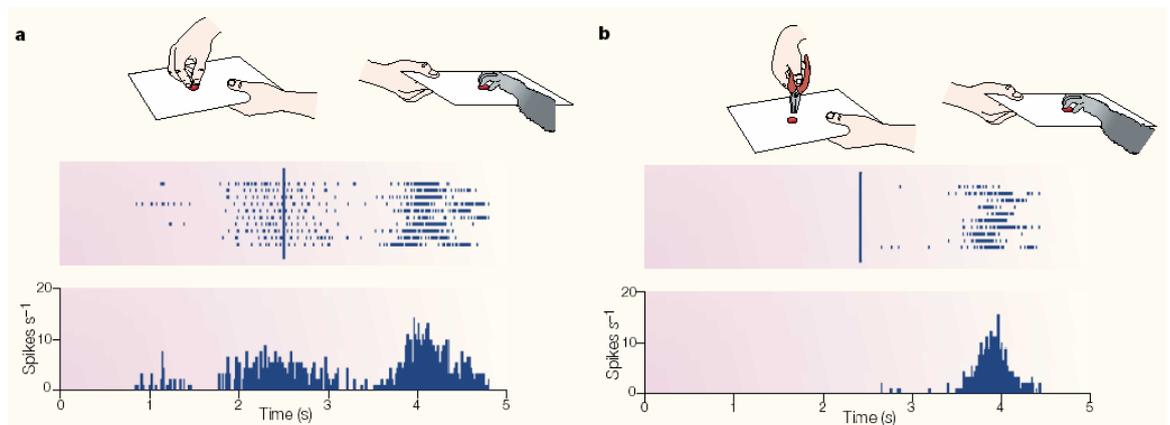
The behavioural studies of action observation and imitation listed above suggest a common representation for action production and perception. Neurons responding during both execution of actions or observation of the same actions performed by another agent have also been found in area F5 of the monkey premotor cortex, and are called “mirror neurons”. These neurons respond even when the object of the observed

action is hidden from the view of the observer (Umiltà et al., 2001), but not when the movement is a pretended manipulation of an object, a mime (Gallese, Fadiga, Fogassi, & Rizzolatti, 1996; Rizzolatti, Fadiga, Gallese, & Fogassi, 1996) (see Figure Intro.14). Some mirror neurons, in addition to responding during observation and execution of actions such as dropping a stick or tearing paper, also respond to the sound of these actions whether presented alone or during observation of the corresponding action, but not to sounds of non-goal-directed actions such as white noise, clicks or monkey calls (Kohler et al., 2002). These results were considered by the authors as further evidence for the coding of action representations by mirror neurons. In another study, Ferrari and colleagues (Ferrari, Gallese, Rizzolatti, & Fogassi, 2003) report that about a third of F5 motor neurons responding to mouth movements also respond during observation of mouth movements: they qualify these neurons as mouth mirror neurons. While the majority (85 %) of these mouth mirror neurons respond to ingestive actions such as grasping, sucking or breaking food, some of them (15 %) respond more strongly during observation of communicative, non-transitive mouth actions such as lip-smacking, lip protrusion or an open-mouth face. While some of the latter neurons responded to execution and observation of exactly the same communicative action, the majority responded during execution of ingestive actions and observation of communicative actions (Ferrari, Gallese, Rizzolatti, & Fogassi, 2003).

Other neurons that respond during observation and execution of actions have also been found in Area 7b (located in the anterior part of the convexity of the anterior inferior parietal cortex) (Fogassi, Gallese, Fadiga, & Rizzolatti, 1998). In the superior temporal sulcus, some of the cells responding during observation of actions appear to respond when the monkey executes an action (Perrett et al., 1989; Wachsmuth, Oram,

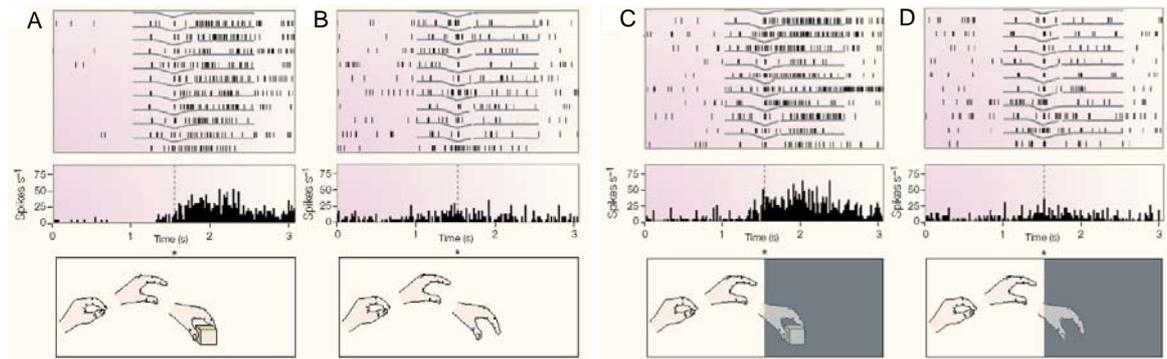
& Perrett, 1994). This might qualify these neurons for the label “mirror neurons”, although they appear to respond only when the monkey can observe his own action, which suggests that it is again the observation of the action (whether performed by the monkey himself or another agent) that plays the most important role in activating these cells (Perrett et al., 1989; Wachsmuth, Oram, & Perrett, 1994).

The way the STS, F5, AIP and Area 7b interact with each other is not fully clear yet. It has been suggested that the frontal system is mainly involved in controlling the interactions of the self with an object (Rizzolatti, Fadiga, Gallese, & Fogassi, 1996), whereas the STS system is more specialised for the detection and recognition of the behaviour of others (Perrett et al., 1990; Puce & Perrett, 2003).



**Figure Intro.13. Visual and motor responses of a mirror neuron in area F5.**

**A)** A piece of food is placed on a tray and presented to the monkey, the experimenter grasps the food, then moves the tray with the food towards the monkey. Activation is seen in F5 neurons while the monkey observes the experimenter’s grasping movements (left side of **a**), and while the same action is performed by the monkey (right side). **B)** Control condition, in which the experimenter grasps the food with pliers. There is no neural response when the monkey observes the movement. Rasters and histograms show activity before and after the point at which the experimenter touched the food (vertical bar). From Rizzolatti, Fogassi and Gallese, 2001.



**Figure Intro.14. Activity of a mirror neuron in F5 in response to action observation in full vision and in hidden conditions.**

The lower part of each panel shows the experimenter's action: the experimenter's hand moving towards an object and grasping it (panels **a** and **c**) or mimicking grasping (panels **b** and **d**). The second factor in this experiment was given by either allowing full vision of the movement (**a** and **b**) or hiding the contact between hand and object (using an opaque sliding screen, illustrated by the grey squares in **c** and **d**). At the beginning of trials with the sliding screen, the monkey was shown whether there was an object present or not before the screen was moved in and the action started. In rasters and histograms, the vertical line shows the point at which the experimenter's hand was closest to the point at which the experimenter's hand began to disappear from the monkey's vision in conditions **c** and **d**. Rasters and histograms are from 10 consecutive trials recorded during each condition. Gray traces above the rasters represent kinematic recordings of the experimenter's hand. The illustrated neuron responded to the observation of grasping and holding in full vision and in the hidden condition (**a** and **c**), in which the interaction between the experimenter's hand and the object occurred behind the opaque screen. The neuronal response was virtually absent in the two conditions in which the observed action was mimicked (**b** and **d**). From Rizzolatti, Fogassi and Gallese, 2001.

### *The superior temporal sulcus*

As discussed above, studies in humans and monkeys have shown that this structure responds during observation of biological motion. Also discussed above were studies in humans showing that this structure also responds during observation of animate-looking, moving geometrical figures. In addition, neurons in the superior temporal sulcus of monkeys respond during observation of static and moving body parts (Puce & Perrett, 2003). Experiments by Perrett and co-workers revealed that the

firing rate of these neurons increase during observation of fingers, hands, chests, faces and mouths, and also during observation of moving faces, limbs and whole bodies. These cells respond to whole-body movements presented in full light, and some of them also respond to the same action when presented as a point-light display. More cells respond to forward-moving whole bodies than backward-moving bodies. In addition, other cells respond to hand-object interactions such as picking, tearing and manipulation of objects, and do not respond to tools only. They also respond more when the interaction between hand and object is causal. STS neurons are also activated by static postures that are the consequence of articulated body motion (Jellema & Perrett, 2003a), or static postures of faces or body that follow actions (Jellema & Perrett, 2003b). Regarding face perception, neurons in anterior STS appear to code facial views (front, side), whereas anterior inferior temporal gyrus neurons appear to code facial identify (Eifuku, De Souza, Tamura, Nishijo, & Ono, 2004).

#### b) Studies in humans

##### *Action observation and execution*

Grezes and colleagues reviewed human neuroimaging studies with PET, fMRI and magnetic electroencephalography (MEG) of observation of actions with or without objects, imitation of actions or observation of objects (Grezes & Decety, 2001). These studies show activations in parietal and premotor areas compatible with the neurophysiological studies. Activated regions include Broca's area in the inferior frontal gyrus, a region thought by some to be a human equivalent of area F5 in the monkey. In an attempt to identify the human equivalent of canonical neurons in the human brain, Grezes and colleagues found activation in the anterior intraparietal

sulcus and the premotor cortex, which could correspond to the AIP-F5 circuit identified in the monkey (Grezes, Armony, Rowe, & Passingham, 2003). The authors report activation increases both during observation and execution of actions (corresponding to the mirror neurons) in the anterior intraparietal sulcus, dorsal premotor cortex, superior temporal sulcus and right parietal operculum. Similar to mirror neurons, the activation of the human frontal operculum during observation and imitation of hand movements is higher when the movement is directed at an object as compared to no object (Buccino et al., 2001; Koski et al., 2002). Imitation and observation of actions showed activations in the ventral premotor cortex (Rizzolatti, Fadiga, Gallese, & Fogassi, 1996; Iacoboni et al., 1999) and the superior temporal sulcus (Iacoboni et al., 2001).

Observation of actions can also influence the output of the motor system. Fadiga and colleagues (Fadiga, Fogassi, Pavesi, & Rizzolatti, 1995) recorded from a muscle (first interosseus) in the hand of subjects while they watched movements either involving or not involving the recorded muscle. Watching movements involving the muscle while receiving trans-cranial magnetic stimulation (TMS) of the part of the motor cortex corresponding to the recorded muscle resulted in greater motor-evoked potentials than watching movements not involving the recorded muscle. Similar results were found in experiments performed by other researchers (Strafella & Paus, 2000; Aziz-Zadeh, Maeda, Zaidel, Mazziotta, & Iacoboni, 2002). In accordance with the responses of mirror neurons to auditory stimulation (Kohler et al., 2002), Aziz-Zadeh and colleagues (Aziz-Zadeh, Iacoboni, Zaidel, Wilson, & Mazziotta, 2004) found sensitivity of the human motor system to action sounds. In this experiment, participants listened to the sounds of a bimanual action (typing on a type-writer) or to sounds of a person walking, while receiving trans-cranial stimulation of the motor

region corresponding to a hand muscle involved in the heard hand action. Greater motor evoked potentials were recorded when participants listened to the sounds of the hand action than when they listened to the sounds of the action performed by the feet. The authors concluded that the motor representation of the action was activated by the action sound, and that this increased the corticospinal excitability of the muscles involved in the heard action (Aziz-Zadeh, Iacoboni, Zaidel, Wilson, & Mazziotta, 2004).

### *STS and social cognition*

In addition to the role of the pSTS in processing of biological motion that has been discussed above, this structure might also be involved in the processing of human faces. Based on PET, fMRI, EEG, MEG experiments and cortical recordings in patients, it has also been proposed that the superior temporal sulcus is part of the network involved in the processing of faces, and particularly in the processing of the changeable aspects of faces, such as eye gaze, mouth movements or face expression (Haxby, Hoffman, & Gobbini, 2002). This brain structure is also believed to be a part of the “social brain”, a network of regions involved in processing social information (Brothers, 1990; Adolphs, 2003). In addition, the superior temporal gyrus in humans has been shown to be activated when subjects make explicit judgements about the trustworthiness of faces (Winston, Strange, O'Doherty, & Dolan, 2002). Also, a greater activation in this region was found while pairs of subjects cooperated in an iterative Prisoner's Dilemma game than when they did not cooperate (Rilling et al., 2002). A recent study showed activation in the superior temporal sulcus when subjects watched faces of willingly cooperating game partners during an Ultimatum Game and no activation when the game partners were forced to play in a particular way (Singer, Kiebel, Winston, Dolan, & Frith, 2004). This was interpreted as showing that only the

observation of faces of intentionally interacting game partners, or partners involved in a socially relevant interaction was able to induce activation in the superior temporal sulcus. In a study of moral dilemmas (Greene, Sommerville, Nystrom, Darley, & Cohen, 2001), activation in the right posterior STS was greater when volunteers read emotionally engaging rather than less engaging moral dilemmas involving social interactions between people leading to serious injury or the death of innocent people. In the emotionally more engaging moral dilemmas, the main protagonist was physically involved in an action rather than only acting through an artifice, which might increase the call on representations of humans and social interactions, and result in greater activation in the posterior STS. In a recent study, short films depicting actors performing diverse every-day actions in their context induced stronger activation when two actors were interacting than when an actor was performing alone in regions including the right pSTS, right inferior frontal gyrus, dorsomedial frontal cortex and bilateral fusiform gyri (Iacoboni et al., 2004).

### c) Anatomical regions of interest in this thesis

In summary, regions that appear particularly interesting for the study of basic processes involved in mentalising include the ventral premotor cortex and the STS. The ventral premotor cortex is interesting for its association with the potential "mirror system" thought to contain representations of actions independent of the actor. The STS might be linked in many ways to mentalising, by its association with processing of: 1) biological motion, 2) moving parts of faces, 3) socially relevant characteristics of faces, 4) other aspects of social paradigms such as actions performed in a social context, and 5) mentalising itself. As a consequence of this list of processes associated with the STS, particular attention to activation of this region will be shown in all

studies of this thesis, with specific, anatomically-based regional analyses in some experiments (see Part 4). Discussion of the STS activation pattern will be part of the discussion of all experiments, and a summary discussion of the STS activation observed in all studies of this thesis will be undertaken in Part 5.

## **Chapter 5. Conclusions**

In summary of this section, I conclude that 1) mentalising is done by matching behaviour to mental states, whether the behaviour is simulated (as proposed in Simulation Theory) or the mental states are inferred (as advocated in Theory Theory) could depend on the situation; 2) goals are important to understand another's actions and the mental states related to them, and might be important even to categorise him as animate; 3) that neural structures could exist that respond to animate motion (the posterior superior temporal sulcus, or pSTS) and to goal-directed action (parieto-premotor circuit, pSTS), and that some brain regions might even respond to both, and also to more complex situations or tasks which can be related to goals and animate agents (pSTS again). Now I will discuss the experiments performed in this thesis, starting with their specific rationale.

## **Part 2: Rationale for experiments**

Experiments in this thesis will deal particularly with the agent detector module, the action observation system, and slightly with the mental state inference system.

### **Chapter 1. General rationale**

In this thesis, I will aim to 1) create simple models of agents based on objective movement parameters (mainly aspects of goal-directed movement) and 2) identify brain regions that respond during observation of such agents. To extend the testing of the role of goals to actions of different types I will also 3) test whether brain regions responding to human action respond differently to goal-directed and non-goal-directed movements.

### **Chapter 2. Specific rationale for experiments involving observation of intentional objects.**

In my cognitive model, the agent identification system is the first step necessary for any subsequent attribution of mental states. A number of studies in infants and adults have shown a role of goal-directed behaviour for the attribution of animacy to moving entities, but no objective mathematical parameter has been associated with the perception of animacy. I will try to develop an algorithm controlling the movements of geometrical shapes with which the appearance of animacy of these objects can be reliably varied, and in which the strength of the percept of animacy can be modified by the value of one parameter. Based on the association between goal-directed movement and the perception of animacy, I will try to use a parameter controlling goal-directed behaviour. Once this algorithm is defined, I will use it in a parametric

fMRI study to see if an increase in this parameter induces an increase in activation in brain structures already associated with the processing of biological motion and complex animations of abstract geometrical objects. This would help to understand which variables are used by the brain during attribution of animacy, and which brain regions are involved in this process. Then, to bring these relatively abstract and simplistic animations closer to previous studies of observation of moving geometrical objects, I will try to make the objects appear more animate by adding other movement characteristics associated with animate agents.

I will focus particularly on one brain structure for these experiments: the cortical region surrounding the posterior part of the superior temporal sulcus and gyrus. This region has been previously associated with processing of biological motion and moving geometrical objects. As a result of this focus, I will add a dimension to my fMRI experiments by asking subjects to perform both an explicit and implicit task using the same stimuli. This is based on previous studies that indicate that activation in the superior temporal gyrus area increases when subjects attend to more "socially relevant" dimensions of a visual display, such as emotion (Narumoto, Okada, Sadato, Fukui, & Yonekura, 2001), trustworthiness (Winston, Strange, O'Doherty, & Dolan, 2002) and contingency between the movements of two objects (Blakemore et al., 2003).

### 2.1. Experiment 1

In this study I aim to identify brain regions whose activity varies with a parameter associated with animate movements of two simple objects. I will present volunteers with short films of two moving disks or balls with different colours that change direction and speed in an unpredictable but smooth manner. These movements

do not seem to follow any obvious physical law, making the objects appear self-propelled (Scholl & Tremoulet, 2000). In addition, the objects will appear to interact with each other, with one object chasing the other, which tries to escape. The level of this interaction will be manipulated by linearly varying the correlation between the movements of the objects, which induces changes in attribution of animacy to the objects. Areas responding in a parametric way to the correlation parameter might thus be involved in the processing of animate movements.

## 2.2. Experiment 2

For this experiment, I hypothesize that observation of a moving geometrical shape chasing another moving shape might induce greater activation increases in the superior temporal gyrus and sulcus when the chasing object appears to understand the goals of the target object and predict its movements rather than simply following the target.

## 2.3. The role of attention

Due to the role of attention on the activation of the neural structures associated with the functions I am interested in, I will ask participants in my experiment to perform two different tasks using the same stimuli. In experiment 1, in one half of the experiment, participants will be asked to rate how much one object appears to chase the other object ("Interactivity-rating task"), and in the other half of the experiment, subjects will be asked to rate how fast the objects move ("Speed-rating task"). In Experiment 2, in one half of the experiment, participants will be asked to decide which strategy the "chasing" object used to catch the "target" object (the "Strategy"

task), in the other half of the experiment, subjects will have to judge whether the chaser was successful in reaching the target objects (the "Outcome" task).

### **Chapter 3. Specific rationale for action observation**

The importance of goals for action understanding and imitation has been described in the Introduction, and the role of goal-directed movements in the identification of living entities has been tested in the previous experiment series. But despite many studies delineating the neural circuits engaged during execution, observation and imitation of action, the role of goals in action observation remains to be fully understood. For example, while object-directed and non-object-directed actions seem to be so different that their execution engages different visual streams, it is unclear whether these types of movements can be differentiated during observation, and whether this is based more on the object presence or on the kinematics of the movement.

To test the relative importance of the object and the movement kinematics for the identification of object-directed movement, I will present healthy volunteers lying in a functional Magnetic Resonance Imaging (fMRI) scanner with four different object manipulations. I will present Mimed and Actual versions of these manipulations, both presented either with or without an object being manipulated. Subjects will have to discriminate between Mimed and Actual kinematics of different hand and arm movements, irrespective of the presence of an object. I will vary the difficulty of the task by using manipulations with more or less important kinematic differences between their Mimed and Actual versions.

To independently manipulate the types of kinematics and the object presence in order to assess their respective effects, I will create high-quality animations based on the recorded movements of an actor, in which I can present either type of kinematics (Mimed or Actual) either with or without object. This design offers in addition the

opportunity to test responses during artificial or realistic combinations of object and movement kinematics.

### **Part 3. Materials and methods**

#### **Chapter 1. MRI and fMRI**

Functional Magnetic Resonance Imaging (or fMRI) is a widely used non-invasive neuroimaging technique and is at the moment the main technique in use at the Functional Imaging Laboratory. But other non-invasive neuroimaging techniques exist, and will be briefly presented here. Electroencephalography (EEG) consists in recording the electrical activity of the human brain using scalp electrodes. It is a traditional and low-cost imaging method, widely used in clinical practice and experimental neuroscience. However, this technique is limited by the inability to precisely localise the neuronal source of the electrical activity. Magnetoencephalography (MEG), a technique which records the minute fluctuations in magnetic fields produced by electrochemical activity in the brain, provides better spatial resolution than scalp-recorded EEG. It is more recent and far more costly than EEG, and not widely available at present. At present, the most widely used techniques for functional neuroimaging are those that measure the metabolic changes indirectly related to neuronal activity. The two principal techniques are Positron Emission Tomography (PET), and functional magnetic resonance imaging (fMRI). PET uses various exogenous radiolabelled substances injected into the bloodstream to measure: regional cerebral blood flow (rCBF) with radioactive H<sub>2</sub>O, glucose consumption by neural tissue with radioactively marked glucose, or the density of receptors for specific neurotransmitters such as Dopamine or Serotonin with radioactive ligands. FMRI relies on changes in magnetic properties of haemoglobin, the Blood Oxygenation Level Dependent (BOLD) signal, to produce a measure of local levels of blood oxygenation. The temporal and spatial resolutions of the BOLD signal are

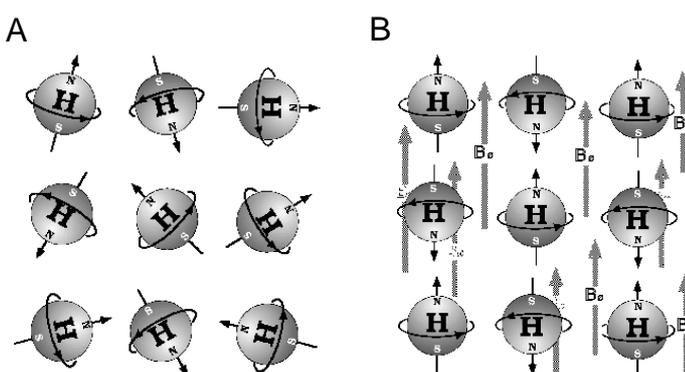
better than that afforded by PET measurements of rCBF. All neuroimaging data in this thesis were acquired by fMRI.

### **1.1. Basic MRI and fMRI physics**

fMRI is a special form of magnetic resonance imaging (MRI), used widely as a diagnostic tool in clinical practice. Both fMRI and MRI rely on the principle of nuclear magnetic resonance (NMR), discovered independently by Bloch and Purcell in the 1940s (Purcell, Torrey, & Pound, 1945; Bloch, Hansen, & Packard, 1946). The first application of NMR as a topographic imaging technique was by Lauterbur (Lauterbur, 1973) who showed that variations in magnetic field can be used to produce two-dimensional pictures of histological specimens, based on the behaviour of hydrogen atoms. Mansfield (Mansfield, 1977) showed how the magnetic signals could be analysed and transformed to an image rapidly and effectively, an essential step for a practical method. These early studies provided a basis for the use of MRI to produce three-dimensional images of the structure of human tissue. Mansfield also showed how extremely rapid imaging could be achieved by very fast gradient variations (so called Echo-Planar Imaging, or EPI). This technique became useful in clinical practice a decade later and is now the technique of choice in MRI and fMRI. Lauterbur and Mansfield together received the Nobel Prize in Medicine in 2003 for their discoveries. The following discussion is mostly based on material from [www.simplyphysics.com](http://www.simplyphysics.com).

## Spin

All atomic nuclei possess a quantum quality called spin. MRI techniques measure the effects of changing the spin of particular atomic nuclei, such as Hydrogen ( $^1\text{H}$ ) and Carbon ( $^{13}\text{C}$ ), which have an odd number of protons. In living organisms, the most abundant source of protons is the hydrogen atom in the form of water. The hydrogen nucleus is positively charged and the spinning motion of this charge induces a local magnetic field. These hydrogen nuclei (or protons) therefore behave like small magnets, i.e. they have a magnetic moment. In the absence of a magnetic field, these individual spins are randomly orientated and the bulk material has no net magnetisation (Figure Physics.1 A). But if an external magnetic field,  $B_0$ , is applied, all individual spins align with the external magnetic field (Figure Physics.1 B). If the spin is not completely aligned with the direction of the magnetic field  $B_0$ , this causes the proton to revolve, or precess, around the field direction (Figure Physics.2). The frequency with which the axis rotates around the field direction is called the resonance (or ‘Larmor’) frequency, and is directly proportional to the field strength  $B_0$ .



**Figure Physics.1. Magnetic properties of hydrogen protons.**

Orientation outside (A) and inside (B) a magnetic field. From [www.simplyphysics.com](http://www.simplyphysics.com).

Quantum mechanics describes that a proton in a magnetic field can have two spin orientations: either aligned with the field, like a compass needle directed towards the magnetic north in the earth’s magnetic field, or against it, which would correspond

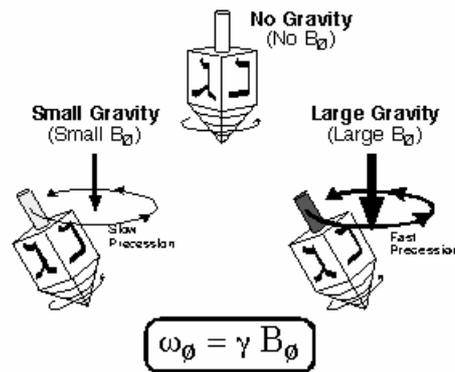
to the compass needle pointing towards the magnetic south. But contrary to compasses in the earth's magnetic field, the amount of protons aligned with the field is only slightly higher than the amount of protons aligned against it, and the difference in numbers (4.5 protons per million at 1.5 Tesla) increases with the strength of the magnetic field, increasing the signal-to-noise ratio. The proton will have different energy levels depending on the orientation of its spin in the magnetic field: when the magnetic moment is aligned with the field, its energy will be lower than when it opposes the field.

The amount of energy required to flip orientations is so small that the normal thermal energy available at room temperature is enough to flip spins. All of the signals generated in MRI are based on small differences between these energy states. The fact that these energy differences are small is one reason why MRI techniques tend to be safe but also why they are typically limited by signal strength.

### Equilibrium magnetisation

Summing over all the nuclei in an object volume gives the net magnetisation for the object. One usually describes the magnetisation by a three dimensional coordinate system where the Z-axis represents the direction of the applied field. At rest in the magnetic field of a MRI scanner, nuclei are either oriented with or against the direction of the magnetic field, with more protons oriented with the field (the lower, or ground, energy state) than against it. The net magnetisation moment in this situation will therefore also be oriented along the direction of the field (i.e., in the Z direction), a state called equilibrium magnetisation. As nuclei rotate or precess (Figure Physics.2), a small part of their rotating magnetisation has a component projecting into the XY plane, but because at rest the magnetic field does not influence them in

the XY plane, all nuclei orient randomly and therefore the net ‘transverse magnetisation’ cancels to zero. The frequency at which these neutrons rotate or precess is called the spin resonance frequency, Larmor frequency or precession frequency. This frequency is specific for the type of nucleus and the field strength, and is 42.56 MHz /Tesla for a proton.



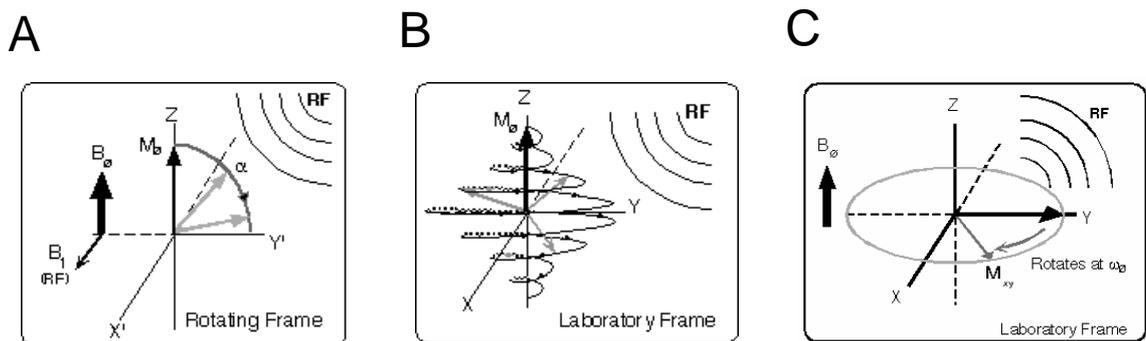
**Figure Physics.2. Example of precession movement.**

A rotating nucleus behaves like a spinning-top (such as the dreidle in this figure), with the frequency of the rotation increasing with the size of the B<sub>0</sub> field. From [www.simplyphysics.com](http://www.simplyphysics.com).

### Radiofrequency magnetic fields

Protons can be excited by an oscillating radiofrequency electromagnetic field (B<sub>1</sub>), applied perpendicularly to the main magnetic field (B<sub>0</sub>). The greatest efficiency for this energy transmission is reached when the radiofrequency of B<sub>1</sub> is equal to the resonance frequency. At the quantum level, a single proton jumps to a higher energy state. Changing the proportion of protons in the different energy states will change the direction of the net magnetisation vector. If protons are moved to the higher energy state until there are equal numbers of high and low energy spins, and they are simultaneously rephased by the RF pulse, then the net magnetisation vector is moved into the transverse plane. At the macroscopic level, to an observer in the external *laboratory frame* of reference, the magnetization vector spirals down towards the XY

plane (Figure Physics.3 A). The tip angle of the magnetisation vector,  $\alpha$ , is a function of the strength and duration of the radio frequency pulse (Figure Physics.3 B). The net magnetisation now also rotates about this new applied radiofrequency magnetic field  $B_1$ , just as they did about the original, much larger field  $B_0$ . During this new rotation, the angle from the original equilibrium direction along the Z-axis increases with time of stimulation by  $B_1$ , going towards the XY plane.



**Figure Physics.3. RF excitation and the effect on proton orientation.**

A proton is subjected to RF energy by a field  $B_1$  perpendicular to  $B_0$ , which tilts the proton's original equilibrium direction ( $M_0$ , in the Z axis) towards the XY plane, by an angle ( $\alpha$  in panel A) dependant on the strength of  $B_1$  and the amount of time it is applied for. For an observer in the laboratory, it would appear to spiral down towards the XY plane. When  $B_1$  is stopped, the proton turns like a spinning-top around Z, with a tilt of  $\alpha$ . The component of  $\alpha$  in the XY plane (the transverse magnetisation, or  $M_{xy}$  in C) gives rise to the NMR signal. From [www.simplyphysics.com](http://www.simplyphysics.com).

By varying the amplitude and duration of the RF exposure, typically delivered in millisecond pulses, any desired angle can be produced. For example, if the  $B_1$  field is applied long enough to equilibrate the spin populations, the net magnetisation vector will be zero and rotated into the XY plane. This is referred to as a  $90^\circ$  pulse, producing a  $90^\circ$  rotation of the net magnetisation, a  $90^\circ$  flip angle. For most angles from the equilibrium longitudinal magnetisation there will be a non-zero component of the magnetisation (the transverse magnetisation) in the XY plane. It is this transverse component that gives rise to a detectable NMR signal (Figure Physics.3.C). In the situation described above, when the  $B_1$  field is turned off after the  $90^\circ$  pulse,

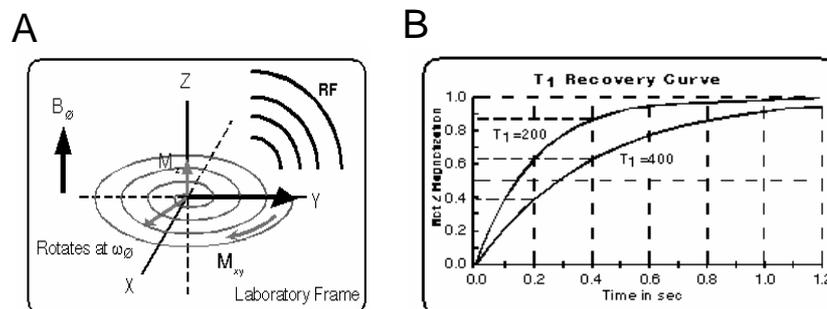
the magnetisation vector will rotate about  $B_0$  in the XY plane with the spin resonance frequency. This is observable because the oscillating magnetic field induces a voltage in a coil positioned in the XY plane (Faraday's Law). MRI systems are designed to measure the transverse magnetisation, so the receiver coils, which may be the same as those used to apply the RF pulses, are sensitive only to the transverse component. The initial amplitude of the detected RF signal is proportional to the number of protons in the sample (the proton density). The greater the proton density, the greater the magnetisation hence the greater the signal detected by the RF coils.

### Relaxation

After the RF pulse is turned off, the spins go from a high energy state to a low energy state, the Z component ( $M_z$ ) of the magnetization is recovering back to  $M_0$  and energy is released as RF emission at the resonance frequency. In the classical description, the  $M_z$  component increases at the expense of the  $M_{xy}$  component. The strength of the signal detected by the RF coils of the scanner, which is directly related to the amount of net transverse magnetisation or  $M_{xy}$  component, will gradually decay to zero. The protons realign with  $M_0$  gradually and in two ways simultaneously: (1) energy is transferred to neighbouring molecules in the surrounding environment by a process called spin-lattice relaxation, bringing the excited protons back to the original field orientation along the Z axis (T1 recovery to thermal equilibrium), and (2) energy is transferred to nearby protons in a process called spin-spin relaxation, which dephases them with respect to each other (T2 and T2\* relaxation). These two types of relaxation are described here.

#### *1) T1 relaxation*

Spin-lattice or T1-relaxation describes how the excited protons return to the original field orientation along the Z-axis (Figure Physics.4 A). Protons that have been excited to the higher energy state return to the lower energy state and reorient with  $M_0$ . The excess energy is dissipated to molecules of the surrounding structure ('lattice') as heat, until a thermal equilibrium is reached. This is an exponential process described by the T1 time constant, which is affected by the composition of the environment (Figure Physics.4 B). For example, the protons in water have a longer T1 than those in fat because the carbon bonds in fat resonate near the Larmor frequency, which facilitates the transfer of energy to the lattice. In the human brain, the different water content of grey and white matter (71% and 84%, respectively) means that the T1 time constants of these two tissues are different and this difference can provide contrast between these two tissues.



**Figure Physics.4. T1 relaxation.**

T1 relaxation is caused by the Mz component of the magnetisation vector realigning with  $M_0$  in the Z axis, i.e. with the  $B_0$  field (A). This relaxation, as it brings the proton into a lower energy state, produces release of energy as RF emission, and as heat to the surrounding tissue or 'lattice'. The T1 signal caused by the RF emission depends on the tissue and decreases exponentially (B). From [www.simplyphysics.com](http://www.simplyphysics.com).

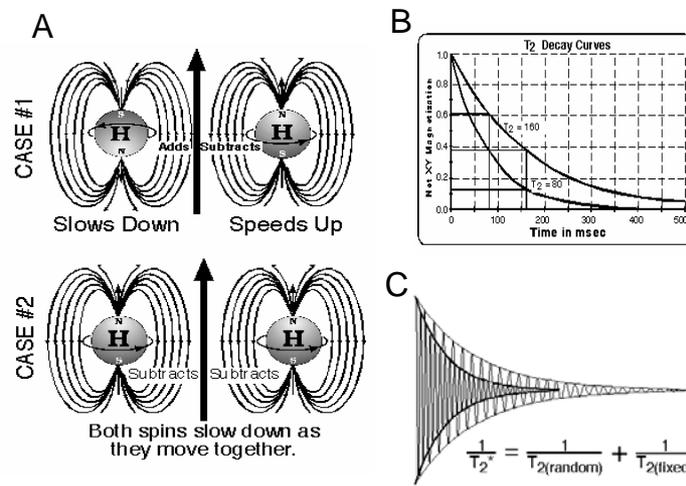
## 2) T2 relaxation

Spin-spin or T2-decay describes how the protons, initially in phase in the XY plane after the RF excitation, begin to dephase. The coherence of the magnetic

moment of the transverse magnetisation  $M_{xy}$  decreases, and energy is exchanged between neighbouring protons, a process called spin-spin (Figure Physics.5 A). This energy transfer causes the rotations of the protons to become desynchronised and have different rotation frequencies: they dephase, which decreases  $M_{xy}$ . This phenomenon is described by the  $T_2$  time constant, which is shorter than the longitudinal relaxation time constant ( $T_2 < T_1$ ) because in addition to energy transfer to the lattice molecules, dephasing represents an extra mechanism of magnetisation cancellation (Figure Physics.5 B).

**$T_2^*$  relaxation:** If the applied magnetic field is non-uniform, dephasing will occur more quickly, as spins in different parts of the object will rotate at different frequencies and quickly lose coherence. This loss of transverse magnetisation due to inhomogeneous fields is often much shorter than the natural  $T_2$  signal decay and is characterised by another exponential time constant,  $T_2'$ . The value of this time constant is determined by the technical implementation of the magnetic field and any field inhomogeneity caused by the properties of the object itself.  $T_2^*$  decay reflects the combination of  $T_2$  and  $T_2'$  signal decays (Figure Physics.5 C). FMRI sequences usually measure  $T_2^*$ . The magnetic field contribution to the inhomogeneity giving rise to the  $T_2'$  signal decay can be refocused with a spin echo, a  $180^\circ$  pulse along an axis in the  $xy$  plane at some time  $t$  after the initial  $90^\circ$  excitation pulse. Following this  $180^\circ$  'echo' pulse, the separate spin signals become more in phase and the signal becomes greater until time  $2t$  when all spins again have the same phase and the  $M_{xy}$  is temporarily again at a maximum. The time  $2t$  is called the echo time (TE). Echoes can be formed without  $180^\circ$  RF pulses by using magnetic field gradients to dephase and then rephase the spins. Adjusting the time of the dephase/rephase gradient balance allows the time of echo formation to be changed. This type of 'gradient echo' is often

used in fast imaging sequences where it can produce faster echoes and use less RF power than spin echoes. All  $T_2^*$  data presented in this thesis were acquired using this gradient-echo technique. Although the  $180^\circ$  pulses cancel out  $T_2'$  effects due to magnetic field inhomogeneities, the recovered  $T_2^*$  signal at the echo is still less than its original height as determined by the proton density. This is due to  $T_2$ -relaxation, the second contribution to  $M_{xy}$  magnetisation dephasing, which cannot be refocused with a spin-echo because it results from spatial and temporal variations in the intrinsic magnetic environment of each spin.



