# ORIGINAL RESEARCH

# Overcoming methodological limitations in theory-ofmind neuroscientific experiments: A voxel based meta-analysis of verbal and non-verbal theory of mind studies

Raul Ando Junior<sup>1</sup>, Alvaro Machado Dias<sup>2</sup>, Adam McNamara<sup>3</sup>

1. Institute of Economy, State University of Campinas, UNICAMP, São Paulo, Brazil. 2. Clinical Neuroscience Lab 'LINC', Federal University of Sao Paulo, Brazil. 3. Department of Psychology, University of Surrey, Guildford, United Kingdom.

**Correspondence:** Prof. Alvaro Machado Dias. Address: Federal University of São Paulo. Pedro de Toledo St., 669 - 3° floor - Vila Clementino - São Paulo, SP, Brazil. Telephone: 55-115-576-4845. Email: alvaromd@usp.br.

| Received: January 13, 2013 | Accepted: May 2, 2013     | Online Published: May 21, 2013 |
|----------------------------|---------------------------|--------------------------------|
| DOI: 10.5430/jbgc.v3n4p27  | URL: http://dx.doi.org/10 | ).5430/jbgc.v3n4p27            |

### Abstract

Components of neuropsychological tests commonly employed by neuroimagers in developing a metric for Theory of Mind (ToM) research are currently extremely mixed. Tests commonly include factors which may be separate from the core ToM component such as emotion, morality and verbal skills. We explore these confounds, and hypothesize that the discrepant conclusions that have arisen in neuroimaging ToM research reflect the degree to which a task is confounded by these factors. To test this hypothesis we conduct a coordinate based activation likelihood estimations (ALE) meta-analysis of neuroimaging studies. The results demonstrate divergences in neuroimaging results are predicted by divergences in the use of verbal or non-verbal tasks. In response to this finding we systematically describe the main dimensions that should be excluded and included in an ideal ToM task drawing on the wide literature surveyed. We argue, good tasks should rely on minimum verbal content; avoid excessive empathic and moral reactions; include human-like figures executing purposeful behaviors; present several standardized phases, with increased levels of difficulty; and incorporate a clear scoring system, so that performance can be categorized. In order to try to accomplish these goals, we introduce the new 'Inverted Comic Strips Task'.

#### Key words

Theory of mind, Neuroimaging meta-analysis, Neuropsychological tasks

# **1** Introduction

The search for brain correlates of social cognition relies on the definition of its essential features. According to a notable review <sup>[1]</sup>, these features fall into two main categories: 'social reinforcements' and 'Theory of Mind' (ToM).

Traditionally, it is said that ToM stands for the general ability to understand someone else's mental content. In some contexts ToM can be conceived to represent a type of 'understanding', that is, something that the person holds and

develops as knowledge in his mind. In other contexts it may be more appropriate to replace the assumption of such clarity by that of the capacity to 'anticipate the development of a behavioral plan' or the unfolding of a teleological sequence of actions. The anticipation of a behavioral plan may occur independent of whether a person holds and is aware of mental representations about the someone else's mental content (for in-depth discussions: <sup>[2, 3]</sup>).

Based on the classic conceptualization of Premack<sup>[4]</sup>, ToM studies can be divided into perceptual, motivational and informational types. The first type basically relates to mental computations of facial expressions, gaze direction (joint attention), and attention to hidden objects. The second relates to behavioral interactions whilst the third comprehends mental representations<sup>[5]</sup>.

Several species show perceptual and motivational ToM capacities <sup>[5]</sup>, while informational ToM is, by definition, representational and (probably) exclusive of humans and, in a much lower extent, other primates <sup>[6]</sup>. Informational ToM involves the most computational demanding cognitive processes and supports the comprehension of several cultural products (e.g., movies, stories), as well as inter-personal role-playing. According to our survey of the literature, informational ToM is the most represented in both cognitive and social neuroscientific literatures <sup>[7]</sup>.

This paper is mainly concerned with the methodological pitfalls that we find to hamper the evolution of ToM studies, in social neuroscience and related domains (e.g., neuropsychiatry). As pointed out in a meta-analysis of ToM schizophrenic studies, "since there is a serious lack of research on the psychometric properties (including construct validity and criterion validity) of the many different theory of mind tasks that have been developed it may not be possible to formulate completely objective inclusion criteria regarding the type of tasks used in the studies" <sup>[8]</sup>.

An overview of the literature on neural correlates of ToM produce mixed results in relation to the existence of a consensus on brain circuits involved in ToM. Some key authors claim that a consensus has emerged, but experimental data does not necessarily confirm their claims. For example, in a authoritative review <sup>[9]</sup> provides an overview of the current field of Social Psychology. It reports a convergence around the perspective that the medial prefrontal cortex (MPFC) is the centre of the core domains: self-concept, attitudes and evaluation, emotional experience, and understanding the mind of others, that is, informational ToM. Aichhorn and collaborators dispute this conclusion stating that "evidence that the core processing for ToM cannot be located in the MPFC came from patients with lesions in relevant areas of the MPFC (...) thus, attention shifted to the TPJ (temporal-parietal junction) a key hub for reasoning about action and appreciating perspective differences" <sup>[10]</sup>. This view is substantiated by a case report discussing the outcome of a rare form of stroke, which affected the MPFC, (a bilateral anterior cerebral artery infarction, without anterior communicating artery aneurysms). The case report concluded that "the regions of the medial frontal lobes damaged in G.T. are not necessary for at least the cognitive aspects of ToM" <sup>[11]</sup>.

There are quite a lot of examples confirming this divergence among tendencies that seem unquestionable at first sight. We hypothesize that this state of things relates both to the fact that (1) a biocomputation as dynamic and complex as ToM probably relies on the activation of brain circuits that cannot be reduced to one area and (2) that excessive variation among applied neuropsychological tasks – particularly in the sense that some are verbal and some non-verbal – leads to the recruitment of brain circuits dedicated to a plethora of mental capacities. These capacities tend to be clustered differently by different researchers, driving the establishment of different conceptions in regard to the biological basis of ToM.

This is a bold perspective which requires investigating, but before we attend to that level it is important to decide between two claims: (1) It is worthwhile to include several different types of tasks in the consolidation of any possible consent on ToM's neural substrate; (2) It is agreeable to isolate informational ToM's essential features and establish the features of choice for standardized ToM tasks. The latter seeks specifically to discriminate ToM from other mental and neurobiological processes.

The bases for the first claim are: (1) The assumption that we are unable to define the exact mental processes by which a third person perspective are generated and attributed. This assumption renders the imposition of clear-cut boundaries between informational ToM and other mental abilities. Thereby limitations on appropriate tasks would be arbitrary and misleading. (2) The perspective that several abilities and capacities (including IQ) are best evaluated by multi-dimensional than by one-dimensional constructs.

The bases for the second claim are: (1) Minimum content and standardization are prerequisite conditions to turn social neuroscience into a widespread inter-cultural discipline with a framework in which to further experiment <sup>[12]</sup>; (2) Informational ToM is not like the IQ, but more akin to one of its dimensions; that is, if we seek to reach the development of some sort of social IQ, we need to convert social abilities into dimensions and this asks for standardization; 3. The definition of straightforward neural correlates of ToM can dramatically improve the evolving field of neuropsychiatric diagnoses and help to differentiate disorders' sub-types, based on unified cognitive-neurobiological evaluation.

