The temporal window of individuation limits visual capacity

Andreas Wutz* and David Melcher

Active Perception Laboratory, Center for Mind/Brain Sciences, University of Trento, Rovereto, Italy

One of the main tasks of vision is to individuate and recognize specific objects. Unlike the detection of basic features, object individuation is strictly limited in capacity. Previous studies of capacity, in terms of subitizing ranges or visual working memory, have emphasized spatial limits in the number of objects that can be apprehended simultaneously. Here, we present psychophysical and electrophysiological evidence that capacity limits depend instead on time. Contrary to what is commonly assumed, subitizing, the reading-out a small set of individual objects, is not an instantaneous process. Instead, individuation capacity increases in steps within the lifetime of visual persistence of the stimulus, suggesting that visual capacity limitations arise as a result of the narrow window of feedforward processing. We characterize this temporal window as coordinating individuation and integration of sensory information over a brief interval of around 100 ms. Neural signatures of integration windows are revealed in reset alpha oscillations shortly after stimulus onset within generators in parietal areas. Our findings suggest that short-lived alpha phase synchronization ($\approx$1 cycle) is key for individuation and integration of visual transients on rapid time scales (<100 ms). Within this time frame intermediate-level vision provides an equilibrium between the competing needs to individuate invariant objects, integrate information about those objects over time, and remain sensitive to dynamic changes in sensory input. We discuss theoretical and practical implications of temporal windows in visual processing, how they create a fundamental capacity limit, and their role in constraining the real-time dynamics of visual processing.

Keywords: visual capacity, temporal window, oscillatory phase synchrony, individuation, integration

INTRODUCTION – VIRTUAL CONTINUITY AND STABILITY OF PERCEPTUAL SPACE AND TIME

The perception system is faced with the task of transforming continuous sensory input into discrete objects and events. It is critical for survival that the perceptual system is sensitive and quickly responsive to changes in the input over time in order, for example, to detect and interpret signals regarding object or self-motion. However, a primary goal of perceptual systems is also to uncover stability in the identity and location of spatiotemporal objects and to integrate information over extended periods of time in order to understand complex phenomena such as biological motion (Neri et al., 1998) or events (Hasson et al., 2008; Lerner et al., 2011; Zacks and Magliano, 2011). Information must be integrated over time to recover the regularities in the world and to use this perception of order to make predictions about the near future (Nastase et al., 2014). Thus, vision in real-time requires a balance combining information over time (in order to integrate motion signals or to keep track of the same spatiotemporal object) and sensitivity to new information.

A simple example of this challenge for a perceptual system is the task of crossing a busy street. Perceiving and predicting the motion of vehicles requires combining information over 100s of milliseconds or even seconds, often including the combination of motion information across occlusion or changes in retinal position caused by eye movements. On the one hand, combining information over a longer time period would likely lead to the best possible estimate of all of the features of the oncoming cars. Nonetheless, the visual system must also provide a good enough estimate of the current location of each vehicle in order to support action. Thus, the perceptual system must optimally balance the competing needs of speed and information: more time yields better information but slows down the ability of the organism to react rapidly to the current state of affairs. It seems likely that the brain provides a compromise by utilizing a hierarchy of different temporal integration windows (Pöppel, 1997, 2009; Hasson et al., 2008; Melcher and Colby, 2008; Holcombe, 2009; Lerner et al., 2011; Masquelier et al., 2011) and by alternating periods of feedforward sampling of new information with feedback/re-entrant processes (Di Lollo et al., 2000; Lamme and Roelfsema, 2000) that create a perceptual synthesis of the disparate sensory information into coherent, stable spatio-temporal entities like objects. Indeed, converging evidence suggests that temporal limits on visual processing can be broadly divided into two groups of perceptual mechanisms (Holcombe, 2009). A fast group comprises processes of feedforward feature detection and works on the scale of some 10s of milliseconds. The second group of visual mechanisms is much slower, taking more than at least 100 ms and operates on more high-level properties, like objects that have been selected and individuated.

Here we consider evidence regarding how the temporal window of object individuation might bridge the gap between fast feedforward sampling of information and slower object-based computations. We start with a selective review of the relevant literature on object individuation, its capacity limits and temporal

limits in visual perception. Then, we describe a methodology to experimentally reduce the effective visual persistence of a visual display in order to more closely map out the time course of object individuation processes. We review our recent behavioral studies using this method to show the unfolding of object individuation and working memory over time. Then we present and discuss magnetencephalography (MEG) evidence regarding the neural correlates of this process, including the possibility that neural synchronization patterns can provide useful information about the nature of integration and individuation. Finally, we discuss the implications of these findings for capacity limits in visual cognition, their relationship with natural vision and oscillatory brain dynamics, and point out some open questions and directions for future research.

INDIVIDUATION MEASURES VISUO-SPATIAL OBJECT PROCESSING

INDIVIDUATION: AN INTERMEDIATE STEP BETWEEN SAMPLING FEATURES AND OBJECTS

Although sensory information seems to extend continuously into perceptual space and time, the content of cognitive operations consists of coherent scenes containing a limited number of discrete and invariant objects in any particular instance (Treisman and Gelade, 1980; Tipper et al., 1990; Kahneman et al., 1992; Baylis and Driver, 1993; Scholl et al., 2001). Such parsing of the sensory environment into elemental perceptual units (Spelke, 1988) provides a link between sensation and cognition that couples perception to the external world, free from an infinite regress of referring to semantic categories (Pylyshyn, 2001). Reading-out objects from feedforward sensory input is called individuation and involves selecting features from a crowded scene, binding them into a unitary spatiotemporal entity and segregating this perceptual unit from other individuals in the image (Treisman and Gelade, 1980; Xu and Chun, 2009). The output of this intermediate-level visual analysis is a stable object-based reference frame in which the different features of a specific location in the scene can be bound together.

Object representations at this stage are suggested to be coarse and contain only minimal feature information. In fact such individual entities do not necessarily provide information about object identity, but can be regarded as a spatio-temporal placeholder of the object in focus until feedback processes fill in content. Several theoretical, psychophysical and neuroimaging studies have emphasized the computational importance and necessity of such incremental object representations in intermediate-level vision, with these entities described as visual indexes (Pylyshyn, 1989), proto objects (Rensink, 2000), or object-files (Kahneman et al., 1992; Xu and Chun, 2009). In its essence an object can be defined as an entity whose recent spatio-temporal history can be reviewed and therefore still can be referred to as the same entity despite of changes in its location over time (Kahneman et al., 1992).

Individuation is an intermediate step in object processing between bottom-up feature detection and the recognition of stable and coherent objects. Object representations at this level of processing are commonly measured with an enumeration task that solely requires knowing whether an object is an individual rather than its identity (which is usually measured with change detection and interpreted as the content of visual working memory). In this review we will map the temporal dynamics of visual object processing as a cascade from (a) sampling a visual signal over (b) a temporal window of ca. 100 ms duration during which a scene is segmented and individuated into (c) stable object-based representations. Our main result characterizes a brief time window of persisting sensory information after stimulus onset that limits object individuation and accounts for capacity limits in visual object processing.

CAPACITY LIMITS IN INDIVIDUATION AND VISUAL MEMORY

Although human cognition is remarkably powerful, its online workspace, working memory, appears to be highly limited in the number of informational units it processes (Miller, 1956; Luck and Vogel, 1997; Cowan, 2000). It is interesting to note that this capacity is linked to cognitive abilities in general. For example, inter-individual variability in measures of fluid intelligence and capacity estimates are highly correlated (Engle et al., 1999; Cowan et al., 2005; Fukuda et al., 2010b) and reduced capacity is often found in patients with neuropsychiatric disorders (Karatekin and Asarnow, 1998; Lee et al., 2010).

Recent evidence suggests that two distinct mechanisms, object individuation and identification, work together in creating these visual object capacity limitations (Xu and Chun, 2009). Individuation appears to be the initial bottleneck in visual object processing from an unlimited in capacity, but fragile, purely bottom-up and in parallel computed sensory representation (iconic memory: Sperling, 1960, 1963; Neisser, 1967) to such a capacity limited, durable and cognitively structured visual store (visual short-term memory: Sperling, 1960, 1963; Phillips and Baddeley, 1971). A subset of these individuated objects are elaborated subsequently during object identification. It is at this stage that identity information becomes available to the observer and the content of the object can be consolidated into durable and reportable representations in visual working memory (Xu and Chun, 2009). As individuation precedes identification, the capacity of the latter has its upper bound in the limit of the former (Melcher and Piazza, 2011; Dempere-Marco et al., 2012; **Figure 1** middle panel). In fact on a single-subject level, estimates of individuation capacity commonly exceed visual memory limits and the two measures tend to be highly correlated (Piazza et al., 2011; **Figure 1** right panel).

It has long been noted that individuation is limited in capacity: we can quickly and effortlessly perceive that there are exactly two items but not that there are exactly eight items (Jevons, 1871; compare **Figure 1** left panel upper row with lower row). Enumeration is equally quick, accurate and effortless within a narrow range of one to four objects. Such small numbers of items are supposedly simultaneously apprehended by a qualitatively distinct mechanism known as "subitizing" (Kaufman et al., 1949). Performance for set-sizes exceeding this range, as measured by reaction time and accuracy, deteriorates with every additional item to be enumerated (**Figure 1** middle panel). This suggests that visual object capacity limits are grounded in this "subitizing" phenomenon and that visual processing beyond this limit has to rely on imprecise estimation or serial and time-consuming counting that requires

FIGURE 1 | Capacity limits in visuo-spatial object processing. Left panel: typical stimuli within experiments on visuo-spatial object processing (individuation, visual memory). Individuation for up to four objects (upper panels) is accurate and fast. Visuo-spatial object processing above this limit requires successive perceptual steps (counting, lower panels). **Middle panel:** visuo-spatial object processing (individuation, visual memory) as a function of set-size. Both tasks show a limit of up to four objects. The inflection point of the sigmoid curve fit to the psychophysical data can be used to estimate individual capacity limits. **Right panel:** single-subject correlation between individuation and visual memory capacity. Limits in visuo-spatial object processing correlate across subjects and individuation usually exceed visual memory limits (Figures adapted with permission from Piazza et al., 2011).

successive perceptual steps. In contrast, "subitizing" is thought to measure visuo-spatial object processing within one single feed-forward processing iteration (for review, see Melcher and Piazza, 2011; Piazza et al., 2011).

