

Intercorporeality and *aida*: Developing an interaction theory of social cognition

Theory & Psychology
2017, Vol. 27(3) 337–353
© The Author(s) 2017



Reprints and permissions:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/0959354317702543
journals.sagepub.com/home/tap



Shogo Tanaka

Tokai University, Japan

Abstract

The aim of this article is to develop an interaction theory (IT) of social cognition. The central issue in the field of social cognition has been theory of mind (ToM), and there has been debate regarding its nature as either theory-theory or as simulation theory. Insights from phenomenology have brought a second-person perspective based on embodied interactions into the debate, thereby forming a third position known as IT. In this article, I examine how IT can be further elaborated by drawing on two phenomenological notions—Merleau-Ponty’s intercorporeality and Kimura’s *aida*. Both of these notions emphasize the sensory-motor, perceptual, and non-conceptual aspects of social understanding and describe a process of interpersonal coordination in which embodied interaction gains autonomy as an emergent system. From this perspective, detailed and nuanced social understanding is made possible through the embodied skill of synchronizing with others.

Keywords

aida, Bin Kimura, interaction theory, intercorporeality, Maurice Merleau-Ponty, phenomenology, social cognition

Theory of mind (ToM), which is generally defined as “the ability to imagine or make deductions about the mental states of other individuals” (American Psychological Association, 2007, p. 935), has long been a central issue in the field of social cognition. Within ToM, the nature of the human ability to understand the minds of others has been a source of debate between proponents of theory-theory (TT) and those of simulation theory (ST) (Davies & Stone, 1995; Doherty, 2009). TT claims that humans practice our understanding of the minds and behaviors of others by referring to common-sense kinds of

Corresponding author:

Shogo Tanaka, Center for Liberal Arts, Tokai University, Kitakaname 4-1-1, Hiratsuka, Kanagawa, 259-1292, Japan.

Email: shg.tanaka@gmail.com

theories of mind, or in other words, folk psychology (e.g., Astington, 1993; Gopnik, 2009). In opposition to TT, ST claims that humans come to understand the minds of others by self-simulating their situation and projecting the results (e.g., Goldman, 2006; Gordon, 1986). In other words, according to TT, humans use objective theory, which is equally applicable to the self and the other, to understand others (Gopnik, 2009). ST holds that humans use subjective simulation, projected from the self to the other. Generally speaking, the former takes an observational, third-person point of view, while the latter takes an introspective, first-person point of view (Fuchs, 2013; Tanaka & Tamachi, 2013). In any case, both theories assume that the mind of the other is private, hidden behind the publicly observable body, and directly inaccessible (Gallagher & Zahavi, 2012).

Proponents of both theories often appeal to subpersonal processes for empirical support, such as the “theory of mind mechanism” (Baron-Cohen, 1995) or the “mirror-neuron as implicit simulation” (Gallese & Goldman, 1998). In such cases, the human capacity for theoretical inference or inner simulation is attributed to brain mechanisms; this causes the debate to become more detached from our direct experience, and can only be solved by neuroscientists. However, the original question in social cognition concerns the ways in which humans understand the minds of others. Even though it is undeniable that neural processes are relevant and critically important, eliminating the *person* who is attempting to understand the other and reducing the question to the *sub-personal* sphere is undesirable.

The phenomenological approach reframes the question by going back to basic experiences in the lifeworld, where humans experience others directly through interaction before running an inference or simulation (Gallagher & Zahavi, 2012). In these experiences, humans directly perceive intentions in another’s actions or emotional states in facial expressions and other bodily movements (Gallagher, 2008a).¹ This idea has, as one of its sources, the following famous statement by Scheler (1948/1954):

For we certainly believe ourselves to be directly acquainted with another person’s joy in his laughter, with his sorrow and pain in his tears, with his shame in his blushing, ... and with the tenor of his thoughts in the sound of his words. If anyone tells me that this is not “perception,” for it cannot be so, in view of the fact that a perception is simply a “complex of physical sensations,” and that there is certainly no sensation of another person’s mind nor any stimulus from such a source, I would beg him to turn aside from such questionable theories and address himself to the phenomenological facts. (pp. 260–261)

Based on this sort of direct perceptual grasp, we practice embodied interactions in the majority of ordinary intersubjective situations (Gallagher, 2004). For example, if a friend points their finger in a certain direction, you will look in that direction. Even during such a minute, nonverbal but embodied interaction, several moments of implicit social understanding occur (you know that your friend found something, you know that your friend wants to bring it to your attention, etc.).

Therefore, insights from phenomenology have introduced a second-person perspective based on embodied interactions to the ToM debate to form a third position referred to as interaction theory (IT; e.g., Fuchs, 2013; Fuchs & De Jaegher, 2009; Gallagher, 2004, 2008b). Fuchs (2013) summarizes IT as follows:

Finally, interaction theory as the most recent approach to social cognition means running the second person route: It is through immediate perception of, and embodied interaction with others that we gain our primary experience of their feelings and intentions, without recourse to inner theories or simulations. This approach focuses on the expressive bodily behavior, inter-bodily resonance, intentions as visible in action and the shared situational context in order to explain social understanding. (p. 656)

In the following sections, I examine how IT can be further elaborated based on Merleau-Ponty's notion of intercorporeality and Kimura's notion of *aida*, both of which introduce rich phenomenological insights into social cognition.

Intercorporeality

As the connection between intercorporeality and social cognition has already been argued (Tanaka, 2015), this section is limited to its theoretical outline. To begin with, the subject of social cognition belongs to the realm of intersubjectivity within phenomenology. Merleau-Ponty (1945/2012) attempted to begin a discussion on intersubjectivity in close connection with the idea of embodiment, as he described human subjectivity as an embodied being. Merleau-Ponty originally proposed the notion of intercorporeality (*intercorporéité*) or “carnal intersubjectivity” (*intersubjectivité charnelle*) in his essay on Husserlian phenomenology (Merleau-Ponty, 1960/1964b).