**Figure Physics.5.  $T_2$  and  $T_2^*$  decay.**

Spin-Spin (A),  $T_2$  (B) and  $T_2^*$  (C) decay. Decay time is again different across tissue types. From [www.simplyphysics.com](http://www.simplyphysics.com).

### Image formation: frequency and phase encoding

A homogeneous field  $B_0$  will affect all protons in a sample with approximately the same magnetic field, and hence the frequencies of their emitted RF signals will all be the same after the excitatory field  $B_0$  is turned off. But to obtain a tomographic image with MRI, one needs to distinguish protons according to their spatial position in the sample. To achieve this, one applies a second magnetic field (called a gradient field) to the sample which varies across the object, such that the resonance

frequencies vary according to the positions of the protons. For clinical whole-body scanners, these gradient fields generally have a maximum linear field strength change per unit distance of 22 mTesla/m. A particular frequency corresponds to the position of the proton in the sample, which depends on the strength and direction of the field, and the amplitude of the signal at that frequency corresponds to the number of protons at the corresponding position. And so, in the words of Lauterbur: the spectrum of a sample placed in a magnetic field gradient is a projection of the spin density along the gradient axis (Lauterbur, 1973). If the gradient field allows coding of positions in one dimension through frequencies, one needs another type of variation of the magnetic field to encode the proton positions in a second dimension: this is achieved by phase-encoding. An applied gradient will cause local magnetisation vectors to rotate with different frequencies depending on their positions within the gradient, increasing the dephasing of transverse magnetisation ( $M_{xy}$ ). If one applies an additional, intermittent magnetic field to the object, one can control the amount of dephasing of the local transverse magnetisation by varying the duration of the applied phase encoding gradient. The frequencies used to dephase the transverse magnetisation can be reconstructed by a series of increasing gradient pulse lengths. Therefore the amplitudes of spin frequencies are determined once more, despite the fact that phase is being manipulated in the second axis. And so, a spin density projection along the frequency encode axis and along the phase encode axis will represent the position of the protons in two dimensions. To create resolution in the third dimension, one can study the sample slice by slice by exciting only one slice at a time by combining the frequency gradient with a pulse of a particular frequency and bandwidth. The typical MR imaging procedure is therefore as follows: first the temporary application of a slice-selection gradient will excite a slice of the object, with application of a

frequency gradient in the X-axis (conventionally). Then follows phase-encoding along an orthogonal axis to X (Y axis).

### Voxels

Step-wise increases in both gradients divide the sample into small cubes, or voxels (volume-elements). All spins in one voxel experience the same frequency and phase encoding. As the signal of a given voxel is the sum of all spin contributions in the voxel, spins within a voxel cannot be distinguished from each other. The resolution of the image depends therefore on the size of the voxels, determined by the step size of the gradients. Increasing the size of the voxel increases its signal and therefore its signal-to-noise ratio, but decreases image resolution. Larger voxels are more likely to encompass groups of spins with very different behaviour, which could evoke a misleading signal (referred to as the partial volume effect). MRI systems operating at higher field strengths benefit from higher signal-to-noise ratios, a characteristic often used to reduce the size of the voxels in an image to increase resolution.

### Image contrast

The contrast of an image is given by the difference in signal intensity between areas caused by the different structures or compositions of an object. The MR signal intensity of a particular voxel is determined by interactions between following factors: proton density, T1 and T2 relaxation times, characteristics of the RF pulse, and magnetic susceptibility arising from other protons and electron clouds in the environment. Some of these factors (most importantly, the reduction of susceptibility artifacts) can be manipulated by the timing of the RF pulses with the pulse sequence

parameters, the "TR" and the "TE". "TR" refers to "repetition time", i.e. the time between two consecutive 90° RF pulses, which is the duration of acquisition of one image volume (all slices), and "TE" refers to echo time. Short TR and TE will emphasise the T1 characteristics of the tissue and produce a "T1 weighted" image, long TR and TE result in a "T2 weighted image".

### Ultrafast MRI sequences: Echo-Planar Imaging

Echo-planar imaging revolutionised MRI by allowing the acquisition of whole brain images in a fraction of a second. This type of imaging sequence, invented by Mansfield in 1977 (Mansfield, 1977), is characterised by Fourier-transforming the acquired data from the time domain into the frequency domain. As two orthogonal gradients are applied during the acquisition of an MRI sequence (frequency encoding and phase encoding), the transformed data are considered to lie in a two-dimensional frequency space, called K space. EPI sequences acquire data from all lines of K space after each RF pulse, whereas other MRI sequences sample only one line per RF pulse. This greatly reduces the acquisition time in EPI and makes it very suitable for dynamic imaging applications such as fMRI. All experiments described in this thesis have used EPI sequences. Two experiments were conducted at the Wellcome Department of Cognitive Neurology, Institute of Neurology, London, UK, using a Siemens VISION system (Siemens, Erlangen), operating at 2T. The third experiment was conducted at the Brain Activity Imaging Center of the Advanced Telecommunications Research laboratories (ATR labs) in Kyoto, Japan, using a 1.5T Shimadzu-Marconi ECLIPSE system. Both systems were used to acquire both T1-weighted anatomical images and gradient-echo echo-planar T2\*-weighted MRI image volumes with blood oxygenation level dependent (BOLD) contrast.



## **1.2. FMRI and neural activity**

### **BOLD contrast in fMRI**

Red blood cells (erythrocytes) transport oxygen from the lungs to organs of the body by binding it to haemoglobin. The haem group of this large protein contains iron that binds oxygen, which alters the magnetic properties of the iron. Iron, without oxygen bound to it, has a net magnetic moment due to its four unpaired electrons (Pauling & Coryall, 1936). This magnetic moment disappears when iron binds oxygen, as the four electrons are redistributed between the iron and the oxygen molecules. Therefore the magnetic state of the blood reflects the level of oxygenation of the local tissue.

As discussed above, the T2\* signal of water protons depends on interactions between the protons (spin-spin) and also on local inhomogeneities of the magnetic field B<sub>0</sub> caused by the magnetic properties of different molecules. As mentioned before, deoxyhaemoglobin is a paramagnetic molecule, and its local magnetic field gradient hastens the decay of transverse magnetisation and consequently shortens the T2\* time constant. Therefore changes in the levels of deoxyhaemoglobin (more precisely, changes in the deoxyhaemoglobin / oxyhaemoglobin ratio) should result in changes in T2\* signal. This effect was demonstrated in animal work carried out by Ogawa and colleagues (Ogawa & Lee, 1990; Ogawa, Lee, Nayak, & Glynn, 1990) and Turner and colleagues (Turner, Le Bihan, Moonen, Despres, & Frank, 1991). Both groups showed that experimental manipulation of the oxy- / deoxyhaemoglobin ratio (usually by hypoxia) produced detectable contrast changes in blood vessels and within the tissue water surrounding vessels. Ogawa and colleagues (Ogawa et al., 1992) and Kwong and colleagues (Kwong et al., 1992) demonstrated that these changes in T2\* signal observed in human blood vessels and tissue were sufficient to

produce a contrast detectable by MRI scanners. This contrast is now widely known as Blood Oxygenation Level Dependent (BOLD) contrast.

### Physiology of BOLD

The relationship between increases in neuronal electrical activity and change in blood oxygenation is still not fully understood (Logothetis & Wandell, 2004). Most researchers agree however on the following causal event sequence. During a task, neuronal activity increases in specific areas of brain grey matter. Oxygen and glucose consumption increases in these areas (Hyder et al., 1997), causing a decrease in blood oxygenation in the capillary bed supplying the active neuronal tissue about 100ms after sensory stimulation (Vanzetta & Grinvald, 1999). This induces a release of vasodilatory factors, which increase blood flow to these capillaries and dilates them (Duelli & Kuschinsky, 1993). After this response the oxygen supply to active tissue begins to exceed consumption and blood oxygenation in the capillaries and draining venules increases (Villringer & Dirnagl, 1995). If the neuronal activity remains high, an equilibrium between vascular supplies and metabolic demands is reached in 1-3 minutes. If on the contrary neuronal activity returns to baseline, blood flow also returns to baseline, but for 30-60 s after blood flow has re-equilibrated the blood volume in the draining venules remains elevated (Mandeville et al., 1999).

From these observations a triphasic model of the BOLD response has emerged. First, the transient decrease in oxygenation due to increased consumption produces a small, transient decrease in the BOLD response, commonly known as the ‘initial dip’. As a result of vascular response to this stimulation, local blood volume then increases, which finally leads to increases in local blood flow 500 ms – 1 s after sensory stimulation. This third phase of the neurovascular response causes four times larger

decrease in deoxyhaemoglobin concentrations than the initial increase due to increased oxygen consumption by the tissues (the initial dip). This decrease in deoxyhemoglobin increases BOLD contrast, which is typically measured in fMRI experiments, and is often termed the Haemodynamic Response Function or HRF. As described, the neurovascular response is delayed in time with respect to the neuronal event, and one typically observes peaks in the BOLD signal 4-6 seconds after a neuronal event, a delay which has to be taken into account during the analysis of fMRI data (see below). This delay has the consequence of dictating the effective temporal resolution of fMRI as it corresponds to a temporal smoothing of the underlying neuronal signal. But assuming that two similar neuronal events following each other by less than 3 seconds invoke a similar HRF and that both HRFs can be summed, experiments with an interstimulus interval of less than the duration of the HRF can be performed. This is the basic assumption underlying event-related experiments as analysed with SPM (see below).

#### Correspondence between neural activity and the BOLD signal

While many studies have now established a link between neural activity and the BOLD signal via brain metabolism, the details of 1) how metabolism correlates with neural signal and 2) how BOLD correlates with metabolism are currently debated topics. Recent results measuring BOLD and mean extra-cellular field potentials show that BOLD signal can be better predicted by the local field potential (LFP) rather than multi-unit activity (MUA). LFP represents slow electric changes, mainly synaptic signals, voltage-dependent membrane oscillations and spike afterpotentials, while MUA measures regional neuronal spiking (for a review and recent data see Logothetis & Wandell, 2004). In addition, as the LFP originates from both inhibitory and

excitatory neuronal activity, regions with more inhibitory activity could show less correlation between BOLD signal and spike rate (mainly excitatory) than regions with mainly excitatory activity. Based on these data, caution is advised in the direct comparison between neurophysiological data, which are mainly based on analysis of action potentials from pyramidal neurons, with the BOLD signal.

As the BOLD signal originates from erythrocytes in capillaries and draining veins, tissue oxygenation is only indirectly measured by current imaging techniques and therefore the maximal resolution of the neural activity obtainable with MRI measures of the BOLD signal is dependent on the local structure and density of the vascular architecture, which varies between regions. New techniques, including near-infrared spectroscopy and MR spectroscopy may enable measures more closely correlated to neural spike rate in the future. As blood occupies only a small fraction of grey matter, BOLD signal changes are of the order of a few percent at best. These small signal changes require the implementation of sophisticated image processing and analysis techniques to ensure that observations reflect true BOLD signal and not noise. These techniques are described in the next section.

## **Chapter 2. FMRI analysis**

### **2.1. Introduction**

FMRI is a method used to acquire information about the activity of the brain while a human or animal subject is subjected to a particular type of stimuli, is performing a particular task or is in any type of state that the experimenter would like to investigate. The fMRI technique permits the measurement of a signal dependent on the blood oxygenation level (the BOLD signal), while the participants are subjected to experimental conditions. The values of the BOLD signal recorded during the different conditions can then be compared, and from these differences changes in brain activation can be inferred. Due to the temporal resolution of the vascular response being in the range of seconds, the main contribution of fMRI data to neuroscience is (at the moment) the localisation of studied processes to specific areas of the brain. With sophisticated analysis methods, particular response profiles over time can be localized also, and interactions between areas during a task can be studied as well with methods such as functional connectivity. In well-designed experiments, the experimenter usually has an "a-priori" hypothesis of the areas of the brain that are likely to show changes in metabolism. This knowledge enables him to optimise the scanning protocol (area of the brain to cover, correction of potential susceptibility or drop-out artefacts by judicious choices of TR, TE and acquisition angle) and perform better-defined analyses. An influential software package for the analysis of fMRI data is the Statistical Parametric Mapping or SPM software, developed at the Wellcome Department of Imaging Neuroscience. All data acquired during the experiments of this thesis were analyzed with SPM. Analysis of data with the SPM approach starts with a series of spatial transformations to 1) correct for head artefacts related to head

movement, 2) to correct for slice acquisition time differences between different parts of the brain, 3) to make each brain fit into a template reference space in order to combine the activation data from a number of subjects together and 4) to impose a known spatial structure on the data. After these steps, a model of the expected BOLD signal changes measured during all conditions of the experiment is created and the data are fitted to this model using the General Linear Model. From the resulting parameter estimates activation maps are created, then tested for statistical significance, including an extra correction for multiple tests guarding against false positive results. All these steps will be detailed below.

## **2.2. Spatial preprocessing**

### Realignment

Even though subjects' heads are routinely immobilised with soft pads, even very experienced subjects move their head about 1 millimetre or so. Realignment uses rigid-body transformations (translation, rotation, etc.) to correct variations in pitch, roll, yaw, and X, Y and Z translation. In this process, the (sum of squared) differences between one image volume and the first image of the series are calculated and iteratively minimised. Once the movement parameters in all 6 directions are defined, the data are resampled using sinc or trilinear interpolation. But these transformations can also simply be stored in individual files corresponding to each image volume and used in the resampling effected during the normalisation step. Without correction, the displacement of excited spins from one spatial location to another from one image volume to the next might be interpreted as variations in activation caused by experimental manipulation, i.e. a movement-related artefact. To account for these movement-related changes in T2\* signal, it is customary to include the realignment parameters calculated during the realignment process into the design matrix at the analysis stage.

### Slice-time correction

An fMRI scanner takes a certain time to acquire one image volume (the repetition time TR, usually ranging from 1-5 seconds) and, in sequential acquisition sequences, slices are acquired in order from top to bottom or bottom to top. As a consequence, slices in different parts of the volume are acquired at different times. If one is interested in a neural event happening at a particular moment in time corresponding to the time of one particular image volume, then only one slice of that

volume will have been acquired at the exact time of interest, the other slices will have been acquired either before or after that time. To compensate for these timing differences, an option is available in SPM to correct for timing differences by Fourier-transforming the data of each volume and resampling them after sinc-interpolation compensating their relative timing. This results in corrected data that appear to have been taken all at the same time. Although this correction was considered to be essential for 1-2 years at the Wellcome Department (about 2000-2002), re-evaluation of its effects has shown that with TRs between 2-3 seconds slice-timing has no discernable effects (unpublished results by Rik Henson). Its use is now dependent on personal choice.

### Normalisation

To enable direct comparison between the data of different subjects, SPM adopts an image deformation approach, which fits each individual brain into a pre-defined template (Friston et al., 1995). A mean functional image is calculated during the realignment step and used to estimate non-linear image warping parameters necessary to fit this mean image into a template EPI image. Here also a minimisation of sum-of-squared differences between the template and the mean subject image is used to update the warping parameters until an optimal solution is found. The first step of the normalisation process consists in determining the optimal 12-parameter affine rigid-body transformations to register the source image to the target. Then a Bayesian framework is used so that the solution of the registration process maximizes the a posteriori probability of it being correct (Ashburner, Neelin, Collins, Evans, & Friston, 1997). More precisely, the process maximises the product of the likelihood function (derived from the residual squared difference) and the prior function (based

on the probability of obtaining a particular set of zooms and shears). After the affine registration step, non-linear deformations are estimated; these are defined by a linear combination of three-dimensional discrete cosine transform basis functions (Ashburner & Friston, 1999). Matching involves simultaneous minimization the membrane energies of the deformation fields and the residual squared differences between the image of interest and the template.

This procedure can also be used to normalise images from other modalities to a template image, such as a co-registered anatomical T1 image from a subject warped to fit the same EPI template used for the functional images. This allows overlay of the results of the SPM-analysis of the functional images onto the structural image of the same subject. During co-registration, the T1 image must be mapped onto the same space as the EPI images, a process also effected by sum-of-square minimisation.

### Spatial smoothing

The normalised images are routinely smoothed spatially with a three-dimensional isotropic Gaussian kernel of typically 5 to 10 mm full width at half maximum. This procedure, although it reduces the precision of spatial localisation by a small amount, compensates the remaining between-subjects variations in gyral and sulcal anatomy that still exist after normalisation and therefore reduces variations in the localisation of activations across subjects, increases the "normality" of the data (parametric statistics can be used only if the data are normally distributed) and increases the validity of the assumption that fMRI data can be considered to be Gaussian Random Fields.

### **2.3. Model-based analysis of fMRI data**

#### Basic description of the SPM approach

The approach used by SPM for the analysis of the data is based on the use of a special case of the General Linear Model. This approach is based on a model of the activation to be expected during the experiment, to which the acquired data are fitted, voxel by voxel. This results in the creation of three-dimensional maps of parameter estimates, which are then contrasted with each other to create statistical parametric maps (SPMs). These SPMs can then be thresholded to user-specified T values for statistical inference, including corrections for multiple comparisons across the whole brain or a specified search area to guard against false positive results, and comparison of effects over groups of subjects. The model of expected activation used in SPM is known as the design matrix.

#### Construction of the model

The model of expected activation is based on the known times of events of interest (presentation of stimuli, press of a button by the subject...) and, in SPM's most widely used option, a defined canonical haemodynamic response function (HRF, see Physics section) describing the change in BOLD signal in response to a given neural event. The HRF is modelled in SPM with a multivariate Taylor expansion of a mixture of gamma-functions (Friston, Josephs, Rees, & Turner, 1998). Higher order basis functions in this expansion include the partial derivative of the HRF with respect to time and dispersion, and allow more flexibility in the characterisation of the haemodynamic response by accommodating variations in dispersion or lag in the haemodynamic response. To create the design matrix, regressors corresponding to time-series of changes in BOLD signal intensity are created for each type of event of

interest (experimental condition) by placing delta or "stick" functions at the time-points corresponding to the events of interest and convolving them with the HRF (one regressor), and if required with the temporal derivative of the HRF (another regressor) and its dispersion derivative (a third regressor for this event type). To weight events of a regressor differently from one another, different parametric modulation options are available, allowing the modelling of linear or exponential time-dependent changes, or allowing the use of any variable with trial-specific values, including outputs from computational models (as was used for example in O'Doherty et al., 2004). To correct for spin-excitation history artefacts (head movement consequences which cannot be corrected during realignment), movement parameters of the head or mouth can be integrated in the model as regressors of no interest (regressors of no interest are characterised in SPM by the fact that they are not created through convolution with a canonical response shape, such as the HRF). Other custom-made regressors that can be included in the model include reaction times or biological parameters such as heart rate or respiratory frequency. This model, once finalised, is called the design matrix (see below).

In event-related experiments such as those performed in this thesis, it is assumed in SPM that the haemodynamic response behaves in linearly additive fashion when two neural events happen in quick succession (Josephs & Henson, 1999). SPM uses an implicit convolution regression model to this end (Friston et al., 1998). This allows modelling of the predicted BOLD response when two events follow each other by less than a few seconds, as is the case in the experiments described in this thesis.

### Eliminating temporal confounds

To exclude confounding drifts in the magnetic field and aliased biorhythms such as respiration frequency or heart rate, a high-pass filter is applied to the data before the model is fitted. The cutoff of this filter is typically chosen to be 128 seconds. To take into account temporal autocorrelation in the serially acquired fMRI data time-series, an auto-regressive model of order 1 + white noise is fitted to the data. In a linear autoregressive model of order  $R$ , a time series  $y_n$  is modelled as a linear combination of  $R$  earlier values in the time series (with lags  $j$  ranging from 1 to  $R$ ), with an error-term  $\varepsilon_n$ , where  $a_j$  is the autoregressive coefficient expressing correlation with the earlier value  $j$  in the time-series:

$$y_n = \sum_{j=1}^R a_j y_{n-j} + \varepsilon_n, \text{ where } \varepsilon_n \approx N(0, \delta^2)$$

In the case used in SPM, the order is 1, and the equation reduces to:

$$y_n = a_1 y_{n-1} + \varepsilon_n$$

In SPM2, the AR-coefficient  $a_1$  and the variance are estimated by using the Restricted Maximum Likelihood (ReML) method. Once these variance parameters are known, the error covariance matrix is used to pre-whiten the data to form Maximum Likelihood (ML) estimates of the parameters. The  $t$  or  $F$  statistic is then formed in the usual way.

### Fitting the model

The general linear model implemented in SPM is a general version of parametric models, of which linear regressions,  $t$ -tests and analyses of variance (ANOVA) are

special cases. The general linear model explains variations in the data ( $Y$ ) in terms of a linear combination of explanatory variables ( $x$ ) plus an error term ( $\varepsilon$ ):

$$Y_j = x_{j1}\beta_1 + \dots + x_{jl}\beta_l + \dots + x_{jL}\beta_L + \varepsilon_j$$

$\beta_l$  are unknown parameters (often referred to as "betas" in the jargon of users of the general linear model) corresponding to each of the  $L$  explanatory variables (the regressors) for the  $j$ th observation (element  $j$  of the data, here timepoint  $j$ , i.e. image volume No.  $j$ ) of  $Y$ . The errors  $\varepsilon$  are assumed to be identically and normally distributed.

One can also express the general linear model in matrix formulation:

$$Y = X\beta + \varepsilon$$

where  $Y$  is the vector of observations ( $Y = [1 \dots j \dots J]^T$ ),  $\varepsilon$  is the vector of error terms ( $\varepsilon = [\varepsilon_1 \dots \varepsilon_j \dots \varepsilon_J]^T$ ) and  $\beta$  is the vector of parameters to be estimated ( $\beta = [\beta_1 \dots \beta_j \dots \beta_L]^T$ ). Matrix  $X$ , of size  $J \times L$ , is the design matrix. This matrix has one row per observation and one column per regressor in the model. The number of parameters  $L$  is (usually) less than the number of observations  $J$  hence the simultaneous equations implied by the general linear model (obtained by expanding the matrix formulation with  $\varepsilon = 0$ ) cannot be solved (it is overdetermined). Therefore, some method is required for estimating parameters that fit the data best ("best fit"), determined by minimisation of squared errors. Using generalised least squares estimators, the relative contributions of each regressor (i.e., the parameter estimate for each regressor) to the experimental variance can be determined.

### t and F-statistics

Using the estimated variances of the estimated parameters, one can make inferences about the parameters using two types of statistical tests. The null hypothesis that all the estimates are zero is tested by an  $F$  statistic, resulting in an SPM( $F$ ). To test whether a particular combination of estimates (a "contrast", or subtraction, such as  $[1 \ 1 \ -1 \ -1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0]$  in a model with 11 regressors) is zero, one can use a  $t$  statistic, resulting in an SPM( $t$ ). The  $t$  statistic is computed by dividing the contrast of parameter estimates by the standard error of that contrast, estimated using the variance of the residuals of the least squares fit.

## **2.4. Inference and the Theory of Gaussian Random Fields**

As described before, SPM calculates the fit of the model (design matrix) for each voxel separately, a mass-univariate approach. Performing say a t statistic on each voxel to assess the effect of a particular contrast poses the problem of multiple comparisons: in classical statistics, one would perform a Bonferroni correction to correct for false positives, i.e. the possibility of getting a significant t statistic by chance. There are two ways to deal with this issue in the case of fMRI data analysed with SPM.

### i) Whole-brain inferences

If the experimenter does not have an a-priori hypothesis or is interested in responses of all voxels in the brain, then a correction for multiple comparisons is applied to the p values resulting from a given test. Given that fMRI data are correlated in space due to the anatomy of the brain (neighbouring voxels have a greater probability of being activated if they surround an activated voxel) and the spatial smoothing performed during pre-processing (used to impose a known smoothness to the data to harmonize an otherwise locally variable smoothness due to brain anatomy), one can consider each EPI image volume of an fMRI experiment as a Gaussian random field, characterised by a certain smoothness (to be calculated for each image). Random field theory provides solutions to control for false positives due to multiple comparisons, and provided that the image is sufficiently smooth this correction is less severe than the Bonferroni correction.

Using the Theory of Gaussian random fields, SPM calculates P values at different levels: set level (the number of activated regions, i.e. clusters above a height and volume threshold), cluster level (number of activated voxels defining a particular

cluster) or voxel level (p value for each voxel within that cluster). The most commonly used levels are the cluster and voxel levels. Knowing the number of clusters and the number of voxels they contain and a  $t$  value superior to a threshold given by the user, the probability  $P$  that a maximum value in a cluster would be greater than that observed under the null hypothesis (when no activation is present) can be tested. SPM approximates  $P$  using the expected Euler characteristic, a topological measure corresponding to the number of peaks minus the number of holes in a field. When high thresholds are used, the expected Euler characteristic corresponds to the number of regions above  $t$ , and gives the probability of the maximum exceeding that threshold. This probability indicates the test level necessary to reject the null hypothesis at that voxel, and corresponds to the adjusted p value at that voxel. To calculate the expected Euler characteristic, one needs to know the smoothness of the data. This is calculated from the residual fields, i.e. the data left after subtraction of the modelled response at each voxel, normalised by the variance at each voxel. This smoothness is represented as the full width at half-maximum (FWHM) of a Gaussian point spread function, and parameterised by a variance-covariance matrix of the spatial partial derivatives. By dividing the volume of the search region by the product of the FWHMs of the smoothness in each dimension, one obtains the number of resolution elements (resels), which indicate the number of independent tests. The expected Euler characteristic depends on the number of clusters and the probability that the cluster will be bigger than  $k$  resels. Less smooth data yield more resels, and therefore the probability to obtain maxima as large as those observed gets bigger, and hence the p values decrease.

## ii) Anatomically constrained inferences

If the experimenter is interested in testing the activity in particular voxels defined before the experiment (an anatomical a-priori hypothesis) which greatly reduces the number of tests performed, then correction for multiple comparisons across the whole brain is inappropriate (too conservative, carries the risk of false negatives) and a less stringent correction is used by correction for multiple comparisons across all voxels of interest. A correction taking into account spatial smoothness can be used if the voxels of interest are defined by their location in a defined search volume (Worsley, Marrett, Neelin, Friston, & Evans, 1996), this is implemented in SPM under the name of Small Volume Correction.

The search volume can be defined in a stored Analyze format image or defined using the SPM interface as a spherical or cubical volume centered on a voxel coordinate. This approach allows testing for significant clusters or voxels within the defined volume (to answer questions such as: "does contrast X yield significant activation increases in the medial orbitofrontal cortex?"). Anatomical definition of a volume can be achieved by drawing the extent of a structure of interest over a reference brain, either from one subject or a mean image from a group of normalised structural brain images. The MRICro software package (Rorden & Brett, 2000) is very useful for this approach and was used repeatedly in the analyses described in this thesis.

In some practical cases however, the number of voxels tested is relatively small and uncorrected p values are reported, albeit at a very stringent threshold (such as  $P < 0.001$ ).

### *Parameter estimate extraction from voxels of interest*

Instead of relying on an SPM-based inference method, a simpler approach consists in extracting the parameter estimates from all voxels defined by a list of voxels for subsequent analysis with a statistical software package such as SPSS. As in the SVC approach described above, the volume of interest can be defined in an Analyse format image or as a 3-D volume centered on a voxel of interest. This approach allows testing for significant clusters or voxels within the defined volume (to answer questions such as: "do the voxels in the medial orbitofrontal cortex show on average a significant activation increase in contrast X?"). A software tool for easy extraction of data from the same set of voxels across a whole group of subjects was created for these analyses (ROItool, unsupported SPM add-on).

### Inferences at the population level: Random Effects

Inferences from the analysis of fMRI time-series can be made either for a single subject or for a representative sample of the population of interest. Inferences of single subject data are drawn from the effect size relative to the within subject variability (fixed effects analysis). This type of analysis is an extension of a case report, commonly used in clinical studies and animal lesion experiments, where an effect is observed in a particular subject and then replicated in further subjects. In the fMRI community, this type of analysis is commonly used for psychophysical studies.

To extend the level of analysis to draw inferences at the level of the population (say, "healthy adults") from which the tested subjects were drawn, one can perform a second level of statistical tests with the results of a comparison of interest in each subject. Such "Random effects" analyses are typically a one-sample t-test testing whether the estimated effect size is significantly greater than zero across all subjects.

This is done by entering the results of the comparison of interest from each subject (a weighted sum of parameter estimate maps from a 1<sup>st</sup> level analysis, a "contrast") into the t-test. As in any t-test, the average effect size is compared against the between subject variability in these contrasts, and the degrees of freedom of the test depend on the number of subjects scanned. The important variables for a multi-subject fMRI experiment are the percent BOLD signal change between conditions, and the variability in percent signal change between subjects. In an effort to find a reasonable, statistically motivated guideline sample size for fMRI experiments, Desmond and Glover (Desmond & Glover, 2002) estimated both the average percent signal change and the average within- and between-subject variability from previously acquired fMRI data. Using these data they performed simulations of experiments to assess variations in statistical power observed with different numbers of trials and subjects, and validated the results of these simulations by comparing them to measured data. They found that for a threshold of  $p=0.05$ , 12 subjects are required to achieve 80 percent power at a single voxel for typical activations. Based on these and other calculations, 12 subjects has become a commonly used sample size in cognitive fMRI studies. All analyses presented in this thesis were random effects analyses with 12 subjects or more.

An interesting alternative to t-tests is available in SPM thanks to recent developments. There are often more than one effect of interest in an fMRI experiment, and instead of running separate one-sample t-tests for each contrast of interest, one can conduct a repeated measures ANOVA, entering one observation for each condition of interest for each subject. Using a single ANOVA, one can then test all contrasts of interest. The statistical hurdle that needs to be overcome for such an analysis is the correction for non-sphericity in the data variance. This stems from the

likely possibility that variance between the data of all conditions tested in one subject is not equal to the variance between the data of one tested condition in all subjects. SPM solves this problem in the following way. The non-sphericity, once determined, can be specified in terms of constraints, which can be used to estimate covariance components and correlations with a ReML (restricted maximum likelihood) algorithm. These ReML estimates are then used to correct for non-sphericity during inference by adjusting the statistics and degrees of freedom appropriately (the parameter estimates will not be affected because SPM uses an ordinary least squares estimator in this case instead of a Gauss-Markov estimator). An ANOVA was used for analysis of Experiment 3 in this thesis.

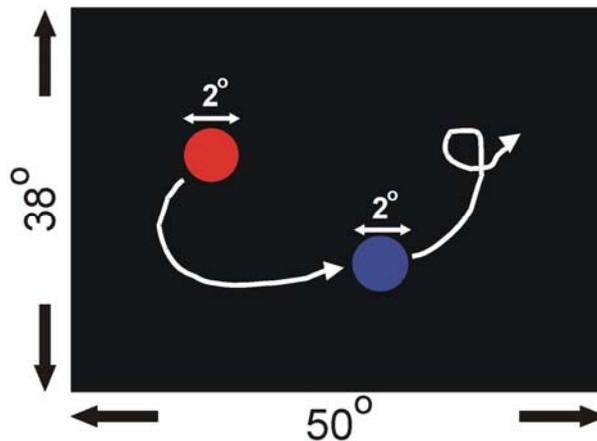
#### Threshold used in this thesis

All activation changes reported in this thesis survived a threshold of  $P < 0.05$  corrected for multiple comparisons across the whole brain or an anatomically constrained search volume. Inferences were drawn mostly at the cluster level. In particular cases of strong a-priori hypotheses, a threshold of  $p < 0.001$  uncorrected was used.

## Part 4. Experiments

### Overview

In the first two experiments, I presented healthy adult volunteers with animations of two round shapes moving in a seemingly animate way (Figure Exp 1.1). For experiment 1, I developed a movement equation in which I could vary the percept of animacy by varying the amount of goal-directed motion expressed by the two abstract moving objects. In the fMRI experiment, subjects had to rate the amount of interaction between the moving objects and their speed (which was manipulated independently) in separate sessions. In experiment 2, I manipulated the strategy used to reach the goal: agents either seemed to use knowledge of the goals attributed to the target object or simply to follow the target object. Stimuli were controlled for speed and quantity of movement, and eye movements were monitored in experiment 1. Experiment 3 was different from Experiments 1 and 2 in that I used animations based on digitised movements of actual human actions, this will be introduced later.



**Figure Exp 1.1. Example of display.**

White arrows represent a segment of the object's motion and were not shown during the experiment.

## **Chapter 1. Experiment 1: Parametric study of animate motion**

### Movement equation

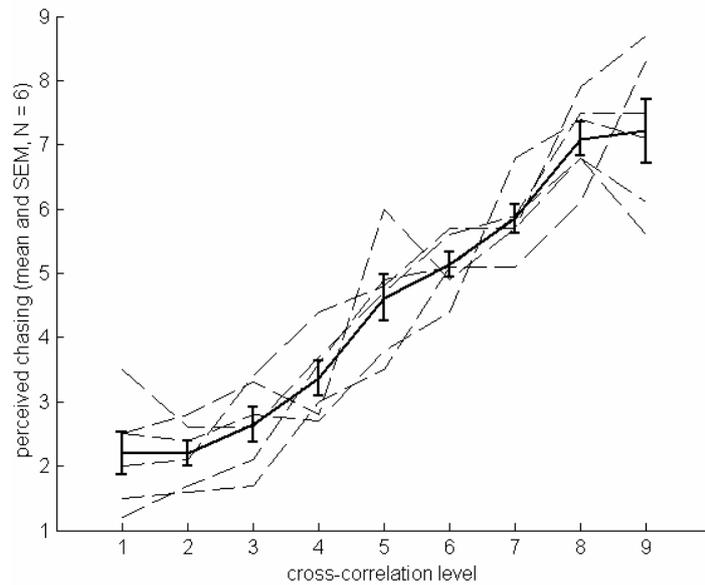
#### *Development*

As described in the Introduction, a number of studies have been performed with infants and adults in order to identify movement characteristics leading to the attribution of animacy to objects in the environment (e.g.: Opfer, 2002; Csibra, 2003; Johnson, 2003). But no studies seem to have formalised an equation producing a movement perceived as animate, or identified objective parameters correlating with animacy, although such a parameter would allow more systematic studies of the perception of animate movements, including parametric brain activation studies.

The theoretical basis for the development of my algorithm were experiments by Michotte (Michotte, 1946), Heider and Simmel (Heider & Simmel, 1944), Csibra (Csibra, 2003), Opfer (Opfer, 2002) and Johnson (Johnson, 2003). Previous neuroimaging work from our group, by Blakemore (Blakemore et al., 2001; Blakemore et al., 2003) and Castelli (Castelli, Happe, Frith, & Frith, 2000), also paved the way. Critically, personal observations by Karl Friston revealed that a Multivariate Autoregressive Model (MAR) embodied in Matrix Exponentials could easily yield movements that appeared animate. It was decided to opt for a display with two objects to simulate a dyadic interaction, and after a prolonged piloting period, it was observed that by varying the amount of randomness in the object motion, and adding a combination of sine waves to give an underlying shape to the movements so that they appeared self-propelled (an important parameter for attribution of animacy, see Tremoulet & Feldman, 2000; Scholl & Tremoulet, 2000), an apparently convincing animation of an interaction between two animate-looking objects could be

produced. The different variables were chosen so that acceleration changes, turning angles, and general shape of the movements appeared biological to the experimenters. By manipulating a cross-correlation matrix such that parameters with opposite signs would drive the movements of the two objects, a chasing scenario could be created: one object moved towards the second object (the target), which moved away from the first (the chaser). The amount of cross-correlation affected the amount of chasing.

A first behavioural pilot study using 9 different levels of chasing revealed that the amount of chasing could be easily determined by 6 subjects naïve to the real purpose of the experiment (see Figure Exp 1.2). Out of these 9 levels, four linearly spaced levels of the cross-correlation parameter were chosen for subsequent animations.

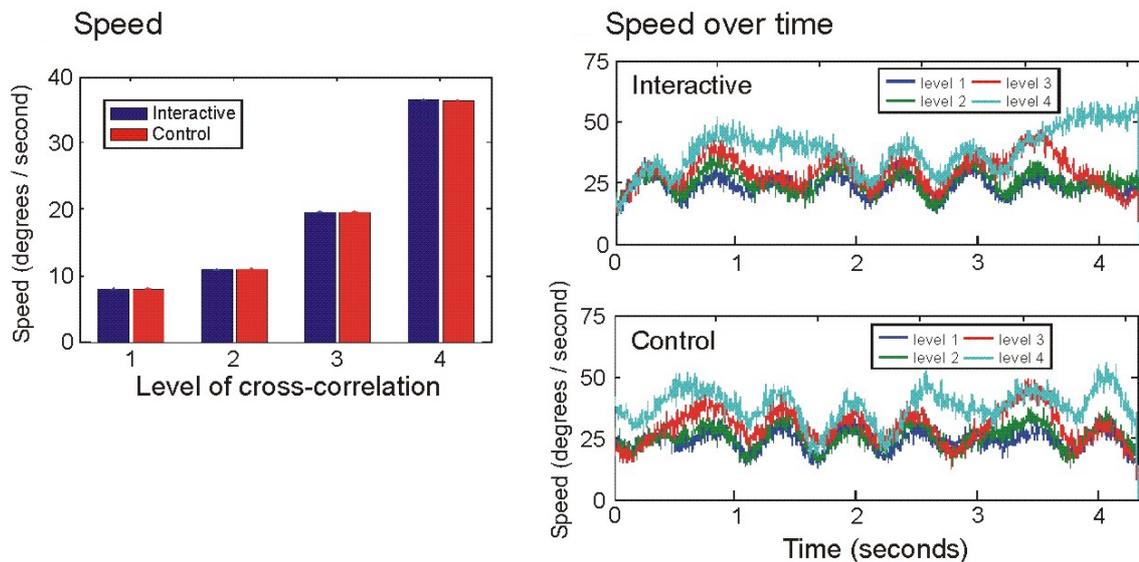


**Figure Exp 1.2. Results from the behavioural pilot study.**

6 Subjects were asked to rate "how much does the red object appear to chase the blue object?" on a continuous scale from 1-9. There were 9 linearly spaced levels of cross-correlation. Plotted are average ratings from 10 trials per cross-correlation level, for each subject (dotted lines) and averaged over subjects (thick line, mean  $\pm$  SEM).

