alone does not define what the functional role of this activity is, but rather that it interacts with other brain regions to control human behavior in the presence of self-related processing [6]. Prior research has shown that activity in the vmPFC upregulates activity in the left superior temporal sulcus, which is thought to reflect social attention. In particular, the strengths of neural coupling between the two regions can predict the biased responses to self-related stimuli in perceptual matching [7]. In parallel, neuropsychological studies have shown that brain damage over the vmPFC abolishes self-biases in memory, whereas brain damage in the dorsal attentional control network that spares the vmPFC causes abnormally large biases toward self-related information due to an exaggerated effect of strong attentional signals [8]. The evidence indicates that self-reference acts as a golden thread linking the vmPFC to other regions to create a specific neural circuit in the brain to support self-reference.

Lane and colleagues [2] propose that recording task-irrelevant resting state and pre-stimulus activity would yield insights into how the self performs this function by identifying the trajectory of the self with little experimental control. Such approaches do contribute to assessing the continuity of consciousness of the self, a topic that has been extensively discussed in cognitive neuroscience [9] as well as philosophy [3]. However, these methods alone are unlikely to pinpoint specific cognitive and neural responses associated with self-reference. By contrast, empirical manipulation provides a way to link self-processing to specific functional or neural processes [10]. For example, the function of self-reference in perception may be targeted by manipulating the luminance of stimuli while controlling other factors. Researchers reported that reduction of stimulus luminance was less detrimental to perceptual sensitivity of shapes

associated with the self than shapes associated with others, suggesting that self-reference can enhance perception [11]. In our view, using various approaches (e.g., empirical manipulation, neural analysis during resting state) within a study would provide us with a better understanding of the functions of self-reference that are intrinsic to perception and cognition over time [12].

[1]Department of Experimental Psychology, University of Oxford, Oxford OX1 3UD, UK

*Correspondence: jie.sui@psy.ox.ac.uk (J. Sui).

References

1. Sui, J. and Humphreys, G.W. (2015) The integrative self: how self-reference integrates perception and memory. *Trends Cogn. Sci.* 19, 719–728
2. Lane, T. et al. The trajectory of self. Trends Cogn. Sci.(in press) http://dx.doi.org/10.1016/j.tics.2016.03.004
3. Dennett, D. (2003) The self as a responding-and-responsible-artifact. *Ann. N. Y. Acad. Sci.* 1001, 39–50
4. Sui, J. et al. (2015) Super-capacity and violations of race independence for self- but not for reward-associated stimuli. *J. Exp. Psychol. Hum. Percept. Perform.* 41, 441–452
5. Northoff, G. et al. (2006) Self-referential processing in our brain–a meta-analysis of imaging studies on the self. *Neuroimage* 31, 440–457
6. Humphreys, G.W. and Sui, J. (2016) Attentional control and the self: the Self-Attention Network (SAN). *Cogn. Neurosci.* 7, 5–17
7. Sui, J. et al. (2013) Coupling social attention to the self forms a network for personal significance. *Proc. Natl Acad. Sci. U. S. A.* 110, 7607–7612
8. Sui, J. et al. (2015) Dissociating hyper- and hypo-self biases to a core self-representation. *Cortex* 70, 202–212
9. Mason, M.F. et al. (2010) Wandering minds: the default network and stimulus-independent thought. *Science* 315, 393–395
10. Cummingham, S.J. et al. (2008) Yours or mine? Ownership and memory. *Conscious. Cogn.* 17, 312–318
11. Sui, J. et al. (2012) Perceptual effects of social salience: evidence from self-prioritization effects on perceptual matching. *J. Exp. Psychol. Hum. Percept. Perform.* 38, 1105–1117
12. Northoff, G. et al. (2010) Rest–stimulus interaction in the brain: a review. *Trends Neurosci.* 33, 277–284

Spotlight

The Importance of Single-Trial Analyses in Cognitive Neuroscience

Mark Stokes[1],* and Eelke Spaak[1]

Theories of working memory typically assume that information is maintained via persistent neural activity. By contrast, Lundqvist *et al.* report that single-trial delay activity is actually 'bursty'; the classic profile of persistent activity is an artefact of trial-wise averaging. Tackling brain–behaviour relationships at the single-trial level is an important future direction for cognitive neuroscience.