THEORIES ABOUT OBJECT PROCESSING CAPACITY LIMITS

In light of its importance for cognitive and perceptual functioning, the search for the root of this capacity limitation is fundamental to the study of visual cognition. There are a number of competing theories for why "subitizing," and individuation in general, is limited to sets of only about three or four items (for review, see Piazza et al., 2011). These theories start with the idea that capacity measures the number of objects individuated "immediately" (Kaufman et al., 1949), as reflected in the root of the word "subitizing" (*subitus*). This capacity is characterized in terms of spatial metaphors such as an index, pointer, or slots. Capacity is thus typically thought of as a limit in spatial resolution, rather than temporal limits. Because of the apparent automaticity and immediateness of processing, several theories assumed an *ad hoc*, direct and continuous indexing between external coordinates and object-files (Pylyshyn, 1989), like focal slots waiting to be filled in with content (Luck and Vogel, 1997; Fukuda et al., 2010a). Since performance tends to deteriorate after around four items (although this does depend on individuals and task), it was proposed that there were four indexes or slots.

Starting with the idea that subitizing is an all-or-none, uniform process might, however, neglect the possibility that capacity is related to the temporal period during which individuation occurs. Individuation is a computationally complex task. Ullman (1984) has characterized vision in terms of serial tasks that involves indexing of salient items, marking previously indexed locations and multiple shifts of the processing focus. In fact, execution of such complex coding in real-time would seem likely to require

the implementation of a specialized routine set-up as a series of elemental operations (Roelfsema et al., 2000). As reviewed in the following section, temporal aspects of visual perception have been studied extensively and show that visual processing is not "immediate" (Kaufman et al., 1949) but always occurs over time. This raises the question of whether these temporal factors, rather than or in addition to spatial factors, might underlie capacity limits.

In terms of time, object individuation is a process that must, as described above, balance between the need for speed and the aim of integrating information over time about salient objects in order to recognize, remember and respond to their properties. This trade-off is apparent in the case of computer vision systems for robotics, in which an exact, metric representation of the environment is computationally expensive and typically too slow to guide behavior in real-time. Computer systems used to drive cars, for example, do not represent in detail the entire visual scene (Bertozzi et al., 2000) because such a complete, metric model cannot be updated in real-time. In the case of the human visual system, one strategy to deal with this trade-off is to individuate and integrate information about a small number of potentially important items within each perceptual cycle.

TEMPORAL RESOLUTION OF VISUO-SPATIAL OBJECT PROCESSING

VISUO-TEMPORAL LIMITS BETWEEN FEATURE DETECTION AND OBJECT-BASED COMPUTATIONS

Temporal resolution refers to the precision of a measurement with respect to time. Estimates of the temporal resolution of vision come from a variety of different tasks but can be divided into two groups of temporal limits: a fast group that operates on the order of 10s of milliseconds and a slower group of visual mechanisms taking more than 100 ms (Holcombe, 2009). The fast temporal limits are usually explained by temporal integration of low-level visual

features (like in the case of flicker fusion or integration masking; Crozier and Wolf, 1941; Kietzman and Sutton, 1968; Scheerer, 1973a,b; Di Lollo and Wilson, 1978; Coltheart, 1980; Enns and Di Lollo, 2000; Breitmeyer and Öğmen, 2006). In contrast, slower temporal limits are usually associated with high-level processing in an object-based frame of reference like in the case of feature conjunctions across space (color-shape: Holcombe and Cavanagh, 2001; or orientation-location: Motoyoshi and Nishida, 2001) or consolidation of objects in visual working memory (Gegenfurtner and Sperling, 1993; Vogel et al., 2006). Unlike the temporal blurring of basic image features, temporal processing limits for this slower group have been suggested to depend on selective attention (Holcombe, 2009). Together these two groups of processes act in concert to create a coherent perceptual impression in time.

Here, we try to combine these two frameworks, temporal resolution and attentional selection. As reviewed above, object individuation appears to be the basic set-up process for object-based representations, introducing selectivity in processing individual properties of a scene. Consistent with this idea recent evidence suggests that "subitizing" and individuation in general, rather than being a pre-attentive indexing mechanism (Trick and Pylyshyn, 1994), requires selective attention (Egeth et al., 2008; Olivers and Watson, 2008; Railo et al., 2008). We show that individuation is limited by temporal integration of sensory information over time and how visual capacity limits arise naturally as a consequence of this integration window. We argue that intermediate-level vision bridges the gap between fast feature detection and slower object-based computations, and that this depends on a temporal integration window that is used to structure and stabilize individual perceptual elements within a sampled sensory image.

TEMPORAL INTEGRATION OF SENSORY PERSISTENCE

Following stimulus onset a briefly presented visual display persists perceptually for a limited temporal window of 80–120 ms (Haber and Standing, 1970; Coltheart, 1980; Di Lollo, 1980). This persisting window acts like a low-pass filter on dynamic aspects of real-time vision, limiting the temporal resolution of perceiving each single visual event. When a second stimulus is presented in rapid succession to a first stimulus, the associated features of both stimulus onsets are partly integrated into a single percept. Such short-lived sensory integration intervals have been described to influence visual perception (Scheerer, 1973a; Enns and Di Lollo, 2000; Breitmeyer and Öğmen, 2006), visual memory (Di Lollo, 1980), and rapid perceptual decision-making (Scharnowski et al., 2009; Rüter et al., 2012).

Important insights into the temporal dynamics of sensory integration have been achieved through the study of visual masking: the reduction of the visibility of one stimulus, called the target, by another stimulus shown before and/or after it, called the mask (Enns and Di Lollo, 2000; Breitmeyer and Öğmen, 2006). It is classically explained in terms of a two-factor theory: integration and interruption masking (Scheerer, 1973a,b). Interruption masking limits more high-level feedback processing after perceptual analysis of the target has largely finished. Integration masking, however, results from short-lived temporal collapsing of feedforward sensory signals, as a consequence of the imprecise temporal resolution of the visual system. Integration of sensory persistence between rapid successive stimuli reduces the time to access the sensory trace of each single stimulus. Hence integration masking degrades visual performance by fractionating the sensory persistence of the target display and limiting its effective presentation time. Integration masking is very effectively implemented with a specific forward masking technique that makes it possible to quantitatively change the duration of sensory persistence and the degree of temporal integration by varying the onset asynchrony between the first and second display (Di Lollo, 1980; Wutz et al., 2012; **Figure 2**).

Mask and target elements share the same physical properties, in order to equate stimulus energy from both visual events. The only physical difference between mask and target constitutes their temporal onset asynchrony. Temporal integration of mask and target features occurs for stimulus onset asynchronies (SOAs) shorter than around 100 ms, because of smearing of sensory persistence triggered at each onset. For SOAs exceeding this critical time frame, mask and target persistence segregate in time and the sensory trace of the target display can be read-out. In this way, varying the SOA within this integration masking sequence controls the effective presentation time of a visual display by fractionating its sensory trace. We designed this technique to map the temporal dynamics of successive perceptual processes involved in object processing with identical visual stimuli only varying task demands: from basic detection to subsequent individuation and finally identification and consolidation of objects in visual working memory (**Figure 3**).

INTEGRATION WINDOWS LIMIT INDIVIDUATION CAPACITY
INDIVIDUATION CAPACITY INCREASES UNIT BY UNIT WITHIN THE SENSORY WINDOW

Individuation stabilizes visual perception by computing objects. This process is thought to operate within a single glance and is strictly limited in capacity to a small set of around four objects. We tested whether visual object capacity is indeed reached at the very moment a stimulus enters the visual field or instead accumulates with longer viewing time by fractionating a single glance into smaller units. We used an integration masking paradigm (see **Figure 2**) in order to vary the time to access the sensory trace of the to be individuated items and measured individuation performance for different set-sizes. Contrary to what is commonly found in "subitizing" tasks, which has consistently shown highly accurate performance up to around four objects across a wide range of studies (see **Figure 1**), fractionating the sensory persistence of the stimulus with integration masking dramatically reduces individuation capacity. This suggests that reading-out a small set of individual and stable objects is not an instantaneous process (**Figure 3**) but rather evolves over time.

Individuation capacity increases in steps within the lifetime of sensory persistence of the stimulus (**Figure 3**; Wutz et al., 2012). Within integration masking, SOA between mask and target directly reflects effective target persistence and time to read-out individual objects. Temporal integration of visual signals is complete and target information is completely inaccessible if there is

FIGURE 2 | Integration masking sequence. The visual stimuli within this integration masking sequence, a random-line noise mask and the target elements ("X"), are rendered physically indistinguishable (i.e., equal luminance, equal mean line length and width, random spatial position of lines), enforcing integration of physical features via mask-target similarity (Blalock, 2013). Mask and target events also offset together at the same time (gray square at about 60 ms on above scale), so that temporal onset asynchrony between visual stimuli (SOA; 50 ms above) constitutes the only physical difference between mask (green square at 0 ms) and target events (blue square at 50 ms). Temporal integration of mask and target features (red, dashed line) occurs for SOAs shorter than around 100 ms, since masking triggered at mask onset continues for this quasi-constant period of sensory persistence (green, dashed line). With longer SOAs the sensory trace triggered at target onset (blue, dashed line) successively segregates from masking persistence and the therein-contained target information can be read-out for an increasingly longer interval. The maximum time window available for target read-out spans the entire effective target persistence (i.e., without integration from preceding masking persistence) of around 100 ms (see Wutz et al., 2012, 2014; Wutz and Melcher, 2013 for details on the masking sequence).

common stimulus onset (SOA = 0 ms). With increasing SOA, visual signals segregate in time and the read-out of each single sensory trace increases correspondingly. The slopes of individuation across read-out time, however, co-vary with the number of individual objects to be processed. Whereas one object is sufficiently stable within 25 ms, two objects require 50 ms to be individuated. Individuation capacity for four objects, which is the average visuo-spatial capacity limit (**Figure 1**), is asymptotically reached after around 100 ms (**Figure 3** left panel; Wutz et al., 2012). Limiting the effective presentation time with integration masking reveals that processing speed and object capacity interact, rather than a uniform individuation improvement across the "subitizing range" with less temporal limitations. Consistent with this result, interactions between perceptual speed and object selection have also been reported for multiple object tracking (Holcombe and Chen, 2013).