## Speed

After this initial phase, it was observed that an increase in the cross-correlation parameter did not only increase the percept of chasing, but also increased speed of both objects (speed depends on value of the cross-correlation parameter, see Figure Exp 1.3). This is due to a higher a cross-correlation coefficient resulting in greater dependence between the object's movements, making the chasing object follow the target more clearly, and the target to move away from the chaser more clearly. These counteracting effects of increasing the attraction/repulsion behaviour increased the movements of the objects, in terms of speed and area of the screen covered.

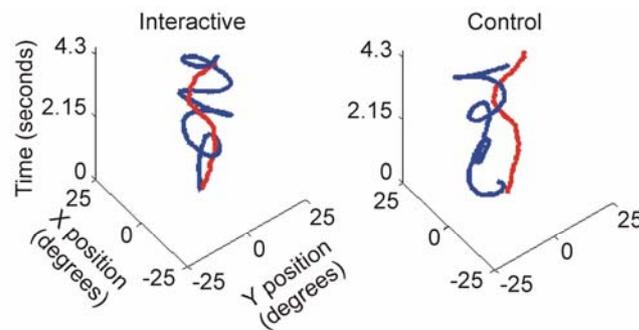


**Figure Exp 1.3. Speed.**

Speed changes of the object's movements due to variations in cross-correlation level. **Left panel:** Average speed (mean and standard error of all trials) of both objects in the 8 different types of films used in the experiment. Speed increased with the level of the cross-correlation parameter, but was almost identical in the Interactive and Control trials. **Right panel:** Average speed of both objects over time in all trials of each level of the cross-correlation parameter, and in their Control conditions: the speed profile over time was very similar in the Interactive and the Control trials.

To control for this potential confound, it was decided to create matched control animations, where the speed and general movement characteristics were conserved but the correlation between the objects' motion was destroyed. To this end, the

movement path of one object (the target, blue object) was inverted in time, and movements in the X and Y axes were swapped. This manipulation destroyed the dependency between the objects but retained exactly the same motion dynamics otherwise (Figure Exp 1.4).



**Figure Exp 1.4. Interactive and Control animations.**

Example of the object's motion in time in an Interactive and in a Control trial: the movement of the red object was kept identical across these trials but the movement of the blue object was changed in the Control conditions, destroying the interactivity between the two objects but conserving the speed of both objects.

### *Pushing and pulling*

A potential confound with chasing objects is that objects that track each other too well might appear to be physically linked to one another in some way, resulting in the impression of objects pushing or pulling each other. This effect is small in my animations because the movements of both objects were not only determined by each other's positions but also had their own movement characteristics. The absence of a pulling or pushing motion is known to reinforce the sense of goal-directed motion (Opfer, 2002). Pulling and pushing can be reduced by 1) a delay between a movement of the target and the movement of the follower and 2) the follower changing its direction differently than the target (Opfer, 2002). My stimuli conformed to both these factors and therefore increased the impression of interactive movements.

### *Final movement equation*

The final equations of motion specified a time-series of positions for each object, where the new position of each object was determined by the previous position of both using a multivariate autoregressive process (MAR). The influence of the objects on one another was parametrically varied according to a cross-correlation parameter, whereas the influence of previous positions of the same object was kept constant. Both objects had their own movement characteristics, such that the blue ball always moved faster than the red ball. The movement equation consisted of a set of differential equations with a cross-correlation matrix containing terms controlling the influence of each object's previous coordinate on its new coordinates (these terms were kept constant during the experiment), and terms controlling the influence of the other object's previous coordinates on the new coordinates of each object. These equations were integrated using matrix exponentials to give a MAR time-series. The cross-correlation parameters were identical for the two objects except for their opposite sign which made one object appear as the chaser and the other as the target. To give the objects a basic movement that appeared biological, the equations of motion included an exogenous component (combination of sine waves with different periods and a small random term). The update equation for each time step was

$$x(t + \Delta t) = \exp(J\Delta t)x(t) + \sum_i \beta_i \sin(t\omega_i) + W\varepsilon(t)$$
$$J = \begin{bmatrix} -0.01 & 0 \\ 0 & -0.01 \end{bmatrix} \otimes \begin{bmatrix} 1 & \rho \\ -\rho & 1 \end{bmatrix}$$

Our exogenous driving terms were controlled by  $\beta = [1/7, 1/10, 1/2, 2/3]$  and  $\omega = [1/100, 1/200, 1/50, 1/40]$ . In this equation,  $x(t)$  are the coordinates of both objects,

$x(t-1)$  are their previous coordinates,  $J$  is the system's Jacobian controlling the dependencies,  $W$  is a constant that scales the random term  $\varepsilon(t) \sim N(0, \Delta t)$  and  $\Delta t$  is the time step between two successive positions.  $J$  was based on a cross-correlation matrix containing the cross-correlation coefficient  $\rho$ , which was modified across conditions and increased the amount of interactivity. This resulted in the impression of the red ball chasing the blue ball, in a parametric fashion.

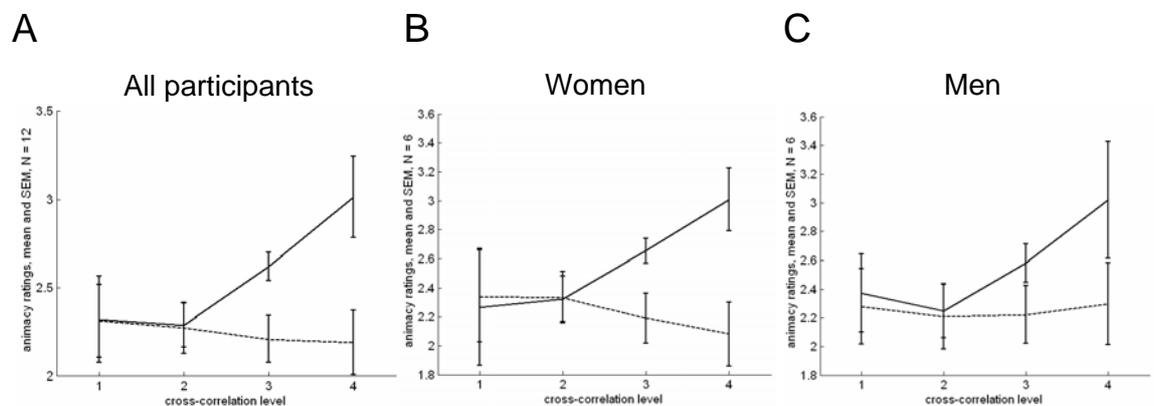
### *Animacy ratings*

Our parameter controlling the amount of animacy in the observed animations being relatively abstract and the stimulus development process being long and involving many steps, it was necessary to make sure that the final animations still appeared animate. To this end, I asked 12 volunteers (6 male and 6 female, aged 23-36), who were unaware of the different experimental conditions, the parameters used and of the aim of the experiment, to judge how "alive" the objects appeared, on a scale of 1-4. Results (Figure Exp 1.5) indicated that animations in the Interact conditions appeared more animate than the Control animations ( $F(1,10) = 11.3$ ,  $p < 0.01$ ; repeated-measures 2-way ANOVA, within-subject factors Cross-correlation level and Interact vs. Control, between-subjects factor gender), and importantly that animacy ratings increased more with the cross-correlation level in the Interact conditions compared to the Control conditions (interaction,  $F(1,10) = 24.3$ ,  $p < 0.001$ ). There was no difference due to gender ( $F(1,10) = 0.001$ ,  $p > 0.5$ ), and no influence of gender on either contrast (interaction Interactive vs. Control by gender:  $F(1,10) = 0.1$ ,  $p > 0.5$ ; 3-way interaction:  $F(1,10) = 1.2$ ,  $p > 0.2$ )<sup>4</sup>. Behavioural reports (see Appendix

---

<sup>4</sup> The design of this experiment was identical to the fMRI protocol: there were eight experimental conditions, 10 repetitions each. Subjects were shown one example of each condition before the experiment, in random order and without explanation. Ratings were averaged over trials and then

1) indicated that many subjects consciously used the chasing behaviour to rate animacy, while others were not aware of this, and others still (small number) discounted fast objects as appearing inanimate and considered "staying together" as relevant factor for animacy. Although subjects did not know at first which strategy to use or even changed strategy, no significant change over repetitions of trials from a condition was found in the ratings (using a repeated-measures ANOVA with factors: cross-correlation level (1-4), interactive vs. control, gender (m/f) and repetition (1-10)).



**Figure Exp 1.5: Animacy ratings.**

From 12 subjects naïve to the purpose and design of the experiment (A), including (B) 6 women and (C) 6 men. Continuous line: ratings for the Interactive conditions (i.e. with chasing); dashed line: ratings for the Control conditions (i.e. with destroyed correlation). See text for statistics.

---

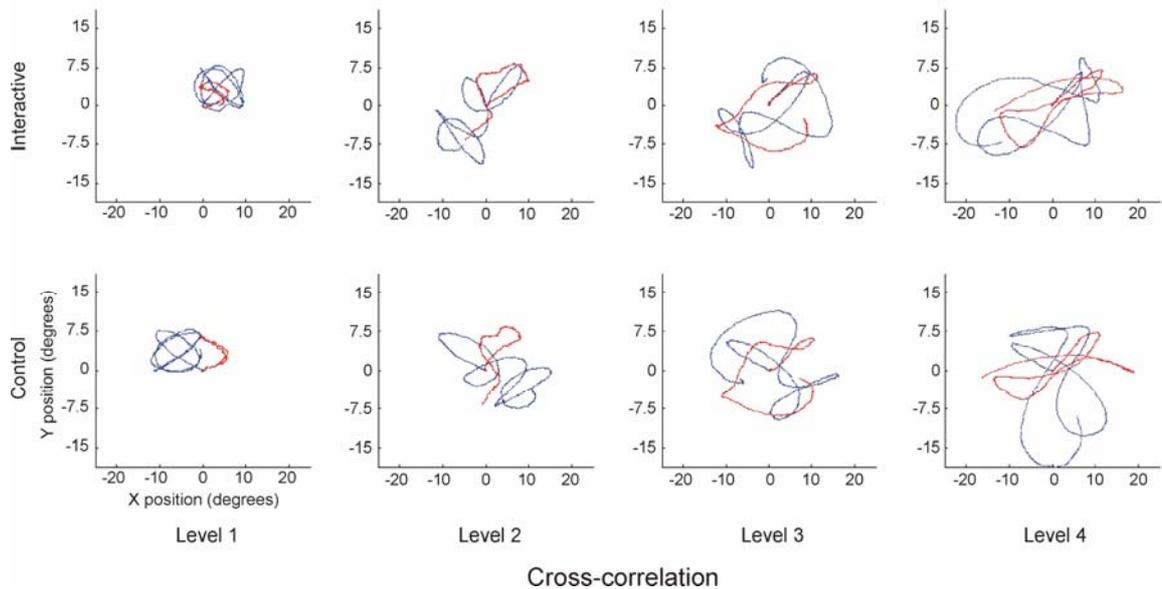
compared across subjects, identically significant effects were found when comparing medians across subjects.

## Methods

### *Design and conditions*

In this experiment I wanted to identify regions whose responses increased in relationship with the amount of interactivity between two objects, a movement characteristic which correlates with attribution of animacy to these objects. I also wanted to test whether these activation changes happen only when subjects pay attention to the interactions between the objects or also when they perform another, incidental task, for which animacy is irrelevant. As increases in interactivity are associated with an increase in the speed of both objects, I used matched control conditions with identical speed and similar, but uncorrelated movements. I therefore used a factorial design with the following three factors: 1) four, linearly increasing levels of interactive motion, 2) two task levels: an interactivity- and a speed-rating task, and 3) two condition levels: Control and Interactive. Combinations of all these factors resulted in 16 different trial types (Figure Exp 1.6), each repeated 10 times during the experiment. The four levels of interactivity were created by manipulating a cross-correlation parameter, which controlled the dependence between object movements. Increasing this parameter increased the objective and subjective interactivity and their speed. I therefore used a matched Control condition for each interactivity level, in which speed was identical but interactivity was destroyed (Figure Exp 1.4). Subjects performed the two tasks on the same stimuli: in the interactivity-rating task, subjects were asked: "how much does the red object follow the blue object, one being the minimum and four the maximum?" In the speed-rating task, the instructions were "how fast do the objects move, one being the minimum and four the maximum?". They responded by pressing one of four buttons on a keypad with the corresponding finger of the right hand. Ratings, response times, eye

movements and brain activation were recorded simultaneously during the experiment. On these four types of dependent variable I assessed the effects of interactivity, and the interaction between task and interactivity.



**Figure Exp 1.6. Example of the 8 conditions.**

Example of movements in the eight different animations used in both tasks of the experiment, yielding 16 different experimental conditions. Plotted are X and Y positions of the red disk (the Chaser) and the blue disk (the Target) in visual degrees. The extent of the screen covered and the movement shapes can be seen, but as the time dimension is lost on these 2-dimensional plots, interactivity cannot be judged.

### *Stimuli and films used*

In the scanner, subjects watched short films (4.3 sec. per film) in which two moving disks appeared to be either interacting or to move independently from each other. The moving objects were two disks or balls, 2 degrees wide on the screen (Figure Exp 1.1). One was red, the other blue, and the background was black. The movement trajectories of the two disks were determined prior to the experiment by an equation of motion implemented in Matlab (The MathWorks, Inc, Natick, MA) and detailed below. This equation made the objects move in a way that appeared biological and, in addition, allowed parametric control of the interactivity between the

two objects by varying the correlation between their movements. The visual display was controlled by in-house presentation software (Cogent 2000, <http://www.vislab.ucl.ac.uk/Cogent2000/index.html>) implemented in Matlab. Films were projected onto an opaque surface in the scanner by an LCD projector, subjects viewed them through a set of mirrors mounted on the headcoil.

### *Eye movements*

Subjects were allowed to move their eyes freely during the experiment. To evaluate potential confounds due to eye movement differences between conditions, I recorded eye movements using an infrared eye-tracking system recording at 60 Hz (ASL Model 504, Applied Science Laboratories, Bedford, MA), with remote, custom-adapted optics for use in the scanner. Reliable eye tracking data throughout the whole scanning session were only available in nine out of twelve subjects as the remaining three subjects wore contact lenses, which created artefacts in the eye-tracker recordings. Eye blinks were removed by eliminating all differences in successive time points more than 3 standard deviations away from the mean difference, which were replaced by the average of the positions immediately before and after the replaced time point. The data in X and Y were mean-corrected, squared, and then summed to yield a measure of total eye-movements (saccades and smooth pursuits were not distinguished) for each trial. The results of this procedure were analysed with the same tests as those used in the analysis of the subjects' ratings.

### *FMRI analysis: SPM whole-brain analysis*

The analysis followed the standard steps as described in the Materials and Methods section of the present thesis. Realignment was standard, no slice-time

correction was used, normalisation was done with resampling to a voxel size of 3 mm<sup>3</sup>. To enhance the signal to noise ratio and enable intersubject functional anatomical comparison the images were smoothed by convolution with a 6-mm full width at half maximum (FWHM) Gaussian kernel. A high-pass filter (using a cut-off of 128 sec) and a correction for temporal auto-correlation in the data (AR 1 + white noise) were applied to accommodate serial correlations. The design matrix of the first-level (single subject) analysis comprised 32 regressors of interest, resulting from the crossing of the following factors: 1) task (interactivity-rating or speed-rating), 2) value of cross-correlation parameter [1, 2, 3 or 4] and 3) condition (Interactive or Control trials), which yields 16 effects of interest; each was modelled using the canonical hemodynamic response function (HRF) and the first temporal derivative. Two covariates of no interest were included, made by convolving the average speed of both objects during each trial with the HRF and its temporal derivative. This enabled us to model separately the effect of the average object speed on the brain activation, which was not an effect of interest in the analysis. Differential realignment parameters were modelled as additional regressors of no interest to model movement-related artifacts. Second-level analyses were performed as T-tests, after smoothing the contrast images from the first-level analysis with an 8-mm FWHM kernel. The threshold used for statistical significance was  $P < 0.05$  corrected for multiple comparisons across the whole brain with inference at the cluster level.

#### *FMRI analysis: regional analyses*

In addition to the classical whole-brain voxel-wise analysis (see Methods section), I performed regional analyses in the following regions of a priori interest: the posterior superior temporal sulci and fusiform gyri of both hemispheres. These

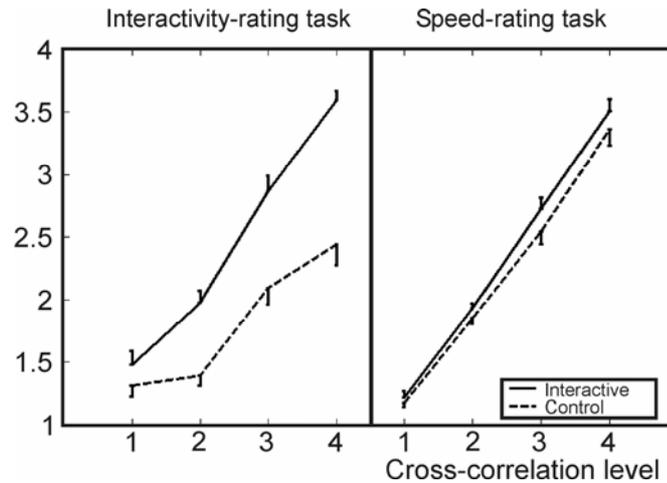
regions of interest (ROI) were based on the average structural image from the participants (see below), determined by reference to a brain atlas (Duvernoy, 1999) and drawn with MRIcro software (Rorden & Brett, 2000) (<http://www.cla.sc.edu/psyc/faculty/rorden/mricro.html>). They encompassed coordinates of activation clusters found in several previous studies (see below). The region of interest in the left fusiform gyrus extended from  $-52$  to  $-17$  mm in the X dimension, from  $-72$  to  $-24$  mm in the Y dimension, and from  $-32$  to  $-7$  mm in the Z dimension. The region of interest in the right fusiform gyrus extended from  $29$  to  $48$  mm in X, from  $-59$  to  $-29$  mm in Y, and from  $-31$  to  $-7$  mm in Z. Volumes were  $8856$  mm<sup>3</sup> (left) and  $8821$  mm<sup>3</sup> (right). The region of interest around the left posterior superior temporal sulcus extended from  $-68$  to  $-53$  mm in X, from  $-66$  to  $-34$  in Y, and from  $-29$  to  $0$  in Z. The region of interest in the right superior temporal sulcus extended from  $50$  to  $70$  mm in X, from  $-65$  to  $-37$  mm in Y, and from  $-29$  to  $-2$  mm in Z. Volumes were  $9234$  mm<sup>3</sup> (left) and  $10422$  mm<sup>3</sup> (right). These regions of interest are displayed in Figure Exp 1.11. Parameter estimates corresponding to the voxels in the ROI were selected with in-house software based on the SPM2 software package. These data were then averaged over all voxels within the ROI and tested for the comparisons of interest (effects of a linear increase in goal-directedness, and task effects on an increase in goal-directedness) using a repeated-measures 3-way ANOVA in a statistics package (SPSS 11.0, SPSS Inc., Chicago, Illinois USA).

## Results

### *Behavioural results*

#### **Ratings**

Analysis of participants' ratings from the interactivity-rating task and the speed-rating task performed in the scanner revealed the following effects (Figure Exp 1.7). In the Interactive trials (where the movements of the two objects were correlated), subjects rated the objects as increasingly interactive with increasing cross-correlation between their movements (repeated-measures ANOVA,  $F(1,11) = 502.7$ ,  $P < 0.001$ ). However, subjects also considered the objects' speed to be increasing with increasing cross-correlation ( $F(1,11) = 446.8$ ,  $P < 0.001$ ). Matched control films, in which the correlation between the objects' movements was destroyed but the speed remained almost identical, elicited no significant difference in measured object speed relative to the Interactive trials:  $F(1,11) = 0.001$ ,  $p > 0.9$ . Figure Exp 1.3). The Control films were rated by the subjects as displaying much less interactivity than the Interactive films (interaction between cross-correlation level and Interactive vs. Control films,  $F(1,11) = 29.4$ ,  $P < 0.001$ ) but as having the same speed ( $F(1,11) = 1.0$ ,  $p > 0.3$ ). These data suggest that increasing the dependency between the movements of our two objects increased the percept of interaction.



**Figure Exp 1.7. Interactivity and Speed ratings.**

Subjects' ratings in the interactivity-rating task (left), and in the speed-rating task (right). On the X axis are the different levels of the cross-correlation parameter that controlled the interactivity between the objects' movements from 1 (minimum) to 4 (maximum), repeated for each task. Broken and unbroken lines correspond respectively to Interactive conditions (with interactive motion) and to their matched Control conditions (where the interactivity between the movements was destroyed, see Methods). On the Y axis are the mean and standard errors (over subjects) of the subjects' ratings: 1 corresponds to the minimum interactivity or speed, 4 corresponds to the maximum.

### Response times

Response times are shown in Table Exp 1.1 and Figure Exp 1.8). Response times were shorter in the Speed-rating task than in the Interactivity-rating task ( $T(1,11) = 3.971$ ,  $p < 0.005$ ). As subjects were free to answer at any point of the trial, I suppose that this is because subjects needed to observe the animations only for a short amount of time to determine the speed of the objects, as the speed did not vary too much during the time course of the animation. For the Interactivity task, a number of cues had to be collected before answering, and these were not distributed equally during the course of the animation. These plots also revealed a "bell-shaped" variation of the response times in the Speed-rating task when plotted with respect to the level of

contingency (see Figure Exp 1.8). I attribute this to the fact that the trials with extreme cross-correlation values were easier to categorise in terms of speed than the "middle" ones, for which subjects had to make a longer decision to attribute the trial to one of the middle categories.

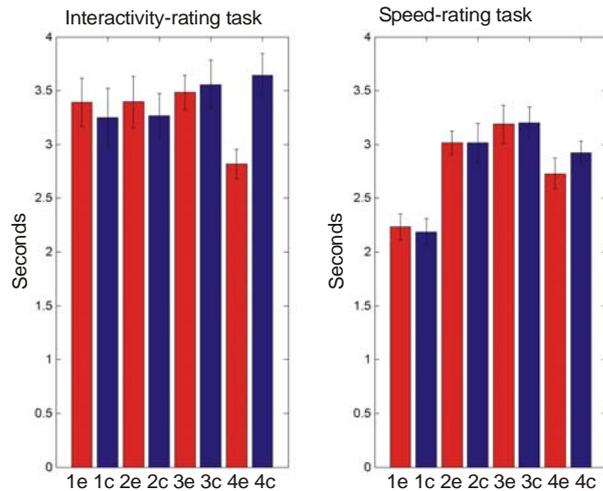
Speed-rating response times increased with cross-correlation level (linear:  $F(1,11) = 34, P < 0.001$ ) but were not different in Interactive or Control trials ( $P > 0.5$ ), and there was no interaction effect between the 2 factors. Plotted, these data show that the significant effect results mainly from the responses in conditions with the smallest cross-correlation level being markedly slower than the others. I suggest that this is because animations with a slow speed were very easy to differentiate from animations with a higher speed (because in contrast to animations with a higher speed, movements were limited to only a relatively small part of the screen, see Figure Exp 1.6), but to differentiate animations with cross-correlation levels 2-4 from each other, subjects had to spend some time to appreciate the movement details before making their response.

Interactivity-rating response times did not vary linearly with cross-correlation level ( $p > 0.5$ ), and were not different in Interactive or Control trials ( $P > 0.5$ ), but responses times in the Control trials increased more with cross-correlation than in Interactive trials (interaction between linear increase in cross-correlation and Interactive vs. Control trials:  $F(1,11) = 25.8, P < 0.001$ ). Again, the plotted data reveal that it is particularly one condition that contributes mostly to the effect: the very fast response in Interactive trials with maximal cross-correlation (Figure Exp 1.8: interactivity-rating, condition 4e). This could very well be explained by the following: when objects moved fast, there were more interactive "events" (due to clear correlation between movements), which would have helped subjects make their mind

up quickly when they saw them. In the matching control trials, subjects did not see these events, and spent some time searching for them before giving a low interactivity rating. With respect to analysis of the fMRI data, in which I used this interaction as contrast of interest, this suggests that the effects I observed might have been slightly different if the subjects had spent the same amount of time on all trials. But as I pooled over both tasks (in the 2-way interaction) or compared between them (3-way interaction), I believe the different effects in the response times of the two tasks compensate each other mostly, thereby reducing false-positives and false-negatives.

**Table Exp 1.1:** Response times (12 subjects)

<b><i>Response times</i></b> (seconds)	<i>Level 1</i>	<i>Level 2</i>	<i>Level 3</i>	<i>Level 4</i>
1) Explicit task				
Interactive	3.4(.22)	3.4(.24)	3.5(.16)	2.8(.13)
Control	3.3(.27)	3.3(.21)	3.6(.23)	3.6(.20)
2) Implicit task				
Interactive	2.2(.12)	3.0(.11)	3.2(.18)	2.7(.14)
Control	2.2(.13)	3.0(.19)	3.2(.15)	2.9(.11)



**Figure Exp 1.8. Response times.**

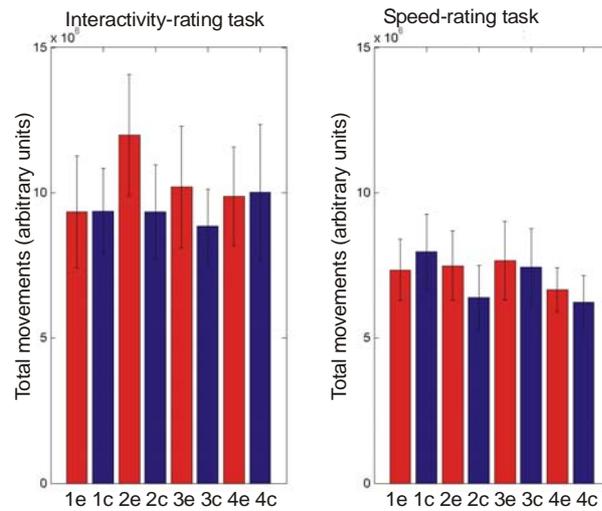
From 12 subjects in all conditions of both tasks, acquired during the fMRI experiment. 1e to 4e correspond to Interactive trials with cross-correlation levels 1 to 4, 1c to 4c correspond to the matched control conditions.

### Eye movements

As subjects were free to move their eyes while watching the animations, I recorded their eye movements to test for differences between conditions. To this effect, I recorded excursion in X and Y and computed the sum-of-squares of total excursion for each condition, for each subject (see Table Exp 1.2 and Figure Exp 1.9). The degree of cross-correlation between the objects' movements did not influence eye movements (repeated-measures 3-way ANOVA,  $F(1,8) = 0.7, p > 0.4$ ), and there were no differences in eye movements between Interactive and Control trials ( $F(1,8) = 2.0, P=0.2$ ). The only factor that significantly influenced eye movements was the task performed by the subjects: subjects moved their eyes less in the speed-rating task than in the interactivity-rating task (interactivity-rating task vs. speed-rating task:  $F(1,8) = 5.8, P=0.04$ ). As I did not compare directly the two tasks, this difference in eye movements did not influence the interpretation of the brain activation analysis.

**Table Exp 1.2:** Eye movements (summed over time, 9 subjects)

<u>Eye movements</u> (arb. units)	Level 1	Level 2	Level 3	Level 4
1) Explicit task				
Interactive	9.3(1.9)	12(2)	10(2)	9.9(1.7)
Control	9.4(1.5)	9.3(1.6)	8.8(1.3)	10(2.3)
2) Implicit task				
Interactive	7.3(1.1)	7.5(1.2)	7.7(1.4)	6.6(0.8)
Control	8.0(1.3)	6.4(1.2)	7.4(1.3)	6.2(0.9)



**Figure Exp 1.9. Eye movements.**

From all conditions of both tasks for 9 subjects, acquired during the fMRI experiment. 1e to 4e correspond to Interactive trials with cross-correlation levels 1 to 4, 1c to 4c correspond to the matched control conditions.

## *Imaging results*

### **Whole-brain**

I used the parametric, factorial design of this experiment to test for activation increases in relation to increases in interactivity displayed by the two objects, and to test how these interaction-dependent responses were affected by the attentional task the subject performed. To discount speed effects I assessed interaction-sensitive activations as the difference between the regression of brain responses on interaction level and the equivalent speed control trials.

#### - Main effects

I initially tested for regions with a linear increase in activation correlating with an increase in the cross-correlation parameter. The regions that showed such an effect were the fusiform gyrus and the medial occipital cortex, in both hemispheres (see Table Exp 1.3). I also looked for regions with a greater activation in Interactive conditions than in Controls, and found significant activation increases in the left medial occipital cortex (Table Exp 1.3). Both of these main effects did not interest us particularly, as the comparison of interest was the interaction between Cross-correlation level and Interactive vs. Control conditions.

**Table Exp 1.3:** Brain activation data, whole-brain voxel-wise analysis. Significant clusters in the comparisons of interest, surviving a threshold of  $P < 0.05$  corrected for multiple comparisons across the whole brain (see Methods), at cluster level (\*) or voxel level (FWE, \*\*).

<i>Structure</i>	<i>Coordinates</i>			<i>Size</i> ( $mm^3$ )	<i>Z score</i>	<i>Brodmann</i> <i>area</i>
Contingency (linear increase)						
L fusiform	-24	-75	-12	7929	5.90**	19

L medial occipital cortex	-6	-84	3		5.39**	17
R medial occipital cortex	3	-87	18		5.14*	18
R fusiform gyrus	27	-48	-6		3.95*	19
Interactive > control						
L medial occipital cortex	-9	-84	-6	263	3.58 *	18
	-21	-93	-9		3.34 *	17
	-18	-84	-18		3.33 *	18

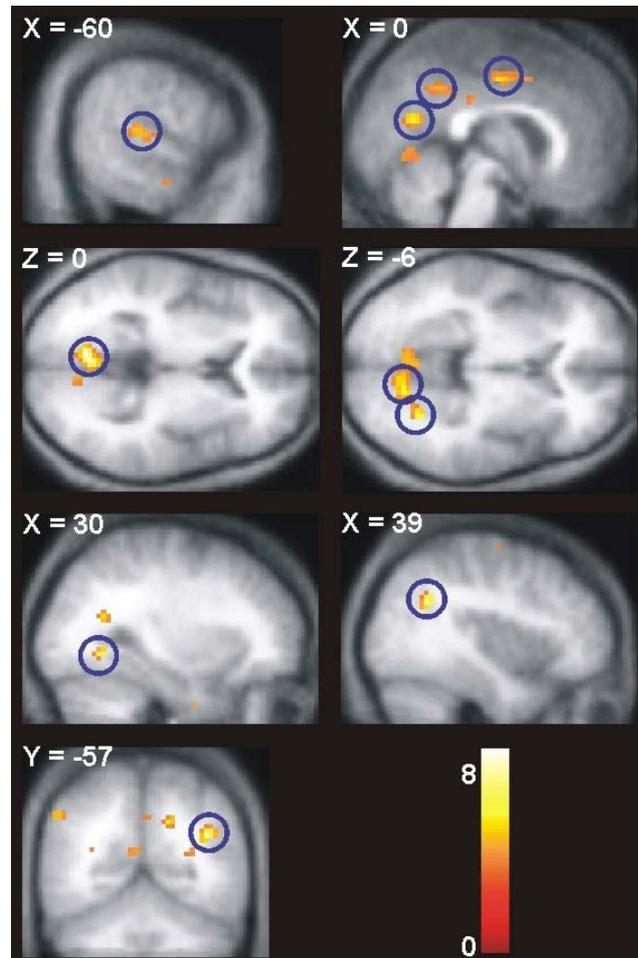
- Effects of a linear increase of interactive motion

The following regions showed response increases with the increase in correlation between the objects' movements in the Interactive condition but no increases in the Control conditions: the right posterior, ascending branch of the superior temporal sulcus in the inferior parietal cortex, the left superior temporal gyrus, the medial occipital cortex in both hemispheres, the right fusiform gyrus, the caudal part of the anterior cingulate gyrus, and the posterior paracingulate gyrus (Figure Exp 1.10, Table Exp 1.4). This was tested formally with a two-way interaction between a linear increase in the cross-correlation factor and the Interactive vs. Control factor.

**Table Exp 1.4:** Interaction. Threshold used was identical to main effect (Table Exp 1.3).

<i>Structure</i>	<i>Coordinates</i>			<i>Size</i> ( <i>mm</i> <sup>3</sup> )	<i>Z score</i>	<i>Brodmann area</i>
Greater linear increase with contingency in the experimental stimuli						
L medial occipital cortex	-6	-69	0	549	4.79 (**)	18

R medial occipital cortex	15	-72	-9		4.26 *	18
R fusiform gyrus	30	-63	-6		4.18 *	19
R asc. STS / inferior parietal	39	-57	27	147	4.51 *	39
R post cingulate	3	0	45	186	4.02 *	24
Medial occipital cortex	0	-66	18	243	3.98 *	23
L STG	-60	-27	9	120	3.63*	41
L inf parietal / angular gyrus	-45	-72	36	78	3.52 ~*	39
Post paracingulate	0	-48	39	117	3.48 *	31



**Figure Exp 1.10. Brain regions whose activation increased linearly with the amount of interactive motion.**

Tested by the interaction between a linear increase in interactivity between objects' movements and Interactive vs. Control trials. For details see Table Exp 1.4. For display, the image is thresholded at  $P < 0.001$  uncorrected for multiple comparisons. Clusters surviving a threshold of  $P < 0.05$  corrected for multiple comparisons across the whole brain are circled in blue, listed in the Table and discussed in the text. The brain slices used for display is the average image from the subjects' normalised structural brain scans (see Methods). Coordinates are in MNI (Montreal Neurological Institute) reference space. The color bar refers to activation intensity expressed in t values.

- Effects of task on observation of interacting movements

I did not find any cluster that showed significantly different activation depending on the task the subjects performed. This was tested with a three-way interaction between a cross-correlation level, Interactive vs. Control and interactivity- vs. speed-rating task.

## Anatomical regions of interest

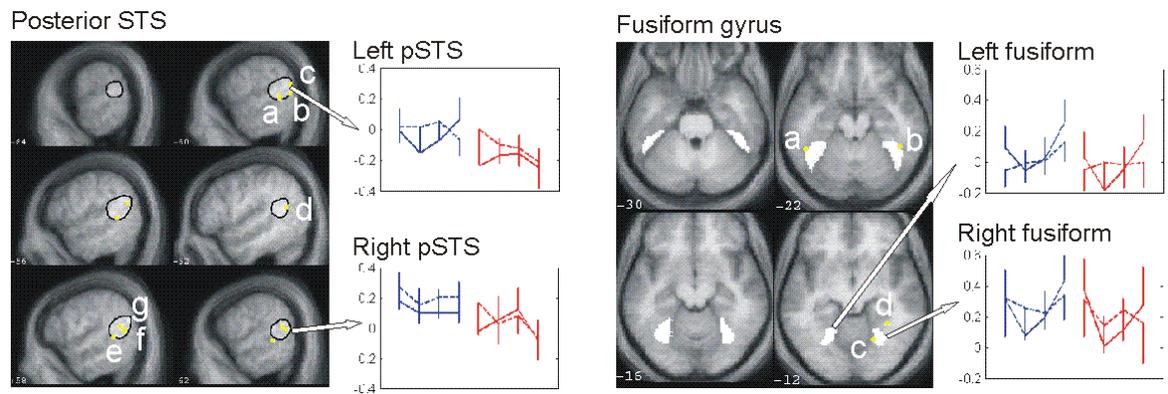
As discussed in Part 1, Section 3.1.2, previous studies using similar types of animations to those used in the present study showed activation in the posterior superior temporal sulcus, the fusiform gyrus, the temporal poles and the media prefrontal cortex (Castelli, Happe, Frith, & Frith, 2000; Blakemore et al., 2003; Schultz et al., 2003). To complement my voxel-wise analysis on the whole brain with analyses based on regions where I might expect activations, I subsequently examined activation in pre-defined anatomical regions of interest: the posterior superior temporal gyrus and fusiform gyrus in both hemispheres (for definition of these areas, see Methods: fMRI analysis, Figure Exp 1.11 and Table Exp 1.5).

**Table Exp 1.5:** Regions of interest, characteristic and significant statistics

<i><b>Region</b></i>	<i>Extent X</i>	<i>Extent Y</i>	<i>Extent Z</i>	<i>Volume (mm<sup>3</sup>)</i>
Fusiform L	-52 to -17	-72 to -24	-32 to -7.2	1104
Fusiform R	29 to 48	-51 to -29	-31 to -7.2	1080
Posterior STS L	-68 to -53	-66 to -34	0 to 29	1053
Posterior STS R	50 to 70	-65 to -37	2 to 29	1212

Activation in the left fusiform region of interest varied with the amount of goal-directed motion displayed by the two objects (quadratic increase with respect to goal-directedness and no increase in the control conditions, tested by interaction between goal-directedness and Interactive vs. control trials:  $F(1,11) = 15$ ,  $P=0.003$ ). Activation in the region of interest in the right fusiform gyrus did not vary significantly with the amount of goal-directedness exhibited by the moving objects. Activation in the region

of interest in the left posterior superior temporal sulcus increased with goal-directedness, but only when subjects rated the interactivity between the objects, not when they rated their speed (3-way interaction between task, contingency level, and Interactive vs. control trials:  $F(1,11)=9.6$ ,  $P<0.01$ ). Activation in the right posterior superior temporal gyrus region of interest was not significantly different across conditions or interactions between them.