Classic models of working memory assume that information is stored via persistent neural activity [1]. Since the earliest neurophysiological experiments in awake, behaving primates [2], researchers have reported evidence that working memory is maintained via persistent delay activity in the prefrontal cortex (PFC). The idea is relatively simple: task-relevant mental representations are kept 'on line' by maintaining corresponding patterns of neural activity for as long as required.

As in most neurophysiological studies, however, classic evidence for persistent activity was essentially based on the results of many individual trials averaged together to form a putative 'representative' estimate of neural activity. Averaging across trials is important to improve the signal-to-noise ratio. If we can assume that each trial is a noisy sample of the true distribution, the average over trials should represent the signal we could expect from
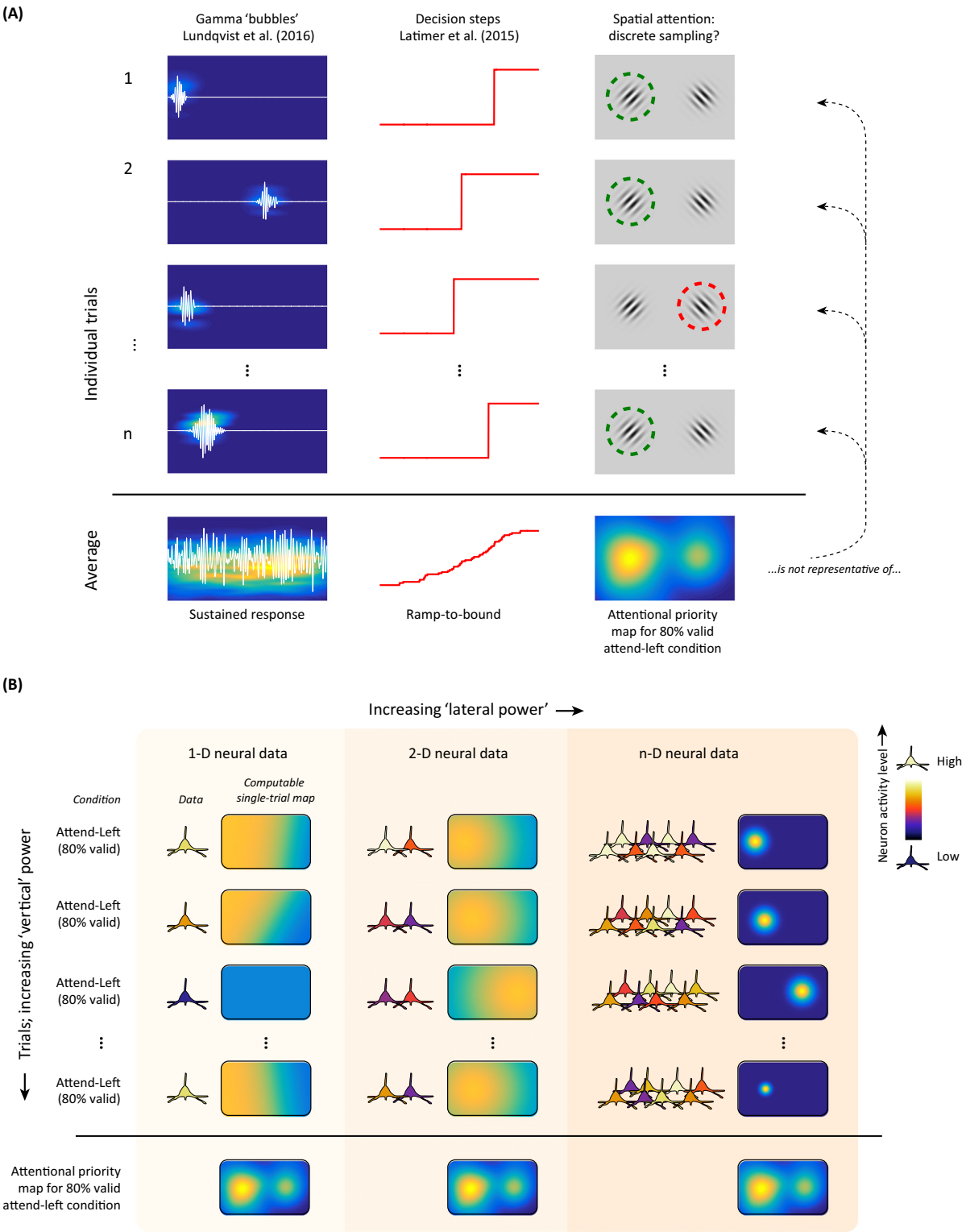
(A)