It is hard to define what the best approach is for all times, but we take for granted that at this moment the second claim should prevail, since it alleviates important limitations that the efforts to establish the neurobiological basis of informational ToM are facing, as suggested above. From that standpoint, the next sections aim to: (1) Test whether the main categorical divergence among ToM tasks – verbal/non-verbal – leads to the recruitment of independent neural circuits, using a meta-analytical approach; (2) Discuss the dimensions that should be included and avoided in the neuropsy-chological tasks designed for neurobiological studies of informational ToM, based on the findings raised in the previous section; 3. Discuss well-known tasks in relation to these paradigms; 4. Introduce a new task, which tries to overcome current methodological limitations.

One methodological trend that emerges from the current neuroscientific literature on ToM regards the lack of discrimination between verbal and non-verbal tasks. This leads to difficulty in establishing valid conclusions on the neural correlates of ToM. For that reason, we conduct a coordinate based activation likelihood estimations (ALE) meta-analysis of neuroimaging studies examining the neural correlates of ToM specifically examining the impact of the verbal nature of the tasks. We focus on the convergences and divergences in reported coordinates of studies employing verbal and non-verbal ToM tasks in normal adults, and include only ToM studies that do not rely on empathy or moral reasoning, since these trends could bias the discrimination of the essential neural correlates of the capacity of taking a third person perspective/prospecting intentional content.

### 2 Material and methods

This section followed to a large extent, methods applied in a previous ALE meta-analysis (Ferreira et al, 2009). In order to collect studies for this section, in 07/20/2010 and 07/21/2010 we systematically searched Medline, ISI Web of Science and Scopus, using the keywords: [("fMRI" or "Voxel" or "neural correlates" or "brain correlates" or "brain imaging") and ("ToM" or "Theory of Mind" or "Social Cognition")], and retrieved all (and only) full experimental papers, written in English, whose abstracts suggested to hold data that meet our purpose. In a second step, we searched the reference lists of these papers and the Sleuth database (http://brainmap.org/sleuth/index.html), for missing studies.

Eligible ToM studies reported ALE coordinates, based upon whole brain analysis, and included normal adults without predefined history of psychiatric or neurological disorders (data from adult controls of papers involving youngsters or psychiatric patients were accepted). Applied neuropsychological ToM tasks had to be clearly described as either verbal or visual; studies involving mixed verbal/non-verbal neuropsychological paradigms, as well as those which addressed matters related to potentially confounding cognitive or affective dimensions (e.g., empathy, moral reasoning) were left aside.

In order to generate our results, reported peak activation of three-dimensional MNI coordinates (MNI stands for: Montreal Neurological Institute; it defines coordenates for brain mapping that are alternative to Talairach atlas) were summed and a modeled activation map was generated for each study, after which, all maps were associated to generate 'Ale scores', further thresholded with a false discovery rate of p < 0.01. Finally, clusters that passed the threshold were overlaid onto an anatomical MNI template, using MRIcron software (http://www.sph.sc.edu/comd/rorden/mricron/), as described elsewhere <sup>[13]</sup>. To normalize spatial coordinates into a single stereotactic space, we transformed MNI into Talairach coordinates, following standard instructions, as described by <sup>[14]</sup> (for details: http://imaging.mrc-cbu.cam.ac.uk/imaging/MniTalairach).

## **3 Results**

Among the near 200 papers that we found to be potentially eligible for this meta-analysis, 9 papers met our inclusion criteria <sup>[15-23]</sup>. Ref. 15 and 21 contributed with two experiments each, the rest of the papers contributed with only one.

The main result of this analysis is that neuroimaging findings using verbal ToM tasks tend to rely on the contribution of only 2 clusters, whose peaks are on BA 9 and 39, while non-verbal tasks render 5 significant clusters, whose peaks are on BA 37, 21 and 22 (see Figure 1). We do confirm Mitchell's claim that the temporal lobe is the core area involved in ToM (BA 21, 22, 37 and 39 are all in the temporal lobe – although only BA 21 is in the middle temporal lobe, contrary to his claim), but we also point to the existence of slight distinctions and partial independence between verbal and non-verbal ToM areas.

| References (Verbal)           | [16] | [17] | [20] | [21] | [23] |
|-------------------------------|------|------|------|------|------|
| Number of Subjects            | 6    | 6    | 10   | 17   | 17   |
| Frontal Lobe                  |      |      |      |      |      |
| Medial Frontal Gyrus, *, L    |      |      | Х    |      | Х    |
| Medial Frontal Gyrus, G, L    | Х    |      |      |      | Х    |
| Medial Frontal Gyrus, G, R    | Х    |      |      |      |      |
| Precentral Gyrus, W, L        |      |      | Х    |      |      |
| Sub-Gyral, W, L               |      |      | Х    |      |      |
| Superior Frontal Gyrus, G, L  |      |      | Х    |      |      |
| Superior Frontal Gyrus, W, R  |      |      | Х    |      |      |
| Limbic Lobe                   |      |      |      |      |      |
| Anterior Cingulate, *, L      | Х    |      |      |      |      |
| Anterior Cingulate, G, L      |      |      |      | Х    |      |
| Parahippocampal Gyrus, G, L   |      |      |      | Х    |      |
| Occipital Lobe                |      |      |      |      |      |
| Middle Temporal Gyrus, W, L   |      |      | Х    |      |      |
| Parietal Lobe                 |      |      |      |      |      |
| Angular Gyrus, W, R           |      |      | Х    |      |      |
| Precuneus, *, L               |      |      |      |      | Х    |
| Temporal Lobe                 |      |      |      |      |      |
| Inferior Temporal Gyrus, *, L |      |      | Х    |      |      |
| Middle Temporal Gyrus, G, R   |      | Х    | Х    |      |      |
| Middle Temporal Gyrus, W, L   |      |      | Х    |      |      |
| Middle Temporal Gyrus, W, R   |      | Х    |      |      |      |
| Sub-Gyral, W, R               | Х    |      |      |      |      |
| Superior Temporal Gyrus, G, L |      |      |      |      | Х    |
| Superior Temporal Gyrus, W, L |      | Х    | Х    |      |      |

