Enacting Intersubjectivity: A Cognitive and Social Perspective on the Study of Interactions
F. Morganti, A. Carassa, G. Riva (Eds.)
Amsterdam, IOS Press, 2008, pp. 49-63

4 Interacting Socially through Embodied Action

Jessica LINDBLOM, Tom ZIEMKE

Abstract. This chapter contrasts traditional, disembodied information-processing approaches to intersubjectivity in socio-cognitive research with more recent, embodied approaches. Based on an analysis of the shortcomings of the former, it focuses on the latter, but also clarifies different notions of embodiment and its role in cognition and social interaction. Integrating a broad range of theoretical perspectives and empirical evidence from mainly social psychology, social neuroscience, embodied linguistics and gesture studies, four fundamental functions of the body in social interaction are identified: (1) the body as a social resonance mechanism, (2) the body as a means and end in communication and social interaction, (3) embodied action and gesture as a 'helping hand' in shaping, expressing and sharing thoughts, and (4) the body as a representational device. The theoretical discussions are illustrated with an example from a case study of in-situ embodied social interaction, with a focus on the importance of crossmodal interaction in the process of scaffolding. It is concluded that the body is of crucial importance in understanding social interaction and cognition in general, and in particular the relational nature of mind and intersubjectivity.

Contents

4.1 Introduction 49
4.2 Disembodied approaches to intersubjectivity 50
4.3 On the embodied nature of social interaction 53
4.4 Illustration and discussion 59
4.5 Conclusions 62

4.6 References 62

4.1 Introduction

The ability to engage in social interaction is a crucial building block of human culture, which is the foundation for the complexity of social life and cognition. In this chapter, we aim to clarify the role and relevance of the body in social interaction, from the perspective of embodied cognitive science.

Broadly speaking, the traditional information-processing approach to intersubjectivity in socio-cognitive research assumes that agents relate to each other in much the same way as they relate to other parts of the external world, i.e. by having more or less explicit internal representations of each other [e.g. 1-3]. This is a centralized view of cognition that considers (social) cognition to take

place inside the skull, with the body only serving as some kind of input and output device, i.e. a physical interface between internal programs (cognitive processes) and external world. We contrast such information-processing approaches to intersubjectivity with embodied and enactive approaches as follows.

Information-processing approaches to intersubjectivity are based on the assumption that the role of the body in social interaction and cognition is merely as a trivial, peripheral 'appendage' to the real intellectual mind. Therefore, bodily aspects are within this theoretical framework frequently addressed in terms of *non-verbal communication*, *nonverbal behavior*, or *body language*. Accordingly, embodied actions such as body posture, gaze and gesture are still commonly considered to be nothing but the visible outcomes of mental intentions and contents which are transmitted from one mind to another.

Gallagher [4] argues that these standard information-processing models of mind as representational, functional and/or computational, despite their complexity, are oversimplified and altogether neglect the many effects of embodiment. The embodied approach, on the other hand, emphasizes the way cognition is shaped by the body and its sensorimotor interaction with the surrounding world [e.g. 4-9]. Hence, this view holds that central to intersubjectivity is first and foremost the experience of being embodied in a social, cultural and material sphere [10]. It might be worth noting that this does not necessarily imply denying mental concepts as such (e.g. beliefs or intentions) altogether, but rather questioning their central underlying role, as stressed in information-processing approaches. Instead, they may be emergent from and grounded in embodied interactions rather than an underlying requirement for cognitive processes [for more details, see 10].

It should also be noted though that there are different views within embodied cognitive science regarding in what sense, or to what extent, cognition is to be considered as embodied [9]. Clark [5], for instance, distinguishes between the positions of *simple embodiment* and *radical embodiment*. According to the former, the traditional foundation of computationalist/functionalist cognitive science can be preserved more or less intact, and embodiment is merely considered a constraint of the 'inner' organization and processing. The radical embodiment position, on the other hand, goes much further and treats the facts of embodiment as a fundamental shift in the explanation of cognition that is "profoundly altering the subject matter and theoretical framework of cognitive science" [5, p. 348]. This chapter is more in line with the latter view.

In a nutshell, the comparison made in this chapter is between these two positions, i.e., information-processing vs. embodied approaches to intersubjectivity. After some brief presentation and critical discussion of the first approach in the next section, this chapter particularly focuses on the second approach, discussing recent work in cognitive science and related disciplines which indicates that the body is of crucial importance in social interaction, cognition and intersubjectivity.

4.2 Disembodied approaches to intersubjectivity

Without reviewing the huge literature on intersubjectivity, one might say that intersubjectivity refers to the manifestation of shared meanings constructed by people in their interactions with each other. Hence, intersubjectivity results in a

basic discrimination between the self and others as well as the ability to compare and project one's own private experiences or cognitive states with those of another person. Broadly speaking, it has been suggested that intersubjectivity is the cradle of social interaction and cognition [11, 12].

It is probably safe to say that within socio-cognitive research, the information-processing view of social interaction still is the most common one, and certainly still the dominant one. A good example of this view is found in *The Encyclopedia of Cognitive Science*, which characterizes social cognition as follows:

"*Social-cognitive research, with its adherence to the information-transmission metaphor, is fundamental to the study of process; that is, social cognition is the part of social psychology that deals with the psychological mechanisms that mediate the individual's response to the social environment. As such, the nature of mental representation and the dynamics of information-processing are central topics of social-cognitive inquiry"* [3, p. 66].

