Decoding the Dynamics of Conscious Perception: The Temporal Generalization Method

Stanislas Dehaene and Jean-Rémi King

Abstract Parsing a cognitive task into a sequence of successive operations is a central problem in cognitive neuroscience. A major advance is now possible thanks to the application of pattern classifiers to time-resolved recordings of brain activity [electro-encephalography (EEG), magneto-encephalography (MEG), or intracranial recordings]. The method determines precisely when a specific mental content becomes explicitly represented in brain activity. Most importantly, the ability of these pattern classifiers to generalize across time and experimental conditions sheds light on the temporal organization of information-processing stages. To illustrate these ideas, we show how the decoding of MEG and EEG recordings can be used to track the fate of conscious and unconscious brain processes during simple masking and auditory novelty tasks. The experimental results yield converging results, suggesting that conscious perception is associated with the late formation of a distributed and stable neural assembly that encodes the content of subjective perception.

Parsing a cognitive task into a sequence of successive operations has been recognized as a central problem ever since the inception of scientific psychology. The Dutch ophthalmologist Franciscus Donders first used mental chronometry to demonstrate that mental operations are slow and can be decomposed into a series of successive stages (Donders 1969). Since then, psychologists have proposed a variety of elegant but indirect methods by which such decomposition could be achieved using behavioral measurements of response times (Pashler 1994; Posner 1978; Sigman and Dehaene 2005; Sternberg 1969, 2001).

The American psychologist and cognitive neuroscientist Michael Posner was among the first to realize that the advent of brain imaging methods provided direct evidence of this classical task-decomposition problem, and he successfully

S. Dehaene (✉)
Collège de France, Paris, France

INSERM-CEA Cognitive Neuroimaging Unit, NeuroSpin Center, Saclay, France
e-mail: stanislas.dehaene@gmail.com

J.-R. King
INSERM-CEA Cognitive Neuroimaging Unit, NeuroSpin Center, Saclay, France

analyzed several tasks such as reading or attention orienting into their component operations (Petersen et al. 1988; Posner and Raichle 1994). Time-resolved methods that capture brain activity at the scale of milliseconds, such as electro- and magneto-encephalography (EEG and MEG) or intracranial recordings, seem particularly well suited to this task-decomposition problem, because they can reveal how the brain activity unfolds over time in different brain areas, each potentially associated with a specific neural code. Yet the amount and the complexity of electrophysiological recordings can rapidly become overwhelming. In particular, it remains difficult to accurately reconstruct the spatial sources of EEG and MEG signals. As a result, the series of operations underlying basic cognitive tasks remain ill-defined in most cases.

Machine learning techniques, combined with high-temporal-resolution brain imaging methods, now provide a new tool with which to address this question. In this chapter, we briefly review a technique that we call the "temporal generalization method" (King and Dehaene 2014), which clarifies how multiple processing stages and their corresponding neural codes unfold over time. We illustrate this method with several examples, and we use them to draw some conclusions about the dynamics of conscious processing.

The Temporal Generalization Method

Contemporary brain imaging techniques such as EEG and MEG typically allow us to simultaneously record a large number of electrophysiological signals from the healthy human brain (e.g., 256 sensors in EEG and 306 sensors in MEG). Similarly, using intracranial electrodes in monkeys or in human patients suffering from epilepsy, hundreds of electrophysiological signals can be acquired at rates of 1 kHz or above. Identifying, from such multidimensional signals, the neuronal representations and computations explicitly recruited at each processing stage can be particularly difficult. For example, reconstructing the neural source of EEG and MEG signals—i.e., determining precisely where in the brain the signals originate— remains a major hurdle. Signals from multiple areas are often superimposed in the recordings from a given sensor and, conversely, the signal from a given brain area simultaneously projects onto multiple sensors.