Incremental individuation of objects within a stimulus' sensory persistence suggests that this temporary integration buffer is functionally critical for object processing. Such an integration interval might reflect the need to equilibrate read-out of invariant and stable perceptual form and almost simultaneously integrate changes in sensory input into a continuous stream of visual impressions. Sensory images that remain stationary within the first 100 ms after

sampling are successively segmented and structured into objects within its sensory persistence. Consequently, visuo-spatial object capacity limitations arise as a result of the narrow integration window bandwidth (**Figure 3**).

The speed of stable information accrual, however, is particularly crucial in case of fast changes in the sampled sensory image (<100 ms). When the sensory signal changes faster than the integration window (<100 ms; change, motion, short SOA masking sequence) individuation capacity is reduced as a function of the rate of sensory change, stabilizing only a subset of objects. This drop in visuo-spatial object processing with higher temporal processing demands balances the needs for perceptual stability in space and continuity in time. One object can already be stabilized in some 10s of milliseconds. In this way at least one object can be selected and further tracked for speeds drawing near the upper temporal limit of visual processing (Kietzman and Sutton, 1968). Structuring an entire scene into multiple objects, however, requires processing over an interval of around 100 ms. We argue that vision uses the time window of sensory persistence following stimulus onset to balance the opposing needs of individuating stable objects and maintaining the temporal resolution necessary to track rapidly changing events.
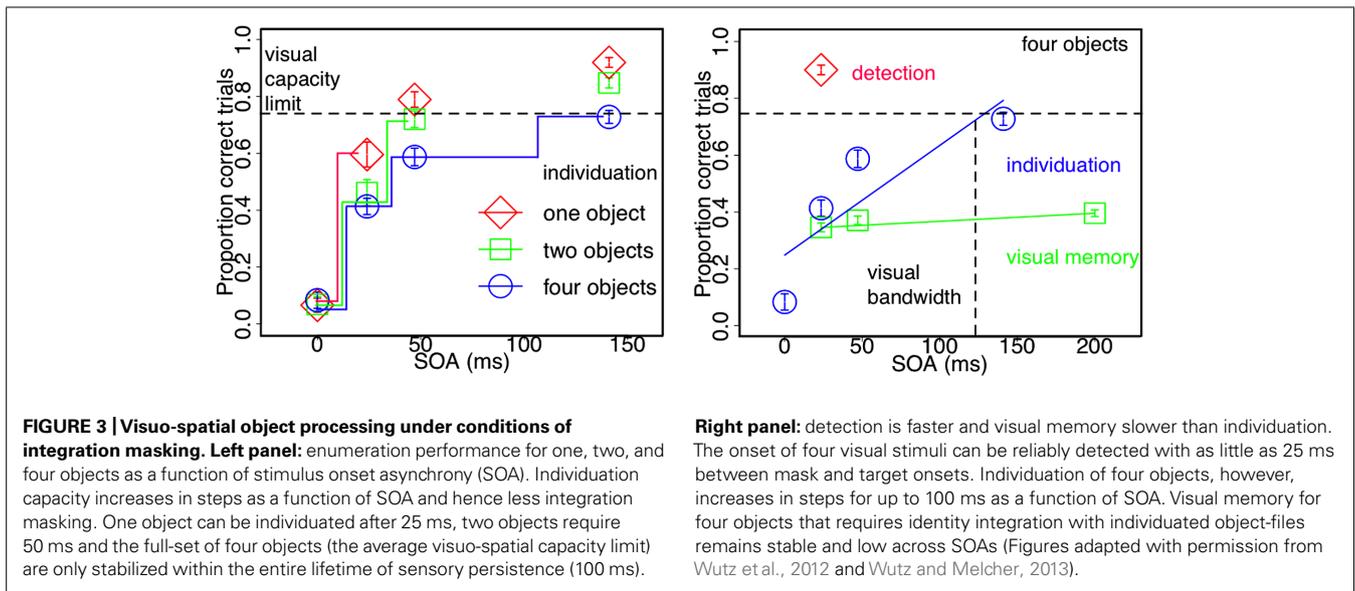
FIGURE 3 | Visuo-spatial object processing under conditions of integration masking. Left panel: enumeration performance for one, two, and four objects as a function of stimulus onset asynchrony (SOA). Individuation capacity increases in steps as a function of SOA and hence less integration masking. One object can be individuated after 25 ms, two objects require 50 ms and the full-set of four objects (the average visuo-spatial capacity limit) are only stabilized within the entire lifetime of sensory persistence (100 ms).

Right panel: detection is faster and visual memory slower than individuation. The onset of four visual stimuli can be reliably detected with as little as 25 ms between mask and target onsets. Individuation of four objects, however, increases in steps for up to 100 ms as a function of SOA. Visual memory for four objects that requires identity integration with individuated object-files remains stable and low across SOAs (Figures adapted with permission from Wutz et al., 2012 and Wutz and Melcher, 2013).

SAMPLING FEATURES IS FASTER, WHILE VISUAL MEMORY IS SLOWER THAN INDIVIDUATION

Temporal buffering of input signals does not necessarily imply that sampling of new information is inhibited completely within this integration interval. In fact, merely detecting a second event requires as little as 25 ms between event onsets (**Figure 3** right panel). Despite this remarkable processing speed, the informational content of such fast feedforward sampling is considered to be virtually unlimited in capacity (Wundt, 1899; Sperling, 1960) and can already involve higher-level visual areas, allowing for rapid scene categorization of natural images (Thorpe et al., 1996; Li et al., 2002), basic image grouping (Field et al., 1993; Roelfsema et al., 2000), visual analysis of scene semantics (but not scene syntax, Vö and Wolfe, 2012; Vō and Wolfe, 2013) or computation of global summary statistics of the raw sensory image. For example, the average size of a set of objects can be computed even when the display changes continuously (Albrecht and Scholl, 2010). Thus certain global properties of the sensory image can be read-out during fast sampling, serving as a layout for visual analysis ("the gist"; Rensink, 2000).

Without translation into a perceptually invariant and stable representation, however, information about individual elements within the sensory image is easily over-written by subsequent input (Wundt, 1899; Sperling, 1960, 1963; Breitmeyer and Öğmen, 2006). Hence, selectivity in spatio-temporal processing does not arise from a failure to sample the sensory image, but reflects subsequent structuring and stabilization of individual perceptual elements (Wutz and Melcher, 2013). Accordingly, individuation (but not basic bottom-up detection) of multiple perceptual elements evokes a set-size specific modulation of the N2pc EEG-component that is commonly assumed to index attentional selection (Mazza and Caramazza, 2011).

This coupling of the spatio-temporal coordinates of the sensory signal to a specific object representation enables identity integration between rapidly sampled content and slowly computed structure. Consequently, visual memory for an entire array of individual elements that requires binding of identity to location remains low throughout the integration bandwidth (**Figure 3** right panel; Wutz and Melcher, 2013). Consistent with the idea that individuation precedes identification, visual working memory performance rises gradually to asymptote under the influence of *backward masking* (Gegenfurtner and Sperling, 1993; Vogel et al., 2006). In contrast to the forward masking paradigm described above, backward masking is thought to reflect a disruption of processing after feedforward perceptual analysis is already completed (Scheerer, 1973a,b) but before consolidating information into visual working memory. This distinction in object processing stages between individuation and identification of objects is further fostered by task-specific activation patterns in parietal areas (Xu and Chun, 2006; Xu, 2007). Within such a "neural object-file" framework multiple visual objects are selected and individuated in an initial feedforward operation involving the inferior intra-parietal sulcus (IPS) and only subsequently identified and maintained in visual working memory (within superior IPS; Xu and Chun, 2009).

Step-wise, feedforward individuation of only a limited number of objects within a temporal buffer limits the temporal dynamics of vision. In real-time processing, however, delayed feedback systems (like the visual system; Felleman and Van Essen, 1991) exhibit asymptotic unstable behavior when confronted with signals with different latencies that have to be combined (Sandberg, 1963). Temporal buffering provides a solution to this problem by synchronizing convergent input streams. In this way, feedback processing, like identification, operates upon the outcome of the whole temporal buffer to ensure spatio-temporally coherent vision. This provides a possible solution to the problem of how to carve continuous sensory input into coherent objects, despite the presence of feedback loops. Temporal windows allow for the read-out of individual elements but also the integration of sensory flux into a dynamic stream of visual impressions (Öğmen, 1993; Wutz and Melcher, 2013).

