Strong Interaction and Self-Agency

Shaun Gallagher *
gallaghr@mail.ucf.edu

ABSTRACT

The interaction theory of social cognition contends that intersubjective interaction is characterized by both immersion and irreducibility. This motivates a question about autonomy and self-agency: If I am always caught up in processes of interaction, and interaction always goes beyond me and my ultimate control, is there any room for self-agency? I outline an answer to this question that points to the importance of communicative and narrative practices.

In regard to social cognition, there has been growing opposition to the standard theory-of-mind (ToM) views, usually referred to as theory theory (TT) and simulation theory (ST). I have defended an alternative approach: "interaction theory" (IT). IT is based on evidence from both phenomenology and developmental psychology, and it offers an alternative to the simulation interpretation of the neuroscience of mirror neurons. An important part of IT is its emphasis on 'strong interaction' (Gallagher in press; also see De Jaeger *et al.* 2010) – a concept of interaction that is a seemingly pervasive feature of intersubjectivity. In this paper I take a closer look at this concept and raise questions about what appears to be a threat to the notion of self-agency. The question is: If we are so interactively interdependent on others in our everyday practical and communicative behaviors, is there any room for autonomy?

Interaction Theory (IT) as an Alternative to TT and ST

In psychology, philosophy of mind, and more recently, in the neurosciences, studies of how one person understands and interrelates with another person

* Department of Philosophy, University of Central Florida – Institute of Simulation and Training – University of Hertfordshire – University of Copenhagen

have been dominated by two main approaches: theory theory and simulation theory. The major tenets of TT are based on scientific experiments that show that children develop an understanding of other minds around the age of four. One version of TT claims that this understanding is based on an innately specified, domain specific mechanism designed for 'reading' other minds (e.g., Baron-Cohen 1995, Leslie 1991). An alternative version claims that the child attains this ability through a course of development in which the child tests and learns from the social environment (e.g., Gopnik and Meltzoff 1997). Common to both versions is the idea that children attain their understanding of other minds through the use of folk or commonsense psychology which they use to make theoretical inferences about certain entities to which they have no access, namely, the mental states of other people. When we make such inferences and attribute specific mental states to others, we are said to be *mentalizing* or *mindreading*.

ST, in contrast, argues that rather than theorizing or making inferences about the other person's mind, we use our own mental experience as an internal model for the other mind (e.g., Gordon 1986, 1995; Heal 1986, 1998a, 1998b). To understand the other person, I simulate the thoughts or feelings that I would experience *if I were in the situation of the other*, exploiting my own motivational and emotional resources. I imagine what must be going on in the other person's mind; or I create in my own mind pretend beliefs, desires or strategies that I use to understand the other's behavior. My source for these simulations is not a theory that I have. Rather, I have a real model of the mind at my immediate disposal, that is, I have *my own mind*, and I can use it to generate and run simulations. I simply run through the sequence or pattern of behavior or the decision-making process that I would engage in if I were faced with the situation in question. I do it 'off line', however. That is, my imaginary rehearsal does not lead to actualizing the behavior on my part. Finally, I attribute this pattern to the other person who is actually in that situation.

Despite extensive debates between proponents of TT and ST, respectively, TT and ST share three basic suppositions. The three suppositions are these.

(1) The problem of social cognition is due to the lack of access that we have to the other person's mental states. Since we cannot directly perceive the other's thoughts, feelings, or intentions, we need some extra-perceptual cognitive process (mindreading or mentalizing) that will allow us to infer or simulate what they are.

(2) Our normal everyday stance toward the other person is a third-person, observational stance. Based on what we observe we use mindreading to *explain* or *predict* their behaviors.

(3) These mentalizing processes constitute our primary and pervasive way of understanding others.

There are also a number of unsolved problems associated with these ToM approaches. I won't go into detail here, but I'll give a brief indication of some of these problems.[1] First, some (but not all) theorists claim the process of theoretical inference or simulation is conscious or introspective (e.g., Goldman 1995; Goldman 2006, p. 147); but there is no phenomenological evidence that this is so, and there should be if the process is both consciously explicit and pervasive. That is, we should be able to catch ourselves in the act, but we don't. The second problem is what I refer to as the starting problem, a version of the frame problem. For TT, the question is how do I know what piece of folk psychology (what rule, or what platitude) actually applies to the case at hand. For ST, one can see the problem clearly in the following description of a simulation routine provided by Nichols and Stich:

> The basic idea of what we call the 'off-line simulation theory' is that in predicting and explaining people's behavior we take our own decision making system 'off-line', supply it with 'pretend' inputs that have the same content as the beliefs and desires of the person whose behavior we're concerned with, and let it make a decision on what to do. (Nichols and Stich 2003, pp. 39-40)

Simulation as a form of mindreading is supposed to provide insight into the beliefs and desires of the other person, but it seems that we need to know the content of those mental states in order to do the simulation. Neither TT nor ST provide a good answer to the starting problem.