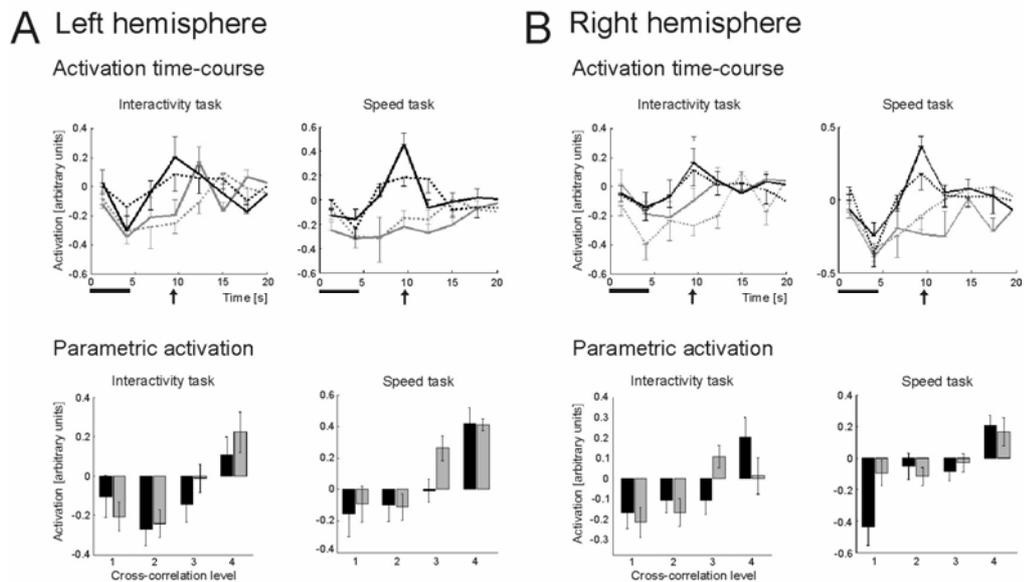
**Figure Exp 1.11. VOI anatomy and data.**

Volumes of interest in the posterior superior temporal sulcus. Encompassing previous studies of animate- and intentional-looking moving abstract objects. Yellow dots represent the maximum intensities of activations found in following two structures: 1) superior temporal sulcus (lettered from a to g, more than one coordinate per study): Blakemore et al, 2001; Blakemore et al, 2003; Schultz et al, 2003; Castelli et al, 2000; 2) Fusiform gyrus (lettered from a to d): Winston et al, 2002; Blakemore et al, 2003; Schultz et al, 2003.

### Activation time-course

Activation time-courses in the superior temporal sulcus showed peaks 10-12 seconds after animation onset (Figure Exp 1.12 A and B, top panels). Activation peaks were higher when the objects were interacting than during the control trials (Mann-Whitney test, interactivity-rating task, cross-correlation level 1:  $U = 42$ , speed-rating task, cross-correlation level 4:  $U = 29$ , both  $P<0.05$  one-tailed). Graphic plots demonstrated the parametric increase of activations in relation to cross-correlation

(Figure Exp 1.12 A and B, bottom: size of activation 9.6 seconds after animation onset).

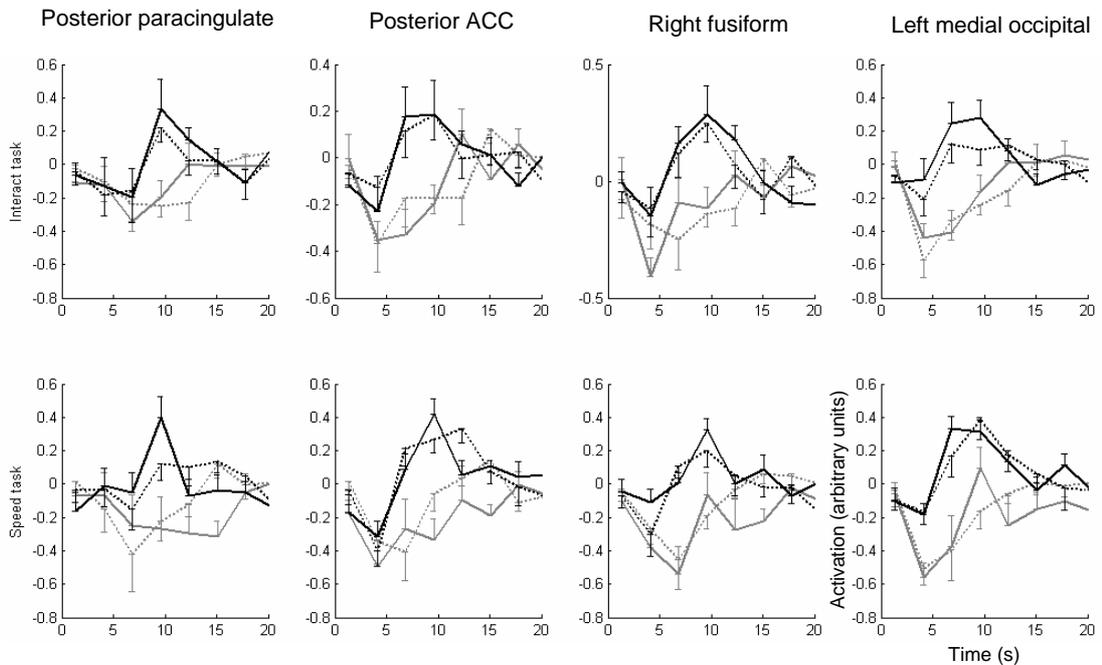


**Figure Exp 1.12. Parameter estimates and time-course data pSTS.**

Activation in the left (panel **A**) and right (panel **B**) posterior superior temporal sulcus during both interactivity-rating and speed-rating tasks. Data were extracted from peak voxels of clusters identified as significantly activated by the SPM analysis, coordinates were:  $-60 -27 -9$  (Left hemisphere, panel **A**) and  $39 -57 27$  (Right hemisphere, panel **B**). X,Y,Z coordinates are in the Montreal Neurological Institute reference frame. Top half of panels **A** and **B**: peri-stimulus time-courses. The black bar below each plot represents time during which the moving objects were shown. Time-point used for parametric activation plot (lower half of figure) is marked by an arrow. Black and grey lines represent trials with cross-correlation level 4 and 1, respectively. Continuous lines represent trials with interacting objects, dashed lines are the matched control conditions. Error bars represent standard error of the mean across subjects. Bottom half of panels **A** and **B**: parametric activation increase with increasing cross-correlation coefficient. Black bars represent trials with interacting objects, grey bars are matched controls. Error bars represent standard errors of the mean. Data are from 4th sampled time-point in top half of pane, indicated by arrows in upper half of the panel (9.6 seconds after onset of animation).

Time-courses in the other areas identified by the SPM analysis were similar, and almost all showed significantly greater activation when the objects were interacting than during the control animations where the interactions were destroyed (Figure Exp

1.13; posterior paracingulate  $U = 41$ ,  $P < 0.04$ ; caudal part of the anterior cingulate  $U = 39$ ,  $P = 0.03$ ; right fusiform gyrus: NS in this analysis but very similar; medial occipital cortex  $U = 42$ ,  $P < 0.05$ ; all tests one-tailed). As the haemodynamic response is expected to peak about 6 seconds after a neuronal event and our animations lasted for 4.3 seconds (black bar in Figure Exp 1.12, top), activation should be highest between 6 to 10 seconds after animation onset. While activation in the right fusiform and in the superior temporal gyrus of both hemispheres peaked at 10-12 seconds after animation onset, the local maximum in the medial occipital gyrus peaked at 7-10 seconds. The later peaks I observed in the superior temporal and fusiform gyri could correspond to a cumulative increase in BOLD response during sustained neuronal activation while the objects were moving. The later peaks may also reflect continued processing even after the objects had disappeared, a process which might not be expressed in the medial occipital cortex.



**Figure Exp 1.13. Time-course of activation in other regions identified in the Interaction contrast (see Table Exp 1.4).**

Data from Interact task and Speed task, respectively in upper and lower row. Coordinates: posterior paracingulate cortex: [0 -48 39], posterior part of anterior cingulate cortex (posterior ACC): [3 0 45], right fusiform gyrus: [30 -63 -6], left medial occipital cortex: [-6 -69 0]. See text above for details and statistics, and Figure Exp 1.12 for additional information.

### Discussion

This study shows that activation in the superior temporal gyrus and in the cortex surrounding the superior temporal sulcus increases linearly with the interactivity between the movements of two abstract, self-propelled objects. Directing attention to the interactions rather than to their speed had no significant impact on activation increases. These cortical structures are known to respond during observation of biological motion, such as point-light walkers (Allison, Puce, & McCarthy, 2000; Pelphrey et al., 2003).

*Does the cortex in the superior temporal sulcus and gyrus respond to objective movement characteristics or to categorical percepts?*

In previous studies, using abstract moving "agents", a human controlled the movements of the abstract shapes to create the desired effects (Heider & Simmel, 1944; Castelli, Happe, Frith, & Frith, 2000; Blythe, Todd, & Miller, 1999; Schultz et al., 2003; Blakemore et al., 2003). Observers can reliably identify which behaviour an animator intends to give to the objects he controls (Blythe, Todd, & Miller, 1999). Animated cartoons on television also speak to the capacity to reliably evoke the impression of complex intentional behaviour in abstract moving shapes. Although it is very interesting that such stimuli induce activation in the superior temporal region and in the fusiform gyrus (Blakemore et al., 2001; Blakemore et al., 2003; Castelli, Happe, Frith, & Frith, 2000; Schultz et al., 2003), it is not clear whether (or which of) these structures respond to objective characteristics of the objects' movements that indicate they might be alive (such as parameters of self-propelled or goal-directed motion), or whether these regions simply respond to anything that appears animate. In my study, I selectively increased the interactivity between the movements of two abstract objects. My results show that activation in certain parts of the superior temporal gyrus and the cortex surrounding the superior temporal sulcus increases in proportion to the amount of interactivity. This suggests that the superior temporal sulcus region is involved in processing movement characteristics that characterise living beings rather than simply responding to the presence of living beings.

#### *Superior temporal sulcus region and social information*

Together with the amygdala, the fusiform gyrus, the orbital and the medial frontal cortex, the superior temporal sulcus region is thought to be part of a network forming the "social brain" (Brothers, 1990; Adolphs, 2003). Neurophysiological recordings in the posterior superior temporal area of monkeys revealed cells that

respond during observation of biological motion such as a walking person, moving parts of a face such as mouth and eyes, or whole faces (Desimone, 1991). Neuroimaging studies of human volunteers yielded similar results (Puce & Perrett, 2003). Pictures and words referring to animals also induce activation increases in the superior temporal area (Martin, Ungerleider, & Haxby, 2000; Kanwisher, McDermott, & Chun, 1997). But this area of the brain is also associated with higher cognitive functions, such as the attribution of mental states to other living organisms (Frith & Frith, 1999; Frith & Frith, 2003; Saxe & Kanwisher, 2003) and even to moving abstract shapes (Castelli, Happe, Frith, & Frith, 2000; Blakemore et al., 2001; Schultz et al., 2003; Blakemore et al., 2003). Activity in the posterior superior temporal sulcus region has previously been found during imitation of human actions (Iacoboni et al., 2001), which is performed mostly by imitating an action's goals rather than the actual movements (Wohlschlaeger, Gattis, & Bekkering, 2003; Koski, Iacoboni, Dubeau, Woods, & Mazziotta, 2003). This brain region is also thought to extract intentional cues from goal-oriented human behaviour (Toni, Thoenissen, & Zilles, 2001). My results show that this region is engaged by the observation of objects moving in a self-propelled and interactive way, which are movement characteristics associated with animate beings. Therefore my results are compatible with a role of the superior temporal region in identifying animate agents in the environment based on their movement, a necessary step in the processing of social information.

### *Animacy and agency*

Other interesting studies that it might be relevant to compare the present results with are two recent studies of agency by Farrer and colleagues (Farrer & Frith, 2002; Farrer et al., 2003), in which a region in the right angular gyrus responded when subjects attributed actions they saw to a person different from themselves. These activations are located only a few millimeters away from the activation I labelled right pSTS. With respect to the idea that the angular gyrus is activated when subjects attribute actions they see to an agent different than themselves, one might interpret our current results in the sense that with increasing clarity of the observed interaction between the moving objects, the perception of another agent also increased, therefore increasing the activation in the area thought to be involved in attribution of agency.

*Other regions engaged by interacting objects: fusiform gyrus and medial occipital cortex*

Significant activation increases corresponding to increasing interactivity were also found in the fusiform gyrus and the medial occipital cortex. Previous studies showed activation increases in the fusiform gyrus during presentation of pictures of faces (Haxby et al., 1994; Puce, Allison, Gore, & McCarthy, 1995; Kanwisher, McDermott, & Chun, 1997) and of living beings (Chao, Martin, & Haxby, 1999; Chao, Haxby, & Martin, 1999), and during observation of moving, abstract "agents" (Castelli, Happe, Frith, & Frith, 2000; Schultz et al., 2003). Activation in the medial occipital cortex is also known to increase during observation and naming of pictures of animals but not tools (Perani et al., 1995; Martin, Wiggs, Ungerleider, & Haxby, 1996; Damasio, Grabowski, Tranel, Hichwa, & Damasio, 1996; Martin, Ungerleider, & Haxby, 2000). Lesions of the medial occipital lobe (particularly in the left hemisphere) are associated with a specific semantic knowledge deficit for animals

(Nielsen, 1958; Tranel, Damasio, & Damasio, 1997). As reviewed above, goal-directed motion and self-propelled movement appear to be the main cues for the attribution of animacy to abstract objects. It is therefore possible that activity in the fusiform gyrus and the medial occipital cortex increased with increasing interactivity because the moving objects appeared increasingly animate and animal-like.

## **Chapter 2. Experiment 2: Study of goal-attribution in a chasing paradigm**

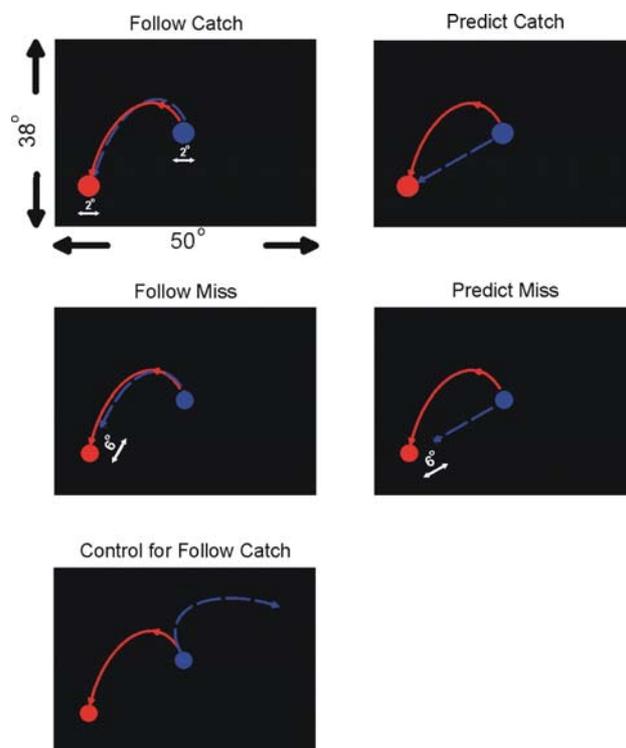
### Methods

#### *Design, conditions and tasks*

In this experiment, I wanted to test whether moving objects elicit stronger activation in the posterior superior temporal sulcus and gyrus when they appear to attribute goals to a target object than when they appear to chase their target without attributing goals to it. I used a chasing situation where one object tried to catch another object by using one of two strategies. The chasing object appeared to attribute a goal to the target object by heading towards the end-point of the target's position, as if it had predicted where the target would go ("Predict" strategy), or simply followed it (no goal attribution, "Follow" strategy). I used matched control conditions without chasing for each experimental condition to control for differences in object motion across conditions.

I also wanted to test whether activation changes happen only when subjects pay attention to the strategy used by the objects or also when they perform another, incidental task. To this end, I crossed the Strategy factor described above with an Outcome factor: the chasing object reached the target object in only 50% of cases, allowing the performance of an incidental task using the same animations. I therefore had two tasks, as follows: 1) in the Strategy task, subjects categorised the chaser's strategy into "following" or "using prediction of the trajectory endpoint"; 2) in the Outcome task (the incidental task), they had to decide whether the chaser caught the target at the end of the trial or not. During the control trials, subjects were asked to press one of the two response buttons, at random. I therefore used a factorial design with the following four factors: 1) two strategies for the chasing object: predicting or

following, 2) two outcome levels: target caught or missed, 3) two task levels: judging strategy or judging outcome, and 4) two condition levels: chase and control (no chase). For examples see Figure Exp 2.1. Combinations of all these factors resulted in 16 different trial types, each repeated 24 times during the experiment, for a total scanning time of 19.5 minutes. Subjects responded by pressing one of two buttons on a keypad with the corresponding finger of the right hand. Ratings, response times and brain activation were recorded simultaneously during the experiment. On these two types of dependent variable I assessed the effects of strategy and the interaction between task and strategy.



**Figure Exp 2.1. Design and stimuli.**

The open red arrow indicates time at which the blue object started to move. Four experimental conditions come from the combination of two factors in the experiment: Predict strategy vs. Follow strategy and Catch outcome vs. No catch outcome. Below on the left is the specific control condition for the Follow Catch condition; such specific control conditions existed for all four experimental conditions, resulting in 8 conditions in total. Subjects categorised either the strategy or the outcome of the blue object's movement in separate tasks on the same animations, therefore the total number of conditions x task combinations in the experiment is 16. The objects moved to any of the four corners of the screen, with equal frequency.

### *Stimuli*

Participants in the scanner watched short animations (4.2 sec. per animation) in which two moving disks appeared to be either interacting or to move independently from each other (Figure Exp 2.1). The moving objects were two disks or balls, with a width on the screen in the scanner of 2 degrees of visual angle (Figure Exp 2.1, top left). One was coloured bright red, the other bright blue, and the background was black. The movement trajectories of the two disks were determined prior to the experiment by an equation of motion implemented in Matlab (The MathWorks, Inc, Natick, MA). The visual display was controlled by in-house presentation software (Cogent 2000, <http://www.vislab.ucl.ac.uk/Cogent2000/index.html>) implemented in Matlab. Animations were projected onto an opaque surface in the scanner by an LCD projector, subjects viewed them through a mirror mounted on the headcoil.

In all the animations, the red object (the target) started to move first, describing an arc from the center of the screen that ended in one of the four corners of the screen (the example shown in Figure Exp 2.1 ends in the bottom left corner; during the experiment, an equal number of trials with endpoint in each corner were shown). The blue object (the chaser) started 0.93 seconds after the red object, and in half the trials it tried to catch it (Chasing trials), in the other half of the trials it moved in the opposite direction from the red object (Control trials). When trying to catch the red object, the blue object either simply followed the target's trajectory or it went directly to the endpoint of the target's movement, as if it had known the end-point of the red object's trajectory. In half of the Chasing trials the blue object caught the red object, in the other half it missed it by 6 degrees.

### *FMRI analysis*

The analysis followed the standard steps as described in the Materials and Methods section of the present thesis. Realignment was standard, no slice-time correction was used, and normalisation was done with resampling to a voxel size of 2 mm<sup>3</sup>. To enhance the signal to noise ratio and enable intersubject functional anatomical comparison the images were smoothed by convolution with a 6-mm full width at half maximum (FWHM) Gaussian kernel. A high-pass filter (using a cut-off of 128 sec) and a correction for temporal auto-correlation in the data (AR 1 + white noise) were applied to accommodate serial correlations. The design matrix of the first-level (single subject) analysis comprised 32 regressors of interest, resulting from the crossing of the following factors: i) following or predicting strategy of the chaser, ii) successful or unsuccessful outcome iii) task (strategy-categorisation or outcome-categorisation), or iv) animations with chasing or controls, which yields 16 effects of interest; each was modelled using the canonical hemodynamic response function (HRF) and the first temporal derivative. Differential realignment parameters were modelled as regressors of no interest to model movement-related artifacts. Second-level analyses were performed as T-tests, after smoothing the contrast images from the first-level analysis with an 8-mm FWHM kernel. For the analysis of task effects, I was particularly interested in identifying regions that showed greater activation in the explicit task than in the implicit task. To assess these effects, I used both a 3-way interaction between all factors in the experiment, and an exclusive masking procedure, which enables more sensitive comparisons between contrasts of interest by allowing variable thresholding of the contrast used as a mask. Both analyses are standard procedures implemented in SPM2. For the exclusive masking analysis, I first assessed separately the interaction between predictive vs. following strategy and chase vs.

control trials in the strategy task and in the outcome task. I then masked the results of this contrast in the Strategy task with the results of the same contrast in the Outcome, thresholding the mask at  $P=0.05$  uncorrected (the default in SPM2 for masking procedures). To identify regions more active in the implicit than the explicit task, I reversed the mask and contrast of interest, also thresholding the mask at  $P=0.05$  uncorrected. Similar results in both masking procedures were obtained with all thresholds below  $P=0.15$ , discussed and presented in Figure Exp 2.5 are analyses using the mask thresholded at  $P=0.05$ .

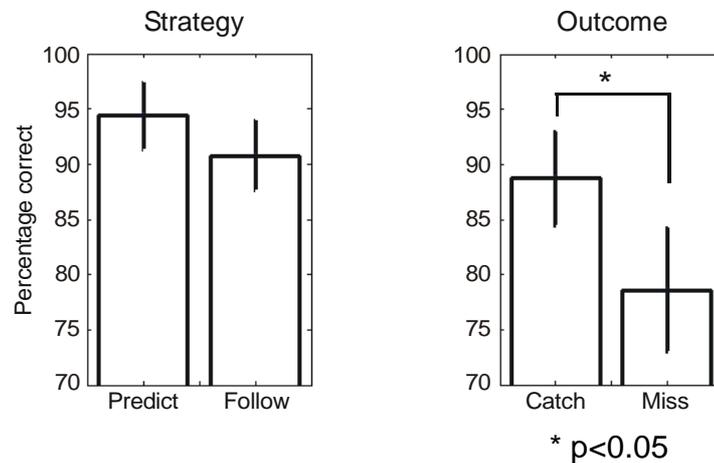
## Results

### *Behavioural results*

#### **Ratings**

Analysis of participants' ratings from the Strategy task and the Outcome task performed in the scanner showed that participants performed well in both the Strategy and the Outcome task. Subjects had to perform the tasks only when a chase between the two objects took place, and they correctly identified the strategy of the chasing object in 94 percent of Predict trials and 91 percent of Follow trials ( $\pm 3$  percent SEM over subjects). Positive outcome (i.e. a successful chase) was identified correctly in 89 percent of trials ( $\pm 4$ ) and negative outcome (i.e. an unsuccessful chase) was identified correctly in 79 percent of trials ( $\pm 5.6$ ). There were no significant differences between ratings of Predict and Follow trials (paired T-test,  $p>0.05$ ), but Positive Outcome trials were more often identified correctly than Negative Outcome trials (paired T-test,  $P<0.05$ ). This difference probably stems from subjects responding too quickly and mis-categorising some Negative Outcome trials as

Positive Outcome trials. This difference does not influence the interpretation of the brain activation analysis because trials with different outcomes were not compared with each other and only served as an implicit task.



**Figure Exp 2.2. Ratings in the Strategy and the Outcome task.**

Acquired during the fMRI experiment.  $N = 14$ , error bars represent S.E.M.

### **Response times**

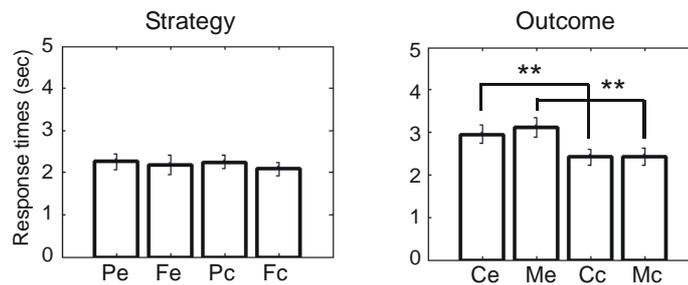
Response times (Figure Exp 2.3) were shorter in the Strategy task than in the Outcome task ( $t(1,13) = -4.684$ ,  $P < 0.001$ ; paired T-test). This was to be expected: subjects were allowed to respond as soon as they had made up their mind in the trial, which would take longer in the Outcome task. In this task, they had to wait until the end of the trial to see if the Chaser reached the Target, while in the Strategy task, the Chaser's strategy was apparent from early on.

In the Strategy task, response times were slightly longer in the Predict than in the Follow trials ( $F(1,13) = 8.3$ ,  $P < 0.05$ ; repeated-measures 2-way ANOVA, factors = Predict vs. Follow, and Experimental vs. Control), but there was no difference between Experimental and Control trials ( $p > 0.5$ ), and no interaction between Experimental vs. Control and Predict vs. Follow factors ( $p > 0.2$ ). Planned post-hoc

tests showed no difference between Predict experimental vs. Predict control or between Follow experimental vs. Follow control (all  $p > 0.05$ ). The difference between Predict and Follow trials could be explained by a difference in difficulty (subjects had a tendency to be better at correctly identifying Predict than Follow trials, see Figure Exp 2.2). As there is no significant interaction between factors, these results do not affect the interpretation of my fMRI data, as I focused my analysis precisely on the interaction contrast (see fMRI results below).

In the Outcome task, response times were similar in the Catch and Miss trials, ( $p > 0.05$ ; repeated-measures 2-way ANOVA, factors = Catch vs. Miss, and Experimental vs. Control), but there was a difference between Experimental and Control trials ( $F(1,13) = 21.1, P = 0.01$ ), and a significant interaction between the 2 factors ( $F(1,13) = 9.0, P < 0.05$ ). Planned post-hoc tests showed that response times in Catch Experimental were longer than in Catch Control ( $F(1,13) = 21.8, P < 0.001$ ); as was the case in the comparison Miss Experimental vs. Miss Control ( $F(1,13) = 30.0, P < 0.001$ ). That subjects took longer to respond in the Experimental than the Control trials for both Catch and Miss trials is most probably related to the fact mentioned above: in the Experimental trials, they had to wait until almost the end of the trial before deciding whether the Target was reached, but not so in the Control trials where objects were not chasing each other. The interaction can be explained by the same effect as seen in the task performance data: subjects probably found it more difficult to identify Catch than Miss trials (they miscategorised more Miss as Catch trials than the opposite), which can explain why they spent more time before making up their mind during Experimental Miss trials. Again, these effects did not interest us because they do not influence the interpretation of the brain activation analysis because trials

with different outcomes were not compared with each other and only served as an implicit task (I pooled over Catch and Miss trials, see below).



**Figure Exp 2.3. Response times of the 14 participants.**

Acquired during both tasks performed during the fMRI experiment. Error bars represent S.E.M.

### *FMRI results*

#### **1) Voxel-wise analysis in the superior temporal sulcus and gyrus**

I used the factorial design of this experiment to test for activation differences due to the strategy used by one object to chase another object and to test how these activation differences were affected by the attentional task the subject performed. To discount object motion effects I assessed interaction-sensitive activations as the difference between the activation increases due to a Predict vs. a Follow strategy in Chase trials and the same activation increases in the equivalent movement Control trials. Based on previous studies of biological motion and interacting geometrical shapes, I was particularly interested in the activation of the cortex in the posterior part of the superior temporal sulcus and gyrus of both hemispheres.

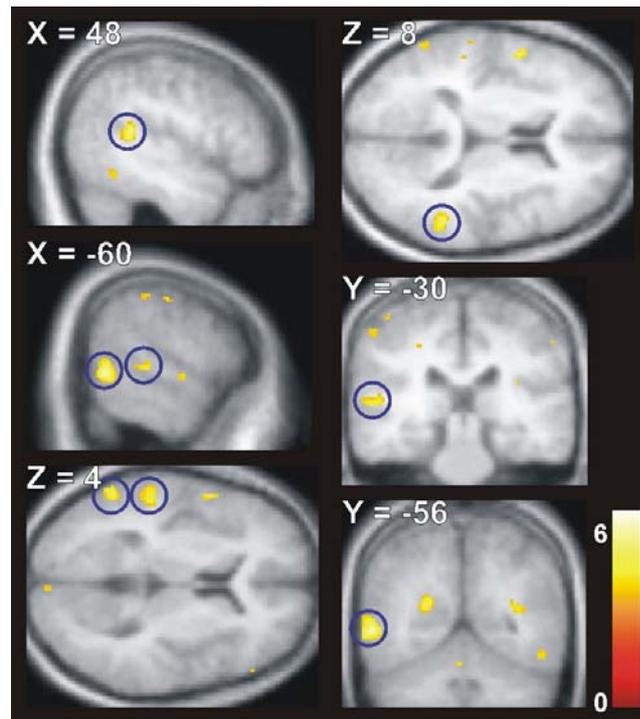
##### **- Effects of Strategy**

Activation increases were found in the superior temporal gyrus and the superior temporal sulcus of both hemispheres when subjects observed a chasing object with a Predict strategy compared to observation of a chasing object with a Follow strategy (Figure Exp 2.4). This was tested formally with a two-way interaction between

Predict vs. Follow strategies and Chase vs. Control trials to discount differences in object movement. The opposite contrast yielded no significant activation increases in the superior temporal sulcus or gyrus.

**Table Exp 2.1.** fMRI data, voxel-wise analysis. Significant clusters in the comparisons of interest, surviving a threshold of  $P < 0.05$  corrected for multiple comparisons across voxels in the superior temporal sulcus or gyrus (see Methods).

<i>Structure</i>	<i>Coordinates</i>			<i>Size (mm<sup>3</sup>)</i>	<i>Z score</i>
Predict > Follow strategy * Chase vs. Control					
R superior temporal gyrus	48	-44	12	600	4.48
L middle temporal gyrus	-60	-56	4	176	3.83
L superior temporal gyrus	-56	-30	4	584	3.58
Predict > Follow strategy * Chase vs. Control, Strategy task only					
L superior temporal gyrus	-54	-34	4	936	4.2
Predict > Follow strategy * Chase vs. Control, Strategy task masked by Outcome task					
L superior temporal gyrus	-54	-34	4	784	4.2



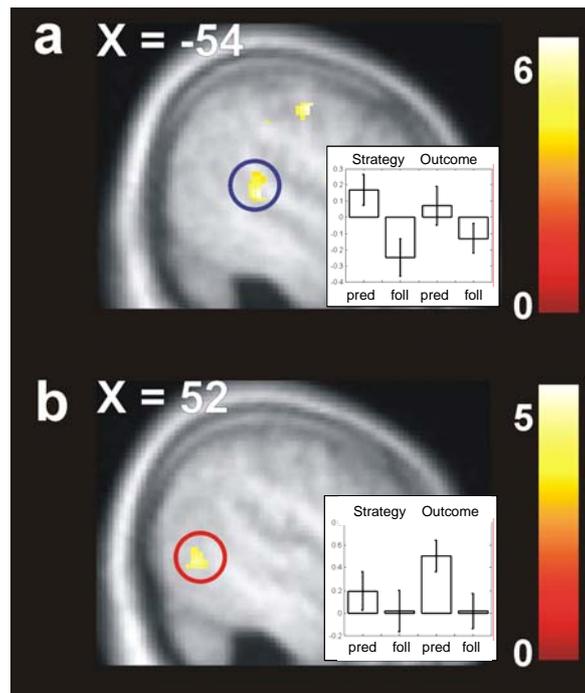
**Figure Exp 2.4. Brain areas responding more to Predict than Follow trials, irrespective of task.**

Clusters in the superior temporal sulcus and gyrus whose activation was significantly greater during observation of a Predict strategy than a Follow strategy. Tested formally with a two-way interaction between Predict vs. Follow strategies and Chase vs. Control trials. For details see Table Exp 3.1. For display, the image is thresholded at  $P < 0.001$  uncorrected for multiple comparisons. Clusters surviving a threshold of  $P < 0.05$  corrected for multiple comparisons across anatomical search regions are circled in blue, listed in Table Exp 3.1 and discussed in the text. The brain slices used for display is the average image from the subjects' normalised structural brain scans (see Methods). Coordinates are in MNI (Montreal Neurological Institute) reference space. The color bar refers to activation intensity expressed in  $t$  values.

#### - Effects of task on strategy

When tested as interactions between Strategy and Task, no cluster showed significantly different activation depending on the task the subjects performed. This was tested by the following three-way interactions: interaction between Task (Strategy vs. Outcome), Strategy (Predict vs. Follow) and Chase vs. Control trial, and interaction between Task (Outcome vs. Strategy), Strategy (Predict vs. Follow) and Chase vs. Control trial. However, a cluster of voxels in the left posterior superior

temporal sulcus was found whose activation was greater during observation of a Predict vs. a Follow strategy when subjects performed the Strategy categorisation task, but was not significantly activated when subjects performed the Outcome task. When I masked (exclusive masking) the results of the comparison in the Strategy task with those of the Outcome task, the cluster in the left posterior superior temporal sulcus identified in the Strategy task remained significantly activated (Figure Exp 2.5 a). This suggests that voxels in this cluster respond only when subjects performed the Strategy task. Masking the results of the comparison in the Outcome task with those of the Strategy task yielded a cluster of voxels in the right middle temporal gyrus, located caudal and ventral to the cluster of activation in the posterior superior temporal sulcus mentioned above. This cluster was not located within the superior temporal sulcus and gyrus search regions (Figure Exp 2.5 b).



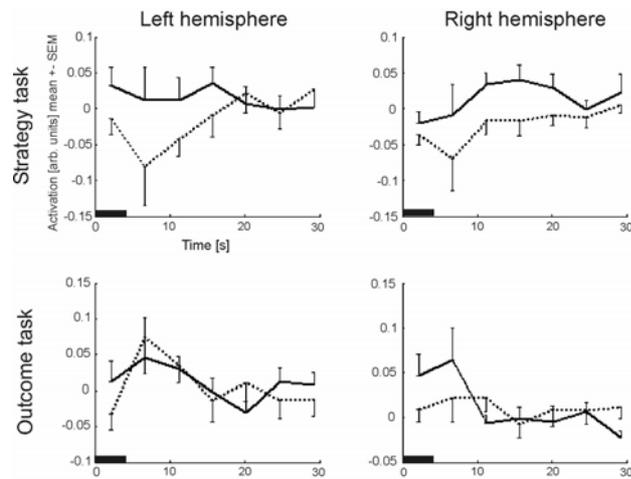
**Figure Exp 2.5. Brain regions: task effects.**

Clusters in the superior temporal sulcus and gyrus whose activation was significantly greater during observation of a Predict strategy than a Follow strategy and was affected by the subjects' task. A: activations in the Strategy task masked with activation in the Outcome task (exclusive mask, threshold of mask:  $P=0.05$  uncorrected). B: activations in the Outcome task masked with activation in the Strategy task (exclusive mask, threshold of mask:  $P=0.05$  uncorrected). The cluster circled in blue survived a threshold of  $P<0.05$  corrected for multiple comparisons across anatomical search regions, was listed in Table Exp 3.1 and discussed in the text. The cluster circled in red was activated to a similar degree but was not located within the search regions. The brain slices, coordinate system and color bar index are similar to Figure Exp 2.4.

## 2) Activation time-courses

Activation time-courses in the voxels of the superior temporal sulcus and gyrus with strongest activation increase during observation of a Predict vs. a Follow strategy confirmed the results of the voxel-wise analysis (Figure Exp 2.6). In the superior temporal sulcus in the left hemisphere, activation during the initial phase of the trial was higher in Predict trials than in Follow trials, but only when subjects performed the Strategy task (repeated-measures 2-way ANOVA: interaction between

Strategy and Time:  $F(1,13)=7.8$ ,  $P=0.015$ ). In the right hemisphere, activation during the whole trial was higher in Predict trials than in Follow trials when subjects performed the Strategy task (main effect of Strategy:  $F(1,13)=4.9$ ,  $P=0.045$ ), and was only initially higher when subjects performed the Outcome task (interaction between Strategy and Time:  $F(1,13)=5.6$ ,  $P=0.035$ ).



**Figure Exp 2.6. Activation time-courses in the superior temporal gyrus in both Strategy and Outcome tasks.**

Solid lines refer to trials in which the chasing object used a Predict strategy, dotted lines refers to trials in which the chasing object used a Follow strategy. The thick black line on the horizontal axis represents the duration of the animation. Activation is displayed in arbitrary units, 0 refers to average activation during the whole course of the experiment. Error bars represent standard errors of the mean across subjects. In the left hemisphere, activation during the initial phase of the trial was higher in Predict trials than in Follow trials, but only when subjects performed the Strategy task (repeated-measures 2-way ANOVA: interaction between Strategy and Time:  $F(1,13)=7.8$ ,  $P=0.015$ ). In the right hemisphere, activation during the whole trial was higher in Predict trials than in Follow trials when subjects performed the Strategy task (main effect of Strategy:  $F(1,13)=4.9$ ,  $P=0.045$ ), and was only initially higher when subjects performed the Outcome task (interaction between Strategy and Time:  $F(1,13)=5.6$ ,  $P=0.035$ ). Data are from the peaks of activation (coordinates: left hemisphere: -54 -34 4, right hemisphere: 48 -44 12, Montreal Neurological Institute reference frame) of the significant clusters in the superior temporal gyrus, identified with the Predict vs. Follow \* Chase vs. Control contrast.