In a similar vein, Singer, Wolpert and Frith [13] claim:

"… the study of social interaction involves by definition a bi-directional perspective and is concerned with the question of how two minds shape each other mutually through reciprocal interactions. To understand interactive minds we have to understand how thoughts, feelings, intentions, and beliefs can be transmitted from one mind to the other … how to communicate these thoughts" [13, p. xvii].

From the above definitions, to mention just a few, it is obvious that much research in the social domain takes an information-processing approach [e.g. 1-3, 13]. However, criticism against this view has been put forward by a number of researchers [14-18]. Gibbs [14], for instance, addresses two major problems with the traditional view.

Firstly, the traditional view of human intentions as exclusively private mental states in individual minds ignores the dynamic, interactive nature of intentional action. Generally speaking, there is a separation between beliefs, intentions, etc., and the actual behavior of social interactions. This implies that social interaction is a rather passive process between two Cartesian minds, as Gallagher [4] puts it. According to Shanker and King [15], the information-transmission metaphor fails to reveal the full story of social interaction, because it significantly *oversimplifies* and *misrepresents* what actually happens in social interaction. They stress that such interactions cannot be reduced to so-called 'social information transfer'. The main point here is that information is not a predefined and discrete entity which can be sent, through signals, from one agent across time and space to another agent in the form of internal mental representations.

However, a stimulating shift has occurred in socio-cognitive research, from mainly using the information transmission metaphor, to applying the so-called dance metaphor [10, 15, 16]. Broadly speaking, the dance metaphor focuses on *dynamically* emerging, creative co-regulated interaction in a particular social situation, instead of discrete and linear processes as in the information-transmission metaphor. In other words, the dance metaphor focuses on the emergence of information in the dyad between embodied agents. As stressed by Ingold [17, p. 627], "such a shift will … release biological and psychological

studies of communication from the straightjacket of hard-core cognitivism". In line with the dance metaphor, the *distributed cognition* approach [19] expands the unit of analysis and focuses on "real-time" interactions between the various interactants and their environment, instead of focusing on mental structures in individual minds. This means, contrary to viewing cognition as mainly internal processes, social interactions are considered to be *directly observable* cognitive events. With this crucial change in perspective, much of cognition previously hidden 'inside' the skull has now become apparent. This means, information is neither pre-given nor hidden internally, but can emerge in the interaction and be manifested in visible embodied actions.

Secondly, the work of cultural anthropologists addresses another problem with the traditional view. The underlying assumption in the traditional view is not shared across different cultures [14], but the focus on individuals' intentions by rather reflects a Western white middle-class bias regarding the nature of selfhood than a universal phenomenon. It therefore might be argued that individual intentionality is one of the 'holy cows' of Western thought which overemphasizes the individual's psychological state at the expense of the social context in which the actions unfold [14]. The study of the social context, which we refer to as *relational*, has strong historical roots in the work of Mead [20] and Vygotsky [21].

Vygotsky's [21] example of the development of pointing in the child illustrates the relational aspect of social interaction. He claimed that what an observer might perceive as pointing initially is only a simple and incomplete grasping movement directed toward a desired object, and nothing more. When the caretaker comes to help the child, the meaning of the gesture situation itself changes as the child's failed reaching attempt provokes a reaction, not from the desired object, but from the other person. The individual movement 'in itself' in its social context becomes a gesture 'for-others'. The caretaker interprets the child's reaching movement as a kind of pointing gesture, i.e. a socially meaningful communicative act, whereas the child itself at the time is not actually aware of its communication ability. However, after a while the child also becomes aware of the communicative function of its movements, and then begins addressing its gestures towards other people, rather than the object of interest that was its primary focus initially. Thus, "*the grasping movement changes to the act of pointing"* [21, p. 56]. This means, the intention of pointing initially does not reside within the child's individual mind, but emerges as an outcome of their on-going social interactions. Accordingly, by treating children as intentional beings, caregivers 'bootstrap' and scaffold them into a socio-cultural environment, which partly rests on the 'illusion of intentionality'.

Another criticism against the traditional view, not addressed by [14], is its *biological implausibility* and disembodiment. Maturana and Varela [18, p. 196], for instance, pointed out that the traditional metaphor of communication is wrong, since "biologically, there is no 'transmitted information' in communication". A similar argument was put forward by Fogel [16, p. 76] who stated that "information is created in the interface between perception and action … It is that last point, the salience of the body … that is missing in many theories of meaning".

Taken together, there is a need for alternative explanations of social interaction that address the issue from an embodied perspective, since the traditional view, in a nutshell, can be regarded as a disembodied sender-receiver explanation of pre-given information, missing contextual and bodily aspects. The next sections

elaborate in some more detail on these objections, particularly regarding why it might be more fruitful to consider humans as embodied cognizers situated in a social, cultural, and material sphere, and why that might be crucial for social interaction and cognition.