Gamma 'bubbles'
Lundqvist et al. (2016)

Decision steps
Latimer et al. (2015)

Spatial attention:
discrete sampling?

Individual trials

1

2

⋮

n

Average

Sustained response

Ramp-to-bound

Attentional priority
map for 80% valid
attend-left condition

...is not representative of...

(B)

Increasing 'lateral power' ⟶

1-D neural data

2-D neural data

n-D neural data

Neuron activity level ⟶

High

Low

Trials; increasing 'vertical' power

Condition

Data

Computable
single-trial map

Attend-Left
(80% valid)

Attend-Left
(80% valid)

Attend-Left
(80% valid)

⋮

Attend-Left
(80% valid)

Attentional priority
map for 80% valid
attend-left condition

Trends in Cognitive Sciences

Figure 1. Unscrambling the Egg: Advances in Cognitive Neuroscience Will Critically Depend on Better Methods to Characterise Single-Trial Neural Dynamics. (A) Combining single trials through computing the mean does not necessarily result in a representative average. Left: Gamma-frequency 'bubbles' [3] (time–frequency representations of power, with traces superimposed) in the prefrontal cortex during individual trials of working memory maintenance activity. The average shows a familiar

(Figure legend continued on the bottom of the next page.)

a single noiseless trial. However, the brain does not operate according to this construct of the 'average response'. Real-world interaction requires real-time perception, encoding, and decision making. To understand the neural basis of behaviour, including working memory, we need to understand the neural dynamics that unfold within a single trial.

In a recent paper, Lunqvist *et al.* [3] developed a novel method for characterising trial-wise neural dynamics during working memory. Specifically, they were interested in whether previous evidence for sustained high-frequency (45–100 Hz, gamma band) activity during a working memory delay, as recorded from local field potentials (LPFs) in the primate PFC, is evident at the single trial or whether the trial-wise dynamics are qualitatively different from the average response. To test this, they developed a 'burstiness' metric to quantify the temporal structure of gamma activity on single trials; that is, before averaging over trials for statistical analyses. Critically, by averaging over this second-order characteristic (i.e., burstiness), they provide new evidence that the unbroken persistent activity apparent in the trial-averaged representation comprises intermittent bursts of gamma activity (Figure 1A, left panel). The authors conclude that this pattern of stochastic bursting is not consistent with standard models of working memory that posit an unbroken chain of persistent firing (see also [4]) but instead favours 'activity-silent' models in which memories are stored in hidden neural states, such as rapid short-term synaptic plasticity [5].

Although these finding are clearly important for models of working memory [6], the

trial-wise approach also has important broader implications for cognitive neuroscience. To properly understand brain–behaviour relationships, we need to consider the neural dynamics as they unfold on a single trial. Trial averaging can misrepresent neural dynamics in many ways. For example, a recent study showed that even stimulus-driven activity in the gamma band is intermittent (or 'bursty') on single trials, despite the robust sustained profile apparent in the trial average [7]. Similarly, in the decision-making literature, single-trial analyses have revealed that activity in parietal neurons displays discrete transitions on single trials rather than the gradual accumulation of decision-related information observed in the trial-average response [8] (Figure 1A, middle panel). These types of approaches could also be used to address numerous longstanding debates, such as graded versus discrete allocation of attention (Figure 1A, right panel).

