Review

Attractors and noise: Twin drivers of decisions and multistability

Jochen Braun a,⁎, Maurizio Mattia b,c

a Cognitive Biology Lab, University of Magdeburg, Germany
b Department of Tecnologies and Health, Istituto Superiore di Sanita, Rome, Italy
c Sezione di Roma 1 Sanita, Istituto Nazionale di Fisica Nucleare, Rome, Italy

Abstract

Perceptual decisions are made not only during goal-directed behavior such as choice tasks, but also occur spontaneously while multistable stimuli are being viewed. In both contexts, the formation of a perceptual decision is best captured by noisy attractor dynamics. Noise-driven attractor transitions can accommodate a wide range of timescales and a hierarchical arrangement with "nested attractors" harbors even more dynamical possibilities. The attractor framework seems particularly promising for understanding higher-level mental states that combine heterogeneous information from a distributed set of brain areas.

© 2010 Elsevier Inc. All rights reserved.

Contents

Introduction ................................................................. 740
Spontaneous activity ............................................................ 741
Optical imaging ............................................................ 741
Magnetic-resonance imaging ............................................... 742
Avalanches ............................................................... 742
Summary .............................................................. 742
Perceptual decisions ........................................................... 742
Two choice tasks ............................................................ 742
Neural correlates ............................................................ 743
Attractor models ............................................................ 743
Summary .............................................................. 744
Multistable perception ........................................................... 744
Dynamical characteristics .................................................. 745
Neural correlates ............................................................ 745
Computational accounts ..................................................... 746
Nested attractors ............................................................ 746
Summary .............................................................. 746
Conclusions ............................................................... 747
Acknowledgments ............................................................. 748
References ................................................................. 748

Introduction

Brain activity is nothing if not dynamic. At whatever scale of volume or time one cares to examine it, brain tissue ceaselessly produces waves, bursts, oscillations, sudden transitions, spindles, fluctuations, transients, and many other dynamic patterns of activity.

⁎ Corresponding author.
E-mail addresses: jochen.braun@ovgu.de (J. Braun), mattia@iss.infn.it (M. Mattia).
URL: http://kobi.nat.uni-magdeburg.de (J. Braun).

1053-8119/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
So it seems mildly paradoxical that some aspects of brain function may depend on the existence of decidedly “undynamic” states, namely, on stable patterns of reverberating activity that sustain and support themselves, at least for some time, against the relentless onslaught from the rest of the brain. Starting from early seminal intuitions (De Noo, 1938; Hebb, 1949), reverberating patterns of activity, also called “attractor states,” have been considered as a possible mechanism for various cognitive processes and functions, among them working memory (Zipser et al., 1993; Amit and Brunel, 1997; Amit and Mongillo, 2003; Del Giudice et al., 2003), recall of long-term memory (Hopfield, 1982; Amit, 1995; Hasselmo and McCalland, 1999), attentional selection (Deco and Rolls, 2005a), rule-based choice behavior (Vasilaki et al., 2009; Fusi et al., 2007) and, most recently, the formation of perceptual states (Wong et al., 2007; Furman and Wang, 2008).

Of course, “attractor states” are a theoretical notion, not an empirical finding. When neuronal activity is described at an appropriate level of abstraction, simulations of populations of spiking neurons capture the collective dynamics that is generated by recurrent interactions between such populations. The existence of attractor states is revealed when a reduced version of the spiking network is analyzed with so-called mean-field techniques (Amit and Brunel, 1997; Brunel and Wang, 2001; Renart et al., 2003). These methods are borrowed from statistical physics and, when applied to networks of formal and spiking neuron models (Hopfield, 1982; Amit and Brunel, 1997), identify sets of average activity levels at which the various interactions between populations of neurons exactly balance each other and thus create a collective steady-state. The charm of this approach is that the properties of these models, and the conditions needed to support such dynamical regimes, can be tested at very different levels of experimental analysis: the biophysical parameters of neurons and synapses, the spiking activity of single neurons and of cell assemblies, the aggregate metabolic demand of neural tissue, the timeevolution of cognitive processes and, indeed, the animal’s behavior (Deco et al., 2009).

What would attractor states “look like” in the brain? Their stability is guaranteed only for idealized networks with infinitely many neurons. In the brain, where neuron numbers are finite, spontaneous activity fluctuations would destabilize and, sooner or later, overthrow any self-sustaining pattern of activity. Accordingly, an attractor state should remain stable up to the time-scale of cognitive processes and should terminate due to spontaneous activity fluctuations. In addition, neuronal populations participating in an attractor state should exhibit stereotypical activity levels so that the trial-to-trial variability should remain stable up to the time-scale of cognitive processes and, indeed, the animal’s behavior (Deco et al., 2009).

Our perceptions and actions vary slightly even under identical conditions. This reflects the fact that brain activity fluctuates independently of external factors. At the level of individual neurons, the precise timing of spikes varies because of channel noise and variability in the mechanisms of synaptic transmissions (Shadlen and Newsome, 1998; Faisal et al., 2008). In addition, in neuronal populations of finite size, the variability of individual spike times will result in substantial activity fluctuations at the population level (Brunel and Hakim, 1999; Spiridon and Gerstner, 1999; Mattia and Del Giudice, 2002). This provides a “noisy background” in the synaptic input received by each neuron which further varies spike timing. Of course, this background activity may not be truly random and may well conceal meaningful structure (Stevens and Zador, 1998; Stuart and Hauser, 2001; Shu et al., 2003).

