Canadian Journal of Philosophy
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/rcjp20

Redrawing the Map and Resetting the Time: Phenomenology and the Cognitive Sciences

SHAUN GALLAGHER & FRANCISCO J. VARELA

Published online: 01 Jul 2013.

To cite this article: SHAUN GALLAGHER & FRANCISCO J. VARELA (2003) Redrawing the Map and Resetting the Time: Phenomenology and the Cognitive Sciences, Canadian Journal of Philosophy, 33:sup1, 93-132

To link to this article: http://dx.doi.org/10.1080/00455091.2003.10717596

CANADIAN JOURNAL OF PHILOSOPHY
Supplementary Volume 29

Redrawing the Map and Resetting the Time: Phenomenology and the Cognitive Sciences[1]

SHAUN GALLAGHER AND FRANCISCO J. VARELA[2]

In recent years there has been some hard-won but still limited agreement that phenomenology can be of central and positive importance to the cognitive sciences. This realization comes in the wake of dismissive gestures made by philosophers of mind who mistakenly associate phenomenological method with untrained psychological introspection (e.g., Dennett 1991). For very different reasons, resistance is also found on the phenomenological side of this issue. There are many thinkers well versed in the Husserlian tradition who are not willing to consider the validity of a naturalistic science of mind. For them cognitive science is too computational or too reductionistic to be seriously considered as capable of explaining experience or consciousness.[3] In some cases, when phenomenologists have seriously engaged the project of the cognitive sciences, rather than pursing a positive rapprochement with this project, they have been satisfied in drawing critical lines that identify its limitations.

---

1 Originally published in electronic form in *The Reach of Reflection: The Future of Phenomenology*, ed. S. Crowell, L. Embree and S. J. Julian (17–45). Electron Press (at http://www.electronpress.com/reach.asp).

2 Francisco Varela died on May 28, 2001.

3 Signals of this kind of suspicion have been sent recently by Paul Ricoeur in his conversation with Jean-Pierre Changeux (Changeux and Ricoeur 2000). Ricoeur suggests that phenomenology stands opposed to the cognitive sciences.

On the one hand, such negative attitudes are understandable from the perspective of the Husserlian rejection of naturalism, or from strong emphasis on the transcendental current in phenomenology. On the other hand, it is possible to challenge these attitudes from perspectives similar to the one taken by Merleau-Ponty (1962, 1964), who integrated phenomenological analyses with considerations drawn from the empirical sciences of psychology and neurology long before cognitive science was defined as such. In constructing a cross-disciplinary tradition like the one Merleau-Ponty inspires, thinkers in the relevant disciplines have to wrestle with a variety of issues, including the issue of naturalism. In this respect, however, natural scientists, more readily than phenomenologists, have come to acknowledge that phenomenology is directly relevant for a natural scientific understanding of cognition (e.g., Varela 1996; Varela et al. 1991). Even the hardest of the hard scientists have made recent peace offerings to phenomenology. For example, the neuroscientist Jean-Pierre Changeux, in his conversation with Paul Ricoeur, declares that his purpose "is not to go to war against phenomenology; to the contrary, [he wants] to see what constructive contribution it can make to our knowledge of the psyche, acting in concert with the neurosciences" (Changeux and Ricoeur 2000, 85). And Alain Berthoz, a neuroscientist who studies motor and perceptual systems, does not hesitate to invoke Husserl's analysis of time-consciousness in his explication of anticipatory aspects of motor control (Berthoz 2000, 16).[4]

In this chapter we explore the various ways in which phenomenology and the cognitive sciences can come together in a positive and productive exchange. In the first part, after some brief remarks about the nature of the cognitive sciences and the problem of naturalization, we begin by mapping out several issues that would benefit from this exchange. In the second part of the chapter we ask if this cross-disciplinary approach can address one of the most basic problems defined by

4 Both Changeux and Berthoz are at the College de France, where Merleau-Ponty has seemingly had a lasting influence.

Husserlian phenomenology, the problem of time-consciousness, and whether Husserl's analysis of this theme has anything to contribute to the cognitive sciences.

Part I: Defining the Issues

A Different Cognitive Science and a Different Phenomenology

If one begins by thinking of cognitive science as it was first formulated in opposition to behaviourism, in terms of computational analysis and information processing, it is difficult to see how phenomenology might participate in the "Cognitive Revolution." On this formulation, the scientific study of cognition is a study of how the subpersonal, non-phenomenological mind manipulates discrete symbols according to a set of syntactical procedures, and how this might be cashed out in neurological terms. This, however, is no longer the current view of cognitive science. Faced with a variety of problems implicit in this view, the cognitive revolution took a different turn in the late 1980s. This corresponded to a new emphasis on neuroscience, and connectionism, which challenged the prevailing computational orthodoxy by introducing an approach based on nonlinear dynamical systems (see, e.g., Port and van Gelder 1995). With this formulation there was a shift away from an emphasis on reductionism to an emphasis on the notions of emergence and self-organization. The question was how higher-level personal structures emerged from lower-level subpersonal, self-organizing processes.

This turn in the fortunes of cognitive science also motivated a new interest in consciousness. It is nothing short of ironic that just when many phenomenologists were trading in their volumes of Husserl and Sartre for the texts of poststructural analysis, and thus abandoning the very notion of consciousness, philosophers of mind, who had started their work on ground circumscribed by Ryle's behaviouristic denial of consciousness, were beginning to explore the territory left behind by the phenomenologists. And while this continental drift away from consciousness was motivating a remapping of philosophical interests,

theorists in the cognitive sciences were working their way toward a necessary rendezvous with phenomenology.

The current situation in the cognitive sciences is characterized by a growing interest in the ecological-embodied-enactive approach (Bermudez et al. 1995; Clark 1997; Varela et al. 1991). This approach takes up the connectionist emphasis on dynamical mechanisms and self-organizing emergence, but it further insists that cognition is best characterized as belonging to embodied, situated agents – agents who are *in-the-world*. On this understanding of the cognitive sciences, just as neuroscientists and neuropsychologists work together with researchers in artificial intelligence and robotics, so also phenomenologists and philosophers of mind work together with the empirical scientists in order to develop a fuller and more holistic view of cognitive life – a life that is not just the life of the mind, but of an embodied, ecologically situated, enactive agent.

This recent redefinition of the cognitive sciences, if it is to include a place for phenomenology, requires that we also conceive of phenomenology in a different way. Or at least we need to see that there is a section of the phenomenological map that can be redrawn along lines that reach across the theoretical divides that separate phenomenology from the sciences. One way to think of this is to think of *naturalizing phenomenology*. For many phenomenologists, this will seem self-contradictory, an antilogy. Phenomenology just is, by definition, non-naturalistic. For many others, the difficult question is how it might be accomplished without it losing the specificity of phenomenology. Everything, however, depends on what one means by naturalization. There is no question of considering here all possible proposals in this regard (see Roy et al. 1999, for a more detailed analysis). But consider two among many.

(1) The subjective data developed in phenomenology should be made objective, and thus amenable to scientific analysis. This suggestion is similar to Nagel's (1974) idea of an "objective phenomenology" that would allow for a level of abstraction from the particularism of individual reports, or to Dennett's (1991) idea of a "heterophenomenology" that would treat phenomenological reports as part of the objective data of science.

(2) Naturalization in the minimal sense means "not being committed to a dualistic kind of ontology" (Roy et al. 1999, 19). This includes the idea that phenomenology has to be explanatory and not just descriptive. Phenomenology would help to resolve the "explanatory gap," and contribute to an explanation of how brain and bodily processes can give rise to phenomenological properties that are not non-physical properties.

The first proposal is problematic in several ways. It equates phenomenology with folk psychology and understands phenomenological data to mean anything that a subject happens to report. In effect, because there is no concern about phenomenological method, this kind of approach, whether it follows a reductionist strategy or pursues the intentional stance, fails to take phenomenology seriously. Naturalization means, in this case, getting rid of phenomenology. Furthermore, if in such naturalizing strategies one does not employ phenomenological method, the very objectivity that is sought will be seriously compromised. For example, in translating a subject's first-person experience into third-person data, it would not do for the scientist to rely on her own subjective experience as an interpretational guide, since this would simply lead to a pollution of the subject's first-person data with the scientist's first-person data. A scientist would have to resort to objectively formalized meanings established within the framework of behavioural science in order to properly interpret the subject's reports. In that case, however, one needs to ask where these formalized meanings (generalizations and abstractions) originate. One comes quickly to the realization that at some point a controlled form of phenomenological experience is required to justify the formalized meanings used as interpretational guides. In effect, an objective interpretational framework depends upon a reflective, methodically guided phenomenological analysis, without which the procedure may simply impose the results of previous uncontrolled and anonymous phenomenological exercises (see Gallagher 1997).

The second proposal, as it is worked out by Roy et al. (1999), requires a phenomenological practice guided by method. What allows Husserlian phenomenology to escape from a naturalistic framework (the natural

attitude) is a change of attitude achieved through a methodical practice (the phenomenological reduction). To move in the opposite direction, that is, to bring phenomenology to bear on the naturalistic enterprise of the cognitive sciences, involves another change of attitude. This does not mean abandoning phenomenological methods, but taking what we learn about first-person experience within the phenomenological attitude and using it in the context of naturalistic explanation. Although Husserl defined phenomenology as a non-naturalistic discipline, the idea that the results of his transcendental science might inform the natural sciences is not inconsistent with his own intent. He suggested, quite clearly, that "every analysis or theory of transcendental phenomenology – including ... the theory of the transcendental constitution of an objective world – can be developed in the natural realm, by giving up the transcendental attitude" (1970, §57).