Exploring single-trial dynamics is not trivial. Any measure of neural activity is subject to noise; therefore, we need multiple samples for any analysis to have sufficient statistical power. We typically achieve statistical power by combining information over trials (which one might call 'vertical power'), but this can easily smooth over important heterogeneity. In [3], the key development was to develop first-level metrics (burstiness) that can be averaged across trials for statistical inference without losing the essential structure of interest. However, this approach critically depends on an *a priori* model of the single-trial phenomenon (e.g., burstiness in [3], discrete state transitions in [8]). Future methodological developments will allow much more general approaches to single-trial analyses.

For example, increasing the number of simultaneous measurements (which we propose calling 'lateral power') will be crucial in more detailed examinations of single-trial dynamics (Figure 1B). Boosting the number of samples within a single trial will provide qualitatively new insights into the neural dynamics that underpin real-time behaviour.

The prospects of rich single-trial analyses are especially exciting with the ever-expanding lateral power in neurophysiology (e.g., multielectrode arrays, calcium imaging) and with multichannel methods with high temporal resolution in humans, such as electrocorticography (ECoG) in patients and magnetoencephalography (MEG) in healthy participants. Multivariate decoding techniques from the field of machine learning provide particularly powerful approaches to leverage the lateral power of high-dimensional neural recordings. These allow increasingly detailed characterisation of single-trial information, crucial not only for our understanding of real-time cognition but also for the development of robust brain–computer interfaces. It is becoming increasingly evident that neural activity is inherently high dimensional, especially in the PFC, where mixed selectivity expands the potential coding space for flexible cognition [9]. High lateral power allows us to sample this dimensionality in real-time, rather than infer coding diversity from the 'representative trial' derived from multiple repetitions of potentially heterogeneous events.

To conclude, a basic assumption in many empirical sciences is that averaging over repeated observations allows us to combine equivalent signals while cancelling

sustained gamma response but qualitatively misrepresents the single-trial dynamics. Middle: Neurons in the parietal cortex display discrete steps reflecting the time of sensory decisions [8]. The average response shows a classic ramp-to-bound process for the decision. Right: When subjects are cued to attend to the left visual field with 80% validity (i.e., a target will appear at the cued location on 80% of trials), spatial attention might be distributed in a continuous fashion throughout the visual field (as in the average, bottom), but such an average response is also consistent with individual trials showing a discrete sampling of visual space (attend left on 80% of trials and right on the other 20%). (B) Typically, statistical power refers to adding repetitions (i.e., trials) over which to average data ('vertical' power, reflected in the vertical dimension here). One can instead think of 'lateral' power: adding more measurements per experimental trial (horizontal dimension). Increasing lateral power is essential for characterising the neural dynamics on a single-trial level. To illustrate, several hypothetical trials of the cued-attention task depicted in (A) (right column) are shown in (B) as a function of increasing lateral power. Neuron colour indicates activity level.

out random noise. However, not all trial-wise variation is noise. In many circumstances, trial-wise averaging will also cancel out important signals. Consider Galton's [10] bean machine (or 'quincunx'): a marble falls down a board with many pins and with each hit of a pin the marble has a 50% chance of falling to either side of that pin. At the bottom of the board are several buckets, one of which will catch the marble at the end of its fall. By the central limit theorem, the distribution of marbles in the buckets will tend towards a Gaussian distribution. While this is a useful and important characterisation of the data, it does not help us understand how any particular marble ended up in its particular bucket; the history of the individual marble is lost in the average. For the next big step forward in cognitive neuroscience, we need to focus on the equivalent of individual marbles hitting individual pins: real-time single-trial dynamics.

[1]University of Oxford, Oxford, UK

*Correspondence: mark.stokes@ohba.ox.ac.uk (M. Stokes).