Merleau-Ponty's aim to connect intersubjectivity with embodiment becomes clear if one notices that the problem regarding the minds of others lies in the background. In another text, referring to the development of infant social cognition, Merleau-Ponty states, “We must abandon the fundamental prejudice according to which the psyche is that which is accessible only to myself and cannot be seen from outside” (1951/1964a, p. 116). Unless we do so, the mind of the other appears as something directly inaccessible, and inference from analogy must be adopted to understand it. However, as we have already seen, this corresponds exactly with the problem included in the conceptual framework of ToM. Once the mind of the other is considered private and hidden, certain indirect means must be employed to access it, regardless of theory or simulation. Intercorporeality focuses on the relationship between one's own body and that of the other in order to illuminate intersubjectivity and social understanding in an alternative manner.

At its core, the notion of intercorporeality contains a perception–action loop between the self and the other (Tanaka, 2014, 2015). The self's perception of the other's action prompts the same action in the self (e.g., contagious yawning) or the same action possibility (e.g., smiling). Conversely, the self's action prompts the same action, or its possibility, in the other's body. In terms of social cognition, through this resonance between bodies, we can directly grasp the intention of another's action. For the self, perceiving the other's action involves potentially performing the same action.² Therefore, through our motor capacity, we understand the meaning of another's action (Kono, 2005). Our basic ability to understand others is perceptual, sensorimotor, and non-conceptual (Gallagher, 2004). The most primary form of social understanding involves directly grasping another's actions through one's own body, and finding one's own possibility

of action in another's body. This bodily understanding precedes the theoretical inferences or inner simulations put forward in the theories of mind.³ This basic interpretation of intercorporeality corresponds well with empirical findings on the mirror neuron system (Tanaka, 2015), which provides the neural basis for the perception–action resonance between the self and the other (Rizzolatti & Craighero, 2004; Rizzolatti & Sinigaglia, 2008).

At the observable level of behavior, intercorporeality appears in two different patterns, generally referred to as “matching and meshing” in communication research (Knapp & Hall, 2010; Tanaka, 2014). Both matching and meshing, formally called “behavior matching” and “interactional synchrony” (Bernieri & Rosenthal, 1991), appear as nonverbal behaviors in interpersonal communications, but in a different way.

First, behavior matching appears as a similarity in bodily actions, such as facial expressions, postures, gestures, and vocalizations, most of which are unintended and non-conscious (Nagaoka, 2006). An infant's strong tendency to cry in response to another infant crying (Simner, 1971), or the postural congruence often observed in pairs or groups during communication (Schefflen, 1964), are classic findings in empirical research. These situations provide an opportunity for the self to live the same bodily intentionality of the other by going through the same actions, such as *at what* the other laughs, *for what* the other distorts the face, *to whom* the other speaks in a cheerful tone. It is possible to say that this aspect of intercorporeality forms the underlying process of empathy because empathy is generally defined as understanding another person based on vicarious experiences. Hatfield, Cacioppo, and Rapson (1993) employ the idea of “emotional contagion” to suggest that shared intentionality between two lived bodies offers an opportunity for the self and the other to enter into the same emotional state.

Second, interactional synchrony appears as a smooth coordination of each other's actions in communications (Bernieri & Rosenthal, 1991). Synchrony constitutes another phenomenal aspect of intercorporeality because the perception–action loop between the self and the other does not always appear as a mirroring behavior. Rather, it appears largely as embodied interactions of action and reaction (Tanaka, 2014). As we have already seen, in perceiving another's action, humans immediately grasp the intention through our motor capacity and react in response to that intention. In daily interactions with others, humans more often show a meaningful reaction than take a similar action. For example, if a speaker lowers their voice and starts to whisper, the listener will naturally lean closer to the speaker in order to identify what is being said. If an interaction partner hands a note to another person, that person will hold out their hand to receive it without thinking. The reaction to the previous action then prompts a subsequent reaction, and thus, the process continues. In other words, we mesh the flow of actions with one another in communication, as if we were dancing or playing music together. This type of well-timed, meaningful interpersonal coordination, the basis of which is the rhythmical circulation of action and reaction between the self and the other, is referred to as interactional synchrony.

For both cases, the main point is that the intentions and emotions involved in another's actions are expressive enough to be perceived without representation and immediately connected to the self's actions. Merleau-Ponty (1960/1964b) also states,

The others are also there ... not as spirits nor even as “mental activities” from the start, but as such that, for example, as we confront in anger or in love, the faces, gestures, or words to which those of our own respond without interposed deliberation, ... each one of us is pregnant with the others and confirmed by them in his body. (p. 181, Author’s Trans.)

In the most fundamental sense, the role of social understanding is not to represent the other’s “mind” or “mental states” behind the observable body; both the perception–action loop and the action–reaction loop between the self and the other are already intersubjectively meaningful. The notion of intercorporeality holds that intersubjectivity does not start as the relation between my mind and that of another, but rather as my body and that of another.

Enactive intersubjectivity

From the enactive point of view, perception is not a process of passively receiving stimuli from the environment and converting them into inner representations; for cognitive agents, it is a process of exploring possible actions toward the environment based on embodied skills and generating meaningful interactions with the environment. In other words, perception itself is a potential action (Miyahara, 2014; Noë, 2004; Varela, Thompson, & Rosch, 1991).