Machine learning techniques can help overcome these difficulties (Fig. 1). The idea is to provide a time slice of electrophysiological signals to a machine-learning algorithm that learns to extract, from this raw signal, information about a specific aspect of the stimulus. For instance, one can ask the algorithm to look for information about whether the visual stimulus was a vertical or a horizontal bar, whether a sound was rare or frequent, whether the subject responded with the right or the left hand, etc. If we train one such classifier for each time point t (or for a time window centered on time t), we obtain a series of classifiers whose performance traces a curve that tells us how accurately the corresponding parameter can be decoded at each moment in time. Typically, this curve remains at chance level before the onset of the stimulus, then quickly rises, and finally decays (Fig. 1).
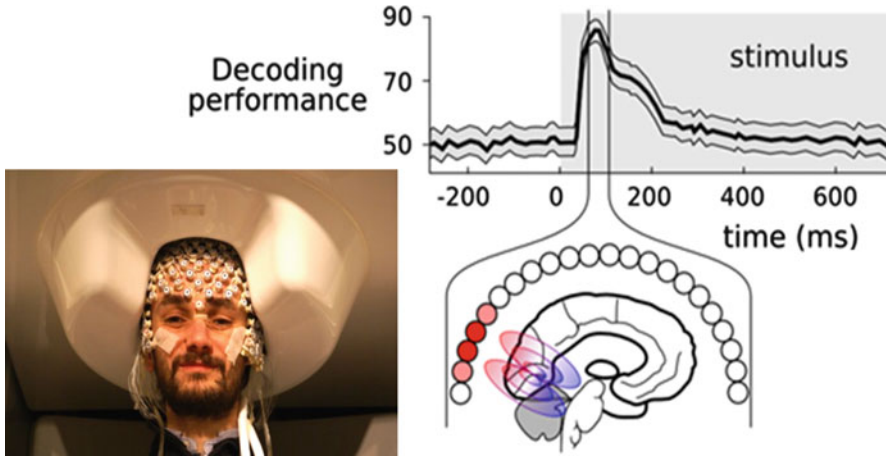
Fig. 1 Principle of temporal decoding (from King 2014). On each trial we simultaneously recorded a large number of brain signals (e.g., 256 EEG and 306 MEG signals). Using the data from a single time point t, or from a time window centered on time t, we could train a Support Vector Machine (SVM) to decode one aspect of the stimulus (for instance, the orientation of a grid on the subject's retina). The time course of decoding performance reveals the dynamics with which the information is represented in the brain. How a decoder trained at time t generalizes to data from another time point, t', reveals whether the neural code changes over time

The decoding curves tracking distinct features of the current trial typically rise and fall at different times, thus providing precious indications about when, and in which order, the respective representations begin to be explicitly coded in brain activity. For example, Fig. 2 illustrates how we decoded the time course of perceptual, motor, intentional and meta-cognitive error-detection processes from the very same MEG/EEG signal (Charles et al. 2014; for another application to the stages of invariant visual recognition, see Isik et al. 2014).

In addition to tackling the *when* question, machine learning may also tell us for *how long* a given neural code is activated and whether it recurs over time. To this aim, we asked how a pattern classifier trained at time t generalizes to data from another time point t'. This approach results in a temporal generalization matrix that contains a vast amount of detail about the dynamics of neural codes (King and Dehaene 2014). If the same neural code is active at times t and t', then a classifier trained at time t should generalize to the other time, t'. If, however, the information is passed on to a series of successive stages, each with its own coding scheme, then such generalization across time should fail, and classifiers trained at different time points will be distinct from each other. More generally, the shape of the temporal generalization matrix, which encodes the success in training at time t and testing at time t' for all combinations of t and t', can provide a considerable amount of information about the time course of coding stages. For instance, it can reveal whether and when a given neural code recurs, how long it lasts, and whether its scalp topography reverses or oscillates. When comparing two experimental conditions C and C', it can also reveal whether and when the series of unfolding stages

Decoder for:
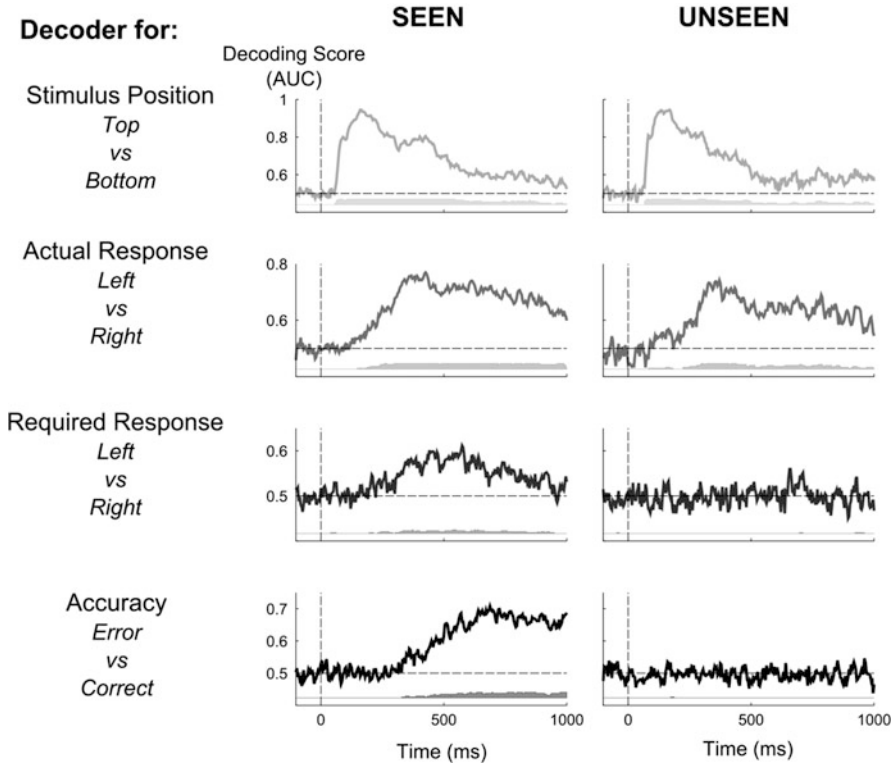
Fig. 2 Example of temporal decoding (from Charles et al. 2014). Distinct decoders were trained to extract four different properties of an unfolding trial from the same MEG and EEG signals: the position of a visual target on screen, the motor response made by the subject, the response that he should have made, and whether the response made was correct or erroneous. Note how those four distinct properties successively emerge in brain signals, from *left* to *right*. The target was masked, such that subjects occasionally reported it as "unseen" (*right column*). In this case, stimulus position and motor response could be decoded, but the brain seemed to fail to record either the required response or the accuracy of the motor response