NEURAL MECHANISMS: ALPHA PHASE SYNCHRONIZES INDIVIDUATION AND INTEGRATION

It has been suggested that implementation of integration windows within perceptual processing might involve brain oscillations (Varela et al., 1981; Pöppel and Logothetis, 1986; Dehaene, 1993). Numerous studies have shown that the temporal relation between sensory stimuli and neural oscillations can alter the perceptual outcome. For example, psychophysical threshold estimates have been shown to vary with the phase of ongoing oscillatory activity (Busch et al., 2009; Mathewson et al., 2009) and recent evidence suggests even a causal link between the two (Neuling et al., 2012). Moreover, perceived simultaneity and sequentiality of apparent motion percepts depend on the phase of the occipital alpha rhythm (Varela et al., 1981; Gho and Varela, 1988). Such periodic fluctuations have previously been described as rhythmic background sampling of the sensory surrounding (VanRullen et al., 2007; Busch and VanRullen, 2010). These results suggest that oscillations impose a "perceptual frame" on feedforward processing such that integration and individuation of sensory signals depends on its periodic phase.

One key characteristic of brain oscillations is robust phase synchronization to transient input (Buzsáki and Draguhn, 2004). In addition to effects of ongoing oscillations prior to stimulus onset, stimulus evoked synchronization patterns might reveal how phase information influences perceptual integration. In this view, external stimulation results in a "reset" of functionally relevant oscillatory patterns such that their phase synchronization is locked to stimulus onset. Resets might in particular occur in response to transient sensory change, like saccadic eye movements or real-world transitions (i.e., stimulus onset). In fact, evoked responses to successfully detected and entirely missed stimuli differ extensively (Busch et al., 2009) and alpha phase-locking accounts for individual differences in a rapid visual discrimination task (Hanslmayr et al., 2005). Likewise, reset cyclic patterns in visual task performance have been reported in response to sudden flash events (Landau and Fries, 2012) or auditory sounds (Romei et al., 2012). Moreover, electro-cortical stimulation studies demonstrated a causal link between phase resets and perceptual performance by showing that repetitive transcranial magnetic stimulation (TMS) at 10 Hz synchronizes natural alpha oscillations (Thut et al., 2011) and biases spatial selection in visual tasks (Romei et al., 2010, 2011).

In support for the idea of a link between phase synchronization and temporal integration windows, we have demonstrated that the perceptual outcome of integration masking depends on short-lived alpha phase synchrony over parietal sensors measured with MEG (Wutz et al., 2014). We contrasted trials in which observers accurately individuated low set-sizes of target items (up to 3) from masking persistence with trials in which mask and target elements integrated in time and individuation failed (see **Figure 2**). Correct individuation is accompanied by a reset selectively synchronizing alpha oscillations within a temporal window of around 100 ms (so for approximately one alpha cycle) shortly after onset of the masking sequence (**Figure 4**).

It is important to note that alpha phase synchrony reset by the masking sequence only distinguishes between individuation and integration of visual transients on rapid time scales (<100 ms; short SOA trials). Segregating sensory changes exceeding this critical time frame (long SOA trials) instead depends on slower beta power modulations prior to stimulus onset (Wutz et al., 2014). The time course of the alpha phase synchrony reset ($\approx$100 ms; $\approx$one alpha cycle) is consistent with the perceptual effects of integration masking (Enns and Di Lollo, 2000; Breitmeyer and Öğmen, 2006; Wutz et al., 2012). These results suggest that short-lived alpha synchronization is in particular key for perceptual processing of fast sensory changes. Precise phase coding within this integration cycle (through e.g., eigenfrequency damped oscillations; Buzsáki and Draguhn, 2004) in response to sensory transitions might balance individuation of perceptual elements and integration of sensory flux to guarantee spatio-temporal coherent perceptual outcomes.

IMPLICATIONS AND FUTURE DIRECTIONS

THE MAGIC WINDOW: TIME AND CAPACITY LIMITS

Following Miller's (1956) seminal paper discussing the "magic number" of 5–7 objects, the nature of these capacity limits has been a matter of intensive debate. Although a review of this extensive literature is beyond the scope here (for review see Cowan, 2000), it is important to note that the role of *time* in capacity limits has been almost neglected in any of the major theories. As described above, limits in the capacity of object individuation can be explained by the limited duration of visual persistence and the cycle of feedforward and feedback processing: in other words, temporal, rather than spatial, bandwidth. One advantage of a temporal window explanation of capacity is that capacity limits emerge naturally out of the rate of object individuation within this window of persistence, without the need to posit any *ad hoc* mechanisms.

In terms of neural implementations, the MEG evidence reported here, as well as related neuroimaging studies (Todd and Marois, 2004; Knops et al., 2014) suggest that neurons in posterior parietal cortex (PPC) may be involved in the individuation of objects. Specifically, capacity limits may reflect the spatial and temporal nature of attentional priority (saliency) maps in PPC (Melcher and Piazza, 2011; Franconeri et al., 2013; Knops et al., 2014). Unlike the priority maps in early visual areas (Zhang et al., 2012), attention priority maps in parietal cortex are thought to integrate bottom-up and top-down saliency estimates for objects over time (Bogler et al., 2011), allowing for object information to be accumulated and maintained (Mirpour et al., 2009; van Koningsbruggen et al., 2010). The results reviewed here emphasize the temporal aspects of the individuation process in determining attentional priority and capacity.

INTERACTION WITH NATURAL VISION: RETINOTOPY AND VISUAL STABILITY

A fundamental challenge for the perception of coherent spatiotemporal objects is that objects move and so do our sensory receptors. In retinotopic space, object motion would be expected to create smear within the image plane along the motion path and blurry object representations (the so-called "moving ghost problem"; Öğmen and Herzog, 2010). Whereas motion smear can be reduced

FIGURE 4 | Individuation and integration of visual signals depends on short-lived alpha phase synchrony shortly after stimulus onset. Phase synchrony [measured with inter-trial coherence (ITC); Makeig et al., 2004; also called phase-locking factor (PLF); Tallon-Baudry et al., 1996] is higher within correctly individuated trials compared to incorrect integration. Phase synchronization in response to the masking sequence is short-lived (~100 ms; lower panel) and selective for alpha oscillations (8–12 Hz; left panel). Neural generators of this effect are located in mostly left-hemispheric parietal areas (peak difference: left inferior parietal; localized using a linear constrained minimum variance (LCMV) beamformer algorithm; Van Veen et al., 1997). In particular, phase synchrony distinguishes between individuation and integration only for short SOA trials, in which temporal integration of rapid transients occurs; but not for long SOA trials, in which sensory changes exceed the critical integration time frame (Figure adapted with permission from Wutz et al., 2014).

by mechanisms similar to meta-contrast masking (Chen et al., 1995; Purushothaman et al., 1998), the read-out of moving objects would still result in fuzzy perceptual form computations. In order to avoid such "ghost-like" appearances the visual system might rely on motion segmentation when computing non-retinotopic representations (Öğmen and Herzog, 2010). This development of non-retinotopic representations necessitates integration over a temporal interval on the order of 100–150 ms (Öğmen et al., 2006; Öğmen and Herzog, 2010; Otto et al., 2010). Temporal integration of feature persistence over this temporal interval has also been implicated in the use of spatial cues for motion direction in natural images (a "motion streak"; Geisler, 1999). In general, perceptual mechanisms responsible for motion and clear, un-smeared objects share functional characteristics and are capable of analyzing form and motion concurrently (Ramachandran et al., 1974; Burr, 1980; Burr et al., 1986), fostering the close link between object form and motion perception, and temporal integration over an interval of ca. 100 ms of image persistence.

The temporal window of individuation reviewed here might serve as a buffer to translate fast retinotopic representations into stable, but slower non-retinotopic (including spatiotopic, frame-based or object-based: Melcher, 2008; Lin and He, 2012) representations that are of particular importance when objects move or change quickly. Perceiving an object as an individual within a crowded scene requires the observer to represent an object's spatiotemporal coordinates distinct from the background and from other individuals in the image. Such a structured perceptual representation contains information about sensory input that is invariant to its absolute retinotopic coordinates and gives rise to non-retinotopic form. Static input remains long enough on a well-defined location in the image, so that its associated features can be firmly attached to this location and capacity limits arise as a function of individuated locations within the image persistence. In case of fast changes in the image plane, however, only a subset of locations can be selected and individuated into non-retinotopic representations. In this way the need for higher temporal resolution balances with limits in the computation of stable non-retinotopic individuals in each single instance. Such an equilibrium might be essential in mediating between stable object and dynamic motion perception with minimal motion smear in the image plane.

Likewise, eye and head movements create a change in the retinal input and thus, potentially, a source of confusion when integrating information over time (for a discussion of the similarity between the effects of object and eye motion, see: Ağaoğlu et al., 2012). Typically, stable eye fixation periods last on the order of 150–300 ms in reading and natural viewing tasks (for review see Rayner, 1998). The external world seems stable despite these dramatic spatiotemporal disruptions in sensory information, perhaps relying on non-retinotopic object representations (Melcher and Colby, 2008; Burr and Morrone, 2011; Melcher, 2011).

We speculate that the visual system might deal with the problems of object and self-motion in a similar way, involving at least two stages of processing (see also Otto et al., 2010). At the first stage, relatively brief visual integration windows, such as those in visual masking studied here, combine information in a retinotopic manner over a time course that allows for feedforward processing. This time window is used to successively individuate spatio-temporal elements and hence stabilize sensory input. It is not coincidental, then, that the most brief eye fixations found in reading and natural viewing and intermediate-level visual integration windows would be of similar minimum durations since the goal of each new fixation is to sample part of the visual scene in order to individuate the most relevant objects. It would not make sense to move the eye before all of the information is sampled up to the level of object individuation, or to "mis-align" this integration window so that the saccade occurs right in the middle (integrating information during individuation from two different spatial locations). Moreover, the complete cycle of feedforward and feedback processing would tend to exhaust all of the useful information available from the fovea, making long fixation durations inefficient unless the information of the retina was dynamic or difficult to resolve.

At the second stage, however, information about the same object should be combined over time, over a longer time window and a non-retinotopic spatial reference frame. Accurate perception of object motion relies on non-retinotopic form computation (Öğmen et al., 2006). Likewise, there are a growing number of examples of spatiotopic perceptual effects across eye movements (for review, see Melcher and Colby, 2008; Burr and Morrone, 2011; Melcher, 2011) and there is converging evidence that this involves time scales of several hundred milliseconds (Zimmermann et al., 2013a,b). Overall, these studies suggest that there are both relatively brief, retinotopic integration windows and longer, spatiotopic windows.