In the context of interpersonal communication, perceiving another person’s action involves a potential reaction for the self. Just as a chair provides the affordance to sit or a stick provides the affordance to grab (Gibson, 1979), the action of the other provides social affordances (Gallagher, 2012; Kono, 2011). According to Condon and Sander (1974), even two-week-old infants can synchronize movements of their hands, head, and legs to an adult’s speech patterns. From the very beginning of life, the other’s action is perceived as something that affords a related reaction in the self. Thus, in our daily communication practices, the other’s action affords the self to react in a certain manner; then, in turn, the self’s action is also perceived as one that affords the other an opportunity to react, and so on. For instance, when I see another person reach their hand toward a cup situated a little too far away on the table, I pass the cup to them; in response to my reaction, they receive it. Humans know how to interact with each other in a spontaneous manner, before interactions become mediated through sociocultural norms.⁴

In line with and extending this view, Fuchs and De Jaegher (2009) proposed the idea of “enactive intersubjectivity,” a variation of IT in which social understanding is thought to be created through coordinated interactions between two embodied agents. From a phenomenological perspective, Fuchs and De Jaegher claim that this can be described as “mutual incorporation.” Merleau-Ponty’s (1945/2012) work shows that the lived body has a pervasive tendency to incorporate instruments into the body schema. For example, when one is accustomed to driving a car, they can have a feeling of extension from fender to fender or sense the road surface at the end of the tires, as if the car were a natural part of the body. Upon learning how to interact with the environment using a new instrument, the body incorporates this into the body schema, bringing about a change in sensitivity and perception (also see Tanaka, 2011).⁵ This phenomenon can be called “unidirectional incorporation,” in which the lived body unilaterally incorporates the instrument.

In contrast to unidirectional incorporation, mutual incorporation is “the reciprocal interaction of two agents in which each lived body reaches out to embody the other,” characterized as “coordination with” (Fuchs & De Jaegher, 2009, p. 474).⁶ In relation to social cognition, Fuchs and De Jaegher describe eye contact between the self and the other as an example of this. When another person stares at me, I may feel the gaze as a pull, suction, or an arrow. In response, I look back into the other person’s eyes in a manner corresponding with my feelings; this reaction to the other person’s gaze immediately affords them to take subsequent action, and so on. In the momentary meeting of gazes, humans do not represent or simulate the other’s mental state, but rather directly perceive it as emotions such as anger, curiosity, or surprise. On one hand, the effect of the gaze brings forth a bodily feeling as being threatened, attracted, or ashamed, and directly prompts a reaction in response to that effect on the other. In this case, it is possible to say that the other person’s body becomes incorporated into my body schema in the sense that my perception and action are closely coordinated with the presence of it, and vice versa; this is the process of mutual incorporation.

Of course, the body of the self and that of the other are not mutually incorporated from the start. Through the oscillation between matches and mismatches, in-phase and phase-delayed states, two bodies coordinate with each other and start to perform in a synchronized manner. The operative intentionality (Merleau-Ponty, 1945/2012) of each body is partially decentered and starts to belong to a place between the self and the other.⁷ In other words, the interaction between two lived bodies gains a sort of *autonomy* as if it had a life of its own. Fuchs and De Jaegher (2009) describe this process as follows:

When two individuals interact in this way, the coordination of their body movements, utterances, gestures, gazes, etc. can gain such momentum that it overrides the individual intentions, and common sense-making emerges. ... The “in-between” becomes the source of the operative intentionality of both partners. Each of them behaves and experiences differently from how they would do outside of the process, and meaning is co-created in a way not necessarily attributable to either of them. (p. 476)

Here, a dyadic system emerges through the intercorporeal loop of action and reaction between the self and the other. The interaction itself likely overrides each party’s personal intention due to the impossibility of controlling the entire course of interaction for the self or the other. From the moment of emergence, the self and the other enter into a relatively unpredictable phase guided by the interaction itself.

As an example, consider a rally in a table tennis match. Each stroke rapidly and often fundamentally changes the game. After, or even before, perceiving their opponent’s striking action, a player starts to move to strike back. At the moment of each stroke, the player experiences operative intentionality at the individual level in accordance with the intention of beating the opponent. However, given the shared goal of playing the match together, the operative intentionality of both partners is “in-between.” The dyadic system of the rally oscillates between two states: one that is similar to a unified circle, with only one center, and another that is similar to an oval, with two distributed centers. Through oscillation, the rally creates performative meanings, such as monotonous, dynamic, hectic, and seesaw, all of which are intersubjectively shared between the players.

The notion of *aida*

What Fuchs and De Jaegher refer to as “in-between” in the passage quoted above is akin to the notion of “*aida*” proposed by Bin Kimura (1988/2000). As a phenomenological psychiatrist, Kimura (1981) first introduced this notion to describe schizophrenia as a particular disturbance in the interpersonal sphere, and then developed it to describe the structure of intersubjectivity.

However, before exploring intersubjectivity with the notion of *aida*, it is necessary to briefly summarize Kimura’s peculiar view on subjectivity. Heavily influenced by the ideas of the physiologist Weizsäcker (1940), Kimura (1988/2000) tried to relocate human subjectivity based on that of living organisms (especially those of animals). Regardless of whether it is conscious, a living organism maintains its subjectivity in relation to its environment, especially through spontaneous movements. By definition, that which is alive as an animal has the capacity to move spontaneously. When it stops moving, the organism appears to be dead. Originally, movement is something that occurs in response to the surrounding environment and informs perceptions of the environment. Movement and perception are entangled, co-constituting what Weizsäcker (1940) called the “gestalt cycle” (*Gestaltkreis*), in which what an organism perceives is permeated by how it moves, and how it moves anticipates what it perceives. For example, when I see a ball rolling toward me, I may reach for it. My possible movements, such as reaching, grabbing, or lifting, permeate my perception of the shape, speed, and direction of the moving ball. My movement of reaching for the ball anticipates my perception of it in hand-shaping, velocity, and angle.