A third problem concerns diversity and applies specifically to ST. Keysers and Gazzola describe simulation in the following way:

> In [simulation] cases, observing what other people do or feel is therefore transformed into an inner representation of what we would do or feel in a similar situation – as if we would be in the skin of the person we observe. (Keysers and Gazzola 2006, p. 390)

But how does knowing what we would do help us know what someone else

[1] See Gallagher 2005 and 2007 for more detail.

would do? Indeed, many times we are in a situation where we see what someone is doing, and know that we would do it differently, or perhaps, not do it at all. A fourth problem concerns development. The kind of inferential or simulation processes found in explicit versions of TT and ST are too cognitively complex to account for the infant's ability to understand the intentions of others. Yet, as we'll see below, there is a large amount of evidence to support the idea that young infants are able to grasp the intentions of others.

The developmental problem is addressed by a recent version of ST that relies on an interpretation of mirror neuron (MN) activation as a form of simulation. In this case, simulation is said to be fast and automatic. If MNs are active in young infants, then the developmental problem does not apply to this version of ST. Since activation of MNs are non-conscious, the issue of phenomenological evidence is irrelevant, and there is no starting problem. So-called neuronal ST, then solves all of the above problems except perhaps the diversity problem. But there are other problems involved in neural ST. One concerns the fact that simulation is originally defined as involving pretense. As Nichols and Stich make clear in the above quote, simulation involves the use of pretend beliefs and desires. We pretend to be in the other person's shoes in order to run the routine. But the notion of pretense does not apply in the case of MNs. Indeed, most theorists claim that MNs are neutral with respect to who the agent is, and agent-neutrality is not consistent with the notion of pretense. MNs can't account for me pretending to be you if in fact there is no distinction between me and you at that level. As a result, there have been attempts to shift the definition of simulation to involve a simple matching (e.g., Goldman 2006, Rizzolatti *et al.* 2001). My motor system is said to go into a state matching yours when I see you perform an action. But the neurological details do not bear this out[2], and it seems counter-intuitive if we think of how we interact with others. In the majority of cases we are not imitating or mimicking others; rather, our motor systems are busy supporting responses or complementary actions.

This is not an exhaustive list of problems with TT and ST, but it should be sufficient to see why we might want to find a better account of social cognition. Interaction theory is proposed as that better account. IT challenges the three suppositions associated with ToM approaches. In their place IT argues for the following propositions.

[2] See, e.g., Catmur *et al.* 2007, Dinstein *et al.* 2008.

(1) Other minds are not hidden away and inaccessible. The other person's intentions, emotions, and dispositions are expressed in their embodied behavior. In most cases of everyday interaction no inference or projection to mental states beyond those expressions and behaviors is necessary.

(2) Our normal everyday stance toward the other person is not third-person, detached observation; it is second-person interaction. We are not primarily spectators or observers of other people's actions; for the most part we are interacting with them in some communicative action, on some project, in some pre-defined relation; or we are treating them as potential interactors.

(3) Our primary and pervasive way of understanding others does not involve mentalizing or mindreading; in fact, these are rare and specialized abilities that we develop only on the basis of a more embodied engagement with others.

IT emphasizes the role of communicative and narrative practices, and it appeals to evidence from developmental studies, starting with primary and secondary intersubjectivity (Trevarthen 1979; Trevarthen and Hubley 1978).

- Primary intersubjectivity (starting from birth) – Sensory-motor abilities – enactive perceptual capacities in processes of interaction
- Secondary intersubjectivity (starting around 1 year of age) – joint attention, shared contexts, pragmatic engagements, acting *with* others
- Communicative and narrative competencies (starting from 2-4 years) – communicative and narrative practices that represent intersubjective interactions, motives, and reasons and provide a more nuanced and sophisticated social understanding.

In this paper I will begin with a focus on primary and secondary intersubjectivity, but I'll return the issue of narrative competence. I take this strategy because I first want to focus on the nature of interaction itself, and most of the essential aspects can be grasped in primary and secondary intersubjectivity. When it comes to the question of self-agency, however, narrative will be shown to play an important role.

The Development of Interaction

Primary intersubjectivity consists of the innate or early-developing sensory-motor capacities that bring us into relation with others and allow us to interact with them. These capacities are manifested at the level of perceptual experience -- we *see* or more generally *perceive* in the other person's bodily movements, gestures, facial expressions, eye direction, etc. what they intend and what they feel, and we respond with our own bodily movements, gestures, facial expressions, gaze, etc. From birth the infant is pulled into these interactive processes. This can be seen in the very early behavior of the newborn. Infants from birth are capable of perceiving and imitating facial gestures presented by another (Meltzoff and Moore 1977, 1994). Importantly, this kind of imitation is not an automatic or mechanical procedure; Csibra and Gergely (2009) have shown, for example, that the infant is more likely to imitate only if the other person is attending to it.

Primary intersubjectivity can be specified in more detail as the infant develops. At 2 months, for example, infants are able to follow the gaze of the other person, to see that the other person is looking in a certain direction, and to sense what the other person sees (which is sometimes the infant herself), in a way that throws the intention of the other person into relief (Baron-Cohen 1995; Maurer and Barrera 1981). In addition, second-person *interaction* is evidenced by the timing and emotional response of infants' behavior. Infants «vocalize and gesture in a way that seems [affectively and temporally] 'tuned' to the vocalizations and gestures of the other person» (Gopnik and Meltzoff 1997, p. 131). At 5-7 months, infants are able to detect correspondences between visual and auditory information that specify the expression of emotions (Walker 1982; Hobson 1993, 2002). At 6 months infants start to perceive grasping as goal directed, and at 10-11 months infants are able to parse some kinds of continuous action according to intentional boundaries (Baldwin and Baird 2001; Baird and Baldwin 2001; Woodward and Sommerville 2000). They start to perceive various movements of the head, the mouth, the hands, and more general body movements as meaningful, goal-directed movements (Senju *et al.* 2006).