Optical imaging

In visual cortex of cats, the spontaneous activity revealed by optical imaging exhibits an intriguing spatial and temporal structure, switching between patterns of activity that resemble the activity patterns evoked by stimulation (Arieli et al., 1996; Slovin et al., 2002; Kenet et al., 2003), electrical recordings (Tsodyks et al., 1999; Petersen et al., 2003; Fiser et al., 2004; Freyer et al., 2009), calcium-imaging of individual neurons (Ikegaya et al., 2004; Maclean et al., 2005; Luczak et al., 2009), and magnetic-resonance imaging (Biswal et al., 1995; Fox et al., 2005). Because it is encountered in anaesthetized and in alert (but resting) animals as well as in brain slices, spontaneous activity is thought to reflect intrinsic properties of neural populations and networks.
spontaneous activity and thalamic stimulation elicit precise spatio-temporal patterns of neuronal spiking that are almost indistinguishable (Ikegaya et al., 2004; MacLean et al., 2005; Luczak et al., 2009). On the basis of these observations, it has been suggested that the internal connectivity of sensory cortices prescribes a fixed repertoire of attractor states and that spontaneous activity comprises serial transitions between these stereotypical states (Singh et al., 2008; Ringach, 2009).

**Magnetic-resonance imaging**

At larger scales and lower frequencies, the resting brain exhibits correlated fluctuations in the blood-oxygen-level dependent (BOLD) signal (Biswal et al., 1995; Fox et al., 2005; Fox and Raichle, 2007), which in turn may reflect fluctuations in the local field-potential generated by neuronal activity (Logothetis and Wandell, 2004; Shmuel et al., 2006; Nir et al., 2007; Goense and Logothetis, 2008; Sirotin and Das, 2009; Angenstein et al., 2009). Typically, slow fluctuations are correlated in groups of areas that are functionally related, such as bilateral motor cortex, visual cortex, oculomotor areas, attention-related areas, areas related to language, episodic memory areas, and others (Vincent et al., 2006; Fox et al., 2006; Vincent et al., 2007; Moeller et al., 2009). Interestingly, the pattern of BOLD signal increases and decreases, which is typically associated with certain cognitive tasks, is mirrored in a pattern of positive and negative correlations in slow BOLD fluctuations (Laufs et al., 2003; Fox et al., 2005). The areas negatively correlated with task performance are thought to form a “default network” (Raichle et al., 2001; Buckner et al., 2008). One study has compared slow fluctuations with anatomical connectivity, finding that pairs of correlated areas are often (though not always) linked by direct anatomical projections (Vincent et al., 2007). This would seem to imply that either mono- or polysynaptic projections may underlie correlated fluctuations.

**Avalanches**

A particularly intriguing aspect of spontaneous activity are “neuronal avalanches” (Plenz and Thiagarajan, 2007; Pajevic and Plenz, 2009). Synchronous firing in upper cortical layers, which negatively reflects the local field potential, propagates to other cortical sites, triggering a chain reaction of such synchronization episodes. With the help of multielectrode arrays, the statistics of such “avalanches” has been studied in the cortex of rodents and monkeys, both in culture and in vivo (Beggs and Plenz, 2003, 2004; Stewart and Plenz, 2008; Gireesh and Plenz, 2008; Petermann et al., 2009). To a good approximation, the probability of an avalanche of a given size and duration follows a power law distribution, with an exponent of -3/2 for size and of -2 for duration. This scale invariance suggests that avalanches exhibit a fractal organization and that avalanche patterns are maximally diverse in size and duration. Avalanche dynamics appears to be at least partly deterministic, as complex avalanche patterns recur with “millisecond precision” (Beggs and Plenz, 2004; Stewart and Plenz, 2008). A possible theoretical explanation for the empirically observed power law exponents is that cortical dynamics corresponds to a “critical phase transition,” at which spatiotemporal patterns are maximally diverse. There are a number of ways in which critical dynamics could be maintained, among them short-term plasticity and small-world connectivity (Bienenstock and Lehmann, 1998; Levina et al., 2007; Teramae and Fukai, 2007; Abbott and Rohrkemper, 2007).

**Summary**

The nature and origin of spontaneous activity in the brain is only partially understood. However, there is converging evidence that spontaneous activity is not random but consists in part of correlated patterns of activity. Similar correlated patterns are observed in response to sensory stimulation in the context of mental activity. As has been pointed out by several authors, these findings are highly suggestive of attractor states (Plenz and Thiagarajan, 2007; Ringach, 2009). Particularly intriguing is the finding that correlated patterns occur at different spatial and temporal scales, which would be consistent with the notion of “nested attractors” (see below). The most detailed observations in this regard concern patterns of synchronization events that deflect the local field potential. However, it remains to be seen how this particular type of activity relates to other forms of spontaneous activity in the brain.