Kimura (1988/2000) emphasizes that human subjectivity also has its origin in the gestalt cycle; thus, the most fundamental aspect of subjectivity is not found in self-consciousness or transcendental ego, but in spontaneous movements directed toward the environment. In the context of our discussion, Kimura’s view on subjectivity is quite *enactive*. As such, human subjectivity that appears as the conscious self maintains its coherence through appropriate actions toward the environment. When the environment is relatively stable, the self keeps its identity through habitual actions. Conversely, when the environment changes, the self must take suitable actions in response. For example, if one moves to a new place with a different climate, one must change their clothing to successfully adapt. By taking necessary actions corresponding with environmental changes, the self may go through qualitative or discontinuous changes; however, this does not mean that the self loses its identity. Through enactive subjectivity, the self can keep its identity only by changing its way of acting in response to environmental change. A paradox is originally contained in the structure of the self-identity.⁸

In its most basic usage, “*aida*” in Japanese refers to the spatial distance between two things or the temporal distance between two events. As it merely refers to distance, *aida* itself is nothing. This idea is meaningful from the relational point of view. Based on the implications of the word, Kimura does not focus on distance, but rather on “betweenness,” which he takes to comprise the interconnection between the living organism and the environment. He writes, “What we call ‘the self’ ... is nothing but the principle of connection that is working ‘between (*aida*)’ us and the world” (1988/2000, p. 85, Author’s Trans.). As we saw above, human subjectivity that appears as the self is not an

autonomous entity; rather, it maintains itself in relation to the surrounding environment through adaptive actions. In this regard, Kimura's *aida* primarily refers to the interface between body and environment, which makes the subjectivity of living organisms, including human beings, possible. I refer to this aspect as "subjective *aida*," in the sense that it refers to the place where the embodied subjectivity comes into being.

Kimura (1988/2000, pp. 38ff) extends this view on *aida* and subjectivity to the realm of intersubjectivity. For human beings, the environment includes not only the natural, but also the social. Thus, environmental changes include encounters with new people or groups, and the self is only able to maintain its identity by coping with interpersonal changes through appropriate social actions. The experience of engaging with new people may change the surface character of the self, but it is a necessary process to maintain a fundamental identity as the same self. Here again, the self maintains its identity by changing its way of acting toward the environment.

In the context of interpersonal and social engagement, *aida* means "between person and person." As is the case for the subjective sphere, *aida* in the interpersonal sphere is also the interface that makes subjectivity possible. However, there is a difference: this interface between person and person makes the subjectivity of both the self and the other possible. When a certain subjectivity appears as a self that is different from others, another subjectivity must delineate the self as the self (see also Kimura, 1972/2002). *Aida* involves such betweenness, through which both the self and the other come into being. This second aspect of *aida* is fittingly termed "intersubjective *aida*," which is an opening of the self and the other. Through this source, I become aware of myself and you become aware of yourself.

Considering the intersubjective *aida*, Kimura (1988/2000, pp. 32ff) describes the experience of a musical ensemble. Ideally, an ensemble would not be guided by the musical score or led by one expert in the group; rather, each player would perform their own part equally and spontaneously, and the sum of the performance would form harmonious music as a whole. On one hand, each player creates a sound based on feedback regarding the part of the music already played (retention) and feedforward regarding the part yet to be played (protection). On the other hand, all the individual performances interlace with the tempo markings, accents, and melodies well enough to form a unified musical performance. Therefore, the music is not only heard in each player's subjective *aida*, but also echoes in the intersubjective *aida*. Kimura (1988/2000) describes this as follows:

The music echoing in the virtual space of "*aida*," which is at the same time the interior and the exterior of each participant, now has organic life of its own, accompanied with its autonomy beyond each player's individual will. (p. 42, Author's Trans.)

As mentioned earlier, what Kimura calls "*aida*" in this passage obviously corresponds with Fuchs and De Jaegher's "in-between." Both cases claim that through interpersonal coordination, that is, through the matching and meshing of each participant's actions, the interaction itself gains autonomy as an emergent system and begins to create intersubjectively sharable meaning, such as music. It is notable that both theories find the origin of intersubjective meaning in the coordination and interlacing of embodied actions between the self and the other.

How autonomy unfolds

We should now focus on and describe how the autonomous process unfolds in embodied interactions. Once the self and the other are mutually incorporated and the intersubjective *aida* gains autonomy, the interaction enters a new phase. Kimura (1988/2000, pp. 43ff) introduces the concept of “*ma*”—the auto-productive function of *aida* that regulates the course of interactions—in explaining this point.

In the case of a musical ensemble, each player’s perception of the music triggers this function. For instance, each player performs an individual part based on auditory feedback and feedforward, and such auditory perception not only includes the sounds of one’s own instrument, but also the sum total of the sounds constituting the music. Otherwise, it would be impossible for the ensemble to play together. Furthermore, the parts of the music already played naturally involve mismatches at the outset, but serve as auditory feedback and the basis of feedforward, which guides each player to coordinate their performance in a more efficient manner. When the performances are coordinated well enough to maintain autonomy, each player’s performance then matches what is expected in the ensemble. In this regard, the sound of each instrument conforms to the music as if it were an indispensable part of one living body. Here we recall Merleau-Ponty’s phrase, “He and I are like organs of one single intercorporeality” (1960/1964b, p. 168).

Thus, it is possible to say that “*ma*” is a principle that guides each participant’s action within the framework of autonomous interactions through each participant’s perception of the entire interaction. According to Kimura (1988/2000),

The whole music gains the auto-productive autonomy that is independent of each player’s will, and “anticipates” in itself the sound to be forthcoming. Each player seems to follow this “anticipation” in a manner of fulfilling it. (p. 52, Author’s Trans.)

We can also state that *ma* is the “inter-subjectivity” (i.e., the subjectivity of the interpersonal sphere) that operates through each participant’s operative intentionality and regulates the course of interactions. In a highly coordinated case, each interactant experiences actions not only spontaneously, but also in accordance with the course of the forthcoming interaction. Yet, typically, each interactant cyclically experiences congruence and incongruence between one’s own action and that which is anticipated among the group.⁹

It is not by chance that Kimura uses the example of a musical ensemble to discuss the intersubjective *aida* and its operation of *ma* (in fact, he spends almost 20 pages describing and considering this experience). As is well known, music is always experienced with a certain emotional tone. For example, the dark, creeping music used in horror movies often provokes chilling fear or paralysis, and upbeat music filling a dance hall often inspires vibrancy or joyful movement. As Krueger (2014) notes, music has its own affordances and offers the perceiver possibilities of action that is accompanied by emotions. In contrast to the affordances of instruments, music affords possibilities of entrainment or *synchronized interaction* with the environment through auditory perception. Therefore, especially in an ensemble, the perceived music offers each player the possibility of entraining and synchronizing with each other.