was delayed, interrupted or reorganized (for detailed discussion, see King and Dehaene 2014).

Advantages of Multivariate Decoding Methods

Temporal decoding and temporal generalization are powerful multivariate methods that present several advantages over classical univariate methods for the characterization of brain activity:

– Within each subject, machine learning methods search for an optimal combination of brain signals that reflects a certain psychological variable. By combining

multiple sensors, the noise level can be drastically reduced, thus optimizing the detection of a significant effect. This technique is particularly useful when working with brain-lesioned patients in whom the topography of brain signals may be distorted; the software essentially replaces the experimenter in searching for significant brain signals (King et al. 2013a).

– "Double-dipping," i.e., using the same data for inference and for confirmatory purposes, a problem that often plagues brain-imaging research (Kriegeskorte et al. 2009), can be largely circumvented in computer-based inference by leaving a subset of the data out of the training database and using it specifically to independently test the classifier.

– Hundreds of brain sensors are summarized into a single time curve that "projects" the data back onto a psychological space of interest. By identifying a near-optimal spatial filter, this aspect of the method simultaneously bypasses the complex problems of source reconstruction and of statistical correction for multiple comparisons across hundreds of sensors and provides cognitive scientists with immediately interpretable signals.

– Finally, because distinct classifiers are trained for different subjects, and only the projections back to psychological space are averaged across subjects, the method naturally takes into account inter-individual variability in brain topography. In this respect, the method makes fewer assumptions than classical univariate methods that implicitly rest on the dubious assumption that different subjects share a similar topography over EEG or MEG sensors. In the decoding approach, we do not average sensor-level data but only their projection onto a psychological dimension that is likely to be shared across subjects.

A drawback of the decoding method is that we cannot be sure that the features that we decode from brain signals are actually being used by the brain itself for its internal computations. For all we know, we could be decoding the brain's equivalent of the steam cloud arising from a locomotive—a side effect rather than a causally relevant signal. To mitigate this problem, we restrict ourselves to the use of linear classifiers such as a linear Support Vector Machine (SVM). In this way, we can at least increase our confidence in the fact that the decoding algorithm focuses on *explicit* neural codes. A neural code for a feature f may be considered as "explicit" when f can be reconstructed from the neural signals using a simple linear transformation. For instance, the presence of faces versus other visual categories is explicitly represented in inferotemporal cortex because many neurons fire selectively to faces, and thus a simple averaging operation suffices to discriminate faces from non-faces (Tsao et al. 2006). This definition of "explicit representation" ensures that the brain has performed a sufficient amount of preprocessing to attain a level of representation that can be easily extracted and manipulated at the next stage of neural processing, either by single neurons or by neuronal assemblies. If we used sophisticated non-linear classifiers such as "deep" convolutional neural networks (LeCun et al. 2015), we could, at least in principle, decode any visual information from the primary visual area V1, but this would be uninformative about when, how and even whether the brain itself explicitly represents this

information. By using linear classifiers, we ensure that we only decode explicit neural signals. It should be kept in mind, however, that the identification of explicit representations with linearly separable ones is a working hypothesis that remains under-investigated. More generally, it is particularly difficult to determine whether and how brain responses play a causal role in behavior and subjective perception (see, e.g., Rangarajan et al. 2014). Beyond decoding analyses, this ambiguity is in fact intrinsic to any non-causal brain-behavior correlation method.

