Higher-order theory of mind in Tacit Communication Game

Harmen de Weerd[1], Rineke Verbrugge[1], and Bart Verheij[1,2]

[1] Institute of Artificial Intelligence, Faculty of Mathematics and Natural Sciences, University of Groningen, The Netherlands
[2] CodeX, Stanford University, California, United States

Abstract
To understand and predict the behaviour of others, people regularly reason about others' beliefs, goals, and intentions. People can even use this theory of mind recursively, and form beliefs about the way others in turn reason about the beliefs, goals, and intentions of others. Although the evolutionary origins of this cognitively demanding ability are unknown, the Vygotskian intelligence hypothesis suggests that higher-order theory of mind allows individuals to cooperate more effectively. In this paper, we investigate this hypothesis through the Tacit Communication Game. In this game, two agents cooperate to set up novel communication in which a Sender agent communicates the goal to the Receiver agent. By simulating interactions between agents that differ in their theory of mind abilities, we determine to what extent higher orders of theory of mind help agents to set up communication. Our results show that first-order and second-order theory of mind can allow agents to set up communication more quickly, but also that the effectiveness of higher orders of theory of mind depends on the role of the agent. Additionally, we find that in some cases, agents cooperate more effectively if they reason at lower orders of theory of mind.

Keywords: theory of mind, depth of reasoning, evolution, learning, level-k, fictitious play, cognitive hierarchy, communication, simulation

1 Introduction

While engaging in everyday activities, people regularly make use of *theory of mind* [32], and reason about what others know and believe. For example, people reason about the reasons that others might have to behave the way they do, and distinguish actions that are intentional from those that are accidental. People are also able to use this ability recursively, and use *second-order theory of mind* to reason about the way others reason about beliefs and goals. This allows them to understand sentences such as 'Alice believes that Bob doesn't know that Carol is throwing him a surprise party', and make predictions about how this knowledge influences the behaviour of Alice.

The human ability to make use of higher-order (i.e. at least second-order) theory of mind has been demonstrated both in tasks that require explicit reasoning using second-order beliefs

[1, 31] as well as in strategic games [20, 24]. However, the use of higher-order theory of mind appears to be a uniquely human ability; whether any non-human species is able to make use of theory of mind of any kind is under debate [6, 30, 39, 42, 23]. This suggests that there may be specific settings in which the ability to reason about the unobservable mental states of others is evolutionarily advantageous. Such settings would support the emergence of higher-order theory of mind, despite the high cognitive demands of such an ability.

According to the Vygotskian intelligence hypothesis [44, 25], cooperative social interactions play a crucial role in the development of human cognitive skills such as theory of mind. Tomasello et al. [40] propose that the uniquely human aspects of cognition that require higher-order theory of mind, such as constructing shared goals and joint intentions, developed because of the need for social cooperation. They state that higher-order theory of mind allows individuals to achieve levels of cooperation beyond those that are achieved by individuals that are unable to reason about the minds of others, and that are therefore unable to construct shared goals and joint intentions [1].

Although the Vygotskian intelligence hypothesis suggests that theory of mind is necessary in some cooperative settings, simulation studies have shown that many forms of cooperation can evolve using simple mechanisms [4, 29, 34, 17, 41]. Many animals are known to engage in cooperative interactions without relying on higher-order theory of mind [45, 14, 7, 39]. Even highly organized and complex cooperative behaviour such as cooperative hunting by lions, wolves, and chimpanzees can be described using fixed roles that rely on simple cues, without the need to construct joint intentions [40, 39, 26]. In fact, even when cooperation is risky, simple punishment strategies are enough to stabilize cooperation [5].

Since many forms of cooperation can be stabilized without the use of theory of mind, we instead focus on the process of establishing and coordinating cooperation between agents. We investigate a particular form of cooperation through the Tacit Communication Game [11, 27, 2]. In this game, a pair of players needs to set up communication that allows one player to inform the other player about the goal of the game. We make use of agent-based computational models to investigate how higher-order theory of mind can help agents to set up communication. Agent-based models have previously been used to explore the origins of communication [37, 33, 9, 38], as well as to determine the effectiveness of higher order of theory of mind in competitive settings [13] and mixed-motive settings [12]. By simulating interactions between computational agents, we can determine how the theory of mind abilities of these agents influences their performance.

The remainder of this paper is set up as follows. In Section 2, we present one particular way in which theory of mind can be helpful in communication. Section 3 introduces the Tacit Communication Game. In Section 4, we describe the agent model and show how the ability to reason at higher orders of theory of mind changes the way agents play the Tacit Communication Game. The details and the results of the simulated interactions between agents are found in Section 6. Section 7 discusses our results, and compares these results to related work.

2 Theory of mind in communication

The Vygotskian intelligence hypothesis suggests that there are forms of cooperation that support the emergence of higher-order theory of mind. One way in which theory of mind may be beneficial for social cooperation is through communication. Bidirectional optimality theory [3] suggests that in order to establish proper communication, the speaker has to take into account the perspective of the hearer, while the hearer has to take into account the perspective

[1] For the computational complexity of joint intentions, see Dziubiński et al. [15].

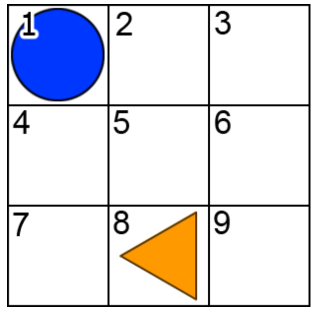


Figure 1: Example of a goal configuration in the Tacit Communication Game. In this example, the Sender has to position his (dark blue) token on tile 1 at the top left location, while the Receiver has to position and orient her (light orange) token as shown on tile 8, at the central bottom location.

of the speaker. This may require a recursive mechanism of bidirectional optimization similar to second-order theory of mind [16]. In this system, a hearer decides on the correct interpretation of a sentence by making use of second-order theory of mind, by considering how the speaker would predict the hearer to interpret the sentence. De Hoop and Krämer [10] argue that an inability to optimize bidirectionally is the reason for children to make mistakes in the interpretation of the following Dutch sentence S:

S: Er ging twee keer een meisje van de glijbaan af.
there went two time a girl of the slide down
"Twice a girl went down the slide."

Sentence S allows for two different interpretations. The sentence can be interpreted as a single girl that went down the slide twice, or two separate girls that went down the slide once. Although Dutch children accept both interpretations, adults prefer the interpretation in which two separate girls went down the slide once.

According to De Hoop and Krämer [10], bidirectional optimization accounts for adults' interpretation of the indefinite subject *een meisje* ('a girl') in sentence S. The preferred interpretation is that the indefinite subject *een meisje* refers to a particular girl rather than to any girl. However, the canonical word order in Dutch sentences has the subject appear in the initial position (as in "Een meisje ging twee keer van de glijbaan af"), while the subject in sentence S appears sentence-internally. Since the canonical word order expresses the preferred referential meaning, this reading is blocked for sentence S. De Hoop and Krämer [10] argue that to arrive at this conclusion, the hearer has to reason that if the speaker intended to express the preferred referential meaning, he would have used the canonical word order. That is, the hearer reasons – though typically subconsciously – that the speaker believes that if the speaker uses the canonical word order, the hearer will interpret this as expressing the preferred referential meaning. As a result, the hearer believes that the speaker will choose another, non-canonical, form to express the non-canonical meaning.