**Perceptual decisions**

It has long been apparent that perceptual performance is probabilistic. When observers try to distinguish between sensory events, they do not succeed or fail consistently. Instead, they succeed with a probability that increases with the physical difference between the events. This probabilistic performance is thought to reflect the presence of “internal noise,” which forms a basic ingredient of quantitative models of perceptual decisions (Green and Swets, 1966). A systematic analysis of response times reveals further particulars about the probabilistic processes that lead to perceptual decisions (Smith and Ratcliff, 2004; Bogacz, 2007). Typically, response times lengthen with perceptual difficulty and are distributed with a positive skew. In addition, response times increase when the observer tries to decide with a higher degree of confidence (“speed-accuracy trade-off”). These observations suggest that noisy perceptual information accumulates gradually until it reaches a criterion level, at which point a response can be initiated with the desired degree of confidence. Indeed, integration-to-bound of a noisy sensory signal affords an excellent phenomenological account of behavioral observations such as response time distributions and their dependence on task variables (Ratcliff and McKoon, 2008; Ratcliff and StarNS, 2009).

Over the last several decades, single-unit recording studies in monkeys have identified several brain structures where neuronal activity relates to perceptual decision making. This has revealed much about the nature of sensory representations and sensory working memory, and about the role of motor structures in perceptual decisions. In addition, this neuropsychological work has spurred theoretical studies that have formulated plausible mechanistic accounts of perceptual decision making. Our brief summary of this experimental and theoretical work will focus on two series of groundbreaking studies of visual and somatosensory perception in monkeys. The principal findings of these studies with monkeys are likely to apply equally to perceptual decision making by humans (Heekeren et al., 2008).

**Two choice tasks**

To trace the neural correlates of visual decision making, Newsome, Shadlen and colleagues have trained macaques to discriminate the predominant direction-of-motion of a field of moving dots (Newsome et al., 1989; Gold and Shadlen, 2007). While viewing the dot field, the animals maintained eye fixation, but afterwards shifted their eyes to an appropriate response target, thereby indicating which direction-of-motion they had perceived. The perceptual difficulty of the task depended on how many moved in the predominant direction (as opposed to random directions). Regardless of difficulty, however, every perceptual decision culminated in a stereotypical motor action (i.e., an eye movement).

Pursuing a similar research programme for somatosensory decisions, Romero and colleagues have trained macaques to discriminate vibrotactile frequency (Romo et al., 2000; Romo and Salinas, 2003). In this behavioral paradigm, the animals sensed two sequential mechanical vibrations with a restrained hand, before indicating which
vibration they had perceived as the higher frequency by pressing an appropriate button with their free hand. Perceptual difficulty depended on the difference between the two stimulus frequencies. Note that this task also engaged working memory, as the animals had to compare the remembered frequency of the first stimulus with the perceived frequency of the second. Once again, every perceptual decision was expressed through a stereotypical action (i.e., a button press).

**Neural correlates**

In both the visual and the somatosensory domain, it was possible to link neuronal activity in sensory cortical areas to the perceptual decision reached by the animal. For example, when the activity of single neurons in visual area MT was recorded while the monkey performed the direction-of-motion discrimination described above, this activity was as informative about the physical stimulus as the monkey's response (Newsome et al., 1989; Britten et al., 1992). When neurons in area MT were stimulated electrically while the monkey performs this task, the animal was more likely to report the direction-of-motion to which the stimulated neurons were tuned, suggesting that the activity of these neurons were causally linked to the monkey's perception of visual motion (Salzman et al., 1990, 1992). Similarly, the activity of single neurons in somatosensory cortex (S1) correlated trial-by-trial with the monkeys' discrimination of vibrotactile frequency (Salinas et al., 2000). When one of the two sequential mechanical vibrations was replaced by direct microstimulation of S1 neurons, the animals continued to perform normally, as if the direct brain stimulation had elicited a comparable perceptual experience to skin vibration (Romo et al., 1998, 2000). Thus there can be little doubt that activity in visual or somatosensory cortex directly informs the corresponding perceptual decisions.

In both behavioral paradigms in question, the animals withheld their motor response until sufficient sensory information had accumulated or sensory stimulation had ceased. Presumably, the animals' decision matured during this period of temporary inactivity, which was terminated by a stereotyped motor response. Indeed, it proved possible in both paradigms to trace the process of "decision formation" to neuronal activity of motor planning areas.

For example, visual decisions were evident in areas known to be involved in the planning and execution of eye movements, most particularly in the lateral intraparietal cortex (LIP), but also in dorsolateral prefrontal cortex (Kim and Shadlen, 1999), frontal eye fields (FEF) (Hanes and Schall, 1996) and superior colliculus (SC) (Thevarajah et al., 2009). While a decision was being formed, activity increased ramp-like in LIP neurons selective for the eye movement that was eventually executed (Roitman and Shadlen, 2002). The rate of increase depended on the quality of visual evidence, becoming steeper for "easier" discriminations. When the visual evidence was transiently enhanced or degraded, the activity of LIP neurons appeared to reflect the time integral of visual information (Huk and Shadlen, 2005). That this LIP activity was linked causally to decision formation was confirmed by microstimulation (Hanks et al., 2006). The accumulation of visual information in LIP clearly concerned neurons involved in saccade planning: when a conflicting eye movement was evoked by microstimulation of frontal eye fields (FEF), the evoked movement was deflected in the direction of the visual decision (Gold and Shadlen, 2000, 2003). Even when the animals maintained the decision in working memory before performing an eye movement, some neurons in superior colliculus (SC) continued to reflect the perceptual decision (Horwitz et al., 2004).