Roy et al. (1999) develop one possible program that would lead to a naturalization of phenomenology in this sense. The issue is how to transpose the results from phenomenological, methodical analysis to the natural context, without stopping at statements of mere correlation. They propose a re-categorization of phenomena at a level of abstraction sufficient to allow for the recognition of common properties between phenomenological data and objective data developed in the sciences. The task would involve moving to a level of explanation that would be abstract enough to embrace all of the data. One possibility would involve a mathematical interpretation, a transformation of concepts into algorithms similar to transformations of this kind found in the physical sciences. In effect, if one could develop a formal language to express phenomenological findings, using perhaps a phenomenological notation of the sort suggested by Marbach (1993), the task would be to integrate it with a similar formal expression of physical processes. The appeal to mathematics is an appeal to formalized and intersubjectively verifiable meanings within a common language.

Another possible route to naturalization is to view phenomenology and the cognitive sciences as mutually constraining (Varela 1996; Gallagher 1997). For example, there are explanations of voluntary action that are developed at the level of cognitive mechanisms. A cognitive mechanism is usually thought of as a heuristic that will eventually be cashed out in terms of neurological processes. Can such an explanation

succeed if it fails to make sense out of the rich phenomenological experience that accompanies voluntary action? This does not mean that cognitive explanations have to identify physical processes that are isomorphic with phenomenological data; but at a minimum, if a cognitive explanation implies or requires a phenomenological correlation that is unlikely or impossible, some negotiation has to take place between the two levels of description. And this does not necessarily mean that phenomenology will either win or lose. It is quite possible that the mutual constraint situation will lead to a productive mutual enlightenment, where progress in the cognitive sciences will motivate a more finely detailed phenomenological description developed under the regime of phenomenological reduction, and a more detailed phenomenology will contribute to defining an empirical research program. A specific example concerning preparation strategies and visual perception has been provided recently (Lutz et al. 2002). In this study phenomenological reduction was used as the basis for selecting electrical events. This, in turn, validated the various phenomenological clusters of reports provided by the subjects.

Although we have invoked phenomenological reduction as crucial, it is also true that there is little agreement and even less explicit development of the *pragmatics* that should define its role in any non-reductionist naturalization project. In other words, explicit accounts of how to carry out the gestures involved in the *epoché* have to be developed, and how these are to be linked to intersubjective validation. This is a central issue, but it will not concern us here (for recent attempts towards a disciplined phenomenological pragmatics see Depraz et al. 2000; 2003).

Back to the Issues Themselves

Consider some of the problems that concern phenomenologists. How do I perceive space? How is perception different from memory, or dream-consciousness, or fantasy? When I remember or imagine something, does my thinking consist of images, or is my thinking more sentential? Does consciousness have a formal structure independently of its contents? In normal voluntary movement, to what extent, and in what sense

modalities, am I aware of my body? As I move through my immediate environment, of what am I aware? How do I understand what another person is thinking?

Although philosophers of mind work in a very different tradition, they are interested in the same kinds of problems. Not unlike Husserl, many of them will attempt to resolve these issues by appealing directly to experience. Others, not unlike Merleau-Ponty, will consider these issues in the light of what empirical studies have to say about them. Indeed, it is equally the case that certain empirical scientists are interested in exactly the same issues. Moreover, all of these groups share a common goal. They want to attend to the issues themselves; they want to understand the basics of human experience. What is it that keeps us from searching together? Perhaps the kinds of answers that we seek are different. Nonetheless, it seems clear that a more complete understanding of these things can be developed if we can see how the various answers line up with each other.

It will be impossible to provide here a complete inventory of all the many issues that are of common concern among phenomenologists, philosophers of mind, and cognitive scientists. But we can suggest, in regard to a few of the issues, how phenomenology can contribute to and learn from the other disciplines. To that end, we will discuss two sets of issues that are central to the concerns of phenomenologists and cognitive scientists: embodied self-consciousness and intersubjectivity.

Embodied Self-Consciousness

Under the title of self-consciousness there are a large number of specific problems. Even if we focus on a minimal sense of embodied self-consciousness, the issue is complex. Distinctions that are familiar to phenomenologists, such as those between objective body (*Körper*) and lived body (*Leib*), are not always acknowledged in the cognitive sciences. Yet very detailed discussions concerning the nature of proprioception and body schemas, found in the cognitive sciences, can enrich the phenomenological distinction. Clarifications can be introduced on both sides. For example, a distinction between body image and body schema, which is often lost within phenomenological discussions, can

be made clear by considering pathological cases involving neurological damage and loss of proprioception. Such cases can motivate a more detailed phenomenological analysis which can, in turn, contribute to further clarifications of the empirical cases (e.g., Gallagher and Cole 1995). The outcome of such clarifications informs philosophical discussions of primary (non-conceptual, pre-reflective) self-consciousness as it develops in the bodily experience of early infancy. Such themes were important to Merleau-Ponty (1962, 1964) and continue to be studied today by both philosophers and developmental psychologists (Gallagher and Meltzoff 1996).

Other related distinctions call for phenomenological clarification. For example, discussions in philosophy of mind often focus on the concept of the sense of ownership for movement, action, and thought (see, for example, Campbell 1999a). One finds, however, that the concept of ownership is complex. It involves distinguishing between senses of ownership for one's own body, for movement, and for action. Action also involves the notion of a sense of agency, which can be distinguished from the sense of ownership. Yet one does not find this distinction carefully made in the philosophical discussions. A phenomenological analysis of the difference between active and passive movement helps to make the distinction.

One would begin this analysis by bracketing all scientific theories about movement and motor control and by attending directly to one's own experience. If we define the *sense of agency* as the sense that I am the one who is causing or generating an action, and the *sense of ownership* as the sense that I am the one who is undergoing an experience, then in the phenomenology of voluntary or willed action, these two senses seem to be indistinguishable. When I intentionally reach for a cup and grasp it, I know this to be my own action. Agency coincides with ownership. This coincidence may be what leads philosophers to think of ownership of action in terms of agency: that the owner of an action is the person who is, in a particular way, causally involved in the production of that action. In the case of *involuntary* action, however, it is quite possible to distinguish, phenomenologically, between sense of agency and sense of ownership. I may have a sense that I am the one who is moving or is being moved, and thereby acknowledge ownership of the movement. I can self-ascribe it as *my* movement. At the same time, I may not have

a sense of causing or controlling the movement, that is, no sense of agency. The agent of the movement is someone else – the person who pushed me from behind, the physician who is manipulating my arm in a medical examination, etc. My claim of ownership (my self-ascription that I am the one who is undergoing such experiences) is perfectly consistent with my lack of a sense of agency (Gallagher 2000).

Working out a simple phenomenological distinction, however, cannot be the end of the story. If the distinction is going to do anything for philosophers in their analysis of action, or for scientists in their analysis of motor control, one needs to carry this distinction back into the empirical discussions. For example, can one find empirical studies that confirm the distinction? If one can, it provides empirical confirmation of the phenomenological analysis, and at the same time, it offers a clear distinction to guide further scientific research. In the scientific literature one can indeed find evidence that supports the phenomenological distinction. Experiments involving certain pathological cases demonstrate a clear dissociation between two different systems of motor control:

(1) a sensory-feedback mechanism that compares intended movement with actual movement, by means of visual and proprioceptive feedback; and
(2) a "forward," pre-action mechanism that compares motor intention with motor commands (see, e.g., Fourneret and Jeannerod 1998; Frith and Done 1988).

In some pathological cases the failure of the forward mechanism corresponds to a lack of a sense of agency. For example, a schizophrenic subject who suffers delusions of control complains that *his* hand is moving (that is, he has a sense of ownership for the movement) but that *he* is not moving it (that is, no sense of agency). Such subjects in experimental circumstances are able to control their movement through sensory-feedback but are unable to control it through the quicker forward mechanism (Frith and Done 1988).

What seems to be a correlation between the phenomenological distinction (sense of agency and sense of ownership for movement) and

the neurological distinction (forward control mechanism and sensory-feedback mechanism) calls for further research (see, for example, de Vignemont 2000; Franck et al. 2001). If the correlation holds up, not only will it suggest a neurological basis for these two aspects of bodily self-consciousness, but it will also provide a scientifically justified distinction that will clarify a variety of philosophical discussions that are in great need of clarification.

Knowledge of Others

An equally complex issue has a long history in the phenomenological tradition. We find the question of intersubjectivity central to the thoughts of Scheler, Husserl, Heidegger, Sartre, Merleau-Ponty, Ricoeur, and Levinas, to name a few of the thinkers who have wrestled with this problem. Again, there is an equally rich discussion ongoing in the philosophy of mind and the cognitive sciences. Moreover, there are some interesting parallels to be found in the two discussions. For example, criticisms of the notion of analogical inference – that is, that I understand the other as a mind only by an analogy I draw between her bodily behaviour and my own – have been outlined in both traditions. Both traditions are also marked by competing theories and the lack of consensus. On the phenomenological side, there are proponents of empathy and those who go beyond empathy (see Zahavi 2001, for a good review). On the cognitive side, there are those who favour "theory theory" (e.g., Baron-Cohen 1995; Leslie 2000), and those who champion the simulation approach (e.g., Gordon 1986, 1995; Goldman 1989). Rarely, however, does one find any crossover between the insights developed in the phenomenological discussions and the theories that make up the cognitive approach. Ironically, it is as if one discussion of knowing the other does not know about the other discussion (but see the essays in Thompson 2001 for an excellent exception to this problem).