3 Game setting

To determine the effectiveness of higher-order theory of mind in cooperative settings, we compare performance of computational agents in the setting of the Tacit Communication Game (TCG) [11, 27, 2]. The Tacit Communication Game is a purely cooperative form of signaling game [22, 35, 36]. The game is played by two players, one of which is assigned the role of the *Sender* while the other player is assigned the role of the *Receiver*. To avoid ambiguity, we will refer to the Sender as if he were male, and to the Receiver as if she were female.

The Sender and the Receiver both control a separate token, which are initially placed in the center of a 3 by 3 grid. The game is separated into two phases. During the first phase, only the Sender is allowed to move his token, while only the Receiver is allowed to move her token during the second phase. During their own phase, players can move their own token horizontally and vertically, as well as rotate their own token any number of times. The object of the game is for both players to move and rotate their token so that it matches a given goal configuration when the trial ends. An example of a possible goal configuration is shown in Figure 1. In this situation, the final position of the Sender's dark blue circle token should be on tile 1, while the Receiver should place her light orange triangular token on tile 8, in the shown orientation.

Crucially, the Sender is the only player who knows the goal configuration. Since the Sender can only move during the first phase, the Sender has to decide how to match his own token to the goal configuration, but also how to communicate to the Receiver what the position and orientation of her token is in the goal configuration. Note that the Sender and the Receiver are not allowed to communicate outside of the game. That is, the only way for the Sender to communicate the goal configuration to the Receiver is through movement and rotation of his token.

Once the Sender and the Receiver have moved their respective tokens, the trials ends. At the end of a trial, both the Sender and the Receiver learn whether they were successful or unsuccessful in matching the goal configuration, without further feedback. Therefore, when a trial ends unsuccessfully, the Receiver is left unaware of the actual goal configuration.

De Ruiter et al. [11] report that human participants are highly proficient in communicating the goal configuration. Pairs of participants successfully matched the goal configuration in over 95% of the conventional communicative trials, in which the Sender and the Receiver had a token of the same shape. Although some variation in participant strategies exists, the Sender typically matches his token to the goal configuration of the Receiver's token first and pause before he matches his token to the goal configuration.

Participants experience more difficulty in unconventional communicative trails, where the Sender controlled a token that had a shape that could not match the orientation of the token of the Receiver. Figure 1 shows an example of an unconventional communicative trial, where the Sender has to communicate the position and orientation of the Receiver's triangular token. However, since the Sender controls a circle himself, he cannot use the orientation of his own token to communicate the orientation of the token of the Receiver. Participant strategies in these trials vary more than in conventional communicative trials. The most common strategy is for the Sender to move back and forth ('wiggling') between two tiles to indicate the orientation of the Responder's token. However, this strategy cannot be used for all possible orientations of the Receiver's token, while other, more successful Sender strategies are not always understood. Nevertheless, participants scored significantly above chance levels in unconventional communicative trials as well.

In this paper, we only consider conventional communicative trials, in which both the Sender and the Receiver control a circle token. That is, agents only need to match the location of their respective tokens to the goal configuration. The Sender sends a *message* by performing a sequence of *movements* and *pauses* on the game board. The Receiver responds by selecting the tile she wants to position her token on. We assume that all actions take the same amount of time. However, the Sender can vary the speed at which he performs actions through the use of a *pause* action. More formally, the set of all possible messages M is such that

1. atomic actions 'u', 'd', 'l', 'r', and '.' are messages in M;

2. if m and n are messages in M, then the concatenation mn is also a message in M.

The atomic actions 'u', 'd', 'l', and 'r' correspond to the Sender moving his token up, down, left, and right, respectively. The action '.' corresponds to the Sender pausing, which does not affect the position of his token.

Since the Tacit Communication Game is a purely cooperative game, once the Sender and the Receiver have successfully used some message m to communicate a given goal configuration l, both players have an incentive to associate that message m with goal configuration l in the future as well. In game-theoretic terms, each successful communication strategy corresponds to a Nash equilibrium in the Tacit Communication Game. Once such a Nash equilibrium has been selected, theory of mind is unlikely to have any role in stabilizing cooperation in the Tacit Communication Game. However, there are many such Nash equilibria that a pair of agents may choose from. Theory of mind may have a role in the process of selecting such a Nash equilibrium, rather than stabilizing the choice of Nash equilibrium once it has been chosen. We hypothesize that higher orders of theory of mind allow agents to establish communication faster. That is, we expect that agents that make use of higher orders of theory of mind need fewer trials to select a Nash equilibrium than agents without the ability to reason about the mental content of others.

In terms of classical game theory, the Tacit Communication Game can be thought of as a coordination game. The Tacit Communication Game allows for many different Nash equilibria, each of which is a possible pair of communication strategies that solves the coordination problem. However, unlike these Nash equilibria, the communication strategies constructed by human participants show a remarkable similarity. We therefore investigate whether a bias in the population to prefer certain messages over others influences the effectiveness of theory of mind. In Setting 1, Senders prefer to send shorter messages over longer messages; a Sender that believes two messages m_1 and m_2 to be equally successful prefers to send the shorter message. In Setting 2, Senders have a bias towards more human-like strategies. In this setting, Senders prefer shorter messages that pass through the goal location of the Receiver.

4 Theory of mind in the Tacit Communication Game

In this section, we describe how agents can make use of theory of mind while they play the Tacit Communication Game. These agents make use of simulation-theory of mind [8, 28, 21], and are designed similarly to those we used to determine the effectiveness of higher orders of theory of mind in competitive settings [13] and mixed-motive settings [12]. Using simulation-theory of mind, an agent can take the perspective of his partner and determine how the agent would act himself if he had been in that situation. Using the implicit assumption that the thought process of any other agent can be accurately modeled by his own thought process, the agent predicts that his partner will make the same decision the agent would have made himself if the roles had been reversed.

In the following subsections, we describe how this process of perspective taking results in different behaviour for agents of different orders of theory of mind playing the Tacit Communication Game. The formal description of these theory of mind agents is presented in Section 5. In the remainder, we will speak of a ToM_k agent to indicate an agent that has the ability to use theory of mind up to and including the k-th order, but not beyond.

4.1 Zero-order theory of mind

Zero-order theory of mind (ToM_0) agents are unable to reason about the mental content of others. In particular, this means that ToM_0 agents cannot represent the fact that their partner has the same goal as they have: to match his or her token to the goal configuration. Instead, a ToM_0 agent attempts to find out what behaviour is more likely to result in both agents matching their respective tokens to the goal configuration.

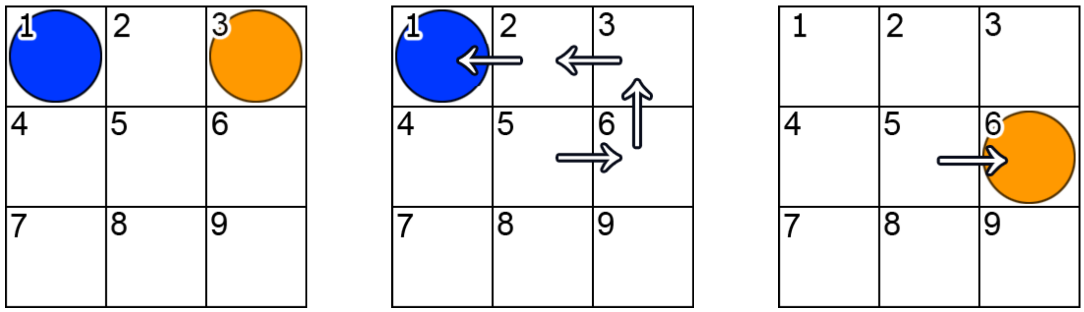
The ToM_0 Sender believes that the behaviour of the Receiver may depend on his own behaviour. He believes that it is his task to find out what actions he should perform to make sure that the Receiver matches her token to the goal configuration. In essence, the ToM_0 Sender views the Tacit Communication Game as the task of learning to operate a static machine. While playing the game, the ToM_0 Sender keeps a list of messages, which specify his own actions and the corresponding reactions of the Receiver. In each trial, the ToM_0 Sender selects the message which he believes most strongly will result in the Receiver matching her token to the goal configuration.

