

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Primitive vs. advanced Theory-of-Mind: Defending and reformulating the difference

Teresa Bejarano. <https://orcid.org/0000-0002-0037-549X>

Department of Philosophy, Logic and Philosophy of Science. University of Sevilla (Spain).

tebefer@us.es

Primitive vs. advanced Theory-of-Mind: Defending and reformulating the difference

Abstract. Can we nowadays keep a qualitative difference between the primitive and advanced Theory-of-Mind? The old criteria have become blurry. In addition, it is clear that in ‘apes’ lifestyle’ it is not necessary to use the communicative-cognitive basic abilities which became indispensable in ‘the new lifestyle’. Thus, it is usual to conclude that apes would have to some degree such abilities. However, this article tries to reformulate and defend that qualitative difference. Thus, after underlining the contrast between two kinds of mental states (‘contents’ and ‘expectations’), I apply it to the detection of foreign mental states as well. Then, three points are proposed: First, ‘vicarious expectations’ sustain the primitive ToM; second, a subject can have no expectation of inner states which are intrinsically impossible for him; third, the state of interacting with ourselves as with a different person –e.g., the thinking what others think of us– cannot be a vicarious expectation of ours, but it requires the estimation of foreign contents. From this hypothesis, I deduce that vicarious expectations are unable to sustain self-conscious emotions or the really effective reception of pointing gestures. These abilities could appear only when ‘the estimation of foreign contents’ –i.e., the origin of the advanced ToM– arose.

Keywords: foreign mental contents; human sclera; reception of pointing gestures; self-conscious emotions; Theory-of-Mind; vicarious expectations

1) Introduction

Can we nowadays keep a qualitative contrast between the primitive mode of ‘Theory of Mind’ (ToM), and the advanced one (which was previously only attributed to humans of 4 years of age and older)? The criteria which still in 2003 separated clearly one from the other have become blurry. Likewise, the human uniqueness which was attributed to some very basic communicative abilities is being put in check by observations and experiments on animals. Moreover, the depreciation which used to be thrown on the success of trained apes is reduced to silence: Do we not also teach children those same abilities for months? At a deeper level, changes have also occurred. Now (probably thanks to the new theories of evolution, which emphasise development and cultural niche) the differences between apes’ lifestyle and “the new, human lifestyle” are center stage. Consequently, it has become clear that in apes’ lifestyle it is not necessary to use the communicative and cognitive basic abilities which are indispensable in the new lifestyle. From all this, it is usual to conclude that apes would have to some degree such abilities, and that these, therefore, would not be uniquely human.

But that conclusion is not necessary. We can alternatively think that “the new lifestyle” led to the uniquely human, advanced mode of ToM. This is the possibility that this article tries to explore.

Section 2 briefly presents the descriptions which in 2003 were published of – primitive and advanced – ‘Theory-of-Mind’ (ToM). Section 3 focuses on Tomasello 2018 and Southgate 2020, who attempt to accommodate the new data regarding the abilities of ToM in infants and apes without having to give up on the separation of the two modes of the ToM. I share their goal, but not their proposals. Likewise, I step away from the attempt made by Lurz et al. 2022. In 4, after highlighting the continued lack of consensus regarding the format in which goals or expectations are activated in non-human consciousness, I, in order to describe such format, choose the metaphor ‘well-

defined, empty profiles': Such emptiness would get their automatic, non-costly differentiation from current mental contents. This hypothesis can be applied to 'vicarious expectations' as well. This very special kind of expectations is what, in my view, allows apes to really go (Karg et al. 2015) beyond the old description –around 2003– of their ability to estimate what the other sees. In **5**, I propose that the difference between the primitive and advanced ToM rests on the contrast 'vicarious expectations' vs. 'foreign contents'. Section **6** focuses on self-conscious emotions, which require the ability of detecting foreign mental contents and are essential in the new lifestyle. In **7**, it is specified that the strict conditions for the very origin of 'foreign mental contents' are not necessary for the subsequent development of the advanced ToM. Section **8** proposes that the effective ('unified', I call it) reception of pointing gestures –also essential in the new lifestyle– similarly requires the detection of foreign mental contents. In **9**, I argue that the human white-of-eye is a strong facilitator of that detection. Finally, Section **10** deals with the falsifiability of the proposals of the article.

2) The old description of the two modes of ToM

For the authors that accepted ToM around 2003, the primitive mode was described as the ability to know what the other sees (/ does not see) –or has (/ has not) seen immediately before. This ability is possessed not only by children much younger than 4 years, but also (as Tomasello, Call, Hare, 2003 showed) by chimpanzees. These results, soon extended to goats or ravens (see e.g. Bugnyar & Heinrich 2005, and Bugnyar et al. 2016), were explained by a very simple mechanism: The subject both tracks a line from the other until the relevant object and is aware of the (possible) opaque barriers obstructing that straight line.

The advanced mode was linked to the ability of attributing 'false beliefs' to others. The early tests of 'false belief' show a video in which, for example, a child

(Maxi) puts his marble inside a vase and then leaves; afterwards, his mother puts the marble inside his toy box and leaves. Right then, Maxi comes back and the experimenter asks the children who have seen the video, ‘Where will Maxi look for his marble?’ The answers coming from children under 4 do not show the false belief which Maxi is bound to have, but their own knowledge. Within this general framework, the implicit knowledge of somebody else’s false beliefs (which was observed –Clements & Perner 1994– in many 3-year olds that gave the wrong explicit answer) did not seem to disturb the mentioned descriptions of the two modes of ToM.

3) Discussing some proposals about the difference between the primitive and advanced ToM

Since Onishi & Baillargeon 2005, numerous results in non-verbal tests have been offered in favor of the estimation of false-belief by infants. That type of tests was later applied to great apes, who achieved similar results –Krupenye et al. 2016, Kano et al. 2017. Certainly, the percentage rate of success in those cases is smaller, more variable than the rate obtained in verbal tests. However, in my view, it is undeniable that there are new data.

But, is the success of non-verbal tests based on the same resource which supports traditional tests? Before moving on to my proposal, let’s focus on some of the theories about that matter. Tomasello, Southgate, and (in a rather inverse direction) Lurz want to separate the new data from what is achieved in the advanced ToM. I am close to their goal, but not to their proposals.

According to Tomasello 2018, the infant grasps others’ beliefs because he “disregards his own (diverging) knowledge”. In my view, such formulation is not convincing, since disregarding the knowledge of the situation in which we find ourselves would be at any age a very inconvenient type of inattention. But it is also true

that, as Tomasello argues, if one's own mental content, instead of being disregarded, is simultaneously carried with somebody else's content in one's own mind, then the two contents must be distinguished and compared by the subject, and thus, we would be identifying the primitive mode with the advanced one –an identification which I am opposed to.

Let's look at other account, which, being relatively similar to Tomasello 2018, is more recent and elaborate. Southgate 2020 (who, unlike Tomasello, doesn't mention the experiments about 'foreign false beliefs' in apes) proposes that "human infants have an altercentric bias, which results from a combination of the value that human cognition places on others, and an absence of a competing self-perspective", and that such bias causes that the particular events that are not co-witnessed with the protagonist of the play are encoded with less strength. This is what explains, according to Southgate, infants' successes in non-verbal tests of false belief.

I will start by saying that I like the idea that "for infants, altercentrism is beneficial". However their tendency towards altercentrism can, without losing its attractiveness, be reformulated in a different way –infants very often produce 'vicarious expectations'. In addition, I reject the alleged "weakness of self-perspective", the same as Tomasello's proposal that the infant "disregards his own (diverging) knowledge". Note that typical perceptions are evolutionarily much older than altercentrism and used at any age much more frequently. Thus, it is unlikely that the degree of conservatism that evolution necessarily includes fails there.