In general, what occurs in embodied interactions of interpersonal communication is quite similar to that which occurs in a musical ensemble.¹⁰ As shown in the case of interactional synchrony (Bernieri & Rosenthal, 1991), listeners slightly coordinate their movements to the speaker's utterance rhythm, the speaker and listener repeat turn-takings in a certain tempo, this meshing of bodily gestures creates a beat of interactions, and each interactant's speech accent and intonation provides a melodic element. In addition to the exchange of meaningful messages via language, these embodied interactions as nonverbal behaviors among participants often generate emotional tones in the interpersonal field, such as convivial, collaborative, cohesive, confrontational, and competitive (Tanaka, 2014, 2015). Therefore, it is possible to say that the entire course of communicative interactions has musical features in itself and involves musical affordances.

The musical features and emotional tone to fulfill the intersubjective *aida* offers possibilities for participants to synchronize their embodied interactions. When coordinated well enough to gain autonomy, this process involves the operation of *ma*. Like players in the same musical ensemble, all participants can anticipate to some degree the subsequent phase of the interaction process. For example, in a football game, members of the same team can simultaneously share the team's intention of defense, offense, and conversion. At a minimum, someone passes the ball accurately toward a particular place, where another player simultaneously makes a dash for it. Everything happens as if the entire course of interaction has its own subjectivity, and each participant experiences their own intention of action in accordance with the operation of *ma*.

Once autonomy is established, embodied interactions among participants tend to be practiced implicitly. However, as Crossley (1996) suggests, linguistic conversation adds a certain element of *reflexivity* to the course of interactions because the speaker not only asks questions and anticipates the answers of the interlocutor, but can also answer their own questions independently. In essence, such interventions through language help clarify the intention of each participant's action and mesh the flow of interactions more closely or in a more reflective manner. Furthermore, through verbal communication, participants are open to change the emotional tone prevailing in the interpersonal field (e.g., easing tension by making a joke). In any case, language use brings the possibility of intervention from the meta-perspective, resulting in the sense of mutual understanding becoming explicit rather than implicit.

In the case of dyadic interaction, all of the processes described thus far occur like a kind of mind-reading because two interactants can foresee, to a certain extent, each other's subsequent actions, including utterances. However, it should be noted that they do so not by employing theoretical inference or inner simulation, but rather by participating in the intersubjective *aida* and synchronizing within its range. As far as the embodied interaction moving forward with autonomy, *ma* operates in an auto-productive manner, and the interactants can naturally anticipate what is expected to occur in the subsequent moment, even though the whole course of interaction stays relatively unpredictable. On one hand, each interactant maintains their freedom as an individual in terms of whether to fulfill this anticipation, while on the other, the subsequent action is readable for the interaction partner as far as it can be anticipated. When synchronized, we know what the other person is intending, feeling, or thinking, and we know how to react in response.

This manner of embodied knowing involves a connection between embodied interaction, theoretical inference, and inner simulation. When the interaction process unfolds according to external rules (e.g., a music score), we can naturally infer the other person's subsequent action and intention on that basis. When the situation involves the different roles of the self and the other (e.g., speaker–listener, perceiver–perceived), we tend to simulate the other person's perceptions and feelings, which may be different from our own. However, mind-reading of this kind requires social expertise and is primarily based on the embodied skill to synchronize interactions with others (Michael, Christensen, & Overgaard, 2014). Without this embodied skill at its base, no higher social cognition, such as interpretations of the other's thoughts, would be present.

Minimum normativity of social perception

As has already been seen, when the intersubjective *aida* maintains its autonomy, it operates in an auto-productive manner. Through ongoing embodied interactions, the process itself implies the subsequent phase, and the participants implicitly know what is expected in the situation. In other words, the participants implicitly know what is *not* expected in the same situation. For example, consider the still face experiment (Tronick, Als, Adamson, Wise, & Brazelton, 1978). When an infant is forced to interact with a nonresponsive, expressionless mother, they rapidly grow wary and attempt to interact in a usual reciprocal pattern. From our viewpoint, the infant demands that the mother perform as expected in the situation because the embodied interaction between them has already established its autonomy, and the infant has been informed of what is expected to happen through the interaction process itself.

Therefore, it is possible to say that the emergent system involves its own *minimum normativity* through interactions, distinguishing each participant's actions into two features, such as expected or unexpected, appropriate or inappropriate, desirable or undesirable, natural or unnatural, suitable or unsuitable, and acceptable or unacceptable. As long as the emergent system maintains its auto-productive function, participants can naturally perceive each other's actions with a certain quality without being mediated by judgment. In terms of social cognition, the direct social perception (Gallagher, 2008a) not only includes the other's intention of action or emotion, but also certain qualities of action that derive from the normativity of the autonomous interaction. For example, I can perceive not only the joy of a friend in their smile, but also the smile as an exaggerated action, as in pretending to laugh (this is perceived as an unnatural action). I perceive not only another person's intention of kicking a ball in their movement, but also that their kick is strong enough to pass the ball to my position (this is perceived as an appropriate action). I do not project judgment onto the perception; instead, the perception itself involves a certain quality according to the shared normativity among interactants.

Of course, every participant brings different social skills, personal sensitivities, and sociocultural backgrounds to an interaction, and all of these factors can affect perceptions toward the other. However, after bracketing all these conditions, we can still conceive of the normativity of the emergent system from the viewpoint of its functionality. Some actions facilitate embodied interactions better than others, some actions suit interpersonal moods, and some actions are better for achieving shared goals. For example,

when playing peekaboo, after hiding and showing a smiling face several times, the autonomy of play emerges through dyadic interactions. In such a situation, a slight delay will facilitate subsequent interactions for stimulating a baby's curiosity better than regularly repeated temporality, and exaggerated facial expressions will suit the playful mood much better than an average smile. Some actions function better than others within the range of established intersubjectivity among the participants.