Animals tended to initiate the eye movement when LIP activity reached a certain criterion level, suggesting that this level marked the completion of the decision process (Roitman and Shadlen, 2002; Huk and Shadlen, 2005). This was true even when animals were permitted to respond prematurely (Kiani et al., 2008). Interestingly, the criterion level appeared to remain similar when task difficulty was varied by degrading the visual evidence or by introducing additional response choices (Huk and Shadlen, 2005; Churchland et al., 2008). Criterion activity remained constant also in the FEF in a countermanding context, in which the animal decided to suppress a response (Hanes and Schall, 1996). A fascinating hint as to what LIP activity may contribute to the animals' subjective experience is offered by the observation that LIP activity correlates trial-by-trial with the animals' subjective confidence in a particular visual decision, as indexed by a post-decision wagering behavior (Kiani and Shadlen, 2009).

In the case of the vibrotactile frequency discrimination, the decision emerged in the activity of areas involved in planning hand and finger movements. During the delay period following the first mechanical vibration, neuronal activity in inferior prefrontal, medial premotor, and ventral premotor cortex represented the frequency of this remembered stimulus (Romo et al., 1999; Hernandez et al., 2002; Romo et al., 2004). With the onset of the second mechanical vibration, the activity initially reflected in a graded fashion the difference between remembered and current frequencies, but then appeared to approach one of two categorical levels corresponding to the binary decision outcome. This comparison and decision process was also evident, albeit to a lesser degree, in somatosensory cortex S2 (Romo et al., 2002). In primary motor cortex, neuronal activity presaged the decision outcome, but did not seem to participate in the comparison of remembered and current stimuli (Romo et al., 2004). In summary, the decision process during this task reached its culmination almost simultaneously in distributed network of sensory and premotor areas.

To better distinguish sensory and decision processing, Romo and colleagues modified the behavioral task and trained animals to report the presence or absence of a single mechanical vibration of varying intensity (de Lafuente and Romo, 2005, 2006). Whereas neural activity in somatosensory area S1 followed vibration intensity in a graded fashion, activity in medial premotor cortex reflected the categorical decision (presence or absence) reached in each trial, largely independently of stimulus intensity. Comparing decision-related activity across eight cortical areas, the authors could show that covariance with stimulus intensity decreased, and covariance with decision outcome increased, gradually from sensory, to premotor, to motor areas (de Lafuente and Romo, 2006). In premotor areas, decision-related activity increased steadily during stimulus presentation, consistent with a gradual accumulation of evidence, and trial-by-trial fluctuations predicted the decision outcome. Interestingly, the decision-related activity in premotor areas remained largely unaffected when the animal was not required to produce a response.

**Attractor models**

What computational machinery could underlie this behavioral and neurophysiological evidence? A plausible decision circuit must combine several features. Firstly, it must categorize perceptual information by choosing a stereotypical response from a discrete number of options. Such winner-take-all behavior implies the reduction of a high-dimensional stimulus space to the task-relevant dimensions. Secondly, a decision circuit must accumulate sensory evidence over time and must allow decision speed to be traded for decision accuracy. Finally, it must feature some intrinsic variability such as to respond to weak evidence in a stochastic manner and to account for psychometric functions.

Attractor models of spiking neuron networks fulfill the above requirements and reproduce much of the available psychophysical and neurophysiological experimental evidence (Wang, 2002; Deco et al., 2007; Wang, 2008), as shown in Fig. 1. In these models, separate pools of excitatory neurons drive each response and interact competitively via inhibitory interneurons (Amit and Brunel, 1997). When recurrent couplings are strong, local reverberation can amplify
activity in any pool that is slightly favored by sensory input. As its activity grows, the activity of other pools is suppressed, ensuring a single winner. A criterion level can be set by providing additional, symmetric input to all pools.

The attractor decision is well described by the energy landscape metaphor. Stimulus onset destabilizes the low-activity state and forces the system to choose between two high-activity decision states. The ball starts on a plateau between two valleys and rolls in random directions, propelled by intrinsic fluctuations. On a broad and gently sloping plateau, the ball’s noisy trajectory gradually reveals the input bias (high criterion). On a narrow plateau, the outcome is determined largely by fluctuations (low criterion). Strength of sensory input is reflected in the incline of the plateau. An unambiguous input creates a steep slope and ensures a fast deterministic outcome (Deco et al., 2007; Wang, 2008). Interestingly, even a weak stimulus, which fails to destabilize the resting state, can result in a decision, provided there is sufficient time for an escape driven by activity fluctuations (Martí et al., 2008; Durstewitz and Deco, 2008).

The attractor framework is general and can readily be extended to more complex situations, such as decisions involving more than two choices (Furman and Wang, 2008; Albantakis and Deco, 2009). Even the comparison of two sequential stimuli (i.e., vibrotactile frequency discrimination) can be achieved. To this end, Machens et al. (2005) introduced a multi-modular network forming a “line attractor,” which could hold a first stimulus in working memory. When the second stimulus appeared, the energy landscape changed into bistable configuration, in which the system was driven by the difference between the two stimuli (Brody et al., 2003).