Let's also focus on Lurz et al. 2022. This article –very different from Tomasello's or Southgate's ones— proposes that apes' success can be explained in "a simple way: Apes don't use metarepresentations, but they merely simulate (/imagine) to believe what the other agent believes". But note that this simulated belief or "low-level

simulation” (as say Lurz et al.) requires both to attend to two contents about the same thing, and to distinguish each from the other. Thus, this task, as implicit as it may be, is not “a really simpler model”, as these authors defend, but is still a metarepresentation.¹

4) Vicarious expectations sustain the primitive ToM

After having criticized these three articles, can we keep the idea of a qualitative difference between the primitive and advanced ToM?² In order to try to answer this question, let’s begin by attending to new experiments –Karg et al. 2015– which show that apes’ ability to estimate what the other sees (or does not see) goes well beyond the results offered in Tomasello et al. 2003.

Let us think of an ape that must estimate if his peer sees or does not see the object that he –the subject– has previously seen. But, at that moment, foliage prevents the *estimator-subject* from seeing the object. To solve this obstacle, he certainly could move. However, since apes –heavy and lacking wings– can often find themselves at different heights from each other, they would take too long to reach a location which would allow them to see the visual field of their peer. This is the situation that was reproduced by the new experiments.

How do apes succeed there? Probably they 1) activate their own expectations about what they would see if they were in the same location as their peer and 2) process such expectations as belonging to the observed peer. These would be expectations of a special –vicarious– kind. But, what is a vicarious expectation?

Let us begin by attending to expectation in general, which is a vital resource to guide behavior and learning in non-human animals. The matter is how expectations and

1. However, with Lurz et al., I agree that apes’ ability in those tests is related to “automatic affective empathy” (in my words –Bejarano 2022–, ‘vicarious expectations’ are related to ‘spontaneous altruism’.)
2. In Developmental Psychology, various authors (recent examples: Poulin-Dubois et al. 2022, or Barone et al. 2022) are offering data favorable to such difference.

goals act in non-human consciousness (and maybe in our most spontaneous mental processes too) while they are marked as absent. Probably, instead of invoking the idea that the animal agent has a mental representation (/simulation / evocation / off-line copy) of them, it would be helpful to understand such ‘presence of absent elements’ in a broader, less demanding mode. Therefore, while at a regrettably metaphoric level, we could speak of well-defined but empty profiles hierarchically arranged according to its greater or lesser degree of dependency on learning.

Okasha, 2022 (despite recognizing that “when we attribute a goal in some different biological contexts, such attributions merely reflect an anthropomorphic bias”) claims that ‘the mental representations of goal’ in avian and mammal species are objective facts, and he justifies such claim “on grounds of (their) evolutionary continuity and neurophysiological similarity (with humans)”. But I doubt that those “grounds” are enough of a guarantee, and suggest the following alternative. It was the new lifestyle –the new style of cooperation– what made the full representation of goals advantageous. Individuals then have to communicate their goals to their *group*, so that the group can cooperate towards reaching that goal.³ This need for producing and understanding such communications makes evocation (/simulation) of the goal appear. ‘Well-defined, empty profiles’, which were sufficient in the old lifestyle, no longer are.

So far, we have dealt with expectation in general. Can the “well-defined but empty profile” be applied to vicarious expectations? Such application seems plausible. Anyway, in order to favor the affirmative answer, I will try to show that these special expectations derive quite directly from a particular non-vicarious expectation.

3. Such communications would use increasingly ‘cultural’ gestures/calls, which however would still be unable to signal any differentiated element in the desired situation.

Let's focus on macaques' 'mirroring', and, firstly, on its origin. Hands, unlike the other parts of primate or non-primate bodies, are totally visible for their possessor, and, in addition, they must be observed during the actions of grasping. Thus, the sensory (proprioceptive/tactile) feedback of any grasping will end up being connected with the visual perception of that movement: Keysers & Perrett 2004. (From this –in my view, attractive– hypothesis, I deduce that the 'audio-vocal mirror-neurons of birds' cannot be mirror-neurons. Note that, while learning the dialect, the bird does not sing yet. Therefore, the externally perceived model is stored without any connection with inner sensations.) It is convenient to underline that, while visual / proprioceptive connection is forming in a macaque, it is still a non-vicarious expectation: It is the subject's grasping that activates in him –in the subject– the expectation of the two versions –visual and inner– of the adequate 'feedback'.

But, when the visual version is given without the corresponding inner sensations (i.e., when it is someone else's hand), the subject will have to disengage from himself the hand that is in sight. This is confirmed by the results of all the rubber-hand experiments (e.g., Pfister et al. 2022: "A single tactile stimulus applied to the rubber hand –but not to the real hand– triggers substantial and immediate disengagement".) Such disengagement does not only involve the hand at sight, but also the proprioceptive/tactile expectations which the observed individual would have activated in the subject, and which now needs to process as 'belonging to other'. It is then when vicarious expectations would arise. In short, while it is typically emphasised that "mirror-neurons map *other-related* information *onto self-related* brain structures" (Bonini et al. 2022, the most recent example), I underline the later, inverse mapping.

Certainly the vicarious expectations above attributed to apes concern the entire body, not only the hand. However, this is an irrelevant difference: Piaget 1945 shows

that it is starting from the hands how the child builds correspondences between his own body and other bodies. Therefore, if the hypothesis turns out to be correct, we could deduce the desired conclusion –i.e., vicarious expectations are directly derived from non-vicarious expectation.

What do I get from all these hypotheses? If vicarious expectations –instead of requiring imagined (/simulated /evoked /off-line) representations– are ‘well-defined but empty states’, then the subject has no need to differentiate vicarious expectations from his own contents. Such differentiation occurs automatically (i.e., without no type of metarepresentation), since contents are full states.

5) My proposal on the difference between the primitive and advanced ToM

Despite rejecting Tomasello’s (2018) idea that infants and apes “disregard their own diverging knowledge”, I accept that the union of “inattention to one’s own mental elements” and “attention to somebody else’s ones” characterises the primitive ToM. But I propose that such inattention and such attention take place, not at the content-level, but at the expectation-level. It is the subject’s own expectation that is disregarded when the subject activates vicarious expectations and encodes them as ‘belonging to other’.⁴ This is my description of the primitive ToM.

This proposal can explain why, in non-verbal tests of false belief, apes’ and infants’ successes are limited. Those tests mainly require vicarious expectations, i.e. an easy requirement. But the two scenes and the consequent demand on working-memory can often provoke in prelinguistic subjects (vs. Clements & Perner’s 3-year olds) a great difficulty. Note that developmentally —and probably evolutionarily— the reception of multiple-word messages causes a great expansion of working-memory.

4. Such disregarding can only take place when the subject is behaviourally inactive. This fact might be relevant regarding the limitations of spontaneous altruism (cf. below –8.3.1–, and previous note 1.)

Let us recapitulate. According to the proposal, the contrast ‘primitive vs. advanced ToM’ equals the contrast ‘(empty, easy) vicarious expectations vs. (full, more difficult) foreign contents’. But then, if vicarious expectations are efficient and useful, what made the estimation of foreign mental contents originally advantageous? I assume the following three points. First, in order to support the primitive mode of ToM, vicarious expectations are sufficient. Second, the nature of these implies that I can only have expectations of inner states which are intrinsically possible for me.⁵ Third, the state of interacting with me as with a different person cannot be a vicarious expectation of mine. From this hypothesis, we can deduce that the ability of detecting (/estimating) foreign contents arose when the ability of detecting foreign mental states that *involve oneself* became an advantageous one.

6) Self-conscious emotions

Thus, I propose that “the thinking what others think *of us*” (Darwin, 1872 about blush; my emphasis) necessarily requires the advanced mode of ToM. But, beyond blush, we can focus on ‘self-conscious emotions’ (/“self-other-conscious emotions”: Reddy, 2010) –embarrassment, shame, guilt and pride (Lewis, 2000). In these, originally at least, the subject needs to see both the facial expression of the other and his gaze on himself (on the subject). But the crucial feature of self-conscious emotions (which distinguishes them from fear, e.g.) is as mentioned above. Indeed, when we experience those emotions, the contents of the foreign mind become more real, more relevant for us than any other reality in our surroundings.⁶ In human brains, is there something characteristic

5. Obviously, any mammal has non-vicarious expectations about the behaviour of animals that are very different to him. But this does not clash at all with my second point.
6. Probably the understanding (to some degree, at least) of the similarity between one’s own interiority and foreign interiority arises together with the detection/estimation of foreign contents. (Bejarano 2022.)

of –and common to– the four? Hopefully this can be answered (affirmatively, I bet) soon.

