



Metarepresenting in communication

Nima Mussavifard¹

Received: 1 June 2023 / Accepted: 3 November 2023 / Published online: 15 November 2023
© The Author(s) 2023

Abstract

The standard cognitive explanation for the emergence of human communication is that it rests largely on the expression and attribution of communicative intentions which are, in turn, enabled by complex metarepresentations of mental states. This complexity is at odds with the limited metapsychological abilities of infants. But mentalistic metarepresentations are neither necessary nor sufficient in explaining communication. Coded ostensive signals (e.g., eye contact) and established channels (e.g., speech) allow that communicative episodes be identified through decoding rather than metarepresentational inferences. Thus, some metarepresentations may be unnecessary. However, metapsychology is also insufficient for explaining communication: the logic of instrumental actions permits interpreting their effect as following from intentions, yet the effect of communicative actions is often unavailable for inferring meaning. Moreover, current evidence for the developmental trajectory of communication and mental state attribution does not support the emergence of the former from the latter. My proposal is that our primitive concept of communication targets, instead, representational action. When we communicate, we typically convey a propositional content that is detached from our acts—a property absent in ordinary goal-directed actions. This view additionally raises the possibility that metarepresentational capacities evolved for representing external, communicative representations and were only later exapted for other purposes.

Keywords Ostensive communication · Grice · Metarepresentation · External representation · Evolution · Natural pedagogy

✉ Nima Mussavifard
mussavifard@gmail.com

¹ Department of Cognitive Science, Central European University, Vienna, Austria

1 Introduction

Humans are unique in their metarepresentational capacities (Cosmides & Tooby, 2000; Sperber, 2000; Suddendorf, 1999). Metarepresentation allows us not only to represent the world directly, as other animals do, but also to represent, reason about, and learn from other representations. These embedded representations could be of a mental nature. Metarepresenting mental representations enables predicting and explaining the behaviors of other organisms, typically by postulating beliefs and desires that could act as the causal drivers of those behaviors (Dennett, 1987). It was in the context of such “theory of mind¹” (Premack & Woodruff, 1978) capacities that the term metarepresentation gained currency (Pylyshyn, 1978). It is, thus, no wonder that metarepresentation is often taken as synonymous with theory of mind. Yet, undeniably, external representations, too, can be objects of representing minds, and humans are no less skillful at those. From markings (Ittelson, 1996), animations (Revenu & Csibra, 2021), and other types of depiction (Clark, 2016) to the more regularly used words and gestures, we exploit perceivable means to convey relevant information to one another. Both types of metarepresentation develop early and are instrumental in our social lives.

Another seemingly unique feat, *ostensive* (or *Gricean*) *communication* (Sperber & Wilson, 1995) is widely acknowledged to be linked to metarepresentation (Malle, 2002; Woensdregt & Smith, 2017). This is a communicative system that utilizes an open-ended range of entities (e.g., actions and objects) to transmit an open-ended range of messages². We can, for example, turn any object-directed action into a demonstration on the fly and thereby convey numerous ways of handling the object (e.g., see Király et al., 2013). This open-endedness likely implies that the same productive system is at work across modalities, giving rise to conventional and ad-hoc means of communication.

Since Grice’s landmark paper (1957), *intentionalism* (Harris, 2019) in its various forms became the dominant account of human communication in philosophy, linguistics, and cognitive science. According to this account, communication largely consists in the expression and attribution of communicative intentions by interlocutors. Mentalistic metarepresentation was then expectedly taken to be at play in communication, explaining both its development and evolution (Scott-Phillips, 2014; Sperber & Wilson, 2002). But this approach has proved to be fraught with difficulties, mainly having to do with the complexity of the cognitive processes it accredits to infants and ancestral hominins. According to an opposing camp (e.g., Geurts, 2021), on the other hand, this picture takes things reversely: it is language and the capacities it affords that permit the development and evolution of metarepresentation.

Therefore, a central question in the study of human communication is the causal direction between mentalizing, on the one hand, and language and communication,

¹ In this paper, I use theory of mind, metapsychology, mindreading, and mentalizing more or less synonymously to refer to metarepresentation of mental states.

² Open-ended communication is only the output of the system and, as such, the explanandum in my account. But I will provide functional definitions of ostensive communication, including what enables open-endedness, in § 3 and § 5.

on the other. What is often missing in this debate is the role of representations of a public nature. Whichever stance we take on the ontogenetic and phylogenetic origins of human communication, we must offer an account of metarepresentation *in* communication, namely the necessary employment of external representations for the communicative transfer of information. This neglect, I will argue in this paper, leads to problems deeper than mere cognitive load and complexity.

After reviewing the predominantly Gricean approaches, the different roles they impute to metarepresentation in human communication, and the common problems that stem from them (§ 2), I discuss how in many contexts higher-order metarepresentation of mental states might not be necessary (§ 3) and, more crucially, why it is not sufficient in explaining the emergence of human communication (§ 4). Then I sketch out an alternative account of the developmental and evolutionary origin of human communication according to which ostensive communication involves a primitive concept that is irreducible to a configuration of mentalistic propositional attitudes and requires instead a capacity to identify external representations and their detached content (§ 5). Lastly, I speculate that, if human communication does not necessitate mentalizing, metarepresentation might have evolved to enable the use of communicative, external representations (§ 6).

2 Metarepresentation in the Gricean approach

In his 1957 paper, Grice noted a fundamental difference between our uses of the word “meaning”. For example, when we say,

The smoke means fire.

we are concerned with a factive sense of the word, to the effect that “ x means p ” entails p . If the smoke means fire, then what we have is indeed fire. He labeled this use of the word *natural meaning*. In contrast, when we say,

Her utterance means “there is fire”.

we are using the word in an entirely different sense. Here we cannot unproblematically conclude p from “ x means that p .” Such non-factive uses involve what Grice termed *nonnatural meaning* (meaning_{NN}, for short), and this is the sense that theories of human communication must address. Grice’s first proposal for analyzing meaning_{NN} was that “ x meant_{NN} something” could be taken to be true if x was intended by the utterer U to induce a belief in an audience. (Note that it is here that the Gricean reduction of the concept of communication to intentions and beliefs originates.) Grice notes immediately that this is not sufficient. For instance, if A leaves B’s handkerchief by the scene of a murder to induce in a detective the belief that B is the murderer, we cannot say that the handkerchief meant_{NN} anything. This is a case of hidden authorship, rather than communication (Grosse et al., 2013; Tomasello, 2008). For communication to take place, the intention behind the act must be overt (Strawson, 1964). (Note now that overttness, including the problems associated with

it, applies only if one adheres to the above-mentioned reduction.) Thus, “ x means_{NN} something” would be true if somebody intended the utterance of x to produce a certain effect in an audience by means of the recognition of this intention. This seems to involve a self-referential intention (Bach, 1987). It sounds as though the audience should already know what U ’s intention is in order to recognize it. Likely in fear of such a “reflexive paradox” (Grice, 1957) and following Strawson’s suggestion, Grice later (1969) formulated his analysis in terms of “iterative” intentions (Bach, 2012). Therefore,

U meant something by uttering x if and only if, for some audience A , U uttered x intending.

- (I) A to produce a particular response r
- (II) A to think (recognize) that U intends (I)
- (III) A to fulfil (I) on the basis of his fulfillment of (II).

Condition (III) is added to rule out what Grice took to be counterexamples to a simpler formulation: King Herod presents Salome with the head of Saint John the Baptist on a charger intending her to believe that Saint John is dead and intending her to recognize this intention; and someone shows his bandaged leg to convey that he has a bandaged leg (as opposed to, say, he cannot play squash). Grice wanted neither of these to be included by his concept of meaning_{NN}, for he believed that in such cases the attribution of intention is incidental to the response—that is, the evidence itself is enough to make the required inferences. Besides these, Grice’s definition gave rise to an industry of counterexamples to the formulation, most notably by Strawson (1964) and Schiffer (1972). I will return to these below.

In developing *relevance theory*, Sperber and Wilson (1995; Wilson & Sperber, 2012) placed Gricean pragmatics within the framework of the emerging cognitive sciences and so spelled out most explicitly the representational requirements of communication. Rejecting the necessity of sub-intention (III) in Grice’s formulation, they suggest that at least two intentions are expressed in communication. The first intention, corresponding to Grice’s sub-intention (I), is:

informative intention: to inform an audience of something.

This is embedded in a second-order intention, corresponding to Grice’s sub-intention (II):

communicative intention: to inform the audience of one’s informative intention.

They point out two potential difficulties in processing these intentions (Sperber, 2000; Sperber & Wilson, 2002): Firstly, the inferences required for interpreting utterances, as envisaged by traditional Gricean accounts, seem to demand complex belief-desire reasoning on the part of the audience. In his other seminal work, Grice (1975) suggested that this is done by the expectation that interlocutors observe a *cooperative principle* which allows the audience to eliminate interpretations that are incompatible

with the principle—a rational process which necessitates sophisticated metapsychology. Secondly, specifying the processes involved in the recognition of communicative intentions reveals several layers of metarepresentation. As an example, when Mary is ostensibly picking berries to convey to Peter that the berries are edible, to fully grasp this, Peter must entertain a fourth-order metarepresentation (Sperber, 2000):

Mary intends₄
 Peter should believe₃
 Mary intends₂
 Peter should believe₁
 these berries are edible.

This metarepresentational format is what permits Mary’s informative intention to be overt (or “mutually manifest”). Moreover, because of their inherent ambiguity, our utterances underdetermine the intended meaning, that is, there is a systematic gap between what is explicitly said and what is intended. This underdetermination would be resolved if the utterance is taken only as a piece of evidence for the intended meaning. The audience ultimately forms a metarepresentation of the speaker’s meaning in the interpretive process. And yet another reason for such high orders of metarepresentation is that, if there is doubt in the competence and benevolence of the communicator, the audience can still interpret the communication appropriately. In such cases, although the informative intention may not succeed and, as a result, the audience does not believe the content, he will nevertheless recognize the communicative intention. That is, metarepresentation separates comprehension from acceptance (Heintz & Scott-Phillips, 2023). However, if the communicator is trustworthy, there would be no need for such a metarepresentational distance (Sperber, 1994b).

Metarepresentation enjoys thus a central role in relevance theory; so much so, indeed, that “the very act of ostensive communication, in both production and comprehension, is an exercise in reading others’ minds” (Scott-Phillips, 2014). Therefore, the causal direction clearly goes from (mentalistic) metarepresentation to ostensive communication. Consequently, the reason humans but no other great apes can communicate ostensively (but for a recent defense of “proto-ostension” in non-human primates see Sperber, 2019) is mainly a metapsychological advantage (Scott-Phillips, 2015). Being more than a code-based signaling system, it is argued, language could not have emerged without the metapsychological capability to attribute higher-order informative intentions. However, the complexity of the necessary computations, on the one hand, and the ease and early developmental emergence of communication, on the other, suggest that a sub-module of our mindreading mechanism might have evolved to exploit the regularities in the domain of communication (Sperber, 2000; Sperber & Wilson, 2002). Particularly, this comprehension module is licensed to assume that the utterance is optimally relevant, that is, it is sufficiently relevant to be worth the audience’s attention and is as relevant as is compatible with the utterer’s abilities and preferences. Then, the audience can follow a path of least effort in constructing interpretive hypotheses and stop when his expectations of relevance are

satisfied. As such, although the output of this process may be attributed as a metarepresentation of the speaker's meaning, the process need not be metarepresentational.