Developmental studies show the very early appearance of, and the importance of, interactive attunement in the form of timing and coordination in the intersubjective context. In still face experiments, for example, infants are engaged in a normal face-to-face interaction with an adult for 1 to 2 minutes,

followed by the adult assuming a neutral facial expression. This is followed by another normal face-to-face interaction. Infants between 3 and 6 months become visibly discouraged and upset during the still face period (Tronick 2007, Tronick *et al.* 1978). The importance of interactive touch has also been demonstrated in the still-person effect (Muir 2002).

Murray and Trevarthen (1985) have also shown the importance of the mother's live interaction with 2-month old infants in their double TV monitor experiment where mother and infant interact by means of a live television link. The infants engage in lively interaction in this situation. When presented with a recorded replay of their mother's previous actions, however, they quickly disengage and become distracted and upset. These results have been replicated, eliminating alternative explanations such as infants' fatigue or memory problems (Nadel *et al.* 1999, Stormark and Braarud 2004).

Primary intersubjectivity is not something that disappears after the first year of life. It is not a stage that we leave behind, and it is not, as Greg Currie suggests, a set of precursor states «that underpin early intersubjective understanding, and *make way* for the development of later theorizing or simulation» (Currie 2008, p. 212, *my emphasis*).[3] Rather, citing both behavioral and phenomenological evidence, IT argues that we don't leave primary intersubjectivity behind; the processes involved here don't "make way" for the purportedly more sophisticated mindreading processes – these embodied interactive processes continue to characterize our everyday encounters even as adults. That is, we continue to understand others in strong interactional terms, facilitated by our recognition of facial expressions, gestures, postures, and actions as meaningful.

Scientific experiments bear this out. Point-light experiments (actors in the dark wearing point lights on their joints, presenting abstract outlines of emotional and action postures), for example, show that not only children (although not autistic children) but also adults perceive emotion even in movement that offers minimal information (Hobson and Lee 1999, Dittrich *et al.* 1996). Close analysis of facial expression, gesture and action in everyday contexts shows that as adults we continue to rely on embodied interactive abilities to understand the intentions and actions of others and to accomplish interactive tasks (Lindblom 2007, Lindblom and Ziemke 2007).

[3] Cf. Baron-Cohen 1991 and 1995.

By the end of the first year of life, infants have a non-mentalizing, perceptually-based, embodied and pragmatic understanding of the intentions and dispositions of other persons. With the advent of joint attention (at around 9 months) and secondary intersubjectivity (at around 1 year) infants start to use context and enter into situations of participatory sense-making (De Jaegher and Di Paolo 2007). That is, infants begin to co-constitute the meaning of the world in their interactions with others. We start to understand the world through our interactions with others, and we gain a more nuanced understanding of others by situating their actions in contexts that are defined by both pragmatic tasks and cultural practices.

Meaning and emotional significance is co-constituted in the interaction – not in the private confines of one or the other's head. The analyses of social interactions in shared activities, in working together, in communicative practices, and so on, show that agents unconsciously coordinate their movements, gestures, and speech acts (Issartel *et al.* 2007, Kendon 1990, Lindblom 2007). In the contextualized practices of secondary intersubjectivity timing and emotional attunement continue to be important as we coordinate our perception-action sequences; our movements are coupled with changes in velocity, direction and intonation of the movements and utterances of the speaker.

The kind of embodied and contextualized interaction that we find in primary and secondary intersubjectivity is what I am calling 'strong interaction'. In strong interaction, our movements are often synchronized in resonance with others, following either in-phase or phase-delayed behaviour, and in rhythmic co-variation of gestures, facial or vocal expressions (Fuchs and De Jaegher 2009). This kind of intersubjective interaction involves coordination but does not imply perfect synchronization. Non-autistic infants from 3-months of age prefer slight modulations (time-delays) and imperfect contingency in responses (Gergely 2001). As De Jaegher (2008) suggests, continuous movements between synchronised, desynchronised and the states in-between, drive the process. Attunement, loss of attunement, and re-established attunement maintain both differentiation and connection.

A Closer Look at Interaction

I want to focus on two aspects of strong interaction: immersion and irreducibility. The first involves the idea that interaction is not something that we decide to enter into. Rather, it is, as the existentialists might say, something that we are *thrown* into, before anything like a decision is even possible. This is closely tied to the fact that interaction is primarily, that is, from the very beginning, embodied – a fact (or the facticity) of our physical nature, and specifically, of the kind of body that we have and the contingencies of our earliest existence. There is, in effect, no scientific mystery to this phenomenon, even if in everyday experience it seems a mystery in terms of why for the most part we cannot help but engage in it. The second aspect involves the idea that strong interaction is irreducible to the individuals involved.

I start with a question related to the first aspect, namely the question of the origin of interaction. Merleau-Ponty points to the bodily nature of interaction with his concept of intercorporeity (Merleau-Ponty 1962, p. 352). What I want to suggest is that intersubjective interaction ultimately derives from a more primary intercorporeal interaction.