Self-conscious emotions are extremely useful in ‘the new lifestyle’. This is a cooperative, communicative lifestyle. Consequently, the care of one’s own reputation became crucial: Leary, 2004; Sznycer, 2019.

Baumard et al. 2013, who focus on “competition to be chosen as a partner in cooperative ventures”, practically identifies the care of reputation with the habit of refraining from “blatantly selfish actions”.⁷ This refraining is certainly essential in the care of reputation. However, even in “cooperative ventures” other aspects are important –e.g., the reputation concerning good communicative abilities. In addition, beyond cooperative ventures, there are (see Crespi et al. 2022) other “arenas of runaway social selection” where reputation is equally crucial.

But let’s pay attention to a slightly different usefulness of self-conscious emotions, which has perhaps been underlined less. ‘The new lifestyle’ requires also the “deliberate practice” –Ericsson 2002; Rossano 2003– that is necessary to achieve any kind of cultural expertise. Here, self-conscious emotions intervene again. Experts arouse general admiration. (About the two types of admiration –for skill and for moral virtue–: Algoe & Haidt 2009. About admiration –vs. envy– for experts: Onu et al. 2016.) Therefore, experts experience pride: See Sznycer & Cohen 2021. And these attractive results can sustain, at least in some of the admirers, prolonged, effortful acquisitions of expertise. Model expertise, despite not influencing automatic imitation (Nevejans & Cracco 2022), can cause desire to acquire such expertise, and in that causality, admiration is more decisive than ‘prestige bias’ (Chellappoo 2020).

7. Baumard et al. really propose: “The best care” is “the *genuinely moral* habit”. But I shall not comment this here.

In short, self-conscious emotions sustain self-control, which is necessary in different aspects of the new lifestyle.⁸ Certainly, self-control is subsequently bolstered by ‘speech directed to oneself’, and it can become an intrapersonal process. However, self-conscious emotions were crucial for the enormous growth of self-control since apes to humans. This adaptive advantage was –I propose– linked to the origin of ‘the ability of estimating foreign contents’.

But two clarifications must be made. Firstly, the ‘thinking of foreign mental states which *involve oneself*’ is a requirement only for the origin of ‘the advanced ToM’, not for its later uses. Secondly, it is convenient to ask ourselves whether the advanced ToM could have another extremely basic function apart from the activation of self-conscious emotions.

7) The advanced ToM beyond its origin

The process of thinking of ‘foreign mental states which involve oneself’ is, in my view, a requirement only for the origin of the advanced ToM. In fact, I propose that, once the ability to think a ‘second line’ of contents becomes strong, the advanced ToM can carry complex functions which do not fulfil that requirement. Such functions sometimes use foreign but non-interactive contents, as in verbal false-belief tests, which involve –I borrow Dor’s 2016 words– “a non-dialogic capacity of mind-reading”.

Note that in those verbal tests, the communicative interaction, far from being between the subject who attributes the mental content and the ‘attributee’, is reduced to that which is established between child and experimenter. Regarding this feature of verbal tests (of false belief), Gallagher 2015 states: “Given the specific attraction of the

8. ‘Self-control’ (Shilton et al. 2020), or ‘self-domestication’? I can only say that the connotations of the term ‘self-domestication’ are less suitable for a capacity that, “even when it takes us to meekness, means the strength and power to use one’s energy for one’s purposes”: Roszak 2022. (This author uses, not “self-control”, but “fortitude” or “resilience”, but these are terms that I can’t use: They include morality, while, in my view, self-control is not necessarily moral.)

second-person interaction (versus third-person perspective), the saliency of the interaction with the experimenter takes precedence over the third-person task".⁹ Barone & Gomila 2020 elaborate that view: "Second-person attributions of false belief" (unlike third-person attributions –'The Ancients believed that the earth was flat', for example) "are transparent, extensional, nonpropositional and implicit". Nonetheless, according to my proposal, if "second-person attributions" go beyond vicarious expectations and really attribute a full content, then they (despite their greater ease¹⁰) are included within 'the advanced ToM'.

Other times, non-original functions of the advanced ToM use non-foreign contents. These contents are either the subject's beliefs which he no longer holds, or possible contents (in any of the senses of 'possible'). However the advanced, uniquely human ToM originally arose –according to my proposal– from a directly relational, interpersonal process, which requires neither language nor experience with narratives. Thus, the original 'estimation of foreign mental contents' is what cognitive archeologists Foley & Mirazón 2020 recommend to look for, namely, a "component attribute" (vs. 'compound concept').

9. Regarding first-person beliefs, if it is required that they possess the sense that habitually is activated in second- or third-person attributions ('believe that *p*' vs. 'know that *p*'), then we must say that originally, such first-person beliefs did not exist. In the beginning, for subjects, their non-outdated beliefs are (Phillips et al. 2021) just the reality. Therefore the concept of belief emerges in an interpersonal way. Likewise, the called 'animal meta-cognition in great apes' (summarized in Tomasello 2022) is not a judgement on one's own contents, but a hesitation about one's own expectations. The intrapersonal meta-cognition is a very late ability even in humans.
10. I accept this greater ease: Precisely, in my view, even pre-syntactic 'requests for a certain object' or 'calls to a certain individual' (which would use pre-words always linked to conative function and conative intonation) could reveal the speaker's false beliefs to the listener, and therefore could provoke the pre-grammatical (theme / rheme) syntax: See Bejarano 2011 (chapter 10), and Bejarano 2014.

8) Can the adequate reception of pointing be another extremely basic function of the advanced ToM?

The need to think foreign thoughts and, at the same time, realise that those thoughts are related to oneself –i.e., the need which has been detected in self-conscious emotions– can maybe be found in the effective reception of pointing gestures too.¹¹

8.1) A first obstacle: Pointing in apes

On the one hand, I have proposed that the advanced ToM is uniquely human. On the other hand, we know that many chimpanzees raised by humans have been taught to produce pointing gestures and to understand them (even the declarative type: Lyn et al., 2011). What answer can I give to this?

I will begin by admitting that similarly “human children display this ability to use communicative cues only after many months of intensive exposure to cultural environments characterized by frequent referential signalling, both verbally and nonverbally” (Clark et al. 2019). In addition, I admit that the absence of pointing is not at all harmful in “apes’ lifestyle”. From those statements, some authors conclude that in non-human primates that ability would be present, although scarcely exercised or developed. See Vasilieva, 2019: “Not only the presence /absence of a trait, but whether it manifests in animals to the same degree as in humans, is equally important for our understanding of trait evolution.” Heintz & Scott-Phillips (2022) offer the following analogy: “Human bodies are not especially well-suited to swing from trees. However, there is no absolute barrier.”

11. The influence of pointing gestures in ontogenesis of language (Southgate et al. 2007) can be exported to evolution. Pointing gestures, in my view, caused the intermediate level between the levels respectively focused by the previous notes 3 and 10. In other words, through pointing, the learning of meanings with which you can ask for a certain thing or call a certain individual becomes possible.

But, according to my proposal, it is only the really effective, ‘unified’ reception of pointing gestures that is uniquely human. Certainly, in this way, I place as vital criterion a process which is still unobservable, which may seem as a withdrawal towards “untestability with scientific methods” (Leavens, 2021). However, as it will be seen, the proposal is connected with some facts and with several potential experiments and research.

8.2) Authors who, when dealing with pointing in apes, have focused on reception.

The focus on reception is not new. Moore, 2013 focuses on the receptive failure of apes, and proposes that “since pointing gestures provide poor evidence for a speaker’s message, they exceed the pragmatic capacity of apes”. Likewise, Morrison 2020 emphasises the ambiguity and necessary disambiguation of pointing gestures. I agree with these claims. But my proposal is different.