4.3 On the embodied nature of social interaction

Many recent findings in cognitive science and related disciplines indicate that the body has several important roles in social interaction and cognition. Here we briefly address different perspectives and empirical findings, ranging from disciplines such as social psychology, phenomenology, social neuroscience, and gesture studies to linguistics. These findings are then generalized to four fundamental functions of the body in social interaction [10, 22].

4.3.1 Social embodiment effects

Semin and Smith [23] point out that empirical findings in social psychology and current research on embodied cognition have a lot in common, given that several interesting phenomena in social psychology can be explained from an embodied perspective. Barsalou et al. [24], for example, have identified the following four kinds of social embodiment effects for which there is plenty of empirical evidence (for details see [24] and [25]) and the many references therein).

Firstly, *perceived social stimuli do not only produce cognitive states, but also bodily states*. For example, it has been reported that high school students who received good grades in an exam adopted a more erect posture than students who received poor grades. In another experiment, subjects primed with concepts commonly associated with *elderly* people (e.g., 'gray', 'bingo', 'wrinkles') exhibited embodiment effects such as slower movement when leaving the experimental lab, as compared to a control group primed with neutral words. Several other studies also show similar effects.

Secondly, *the observation of bodily states in others often results in bodily mimicry in the observer*. People often mimic behaviors, and subjects often mimic an experimenter's actual behavior, e.g. rubbing the nose or shaking a foot. Subjects also tend to mimic observed facial expressions, which is widely documented in the literature.

Thirdly, *bodily states produce affective states*, which means that embodiment not only facilitates a response to social stimuli but also produces tentative stimuli. For example, subjects rated cartoons differently when holding a pen between their lips than when holding it between their teeth. The latter triggered the same musculature as smiling, which made the subjects rate the cartoons as funnier, whereas holding the pen between the lips activated the same muscles as frowning and consequently had the opposite effect. Moreover, bodily postures influence the subjects' affective state; e.g., subjects in an upright position experienced more pride than subjects in a slumped position.

Fourthly, *compatibility between bodily and cognitive states enhances performance*. For instance, several motor performance compatibility effects have been reported in experiments in which subjects responded faster to 'positive'

words (e.g. ‘love’) than ‘negative’ words (e.g. ‘hate’) when asked to pull a lever towards them.

In more recent work, some of the abovementioned researchers focus explicitly on traditional conceptions in social psychology, such as attitudes, social perception, and emotions. Niedenthal et al. [25], for example, emphasize that empirical studies show that bodily postures and motoric activities, such as nodding heads (in agreement) or shaking heads (in disagreement) are related with positive or negative preferences and action predispositions toward objects. Similarly, others studied how such head movements influence the attitudes towards a pen placed on the table in front of the participants during the cover story of testing head phones. Afterwards, a naïve experimenter offered the ‘old’ pen that had been placed on the table during the experiment or a ‘new’ pen the subjects had not seen before. Depending on the performed head movements, i.e., nodding in agreement or shaking in disagreement, the participants favored the pen that correlated with the developed attitude. In other words, the nodding participants chose the ‘old’ pen, whereas the head-shaking participant preferred the ‘new’ one.

These examples, as well as many other studies, demonstrate that there is a strong relation between embodied and cognitive states in social interaction. In short, the bi-directional swapping between these states occurs automatically without higher knowledge structures. These findings suggest that the body might be used as a resonance mechanism in the process of perceiving others, and it has been suggested that so-called mirror neurons function as the neurobiological underpinning for these social embodiment effects, as discussed in more detail in the following.

4.3.2 Social neuroscience

Recent findings in social neuroscience provide strong evidence for an embodied interpretation of intersubjectivity. For instance, *simulation theories* and work on *mirror neurons* are good examples of more radically embodied views (in Clark’s above sense). In short, the simulation account argues that cognitive processes are achieved by the reactivation of the same neural structures used for physically sensing, moving and acting in the environment, but also in *social interaction* and *cognition* [26-31].

Proponents of simulation theories hold that social understanding essentially is the ability to project oneself into another person’s point of view; *simulating* what it is like to be in the other person’s situation [28-32]. In short, the simulation account argues that cognitive processes are achieved by the reactivation of the same neural structures used for physically sensing, moving and manipulating the environment, but also the conceptualization and understanding of *intersubjectivity* and *language.* The capacity to simulate requires an ability to imitate the ‘inner states’ of another person and it has been supposed that the body and its sensorimotor processes can be used as a linking device when perceiving others[1]. Gallese’s [31] theory of the *shared manifold* of intersubjectivity, for example, proposes that all kinds of

[1] It should be noted, however, that this view should not be misinterpreted as claiming there is a direct correlation between so-called ‘objective’ neurological states in the brain and ‘subjective’ phenomenological experience On the contrary, as pointed out by Gallagher [4], bridging the troubled water of social cognitive neuroscience and phenomenology through a direct mapping is no viable approach, because “there is no short cut that can bypass the effects of embodiment” [4, p. 244].

interpersonal relations depend, at a basic level, on the foundation of a shared manifold space characterized by routines of embodied simulations. Simulation is presumably accomplished through the "sharing" of neural mechanisms between sensorimotor processes and higher-level processes. Such accounts show that the traditional strong division between perception and action as well as between sensorimotor and cognitive processes, might need to be revised. According to Svensson [32], an important factor in understanding the embodiment of higher-level cognition, is to consider embodied simulation as offline representation. This means, embodied simulation processes can function as offline representations, i.e., generally speaking, the internal replication of agent-environment interaction for issues beyond the "here and now". [32].