Recently there have been several other attempts at either simplifying or qualifying the standard formulations to allow that prelinguistic human and non-human animals possess some form of ostensive communication. In one such attempt, Breheny (2006) states the developmental dilemma as a conflict between the assumption that communication involves attributing propositional attitudes like intentions and beliefs and the then widely acknowledged finding that, although capable of adult-like communication, children below the age of four are unable to represent propositional attitudes—as evidenced by their failure to appreciate others' false beliefs (Wellman et al., 2001; Wimmer & Perner, 1983). This appears especially problematic for accounts that place mindreading at the center of language development (Bloom, 2000; Tomasello, 2008). Drawing on some of the ideas in relevance theory, he proposes that basic communication might not require mentalizing, but rather the ability to recognize instances of communication based on an action concept, and the use of a relevance-guided procedure to identify the referent. But an alternative account of children's mindreading, according to which their failure at standard false-belief tasks is due to processing demands rather than a conceptual shortcoming (Fodor, 1992; Leslie et al., 2004), has received more empirical support in recent years (Onishi & Baillargeon, 2005; Surian et al., 2007; but see Rakoczy & Behne, 2019). Thus, one strategy has been to use these results to deny that the complexity of communicative intentions poses a fatal blow to Gricean approaches (Thompson, 2014).

Another strategy for mitigating the inferential complexity is to hold that, while the Gricean approach is generally correct, one can reduce the representations involved in the analysis without missing the main insights. Like Breheny, Moore (2014, 2016, 2018) finds the dependence of the dominant views on a concept of belief problematic, as this could imply that neither infants nor non-human primates communicate ostensively. The issue is partially resolved in directive communication which is aimed at producing an action in the audience. However, informative communication too may not always necessitate changing belief-states: communication can be directed at changing perceptual states such as seeing. As for the metarepresentational complexity of communicative intentions, Moore remarks that even ten-year-old children have difficulty with entertaining fourth-order metarepresentations. Any Gricean account of the origin of human communication must therefore explain how creatures without such sophisticated capacities can nevertheless develop ostensive communication. Following Gómez (1994, 1996), Moore's response to this challenge is that a minimal Gricean account without the demanding metarepresentations is possible if one admits a functional reading of Grice's formulation. According to such a reading, the first clause in the formulation can be enacted by *sign production*, which refers to communicative behavior (e.g., pointing) with the purpose of eliciting some response *r* from the audience. The second clause can then be enacted by functionally separate *acts of address*. These are performed to direct the audience's attention to the action (e.g., through eye contact), thus fulfilling the overtness requirement of Gricean communication. This functional separation considerably reduces the number of metarepresentations in the analysis. Adapting the above example, in addressing her gesture to Peter's attention through, say, an attention-getter:

Mary intends₁
Peter attends and responds to her gesture.

And in pointing at berries:

Mary intends₁
Peter looks at berries.

Therefore, two pairs of metarepresentation would be enough for comprehending such forms of communication. As the communicator does not need to represent her own intention, producing minimally Gricean acts does not involve metarepresentations at all. In a similar vein, Tomasello (2008) analyses communication in terms of two intentions: the *referential intention* that the audience attend to something and the *social intention* that the audience know or do something because of that attending. Since these intentions do not target beliefs, they are likely not metarepresentational. Both these minimal accounts may therefore permit the emergence of communication in preverbal infants and possibly non-human primates. If so, one could be justified in positing that sophisticated forms of metarepresentation evolved culturally from simpler forms by exploiting linguistic tools (Moore, 2021).

Another strategy in tackling the complexity of Gricean communication is to argue that there exist intermediate cases that share some features with full-blown ostensive communication but are simpler in structure (e.g., Green, 2019; Planer, 2017a, b). Identifying these intermediate cases, the argument goes, makes the task of explaining the transition from coded animal signals to mentalistic communication much easier. Thus, while postulating non-Gricean intermediate links, many of these accounts take issue more with the Gricean nature of early communication than with the validity of the Gricean analysis itself. As will become clear, I take the opposite path: while Grice's insights about the explanandum were correct, his intentionalist analysis does not provide a sufficient account of the emergence of human communication.

But, besides the present account, several alternatives to the Gricean picture have been proposed (Armstrong, 2023; Bar-On, 2013; Geurts, 2022; Millikan, 1987). Bar-On, for instance, suggests that expressive behaviors (e.g., distress calls) permit organisms to share their states of mind about the world, absent the Gricean intentions. These behaviors may then have been replaced by linguistic expressive vehicles. And Armstrong proposes *minded communication* as a system that can functionally coordinate communicators' mental states without requiring Gricean metarepresentations. This system could, nonetheless, provide a platform for the evolution of metarepresentation (cf. § 6). However, rather than mental representations, my account focuses on representational action as an evolutionarily novel concept that enables open-ended communication. Of course, such action must be mentally represented by the interactants, but its primary function need not be expressing or changing mental states.

To summarize, Grice accounted for the distinction between natural and nonnatural meaning by suggesting that human communication involves complex, overt intentions. Such intentions in turn build upon multi-layered metarepresentations of mental states. The complexity of such metarepresentational inferences creates difficulties for psychologically plausible accounts of the emergence of human communication

in ontogeny and phylogeny. These difficulties have usually been addressed by qualifying or minimizing the representational demands of communication. In the next sections, I will discuss the degree and kind of metarepresentations required for a parsimonious account of the concept of communication.

3 Higher-order metarepresentations are not necessary

The higher-order intentions in Gricean formulations serve mainly to distinguish communicative from non-communicative behavior (e.g., hidden authorship). The audience should, for instance, infer that the communicator wants him to believe that she has a message (i.e., an informative intention) for him. Therefore, most of the complexity of the representations stems from such inferences to the communicative nature of behavior. However, proponents of the theory of *natural pedagogy* have argued that human children are innately sensitive to a set of *ostensive signals* which, among other things, indicate to them the occurrence of communicative episodes (Csibra, 2010; Csibra & Gergely, 2006, 2009, 2011). These signals (e.g., eye contact, infant-directed speech, and contingent responsivity) allow caregivers to impart useful knowledge to infants by exploiting the infants' tendency to interpret such episodes as involving generic information. One of the implications of these signals is that, rather than metarepresentationally *infer* the presence of communication, infants can simply *decode* the signals to identify communication (Csibra, 2010). They would subsequently need to infer only the message behind the communicative act. Although later in development they might also be able to postulate that the communicator has an intention to communicate, this is arguably not always needed. If so, then the requisite metarepresentations in the Gricean analysis would be significantly reduced (see also Moore, 2014). But pedagogy is not the only context in which humans can directly recognize communication. Adult interactions, too, sometimes involve ostensive signals, such as eye contact, for rendering the behavior communicative (e.g., while demonstrating or pantomiming).

Therefore, a functional reading of Grice's second clause, which (contra Moore, 2016, 2018) considers its role in distinguishing communication from non-communicative behavior, leads us to parsimoniously define ostension as the flexible marking of entities (e.g., actions and objects) as communicative—a function which, after Grice's $\text{meaning}_{\text{NN}}$, we can call $\text{marking}_{\text{NN}}$ ³. Ostensive communication would then be a communicative system that makes use of $\text{marking}_{\text{NN}}$. This capacity, which enables the open-ended production of novel communicative means, permits the function of ostension to be fulfilled without necessarily drawing on metapsychology. Decoding ostensive signals, is one mechanism for fulfilling this. However, the “communicative presumption” (Bach & Harnish, 1979) is at work whenever we use an established channel of communication. These channels, like the spoken and gestural modality, might be developmentally established through their co-occurrence with ostensive signals. In effect, in many, if not most, contexts, we can take the presence of commu-

³ As will become clear below, the nonnatural feature of communication is derived from the representational quality of the ostensively marked entities, not the marking itself.

nication for granted without needing to call on complex inferences. Thus, in development and evolution, humans could have utilized ostensive signals to mark_{NN} entities both for using ad-hoc means and for establishing communicative channels—thereby bootstrapping the emergence of communication.

The existence of specialized ostensive signals can ultimately simplify the Gricean analysis of communication into something like the following:

Mary (ostensively) intends₂
 Peter believes₁
 these berries are edible.

What is needed for this simpler formulation to work is that, although temporally and procedurally separate, the decoding of ostensive signals is conceptually linked to the interpretation of the content (Csibra, 2010).

The separation of the functional counterparts of Grice's clause (I) and (II) might be not only a *possibility* but a *necessity* for the emergence of ostensive communication. This is because embedding the informative intention within the communicative intention, as standard accounts do, makes a full grasp of the communicative intention impossible without also figuring out the informative intention. As young infants might not yet have the world knowledge or cognitive wherewithal to infer the content of the latter intention, they would not be able to infer the former either. Consequently, neither could develop. One solution would be to hold that infants represent the informative intention embedded as a placeholder, specifying the content type, within the communicative intention. But if the metarepresentational complexity in more developed communication does create explanatory problems for a psychologically plausible account, entertaining complex metarepresentations about an empty first-order content should be seen as an even more formidable task for infants. (Consider: Mary intends Peter believes Mary intends Peter believes something.) If, instead, recognizing communication and inferring the content are carried out by two separate processes, the infant can do the former without the latter. Thus, while the two must be conceptually linked, they may need to be (meta)representationally separate.

Although what adult interlocutors explicitly communicate often, if not always, underdetermines what they intend to convey, it does not follow that the audience must necessarily consider the communicator and her intention in the inferential processes (see also Recanati, 2002). There are cases in which the specific mental state of the communicator is not relevant to the interpretation (Geurts, 2019). For instance, when a priest utters "I hereby pronounce you husband and wife," (or in similar conventional speech acts) it does not usually matter what he intended by this. Likewise, when we interpret the meaning of an unfamiliar road sign, the communicative intentions of the person(s) who installed it may not contribute much (Sterelny, 2017)—assuming that those are accessible at all. But these could be viewed as exceptions to an otherwise interpersonal form of communication in which the identity and the mental states of both the communicator and the audience are inferable and contribute

to the interpretation⁴. While this is certainly true, it would be a mistake to conclude that the mechanisms involved in this final product are identical to the ones responsible for their emergence in development and evolutionary history (see also Bar-On, 2013). Again, some of the findings in the natural pedagogy framework could shed light on this.