Clearly, the attractor framework closely parallels phenomenological models (“integration-to-bound”) of perceptual decisions (Usher and McClelland, 2001; Ratcliff and McKoon, 2008). Indeed, the attractor account is formally equivalent to a unidimensional nonlinear diffusion with threshold (Roxin and Ledberg, 2008), provided the dynamic regime remains close to the critical point at which the resting state is destabilized.

Summary

In summary, the remarkable body of work on perceptual decisions seamlessly integrates behavioral evidence, single-unit activity, and theory of spiking neuron populations. Interacting neural populations in cortical areas related to motor planning exhibit distinct steady-states of activity that correspond to alternative decision outcomes. The integration of noisy sensory information is accomplished by a stochastic transition to one of these attractor states. Importantly, the decision-related activity is observed even when the response is delayed or suppressed (de Lafuente and Romo, 2006). It should be kept in mind, however, that the evidence for attractor-based decisions draws on behavioral situations in which each perceptual choice is associated consistently with a specific motor response. When animals are taught a response rule that varies from trial to trial, decision-related activity is no longer evident in motor planning areas (Gold and Shadlen, 2003).

Multistable perception

With many displays, prolonged viewing does not produce a stable visual experience but provokes from time to time a discrete change in appearance. This phenomenon is called bistable or multistable perception, depending on whether two or more alternative appearances are observed. Well-known examples are the Necker cube, the perception of depth-from-motion, or binocular rivalry (At奈eve, 1971; Leopold and Logothetis, 1999). Multistable phenomena are not restricted to the visual domain and occur also in auditory and tactile perception (Pressnitzer and Hupe, 2006; Carter et al., 2008).

Multistability is related to perceptual illusions and the inferential nature of perception (Kersten et al., 2004): it occurs when the visual system relies on prior knowledge of the natural environment and of image formation to either infer additional object properties that are not present in the retinal image, or to suppress retinal information that is difficult to reconcile with prior knowledge. For example, the visual system infers an illusionary (but compelling) appearance of depth from certain patterns of shading, line junctions, or motion flow. When the situation is ambiguous and admits two interpretations, a bistable appearance of depth results. An even more dramatic form of multistability occurs with “unnatural” scenes that are unlikely to be encountered in the physical world. In these cases, the implausible information intermittently disappears from awareness, resulting in binocular or monocular rivalry (Campbell and Howell, 1972; Bonneh et al., 2001; Blake and Logothetis, 2002).
In what sense can multistable perception be compared to a decision process? In a perceptual decision, sensory information is classified into distinct categories and the outcome is probabilistic. In the case of a multistable display, the possible outcomes (i.e., appearances) similarly fall into distinct categories and are probabilistic. This is true both when a display is first presented and when it is viewed continuously so that its appearance changes from time to time. Typically, the transition between appearances is marked by a brief episode of indeterminate, mixed, or intermediate appearance. Transitions are probabilistic in the sense that every possible pairing “from appearance X to appearance Y” is typically observed, although not all pairings with equal frequency (Suzuki and Grabowecky, 2002).

Even “unsuccessful” transitions (“from appearance X to appearance X”) do occur (Brascamp et al., 2006).

**Dynamical characteristics**

The main difference to the acute perceptual decisions discussed in the previous section is that, in the case of multistable phenomena, stimulation is continuous and constant and that the decision is periodically renewed. Unsurprisingly, the appearance after each transition is influenced strongly by the preceding appearances. The strongest influence of this kind is visual adaptation (Blake, 1989; Petersik, 2002; Blake et al., 2003; Kanai and Verstraten, 2005; Pearson and Clifford, 2005), which biases a renewed decision against earlier appearances and ensures that “unsuccessful” transitions are relatively infrequent. However, in those (rare) cases in which the alternative appearances are adapted to a comparable degree, “unsuccessful” and “successful” transitions occur comparably often (Pastukhov et al., 2009). Another, weaker but longer-lasting influence is a visual facilitation (Brascamp et al., 2008; Pastukhov and Braun, 2008), which biases a renewed decision in favour of earlier appearances. For example, when the stimulus presentation is periodically interrupted (in order to permit recovery from adaptation), the long-lasting facilitation allows the current appearance to persist across the interruption (Leopold et al., 2002; Maier et al., 2003). This facilitation stabilizes perception considerably, slowing or even arresting changes of appearance for intermittently presented displays.

Multistable perception lacks a characteristic timescale in that average duration of one appearance (“dominance time”) ranges at least from seconds to minutes. Average dominance times vary both between individual observers (Aafjes et al., 1966; Medith, 1967) and between different bistable displays (Hupe and Rubin, 2004; Van Ee, 2005; Brascamp et al., 2005; Sheppard and Pettigrew, 2006). Even for the same observer and same display, average dominance times vary substantially with stimulus intensity (Bossink et al., 1993; Brascamp et al., 2006), with attention (Meng and Tong, 2004; Mitchell et al., 2004; Paffen et al., 2006; Pastukhov and Braun, 2007), and when a display is periodically interrupted (Orbach et al., 1963; Leopold et al., 2002; Maier et al., 2003). In some cases, the average dominance time experienced by a given observer on a given display may differ by two orders of magnitude for different stimulus regimes (Pastukhov and Braun, 2007).