Many of the claims made on the cognitive side, under the heading of "theory of mind," are based on neurological studies and experiments performed in the context of developmental psychology. The interpretation of these scientific studies, however, could benefit greatly from the guidance of phenomenology. Many of the assumptions made

in the cognitive interpretations do not hold up to phenomenological analysis. For example, if we are to accept the cognitive explanations (whether they are provided by theory theorists or simulation theorists), then we would have to accept the idea that our primary interactions with others involve attempts to *explain* their mental states or to *predict* their behaviours. For example, theory theorists cite compelling evidence from false belief experiments with young children (e.g., Wimmer and Perner 1983). The experiments show that children at three years of age are unable to offer explanations or predictions that would distinguish their own point of view from the perspective of others. At ages four and above, however, most children are able to understand and explain false beliefs in others and are thus able to predict the other's behaviour. These studies are interpreted (and often designed) to support the idea that children develop a theory of mind – that is, a theory about the mental states and behaviours of others – and at a certain point in development are able to use this theory to explain the other's beliefs and, on that basis, predict their behaviour.

In many authors this idea is then generalized to the claim that, once formed, we use theory of mind as our primary means for understanding other persons (e.g., Changeux in Changeux and Ricoeur 2000, 154–57; Tooby and Cosmides 1995). For some theorists, the mentalistic framework seems to be the only possible framework for understanding others, and they consider it "our natural way of understanding the social environment" (e.g., Baron-Cohen 1995, 3–4).

Can phenomenologists, who have themselves struggled with the assignment of primacy to mentalistic frameworks, offer any insights that might be useful in this context? We think there is good phenomenological evidence to the effect that, in the majority of intersubjective situations, the human subject does not normally posit a theoretical entity, called a mental state, and then attribute it to the other person (this has been a long-standing line of thought in phenomenology; see e.g. Depraz 1995; Gallagher 2001). We do not interact with the other by conceiving of her mind as a set of *cogitationes* closed up in immanence (Merleau-Ponty 1962, 353). Both theory theory and simulation theory conceive of communicative interaction between two people as a process that takes place between two Cartesian minds. Children with a theory of mind supposedly "see people as living their lives within a world of mental

content that determines how they behave in the world of real objects and acts." They construe "people's real-world actions as *inevitably* filtered through representations of the world rather than linked to the world directly" (Wellman 1993, 31–32). This assumes that one's understanding involves a retreat into a realm of *theoria* or *simulacra*, into a set of internal mental operations that come to be expressed (externalized) in speech, gesture, or interaction. If, in contrast, along with Merleau-Ponty, we think of communicative interaction as being accomplished in the very action of communication, in the speech, gesture, and interaction itself, then the idea that the understanding of another person involves an attempt to theorize about an unseen mental state, or to "mind-read," is problematic.

This is not to say that one cannot find the proper resources within the cognitive sciences that would correct the mentalistic approach of theory of mind. A large body of research into "primary intersubjectivity" (Trevarthan 1979) shows the importance of certain embodied practices – practices that are emotional, perceptual, and nonconceptual – for an understanding of another person's intentions long before the child reaches four years of age. These include the ability of infants to track the other person's eyes and to participate in shared attention behaviours. It is also clear that various movements of the head, the mouth, the hands, and more general body movements are perceived as meaningful (as being goal-directed) and are important for a perceptual (non-conceptual) understanding of the intentions and dispositions of other persons as well as for social reinforcement (see review by Allison, Puce, and McCarthy 2000). There is also evidence for affective and temporal coordination between the gestures and expressions of the infant and those of the other persons with whom they interact. Infants "vocalize and gesture in a way that seems 'tuned' [affectively and temporally] to the vocalizations and gestures of the other person" (Gopnik and Meltzoff 1997, 131). Importantly, the perception of emotion in other people's movement is a perception of an embodied comportment, rather than a theory or simulation of an emotional state. Moore et al. (1997) have demonstrated the emotional nature of human movement using point-lights attached to various body joints. Subjects view the abstract but clearly embodied movement in a darkened room and are able to identify the emotion that is being represented. The emotional

states of others are thus not mental attributes that we have to infer. Rather, I perceive the emotion in the movement and expression of the other's body.

Indeed, phenomenologists have a great deal to gain by taking seriously recent discoveries in developmental studies and in neuroscientific research. Studies of "primary intersubjectivity" have a great deal to offer in support of or as corrective of phenomenological analyses. The discovery that newborns are capable of a certain kind of imitation has implications that would push many of Merleau-Ponty's conclusions back to earlier stages of development (see Gallagher and Meltzoff 1996). The discovery of mirror neurons in the premotor cortex and Broca's area has important implications for our understanding of intersubjective perception, as well as the development of gesture and language. All of these scientific studies are suggestive for phenomenology, and, in turn, phenomenology can play an essential role in working out the proper interpretations and implications of these studies.

Rather than simply naming more examples of areas in which phenomenology can work together with the cognitive sciences, we want to turn to an issue that many phenomenologists consider to be fundamental for an understanding of consciousness, namely, the problem of time-consciousness. We want to show that with respect to this specific problem, phenomenology might offer correctives to various cognitive analyses, but also that phenomenology might benefit from some of the more sophisticated cognitive approaches.

Part II: A Resetting of Time-Consciousness

In the cognitive sciences today, studies that address temporality and cognition are usually cast in terms of working (short-term) memory and the "binding" problem (that is, the problem of how the distributed processing of information in the brain can give rise to a unified percept or action). These are specialized problems and for the most part are addressed at levels of sub-personal cognitive mechanisms or in terms of neurological processes. In such cases, the temporality problem seems to be located at a level of analysis that is isolated from more general phenomenological concerns. The notion that the temporality

of consciousness is an essential feature that ought to be integrated in a large variety of cognitive analyses is often lost. For this reason cognitive analyses are often static. Furthermore, this forgetfulness of the temporal nature of consciousness leads to accounts of cognition that are phenomenologically problematic.

In this part we want to pursue a twofold thesis. First, we want to show that the phenomenology of time-consciousness can resolve certain problems found in static cognitive accounts of experience. Second, we want to suggest that insights developed in the study of cognitive dynamics can contribute to a better understanding of time-consciousness.

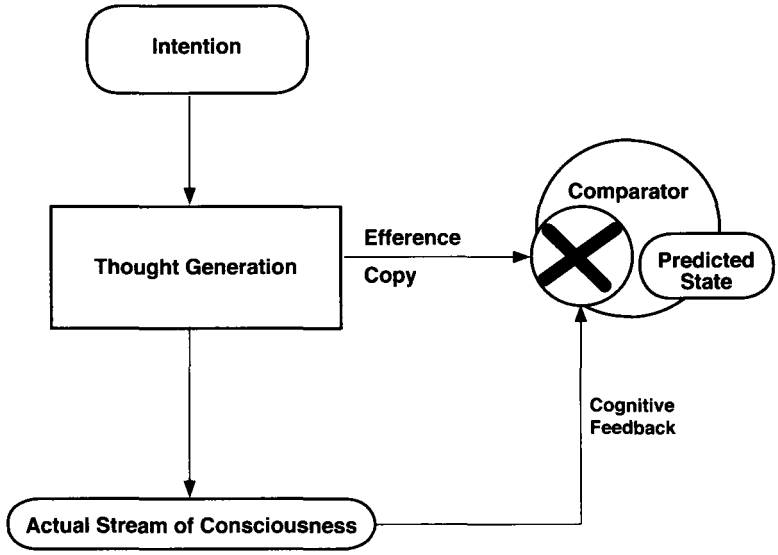
A Cognitive Model and Some Phenomenological Problems

We begin by considering a model of cognition developed by Christopher Frith (1992) in his influential analysis of schizophrenia. Following Feinberg (1978), Frith applies an explanation developed in an analysis of the self-monitoring of motor action to the analysis of mental experience. His hypothesis is that positive symptoms of schizophrenia, such as thought-insertion, delusions of control, and auditory hallucination, involve a failure of self-monitoring which does not occur in normal subjects. Frith provides an account of self-monitoring in terms of subpersonal cognitive mechanisms ultimately intended to be cashed out in terms of neurophysiology. Frith explains the phenomenology of thought insertion in the following way.

> Thinking, like all our actions, is normally accompanied by a sense of effort and deliberate choice as we move from one thought to the next. If we found ourselves thinking without any awareness of the sense of effort that reflects central monitoring, we might well experience these thoughts as alien and, thus, being inserted into our minds. (Frith 1992, 81).

The Feinberg-Frith model assumes that thinking is a kind of action, and that, similar to instances of motor action, it is normally accompanied by an effortful intention. According to this account, the intention to think is the element that guarantees my *sense of agency* for the thought. But

Figure 1: The Feinberg-Frith Model



this guarantee works only through mechanisms that operate on a subpersonal, non-conscious level. The intentional generation of thought generates not only the conscious thought, but, on the subpersonal level, an extra signal, an "efference copy," that is sent to a comparator mechanism. The comparator acts as a central monitor, which registers the occurrence of the actual thought, thus verifying that the intention and the thought match (see figure 1).