Once the ToM_0 Sender has sent the message, he observes the reaction of the Receiver, and updates his beliefs accordingly. Note that the ToM_0 Sender makes the implicit assumption that the Receiver does not change her behaviour after observing the outcome of a game. Like learning how to operate a static machine, the ToM_0 Sender assumes that he should find the right message that will cause the Responder to perform the correct action.

The way a ToM_0 Sender views the Tacit Communication Game is similar to the way someone would learn how to operate a coffee machine with unlabeled buttons. The operator randomly tries buttons and determines what the outcome of pressing that button is. If, for example, the operator presses some button b and the machine dispenses tea, the operator believes that if he were to press the same button b again, the machine would dispense tea again. In particular, this belief is independent of whether or not the operator intended to get tea from the machine.

Similarly to the ToM_0 Sender, the ToM_0 Receiver is unable to consider the goals of her partner. Instead, the ToM_0 Receiver forms beliefs about how the actions of the Sender relate to the goal configuration. During every trial, the Receiver attempts to find the actions of the Sender in her memory and matches her token to the goal configuration she believes to correspond to the message. After observing the outcome of the game, the ToM_0 Receiver updates her memory accordingly.

Example 1. When a ToM_0 Sender and a ToM_0 Receiver meet each other in the Tacit Communication Game for the first time, neither agent has a clear idea of how the other will behave. When the Sender observes the first goal configuration, such as the one depicted in Figure 2a, he will select a new message that will result in him matching his token to the goal configuration, and record how the Responder reacts. In the case of Figure 2b, we assume that the Sender has chosen to send the message 'rull'. That is, the Sender first moves his token right, up, left, and left. If the Responder reacts by correctly guessing her goal location and moves her token to tile 3, both agents will associate the message 'rull' with the goal configuration of Figure 2a. In this case, whenever the goal configuration of Figure 2a appears, the Sender will send the message



(a) Goal configuration (b) Sender actions (c) Receiver actions

Figure 2: Example of moves in a Tacit Communication Game. The Sender observes the goal configuration (a), and is the first to move his (dark blue) token across the board (b). Once the Sender has moved, the Receiver can move her (light orange) token (c).

'rull' and the Receiver will correctly move to her goal location.

However, when the Receiver guesses incorrectly, the beliefs of the agents diverge. Suppose that the Receiver moves to tile 6 as shown in Figure 2c. The ToM_0 Sender now believes that if he were to send the message 'rull' again, the Receiver will move to tile 6 again. However, the ToM_0 Receiver believes that the message 'rull' is not associated to tile 6, and will therefore move to a different location. Thus, in this case the Sender and Receiver communicated at 'cross-purposes'.

4.2 First-order theory of mind

In addition to zero-order beliefs, a first-order theory of mind (ToM_1) agent also considers the Tacit Communication Game from the perspective of his or her partner. A ToM_1 Sender considers the possibility that the Receiver is a ToM_0 agent who wants to match her token to the goal configuration, but does not know the goal configuration. Whenever a ToM_1 Sender considers sending some message m, he puts himself in the position of the Receiver, and determines how he would interpret the message m if he were a ToM_0 agent.

Similarly, a ToM_1 Receiver considers the possibility that the Sender is a ToM_0 agent who wants the Receiver to match her token to the goal configuration, and who attempts to manipulate her behaviour accordingly. When interpreting a message m, the ToM_1 Receiver puts herself in the position of the Sender, and determines what message she would have sent in his place. In Setting 2, where Senders prefer to send the shortest message that passes through the goal location of the Receiver, this gives the ToM_1 Receiver information about her goal location.

First-order theory of mind can be particularly helpful when interpreting a novel message m. By placing herself in the position of the Sender, the ToM_1 Receiver believes that the Sender would only have sent a novel message if there was no suitable existing message m' to communicate the goal configuration. That is, if the agents have already successfully used message m' to communicate goal configuration l', then the Sender would not use a novel message m to communicate the same goal configuration l'. This use of theory of mind reasoning results in a thought process that is very close to bidirectional optimality theory [3] (see Section 2), except that in our case, second-order theory of mind is not needed for the agent to reason this way. This is because in the concrete game setting of the Tacit Communication Game, Senders can eventually reason about the actions of the Receiver ('the Sender believes that I will select tile 5') rather than the knowledge of Receivers ('the Sender believes that I will know my goal location is tile 5').

Although a ToM_1 agent is capable of considering his or her partner as a ToM_0 agent, repeated interactions with the same partner may lead the ToM_1 agent to realize that the predictions made by the use of first-order theory of mind are inaccurate. In this case, a ToM_1 agent may decide to behave as if he were a ToM_0 agent instead.

Example 2. Consider the Tacit Communication Game as shown in Figure 2. In this game, the Sender learns the goal configuration (Figure 2a) and sends the message 'rull' (Figure 2b). A ToM_1 Receiver interprets this message by putting herself in the position of the Sender, and determining in what situation she would have sent the same message. If the Receiver has not seen the message 'rull' before, this gives the ToM_1 Receiver information about her goal location in Setting 2. Since all Senders in Setting 2 prefer to send short messages that pass through the goal location of the Receiver, the ToM_1 Receiver believes that any new message she sees passes through her goal location. For example, the ToM_1 Receiver believes it is impossible that her goal location is tile 9, since the Sender's token did not pass through tile 9.

Note that although by sending the message 'rull', the Sender's token passes over tile 2, the ToM_1 Receiver believes that it is unlikely that her goal location is tile 2. By placing herself in the position of the Sender, she believes that if her goal location would have been tile 2, she would have preferred to send the message 'ul'. This message also passes through her goal location, but it is shorter than 'rull'.

Suppose that the Receiver responds by moving to tile 6 (Figure 2c), failing to match her token to the goal configuration. A ToM_1 Receiver believes that once she has responded to a certain message m by going to location l, the Sender will believe that she will go to location l whenever he sends message m in the future. After the events shown in Figure 2, a ToM_1 Receiver therefore believes that if the Sender were to send the message 'rull' again, he expects her to go to tile 6. On the other hand, if the Sender were to send a different message such as 'ul', she believes that the Sender wants her to move to a tile different than tile 6.

The use of theory of mind thus allows the ToM_1 Receiver to circumvent the problem of limited feedback. Even though the trial ended unsuccessfully, and the ToM_1 Receiver does not know what the original intention of the message 'rull' was, the Receiver believes that in the future, the Sender will use the message 'rull' to indicate that the goal location of the Receiver is tile 6.