A Test Using Auditory Novelty Signals

Figure 3 illustrates an application of the temporal generalization method to auditory novelty detection in the local/global paradigm (Bekinschtein et al. 2009; King et al. 2013a). This paradigm aims to separate two types of brain signals evoked by the violation of two types of auditory expectations: (1) automatic detection of unexpected sounds, and (2) conscious detection of unexpected sound sequences. As we shall see, the temporal generalization analysis separates these two intermingled signals, facilitates their detection, and shows that their temporal dynamics differ radically.

We recorded MEG and EEG signals while human subjects heard sequences of five repeated sounds (Wacongne et al. 2011). Sometimes the auditory sequence ended with a different sound. This unexpected local violation generated a local mismatch response, arising primarily from auditory cortex (Bekinschtein et al. 2009). Furthermore, in each block, the entire sequence was repeated several times and, occasionally, was violated by presenting a rare instance of a distinct sequence. The difference between rare and frequent sequences generated a global novelty response, arising from distributed brain areas including associative areas of parietal and prefrontal cortex, and associated with a P3b component of the event-related potential (Bekinschtein et al. 2009).

Temporal decoding allowed us to track the corresponding novelty signals in the brain. First, classifiers could be trained to discriminate whether the fifth sound was repeated or deviant (local mismatch). Above-chance decoding scores could be observed during a time window ~100–400 ms after the deviant sound. Crucially, the temporal generalization matrix revealed that this long period did not correspond to a single neural code (Fig. 3b). A diagonal generalization pattern indeed suggested that error signals changed over time as they propagated through a hierarchy of distinct brain areas. There were even periods of below-chance generalization (marked in blue in Fig. 3b), indicating that the spatial pattern of brain activity observed at time t tended to reverse at time t', possibly due to top-down inputs to the same brain area that have been postulated to play a role in cancelling out the bottom-up error signals (Friston 2005).