References

1. Sreenivasan, K.K. et al. (2014) Revisiting the role of persistent neural activity during working memory. *Trends Cogn. Sci.* 18, 82–89
2. Fuster, J.M. and Alexander, G.E. (1971) Neuron activity related to short-term memory. *Science* 173, 652–654
3. Lundqvist, M. et al. (2016) Gamma and beta bursts underlie working memory. *Neuron* 90, 152–164
4. Shafi, M. et al. (2007) Variability in neuronal activity in primate cortex during working memory tasks. *Neuroscience* 146, 1082–1108
5. Sandberg, A. et al. (2003) A working memory model based on fast Hebbian learning. *Network* 14, 789–802
6. Stokes, M.G. (2015) Activity-silent" working memory in prefrontal cortex: a dynamic coding framework. *Trends Cogn. Sci.* 19, 394–405
7. Lowet, E. et al. (2016) Areas V1 and V2 show microsaccade-related 3-4-Hz covariation in gamma power and frequency. *Eur. J. Neurosci.* 43, 1286–1296
8. Latimer, K.W. et al. (2015) Single-trial spike trains in parietal cortex reveal discrete steps during decision-making. *Science* 349, 184–187
9. Rigotti, M. et al. (2013) The importance of mixed selectivity in complex cognitive tasks. *Nature* 497, 585–590
10. Galton, F. (1894) *Natural Inheritance,* pp. 63–65, Macmillan

Forum
Conjuring Deceptions: Fooling the Eye or Fooling the Mind?

Vebjørn Ekroll[1,@,*] and Johan Wagemans[1,@]

Currently, we see the contours of a new research program emerging, where cognitive scientists study what magicians do and why it works. This research program may aid us in formulating interesting questions about central aspects of human experience and in gaining new perspectives on the relation between perception and cognition.

The study of illusions as a window into the mechanisms underlying perception and cognition is one of the main methodological tools of cognitive science. Thus, the stunning illusions produced by professional magicians are of great interest to cognitive scientists [1–3]. Perhaps even more importantly, however, thinking about magic from the practical perspective of the magician may reveal theoretically interesting questions that are otherwise less obvious.

Magicians are sometimes referred to as illusionists, but their ultimate aim is to design miracles, not mere illusions [4]. That is, the magician's first question is how they can create the illusion of impossibility [3,4]. Relatedly, the magician's second question is how they can make sure that nobody is able to figure out how it was done. That is, they are essentially aiming to construct a problem that is as difficult to solve as possible, given the fundamental principles of human problem solving [5].

These two simple practical questions are probably not frequently pondered in academic cognitive science, but they may turn out to be of considerable heuristic value.

The Formation of Mutually Contradictory Beliefs

Experiencing something as impossible implies that one must hold two beliefs that contradict each other. Furthermore, if either of them is not experienced as absolutely certain, the impossibility is ruined. Thus, the magician's first question immediately taps into the question of what are the sources of strong beliefs. Furthermore, since two contradicting beliefs implies that one of them must be wrong, the magician's first question raises the question of how false, yet strong, convictions are best created.

Cognitively Impenetrable Illusions

Being impervious to conscious reasoning (cognitively impenetrable) is considered to be a major hallmark of perceptual illusions [6,7]. Hence, it seems natural to think of perceptual illusions as the ideal answer to the magician's second question.

Yet, in current discussions of the factors underlying the art of conjuring, the classical perceptual illusions have a relatively minor role [1,3]. How can this be reconciled with the above reasoning? We believe that the contradiction is only apparent: magicians do in fact rely on perceptual illusions to a considerable extent, but our traditional scientific notions about what should count as a perceptual illusion are too limited for two reasons. First, some of the achievements of the perceptual system, such as inferences about causality, intentions, reality, or occluded scene regions, are so staggering that it is difficult to envision that they are not due to more 'intelligent' high-level cognition. Second, the introspectively obvious fact that we consciously reason about these things suggests that there is no role left for the perceptual system in making inferences about them.