For a ToM_1 Sender, the events shown in Figure 2 result in a conflict between his zero-order and first-order beliefs about the behaviour of the Receiver. According to his zero-order beliefs, whenever the Sender would choose to send the message 'rull' again, the Receiver will respond by moving to tile 6 again. However, the Sender's first-order theory of mind predicts that the Receiver could select any tile other than tile 6. Different orders of theory of mind therefore result in different hypotheses about the behaviour of the Receiver. Whether the ToM_1 Sender decides to act of his zero-order or on his first-order beliefs is based on the accuracy of these predictions. If first-order theory of mind accurately predicted the actions of the Receiver in previous interactions, the ToM_1 Sender relies on his first-order beliefs rather than his zero-order beliefs.

4.3 Higher orders theory of mind

An agent capable of using theory of mind order beyond the first not only takes the perspective of his partner into account, but also considers the possibility that his partner is taking the perspective of the agent himself into account. For each additional order of theory of mind k, a ToM_k agent maintains an additional model of his or her partner as a ToM_{k-1} agent. A ToM_k agent is therefore limited in the maximum depth of recursive reasoning about the mental content of others. In our setup, a Sender and Receiver therefore cannot achieve *common knowledge of rationality*, which is a typical assumption in game-theoretic models [19, 43].

Note that when one agent is reasoning at kth-order theory of mind, the optimal response of the other agent is to reason at $(k+1)$st order theory of mind. In general, reasoning at kth-order theory of mind is not optimal for an agent if her or her partner reasons at any order other than $(k-1)$st order theory of mind. In our agent model, a ToM_k agent is not restricted to reason only at kth-order theory of mind, but a ToM_k agent can reason at any order of theory of mind up to and including kth-order. While playing the Tacit Communication Game, agents form beliefs about the order of theory of mind at which their partner is reasoning, and select the order at which they reason themselves accordingly.

Example 3. Second-order theory of mind may help a ToM_2 Sender to select the message he wants to send. This is especially true for ToM_2 Senders in Setting 2, where Senders prefer to send short messages that pass through the goal location of the Receiver.

Suppose the goal configuration is as shown in Figure 2a, and the ToM_2 Sender considers sending the message 'rull' (Figure 2b). Following the reasoning outlines in Example 2, a ToM_2 Sender in Setting 2 believes that if he were to send message 'rull', the Receiver will believe her goal location is either tile 3 or tile 6. For any other goal location of the Receiver, the ToM_2 Sender believes that a ToM_1 Receiver believes that a ToM_0 Sender would prefer to send a different message. This results in the ToM_2 Sender believing that there is a 50% probability that the Receiver will select the correct tile 3 if he were to send message 'rull'.

If, on the other hand, the ToM_2 Sender were to send message 'urll', he believes that the Receiver would only consider tile 3 as her goal location. For every other goal location of the Receiver, the ToM_2 Sender believes that a ToM_1 Receiver would expect a different message. Second-order theory of mind can therefore help the Sender in selecting unambiguous messages.

5 Mathematical model of theory of mind

In the previous section, we described how theory of mind agents can make use of simulation-theory of mind while playing the Tacit Communication Game. In this section, we present the implementation of these computational agents.

The agents described in this section are inspired by the theory of mind agents that we used to investigate the effectiveness of theory of mind in competitive settings [13] and mixed-motive settings [12]. In contrast to previous work, in which theory of mind agents always performed a similar role within a game, the Sender and Receiver role in the Tacit Communication Game are quite different.

In our description below, the Tacit Communication Game consists of

- a set of possible goal locations L,
- a set of possible messages M, and

- a function $r : M \to L$ so that the location of the Sender after sending message $m \in M$ is $r(m)$.

The game consists of a number of individual trials that additionally specify the goal locations l_S and l_R for the Sender and Receiver respectively.

In each trial within the Tacit Communication Game, the Sender observes the goal locations l_S and l_R, and selects a message $m \in M$ to send to the Receiver. The Receiver observes the message m and selects a location $l \in L$ as a response. The trial ends successfully if $r(m) = l_S$ and $l = l_R$. In all other cases, the trial ends unsuccessfully.

5.1 Zero-order theory of mind model

While playing the Tacit Communication Game, a *ToM*$_0$ Sender does not form explicit beliefs about the mental content of the Receiver. Instead, the *ToM*$_0$ Sender models the behaviour of the Receiver through zero-order beliefs $b^{(0)}_m : L \to [0, 1]$, so that the Sender believes that the probability that the Receiver will go to goal location $l \in L$ after he has sent message $m \in M$ is $b^{(0)}_m(l)$. Initially, the Sender does not know how the Receiver will react to a given message m (i.e. $b^{(0)}_m(l) = \frac{1}{|L|}$). Over repeated trials, the *ToM*$_0$ Sender updates these beliefs to reflect the behaviour of the Receiver.

After observing the goal configuration, the *ToM*$_0$ Sender determines the set of messages $\mu^{(0)}(l_S, l_R)$ that cause the Sender's token to end up on his goal location l_S while maximizing his belief that the Receiver will select her goal location l_R. That is,

$$\beta^{(0)}(l_S, l_R) = \max_{m \in M, r(m)=l_S} b^{(0)}_m(l_R) \quad (1)$$

$$\mu^{(0)}(l_S, l_R) = \left\{ m \in M \,\middle|\, r(m) = l_S, b^{(0)}_m(l_R) \geq \beta^{(0)}(l_S, l_R) \right\}. \quad (2)$$

In addition, we assume that the Sender has complete preferences over messages, and compare how the choice of messages due to different preferences change the effectiveness of theory of mind agents. In Setting 1, Senders have preference relation that prefers shorter over longer messages, irrespective of the goal configuration. That is, the Sender will randomly select one of the messages of $\mu^{(0)}(l_S, l_R)$ that minimize the length of the message.

In Setting 2, Senders have a preference relation that is meant to be closer to the way human participants choose messages to send [11, 27, 2]. Under this preference relation, the Sender prefers to send shorter messages that have passed through the Receiver's goal location l_R.

Similar to the *ToM*$_0$ Sender, a *ToM*$_0$ Receiver forms zero-order beliefs $b^{(0)}_m : L \to [0, 1]$. When the Receiver observes a message $m \in M$, she believes that the probability that selecting location $l \in L$ will be successful is $b^{(0)}_m(l)$. In the second phase of the Tacit Communication Game, the *ToM*$_0$ Receiver therefore selects one of the locations $\lambda^{(0)}(m)$ that maximize the agent's zero-order beliefs. That is, the Receiver selects a location from the set at random:

$$\lambda^{(0)}(m) = \left\{ l \in L \,\middle|\, b^{(0)}_m(l) \geq \max_{l' \in L} b^{(0)}_m(l') \right\}. \quad (3)$$

Once the Sender has sent message m^* and the Receiver has selected location l^*, the agents update their beliefs. The Sender believes that if he were to send the message m^* again, the Receiver will respond the same way as before and select location l^* again. That is, the Sender sets $b^{(0)}_{m^*}(l^*) = 1$ and $b^{(0)}_{m^*}(l) = 0$ for all $l \neq l^*$.

In contrast, the beliefs of the Receiver reflect the outcome of selecting l^* in response to m^*. If the trial ended successfully, the Receiver believes that the correct response to m^* is l^*. That is, the Receiver sets $b^{(0)}_{m^*}(l^*) = 1$ and $b^{(0)}_{m^*}(l) = 0$ for all $l \neq l^*$. Note that in this case, the Receiver and the Sender end up having the same zero-order beliefs. However, when a trial ends unsuccessfully, the Receiver believes that the correct response to m^* is not l^*. In this case, she sets $b^{(0)}_{m^*}(l^*) = 0$ and normalizes her beliefs $b^{(0)}_{m^*}$.