Table 1. Reported ROIs among neuroimage experiments (VBM-based) based on Verbal ToM tasks

| Table 2. | Reported | ROIs among n | euroimage exp | periments ( | VBM-based | ) based or | n Non-Ve | erbal Tol | M tasks |
|----------|----------|--------------|---------------|-------------|-----------|------------|----------|-----------|---------|
|          |          |              |               |             |           | /          |          |           |         |

| References (Non-Verbal)        | [15] | [15] | [18] | [19] | [21] | [22] |
|--------------------------------|------|------|------|------|------|------|
| Number of Subjects             | 8    | 8    | 12   | 28   | 17   | 13   |
| Anterior Lobe                  |      |      |      |      |      |      |
| Culmen, G, R                   |      |      | Х    |      |      |      |
| Frontal Lobe                   |      |      |      |      |      |      |
| Inferior Frontal Gyrus, G, R   |      |      |      |      | Х    |      |
| Inferior Frontal Gyrus, W, L   |      |      | Х    |      | Х    | Х    |
| Inferior Frontal Gyrus, W, R   | Х    |      | Х    | Х    |      | Х    |
| Medial Frontal Gyrus, *, R     |      |      |      | Х    |      | Х    |
| Medial Frontal Gyrus, G, R     |      |      | Х    |      |      |      |
| Medial Frontal Gyrus, W, L     |      |      |      |      | Х    |      |
| Medial Frontal Gyrus, W, R     | Х    |      |      |      |      |      |
| Middle Frontal Gyrus, G, L     |      |      |      |      |      | Х    |
| Middle Frontal Gyrus, W, R     |      |      |      |      |      | Х    |
| Precentral Gyrus, *, L         |      |      |      |      |      | Х    |
| Precentral Gyrus, W, L         |      |      |      |      |      | Х    |
| Sub-Gyral, W, R                |      |      |      |      |      | Х    |
| Superior Frontal Gyrus, *, R   |      |      | Х    |      |      |      |
| Superior Frontal Gyrus, G, L   |      |      | Х    |      |      |      |
| Superior Frontal Gyrus, W, R   |      |      |      |      | Х    |      |
| Limbic Lobe                    |      |      |      |      |      |      |
| Anterior Cingulate, W, L       |      | Х    |      |      |      |      |
| Parahippocampal Gyrus, G, R    |      |      |      |      | Х    |      |
| Parahippocampal Gyrus, W, L    |      |      |      |      | Х    |      |
| Posterior Cingulate, G, L      |      |      |      |      |      | Х    |
| Occipital Lobe                 |      |      |      |      |      |      |
| Cuneus, G, L                   |      |      |      |      |      | Х    |
| Cuneus, G, R                   |      | Х    |      |      |      |      |
| Fusiform Gyrus, W, R           |      | Х    |      |      |      |      |
| Inferior Occipital Gyrus, W, R |      |      |      | Х    |      |      |
| Lingual Gyrus, G, L            |      |      |      | Х    |      |      |
| Lingual Gyrus, W, L            |      |      |      |      |      | Х    |
| Middle Occipital Gyrus, W, L   |      |      |      |      | Х    |      |
| Parietal Lobe                  |      |      |      |      |      |      |
| Inferior Parietal Lobule, G, L |      |      | Х    |      |      |      |
| Inferior Parietal Lobule, W, R |      |      | Х    |      |      | Х    |
| Precuneus, G, R                |      |      |      |      | Х    | Х    |
| Precuneus, W, R                |      |      | Х    |      |      |      |
| Pons                           |      |      |      |      |      |      |
| *, *, Right Brainstem          |      |      | Х    |      |      |      |
| Posterior Lobe                 |      |      |      |      |      |      |
| Pyramis, G, Left Cerebellum    |      |      | Х    |      |      | Х    |
| Tuber, G, Right Cerebellum     |      |      |      |      |      | Х    |
| Sub-lobar                      |      |      |      |      |      |      |
| Extra-Nuclear, W, L            |      | Х    | Х    |      |      |      |
| Extra-Nuclear, W, R            |      |      | Х    |      |      |      |
| Temporal Lobe                  |      |      |      |      |      |      |
| Fusiform Gyrus, *, R           |      | Х    |      |      |      |      |
| Fusiform Gyrus, G, L           |      |      |      |      |      | Х    |
| Fusiform Gyrus, G, R           |      |      |      |      | Х    | Х    |
| Fusiform Gyrus, W, L           |      |      | Х    |      |      |      |
| Int. Temporal Gyrus, W, R      |      | Х    |      |      |      |      |
| Middle Temporal Gyrus, G, R    |      |      | Х    |      | X    | X    |
| Middle Temporal Gyrus, W, L    |      |      |      |      | Х    | Х    |



Figure 1. Significant clusters for verbal and non-verbal ToM

# 4 Discussion

The findings above suggest that whenever one relies on ToM verbal tasks, one will find a much clearer picture than if placing the focus on findings emerging from non-verbal tasks. Additionally, the fall out of studies when applying rigorous inclusion criteria must be noted to be exceptionally high. The high fall out renders the meta-analysis less statistically powerful but additionally represents a compelling argument in favor for the standardization of ToM neuropsychological tasks. From that standpoint it is important to ask: what mental capacities should be avoided and included in ToM tasks?

#### 4.1 Disentangling potentially confounding abilities

Our review of the literature reveals that the mental capacities that introduce the most prominent dilemmas in ToM neuroscientific studies are language, empathy, and moral cognition. This is not to say that informational ToM holds back subtle relations with other mental capacities and abilities (e.g. meta-cognition <sup>[24]</sup>), but solely that these tend not to be as troublesome, either because they appear less, or because they can be avoided with simple measures, contrary to the capacities discussed hereunder.

#### 4.1.1 Language

As any type of relation between thought and language, the relation between the latter and ToM is highly labyrinthine and cannot be disentangled here. As long as it is assumed that what we call 'mental content' can be reduced to 'propositional content' without loss of generality <sup>[25]</sup>, ToM can be said to be structured as is language (carrying syntactic properties, obeying principles associated to natural languages). Conversely, the denial of this standpoint endorses the independence of

these processes, pointing to the importance of avoiding taking one for the other (for a recent and thorough discussion of the relations between social cognition and language<sup>[26]</sup>).

We will not define a specific position in this debate (as it is needless here), but rather emphasize the benefit of avoiding methodological issues, at this moment. A remarkable finding is that the autistic ToM impairments are not equally distributed throughout the available neuropsychological tasks, but considerably more prominent in verbal than in visual assignments <sup>[27]</sup>. Persons who were born deaf also present ToM impairments <sup>[28]</sup>, although their (limited) ToM ability evolves during life following a regular age-related schedule <sup>[29]</sup>. When controlled for other variables, these individuals show a considerably better performance in non-verbal than in verbal tasks <sup>[30]</sup>.

Based on these and other similar findings, it is clear that researchers should have a judicious position in relation to the acceptable verbal requirements of informational ToM tasks and their instructions. Ideally, studies should all be based on non-verbal or nearly non-verbal instruments. This precaution could not only enhance clarity of neuroscientific studies, but aid the development of the studies of ToM-language developmental interface <sup>[31]</sup>.

#### 4.1.2 Empathy

In the same sense that ToM defines subtle relations with natural language processing, it defines subtle relations with empathy—although, these point in the opposite direction. The English word 'empathy' comes from the German Einfühlung, which literally means 'feeling into' <sup>[32]</sup>. Empathy stands for a non-declarative trait that has been defined in many ways, most of which can be related to two basic capacities: 1) cognitive role taking and imitation <sup>[33]</sup>. Affective role taking ("to empathize means to feel the same emotion as another" <sup>[34]</sup>); 2) concerns about someone else's feelings ("experiencing an affective response congruent with the other's well being" <sup>[34]</sup>). According to the general guidelines of the former, ToM and empathy stand for indissociable capacities; according to the latter, they represent different mental capacities and should be discerned. To the extent that this last perspective holds, self vs. other tasks should not be assumed as standard ToM tasks.

