Introduction

> "The relation of experience to time has not been profoundly studied. Its objects are given as being of the present, but the part of time referred to by the datum is a very different thing from the conterminous of the past and future which philosophy denotes by the name Present. The present to which the datum refers is really a part of the past — a recent past — delusively given as being a time that intervenes between the past and the future. Let it be named the *specious* present, and let the past, that is given as being the past, be known as the obvious past. All the notes of a bar of a song seem to the listener to be contained in the present. All the changes of place of a meteor seem to the beholder to be contained in the present. At the instant of the termination of such series, no part of the time measured by them seems to be a past."
>
> (William James, *The Principles of Psychology*, ascribed to E.R. Kelly)

Intuition has it that one's phenomenally experienced existence is confined to a point in time — the present moment, as it were, at least on those rare occasions when the mind manages to refrain from wandering in time and space to virtual worlds of its own making. Although the conception of an instantaneous present to which experience is confined is readily dispelled by the kind of analysis that prompted William James to call it "specious," it is still acceptable for a theory of consciousness to be summarily exempted from addressing the issue of the temporality of phenomenal experience. The chapters comprising this volume represent a collective attempt on the part of their authors to redress this metatheoretical anomaly, by exploring the implications of the idea that phenomenal experience is best understood — indeed, can only be understood — as being in time.

The exclusive theoretical focus on temporally punctate events, such as the vaunted instantaneous phenomenal percept, which is induced by the onset of an external stimulus and which "enters consciousness," after some delay, at a definite moment in time, is typically accompanied by a tacit assumption that our stream of consciousness (curiously enough, another Jamesian coinage) plays out like a movie. Movies, of course, consist of series of static snapshots that follow each other at short enough intervals so as to appear continuous. The analogy between conscious experience and a movie is, however, problematic, in that it puts off the temporal continuity issue rather than resolving it. Indeed, the seamlessness of motion picture perception does not inhere in the collection of static images that comprise it, but rather is the outcome of a laborious process that is not felt by the viewer as such only because his or her brain happens to be good at the kind of spatiotemporal interpolation that underlies the illusion of motion continuity.

Given that letting our experience be up to someone else would make no sense, a theory of experience must be *intrinsic* to the experiencer. In particular, it may not posit anything like the interpretation of the experiencer's brain activity by an external observer. Thus, a "cinematographic" theory of experience must include an explanation of how the brain juggles the dual tasks of (i) generating the series of frozen states corresponding to the phenomenal moments and (ii) processing the succession of its own states so as to imbue its experience with temporality. A radical alternative to this dual-task approach would be to identify experience with time-extended, continuous, dynamical brain processes to begin with, rather than assuming that continuity stems from self-observation. The four chapters in the first part of this book outline these competing theoretical approaches and examine their computational and neurobiological implications.

Part I. The dynamics of experience: possible brain mechanisms and computational considerations (Lloyd; Malach; Merker; Edelman & Fekete)

> "Time is the substance of which I am made. Time is a river that sweeps me along, but I am the river; it is a tiger that mangles me, but I am the tiger; it is a fire that consumes me, but I am the fire."
>
> (Jorge Luis Borges, *The New Refutation of Time*)

Our conscious experience results to a large extent from our brain's interaction with the environment. Both brain activity and the consciousness that it gives rise to change in time. Therefore, the relationship between changes in brain activity and changes in phenomenal experience are crucial for understanding and modeling the latter.

Lloyd, who advocates a dynamical systems approach to these issues, introduces it by the way of Husserl's phenomenology, which stresses both the fundamentally flowing aspects of experience and its invariant aspects, such as the persistence of objects and entities. The dynamical systems perspective emerges as a middle way between those phenomenological frameworks that neglect the stationary on the one hand, and on the other hand the traditional cognitive sciences that neglect the temporal.

Under this perspective, an experience is a region in neural activation space in which the trajectories associated with a type of experience reside. As an example, Lloyd looks for the signature of such phenomenal structure in fMRI data from a virtual reality experiment that consisted of interlaced driving, passive viewing, and fixating epochs. Using independent component analysis (ICA) and multidimensional scaling (MDS), he shows that brain activation states cluster by epoch type. As revealed by hierarchical cluster analysis, brain activation is structured at multiple scales: as the number of clusters increases, progressively finer temporal structure in activity manifests itself. At such finer scales, the analysis resolves the brain's coherent wandering through activation space, which is seen to slow down during fixations.