5.2 Theory of mind model

In our agent model, theory of mind allows agents to view the game from the perspective of their partner, and determine what their action would have been if the roles had been reversed. Each theory of mind agent contains a model of a ToM_0 Sender as well as a ToM_0 Receiver. Beliefs of higher orders of theory of mind are determined by taking into account the information of lower orders of theory of mind. For example, a ToM_k Sender constructs kth-order beliefs $b^{(k)}_m(l)$ for the probability that the Receiver will select location l in response to message m by determining what locations he would select himself if he had been a ToM_{k-1} Receiver.

Note that ToM_k agents in our agent model can reason at any order of theory of mind up to and including kth-order theory of mind. That is, a ToM_k agent can change the order of theory of mind he reasons at, for example if he comes to believe that his partner is not reasoning at $(k-1)$st-order theory of mind. To this end, an agent specifies a confidence $c_k \in [0, 1]$ for each order of theory of mind available to the agent, which represents how much confidence the agent has in predictions based on kth-order theory of mind relative to lower orders of theory of mind.

A ToM_k Sender constructs his kth-order beliefs by looking at the game from the perspective of a ToM_{k-1} Receiver, and determining how likely it would be that he would select location l after observing messages m. Since a ToM_{k-1} Receiver responds to message m by randomly selecting a location from the set $\lambda^{(k-1)}(m)$, a ToM_k Sender's kth-order beliefs are given by

$$b^{(k)}_m(l) = \begin{cases} 0 & \text{if } l \notin \lambda^{(k-1)}(m) \\ \frac{1}{|\lambda^{(k-1)}(m)|} & \text{if } l \in \lambda^{(k-1)}(m). \end{cases} \tag{4}$$

These beliefs are then combined with lower-order beliefs with the help of the confidence c_k in kth-order theory of mind, according to

$$B^{(k)}_m(l) = c_k \cdot b^{(k)}_m(l) + (1 - c_k) \cdot B^{(k-1)}_m(l), \quad \text{where} \quad B^{(0)}_m(l) = b^{(0)}_m(l). \tag{5}$$

The ToM_k Sender then determines which messages maximize his combined belief that the Receiver will respond by choosing her actual goal location l_R, analogous to the way a ToM_0 Sender does:

$$\beta^{(k)}(l_S, l_R) = \max_{m \in M, r(m) = l_S} B^{(k)}_m(l_R) \tag{6}$$

$$\mu^{(k)}(l_S, l_R) = \left\{ m \in M \,\middle|\, r(m) = l_S, B^{(k)}_m(l_R) = \beta^{(k)}(l_S, l_R) \right\}. \tag{7}$$

The ToM_k Sender selects a message from the set $\mu^{(k)}(l_S, l_R)$ according to his preferences. In Setting 1, the ToM_k Sender will therefore select a message from $\mu^{(k)}(l_S, l_R)$ that minimizes the length of the message. In Setting 2, however, the ToM_k Sender randomly selects a message from $\mu^{(k)}(l_S, l_R)$ that minimizes the length of those messages in $\mu^{(k)}(l_S, l_R)$ that pass through the goal location l_R of the Receiver.

Similarly, a ToM_k Receiver interprets an observed message m by placing herself in the position of a ToM_{k-1} Sender, and determining for what goal configurations she would have sent the message m. The ToM_k Receiver's kth-order beliefs are constructed through

$$b_m^{(k)}(l) = \begin{cases} 0 & \text{if } m \notin \mu^{(k-1)}(l_S, l_R) \\ 0 & \text{if there is a } m' \in \mu^{(k-1)}(l_S, l_R) \text{ such that } m' > m \\ \gamma & \text{otherwise,} \end{cases} \tag{8}$$

where γ is a normalizing constant. Analogous to the ToM_k Sender, the ToM_k Receiver combines her kth-order beliefs with lower-order beliefs according to Equation (5). The ToM_k Receiver then selects a location that maximizes his combined beliefs $B_m^{(k)}$. That is, she randomly selects a location from the set

$$\lambda^{(k)}(m) = \left\{ l \in L \,\middle|\, B_m^{(k)}(l) \geq \max_{l' \in L} b_m^{(k)}(l') \right\}. \tag{9}$$

Once the Sender has sent message m and the Receiver has selected a location l, both agents update their confidences c_k in the different orders of theory of mind available to them according to the accuracy of predicted behaviour. If the Sender and the Receiver successfully matched their respective tokens to the goal configuration, a ToM_k Receiver updates her confidence c_k in kth-order theory of mind according to

$$c_k := \begin{cases} 1 & \text{if } l \in \lambda^{(k)}(m) \\ 0 & \text{otherwise.} \end{cases} \tag{10}$$

That is, if kth-order theory of mind accurately predicted that location l would be correct, the ToM_k Receiver is highly confident that application of kth-order theory of mind will be successful in the future. However, if kth-order theory of mind failed to predict that location l would be correct, the ToM_k Receiver loses confidence in application of kth-order theory of mind.

Due to the limited feedback available to the Receiver, if the Sender and the Receiver failed to match their respective tokens to the goal configuration, the Receiver obtains little information about whether or not kth-order theory of mind accurately predicts the behaviour of the Sender. Nonetheless, if kth-order theory of mind predicted that location l was the only possible correct location (i.e. $\lambda^{(k)}(m) = \{l\}$), a ToM_k Receiver loses all confidence in kth-order theory of mind.

For a ToM_k Sender, his confidence c_k in kth-order theory of mind is updated according to

$$c_k := \begin{cases} (1-f) \cdot c_k + f & \text{if } l \in \lambda^{(k-1)}(m) \\ (1-f) \cdot c_k & \text{otherwise,} \end{cases} \tag{11}$$

where

$$f = \frac{10 - \left|\lambda^{(k-1)}(m)\right|}{9}. \tag{12}$$

In this update, the confidence c_k in kth-order theory of mind is increased if it accurately predicted that the Receiver might have chosen location l, while the confidence c_k is decreased if the Sender's kth-order theory of mind considered that l was not a likely location for the Receiver to choose. The strength f of the update is determined by how strongly kth-order theory of mind predicted the behaviour of the Receiver. If any action of the Receiver is possible according to kth-order theory of mind ($|\lambda^{(k-1)}(m)| = 9$), the confidence c_k will only be changed slightly. However, if kth-order theory of mind makes a strong prediction about Receiver behaviour ($|\lambda^{(k-1)}(m)| = 1$), the confidence c_k will be adjusted by a large margin.

		Receiver		
		ToM_0	ToM_1	ToM_2
	ToM_0	771.3 (3.16)	49.4 (0.08)	290.3 (1.42)
Sender	ToM_1	220.6 (0.48)	139.1 (2.01)	107.5 (0.28)
	ToM_2	258.3 (0.51)	51.6 (0.09)	144.2 (0.42)

Table 1: Average number of unsuccessful trials before a pair of agents establishes communication in the Tacit Communication Game in Setting 1 (standard error in parentheses). In this setting, Senders prefer to send messages of shorter length, irrespective of the goal configuration. Results are shown for Senders and Receivers of various orders of theory of mind.