If the efference copy is somehow blocked from reaching the central monitoring mechanism, thought occurs which seems, to the subject, not to be generated by the subject himself. If the efference copy is blocked or goes astray, or is not properly generated, thinking still occurs, but it is not registered as under the subject's control (intention and thinking fail to match) – it appears to be an alien or inserted thought.

Gallagher (2000) has shown that several aspects of the Feinberg-Frith model are phenomenologically problematic. A first set of problems pertains to Frith's characterization of the intention to think. What role

does something like an "intention to think" or the efference copy of that intention play in the case of thinking or conscious experience? It is difficult to conceive of an intention to think prior to thinking itself, unless it is a conscious preparation, as when I might decide to sit down and start thinking about this issue. In that case, however, the intention to think is itself a case of thinking, and we are threatened by an infinite regress: Do I require an intention to think in order to intend to think?

Frith speaks of a *conscious* feeling of effort for a willed intention to think, and he equates this with a *conscious* monitoring of efference copy (1992, 86). This description, which relies not just on an intention to think, but on a "metarepresentation" of the intention to think, fails to capture a built-in sense of agency for thought. Metarepresentation is a second-order reflective consciousness, "the ability to reflect upon how we represent the world and our thoughts." According to Frith, this is part of what it means to monitor our actions and thoughts, and, he claims, it is precisely what is missing or disrupted in the schizophrenic's experience.[5]

Most cases of normal thinking, however, are neither prefaced by conscious intentions to think, nor followed by an introspective metarepresentation. Furthermore, if Frith were to remain consistent with his own model, then metarepresentational introspection is itself a thinking process. It should generate its own efference copy, which would have to be matched on top of the original match. It would thus add an extra level of consciousness to the comparator's verification process and threaten an infinite regress once again.

In the normal phenomenology, at least in the large majority of cases, there is not first an intention and then a thinking, nor thinking plus a concurrent but separate awareness of intention to think. John Campbell (1999a), who considers some version of the Feinberg-Frith account to be the most parsimonious one available, suggests, in contrast to Frith's characterization, that efference copy is not itself available

5 We note that this view of consciousness is somewhat similar to David Rosenthal's "higher-order thought" model (Rosenthal 1997). For a phenomenological critique of this model, see Zahavi and Parnas (1999).

to consciousness. Campbell, however, retains the idea that efference copy matches up with thought itself (introspectively observed) at the comparator in order to verify that the thought is one's own.[6]

It is difficult to understand why something like *efference copy* is really necessary in the case of conscious thinking. In the model of visuo-motor control, from which Frith takes the notion of efference copy, the latter serves a pragmatic or executive function rather than a verificational one. In the case of self-movement, the motor system sends efference copy of a motor command to the visual and vestibular systems, informing those systems to make adjustments, with the very practical effects of stabilizing the visual field. In that case the function of efference copy is to inform the visual and vestibular systems that the organism, rather than the world, is moving. Its purpose is visuo-motor control, not verification that movement is taking place. Is this type of executive function necessary for the thinking process?

Although one can distinguish different cognitive systems – the memory system, the perceptual system, and so on, the Feinberg-Frith model does not assign to efference copy a communicative role among these systems. Campbell suggests, following Feinberg, that efference copy has the pragmatic function of keeping thoughts on track, checking "that the thoughts you actually execute form coherent trains of thought" (1999a, 616). To keep thoughts coherent and on track, however, means keeping them on a semantic track, that is, on a certain track of meaning. Why assign this task to a subpersonal, non-semantic mechanism when,

6 Campbell describes the comparator process as involving a form of introspection: "it is the match between the thought detected by introspection, and the content of the efferent copy picked up by the comparator, that is responsible for the sense of ownership of the thought" (1999b). He also states: "You have knowledge of the content of the thought only through introspection. The content of the efferent copy is not itself conscious. But it is match at the monitor between the thought of which you have introspective knowledge and the efferent copy that is responsible for the sense of being the agent of that thought. It is a disturbance in that mechanism that is responsible for the schizophrenic finding that he is introspectively aware of a thought without having the sense of being the agent of that thought."

simply put, we are consciously aware of our thoughts and can keep track of them, and keep them on track, at a conscious level? Nor is it clear that we would ordinarily need a second-order or metarepresentational consciousness to keep our first-order thoughts on track, or to verify that I myself am doing the thinking.

A further set of problems encountered in Frith's model involves its static nature (Gallagher 2000). Frith takes no account of the temporal flow-structure of thought. Although he does acknowledge that the subpersonal comparator mechanisms involve issues of timing (for example, arrival time of efference copy at the comparator relative to registration of the conscious thought), he does not consider the temporal structure of the thinking itself. On a more adequate version of his model, the temporal structure of consciousness would introduce important constraints on the operations of the central monitor.

One motivation for incorporating temporality into Frith's model is to provide an explanation of why the central monitor might fail to register thought as self-generated in some instances but not in all instances. We will refer to this as *the problem of the episodic nature of positive symptoms*. Frith's description of the neurophysiology associated with the positive symptoms of schizophrenia does not address this issue. "Positive symptoms occur because the brain structures responsible for willed actions no longer send corollary discharges to the posterior parts of the brain concerned with perception. This would be caused by disconnections between these brain regions" (1992, 93). One would need to explain why these disconnections manifest themselves in some patients episodically, but not at all times. Simply put, not all of the schizophrenic's thoughts are experienced as inserted thoughts.[7]

7 That this is the case is clear, not only from empirical reports by patients, but by logical necessity. The subject's complaint that various thoughts are inserted depends on a necessary contrast between thoughts that seem inserted and those that do not seem inserted – and at a minimum, the thoughts that constitute the subject's complaint cannot seem inserted. If all thoughts were experienced as inserted by others, the subject would not be able to complain "in his own voice," so to speak.

In this case, the phenomenology obviously places constraints on the cognitive explanation.

A second problem for Frith's model involves the selectivity of positive symptoms. In this regard, in cases of thought insertion, specific kinds of thought contents, but not all thought contents, appear to be thought inserted. No explanation that remains totally on the subpersonal level will be able to explain this selectivity. It is not simply that in experiences of thought insertion patients occasionally experience thoughts coming into their minds from an outside source. Rather, their experiences are very specific, and are sometimes associated with specific others. Such thoughts seem to have a certain semantic and experiential consistency that cannot be adequately explained by the disruption of subpersonal processes alone. For example, a schizophrenic will report that thoughts are being inserted by a particular person and that they are always about a specified topic or that, in auditory hallucination, the voice always seems to say the same sort of thing.

A Frithian model that includes the notion of a temporal stream of thought, however, could explain problems with the sense of agency in terms of a lack of temporal synchrony at the subpersonal level – between the stream of thought as represented at the comparator, and efference copy. To explain the episodic nature of the symptoms one could appeal to certain events on the personal or experiential level that would motivate the desynchronization (Gallagher 2000). The personal-level motivation, the description of which is best captured with the help of a phenomenological analysis, might then be understood as part of the explanation for the selectivity of thought insertions.

But even in a less static version of Frith's model, the phenomenological problems involving intention to think and efference copy, as we outlined above, would still remain. In contrast, we want to suggest that a phenomenological model of the sort introduced by Husserl in his analysis of time-consciousness, provides a much more parsimonious account of intentionality and self-monitoring, one, moreover, that would allow an explanation of the episodic and selective nature of inserted thought.

A Phenomenological Model

The Feinberg-Frith model of cognition borrows heavily from explanations of motor control in terms of efference copy and comparators. It is important to note, however, that intentional aspects of motor action, and the generation of the sense of agency for such action, are normally experienced as *intrinsic* to the action. They are phenomenologically indistinguishable properties of the acting itself (Marcel 2003; Gallagher and Marcel 1999). There is good evidence from the study of motor action, for example, that an intrinsic sense of agency is based on "forward," anticipatory processes that occur prior to the action (Georgieff and Jeannerod 1998; Haggard and Eimer 1999; Haggard and Magno 1999). Alain Berthoz, in his recent work on movement, makes much of the ubiquity of such anticipatory mechanisms in the sensory-motor systems. Anticipation is "an essential characteristic of their functioning" and serves our capacity to reorganize our actions in line with events that are yet to happen (Berthoz 2000, 25). The neurological and behavioural evidence suggests that the sense of agency for action, which goes awry in pathological symptoms such as delusions of control, is not based on a *post factum* verification, or a comparator function occurring subsequent to the action or thought. Rather, the sense of agency is generated in an executive or control function that anticipates action.

As we have indicated, efference copy may indeed play an important practical (executive) role in the case of visuo-motor systems, but it is not clear what role it would play in the stream of thought. Alternative and more parsimonious explanations for a sense of agency that is *intrinsic* to thought, and for the loss of the sense of agency in schizophrenic thought insertion, can be advanced by employing Edmund Husserl's model of the retentional-protentional structure of time-consciousness.