This point is also interesting because it touches the Theory Theory (TT) – Simulation Theory (ST) discussion. For instance, studies inspired by ST show that exposure to intentional acts leads to inter-personal synchronous activity of a specific group of neurons <sup>[35-37]</sup>, in both non-human primates (found in the premotor <sup>[38]</sup> and parietal cortex <sup>[39]</sup>) and humans (specially in Broadman areas 44 and 46 <sup>[40-42]</sup>), supporting imitation, 'role playing', and other capacities that can be considered within the scope of both ToM and cognitive empathy <sup>[43]</sup>.

On the other hand, when we consider that the naïve picture that people form about prospection (their mental representation of extrinsic content) and the object of prospection in someone else's mind, the phenomenological difference of the two is unveiled, since empathy presumes adherence to these representations. We cannot have empathy without some level of existential adherence to the prospected subject's condition, but this is perfectly feasible in relation to stricto sensu informational ToM. As one may note, this conception also applies to antipathy/averseness.

Empathy is a holistic (embodied) experience that binds one person to another; it carries a pre-representational process (grounding it emotionally) that seems to be precedent to the formation of third-person perspectives, and that can bias the prospection of mental content. This relative independence may explain why a selective brain lesion (MPFC) may simultaneously lead to "blunting empathy" and "intact third person perspective taking <sup>[11]</sup>.

In practice this is very simple: neuropsychological tasks should avoid evoking compassion, avoidance, and any type of affective adherence to intentional content, or the neurobiological correlates of performance will reflect much more than the fundamental properties of informational ToM.

#### 4.1.3 Moral reasoning

Moral cognitions are biocomputations whereby people react to situations constrained by socially-defined rules of behavior. These computations occupy a central position in the history of western philosophy and have been recently reexamined in neuroscientific experiments exploring moral dilemmas.

The available dilemmas can be divided into personal and impersonal <sup>[44]</sup>. The former relate to attitudes that affect specific individuals, the latter relate to attitudes that affect groups. Together they represent a very complicated methodological bearing upon ToM studies <sup>[45]</sup>. They introduce methodological issues that have analogous relations with those considered in relation to empathy.

The limitations of including moral feelings in informational ToM tasks begin with the fact that morality varies between people and between cultures, regardless of the existence of an evolutionary basis for general tendencies [46]. Certain attitudes that are generally considered to be either amoral or immoral in one society are not in another. In practice this raises a note of caution in relation to the several ToM tasks that rely on violations of social norms.

Second, it is paramount to bear in mind that social values regulate the attribution of intentions. For instance, in a series of well-known experiments, it was demonstrated that people tend to attribute intentions to others based on the "goodness" or "badness" of the side effects of their actions, with a bias toward intentional attribution to actions that produce deleterious side effects <sup>[47-49]</sup>. When discussing empathy we note that immediate empathic/avoidance reactions to the represented characters should be refrained from being used in informational ToM tasks. We also suggest that 'judicious reactive reactions' should be as avoided as well.

It is interesting to note that this matter is not solely related to abstract moral constructs (the rightness/wrongness of thirdperson actions), but rather becomes especially prominent when personal interests are taken into consideration. Personal interests tend to arouse even more intense and immediate reactions. In that sense, our third note of caution extends to tasks that manipulate expected values, which cast feelings and attributional schemas and thereby bias the circumscription of informational ToM.

For example, in an authoritative review, Behrens and collaborators <sup>[1]</sup> emphasized that the Ultimatum game (UG) plays a prominent role in ToM studies. The UG is a zero-sum, two-player game, wherein participants have to maximize gains through the division of a predefined amount of money <sup>[50]</sup>. There is no lower limit to the offers, but every time that one is rejected, both players receive nothing. In that sense, it is suggestive that both players should reason in terms of the opponent's strategy, and thus in terms of their mental content.

Based on utilitarian perspectives, the best choice for someone who is playing the responder is to accept any amount of money, while the best strategy for someone playing the proposer is to offer very low values. However, studies show that the average person is not as rational as possible: on average, 80% of the offers below 20% of the total amount are rejected. At the same time, proposers tend spontaneously to offer around 40% of the total sum to the other player. Both these tendencies neglect the rational maximization of the Expected Utility (EU) <sup>[51]</sup>. Contrary to human behavior, chimpanzees tend to present very low offers, and to accept anything above zero in the adapted version of the game, which characterizes them as 'rational maximizers' <sup>[52]</sup>; for a discussion on the meaning of EU and its role in human cognition <sup>[53]</sup>.

It is pretty clear that the human tendency to neglect the maximization of EU is not some kind of misguided reasoning tendency <sup>[54]</sup>, but rather a socially driven reaction to inequity (for a discussion <sup>[55]</sup>; for a comparative analysis of fairness in men and apes <sup>[56]</sup>).

In sum, there are actually two epistemological perspectives associating the mentalization of intentions with the UG. The first relates to the effort of mapping another player's intentions, in order to generate self-centered strategies to cope with

those intentions. The second relates to the trade-off between the tendency to maximize EU and social interactions that lead to potential severances in reaction to unfairness (for a discussion <sup>[57]</sup>).

At the heart of the difference between the two and the complexity to disentangle them both in phenomenological and neurobiological levels is that they elicit distinct reactions to the prospection of someone else's intention. The association of deeply rooted values and personal interests merge with the prospection of mental content in affective components. This leads to inevitable bias during inter and intra-cultural judgments.

In the context of non-cooperative games, moral cognitions represent post-representational computations par excellence. They follow and bias the circumscription of informational ToM computations. For this reason we consider that morality and empathy define opposite relations with informational ToM. This is also the reason we consider that an ideal ToM task should be carefully designed to avoid moral forces and expected values, as a whole.

As one may note, the most immediate and noteworthy consequence of this proposal is the indication to leave aside the most prominent tasks of the current time, non-cooperative games, at this initial phase of development of the field. Secondly, after incorporating all the aforementioned points into a crafted ToM task design, one may strengthen the specific line of studies on the neurobiological interplay between ToM and moral cognition, which has already reached very interesting conclusions <sup>[58]</sup>.

# 4.2 What should be included in the scope of minimum content tasks of informational ToM?

Now that we have defined what dimension should not be included as part of minimum content ToM tasks, it is time to outline what they should include.

Starting with a task that is non-verbal, non-moral, and that does not evoke strong empathic/antipathic reactions, it is sensible to assume the presence of purposeful behaviors, as its first sine qua non condition. An informational ToM task must present actions that affect a state of things (change reality), inasmuch that a group of people (controls) can assume the existence of a natural association between these actions and the unfolding of a mental plan.

This axiom is in line with the canonical assumption that ToM relates to the representation of authors, actions, and their relations <sup>[59]</sup>. The task should emphasize and measure the putative role of the relation between concatenated mental representations and specific effects on the external reality. The dynamics of the external reality should follow the progression of the concatenated mental representations. In a deeper sense, one may say that the unfolding of a behavioral sequence should be naturally characterized as a world-to-mind intentional investment <sup>[60]</sup>), provided that the capacity to build this correlation using predefined resources (e.g., something like pieces of a puzzle) can be evaluated and performance categorized.

Another axiomatic perspective, emerges from findings related to selective activity of brain networks to flesh and blood agents, in the context of purposeful interactions <sup>[61]</sup>. The authors applied several classic two-players ToM tasks (including the prisoner dilemma and the UG) to each of the participants. Participants were asked to play against several types of machines, each of which endowed with an external design relatively closer to a human body (a computer, a functional robot, an anthropomorphic robot, and a real human being), and found that the enhancement of the anthropomorphic properties of automata was directly related with reported motivation, as much as with blood flow perfusion in brain areas traditionally related to social cognition (e.g. medial prefrontal cortex; right temporo-parietal junction).