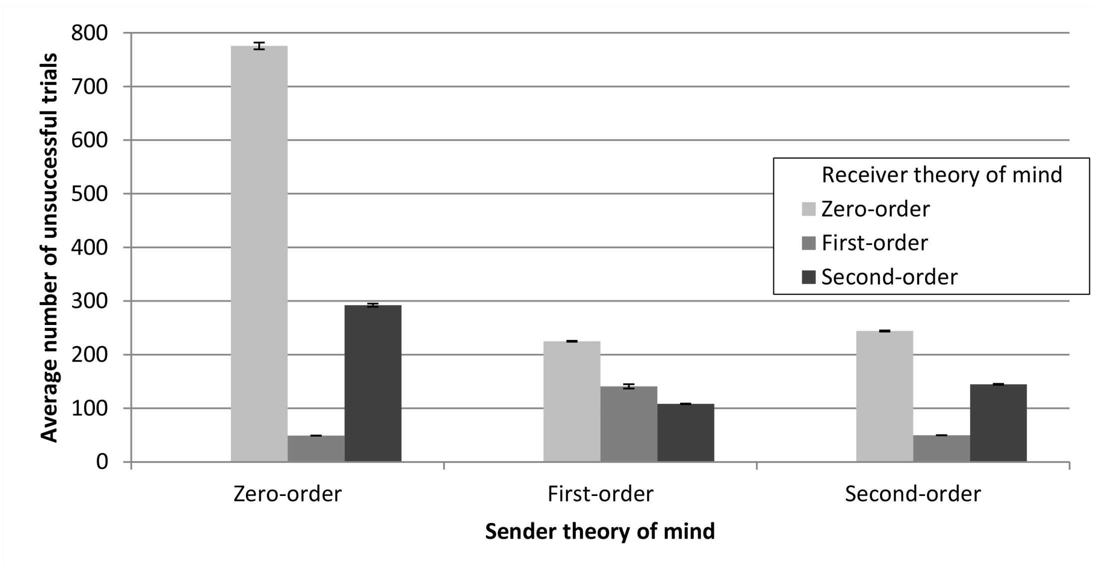


Figure 3: Average number of unsuccessful trials before a pair of agents establishes communication in the Tacit Communication Game in Setting 1. In this setting, Senders prefer to send messages of shorter length, irrespective of the goal configuration. Results are shown for Senders and Receivers of various orders of theory of mind. Brackets indicate standard errors of the results.

6 Results

We performed simulations in which the theory of mind agents described in Section 4 and Section 5 played the Tacit Communication Game. Although the mathematical model of Section 5 does not restrict the length of a message and would allow a message of infinite length, the messages used in our simulations were limited to have no more than 8 actions. Since Senders in our agent model always preferred to send shorter messages over longer messages, this restriction did not affect agent performance: throughout the simulation experiments, no Sender chose to send a message of more than 6 actions.

Each game consisted of a number of trials in which the Sender observed a randomly selected goal configuration, the Sender and Receiver moved their tokens in turn, and the agents were informed whether the trial ended successfully or unsuccessfully. The game continued with the same pair of agents until the agents had successfully matched their tokens to each possible goal configuration. At that moment, the pair of agents was said to have successfully established communication. Following the experimental setup with human participants [11, 27, 2], we excluded goal configurations in which either player had his or her goal location in the center, as well as goal configurations in which the Sender and the Receiver had the same goal location. That is, agents only encountered 56 of the possible 81 different goal configurations. However, agents did not know this restriction in the selection of goal configurations; agents believed that all 81 goal configurations were possible.

We compare two simulation settings that differ in the way the Sender selects which message he wants to send. In Setting 1, Senders prefer to send shorter messages rather than longer messages. In Setting 2, Senders prefer to send a shorter message that causes their token to pass through the goal location of the Receiver at some point. The preferences in Setting 2 were selected to be closer to the way human Senders choose messages.

We calculate the performance of a pair of agents by the number of unsuccessful trials before the pair establishes communication, averaged over 20,000 trials. Table 1 and Figure 3 show how performance varies with the order of theory of mind of the Sender and the Receiver in Setting 1, in which the Sender has a preference to send shorter messages.

Figure 3 shows that even without the use of theory of mind, agents can establish communication. However, if a ToM_0 Sender is paired with a ToM_0 Receiver, an average of 771 trials end unsuccessfully before communication is established. This means that a pair of ToM_0 agents performs worse than chance; they make an average 13 errors on a given goal configuration before they agree on how that goal configuration will be communicated. This is due to the cross-purposes communication mentioned in Example 1.

Figure 3 also shows that the ability to make use of first-order theory of mind has a strong positive effect on performance. Because the Sender and the Receiver fulfill very different roles in the Tacit Communication Game, theory of mind also benefits the Sender and the Receiver to a different degree. When a ToM_1 Sender is paired with a ToM_0 Receiver, approximately 220 trials end unsuccessfully before communication is established. This combination of a ToM_1 Sender and a ToM_0 Receiver therefore performs approximately on chance level by making an average 4 errors per goal configuration.

The combination of a ToM_0 Sender with a ToM_1 Receiver performs better, with an average 49 unsuccessful trials. This combination of a ToM_0 Sender and ToM_1 Receiver thus makes less than 1 error on average per goal configuration. In fact, the best performance as shown by Figure 3 is achieved by the a ToM_0 Sender paired with a ToM_1 Receiver. Compared to this situation, an increase in the theory of mind abilities of either agent results in a decrease in performance. In particular, this means that a ToM_1 Sender is better off not using his theory of mind if the Receiver has the ability to make use of theory of mind as well.

Unlike human participants, different pairs of agents in Setting 1 end up with very different ways of communicating a goal configuration. Since Senders in Setting 1 prefer shorter messages, the messages that agents end up using to communicate goals are typically short; most goal configurations are communicated with only 2 or 3 moves of the Sender. For example, a typical Sender may use the message 'ul' to communicate that the goal location of the Receiver is location 3, while the message 'lu' communicates that the Receiver's goal location is location 5.

In Setting 2, Senders prefer to send short messages that visit the goal location of the Receiver. Table 2 and Figure 4 show how performance in this setting varies with the theory of mind abilities of the agents. When the Receiver is a ToM_0 agent, the results in Figure 4 are the same

		Receiver		
		ToM_0	ToM_1	ToM_2
	ToM_0	776.4 (3.28)	6.8 (0.04)	108.9 (1.42)
Sender	ToM_1	224.8 (0.50)	84.6 (1.35)	12.9 (0.19)
	ToM_2	249.2 (0.47)	0.0 (0.00)	0.0 (0.00)

Table 2: Average number of unsuccessful trials before a pair of agents establishes communication in the Tacit Communication Game in Setting 2 (standard error in parentheses). In this setting, Senders prefer to send messages of shorter length in which their token passes through the goal location of the Receiver. Results are shown for Senders and Receivers of various orders of theory of mind.