We know the principle from neuroscience, *movement influences morphology* (Edelman 1992, Sheets-Johnstone 1998): brain development *results* from the system as a whole adapting to new levels of organization at more peripheral levels, rather than the neurological developments unfolding to 'allow' increasing proprioceptive capacities (Van der Meer and Van der Weel 1995). Consider the variety of developmental processes that follow this principle. For example, there is good evidence that both (1) a primitive proprioceptive registration of one's bodily movement, and (2) a differentiation between self and non-self develop prenatally (see Gallagher 1996). For example, proprioceptors in the muscles (muscle spindles) first appear at 9 weeks gestational age (Humphrey 1964); parts of the vestibular system develop as early as the fourth month of gestation (Jouen and Gapenne 1995); and cortical connections necessary for body-schematic proprioceptive processes are in place by 26 weeks gestational age. In addition, the differentiation between self and non-self in the later-term fetus is evidenced across a number of studies of fetal behavioral reaction to various stimuli. In response to auditory stimuli, as early as 24 weeks gestational age, fetal heart rate changes; and after 25 weeks, the fetus responds by blinking its eyes or moving its limbs. Cortical response to such stimuli has been demonstrated in

premature infants between 24-29 weeks gestational age (Fifer and Moon 1988). Differential responsiveness in the late-term fetus, signals a preference for some sounds (such as the mother's voice) rather than others (DeCasper and Spence 1986). Bright light directed on the lower abdomen of the mother in the third trimester can elicit fetal eye blinks (Emory and Toomey 1988), and fetal facial movements prompted by music or voice may be indicative of a similar differential awareness (Birnholz 1988). And we know that what Aristotle called the most basic sense, the tactile sense, develops early in the fetus, with cortical pathways intact by 20-24 weeks gestational age, with a differential registration between self-touching and being touched even earlier (Glass 2005).

Even before the development of full-fledged proprioceptive and tactile senses, however, the fetus is already moving. At twelve weeks gestational age, there is evidence of spontaneous and repetitious movements – e.g., movement of the hand to the mouth occurs multiple times an hour from this time (De Vries *et al.* 1982; Tajani and Ianniruberto 1990). At ten weeks gestational age fetuses display structured bodily movements which they develop through habituation (Krasnegor *et al.* 1998); for example, regular mouth opening and closing, swallowing, and movement in response to stimuli such as the mother's laugh or cough.

> The first movements to occur are sideward bendings of the head. [...] At 9-10 weeks [gestational] age complex and generalized movements occur. These are the so-called general movements [...] and the startles. Both include the whole body, but the general movements are slower and have a complex sequence of involved body parts, while the startle is a quick, phasic movement of all limbs and trunk and neck. (Prechtl 2001)

Two kinds of movement are involved here: early fetal movement, which is spontaneous and repetitive and starts out as a reflex that unfolds genetically (De Vries *et al.* 1982); and early fetal movement that appears regulated and practiced – i.e., non-reflex (Krasnegor *et al.* 1998) – and that starts out as a response to stimuli. Setting aside the question of which of these come first, we can say that at some point in early fetal motility responsive movement comes along.[4] The question is: To what is this movement a response? What is the

[4] I note here a recent study by Zoia *et al.* (2007) on intentional or directed movement in the fetus. Zoia *et al.* examined kinematic patterns of foetal movements showing that at 22 weeks hand to mouth and hand to head movement involved straighter and more accurately aimed trajectories with acceleration and deceleration phases consistent with target size and sensitivity. Thus, «by 22 weeks of gestation the movements seem to show the recognizable form of intentional actions, with kinematic

origination of this movement that helps to set the train of development in motion? The answer is that this kind of movement is a reaction to the mother's bodily movement – a kind of intercorporeal interaction.

> It is likely that these earliest regulated movements, which are prior to proprioceptive capacity, are a response within and to, the maternal body in *her* regulated and habituated, body schematic movement. [...] Add to physical movement the regular maternal heart beat, digestion, and breathing and we can see that the intrauterine world is not only a moving but quite rhythmic or regulated animate world. (Lymer 2010, p. 230)

This is not yet *intersubjective* interaction (the mother may not even know she's pregnant this early; and there is no claim that the fetus is an experiencing subject), but it is an *intercorporeal* interaction – a non-conscious motor coupling between mother and fetus driven toward and then driven by proprioception and touch. The point I want to make here is that whatever the moment of the awakening of consciousness – whether that is prenatal (at around 26 weeks gestational age) or later than that – and wherever we might locate the earliest aspect of self-awareness, this kind of intercorporeal interaction predates that, so that we find ourselves already immersed in interactive processes that prefigure the intersubjective ones found in primary intersubjectivity.

To this immersion I want to add that the primary and secondary intersubjective interactions that we find in infancy are more than capacities or mechanisms that belong to the individuals involved in interaction. They are not based simply on "first-order *mechanisms*" (Buckner *et al.* 2009) that we find in each individual, because they are not reducible to the sum of individual capacities (De Jaegher and Di Paolo 2007, De Jaegher *et al.* 2010). In the case of intersubjective interaction, $1 + 1 > 2$. This is what De Jaegher and Di Paolo (2007) mean by saying that interaction has some degree of autonomy. The interaction in intersubjective contexts goes beyond each participant; it results in something (the creation of meaning) that goes beyond what each individual qua individual can bring to the process – just as when two people dance the tango, something dynamic is created that neither one could create on their

patterns that depend on the goal of the action, suggesting a surprisingly advanced level of motor planning» (Zoia *et al.* 2007, p, 217). Also see Becher 2004 : «Purposive movement depends on brain maturation. This begins at about 18 weeks' gestation and progressively replaces reflex movements, which disappear by about 8 months after birth [...]»

own. Moreover, as we have just seen in regard to the origins of interaction, we are in the tango before we even know it.