Another outstanding characteristic of multistable perception is the highly irregular timing of changes in appearance: individual dominance times scatter widely around their average value and the distribution of dominance times typically approximates a Gamma function (Levelt, 1967; Borsellino et al., 1972; Murata et al., 2003; Brascamp et al., 2005). Independently of the average dominance time, many multistable displays exhibit Gamma functions of comparable shape. Specifically, the coefficient of variation lies generally between 0.4 and 0.6, as shown in Fig. 2. This striking correlation between the mean value and the variance of dominance times constitutes a “scalar property” and suggests that both measures have the same mechanistic origin (Gigante et al., 2009).

**Neural correlates**

The neural basis of multistable visual perception has been reviewed very recently (Sterzer et al., 2009). In the case of binocular rivalry, fMRI studies with human observers show that BOLD signals in primary visual cortex closely mirror both temporal and spatial shifts in subjective appearance (Polonsky et al., 2000; Tong and Engel, 2001; Lee et al., 2005). Even BOLD signals in the lateral geniculate nucleus reflect fluctuations of awareness (Haynes et al., 2005; Wunderlich et al., 2005). In primates, these BOLD signals were shown to correlate more closely with local field potentials than with spiking activity (Wilke et al., 2006; Maier et al., 2008), explaining the discrepancy to earlier neurophysiological studies (Leopold and Logothetis, 1999). In extrastriate visual cortex, too, the amplitude of neural activity as measured electrophysiologically in monkeys or with functional imaging in humans, follows the reversals of subjective appearance (Blake and Logothetis, 2002; Tong et al., 2006). Interestingly, more subtle patterns of activity continue to carry information about the stimuli that are suppressed from awareness (Fang and He, 2005; Sterzer et al., 2008). The privileged representation of visual information that dominates awareness in higher visual cortex has been confirmed also with other types of multistable displays. For example, changes in the appearance of bistable motion are reflected in the activity of motion-selective visual areas (Dodd et al., 2001; Castelo-Branco et al., 2002; Muckli et al., 2002; Brouwer and van Ee, 2007) and episodes of motion-induced blindness (a type of monocular rivalry between stationary and moving stimulus elements) are associated with antagonistic increases and decreases of activity in cortical areas responsive to stationary or moving stimuli, respectively (Donner et al., 2008). Studies in which multistable displays are presented intermittently further strengthen the link between subjective appearance and activity in higher visual areas (Sterzer et al., 2008; Hesselmann et al., 2008): in these areas, baseline activity prior to stimulus onset is predictive of stimulus appearance after the onset.

Particularly interesting from a dynamical systems perspective is the fact that changes in the appearance of a multistable display are associated with transient increases in BOLD activity (Sterzer et al.,
These transient increases are observed not only in visual cortex (Sterzer et al., 2002, 2003), but also in the parietal and frontal areas that have also been associated with visual attention and working memory (Lumer et al., 1998; Sterzer and Kleinschmidt, 2007). Whether the transient activation of parietal and frontal areas is a cause or a consequence of the changing appearance remains unclear.

Tentative support for a causal contribution is furnished by reports that parietal/frontal activity transients precede both the change in appearance and the activity transients in visual cortex (Sterzer and Kleinschmidt, 2007; Britz et al., 2009). This should not be taken to imply, however, that attentional mechanisms are responsible for changes in appearance: voluntary control over multistable perception is typically poor (Van Ee, 2005; Hancock and Andrews, 2007) and multistable displays change their appearance even when attention is tightly controlled (Pastukhov and Braun, 2007; Lee et al., 2007).

**Computational accounts**

Several observations about multistable phenomena are suggestive of fluctuation-driven transitions between attractor states. For example, experiments with quadra-stable displays show that the combined appearance of different images parts is often “trapped” in certain preferred configurations (Suzuki and Grabowecky, 2002). Direct evidence for an attractor mechanism comes from a demonstration of stochastic resonance in binocular rivalry (Kim et al., 2006). In this study, antagonistic contrast modulation of the rivaling stimuli combines with spontaneous activity fluctuations to produce discrete “resonance peaks” in the distribution of dominance times. This allows the amplitude of the spontaneous fluctuations to be estimated as the equivalent of a 30% modulation of stimulus contrast. A related finding is that perceptual learning accelerates rivalrous alternations of appearance approximately three-fold (Suzuki and Grabowecky, 2007). This suggests that the neural plasticity induced by training smooths and simplifies the energy landscape, leaving merely two distinct but shallow wells.

Various computational accounts have been proposed for multistable phenomena. While some have been phenomenological and others have detailed populations of spiking neurons, many of these models propose essentially a noisy oscillator circuit (Lehky, 1988; Mueller, 1990; Kalarickal and Marshall, 2000; Laing and Chow, 2002; Wilson, 2003, 2007; Freeman, 2005; Shpiro et al., 2009). Just as in perceptual decision models, mutual inhibition ensures that one of the alternative stimulus representations emerges dominant. A habituating process such as spike-frequency adaptation (Schwindt et al., 1989; Sanchez-Vives et al., 2000) or short-term synaptic depression (Abbott et al., 1997; Tsodyks and Markram, 1997) curtails this dominance and forces a perceptual reversal. Finally, intrinsic noise introduces some variability to the dominance times (Lehky, 1995; Brascamp et al., 2006). Oscillator models predict a characteristic dominance time that is set by the habituation.