## Discussion

This study shows that activation in the superior temporal gyrus and in the cortex surrounding the superior temporal sulcus increases when a simple object appears to chase another object by understanding the target's goal and predicting its movement instead of simply following it. Directing attention to the object's strategy rather than to the outcome of its chase further increased activation in the left superior temporal gyrus.

### *Attribution of intentions and goal-directed movements and the superior temporal sulcus and gyrus*

A number of studies have shown activation increases in the posterior part of the superior temporal sulcus and gyrus (among other regions) during attribution of mental states to other agents, also called mentalising or theory-of-mind (Frith & Frith, 2003). The exact mechanism on which this ability is based is not fully determined yet, but intuitively, a necessary first step towards the attribution of mental states to an agent could be the attribution of goals to this agent. When moving objects appear to follow goals, they are more likely to be considered animate by human observers (Opfer, 2002; Csibra, 2003). Previous studies show that activation in the superior temporal sulcus and gyrus increases during observation of geometrical shapes if their movements appear goal-directed or to have intentions (Castelli, Happe, Frith, & Frith, 2000). These data suggest that this part of the cortex is sensitive to movement characteristics associated with living entities, especially goal-directed movements. In my present study, participants observed an object understanding the goals of another object and predicting its movements or simply following it in order to catch it. As predicting the movements of a target object is a characteristic associated with a potential agent, the greater activation in the superior temporal sulcus and gyrus when

the chasing object predicted the movements of the target object compared to simply following it suggests that this brain region plays a role in the identification of intentional agents (i.e. agents that appear to have intentions).

#### *Attention effects*

The cluster identified in the left superior temporal gyrus showed greater activation increases in Predict trials compared to Follow trials when subjects tried to identify the strategy of the chaser than when they assessed the outcome of the chase (Figure Exp 2.5a). As discussed in the introduction, activation increases in the superior temporal gyrus area have been reported in the right hemisphere when attending to the emotion or the trustworthiness of faces, and in the left hemisphere when attending to contingency between the movements of abstract objects (Blakemore et al., 2003). Blakemore and colleagues suggested that the attention effects they observed could be due to top-down effects related to the search for agents, rather than bottom-up detection on the basis of visual cues. In my case, attending to the strategy of the moving agents, while not being exactly equivalent to searching for agents, cued subjects into paying attention to characteristics of living organisms, rather than simply comparing the positions of the chaser to the position of the target at the end of the trial in the Outcome task. I suggest that the active process of searching either for agents or for characteristics of their movements can increase or prolong neural responses in areas already sensitive to entities displaying potentially goal-directed behaviour.

*Time-course: early activations vs. constant*

In the left hemisphere, activation during the initial phase of the trial was higher in Predict trials than in Follow trials when subjects performed the Strategy task; in the right hemisphere, a similar pattern was observed during the Outcome task (Figure Exp 2.6). Activation in the right hemisphere was higher in Predict trials than in Follow trials during the whole trial when subjects performed the Strategy task. In the two cases where the activation was only higher in Predict than Follow trials at the beginning of the trial, plots of the data show that the activation difference was present at 6.75 seconds after trial onset (Figure Exp 2.6, top left and bottom right), which is compatible with a neural event occurring at the onset of the trial. At the next sampled time point (11.25 seconds), the activation difference in the left hemisphere was less important and it disappeared in the right hemisphere. I suggest that this reflects an initial and automatic response of the cortex to our animations, which is increased and/or prolonged when subjects pay attention to a socially relevant dimension of the objects' movement. The time-course analyses suggest that this response is stronger in the right hemisphere, but as the SPM analyses do not confirm this tendency, I cannot argue strongly in favour of hemispheric specialisation in this task.

*No activation increases in the medial prefrontal cortex*

It is interesting to note that I did not find activation increases in the medial prefrontal cortex or the temporal poles in all the contrasts I tested, areas which, together with the superior temporal sulcus, are thought to be involved in the attribution of mental states (Frith & Frith, 2003). Previous PET and fMRI studies using complex animations leading to attribution of mental states have shown activation in these areas (Castelli, Happe, Frith, & Frith, 2000; Schultz et al., 2003), but studies using simple animations have not (Blakemore et al., 2001; Blakemore et

al., 2003). The latter studies used animations with objects interacting causally, appearing animate, and/or interacting contingently, which are not thought to lead to attribution of mental states. In my present study, observers watching an object attributing a goal to another object did not show a significantly greater activation in the medial prefrontal cortex than when watching an object following another object, regardless of the task the subjects were performing. I suggest that subjects solved this simple task with only a small mentalising effort or none at all, which was not sufficient to induce activation in the medial prefrontal cortex. While activation increase in the superior temporal sulcus could be related to mentalising, I would rather suggest that it is due to detection of entities displaying goal-directed behaviour or appearing to have intentions, which could explain why activation in the medial prefrontal cortex was not increased. This view is supported by informal reports from my subjects, and is consistent with results and interpretation of a previous study (Blakemore et al., 2003).

#### *Differences in movement paths*

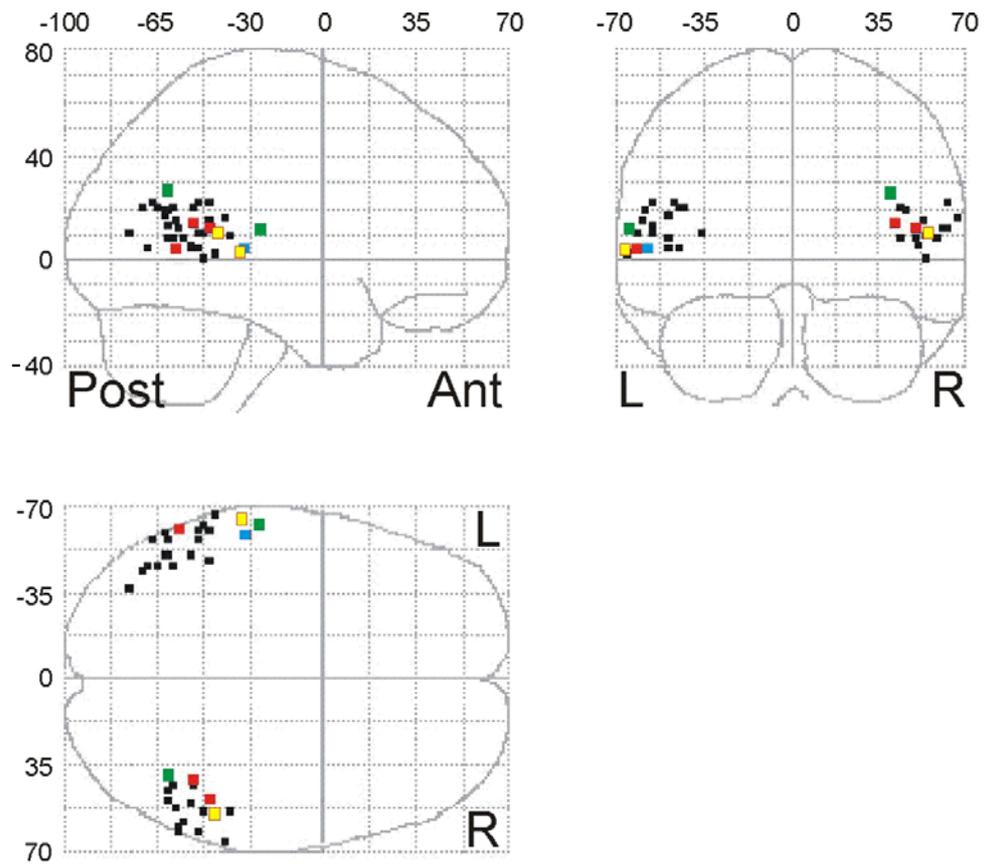
A possible alternative explanation of activation differences caused by the observation of objects with different movement paths could be the difference in complexity between the paths. In the present experiment, this was controlled by matched control conditions in which the paths of both chasing and target objects were identical to the paths in the chasing trials except for the fact that the chasing object did not move towards the target object but in the opposite direction. This control condition only affected the contingency between the objects' paths and not the movement paths themselves. Objects moving in opposite directions could be seen to be avoiding instead of approaching each other, and therefore represent another form of social interaction, which could also induce activation in the superior temporal

sulcus. Although this is a very valid possibility, I observed activation increases in the superior temporal sulcus and gyrus not only in the interaction between Chase vs. Control and Predict vs. Follow strategy, but also in the comparison of all Chase with all Control trials (data not shown). Also, within Chase trials, the activation was stronger when the chasing object appeared to attribute a goal to the target object than when it just followed it. All these data support my interpretation suggesting that goal attribution by the Chaser induces activation increases in the superior temporal sulcus and gyrus, and that attention to the object's strategy further increases the activation.

*Specialisation along the superior temporal sulcus?*

The cortex in the superior temporal sulcus and gyrus responds to i) the attribution of mental states to others, ii) the identification of biological motion and iii) the response to animate-looking, moving geometrical shapes. Whether different parts of the cortex in the sulcus and gyrus are specialised for each of these functions or not is not yet known, although recent unpublished results from a study by Saxe, Xiao, Kovacs, Perrett and Kanwisher suggest that different parts of the pSTS area respond to biological motion or to false belief stories (discussed in Saxe & Kanwisher, 2003). Against the idea of specialisation is the fact that activation increases during observation of biological motion have been found all along the superior temporal sulcus (Puce & Perrett, 2003) and do not appear organised in a particular way. Instead, a number of studies of all three types of processes listed above have yielded activation in the same area of the superior temporal gyrus and sulcus (Figure Exp 2.7).

The activation increases found in the present study are also located in the pSTS area. This suggests that this region of the cortex could be involved in a common aspect of these three tasks, for example the response to goal-directed movements, whether expressed by moving abstract shapes or human beings.



**Figure Exp 2.7. Comparison between STS/STG activation in all 3 experiments.**

Experiments 3 in yellow with red contour, experiment 1 in green and experiment 2 in blue and red, respectively for task-dependent and non-task-dependent regions, and previous studies (in black) in following topics: mentalising (Fletcher et al., 1995; Goel, Grafman, & Hallett, 1995; Gallagher et al., 2000; Brunet, Sarfati, Hardy-Bayle, & Decety, 2000; Vogeley et al., 2001), observation of animate agents (Castelli, Happe, Frith, & Frith, 2000; Schultz et al., 2003), biological motion (Bonda, Petrides, Ostry, & Evans, 1996; Grossman et al., 2000; Puce, Allison, Bentin, Gore, & McCarthy, 1998; Grezes & Costes, 1998; Grezes et al., 2001; Wicker, Michel, Henaff, & Decety, 1998; Hoffman & Haxby, 2000; Campbell et al., 2001) and detection of change (Downar, Crawley, Mikulis, & Davis, 2000). Displayed on Maximum Intensity Projection glass brain with SPM2. Full coordinates used: Experiment 1: [-60 -27 9; 39 -57 27], experiment 2: [-60 -56 4; -54 -34 4; 48 -44 12], Fletcher 1995: [-44 -58 20; 42 -50 20], Goel 1995: [-42 -70 20; -44 -64 20], Gallagher 2000: [-54 -66 22; 60 -48 22], Castelli 2000: [-58 -48 4; 60 -56 12], Brunet 2000: [-64 -42 2; 52 -46 0], Vogeley 2001: [-46 -44 22; 58 -56 12], Schultz 2003: [-57 -61 19; 51 -57 15], Grezes 2001: [-35 -75 10], Bonda 1996: [-48 -61 17; 56 -54 8], Grossman 2000: [-54 -60 13; 44 -60 19], Puce 1998: [-48 -51 4; 49 -51 5], Campbell 2001: [-58 -44 15; 52 -36 9], Grezes 1998: [-60 -46 10; 64 -38 16], Wicker 1998: [-44 -68 4; 42 -58 8], Hoffman 2000: [-48 -60 8; 48 -60 8], Downar 2000: [-54 -48 10].

### **Chapter 3. Experiment 3: action observation**

#### **3.1. Introduction and overview**

When someone mimes or pretends to perform an action, her or his movements are not directed towards a present object, and are thought to be based on stored representations of the target object rather than perceptual input (Goodale, Jakobson, & Keillor, 1994). The following data suggest that mimed movements are controlled by different processes from object-directed actions. Mimed movements can have different kinematics from object-directed or actual movements (Goodale, Jakobson, & Keillor, 1994) and are much less affected by visual illusions (Westwood, Chapman, & Roy, 2000) than are actual movements (Aglioti, DeSouza, & Goodale, 1995; Haffenden & Goodale, 1998; Ellis, Flanagan, & Lederman, 1999; Flanagan & Beltzner, 2000). These data, and the report of a patient with ventral visual stream lesions (James, Culham, Humphrey, Milner, & Goodale, 2003) who is unable to perform mimed movements based on perceptual cues (Goodale, Jakobson, & Keillor, 1994), suggest that neural structures underlying control of mimed actions are located in the ventral rather than the dorsal visual stream (Milner & Goodale, 1995; Westwood, Chapman, & Roy, 2000). Thus, the fact that actual and mimed movements differ based on their object-directedness might determine their control by different visual streams.

Observing and imitating object-directed actions engages ventral and dorsal parietal, ventral and dorsal premotor, and superior temporal areas (Iacoboni et al., 2001; Buccino et al., 2001; Koski, Iacoboni, Dubeau, Woods, & Mazziotta, 2003; Grezes & Decety, 2001; Grezes, Armony, Rowe, & Passingham, 2003). A growing wealth of data suggests that the superior temporal sulcus is involved in the processing

of information relevant to social communication and interaction, including goal-directed behaviour (displayed as humans reaching and grasping for objects or as interacting abstract objects), biological motion, body parts, faces, eye gaze (for reviews, see: Puce & Perrett, 2003; Haxby, Hoffman, & Gobbini, 2002; Adolphs, 2003). While all these types of information involve some kind of motion, this region of the brain is also involved in more cognitive tasks, such as mentalising (Frith & Frith, 2003). The inferior premotor area responds also during observation of object-directed movements made by humans, and also during their execution (Iacoboni et al., 1999; Koski, Iacoboni, Dubeau, Woods, & Mazziotta, 2003). In monkeys, neurons that respond during observation of reaching and grasping actions have been found in the superior temporal sulcus (Jellema, Baker, Wicker, & Perrett, 2000), and neurons responding both during observation and execution of object-directed action (Mirror Neurons) have been described in the ventral premotor (di Pellegrino, Fadiga, Fogassi, Gallese, & Rizzolatti, 1992; Gallese, Fadiga, Fogassi, & Rizzolatti, 1996; Umiltà et al., 2001) and parietal cortices (Fogassi, Gallese, Fadiga, & Rizzolatti, 1998; Gallese, Fogassi, Fadiga, & Rizzolatti, 2002). Recent data suggest that the Mirror Neurons respond differently to object-directed and non-object-directed actions (i.e. a mimed movement) on the basis of the movement kinematics rather than the object presence (Umiltà et al., 2001, see Figure Intro.9).

To test the relative importance of the object and the movement kinematics for the identification of object-directed movement, I presented healthy volunteers lying in a functional Magnetic Resonance Imaging (fMRI) scanner with four different object manipulations. I presented Mimed and Actual versions of these manipulations, both presented either with or without an object being manipulated. Subjects had to discriminate between Mimed and Actual kinematics of different hand and arm

movements, irrespective of the presence of an object. I varied the difficulty of the task by using manipulations with more or less important kinematic differences between their Mimed and Actual versions.

To independently manipulate the types of kinematics and the object presence in order to assess their respective effects, I created high-quality animations based on the recorded movements of an actor, in which I could present either type of kinematics (Mimed or Actual) either with or without object. This design offered in addition the opportunity to test responses during artificial or realistic combinations of object and movement kinematics.

### **3.2. The fMRI experiment**

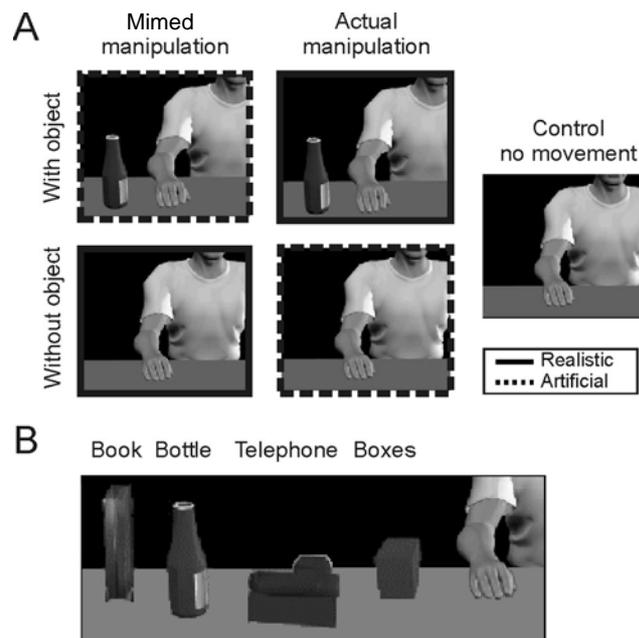
#### Methods

##### *Subjects*

Fourteen right-handed human volunteers (eleven Males and three Females, aged 21 to 40, average age 29.2 years) participated in the study. Two participants had problems understanding the instructions and were excluded from the study. All participants gave full written informed consent and the study was approved by the local ethics committee.

I independently manipulated the two main factors that could provide information during the observation of object-directed movements: the kinematics of the movement and the object presence. I presented two types of movements – Actual manipulation of an object and Mimed manipulation of the same object – which differed exclusively in movement kinematics. I also manipulated the visual presence of the target object in the animation – object present or absent – which was manipulated completely independently of movement kinematics. I combined these two factors to create a (2x2)

design resulting in four experimental conditions (Figure Exp 3.1A). For example, one of the four possible conditions showed an Actual manipulation of an object (based on the recorded movement kinematics of an actor really manipulating an object, see details below) but presented without object (object was visually absent from the animation scene) (Figure Exp 3.1A bottom right). Obviously, only two of the four animations represented realistic situations, namely the kinematics of an actual manipulation movement presented with an object, and the kinematics of a mimed manipulation movement presented without object. The two other conditions represented artificial situations that cannot be encountered in reality. The control condition was a static image of the same actor without an object (Figure Exp 3.1A far right). I used in-house stimulus presentation software (Cogent 2000, <http://www.vislab.ucl.ac.uk/Cogent2000/index.html>) implemented in Matlab (The MathWorks Inc, Natick, MA) to present the animations and record button presses.



**Figure Exp 3.1. Experimental design and stimuli.**

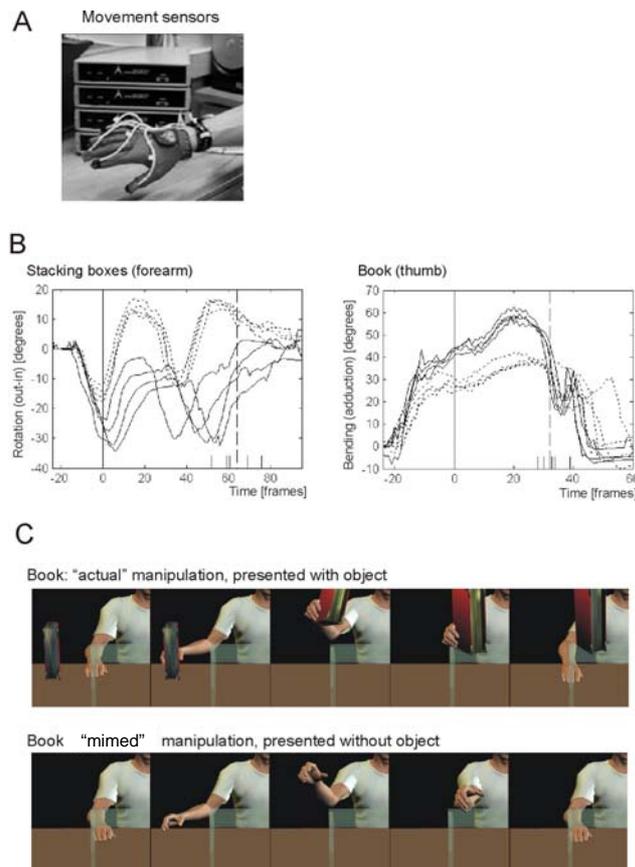
**A** The experimental conditions were determined by a 2x2 factorial design, with factors: movement kinematics (X axis of the Figure, Mimed or Actual) and object presence (Y axis), resulting in four experimental conditions. Two of these conditions represented realistic situations (Figure Exp 3.2 C), the other two represented artificial situations. The Control condition was a static actor without object, not changing during the whole trial. **B** Objects manipulated by the actor were a heavy book, a ketchup bottle, a stack of two blocks and a telephone, shown here together with the virtual actor's hand for size comparison. The animations were presented in colour in the scanner.

### *Recording of the movement kinematics*

Before the experiment, and in the absence of the subjects, I digitised (see Motion capture system, below) the movement kinematics of an actor performing (1) a manipulation of an object that was physically present and visible to the actor (an actual manipulation), and (2) a movement as similar as possible to (1) but executed in the absence of any target object (a mimed manipulation). The recorded movement kinematics were then imported into an animation program (Poser 4, Curious Labs Inc, Santa Cruz, California, [www.curiouslabs.com](http://www.curiouslabs.com)) to animate a virtual actor.

### *Motion capture system*

The movements of the actor's arm and hand were recorded using a magnetic motion tracking system sampled at 100 Hz (miniBird sensors, Ascension Technology Corp., Burlington, Vermont, [www.ascension-tech.com](http://www.ascension-tech.com)). A total of eight sensors were used, attached to the terminal segments of each digit, the dorsal hand surface, the forearm and the upper arm (see Figure Exp 3.2 A). The miniBird system reports 6 degrees of freedom per sensor (3 translations and 3 rotations) by detecting currents induced in the sensors by a magnetic field generated by a stationary transmitter. The joint angles of the digits were calculated from the positions and rotations of the terminal segments of each digit relative to the hand. The joint angles of the arm were calculated directly from the rotations of the associated arm sensors. The system was calibrated before each recording session by the actor adopting a series of standard postures. Movement recordings consisting of sequential frames of rotation angles for all joints of the hand and arm were exported offline into Poser to animate a virtual actor.



**Figure Exp 3.2. Movement kinematics and animation.**

**A** Movement capture sensors were placed at the fingertips and dorsal aspect of the hand, and on the lower and upper arm. **B** Examples of reconstructed joint angle time-series from two different joints and for two different manipulation types, with recordings in Mimed and Actual conditions (respectively dotted and continuous lines, 4 traces each, to create 8 different animations per manipulation type). These time-courses were reconstructed from the positions and rotations of the movement captors, using an inverse model algorithm using the anatomical characteristics of the virtual actor. 31 joint measures were acquired this way. Vertical line at time 0 and short lines respectively represent initial contact and final release of the object(s) by the recorded actor. **C** Corresponding screen-shots from animations with Actual kinematics presented with object (top) and Mimed kinematics presented without object (bottom).

*Differences between Mimed and Actual movement kinematics*

Actual and mimed movements of the actor differed from each other by the movements of some of the joints of the arm and hand. I assessed the differences in movement kinematics between Mimed and Actual versions for each of the manipulation types, both during the period in which the actor's hand was in contact

with the manipulated object and the whole duration of the manipulation, for the 30 joint variables for which I had movement data. For each of the four object manipulations, joint movements in the four "mimed" and the four "actual" trials (area under the curve of the movement of each joint in "actual" and "mimed" object manipulations, see example in Figure Exp 3.2 B) were ranked and compared in multivariate ANOVAs<sup>5</sup> in SPSS (one ANOVA per object, fixed factors: actual vs. mimed manipulation, four recordings each. Dependent variables were 30 joint movement measures: X, Y and Z rotations where applicable to the following joints: shoulder, forearm, phalanges 1 and 2 of the thumb, and all 3 phalanges of index, middle, ring and little fingers).

The number of joints with significantly different movement profiles ( $P < 0.05$ ) varied depending on the time window within which the movement kinematics were analysed, and depending on the manipulation type (Table Exp 3.1). The number of joint variables (out of 30) and percentage that showed significant differences between Mimed and Actual movements both during the period in which the object was in contact with the actor's hand and also for the whole length of the movement was 15 for Book manipulations (50%), 0 for Bottle, 3 for Phone (10%) and 10 for Boxes (33%). In addition, 5 joint variables (17%) showed differences only when the object was in contact with the actor's hand in Book manipulations, while this number was 10 for Bottle manipulations (33%), 0 for Phone and 5 (17%) for Boxes. Only Phone manipulations showed significant differences in one joint variable when analysed during the whole movement but not when analysed during the contact phase between object and hand.

---

<sup>5</sup> To account for possible non-normality in the joint movement data, a non-parametric based approach was used, consisting of rank-transforming the data before applying parametric tests (Conover & Iman, 1981).



Each animation lasted about 4 sec and was presented at 30 frames/sec. The animations were projected onto a screen in the scanner, situated about 10 cm away from the subjects' face. 12 trials of each condition per manipulation type were presented, for a total of 192 trials. In addition there were 64 control trials (static person without object, as described above). All subjects saw all animations, in different randomised orders. One experimental session lasted 26 min.

### *Task*

Prior to the experiment, subjects were told about the task and how the stimuli were made, then shown examples of all conditions of the experiment. In a two-alternative forced-choice decision task, the participants were asked to categorise each trial into actual or mimed manipulation, depending on the movement kinematics displayed in the animation, without paying attention to the presence of the object. Participants pressed with their right hand one button of a response box if they thought the movement was an actual manipulation, and another button if they considered it to be a mimed manipulation. They were instructed to respond as soon as they had reached their decision. No responses were given during the control trials.

### *Image acquisition*

A Siemens VISION System (Siemens, Erlangen, Germany), operating at 2 Tesla, acquired both T1-weighted anatomical images and gradient-echo echoplanar T2\*-weighted MRI images with blood oxygenation level dependent (BOLD) contrast. The scanning sequence was a trajectory-based reconstruction sequence with repetition time of 2736 ms and echo time of 35 ms. Each volume, positioned to cover the whole brain, comprised 36 sagittal slices, with an isotropic in-plane resolution of 3 mm, a slice thickness of 3 mm and a 1 mm interval between slices. For each subject, 648

volumes were acquired in one session of 29 minutes, including 5 subsequently discarded “dummy” volumes at the start of the session to allow for T1 equilibration effects. A structural MR image was acquired for each subject (modified MP RAGE sequence; (Deichmann, Good, Josephs, Ashburner, & Turner, 2000); parameters were: TR = 11ms, TE = 4 ms, Flip Angle 12 degrees, image matrix 256 pixels (Read) x 224 pixels (Phase), voxel size 1 x 1 x 1 mm, 176 slices per volume).

### *fMRI data analysis*

The (2x2) design with factors movement type (based on kinematics) and visual object presence was used for the fMRI data analysis to assess the main effects of object presence and of type of kinematics, and the interaction between the two. These effects were calculated for each type of manipulation (i.e. Book, Bottle, Phone and Boxes).

The analysis followed the standard steps as described in the Materials and Methods section of the present thesis. Realignment was standard, no slice-time correction was used, normalisation was done with resampling to a voxel size of 3 mm<sup>3</sup>. To enhance the signal to noise ratio and enable intersubject functional anatomical comparison the images were smoothed by convolution with a 6-mm full width at half maximum (FWHM) Gaussian kernel. A high-pass filter (using a cut-off of 128 sec) and a correction for temporal auto-correlation in the data (AR 1 + white noise) were applied to accommodate serial correlations. My model included two regressors for each of the four trial types: one regressor for correctly categorised and one regressor for incorrectly categorised trials. Thus the design matrix contained twelve regressors of interest: actual or mimed manipulation, each presented with or without object, for each manipulation type. Each of these conditions was modelled as a variable-length epoch from stimulus presentation onset to the subject's button press

by a series of delta (“stick”) functions. Each was modelled using the canonical hemodynamic response function (HRF). The control trials and intertrial intervals were modelled implicitly in the session effect regressor. Differential realignment parameters were modelled as regressors of no interest to model movement-related artifacts. Second-level analyses were performed in a 3-way ANOVA (Henson and Penny, in preparation), using correction for unequal between- and within-subject variance components (non-sphericity correction), after smoothing the contrast images from the first-level analysis with an 8-mm FWHM kernel. The threshold used for statistical significance was  $P < 0.05$  corrected for multiple correction across the whole brain, with inferences at the voxel or cluster level (based on a  $P < 0.001$  uncorrected map).

#### *Image used for display*

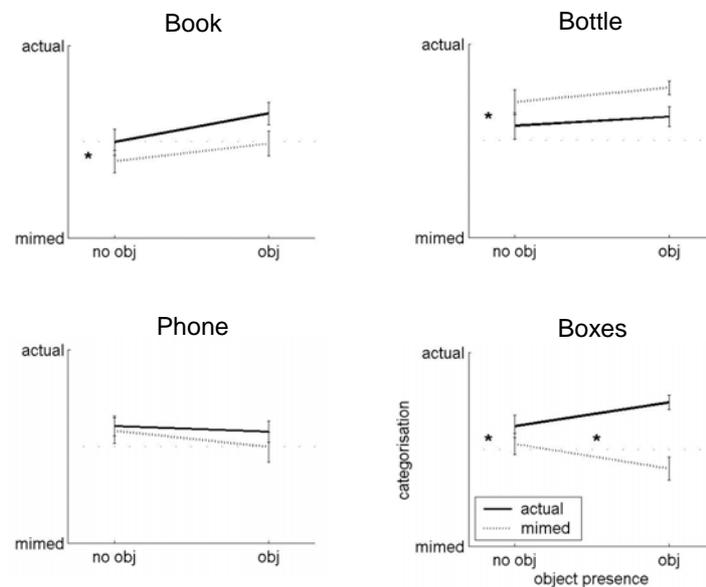
The mean image used for display in the Figures was calculated by averaging the twelve subjects’ structural images that were previously coregistered with the mean functional image of the same subject and normalised to the standard Montreal Neurological Institute (MNI) space. Anatomical structures were identified with brain atlases by Duvernoy (Duvernoy, 1999).

### Results

#### *Behaviour*

Using the (2x2) experimental design, I assessed participants’ discrimination between Actual and Mimed manipulation movements in the animations. I also assessed the effects of Object presence and the interaction between Object presence and type of Movement dynamics, separately for each of the 4 types of object

manipulation presented (Book, Bottle, Phone and Boxes; see Figure Exp 3.3). Subjects could correctly categorise Mimed and Actual versions of Book and Boxes manipulations (repeated-measures 2-way ANOVA (Mimed vs. Actual, Object vs. No object),  $F(1,11) = 7$ ,  $P < 0.05$  for Book and  $F(1,11) = 11.6$ ,  $P < 0.01$  for Boxes) For the Bottle manipulations, they could discriminate between Mimed and Actual versions ( $F(1,11) = 6.2$ ,  $P < 0.05$ ), but systematically mistook Mimed for Actual and Actual for Mimed. They could not discriminate between Mimed and Actual versions of the Phone manipulations at all ( $F(1,11) = 0.6$ ,  $p > 0.5$ ). The presence of an object increased categorisation performance only for the Boxes manipulation ( $F(1,11) = 9.2$ ,  $P < 0.05$ ).

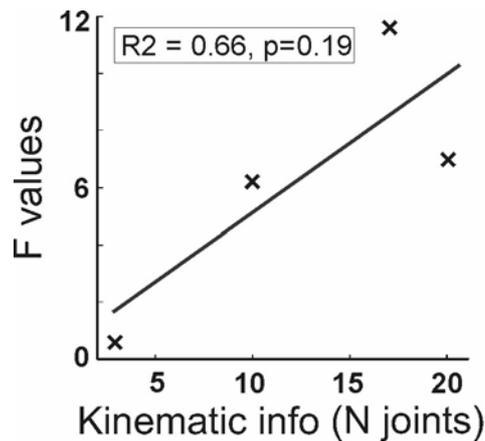


**Figure Exp 3.3. Behavioural results.**

Continuous and broken lines represent data from conditions with Actual and Mimed kinematics respectively. On X axis are trials with and without objects, on Y axis are subject's average categorisation ratings (+ SEM). Stars to the left of a plot indicate significant main effects of movement kinematics (2-way repeated measures ANOVA  $F(1,11)$ ,  $P < 0.05$ ), which indicate that subjects could discriminate the objects. The star at the middle of the bottom right plot indicates a significant interaction between Object presence and Kinematics: effectively, subjects' categorisation was better when the animation contained an object.

Comparing these results to the statistics on objective differences in the movement kinematics allows interesting interpretations. The object manipulation

types that were correctly categorised (Book and Boxes) were those that had the most joint variables with significant differences in movement kinematics (respectively 66 and 57% of joint variables showed significant differences when analysed during manipulation). Bottle manipulations, in which subjects could discriminate between Mimed and Actual versions but categorised them incorrectly, had only 33 % of joint measures that showed significant differences in movement kinematics, and in contrast with all other manipulation types, these differences were only found in the object-hand contact period. Phone manipulations, which subjects could not discriminate at all, had the smallest number of joint variables with significant differences between Mimed and Actual movement versions (13%). One could therefore conclude that 1) over 33% of joint variables with significant differences between Mimed and Actual allowed discrimination, but correct categorisation was possible only when over 50% of joint variables showed differences, or when these were found both during the whole movement and the contact phase between object and hand, not only the latter. Behavioural results have a (non-significant) tendency to correlate with the number of joint variables with significant differences between movement versions, when analysed for the contact period between object and hand (see Figure Exp 3.4).



**Figure Exp 3.4. Correlation between kinematic information and behaviour.**

Correlation between kinematic information in all four objects (measured as number of joint variables that showed significant difference between Mimed and Actual kinematics during object manipulation) and behaviour (measured by F values from categorisation task, normalised data): not significant but a trend is visible.

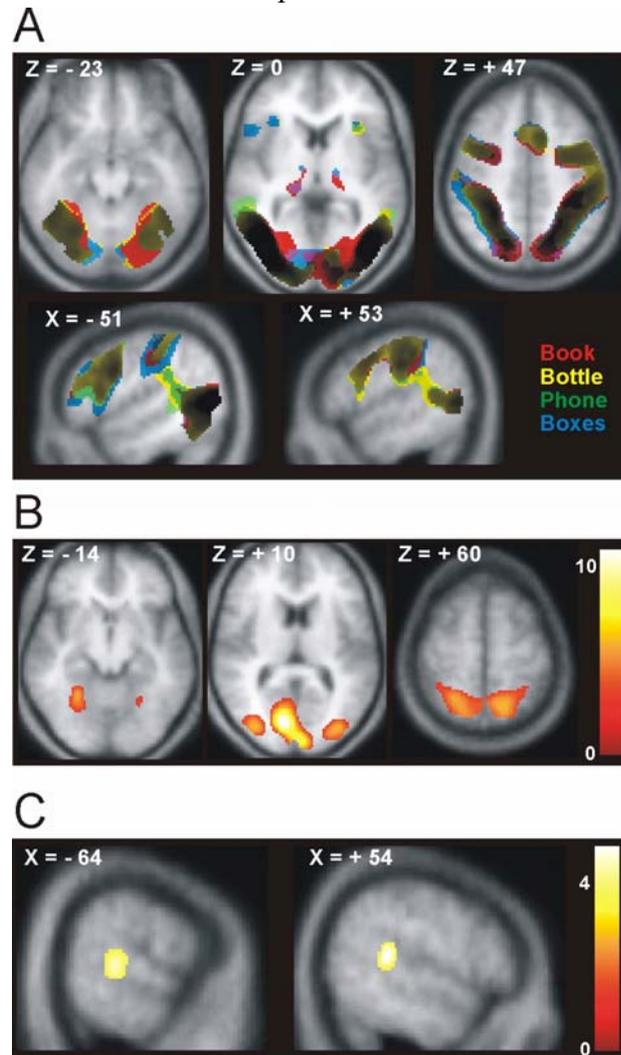
*fMRI analysis: 1) whole-brain analysis*

For this analysis, all animations of each manipulation type were compared against the resting actor, then both the main effects of movement type and the main effects of object presence were tested, and then the interaction between them.

**All animations vs. resting actor**

Animations of all types of manipulation vs. a resting actor without object showed significant activation increases (all at  $P < 0.05$  corrected for multiple comparisons across all voxels of the brain) in dorsal and ventral parietal cortices, fusiform gyrus, lateral occipital areas, posterior part of the superior temporal sulcus, ventral and dorsal premotor cortices, SMA/pre-SMA, and thalamus, all in both hemispheres (Figure Exp 3.4 A, Table Exp 3.2). Book and Boxes induced significantly greater activation than Bottle and Phone (all at  $P < 0.05$  corrected for multiple comparisons across all voxels of the brain) in medial occipital cortex, fusiform gyrus, lateral occipital cortex, and dorsal parietal cortex, all in both

hemispheres (Figure Exp 3.4 B, Table Exp 3.3). Bottle and Phone induced greater activation than Book and Boxes in bilateral pSTS (Figure Exp 3.4 C, Table Exp 3.4) (only at the slightly lower thresholds of  $P < 0.0001$  uncorrected,  $P < 0.06$  corrected for multiple comparisons across the whole brain,  $P < 0.01$  corrected for multiple comparisons across all voxels of the posterior STS as defined anatomically).



**Figure Exp 3.5. Brain activation during all conditions with a moving actor vs. the resting actor image.**

**A** For each manipulation type, areas commonly activated by all manipulation types are in green-brown. Image thresholded at  $P < 0.05$  corrected for multiple comparisons across the whole brain. **B** Areas more activated by manipulation types Book and Boxes (correctly categorised by the participants) than Phone and Bottle (incorrectly categorised). Threshold:  $P < 0.05$  corrected (whole brain) **C** Greater activation in the posterior superior temporal sulcus during observation of Phone and Bottle than Book and Boxes manipulations. No other regions showed significant activation increase in this contrast. Threshold:  $P < 0.05$  corrected (pSTS search region;  $P = 0.054$  corrected for whole brain).