Multiscale analysis thus brings out both the stable brain states sought after by cognitive sciences and the dynamic ebb and flow of brain activation expected from phenomenology. Inspired by the tri-partite phenomenology of Husserl, according to which the present enfolds in it elements of the past and the future, Lloyd uses machine learning techniques to show that each state can indeed be used to predict following and preceding states. As one would expect, the accuracy of such prediction decreases, the further away one is from the reference "now" moment. A simulated "lesion analysis" suggests that this temporal structure is spatially distributed in the brain — again, as expected from a putative substrate of a fundamental facet of experience. Lloyd ends his chapter by noting that in a successful theory of consciousness higher order properties must be explicitly modeled and suggests that explicit representations of activation-space trajectories may underlie many such properties.

Malach's chapter proposes an account of phenomenal awareness in terms of the processing of ambiguous sensory inputs. Malach argues that the ambiguity is resolved (sometimes improperly, as in the case of visual illusions) through interaction between the input and information already stored in the cortical visual system. The process of ambiguity resolution is hierarchical, with most anterior regions of the cortex giving rise to the full holistic percept. The mechanism whereby the input interacts with the hierarchy of "templates" is identified with recurrent activation of small neural assemblies, or local neuronal "ignitions."

In support of this explanation, Malach cites experimental evidence of high firing rates, sustained activity, and dense local connectivity — three prerequisites of the ignition account that are indeed found in cortical circuits. A key prerequisite — evidence of a tie-in between conscious awareness and local reverberatory activity — is, however, still heavily debated. The local ignition hypothesis will, therefore, stand or fall, depending on whether or not such evidence solidifies.

The account of phenomenal awareness developed by **Merker** in his chapter also hinges on the observation that sensory information is inherently ambiguous and that, correspondingly, its processing by the brain must be probabilistic and must eventually resolve the uncertainty, given that typical perceptual states are unambiguous. This resolution process is posited to occur outside consciousness, with only its outcome becoming phenomenally available, in the form of the unambiguous global best estimate of the current scene as a whole. On Merker's account, the uncertainty resolution is orchestrated by subcortical circuits, which complement in this respect the inherently probabilistic modus operandi of the cortex.

The disambiguation process is complicated by the time constraints to which it is subject. The funneling of massive amounts of information into a unified flow of consciousness requires constant buffering. Only when this information is integrated into a holistic, dynamic, ego-centric model of body-world interactions does the mosaic of representations become conscious. The buffer thus serves to integrate the disambiguated stream of probabilistic data, creating the specious present. Merker shows that fitting candidates for this function are the higher order nuclei of the dorsal thalamus, specifically, the caudal reaches of the dorsal pulvinar.

Edelman and Fekete discuss some of the constraints that a computational model of experience must satisfy if it is to deal properly with time. Specifically, they focus on the interaction between the computational requirements of tractability and timeliness on the one hand and, on the other hand, the fundamental characteristic of phenomenality: its autonomy, that is, the requirement that putative theoretical constructs in their entirety be intrinsically computable by the brain itself, without recourse to an external observer. Edelman and Fekete argue that the common move in the sciences of the mind – equating mental content with the *instantaneous* state of the brain – is fundamentally flawed: due to physical constraints such as speed of communication and the computational intractability of asynchronous distributed consensus, an instantaneous state can only be a figment of the external observer's imagination.

In comparison, the *dynamics* of the system – that is, its trajectory through the space of its possible states – is no more and no less than what the system actually computes. Consequently, trajectories through a system's state space seem to be the only viable theoretical primitive for modeling experience (cf. Wiese and Metzinger, this volume). The theoretical benefits of this stance are immediate, and include accounting for the role of silent units in establishing concurrent experience, the ability to intrinsically (and counterfactually) define the confines of a single system, and of course the possibility of offering an intrinsic account of the time-infused essence of experience. In this light, it becomes clear that the popular notion of attractor dynamics as a building block of cognition (and of consciousness) is flawed: outside of strict stimulus-response scenarios within the confines of a laboratory it is simply not applicable. In contrast, a richer conception of dynamics, such as chaotic itinerancy, which uses properties of trajectory spaces as explanatory primitives, seems to be up to the task of explaining being in time.