In line with this perspective, a recent study concluded that BOLD activation in brain areas commonly related to social cognition (e.g., right temporoparietal junction (rTPJ), anterior cingulate cortex (ACC), right superior temporal sulcus (rSTS), ventral striatum, and amygdala) was more intense when the subjects engaged in live social interactions

(cooperative games) with the experimenter (through a live feed video), than when they watched a recording of this very process <sup>[62]</sup>.

As one may note, this means that the characters of a ToM task should have well-defined anthropological characteristics.

A third perspective, is represented by the fact that the process of prospecting other people's mental content is not equally arduous along the different real-world situations, due to the presence of variables that ease or hamper it. Both the behavior of prospected agents and external scenarios can be manipulated in order to produce task phases of different levels of difficulty. This does not represent an essential condition of all tasks, but certainly should be taken as part of the scope of minimum content instruments.

We assume this aspect to be very important, since it represents the easiest pathway to the development of a ToM maturity index, as well as to the possibility of investigating the neurobiological correlates of progressive increases in ToM demands, under a standardized procedure. It is interesting to note the absence of studies addressing the question: 'what happens in the brain when social demands increase?'.

A forth, non-obligatory, but very important dimension emerges from a facet of real-world social interactions that is hardly represented in either neurobiological or psychological literature. If a prospector were to adhere intensively to mental representations of other people's purposes, he/she will tend to decrease his/her chances of success in a rapidly changing social environment. Reevaluation of contextual clues tends to affect the purposefulness of actions, pushing the necessity to constantly reconfigure projections of intentional content. In practice, this perspective can be satisfied by the exploitation of the principle that 'what seems to be at first, might not prove to be at last', as frequently proves to be the case, in the context of real world intentional interplays.

#### 4.3 Available ToM Tasks and the New Inverted Comic Strips Task (ICST)

In the previous sections we presented the main dimensions that should be avoided and included in the scope of informational ToM minimum content neuropsychological tasks. Herein we will review the current tasks in light of these dimensions and introduce a new task, which is expected to serve as an example of how to overcome current limitations. To sum up, the perspectives that were raised are:

- Minimum verbal content, in order to avoid biases represented by individual differences in linguistic abilities and disorders that affect verbal capacities.
- Avoidance of empathic/antipathic content, which cannot be easily disentangled from prospection.
- Avoidance of moral forces and non-cooperative games, for the same reason given above.
- Purposefulness, as the base for the prospection of intentional mental content.
- Use of human characters; people tend to be less motivated to think about the mental content of entities that are assumed as not having a mind.
- Phases of increasing complexity, in order to ease the way to individual differences (and classification) and patterns of brain activation associated with cognitive demands in standardized procedures (non-obligatory, but certainly strategic).
- Ambiguity disengagement (non-obligatory, but certainly strategic).

In light of these dimensions, the establishment of what tasks to review was determined as follows: First we considered tasks that were reported in the meta-analysis on informational ToM, indexed in PubMed; Second we searched recent neuroimage studies as indexed in several databases.

An initial literature search was made in PubMed using the keywords "Theory of Mind" and "Meta-analysis"; the full papers were retrieved and searched for the tasks that were used (some of these papers are related to neurobiological studies, some do not). These tasks were separated in agreement with their general descriptive features. After that procedure was concluded, a complementary search was made in PubMed, Google Scholar, and PsycINFO using the keywords "fMRI or magnetic stimulation" and "Theory of Mind", for the period of 2009/07/01 to 2009/08/01. All tasks were analyzed and the ones that were not related to informational ToM were not considered; the others were also separated in accordance with their general features.

Our results indicate that in 09/22/2009 the keywords 'Theory of Mind and Meta-analysis' retrieve 21 publications in Pubmed, among which only 8 are real meta-analysis of ToM studies, while only 6 presented a discussion of neuropsy-chological tasks. Among these publications:

Table 3. Current informational ToM tasks in meta-analytical studies

| All 6 mentioned a verbal version of the false belief task <sup>[63-68]</sup>                                                                                    |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 4 mentioned 'intention inference' from a verbal history <sup>[64, 65, 68, 69]</sup>                                                                             |
| Other tasks, like the Faux Pas task (social impropriety) <sup>[64]</sup> and the interpretation of face expression <sup>[70]</sup> , where mentioned only once. |

During the month of July (2009), 10 neuroimage studies were published, using the following tasks:

**Table 4.** Current informational ToM tasks used in social neuroscience

| False belief tasks <sup>[71-74]</sup>                                                                                               |
|-------------------------------------------------------------------------------------------------------------------------------------|
| Intention inference from a verbal history <sup>[75]</sup> ; completing stories <sup>[76]</sup> ; Strange story task <sup>[72]</sup> |
| Depicting the intention of characters of a comic strip <sup>[77, 78]</sup>                                                          |
| Sentence analysis in regard to mental states <sup>[79]</sup>                                                                        |
| Faux pas recognition test <sup>[72]</sup>                                                                                           |

According to our survey, none of these or other available tasks can be used to classify performance on a quantitative basis, and despite the fact that several tasks present multiple phases, none present phases based on incremental levels of difficulty—in accordance with the fact that none manipulates behaviors and scenarios in order to make the prospection of mental content increasingly harder.

Moreover, the two most applied type of task (false belief and intention inference from verbal histories) are 100% verbal. This presents potential difficulties for individuals with verbal difficulties as well as methodological problems in regards to identifying brain correlates for ToM. This becomes especially true should one need to compare differentially affected and non-affected individuals. The faux pas task has an unavoidable moral component, while 'interpretation of face expression' is not primarily related with informational ToM, but rather with perceptual ToM <sup>[4]</sup>.

As previously argued, this gap must be filled. Therefore we introduce the new 'Inverted Comic Strips Task' ('ICST'), as an example of a task capable of offering the most fundamental features that were found to be lacking in previous tasks: essential content, phases of increasing complexity, and a scoring system that allows a quantitatively based classification of performance. The task is also visually attractive (aiming to increase adherence in a fashion that does not introduce bias), with very simple instructions.

The ICST is a software-based task that comprises a number of comic strips, each of which presenting a purposeful social interaction that is expected to be comprehended in the inverted direction (from the end to the beginning of the story), using the minimum amount of information/frames as possible. Each strip is presented in a separated screen, so that only the conclusion of the story can be viewed in the first screen. This conclusion always consists of two main characters and a speech balloon, which defines the story's ultimate sense, based on the perspective that all the action follows from the unfolding of mental plans.



Figure 2. The inverted comic strip task

(a) Trial structure for task and score for that trial if participant is correct (green) or incorrect (red) in selecting response at each stage. Participant's are presented with the single image (b) from the cartoon series and have to determine which of the two characters is speaking. If a participant is unsure, they may 'buy' (cost = 1 point / follow black arrow) another sequential part of the strip (c-e) to gain more information. A participant may choose which character is speaking at any point and exit the trial (red arrow), participants are not required to 'buy' more information. Participants choose the character who is speaking by using a mouse to place the speech bubble over the character and then pressing the button marked "I'm done". Highest scores are attained by correctly attributing the statement to the correct character as early in the series as possible. The lowest scores are derived from making errors with plenty of information. Strips are intuitive and fun, the humorous 'cartoon' component of the strips greatly assists in keeping them morally neutral.