So, not just in its origins, but as an ongoing process, interaction has a certain kind of irreducibility; it goes beyond the individual participants. In cases were one person is totally in control of the other person (if total control is ever possible), there is no interaction in this specific sense. The characteristics of immersion and irreducibility motivate the question about individual autonomy – self-agency. Merleau-Ponty talks about the infant getting caught up in the "whirlwind of language" – but prior to that the infant is caught up in the whirlwind of interaction – and even as adults we remain in that whirlwind. And within that whirlwind, does the irreducibility of interaction leave any room for self-agency or individual autonomy? If I, always already, even before birth, am caught up in a whirlwind of interaction, and that interaction always goes beyond me and my ultimate control, is there really any room for self-agency?

Self-Agency and the Narrative Self

There are current lively debates about self-agency and related concepts of freedom, free will, intention formation, and the sense of agency, with a variety of positions being staked out. From materialist and reductionist perspectives numerous theorists argue that self-agency is an illusion. They point to neuroscientific data (e.g., the Libet experiments that seem to show that the brain knows what we are going to do before we, as conscious individuals, do) or to the results of psychological experiments (e.g., Wegner 2002, Pockett 2006); or they suggest that if we do have free will, we need a subpersonal explanation of it that shows how it is generated in the individual brain (Spence 1996). Those who defend free will also often appeal to processes that are *in the head* (intention formation, reflective decision-making, or the phenomenological sense of agency, e.g., Pacherie 2008, Stephens and Graham 2000), or to mental causation, (Searle 1983, Lowe 1999). These approaches – whether they dismiss or defend the notion of free will – follow a traditional view that conceives of self-agency (or the lack of it) as a matter of individual subjectivity. Free will is either in the individual system or it is not. Even those theories that take social phenomena into account often use the individual as a measure of whether free will exists. For example, social

determinists argue that individual free will doesn't exist precisely because the individual is fully determined by our social interactions, cultural forces, etc.

In general, then, discussions of freedom/free will/self-agency focus on the individual – the question is framed that way, for example, if we ask about individual autonomy. I want to suggest, however, that in response to the question about self-agency, motivated by the account of strong interaction, we can conceive of self-agency in different terms by conceiving of the agent as something other than an individual who either has or does not have free will. If we view the agent as someone who emerges from intercorporeal interactions, and develops in social interactions with others, then we have a good model for speaking about self-agency in a system that is not reducible to a simple individual. On this model, self-agency – and a proper sense of freedom (which comes along with a proper sense of responsibility) – can be found only in the context of social interaction, where our intentions are formed in or out of our interactions with others.

Clearly, we learn to act from watching and interacting with others as they act in the world. We learn our own action-possibilities from others. Through our interactions with others we generate shared intentions and form our own intentions out of the same fabric. In this context, how can we explain self-agency?

It is at this point that I want to point out the importance of that aspect of interaction theory that involves communicative and narrative competencies (Gallagher and Hutto 2008). Beyond the processes of primary and secondary intersubjectivity, communicative and narrative practices allow for a certain volitional space to open up – the possibility of taking a critical perspective on ourselves. Narrative allows us to reflectively locate our interactions in what Bruner calls the 'landscape of action' and 'the landscape of consciousness' (Bruner 1986). That is, through narrative, we can reflect on our actions and interactions, and on what our motives for such actions might be.

In this process, and specifically in autobiographical (or self-) narrative, narrative distance, a concept that goes back to Aristotle's *Poetics*, is established between the self who narrates and the self who is narrated. This distance allows for the possibility of what Harry Frankfurt (1971) calls second-order volitions – that is, volitions in which we consider or evaluate our own first-order action volitions. On Frankfurt's view, this capacity for second-order volitions, or what Charles Taylor (1989) calls the possibility for a *strong evaluation* of our own desires, is essential for moral personhood. From an

interactionist perspective, this is possible only as a result of social interaction processes, in social settings where we act and interact, and where we exercise our communicative and narrative practices.

What we call autonomy, then, is not constituted in just an internal intra-individual negotiation made by an agent with respect to herself, but is inextricably interwoven into and out of our relationships with others. In this regard, self-agency becomes a matter of degree rather than an all or nothing issue. Some people arrange their lives with others, or find themselves in such arrangements, so that they have a high degree of freedom – a greater range of possibilities than others who find themselves in social relationships, or cultures, or institutions where they are prevented from acting freely.

There is nothing new in this thought: our social interactions and arrangements are such that they either promote freedom or prevent it. Whatever self-agency is, it's weaved out of this fabric of interaction; not a characteristic of the individual; but a characteristic of a set of relationships. In some of my interactions I am freer than in others. Some arrangements support self-agency, and some do not. I could say, without contradiction, that I am free and I am not free – but only in the sense that my self-agency is constituted in my different relations *differently*.