Part II. The dynamics of experience: theoretical implications (Fekete & Edelman; van Leeuwen & Smit; Huette & Spivey)

> "Every instant is autonomous. Not vengeance nor pardon nor jails nor even oblivion can modify the invulnerable past. No less vain to my mind are hope and fear, for they always refer to future events, that is, to events which will not happen to us, who are the diminutive present. They tell me that the present, the "specious present" of the psychologists, lasts between several seconds and the smallest fraction of a second, which is also how long the history of the universe lasts. Or better, there is no such thing as "the life of a man," nor even "one night in his life." Each moment we live exists, not the imaginary sum of those moments."
>
> (Jorge Luis Borges, *A New Refutation of Time*)

Is every instant indeed autonomous, as Borges would have it? Is there not a bridging process that links together the instances into a chain that is the flow of consciousness? This part of the book looks at how the flow of time in phenomenal experience is addressed by several dynamical computational models. In these models, phenomenal awareness is modeled in terms of the system's state-space trajectories, with a special focus on the structure of this space.

Fekete and Edelman examine the ability of one class of such spaces, those that arise in digital computer systems, to give rise to phenomenal experience. They begin by analyzing the familiar digital replacement scenario — a thought experiment in which the brain is replaced, neuron by neuron, with digital chips. A careful, physiologically informed analysis of this scenario reveals that the purported argument for the equivalence of the resulting digital "brain" to the original – the principle of organizational invariance (OI) – falls short on two counts. First, it relies on an inadequate mode of abstraction, which only insists on preserving input/output relations when replacing a functional unit such as a neuron. Second, it depends critically on the notion of sufficient grain for modeling (in space and time) — an arbitrary move that is shown to result in inconsistency.

Fekete and Edelman next analyze the possibility of computer-simulated neural networks as a putatively sufficient substrate for experience. They show that if implementation details are properly considered, the claim that a digital computer realizes a particular dynamical system as specified by a given set of equations is essentially a matter of external attribution, thus failing the autonomy requirement of theories of experience. Among the reasons for this failure are the inability of finite discrete systems (as opposed to analog systems) to intrinsically represent numbers, and the lack of physical (and hence ontological and epistemological) distinction within a simulation between dynamical variables and causal interactions/forces that operate on those variables, as both are represented (instantiated) in the exact same format (bits). Further analysis suggests that the minimal requirement for a machine to be capable of realizing experience is to be an open, analog, time continuous dynamical system — unlike digital computers, which fail to capture the intrinsic temporality of consciousness.

Van Leeuwen and Smit's chapter puts forward the hypothesis that the mental processes occurring in the common situation where the perception is underdetermined by sensory data are strongly affected by mind wandering, which makes perception more dependent on past history and variable over time. This view, which may be compared to that of Lloyd (this volume), stands in contrast to models of perceptual processing that focus exclusively on invariance. Much like Merker and Malach (this volume) ,they see the importance for this process in giving the perceptual system flexibility. The posited dynamical process underlies consciousness but is not conscious in itself and therefore should be examined by psychophysics, not introspection.

To study this process, Van Leeuwen and Smit sought patterns in the dynamics of perceptions reported by subjects who viewed Necker cube stimuli, which they augmented with EEG data. In addition, they also looked at the subjects during rest, and they were able to identify patterns of synchrony between different parts of the brain, in the alpha frequency band of the EEG potential. When the observers were engaged in a task, these patterns were replaced by synchrony events in the beta and gamma bands. Van Leeuwen and Smit suggest that the alpha-band activity corresponds to non-conscious wandering that leads up dynamically to the emergence of experienced, phenomenal states.