In many studies on the development of communication, the identity of the communicator is not revealed (e.g., a recorded voice is played back). Children, nevertheless, seem to interpret the communicative acts appropriately, expanding our knowledge on the development of communication. Such “depersonalized” communication could, of course, be merely an artifact of experimental design. However, Egyed et al. (2013) have conducted a study that could address the issue more directly (see also Novack et al., 2014): Eighteen-month-old infants saw an experimenter show a positive emotional expression toward one object and a negative emotional expression toward another. These were preceded either by ostensive signals or no communication. In the test phase, either the same or a new experimenter requested the infant to hand them one of the objects. The results suggested that in ostensive contexts infants interpreted the expression as communicating generalizable object-directed knowledge, encouraging them to hand the *new* experimenter the positively valenced object. Although one can still argue that the experimenters’ communicative intentions are considered by the infants, it could alternatively imply that the identity, and thus the specific mental state, of the communicator are less relevant for infants in the interpretive process than the communicative act itself and what it conveys about the referent (see also Topál et al., 2009). Here communication seems to work more like the road sign example, creating a dilemma for intentionalism: either this is not a case of non-natural meaning or the mentalistic reduction must be dropped altogether. If communicative acts are mostly interpreted this way early in infancy and if these pedagogical scenarios reflect the context in which humans evolved their communicative system, then the emergence of communication might not be as dependent on the attribution of mental states as intentionalism suggests. Such a pedagogical origin for the evolution of ostensive communication (defended extensively in Csibra & Gergely, 2006; Laland, 2017) is especially attractive because, contrary to most other hypotheses, it proposes a unique selection pressure (i.e., teaching variable cultural and technological knowledge) for the unique ability in humans to communicate open-endedly.

Beside representing other representations, another feature of metarepresentations is suggested to be that they decouple the metarepresented content from the rest of the cognitive system, adding to it information that specifies the kind and source of the representation (Cosmides & Tooby, 2000; Leslie, 1987; see also § 6). Decoupling would allow cognizers to make inferences about the content within its relevant scope without committing errors originating from confusing metarepresented and first-order representations. This is critical for attributing mental states, since otherwise one would take someone else’s beliefs and intentions as their own and behave maladaptively (Cosmides & Tooby, 2000). This might also be crucial for adult communi-

⁴ Alternatively, one could argue that these are not exceptions but involve attribution of general intentions without an identified subject. My main point is not that these are definitive counterexamples, but rather that these *can* be given a different, more parsimonious, interpretation without appealing to intentions.

cation in which the communicative nature of the representation and who produced it could potentially affect not only how one interprets it but also the value one attaches to it—as the representation could be mistaken or deceptive. However, in parent–infant interactions, trust in the benevolence and competence of the communicator is built into the kin-based organization of the interaction. A pedagogical scenario for the origin of ostensive communication (or indeed any scenario with sufficient convergence of interests) would, therefore, allow that one treats the communicated representation as equivalent, if not superior (Marno & Csibra, 2015; Topál et al., 2008), to information obtained from perception and first-hand experience. Representing the self’s belief as the target of communication would also be mostly unnecessary, for, besides its reliability, the communication accompanying infant-directed signals is naturally meant for the infant. If so, the representations required for the emergence of a concept of communication would be minimized considerably.

4 Metarepresentations (of mental representations) are not sufficient

As mentioned before, Grice’s formulation of the intentional structure of communication has generated a host of counterexamples targeting either its necessity or sufficiency. For instance, it has been argued that torturing has the same structure as $\text{meaning}_{\text{NN}}$: the torturer intends the audience to produce a response r , to recognize this intention, and to produce r based on this recognition (Grice, 1969). However, a more widely-discussed type of counterexample was introduced by Strawson (1964): Mary wants Peter to mend her broken hairdryer. Thus, she pretends to mend the pieces together, hoping that Peter notices this and helps out. She intends Peter to realize that she intends him to help, but she does not intend him to realize that she intends him to realize that she intends him to help (example from Sperber & Wilson, 1995). Here, although all Grice’s clauses seem to be fulfilled, advocates of the general analysis do not wish to consider it as a case of $\text{meaning}_{\text{NN}}$, because, intuitively, this does not involve communication. As a result, the sufficiency of the analysis is threatened. Thus, Strawson proposed that the communicator “should not only intend A to recognize his intention to get A to think that p , but that he should also *intend A to recognize his intention to get A to recognize his intention to get A to think that p* ” (p. 447). However, he noted correctly that even this condition is unlikely to rule out further counterexamples. No matter how many such intentions we add to the formulation, there will be counterexamples in which the actor has a further deceptive intention to hide the lower-order intention. This will result in an infinite regress of intentions (and metarepresentations)—undesirable for a cognitive account.

One type of measure to root out these “sneaky intentions” (Grice, 1969) has been to introduce a condition that bars deception. For instance, Neale (1992) adds to clauses (I) and (II) a further clause that U does not intend A to be deceived about U ’s intentions (I) and (II) (see also Grice, 1969; Moore, 2016). The issue with this measure is that, while it may be a suitable solution for conceptual analysis, it does not provide a plausible cognitive account of communication. It appears unlikely that every time the audience is addressed in communication, he considers the absence of deceptive intentions. And it is even more unlikely that the less sophisticated individuals (i.e., ear-

lier hominins and infants) developing communication could entertain such thoughts. Plausibly, considering this kind of deception is an exception to the default interpretation of communication as being honest with respect to its communicative nature. Of course, the clause is not suggested to be represented by interlocutors. But the point is that it fails to offer a sufficient cognitive account.

Another measure, advocated by Schiffer (1972), is introducing a mutual-knowledge condition. U and A mutually know that p , if and only if U knows that p , A knows that p , U knows that A knows that p , A knows that U knows that p , and so on ad infinitum. Schiffer believes that it is the absence of mutual knowledge of this form that produces the deceptive counterexamples. For communication to be properly overt, the intentions involved must be mutually known between the communicator and the audience. However, as a cognitive account, this appears to replace the “vertical” regress of the communicator’s mental states with a “horizontal” regress of both interlocutors’ mutual mental states. That is, to mutually know intentions, interlocutors seem to need representations of infinite knowledge states about knowledge states. So construed, mutual knowledge, too, is deemed unable to provide a psychologically plausible explanation of communication (Sperber & Wilson, 1995).

Various attempts have been made to retain the strength of mutual knowledge without adhering to the apparent regress. For example, Tomasello (2008) argues that, based on our contextual needs, we only compute some of the recursive representations and we often access them using simple heuristics. But Sperber and Wilson (1995) suggest using, instead, the weaker concept of *mutual manifestness*. According to them, “[a] proposition is manifest to an individual at a given time to the extent that he is likely to some positive degree to entertain it and accept it as true” (Sperber & Wilson, 2015, p. 134). In a mutual cognitive environment, propositions that identify individuals who share that environment are *mutually* manifest as well. Similarly, Geurts (2018, 2019) suggests substituting mutual knowledge with the less demanding normative notions of “reasons to believe” (Lewis, 1969) or “mutual commitment”—both of which take the iterative structure of common ground to be a chain of implications rather than actual cognitive processes.

But whichever view on mutual knowledge we end up accepting, measures such as an anti-deception condition or a mutual-knowledge condition have been proposed to rule out counterexamples to the conceptual analysis of the notion of meaning_{NN}—aimed ultimately at spelling out its necessary and sufficient conditions. Consequently, their application for providing models of cognition should be approached with caution (see also Scarafone & Michael, 2022). Firstly, accounting for individual communicative cognition by appealing to such interpersonal concepts is not the only option. As argued in the previous section, it is possible that, early in infancy, humans are endowed with an inferential communication which does not necessitate complex mindreading. Accordingly, if we have a sufficient account of communicative *actions* independently of the propositional attitudes that cause or follow them (see § 5), then individuals can be suggested to develop this kind of action and its relevant concept (cf. Breheny, 2006) before the mastery of its interpersonal consequences. For instance, realizing that communicated content (as opposed to, say, hidden authorship) is in common ground or that it commits the participants to act accordingly can emerge later than the ability to produce or understand communicative acts.

Secondly and relatedly, as we saw with the torture example, recognizing that the act is produced with a similar mental set-up to communication is not sufficient to render the act communicative. Even if communication is mentalistic through and through, one would still need a proper concept of the type of action that licenses such mental state attributions. As an analogy, without a concept of goal-directed instrumental actions (e.g., knowing that they cause changes in the external world), attributing intentions would be insufficient for making the relevant predictions. Likewise, mentalizing in communication should follow a concept of what constitutes proper communicative acts (e.g., they are marked_{NN}) and their goals (e.g., they aim to convey information). Having a convincing account of this type of action, we can then turn to investigating whether and how attributing mental states contributes to its understanding. In contrast to such an action-based account, the standard approach attempts to characterize communicative cognition largely, if not solely, by specifying the propositional attitudes that are deemed necessary. These postulated mental states are then weighed against unspecified intuitions as to what action should or should not count as communication proper (e.g., the intuition that hidden authorship or torture should be excluded). I believe that this is the wrong explanatory direction. A sounder approach would be to directly address those intuitions by offering an account of communicative action. Issues of the psychological plausibility of complex metarepresentations and mutual knowledge arise arguably because the behavioral is explained solely in terms of the psychological⁵. And when this fails, ever more complex representations are added to compensate for the failure. If one focuses on the action, issues related to the overtness of mental states might not emerge at all.

In Gricean formulations, communicative representations consist in a configuration of intentions and beliefs. This implies that by developing the two concepts of intention and belief and linking them humans come to possess a metarepresentational complex that enables communication. More specifically, the reduction involves casting the problem of explaining communication to another level, where, among the set of possible intentions, there is a type of belief-inducing intentions. Granting the validity of this reduction, we can make a prediction regarding the developmental trajectory of the concept of communication: (1) infants develop the ability to attribute intentions and beliefs; (2) they link these two propositional attitudes; (3) they make use of the latter link to express and attribute informative intentions; and (4) they develop a higher-order informative intention that allows them to express and attribute communicative intentions. This would be ostensive communication proper. The problem with this prediction is already visible in (1), as from early in infancy humans show a rich and flexible understanding of communication (Bohn & Frank, 2019; Csibra, 2010; Vouloumanos et al., 2014). Thus, communication emerges alongside, if not earlier than, mindreading capacities and its core features are unlikely to be dependent on such mentalistic concepts. The link in (2) is also problematic in explaining the development of communication. On the one hand, this would delay the emergence of communication even further in development. On the other hand, prominent

⁵ The point is not that behavior cannot be explained by reference to the mental states of the producer or interpreter. Indeed, this is the goal of cognitive accounts of action. The point is rather that such a cognitive account should clearly specify the entity at which the mental states are directed.

theories of conceptual change (Carey, 2009; Spelke, 2003; Xu, 2019) take language to be at the center of this process, either through the structures it generates in the mind or through its compositional semantics that permits combining information across concepts. However, since the hypothetical conceptual change that would link the concepts of intention and belief is itself postulated to explain communication, linguistic communication is clearly unavailable to the process. One must therefore specify how this process occurs (e.g., see Gopnik, 2011). Regarding stages (3) and (4), it is question-begging, and as yet empirically unsupported, that the expression and attribution of informative intentions precedes communication. Indeed, it appears more likely that absent or hidden authorship develop as offshoots of communication, that is, as informative behaviors that suppress, or leave out, communicative cues. Therefore, on ontogenetic grounds, belief-inducing intentions fail to account for the early emergence of the concept of communication.

That communication does not emerge from metapsychological representations alone becomes more evident when we apply the logic of the latter to the interpretation of communication. In attribution of ordinary, instrumental intentions, we predict the behavior of agents to follow from the content of their conative propositional attitudes. For example, when we see Mary walking in the direction of her house, we attribute to her the intention to go home, and we predict her behavior accordingly:

Mary intends [Mary goes home] → Mary goes home.