It's also the case that certain interactions can make one participant free and the other a slave. So the question that derives directly from conceiving of intersubjective interaction as a primary force in shaping our cognitive, emotional, and social life is not the *metaphysical* question: Do I as an individual have free will? It is rather the *political* question: who is free (or more free) and who is not, and why? The political question is a pragmatic and critical one, because we can ask why, and motivate change.

Conclusion

With regard to discussions of social cognition, shifting away from theory-of-mind approaches, such as theory theory and simulation theory, and taking up the interaction theory and the emphasis on intersubjective interaction also involves shifting away from conceptions of self-agency that are reducible to neural or mental or strictly individual processes framed in terms of mental causation. I've suggested that self-agency is a matter of degree and that it can be won or lost in the varying contexts of interaction – contexts from which I can

distance myself through a narrative process that allows for strong evaluation. Accordingly, self-agency and related phenomena such as free will and intention formation – these are not things that pertain strictly to an individual; rather, they are constituted in interaction and in communicative and narrative practices.

REFERENCES

Baird, J. A., & Baldwin, D. A. (2001). Making sense of human behavior: Action parsing and intentional inference. In B. F. Malle, L. J. Moses & D. A. Baldwin (Eds.), *Intentions and Intentionality: Foundations of Social Cognition*, (pp. 193-206). Cambridge, MA: MIT Press.

Baldwin, D. A., & Baird, J. A. (2001). Discerning intentions in dynamic human action. *Trends in Cognitive Science, 5*(4), 171-178.

Banks, W., Pockett, S., & Gallagher, S. (2006). *Does Consciousness Cause Behavior? An Investigation of the Nature of Volition*, (pp. 109-124). Cambridge, MA: MIT Press.

Baron-Cohen, S. (1991). Precursors to a theory of mind: Understanding attention in others. In A. Whiten (Ed.), *Natural Theories of Mind: Evolution, Development and Simulation of Everyday Mindreading*, (pp. 233-251). Cambridge, MA: Basil Blackwell.

Baron-Cohen, S. (1995). *Mindblindness: An Essay on Autism and Theory of Mind*. Cambridge, MA: MIT Press.

Becher, J.-C. (2004). Insights into early fetal development. *Behind the Medical Headlines*. (Royal College of Physicians of Edinburgh and Royal College of Physicians and Surgeons of Glasgow October 2004).

Birnholz, J. C. (1988). On observing the human fetus. In W. P. Smotherman & S. R. Robinson (Eds.), *Behavior of the Fetus*, (pp. 47-60). Caldwell, NJ: Telford Press.

Bruner, J. (1986). *Actual Minds, Possible Worlds*. Cambridge, MA: Harvard University Press.

Buckner, C., Shriver, A., Crowley, S., & Allen, C. (2009). How 'weak' mindreaders inherited the earth. *Behavioral and Brain Sciences, 32*(2), 140-141.

Catmur, C., Walsh, V., & Heyes, C. (2007). Sensorimotor learning configures the human mirror system. *Current Biology, 17*(17), 1527-1531.

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences, 13*(4), 148-153.

Currie, G. (2008). Some ways of understanding people. *Philosophical Explorations, 11*(3), 211-218.

De Casper, A. J., & Spence, M. J. (1986). Prenatal maternal speech influences newborns' perception of speech sounds. *Infant Behavior and Development, 9*, 133-150.

De Jaegher, H. (2008). Social understanding through direct perception? Yes, by interacting. *Consciousness and Cognition, 18*(2), 535-542.

De Jaegher, H., & Di Paolo, E. (2007). Participatory sense-making: An enactive approach to social cognition. *Phenomenology and the Cognitive Sciences, 6*(4), 485-507.

De Jaegher, H., Di Paolo, E., & Gallagher, S. (2010). Does social interaction constitute social cognition? *Trends in Cognitive Sciences, 14*(10): 441-447.

De Vries, J. I. P., Visser, G. H. A., & Prechtl, H. F. R. (1982). The emergence of fetal behaviour: I. Qualitative aspects. *Early Human Development, 7*, 301-322.

Dinstein, I., Thomas, C., Behrmann, M., & Heeger, D. J. (2008). A mirror up to nature. *Current Biology, 18*(1), R13-R18.

Dittrich, W. H., Troscianko, T., Lea, S. E. G., & Morgan, D. (1996). Perception of emotion from dynamic point-light displays represented in dance. *Perception, 25*, 727-738.

Edelman, G. (1992). *Bright Air, Brilliant Fire*. New York: Basic Books.

Emory, E. K., & Toomey, K. A. (1988). Environmental stimulation and human fetal responsivity in late pregnancy. In W. P. Smotherman & S.

R. Robinson (Eds.), *Behavior of the Fetus*, (pp. 141-161). Caldwell, NJ: Telford Press.

Fifer, W. P., & Moon, C. (1988). Auditory experience in the fetus. In W. P. Smotherman & S. R. Robinson (Eds.), *Behavior of the Fetus*, (pp. 175-188). Caldwell, NJ: Telford Press.

Frankfurt, H. (1971). Freedom of the will and the concept of a person. *Journal of Philosophy, 68*(1), 5-20.