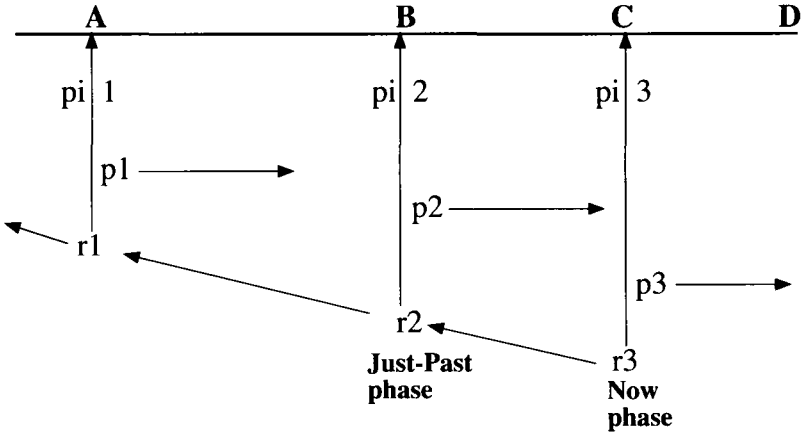
My conscious experience includes a sense of what I have just been thinking (perceiving, remembering, etc.) and a sense that this thinking (perceiving, remembering, etc.) will continue in either a determinate or indeterminate way. This phenomenological temporal sense is based on retentional and protentional dynamics, following the logic of time-consciousness that Husserl outlines.

Husserl's analysis of time-consciousness not only explains how the experience of temporal objects is possible, given an *enduring* act of

consciousness, it also explain how consciousness unifies *itself* across time.[8] In figure 2, the horizontal line ABCD represents a temporal object such as a melody of several notes. The vertical lines represent abstract momentary phases of an enduring act of consciousness. Each phase is structured by three functions. First, a *primal impression* (pi), which allows for the consciousness of an object (a musical note, for example) that is simultaneous with the current phase of consciousness. Second, a *retention* (r), which retains previous phases of consciousness and their intentional content. And third, a *protention* (p), which anticipates experience which is just about to happen.

Each phase of consciousness involves a retention of the previous phase of consciousness. Since the previous phase includes its own retention of a previous phase, there is a retentional continuum that stretches back through prior experience. There are two important

Figure 2: Time-Consciousness according to Husserl



8 A more detailed account can be found in Husserl (1991). For an extended analysis of Husserl's model and its similarities and differences from James's notion of the specious present, see Gallagher (1998).

aspects to this retentional continuity. The first, the "longitudinal intentionality" (*Längsintentionalität*) of retention, provides for the intentional unification of consciousness itself since retention is the retention of previous phases of consciousness. Second, since the prior phases of consciousness contain their respective primal impressions of the experienced object, the continuity of that experienced object is also established. Husserl refers to this as the "transverse intentionality" (*Querintentionalität*) of retention (Husserl 1991, 85).

Retention, for example, keeps the intentional sense of the words of a sentence available even after the words are no longer sounded. Moreover, implicit in this retentional function is the sense that *I* am the one who has just said or heard these words. The words do not become part of a free-floating anonymity, they remain part of the sentence that *I* am in the process of uttering or hearing. Furthermore, if I am in the process of uttering a sentence, I have some anticipatory sense of where the sentence is going, or at the very least, that the sentence is heading to some kind of ending. This sense of knowing where the sentence (the thought) is heading, even if not completely definite, seems essential to the experience I have of speaking in a meaningful way.

The protentional aspect of consciousness provides it with this intentional anticipation of something about to happen. Husserl points out that protention allows for the experience of surprise. If I am listening to a favourite melody and someone hits the wrong note, I am surprised or disappointed. If someone fails to complete a sentence, I experience a sense of incompleteness, precisely because consciousness involves an anticipation of what is to come next, and in these cases, what actually happens fails to match my anticipation. The content of protention, however, is not always completely determinate and may approach the most general sense of "something (without specification) has to happen next."

Husserl's analysis of protention doesn't go much further. As we saw above, however, in his analysis of retention he suggests that there is a double intentionality – one aspect that is directed at the content of experience and another that is directed at consciousness itself. In listening to a melody, I am not only aware of the melody, I am implicitly aware of myself as I am aware of the melody. This implicit, longitudinal intentionality is a non-observational, pre-reflective

awareness of my own flowing consciousness, which delivers an implicit sense that this experience is part of my stream of consciousness. This sense of ownership for the experience involves no reflective, second-order, metacognition.

Although protention is asymmetrical with retention in many regards (Gallagher 1998; Varela 1999), there is clearly a longitudinal aspect to protention. That is, my anticipatory sense of the next note of the melody, or of where my sentence is going, or that I will continue to think, etc., is also, implicitly, an anticipatory sense that these experiences will be experiences *for me*, or that *I* will be the one listening, speaking, or thinking. In effect, protention involves a projective sense of what *I* am about to do or experience. Indeed, in contrast to the indeterminate sense of what the content of protention may be, the anticipatory sense of self is relatively determinate.

In the normal case, as we indicated, the *sense of agency* with respect to my own thought does not develop retrospectively, as if I must stop the process and, in a Frithian metarepresentation, think whether I am really the one who is thinking. Rather, taking the clue from the importance of anticipation in motor action, we can pursue the idea that the sense of agency involves an implicit anticipatory moment in thinking itself.[9] As Husserl's analysis shows, this anticipatory aspect is part of the very structure of consciousness, rather than a second-order retrospection or verification. This suggests that protention plays a role in providing a sense of agency in the cognitive domain.

Consider, first, that thought may be generated by the subject in a willed and controlled fashion. Telling a story is a good example. I follow the well-known plot, I have a sense of where I am going in the narrative,

9 Following clues from the analysis of motor action, Gallagher (2000) suggests that the dynamics of *protention* underlie the sense of agency for thought, or more precisely, that the protentional registration is a necessary but not a sufficient condition for the sense of agency. The function of *retention*, on the other hand, provides a sense of ownership for thought. A failed protentional mechanism may explain difficulty in generating spontaneous actions, a negative symptom in schizophrenia, as well as difficulties in performing self-directed search (see Frith 1992, 48).

and I push the thinking process along from one step to another in a controlled manner. In this case the protentional aspect of consciousness operates to give me a sense of where my thinking is going *in its very making*, that is, as it is being generated and developed. It provides a sense that the thinking process is being generated in my own stream of consciousness and, to some degree, under my control. Protention normally puts me in the forefront of my thoughts and allows me to take up these thoughts as my own product, as they develop.

A second kind of thinking may be more passive. Unbidden thoughts, memories, fantasies may invade my current stream of consciousness. These are thoughts that I do not intend, although my sense is that these thoughts are coming from myself, rather than from some alien source. They are not only part of my stream of consciousness, but, despite the fact that I am not willing them, and may even be resisting them, they seem to be generated within my own cognitive experience. Within the framework of an unwanted memory or an unwelcome fantasy, protention provides a sense of where the thoughts are coming from and where they are heading, as they are being passively generated. Protention may also provide a sense of *not knowing* where we are heading, a sense of uncertainty, or indeterminacy with respect to where such thoughts will lead. Even in such indeterminacy, I have a sense that they are originating and developing within my stream of consciousness – a sense of *passive generation*.

What would our experience be like if the protentional aspect of consciousness disappeared? In the case of passive, unbidden thoughts, thinking would continue to happen, but there would be neither a sense of agency, nor a sense that these thoughts were being passively generated in my cognitive system, even though they were appearing in my stream of consciousness. Without protention, thought continues, but it appears already made, not generated in my own stream of consciousness. These thoughts would appear as if from nowhere, suddenly and unexpectedly. I would be able to make sense of them only in their retentional train, in retrospect, but not as something self-generated.

This may even be the case with intended thoughts for which I would normally have a sense of agency. Without protention, whatever intention I may have, my sense of what I am about to do or think is disrupted. My non-observational, pre-reflective sense of agency, which I

normally experience within a protentional framework, will be missing due to the lack of protention. In this case, I will experience thoughts that seem to anticipate what I *would have* thought. This is consistent with the reports of some schizophrenic patients, that the external force responsible for the inserted thoughts seems to know what they intend to think before they actually think it (Spence 1996). The thought seems to match up with their own intention, but it still seems to them that they are not the agentive cause of the thought.

Without protention, in such cases, thoughts will continue to occur within the stream of consciousness. But they are not experienced *in the making*. Schizophrenics experience what is actually their own thinking, but as thinking that is not generated by them, a thinking that is *already made* or pre-formed for them. This primary failure of protention may then motivate a metarepresentational reflection, an introspection that becomes the hyperreflection characteristic of schizophrenic experience (Sass 1998). In metarepresentation the patient may start to ascribe the seemingly alien thought to some particular force or individual and report that it is inserted (Spence 1996).

Appealing to the Husserlian model of time-consciousness as part of an explanation of intentionality and self-agency for thought, and of their failure in cases of thought insertion, provides a much more parsimonious account than the Feinberg-Frith model. It allows for an *implicit* sense of agency that corresponds to the normal phenomenology of action and thought and the failure of that implicit sense in pathological cases, and it does not require an extra system involving efference copy and comparators. There are still two issues to be resolved. First, can this account be made consistent with the episodic and selective nature of inserted thought. And second, what explains the failure of protention? Both of these questions, however, lead us back to the cognitive sciences.