Second, the global effect was marked by a completely distinct pattern of temporal generalization. From about 150 ms on, classifiers could discriminate whether the global sequence was frequent or rare. The results demonstrated a
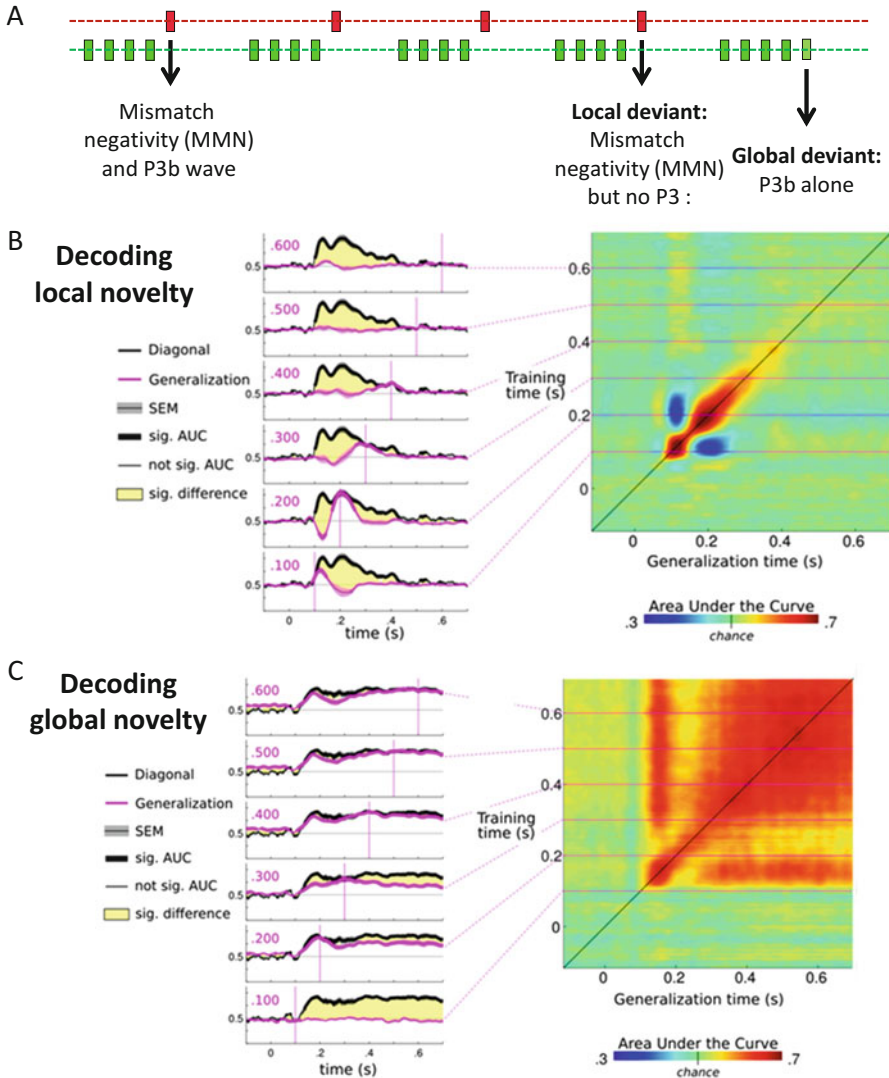
Fig. 3 Temporal decoding applied to an auditory violation paradigm, the local/global paradigm (from King et al. 2013a). (**a**) Experimental design: sequences of five sounds sometimes end with a different sound, generating a local mismatch response. Furthermore, the entire sequence is repeated and occasionally violated, generating a global novelty response (associated with a P3b component of the event-related potential). (**b, c**) Results using temporal decoding. A decoder for the local effect (**b**) is trained to discriminate whether the fifth sound is repeated or different. This is reflected in a diagonal pattern, suggesting the propagation of error signals through a hierarchy of distinct brain areas. Below-chance generalization (in *blue*) indicates that the spatial pattern observed at time t tends to reverse at time t'. A decoder for the global effect (**c**) is trained to discriminate whether the global sequence is frequent or rare. This is reflected primarily in a square pattern, indicating a stable neural pattern that extends to the next trial. In all graphs, t = 0 marks the onset of the fifth sound
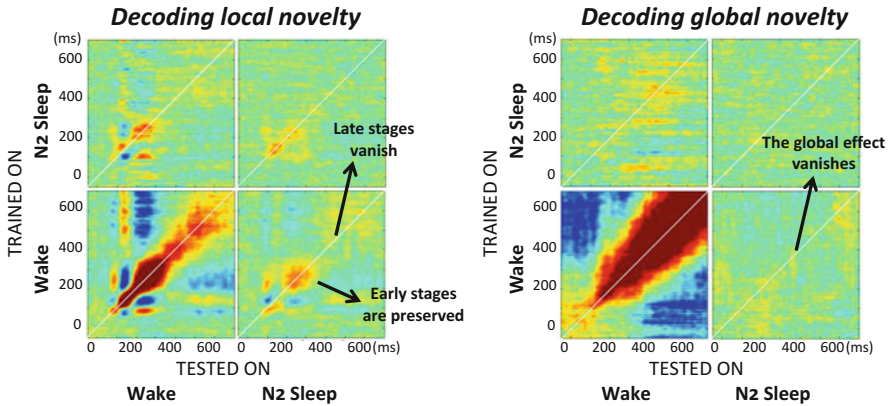
Fig. 4 Generalization of decoding across two experimental conditions, wakefulness and sleep, can reveal which processing stages are preserved or deleted (from Strauss et al. 2015). Subjects were tested with the same local/global paradigm as in Fig. 2 while they fell asleep in the MEG scanner. The local effect was partially preserved during sleep (*left*): between about 100 and 300 ms, a decoder could be trained during wake and generalize to sleep, or vice versa. Note that all late components and, interestingly, off-diagonal below-chance components vanished during sleep. As concerns the global effect (*right*), it completely vanished during sleep

square pattern of temporal generalization (Fig. 3c), indicating that the violation of global sequence expectations evoked a single and largely stable pattern of neural activity (with only a small enhancement on the diagonal, indicating a slow change in neural coding).

Further research showed that the late global response is a plausible marker of conscious processing (Dehaene and Changeux 2011): if processing reaches this level of complexity, whereby the present sequence is represented and compared to those heard several seconds earlier, then the person is consciously representing the deviant sequence and can later report it (Bekinschtein et al. 2009). Inattention abolishes the late global response but not the early local response. So does sleep: as soon as a person falls asleep and ceases to respond to the global deviants, the global response vanishes whereas the local response remains partially preserved, at least in its initial components (Fig. 4; see Strauss et al. 2015).

The disappearance of late and top-down processing stages seems to be a general characteristic of the loss of consciousness (for review, see Dehaene and Changeux 2011). In the local/global paradigm, when patients fall into a vegetative state or in a coma, the global effect vanishes whereas the local effect remains preserved. The global effect may therefore be used as a "signature" of conscious processing, useful to detect that consciousness is in fact preserved in a subset of patients in apparent vegetative state. In such patients, the temporal decoding method can optimize the detection of a global effect, even in the presence of delays or topographical distortions due to brain and skull lesions (King et al. 2013a). Unfortunately, the global effect is not a very sensitive signature of consciousness, because it may remain undetectable in some patients who are demonstrably conscious yet unable to

attend or whose EEG signals are contaminated by noise. When the global effect is present, however, it is likely that the patient is conscious or will quickly recover consciousness (Faugeras et al. 2011, 2012). Therefore, the decoding of the global effect adds to the panoply of recent EEG-based mathematical measures that, collectively, contribute to the accurate classification of disorders of consciousness in behaviorally unresponsive patients (King et al. 2013b; Sitt et al. 2014).