**Table Exp 3.2.** Brain activation 1. Regions commonly activated for all types of object manipulations (all survive  $p \text{ corr} < 0.05$  FWE at voxel level and are  $> 5$  voxels in size).

<i>Anatomy</i>	<i>Area</i>	<i>Side</i>	<i>X</i>	<i>Y</i>	<i>Z</i>	<i>Z score</i>
Middle Occipital Gyrus	BA 18	R	44	-76	0	Inf
Inf. Temporal Gyrus	BA 18	L	-48	-74	-2	Inf
Lingual Gyrus	BA 18	R	28	-72	-12	Inf
Middle Occipital Gyrus	BA 18	R	24	-92	18	Inf
Middle Occipital Gyrus	BA 18	L	-30	-92	14	Inf
Cuneus	BA 17-18	R	10	-98	4	Inf
Precuneus	BA 19	L	-26	-82	36	Inf
Fusiform Gyrus	BA 37	R	38	-50	-22	Inf
Mid. Temporal Gyrus	BA 22 - 21	R	54	-50	6	6.02
Sup. Temporal Gyrus	BA 42	L	-60	-36	16	5.62
Inferior Parietal Lobule	BA 40	R	46	-30	20	4.54
Inferior Parietal Lobule	BA 40	R	34	-48	58	Inf
Inferior Parietal Lobule	BA 40	L	-38	-44	58	Inf
Sup. Parietal Lobule	BA 7	R	18	-66	54	Inf
Precuneus / SPL	BA 7	L	-18	-72	52	Inf
Postcentral Gyrus	BA 2-3	R	50	-26	42	Inf
Inferior Frontal Gyrus	BA 47	R	30	24	-4	4.63
Inferior Frontal Gyrus	BA 47	L	-32	26	-8	6.43
Inferior Frontal Gyrus	BA 9	L	-54	12	32	Inf
Inferior Frontal Gyrus	BA 45	L	-52	18	12	5.59

Inferior Frontal Gyrus	BA 46	L	-52	42	8	4.49
Medial Frontal Gyrus	BA 6	R	4	2	56	7.36
Middle Frontal Gyrus	BA 6	R	24	-6	58	Inf
Precentral Gyrus	BA 6 or 9	R	50	2	40	Inf
Precentral Gyrus	BA 6	L	-64	-14	40	5.77
Middle Frontal Gyrus	BA 6	L	-32	-2	52	Inf
Pulvinar		L	-20	-28	2	5.34
Thalamus		R	14	-20	4	4.89
Cerebellum		L	-14	-76	-50	4.93
Cerebellum		L	-20	-78	-18	5.9
Cerebellum		R	4	-76	-36	5.31

**Table Exp 3.3.** Brain activation 2. Regions showing activation differences depending on the type of manipulation (during all conditions; all survive p corr <0.05 FWE corrected for voxels of the whole brain or the pSTS).

1) Greater activation during Book and Boxes displays than Bottle and Phone.

<i>Anatomy</i>	<i>Area</i>	<i>Side</i>	<i>X</i>	<i>Y</i>	<i>Z</i>	<i>Z score</i>
Cuneus	BA 17	L	-12	-82	10	Inf
Fusiform Gyrus	BA 19	L	-28	-56	-10	7.28
Middle Occipital Gyrus	BA 19	L	-36	-86	12	6.69
Middle Occipital Gyrus	BA 19	R	34	-86	14	7.12
Precuneus / SPL	BA 7	R	12	-64	62	6.69
Sup. Parietal Lobule	BA 7	L	-22	-60	64	6.45
Postcentral Gyrus	BA 2	R	36	-32	42	5.01

Cerebellum: Declive		R	24	-58	-16	4.68
---------------------	--	---	----	-----	-----	------

2) Greater activation during Bottle and Phone displays than Book and Boxes.

<i>Anatomy</i>	<i>Area</i>	<i>Side</i>	<i>X</i>	<i>Y</i>	<i>Z</i>	<i>Z score</i>
Posterior STS	BA 22	R	54	-42	12	4.42
Posterior STS	BA 22	L	-64	-34	4	4

### **Main effect 1: Effect of movement type**

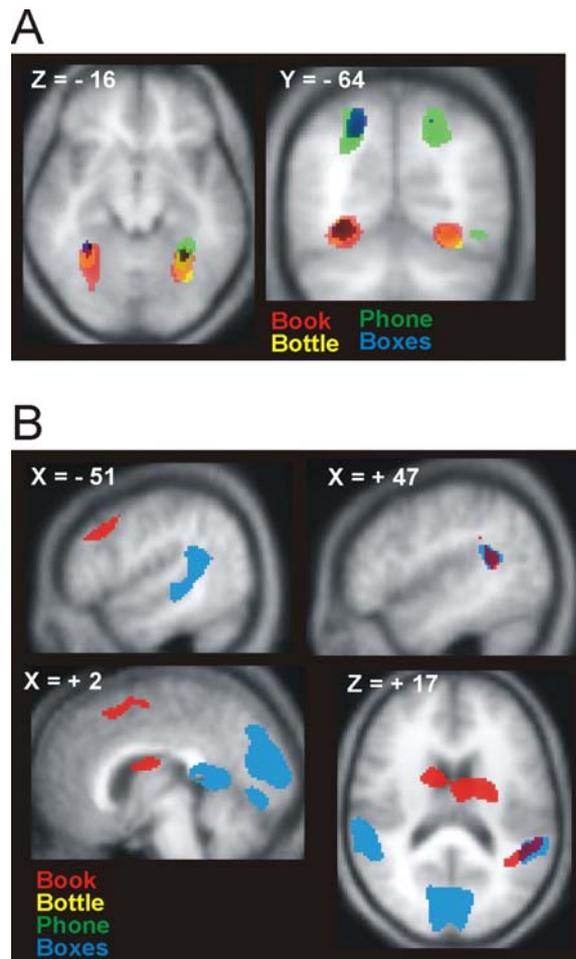
There were no significant activation differences between animations with Mimed and Actual movement kinematics, for any type of manipulation.

### **Main effect 2: Effect of object presence on manipulation of all types**

Observation of animations where the manipulation movement was shown with an object compared to animations without an object (Figure Exp 3.5 A) induced stronger activation in fusiform (left hemisphere for all manipulations, only at uncorrected threshold for Phone; right hemisphere for all, with Boxes and Phone at uncorrected threshold only), dorsal parietal cortex (left hemisphere: Phone and Boxes, right hemisphere: Phone, Boxes at uncorrected threshold) and middle occipital gyrus (left and right hemisphere, both for Phone and Boxes, Bottle at  $P < 0.001$  uncorrected threshold only). In Book there was a significant activation in left posterior medial occipital cortex.

Animations where the manipulation movement was shown without an object induced significantly stronger activation only for the Book and Boxes manipulations when compared with animations presented with an object (Figure Exp 3.5 B). The activated regions were mostly different between Boxes and Book; the only common

region with significantly increased activation was right posterior STG. Other regions significantly active in Book where: left dorsal premotor cortex, medial superior frontal cortex and the body of the caudate nucleus. Other regions significantly active in Boxes where: left posterior STS and STG, posterior cingulate and medial occipital cortex.



**Figure Exp 3.6. Effects of object presence.**

**A** Brain regions more active in conditions with objects than conditions without objects, for each type of object manipulation. Overlaps are found in the fusiform gyrus and the lateral occipital cortex. Image thresholded at  $P < 0.05$  corrected for multiple comparisons across the whole brain. **B** Brain regions with higher activation in trials without objects compared to trials with objects, for all object manipulation types. Only Book and Boxes showed significant activation in this contrast. Threshold:  $P < 0.05$  corrected (cluster level, identified at  $P < 0.001$  uncorrected).

**Table Exp 3.4.** Brain activation 3. Regions showing activation differences during conditions with Object and Without object (\* survive  $P < 0.05$  whole-brain corrected at voxel level, other survive correction at cluster level only). H refers to Hemisphere, BA to Brodmann Area.

1) Greater activation during Object than No object displays.

<i>Anatomy</i>	<i>H</i>	<i>BA</i>	<i>X</i>	<i>+ -</i>	<i>Y</i>	<i>+ -</i>	<i>Z</i>	<i>+ -</i>	<i>Objects (Z score)</i>
R fusiform	R	19/3 7	36	(9)	58	(10)	-15	(5)	Bk(5.7)* Bt(4.4)* Ph(3.4) Bx(3.5)
L fusiform	L	19/3 7	-30	(2)	-53	(8)	-13	(3)	Bk(7.1)* Bt(4.9)* Ph(3.9) Bx(4.7)
R mOccG	R	19	31	(5)	-90	(2)	15	(1)	Bt(3.2) Ph(5.3)* Bx(3.9)
L mOccG	L	19	-33	(3)	-90	(4)	16	(2)	Bt(3.4) Ph(3.6) Bx(4.5)
R SPL	R	7	24	(6)	-61	(11)	55	(11)	Ph(4.7)* Bx(3.5)
L SPL	L	7	-25	(1)	-66	(0)	46	(4)	Ph(4.3)* Bx(4.3)
R IPL	R	40	60	(0)	-34	(0)	44	(0)	Ph(3.1)

2) Greater activation during No object than Object displays.

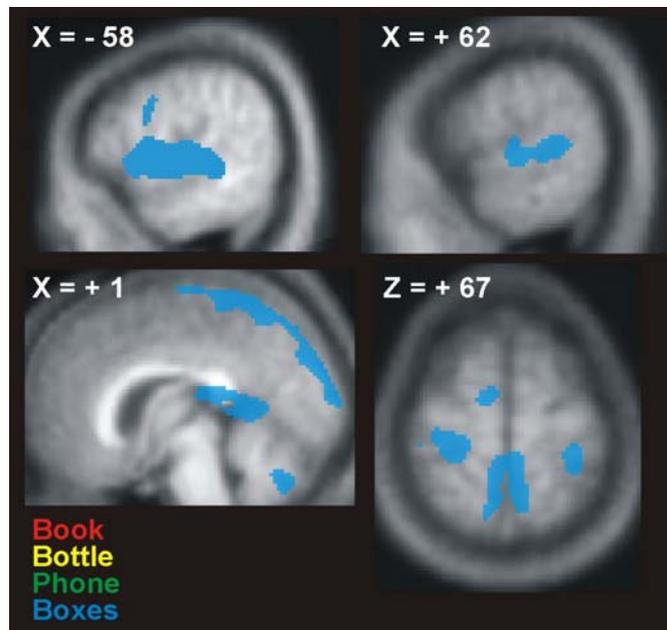
<i>Anatomy</i>	<i>H</i>	<i>BA</i>	<i>X</i>	<i>+ -</i>	<i>Y</i>	<i>+ -</i>	<i>Z</i>	<i>+ -</i>	<i>Objects (Z score)</i>
Mid.-Sup. Front Gyr.	L	8-9	-36		26		50		Bk(4)*
Inf. Parietal Lobule	R	40	54		-40		24		Bk(3.9)*
Sup. Temporal Gyrus	R	22	49	(5)	-47	(1)	14	(2)	Bk(3.7)* Bx(4.3)*
Inferior Frontal Gyrus	L	47	-56		34		0		Bk(3.8)

Middle Frontal Gyrus	R	9	44		16		34		Bk(3.6)
Middle Frontal Gyrus	R	10	40		40		26		Bk(3.3)
Middle Frontal Gyrus	L	6	-56		2		46		Bx(3.2)
Inferior Frontal Gyrus	R	46-45	44		26		14		Bk(3.4)
Inferior Frontal Gyrus	L	45-44	-48		18		16		Bk(3.6)
Precentral Gyrus	L	6	-38		-8		60		Bk(3.2)
Medial Frontal Gyrus	R	6	4		16		52		Bx(3.7)*
Mid. Temporal Gyrus	L	21	-62		-38		-4		Bx(4.5)*
Mid. Temporal Gyrus	R	21	60		-30		-8		Bx(4.1)
Sup. Temporal Gyrus	L	38	-58		14		-4		Bx(3.5)

**Interaction: Regions responding differently to artificial or realistic movement-object combinations.**

Combinations of movement and object that represent realistic situations (i.e. a Mimed movement presented without object and an Actual movement presented with an object) induced greater activation (Figure Exp 3.6, Table Exp 3.5) than artificial situations (Mimed movement with object and Actual movement without object) in a number of regions, but again only for Book and Boxes manipulations (although activation for Book manipulations only survived a threshold of  $P < 0.0001$  uncorrected). The regions with increased activation were again mostly different between Boxes and Book; the only region with overlapping activation increases was located in the inferior left precentral gyrus. Other regions were significantly active for the Boxes manipulations: bilateral pSTG extending into pSTS on the left, posterior cingulate, posterior medial occipital cortex, superior medial parietal cortex and left superior precentral gyrus.

There were no regions with significantly higher activation in the opposite contrast.



**Figure Exp 3.7. Congruence effects.**

Brain regions with higher activation during Realistic than Artificial movement / object combinations. This contrast compared conditions Mimed No object and Real Object with Mimed Object and Real No object. Only Boxes showed significant activations in this contrast. Threshold:  $P < 0.05$  corrected (cluster level, based on  $P < 0.001$  uncorrected).

**Table Exp 3.5. Brain activation 4. Regions showing greater activation during Realistic (congruent) vs. Artificial (incongruent) movement / object combinations. Threshold used was  $P < 0.05$  corrected at the cluster level (based on  $P < 0.001$  uncorrected map).**

<i>Anatomy</i>	<i>Side</i>	<i>Area</i>	<i>X</i>	<i>Y</i>	<i>Z</i>	<i>Object (Z score)</i>
Putamen	L		-14	16	-4	Bx(3.2)
Thalamus	L		-16	-30	6	Bx(3.7)
Insula	R	BA 13	46	0	4	Bx(3.5)
Precentral Gyrus	L	BA 4	-32	-34	62	Bx(3.4)
Precentral Gyrus	L	BA 4	-60	-2	18	Bk(3.9)**

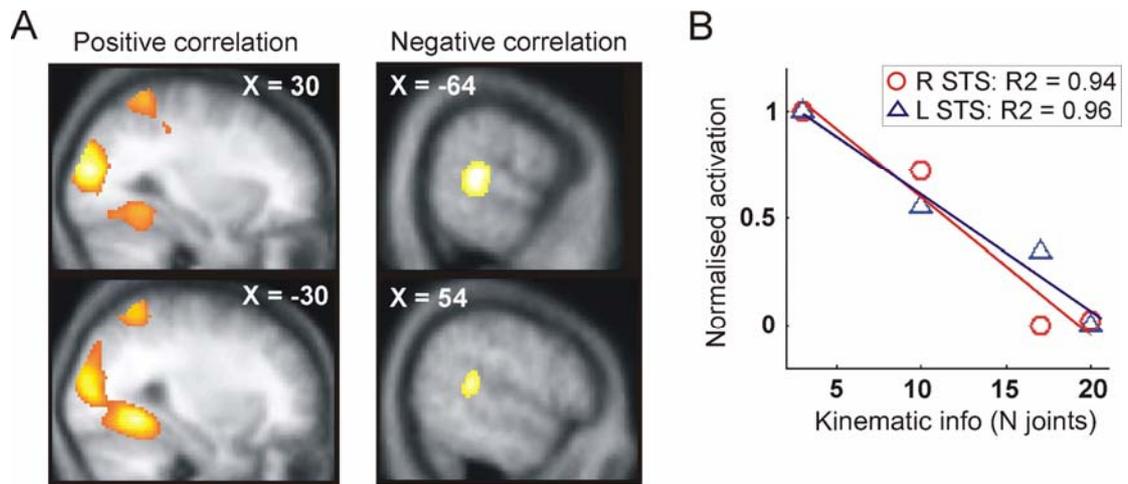
Precentral Gyrus	R	BA 4	58	-6	20	Bk(3.7)*
Middle Frontal Gyrus	L	BA 6	-42	0	54	Bx(3.4)
Superior Frontal Gyrus	L	BA 6	-22	-8	72	Bx(3.3)
Superior Frontal Gyrus	R	BA 6	16	-14	72	Bx(3.2)
Postcentral Gyrus	R	BA 43	58	-18	14	Bx(3.3)
Postcentral Gyrus	R	BA 7	4	-48	70	Bx(4.1)
Postcentral Gyrus	L	BA 7	-8	-52	70	Bx(3.9)
Postcentral Gyrus	L	BA 3	-44	-28	60	Bx(3.4)
Postcentral Gyrus	R	BA 3	28	-32	70	Bx(3.4)
Inferior Parietal Lobule	L	BA 40	-50	-30	26	Bx(3.5)
Sup. Temporal Gyrus	R	BA 22	60	-42	8	Bx(3.7)
Mid. Temporal Gyrus	L	BA 39	-34	-64	22	Bk(4.4)
Mid. Temporal Gyrus	R	BA 39	44	-48	6	Bx(3.7)

*fMRI analysis 2: Correlation with kinematic information*

As discrimination performance showed a trend to correlate with kinematic information (Figure Exp 3.4), and given that correctly categorised animations induced different activation than incorrectly categorised animations (Figure Exp 3.5 B and C), correlation between kinematic information and brain activation was tested directly in SPM. Regression analyses were performed on contrast images from each manipulation type comparing animations across all experimental conditions with the resting actor, using as explanatory variables the number of joint variables with significant differences in kinematics between Mimed and Actual movement versions (Table Exp 3.1). I looked for voxels which showed positive or negative correlation

with kinematic information. A positive correlation was found in occipital and parietal regions, including posterior medial occipital cortex, lateral occipital areas, fusiform gyrus, and medial superior dorsal parietal cortex. The only brain regions that showed a significant inverse correlation with kinematic information were clusters in bilateral pSTS, and a small cluster in left medial occipital cortex (Figure Exp 3.8 A).

Given the evidence on the role of the pSTS during action observation and biological motion and given the results of the experiments on observation of animate-looking, moving objects, I decided to complement the analysis performed within SPM with a region-of-interest analysis on the pSTS. I averaged the parameter estimates from the clusters of voxels in the pSTS responding more strongly to incorrectly discriminated than correctly discriminated animations (see Figure Exp 3.5 C) and correlated this average activation with the number of joints showing significant differences in kinematic information between Mimed and Actual manipulation versions (Figure Exp 3.8 B). Activation in the pSTS of both hemispheres showed a strong inverse correlation with kinematic information (Left:  $R^2 = 0.96$ ,  $p = 0.02$ ; Right:  $R^2 = 0.94$ ,  $p = 0.03$ ), thereby confirming the results of the regression analysis performed directly in SPM.



**Figure Exp 3.8. Correlation between stimuli, and pSTS activation.**

**A** Correlation between kinematic information in all four objects (measured as number of joint variables that showed significant difference between Mimed and Actual kinematics during object manipulation) and brain activation performed in SPM. Left part: positive correlation (Thresholded at  $p < 0.05$  corrected for multiple comparisons across all the voxels of the brain). Right part: negative correlation. Only the pSTS and a small cluster in left medial occipital cortex (not shown) showed significant inverse correlation (Threshold:  $p < 0.001$  uncorrected). **B** Correlation between kinematic information in all four objects and the activation in posterior STS (normalised data, Right hemisphere in red, left hemisphere in blue).

### Discussion

This study yielded the following behavioural and brain activation results. Analysis of behaviour showed that subjects were able to correctly categorise half of the presented manipulations observed into Mimed and Actual movements on the basis of their kinematics, while the presence of an object irrelevant to the task increased their performance in one occasion. This categorisation was made correctly only when 1) more than a third of the movement parameters show significant objective differences between Mimed and Actual kinematics and 2) when these differences are present during more than just the contact phase between hand and object. Discrimination performance showed a correlative trend with kinematic information.

Analysis of brain activation showed that comparing all types of manipulation with a resting actor reveals significant activation increases in brain regions known to be involved in action observation (premotor cortex, STS) and in primary and secondary visual regions (including fusiform gyrus, LOC, STS and dorsal parietal cortex). Activation in some of these areas was dependent on kinematic information available for the task at hand: activation was greater when there was more kinematic information in many visual areas (posterior medial occipital cortex, fusiform gyrus, LOC, dorsal parietal cortex), and activation in the pSTS correlated inversely with kinematic information. When subjects observed object manipulations they could categorise correctly into Mimed and Actual, activation in the right posterior STG was increased when no object was presented in the animation compared to when an object was present.

I propose that these data suggest a role of the superior temporal sulcus in the extraction of task-relevant biological motion information from observed human movements, with activation compensating for task difficulty (caused in the present study by less kinematic differences between the movements to categorise or no information about the movement of the object during the manipulation). This is in opposition to other brain areas responding during action observation, such as the other parts of the visual system whose activation appears to increase with kinematic information, and to the premotor cortex whose activity does not seem to be influenced by kinematic information.

### *Stimuli*

Our behavioural results show that when 33% or more joint variables showed differences between Mimed and Actual movements, subjects were able to discriminate between these two movement versions, but only the manipulations with

57 % or more joint variables showing significant differences were discriminated correctly. In the correctly discriminated manipulations, three-quarters or more of the variables showed significant differences between Mimed and Actual versions during the whole duration of the movement. Therefore, necessary for a correct categorisation were significant differences between Mimed and Actual movements in a certain number of joints, during a period longer than just the contact phase between the object and the actor's hand. Interestingly, the presence of the object had only a mild effect on performance: it increased only for Boxes, where categorisation was already successful in their absence. Previous studies showed that movement kinematics of Mimed and Actual movements can be different (Goodale, Jakobson, & Keillor, 1994), and also show less variation between replications of a particular movement by one person than do Actual movements (Mason, Gomez, & Ebner, 2001). Studies also showed that observers can tell when people experience violations of their expectations about target objects, such as a box being lighter or heavier than expected when picked up (Runeson & Frykholm, 1983; Grezes, Frith, & Passingham, 2004). The present study shows that subtle differences in movement kinematics, probably resulting from the effect of the object on the movement, can be picked up by subjects, if they exist in a sufficient number of the joints involved in the movement and if they last a certain time. The importance of the part of the movement before contact with the object (the reaching part) as a defining part of the action can be seen from studies which showed that two principal components can explain more than 80 percent of the variation of hand and finger joints during a variety of reaching and grasp actions (Santello, Flanders, & Soechting, 1998; Mason, Gomez, & Ebner, 2001), with the second of these components corresponding to the reaching and preshaping phase of the hand. It is therefore not surprising that differences between kinematics of Mimed and Actual

movements need to be found also in this part of the movement for correct identification by observers.

Convincing pantomime artists, who create the impression that they are manipulating an object that is not actually present, probably base their performance on their capacity to minimise the kinematic differences between their mime and the actual manipulations of an object, resulting for observers in the impression that the object actually had an effect on the pantomime's movements. Recording movement kinematics from such artists could therefore be an interesting extension to my behavioural results.

#### *STS and processing of biological motion*

As discussed in Part 1, Chapter 3, Section 3.2.2, there is a growing amount of data suggesting that the region surrounding the posterior STS and STG is implicated in action observation and observation of biological motion. The present study shows that activity in this area increases with task difficulty during an action-observation task (note: this cannot simply reflect a time-on-task effect, as activity was modelled only from stimulus onset to subject's response). Particularly, I found the following effects. First, activation was greater during observation of manipulations for which Mimed and Actual versions show little kinematic differences, which were also the movements that subjects had more trouble categorising into Mimed or Actual. This suggests that the pSTS "works harder" to extract the movement information contained in the stimuli and relevant to the task.

Second, during observation of the two types of manipulation that subjects could discriminate well, activation in the pSTS was greater when there was no object present in the animation. The behavioural data show that the presence of an object increased subjects' Mimed vs. Actual discrimination of the Boxes manipulations.

Animations without objects appear to be more difficult to categorise than animations with objects, which might be due to the absence of information available about the way the object moved in response to the actor's movement; this information might have informed subjects about the type of movement they were observing. These two facts suggest that the more information was available to the subjects to solve their task, the better they succeeded and the less pSTS recruitment was necessary (Figure Exp 3.9). This is in opposition to the responses of many other visual areas also responding to animations of human actions, whose activation increased with available kinematic information. This might suggest that these regions, being less specialised in the analysis of biological motion than the pSTS, simply respond more when there are more visual information available to solve the task.

I believe that these data consolidate the link between processing of biological movements and the pSTS (for a recent review see Puce & Perrett, 2003). Discriminating between subtly different movements could be useful for the detection of non-verbal communication cues and the identification of mental states of the observed people. This would conform to the role of the pSTS in the processing of biological motion on the basis of which we make social attributions (Adolphs, 2003). How certain regions could be specialised for the processing of moving or changing aspects of a particular type of socially relevant stimuli has been described in detail in a recent model of face perception (Haxby, Hoffman, & Gobbini, 2002).

#### *STS and attribution of mental states*

A function possibly related to action observation and the processing of biological motion and associated with the pSTS is the attribution of mental states to other agents, as seen in a variety of tasks and substrates (for a review, see Frith &

Frith, 2003). A recent study combining mental state attribution and action observation showed that activity in the pSTS and parts of the premotor and orbitofrontal cortex was greater when subjects thought that a person whose action they observed had a false belief about the object he or she was manipulating (Grezes, Frith, & Passingham, 2004). In the present experiment, I did not find differences in activation between observation of Mimed or Actual movements. Although this could be seen as clashing with the results of the study by Grezes and colleagues, I believe that the task and stimuli I used did not cue subjects into thinking about the mental states of the actor executing the action. In my experiment, subjects were not told that the actor had particular beliefs about the object he was to manipulate: they were only told that the actor had mimed an action while observing the object, or had really performed the action with the object. By contrast, in the study by Grezes et al, participants were told that the person they were observing had sometimes been misled as to the weight of the object they had to lift. This difference could have led subjects in my experiment to pay more attention to the differences in kinematics than to differences in mental states of the actor. This could explain why I found that activation in the pSTS varied depending on the amount of kinematic differences between Mimed and Actual movements the subjects had to observe, but not depending on the type of movement kinematics themselves (Mimed or Actual).

#### *STS and Goal-directed action*

The posterior STS or TPJ is also known to be activated during execution and imitation of goal-directed actions in humans (Winstein, Grafton, & Pohl, 1997; Jacoboni et al., 2001). In monkeys also, neurons responding during observation of actions have been found in this brain region (Jellema, Baker, Wicker, & Perrett,

2000). As discussed above, I did not find regions with activation differences between Mimed and Actual conditions.

I do not think that this absence of an effect necessarily reduces the association between pSTS and goal-directed action. Goals are defined differently by different authors, as described in a recent study by Koski and colleagues (Koski, Iacoboni, Dubeau, Woods, & Mazziotta, 2003). The goal-directed theory of imitation (Bekkering, Wohlschlaeger, & Gattis, 2002) defines goals as physical objects that can be targets for reaching and grasping movements, but also as a representation of the goal in neural codes, in a “functional mechanism necessary to initiate an imitative action” (Koski, Iacoboni, Dubeau, Woods, & Mazziotta, 2003). Tomasello separates action goals from the means to achieve them (Tomasello, 1999), and Travis defines a goal as a “mental state representing a desired state of affairs in the world” (Travis, 1997). Therefore, depending on the definition of goal, a mimed movement can be considered goal-directed or not. If one considers a mimed movement as goal-directed (after all, one recognises the object or the goal of a mimed action, which is really what makes the mimed movement a mime, or one could say that the end-state of the mimed movement is the goal of the movement), then the absence of an activation difference in the pSTS between mimed and actual movements could be due to the fact that there was a goal in both types of movements. If one defines mimed movements as not goal-directed however, the present data can be seen to show the role of the pSTS in extracting information from observed movements in order to discriminate them in terms of goal-directedness. This would be supported by greater activation in the pSTS during observation of movements that contained less information necessary for the discrimination into Mimed and Actual. Whether a mimed movement is a goal-directed movement or not does not influence the interpretation of the correlation between

activation and kinematic information very much: in both cases, the pSTS could be involved in the extraction of task-relevant kinematic information. This result is certainly more interesting than the absence of an activation difference between observation of mimed and actual movements.

#### *Ventral vs. dorsal stream*

As discussed in the introduction, the ventral stream is thought to control mimed movements, at least to observed objects, while object-directed actions are thought to be controlled by the dorsal stream. Therefore it could have been possible that the two visual streams would be activated differently also during *observation* of Mimed and Actual actions. My results show that there was no significant difference in activation in these two streams between observation of Mimed and Actual movements. Rather, there was stronger activation in parts of both visual streams during observation of correctly discriminated object manipulations and manipulations with more kinematic information available for the task (medial dorsal parietal cortex, lateral and medial occipital cortex, and fusiform gyri, all bilaterally, see Figures Exp 3.5 B and Exp 3.9 A), and stronger activation in bilateral pSTS during observation of less well discriminated object manipulations or manipulation with less kinematic information available for the task (Figures Exp 3.5 C and Exp 3.9 A). Therefore, it appears that the difference in visual stream involvement that is found for the execution of mimed or actual actions is not found for observation of these actions.

#### *Activation of action-observation areas with animations*

Although a previous study suggested that only the observation of physically present actors can lead to activation of all parts of the action observation system (Perani et al., 2001), I did find activation in all the areas classically associated with action observation. Although there was some variation in activation in some of these areas depending on the manipulation observed (Figure Exp 3.4 A and B), strong and very similar activation patterns in these areas were found for all four types of manipulation.

In contrast with my hypothesis, regions of the ventral premotor cortex that could be analogues of monkey area F5 and could contain Mirror Neurons (Areas 44 and/or 45, see for example Grezes, Armony, Rowe, & Passingham, 2003) did not show significant differences between observation of Mimed or Actual movement versions or conditions with Object present or absent, for either type of object manipulation. This is in contrast with neurophysiological findings (Umiltà et al., 2001, see Figure Intro.8) which show that Mirror Neurons respond during object-directed actions even when the grasped object and the contact between the hand and object are out of view. This was taken to imply that Mirror Neurons might identify the object-directedness of the movement through analysis of the movement kinematics, rather than attending to object presence. My present results could suggest either that 1) in humans, the response of Mirror Neuron-analogue cells is different than in monkeys, or 2) cells responding like monkey Mirror Neurons are mixed among other cells responding differently and the resolution of fMRI cannot distinguish them, or 3) there is inhibitory neural activity in the tissue surrounding Mirror Neuron-like cells during conditions with Mimed kinematics, which would also lead to an increase in metabolic rate and therefore to an increase in BOLD signal (Logothetis & Wandell, 2004), and therefore no difference was found. Alternatively, Mirror Neuron-like cells are more

sensitive to the realism of the stimuli than other structures involved in action observation in the human, and are therefore not triggered by my virtual reality animations.

*Realistic vs. artificial movement-object combinations*

Only for animations depicting the manipulation of the two boxes, observation of Realistic object and movement kinematics combinations resulted in higher activation than during artificial combinations in the following brain regions: in the pSTS of both hemispheres, anterior superior STS/inferior frontal gyrus (BA 22/44/BA6), anterior dorsal (BA 1 and BA 3) and medial dorsal parietal cortex, posterior cingulate gyrus and medial occipital cortex. Realistic vs. Artificial combinations correspond to the effect of object presence on Mimed vs. Actual difference. Animations of the boxes manipulations were the only type of manipulation in which the object presence had a significant effect on Mimed/Actual discrimination performance (see Figure Exp 3.3). Therefore the effect of object presence on brain activation is paralleled in the subjects' performance. Whether these two observations are causally correlated, and in which way, could not be determined in this study.

## **Part 5. General discussion**

### **1. So what is this thesis all about?**

The purpose of this thesis was to study the neural correlates of a basic cognitive component of the mentalising ability, also known as Theory-of-Mind or ToM. ToM is defined as the process of representing and manipulating mental states of the self or other people. Previous neuroimaging studies from different research groups have shown that 3 brain structures seem to be commonly activated in different tasks involving mentalising: the medial prefrontal cortex (mPFC), the temporal pole, particularly in the left hemisphere (TMPp) and the posterior superior temporal sulcus (STS or pSTS). Although all 3 regions are known to show activation also during other processes that could be related to or involved in mentalising, the relative contributions of these regions to mentalising are speculative. The work in this thesis was concentrated on a first step judged to be essential for ToM: the identification of living entities in the environment.

This process was approached by studying types of movement particular to living entities, because on the basis of developmental psychology studies, actions of living entities appear to be very characteristic for living organisms and important for the attribution of animacy. Also, arguably the most important use for mentalising is the explanation and prediction of the behaviour of agents, which adds relevance to the study of animate movement rather than shape, which was not studied in this experiment series.

I suggest that a neural system specialised in the determination of what is animate and what is not could rely on the identification of common characteristics in all types of animate motion, which could be hard-wired and present at birth or acquired by

generalisation from the observation of diverse animate entities during development. A neural system involved in detecting animate things should thus respond to these characteristics. Developmental studies suggest that goal-directed and self-propelled (also called internally-generated, see Premack and Premack, 1995) movements are the most important movement characteristics for the attribution of animacy.

As one of the 3 neural structures commonly activated during mentalising tasks, the superior temporal sulcus, is known to respond during the observation of biological motion, I set out to test how the posterior STS responds to simple animations displaying characteristics of animate motion, and see if this response is dependent of attention or not. To this end, I first developed an algorithm for the motion of two interacting, simple and abstract shapes, in which the variation of one parameter could affect attribution of animacy by naïve observers.

Then, in an fMRI experiment, I looked for brain regions showing activation correlating with the value of this parameter: they included the pSTS and other regions associated with processing of living entities: the fusiform gyrus and the medial occipital cortex, and two regions in the cingulate and posterior paracingulate cortex. My results further suggest that this response is independent of the subject's focus of attention. In a second experiment, in which I introduced a more complex parameter also characteristic of animate movements and related to mentalising, I showed that the pSTS responded again when the abstract object displayed more animate behaviour, but this time, this response was increased when subjects paid attention to the parameter associated with animacy. This suggests that simple and fundamental movement characteristics of animate motion may activate the pSTS automatically, while this structure responds to more complex characteristics, indicative of representations of goals, only when subjects look for them. While the former response

is sensible for an automatic system detecting animate entities, the latter corresponds more to a mentalising-like process, which is most probably a conscious process.

Both these experiments used moving, abstract objects. However, mentalising is performed mostly in contact with other complex living animals, mostly human beings. I therefore extended the study of biological movement to human actions in another fMRI experiment, which showed that pSTS responds not only to biological movements of different kinds, but also analyses these movements in order to understand them by "extracting" kinematic information from observed human movements. The subject's task was to determine if the observed movement was mimed or not, which can be interpreted as goal-directed or not, depending on the definition of goals one uses (see Introduction). The results showed that activity in pSTS was higher when less kinematic information was present, suggesting that the activity compensated for task difficulty. Evidently, evaluating whether an observed movement is goal-directed or not is important for the identification of the mental state of the agent performing the action, and therefore activity in the pSTS appears again to be related to the mentalising process.

## **2. Activation in pSTS during observation of abstract and human movements**

I would like to propose that pSTS activation in all experiments can be explained by a role of this region in detection and analysis of biological motion. PSTS 1) shows higher activation when observed movements look more animate (i.e. more goal-directed vs. less goal-directed moving abstract shapes, shapes appearing to attribute goals to target objects vs. simply following, human actions vs. resting actor), and 2) when observing different movements of equal animacy, pSTS activation adapts to demands of a task that requires extraction of information from the observed

movements. These roles are compatible with previous studies of biological motion and as a part of the neural circuitry involved in mentalising.

With respect to goal-directedness, we cannot say that pSTS activation is always greater when goals are present. While in both experiments involving abstract objects pSTS activation was greater when objects showed more goal-directedness or appeared to use representations of goals of a target object, in the action observation experiment, pSTS activation was greater when there was less difference in kinematics between the goal-directed and the non-goal-directed movement versions, but showed no systematic activation difference depending on goal-directedness itself. While these responses are slightly different, both suggest detection and processing of movement cues indicative of goal-directed movements. I suggested above that this could be due to different processes both performed by the pSTS: 1) detection and 2) analysis of biological movements. The difference could also be due to different levels of expertise of pSTS in abstract moving object movements and human actions, leading to slightly different functions depending on the type of biological movements presented and the task to be performed on them.

This pattern of response by the pSTS reminds of the response of the Visual Word Form Area (VWFA) in the left fusiform gyrus (Cohen et al., 2000; Cohen et al., 2002), whose activity is greater during observation of words rather than consonant strings. But similarly to the activity in the pSTS found in action-categorisation experiment of this thesis, activity in the VWFA is higher when pseudowords or words of low frequency rather than high-frequency words are presented (Kronbichler et al., 2004). This could reflect a similar increase in "effort" of the neuronal population in trying to identify which word is presented, once the stimulus is close enough to a real word (i.e. it contains vowels and consonants). I would argue that this response is

similar to that of the pSTS in the action-categorisation task of the thesis, in which, once a visual stimulus has been identified as biological with reasonable certainty, a greater "effort" is required when less kinematic information is available to solve the task to be performed on the biological movements.

The involvement of pSTS in social cognition, including analysis of changeable parts of faces and moving human bodies, suggests that this structure is specialised for processing of socially relevant motion data. That it responds also when abstract objects appear to interact in a fashion evocative of human- or animal-like behaviour is interesting, but one should not forget that such animations are abstract and thus not encountered often in normal life, and that therefore the pSTS has not much direct experience in viewing it. If pSTS has evolved as a structure enabling recognition and processing of social information in conspecifics (such as face and body movements), then it might be expected to respond differently to such information rather than to moving abstract, artificial stimuli, however sophisticated they are. Particularly, observing human actions and trying to find which goal they seek to attain is an everyday process for us, and the pSTS, being so specialised for human motion, could adapt its involvement depending on task difficulty in order to supply other brain areas specialised in action observation (such as the premotor cortex or intraparietal sulcus) with enough information so that they can process the action. Detecting animate motion cues in all types of movements (even abstract) is another process, with a different purpose: more like a warning system ("there is maybe something living over there, let's keep an eye on it"), which would be expected to respond differently to clear human motion compared to abstract motion.