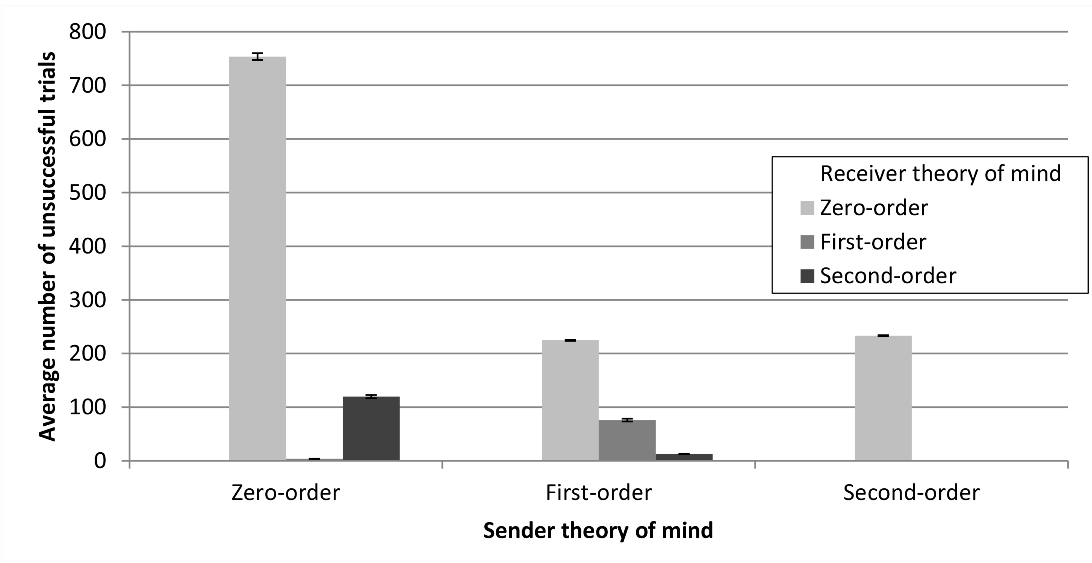


Figure 4: Average number of unsuccessful trials before a pair of agents establishes communication in the Tacit Communication Game in Setting 2. In this setting, Senders prefer to send messages of shorter length in which their token passes through the goal location of the Receiver. Results are shown for Senders and Receivers of various orders of theory of mind. Brackets indicate standard errors of the results. Note that when a ToM_2 Sender is paired with a ToM_1 or ToM_2 Receiver, no trials end unsuccessfully.

as those in Figure 3. That is, even though the messages that ToM_0 Senders send include more information about the Receiver's goal location than messages sent in Setting 1, this alone is not enough to increase performance.

However, when the Receiver can reason using about the mental content of the Sender, the results change. In Setting 2, where Senders prefer to send messages in which their token passes through the goal location of the Receiver, the combination of a ToM_0 Sender with a ToM_1 Receiver performs even better than the same combination in Setting 1. On average, only 7 trials end unsuccessfully before communication is established. By placing herself in the position of the Sender, a ToM_1 Receiver believes that the Sender will have passed through her goal location while sending the message. Although this additional information does not uniquely identify the Receiver's goal location, it does reduce the number of unsuccessful trials from 49 to 7.

Secondly, Figure 4 shows that second-order theory of mind can be helpful. When a ToM_2 Sender is paired with a ToM_1 Receiver, no trial ends unsuccessfully. That is, the ToM_2 Sender consistently selects messages that the Receiver interprets correctly. This is because the ToM_2 Sender reasons about the way a ToM_1 Receiver will interpret his message, which allows him to avoid ambiguous messages.

The messages that pairs of agents end up using to communicate goal configurations in Setting 2 resemble the messages sent by human participants in the Tacit Communication Game more closely than the messages in Setting 1. Like human participants, Senders move towards the goal location of the Receiver before going to their own goal location. However, instead of pausing at the goal location for the Receiver, Senders select messages that unambiguously select one of the tiles as the goal location for the Receiver. That is, whereas a human Sender may send the message 'ul.rr' to communicate that the Receiver's goal location is tile 1, our Sender agent would prefer to send 'ulrr' because he considers that message to be more efficient. However, although agents generally avoid the use of the pause action, they still use this action when communication is not immediately successful.

7 Discussion

Humans appear to be the only species capable of making use of higher-order theory of mind. The Vygotskian intelligence hypothesis argues that although this ability is cognitively demanding, it is evolutionarily advantageous to have higher-order theory of mind because of its beneficial effects on cooperation. In this paper, we simulated interactions between computational agents in a cooperative communication task to determine to what extent the use of higher orders of theory of mind leads to better cooperation.

Our results show that even in situations in which cooperation can in principle be maintained without the use of theory of mind, reasoning about the minds of others can be helpful in establishing cooperation much more efficiently. Within the setting of the Tacit Communication Game, first-order theory of mind helps Receivers to interpret messages from Senders, and sharply reduces the number of trials needed to set up communication. However, although first-order theory of mind also benefits the Sender when the Receiver is unable to reason about the mental content of others, the beneficial effect on performance is not as high as for Receivers.

We find that the effectiveness of higher orders theory of mind depends on the way a Sender chooses among messages he considers to be equally likely to succeed. If Senders choose randomly in such a case, we find no further benefit for the use of higher orders of theory of mind. However, if agents experience a bias so that private knowledge influences the way they choose, higher-order theory of mind can further increase the effectiveness of cooperation. Our results show that in some cases, such biases can allow agents to establish communication without error.

The high proficiency of human participants in the Tacit Communication Game suggests that humans do not rely purely on zero-order theory of mind, but also take the perspective of their partner into consideration from the very start of the game. Apparently, the instructions of the Tacit Communication Game, which explicitly mention the goal of the other player, are enough to trigger theory of mind reasoning in participants and reason about the ways information can be conveyed using only simple movements. The low variation in strategies employed by human participants in the Tacit Communication Game suggests that in conventional communicative trials like the ones we investigate in this paper, participants are also biased towards certain solutions. Although we did not investigate the origins of these biases in this paper, human Senders may use analogies [18] to select what message to send. The Sender would prefer messages if he can identify an analogy with the actions he wants the Receiver to perform.

Our results show that under the right circumstances, the use of higher-order theory of mind can further reduce the number of incorrectly interpreted messages by allowing Senders to select messages that Receivers interpret correctly the first time they see them. In some cases, this can allow higher-order theory of mind agent to communicate about goals they have never seen before and using messages they have never used before. In the extreme case, the use of higher-order theory of mind may have an additional benefit not explored in this paper; if novel messages from Senders are understood by Receivers most of the time, agents need less memory to communicate. Rather than remembering all messages they have ever encountered, agents can suffice by remembering those messages in which their prediction did not match the outcome.

In our agent model, we assumed agents start playing the Tacit Communication Game without any pre-existing knowledge about the game. In future work, we intend to determine how such knowledge influences our results. In particular, we aim to incorporate the instructions given to human participants in the Tacit Communication Game more closely.

In this paper, we constructed agents that make use of human-like strategies while playing the Tacit Communication Game. In future work, we aim to investigate human strategies more directly by letting participants play the Tacit Communication Game with computational agents. In particular, we intend to find out whether human participants can learn to communicate with agent Senders that have no bias towards human-like strategies, and in what way human Senders change their behaviour when confronted with Receivers that misinterpret their messages. This way, we aim to find out to what extent people make use of theory of mind when faced with the task of setting up a new way of communicating with a stranger like in the Tacit Communication Game.

Acknowledgments

This work was supported by the Netherlands Organisation for Scientific Research (NWO) Vici grant NWO 277-80-001, awarded to Rineke Verbrugge for the project 'Cognitive systems in interaction: Logical and computational models of higher-order social cognition'. We would like to thank Iris van Rooij, Mark Blokpoel, and Ivan Toni for a number of illuminating discussions on the Tacit Communication Game and for the inspiring cooperation.

References

[1] I. Apperly. *Mindreaders: The Cognitive Basis of "Theory of Mind"*. Psychology Press, Hove, UK, 2011.

[2] Mark Blokpoel, Marlieke van Kesteren, Arjen Stolk, Pim Haselager, Ivan Toni, and Iris Van Rooij. Recipient design in human communication: Simple heuristics or perspective taking? *Frontiers in Human Neuroscience*, 6:253, 2012.