Therefore, for me, Lyn & Christopher, 2018 is a more useful study. These authors list three conditions in which the experimenter may point out and whose reception by apes is differently successful: “i) Proximal-Proximal: The choice items are close together and the point is close to the correct item. ii) Proximal-Distal: The choice items are close together, but the point is further away. iii) Distal-Distal: The choice items are further apart, and the point is therefore necessarily further away.”

According to those authors, in Proximal-Proximal and also in Distal-Distal, point-following can be achieved by simple mechanisms. However, “in Proximal-Distal, the best predictor of success is ontogenetically previous human social contact”. I would underline the fact that it is just in Proximal-Distal where the direction of head of the producer (that is, the cue that chimpanzees use to estimate what others can see: Tomasello et al. 2007) is unable to signal the object.

8.3) Apes' and humans' reception of pointing gestures

8.3.1)

Before focusing on this contrast, it is convenient to go over spontaneous production in apes. Leavens et al., 2005: “Unlearned (i.e., with no explicit training whatsoever) captive chimpanzees frequently point to unreachable foods. These are communicative signals because apes will not reach towards obviously unreachable food if there is nobody around to see them do it”. In addition, in those chimpanzees repeated gaze-alternation between the food and the experimenter was significantly associated with their pointing gestures. Since then, Leavens and other authors began to ask themselves whether conditions similar to the ones (cage and benevolent recipient) which in those observations were considered as decisive appeared in wild chimpanzees too. Hobaiter et al., 2014 propose the following: “Wild chimpanzees experience few physical barriers, but the presence of a dominant, unrelated chimpanzee monopolizing a particular resource may be a greater barrier to a young chimpanzee’s access than bars on a cage. To overcome this challenge, a juvenile’s only resource is another chimpanzee, mainly its mother.” Thus, they found a case in the jungle which they classified as “possibly deictic”. The possible conclusion from all this is that wild chimpanzees can use this type of production with their conspecifics to achieve their goals.

Nevertheless, in order for that production to be a useful resource in the wild, it is necessary for recipients to understand (at least occasionally) the desire of the producer and to deliver, selflessly, the desired object. Animal altruism has been much discussed: e.g., Rendall et al. (2009) versus De Waal (2009). But I do not discard it, as long as it

does not cross the limits of the (always narrow) ‘spontaneous altruism’.¹² Thus, let’s focus on recipient-chimpanzee’s understanding of the producer’s desire. Is that understanding equivalent to our understanding of pointing gestures?

8.3.2)

Before giving an answer, let’s attend to yet another issue. Heintz & Scott-Phillips (2022b) presented a general framework of intentional communicative production in humans and apes. Human peculiarity, according to these authors, would begin with “Gricean communication”. Grice, 1957: “The producer says *p* intending 1) that the recipients believe *p* or do *p* and 2) that the recipients recognize his/her intention (1)”. Or, in Heintz & Scott-Phillips words, “Gricean communication is characterised by intentional manipulation of attention towards one’s own *informative* (declarative or imperative) intentions”.

Regarding production, I would like to discuss the decision by Heintz & Scott-Phillips to place apes outside “Gricean communication”. Remember that in Leavens’ “untrained chimpanzees”, pointing gestures were significantly associated with gazes towards the addressee. With that gaze, chimpanzees likely wanted to show their informative intentions. More in general (and with a clarification on Grice himself), intentionally communicative producers need to perform an additional process to hide, not to show, their informative intentions.¹³

12. About ‘spontaneous altruism’: Tomasello 2012, Rand et al. 2012, and, especially, “self-other merging” (Miyazono & Inarimori 2021). About the particular (probably, more primitive) type of altruism that, “connected to reactive, non-cognitive fear circuits, helps others under threat” (for instance, in social hunt): Vieira et al., 2020, Vieira & Olsson, 2022.
13. So, when Grice proposes his “2”, it is just his example (handkerchief, murder) of “absence of 2” (or “sophisticated hidden authorship”: Moore, 2015) that really persuades us. (Cf. Geurts, 2019.)

Other article –Warren & Call, 2022– also lists the evolutionarily different types of communication, but mainly focuses on “inferential communication, where individuals can spontaneously invent gestures that others might be able to comprehend”, and suggests (see also Tomasello & Call, 2019) that this is the communication which apes could sometimes reach. Warren & Call acknowledge: “Whether apes can use inference in communicative situations is unclear.” But the imagined situations that they describe will undoubtedly inspire experiments.

8.3.3)

Focusing finally on the reception of pointing gestures in chimpanzees, I begin by highlighting that the communicative value of ‘gazes towards the addressee’ is understood. Indeed “the sensitivity to being watched is both innate and shared by most vertebrates” (Klein et al. 2009). Thus, in the species that are able to perform ‘recipient-directed’ communication, recipients of that gaze understand that they are the addressees of this innate communicative resource.¹⁴

However, in the chimpanzee-recipient such communicative value is not applied –this is my proposal– to the other element produced by Leavens’ untrained chimpanzees, that is, to the combination ‘gaze towards the object and hand/arm movements’. It is fair to specify up to which point this description of non-human reception of pointing gestures seems implausible to human intuition. The producer, both before and after making movements in a certain direction with his arm and head, communicates with the recipient by means of eye-contact. Why would the recipient not understand that the producer’s movements are communicative, or, in other words, that the communicative value of eye-contact is applied to those movements and give them

14. In chimpanzees, eye-contact is a friendly communicative resource; in gorillas, it communicates mild threat.

an informative (imperative or declarative) function? For humans, that *unification of the two consecutive instants* is obliged and unstoppable, I acknowledge it. But I hold on to the idea that the confidence in our intuition must be ready to withdraw if there are reasons for it, or, more concretely, if the anti-intuitive proposal can explain more facts.

To start with, let's note that the cage (Leavens 2005) or the dominating individual (Hobaiter et al. 2014) make the chimpanzee's gesture non-absurd for conspecifics even if it is not interpreted as communicative. On the contrary, our human pointing gestures can be considered closer to communicative pantomimes. Tomasello, 2008 stresses how strange any pantomime can be for a recipient if the gestures involved are not interpreted as being communicative ("the recipient will see my iconic gestures as some kind of strangely misplaced instrumental action"), but he never says anything of the sort about pointing. However in my view, there is some similarity between the two cases.¹⁵.

I have already explained my proposal on the key difference of human (versus non-human) reception of pointing gestures.¹⁶ But, in order to elaborate further on that difference, it is convenient to underline that the alternation of gaze between the object and the addressee is a vital feature in the reception of pointing. Paulus & Fikkert 2013 show that the necessary and sufficient element for infants to first understand pointing

15. According to Tomasello & Call, 2019, "attention-getters, since they manipulate attention of addressees, evolutionarily precede pointing gestures, while intention-movements, since they manipulate the imagination, were transformed into pantomiming". I agree with such difference, but my interest is now in the mentioned similarity.

16. What about the dogs? "Eye-contact is the major cue that dogs use to determine when human pointing is intended for them": (Kaminski et al. 2012; Téglás et al., 2012). The original basis is "social hunt" (Zuberbühler, 2008). The dominant wolf must decide on one particular prey individual and signal it. This communication is likely realised with an innate gesture that pre-activates in the members of the herd a plan of attack against the signalled prey. So, subordinate wolves will start executing that plan when some dominant one makes eye-contact with them. In short, I am suggesting here, again, a non-unified reception –a reception of two separated, innate signals. In addition, of course, "sensitivity to human gestures of pointing was selected by breeders in domesticated dogs" (Hare et al. 2002). Likewise, ontogeny must be taken into consideration (Clark et al. 2019).

gestures is not the hand-movement (or its situational / cultural variations: Cooperrider & Slotta, 2018), but the alternation between ‘gaze towards the object’ and ‘gaze towards the recipient’.