These differences in expertise might also lead to a different effect and role of goals. Goal-directed motion might be a cue for attribution of animacy in the motion of

simple-looking entities (such as small insects, animals very far away), where almost no information about animacy is given by their aspect. But in clearly human movements located very close to the observer, goal-directed movements are more subtle cues about the intentions of the person, or even substrates for imitation (see role of goals in imitation discussed in Part 1, Chapter 3, Section 3.2.1). While there are obvious links to intentionality in both cases, the different information content and use of it could lead to different involvement of the pSTS. In this view, it can even appear surprising and interesting that pSTS activation is connected to goals in both cases, albeit slightly differently.

It needs to be clarified that there are two categories of visual cues that can carry information about animacy: the movement and the aspect of an object (Gelman, Durgin, & Kaufman, 1995; Opfer, 2002; Johnson, 2003), of which only movement was studied in the experiments of this thesis. While I suggest that my data show a role of the pSTS in detection and analysis of movement characteristics of animate entities, it is possible that cues for animacy in the aspect of objects are processed by a different area, for example the fusiform gyrus, other parts of the LOC (such as the extra-striate body area, see Downing, Jiang, Shuman, & Kanwisher, 2001) or the medial occipital cortex. This network might have distributed specialisations, as has been proposed for processing of faces: static aspects might be processed in the lateral fusiform gyrus, whereas changing aspects by the STS (Haxby, Hoffman, & Gobbini, 2002). This interesting question needs to be addressed in further studies.

### **3. Other functions of the superior temporal sulcus and gyrus**

In addition to mentalising, processing of animate motion, faces and other socially relevant stimuli, the posterior part of the STS responds also to changes in

visual, auditory or sensory stimulation (Downar, Crawley, Mikulis, & Davis, 2000) independently of the context of the task (Downar, Crawley, Mikulis, & Davis, 2001), and even when there is no task to perform (Downar, Crawley, Mikulis, & Davis, 2002). It seems to be also involved in detection of visual oddballs (differing in form, location, or both) (Marois, Leung, & Gore, 2000), and could be involved in reorienting to visual events appearing at unexpected locations (Corbetta, Kincade, Ollinger, McAvoy, & Shulman, 2000).

I do not think that novelty or the presence of a change in the stimuli presented in the experiments of this thesis can explain the pSTS activations that were observed. In all studies, an equal number of trials were presented in the conditions that were compared to one another, and these were fully randomised (differently for each subject). Also, in experiments 1 and 2, a closely-matched control condition was used for each experimental condition, and always included in the contrasts of interest. These control conditions were presented as frequently as the experimental conditions, were also randomised, and could be considered to differ in terms of “novelty” and “change” only with great difficulty<sup>6</sup>. In experiment 2, one region showed different responses depending on the task performed, which should not affect novelty or change in stimulation. In experiments 1 and 3, novelty or change could hardly explain the correlations observed between stimulus characteristics and STS activation. I feel therefore confident to propose a different role for the pSTS than novelty or change

---

<sup>6</sup> An interesting idea would be to suggest that animate motion appears is characterised by being more novel or to contain more motion oddballs than inanimate motion. Indeed, one might consider that the aspects of the motion that cannot be explained by simple physical laws (see Part 1, Chapter 3, Section 3.1.1) or that are otherwise unexpected are critical for the movement to appear biological, and one might therefore consider these events as motion oddballs or novel trajectories. While this idea deserves further study, it is difficult to think that novelty is the only characteristic of biological motion given the numerous psychophysical studies demonstrating advanced identification of complex biological motion with very minimal stimuli (see Part 1, Chapter 3, Section 3.1.2), which very strongly suggests that some form of knowledge about biological motion is accessed and therefore stored in some form. The existence of such stored knowledge of biological movements appears difficult to reconcile with the postulated importance of novelty.

detection. Evidently, this structure might very well be involved in both types of processes, and maybe more.

The pSTS might also be involved in integration of sensory information within and between modalities (auditory and visual) about complex objects, such as pictures and sounds of animals and tools and videos of moving tools (Beauchamp, Lee, Argall, & Martin, 2004). Integrating such multisensory information could correspond to the role of the superior temporal polysensory area (STP) of the monkey (see Introduction). In all my experiments, only visual stimulation was used in all conditions, and conditions all involved complex motion. As some data suggest that biological motion could be a different kind of motion (see Part 1, Chapter 3, Section 3.1.2), one could argue that more types of visual information are contained in my experimental than my control stimuli. Therefore the activation increases would also have happened if another type of information would have been added to the control conditions instead of biological motion. But the parametric response observed in the pSTS in experiment 3 was *inversely* correlated to differences in objective motion information, and there was no evident difference in “biologicalness” between the objects used. Therefore, again, I think that the interpretation advanced for my activations, based on the role of the STS in processing of biological motion, is fitting the data best.

A recent study suggests that the pSTS is implicated in the learning of a tracking task (Maquet, Schwartz, Passingham, & Frith, 2003). In this study, a group of volunteers were asked perform a pursuit task in which they had to track a disk on a computer screen with a joystick. Unbeknownst to them, the motion trajectory of this disk in the horizontal axis was predictable, but the movements in the vertical axis were unpredictable. Half the participants were sleep-deprived in the following night,

whereas the other half was allowed to sleep normally. When the sleep-deprived participants had recovered from their lack of sleep 3 days later, both groups were retested on the tracking task using movements with trajectories predictable in the learned axis or the other (unlearned) axis. In the group that was allowed to sleep, performance and activation in right STS showed a greater difference between the learned and the new trajectories than in the sleep-deprived participants, and greater connectivity was observed between the cerebellum and the right STS and between the supplementary eye field and the frontal eye field.

As the STS activation observed by Maquet and colleagues falls in the same range as the activations related to observation of animate and biological motion described in this thesis, one could suggest a common process between learning of the pursuit task and expertise in biological motion. One way in which the STS could be involved in the acquisition of complex tracking skills is to extract predictable patterns from observed motion. As the trajectories used in the study by Maquet and colleagues had not been observed before, activation increased after exposition to them. As humans are exposed to biological motion during all their life and pay great attention to it because it is so relevant for their life, one could suggest that the STS has had plenty of time to extract and store its invariant characteristics. These stored characteristics could be activated during exposure to biological motion, and explain activation increases and psychophysical performance when point-light displays are compared to other types of complex but unlearned visual motion. In this view, the STS could be considered as a module specialised in the analysis of complex visual motion and storage of their characteristics, which has been highly trained on biological movements during the life of the individual.

#### **4. Potential confounding factors**

While great care was taken in all experiments to minimise confounding factors or their effects in all experiments, some might of course remain. Here is a review of some of these confounds, their origin and possible solutions for follow-up studies.

In Experiment 1, the cross-correlation level affected speed (Figure Exp 1.3), movement quantity and the extent of the screen covered by the moving objects (Figure Exp 1.7). These factors were taken account of in the design by creating matched Control conditions, in which these variables were almost identical to those in the Interactive conditions, while the chasing behaviour between the objects was destroyed. But despite this fact, interactivity ratings of the control animations (Figure Exp 1.8) showed that some cues leading to attribution of goal-directed behaviour might have remained present (ratings of interactivity increased with the cross-correlation coefficient, even when the interaction between the objects was destroyed), while animacy ratings of the same control animations revealed that these controls did not appear animate (Figure Exp 1.6). This suggests that other influences might also have affected the percept of animacy. Behavioural reports suggest that the number of crossings of the objects could have made the objects appear goal-directed, and crossings were more frequent when the objects moved more (i.e. higher cross-correlation level). In addition, some subjects in the animacy attribution experiment reported that objects in fast-moving animations did not appear animate, which further implicates that the effect of speed should have been controlled better. Nevertheless, on the whole, behavioural and fMRI results are significant and consistent with the original hypotheses, which suggests that, although normalisation of the animations for speed would have been ideal, the role of speed was probably controlled reasonably well.

In Experiment 2, the movements of the objects were very simple, following arcs or straight lines. This was chosen to make the path differences clearer and to simplify subjects' tasks, but might have made the objects appear less animate than those of Experiment 1. One could have used a modified version of the algorithm developed in Experiment 1, but this appeared quite complex during stimulus development and testing, and for the sakes of simplicity and speed it was decided to concentrate on the main factor of Strategy. The movements of the objects nevertheless contained characteristics associated with animacy (see Figure Exp 2.1): one appeared to chase the other, the distance between them was not constant and the movements did not start at the same time, thereby minimising the impression of pushing and pulling, and the chaser stopped its movement when reaching the target, further increasing the impression of goal-directed movements (Opfer, 2002).

In Experiment 3, the choice of object manipulation types could have been different, to maximise differences between mimed and actual movement kinematics, which might have isolated brain regions responding more during observation of mimed or actual actions. Alternatively, the number of object manipulations used could have been increased, to reinforce the correlation between brain kinematics information, behaviour and brain activation. Either possibility would have involved the recording of more movements and the analysis of their kinematics, which was technically very complex and time-consuming. The technical set-up was quite fragile, and an intrinsic design fault in the commercial recording equipment led Daniel Wolpert's group to discontinue development of the recording system. The consequences of the design fault were corrected by manual editing of the kinematic recordings for the experiment described in this thesis, which was again time-consuming but led to acceptable results. Therefore, instead of recording and analysing

many movements, preliminary tests using video recordings of mimed and actual versions of different object manipulation movements were performed, using a number of different objects, including: a cup (lifting, drinking, filling), a jar (lifting), a hammer and nail, different boxes (various lifting and placing movements), paper (folding and tearing), a pen (writing on paper), and the four objects and manipulations used in the experiment. The movements digitised and used in the experiment were chosen following the assumption that they would show the most differences in kinematics between mimed and actual versions and were the easiest to categorise for observers. All these factors led to the design and stimuli used here, which nevertheless yielded interesting data. Other alternatives might be used to avoid the technical difficulties of this digitising equipment: use of an Opto-Track system (set of cameras recording the 3-D position of captors in space, at high speed), but this has the disadvantage that some of the captors might be hidden from the camera in parts of the movements. Or using video recordings, a manipulated object could be edited out from an Actual recording, but adding the object realistically into a Mimed recording appears very difficult. The use of a high-speed video camera together with a movement digitising system recently led to an interesting idea in Daniel Wolpert's laboratory: one could resample the frames of a high-frequency video recording using the digitised movements, thereby systematically altering the movement kinematics of a recorded movement. These methods are currently explored in Daniel Wolpert's laboratory and might lead to fMRI studies in the future.

### **5. How the 3 mentalising regions could interact**

The three brain regions systematically involved in mentalising are the medial prefrontal cortex (mPFC), the superior temporal sulcus (STS) and the temporal pole

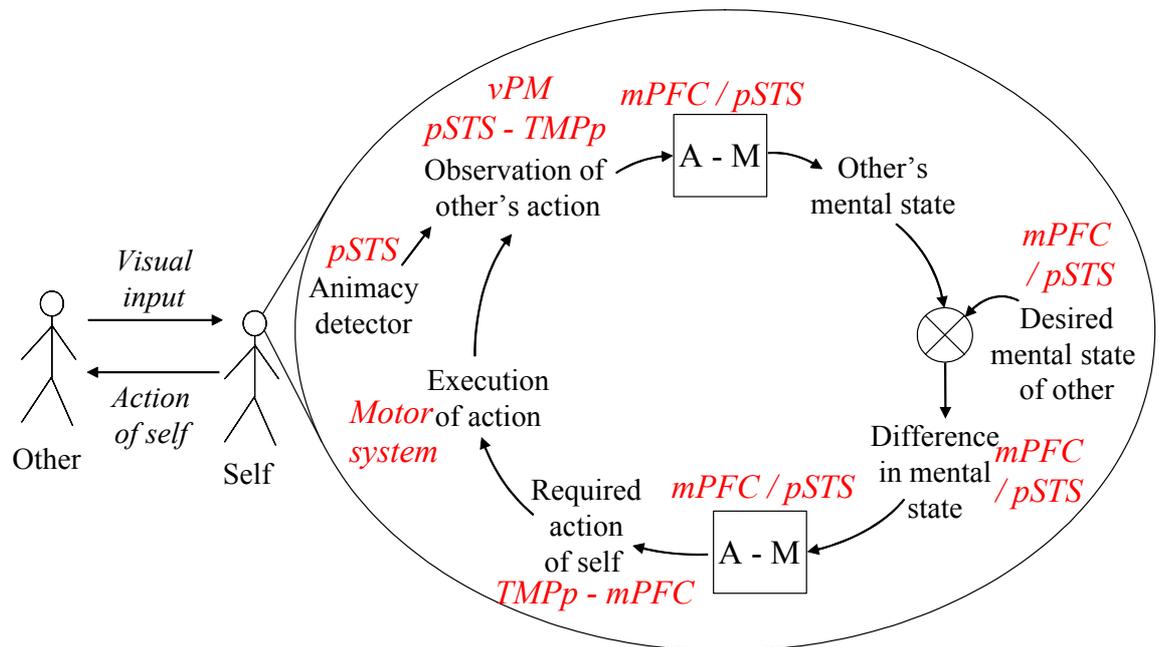
(TMPp). The mPFC is connected both to STS and TPMp (Bachevalier, 1997). Results from the studies performed in this thesis support the role of the pSTS in the detection and analysis of goal-directed movements, which are characteristics of animate agents. This is important for identification of animate entities, whose behaviour can often be explained by mental states, whether they actually have them or not. This is also important because understanding of the movements of animate entities is an essential component of the mentalising ability itself.

I propose that the pSTS, once it has identified a potentially living organism, observes its actions (together with the ventral premotor cortex, vPM) and extracts from its movements cues as to which goal underlies the organism's actions. It then communicates this information to the TMPp (either directly or via the mPFC), which tries to find a matching behavioural script for the successive actions observed. This is useful in predicting what the organism might do next. If information continues to be received from the pSTS, the actual behaviour of the entity can be compared to the behaviour predicted by the script (either in the pSTS, the TMPp or in another brain structure), and the mismatches registered. Both the script information from TMPp and the movement and animacy information from pSTS would then be transmitted to the mPFC and/or the STS<sup>7</sup>, which would try to identify possible mental states explaining the behaviour (see description of the A-M module in Part 1, Chapter 2, Section 2.3). The hypothetical mental state would then be compared to the desired mental state of the agent if there is one. Based on the discrepancy, an action could be selected to change the mental state (again using the A-M model), dependent on the script currently in context (TMPp, see just above). The selected action would then be

---

<sup>7</sup> See the latest studies suggesting that STS and not MPFC is essential for mentalising: neuroimaging (Saxe & Kanwisher, 2003) and lesion studies (Bird, Castelli, Malik, Frith, & Husain, 2004; Samson, Apperly, Chiavarino, & Humphreys, 2004).

planned and executed by the various components of the motor system (Figure Discussion.1)



**Figure Discussion.1. Possible neural correlates of ToM cognitive components.**

The cognitive model was described in the Introduction. See text for details.

## **6. What comes next? Future directions**

Many open questions remain at the end of this thesis, both about perception of animacy and mentalising. Future experiments should increase understanding of the necessary and sufficient characteristics leading to attribution of animacy, through psychophysical experiments using tightly controlled stimuli, varying both movement and aspect of the objects, and assessing the relative importance of movement and aspect in different situations. Also, it is not clear how much of these characteristics are common across different human cultures, or between humans and diverse other animals. While neurophysiological and fMRI data suggest that the STS of rhesus macaques also responds during observation of moving and immobile animals (Logothetis, Guggenberger, Peled, & Pauls, 1999), what about other animals? Do they

differentiate between animate and inanimate entities like we do, is that as evolutionary useful as we might think?

The ontogenic development of animacy attribution is potentially very interesting: are characteristics of animate motion hard-coded in brain structures (such as the pSTS), or do they crystallise after repeated expositions during development? Then would restricted exposure to particular types of animate motion affect animacy attribution, and maybe also brain activity? How about the connections to development of other cognitive processes, such as executive function? Does the cultural definition of what is alive and what is not affect animacy attribution? If yes, does it influence psychophysical performance in experiments with very simple stimuli such as used in this thesis? If computers become more complex, one can wonder if they will one day be considered really animate, beyond the colloquial use of mentalistic terms to describe their behaviour ("My computer really wants to annoy me today: look what it did to my thesis file!"). But this might rely upon a different process than the observation of human actions for example, whose "biologicalness" appears to play a decisive role in activating the action representation system (Castiello, 2003).

Regarding the neural basis of mentalising, future developments could involve better cognitive models of mentalising, from which one could derive tasks that engage only some of the mentalising areas. This could lead to better understanding of the relative contributions of all 3 areas known to be involved in mentalising, and maybe show how other regions involved in social cognition (orbitofrontal cortex, amygdala), complex visual processing (elements of the lateral occipital complex) or action observation (premotor cortex, intraparietal sulcus) can contribute in particular situations. The recent development of more sophisticated neuroimaging data analysis tools and techniques, such as combined EEG and fMRI or connectivity software tools

such as SPM's Dynamic Causal Modelling (DCM), could also reveal the time-course of engagement of mentalising areas and how these regions interact with each other in the normal brain.

## **7. Conclusion**

In conclusion of the thesis, I propose that: 1) detecting animate movements is important for mentalising; 2) as the movement of animate entities is characterised (at least partly) by being goal-directed, brain regions involved in identification of such animate entities could be detecting and analysing goal-directed motion, 3) my results confirm this hypothesis: the pSTS is involved during detection and analysis of goal-directed motion, whether during observation of movements of abstract objects or humans. As this brain region is a part of the mentalising network proposed by some authors, this might explain how it contributes to the whole process of mentalising.

## **Part 6. Appendices**

### **Appendix 1: Behavioural reports of animacy testing in Experiment 1.**

Question asked was: "how alive do the objects look to you? Please rate on a scale from 1-4." It was suggested that subjects might feel a bit lost at the beginning of the experiment, but they should not worry and do the task anyway. They were encouraged to use the whole scale. Animations and experimental protocols used were the same as in the fMRI experiment, but subjects were different. Participants were neither aware of the experimental conditions nor of the purpose or the critical parameters of the experiment. One participant had participated in a very early stage pilot study.

#### **Subject 1**

Looked at the objects chasing each other, the more they chased each other, the more alive they appeared.

#### **Subject 2**

Was a bit lost at the beginning of the experiment; was using chasing at the end of the experiment; quantity of movement was also a clue for animacy: the more they moved, the more animate they appeared.

#### **Subject 3**

Objects appeared more alive when one object appeared to interact with the other; this happened when, after the objects crossed each other, at least one of them came back towards the other object. This was more apparent when the objects were slow, and they therefore appeared more animate.

#### Subject 4

Was looking at chasing at the end of the experiment as a criterion.

#### Subject 5

Did not really know what rule to follow. *NOTE: ratings nevertheless showed a (small) difference between Interactive and Control.*

#### Subject 6

Found the task difficult, was paying attention independently to the movements of both objects at the beginning but very soon focused on their interaction. Used chasing behaviour as critical parameter.

#### Subject 7

Considered that objects appeared more alive when they stayed close to each other, which showed that they cared for each other.

#### Subject 8

Found it difficult, was looking at chasing at the end of the experiment to determine animacy.

#### Subject 9

Found it difficult, used 3 different strategies: first he focussed on playfulness of both objects, then on how much one follows the other, then on both factors together.

#### Subject 10

Found it difficult, none looked particularly animate, used spatial extent as criterion for animacy and other complex mathematical measures. *Note: used only the two middle values of the scale.*

### Subject 11

#### PARTICIPATED IN PILOTS

Found that the faster the objects moved, the easier it was to rate their animacy (thinks his response times are lower with higher speed). Used how much they followed each other and playfulness and jerkiness as criteria for animacy. *Note: he remembered having participated in the pilot experiment. His ratings were very much like the average plot, as were some of the other subjects' ratings, suggesting not much influence of having seen the stimuli before.*

### Subject 12

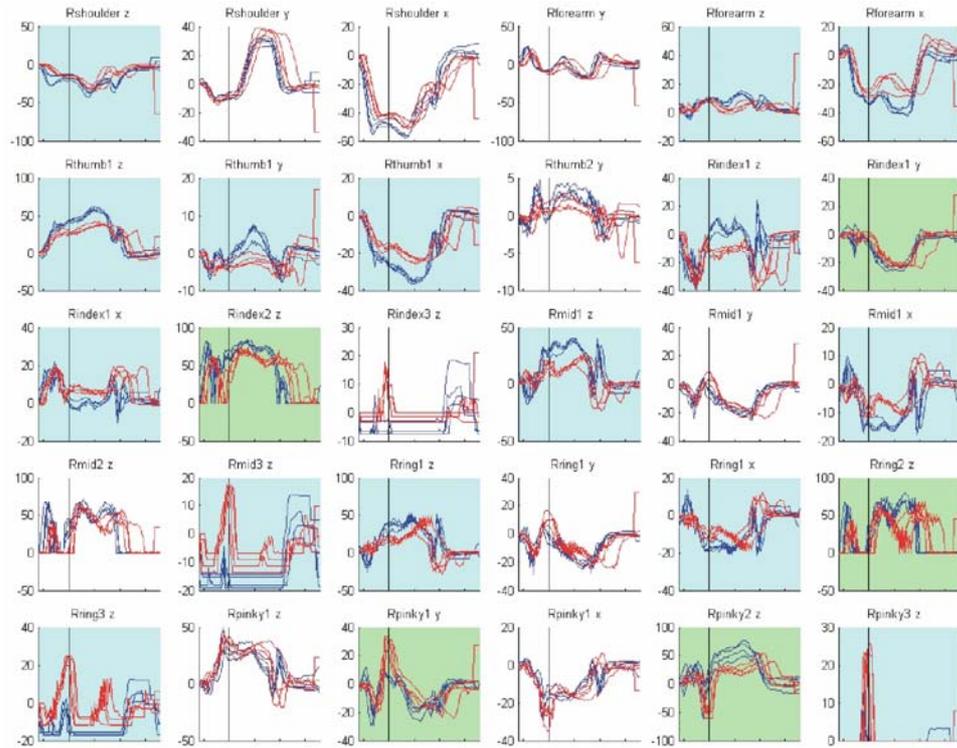
Discounted fast objects as appearing too artificial; otherwise used following and playfulness as index of animacy. The animations made him think of a mother and a child interacting: the child runs around the mother, strays around a bit but always comes back.

### **Appendix 2: kinematic recordings of object manipulations used in Experiment 3**

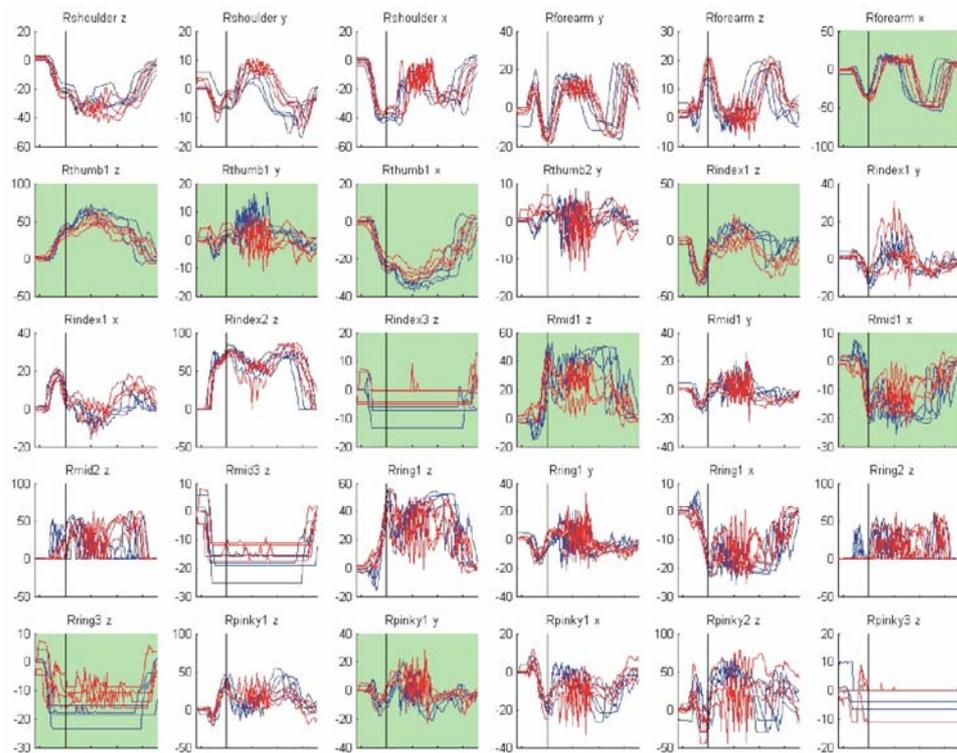
Blue traces represent Actual movement kinematics (4 recordings per manipulation type), red traces represent Mimed movement kinematics (4 recordings per manipulation type). Plots of joint variables with significant differences in kinematics between Actual and Mimed as analysed during the whole movement only have a dark blue background, those with significant differences as analysed during the

manipulation phase of the movement have a green background, and those with significant differences in both forms of analysis have a light blue background.

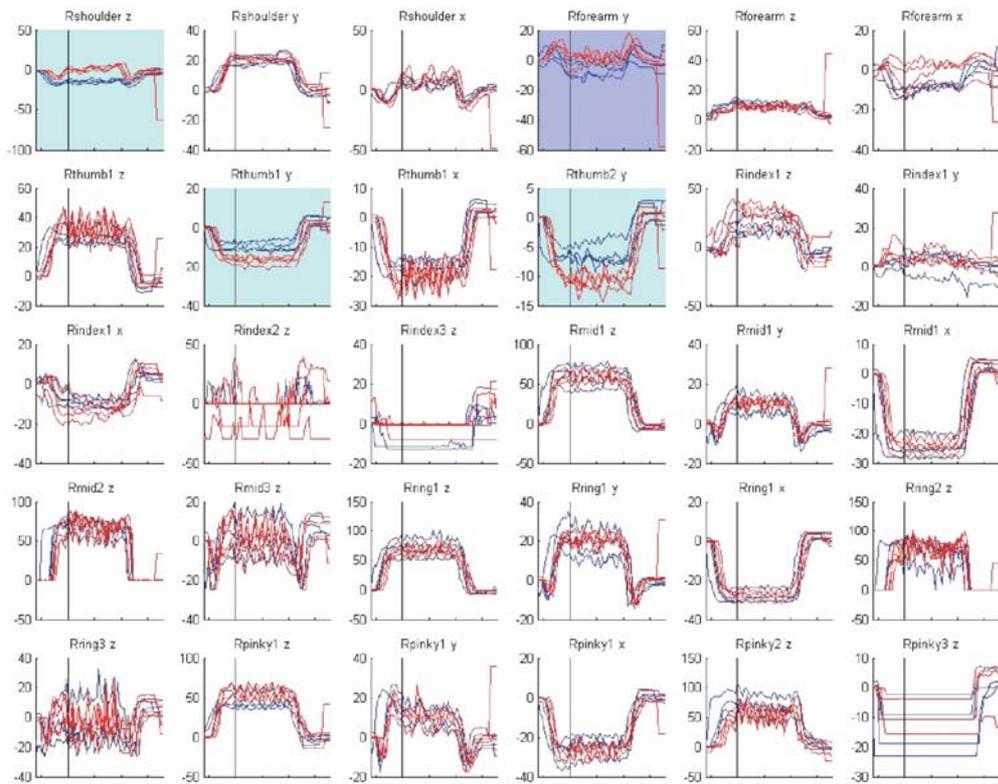
### Book



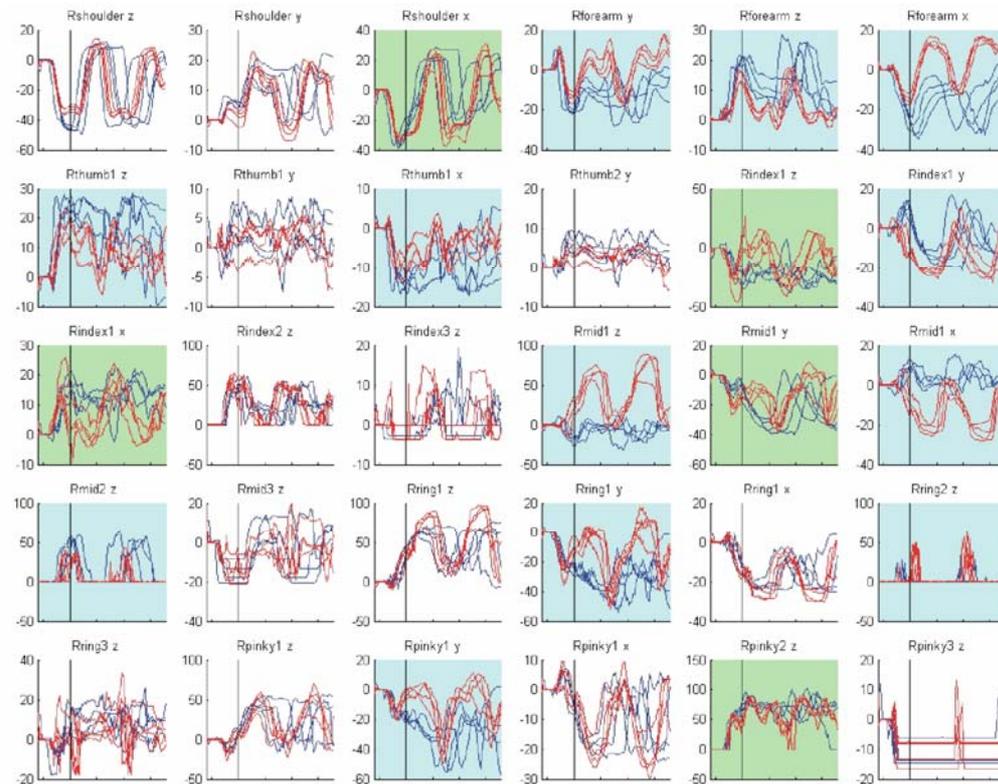
### Bottle



## Phone



## Boxes



## **Part 7. References**

### Reference List

Adolphs, R. (1999). Social cognition and the human brain. *Trends Cogn Sci.*, 3, 469-479.

Adolphs, R. (2003). Cognitive neuroscience of human social behaviour. *Nat.Rev.Neurosci.*, 4, 165-178.

Aggleton, J. P., Burton, M. J., & Passingham, R. E. (1980). Cortical and subcortical afferents to the amygdala of the rhesus monkey (*Macaca mulatta*). *Brain Res.*, 190, 347-368.

Aglioti, S., DeSouza, J. F., & Goodale, M. A. (1995). Size-contrast illusions deceive the eye but not the hand. *Curr.Biol.*, 5, 679-685.

Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: role of the STS region. *Trends Cogn Sci.*, 4, 267-278.

Allman, J. M., Hakeem, A., Erwin, J. M., Nimchinsky, E. A., & Hof, P. (2001). The anterior cingulate cortex - the evolution of an interface between emotion and cognition. *Unity Know.Converg.Nat.Hum.Sci.*, 953, 107-117.

Ashburner, J. & Friston, K. J. (1999). Nonlinear spatial normalisation using basis functions. *Human Brain Mapping*, 7, 254-266.

Ashburner, J., Neelin, P., Collins, D. L., Evans, A. C., & Friston, K. J. (1997). Incorporating prior knowledge into image registration. *Neuroimage.*, 6, 344-352.

Aziz-Zadeh, L., Iacoboni, M., Zaidel, E., Wilson, S., & Mazziotta, J. (2004).  
Left hemisphere motor facilitation in response to manual action sounds.  
*Eur.J.Neurosci.*, *19*, 2609-2612.

Aziz-Zadeh, L., Maeda, F., Zaidel, E., Mazziotta, J., & Iacoboni, M. (2002).  
Lateralization in motor facilitation during action observation: a TMS study. *Exp.Brain  
Res.*, *144*, 127-131.

Bachevalier, J., Meunier, M., Lu, M. X., & Ungerleider, L. G. (1997).  
Thalamic and temporal cortex input to medial prefrontal cortex in rhesus monkeys.  
*Exp.Brain Res.*, *115*, 430-444.

Barbas, H. (1988). Anatomic organization of basoventral and mediodorsal  
visual recipient prefrontal regions in the rhesus monkey. *J.Comp Neurol.*, *276*, 313-  
342.

Baron-Cohen, S. (1995). *Mindblindness: an essay on autism and theory of  
mind*. Cambridge, MA: MIT Press.

Baron-Cohen, S., Jolliffe, T., Mortimore, C., & Robertson, M. (1997). Another  
advanced test of theory of mind: evidence from very high functioning adults with  
autism or Asperger syndrome. *J.Child Psychol.Psychiat.*, *38*, 813-822.

Baron-Cohen, S., Leslie, A., & Frith, U. (1985). Does the autistic child have a  
theory of mind? *Cognition*, *21*, 37-46.

Barresi, J. & Moore, C. (1996). Intentional relations and social understanding.  
*Behavioural and Brain Sciences*, *19*, 107-122.

Bassili, J. N. (1976). Temporal and spatial contingencies in the perception of social events. *Journal of personality and social psychology*, 33, 680-685.

Beauchamp, M. S., Lee, K. E., Argall, B. D., & Martin, A. (2004). Integration of auditory and visual information about objects in superior temporal sulcus. *Neuron*, 41, 809-823.

Beauchamp, M. S., Lee, K. E., Haxby, J. V., & Martin, A. (2002). Parallel visual motion processing streams for manipulable objects and human movements. *Neuron*, 34, 149-159.

Beauchamp, M. S., Lee, K. E., Haxby, J. V., & Martin, A. (2003). fMRI responses to video and point-light displays of moving humans and manipulable objects. *J.Cogn Neurosci.*, 15, 991-1001.

Bekkering, H., Wohlschlaeger, A., & Gattis, M. (2002). Imitation of gestures in children is goal-directed. *Quarterly Journal of Experimental Psychology*, 53, 153-164.

Belin, P., Zatorre, R. J., Lafaille, P., Ahad, P., & Pike, B. (2000). Voice-selective areas in human auditory cortex. *Nature*, 403, 309-312.

Bennett, J. (1978). Some remarks about concepts. *Behavioural and Brain Sciences*, 1, 557-560.

Berthoz, S., Armony, J. L., Blair, J. L. R., & Dolan, R. J. (2002). An fMRI study of intentional and unintentional (embarassing) violations of social norms. *Brain*, 125, 1696-1708.

Bird, C., Castelli, F., Malik, O., Frith, U., & Husain, M. (2004). The impact of extensive medial frontal lobe damage on 'Theory of Mind' and cognition. *Brain*, *127*, 914-928.

Blakemore, S. J., Boyer, P., Pachot-Clouard, M., Meltzoff, A., Segebarth, C., & Decety, J. (2003). The detection of contingency and animacy from simple animations in the human brain. *Cereb.Cortex*, *13*, 837-844.

Blakemore, S. J. & Decety, J. (2001). From the perception of action to the understanding of intention. *Nat.Rev Neurosci.*, *2*, 561-567.

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

Blakemore, S. J., Frith, C. D., & Wolpert, D. M. (1999). Spatio-temporal prediction modulates the perception of self-produced stimuli. *J.Cogn Neurosci.*, *11*, 551-559.

Bloch, F., Hansen, W. W., & Packard, M. (1946). The nuclear induction experiment. *Phys.Rev.*, *70*, 474-485.

Blythe, P. W., Todd, P. M., & Miller, G. F. (1999). How motion reveals intention: categorising social interactions. In G.Gigerenzer & P. M. Todd (Eds.), *Simple heuristics that make us smart* (pp. 257-285). Oxford: Oxford University Press.

Bonda, E., Petrides, M., Ostry, D., & Evans, A. (1996). Specific involvement of human parietal systems and the amygdala in the perception of biological motion. *J.Neurosci.*, *16*, 3737-3744.

Boussaoud, D., Ungerleider, L. G., & Desimone, R. (1990). Pathways for motion analysis: cortical connections of the medial superior temporal and fundus of the superior temporal visual areas in the macaque. *J.Comp Neurol.*, 296, 462-495.

Brass, M., Bekkering, H., & Prinz, W. (2001). Movement observation affects movement execution in a simple response task. *Acta Psychologica*, 106, 3-22.

Brothers, L. (1990). The social brain: a project for integrating primate behaviour and neurophysiology in a new domain. *Concepts in Neuroscience*, 1, 27-151.

Brunet, E., Sarfati, Y., Hardy-Bayle, M.-C., & Decety, J. (2000). A PET investigation of the attribution of mental intentions with a nonverbal task. *Neuroimage.*, 11, 157-166.

Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., Seitz, R. J., Zilles, K., Rizzolatti, G., & Freund, H. J. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *Eur.J.Neurosci.*, 13, 400-404.

Bush, G., Luu, P., & Posner, M. I. (2000). Cognitive and emotional influences in the anterior cingulate cortex. *Trends Cogn Sci.*, 4, 215-222.

Campbell, R., MacSweeney, M., Surguladze, S., Calvert, G., McGuire, P., Suckling, J., Brammer, M. J., & David, A. S. (2001). Cortical substrates for the perception of face actions: an fMRI study of the specificity of activation for seen speech and for meaningless lower-face acts (gurning). *Brain Res.Cogn Brain Res.*, 12, 233-243.

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

Castiello, U. (2003). Understanding other people's actions: intention and attention. *J.Exp.Psychol.Hum.Percept.Perform.*, *29*, 416-430.

Cavada, C. & Goldman-Rakic, P. S. (1989). Posterior parietal cortex in rhesus monkey: II. Evidence for segregated corticocortical networks linking sensory and limbic areas with the frontal lobe. *J.Comp Neurol.*, *287*, 422-445.

Chao, L. L., Haxby, J. V., & Martin, A. (1999). Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. *Nat.Neurosci.*, *2*, 913-919.

Chao, L. L., Martin, A., & Haxby, J. V. (1999). Are face-responsive regions selective only for faces? *Neuroreport*, *10*, 2945-2950.