## Late Metastable Activity as a Signature of Consciousness

Why does the global response to auditory novelty track conscious processing? We have hypothesized that conscious perception corresponds to the entry of information into a global neuronal workspace (GNW), based on distributed associative areas of the parietal, temporal and prefrontal cortices, that stabilizes information over time and broadcasts it to additional processing stages (Dehaene and Naccache 2001; Dehaene et al. 2003, 2006). Even if the incoming sensory information is very brief, the GNW transforms and stabilizes its representation for a period of a few hundreds of milliseconds, as long as is necessary to achieve the organism's current goals Such a representation has been called "metastable" (Dehaene et al. 2003) by analogy with the physics of low-energy attractor states, where metastability is defined as "the phenomenon when a system spends an extended time in a configuration other than the system's state of least energy" (Wikipedia). Similarly, conscious representations are thought to rely on brain signals that persist for a long duration, yet without being fully stable because they can be suddenly replaced as soon as a new mental object becomes the focus of conscious thought.

The brain activity evoked by global auditory violations in the local/global paradigm fits with this hypothesis. First, this signal is only present in conscious subjects who can explicitly report the presence of deviant sequences. Furthermore, this signal is late, distributed in many high-level association areas including prefrontal cortex, and stable for an extended period of time (Bekinschtein et al. 2009). The latter point is particularly evident in temporal generalization matrices, which show that the global effect, although triggered by a transient auditory signal (a single 150-ms tone), is reflected in a late and approximately square (Fig. 3) or thick-diagonal (Fig. 4) pattern of decoding Such a pattern indicates that the evoked neural pattern is stable over a long time period. Our results indicate that the neural activation pattern can be either quasi-stable for hundreds of milliseconds (as occurs in Fig. 3, where subjects simply had the instruction to attend to the stimuli), or slowly changing with considerable temporal overlap among successive neural codes (as occurs in Fig. 4, where subjects were instructed to perform a motor response to global deviants, thus enforcing a series of additional decision, response and monitoring stages).

Many additional paradigms have revealed that conscious access is associated with an amplification of incoming information, its transformation into a metastable representation, and its efficient propagation to subsequent processing stages (Del

Cul et al. 2007; Kouider et al. 2013; Salti et al. 2015; Schurger et al. 2015; Sergent et al. 2005). For example, Fig. 5 shows the results of temporal decoding applied to a classical masking paradigm, in which a digit is made invisible by following it at a short latency with a "mask" made up of letters surrounding the digit's position (Charles et al. 2014; Del Cul et al. 2007). At short delays, subjects report the absence of a digit even when it is physically present on screen. Nevertheless, a pattern classifier can be trained to discriminate digit-present and digit-absent trials (thus decoding, from the subject's brain, a piece of information that the subject himself ignores). The classifier for subliminal digits presents a sharp diagonal pattern (Fig. 5), indicating that the digit traverses a series of transient coding stages without ever stabilizing into a long-lasting activation. When the digit is seen, however, a square pattern of temporal generalization can be observed, suggesting a metastable representation of the digit's presence. A similar difference in metastability can be observed when sorting physically identical threshold trials (SOA = 50 ms) into those that were subjectively reported as seen or unseen (Fig. 5).