Huette and Spivey argue that before consciousness can be understood, it must be modeled mathematically. In our everyday life we experience a stream of consciousness that is always an admixture of concepts, percepts, and emotions, all of which are highly interactive, content-sensitive and fleeting. That none of these are experienced purely in isolation suggests that consciousness should be modeled as a fuzzy dynamical system, in which experience is equated with the trajectory though a mental state space. This space is populated by entities such as concepts and percepts – regions with attracting properties. When experience – the system's trajectory – passes near an attractor, the corresponding concept becomes part of it, taking the form of a function of thousands of variables that span multiple time scales.

Huette and Spivey draw a distinction between consciousness, or being in time, and awareness, which consists of concepts and percepts. In the process of maturation, the brain learns to constrain the impact of the environment so as to form an increasingly clustered mental space (clusters being attracting regions), which results in increasingly structured trajectories – that is, phenomenal flow. In this fuzzy representation space, the most travelled paths between concepts form "cylindrical" sets – bundles of trajectories that over time acquire attracting properties of their own. The distance of the trajectory to the core of a concept is what determines awareness. In early stages of development, these cylinder sets, representing structured thought or action, are sharply delimited and distinct, and so is awareness. With time, the trajectories tend to "straighten out" as shortcuts result from striving for efficiency. Thus, the authors predict that consciousness, which is best understood as being in time, strengthens with life experience, as mental space becomes more clustered and cylindrical sets more well formed, while awareness peaks at a young age, then declines, as mental economy takes over.

Part III. The dynamics of experience: a philosophical perspective (Yoshimi; Wiese & Metzinger; Brown; Dale, Tollefsen & Kello)

> "No man ever steps in the same river twice, for it's not the same river and he's not the same man."
>
> (Plato, *Cratylus*, ascribed to Heraclitus)

The last part of the book offers several philosophical and methodological perspectives on the core questions that have been brought up so far. Does a model have to have a temporal aspect in order to accurately capture consciousness? Or perhaps a set of stimulus-response associations can describe consciousness reasonably well, at least in some cases?

Yoshimi sets out to illustrate how Husserl's theory of consciousness is naturally expressed in terms of dynamical system theory, and hence offers a rich source of ideas for the empirical study of experience. He suggests that the basic construct, from which all other Husserlian constructs can be derived, is that of a phenomenological state space — the space of all possible conscious states C. The stream of consciousness is then a succession of states in C. The question arises as to what might be the principles according to which C is organized. According to the mereology — calculus of parts and wholes — proposed by Husserl, C can be broken down into subspaces, corresponding, e.g., to aspects of experience such as vision and sound (cf. Wiese and Metzinger, this volume). Our experience of external reality then results from an inner model (horizon) of the world, in which our knowledge, our past (immediate and distant), and our dispositions coalesce to constrain ("constitute") possible phenomenal states.

This model, the horizon, is in fact a structure in C, which varies in its degrees of articulation: regions reflecting paucity of experience are less articulated, but can become more so with learning. These observations can be formalized in terms of a three-argument "expectation function," which maps a person's overall background knowledge, bodily movements, and current visual experience (say) into a probability distribution on V (a subset of C). This function determines the degree to which unfolding events feel surprising or familiar to us, and hence leads to elaboration of the horizon to the extent we are surprised. From a geometric standpoint, it can be thought of as "highlighting" a structure in C holding the possible experiences, given a context. Husserl tried to offer a more detailed analysis along these lines in terms of transcendental-eidetic "laws" – invariances that enable experience in various domains (e.g., possible color experiences, conversations, etc.). Such domains are structures in C, and can be probed using "free variation" to discover the rules that govern them, by varying an exemplar pertaining to the domain, so as to infer invariance. Yoshimi discusses several examples of this principle, such as the eidetic laws regarding the experience of movement, and describes his research in applying these notions to the study of embodied neural networks.

Wiese and Metzinger catalogue the analytical building blocks for a phenomenological theory of conscious experience. They explore the complex relationships between unity and temporal continuity found in experience, by the way of multiple constraint satisfaction across different levels and modes of analysis. The constraints derived from unity and temporal continuity involve mereotopological notions (e.g., parthood from mereology and openness from topology). The authors suggest that mereotopology is a fruitful theoretical approach that can inform empirical investigation, in particular enabling integration of data and theory at different levels. As a case study, they describe a critical analysis of the notion of "duration of experience." Finally, the authors propose a representational architecture that affords a theoretical framework for understanding the unity of consciousness. In this architecture, a representational system comprises subsystems representing objects, but also the representational activity of the other subsystems (and through that, their own activity). This dynamic coupling of local and meta representations leads to representational fusion, which underlies the global unity of experience. Wiese and Metzinger argue that this globally coherent temporally extended state corresponds to the phenomenon of the specious present.