Fuchs, T., & De Jaegher, H. (2009). Enactive intersubjectivity: Participatory sense-making and mutual incorporation. *Phenomenology and the Cognitive Sciences, 8*(4), 465-486.

Gallagher, S. (1996). The moral significance of primitive self-consciousness. *Ethics, 107*(1), 129-140

Gallagher, S. (2005). *How the Body Shapes the Mind.* Oxford: Oxford University Press/Clarendon Press

Gallagher, S. (2007). Simulation trouble. *Social Neuroscience, 2*(3-4), 353-365.

Gallagher, S. (in press). Narrative competency and the massive hermeneutical background. In P. Fairfield (Ed.), *Hermeneutics in Education*. New York: Continuum.

Gallagher, S., & Hutto, D. (2008). Understanding others through primary interaction and narrative practice. In J. Zlatev, T. Racine, C. Sinha & E. Itkonen (Eds.), *The Shared Mind: Perspectives on Intersubjectivity*, (pp. 17-38). Amsterdam: John Benjamins.

Gergely, G. (2001). The obscure object of desire: 'Nearly, but clearly not, like me': Contingency preference in normal children versus children with autism. *Bulletin of the Menninger Clinic, 65*(3), 411-426.

Glass, P. (2005). The vulnerable neonate and the neonatal intensive care environment. In G. B. Avery, M. G. MacDonald, M. M. K. Seshia & M. D. Mullett (Eds.), *Avery's Neonatology: Pathophysiology & Management of the Newborn*, (pp. 111-129). Philadelphia: Lippencott, Williams and Wilkins.

Goldman, A. I. (1995). Desire, intention and the simulation theory. In B. F. Malle, L. J. Moses & D. A. Baldwin (Eds.), *Intentions and Intentionality: Foundations of Social Cognition*, (pp. 207-224). Cambridge, MA: MIT Press.

Goldman, A. I. (2006). *Simulating Minds: The Philosophy, Psychology, and Neuroscience of Mindreading*. New York: Oxford University Press.

Gopnik, A., & Meltzoff, A. N. (1997). *Words, Thoughts, and Theories.* Cambridge, MA: MIT Press.

Gordon, R. M. (1986). Folk psychology as simulation. *Mind and Language, 1*(2), 158-171.

Gordon, R. M. (1995). Developing commonsense psychology: Experimental data and philosophical data. Paper presented at the APA Eastern Division Symposium on Children's Theory of Mind, 12/27/95. (http://www.umsl.edu/~philo/Mind_Seminar/New%20Pages/papers/Gordon/apakids9.htm)

Heal, J. (1986). Replication and functionalism. In J. Butterfield (Ed.), *Language, Mind, and Logic*, (pp. 45-59). Cambridge: Cambridge University Press.

Heal, J. (1998a). Co-cognition and off-line simulation: Two ways of understanding the simulation approach. *Mind and Language, 13*, 477-498.

Heal, J. (1998b). Understanding other minds from the inside. In A. O'Hear (Ed.), *Current Issues in Philosophy of Mind*. Cambridge: Cambridge University Press.

Hobson, P. (1993). The emotional origins of social understanding. *Philosophical Psychology, 6*(3), 227-249.

Hobson, P. (2002). *The Cradle of Thought*. London: Macmillan.

Hobson, P., & Lee, A. (1999). Imitation and identification in autism. *Journal of Child Psychology and Psychiatry, 40*(4), 649-659.

Humphrey, T. (1964). Some correlations between the appearance of human fetal reflexes and the development of the nervous system. *Progress in Brain Research, 4*, 93-135.

Issartel, J., Marin, L., & Cadopi, M. (2007). Unintended interpersonal coordination: 'Can we march to the beat of our own drum'? *Neuroscience Letters, 411*(3), 174-179.

Jouen, F., & Gapenne, O. (1995). Interactions between the vestibular and visual systems in the neonate. In P. Rochat (Ed.), *The Self in Infancy: Theory and Research*, (pp. 277-301). Amsterdam: Elsevier Science.

Kendon, A. (1990). *Conducting Interaction: Patterns of Behavior in Focused Encounters*. Cambridge: Cambridge University Press.

Keysers, C., & Gazzola, V.(2006). Towards a unifying neural theory of social cognition. In S. Anders, G. Ende, M. Junghofer & J. Kissler (Eds.), *Understanding Emotions*, (pp. 379-402). Amsterdam: Elsevier.

Krasnegor, N. A., Fifer, W., Maulik, D., McNellis, D., Romero, R., & Smotherman, W. (1998). Fetal behavioral development: Measurement of habituation, state transitions, and movement to assess fetal well being and to predict outcome. *The Journal Maternal-Foetal Investigation, 8*, 51–57.

Leslie, A. (1991). The theory of mind impairment in autism: Evidence for a modular mechanism of development? In A. Whiten (Ed.), *Natural Theories of Mind: Evolution, Development and Simulation of Everyday Mindreading*, (pp. 63-78). Oxford: Blackwell.

Lindblom, J. (2007). Minding the body: Interacting socially through embodied action. *Linköping Studies in Science and Technology*. Dissertation No. 1112.

Lindblom, J., & Ziemke, T. (2007). Embodiment and social interaction: Implications for cognitive science. In T. Ziemke, J. Zlatev & R. Frank (Eds.), *Body, Language, and Mind: Embodiment*, (pp.129-162). Berlin: Mouton de Gruyter.