Cohen, L., Dehaene, S., Naccache, L., Lehericy, S., Dehaene-Lambertz, G., Henaff, M. A., & Michel, F. (2000). The visual word form area: spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain*, *123 ( Pt 2)*, 291-307.

Cohen, L., Lehericy, S., Chochon, F., Lemer, C., Rivaud, S., & Dehaene, S. (2002). Language-specific tuning of visual cortex? Functional properties of the Visual Word Form Area. *Brain*, *125*, 1054-1069.

Conover, W. J. & Iman, R. L. (1981). Rank transformation as a bridge between parametric and nonparametric statistics. *The American Statistician*, *35*, 124-129.

Corbetta, M., Kincade, J. M., Ollinger, J. M., McAvoy, M. P., & Shulman, G. L. (2000). Voluntary orienting is dissociated from target detection in human posterior parietal cortex. *Nat.Neurosci.*, *3*, 292-297.

Csibra, G. (2003). Teleological and referential understanding of action in infancy. *Philos.Trans.R.Soc.Lond B Biol.Sci.*, *358*, 447-458.

Csibra, G., Gergely, G., Biro, S., Koos, O., & Brockbank, M. (1999). Goal attribution without agency cues: the perception of 'pure reason' in infancy. *Cognition*, *72*, 237-267.

Damasio, H., Grabowski, T. J., Tranel, D., Hichwa, R. D., & Damasio, A. R. (1996). A neural basis for lexical retrieval. *Nature*, *380*, 499-505.

Deichmann, R., Good, C. D., Josephs, O., Ashburner, J., & Turner, R. (2000). Optimization of 3-D MP-RAGE sequences for structural brain imaging. *Neuroimage.*, *12*, 112-127.

Dennett, D. (1978). Beliefs about beliefs. *Behavioural and Brain Sciences*, *1*, 568-570.

Desimone, R. (1991). Face-selective cells in the temporal cortex of the monkey. *J.Cogn Neurosci.*, *3*, 1-8.

Desmond, J. E. & Glover, G. H. (2002). Estimating sample size in functional MRI (fMRI) neuroimaging studies: statistical power analyses. *J Neurosci Methods*, *118*, 115-128.

Devinsky, O., Morrell, M. J., & Vogt, B. A. (1995). Contribution of anterior cingulate cortex to behaviour. *Brain*, *118*, 279-306.

- di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: a neurophysiological study. *Exp. Brain Res.*, *91*, 176-180.
- Dickinson, A. R. & Balleine, B. W. (2000). Causal cognition and goal-directed action. In C. Heyes & L. Huber (Eds.), *Evolution of Cognition* (pp. 185-204). Cambridge, MA: MIT Press.
- Dittrich, W. H. & Lea, S. E. (1994). Visual perception of intentional motion. *Perception*, *23*, 253-268.
- Downar, J., Crawley, A. P., Mikulis, D. J., & Davis, K. D. (2000). A multimodal cortical network for the detection of changes in the sensory environment. *Nat. Neurosci.*, *3*, 277-283.
- Downar, J., Crawley, A. P., Mikulis, D. J., & Davis, K. D. (2001). The effect of task-relevance on the cortical response to changes in visual and auditory stimuli: an event-related fMRI study. *Neuroimage.*, *14*, 1256-1267.
- Downar, J., Crawley, A. P., Mikulis, D. J., & Davis, K. D. (2002). A cortical network sensitive to stimulus salience in a neutral behavioural context across multiple sensory modalities. *J Neurophysiol.*, *87*, 615-620.
- Downing, P. E., Jiang, Y., Shuman, M., & Kanwisher, N. (2001). A cortical area selective for visual processing of the human body. *Science*, *293*, 2470-2473.
- Duelli, R. & Kuschinsky, W. (1993). Changes in brain capillary diameter during hypocapnia and hypercapnia. *J Cereb Blood Flow Metab*, *13*, 1025-1028.
- Duvernoy, H. M. (1999). *The Human Brain*. Vienna: Springer Verlag.

- Eifuku, S., De Souza, W. C., Tamura, R., Nishijo, H., & Ono, T. (2004). Neuronal correlates of face identification in the monkey anterior temporal cortical areas. *J Neurophysiol.*, *91*, 358-371.
- Ellis, R. R., Flanagan, J. R., & Lederman, S. J. (1999). The influence of visual illusions on grasp position. *Exp. Brain Res.*, *125*, 109-114.
- Fadiga, L., Fogassi, L., Pavesi, G., & Rizzolatti, G. (1995). Motor facilitation during action observation: a magnetic stimulation study. *J. Neurophysiol.*, *73*, 2608-2611.
- Fagg, A. H. & Arbib, M. A. (1998). Modeling parietal-premotor interactions in primate control of grasping. *Neural Netw.*, *11*, 1277-1303.
- Farrer, C., Franck, N., Georgieff, N., Frith, C. D., Decety, J., & Jeannerod, M. (2003). Modulating the experience of agency: a positron emission tomography study. *Neuroimage.*, *18*, 324-333.
- Farrer, C. & Frith, C. D. (2002). Experiencing oneself vs another person as being the cause of an action: the neural correlates of the experience of agency. *Neuroimage.*, *15*, 596-603.
- Felleman, D. J. & Van Essen, D. C. (1991). Distributed hierarchical processing in primate cerebral cortex. *Cereb. Cortex*, *1*, 1-47.
- Ferrari, P. F., Gallese, V., Rizzolatti, G., & Fogassi, L. (2003). Mirror neurons responding to the observation of ingestive and communicative mouth actions in the monkey ventral premotor cortex. *Eur. J. Neurosci.*, *17*, 1703-1714.

Ferstl, E. C. & von Cramon, D. Y. (2002). What does the frontomedian cortex contribute to language processing: coherence or theory of mind? *Neuroimage.*, 17, 1599-1612.

Flanagan, J. R. & Beltzner, M. A. (2000). Independence of perceptual and sensorimotor predictions in the size-weight illusion. *Nat.Neurosci.*, 3, 737-741.

Fletcher, P. C., Happe, F., Frith, U., Baker, S. C., Dolan, R. J., Frackowiak, R. S., & Frith, C. D. (1995). Other minds in the brain: a functional imaging study of 'theory of mind' in story comprehension. *Cognition*, 44, 283-296.

Fogassi, L., Gallese, V., Fadiga, L., & Rizzolatti, G. (1998). Neurons responding to the sight of goal-directed hand/arm actions in the parietal area PF7b of the macaque monkey. *Soc.Neurosci.Abs.* 24, 257.5.

Ref Type: Abstract

Friston, K. J., Ashburner, J., Frith, C. D., Poline, J. B., Heather, J. D., & Frackowiak, R. S. (1995). Spatial registration and normalisation of images. *Hum.Brain Mapp.*, 2, 165-189.

Friston, K. J., Fletcher, P. C., Josephs, O., Holmes, A. P., Rugg, M. D., & Turner, R. (1998). Event-related fMRI: Characterizing differential responses. *Neuroimage.*, 7, 30-40.

Friston, K. J., Josephs, O., Rees, G., & Turner, R. (1998). Nonlinear event-related responses in fMRI. *Mag Res Med*, 39, 41-52.

Frith, C. D. & Frith, U. (1999). Interacting minds--a biological basis. *Science*, 286, 1692-1695.

- Frith, U. (2001). Mind blindness and the brain in autism. *Neuron*, 32, 969-979.
- Frith, U. & Frith, C. D. (2003). Development and neurophysiology of mentalizing. *Philos.Trans.R.Soc.Lond B Biol.Sci.*, 358, 459-473.
- Gallagher, H. L. & Frith, C. D. (2003). Functional imaging of 'theory of mind'. *Trends Cogn Sci.*, 7, 77-83.
- Gallagher, H. L., Happe, F., Brunswick, N., Fletcher, P. C., Frith, U., & Frith, C. D. (2000). Reading the mind in cartoons and stories: an fMRI study of 'theory of mind' in verbal and nonverbal tasks. *Neuropsychologia*, 38, 11-21.
- Gallagher, H. L., Jack, A. I., Roepstorff, A., & Frith, C. D. (2002). Imaging the intentional stance in a competitive game. *Neuroimage.*, 16, 814-821.
- Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, 119 ( Pt 2), 593-609.
- Gallese, V., Fogassi, L., Fadiga, L., & Rizzolatti, G. (2002). Action representation in the inferior parietal lobule. In W.Prinz & B. Hommel (Eds.), *Attention and Performance* (XIX ed., Oxford: Oxford University Press.
- Gallese, V. & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends Cogn Sci.*, 2, 493-501.
- Gelman, R., Durgin, F., & Kaufman, L. (1995). Distinguishing between animates and inanimates: not by motion alone. In S.Sperber, D. Premack, & A. Premack (Eds.), *Causal cogniition: a multi-disciplinary debate* ( Cambridge: Oxford University Press.

Gergely, G., Bekkering, H., & Kiraly, I. (2002). Rational imitation in preverbal infants. *Nature*, *415*, 755.

Gergely, G. & Csibra, G. (1997). Teleological reasoning in infancy: the infant's naive theory of rational action. A reply to Premack and Premack. *Cognition*, *63*, 227-233.

Gergely, G. & Csibra, G. (2003). Teleological reasoning in infancy: the naive theory of rational action. *Trends Cogn Sci.*, *7*, 287-292.

Gergely, G., Nadasdy, Z., Csibra, G., & Biro, S. (1995). Taking the intentional stance at 12 months of age. *Cognition*, *56*, 165-193.

Giese, M. A. & Poggio, T. (2003). Neural mechanisms for the recognition of biological movements. *Nat.Rev.Neurosci.*, *4*, 179-192.

Godschalk, M., Lemon, R. N., Kuypers, H. G., & Ronday, H. K. (1984). Cortical afferents and efferents of monkey postarcuate area: an anatomical and electrophysiological study. *Exp.Brain Res.*, *56*, 410-424.

Goel, V., Grafman, J. N. S., & Hallett, M. (1995). Modelling other minds. *Neuroreport*, *6*, 1741-1746.

Goldman, A. (1989). Interpretation psychologized. *Mind Lang*, *4*, 161-185.

Goldman, A. (1993). Consciousness, folk psychology and cognitive science. *Consciousness and Cognition*, *2*, 364-382.

Goodale, M. A., Jakobson, L. S., & Keillor, J. M. (1994). Differences in the visual control of pantomimed and natural grasping movements. *Neuropsychologia*, 32, 1159-1178.

Gopnik, A. (1993). How we know our minds: the illusion of first-person knowledge of intentionality. *Behav. Brain Res.*, 16, 1-14.

Gopnik, A. & Astington, J. W. (1988). Children's understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction. *Child Dev.*, 59, 26-37.

Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgement. *Science*, 293, 2105-2108.

Greenwald, A. G. (1970). Sensory feedback mechanisms in performance control: with special reference to the ideomotor principle. *Psychol. Rev.*, 77, 73-99.

Gregory, C., Lough, S., Stone, V. E., Erzinclioglu, S., Martin, L., Baron-Cohen, S., & Hodges, J. R. (2002). Theory of mind in patients with frontal variant frontotemporal dementia and Alzheimer's disease: theoretical and practical implications. *Brain*, 125, 752-764.

Grezes, J., Armony, J. L., Rowe, J., & Passingham, R. E. (2003). Activations related to "mirror" and "canonical" neurones in the human brain: an fMRI study. *Neuroimage.*, 18, 928-937.

Grezes, J. & Costes, N. (1998). Top-down effect of strategy on the perception of biological motion: a PET investigation. *Cognitive Neuropsychology*, 15, 582.

Grezes, J. & Decety, J. (2001). Functional anatomy of execution, mental simulation, observation, and verb generation of actions: a meta-analysis. *Hum. Brain Mapp.*, *12*, 1-19.

Grezes, J., Fonlupt, P., Bertenthal, B., Delon-Martin, C., Segebarth, C., & Decety, J. (2001). Does perception of biological motion rely on specific brain regions? *Neuroimage.*, *13*, 775-785.

Grezes, J., Frith, C. D., & Passingham, R. E. (2004). Inferring false beliefs from the actions of oneself and others: an fMRI study. *Neuroimage.*, *21*, 744-750.

Grossman, E., Donnelly, M., Price, R., Pickens, D., Morgan, V., Neighbor, G., & Blake, R. (2000). Brain areas involved in perception of biological motion. *J. Cogn Neurosci.*, *12*, 711-720.

Haffenden, A. M. & Goodale, M. A. (1998). The effect of pictorial illusion on prehension and perception. *J. Cogn Neurosci.*, *10*, 122-136.

Harman, G. (1978). Studying the chimpanzee's theory of mind. *Behavioural and Brain Sciences*, *1*, 591.

Haruno, M., Wolpert, D. M., & Kawato, M. (2001). Mosaic model for sensorimotor learning and control. *Neural Comput.*, *13*, 2201-2220.

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2002). Human neural systems for face recognition and social communication. *Biol. Psychiatry*, *51*, 59-67.

Haxby, J. V., Horwitz, B., Ungerleider, L. G., Maisog, J. M., Pietrini, P., & Grady, C. L. (1994). The functional organization of human extrastriate cortex: a PET-rCBF study of selective attention to faces and locations. *J. Neurosci.*, *14*, 6336-6353.

Heider, F. & Simmel, M. (1944). An experimental study of apparent behaviour. *Am.J.Psychol.*, 57, 243-259.

Hoffman, E. A. & Haxby, J. V. (2000). Distinct representations of gaze and identity in the distributed human neural system for face perception. *Nat.Neurosci.*, 3, 80-84.

Howard, R. J., Brammer, M., Wright, I., Woodruff, P. W., Bullmore, E. T., & Zeki, S. (1996). A direct demonstration of functional specialization within motion-related visual and auditory cortex of the human brain. *Curr.Biol.*, 6, 1015-1019.

Hyder, F., Rothman, D. L., Mason, G. F., Rangarajan, A., Behar, K. L., & Shulman, R. G. (1997). Oxidative glucose metabolism in rat brain during single forepaw stimulation: a spatially localised <sup>1</sup>H[<sup>13</sup>C] nuclear magnetic resonance imaging study. *J Cereb Blood Flow Metab*, 17, 1040-1047.

Iacoboni, M., Koski, L. M., Brass, M., Bekkering, H., Woods, R. P., Dubeau, M. C., Mazziotta, J. C., & Rizzolatti, G. (2001). Reafferent copies of imitated actions in the right superior temporal cortex. *Proc.Natl.Acad.Sci.U.S.A*, 98, 13995-13999.

Iacoboni, M., Lieberman, M. D., Knowlton, B. J., Molnar-Szakacs, I., Moritz, M., Throop, C. J., & Fiske, A. P. (2004). Watching social interactions produces dorsomedial prefrontal and medial parietal BOLD fMRI signal increases compared to a resting baseline. *Neuroimage.*, 21, 1167-1173.

Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazziotta, J. C., & Rizzolatti, G. (1999). Cortical mechanisms of human imitation. *Science*, 286, 2526-2528.

James, T. W., Culham, J., Humphrey, G. K., Milner, A. D., & Goodale, M. A. (2003). Ventral occipital lesions impair object recognition but not object-directed grasping: an fMRI study. *Brain*, *126*, 2463-2475.

James, W. (1890). *The principles of psychology*. (Paperback, 1983 ed.) Harvard University Press.

Jellema, T., Baker, C. I., Wicker, B., & Perrett, D. I. (2000). Neural representation for the perception of the intentionality of actions. *Brain Cogn*, *44*, 280-302.

Jellema, T. & Perrett, D. I. (2003a). Cells in monkey STS responsive to articulated body motions and consequent static posture: a case of implied motion? *Neuropsychologia*, *41*, 1728-1737.

Jellema, T. & Perrett, D. I. (2003b). Perceptual history influences neural responses to face and body postures. *J.Cogn Neurosci.*, *15*, 961-971.

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception and Psychophysics*, *14*, 201-211.

Johnson, S. C. (2003). Detecting agents. *Philos.Trans.R.Soc.Lond B Biol.Sci.*, *358*, 549-559.

Josephs, O. & Henson, R. N. A. (1999). Event-related functional magnetic resonance imaging: modelling, inference and optimization. *Proc.R.Soc.Lond B Biol.Sci.*, *354*, 1228.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J.Neurosci.*, *17*, 4302-4311.

Karnath, H.-O. (2001). New insights into the functions of the superior temporal cortex. *Nat.Rev Neurosci.*, *2*, 568-576.

Kerzel, D., Hecht, H., & Kim, N.-G. (1999). Image velocity, not Tau, explains arrival-time judgments from global optical flow. *J.Exp.Psychol.Hum.Percept.Perform.*, *25*, 1540-1555.

Kilner, J. M., Paulignan, Y., & Blakemore, S. J. (2003). An interference effect of observed biological movement on action. *Curr.Biol.*, *13*, 522-525.

Kohler, E., Keysers, C., Umiltà, M. A., Fogassi, L., Gallese, V., & Rizzolatti, G. (2002). Hearing sounds, understanding actions: action representation in mirror neurons. *Science*, *297*, 846-848.

Koski, L., Iacoboni, M., Dubeau, M. C., Woods, R. P., & Mazziotta, J. C. (2003). Modulation of cortical activity during different imitative behaviors. *J.Neurophysiol.*, *89*, 460-471.

Koski, L., Wohlschläger, A., Bekkering, H., Woods, R. P., Dubeau, M. C., Mazziotta, J. C., & Iacoboni, M. (2002). Modulation of motor and premotor activity during imitation of target-directed actions. *Cereb.Cortex*, *12*, 847-855.

Kronbichler, M., Hutzler, F., Wimmer, H., Mair, A., Staffen, W., & Ladurner, G. (2004). The visual word form area and the frequency with which words are encountered: evidence from a parametric fMRI study. *Neuroimage*, *21*, 946-953.

Kurata, K. (1991). Corticocortical inputs to the dorsal and ventral aspects of the premotor cortex of macaque monkeys. *Neurosci.Res.*, *12*, 263-280.

Kwong, K. K., Belliveau, J. W., Chesler, D. A., Goldberg, I. E., Weiskopf, R. M., Poncelet, B. P., Kennedy, D. N., Hoppel, B. E., Cohen, M. S., Turner, R., Cheng, H.-M., Brady, T. J., & Rosen, B. R. (1992). Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *Proc.Natl.Acad.Sci.U.S.A*, *89*, 5675-5679.

Lacquaniti, F., Terzuolo, C., & Viviani, P. (1983). The law relating kinematic and figural aspects of drawing movements. *Acta Psychologica*, *54*, 115-130.

Lauterbur, P. C. (1973). Image formation by induced local interactions: examples employing the nuclear magnetic resonance. *Nature*, *242*, 190-191.

Leslie, A. & Thaiss, L. (1992). Domain specificity in conceptual development: Neuropsychological evidence from autism. *Cognition*, *43*, 225-251.

Logothetis, N. K., Guggenberger, H., Peled, S., & Pauls, J. (1999). Functional imaging of the monkey brain. *Nat.Neurosci.*, *2*, 555-562.

Logothetis, N. K. & Wandell, B. A. (2004). Interpreting the BOLD signal. *Annu Rev Physiol*, *66*, 735-769.

Lotze, R. H. (1852). *Medizinische Psychologie oder Physiologie der Seele*. (reprint 1966 ed.) Amsterdam: E.J. Bonset.

Luppino, G., Murata, A., Govoni, P., & Matelli, M. (1999). Largely segregated parietofrontal connections linking rostral intraparietal cortex (areas AIP and VIP) and the ventral premotor cortex (areas F5 and F4). *Exp.Brain Res.*, *128*, 181-187.

Mandeville, J. B., Marota, J. J., Ayata, C., Zaharchuk, G., Moskowitz, M. A., Rosen, B. R., & Weiskopf, R. M. (1999). Evidence of a cerebrovascular postarteriole windkessel with delayed compliance. *J Cereb Blood Flow Metab*, *19*, 679-689.

Mansfield, P. (1977). Multi-planar image formation using NMR spin echoes. *J.Physiol*, *10*, 349-352.

Maquet, P., Schwartz, S., Passingham, R., & Frith, C. (2003). Sleep-related consolidation of a visuomotor skill: brain mechanisms as assessed by functional magnetic resonance imaging. *J.Neurosci.*, *23*, 1432-1440.

Marois, R., Leung, H. C., & Gore, J. C. (2000). A stimulus-driven approach to object identity and location processing in the human brain. *Neuron*, *25*, 717-728.

Martin, A., Ungerleider, L. G., & Haxby, J. V. (2000). Category specificity and the brain: the sensory/motor model of semantic representations of objects. In M.S.Gazzaniga (Ed.), *The new cognitive neurosciences* (second edition ed., pp. 1023-1036). Cambridge, MA: MIT Press.

Martin, A., Wiggs, C. L., Ungerleider, L. G., & Haxby, J. V. (1996). Neural correlates of category-specific knowledge. *Nature*, *379*, 649-652.

Mason, C. R., Gomez, J. E., & Ebner, T. J. (2001). Hand synergies during reach-to-grasp. *J Neurophysiol.*, *86*, 2896-2910.

McLeod, P., Dittrich, W. H., Driver, J., Perrett, D., & Zihl, J. (1996). Preserved and impaired detection of structure from motion by a "motion-blind" patient. *Visual Cognition*, *3*, 363-391.

Meltzoff, A. (1995). Understanding the intentions of others: re-enactment of intended acts by 18-month-old children. *Developmental Psychology*, 31, 838-850.

Meltzoff, A. & Moore, C. (1977). Imitation of facial and manual gestures by human neonates. *Science*, 198, 75-78.

Meltzoff, A. & Moore, M. K. (1997). Explaining facial imitation: a theoretical model. *Early Dev. Parenting*, 6, 179-192.

Michotte, A. (1946). *La perception de la causalite, etudes psychologiques VI*. Louvain, France: Institut superieur de philosophie.

Milner, A. D. & Goodale, M. A. (1995). *The visual brain in action*. Oxford: Oxford University Press.

Moran, M. A., Mufson, E. J., & Mesulam, M. M. (1985). Neural inputs into the temporopolar cortex of the rhesus monkey. *Journal of Computational Neurology*, 256, 88-103.

Murata, A., Fadiga, L., Fogassi, L., Gallese, V., Raos, V., & Rizzolatti, G. (1997). Object representation in the ventral premotor cortex (area F5) of the monkey. *J. Neurophysiol.*, 78, 2226-2230.

Murata, A., Gallese, V., Luppino, G., Kaseda, M., & Sakata, H. (2000). Selectivity for the shape, size, and orientation of objects for grasping in neurons of monkey parietal area AIP. *J. Neurophysiol.*, 83, 2580-2601.

Narumoto, J., Okada, T., Sadato, N., Fukui, K., & Yonekura, Y. (2001). Attention to emotion modulates fMRI activity in human right superior temporal sulcus. *Brain Res. Cogn Brain Res.*, 12, 225-231.

Neri, P., Morrone, M. C., & Burr, D. C. (1998). Seeing biological motion. *Nature*, 395, 894-896.

Nielsen, J. M. (1958). *Memory and amnesia*. Los Angeles: San Lucas Press.

Nimchinsky, E. A., Gilissen, E., Allman, J. M., Perl, D. P., Erwin, J. M., & Hof, P. (1999). A neuronal morphologic type unique to humans and great apes. *Proc.Natl.Acad.Sci.U.S.A*, 96, 5268-5273.

O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K. J., & Dolan, R. J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science*, 304, 452-454.

Ogawa, S. & Lee, T.-M. (1990). Magnetic resonance imaging of blood vessels at high fields: in vivo and in vitro measurements and image simulation. *Mag Res Med*, 16, 9-18.

Ogawa, S., Lee, T.-M., Nayak, A. S., & Glynn, P. (1990). Oxygenation-sensitive contrast in magnetic resonance image of rodent brain at high magnetic fields. *Mag Res Med*, 14, 68-78.

Ogawa, S., Tank, D. W., Menon, R., Ellerman, J. M., Kim, S.-G., Merkle, H., & Ugurbil, K. (1992). Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proc.Natl.Acad.Sci.U.S.A*, 89, 5951-5955.

Opfer, J. E. (2002). Identifying living and sentient kinds from dynamic information: the case of goal-directed versus aimless autonomous movement in conceptual change. *Cognition*, 86, 97-122.

Pauling, L. & Coryall, C. D. (1936). The magnetic properties and structure of hemoglobin, oxyhemoglobin and carbonoxyhemoglobin. *Proc.Natl.Acad.Sci.U.S.A.*, 22, 210-216.

Pelphrey, K. A., Mitchell, T. V., McKeown, M. J., Goldstein, J., Allison, T., & McCarthy, G. (2003). Brain Activity Evoked by the Perception of Human Walking: Controlling for Meaningful Coherent Motion. *Journal of Neuroscience*, 23, 6819-6825.

Perani, D., Cappa, S. F., Bettinardi, V., Bressi, S., Gorno-Tempini, M., Matarrese, M., & Fazio, F. (1995). Different neural systems for the recognition of animals and man-made tools. *Neuroreport*, 6, 1637-1641.

Perani, D., Fazio, F., Borghese, N. A., Tettamanti, M., Ferrari, S., Decety, J., & Gilardi, M. C. (2001). Different brain correlates for watching real and virtual hand actions. *Neuroimage.*, 14, 749-758.

Perner, J. & Lang, B. (1999). Development of theory of mind and executive control. *Trends Cogn Sci.*, 3, 337-343.

Perrett, D. I., Harries, M. H., Bevan, R., Thomas, S., Benson, P. J., Mistlin, A. J., Chitty, A. J., Hietanen, J. K., & Ortega, J. E. (1989). Frameworks of analysis for the neural representation of animate objects and actions. *J.Exp.Biol.* 146, 87-113.

Ref Type: Abstract

Perrett, D. I., Harries, M. H., Mistlin, A. J., Hietanen, J. K., Benson, P. J., Bevan, R., Thomas, S., Oram, M. W., Ortega, J. E., & Brierly, K. (1990). Social signals analyzed at the single cell level: Someone is looking at me, something touched me, something moved. *Int.J.Comp.Psychol.* 4, 25-55.

Ref Type: Abstract

Picard, N. & Strick, P. L. (2001). Imaging the premotor areas.  
*Curr.Opin.Neurobiol.*, 11, 663-672.

Povinelli, D. J. & Vonk, J. (2003). Chimpanzee minds: suspiciously human?  
*Trends Cogn Sci.*, 7, 157-160.

Premack, D. (1990). The infant's theory of self-propelled objects. *Cognition*,  
36, 1-16.

Premack, D. & Premack, A. J. (1995). Origins of human social competence. In  
M.S.Gazzaniga (Ed.), *The Cognitive Neurosciences* (pp. 205-218). Cambridge, MA:  
MIT Press/ Bradford Books.

Premack, D. & Woodruff, G. (1978). Does the chimpanzee have a theory of  
mind? *Behavioural and Brain Sciences*, 4, 515-526.

Prinz, W. (2002). Experimental approaches to imitation. In A.Meltzoff & W.  
Prinz (Eds.), *The imitative mind: Development, evolution, and brain bases* (pp. 143-  
162). Cambridge,UK: Cambridge University Press.

Puce, A., Allison, T., Bentin, S., Gore, J. C., & McCarthy, G. (1998).  
Temporal cortex activation in humans viewing eye and mouth movements.  
*J.Neurosci.*, 18, 2188-2199.

Puce, A., Allison, T., Gore, J. C., & McCarthy, G. (1995). Face-sensitive  
regions in human extrastriate cortex studied by functional MRI. *J.Neurophysiol.*, 74,  
1192-1199.

Puce, A. & Perrett, D. (2003). Electrophysiology and brain imaging of biological motion. *Philos.Trans.R.Soc.Lond B Biol.Sci.*, 358, 435-445.

Purcell, E. M., Torrey, H. C., & Pound, R. V. (1945). Resonance absorption by nuclear magnetic moment. *Phys.Rev.*, 69, 37-38.

Rilling, J., Gutman, D., Zeh, T., Pagnoni, G., Berns, G., & Kilts, C. (2002). A neural basis for social cooperation. *Neuron*, 35, 395-405.

Rizzolatti, G., Camarda, R., Fogassi, L., Gentilucci, M., Luppino, G., & Matelli, M. (1988). Functional organization of inferior area 6 in the macaque monkey. II. Area F5 and the control of distal movements. *Exp.Brain Res.*, 71, 491-507.

Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Brain Res.Cogn Brain Res.*, 3, 131-141.

Rizzolatti, G., Fogassi, L., & Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nat.Rev.Neurosci.*, 2, 661-670.

Rorden, C. & Brett, M. (2000). Stereotaxic display of brain lesions. *Behav.Neurol.*, 12, 191-200.

Rowe, A. D., Bullock, P. R., Polkey, C. E., & Morris, R. G. (2001). 'Theory of mind' impairments and their relationship to executive functioning following frontal lobe excisions. *Brain*, 124, 600-616.

Runeson, S. & Frykholm, G. (1983). Kinematic specification of dynamics as an informational basis for person and action perception: Expectation, gender recognition, and deceptive intention. *J.Exp.Psychol.Gen.*, 112, 585-615.

Sakata, H., Taira, M., Kusunoki, M., Murata, A., & Tanaka, Y. (1997). The TINS Lecture. The parietal association cortex in depth perception and visual control of hand action. *Trends Neurosci.*, *20*, 350-357.

Sakata, H., Taira, M., Murata, A., & Mine, S. (1995). Neural mechanisms of visual guidance of hand action in the parietal cortex of the monkey. *Cereb.Cortex*, *5*, 429-438.

Samson, D., Apperly, I. A., Chiavarino, C., & Humphreys, G. W. (2004). Left temporoparietal junction is necessary for representing someone else's belief. *Nat.Neurosci.*, *Advance online publication*, 1-2.

Santello, M., Flanders, M., & Soechting, J. F. (1998). Postural hand synergies for tool use. *J.Neurosci.*, *18*, 10105-10115.

Saxe, R. & Kanwisher, N. (2003). People thinking about thinking people. The role of the temporo-parietal junction in "theory of mind". *Neuroimage.*, *19*, 1835-1842.

Schenk, T. & Zihl, J. (1997). Visual motion perception after brain damage: II. Deficits in form-from-motion perception. *Neuropsychologia*, *35*, 1299-1310.

Scholl, B. J. & Tremoulet, P. D. (2000). Perceptual causality and animacy. *Trends Cogn Sci.*, *4*, 299-309.

Schultz, R. T., Grelotti, D. J., Klin, A., Kleinman, J., Van der, G. C., Marois, R., & Skudlarski, P. (2003). The role of the fusiform face area in social cognition: implications for the pathobiology of autism. *Philos.Trans.R.Soc.Lond B Biol.Sci.*, *358*, 415-427.

- Searle, J. R. (1983). *Intentionality: an essay in the philosophy of mind*. Cambridge: Cambridge University Press.
- Shidara, M., Kawano, K., Gomi, H., & Kawato, M. (1993). Inverse-dynamics encoding of eye movement by Purkinje cells in the cerebellum. *Nature*, *365*, 50-52.
- Siegal, M. & Varley, R. (2002). Neural systems involved in theory of mind. *Nat.Rev.Neurosci.*, *3*, 462-471.
- Singer, T., Kiebel, S. J., Winston, J. S., Dolan, R. J., & Frith, C. D. (2004). Brain responses to the acquired moral status of faces. *Neuron*, *41*, 653-662.
- Sperber, D. & Wilson, D. (1986). *Relevance: communication and cognition*. Oxford: Blackwell.
- Stone, V. E., Baron-Cohen, S., & Knight, R. T. (1998). Frontal contribution to Theory of Mind. *J.Cogn Neurosci.*, *10*, 640-656.
- Strafella, A. P. & Paus, T. (2000). Modulation of cortical excitability during action observation: a transcranial magnetic stimulation study. *Neuroreport*, *11*, 2289-2292.
- Stuss, D. T., Gallup, G. G. Jr., & Alexander, M. P. (2001). The frontal lobes are necessary for 'theory of mind'. *Brain*, *124*, 279-286.
- Tai, Y. F., Scherfler, C., Brooks, D. J., Sawamoto, N., & Castiello, U. (2004). The human premotor cortex is 'mirror' only for biological actions. *Curr.Biol.*, *14*, 117-120.

Thornton, I. M., Rensink, R. A., & Shiffrar, M. (2002). Active versus passive processing of biological motion. *Perception, 31*, 837-853.

Tomasello, M. (1999). *The cultural origins of human cognition*. Cambridge, MA: Harvard University Press.

Tomasello, M., Call, J., & Hare, B. (2003). Chimpanzees understand psychological states - the question is which ones and to what extent. *Trends Cogn Sci., 7*, 153-156.

Toni, I., Thoenissen, D., & Zilles, K. (2001). Movement preparation and motor intention. *Neuroimage., 14*, S110-S117.

Tranel, D., Damasio, H., & Damasio, A. R. (1997). A neural basis for the retrieval of conceptual knowledge. *Neuropsychologia, 35*, 1319-1327.

Travis, L. L. (1997). Goal-based organization of event memory in toddlers. In P.W.van den Broek, P. J. Bauer, & T. Bourg (Eds.), *Developmental spans in event comprehension and representation: bridging fictional and actual events* (pp. 111-138). Mahwah, NJ: Erlbaum.

Tremoulet, P. D. & Feldman, J. (2000). Perception of animacy from the motion of a single object. *Perception, 29*, 943-951.

Turner, R., Le Bihan, D., Moonen, C. T., Despres, D., & Frank, J. (1991). *Mag Res Med, 22*, 159-166.

Umiltà, M. A., Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keysers, C., & Rizzolatti, G. (2001). I know what you are doing: a neurophysiological study. *Neuron, 31*, 155-165.

Vaina, L. M., Lemay, M., Bienfang, D. C., Choi, A. Y., & Nakayama, K. (1990). Intact "biological motion" and "structure from motion" perception in a patient with impaired motion mechanisms: a case study. *Vis.Neurosci.*, *5*, 353-369.

Vaina, L. M., Solomon, J., Chowdhury, S., Sinha, P., & Belliveau, J. W. (2001). Functional neuroanatomy of biological motion perception in humans. *Proc.Natl.Acad.Sci.U.S.A*, *98*, 11656-11661.

Vanzetta, I. & Grinvald, A. (1999). Increased cortical oxidative metabolism due to sensory stimulation: implications for functional brain imaging. *Science*, *286*, 1555-1558.

Villringer, A. & Dirnagl, U. (1995). Coupling of brain activity and cerebral blood flow: basis of functional neuroimaging. *Cerebrovasc Brain Metab Rev*, *7*, 240-276.

Vogeley, K., Bussfeld, P., Newen, A., Hermann, S., Happe, F., Falkai, P., Maier, W., Shah, N. J., Fink, G. R., & Zilles, K. (2001). Mind reading: neural mechanisms of theory of mind and self-perspective. *Neuroimage.*, *14*, 181.

von Kriegstein, K., Eger, E., Kleinschmidt, A., & Giraud, A.-L. (2003). Modulation of neural responses to speech by directing attention to voices or verbal content. *Cognit.Brain Res.*, *17*, 48-55.

Wachsmuth, E., Oram, M. W., & Perrett, D. I. (1994). Recognition of objects and their component parts: responses of single units in the temporal cortex of the macaque. *Cereb Cortex*, *4*, 509-522.

- Warren, J. D., Uppenkamp, S., Patterson, R. D., & Griffiths, T. D. (2003). Separating pitch chroma and pitch height in the human brain. *Proc.Natl.Acad.Sci.U.S.A*, *100*, 10038-10042.
- Weiskrantz, L., Elliott, J., & Darlington, C. (1971). Preliminary observations on tickling oneself. *Nature*, *230*, 598-599.
- Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: the truth about false belief. *Child Development*, *5*, 655-684.
- Westwood, D. A., Chapman, C. D., & Roy, E. A. (2000). Pantomimed actions may be controlled by the ventral visual stream. *Exp.Brain Res.*, *130*, 545-548.
- Wicker, B., Michel, F., Henaff, M., & Decety, J. (1998). Brain regions involved in the perception of gaze: a PET study. *Neuroimage.*, *8*, 221-227.
- Wimmer, H. & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, *13*, 103-128.
- Winstein, C. J., Grafton, S. T., & Pohl, P. S. (1997). Motor task difficulty and brain activity: investigation of goal-directed reciprocal aiming using positron emission tomography. *J.Neurophysiol.*, *77*, 1581-1594.
- Winston, J. S., Strange, B. A., O'Doherty, J., & Dolan, R. J. (2002). Automatic and intentional brain responses during evaluation of trustworthiness of faces. *Nat.Neurosci.*, *5*, 277-283.

Wohlschlaeger, A., Gattis, M., & Bekkering, H. (2003). Action generation and action perception in imitation: an instance of the ideomotor principle.

*Philos.Trans.R.Soc.Lond B Biol.Sci.*, 358, 501-515.

Wolpert, D. M., Doya, K., & Kawato, M. (2003). A unifying computational framework for motor control and social interaction. *Philos.Trans.R.Soc.Lond B Biol.Sci.*, 358, 593-602.

Wolpert, D. M., Ghahramani, Z., & Flanagan, J. R. (2001). Perspectives and problems in motor learning. *Trends Cogn Sci.*, 5, 487-494.

Wolpert, D. M. & Kawato, M. (1998). Multiple paired forward and inverse models for motor control. *Neural Netw.*, 11, 1317-1329.

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition*, 69, 1-34.

Woodward, A. L. & Sommerville, J. A. (2000). Twelve-month-old infants interpret action in context. *Psychological Science*, 11, 73-77.

Worsley, K. J., Marrett, P., Neelin, P., Friston, K. J., & Evans, A. C. (1996). A unified statistical approach for determining significant signals in images of cerebral activation. *Hum.Brain Mapp.*, 4, 58-73.

Zatorre, R. J., Bouffard, M., & Belin, P. (2004). Sensitivity to auditory object features in human temporal neocortex. *J.Neurosci.*, 24, 3637-3642.