This is possible because we have a good understanding of the means she can take to obtain her goal. For instance, we consult our knowledge that she lives nearby to infer that she will walk home. Clearly, the same procedure cannot be used to predict communicative action:

*Mary intends [Peter believes *p*] → Peter believes *p*.

Thus, Mary's intention that Peter believe something (e.g., that the berries are edible) is not sufficient. (After all, she does not possess telepathic abilities.) This is possibly because the subject of the embedded proposition is another person. So, a more proper analogy might be the following:

Mary intends [Peter sits down] → Mary makes Peter sit down.

Again, this inference is meaningful because we have a good grasp of the physical possibilities and constraints of human action. We subconsciously consult our knowledge to yield the conclusion that Mary pushes Peter down. But:

Mary intends [Peter believes *p*] → Mary makes Peter believe *p*.

Although this time the inference is not wrong, it is trivial because the (communicative) action that leads to Peter's belief is the very thing we would like to predict. Since the range of communicative means is much broader (see also Sperber & Wilson, 2002), we cannot as easily predict them. And more importantly, the infant simply

does not have prior access to them. The resources that are available for interpreting instrumental goals are unavailable in communication. These include, among other things, efficiency and simulation. The infant can predict behavior by assuming that agents choose the most efficient means given the environmental constraints (Gergely & Csibra, 2003). But efficiency is less relevant to communicative behavior (but see Bohn & Frank, 2019). For instance, the English word “tree” in itself is no more efficient or rational for denoting the concept TREE than the German word “Baum”. Simulation too is unhelpful, because the infant cannot rely on her own limited repertoire to predict others’ communicative behavior. Therefore, the above schema for action prediction is unlikely to be sufficiently useful for comprehending, and learning about, communicative behavior.

A similar difficulty emerges for action explanation. In explaining an action, we rely on the effect that the action caused to infer a corresponding intention. Thus:

Mary made Peter sit down → Mary intended [Peter sits down].

Yet this is not possible for explaining most communicative actions:

*Mary made Peter believe p → Mary intended [Peter believes p].

It is not possible in third-personal communication since, unlike instrumental actions, there is often no observable change in the audience, and we obviously do not have access to their beliefs. Considering the central role of declarative, as opposed to imperative, communication for humans (Tomasello, 2008), the problem becomes even more striking. In second-personal communication, the audience cannot start from the belief they form as a result of being addressed by an act, and then attribute that as the intention of the communicator. Recognizing this intention, on the standard accounts, is itself the very goal of the communicator (Sperber & Wilson, 2002). Faced with novel words or gestures, as infants are, there is no obvious way of first forming a belief so as to interpret it as the intention of the communicator or the meaning of the utterance. A relevance-guided comprehension module may avoid some of the problems with intention attribution. However, attributing the inference of such a module as the communicator’s intention would imply metacognitive access to the conclusion, complicating the developmental trajectory even more.

As a result, the standard treatment of utterances as just another case of instrumental action and the underlying processes as cases of mentalistic inferences is not a sufficient approach and is unlikely to explain the emergence of human communication. It might be more fruitful in accounting for language *use*, in which there is a largely established channel in place and interlocutors can use the tools at their disposal to manipulate one another’s mental states and actions. It could also prove helpful for explaining how non-human primates use their signals which are relatively preestablished and limited but utilized flexibly to achieve various goals (Byrne et al., 2017). Human communication, however, is marked not only by flexible use but also flexible production of ad-hoc communicative means (e.g., pantomiming and demonstration) or entirely learned devices (e.g., words and gestures). Then, an account is in order

which explains the unique features of human communication and the underpinning cognitive representations that make the system possible.

5 Communication as representational action

So far, I have claimed that we cannot unproblematically analyze communicative cognition into a configuration of intentions and beliefs. Ostensive communication may alternatively involve an irreducible, primitive concept that enables comprehending, learning about, and learning from communicative episodes. Thus, our example could be further simplified to only one metarepresentation:

Mary communicates₁
these berries are edible.

And sometimes, we saw, even the identity of the communicator may not be relevant. Now the question is: what is communication if not the expression and attribution of intentions? Perhaps the main phenomenon that a recourse to metapsychology makes us neglect is informative behavior. The presence of this class of behavior is, of course, not denied, but rather taken for granted. However, this should arguably be the central question for any account of the development and evolution of human-specific communication. From early in life, we humans, but apparently no other primate, see a class of behavior as having an informational function. This is unlikely to emerge solely from metarepresentations of mental states, for these are targeted at the relation between other minds and the external world. And besides, they typically serve to explain and predict actions that bring about a perceivable change in the world, not behavior that informs. A primate, armed with the most sophisticated metarepresentational capacities, would still struggle to understand why another agent moves its arms around in a strange, ineffective fashion. (Possibly it would think that the agent is desperately trying to drive away an insect, and so would go on with its business.) To us, however, it appears very trivial that a pantomiming action is informing about what it resembles—so much so that we might come to think that no extra cognitive mechanism is required for it. But no rationalization of otherwise instrumental behavior is likely to lead one to its representational nature if one does not already possess the relevant concept for interpreting informative behavior. Moreover, while, as shown above, it is possible to envisage a non-mentalistic form of inferential communication, there cannot be any realistic account of human communication that does not involve informative behavior—that is, the manipulation of external stimuli for the purpose of conveying information. Therefore, an account of this class of behavior and its representation in the mind takes explanatory precedence over mental state attribution.

But what does it mean for an action to have an informational function? Everything in our environment is a potential source of information, including other people's behavior. Thus, ostensive communication is not special in transmitting information. To answer the question, we could go back to Grice. As I mentioned above, while Grice's analysis of meaning and communication might have been mistaken, he seems to have had the right insight about what should count as meaning_{NN}. We saw that,

read functionally, clause (II) deals with the distinctive quality in human communication of marking_{NN}. Clause (III), on the other hand, requires the audience to produce *r* based on the fulfillment (or recognition) of the sub-intention in (II). Grice's insight was that if this is not the case, then what we have is an instance of showing—that is, “deliberately and openly letting someone know” (Grice, 1957). He believed that such cases (e.g., when you show a bandaged leg to convey that you have a bandaged leg) should not be considered meaning_{NN}. In these cases, as opposed to cases of “telling”, the required inference can be made based on the observable evidence and regardless of the purpose of communication. Of course, in keeping with his commitment to intentionalism, Grice's reasoning for this was that the inference be based on the recognition of the communicator's intention. However, it is possible to take a non-intentionalist lesson from this.

In paradigm cases of telling (e.g., in linguistic utterances and gestures), the relation between the communicative medium and the content is clearly one of representationality. This means that in order to arrive at the content of a linguistic utterance we cannot rely solely on the utterance and its physical (i.e., auditory and articulatory) features, but we need to infer a detached propositional content that the utterance is representing. That this is the case is most striking in depictions (Clark, 2016) because, despite perceptual similarities between the medium and the content, the depiction is meaningful only in relation to what it represents. Otherwise, a realistic drawing of a cup is just a mark on a piece of paper without the affordances of an actual cup. Although perhaps less striking, this is at the heart of most of our communication. The word “cup” (or miming the affordance of a cup) is not an actual cup—nor should it resemble one. How it relates to a particular cup or a CUP concept is through representing. The dependence on detached contents becomes more discernible when we look at the acquisition history of the symbol. That is, although the relation between “cup” and its meaning seems to be direct, hearing the word for the first time we must infer what it is a representation of. This is not the case in the bandaged leg example. There, drawing attention to the bandage does not represent anything (unless used to communicate something else). The bandaged leg is simply a bandaged leg.

At first blush, the theoretical reliance on representationality might appear to rule out a wide range of human communication. Is pointing to a cup not ostensive communication? Sometimes we point at a cup to evoke the kind associated with the proximal referent—for example, when we ask for any member of the same kind as the indicated cup (meaning: “Bring me *a* cup!”). Here, the distal referent (i.e., the kind) is detached from the communicative medium. But even when we are pointing at a cup to convey something about that very cup, we are seldom merely drawing attention to it. The same holds when we show a bandaged leg. Often, we point at something to lead the audience to inferences that they would not otherwise make. In most such cases, even if the referent is not detached from our communication, that is, it is part of the proximal medium we utilize to inform, the predicate is nevertheless detached from it. By pointing at the cup, we request our audience, say, to FILL it. What is distinctive about human pointing is not only that it can refer to things. Even some species of fish might be able to do so (Moore, 2018; Vail et al., 2013). Human pointing is special for its capacity to invite detached inferences (see also Tomasello, 2008). Hence, reference in pointing should not be considered in isolation from predication.

Instead, pointing should be seen as an “utterance” with a full propositional content. Or take cases of demonstration. Sometimes, both the referent and the predicate fall out of the scope of the communicative medium. For example, by manipulating some object-props, we denote an action kind that can be performed on an object kind, none of which are by definition present here and now. And even when we demonstrate an action on a particular object (e.g., a unique machine), the predicate (e.g., the action kind that the learner must perform) is represented by the communicative act.

Representationality can, nonetheless, rule out the familiar counterexamples. Torture is not a representational action, even if its intentional structure turns out to be akin to communication. It is, alas, an instrumental action taken to bring about a change (albeit peculiar) in another person. Examples of hidden authorship, too, involve (at least for the audience) non-representational action. Regardless of the mental states of the actor, the audience treats planted evidence as he would do, were it not slyly arranged by someone else. Yet when the audience takes some arrangement to be communicative, he interprets it as representing a detached content. A key on the table would then not only be a key, but also represent a content that is to be inferred—say, that the audience should lock the door using that key. As I said, Grice wished also to rule out showing in his account of meaning_{NN} for its apparent natural feature—although he included cases like nonspontaneous frowning (Grice, 1982). This move has been criticized as unnecessarily excluding important forms of communication (Neale, 1992; Sperber & Wilson, 2015). One way of approaching the question would be to hold that admitting showing in our account achieves inclusivity at the expense of explanatory power. If we allow examples like the bandaged leg, we might miss the crucial feature in most human communication (e.g., in linguistic utterances) of informing about something only indirectly. However, as I mentioned, we rarely communicate like the limiting case of the bandaged leg. This is also reflected in other approaches to communication which require utterances to convey relevant information (Grice, 1975; Sperber & Wilson, 1995). The relevance expectation often leads the audience to seek information that is not readily available. Besides, the function for which a cognitive system has evolved need not entirely overlap with how it can actually be used (see also Sperber, 1994a; Sperber & Hirschfeld, 2004). While, as I claim, our communicative concept may be geared to representational action, it can be exploited also to direct attention to stimuli. Then, even linguistic utterances, as paradigm cases of representational communication, could be used to manipulate attention: when we say “I am here!” to a friend who is looking for us in a crowd, regardless of the semantics of the constituent words, our voice is likely to attract her attention and lead her to the same conclusion (Recanati, 1986). The present account would also incorporate examples like nonspontaneous frowning: how it transmits the information is not necessarily or only the communicative intention behind it, but also its representational relation to its content (i.e., actual frowns and their implication). Lastly, soliloquy (i.e., communication without an audience) can create explanatory problems for intentionalist accounts, as there is no audience in whom you can intend to induce a response or belief (but see Harris, 2019). However, you can perform a representational action whether or not it is addressed to anyone (Searle, 1983).