One clear hypothesis from this idea is that retinotopic temporal integration windows should be reset by saccades and aligned to new eye fixations. As described above, it would be problematic if the basic object individuation process combined information from different spatial locations due to a saccadic eye movement changing retinal position during the integration window. Some evidence for a reset in the window of object individuation comes from studies of masking. Visual persistence, as measured by the missing dot task (Di Lollo, 1980), does not continue across saccades (Bridgeman and Mayer, 1983; Jonides et al., 1983) and masking can be disrupted by the intention to make a saccade (De Pisapia et al., 2010). On the other hand, the much longer temporal integration windows involved in apparent motion, over

100s of milliseconds, do not seem to be disrupted by saccades (Fracasso et al., 2010; Melcher and Fracasso, 2012). Further studies are needed to precisely define the relationship between fixation onset and the temporal windows of object individuation. The exact timing of temporal integration windows relative to eye movements might play a critical role for the impression of visual stability on rapid time scales. Such fast, feedforward computations might still involve retinotopic coordinates and therefore require saccadic remapping. However, much of the impression of visual stability might involve longer time windows that are not entirely retinotopic and thus do not require saccadic remapping.

NEURAL SYNCHRONIZATION COORDINATES FEEDFORWARD AND FEEDBACK OBJECT PROCESSING

We have reported that the short-lived alpha phase synchronization reset by stimulus onset predicts perceptual performance on an integration task. Time- and frequency characteristics of this effect (100 ms at 10 Hz) point to an alpha phase reset involved in feedforward individuation of objects. This is in line with classical findings identifying partly reset alpha oscillations in event-related potential (ERP) signatures (especially in the N1 component, Makeig et al., 2002). The functional role of alpha oscillations in perception and cognition are debated. Recent advances, however, have associated alpha phase information with the selection and recognition of object representations (for review see Palva and Palva, 2007). In support of this view, but in contrast to spatial or numerical limits in object segmentation, we propose an account based on temporal bandwidth in which phase-locking couples external signals to alpha integration cycles. Processing limits might then arise as a result of feedforward encoding within one synchronized cycle. A temporal window model based on neural synchronization patterns has several interesting functional characteristics that could coordinate feedforward and feedback object processing.

Synchronous coupling to oscillatory dynamics can structure processing into cyclic time windows for coherent integration of convergent inputs that arrive with different latencies (Buzsáki and Draguhn, 2004). In this way alpha oscillatory cycles might reflect temporal reference frames as elementary building blocks in feedforward processing. In fact, alpha cycles have been previously discussed as segmenting input into discrete snapshots of ∼100 ms (VanRullen and Koch, 2003). In line with this view illusory motion reversals in the continuous wagon wheel illusion are most prominent at wheel-motion frequencies around 10 Hz and are correlated with alpha band amplitude in the ongoing EEG trace (VanRullen et al., 2005, 2006).

VanRullen and Koch (2003) also suggested a possible way to read-out object information within such a temporal window that might involve coupled networks of nested oscillatory sub-cycles (coding individual content) within slow-wave carriers (defining the temporal reference frame). Such neural networks are capable of representing individual information by means of frequency-division multiplexing (Lisman and Idiart, 1995). Especially, phase-amplitude coupling between α- and γ-frequency bands could prioritize the selection of multiple visual objects (Jensen et al., 2012, 2014). In this view the selection of individual

items might be regulated via timed release of inhibition within one alpha cycle (Van Rullen and Thorpe, 2001; Klimesch et al., 2007; Jensen et al., 2012). Indeed, neural network dynamics of individuation can be modeled based on inhibition between competing items in a saliency map (Knops et al., 2014). Whereas multiplex coding is a well-established principle of neural function (O'Keefe and Recce, 1993; Kayser et al., 2009; Siegel et al., 2009), future work is needed to determine its functional significance for human visual cognition. Our results support the view that oscillatory synchronization might represent multiplexed phase coding and suggest that object capacity limits can arise, not only by the read-out speed of individual elements, but also from the bandwidth of the carrier function.

Importantly, integration windows can help to coordinate visual processing dynamically, because phase synchronization occurs in response to internal or external changes in input (via phase resetting; Buzsáki and Draguhn, 2004; Buzsáki, 2006; for review see Thut et al., 2012). In this way, brief phase synchronization might contribute to the rapid coordination of distributed neuronal populations (like the retinotopically organized areas along the visual hierarchy; von der Malsburg, 1981; von der Malsburg and Schneider, 1986; Singer and Gray, 1995; Fries, 2005). This might be important in order to cope with the combinatorial complexity of crowded visual scenes that contain individual elements that can consist of a nearly infinite number of feature combinations and can appear at any given moment in time or spatial location. This flexibility in combining arbitrarily complex features over space and time would seem to require neural network communication. In line with this view, phase synchronization has been hypothesized to sub-serve cross-modal integration or feature binding and to gate the information flow between local neuronal ensembles (Singer, 1999; Salinas and Sejnowski, 2001). Consistent with this idea, phase synchrony between distributed processing sites has been demonstrated to predispose visual perception (Hipp et al., 2011), route selective attention (Siegel et al., 2008; for review see Womelsdorf and Fries, 2007), predict individual working memory capacity (Palva et al., 2010) and reflect higher-level temporal processing limits (Gross et al., 2004).

Our results reveal wide spread synchronization patterns in parietal cortices locked to stimulus onset already at the level of object segmentation. We argue that vision makes use of phase synchronization as a temporal reference frame in which distributed processing can be orchestrated and aligned to input transitions. Reset synchronization patterns might therefore coordinate feed-forward and feedback mechanisms involved in encoding complex and dynamic visual scenes with nearly real-time speeds. In this framework, temporal windows might reflect a neural strategy for coherent perception of objects in space and time.

CONCLUSION

As described above, there is accumulating psychophysical and electrophysiological evidence for an intermediate-level temporal window involved in the individuation of a small number of relevant objects in a scene. Individuation capacity increases in steps within the lifetime of visual persistence of the stimulus, suggesting that visual capacity limitations arise as a result of the narrow temporal window of sensory persistence. In contrast to the main theories based on spatial slots or finite spatial resources, these findings suggest that *time* is the critical factor in the emergence of capacity limits. In this way, capacity limits can be seen as a result of the need of the visual system to coordinate feedforward and feedback processes. The cycle of feedforward and feedback processing reflects a compromise between the competing needs of a perceptual system to integrate information over extended periods of time (to get a better estimate of stable object and event properties) and sensitivity to changes in the environment.

REFERENCES

Ağaoğlu, M. N., Herzog, M. H., and Öğmen, H. (2012). Non-retinotopic feature processing in the absence of retinotopic spatial layout and the construction of perceptual space from motion. *Vision Res.* 71, 10–17. doi: 10.1016/j.visres.2012.08.009

Albrecht, A. R., and Scholl, B. J. (2010). Perceptually averaging in a continuous visual world: extracting statistical summary representations over time. *Psychol. Sci.* 21, 560–567. doi: 10.1177/0956797610363543

Baylis, G. C., and Driver, J. (1993). Visual attention and objects: evidence for hierarchical coding of location. *J. Exp. Psychol. Hum. Percept. Perform.* 19:451. doi: 10.1037/0096-1523.19.3.451

Bertozzi, M., Broggi, A., and Fascioli, A. (2000). Vision-based intelligent vehicles: state of the art and perspectives. *Rob. Auton. Syst.* 32, 1–16. doi: 10.1016/S0921-8890(99)00125-6

Blalock, L. D. (2013). Mask similarity impacts short-term consolidation in visual working memory. *Psychon. Bull. Rev.* 20, 1290–1295. doi: 10.3758/s13423-013-0461-9

Bogler, C., Bode, S., and Haynes, J. D. (2011). Decoding successive computational stages of saliency processing. *Curr. Biol.* 21, 1667–1671. doi: 10.1016/j.cub.2011.08.039

Breitmeyer, B. G., and Öğmen, H. (2006). *Visual Masking: Time Slices Through Conscious and Unconscious Vision.* New York, NY: Oxford University Press. doi: 10.1093/acprof:oso/9780198530671.001.0001

Bridgeman, B., and Mayer, M. (1983). Failure to integrate visual information from successive fixations. *Bull. Psychon. Soc.* 21, 285–286. doi: 10.3758/BF03334711

Burr, D. (1980). Motion smear. *Nature* 284, 164–165. doi: 10.1038/284164a0

Burr, D. C., and Morrone, M. C. (2011). Spatiotopic coding and remapping in humans. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 366, 504–515. doi: 10.1098/rstb.2010.0244

Burr, D. C., Ross, J., and Morrone, M. C. (1986). Seeing objects in motion. *Proc. R. Soc. Lond. B Biol. Sci.* 227, 249–265. doi: 10.1098/rspb.1986.0022

Busch, N. A., Dubois, J., and Van Rullen, R. (2009). The phase of ongoing EEG oscillations predicts visual perception. *J. Neurosci.* 29, 7869–7876. doi: 10.1523/JNEUROSCI.0113-09.2009

Busch, N. A., and VanRullen, R. (2010). Spontaneous EEG oscillations reveal periodic sampling of visual attention. *Proc. Natl. Acad. Sci. U.S.A.* 107, 16048–16053. doi: 10.1073/pnas.1004801107

Buzsáki, G. (2006). *Rhythms of the Brain.* New York, NY: Oxford University Press. doi: 10.1093/acprof:oso/9780195301069.001.0001

Buzsáki, G., and Draguhn, A. (2004). Neuronal oscillations in cortical networks. *Science* 304, 1926–1929. doi: 10.1126/science.1099745