Brown begins by noting that several types of consciousness have been identified by research: creature consciousness (a creature being awake and responding to stimuli), transitive consciousness (being conscious of things and events), state consciousness, access consciousness, and phenomenal consciousness. He suggests that creature consciousness can be understood in terms of the AIM model (Hobson et al., 2000), according to which the global chemical state of the brain (levels of neuromodulation) realizes a low dimensional state space in which deep sleep, dreaming, and wakefulness occupy distinct regions (as do mood disorders, e.g., depression). On this view, creature consciousness is a state of the brain, while a conscious state, being a particular, is a brain state.

Brown then goes on to argue that all other forms of consciousness (brain states) can be reduced to transitive consciousness, which is understood in terms of synchronized neural activity, with high frequency synchrony being essential to the brains representational vocabulary (cf. malach this volume), and low frequency synchrony serving as a carrier to disseminate information between brain regions (cf. van Leeuwen and Smit, this volume). He claims that experimental evidence supports the Higher-Order Representation Of a Representation (HOROR) theory of phenomenal consciousness, which shares similarities with both lower and higher order theories. As evidence for this idea, Brown cites the impact of the acquisition of new concepts on experience (e.g., in wine tasting), suggesting that all phenomenal content derives from concepts and categories. Another line of supporting evidence comes from experiments indicating that in detection tasks subjects report unattended stimuli as more visible than attended ones with the same first-order representational strength (as quantified by signal detection sensitivity). This leads to a claim that the higher order representation in itself suffices for experience, resulting in a kind of physical identity theory. Brown dismisses some of the familiar arguments against this move based on conceivability (e.g., zombie arguments), by applying them to physical properties, which turn out to be equally susceptible to such an attack. Thus, as it is synchrony that individuates states of consciousness (states of the brain), and as synchrony is a temporally extended property, experience is seen as inherently being in time.

In the closing chapter, **Dale, Tollefsen, and Kello** review and try to integrate current consciousness research and theory. They identify several distinct avenues through which the problem has been approached and argue that as consciousness is a complex phenomenon, it will require a pluralistic explanation. In particular, the multiscale property of phenomenal experience allows numerous theories to have each an explanatory role, depending on one's meaning of "experience," and on the measurement grain size chosen for the analysis. The authors argue that such a multiscale account of experience is best articulated through the language of dynamics and complexity science. They suggest that explanatory coherence will be achieved though investigation of the temporally extended facets of experience expressed in its sustained nature, which is nevertheless constituted by interacting parts (content) at multiple scales in space and time. Moreover, they argue that the patterns of complexity found at one level – such as that of neuronal dynamics – cannot explain consciousness, as it recurs in various unrelated physical phenomena. Rather, it is the unique multiscale organization of the dynamics introduced by coupling with our physical and social environs, resulting from problem solving at multiple scales, which is constitutive of consciousness. Thus, they suggest that a promising way of moving forward is through dynamical network modeling that explicitly allows for a hierarchy of spatiotemporal levels.

Conclusion

The contributions to this volume differ in their theoretical outlook and in their methodology, which ranges from neuroanatomy, electrophysiology, and functional imaging, through computational analysis and behavioral experimentation, to philosophical analysis. As such, they offer a wide variety of perspectives on the issues about which the authors are in agreement: the dynamical nature of phenomenal experience, and the need to understand and model it as such. Given the profundity and the difficulty of the fundamental questions that arise in consciousness studies, their eventual resolution will require a concerted, interdisciplinary effort on the part of the entire community of researchers. We hope that this book will help generate the right kind of dynamics in this most exciting field of philosophical and scientific endeavor, and look forward to the phenomenal experience of watching this dynamics play out.

Shimon Edelman

Tomer Fekete

Neta Zach

February 2012