The point is not that communication is the only domain in which we make inferences about things detached from the perceivable stimuli (see also Gärdenfors, 1995;

Planer, 2021). This happens almost all the time—for instance, when we infer the presence of things that we cannot currently see. Nor does it mean simply that ostensive communication evokes representations in our minds. Some alarm calls in birds and chimpanzees appear to evoke a mental representation corresponding to their (functional) referent (Sato et al., 2022; Suzuki, 2018). What is unique about ostensive communication is that it creates in our mind the *expectation* that the stimulus is representational. It relies, thus, on understanding a class of perceptual stimuli as possessing “aboutness”. This expectation permits communicators to provide evidence, through marked_{NN} entities, for unlimited contents, hoping that the audience will realize this and seek those contents. It also creates a separation between recognition and content, because the audience (e.g., infants) can recognize something as representational and learn about its features, without yet having access to the represented.

Animal signals may also be construed as representational in a functional sense, as they have evolved to stand for various meanings. Moreover, enculturated apes can acquire multiple language-like signs (reviewed in Gillespie-Lynch et al., 2014). Yet, it is doubtful that other species possess any matching disposition to attribute representationality to unfamiliar stimuli in novel channels (see also Novack & Waxman, 2020; Warren & Call, 2022). Humans, on the contrary, interpret novel communicative behavior in various modalities as representational from very early in infancy (Ferguson & Waxman, 2016; Novack et al., 2014; Tauzin & Gergely, 2018). Therefore, one of the crucial distinctions between human and non-human communication (across development) is that the former uses entities that trigger a search for content, even when both the entities and their contents are unspecified in the repertoire. As a result, human communication seems to require a specialized concept targeting variable representational action. This is arguably not necessary for non-human great apes. Although they can exploit various social and contextual cues to augment their communication (Bohn & Frank, 2019), their signals are constrained by the timescales of phylogenetic (Byrne et al., 2017) or ontogenetic (Tomasello & Call, 2019) ritualization. Hence, distinct processes may be responsible for the development and use of distinct signals. In contrast, both the size of our learned symbolic repertoire and the possibility to use ad-hoc means necessitate a broad and flexible conceptual understanding of external representations that is likely specific to humans.

The representational understanding of communicative acts provides an extraordinary possibility for humans to link concepts to other concepts or entities—with potential consequences for conceptual development. By drawing on the representational nature of communication, you can spontaneously specify a referent, perform an action that calls to mind your concept of choice, and connect the concept to the referent. In this way, you can communicate that you want your phone to be brought to you or even suggest (in pretense) that a banana is your phone (Leslie, 1987). Thus, due to their representational nature, our communicative acts permit us to establish arbitrary (or nonnatural) informational links. We can, therefore, call this representing function *informing_{NN}*. Inferences in other domains arguably draw on the existence of preestablished informational links, based, for example, on statistical or nomic relations. Smoke means fire because there is a causal and statistical relation between the two (see also Piccinini & Scarantino, 2011; Scarantino & Piccinini, 2010). An alarm call means snake, due to genetically and/or statistically encoded associations. How-

ever, ostensive communication *is* a way of establishing informational relations. As such, its functioning does not require (although it can use) preexisting informational links between the communicative action and the referent. We can, thus, account for the difference between what Grice called natural and nonnatural meaning without appealing to the intentions behind them.

When we communicate, we use one entity E_1 (typically an action) to inform about another entity E_2 (i.e., the referent). This way we set up an asymmetrical informational link between E_1 and E_2 , such that E_1 can be used to draw inferences about E_2 —but not vice versa. Consequently, the audience's task is to identify in the cognitive environment the scope of the representational medium, on the one hand, and the scope of the representational content, on the other. Specifying these representational scopes is not a trivial matter, of course. It is likely that early in development these are largely prespecified. For instance, the action may be taken to designate the predicate (most evidently in demonstrations) and the object may be construed as the referent. By building on this action-object link and the iconic features of the action, the infant can both acquire knowledge and bootstrap the development of the communicative system. Later in ontogeny, depending on the context (and, of course, the communicator's inferred mental state), the scope can vary dramatically. Sometimes the object is part of the communicative medium, sometimes it is not (as in displaced reference). Sometimes the action informs us about an entity, sometimes (as in the example with the key) we do not observe any action whatsoever.

The communicative system, then, typically functions along these lines in comprehension: (1) communicative episodes are detected through recognizing ostensive stimuli (i.e., markers_{NN}); (2) the referent is identified (mostly through following deictic gestures); (3) an informational link is established between the communicative act, a conceptual predicate, and the referent; (4) inferences are drawn about the referent. Let us see how this works. In demonstrations, after detecting ostension, the infant identifies the objects as the proximal referents, by following both the adult's gaze and her manual handling of the objects (Pomiczowska & Csibra, 2022). A predicate placeholder is generated, filled by conceptual information drawn from the spatiotemporal features of the object manipulation, and linked to the referent objects. Finally, the infant makes inferences about the referents. The resulting effect is thus predicated as the function of an object (e.g., "A opens bottles.") or the final arrangement is ascribed as a relation between two objects (e.g., "A goes on top of B."). Sometimes, however, the predicate is not iconically represented in the action. In cases of pointing, for example, infants must fill the predicate placeholder based on the context (or even on the referent). Detecting communication, they use the adult's pointing gesture to identify the referent, they generate a predicate placeholder and fill it with conceptual information from the context (e.g., GIVE in a play context), and use this to yield inferences about the referent (e.g., that they must give the object to their mother). And still sometimes, as in linguistic communication, the predicate might be codified in the action.

Note that such inferences might not have been made outside the domain of communication. Observing someone perform a purely instrumental action on an object, infants may or may not draw those conceptual inferences. (They may, for instance, encode it as a transient relation between that specific action and the object.) Or seeing

their mother search for an object, they might eventually realize that she wishes the object to be handed to her. However, by exploiting the above-mentioned procedure, communication can constrain and secure the necessary inferences. While leaving the key on the table could remind your flatmates to lock the door, if you do it conspicuously to tap into their communicative concept, you secure your intended inference. They will now take it to represent a detached content by, say, ascribing a predicate (e.g., LOCK) to an implied referent (e.g., the key or the door). This is, of course, an atypical example, where there is only a trace of the communicative action. As a perhaps more typical example, the caregiver demonstrates an action to secure in the infant's mind an informational link between the action and the referent, leading the infant to otherwise opaque inferences about the object kind (Csibra & Gergely, 2011). And in linguistic communication, one can even exploit the representational nature of communicative action to create, and inform about, fictional entities. Furthermore, whereas information use is abundant across cognitive domains, "referential information" (Scarantino & Clay, 2015), in which one entity is directly stipulated to inform about another entity, seems unique to communication.

Thus, the use of representational means to convey detached propositional contents or simply informing_{NN} enables an inferential system that does not necessitate attributing mental states (cf. Armstrong, 2023; Bar-On, 2013). The present action-based account can then cover a similar range of communicative behaviors (i.e., the explanandum) to intentionalist accounts without appealing to communicative intentions. Specifically, Gricean formulations have a two-tiered nature: a first-order intention to induce a belief or response and a second-order intention that the first intention be recognized. In my account, informing_{NN} or representation corresponds functionally to the former intention, whereas marking_{NN} corresponds functionally to the second-order intention. Thus, in a more complete definition, ostensive communication is a system that involves marking_{NN} and informing_{NN}. Since these are functions rather than entities, they can be implemented in the same act or in separate acts. For instance, by shaking an empty glass you can both signal that your act is communicative (in other words, representational) and that the glass should be refilled. But you can also make eye contact (the marker_{NN}) and raise your glass (the informer_{NN}) to convey the same content. Focus on the functions, moreover, helps avoid the complex representational requirements. Although functionally it has a similarly two-tiered structure, my formulation needs only one level of metarepresentation: a representation of the external representation. Accordingly, for minimally successful communication, the communicator must manipulate external entities (typically actions and objects) in various ways to provide evidence for the represented, and she must mark_{NN} these entities appropriately. "Uptake" (Austin, 1962) takes place, on the one hand, when the audience recognizes the marking_{NN} and so identifies the entities as representational, and, on the other, when he infers the represented content.

Offering an action-based account is not to deny the importance of mentalizing in communication. Often, we rely on the intentions ("representing intentions"; Searle, 1983) of the communicator to identify the referent or predicate. And sometimes the content of the utterance is itself a mental state (e.g., in expressive communication). However, the attribution of mental states would be unhelpful in communication if the representational structure were not in place. With it, one can comprehend communi-

cation both when the mental state is irrelevant or unavailable and when it is necessary to arrive at the right interpretation. The schema of intention attribution failed to predict and explain communication (§ 4) because communicative acts involve a different type of goal. The goal in instrumental actions is typically a two-place relation between an action a and a change of state b : $G_i(a, b)$. However, in communication the goal is a three-place relation between an action x , what it represents y , and (sometimes) the change of state (e.g., belief or action) z : $G_c(x, y, z)$. Thus, with respect to the goal, communication involves components that are distinct from instrumental, goal-directed action.

To sum up, human communication is characterized by marking_{NN} and informing_{NN}. Marking_{NN} allows us to open-endedly mark entities as communicative through ostensive stimuli—thus, enabling open-endedness of medium. Informing_{NN} is about how the message in the marked_{NN} entity is communicated. The informational relation between the communicative action and its message is one of representationality, which involves a detached propositional (i.e., with a predicate-argument structure) content. By manipulating external entities in various ways and establishing arbitrary informational links between represented concepts and referents, humans can open-endedly convey information to one another. External representations, therefore, enable open-endedness of content. Although this communicative system can utilize postulated mental states to home in on the content, this process involves components that are distinct from those of ordinary goal-directed actions. We thus have an account of communication that, although certainly sketchy at this point, takes account both of the action (i.e., communicatively marked, representational action) and its cognitive underpinning.

6 The evolution of metarepresentation

Evolutionary theories of the origin of metarepresentational capacities can be classified into three groups: first, theories that propose metarepresentation evolved to solve mostly individual, rather than social, problems such as metacognition (Couchman et al., 2009) and decoupling (Cosmides & Tooby, 2000); second, theories that suggest metarepresentation originated in social cognition (Baron-Cohen, 1999; Byrne & Whiten, 1997; Sperber, 2000); and third, theories that claim human-like metarepresentation evolved culturally in language (Geurts, 2021; Moore, 2021).

Prominent among the first strand of theories is Cosmides and Tooby's (2000) suggestion that metarepresentation is an adaptation to the "cognitive niche". This is an adaptive mode that involves increasing use of contingent information for the regulation of improvised behavior (see also Pinker, 2010). Through decoupling (Leslie, 1987), a metarepresented content is quarantined to allow inferences that are valid within the relevant scope but harmful if applied outside of it. These representations are stored with source tags indicating how they have been obtained (e.g., self vs. other). Subsequent information about the source (e.g., its reliability) may affect the truth-status of the representation and promote it in the cognitive architecture. Such metarepresentational capacities are useful not only for solving socio-cognitive problems but also, among other things, for planning and episodic memory.