Attractor dynamics offers an alternative mechanism for multistability. Both in phenomenological models (Kim et al., 2006; Shpiro et al., 2009) and in detailed networks of formal (Riani and Simonotto, 1994) and spiking (Salinas, 2003; Moreno-Bote et al., 2007) neurons, each stable appearance represents an attractor state of the system dynamics: a valley in an energy landscape. Subjective appearance changes when the barrier between two attractors is crossed. Spontaneous activity fluctuations provide the random “kicks” that are needed to escape from an attractor state. By varying network parameters, the strength of the “kicks” and the height of the barriers can be adjusted such as to obtain very low escape rates (Miller and Wang, 2006; Marti et al., 2008), as is well known from “noise-driven transitions” in multistable systems (Hanggi et al., 1990). In contrast to oscillator models, this provides a natural way for obtaining a range of dominance periods from the same set of neuronal time-constants (Salinas, 2003). However, purely noise-driven transitions predict exponential distributions of dominance times (Moreno-Bote et al., 2007; Marti et al., 2008). To obtain the Gamma-like distributions that characterize multistable phenomena, the attractor framework can be extended by a habituating process (Moreno-Bote et al., 2007; Shpiro et al., 2009). In this framework, habituation destabilizes the dominant state by raising its valley in the energy landscape. Interestingly, the empirical dominance time statistics is reproduced best when the adaptation strength is such that the system operates just outside the boundary of the oscillatory regime (Shpiro et al., 2009).

**Nested attractors**

Although attractor networks with habituation account adequately for individual instances of multistable dynamics, this framework is too simplistic to capture the dynamics under more general conditions. The “scaling” of the variability with the mean of the dominance times (mentioned above) presents a particular difficulty. Similar “scalar properties” have been observed in other cognitive tasks involving long timescales such as judgments of interval duration (Gibbon et al., 1997). In the latter case, the scalar statistics of interval judgments can be reproduced by pools of stochastic switches that accumulate over time until a threshold is reached (Okamoto and Fukai, 2001; Okamoto et al., 2007; Mattia, 2008; Almeida and Ledberg, 2009). Interestingly, such a “stochastic integration” can also explain the Gamma distribution of dominance periods (Lehky, 1988). Recently, the attractor framework and stochastic integration have been combined into a more general account of multistable perception that explains the scalar property of dominance statistics (Gigante et al., 2009). In this approach, perception escapes from an attractor when sufficiently many “flips” have occurred in an underlying pool of switches. In the energy landscape metaphor, the valley floor now consists of discrete wells and the monotonic climbing of a barrier is replaced by multiple discrete hops between such wells. This type of dynamics has been termed “chaotic itinerary” (Tsuda, 2001; Durstewitz and Deco, 2008). If each stochastic switch is implemented as a bistable network (as described above), then the perceptual state is represented by an attractor of attractors, as illustrated in Fig. 3.

A “nested attractor” would explain not only the scalar property, but would also offer a plausible mechanism for another puzzling aspect of multistable dynamics. As mentioned, the intermittent presentation of a multistable display stabilizes appearance (Orbach et al., 1963; Leopold et al., 2002; Maier et al., 2003), revealing a long-lasting facilitatory effect of earlier perceptual states (Brascamp et al., 2008; Pastukhov and Braun, 2008). Noisy oscillators address this phenomenon by postulating additional habituating processes at longer timescales (Noest et al., 2007; Brascamp et al., 2008; Wilson, 2007). A hierarchy of stochastic integrators operating at different timescales is a promising alternative (Gigante et al., 2009). In this scenario, the top-level attractor that represents the dominant appearance has nesting within a variety of lower-level attractors integrating both sensory input (on faster time-scales) and dominant appearance (on slower time-scales). The appearance-driven attractor pools provide a memory trace which biases perceptual transitions in a history-dependent manner. Thus, subjective appearance is based on a distributed representation and reflects both active sensory pools and active memory traces.

**Summary**

In summary, multistable perception represents a sequence of stochastic perceptual decisions that are variously biased by the current sensory input and by the outcome of earlier decisions. Scalar properties are a prominent feature of this multistable dynamics. Transitions between one appearance and another are reflected in the activity of those visual cortical areas that represent the visual attributes involved. No “appearance-related” activity is evident in
premotor areas that would compare to the “decision-related” activity in perceptual choice tasks. However, “transition-related” activity does occur outside visual cortex in certain parietal and frontal areas, which are generally associated with attentional and working-memory functions. Many aspects of multistable dynamics are captured by the kind of noisy attractor transitions that have also been proposed for perceptual choice tasks. However, the scalar properties and history-dependence of multistable dynamics suggest a less simplistic “nested attractor” picture, which could be the hallmark of a highly distributed representation.

Conclusions

The neural mechanisms underlying perceptual decisions can be studied profitably with at least two paradigms. With perceptual choice tasks, the flow of sensory information and the formation of a perceptual decision can be traced in exquisite detail (i.e., in the activity of single neurons) through several cortical stages, which combine sensory representations, working-memory and decision representations, and representations of intended motor actions to varying proportions. The great strength of this approach lies in the fact that it allows decision-related activity to be identified and localized to premotor areas. The price for obtaining this neuroanatomical focus of decision-related activity is that the perceptual choice must be consistently associated with a stereotypical motor response. When this association is broken, the decision-related activity presumably becomes distributed and correspondingly more difficult to identify.