A Dynamic Model of Protention

There is good empirical evidence to support the idea of a breakdown in the protentional function in schizophrenia. Specifically, research into various aspects concerning schizophrenia and temporality supports this interpretation. These aspects include:

(1) Inability to act toward the future, and a feeling that the future is a repetition of the past (Minkowski 1933)

(2) Difficulties in indexing events in time, with positive correlation to inner-outer confusions (manifested in symptoms such as auditory hallucinations, feelings of being influenced, delusional perceptions, and so forth) (Melges 1982; Melges and Freeman 1977)

(3) Curtailment of future time-perspective (Wallace 1956; Dilling and Rabin 1967)

(4) Difficulty planning and initiating action (Levin 1984)

(5) Problems with temporal organization (Klonoff et al. 1970; DePue et al. 1975)

(6) Difficulties with experienced continuity (Pöppel 1994).

(7) Slowing of temporal processing of different sense modalities, leading to a form of "temporal diplopia" in which consciousness does not seem to coincide with itself. (Pöppel 1994, 192)

These "impairments of self-temporalization" (Bovet and Parnas 1993, 584) are consistent with problems concerning the protentional aspect of experience and are linked with the same neurological dysfunctions involved in the schizophrenic's voluntary movement (Singh et al. 1992; also see Graybiel 1997).

What could cause the failure of protention? There are several ways that one can approach this question. In terms of traditional cognitive science, the task would be to identify a certain set of mechanisms on the subpersonal level. Some aspect of information processing in the neuronal mechanisms that underpin working memory might be responsible for the disruption of protentional consciousness. This approach would be quite consistent with Frith's attempt to identify the cognitive mechanisms responsible for the positive symptoms of schizophrenia. That is, one would attempt to identify specific dysfunctions or disconnections in brain structures responsible for

delivery of efference copy to relevant parts of the brain. Such a subpersonal account, however, would necessarily be incomplete unless one could explain why these mechanisms work in some cases but not in others (the problems of the selective and episodic nature of positive symptoms) and how they work with complex psychodynamic processes (such as repression and obsession) in psychopathology.

Insofar as this approach casts the problem in strict information-processing terms (the slowing and discordance of temporal processing) it also ignores an important phenomenological feature of such experience – a feature that may, in turn, play an important role in both positive and negative symptoms of schizophrenia. Clearly, a certain *affective disconcertion* is an integral part of the phenomenological picture associated with these impairments of self-temporalization. That is, the schizophrenic phenomenology is not simply a structural or logical problem that can be captured in terms of a deficiency in information processing. If the subject's experiences can be characterized in some cases as involving a logical "asynchronicity" or a confusion of thought, these cognitive problems are importantly not affectively neutral (even if the content of thought is not emotionally charged) but are characterized by specific affective dispositions, which in some cases involve a lack of or alienation from affect.

This idea links up closely to an explication of protention in terms of *affective tone* (Varela 1999, 2000; Varela and Depraz 2000). On this view, protention is constitutionally involved with an affective tensity or, from a different perspective, a readiness for action. Protention, a formal structure of noetic intentionality, according to Husserl, is directed toward a certain kind of content, which, most formally stated, is *the not yet*. The "not yet" is always suffused with affect at the same time that it is conditioned by the emotional tone that accompanies the flow of experience. Protention involves something like a prediction of the unpredictable. Through protention, one opens oneself to the *not yet* and is in that sense as much self-affected as affected by its content. As Husserl puts it, consciousness "is affected by that which consciousness is conscious of, it trails affect; it is attracted, held, and taken in by that which affects it" (MS C III/I). Depraz explains this in terms of Husserl's emphasis on instinctive intentionality: "Affect is there before being there for me in full consciousness: I am affected

before knowing that I am affected. It is in that sense that affect can be said to be primordial" (1994, 75)

We can think of this also in Heideggerian terms: one finds oneself always in a certain disposition (*Befindlichkeit*), lived through as a certain mood, for example, and often this disposition is experientially transparent. That is, one lives through it prenoetically, as a pre-reflective being-in-the-world. At a different level of description, affect has a deeply rooted biological basis; affective disposition is tied to certain neuronal dispositions. Affective tonality, experiential transparency, and readiness for action can be entirely reordered by neuronal events and the balance of neurotransmitters. Such a reordering can sometimes leave a subject specifically indisposed for experience or action.

Is it possible to find a common, albeit abstract, level of description that would capture both the dynamics of neuronal processes and the dynamics of the retentional-protentional flow of time-consciousness? A common feature is that both domains involve self-organizing dynamics (self-constituting processes) which are defined by boundary conditions and initial conditions. Affective tonality can be seen as a major boundary and initial condition for neurodynamics in the temporal flow embodied in brain integration mechanisms (Varela 1999).

A neurodynamical account postulates the following model. Every cognitive act, from perceptuo-motor behaviour to human reasoning, arises through the concurrent participation of several functionally distinct and topographically distributed regions of the brain and their sensori-motor embodiment (Varela et al. 2001). The task of integrating these different neuronal components involves a process that is set out in a complex temporal framework of three different scales of duration (Pöppel 1988; Varela et al. 1991; Varela 1999), the first two of which are directly relevant here.

(1) elementary events (the 1/10 scale varying between 10–100 msecs)
(2) relaxation time for broader integration (the 1 scale, varying from 0.5 to 3 secs)
(3) descriptive-narrative assessments (the 10 scale involving memory)

The various neuronal processes in need of integration at the level of the second scale require a temporal frame or window that defines the duration of the lived present. The best way to understand the recursive structuring of these temporal scales is on the model of nonlinear dynamics.

Evidence for the first scale is found in the so-called fusion interval of various sensory systems: the minimum amount of time needed for two stimuli to be perceived as non-simultaneous, a threshold which varies with each sensory modality. The elementary sensory-motor events that constitute experience (corresponding to what Husserl would call hyletic data) can be grounded in the intrinsic cellular rhythms of neuronal discharges within the range of 10 msec (the rhythms of bursting interneurons) to 100 msec (the duration of an EPSP/IPSP sequence in a cortical pyramidal neuron). These elementary events are then integrated into the second scale, corresponding to the living or "specious" present, the level of a fully constituted, normal cognitive operation. Neuroscience explains this integration in terms of cell assemblies, distributed subsets of neurons with strong reciprocal connections (see Varela 1995; Varela et al. 2001). In terms of a dynamical systems model, the cell assembly must have a relaxation time followed by a bifurcation or phase transition, that is, a time of emergence within which it arises, flourishes, and subsides, only to begin another cycle. Integration occurs because neural activity forms transient aggregates of phase-locked signals coming from multiple regions. Synchrony (via phase-locking) must *per force* occur at a rate sufficiently high so that there is enough time for the integration to hold together within the constraints of transmissions times. In brief, we have neuronal-level constitutive events, which have a duration on the 1/10 scale, forming aggregates that manifest themselves as incompressible but complete cognitive acts on the 1 scale. This completion time is dynamically dependent on a number of dispersed assemblies and not on a fixed integration period; in other words it is the basis of the origin of experienced duration without an external or internal ticking clock. This temporal window is necessarily flexible (0.5 to 3 secs) depending on a number of factors: context, fatigue, sensory modality, age, and so on. This integration-relaxation process at the 1 scale level,

corresponds to the living present and allows for an integration that is describable in terms of the retentional-protentional structure.

The kind of self-organization that underlies the emergence of neural assemblies thus involves a component level (at the 1/10 scale) that cashes out in terms of single or groups of nonlinear oscillators. These oscillators enter into a synchrony that is registered as a collective indicator or variable, a relative phase. This collective variable manifests itself at a global level as a cognitive action or behaviour. The self-organization involved here is not an abstract computation, but an embodied behaviour subject to initial conditions (characterized, for example, as what the experiencing subject intends to do or has just done), and non-specific parameters (for example, changes in perceptual conditions, attentional modulation). In this regard there are specific local-global interdependencies. The emerging behaviour or experience cannot be understood independently of the elementary components (for example, hyletic data generated by the organism's interaction with the physical environment); the components attain relevance through their relation with their global correlate.

The fact that an assembly of coupled oscillators attains a transient synchrony and that it happens within a certain temporal window is the explicit substrate of the living present. The dynamical models and the data show that this synchronization is dynamically unstable and will thus constantly and successively give rise to new assemblies (these transformations define the trajectories of the system). Each emergence bifurcates from the previous ones determined by its initial and boundary conditions. Thus the preceding emergence is still present in the succeeding one as the trace of the dynamical trajectory (*retention* on the phenomenological level). The order parameters (initial conditions and boundary conditions) are important here. They are defined by the embodiment and experiential context of the action, behaviour, or cognitive act. The boundary conditions shape the action at the global level and include the contextual setting of the task performed, as well as the independent modulations arising from the contextual setting where the action occurs (i.e., new stimuli or endogenous changes in motivation) (Varela 1999).

The dynamical system described here does not accord with the classical notion of stability that derives from a mechanical picture

of the world, or a computational picture of cognition. Stability in the latter case means that initial and boundary conditions lead to trajectories concentrated in a small region of phase space where the system remains, a point attractor or a limit cycle. In contrast, biological systems demonstrate *instability* as the basis of normal functioning – constitutional instabilities are the norm (see Varela 1999 for a summary of the empirical evidence). This instability accounts for the formal flow property of experience. Nonlinear systems provide a self-movement that does not depend (within a range of parameters) on the content of the system. Whether the experiential content of my visual percept is a person or a pyramid, the intrinsic or immanent motion is generically the same, a self-propelled motion. The self-constituting flow of consciousness involves a perpetual change punctuated by transient aggregates underlying momentary acts (at the 1 scale of duration). Changes in initial and boundary conditions drive this flow by motivating transformations to new dynamical phases, in a way that is not predictable along pre-determined trajectories.