But a more popular approach to the evolution of metarepresentation views it as an adaptation to a “socio-cognitive niche” (Whiten & Erdal, 2012). According to this view, the evolution of distinctive cognitive abilities in primates (also called “Machiavellian intelligence”) is largely determined by living in large, semi-permanent groups of long-lived individuals and the problems it poses (Byrne, 1996; Byrne & Whiten, 1997). This environment favors, on the one hand, the use of deception to achieve individual benefits, and, on the other, cooperation and coalition building. This causes an arms race between the social skills of those seeking higher ranks in the group and those collaborating to counter the alpha’s dominance—a positive feedback loop that leads to ever more complex socio-cognitive adaptations. Metarepresentation is among these adaptations, enabling individuals to interpret the behavior of conspecifics not just as bodily movement but as action guided by beliefs and desires (Sperber, 2000). Such reasoning helps individuals to protect themselves from others, to exploit them, and to collaborate with them. Hence, metarepresentation could have evolved independently of communication in response to social selection pressures (Scott-Phillips, 2014; Sperber, 2000; but see Armstrong, 2023). As a result, communication could emerge, in ontogeny and phylogeny alike, on the back of psychological metarepresentation (Baron-Cohen, 1999). Evidence for preverbal infants’ sensitivity to false beliefs in non-linguistic tasks provides some support for this (Kovács et al., 2010; Onishi & Baillargeon, 2005; Surian et al., 2007). Perhaps the main challenge for this approach is to explain how and why humans transitioned from a “perception-goal psychology” (Call & Tomasello, 2008), characteristic of non-human primates, to a belief-desire psychology, with its unique recursive structure, or, in other words, how higher-order representations emerged from primary representations.

Advocates of the third approach attempt to address the latter challenge. According to them, natural languages provide humans with representational tools that also enable expressing and entertaining propositional attitudes. These tools include recursion (e.g., embedding a sentence within a sentence), negation, and evidential marking (De Villiers, 2013; Moore, 2021). Besides, language reveals the logical structure of propositional attitudes which allows contrasting and combining them with the content of other propositional attitudes to yield further inferences (Bermúdez, 2017; Moore, 2021). Since we represent others’ mental states to use them in our own conscious practical decision-making, it is argued, they must be consciously accessible representations of an external language rather than sentences in a language of thought (Bermúdez, 2017). If true, this dependence on natural languages could imply that full-blown metarepresentations are the outcome of relatively recent cultural, rather than biological, evolution. Geurts (2021) notes that in many languages, quotative verbs are also used for attributing mental states, including beliefs and intentions. Thus, one can imagine an evolutionary trajectory from quotation to the public practice of attributing mental states (corresponding to the quoted expression) and eventually to implicit mental state attribution. Similarly, Moore (2021) suggests that human-specific forms of metapsychology are linguistically constructed folk models of the human mind which have been invented and modified for various purposes. According to him, beside quotation, a major source for propositional attitude concepts is perception verbs such as “see” and “hear” which have culturally evolved to attain cognitive senses. Proponents of this approach should explain how an inferential communica-

tive system that went beyond the simple coded signaling of non-human animals could get off the ground absent a sophisticated metapsychology (Sperber, 2000). More crucially, linguistic expressions are themselves external, representational devices that necessitate metarepresentation. Therefore, their existence cannot be taken for granted in accounting for the evolution of metarepresentation.

In the absence of direct paleoanthropological evidence for the emergence of meta-representation, any hypothesis about its evolutionary root will have to be largely speculative—unless, of course, we obtain sufficient evidence for metarepresentation in non-human primates (see also Krupenye et al., 2016). However, the plausible alternatives would still be worth considering. One of the main reasons for the popularity of the idea that psychological metarepresentation precedes communication (e.g., Baron-Cohen, 1999) is the assumption that the latter requires the former. However, if my account of communication is correct and communicative cognition emerges independently of mentalizing, we can also envisage the opposite direction: metarepresentation evolved to enable external, communicative representations and only later was it exapted for postulating mental states to interpret instrumental behavior.

This proposal is explanatorily more powerful than the language-first proposal, as it does not take linguistic representations for granted. Rather, it proposes a representational system that, on the one hand, permitted both linguistic and non-linguistic external representations, and, on the other, provided a platform for metapsychological representations (for a similar view see Armstrong, 2023). Like the language-first proposal, however, it suggests an ecology in which there are perceivable objects in the world that can promote incremental evolution from organisms lacking metarepresentation to ones with increasingly sophisticated metarepresentational capacities which they can exploit for various communicative, metacognitive, and metapsychological functions.

Consider two types of representing the representational medium: in the first (admittedly very shallow) type, the representational medium is represented and used in learning, but only as a non-representational entity; in the second (full-fledged) type, the representational medium is represented as a representation proper, that is, with a representational content. The first type is clearly simpler and can potentially support the evolution of the more sophisticated type. However, this is only possible with public representations, for mental representations are not available to perception; and even if they were somehow inferable, their representation as non-representational objects would be futile. As a result, taking metapsychology as the original function demands an evolutionary leap from organisms capable of only primary representations to ones with the ability to postulate abstract, higher-order constructs (i.e., mental states) directed at similarly abstract contents (e.g., a false belief-content or a future state of affairs).

The hypothesis that human communication evolved mainly for teaching knowledge to the offspring, plausible in its own right, offers the intriguing possibility of a scenario in which communicators *use* external entities representationally to transmit information about kinds, in the absence of specialized cognitive mechanisms. The teachers could, say, perform fitness-relevant actions (e.g., knapping flints) in the presence of their children. This already provides a suitable environment for learning. They could, additionally, monitor their children's gaze or emit

sounds to secure their attention. Because these cues (i.e., eye contact and vocalizations) are associated with adaptive information, children would benefit from evolving a preference for them—eventually promoting them to child-directed, ostensive signals. Moreover, as these proto-demonstrations often entail generalizable knowledge, it would be advantageous for the learner to develop a cognitive shortcut from the ostensibly marked action-object pair to the respective kinds, that is, interpret the demonstration as *representing* a generic predicate on an object kind. This would, thus, be an ecology in which actions have an aboutness—a feature which can promote specialized cognitive mechanisms capable of utilizing it efficiently. (The attraction of this scenario notwithstanding, one can imagine further alternative scenarios where actors manipulate public stimuli to convey information, and observers subsequently evolve a conceptual framework in which to make better sense of the stimuli.)

In addition, metarepresentation has some features that are instrumental for mentalizing but not for representing external representations. Decoupling, for instance, is an integral feature of mental state attribution, for otherwise someone else's belief or desire would be detrimentally taken as one's own. In a kin-selected communicative system (Fitch, 2004, 2007), however, you would be safe to encode and store the transmitted knowledge without needing to quarantine it. Likewise, if the parent is both competent and benevolent, you would not need to worry about whether the representation misrepresents the content—suggested to be necessary for a full understanding of representations (Perner, 1991).

Thus, at least in theory, metarepresenting representations of a public nature can have a relatively simpler structure. This, in turn, enables an incremental evolutionary trajectory from a non-representational understanding of representational entities to competence with recursive representations (i.e., representations of representations). Once recursive representation was in place for communication between kin, it could further be used to transmit information to non-kin. This extension in use would, however, necessitate some form of decoupling. Decoupled, recursive representations would not only enable communication but could be exploited also to attribute the mental states corresponding to the utterances. This representational format could ultimately be applied in domains where the medium is abstract and can only be contextually inferred. In this way, instrumental action interpretation would be significantly enhanced, for now the causes for behavior can be expanded beyond what is immediately observable. This application could be through an intermediary factor like language, but it could occur more directly through repurposing the same representational structure. Moreover, communication itself creates strong pressure for the evolution of increasingly sophisticated mindreading, augmenting the resources for successful communication.

One implication of considering external, communicative representations as the original domain of metarepresentation could be that what set apart human communication from the limited communication of other great apes was not cognitive constraints in recursive mindreading, but rather an environment that favored ever more flexible communication. Since other animals face only a limited range of prespecified signals in their lifetime, adequate for the recurring problems which drove their evolution, they do not need an encompassing “naïve

signaling theory”, that is, a metarepresentational concept of communication. However, an environment that involves ever-changing and cognitively opaque knowledge and technology, like the one our hominin ancestors inhabited (Boyd & Silk, 2014; Shea, 2016), fosters an open-ended system for the faithful transfer of information. Such an environment creates a “representational niche” in which it is beneficial to interpret (communicative) action as representationally conveying information that is applicable beyond the locally perceived behavior to displaced conceptual entities. As emphasized, this scenario is inevitably speculative. But if plausible, it can potentially change our perspective on the evolution of uniquely human forms of communication and perhaps social cognition.

7 Final remark

Human communication is commonly understood in terms of the intentional structure that is at play in its production and comprehension. This account has several shortcomings in explaining communicative cognition which were discussed above. Particularly, complex metarepresentations of intentions and beliefs are neither necessary nor sufficient in accounting for the design features of ostensive communication as a behavior and as a cognitive concept. Chiefly, the standard account may lead to overlooking informative (i.e., representational) action—common across our diverse uses of communication. As an evolutionary account, intentionalism has arguably hindered progress in comparative research. Firstly, emphasizing the intentions behind human communication obscures what is truly unique about it and leads one to seek its origin in the intentionality of primate communication (e.g., Zuberbühler, 2018)—rather than, for instance, behavior that is potentially homologous with representational communication. Secondly, as the purported mental states are inaccessible, devising paradigms in which to test similar traits in non-human animals will prove difficult, if not impossible. However, if my action-based account is plausible, one may conduct studies which can empirically test whether other primates are capable of flexibly marking entities as communicative, and, more pertinent to the present paper, whether they interpret unfamiliar stimuli as representing a detached content. Such studies will likely move research on the evolutionary origin of human communication forward and shed light on what genuinely separates (or unifies) our interactions and those of our primate cousins.

Acknowledgements I thank Gergely Csibra and Dan Sperber for discussions that led to the current paper. I also thank Richard Moore, Anne Reboul, Thom Scott-Phillips, Nausicaa Pouscoulous, Martin Dockendorff, Antonio Scarafone, and Barbu Revencu for their useful comments. Finally, I thank Francesca Bonalumi and Simily Sabu for helping me with the preparation of the manuscript.

Funding Open access funding provided by Central European University Private University. This work was supported by a European Research Council Advanced Investigator Grant (#742231 “PARTNERS”).