Such an understanding may rely on a resonance mechanism, being part of a particular type of visuo-motor neurons found in pre-motor cortex of the macaque monkey brain, so-called *mirror neurons*, which exemplify how perception, action, and social cognition, might come together at the level of single neurons. Mirror neurons are located in area F5 in the monkey brain and become activated both when *performing* specific goal-directed hand (and mouth) movements and when *observing* or *hearing* about the same actions [26-28]. Since mirror neurons respond in both conditions, it has been argued that the mirror system functions as a kind of action representation, linking 'action' and 'action-perception'. Such a mirroring mechanism might enable an agent to understand the meaning of the observed action by embodied reactivation. This means, even while only *observing* the actions of another individual, a neural 'triggering' event in fact takes place in the observer. Accordingly, the linking between action and perception offers an 'intuitive' first-person understanding of the observed action.

In later studies, mirror neurons have been investigated under two conditions, namely hidden and full visual scenes [33]. In the visual condition, the monkey was able to see the entire action, for example, a hand-grasping movement. In the hidden condition, the same action was carried out, but its crucial and final part, i.e., the interaction with the actual object, was invisible, and the monkey merely 'knew' that the target object was present. The result, however, demonstrated that more than half of the mirror neurons responded in the hidden condition [33]. This implies that the intention behind the action actually was mediated, despite the fact that the monkey did not see the actual hand-object interaction. That is, the goal of the action was still hinted at, given that the gap of missing visual information is filled by reactivating the complete action. This means, the mirror neurons are able to compensate for the missing 'information', and still seem to interpret the actual goal of the action.

More recent work on the activation of the mirror neuron system has been performed in specific contexts (such as before and after drinking tea) [34]. The study indicates that a certain kind of mirror neurons, so called *logically related* mirror neurons, may constitute the foundation for intentionality. Traditionally, the description of an action and the interpretation of the reason why that particular action is performed have been considered to rely on two different mechanisms. The mirror neuron system, however, provides an alternative solution, given that logically related mirror neurons automatically code the motor acts that are most expected to follow the observed action in a particular context [34]. This means, the ability to infer the forthcoming new goal is already 'there' in the mirror neuron system and explaining intentionality by two different mechanisms is both

unnecessary and biologically implausible. Another study [35] implies that information of intentions might be conveyed also by the grasping action itself. Their data suggests that the human mirror neuron system uses both contextual and action-type information (precision-grip vs. whole-hand prehension) to predict others' intentions. Furthermore, there are indications that mirror neuron activity is linked to social competence [35]. Hence, it has been speculated that the mirror system might be a basic mechanism necessary for imitation and attributing mental states to others [26-35]. This implies that during the course of ontogeny, the mirror neuron system and simulation processes might develop further, through maturation as well as socially scaffolded interaction, to more advanced forms.

Based on these findings, Gallese stresses "there is now enough empirical evidence to reject a disembodied theory of the mind as biologically implausible" [32, p. 166]. All in all, the consideration of the mirror neuron system and simulation theories as the neurobiological underpinning of social interaction and cognition provides significant examples of more 'radically' embodied views of intersubjectivity.

4.3.3 Embodied linguistics

In addition to action recognition, mirror neurons are also considered to be involved in more complex social actions, such as gesture and language. Rizzolatti and Arbib [27], for instance, suggest that the human communicative and linguistic capacity is a natural extension of action recognition based on mirror neuron mechanisms. This provides a tentative explanation of why and how the human Broca's area, involved in gesture and language processes, emerged from area F5 in the monkey brain. Arbib [36], for instance, suggests that the mirror system provides the causal mechanism for basic intentional interaction and thus might constitute the foundation of human language. As Rizzolatti [26] points out, however, it is obvious that the mirror neuron mechanism itself is unable to explain the whole complexity of speech and human language, but it actually clarifies one of the fundamental aspects of intersubjectivity, namely how interacting partners are able to share the communicated meaning of a dialogue. In other words, the epistemological divide (i.e., verbal versus non-verbal interaction) in linguistics may be bridged from an embodied perspective.

Several researchers have demonstrated converging empirical evidence which suggests that the systems of hand and mouth movements are not two separate systems. Rather, they should be viewed as an integrated communicative "speech-language-gesture" system, linking action, thought and cognition. McNeill [37], for example, proposed that speech and gesture form a single system of communication, grounded in a common underlying thought process, emphasizing that "[g]estures do not just reflect thought but *have an impact on thought.* Gestures, together with language, *help constitute thought.*…Gestures occur, according to this way of thinking, because they are a part of the speaker's ongoing thought process" [37, p. 245].

Despite the close connection between gesture and speech in language, they generally differ in *how* they carry meaning [4, 37, 38]. Gesture offers alternative ways of expressing ideas that are hard to articulate in speech, as well as when there is no proper word at hand for the actual meaning to be conveyed. Furthermore, gestures can present different pieces of information simultaneously, which in

speech would need to be expressed sequentially. Goldin-Meadow also discovered that speech and gesture convey *different* information, but not necessarily conflicting meaning [38].