Concerning the normativity of perception, it is helpful to remember that we perceive objects with certain qualities based on implicit standards, such as heavy or light, hot or cold, and bright or dark, not only in social perception, but also in object perception. From Merleau-Ponty's perspective (1945/2012), these qualities are not derived from judgment, but rather, they are an inherent part of perception itself. The perceiver's body inhabits the world and anchors itself in a certain environment, and the perceptual field is differentiated and organized through interactions between the perceiver and that environment. Due to these interactions, the perception itself becomes attuned to an implicit standard. According to this standard, we can perceive a particular object as "heavy" rather than "light," or as "warm" rather than "cold." A sort of normativity also exists in object perception, which is instituted between the body and the environment through interactions (see also Kono, 2000). We can directly perceive the characteristic qualities of the object in a differentiated way based on this implicit normativity.

Returning to social perception, a further step in knowing another person beyond direct perception of their intentions and emotions is becoming aware of the qualities involved in the perceptions of their bodily actions. Once autonomy is established through embodied interactions, one can perceive another's actions with certain qualities according to implicit normativity. The other's facial expressions may be natural, forced, relaxed, innocent, or artificial; bodily movements and gestures may be smooth, exaggerated, awkward, or graceful; and voices may be high-pitched, gloomy, loud, or lively. All these qualities are not merely subjective judgments; they are derived from the normativity of autonomous interactions and perceived as subtle deviations from it. Some of these qualities may be expected, appropriate, desirable, natural, suitable, or acceptable to facilitate interaction process, whereas others may not.

It is important to remember that the emergent system among participants has its origin in the matching and synchrony of their actions. As a contrasting example, we can think of a person standing in front of a mirror. In this case, every single action that person makes is perfectly matched and synchronized with the body reflected in the mirror. However, it is impossible to recognize another person in the mirror because the body lacks *otherness*. Originally, a body that appears as another person is one that acts differently from the self. Intersubjectivity is constituted between the body of the self and that of the other when they interact with each other. On one hand, there must be a difference between two bodies, but on the other, they must be able to match and mesh their actions. The implicit normativity of social perception is created through behavior matching and interactional synchrony between two (or more) differently moving bodies.

Therefore, in principle, perceiving the qualities of another person's actions as subtle deviations from normativity is being aware of mismatches and desynchronizations during the shared interaction process. Once the autonomous process of embodied interactions is established between the self and the other, this becomes an important key to knowing the mental state of the other. We can describe various examples commonly

experienced in the lifeworld. When conversation is maintained in a certain rhythm, a delayed utterance can mean that the other is being careful when saying something important or trying to attract your attention. Similarly, when the utterance is repeated in a certain volume of voice, a louder voice can mean that the other is trying to stress the statement. If ordinary interactions are established at a close distance, moving further away during a conversation can mean that the other is trying to keep distance from you or hesitates to participate in the conversation, and so on.

As we have already seen, every participant brings different skills and sociocultural backgrounds into the interaction; therefore, all of these examples are open to diverse and complex interpretations. These interpretations can be a connecting point between interpersonal interactions, folk psychological theories, and inner simulative processes. However, this does not necessarily mean that these interpretations require theoretical inferences or simulations. As long as the interaction maintains its autonomy, there is shared social context among interactants, according to which the meanings of perceived qualities of another's actions are determined. Based on this assumption, mismatched or desynchronized actions, which are perceived with certain qualities, indicate something that can be attributed to the other person's mental state. These actions are the primary index for showing what they are intending, feeling, or thinking internally. In this regard, the other mind is the other body whose actions are matched but mismatched and synchronized but desynchronized with my actions.

Conclusion

I have attempted to develop an IT of social cognition. First, according to a phenomenological perspective, the question regarding ToM was reframed not by focusing on the other mind, but on the other *person*. Appearing as persons, the self and the other are socially engaged through embodied interactions in which the other's intentions and emotions are directly perceived (Gallagher, 2004, 2008a, 2008b; Gallagher & Zahavi, 2012). Through intercorporeality, perceiving the other's actions prompts the possibility of the same action and a closely related reaction in the self, and vice versa. In other words, the self and the other *enact* one another through intercorporeality, and such reciprocal interaction gains its autonomy as an emergent system based on the coordination of nonverbal behaviors and utterances (Fuchs & De Jaegher, 2009; Tanaka, 2014).

Kimura's argument regarding *aida*, and especially the notion of *ma*, provides insight into how the embodied interaction process unfolds after gaining its autonomy. *Ma* is an auto-productive function that regulates the course of interaction within the framework of shared intersubjectivity among the interactants, a process comparable to a musical ensemble. The interactants implicitly know what they are expected to do in the situation because the interaction process itself indicates a subsequent phase (Kimura, 1988/2000). As far as the interaction moving forward with autonomy, interactants practice a type of mind-reading through which they can read not only the actions of the other, but also what they are intending, feeling, or thinking. However, this mind-reading is not necessarily based on theory or simulation, but rather on the embodied skill to synchronize with others in interaction.

In my view, it is possible to find, in the process of autonomous interaction, a source of normativity of social perception that distinguishes each participant's actions into two

features, such as expected/unexpected. According to this normativity, interactants can directly perceive each other's actions with a certain quality, such as "natural" rather than "unnatural," or "appropriate" rather than "inappropriate." Reflecting that normativity itself has its origin in the matching and synchrony of embodied interactions, the perceived qualities of the other's actions are partly derived from mismatches and desynchronizations, which may indicate what the other person is intending, feeling, or thinking behind the scenes.

As is claimed by IT, in principle, the mind of the other is not hidden (Gallagher, 2004, 2008a; Kono, 2005) but rather expressed through actions and present within the body. However, this does not mean that situations exist in which the other's mental state is hidden behind the body; rather, these typically appear as mismatched or desynchronized actions during coordinated interactions. They can be perceived with certain qualities and their meanings can also be determined within the context of the interaction itself. Again, the mental state of the other is not hidden or private; it manifests itself through the malfunction of embodied interactions. As such, from the viewpoint of embodied interaction, it is understandable.