[3] R. Blutner. Some aspects of optimality in natural language interpretation. *Journal of Semantics*, 17(3):189–216, 2000.

[4] Robert Boyd, Herbert Gintis, Samuel Bowles, and Peter J Richerson. The evolution of altruistic punishment. *Proceedings of the National Academy of Sciences*, 100(6):3531–3535, 2003.

[5] Robert Boyd and Peter J Richerson. Punishment allows the evolution of cooperation (or anything else) in sizable groups. *Ethology and Sociobiology*, 13(3):171–195, 1992.

[6] N.S. Clayton, J.M. Dally, and N.J. Emery. Social cognition by food-caching corvids. The western scrub-jay as a natural psychologist. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1480):507–522, 2007.

[7] Bernard J Crespi. The evolution of social behavior in microorganisms. *Trends in Ecology & Evolution*, 16(4):178–183, 2001.

[8] M. Davies. The mental simulation debate. *Philosophical Issues*, 5:189–218, 1994.

[9] P. De Bie, T. Scott-Phillips, S. Kirby, and B. Verheij. Using software agents to investigate the interactive origins of communication systems. In *The Evolution of Language: Proceedings of the 8th International Conference (EVOLANG8)*, pages 393–394. World Scientific, 2010.

[10] H. De Hoop and I. Krämer. Children's optimal interpretations of indefinite subjects and objects. *Language Acquisition*, 13(2):103–123, 2006.

[11] Jan Peter De Ruiter, M Noordzij, Sarah Newman-Norlund, Peter Hagoort, and Ivan Toni. On the origin of intentions. In P. Haggard, Y. Rossetti, and M. Kawato, editors, *Attention & Performance XXII: Sensorimotor Foundation of Higher Cognition*, pages 593–610. Oxford University Press, 2007.

[12] H. De Weerd, R. Verbrugge, and B. Verheij. Agent-based models for higher-order theory of mind. In *Advances in Social Simulation, Proceedings of the 9th Conference of the European Social Simulation Association*, volume 229, pages 213–224. Springer, Berlin, 2014.

[13] Harmen De Weerd, Rineke Verbrugge, and Bart Verheij. How much does it help to know what she knows you know? An agent-based simulation study. *Artificial Intelligence*, 199–200:67–92, 2013.

[14] Lee Alan Dugatkin. *Cooperation among Animals: An Evolutionary Perspective.* Oxford University Press, Oxford, UK, 1997.

[15] Marcin Dziubiński, Rineke Verbrugge, and Barbara Dunin-Kęplicz. Complexity issues in multiagent logics. *Fundamenta Informaticae*, 75(1):239–262, 2007.

[16] L. Flobbe, R. Verbrugge, P. Hendriks, and I. Krämer. Children's application of theory of mind in reasoning and language. *Journal of Logic, Language and Information*, 17(4):417–442, 2008.

[17] Peter Gärdenfors. The cognitive and communicative demands on cooperation. In Jan van Eijck and Rineke Verbrugge, editors, *Games, Actions, and Social Software*, volume 7010, pages 164–183. Springer, Berlin, 2012.

[18] Dedre Gentner. Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, 7(2):155–170, 1983.

[19] Herbert Gintis. *The Bounds of Reason: Game Theory and the Unification of the Behavioral Sciences*. Princeton University Press, Princeton (NJ), 2009.

[20] T. Hedden and J. Zhang. What do you think I think you think?: Strategic reasoning in matrix games. *Cognition*, 85(1):1–36, 2002.

[21] S. Hurley. The shared circuits model (SCM): How control, mirroring, and simulation can enable imitation, deliberation, and mindreading. *Behavioral and Brain Sciences*, 31(01):1–22, 2008.

[22] D.K. Lewis. *Convention: A Philosophical Study.* Wiley-Blackwell, Chichester, 1969.

[23] Alia Martin and Laurie R Santos. The origins of belief representation: Monkeys fail to automatically represent others beliefs. *Cognition*, 130(3):300–308, 2014.

[24] B. Meijering, H. van Rijn, N.A. Taatgen, and R. Verbrugge. I do know what you think I think: Second-order theory of mind in strategic games is not that difficult. In *Proceedings of the 33nd Annual Conference of the Cognitive Science Society*, pages 2486–2491. Cognitive Science Society, 2011.

[25] Henrike Moll and Michael Tomasello. Cooperation and human cognition: The Vygotskian intelligence hypothesis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1480):639–648, 2007.

[26] C Muro, R Escobedo, L Spector, and RP Coppinger. Wolf-pack (*Canis lupus*) hunting strategies emerge from simple rules in computational simulations. *Behavioural Processes*, 88(3):192–197,

2011.

[27] Sarah E Newman-Norlund, Matthijs L Noordzij, Roger D Newman-Norlund, Inge AC Volman, Jan Peter De Ruiter, Peter Hagoort, and Ivan Toni. Recipient design in tacit communication. *Cognition*, 111(1):46–54, 2009.

[28] S. Nichols and S.P. Stich. *Mindreading: An Integrated Account of Pretence, Self-awareness, and Understanding Other Minds*. Oxford University Press, USA, 2003.

[29] Martin A Nowak. Five rules for the evolution of cooperation. *Science*, 314(5805):1560–1563, 2006.

[30] D.C. Penn and D.J. Povinelli. On the lack of evidence that non-human animals possess anything remotely resembling a 'theory of mind'. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1480):731–744, 2007.

[31] J. Perner and H. Wimmer. "John thinks that Mary thinks that...". Attribution of second-order beliefs by 5 to 10 year old children. *Journal of Experimental Child Psychology*, 39(3):437–71, 1985.

[32] D. Premack and G. Woodruff. Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, 1(04):515–526, 1978.

[33] Thomas C Scott-Phillips, Simon Kirby, and Graham RS Ritchie. Signalling signalhood and the emergence of communication. *Cognition*, 113(2):226–233, 2009.

[34] Karl Sigmund. *The Calculus of Selfishness*. Princeton University Press, Princeton, NJ, 2010.

[35] Brian Skyrms. *Evolution of the Social Contract*. Cambridge University Press, Cambridge, 1996.

[36] Brian Skyrms. *Signals: Evolution, Learning, and Information*. Oxford University Press, 2010.

[37] Luc Steels. Evolving grounded communication for robots. *Trends in Cognitive Sciences*, 7(7):308–312, 2003.

[38] Luc Steels. Modeling the cultural evolution of language. *Physics of Life Reviews*, 8(4):339–356, 2011.

[39] M. Tomasello. *Why we Cooperate*. MIT Press, Cambridge, MA, 2009.

[40] M. Tomasello, M. Carpenter, J. Call, T. Behne, and H. Moll. Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, 28(05):675–691, 2005.

[41] Daniel J Van der Post, Harmen de Weerd, Rineke Verbrugge, and Charlotte K Hemelrijk. A novel mechanism for a survival advantage of vigilant individuals in groups. *The American Naturalist*, 182(5):682–688, 2013.

[42] E. Van der Vaart, R. Verbrugge, and C.K. Hemelrijk. Corvid re-caching without 'theory of mind': A model. *PLoS ONE*, 7(3):e32904, 2012.

[43] R. Verbrugge. Logic and social cognition: The facts matter, and so do computational models. *Journal of Philosophical Logic*, 38:649–680, 2009.

[44] L.S. Vygotsky and M. Cole. *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press, Cambridge, MA, 1978.

[45] Gerald S Wilkinson. Reciprocal food sharing in the vampire bat. *Nature*, 308(5955):181–184, 1984.