At the commencement of the task, the balloon is not associated with any of the characters, and the objective of the participants is to figure out to who it belongs, using the least number of drawing boxes, each thereof showing pieces of information that ease the process of mapping the characters' intentions. Levels of difficulty are defined by the number of drawing boxes that each story carries. The easier ones carry four while the hardest carry eight (each phase brings one extra drawing box), and the participants are not allowed to guess prior to buying at least one drawing box.

This underlying definition of difficulty relies on the fact that in shorter stories, the conclusion is closer to the beginning. In accordance with that, all drawing boxes cost the same, as much as all phases are scaled by their relative difficulty, which is precisely defined by the number of drawing boxes that the strip contains multiplied by the number of strips of the phase. There are three strips with four drawing boxes, two with six, and one with eight.

This task can be widely applied (non-clinical populations, autism, deaf, schizophrenia) and can be used as a model to alternative neuropsychological instruments (e.g., tasks with greater level of difficulty), in order to respond to the challenge of defining the core neural circuits associated with ToM, categorize performance (which could help us advance toward the development of a social IQ), and explore performance in Western and other cultures, without linguistic, empathic and moral biases.

The ICST also comprises a database to where the collected data can be uploaded (researchers may use it online or offline), for further analysis. This database implements basic statistical procedures (www.r-project.org) and aims to serve researchers with interest of using the test around the globe, as well as to provide automated means to compare performance across different populations.

The English version can be found at: http://hq.oda.mat.br/en, whereas the Portuguese version can be found at: http://hq. oda.mat.br, versions in other languages can be created, in accordance with request.

We aim to make this test part of a collaborative effort toward a better understanding of Theory of Mind and its neural correlates and, in a second moment, to define a suite of neuropsychological tests with the potential to take part in the establishment of a social IQ. We will be happy to assist colleagues with interest in using the ICST.

#### 4.4 Final remarks

This paper points to the existence of limitations among different ToM tasks that are currently been applied in social neuroscientific studies. This perspective is tied to a persuasive argument in favor of the abandonment of certain procedures and the development of a method to classify performance and permit the understanding of the neurobiological correlates of increasing social cognitive demands. Finally, one of the many possible solutions to this situation is presented.

As aforementioned, this is not a proposal for the integration of different behavioral frameworks currently used in ToM experiments, but rather a proposal for the introduction of an exclusive and unified framework. The framework incorporates essential features, in order to reach some consensus about the neurobiology of ToM, as well as aid the intra and inter-cultural expansion of social neuroscience. A long term aim would be to categorize performance and generate a social IQ. This would take the form of a suite of neuropsychological instruments akin to this minimum content proposal, related to all other social capacities that prove to be irreducible to any other.

#### References

- Behrens, T.E.J., L.T. Hunt, and M.F.S. Rushworth, The Computation of Social Behavior. Science, 2009. 324(5931): p. 1160-1164. PMid:19478175 http://dx.doi.org/10.1126/science.1169694
- [2] Wellman, H., The child's theory of mind1992: MIT Press Cambridge, MA.
- [3] Astington, J., The child's discovery of the mind1993: Harvard Univ Pr.
- [4] Premack, D., "Does the chimpanzee have a theory of mind?" revisited, in Machiavellian Intelligence, R.W. Byrne and A. Whiten, Editors. 1988, Oxford University Press: Oxford. p. 160-78.
- [5] Emery, N.J. and N.S. Clayton, Comparative Social Cognition. Annual Review of Psychology, 2009. 60(1): p. 87-113. PMid:18831684 http://dx.doi.org/10.1146/annurev.psych.60.110707.163526
- [6] Hampton, R.R., Rhesus monkeys know when they remember. Proceedings of the National Academy of Sciences of the United States of America, 2001. 98(9): p. 5359-5362. PMid:11274360 http://dx.doi.org/10.1073/pnas.071600998
- [7] Cacioppo, J., Foundations in social neuroscience2002: The MIT Press.
- [8] Sprong, M., et al., Theory of mind in schizophrenia: Meta-analysis. The British Journal of Psychiatry, 2007. 191(1): p. 5-13. PMid:17602119 http://dx.doi.org/10.1192/bjp.bp.107.035899
- [9] Mitchell, J.P., Social psychology as a natural kind. Trends in Cognitive Sciences, 2009. 13(6): p. 246-251. PMid:19427258 http://dx.doi.org/10.1016/j.tics.2009.03.008
- [10] Aichhorn, M., et al., Temporo-parietal Junction Activity in Theory-of-Mind Tasks: Falseness, Beliefs, or Attention. Journal of Cognitive Neuroscience, 2008. 21(6): p. 1179-1192. PMid:18702587 http://dx.doi.org/10.1162/jocn.2009.21082
- [11] Bird, C.M., et al., The impact of extensive medial frontal lobe damage on 'Theory of Mind' and cognition. Brain, 2004. 127(4): p. 914-928. PMid:14998913 http://dx.doi.org/10.1093/brain/awh108
- [12] Dias, A.M., The foundations of neuroanthropology. Frontiers in Evolutionary Neuroscience, 2010. 2.
- [13] Ferreira, L.K., et al., Neurostructural predictors of Alzheimer's disease: A meta-analysis of VBM studies. Neurobiology of Aging, 2009. In Press, Corrected Proof. PMid:20005012
- [14] Chau, W. and A.R. McIntosh, The Talairach coordinate of a point in the MNI space: how to interpret it. Neuroimage, 2005. 25(2):
   p. 408-416. PMid:15784419 http://dx.doi.org/10.1016/j.neuroimage.2004.12.007
- [15] Brunet, E., et al., Abnormalities of brain function during a nonverbal theory of mind task in schizophrenia. Neuropsychologia, 2003. 41(12): p. 1574-1582. http://dx.doi.org/10.1016/S0028-3932(03)00119-2
- [16] Fletcher, P.C., et al., Other minds in the brain: a functional imaging study of "theory of mind" in story comprehension. Cognition, 1995. 57(2): p. 109-128. http://dx.doi.org/10.1016/0010-0277(95)00692-R
- [17] Gallagher, H.L., et al., Reading the mind in cartoons and stories: an fMRI study of 'theory of mind' in verbal and nonverbal tasks. Neuropsychologia, 2000. 38(1): p. 11-21. http://dx.doi.org/10.1016/S0028-3932(99)00053-6
- [18] Gobbini, M.I., et al., Two takes on the social brain: a comparison of theory of mind tasks. J Cogn Neurosci, 2007. 19(11): p. 1803-14. PMid:17958483 http://dx.doi.org/10.1162/jocn.2007.19.11.1803