Acknowledgements

An earlier version of this article was presented at the 39th Annual Conference of the International Merleau-Ponty Circle. I would like to thank the participants of my paper session for their invaluable feedback. I would also like to thank Thomas Fuchs for his comments on the first draft of this paper.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Japan Society for the Promotion of Science [KAKENHI grant nos. 15H03066, 15KK0057 and 15K12634].

Notes

1. There is empirical evidence for this sort of direct perception. For example, an experimental study by Sartori, Becchio, and Castiello (2011) shows that people can distinguish different intentions involved in actions based on bodily movement information.
2. This point is sometimes confused with simulation theory. In particular, a version of simulation theory that equates the function of the mirror neuron system as implicit simulation (e.g., Gallese & Goldman, 1998) might also consider intercorporeality as simulation. However, the concept of "simulation" in simulation theory originally means to pretend as if the self were in the other's situation. In contrast, there is no "as if" relation between the self and the other in the case of intercorporeal resonance. The same action or its possibility is literally shared beyond pretense.
3. This does not mean that interaction theory based on intercorporeality denies the role of theoretical inferences or simulations. Interaction theory puts the primacy of the self-other interaction practiced on a second-person perspective. As derivatives of such interactions, both theoretical inferences and simulations could occur in actual interpersonal situations.

4. This might be a matter for debate. In fact, both social perceptions and interpersonal actions are influenced by and formed through sociocultural norms. Therefore, people are often confused about how to interact with others who are from different social or cultural backgrounds. Yet, even in such cases, people still know whether or not to come closer, talk, smile, make eye contact, and so on. Beyond sociocultural differences, there remains a space for embodied interactions. This article establishes this sort of space as a starting point of embodied interactions and attempts to explain how the minimum of normativity is created through them. The primary interest of this article is to provide a genetic account of the emergence of sociality between the self and the other. Regarding the sociocultural dimension of embodied interactions, an independent discussion is needed apart from this article.
5. In current neuroscience, empirical studies have supported this view. For example, Maravita and Iriki (2004) show how extended motor capability through tool use changes the neural representation of the body.
6. Fuchs and De Jaegher distinguish between “coordination with” and “coordination to” as well as between “mutual incorporation” and “unidirectional incorporation.” Unidirectional incorporation, which is typically seen in the skillful handling of instruments, is characterized as “coordination to” because the embodied agent does not coordinate with another moving agent, but to the static object.
7. Merleau-Ponty (1945/2012) distinguishes between two types of intentionality. The first is “act intentionality,” which refers to conscious judgment of something in the world. The second is “operative intentionality,” which refers to the implicit and pre-reflective unity of body and world. Operative intentionality works through embodied actions toward objects, people, and the environment.
8. Quoting Nishida, a Japanese philosopher who often expressed his fundamental thoughts with paradoxes, Kimura (1988/2000) refers to this paradoxical structure of self-identity as “discontinuity through continuity.”
9. John Shotter (1995) described people’s conversation activity in terms of the concept of “joint action” and provided a similar view to Kimura’s idea on *ma*. According to Shotter, “in joint actions between them, people constitute background situations in which they are accountable to each other in terms related to those situations” (p. 55). This description corresponds well with what is described here as *ma* or “inter-subjectivity.” Shotter also pointed out that “one is involved in and expected to maintain the action in a way quite different from those in third-person roles” (p. 55). From Kimura’s perspective, this happens because the process of self–other interaction itself gains its autonomy and *ma* starts to guide each participant’s actions.
10. For example, assuming that social interactions are organized musically, Erickson (2009) demonstrated how classroom discussions between a teacher and students are musical.

References

- American Psychological Association. (2007). *APA dictionary of psychology*. Washington, DC: Author.
- Astington, J. W. (1993). *The child’s discovery of the mind*. Cambridge, MA: Harvard University Press.
- Baron-Cohen, S. (1995). *Mindblindness: An essay on autism and theory of mind*. Cambridge, MA: The MIT Press.
- Bernieri, F. J., & Rosenthal, R. (1991). Interpersonal coordination: Behavior matching and interactional synchrony. In R. S. Feldman & B. Rimé (Eds.), *Fundamentals of nonverbal behavior* (pp. 401–432). Cambridge, UK: Cambridge University Press.
- Condon, W. S., & Sander, L. W. (1974). Synchrony demonstrated between movements of the neonate and adult speech. *Child Development*, *45*, 456–462.