Lowe, E. J. (1999). Self, agency and mental causation. *Journal of Consciousness Studies, 6*(8-9), 225-239.

Lymer, J. (2010). The phenomenology of the maternal-foetal bond. Ph.D. Dissertation. Wollongong University, Australia.

Maurer, D., & Barrera, M. E. (1981). Infants' perception of natural and distorted arrangements of a schematic face. *Child Development, 52*(1), 196-202.

Meltzoff, A., & Moore, M. K. (1977). Imitation of facial and manual gestures by human neonates. *Science, 198*, 75-78.

Meltzoff, A., & Moore, M. K. (1994). Imitation, memory, and the representation of persons. *Infant Behavior and Development, 17,* 83-99.

Merleau-Ponty, M. (1962). *Phenomenology of Perception.* (Tr. by C. Smith). London: Routledge and Kegan Paul.

Muir, D., & Lee, K. (2003). The still-face effect: Methodological issues and new applications. *Infancy, 4*(4), 483-491.

Murray, L., & Trevarthen, C. (1985). Emotional regulation of interactions between 2-month-olds and their mothers. In T.M. Field & N. A. Fox (Eds.), *Social Perception in Infants*, (pp. 177-197). Norwood, NJ: Ablex.

Nadel, J., Croué, S., Mattlinger, M.-J, Canet, P., Hudelot, C., Lécuyer, C, & Martini, M. (1999). Expectancies for social contingency in 2-month-olds. *Developmental Science, 2*, 164-173.

Nichols, S., & Stich, S. P. (2003). *Mindreading: An Integrated Account of Pretence, Self-Awareness, and Understanding Other Minds.* Oxford: Clarendon Press.

Pacherie, E. (2008). The phenomenology of action: A conceptual framework. *Cognition, 107*(1), 179–217.

Prechtl, H. (2001). Prenatal and early postnatal development of human motor behavior. In A. F. Kalverboer & A. Gramsbergen (Eds.), *Handbook of brain and behaviour in human development*, (pp. 415-418). Dordrecht: Kluwer Academic Publishers.

Rizzolatti, G., Fogassi, L., & Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Review Neuroscience, 2*, 661-670.

Searle, J. (1983). *Intentionality: An Essay in the Philosophy of Mind*. Cambridge: Cambridge University Press.

Senju, A., Johnson, M. H., & Csibra, G. (2006). The development and neural basis of referential gaze perception. *Social Neuroscience, 1*(3-4), 220-234.

Sheets-Johnston, M. (1998). Consciousness: A Natural History. *Journal of Consciousness Studies, 5*(3), 260-294.

Spence, S. A. (1996). Free will in the light of neuropsychiatry. *Philosophy, Psychiatry, and Psychology, 3*(2), 75-90.

Stephens, G. L., & Graham, G. (2000). *When Self-Consciousness Breaks: Alien Voices and Inserted Thoughts*. Cambridge, MA: MIT Press.

Stormark, K. M., & Braarud, H. C. (2004). Infants' sensitivity to social contingency: A "double video" study of face-to-face communication between 2- and 4-month-olds and their mothers. *Infant Behavior and Development, 27*, 195-203.

Tajani, E., & Ianniruberto, A. (1990). The uncovering of fetal competence. In M. Papini, A. Pasquinelli & E. A. Gidoni (Eds.), *Development Handicap and Rehabilitation: Practice and Theory*. Amsterdam: Elsevier Science Publishers.

Taylor, C. (1989). *Sources of the Self*. Cambridge, MA: Harvard University Press.

Trevarthen, C. B. (1979). Communication and cooperation in early infancy: A description of primary intersubjectivity. In M. Bullowa (Ed.), *Before Speech*, (pp. 321-348). Cambridge: Cambridge University Press.

Trevarthen, C., & Hubley, P. (1978). Secondary intersubjectivity: Confidence, confiding and acts of meaning in the first year. In A. Lock (Ed.), *Action, Gesture and Symbol: The Emergence of Language*, (pp. 183-229). London: Academic Press.

Tronick, E. Z. (2007). *The Neurobehavioral and Social Emotional Development of Infants and Children*. New York: Norton.

Tronick, E., Als, H., Adamson, L.,Wise, S., & Brazelton, T. B. (1978). The infants' response to entrapment between contradictory messages in

face-to-face interactions. *Journal of the American Academy of Child Psychiatry, 17*, 1-13.

Van der Meer, A. L., Van der Weel, F. R., & Lee, D. N. (1995). The functional significance of arm movements in neonates. *Science, 267*, 693-695.

Walker, A. S. (1982). Intermodal perception of expressive behaviors by human infants. *Journal of Experimental Child Psychology, 33*, 514-535.

Wegner, D. (2002). *The Illusion of Conscious Will.* Cambridge, MA: MIT Press.

Woodward, A. L., & Sommerville, J. A. (2000). Twelve-month-old infants interpret action in context. *Psychological Science, 11*, 73–77.

Zoia, S., Blason, L., D'Ottavio, G., Bulgheroni, M., Pezzetta, E., Scabar, A., Castiello, U. (2007). Evidence of early development of action planning in the human foetus: A kinematic study. *Experimental Brain Research, 176*(2), 217-226.