On the one hand, the ‘gaze towards the object’ causes the recipient to estimate what the producer sees. On the other hand, the ‘gaze towards the recipient’ informs the recipient that he is being the addressee. (In the beginning –when infants first understand pointing, and likely also in evolution– the ability to understand pointing is just the ability to understand the pointing gestures that are addressed to oneself. “The basic arena for social interaction is the dyad” –Clark, 1996.) Inter-brain consequences of eye-contact are increasingly studied. Di Bernardi Luft et al., 2022: “Inter-brain synchronization mainly flows from leader to follower”, and thus –I add– from the producer of pointing gestures to the addressee. Or, more focused on teaching, Pan et al., 2020. In general, second-person approaches underline eye-contact: Cañigueral et al., 2022. (Eye-contact is also crucial –as indicated above– in the origin of self-conscious emotions.)

But those two instants (‘gaze towards the addressee’ and ‘gaze towards the object’) cannot in any way remain separate, but they must be unified. The recipient does not only have to detect what the producer is looking at, but he must simultaneously understand that what the producer is looking for by looking at the object is to point at the object for him, for the recipient. In other words, the recipient must understand that the innate communicative function of gazes towards him can be applied to gazes or movements that are not innately communicative.

According to my proposal, it is in that unification where the problem arises for apes. Let’s see it. The direction of the producer’s head and hand towards the object is what allows recipient-chimpanzees to activate ‘vicarious expectations’ and attribute

them to the producer. However, because there can never be any kind of expectation of the results of an action intrinsically impossible for the subject, the recipient will be unable to apply to those vicarious expectations an interpersonal communicative function to himself.

Therefore, the unified, fully adequate reception will only be possible by the estimation of mental full contents of the producer. Thus, there would be a common ability to that reception and self-conscious emotions (and also to linguistic reception¹⁷). That ability could be described as the one of ‘remaining in your shoes when you look at me’.

Of course, before the unified, fully effective reception, several behaviors (similar to the one carried out by Leavens’ untrained chimpanzees) could achieve some degree of reception, and could be useful for both producer and recipient. Let’s consider, for instance, the action of pushing a conspecific until we place him so that he can see the relevant object. This type of communicative behaviour would have been multiplied in the beginning of the ‘new, cooperative lifestyle’, without the recipient grasping the communicative function of the pushing yet. But this quasi-reception finally became accessible to natural selection (and ‘coevolution genes / culture’). And so, the effective, unified reception appeared, together with the detection of foreign contents.

Of course, before the unified, fully effective reception, several behaviors (similar to the one carried out by Leavens’ untrained chimpanzees) could achieve some degree of reception, and be useful for both producer and recipient. Let’s consider, for instance, the action of pushing a conspecific until we place him so that he can see the relevant object. This type of communicative behaviour would have been multiplied in

17. *** 2011, chapter 6: My argumentation started by focusing on the reception (also studied by Fernandez-Rubio 2021) of the most egocentric deictics.

the beginning of the ‘new, cooperative lifestyle’, without the recipient grasping the communicative function of the pushing yet. But this quasi-reception finally became accessible to natural selection and to ‘coevolution genes / culture’.¹⁸ And so, the effective, unified reception appeared, together with the detection of foreign contents.

8.4) The great obstacle: Developmental asynchrony

Now it is urgent to meet the following objection to my proposed link ‘human reception of pointing gestures’ / ‘advanced Theory-of-Mind’. Why do then 3-year-old children fail in verbal tests of false belief even if they already have enough mastery of language? Why do those children, who have been adequately receiving pointing gestures for two years, fail?

Let’s remember what was proposed above (section 4). In the reception of pointing gestures, the detection/estimation carried out by the recipient –i.e., the detection of which is the object which the producer looked at– remains involved in the reception of a communication that is directed to him, the recipient. This is the original situation that causes the bankruptcy of vicarious expectations.

However in verbal tests of false belief, that does not happen: There, communicative interaction, far from being between the recipient and the same individual for whom the recipient has made the estimation, is reduced to that which is established between child and experimenter. Thus, when these children receive the experimenter’s question, they must grasp the content of that speech that is addressed to them (therefore, a foreign mental content), and so, they find difficult the task of almost simultaneously focusing on the mental content of the protagonist of the ‘play’ as well.

18. Regarding Theory-of-Mind, Heyes 2018 (see also Moore 2021) emphasizes its learning above all. (As Tramacere & Mafessoni 2022 say, “she seems to disregard coevolution learning / genes”). I propose however that *the detection of foreign contents* involves genetic novelties that arose in the new lifestyle.

This is why, although in their reception of pointing gestures those children are able to detect ‘foreign contents’, they cannot use their ability in the verbal tests.

But there is an additional cause for that asynchrony. The unified reception of pointing gestures is strongly facilitated by a new, genetic resource. Let us see it.

9) The human white sclera and the unified reception of pointing gestures use

Kobayashi & Koshima, 2001 focused on the universally human white sclera, or, more precisely, on both its horizontal enlargement and its depigmentation, and proposed that these human peculiarities enhance “the visibility of eye-gaze orientation”. However, gaze-following –a phylogenetically old ability– is carried out without the help of the white-of-eye. Indeed, Tomasello et al. 2007 showed in apes the reliance on head (vs. eyes) in gaze-following. In a similar way, Chris Moore 2008 concluded from his experiments that when infants first start to follow gaze (at that age –note, please– they are still unable to receive pointing gestures), “they do so on the basis of head direction, not eye direction”.

Tomasello et al. 2007, putting ‘the enhancement of the visibility of eye-orientation’ in the evolutionary context of human cooperativeness, hypothesised that humans evolved such unique eye morphology to facilitate joint attentional and communicative interactions among conspecifics. (See also Wolf et al. 2023.) Yáñez & Gomila 2018, who also underline “the interactional importance of eye-direction”, add something very interesting: “especially when oneself is the focus of that attention”. I will specify this emphasis on cooperation and interaction in order to connect it with my proposal of the ‘unified’, effective reception of pointing gestures. Let’s start by redescribing “the enhanced visibility”.

Mayhew & Gómez 2015, Perea-García et al. 2019 (but see Mearing & Koops 2021) and Caspar et al. 2021 have proposed that the human eye is not necessarily

unique among ape species. But let's focus on horizontal elongation. This feature may have evolved to allow non-arboreal primates to scan their environment widely. Despite that, such elongation together with the universal "totally/bilaterally white sclera" make the location of the iris conspicuous not only in averted but also in direct gaze. In addition, "the eye-outline is easier to see in humans (than in apes) irrespective of skin color" (Kano et al., 2021), and this makes the location of the iris even more conspicuous. Thus, human eyes –this is my point– make *the successive locations (the horizontal travelling)* of the iris conspicuous.

In this way, the visibility of the crucial (remember Paulus & Fikkert 2013) *alternation* between gazes is really enhanced. It may be suggested that, when the producer moves his iris from the 'gaze towards the object' to the 'gaze towards the recipient', that movement is perceived by human recipients like if it was injecting the 'gaze towards the object' –and, consequently, also the result of the estimation carried out by recipients– into the 'gaze towards the recipient', that is, into the communication. In this way, the human eye leads the recipient of pointing gestures to unify the two instants, and, therefore, to detect/estimate the producers' mental states that involve himself –i.e., the recipient– as their addressee, and, thus, to estimate 'foreign (full) contents' and not only 'vicarious (empty) expectations'.

Human sclera is an *anatomical* 'facilitator resource' of an advanced mental process. It is also a *strong* one. These qualifications reinforce the suspicion that the 'unified' reception of pointing gestures, as self-conscious emotions, was a very early function of the detection of foreign mental contents.

Probably, the depigmented sclera could become universal in an evolutionarily very short time, and therefore could emerge in the same species in which the adequate

reception of pointing gestures was beginning to emerge. But, did it happen in Sapiens?¹⁹ Or in Neanderthals / Denisovans? Or in earlier species? According to my proposal, this is an absolutely crucial question. I hope that Paleogenomics and Genomics specialists will answer it soon. (Certainly the depigmented sclera is a very simple feature. However its universality makes, of course, the task difficult.)