Just here, protention plays an important role in the self-movement of the flow. If, as we have suggested, protention is linked to affective tonality (which is reflective of the embodied and contextualized situation), then it helps to define specific boundary and initial conditions for the neurodynamics just described. In the initiation of an intentional cognitive act (for example, I decide to look for a particular object in the environment) I induce a transformation that is coloured by an affective disposition that anticipates the change in perception. In the anticipation of a certain experience, I introduce exogenous order parameters that alter the geometry of the phase space.

Empirical evidence for this can be found in studies of intentional movement. The intention to carry out a movement is coupled with a change in affective tone that varies in degree. One well-known case involves the readiness potential. For a finger movement, a large slow electrical potential can be measured over the entire scalp, preceding by a fraction of a second the beginning of the motion. This is not a correlate of an intention, but it gives some indication of how vast a reconfiguration of a dynamical landscape is involved at the origin of a fully constituted act. Such diffuse effects are in accord with mechanisms associated with neurotransmitters that condition the modes of response at the neuronal level.

On this dynamical view, if protention is linked to affective tonality, then the way to account for the failure of protention is not to search for a particular mechanism (a comparator or misplaced efference copy). Rather, on the neurological level, the sort of mechanism that underlies protention is more appropriately thought of in terms of widely distributed and dynamical processes than in terms of localized functions. As a result, the conceptual framework for thinking about the neurological mechanisms responsible for the symptoms of schizophrenia is quite different from the one involving concepts of comparator, central monitor, efferent copy, etc. Schizophrenic patients feel alienated not just from thought and action; as Louis Sass (private correspondence) points out, they also feel alienated from affects, from their own body and skin, from their own saliva, from their own name, etc. It seems unlikely that all of these phenomena can be explained by problems involving efference copy – problems that may in fact be secondary to a more global dysfunction.[10]

This also means that a disruption of protention is likely to involve widespread cognitive and emotional problems of the sort found in schizophrenia, including incongruity of affect, flat affect (athymia), and "grossly inappropriate affect" (DSM-III-R). Consistent with this picture, premorbid characteristics of schizophrenia patients include difficulties in interpersonal relations, anxiety, neophobia, and defective emotional rapport (Bovet and Parnas 1993).

Importantly, the explanation that links together protentional problems, problems with the implicit sense of agency in cases of inserted thought, and other symptoms of schizophrenia that involve inordinate experiences of time and anomalous affective states, depends on elements in the intentional content of experience, and not simply on disruptions of a subpersonal mechanism. If certain subpersonal

10 Even in the realm of motor action, where such localized mechanisms may be involved in generating the sense of agency, since what is at stake is precisely a *sense* of agency, that is, an *experience* of agency, it is likely, according to the Husserlian model, that the protentional-retentional structure has a role to play in the conscious registration of that sense of agency for movement, or in its failure, for example, in delusions of control.

dynamics are predisposed to malfunction in schizophrenic patients, one possible trigger for this malfunction may be intentional content.

As we saw, a subpersonal explanation does not entirely address the episodic and selective aspects of thought insertion, which may in fact have their proximate cause on the level of semantic/intentional content. There are good arguments and good evidence to show that intentional content has an effect on the temporal structure of experience (Friedman 1990; Gallagher 1998; James 1890). Experience speeds up or slows down according to *what* we are experiencing. Consider, for example, the ordinary experience of how time passes when we are with different people. In some cases time passes too quickly; in other cases too slowly. If boredom can slow the system down and enjoyment and interest speed it up, perhaps anxiety or some experience-related change of affective disposition can cause a disruption of the subpersonal protentional dynamic with a resulting loss of protention in the phenomenological stream. It seems reasonable to propose that a disruption of the protentional dynamic could cause a looping effect that would reinforce the affective trigger. Without protention, for example, it is quite possible that patients would experience others and the world as being invasive, "on top of them," too close, etc., which are, in fact, anxiety-causing experiences, and experiences commonly reported by schizophrenics.

This alternative account of schizophrenic symptoms in terms of a disruption in the protentional dynamic does not need to postulate, in the cognitive domain, comparators or mechanisms involving efference copy. Rather, self-monitoring processes that involve the senses of ownership and self-agency, essential aspects of minimal self-awareness, are built into consciousness as the longitudinal aspects of the retentional-protentional structure. This account requires no mechanisms over and above the mechanisms that constitute the temporal structure of consciousness itself. Furthermore, this account resolves all of the phenomenological problems found in the Feinberg-Frith model. The "intention to think," for example, is not something separate from thinking itself; it is included in the very structure of thought. Accordingly, the schizophrenic does not discover alien thoughts by means of a metarepresentational introspection; rather he will have an

immediate, non-observational sense that something is wrong, a sense that is likely to motivate the hyper-reflective metarepresentations that characterize schizophrenia and that lead to further misattributions of agency.

References

Allison, T., Q. Puce, and G. McCarthy (2000). "Social Perception from Visual Cues: Role of the STS Region," *Trends in Cognitive Science* **4**: 267–78.

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

Bermudez, J., A. J. Marcel, and N. Eilan (1995). *The Body and the Self.* Cambridge: MIT Press.

Berthoz, A. (2000). *The Brain's Sense of Movement.* Cambridge: Harvard University Press.

Bovet, P., and J. Parnas (1993). "Schizophrenic Delusions: A Phenomenological Approach," *Schizophrenia Bulletin* **19**: 579–97.

Campbell, J. (1999a). "Schizophrenia, the Space of Reasons and Thinking as a Motor Process," *The Monist* **82**: 609–25.

Campbell, J. (1999b). "Immunity to Error through Misidentification and the Meaning of a Referring Term," *Philosophical Topics* **26**: 89–104.

Changeux, P., and P. Ricoeur (2000). *What Makes Us Think?* (trans. M. B. DeBevoise). Princeton: Princeton University Press.

Clark, A. (1997). *Being There: Putting Brain, Body, and World Together Again.* Cambridge, MA: MIT Press.

Dennett, D. C. (1991). *Consciousness Explained.* Boston: Little Brown.

Depraz, N. (1994). "Temporalité et affection dan les manuscrits tardifs sur la temporalité (1929–1935) de Husserl," *Alter* **3**: 81–114.

Depraz N. (1995). *Transcendance et incarnation. Le statut de l'intersubjectivité comme altérité à soi chez Husserl.* Paris: Vrin.

Depraz, N., F. J. Varela, and P. Vermersch (2000). "The Gesture of Awareness: An Account of its Structural Dynamics." In M. Velmans, ed., *Investigating Phenomenal Consciousness: New Methodologies and Maps* (121–39). Amsterdam: John Benjamins.

Depraz, N., F. J. Varela, and P. Vermersch (2003). *On Becoming Aware: A Pragmatics of Experiencing.* Amsterdam: John Benjamins.

DePue, R. A., M. D. Dubicki, and T. McCarthy (1975). "Differential Recovery of Intellectual, Associational, and Psychophysiological Functioning in Withdrawal and Active Schizophrenics," *Journal of Abnormal Psychology* **84**: 325–30.

de Vignemont, F. (2000). "When the 'I think' Does Not Accompany My Thoughts." Paper presented at Association for the Scientific Study of Consciousness. Bruxelles (July, 2000). Abstracted in *Consciousness and Cognition* **9**, Part 2.

Dilling, C., and A. Rabin (1967). "Temporal Experience in Depressive States and Schizophrenia." *Journal of Consulting Psychology* **31**: 604–8.

Feinberg, I. (1978). "Efference Copy and Corollary Discharge: Implications for Thinking and its Disorders." *Schizophrenia Bulletin* **4**: 636–40.

Fourneret, P., and M. Jeannerod (1998). "Limited Conscious Monitoring of Motor Performance in Normal Subjects," *Neuropsychologia* **36**: 1133–40.

Franck, N., Farrer, C., Georgieff, N., Marie-Cardine, M., Daléry, J., d'Amato, T., and Jeannerod, M. (2001). "Defective Recognition of One's Own Actions in Patients with Schizophrenia," *American Journal of Psychiatry* **158**: 454–59.

Friedman, W. (1990). *About Time: Inventing the Fourth Dimension*. Cambridge, MA: MIT Press.

Frith, C. D. (1992). *The Cognitive Neuropsychology of Schizophrenia*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Frith, C. D., and D. J. Done (1988). "Towards a Neuropsychology of Schizophrenia," *British Journal of Psychiatry* **153**: 437–43.

Frith, C. D., and U. Frith (1999). "Interacting Minds – A Biological Basis," *Science* **286**: 1692–95.

Gallagher, S. (1998). *The Inordinance of Time*. Evanston: Northwestern University Press.

Gallagher, S. (1997). "Mutual Enlightenment: Recent Phenomenology in Cognitive Science," *Journal of Consciousness Studies* **4(3)**: 195–214.

Gallagher, S. (2000). "Self-Reference and Schizophrenia: A Cognitive Model of Immunity to Error through Misidentification." In D. Zahavi, ed., *Exploring the Self: Philosophical and Psychopathological Perspectives on Self-experience* (203–39). Amsterdam: John Benjamins.

Gallagher, S. (2001). "The Practice of Mind: Theory, Simulation or Primary Interaction." In E. Thompson, ed., *Between Ourselves: Second-Person Issues in the Study of Consciousness*, 83–108. Thorverton, UK: Imprint Academic. Also published in *Journal of Consciousness Studies* **8(5–7)**: 83–108.