Chen, S., Bedell, H. E., and Öğmen, H. (1995). A target in real motion appears blurred in the absence of other proximal moving targets. *Vision Res.* 35, 2315–2328. doi: 10.1016/0042-6989(94)00308-9

Coltheart, M. (1980). Iconic memory and visible persistence. *Percept. Psychophys.* 27, 183–228. doi: 10.3758/BF03204258

Cowan, N. (2000). The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behav. Brain Sci.* 24, 87–185. doi: 10.1017/S0140525X01003922

Cowan, N., Elliott, E. M., Scott Saults, J., Morey, C. C., Mattox, S., Hismjatul-lina, A., et al. (2005). On the capacity of attention: its estimation and its role in working memory and cognitive aptitudes. *Cogn. Psychol.* 51, 42–100. doi: 10.1016/j.cogpsych.2004.12.001

Crozier, W. J., and Wolf, E. (1941). Theory and measurement of visual mechanisms IV. Critical intensities for visual flicker, monocular and binocular. *J. Gen. Physiol.* 24, 505–534. doi: 10.1085/jgp.24.4.505

Dehaene, S. (1993). Temporal oscillations in human perception. *Psychol. Sci.* 4, 264–270. doi: 10.1111/j.1467-9280.1993.tb00273.x

Dempere-Marco, L., Melcher, D., and Deco, G. (2012). Effective visual working memory capacity: an emergent effect from the neural dynamics in an attractor network. *PLoS ONE* 7:e42719. doi: 10.1371/journal.pone.0042719

De Pisapia, N., Kaunitz, L., and Melcher, D. (2010). Backward masking and unmasking across saccadic eye movements. *Curr. Biol.* 20, 613–617. doi: 10.1016/j.cub.2010.01.056

Di Lollo, V. (1980). Temporal integration in visual memory. *J. Exp. Psychol. Gen.* 109, 75–97. doi: 10.1037/0096-3445.109.1.75

Di Lollo, V., Enns, J. T., and Rensink, R. A. (2000). Competition for consciousness among visual events: the psychophysics of reentrant visual processes. *J. Exp. Psychol. Gen.* 129, 481–507. doi: 10.1037/0096-3445.129.4.481

Di Lollo, V., and Wilson, A. E. (1978). Iconic persistence and perceptual moment as determinants of temporal integration in vision. *Vision Res.* 18, 1607–1610. doi: 10.1016/0042-6989(78)90251-1

Egeth, H. E., Leonard, C. J., and Palomares, M. (2008). The role of attention in subitizing: is the magical number 1? *Vis. Cogn.* 16, 463–473. doi: 10.1080/13506280801937939

Engle, R. W., Tuholski, S. W., Laughlin, J. E., and Conway, A. R. (1999). Working memory, short-term memory, and general fluid intelligence: a latent-variable approach. *J. Exp. Psychol. Gen.* 128:309. doi: 10.1037/0096-3445.128.3.309

Enns, J. T., and Di Lollo, V. (2000). What's new in visual masking? *Trends Cogn. Sci.* 4, 345–352. doi: 10.1016/S1364-6613(00)01520-5

Felleman, D. J., and Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cereb. Cortex* 1, 1–47. doi: 10.1093/cercor/1.1.1

Field, D. J., Hayes, A., and Hess, R. F. (1993). Contour integration by the human visual system: evidence for a local "association field." *Vision Res.* 33, 173–193. doi: 10.1016/0042-6989(93)90156-Q

Fracasso, A., Caramazza, A., and Melcher D. (2010). Continuous perception of motion and shape across saccadic eye movements. *J. Vis.* 10:14. doi: 10.1167/10.13.14

Franconeri, S. L., Alvarez, G. A., and Cavanagh, P. (2013). Flexible cognitive resources: competitive content maps for attention and memory. *Trends Cogn. Sci. (Regul. Ed.)* 17, 134–141. doi: 10.1016/j.tics.2013.01.010

Fries, P. (2005). A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends Cogn. Sci. (Regul. Ed.)* 9, 474–480. doi: 10.1016/j.tics.2005.08.011

Fukuda, K., Awh, E., and Vogel, E. K. (2010a). Discrete capacity limits in visual working memory. *Curr. Opin. Neurobiol.* 20, 177–182. doi: 10.1016/j.conb.2010.03.005

Fukuda, K., Vogel, E., Mayr, U., and Awh, E. (2010b). Quantity, not quality: the relationship between fluid intelligence and working memory capacity. *Psychon. Bull. Rev.* 17, 673–679. doi: 10.3758/17.5.673

Gegenfurtner, K. R., and Sperling, G. (1993). Information transfer in iconic memory experiments. *J. Exp. Psychol. Hum. Percept. Perform.* 19, 845–866. doi: 10.1037/0096-1523.19.4.845

Geisler, W. S. (1999). Motion streaks provide a spatial code for motion direction. *Nature* 400, 65–69. doi: 10.1038/21886

Gho, M., and Varela, F. J. (1988). A quantitative assessment of the dependency of the visual temporal frame upon the cortical rhythm. *J. Physiol. (Paris)* 83:95.

Gross, J., Schmitz, F., Schnitzler, I., Kessler, K., Shapiro, K., Hommel, B., et al. (2004). Modulation of long-range neural synchrony reflects temporal limitations of visual attention in humans. *Proc. Natl. Acad. Sci. U.S.A.* 101, 13050–13055. doi: 10.1073/pnas.0404951101

Haber, R. N., and Standing, L. G. (1970). Direct estimates of the apparent duration of a flash. *Can. J. Psychol.* 24:216. doi: 10.1037/h0082858

Hanslmayr, S., Klimesch, W., Sauseng, P., Gruber, W., Doppelmayr, M., Freunberger, R., et al. (2005). Visual discrimination performance is related to decreased alpha amplitude but increased phase locking. *Neurosci. Lett.* 375, 64–68. doi: 10.1016/j.neulet.2004.10.092

Hasson, U., Yang, E., Vallines, I., Heeger, D. J., and Rubin, N. (2008). A hierarchy of temporal receptive windows in human cortex. *J. Neurosci.* 28, 2539–2550. doi: 10.1523/JNEUROSCI.5487-07.2008

Hipp, J. F., Engel, A. K., and Siegel, M. (2011). Oscillatory synchronization in large-scale cortical networks predicts perception. *Neuron* 69, 387–396. doi: 10.1016/j.neuron.2010.12.027

Holcombe, A. O. (2009). Seeing slow and seeing fast: two limits on perception. *Trends Cogn. Sci. (Regul. Ed.)* 13, 216–221. doi: 10.1016/j.tics.2009.02.005

Holcombe, A. O., and Cavanagh, P. (2001). Early binding of feature pairs for visual perception. *Nat. Neurosci.* 4, 127–128. doi: 10.1038/83945

Holcombe, A. O., and Chen, W. Y. (2013). Splitting attention reduces temporal resolution from 7 Hz for tracking one object to < 3 Hz when tracking three. *J. Vis.* 13:12. doi: 10.1167/13.1.12

Jensen, O., Bonnefond, M., and VanRullen, R. (2012). An oscillatory mechanism for prioritizing salient unattended stimuli. *Trends Cogn. Sci. (Regul. Ed.)* 16, 200–206. doi: 10.1016/j.tics.2012.03.002

Jensen, O., Gips, B., Bergmann, T. O., and Bonnefond, M. (2014). Temporal coding organized by coupled alpha and gamma oscillations prioritize visual processing. *Trends Neurosci.* 37, 357–369. doi: 10.1016/j.tins.2014.04.001

Jevons, W. S. (1871). The power of numerical discrimination. *Nature* 3, 363–372. doi: 10.1038/003281a0

Jonides, J., Irwin, D., and Yantis, S. (1983). Failure to integration visual information from successive fixations. *Science* 222:188. doi: 10.1126/science.6623072

Kahneman, D., Treisman, A., and Gibbs, B. J. (1992). The reviewing of object files: object- specific integration of information. *Cogn. Psychol.* 24, 174–219. doi: 10.1016/0010-0285(92)90007-O

Karatekin, C., and Asarnow, R. F. (1998). Working memory in childhood-onset schizophrenia and attention-deficit/hyperactivity disorder. *Psychiatry Res.* 80, 165–176. doi: 10.1016/S0165-1781(98)00061-4

Kaufman, E. L., Lord, M. W., Reese, T., and Volkmann, J. (1949). The discrimination of visual number. *Am. J. Psychol.* 62, 496–525. doi: 10.2307/1418556

Kayser, C., Montemurro, M. A., Logothetis, N. K., and Panzeri, S. (2009). Spike-phase coding boosts and stabilizes information carried by spatial and temporal spike patterns. *Neuron* 61, 597–608. doi: 10.1016/j.neuron.2009.01.008

Kietzman, M. L., and Sutton, S. (1968). The interpretation of two-pulse measures of temporal resolution in vision. *Vision Res.* 8, 287–302. doi: 10.1016/0042-6989(68)90016-3

Klimesch, W., Sauseng, P., and Hanslmayr, S. (2007). EEG alpha oscillations: the inhibition-timing hypothesis. *Brain Res. Rev.* 53, 63–88. doi: 10.1016/j.brainresrev.2006.06.003

Knops, A., Piazza, M., Sengupta, R., Eger, E., and Melcher, D. (2014). A shared, flexible neural map architecture reflects capacity limits in both visual short term memory and enumeration. *J. Neurosci.* 34, 9857–9866. doi: 10.1523/JNEUROSCI.2758-13.2014

Lamme, V. A. F., and Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends Neurosci.* 23, 571–579. doi: 10.1016/S0166-2236(00)01657-X

Landau, A. N., and Fries, P. (2012). Attention samples stimuli rhythmically. *Curr. Biol.* 22, 1000–1004. doi: 10.1016/j.cub.2012.03.054