Furthermore, it has been argued that gestures also have representational properties. Goldin-Meadow [38], for example, emphasizes that the gestures accompanying speech are *symbolic acts that convey meaning*. Gallagher [4] emphasizes the fundamental difference between instrumental acts (e.g., opening a jar or reaching out to pick up a glass), and the generation of a gesture *signifying* the very action of opening a jar or picking up the glass. In other words, the act of gesture achieves an entirely different function than the actual grasping or opening, because those actions have representational content, which is a cognitive and possibly a communicative function that requires the generation and expression of meaning [4]. Accordingly, gesture is a natural part of communication, and enables people to embody their thoughts in action.

Nevertheless, many researchers still overlook the integrated nature of speech and gesture in the evolution of human language [e.g. 27, 39]. McNeill [40], on the contrary, emphasizes the double characteristic of language, i.e., speech and gesture in the course of *joint action* in evolution. Brain area 44 is mainly responsible for the organization of action sequences, whereas area 45 is the part of Broca's area that contains many mirror neurons, which McNeill suggests became self-responding to one's own actions, subsequently imbuing them to contain meaning. During the course of phylogeny, these two systems became co-opted in order to unite gesture and vocalization [40].

According to McNeill, the crucial shift in the function of mirror neurons occurred when they began to respond to significances other than the actions themselves, as a way of co-opting areas 44 and 45 in Broca's area, providing the basis for recognizing the actions of others. In other terms, this co-opted system seems to be part of a circuit for recognizing intentional goal-directed actions from one's own actions or from others. It should be stressed, that McNeill emphasizes the *relational* nature of the mirror neuron system which, in our opinion, is overlooked in many other theories. McNeill refers to Mead whom argued that "[g]estures become significant symbols when they implicitly arouse in an individual making them the same response which they explicitly arouse in other individuals" [40, p. 250]. Hence, meaningfulness emerges from the ability to activate a social reaction of another in yourself, a way of reacting in your own actions similarly to the actions of others, which McNeill denotes *Mead's loop*.

This means, gesturing also has the important role of activating our own mirror neuron system, as well as offering oneself the ability to take the role/perspective of the other simultaneously [40]. Thus, the shift in social interaction, as previously described in Vygotsky's pointing example can partly be explained neurologically by Mead's loop. In a similar vein, Gallagher [41], for example, argues that phenomenologically, when one sees another person's action or gesture, one *directly perceives* or immediately 'sees' the meaning in the action/gesture, without the need to model it at a higher cognitive level. His main point is that the relevant neural systems "are *activated by the other person's action"*. Thus, "the other person *has an effect on us*" [41, p. 8-9]. In other words, Mead's loop creates a connection of gesture to discourse, given that this relational characteristic is also present in speech.

This implies that bodily actions might be of crucial importance in the process of intersubjectivity. From a radically embodied perspective, the activation and/or reactivation of the mirror neuron system, together with other bodily mechanisms, might function as the glue that binds hand, mouth and language together, in a social and cultural sphere.

4.3.4 Four fundamental functions of the body in social interaction

In summary, the work presented in the previous sections offers highly complementary rather than alternative views on the role of embodiment in intersubjectivity. By integrating these perspectives, we can obtain a deeper understanding without bypassing the effects of embodiment. Based on the previous discussions and empirical findings, four fundamental functions of embodiment in social interaction can be identified [for more details, see 10, 21]. It should be noted that these fundamental functions are not fixed, and, to some degree, overlapping.

- *The body functions as a social resonance mechanism.*
- *The body functions as a means and end in communication and social interaction.*
- *Bodily actions and gestures function as a helping hand in shaping, expressing and sharing thoughts.*
- *The body functions as a representational device.*

The body functions as a social resonance mechanism suggests that there is no need to decode or represent embodied social stimuli to more 'advanced' or cognitive states since the bodily states in themselves actually are cognitive states, as the work of Barsalou et al. [24] and others shows [25]. Hence, this first function characterizes how cognitive/bodily states of interacting partners are reflected both in themselves and in-between them at a basic level, during both online and offline interactions. The examples presented in this chapter, as well as other studies, demonstrate there is a strong relation between embodied and cognitive states in social interaction, since the bi-directional exchange between these states as well as between the interacting partners, occur automatically without the involvement of higher knowledge structures.

The body functions as a means and end in communication and social interaction. The suggested linkage between 'action' and 'action-perception' provided by the mirror neuron system implies that the body and its sensorimotor processes are 'cognitive' in themselves. The great benefit of this action-understanding linkage, beside its parsimony, is the inbuilt *dual* ability of grasping both the 'what' and 'why' aspects of the present action, i.e., what the action is about as well as catching the intention behind the movement. Hence, this second function stresses how bodily actions operate both outwardly and inwardly in meaning-making activity, e.g., through *Mead's loop*. The functions of the body as a 'resonance mechanism' and also 'a means and end' might seem quite similar. However, while the function of the body 'as a resonance mechanism' simply means that cognitive and bodily states of the interacting partners are reflected in both themselves and in-between them, it does not explain the relationship between their first-hand and third-hand experiences in social interaction. Instead, viewing the function of the body 'as a means and end' offers a tentative explanation of that particular linkage, thereby unifying the 'inside and 'outside' perspectives of

socially embodied interaction. In other words, the previously portrayed function of embodiment in social interaction mostly stresses 'that' the body and its sensorimotor processes function as a social resonance mechanism, whereas the second function rather focuses on 'how' this is accomplished.