Metastability can also be assessed by other means, for instance, by measuring whether the neural activation "vector" evoked by a given stimulus points in a consistent direction for a long-enough duration (Schurger et al. 2015). Here again, a few hundreds of milliseconds after the onset of a picture, stability was
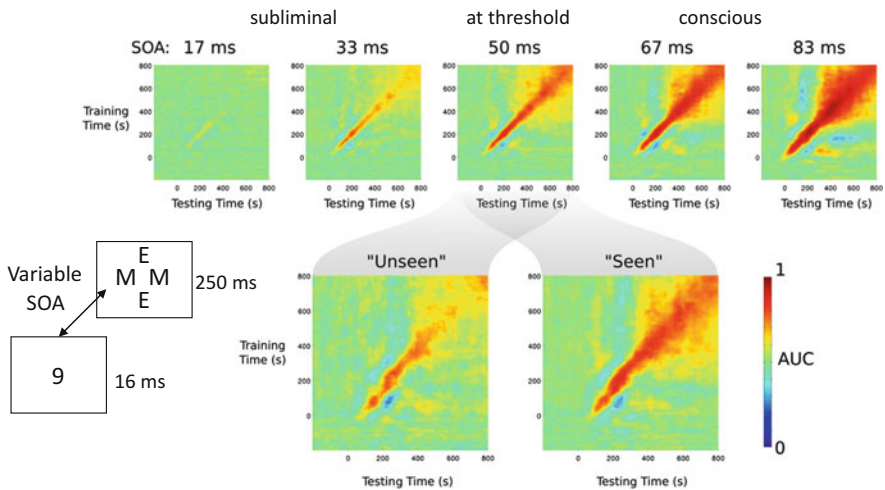
Fig. 5 Decoding reveals the signatures of subliminal and conscious processing in a masking paradigm (data from Charles et al. 2013, 2014). When the stimulus-onset-asynchrony (SOA) between a digit and a letter mask remains below 50 ms, the digit generally remains subjectively invisible. A decoder trained to discriminate digit-present and digit-absent trials decodes only a sharp diagonal pattern, indicating that the digit quickly traverses a series of successive coding stages. When the digit is seen, however, a square pattern of temporal generalization emerges, indicating that a temporally stable representation is achieved. A similar, though more modest difference, can be observed when sorting physically identical threshold trials (SOA = 50 ms) into those that were subjectively reported as seen or unseen

higher when the picture was consciously perceived than when it was unseen. Thus, late metastability consistently appears to be a plausible signature of consciousness.

Conclusion

Determining the sequence of processing stages through which a stimulus passes is an essential goal for cognitive neuroscience. Furthermore, if the GNW theory is correct, assessing whether a brief stimulus reaches a late stage of information processing in which the sensory information is stabilized and is made available to further processors can provide an efficient signature of consciousness. Both of these goals can now be achieved through the use of temporal decoding and of the temporal generalization method. Multivariate decoding of temporal signals provides a sensitive method to probe the time course of information processing. The code is freely available as part of the open-source MNE-Python software (Gramfort et al. 2014; http://martinos.org/mne/). Thus, all of the above techniques can now be readily applied to novel problems.

To summarize, our experimental findings suggest that (1) the initial stages of stimulus-evoked brain activity reflect non-conscious processing and are systematically associated with a "diagonal" pattern of temporal generalization; (2) conscious perception relates to a late period of metastability and slow sequential processing, associated with the ignition of a distributed parietal and prefrontal network and with a temporally extended, "square" pattern of temporal generalization. Recently, we have obtained evidence suggesting that these conclusions may generalize to dual-task paradigms such as the attentional blink and the psychological refractory period (Marti et al. 2015). In the future, it will be essential to determine whether they can be validated in additional paradigms.

References

Bekinschtein TA, Dehaene S, Rohaut B, Tadel F, Cohen L, Naccache L (2009) Neural signature of the conscious processing of auditory regularities. Proc Natl Acad Sci USA 106:1672–1677

Charles L, King J-R, Dehaene S (2014) Decoding the dynamics of action, intention, and error detection for conscious and subliminal stimuli. J Neurosci 34:1158–1170

Charles L, Van Opstal F, Marti S, Dehaene S (2013) Distinct brain mechanisms for conscious versus subliminal error detection. Neuroimage 73:80–94. doi:10.1016/j.neuroimage.2013.01.054

Dehaene S, Changeux JP (2011) Experimental and theoretical approaches to conscious processing. Neuron 70:200–227

Dehaene S, Naccache L (2001) Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. Cognition 79:1–37

Dehaene S, Sergent C, Changeux JP (2003) A neuronal network model linking subjective reports and objective physiological data during conscious perception. Proc Natl Acad Sci USA 100:8520–8525

Dehaene S, Changeux JP, Naccache L, Sackur J, Sergent C (2006) Conscious, preconscious, and subliminal processing: a testable taxonomy. Trends Cogn Sci 10:204–211

Del Cul A, Baillet S, Dehaene S (2007) Brain dynamics underlying the nonlinear threshold for access to consciousness. PLoS Biol 5:e260

Donders FC (1969) On the speed of mental processes (translation). Acta Psychol (Amst) 30:412–431

Faugeras F, Rohaut B, Weiss N, Bekinschtein TA, Galanaud D, Puybasset L, Bolgert F, Sergent C, Cohen L, Dehaene S, Naccache L (2011) Probing consciousness with event-related potentials in the vegetative state. Neurology 77:264–268

Faugeras F, Rohaut B, Weiss N, Bekinschtein T, Galanaud D, Puybasset L, Bolgert F, Sergent C, Cohen L, Dehaene S, Naccache L (2012) Event related potentials elicited by violations of auditory regularities in patients with impaired consciousness. Neuropsychologia 50:403–418