I propose that this very early detection of foreign mental contents and its strong, anatomical facilitator are the decisive, evolutionarily latest basis for our unique abilities. But my proposal can accept either that such basis emerged in Sapiens, or that, on the contrary, in Sapiens, only derivations (see above, section 4) of that detection emerged, while the detection itself had emerged in Neanderthals. I only predict that our white of eye will not be found in earlier hominins. Therefore, while as long as limitations of apes and early hominins are concerned, my proposal possesses strict falsifiability, in this issue, nevertheless, it unfortunately possesses a more blurred one.

10) Summarizing, and looking towards the future

This article has hypothesised that the transition from ‘vicarious expectations’ to ‘foreign mental contents’ can be described as a genomic novelty that appeared in the evolution (rather ‘coevolution genes / culture’) with the ‘new, human lifestyle’: In this lifestyle, self-conscious emotions and the effective reception of pointing, which require the ability of estimating foreign contents, are crucial resources.

19. This is not at all an absurd suggestion. Firstly, within the lineage of Sapiens and even in dates totally within the so-called ‘anatomically modern humans’, there is a marked evolution in the shape of the cranium: See Neubauer et al. 2018 (although, at least since 160.000 b. p., differences with living humans would mainly affect, according to Zollikofer et al. 2022, the face and cranial base). Secondly, regarding our absence of very prominent browbridges – which are present in Neanderthals–, Godinho et al. 2018 reject the old hypotheses on the function of such absence, and suggest “its potential role in social communication”. (Siposova et al. 2018 underline the role of raised, highly mobile eyebrows in ‘the reception of communicative looks’.).

No new empirical result has been offered here. However, the main proposal and each subproposal raise questions: My view of expectations?; apes' vicarious expectations?; the anti-intuitive 'non-unified reception of pointing' in chimpanzees?, etc. Those questions can lead to different experiments and to research in Neuroscience or Paleogenetics, whose results will have an impact on my proposal, in one way or another. But this has already been discussed above. Therefore, I will add only a more personal comment.

I am really looking forward to those results. Even if they discarded my proposals, I would feel that my effort has been useful. However, I cannot obtain such results. I can only –and this is what I do– request them.

References

Algoe, Sara & Haidt, Jonathan (2009). Witnessing excellence in action: The “other-praising” emotions of elevation, gratitude, and admiration. *The Journal of Positive Psychology*, 4(2), 105–127.
<https://doi.org/10.1080/17439760802650519>

Apperly, Ian & Butterfill, Stephen (2009). Do humans have two systems to track beliefs and belief-like states? *Psychological Review*. 116 (4): 953–970. <https://doi.org/10.1037/a0016923>

Bach, Patric; Frank, Cornelia & Kunde, Wilfried (preprint). Why motor imagery isn't really motoric: Towards a reconceptualization in terms of effect-based action control. <http://doi.org/10.31234/osf.io/bkf7>

Barone, Pamela & Gomila, Antoni (2019). Infants' performance in the indirect false belief tasks: A second-person interpretation. *Cognitive science*. <https://doi.org/10.1002/wcs.1551>

Barone, Pamela; Wenzel, Lisa; Proft, Marina & Racoczy, Hannes (2022). Do young children track other's beliefs, or merely their perceptual access? An interactive, anticipatory measure of early theory of mind. *Royal Society Open Science*. <https://royalsocietypublishing.org/doi/10.1098/rsos.211278>

Baumard, Nicolas; André, Jean-Baptiste; Sperber, Dan (2013). A Mutualistic Approach to Morality. The Evolution of Fairness by Partner Choice, *Behavioral and Brain Sciences*, 36, 59-78.
<https://doi.org/10.1017/S0140525X11002202>

Bejarano, Teresa (2011). *Becoming Human: From pointing gestures to syntax*. Benjamins.

Bejarano, Teresa (2014). From *Holophrase to Syntax: Intonation and the Victory of Voice over Gesture, Humana.Mente. Journal of Philosophical Studies* 27, 21-37.
<https://www.humanamente.eu/index.php/HM/article/view/95>

Bejarano, Teresa (2022). The Most Demanding Moral Capacity: Could Evolution Provide Any Base? *Isidorianum*, 31(2), 91-126.
<https://www.sanisidoro.net/publicaciones/index.php/isidorianum/article/view/Bejarano>

Bonini, Cristina; Rotunno, Cristina; Arcuri, Edoardo & Gallese, Vittorio (2023). The mirror mechanism: linking perception and social interaction. *Trends in Cognitive Sciences*. [10.1016/j.tics.2022.12.010](https://doi.org/10.1016/j.tics.2022.12.010)

Bugnyar, Thomas & Heinrich, Bernd (2005). Ravens differentiate between knowledgeable and ignorant competitors. *Proceedings of the Royal Society B*, 272, 1641–1646. <https://doi.org/10.1098/rspb.2005.3144>

Bugnyar, Thomas; Reber, Stephan & Buckner, Cameron (2016). Ravens attribute visual access to unseen competitors. *Nature Communications*. <https://www.nature.com/articles/ncomms10506>

Cañigueral, Roser; Krishnan-Barman, Sujatha; Hamilton, Antonia (2022). Social signalling as a framework for second-person neuroscience. *Psychonomic Bulletin & Review*.
<https://doi.org/10.3758/s13423-022-02103-2>

Chellappoo, Azita (2021). Rethinking Prestige Bias. *Synthese*, 198, 8191–8212.
<https://doi.org/10.1007/s11229-020-02565-8>

Clark, Hannah; Elsherif, Mahmoud & Leavens, David (2019). Ontogeny vs. phylogeny in primate/canid comparisons: a metaanalysis of the object choice task. *Neuroscience and biobehavioral reviews*, 105, 178-189. <https://doi.org/10.1016/j.neubiorev.2019.06.001>

Clark, Herbert. (1996). *Using language*. Cambridge U. P.

Cooperider, Kensy & Slotta, James (2018). The Preference for Pointing With the Hand Is Not Universal. *Cognitive Science* 42(1). <http://dx.doi.org/10.1111/cogs.12585>

Crespi, Bernard; Flinn, Mark; Summers, Kyle (2022). Runaway Social Selection in Human Evolution. *Frontiers in Ecology and Evolution*. <https://doi.org/10.3389/fevo.2022.894506>

Darwin, Charles (1872). *The Expression of the Emotions in Man and Animals*. London, John Murray, 1872.

De Waal, Frans (2010). *The Age of Empathy*. Three Rivers Press.

Di Bernardi Luft, Caroline *et alii* (2022). Social synchronization of brain activity increases during eye-contact. *Communications Biology* 5(1). <http://dx.doi.org/10.1038/s42003-022-03352-6>

Dor, Daniel (2016). From experience to imagination: Language and its evolution as a social communication technology, *Journal of Neurolinguistics*.
<http://dx.doi.org/10.1016/j.jneuroling.2016.10.003>

Ericsson, K. Anders (2002). Attaining excellence through deliberate practice: insights from the study of expert performance. In M. Ferrari (Ed.), *The pursuit of excellence through education*, Erlbaum, Mahwah, NJ (pp. 21-56).

Fernandez-Rubio, Paula (2021). Pragmatic markers: The missing link between language and Theory of Mind. *Synthese*, 199. <https://doi.org/10.1007/s11229-020-02768-z>

Foley, Robert & Mirazón Lahr, Marta (2020). Variable Cognition in the Evolution of Homo: Biology and Behaviour in the African Middle Stone Age. In book: *Landscapes of Human Evolution* (pp. 125-141).

Gallagher, Shaun (2015). The Problem with 3-Year-Olds. *Journal of Consciousness Studies*, 22, 160-182.