Lee, E. Y., Cowan, N., Vogel, E. K., Rolan, T., Valle-Inclán, F., and Hackley, S. A. (2010). Visual working memory deficits in patients with Parkinson's disease are due to both reduced storage capacity and impaired ability to filter out irrelevant information. *Brain* 133, 2677–2689. doi: 10.1093/brain/awq197

Lerner, Y., Honey, C. J., Silbert, L. J., and Hasson, U. (2011). Topographic mapping of a hierarchy of temporal receptive windows using a narrated story. *J. Neurosci.* 31, 2906–2915. doi: 10.1523/JNEUROSCI.3684-10.2011

Li, F. F., VanRullen, R., Koch, C., and Perona, P. (2002). Rapid natural scene categorization in the near absence of attention. *Proc. Natl. Acad. Sci. U.S.A.* 99, 9596–9601. doi: 10.1073/pnas.092277599

Lin, Z., and He, S. (2012). Automatic frame-centered object representation and integration revealed by iconic memory, visual priming, and backward masking. *J. Vis.* 12:24. doi: 10.1167/12.11.24

Lisman, J. E., and Idiart, M. A. (1995). Storage of 7 +/− 2 short-term memories in oscillatory subcycles. *Science* 267, 1512–1515. doi: 10.1126/science.7878473

Luck, S. J., and Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature* 390, 279–281. doi: 10.1038/36846

Makeig, S., Debener, S., Onton, J., and Delorme, A. (2004). Mining event-related brain dynamics. *Trends Cogn. Sci. (Regul. Ed.)* 8, 204–210. doi: 10.1126/science.1066168

Makeig, S., Westerfield, M., Jung, T. P., Enghoff, S., Townsend, J., Courchesne, E., et al. (2002). Dynamic brain sources of visual evoked responses. *Science* 295, 690–694. doi: 10.1126/science.1066168

Mathewson, K. E., Gratton, G., Fabiani, M., Beck, D. M., and Ro, T. (2009). To see or not to see: prestimulus alpha phase predicts visual awareness. *J. Neurosci.* 29, 2725–2732. doi: 10.1523/JNEUROSCI.3963-08.2009

Masquelier, T., Albantakis, L., and Deco, G. (2011). The timing of vision – how neural processing links to different temporal dynamics. *Front. Psychol.* 2:151. doi: 10.3389/fpsyg.2011.00151

Mazza, V., and Caramazza, A. (2011). Temporal brain dynamics of multiple object processing: the flexibility of individuation. *PLoS ONE* 6:e17453. doi: 10.1371/journal.pone.0017453

Melcher, D. (2008). Dynamic, object-based remapping of visual features in trans-saccadic perception. *J. Vis.* 8, 2.1–2.17. doi: 10.1167/8.14.2

Melcher, D. (2011). Visual stability. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 366, 468–475. doi: 10.1098/rstb.2010.0277

Melcher, D., and Colby, C. L. (2008). Trans-saccadic perception. *Trends Cogn. Sci. (Regul. Ed.)* 12, 466–473. doi: 10.1016/j.tics.2008.09.003

Melcher, D., and Fracasso, A. (2012). Remapping of the line motion illusion across eye movements. *Exp. Brain Res.* 218, 503–514. doi: 10.1007/s00221-012-3043-6

Melcher, D., and Piazza, M. (2011). The role of attentional priority and saliency in determining capacity limits in enumeration and visual working memory. *PLoS ONE* 6:e29296. doi: 10.1371/journal.pone.0029296

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychol. Rev.* 63, 81–97. doi: 10.1037/h0043158

Mirpour, K., Arcizet, F., Ong, W. S., and Bisley, J. W. (2009). Been there, seen that: a neural mechanism for performing efficient visual search. *J. Neurophysiol.* 102, 3481–3491. doi: 10.1152/jn.00688.2009

Motoyoshi, I., and Nishida, S. Y. (2001). Temporal resolution of orientation-based texture segregation. *Vision Res.* 41, 2089–2105. doi: 10.1016/S0042-6989(01)00096-7

Nastase, S., Iacovella, V., and Hasson, U. (2014). Uncertainty in visual and auditory series is coded by modality-general and modality-specific neural systems. *Hum. Brain Mapp.* 35, 1111–1128. doi: 10.1002/hbm.22238

Neisser, U. (1967). *Cognitive Psychology.* New York, NY: Appleton-Century-Crofts.

Neri, P., Morrone, M. C., and Burr D. C. (1998). Seeing biological motion. *Nature* 395, 894–896. doi: 10.1038/27661

Neuling, T., Rach, S., Wagner, S., Wolters, C. H., and Herrmann, C. S. (2012). Good vibrations: oscillatory phase shapes perception. *Neuroimage* 63, 771–778. doi: 10.1016/j.neuroimage.2012.07.024

Öğmen, H. (1993). A neural theory of retino-cortical dynamics. *Neural Netw.* 6, 245–273. doi: 10.1016/0893-6080(93)90020-W

Öğmen, H., and Herzog, M. H. (2010). The geometry of visual perception: retinotopic and nonretinotopic representations in the human visual system. *Proc. IEEE Inst. Electr. Electron. Eng.* 98, 479–492. doi: 10.1109/JPROC.2009.2039028

Öğmen, H., Otto, T., and Herzog, M. H. (2006). Perceptual grouping induces non-retinotopic feature attribution in human vision. *Vision Res.* 46, 3234–3242. doi: 10.1016/j.visres.2006.04.007

O'Keefe, J., and Recce, M. L. (1993). Phase relationship between hippocampal place units and the EEG theta rhythm. *Hippocampus* 3, 317–330. doi: 10.1002/hipo.450030307

Olivers, C. N., and Watson, D. G. (2008). Subitizing requires attention. *Vis. Cogn.* 16, 439–462. doi: 10.1080/13506280701825861

Otto, T. U., Öğmen, H., and Herzog, M. H. (2010). Attention and non-retinotopic feature integration. *J. Vis.* 10:8. doi: 10.1167/10.12.8

Palva, J. M., Monto, S., Kulashekhar, S., and Palva, S. (2010). Neuronal synchrony reveals working memory networks and predicts individual memory capacity. *Proc. Natl. Acad. Sci. U.S.A.* 107, 7580–7585. doi: 10.1073/pnas.0913113107

Palva, S., and Palva, J. M. (2007). New vistas for α-frequency band oscillations. *Trends Neurosci.* 30, 150–158. doi: 10.1016/j.tins.2007.02.001

Phillips, W. A., and Baddeley, A. D. (1971). Reaction time and short-term visual memory. *Psychon. Sci.* 22, 73–74. doi: 10.3758/BF03332500

Piazza, M., Fumarola, A., Chinello, A., and Melcher, D. (2011). Subitizing reflects visuo- spatial object individuation capacity. *Cognition* 121, 147–153. doi: 10.1016/j.cognition.2011.05.007

Pöppel, E. (1997). A hierarchical model of temporal perception. *Trends Cogn. Sci. (Regul. Ed.)* 1, 56–61. doi: 10.1016/S1364-6613(97)01008-5

Pöppel, E. (2009). Pre-semantically defined temporal windows for cognitive processing. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364, 1887–1896. doi: 10.1098/rstb.2009.0015

Pöppel, E., and Logothetis, N. (1986). Neuronal oscillations in the human brain. *Naturwissenschaften* 73, 267–268. doi: 10.1007/BF0036778

Purushothaman, G., Öğmen, H., Chen, S., and Bedell, H. E. (1998). Motion deblurring in a neural network model of retino-cortical dynamics. *Vision Res.* 38, 1827–1842. doi: 10.1016/S0042-6989(97)00350-7

Pylyshyn, Z. W. (1989). The role of location indexes in spatial perception: a sketch of the FINST spatial index model. *Cognition* 32, 65–97. doi: 10.1016/0010-0277(89)90014-0

Pylyshyn, Z. W. (2001). Visual indexes, preconceptual objects, and situated vision. *Cognition* 80, 127–158. doi: 10.1016/S0010-0277(00)00156-6

Railo, H., Koivisto, M., Revonsuo, A., and Hannula, M. M. (2008). The role of attention in subitizing. *Cognition* 107, 82–104. doi: 10.1016/j.cognition.2007.08.004

Ramachandran, V. S., Rao, V. M., and Vidyasagar, T. R. (1974). Sharpness constancy during movement perception – short note. *Perception* 3, 97–98. doi: 10.1068/p030097

Rayner K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychol. Bull.* 124, 372–422. doi: 10.1037/0033-2909.124.3.372

Rensink, R. A. (2000). The dynamic representation of scenes. *Vis. Cogn.* 7, 17–42. doi: 10.1080/135062800394667

Roelfsema, P. R., Lamme, V. A. F., and Spekreijse, H. (2000). The implementation of visual routines. *Vision Res.* 40, 1385–1411. doi: 10.1016/S0042-6989(00)00004-3

Romei, V., Driver, J., Schyns, P. G., and Thut, G. (2011). Rhythmic TMS over parietal cortex links distinct brain frequencies to global versus local visual processing. *Curr. Biol.* 21, 334–337. doi: 10.1016/j.cub.2011.01.035

Romei, V., Gross, J., and Thut, G. (2010). On the role of prestimulus alpha rhythms over occipito-parietal areas in visual input regulation: correlation or causation? *J. Neurosci.* 30, 8692–8697. doi: 10.1523/JNEUROSCI.0160-10.2010

Romei, V., Gross, J., and Thut, G. (2012). Sounds reset rhythms of visual cortex and corresponding human visual perception. *Curr. Biol.* 22, 807–813. doi: 10.1016/j.cub.2012.03.025

Rüter, J., Marcille, N., Sprekeler, H., Gerstner, W., and Herzog, M. H. (2012). Paradoxical evidence integration in rapid decision processes. *PLoS Comput. Biol.* 8:e1002382. doi: 10.1371/journal.pcbi.1002382