Declarations

Conflict of interest None.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Armstrong, J. (2023). Communication before communicative intentions. *Noûs*, 57(1), 26–50.
- Austin, J. L. (1962). *How to do things with words*. Oxford University Press.
- Bach, K. (1987). On communicative intentions: A reply to Recanti. *Mind & Language*, 2(2), 141–154. <https://doi.org/10.1111/j.1468-0017.1987.tb00112.x>.
- Bach, K. (2012). Meaning and communication. In G. Russell, & D. G. Fara (Eds.), *Routledge Companion to the Philosophy of Language* (pp. 79–90). Routledge.
- Bach, K., & Harnish, R. M. (1979). *Linguistic communication and Speech acts*. MIT Press.
- Bar-On, D. (2013). Origins of meaning: Must we ‘Go Gricean’? *Mind & Language*, 28(3), 342–375. <https://doi.org/10.1111/mila.12021>.
- Baron-Cohen, S. (1999). The evolution of a theory of mind. In M. C. Corballis, & S. E. G. Lea (Eds.), *The descent of mind: Psychological perspectives on hominid evolution* (pp. 261–277). Oxford University Press.
- Bermúdez, J. L. (2017). Can nonlinguistic animals think about thinking? In K. Andrews, & J. Beck (Eds.), *The Routledge handbook of philosophy of animal minds* (pp. 119–130). Routledge.
- Bloom, P. (2000). *How children learn the meanings of words*. The MIT Press.
- Bohn, M., & Frank, M. C. (2019). The pervasive role of pragmatics in early language. *Annual Review of Developmental Psychology*, 1, 223–249.
- Boyd, R., & Silk, J. B. (2014). *How Humans Evolved: Seventh Edition*. W. W. Norton & Company.
- Breiheny, R. (2006). Communication and folk psychology. *Mind & Language*, 21(1), 74–107. <https://doi.org/10.1111/j.1468-0017.2006.00307.x>.
- Byrne, R. W. (1996). Machiavellian intelligence. *Evolutionary Anthropology: Issues News and Reviews*, 5(5), 172–180. [https://doi.org/10.1002/\(SICI\)1520-6505\(1996\)5:5<172::AID-EVAN6>3.0.CO;2-H](https://doi.org/10.1002/(SICI)1520-6505(1996)5:5<172::AID-EVAN6>3.0.CO;2-H).
- Byrne, R. W., & Whiten, A. (1997). Machiavellian intelligence. In A. Whiten, & R. W. Byrne (Eds.), *Machiavellian Intelligence II: Extensions and evaluations* (pp. 1–23). Cambridge University Press. <https://doi.org/10.1017/CBO9780511525636.002>.
- Byrne, R. W., Cartmill, E., Genty, E., Graham, K. E., Hobaiter, C., & Tanner, J. (2017). Great ape gestures: Intentional communication with a rich set of innate signals. *Animal Cognition*, 20(4), 755–769. <https://doi.org/10.1007/s10071-017-1096-4>.
- Call, J., & Tomasello, M. (2008). Does the chimpanzee have a theory of mind? 30 years later. *Trends in Cognitive Sciences*, 12(5), 187–192. <https://doi.org/10.1016/j.tics.2008.02.010>.
- Carey, S. (2009). *The origin of concepts*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195367638.001.0001>.
- Clark, H. H. (2016). Depicting as a method of communication. *Psychological Review*, 123(3), 324–347. <https://doi.org/10.1037/rev0000026>.
- Cosmides, L., & Tooby, J. (2000). Consider the source: The evolution of adaptations for decoupling and metarepresentation. In D. Sperber (Ed.), *Metarepresentations: A multidisciplinary perspective* (pp. 53–115). Oxford University Press.
- Couchman, J. J., Coutinho, M. V. C., Beran, M. J., & Smith, J. D. (2009). Metacognition is prior. *Behavioral and Brain Sciences*, 32(2), 142–142. <https://doi.org/10.1017/S0140525X09000594>.
- Csibra, G. (2010). Recognizing communicative intentions in infancy. *Mind & Language*, 25(2), 141–168. <https://doi.org/10.1111/j.1468-0017.2009.01384.x>.
- Csibra, G., & Gergely, G. (2006). Social learning and social cognition: The case for pedagogy. In M. H. Johnson & Y. Munakata (Eds.), *Processes of Change in Brain and Cognitive Development. Attention and Performance XXI*, (pp. 249–274). OUP Oxford.

- Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences*, 13(4), 148–153.
- Csibra, G., & Gergely, G. (2011). Natural pedagogy as evolutionary adaptation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366(1567), 1149–1157.
- De Villiers, J. (2013). Language and reasoning about beliefs. In M. Banaji, & S. Gelman (Eds.), *Navigating the Social World* (pp. 96–100). Oxford University Press.
- Dennett, D. C. (1987). *The intentional stance*. MIT Press.
- Egyed, K., Király, I., & Gergely, G. (2013). Communicating shared knowledge in infancy. *Psychological Science*, 24(7), 1348–1353. <https://doi.org/10.1177/0956797612471952>
- Ferguson, B., & Waxman, S. R. (2016). What the [beep]? Six-month-olds link novel communicative signals to meaning. *Cognition*, 146, 185–189. <https://doi.org/10.1016/j.cognition.2015.09.020>
- Fitch, W. T. (2004). Kin selection and “Mother tongues”: A neglected component in language evolution. In D. K. Oller & U. Griebel (Eds.), *Evolution of Communication Systems: A Comparative Approach*, (pp. 275–296). MIT Press.
- Fitch, W. T. (2007). Evolving meaning: The roles of kin selection, allomothering and paternal care in language evolution. In C. Lyon, C. L. Nehaniv, & A. Cangelosi (Eds.), *Emergence of Communication and Language* (pp. 29–51). Springer. https://doi.org/10.1007/978-1-84628-779-4_2
- Fodor, J. A. (1992). A theory of the child’s theory of mind. *Cognition*, 44(3), 283–296. [https://doi.org/10.1016/0010-0277\(92\)90004-2](https://doi.org/10.1016/0010-0277(92)90004-2).
- Gärdenfors, P. (1995). Cued and detached representations in animal cognition. *Behavioural Processes*, 35(1), 263–273. [https://doi.org/10.1016/0376-6357\(95\)00043-7](https://doi.org/10.1016/0376-6357(95)00043-7).
- Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: The naïve theory of rational action. *Trends in Cognitive Sciences*, 7(7), 287–292. [https://doi.org/10.1016/S1364-6613\(03\)00128-1](https://doi.org/10.1016/S1364-6613(03)00128-1).
- Geurts, B. (2018). Convention and common ground. *Mind & Language*, 33(2), 115–129.
- Geurts, B. (2019). Communication as commitment sharing: Speech acts, implicatures, common ground. *Theoretical Linguistics*, 45(1–2), 1–30. <https://doi.org/10.1515/tl-2019-0001>.
- Geurts, B. (2021). First saying, then believing: The pragmatic roots of folk psychology. *Mind & Language*, 36(4), 515–532. <https://doi.org/10.1111/mila.12345>.
- Geurts, B. (2022). Evolutionary pragmatics: From chimp-style communication to human discourse. *Journal of Pragmatics*, 200, 24–34.
- Gillespie-Lynch, K., Greenfield, P. M., Lyn, H., & Savage-Rumbaugh, S. (2014). Gestural and symbolic development among apes and humans: Support for a multimodal theory of language evolution. *Frontiers in Psychology*, 5. <https://www.frontiersin.org/articles/https://doi.org/10.3389/fpsyg.2014.01228>.
- Gómez, J. C. (1994). Mutual awareness in primate communication: A Gricean approach. In S. T. Parker, R. M. Mitchell, & M. L. Boccia (Eds.), *Self-awareness in animals and humans: Developmental perspectives* (pp. 61–80). Cambridge University Press. <https://doi.org/10.1017/CBO9780511565526.007>.
- Gómez, J. C. (1996). Ostensive behavior in great apes: The role of eye contact. In A. E. Russon, K. A. Bard, & S. T. Parker (Eds.), *Reaching into thought: The minds of the great apes* (pp. 131–151). Cambridge University Press.
- Gopnik, A. (2011). A unified account of abstract structure and conceptual change: Probabilistic models and early learning mechanisms. *Behavioral and Brain Sciences*, 34(3), 129.
- Green, M. (2019). Organic Meaning: An Approach to Communication with Minimal Appeal to Minds. In A. Capone, M. Carapezza, & F. Lo Piparo (Eds.), *Further Advances in Pragmatics and Philosophy: Part 2 Theories and Applications* (pp. 211–228). Springer International Publishing. https://doi.org/10.1007/978-3-030-00973-1_12.
- Grice, H. P. (1957). Meaning. *The Philosophical Review*, 66(3), 377–388. <https://doi.org/10.2307/2182440>.
- Grice, H. P. (1969). Utterer’s meaning and intentions. *The Philosophical Review*, 78(2), 147–177.
- Grice, H. P. (1975). Logic and Conversation. In D. Davidson & G. Harman (Eds.), *The Logic of Grammar* (pp. 64–75).
- Grice, H. P. (1982). Meaning revisited. *Mutual Knowledge*, 222–243.
- Grosse, G., Scott-Phillips, T. C., & Tomasello, M. (2013). Three-year-olds hide their communicative intentions in appropriate contexts. *Developmental Psychology*, 49(11), 2095–2101. <https://doi.org/10.1037/a0032017>.
- Harris, D. W. (2019). Intention Recognition as the Mechanism of Human Communication. In A. Sullivan (Ed.), *Sensations, Thoughts, Language: Essays in Honor of Brian Loar*.
- Heintz, C., & Scott-Phillips, T. (2023). Expression unleashed: The evolutionary and cognitive foundations of human communication. *Behavioral and Brain Sciences*, 46, e1.
- Ittelson, W. H. (1996). Visual perception of markings. *Psychonomic Bulletin & Review*, 3(2), 171–187. <https://doi.org/10.3758/BF03212416>.