Geurts, Bart (2019). What's wrong with Gricean pragmatics. <https://doi.org/10.36505/ExLing-2019/10/0001/000363>

Godinho, Ricardo; Spikins, Penny & O'Higgins, Paul (2018). Supraorbital morphology and social dynamics in human evolution. *Nature (Ecology & Evolution)* <https://doi.org/10.1038/s41559-018-0528-0>

Grice, Paul (1957). Meaning. *The Philosophical Review*, 66, 377–88.
<https://semantics.uchicago.edu/kennedy/classes/f09/semprag1/grice57.pdf>

Hare, Brian; Brown, Michelle; Williamson, Christina & Tomasello, Michael (2002). The domestication of social cognition in dogs. *Science*, 298, 1634–1636. <https://doi.org/10.1126/science.1072702>

Heyes, Cecilia (2018). *Cognitive Gadgets: The Cultural Evolution of Thinking*. Harvard UP.

Hobaiter, Catherine; Leavens, David & Byrne, Richard (2014). Deictic gesturing in wild chimpanzees? *Journal of Comparative Psychology*, 128, 82-87. <https://doi.org/10.1037/a0033757>

Kaminski, Juliane; Schulz, Linda & Tomasello, Michael (2012). How dogs know when communication is intended for them. *Developmental Science*, 15(2), 222–232. <https://doi.org/10.1111/j.1467-7687.2011.01120.x>

Kano, Fumihiro; Furuichi, Takeshi; Hashimoto, Chie; Krupenye, Christopher; Leinwand, Jesse; Hopper, Lydia; Martin, Christopher; Otsuka, Ryoma & Tajima, Tomoyuki. (2022). What is unique about the human eye? Comparative image analysis on the external eye morphology of human and nonhuman great apes. <https://doi.org/10.1101/2021.09.21.461196>

Kano, Fumihiro; Krupenye, Christopher; Hirata, Satoshi; Call, Josep & Tomasello, Michael (2017).

Submentalizing Cannot Explain Belief-Based Action Anticipation in Apes. *Trends in cognitive sciences*, 21(9), 633-634. <https://doi.org/10.1016/j.tics.2017.06.011>

Karg, Katia; Schmelz, Martin; Call, Josep; Tomasello, Michael (2015). The goggles experiment: can chimpanzees use self-experience to infer what a competitor can see? *Animal Behavior*, 105, 211–221. <https://doi.org/10.1016/j.anbehav.2015.04.028>

Keysers, Christian & Perrett, David (2004). Demystifying social cognition: a Hebbian perspective. *Trends in cognitive sciences*, 8, 501-507. <http://dx.doi.org/10.1016/j.tics.2004.09.005>

Klein, Jeffrey; Shepherd, Stephen; Platt, Michael (2009). Social attention and the brain. *Current Biology*, 19. <https://doi.org/10.1016/j.cub.2009.08.010>

Kobayashi, Hiromi & Koshima, Shiro (2001). Unique morphology of the human eye and its adaptive meaning. *Journal of Human Evolution*, 40, 419–435. <https://doi.org/10.1006/jhev.2001.0468>

Krupenye, Christopher; Kano, Fumihiro; Hirata, Satoshi; Call, Josep & Tomasello, Michael (2016). Great apes anticipate that other individuals will act according to false beliefs. *Science*. [10.1126/science.aaf8110](https://doi.org/10.1126/science.aaf8110)

Leary, Mark (2004). The sociometer. In R. Baumeister & K. Vohs (eds.), *Handbook of self-regulation* (pp. 373–391). Guilford.

Leavens, David (2021). The Referential Problem Space revisited: An ecological hypothesis of the evolutionary and developmental origins of pointing. *Cognitive science*. <https://doi.org/10.1002/wcs.1554>

Leavens, David; Hopkins, William & Bard, Kim. (2005). Understanding the point of chimpanzee. Epigenesis and Ecological Validity. *Current Directions in Psychological Science* 14(4), 185-189. <https://doi.org/10.1111/j.0963-7214.2005.00361.x>

Lewis, Michael (2000). The emergence of human emotions. In M. Lewis & J. Haviland-Jones (eds.), *Handbook of emotions* (pp. 265–280). Guilford.

Lurz, Robert; Krachun Carla; Marenco, Mary Catherine; Hopkins, William (2022). Do Chimpanzees Predict Others' Behavior by Simulating Their Beliefs? *Animal Behavior and Cognition* 9, 153-175. [10.26451/abc.09.02.01.2022](https://doi.org/10.26451/abc.09.02.01.2022)

Lyn, Heidi; Greenfield, Patricia; Savage-Rumbaugh, Sue; Gillespie-Lynch, Kristen & Hopkins, William (2011). Nonhuman Primates do Declare! A Comparison of Declarative Symbol and Gesture Use in Children, Bonobos, and Chimpanzees. *Language & communication*. <https://doi.org/10.1016/j.langcom.2010.11.001>

Lyn, Heidi & Christopher, Jennie (2018). A point is not a point is not a point: Reinterpreting three basic kinds of pointing comprehension. *Proceedings of Evolang 2018* (pp. 260-263).

https://pure.mpg.de/rest/items/item_3190925_17/component/file_3260022/content

Mayhew, Jessica & Gómez, Juan Carlos (2015). Gorillas with White Sclera. *American Journal of Primatology*, 77, 869–87. <https://doi.org/10.1002/ajp.22411>

Mearing, Alex; Koops, Kathelijne (2021). Quantifying gaze conspicuousness: Are humans distinct from chimpanzees and bonobos? *Journal of Human Evolution*. <https://doi.org/10.1016/j.jhevol.2021.103043>

Miyazono, Kengo & Inarimori, Kiichi (2021). Empathy, Altruism, and Group Identification. *Frontiers in Psychology*, 12. <https://doi.org/10.3389/fpsyg.2021.749315>.

Moore, Chris (2008). The Development of Gaze Following. *Child Development Perspectives*, 66-70.

<https://doi.org/10.1111/j.1750-8606.2008.00052.x>

Moore, Richard (2013). Evidence and Interpretation in Great Ape Gestural Communication. *HUMANA.MENTE Journal of Philosophical Studies*, 6, 27-51.

Moore, Richard (2015). A Common Intentional Framework for Ape and Human Communication. *Current Anthropology* 56(1), 56-80

Moore, Richard (2021). The cultural evolution of mind-modelling. *Synthese*, 199(1), 1751–1776.

<https://doi.org/10.1007/s11229-020-02853-3>

Morrison, Donald (2020). Disambiguated Indexical Pointing as a Tipping Point for the Explosive Emergence of Language Among Human Ancestors. *Biological Theory*, 15, 196–211.

<https://doi.org/10.1007/s13752-020-00355-6>

Nagel, Ernest (1977). Teleology revisited: goal-directed process in biology. *Journal of Philosophy* 74(5): 261-279.

Neubauer, Simon; Hublin, Jean-Jacques & Gunz, Philipp (2018). The evolution of modern human brain shape. *Science Advances*, 4. <http://doi.org/10.1126/sciadv.aoa5961>

Nevejans, Maura & Cracco, Emiel (2022). Model expertise does not influence automatic imitation.

Experimental Brain Research, 240(4), 1267-1277. <http://doi.org/10.1007/s00221-022-06338-2>

Okasha, Samir (2022). Goal Attributions in Biology: Objective Fact, Anthropomorphic Bias, or Valuable Heuristic? [*Okasha, Teleconomy Vienna series FINAL.pdf](#)

Onishi, Kristine & Baillargeon, Renée (2005). Do 15-Month-Old Infants Understand False Beliefs?.