Bodily actions and gesture function as a helping hand in shaping, expressing and sharing thoughts. Besides speech, manual gesture is a significant (embodied) aspect of meaning-making activity, which can provide important information to the listener, since gesture offers speakers the means of expressing thoughts difficult to articulate in speech. Through gesturing, we are able to generate and embody dynamical associations between different matters, which can offer new insights to the present situation or problem at hand. In addition, gesture sometimes serves as an explicit instance of the action-meaning embodied in speech, suggesting that hand movements are physical externalizations of the speaker's ideas.

The body functions as a representational device. In addition to speech, there is the more controversial claim that non-vocal embodied action also has representational properties, where certain kinds of gesture, portraying representational aspects, are the most obvious examples of the body as an external representational device. Furthermore, the internal reactivation of "agent-environment-interaction", in the form of embodied simulation, can be considered as representations in a strict sense. The neurological roots of this ability might be the activity of the mirror neurons, since their linkage between 'action' and 'action-perception' suggests a kind of 'action representation' that is directly enacted in social interaction. Furthermore, since mirror neurons seem to 'understand' the goal of the action, it can be argued that the grasping of the action does not require a declarative understanding, since it is meaningful in itself.

4.4 Illustration and discussion

The previous section has summarized some of the arguments for the view that it is first and foremost the enacted body, and the experiences that come from its situatedness in a social and cultural sphere, that constitute the roots of social interaction and cognition.

In order to further illustrate this view, some frame-by-frame analyzed images from an episode of spontaneous social interaction captured *in situ* are presented in the following. The examples are from the first author's fieldwork on a horse ranch that maintains and preserves Spanish mustang horses [for more details, see 10]. We here briefly illustrate and discuss the role of scaffolding [42-43] (see Figure 1). In this episode the head of the ranch, Bob, is telling the visitors about the different places that nowadays keep herds of horses originating from the ranch's herds. In summary, all his bodily actions, i.e., facial expression, tone of voice, bodily posture, gesture, speech as well as gaze, reflect the significance of cross-modal embodied actions in social interaction. It should be pointed out that throughout the following analysis the earlier identified four functions of the body in social interaction are used as the underpinnings for describing and explaining how embodiment is part and parcel of social interaction and cognition. They are, however, not always mentioned explicitly throughout the analysis, since this

would result in many unnecessary reiterations rather than adequate descriptions and analyses of the interactions at hand.

Throughout the enumeration of different places, Bob uses his fingers as scaffolds in several ways. First, through slight tapping actions, he uses his fingers as a means to put the places in order. For each location, he touches his fingers in a certain order as a way to inform himself to keep track of the places. These tapping actions signify the representational aspect of gesturing and therefore convey meaning in their own way, given that the actual gestures are signs of another aspect other than the actual tapping movement. That is, the tapping actions serve as a way to keep track of the different *locations*. Second, the tapping action is flawlessly integrated with the speech utterance of the name of the location, thus indicating and highlighting the central information in the utterance in both speech and gesture. That is, the two most important aspects of the utterances, the *number* and *location*, are manifested in speech as well as gesture simultaneously.

Figure 1. Bob counting on his fingers

Third, in order to re-enact/remember the different locations, the very actions of moulding, moving, and contacting his fingers facilitate the process of remembering the locations and their names. That is, the act of moving his fingers functions as a way of *shaping* and *expressing* thoughts, given that in the precise moment when he has figured out the name of the location, he stops moving the current finger, and then touches the other hand's finger. That is, the movement *shows the status* of the act of remembering.

Taken together, this example demonstrates the act of scaffolding, by using one's fingers as *representational devices* during the enumeration of the places. In total, there are nine different herds, situated at seven locations. Bob's cognitive strategy of off-loading the act of remembering into a visual and external representational format through embodied actions is very obvious and observable in the above examples. However, the very action and experience of moving/touching his fingers actually brings forth the names, by facilitating the shaping of the numbers and locations, instead of functioning as a way of externalizing the already existing names of the places.

More specifically, this strategy has far more wide reaching consequences than stated above. The explanation thus far focuses on what one might term *individual scaffolding*, but it should be noted that Bob is actually both in relation with himself and with the others. By using his fingers he creates and experiences action-

perception loops through his body, given that the actual movements provide him with the kinesthetic experience of movement as well as the felt sense of touching when he grips and senses his fingers during the enumeration of places, in which every embodied action he makes creates a spectrum of embodied experiences to him. Thus, he constructs and experiences sensory-motor information at the same time through the crossmodal integration of his embodied being. This type of *relational embodied scaffolding* functions as a way of being in dialogue with oneself as well as with others, and is accomplished by the activation of his own mirror neuron system through *Mead's loop*. This means, most of the enumeration and the act of remembering is an *intra*-personal interaction, but there is also an *inter*-personal theme present.