- [19] Kobayashi, C., G.H. Glover, and E. Temple, Children's and adults' neural bases of verbal and nonverbal 'theory of mind'. Neuropsychologia, 2007. 45(7): p. 1522-32. PMid:17208260 http://dx.doi.org/10.1016/j.neuropsychologia.2006.11.017
- [20] Ruby, P. and J. Decety, What you believe versus what you think they believe: A neuroimaging study of conceptual perspective-taking. European Journal of Neuroscience, 2003. 17(11): p. 2475-2480. http://dx.doi.org/10.1046/j.1460-9568.2003.02673.x
- [21] Vanderwal, T., et al., Self, mother and abstract other: An fMRI study of reflective social processing. Neuroimage, 2008. 41(4): p. 1437-1446. PMid:18486489 http://dx.doi.org/10.1016/j.neuroimage.2008.03.058
- [22] Vollm, B.A., et al., Neuronal correlates of theory of mind and empathy: a functional magnetic resonance imaging study in a nonverbal task. Neuroimage, 2006. 29(1): p. 90-8. PMid:16122944 http://dx.doi.org/10.1016/j.neuroimage.2005.07.022
- [23] Young, L., D. Dodell-Feder, and R. Saxe, What gets the attention of the temporo-parietal junction? An fMRI investigation of attention and theory of mind. Neuropsychologia, 2010. http://dx.doi.org/10.1016/j.neuropsychologia.2010.05.012
- [24] Doherty, M. and J. Perner, Metalinguistic awareness and theory of mind: Just two words for the same thing? Cognitive Development. 13(3): p. 279-305. http://dx.doi.org/10.1016/S0885-2014(98)90012-0
- [25] Wittgenstein, L., 1956: Philosophical Investigations1956, New York: MacMillan.
- [26] Fitch, W.T., L. Huber, and T. Bugnyar, Social Cognition and the Evolution of Language: Constructing Cognitive Phylogenies. Neuron, 2010. 65(6): p. 795-814. PMid:20346756 http://dx.doi.org/10.1016/j.neuron.2010.03.011
- [27] Leslie, A.M. and L. Thaiss, Domain specificity in conceptual development: neuropsychological evidence from autism. Cognition, 1992. 43(33): p. 225-251. http://dx.doi.org/10.1016/0010-0277(92)90013-8
- [28] Siegel, D., The Developing Mind: Toward a Neurobiology of Interpersonal Experience1999: Guilford Publications, Inc., Dept. 72, Spring Street, New York, NY 10012; Tel: 800-365-7006 (Toll free); Fax: 212-966-6708; Web site: http://www.guilford.com (\$37.95).
- [29] Woolfe, T., S. Want, and M. Siegal, Signposts to development: Theory of mind in deaf children. Child Development, 2002: p. 768-778. PMid:12038550 http://dx.doi.org/10.1111/1467-8624.00437
- [30] Figueras-Costa, B. and P. Harris, Theory of Mind Development in Deaf Children: A Nonverbal Test of False-Belief Understanding. J. Deaf Stud. Deaf Educ., 2001. 6(2): p. 92-102. PMid:15451854 http://dx.doi.org/10.1093/deafed/6.2.92
- [31] Antonietti, A., O. Sempio, and A. Marchetti, Theory of mind and language in developmental contexts2006: Springer Verlag. http://dx.doi.org/10.1007/b106493
- [32] Preston, S.D. and F.B.M. de Waal, Empathy: Its ultimate and proximate bases. Behavioral and Brain Sciences, 2001. 25(01): p. 1-20. http://dx.doi.org/10.1017/S0140525X02000018
- [33] Deutsch, F. and R.A. Maddle, Empathy: Historic and current conceptualizations, measurement, and a cognitive theoretical perspective. Human Development 1975. 18: p. 267-87. PMid:765261 http://dx.doi.org/10.1159/000271488
- [34] Eisenberg, N. and P.A. Miller, The relation of empathy to prosocial and related behaviors. Psychological Bulletin, 1987. 101(1): p. 91-119. http://dx.doi.org/10.1037/0033-2909.101.1.91
- [35] Gordon, R.M., Folk Psychology as Simulation. Mind & Language, 1986. 1(2): p. 158-171. http://dx.doi.org/10.1111/j.1468-0017.1986.tb00324.x
- [36] Gordon, R.M., The Simulation Theory: Objections and Misconceptions. Mind & Language, 1992. 7(1-2): p. 11-34. http://dx.doi.org/10.1111/j.1468-0017.1992.tb00195.x
- [37] Gordon, R.M., Reply to Stich and Nichols. Mind & Language, 1992. 7(1-2): p. 87-97. http://dx.doi.org/10.1111/j.1468-0017.1992.tb00198.x
- [38] Pellegrino, G., et al., Understanding motor events: a neurophysiological study. Experimental Brain Research, 1992. 91(1): p. 176-180. http://dx.doi.org/10.1007/BF00230027
- [39] Rizzolatti, G. and L. Craighero, The Mirror-Neuron System. Annual Review of Neuroscience, 2004. 27(1): p. 169-192. PMid:15217330 http://dx.doi.org/10.1146/annurev.neuro.27.070203.144230
- [40] Cheng, Y., A.N. Meltzoff, and J. Decety, Motivation modulates the activity of the human mirror-neuron system. Cerebral Cortex, 2007. 17(8): p. 1979-1986. PMid:17077160 http://dx.doi.org/10.1093/cercor/bhl107
- [41] Cheng, Y., et al., Sex differences in the neuroanatomy of human mirror-neuron system: A voxel-based morphometric investigation. Neuroscience, 2009. 158(2): p. 713-720. PMid:19010397 http://dx.doi.org/10.1016/j.neuroscience.2008.10.026
- [42] Gazzola, V., L. Aziz-Zadeh, and C. Keysers, Empathy and the Somatotopic Auditory Mirror System in Humans. Current Biology, 2006. 16(18): p. 1824-1829. PMid:16979560 http://dx.doi.org/10.1016/j.cub.2006.07.072
- [43] Ruby, P. and J. Decety, Effect of subjective perspective taking during simulation of action: a PET investigation of agency. Nat Neurosci, 2001. 4(5): p. 546-550. PMid:11319565
- [44] Greene, J.D., et al., The neural bases of cognitive conflict and control in moral judgment. Neuron, 2004. 44(2): p. 389-400. PMid:15473975 http://dx.doi.org/10.1016/j.neuron.2004.09.027