Science, 308, 255-258. [10.1126/science.1107621](https://doi.org/10.1126/science.1107621)

Onu, Diana; Kessler, Thomas & Smith, Joanne (2016). Admiration: A conceptual review of the knowns and unknowns. *Emotion Review*, 8. <https://doi.org/10.1177/1754073915610438>

Pan, Yeng et al. (2020). Instructor-learner brain coupling discriminates between instructional approaches and predicts learning. *NeuroImage*, 211. <https://doi.org/10.1016/j.neuroimage.2020.116657>

Paulus, Markus & Fikkert, Paula (2013). Conflicting Social Cues: Infants' Reliance on Gaze and Pointing Cues in Word Learning. *Journal of Cognition and Development*, 15, 43-59.
<http://doi.org/10.1080/15248372.2012.698435>

Pfister, Roland; Klaffehn, Annika; Kalckert, Andreas; Kunde, Wilfried & Dignath, David (2021). How to lose a hand: Sensory updating drives disembodyment. *Psychonomic Bulletin & Review*, 28, 827-833.
<https://link.springer.com/article/10.3758/s13423-020-01854-0>

Phillips, Jonathan, et al. (2021). Knowledge before belief. *Behavioral and Brain Sciences*. DOI: [10.1017/S0140525X20000618](https://doi.org/10.1017/S0140525X20000618)

Piaget, Jean (1954). *La formation du symbole chez l'enfant*. Neuchâtel. Delachaux & Niestlé.

Poulin-Dubois, Diane; Goldman, Elizabeth; Meltzer, Alexandra & Psaradellis, Elaine (2023). Discontinuity from implicit to explicit theory of mind from infancy to preschool age. *Cognitive Development*, 65. <http://doi.org/10.1016/j.cogdev.2022.101273>

Rand, David; Greene, Joshua & Nowak, Martin (2012). Spontaneous Giving and Calculated Greed, *Nature* 489, 427-430. <http://dx.doi.org/10.1038/nature11467>

Reddy, Vasudevi (2010). How Infants Know Minds. Harvard University Press.

Rendall, Drew; Owren, Michael & Ryan, Michael (2009). What do animal signals mean? *Animal Behaviour*, 78, 233–240. <http://dx.doi.org/10.1016/j.anbehav.2009.06.007>

Rossano, Matt (2003). Expertise and the evolution of consciousness. *Cognition*, 89, 207-236.
[http://dx.doi.org/10.1016/S0010-0277\(03\)00120-3](http://dx.doi.org/10.1016/S0010-0277(03)00120-3)

Roszak, Piotr (2022). Not Only Coping: Resilience and Its Sources from a Thomistic Perspective. *Journal of Religion and Health*. <http://dx.doi.org/10.1007/s10943-022-01711-5>

Scott-Phillips, Thom & Heintz, Christophe (2022). Expression unleashed: The evolutionary & cognitive foundations of human communication. *Behavioral and Brain Sciences*.
<http://dx.doi.org/10.1017/S0140525X22000012>

Scott-Phillips, Thom & Heintz, Christophe (preprint). Animal communication in linguistic perspective.
<http://dx.doi.org/10.31234/osf.io/vnc5m>

Shilton, Dor; Breski, Mati; Dor, Daniel & Jablonka, Eva (2020). Human Social Evolution: Self-Domestication or Self-Control? *Frontiers in Psychology* 11. <http://dx.doi.org/10.3389/fpsyg.2020.00134>

Siposova, Barbora; Tomasello, Michael & Carpenter, Malinda (2018). Communicative eye contact signals a commitment to cooperate for young children. *Cognition*, 179, 192-201. DOI: <https://doi.org/10.1016/j.cognition.2018.06.010>

Southgate, Victoria (2020). Are infants altercentric? The other and the self in early social cognition. *Psychological Review*, 127(4), 505–523. <https://doi.org/10.1037/rev0000182>

Southgate, Victoria; van Maanen, Catharine & Csibra, Gergely (2007). Infant Pointing: Communication to Cooperate or Communication to Learn? *Child Development* 78(3), 735-40. <https://srcd.onlinelibrary.wiley.com/doi/10.1111/j.1467-8624.2007.01028.x>

Sznycer, Daniel (2019). Forms and Functions of the Self-Conscious Emotions. *Trends in Cognitive Sciences* 23(2). <https://doi.org/10.1016/j.tics.2018.11.007>

Sznycer, Daniel & Cohen, Adam (2021). How pride works. *Evolutionary Human Sciences*, 3, 1-39. <https://doi.org/10.1017/ehs.2021.6>

Teglas, Ernö; Gergely, Anna; Kupan, Krisztina; Miklosi, Adam & Topal, Jozsef (2012). Dogs' gaze following is tuned to human communicative signals. *Current Biology*, 22(3), 209–212. <http://dx.doi.org/10.1016/j.cub.2011.12.018>

Tomasello, Michael (2008). *Origins of human communication*. MIT Press.

Tomasello, Michael (2012). Why be nice? Better not think about it. *Trends in Cognitive Sciences* 16, 580-581: <https://doi.org/10.1016/j.tics.2012.10.006>.

Tomasello, Michael (2018). How Children Come to Understand False Beliefs: A Shared Intentionality Account. *Proceedings of the National Academy of Sciences* 115, 8491-8498. <https://doi.org/10.1073/pnas.1804761115>.

Tomasello, Michael (2022). Social cognition and metacognition in great apes: a theory. *Animal Cognition*. <https://doi.org/10.1007/s10071-022-01662-0>

Tomasello, Michael; Call, Josep; Nagell, Katherine; Olguin, Raquel & Carpenter, Malinda (1994). The learning and use of gestural signals by young chimpanzees: A trans-generational study. *Primates*, 137–154). <https://doi.org/10.1007/BF02382050>

Tomasello, Michael; Call, Josep & Hare, Brian (2003). Chimpanzees understand psychological states – the question is which ones and to what extent. *Trends in Cognitive Sciences* 7, 153-156.

[https://doi.org/10.1016/S1364-6613\(03\)00035-4](https://doi.org/10.1016/S1364-6613(03)00035-4)

Tomasello, Michael; Hare, Brian; Lehmann, Hagen & Call, Josep (2007). Reliance on head versus eyes in the gaze following of great apes and human infants: the cooperative eye hypothesis. *Journal of Human Evolution*, 52, 314–320. <https://doi.org/10.1016/j.jhevol.2006.10.001>

Tomasello, Michael & Call, Josep (2019). Thirty years of great ape gestures. *Animal Cognition*, 22, 461-469. <https://doi.org/10.1007/s10071-018-1167-1>

Tramacere, Antonella & Mafessoni, Fabrizio (2022). Cognitive Twists: The Coevolution of Learning and Genes in Human Cognition. *Review of Philosophy and Psychology*. 10.1007/s13164-022-0670-w

Vasilieva, Olga (2019). Beyond “Uniqueness”: Habitual Traits in the Context of Cognitive-communicative Continuity. *Theoria et Historia Scientiarum* 16, 129.

<https://doi.org/10.12775/th.2019.008>

Vieira, Joana; Schellhaas, Sabine; Enström, Erik & Olsson, Andreas. (2020). Help or flight? Increased threat imminence promotes defensive helping in humans. *Proceedings of the Royal Society B: Biological Sciences*, 287. <https://doi.org/10.1098/rspb.2020.1473>

Vieira, Joana & Olsson, Andreas (preprint). Help or flight: Neural defensive circuits promote helping under threat in humans. <https://doi.org/10.1101/2021.12.17.473167>

Warren, Elisabeth & Call, Josep (2022). Inferential Communication: Bridging the Gap Between Intentional and Ostensive Communication in Non-human Primates. *Frontiers in Psychology*.

<https://doi.org/10.3389/fpsyg.2021.718251>

Wolf, Wouter; Thielhelm, Julia; Tomasello Michael (2023). Five-year-old children show cooperative preferences for faces with white sclera. *Journal of Experimental Child Psychology* 225(5).

[10.1016/j.jecp.2022.105532](https://doi.org/10.1016/j.jecp.2022.105532)

Yáñez, Bernardo & Gomila, Antoni (2018). Evolución de la esclerótica del ojo humano: Una hipótesis social. *Ludus vitalis*, 26, 119-132.

Zollikofer et al. (2022). Endocranial ontogeny and evolution in early Homo sapiens: The evidence from Herto, Ethiopia. *Proceedings of the National Academy of Sciences* 119(32).

<https://doi.org/10.1073/pnas.2123553119>

Zuberbühler, Klaus (2008). Gaze-following. *Current Biology*, 18, R453-5.