- Király, I., Csibra, G., & Gergely, G. (2013). Beyond rational imitation: Learning arbitrary means actions from communicative demonstrations. *Journal of Experimental Child Psychology*, 116(2), 471–486. <https://doi.org/10.1016/j.jecp.2012.12.003>
- Kovács, Á. M., Téglás, E., & Endress, A. D. (2010). The social sense: Susceptibility to others' beliefs in human infants and adults. *Science*, 330(6012), 1830–1834. <https://doi.org/10.1126/science.1190792>.
- Krupenye, C., Kano, F., Hirata, S., Call, J., & Tomasello, M. (2016). Great apes anticipate that other individuals will act according to false beliefs. *Science*, 354(6308), 110–114. <https://doi.org/10.1126/science.aaf8110>.
- Laland, K. (2017). The origins of language in teaching. *Psychonomic Bulletin & Review*, 24(1), 225–231. <https://doi.org/10.3758/s13423-016-1077-7>
- Leslie, A. M. (1987). Pretense and representation: The origins of theory of mind. *Psychological Review*, 94(4), 412–426. <https://doi.org/10.1037/0033-295X.94.4.412>.
- Leslie, A. M., Friedman, O., & German, T. P. (2004). Core mechanisms in 'theory of mind'. *Trends in Cognitive Sciences*, 8(12), 528–533. <https://doi.org/10.1016/j.tics.2004.10.001>.
- Lewis, D. K. (1969). *Convention: A philosophical study*. Wiley-Blackwell.
- Malle, B. F. (2002). The relation between language and theory of mind in development and evolution. In B. F. Malle, & T. Givón (Eds.), *The evolution of Language out of Pre-language* (Vol. 18, pp. 265–284). John Benjamins Publishing.
- Marno, H., & Csibra, G. (2015). Toddlers Favor Communicatively Presented Information over statistical reliability in learning about Artifacts. *PLOS ONE*, 10(3), e0122129. <https://doi.org/10.1371/journal.pone.0122129>.
- Millikan, R. G. (1987). *Language, thought, and other biological categories: New foundations for realism*. MIT press.
- Moore, R. (2014). Ontogenetic constraints on Grice's theory of communication. In D. Matthews (Ed.), *Pragmatic development in First Language Acquisition* (pp. 87–104). John Benjamins Publishing Company.
- Moore, R. (2016). Gricean Communication and Cognitive Development. *The Philosophical Quarterly*, pqw049. <https://doi.org/10.1093/pq/pqw049>.
- Moore, R. (2018). Gricean communication, language development, and animal minds. *Philosophy Compass*, 13(12), e12550. <https://doi.org/10.1111/phc3.12550>.
- Moore, R. (2021). The cultural evolution of mind-modelling. *Synthese*, 199(1), 1751–1776. <https://doi.org/10.1007/s11229-020-02853-3>.
- Neale, S. (1992). Paul Grice and the Philosophy of Language. *Linguistics and Philosophy*, 15(5), 509–559.
- Novack, M. A., & Waxman, S. (2020). Becoming human: Human infants link language and cognition, but what about the other great apes? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 375(1789), 20180408. <https://doi.org/10.1098/rstb.2018.0408>.
- Novack, M. A., Henderson, A. M. E., & Woodward, A. L. (2014). Twelve-Month-Old infants Generalize Novel signed labels, but not preferences across individuals. *Journal of Cognition and Development*, 15(4), 539–550. <https://doi.org/10.1080/15248372.2013.782460>.
- Onishi, K. H., & Baillargeon, R. (2005). Do 15-Month-Old infants understand false beliefs? *Science*, 308(5719), 255–258. <https://doi.org/10.1126/science.1107621>.
- Perner, J. (1991). *Understanding the representational mind*. The MIT Press.
- Piccinini, G., & Scarantino, A. (2011). Information processing, computation, and cognition. *Journal of Biological Physics*, 37(1), 1–38. <https://doi.org/10.1007/s10867-010-9195-3>.
- Pinker, S. (2010). The cognitive niche: Coevolution of intelligence, sociality, and language. *Proceedings of the National Academy of Sciences*, 107(supplement_2), 8993–8999. <https://doi.org/10.1073/pnas.0914630107>.
- Planer, R. J. (2017a). Protolanguage might have evolved before Ostensive Communication. *Biological Theory*, 12(2), 72–84. <https://doi.org/10.1007/s13752-017-0262-x>.
- Planer, R. J. (2017b). Talking about Tools: Did early pleistocene Hominins have a Protolanguage? *Biological Theory*, 12(4), 211–221. <https://doi.org/10.1007/s13752-017-0279-1>.
- Planer, R. J. (2021). What is symbolic cognition? *Topoi*, 40(1), 233–244. <https://doi.org/10.1007/s11245-019-09670-5>.
- Pomiechowska, B., & Csibra, G. (2022). Nonverbal Action Interpretation guides Novel Word Disambiguation in 12-Month-Olds. *Open Mind*, 6, 51–76. https://doi.org/10.1162/opmi_a_00055.
- Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, 1(4), 515–526. <https://doi.org/10.1017/S0140525X00076512>.

- Pylyshyn, Z. W. (1978). When is attribution of beliefs justified? [P&W]. *Behavioral and Brain Sciences*, 1(4), 592–593. <https://doi.org/10.1017/S0140525X00076895>.
- Rakoczy, H., & Behne, T. (2019). Commitment sharing as crucial step toward a developmentally plausible speech act theory? *Theoretical Linguistics*, 45(1–2), 93–97. <https://doi.org/10.1515/tl-2019-0007>.
- Recanati, F. (1986). On defining communicative intentions. *Mind & Language*, 1(3), 213–241. <https://doi.org/10.1111/j.1468-0017.1986.tb00102.x>.
- Recanati, F. (2002). Does linguistic communication rest on inference? *Mind & Language*, 17(1–2), 105–126. <https://doi.org/10.1111/1468-0017.00191>.
- Revcnu, B., & Csibra, G. (2021). For 19-Month-Olds, what happens On-Screen stays On-Screen. *Open Mind*, 5, 71–90. https://doi.org/10.1162/opmi_a_00043.
- Sato, Y., Kano, F., Morimura, N., Tomonaga, M., & Hirata, S. (2022). Chimpanzees (Pan troglodytes) exhibit gaze bias for snakes upon hearing alarm calls. *Journal of Comparative Psychology*, 136(1), 44–53. <https://doi.org/10.1037/com0000305>.
- Scarafone, A., & Michael, J. (2022). Getting ready to share commitments. *Philosophical Topics*, 50(1), 135–160.
- Scarantino, A., & Clay, Z. (2015). Contextually variable signals can be functionally referential. *Animal Behaviour Complete*, (100), e1–e8. <https://doi.org/10.1016/j.anbehav.2014.08.017>.
- Scarantino, A., & Piccinini, G. (2010). Information without Truth. *Metaphilosophy*, 41(3), 313–330. <https://doi.org/10.1111/j.1467-9973.2010.01632.x>.
- Schiffer, S. R. (1972). *Meaning*. Oxford, Clarendon Press.
- Scott-Phillips, T. (2014). *Speaking our minds: Why human communication is different, and how language evolved to make it special*. Bloomsbury Publishing.
- Scott-Phillips, T. (2015). Nonhuman Primate Communication, Pragmatics, and the origins of Language. *Current Anthropology*, 56(1), 56–80. <https://doi.org/10.1086/679674>.
- Searle, J. R. (1983). *Intentionality: An essay in the philosophy of mind*. Cambridge University Press.
- Shea, J. J. (2016). *Stone tools in human evolution: Behavioral differences among technological primates*. Cambridge University Press.
- Spelke, E. S. (2003). What makes us smart? Core Knowledge and Natural Language. In D. Getner, & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of Language and Thought* (pp. 277–311). MIT Press.
- Sperber, D. (1994a). The modularity of thought and the epidemiology of representations. In L. A. Hirschfeld, & S. A. Gelman (Eds.), *Mapping the mind: Domain specificity in Cognition and Culture* (pp. 39–67). Cambridge University Press.
- Sperber, D. (1994b). Understanding Verbal understanding. In J. Khalfa (Ed.), *What is Intelligence?* (pp. 179–198). Cambridge University Press.
- Sperber, D. (2000). Metarepresentations in an evolutionary perspective. In D. Sperber (Ed.), *Metarepresentations: A multidisciplinary perspective* (pp. 117–137). Oxford University Press.
- Sperber, D. (2019). Personal notes on a Shared Trajectory. In K. Scott, B. Clark, & R. Carston (Eds.), *Relevance, Pragmatics and Interpretation* (pp. 13–20). Cambridge University Press. <https://doi.org/10.1017/9781108290593.002>.
- Sperber, D., & Hirschfeld, L. A. (2004). The cognitive foundations of cultural stability and diversity. *Trends in Cognitive Sciences*, 8(1), 40–46. <https://doi.org/10.1016/j.tics.2003.11.002>.
- Sperber, D., & Wilson, D. (1995). *Relevance: Communication and cognition* (2nd ed.). Blackwell Publishing.
- Sperber, D., & Wilson, D. (2002). Pragmatics, modularity and mind-reading. *Mind & Language*, 17(1–2), 3–23. <https://doi.org/10.1111/1468-0017.00186>.
- Sperber, D., & Wilson, D. (2015). Beyond Speaker's meaning. *Croatian Journal of Philosophy*, 15(44), 117–149.
- Sterelny, K. (2017). From code to speaker meaning. *Biology & Philosophy*, 32(6), 819–838. <https://doi.org/10.1007/s10539-017-9597-8>.
- Strawson, P. F. (1964). Intention and convention in Speech acts. *The Philosophical Review*, 73(4), 439–460.
- Suddendorf, T. (1999). The rise of the metamind. In M. C. Corballis, & S. E. G. Lea (Eds.), *The descent of mind: Psychological perspectives on hominid evolution* (pp. 218–260). Oxford University Press.
- Surian, L., Caldi, S., & Sperber, D. (2007). Attribution of beliefs by 13-Month-Old infants. *Psychological Science*, 18(7), 580–586. <https://doi.org/10.1111/j.1467-9280.2007.01943.x>.
- Suzuki, T. (2018). Alarm calls evoke a visual search image of a predator in birds. *Proceedings of the National Academy of Sciences*, 115, 201718884. <https://doi.org/10.1073/pnas.1718884115>.

- Tauzin, T., & Gergely, G. (2018). Communicative mind-reading in preverbal infants. *Scientific Reports*, 8(1), <https://doi.org/10.1038/s41598-018-27804-4>.
- Thompson, J. R. (2014). Meaning and mindreading. *Mind & Language*, 29(2), 167–200. <https://doi.org/10.1111/mila.12046>.
- Tomasello, M. (2008). *Origins of human communication*. MIT Press.
- Tomasello, M., & Call, J. (2019). Thirty years of great ape gestures. *Animal Cognition*, 22(4), 461–469. <https://doi.org/10.1007/s10071-018-1167-1>.
- Topál, J., Gergely, G., Miklósi, Á., Erdőhegyi, Á., & Csibra, G. (2008). Infants' perseverative search errors are Induced by pragmatic misinterpretation. *Science*, 321(5897), 1831–1834. <https://doi.org/10.1126/science.1161437>.
- Topál, J., Gergely, G., Erdőhegyi, Á., Csibra, G., & Miklósi, Á. (2009). Differential Sensitivity to Human Communication in Dogs, wolves, and human infants. *Science*, 325(5945), 1269–1272. <https://doi.org/10.1126/science.1176960>.
- Vail, A. L., Manica, A., & Bshary, R. (2013). Referential gestures in fish collaborative hunting. *Nature Communications*, 4(1), <https://doi.org/10.1038/ncomms2781>.
- Vouloumanos, A., Martin, A., & Onishi, K. H. (2014). Do 6-month-olds understand that speech can communicate? *Developmental Science*, 17(6), 872–879. <https://doi.org/10.1111/desc.12170>.
- Warren, E., & Call, J. (2022). Inferential Communication: Bridging the Gap Between Intentional and Ostensive Communication in Non-human Primates. *Frontiers in Psychology*, 12. <https://www.frontiersin.org/articles/https://doi.org/10.3389/fpsyg.2021.718251>.
- Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development*, 72(3), 655–684. <https://doi.org/10.1111/1467-8624.00304>.
- Whiten, A., & Erdal, D. (2012). The human socio-cognitive niche and its evolutionary origins. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1599), 2119–2129. <https://doi.org/10.1098/rstb.2012.0114>.
- Wilson, D., & Sperber, D. (2012). *Meaning and relevance*. Cambridge University Press.
- Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13(1), 103–128. [https://doi.org/10.1016/0010-0277\(83\)90004-5](https://doi.org/10.1016/0010-0277(83)90004-5).
- Woensdregt, M., & Smith, K. (2017). Pragmatics and Language Evolution. *Oxford Research Encyclopedia of Linguistics*. <https://doi.org/10.1093/acrefore/9780199384655.013.321>.
- Xu, F. (2019). Towards a rational constructivist theory of cognitive development. *Psychological Review*, 126(6), 841–864. <https://doi.org/10.1037/rev0000153>.
- Zuberbühler, K. (2018). Intentional communication in primates. *Revue Tranel*, 68, 69–75.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.