- [45] Knobe, J., Theory of mind and moral cognition: exploring the connections. Trends in Cognitive Sciences, 2005. 9(8): p. 357-359. PMid:16006176 http://dx.doi.org/10.1016/j.tics.2005.06.011
- [46] Baschetti, R., Evolution and morality. The Lancet, 2006. 367(9527): p. 1975-1976. http://dx.doi.org/10.1016/S0140-6736(06)68872-7
- [47] Cokely, E.T. and A. Feltz, Individual differences, judgment biases, and theory-of-mind: Deconstructing the intentional action side effect asymmetry. Journal of Research in Personality, 2009. 43(1): p. 18-24. http://dx.doi.org/10.1016/j.jrp.2008.10.007
- [48] Knobe, J., Intentional action and side effects in ordinary language. Analysis, 2003. 63(3): p. 190. http://dx.doi.org/10.1093/analys/63.3.190
- [49] Knobe, J., Intention, intentional action and moral considerations. Analysis, 2004. 64(2): p. 181. http://dx.doi.org/10.1093/analys/64.2.181
- [50] Rubinstein, A., Perfect Equilibrium in a Bargaining Model. Econometrica, 1982. 50(1): p. 97-109. http://dx.doi.org/10.2307/1912531
- [51] Nowak, M.A., K.M. Page, and K. Sigmund, Fairness Versus Reason in the Ultimatum Game. Science, 2000. 289(5485): p. 1773-1775. PMid:10976075 http://dx.doi.org/10.1126/science.289.5485.1773
- [52] Jensen, K., J. Call, and M. Tomasello, Chimpanzees Are Rational Maximizers in an Ultimatum Game. Science, 2007. 318(5847):
   p. 107-109. PMid:17916736 http://dx.doi.org/10.1126/science.1145850
- [53] Tversky, A. and D. Kahneman, Rational Choice and the Framing of Decisions. The Journal of Business, 1986. 59(s4): p. S251. http://dx.doi.org/10.1086/296365
- [54] Sanfey, A.G., et al., The Neural Basis of Economic Decision-Making in the Ultimatum Game. Science, 2003. 300(5626): p. 1755-1758. PMid:12805551 http://dx.doi.org/10.1126/science.1082976
- [55] Ohmura, Y. and T. Yamagishi, Why do people reject unintended inequity? Responders' rejection in a truncated ultimatum game. Psychol Rep, 2005. 96(2): p. 533-41. http://dx.doi.org/10.2466/pr0.96.2.533-541
- [56] Call, J. and K. Jensen, Chimpanzees may recognize motives and goals, but may not reckon on them. Novartis Found Symp, 2007. 278: p. 56-65; discussion 65-70, 89-96, 216-21.
- [57] Kahneman, D., Maps of Bounded Rationality: Psychology for Behavioral Economics. The American Economic Review, 2003. 93(5): p. 1449-1475. http://dx.doi.org/10.1257/000282803322655392
- [58] Young, L., et al., The neural basis of the interaction between theory of mind and moral judgment. Proceedings of the National Academy of Sciences, 2007. 104(20): p. 8235-8240. PMid:17485679 http://dx.doi.org/10.1073/pnas.0701408104
- [59] Frith, U. and C. Frith, The biological basis of social interaction. Current Directions in Psychological Science, 2001. 10(5): p. 151-155. http://dx.doi.org/10.1111/1467-8721.00137
- [60] Searle, J., Minds, Brains and Science: The 1984 Reith Lectures 1984, Cambridge: Harvard University Press.
- [61] Krach, S., et al., Can machines think? Interaction and perspective taking with robots investigated via fMRI. PLoS One, 2008. 3(7): p. e2597. PMid:18612463 http://dx.doi.org/10.1371/journal.pone.0002597
- [62] Redcay, E., et al., Live face-to-face interaction during fMRI: A new tool for social cognitive neuroscience. Neuroimage, 2010. 50(4): p. 1639-1647. PMid:20096792 http://dx.doi.org/10.1016/j.neuroimage.2010.01.052
- [63] Wellman, H.M. and D. Liu, Scaling of theory-of-mind tasks. Child Dev, 2004. 75(2): p. 523-41. PMid:15056204 http://dx.doi.org/10.1111/j.1467-8624.2004.00691.x
- [64] Bora, E., M. Yucel, and C. Pantelis, Theory of mind impairment in schizophrenia: meta-analysis. Schizophr Res, 2009. 109(1-3): p. 1-9. PMid:19195844 http://dx.doi.org/10.1016/j.schres.2008.12.020
- [65] Sprong, M., et al., Theory of mind in schizophrenia: meta-analysis. Br J Psychiatry, 2007. 191: p. 5-13. PMid:17602119 http://dx.doi.org/10.1192/bjp.bp.107.035899
- [66] Liu, D., et al., Theory of mind development in Chinese children: a meta-analysis of false-belief understanding across cultures and languages. Dev Psychol, 2008. 44(2): p. 523-31. PMid:18331141 http://dx.doi.org/10.1037/0012-1649.44.2.523
- [67] Wellman, H.M., D. Cross, and J. Watson, Meta-analysis of theory-of-mind development: the truth about false belief. Child Dev, 2001. 72(3): p. 655-84. PMid:11405571 http://dx.doi.org/10.1111/1467-8624.00304
- [68] Spreng, R.N., R.A. Mar, and A.S. Kim, The common neural basis of autobiographical memory, prospection, navigation, theory of mind, and the default mode: a quantitative meta-analysis. J Cogn Neurosci, 2009. 21(3): p. 489-510. PMid:18510452 http://dx.doi.org/10.1162/jocn.2008.21029
- [69] Di Martino, A., et al., Functional brain correlates of social and nonsocial processes in autism spectrum disorders: an activation likelihood estimation meta-analysis. Biol Psychiatry, 2009. 65(1): p. 63-74. PMid:18996505 http://dx.doi.org/10.1016/j.biopsych.2008.09.022

- [70] Schulte-Rüther, M., et al., Mirror Neuron and Theory of Mind Mechanisms Involved in Face-to-Face Interactions: A Functional Magnetic Resonance Imaging Approach to Empathy. Journal of Cognitive Neuroscience, 2007. 19(8): p. 1354-1372.
   PMid:17651008 http://dx.doi.org/10.1162/jocn.2007.19.8.1354
- [71] Rasmussen, C., K. Wyper, and V. Talwar, The relation between theory of mind and executive functions in children with fetal alcohol spectrum disorders. Can J Clin Pharmacol, 2009. 16(2): p. e370-80. PMid:19638654
- [72] Umeda, S., M. Mimura, and M. Kato, Acquired personality traits of autism following damage to the medial prefrontal cortex. Soc Neurosci, 2009: p. 1-11.
- [73] Saxe, R.R., et al., Brain Regions for Perceiving and Reasoning About Other People in School-Aged Children. Child Development, 2009. 80(4): p. 1197-1209. PMid:19630902 http://dx.doi.org/10.1111/j.1467-8624.2009.01325.x
- [74] Otsuka, Y., et al., Individual differences in the theory of mind and superior temporal sulcus. Neuroscience Letters, 2009. 463(2): p. 150-153. PMid:19632299 http://dx.doi.org/10.1016/j.neulet.2009.07.064
- [75] Bedny, M., A. Pascual-Leone, and R.R. Saxe, Growing up blind does not change the neural bases of Theory of Mind. Proc Natl Acad Sci U S A, 2009. 106(27): p. 11312-7. PMid:19553210 http://dx.doi.org/10.1073/pnas.0900010106
- [76] Saxe, R.R., et al., Brain regions for perceiving and reasoning about other people in school-aged children. Child Dev, 2009. 80(4): p. 1197-209. PMid:19630902 http://dx.doi.org/10.1111/j.1467-8624.2009.01325.x
- [77] Benedetti, F., et al., Functional and structural brain correlates of theory of mind and empathy deficits in schizophrenia. Schizophr Res, 2009. PMid:19632816 http://dx.doi.org/10.1016/j.schres.2009.06.021
- [78] Russell, T.A., et al., Aspects of social cognition in anorexia nervosa: Affective and cognitive theory of mind. Psychiatry Res, 2009. 168(3): p. 181-5. PMid:19467562 http://dx.doi.org/10.1016/j.psychres.2008.10.028
- [79] Otsuka, Y., et al., Individual differences in the theory of mind and superior temporal sulcus. Neurosci Lett, 2009. PMid:19632299 http://dx.doi.org/10.1016/j.neulet.2009.07.064