With multistable perception, the dominant perceptual appearance correlates with neural activity in those areas of visual cortex that represent the visual attributes in question, at least as judged by magnetic-resonance imaging. In addition, transitions of perceptual appearance are also associated with transient activations in parietal and frontal areas. Perhaps due to its distributed nature, it has proven more difficult to study multistable perception at the level of the activity of single neurons, although recent developments in this respect are extremely encouraging (Wilke et al., 2006; Maier et al., 2008; Leopold et al., 2008). Certainly the flow of information and the causal chain of events is not nearly as well understood as for perceptual choice tasks. Arguably, the ongoing nature of multistable perception offers an advantage in allowing the stochastic dynamics of the decision process to be characterized in great detail. Few other situations come to mind in which the collective dynamics of distributed neural representations is so readily accessible.

In both paradigms, the dynamics of perceptual decisions is best captured by a noisy attractor framework. However, the appropriate mental picture is not a static, double-well energy landscape but rather a sequence of temporary energy landscapes that are constantly overturned by external or internal events (e.g., stimulus onsets, transitions of perceptual appearance) (Deco and Rolls, 2005b; Rigotti et al., 2007). Together with spontaneous activity fluctuations, this results in a complex dynamics of “attractor hopping,” activity transitions through “attractor ruins,” dynamical sequences of metastable activity states, and so on (Mangillo et al., 2003; Destexhe and Contreras, 2006; Durstewitz and Deco, 2008; Deco et al., 2009).

Fig. 3. A nested attractor performing competing stochastic integrations of sensory input, as proposed for bistable perception (Gigante et al., 2009). (a) A pool of bistable stochastic switches (In) represent different sensory inputs (small blobs in gray cloud). Their summed activity $\sum_{\text{in}}$ biases the relative stability of the two global attractor states in the output module (Out). In turn, Out activity feeds back to the input switches, in each case favoring one of the local attractor states. The energy landscape (large blob at bottom) shows two global basins of attraction, each corrugated by several local basins. (b) Example of stochastic dynamics of In (reddish traces) and Out (black trace) attractors, both implemented as in Moreno-Bote et al. (2007). Steady-state activities are close to $-1$ and $+1$. Before a global (Out) transition (vertical dashed lines) can occur, all local (In) switches must have changed state. Equivalently, a global transition occurs whenever the sum $\sum_{\text{in}}$ of local activities reaches a maximum or a minimum (gray trace). (c) Dynamic trajectory of a 500 s simulation in the state space cross-section ($\sum_{\text{in}}, \text{Out}$). The marginal distributions show that dynamics is bistable from a global and multi-stable from a local perspective.
Importantly, the time-scale of attractor transitions may span several orders of magnitude even with identical neuronal components (Miller and Wang, 2006; Marti et al., 2008), allowing the nervous system to match a wide range of environmental processes.

The most appropriate framework for capturing the dynamics of the distributed neural representations underlying multistable perception may be a “nested attractor” (Gigante et al., 2009). In this scenario, a top-level transition is possible only after a sufficient number of lower-level transitions have accumulated. Such a stochastic integration process would explain in a natural way the “scalar property” that characterizes multistable perception and other types of perceptual judgments involving long timescales (Gibbon et al., 1997). Note that the lower-level attractor pools need not be homogeneous. Instead, they might comprise both faster-switching sensory representations and slower-switching representations of memory traces, of anticipat-ed stimulation, of behavioral goals, and so on. In consequence, the top-level transition of perceptual appearance would come to reflect external and internal events represented in a widely distributed set of brain areas.

The coherent picture that emerges is that both goal-directed and spontaneous decisions in perception may be actuated by the same hierarchical and distributed machinery, as summarized in Fig. 4. It is tempting to speculate that these insights into perceptual decision making may reflect a general computational strategy adopted by the brain. A hierarchy of nested attractors would provide a natural substrate for all manner of spatial and/or tempo-erarchies in the neural responses evoked by spatiotemporal sensory patterns (Kiebel et al., 2008). For instance, when an observer watches a silent movie, longer movie segments evoke more reliable responses in higher cortical areas, suggesting a hierarchical representation of temporal structure (Hasson et al., 2008). Similar evidence for hierarchical organization has been reported for acoustic stimulation (Overath et al., 2008). Furthermore, it has been suggested that increasingly abstract behavioral goals are associated with increasingly longer timescales, which in prefrontal cortex appear to be represented in a rostrocaudal gradient (Koechlin and Summerfield, 2007; Badre and D’Esposito, 2009). Accordingly, it seems quite possible that distributed dynamical representations based on noise-driven, nested attractors may be operating in the context of a wide range of cognitive processes. A critical prediction of this point of view is that spontaneous activity should exhibit recurring patterns that are nested over a wide range of spatial and temporal scales, possibly even in a fractal manner. While the available evidence on spontaneous activity is certainly suggestive in this regard, it remains far from conclusive (Plenz and Thiagarajan, 2007). Accordingly, we eagerly await further studies of spontaneous activity that will simultaneously cover multiple spatial and temporal scales.

**Acknowledgments**

The authors thank Paolo Del Giudice, Stefano Fusi, and Guido Gigante for many stimulating discussions. J.B. is supported by the BMBF Bernstein Network of Computational Neuroscience.

**References**