Additional evidence in favor of this interpretation emerges during the frame-by-frame analysis of the actions. The tiny movement and touching of Bob's fingers when he uses them as scaffolds in the naming of the herds, are almost 'invisible' and perceived unconsciously in real time. Due to the ways our cognition is embodied, the body 'knows' and 'grasps' directly what is going on. Thus, the effects of embodiment do the job for us. Indeed, one might ask – *why* do humans perform such actions so frequently when they are almost 'invisible' and therefore not necessarily communicative in any obvious way. What role and relevance do they actually have if they are not used for social interaction and communication in the first place?

Our tentative answer to these questions is that these actions are not first and foremost inter-communicative, but also function inwardly. This means, they are both intra- and inter-communicative, stressing the relational aspect of social interaction and cognition that is profoundly manifested in our embodiment. Thus, the embodied nature of social interaction and cognition unifies the individual and social perspectives. For instance, in Mead's loop, gesture and language are displayed but their *relational* characteristics are the same – they are both external actions that we can act upon in the public sphere and internal embodied actions used to organize and structure our internal and sometimes abstract and decoupled thinking, though still grounded in embodied experience.

Furthermore, Bob's enumeration is a significant example of what Clark [44] refers to as *surrogate situatedness.* According to Clark, "human reason is … disengaged but not disembodied" [44 p. 236] and there is no sharp line between so-called online versus offline cognition, given that both processes are running in parallel. He argues that humans create and use human-built structures in order to transform the space of higher-level cognition, and stresses that "we actively create restricted artificial environments that allow us to deploy basic perception-action-reason routines in the absence of their proper objects" [44, p. 233]. According to Clark, these strategies allow human cognition to be disengaged while at the same time offering a concrete place in which to organize action-perception couplings of an essentially real world-like kind of interaction. Broadly speaking, whenever people act/think/communicate they are always in interaction, either with themselves or with other interactants or objects.

4.5 Conclusions

The aim of this chapter has been to clarify the role and relevance of the body in social interaction and cognition, and we have tried to present an integrated, interdisciplinary theoretical foundation for studying the embodied nature of social interaction and cognition.

What ties all these issues together is the idea of the social mind as being *relational*, 'radically' *embodied*, and situated in the social sphere. That means, the social dimension meets the physiological dimension, thus 'reaping the best of both worlds' without neglecting the effects of embodiment for social interaction. What further unites all these issues, as hand occurs in glove, is how profoundly embodiment shapes social interaction and cognition through unfolding socially embodied actions in social and cultural contexts. As discussed in this chapter, the key to this coherent union is the way our social mind is embodied, a fact that should not be neglected or trivialized.

To summarize, the ways humans are embodied imply that one's own understanding of social interaction is more than the exchange of communication signals between disembodied information-processors. Instead, meaning and intentions are *emergent products* of socially embodied interaction, and in many situations they can be viewed as distributed phenomena rather than as individual private mental acts or properties.

4.6 References

[1] M. Augoustinos & I. Walker, *Social cognition*. London: Sage, 1995.
[2] Z. Kunda, *Social cognition*. Cambridge, MA: MIT Press, 1999.
[3] K. A. Quinn, N. C. Macrae & G. V. Bodenhausen, Social cognition. *Encyclopedia of Cognitive Science*, (pp.66-73). London: Macmillan, 2003.
[4] S. Gallagher, *How the body shapes the mind*. Oxford: Oxford University Press, 2005.
[5] A, Clark, An embodied cognitive science? *Trends in Cognitive Sciences, 3*(9), 345-351, 1999.
[6] R. W. Gibbs, Jr., *Embodiment and cognitive science.* Cambridge: Cambridge University Press, 2006.
[7] G. Lakoff & M. Johnson, *Philosophy in the flesh.* New York: Basic Books, 1999.
[8] F. J. Varela, E, Thompson & E. Rosch, *The embodied mind.* Cambridge: MA: MIT Press, 1991.
[9] T. Ziemke, J. Zlatev & R. Frank (Eds.), *Body, language, and mind: Embodiment,1*. Berlin: Mouton de Gruyter, 2007.
[10] J. Lindblom, *Minding the Body – interacting socially through embodied action.* Doctoral dissertation, University of Linköping/University of Skövde, Sweden. ISBN 978-91-85831-48-7, 2007
[11] S. Bråten, Participants perception of others' acts. *Culture & Psychology, 9* (3), 261-276, 2003.
[12] P. Rochat & T. Striano, Social-cognitive development in the first year. In P. Rochat (Ed.), *Early Social Cognition,* (pp.3-34). Mahwah, NJ: Lawrence Erlbaum, 1999.
[13] T. Singer, D. Wolpert & C. Frith, Introduction. In C. Frith & D. Wolpert, *The Neuroscience of Social Interaction,* (pp. xiii-xxvii). Oxford: Oxford University Press, 2004.
[14] R. W. Gibbs, Jr., Intentions as emergent products of social interactions. In B. F. Malle, L. J. Moses & D. A. Baldwin (Eds.), *Intentions and Intentionality*, 105-122. Cambridge, MA: MIT Press, 2001.
[15] S. G. Shanker & B. J. King, The emergence of a new paradigm in ape language research. *Behavioral & Brain Sciences, 25*, 605-656, 2002.
[16] A. Fogel, *Developing through relationships.* New York: Harvester Wheatsheaf, 1993.
[17] T. Ingold, Communication and communion. *Behavioral & Brain Sciences, 25*(5), 627-628, 2002.
[18] H. Maturana & F.Varela, *The tree of knowledge*. Boston: Shambhala, 1987.
[19] E. Hutchins, *Cognition in the wild.* Cambridge, MA: MIT Press, 1995.
[20] G. H. Mead, *Mind, self and society*. Chicago: Chicago University Press, 1934.