- Crossley, N. (1996). *Intersubjectivity: The fabric of social becoming*. London, UK: SAGE.
- Davies, M., & Stone, T. (1995). *Folk psychology: The theory of mind debate*. Oxford, UK: Blackwell.
- Doherty, M. J. (2009). *Theory of mind: How children understand others' thoughts and feelings*. New York, NY: Psychology Press.
- Erickson, F. (2009). Musicality in talk and listening: A key element in classroom discourse as an environment for learning. In S. Malloch & C. Trevarthen (Eds.), *Communicative musicality* (pp. 449–463). Oxford, UK: Oxford University Press.
- Fuchs, T. (2013). The phenomenology and development of social perspectives. *Phenomenology and the Cognitive Sciences*, 12, 655–683.
- Fuchs, T., & De Jaegher, H. (2009). Enactive intersubjectivity: Participatory sense-making and mutual incorporation. *Phenomenology and the Cognitive Sciences*, 8, 465–486.
- Gallagher, S. (2004). Understanding interpersonal problems in autism: Interaction theory as an alternative to theory of mind. *Philosophy, Psychiatry, & Psychology*, 11, 199–217.
- Gallagher, S. (2008a). Direct perception in the intersubjective context. *Consciousness and Cognition*, 17, 535–543.
- Gallagher, S. (2008b). Inference or interaction: Social cognition without precursors. *Philosophical Explorations*, 11, 163–174.
- Gallagher, S. (2012). *Phenomenology*. London, UK: Palgrave Macmillan.
- Gallagher, S., & Zahavi, D. (2012). *The phenomenological mind* (2nd ed.). London, UK: Routledge.
- Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences*, 2, 493–501.
- Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston, MA: Houghton Mifflin.
- Goldman, A. I. (2006). *Simulating minds: The philosophy, psychology, and neuroscience of mind-reading*. Oxford, UK: Oxford University Press.
- Gopnik, A. (2009). *The philosophical baby*. New York, NY: Farrar, Straus and Giroux.
- Gordon, R. M. (1986). Folk psychology as simulation. *Mind and Language*, 1, 158–171.
- Hatfield, E., Cacioppo, J. T., & Rapson, R. L. (1993). Emotional contagion. *Current Directions in Psychological Science*, 2, 96–99.
- Kimura, B. (1981). *Jiko, aida, jikan* [Self, aida, and time]. Tokyo, Japan: Kobundo.
- Kimura, B. (2000). *L'entre* [The between] (C. Vincent, Trans.). Grenoble, France: Jérôme Million. (Original work published 1988)
- Kimura, B. (2002). *Zwischen Mensch und Mensch* [Between person and person] (E. Weinmayr, Trans.). Darmstadt, Germany: Wissenschaftliche Buchgesellschaft. (Original work published 1972)
- Knapp, M. L., & Hall, J. A. (2010). *Nonverbal communication in human interaction* (7th ed.). Boston, MA: Wadsworth.
- Kono, T. (2000). *Merleau-Ponty no imiron* [The semantics of Maurice Merleau-Ponty]. Tokyo, Japan: Sobunsha.
- Kono, T. (2005). *Kankyo ni hirogaru kokoro: Seitaigakuteki tetsugakuno tenbo* [The mind extended into the environment: A prospect of ecological philosophy]. Tokyo, Japan: Keiso Shobo.
- Kono, T. (2011). *Ishiki ha jitsuzai shinai: Kokoro, chikaku, jiyuu* [Consciousness does not exist: Mind, perception and freedom]. Tokyo, Japan: Kodansha.
- Krueger, J. (2014). Affordances and the musically extended mind. *Frontiers in Psychology*, 4, 1003. Retrieved from <http://journal.frontiersin.org/article/10.3389/fpsyg.2013.01003/>
- Maravita, A., & Iriki, A. (2004). Tools for the body (schema). *Trends in Cognitive Sciences*, 8, 79–86.
- Merleau-Ponty, M. (1964a). The child's relations with others. In *The primacy of perception* (W. Cobb, Trans., pp. 96–155). Evanston, IL: Northwestern University Press. (Original work published 1951)

- Merleau-Ponty, M. (1964b). The philosopher and his shadow. In *Signs* (R. C. McCleary, Trans., pp. 159–181). Evanston, IL: Northwestern University Press. (Original work published 1960)
- Merleau-Ponty, M. (2012). *Phenomenology of perception* (D. A. Landes, Trans.). New York, NY: Routledge. (Original work published 1945)
- Michael, J., Christensen, W., & Overgaard, S. (2014). Mindreading as social expertise. *Synthese*, *191*, 817–840.
- Miyahara, K. (2014). Mienikusa no genshogaku: Enaktivuna chikaku no kagaku nimukete [Phenomenology of suboptimal seeing: Towards an enactive unification of the sciences of perception]. *Moralia*, *20/21*, 211–232.
- Nagaoka, C. (2006). Taijin komyunikesyon niokeru higengokoudou no nisha sougoeikyō nikan-suru kenkyū [Mutual influence of nonverbal behavior in interpersonal communication]. *Japanese Journal of Interpersonal and Social Psychology*, *6*, 101–112.
- Noë, A. (2004). *Action in perception*. Cambridge, MA: The MIT Press.
- Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, *27*, 169–192.
- Rizzolatti, G., & Sinigaglia, C. (2008). *Mirrors in the brain* (F. Anderson, Trans.). Oxford, UK: Oxford University Press.
- Sartori, L., Becchio, C., & Castiello, U. (2011). Cues to intention: The role of movement information. *Cognition*, *119*, 242–252.
- Schefflen, A. E. (1964). The significance of posture in communicative systems. *Psychiatry*, *27*, 316–331.
- Scheler, M. (1954). *The nature of sympathy* (P. Heath, Trans.). London, UK: Routledge & Kegan Paul. (Original work published 1948)
- Shotter, J. (1995). In conversation: Joint action, shared intentionality and ethics. *Theory & Psychology*, *5*, 49–73.
- Simner, M. L. (1971). Newborn's response to the cry of another infant. *Developmental Psychology*, *5*, 136–150.
- Tanaka, S. (2011). The notion of embodied knowledge. In P. Stenner, J. Cromby, J. Motzkau, J. Yen, & Y. Haosheng (Eds.), *Theoretical psychology: Global transformations and challenges* (pp. 149–157). Concord, Canada: Captus University publications.
- Tanaka, S. (2014). Creation between two minded bodies: Intercorporeality and social cognition. *Academic Quarter*, *9*, 265–276.
- Tanaka, S. (2015). Intercorporeality as a theory of social cognition. *Theory & Psychology*, *25*, 455–472.
- Tanaka, S., & Tamachi, M. (2013). A phenomenological view of the theory of mind. *Bulletin of Liberal Arts Education Center, Tokai University*, *33*, 93–100.
- Tronick, E., Als, H., Adamson, L., Wise, S., & Brazelton, T. B. (1978). Infants' response to entrapment between contradictory messages in face-to-face interaction. *Journal of the American Academy of Child and Adolescent Psychiatry*, *17*, 1–13.
- Varela, F., Thompson, E., & Rosch, E. (1991). *The embodied mind: Cognitive science and human experience*. Cambridge, MA: The MIT Press.
- Weizsäcker, V. V. (1940). *Der Gestaltkreis* [The formative cycle]. Stuttgart, Germany: Thieme.

Author biography

Shogo Tanaka is a Professor in the Center for Liberal Arts at Tokai University, Japan. His primary interests are in phenomenological psychology and cognitive science, particularly theoretical issues regarding the embodied mind. His current research focus is to extend the notion of the embodied mind into social cognition, based on Merleau-Ponty's phenomenology. His most recent publication includes "Intercorporeality as a theory of social cognition" (*Theory & Psychology*).