Gallagher, S., and J. Cole (1995). "Body Schema and Body Image in a Deafferented Subject," *Journal of Mind and Behavior* **16**: 369–90.

Gallagher, S., and A. J. Marcel (1999). "The Self in Contextualized Action," *Journal of Consciousness Studies* **6(4)**: 4–30.

Gallagher, S., and A. Meltzoff (1996). "The Earliest Sense of Self and Others: Merleau-Ponty and Recent Developmental Studies," *Philosophical Psychology* **9**: 213–36.

Georgieff, N., and M. Jeannerod (1998). "Beyond Consciousness of External Events: A 'Who' System for Consciousness of Action and Self-Consciousness," *Consciousness and Cognition* **7**: 465–77.

Goldman, A. I. (1989). "Interpretation Psychologized," *Mind and Language* **4**: 161–85.
Gopnik, A., and A. N. Meltzoff (1997). *Words, Thoughts, and Theories*. Cambridge, MA: MIT Press.
Gordon, R. M. (1986). "Folk Psychology as Simulation," *Mind and Language* **1**: 158–71.
Gordon, R. M. (1995). "Simulation without Introspection or Inference from Me to You." In M. Davies and T. Stone, eds. *Mental Simulation: Evaluations and Applications*. Oxford; Blackwell Publishers.
Graybiel, A. M. (1997). "The Basal Ganglia and Cognitive Pattern Generators," *Schizophrenia Bulletin* **23**: 459–69.
Haggard, P., and M. Eimer (1999). "On the Relation between Brain Potentials and the Awareness of Voluntary Movements," *Experimental Brain Research* **126**: 128.
Haggard, P., and E. Magno (1999). "Localising Awareness of Action with Transcranial Magnetic Stimulation," *Experimental Brain Research* **127**: 102.
Husserl, E. (1970). *Cartesian Meditations* (trans. D. Cairns). The Hague: Nijhoff.
Husserl, E. (1991). *On the Phenomenology of the Consciousness of Internal Time (1893–1917)* (trans. J. Brough). *Collected Works IV*. Dordrecht: Kluwer.
James, W. (1890). *The Principles of Psychology*. New York: Dover, 1950.
Klonoff, H., C. Fibiger, and G. H. Hutton (1970). "Neuropsychological Pattern in Chronic Schizophrenia," *Journal of Nervous and Mental Disorder* **150**: 291–300.
Leslie, A. (2000). "'Theory of Mind' as a Mechanism of Selective Attention." In M. Gazzaniga, ed., *The New Cognitive Neurosciences*. Cambridge, MA: MIT Press, 1235–47.
Levin, S. (1984). "Frontal Lobe Dysfunction in Schizophrenia – Eye Movement Impairments," *Journal of Psychiatric Research* **18**: 27–55.
Lutz, A., J.-P. Lachaux, J. Martinerie, and F. J. Varela (2002). "Guiding the Study of Brain Dynamics Using First-Person Data: Synchrony Patterns Correlate with Ongoing Conscious States during a Simple Visual Task," *Proceedings of the National Academy of Sciences USA* **99**: 1586–91.
Marbach, E. (1993). *Mental Representation and Consciousness: Towards a Phenomenological Theory of Representation and Reference*. Dordrecht: Kluwer.
Marcel, A. J. (2003). "The Sense of Agency: Awareness and Ownership of Actions and Intentions." In J. Roessler and N. Eilan, eds., *Agency and Self-Awareness: Issues in Philosophy and Psychology*. Oxford: Oxford University Press.
Melges, F. T. (1982). *Time and the Inner Future: A Temporal Approach to Psychiatric Disorders*. New York: Wiley.
Melges, F. T., and A. M. Freeman (1977). "Temporal Disorganization and Inner-Outer Confusion in Acute Mental Illness," *American Journal of Psychiatry* **134**: 874–77.

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

Merleau-Ponty, M. (1963). *The Structure of Behavior* (trans. A. L. Fisher). Boston: Beacon Press.

Merleau-Ponty, M. (1964). *The Primacy of Perception* (trans. W. Cobb). Evanston: Northwestern University Press.

Minkowski, E. (1933). *Lived Time: Phenomenological and Psychopathological Studies* (trans. N. Metzel). Evanston: Northwestern University Press, 1970.

Moore, D. G., R. P. Hobson, and A. Lee (1997). "Components of Person Perception: An Investigation with Autistic, Non-Autistic Retarded and Typically Developing Children and Adolescents," *British Journal of Developmental Psychology* **15**: 401–23.

Nagel, T. (1974). "What is it like to be a bat?" *Philosophical Review* **82**: 435–50.

Pöppel, E. (1988). *Mindworks: Time and Conscious Experience*. Boston: Harcourt Brace Jovanovich.

Pöppel, E. (1994). "Temporal Mechanisms in Perception." *International Review of Neurobiology* **37**: 185–202.

Port, R., and T. van Gelder, eds. (1995). *Mind as Motion: Explorations in the Dynamics of Cognition*. Cambridge, MA: MIT Press.

Rosenthal, D. (1997). "A Theory of Consciousness." In N. Block et al., eds., *The Nature of Consciousness*. Cambridge, MA: MIT Press.

Rizzolatti, G., Fadiga, L., Matelli, M., Bettinardi, V., Paulesu, E., Perani, D., and Fazio, G. (1996). "Localization of Grasp Representations in Humans by PET: 1. Observation versus Execution," *Experimental Brain Research* **111**: 246–52.

Roy, J-M., Petitot, J., Pachoud, B. and Varela, F.J. (1999). "Beyond the Gap: An Introduction to Naturalizing Phenomenology." In J. Petitot, F.J. Varela, B. Pachoud, and J.-M. Roy, eds. *Naturalizing Phenomenology: Issues in Contemporary Phenomenology and Cognitive Science* (1–83). Stanford: Stanford University Press.

Sass, L. (1998). "Schizophrenia, Self-Consciousness and the Modern Mind," *Journal of Consciousness Studies* **5(5–6)**: 543–65.

Singh, J. R., Knight, R.T., Rosenlicht, N., Kotun, J.M., Beckley, D.J., and Woods, D.L. (1992). "Abnormal Premovement Brain Potentials in Schizophrenia," *Schizophrenia Research* **8**: 31–41.

Spence, S. (1996). "Free Will in the Light of Neuropsychiatry," *Philosophy, Psychiatry, and Psychology* **3**: 75–90.

Thompson, E., ed. (2001). *Between Ourselves: Second-Person Issues in the Study of Consciousness*. Special issue of *Journal of Consciousness Studies* **8(5–7)**.

Tooby, J., and L. Cosmides (1995). "Foreword" to S. Baron-Cohen, *Mindblindness: An Essay on Autism and Theory of Mind* (xi–xviii). Cambridge, MA: MIT Press.

Trevarthen, C. (1979). "Communication and Cooperation in Early Infancy: A Description of Primary Intersubjectivity." In M. Bullowa, ed., *Before Speech*. Cambridge: Cambridge University Press.

Varela, F. J. (1995). "Resonant Cell Assemblies: A New Approach to Cognitive Functioning and Neuronal Synchrony," *Biological Research* **28**: 81–95.

Varela, F. J. (1996). "Neurophenomenology: A Methodological Remedy for the Hard Problem," *Journal of Consciousness Studies* **3(4)**: 330–49.

Varela, F. J. (1999). "The Specious Present: A Neurophenomenology of Time Consciousness." In J. Petitot, F. J. Varela, B. Pachoud, and J.-M. Roy, eds., *Naturalizing Phenomenology: Issues in Contemporary Phenomenology and Cognitive Science* (266–314). Stanford: Stanford University Press.

Varela, F. J., and N. Depraz (2000). "At the Source of Time: Valance and the Constitutional Dynamics of Affect," *Arobase: Journal des lettres et sciences humaines* **4**: 143–66.

Varela, F. J., E. Thompson, and E. Rosch (1991). *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge, MA: MIT Press.

Varela, F. J., J. P. Lachaux, E. Rodriguez, and J. Martinerie (2001). "The Brainweb: Phase-Synchronization and Long-Range Integration," *Nature Reviews Neuroscience* **2**: 229–39.

Wallace, M. (1956). "Future Time Perspectives in Schizophrenia," *Journal of Abnormal and Social Psychology* **52**: 240–45.

Wellman, H. (1993). "Early Understanding of Mind: The Normal Case." In S. Baron-Cohen, H. Tager-Flusberg, C.J. Cohen, and D.J. Cohen, eds. *Understanding Other Minds: Perspectives from Autism* (pp. 10–39). New York: Oxford University Press.

Wimmer, H., and J. Perner (1983). "Beliefs about Beliefs: Representation and Constraining Function of Wrong Beliefs in Young Children's Understanding of Deception," *Cognition* **13**: 103–28.

Zahavi, D. (2001). "Beyond Empathy: Phenomenological Approaches to Intersubjectivity," *Journal of Consciousness Studies* **8(5–7)**: 151–67.

Zahavi, D., and J. Parnas (1999). "Phenomenal Consciousness and Self-Awareness: A Phenomenological Critique of Representational Theory." In S. Gallagher and J. Shear, eds., *Models of the Self* (253–70). Exeter: Imprint Academic.