Salinas, E., and Sejnowski, T. J. (2001). Correlated neuronal activity and the flow of neural information. *Nat. Rev. Neurosci.* 2, 539–550. doi: 10.1038/35086012

Sandberg, L. W. (1963). Signal distortion in nonlinear feedback systems. *Bell Syst. Tech. J.* 42, 2533–2550. doi: 10.1002/j.1538-7305.1963.tb00976.x

Scharnowski, F., Rüter, J., Jolij, J., Hermens, F., Kammer, T., and Herzog, M. H. (2009). Long-lasting modulation of feature integration by transcranial magnetic stimulation. *J. Vis.* 9:1. doi: 10.1167/9.6.1

Scheerer, E. (1973a). Integration, interruption and processing rate in visual backward masking. *Psychol. Forsch.* 36, 71–93. doi: 10.1007/BF00424655

Scheerer, E. (1973b). Integration, interruption and processing rate in visual backward masking. *Psychol. Forsch.* 36, 95–115. doi: 10.1007/BF00424965

Scholl, B. J., Pylyshyn, Z. W., and Feldman, J. (2001). What is a visual object? Evidence from target merging in multiple object tracking. *Cognition* 80, 159–177. doi: 10.1016/S0010-0277(00)00157-8

Siegel, M., Donner, T. H., Oostenveld, R., Fries, P., and Engel, A. K. (2008). Neuronal synchronization along the dorsal visual pathway reflects the focus of spatial attention. *Neuron* 60, 709–719. doi: 10.1016/j.neuron.2008.09.010

Siegel, M., Warden, M. R., and Miller, E. K. (2009). Phase-dependent neuronal coding of objects in short-term memory. *Proc. Natl. Acad. Sci. U.S.A.* 106, 21341–21346. doi: 10.1073/pnas.0908193106

Singer, W. (1999). Neuronal synchrony: a versatile code for the definition of relations? *Neuron* 24, 49–65. doi: 10.1016/S0896-6273(00)80821-1

Singer, W., and Gray, C. M. (1995). Visual feature integration and the temporal correlation hypothesis. *Annu. Rev. Neurosci.* 18, 555–586. doi: 10.1146/annurev.ne.18.030195.003011

Spelke, E. (1988). "Where perceiving ends and thinking begins: the apprehension of objects in infancy," in *Perceptual Development in Infancy,* ed. A. Yonas (Hillsdale, NJ: Erlbaum), 197–234.

Sperling, G. (1960). The information available in brief visual presentations. *Psychol. Monogr. Gen. Appl.* 74, 1–29. doi: 10.1037/h0093759

Sperling, G. (1963). A model for visual memory tasks. *Hum. Factors* 5, 19–31.

Tallon-Baudry, C., Bertrand, O., Delpuech, C., and Pernier, J. (1996). Stimulus specificity of phase-locked and non-phase-locked 40 Hz visual responses in human. *J. Neurosci.* 16, 4240–4249.

Thorpe, S., Fize, D., and Marlot, C. (1996). Speed of processing in the human visual system. *Nature* 381, 520–522. doi: 10.1038/381520a0

Thut, G., Miniussi, C., and Gross, J. (2012). The functional importance of rhythmic activity in the brain. *Curr. Biol.* 22, R658–R663. doi: 10.1016/j.cub.2012.06.061

Thut, G., Veniero, D., Romei, V., Miniussi, C., Schyns, P., and Gross, J. (2011). Rhythmic TMS causes local entrainment of natural oscillatory signatures. *Curr. Biol.* 21, 1176–1185. doi: 10.1016/j.cub.2011.05.049

Tipper, S. P., Brehaut, J. C., and Driver, J. (1990). Selection of moving and static objects for the control of spatially directed action. *J. Exp. Psychol. Hum. Percept. Perform.* 16:492. doi: 10.1037/0096-1523.16.3.492

Todd, J. J., and Marois, R. (2004). Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature* 428, 751–754. doi: 10.1038/nature02466

Treisman, A., and Gelade, G. (1980). A feature integration theory of attention. *Cogn. Psychol.* 12, 97–136. doi: 10.1016/0010-0285(80)90005-5

Trick, L. M., and Pylyshyn, Z. W. (1994). Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. *Psychol. Rev.* 101:80. doi: 10.1037/0033-295X.101.1.80

Ullman, S. (1984). Visual routines. *Cognition* 18, 97–159. doi: 10.1016/0010-0277(84)90023-4

van Koningsbruggen, M. G., Gabay, S., Sapir, A., Henik, A., and Rafal, R. D. (2010). Hemispheric asymmetry in the remapping and maintenance of visual saliency maps: a TMS study. *J. Cogn. Neurosci.* 22, 1730–1738. doi: 10.1162/jocn.2009.21356

VanRullen, R., Carlson, T., and Cavanagh, P. (2007). The blinking spotlight of attention. *Proc. Natl. Acad. Sci. U.S.A.* 104, 19204–19209. doi: 10.1073/pnas.0707316104

VanRullen, R., and Koch, C. (2003). Is perception discrete or continuous? *Trends Cogn. Sci. (Regul. Ed.)* 7, 207–213.

VanRullen, R., Reddy, L., and Koch, C. (2005). Attention-driven discrete sampling of motion perception. *Proc. Natl. Acad. Sci. U.S.A.* 102, 5291–5296. doi: 10.1073/pnas.0409172102

VanRullen, R., Reddy, L., and Koch, C. (2006). The continuous wagon wheel illusion is associated with changes in electroencephalogram power at∼13 Hz. *J. Neurosci.* 26, 502–507. doi: 10.1523/JNEUROSCI.4654-05.2006

Van Rullen, R., and Thorpe, S. J. (2001). Rate coding versus temporal order coding: what the retinal ganglion cells tell the visual cortex. *Neural Comput.* 13, 1255–1283. doi: 10.1162/08997660152002852

Van Veen, B. D., Van Drongelen, W., Yuchtman, M., and Suzuki, A. (1997). Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. *IEEE Trans. Biomed. Eng.* 44, 867–880. doi: 10.1109/10.623056

Varela, F. J., Toro, A., Roy John, E., and Schwartz, E. L. (1981). Perceptual framing and cortical alpha rhythm. *Neuropsychologia* 19, 675–686. doi: 10.1016/0028-3932(81)90005-1

Vö, M., and Wolfe, J. M. (2012). Semantic and syntactic inconsistencies in scenes elicit differential ERP signatures. *J. Vis.* 12, 120–120. doi: 10.1167/12.9.120

Vö, M. L. H., and Wolfe, J. M. (2013). Differential electrophysiological signatures of semantic and syntactic scene processing. *Psychol. Sci.* 24, 1816–1823. doi: 10.1177/0956797613476955

Vogel, E. K., Woodman, G. F., and Luck, S. J. (2006). The time course of consolidation in visual working memory. *J. Exp. Psychol. Hum. Percept. Perform.* 32, 1436–1451. doi: 10.1037/0096-1523.32.6.1436

von der Malsburg, C. (1981). *The Correlation Theory of Brain Function.* Internal Report 81-2. Göttingen: Max-Planck-Institute for Bio-physical Chemistry.

von der Malsburg, C., and Schneider, W. (1986). A neural cocktail-party processor. *Biol. Cybern.* 54, 29–40. doi: 10.1007/BF00337113

Womelsdorf, T., and Fries, P. (2007). The role of neuronal synchronization in selective attention. *Curr. Opin. Neurobiol.* 17, 154–160. doi: 10.1016/j.conb.2007.02.002

Wundt, W. (1899). Zur Kritik tachistoskopischer Versuche. *Philos. Stud.* 15, 287–317.

Wutz, A., Caramazza, A., and Melcher, D. (2012). Rapid enumeration within a fraction of a single glance: the role of visible persistence in object individuation capacity. *Vis. Cogn.* 20, 717–732. doi: 10.1080/13506285.2012.686460

Wutz, A., and Melcher, D. (2013). Temporal buffering and visual capacity: the time course of object formation underlies capacity limits in visual cognition. *Attent. Percept. Psychophys.* 75, 921–933. doi: 10.3758/s13414-013-0454-9

Wutz, A., Weisz, N., Braun, C., and Melcher, D. (2014). Temporal windows in visual processing: 'Pre-stimulus brain state' and 'post-stimulus phase reset' segregate visual transients on different temporal scales. *J. Neurosci.* 34, 1554–1565. doi: 10.1523/JNEUROSCI.3187-13.2014

Xu, Y. (2007). The role of the superior intra-parietal sulcus in supporting visual short-term memory for multi-feature objects. *J. Neurosci.* 27, 11676–11686. doi: 10.1523/JNEUROSCI.3545-07.2007

Xu, Y., and Chun, M. M. (2006). Dissociable neural mechanism supporting visual short-term memory for objects. *Nature* 440, 91–95. doi: 10.1038/nature04262

Xu, Y., and Chun, M. M. (2009). Selecting and perceiving multiple visual objects. *Trends Cogn. Sci. (Regul. Ed.)* 13, 167–174. doi: 10.1016/j.tics.2009.01.008

Zacks, J. M., and Magliano, J. (2011). "Film, narrative, and cognitive neuroscience," in *Art and the Senses*, eds D. Melcher and F. Bacci (New York: Oxford University Press).

Zhang, X., Zhaoping, L., Zhou, T., and Fang F. (2012). Neural activities in V1 create a bottom-up saliency map. *Neuron* 73, 183–192. doi: 10.1016/j.neuron.2011.10.035

Zimmermann, E., Morrone, M., and Burr, D. (2013a). Spatial position information accumulates steadily over time. *J. Neurosci.* 33, 18396–18401. doi: 10.1523/JNEUROSCI.1864-13.2013

Zimmermann, E., Morrone, M. C., Fink, G. R., and Burr, D. (2013b). Spatiotopic neural representations develop slowly across saccades. *Curr. Biol.* 23, R193–R194. doi: 10.1016/j.cub.2013.01.065