Friston K (2005) A theory of cortical responses. Philos Trans R Soc Lond B Biol Sci 360:815–836

Gramfort A, Luessi M, Larson E, Engemann DA, Strohmeier D, Brodbeck C, Parkkonen L, Hämäläinen MS (2014) MNE software for processing MEG and EEG data. Neuroimage 86:446–460

Isik L, Meyers EM, Leibo JZ, Poggio T (2014) The dynamics of invariant object recognition in the human visual system. J Neurophysiol 111:91–102

King J-R (2014) Characterizing electro-magnetic signatures of conscious processing in healthy and impaired human brains. Phd Thesis, University of Paris VI, Paris, France

King J-R, Dehaene S (2014) Characterizing the dynamics of mental representations: the temporal generalization method. Trends Cogn Sci 18:203–210

King JR, Faugeras F, Gramfort A, Schurger A, El Karoui I, Sitt JD, Rohaut B, Wacongne C, Labyt E, Bekinschtein T et al (2013a) Single-trial decoding of auditory novelty responses facilitates the detection of residual consciousness. Neuroimage 83:726–738

King J-R, Sitt JD, Faugeras F, Rohaut B, El Karoui I, Cohen L, Naccache L, Dehaene S (2013b) Information sharing in the brain indexes consciousness in noncommunicative patients. Curr Biol 23:1914–1919

Kouider S, Stahlhut C, Gelskov SV, Barbosa LS, Dutat M, de Gardelle V, Christophe A, Dehaene S, Dehaene-Lambertz G (2013) A neural marker of perceptual consciousness in infants. Science 340:376–380

Kriegeskorte N, Simmons WK, Bellgowan PS, Baker CI (2009) Circular analysis in systems neuroscience: the dangers of double dipping. Nat Neurosci 12:535–540

LeCun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521:436–444

Marti S, King J-R, Dehaene S (2015) Time-resolved decoding of two processing chains during dual-task interference. Neuron 88(6):1297–1307. doi:10.1016/j.neuron.2015.10.040

Pashler H (1994) Dual-task interference in simple tasks: data and theory. Psychol Bull 116:220–244

Petersen SE, Fox PT, Posner MI, Mintun M, Raichle ME (1988) Positron emission tomographic studies of the cortical anatomy of single-word processing. Nature 331:585–589

Posner MI (1978) Chronometric explorations of the mind. Lawrence Erlbaum, Hillsdale, NJ

Posner MI, Raichle ME (1994) Images of mind. Scientific American Library, New York

Rangarajan V, Hermes D, Foster BL, Weiner KS, Jacques C, Grill-Spector K, Parvizi J (2014) Electrical stimulation of the left and right human fusiform gyrus causes different effects in conscious face perception. J Neurosci 34:12828–12836

Salti M, Monto S, Charles L, King J-R, Parkkonen L, Dehaene S (2015) Distinct cortical codes and temporal dynamics for conscious and unconscious percepts. eLife 4

Schurger A, Sarigiannidis I, Naccache L, Sitt JD, Dehaene S (2015) Cortical activity is more stable when sensory stimuli are consciously perceived. Proc Natl Acad Sci USA 112:E2083–E2092

Sergent C, Baillet S, Dehaene S (2005) Timing of the brain events underlying access to consciousness during the attentional blink. Nat Neurosci 8:1391–1400

Sigman M, Dehaene S (2005) Parsing a cognitive task: a characterization of the mind's bottleneck. PLoS Biol 3:e37

Sitt JD, King J-R, El Karoui I, Rohaut B, Faugeras F, Gramfort A, Cohen L, Sigman M, Dehaene S, Naccache L (2014) Large scale screening of neural signatures of consciousness in patients in a vegetative or minimally conscious state. Brain 137:2258–2270

Sternberg S (1969) The discovery of processing stages: extensions of Donders' method. Acta Psychol 30:276–315

Sternberg S (2001) Separate modifiability, mental modules, and the use of pure and composite measures to reveal them. Acta Psychol (Amst) 106:147–246

Strauss M, Sitt JD, King J-R, Elbaz M, Azizi L, Buiatti M, Naccache L, van Wassenhove V, Dehaene S (2015) Disruption of hierarchical predictive coding during sleep. Proc Natl Acad Sci USA 122:E1353–E1362

Tsao DY, Freiwald WA, Tootell RB, Livingstone MS (2006) A cortical region consisting entirely of face-selective cells. Science 311:670–674

Wacongne C, Labyt E, van Wassenhove V, Bekinschtein T, Naccache L, Dehaene S (2011) Evidence for a hierarchy of predictions and prediction errors in human cortex. Proc Natl Acad Sci USA 108:20754–20759