[21] L. S. Vygotsky, *Mind in society*. Cambridge, MA: Harvard University Press, 1978.
[22] J. Lindblom, Embodied action as a 'helping ' hand in social interaction. *Proceedings of the 28th Annual Conference of the Cognitive Science Society*, (pp. 477-482). Mahwah, NJ: Lawrence Erlbaum, 2006.
[23] G. R. Semin & E. R. Smith, Interfaces of social psychology with situated and embodied cognition. *Cognitive Systems Research, 3*(3), 385-396, 2002.
[24] L.W. Barsalou, P. M. Niedenthal, A. K. Barbey & J. A. Ruppert, Social embodiment. In B. H. Ross (Ed.), *The Psychology of Learning and Motivation, 43*, (pp. 43-92). San Diego, CA: Academic Press, 2003.
[25] P. M. Niedenthal, L. W. Barsalou, P. Winkielman, S. Krath-Gruber & F. Ric, Embodiment in attitudes, social perception, and emotion. *Personality and Social Psychology Review, 9*(3), 184-211, 2005.
[26] G. Rizzolatti, The mirror neuron system and its function in humans. *Anatomical Embryology, 210*, 419-421, 2005.
[27] G. Rizzolatti & M. A. Arbib, Language within our grasp. *Trends in Neurosciences, 21*, 188-194, 1998.
[28] V.Gallese, C.Keysers & G.Rizzolatti, A unifying view of the basis of social cognition. *Trends in Cognitive Sciences, 8* (9), 398-403, 2004.
[29] H. Svensson, J. Lindblom & T. Ziemke, Making sense of embodied cognition: simulation theories of shared neural mechanisms for sensorimotor and cognitive processes. In T. Ziemke, J. Zlatev & R. Frank (Eds.) *Body, language, and mind: Embodiment,1*, (pp. 241-270). Berlin: Mouton de Gruyter, 2007.
[30] J. Lindblom & T. Ziemke, Embodiment and social interaction: implications for cognitive science. In T. Ziemke, J. Zlatev & R. Frank (Eds.), *Body, language, and mind: Embodiment, 1*, (pp.129-162). Berlin: Mouton de Gruyter, 2007.
[31] V. Gallese, The manifold nature of interpersonal relations: the quest for a common mechanism. In C. Frith & D. Wolpert (Eds.) *The neuroscience of social interaction: decoding, imitating and influencing the actions of others*, (pp.159-182). Oxford: Oxford University Press, 2004.
[32] H. Svensson, *Embodied simulation as off-line representation.* Licentiate thesis, University of Linköping/University of Skövde, Sweden. ISBN 978-91-85831-83-8, 2007.
[33] M.A.Umiltà, E. Kohler, V. Gallese, L. Fogassi, L. Fadiga, C. Keysers & G. Rizzolatti, (2001) "I know what you are doing": a neurophysiological study. *Neuron, 32*, 91-101, 2001.
[34] M. Iacoboni, I. Molnar-Szakacs, V. Gallese, G. Buccino, J. C. Mazziotta & G. Rizzolatti, Grasping the intentions of others with one's own mirror neuron system. *PLoS Biology, 3*(3):e79, 529-535, 2005.
[35] T. Kaplan & M. Iacoboni, Getting a grip on other minds. *Social Neuroscience, 1*(3-4), 175-183, 2006.
[36] M. A. Arbib, From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. *Behavioral and Brain Sciences, 28*, 105-167, 2005.
[37] D. McNeill, *Hand and mind*. Chicago: Chicago University Press, 1992.
[38] S. Goldin-Meadow, *Hearing gesture*. Cambridge, MA: Belknap Press, 2003.
[39] J. Zlatev, Embodiment, language, and mimesis. In: T. Ziemke J. Zlatev, & R. Frank (Eds.), *Body, language, and mind: Embodiment, 1*, (pp. 297-337). Berlin: Mouton de Gruyter, 2007.
[40] D. McNeill, *Gesture and thought.* Chicago: The University of Chicago Press, 2005.
[41] S. Gallagher, Simulation trouble. *Social Neuroscience, 2*, 1-13, 2007.
[42] D. Wood, J. S. Bruner & G. Ross, The role of tutoring in problem-solving. *Journal of Child Psychology and Psychiatry, 17*, 89-100, 1976.
[43] M. Wilson, Six views of embodied cognition. *Psychonomic Bulletin & Review, 9*(4), 625-636, 2002.
[44] A. Clark, Beyond the flesh: some lessons from a mole cricket. *Artificial Life, 11*, 233-244, 2005.