Cognition 128 (2013) 353–362

Contents lists available at SciVerse ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Human visual processing oscillates: Evidence from a classification image technique

Caroline Blais [a,b,*], Martin Arguin [b], Frédéric Gosselin [b,*]

[a] Département de Psychoéducation et de Psychologie, Université du Québec en Outaouais, Canada
[b] Département de Psychologie, Université de Montréal, Canada

ARTICLE INFO

Article history:
Received 11 May 2011
Revised 16 April 2013
Accepted 28 April 2013
Available online 10 June 2013

Keywords:
Visual perception
Oscillations
Temporal processing
Visual sampling

ABSTRACT

Recent investigations have proposed that visual information may be sampled in a discrete manner, similarly to the snapshots of a camera, but this hypothesis remains controversial. Moreover, assuming a discrete sampling of information, the properties of this sampling—for instance, the frequency at which it operates, and how it synchronizes with the environment—still need to be clarified. We systematically modulated the signal-to-noise ratio of faces through time and examined how it impacted face identification performance. Altogether, our results support the hypothesis of discrete sampling. Furthermore, they suggest that this mechanism may operate at a rate of about 10–15 Hz and that it is synchronized with the onset of the stimulus.

1. Introduction

When a stimulus is processed by the visual system, the temporal unfolding of visual information extraction may assume different profiles. Despite the fact that we experience a continuous flow of information when we look at the world surrounding us, recent evidence suggests that information is processed in a discrete manner, such that information extraction occurs in distinct moments, in a way similar to the snapshots of a camera (Busch, Dubois, & VanRullen, 2009; Busch & VanRullen, 2010; Landau & Fries, 2012; Mathewson, Fabiani, Gratton, Beck, & Lleras, 2010; Mathewson, Gratton, Fabiani, Beck, & Ro, 2009; Mathewson et al., 2012; Rohenkohl, Cravo, Wyart, & Nobre, 2012; VanRullen & Koch, 2003; VanRullen, Reddy, & Koch, 2005). Here, we propose to test the periodicity of perception by systematically modulating the signal-to-noise ratio of stimuli through time and by examining how it impacts performance. If periodicity is found, this technique will allow us to further characterize its temporal properties.

The empirical evidence in support of the hypothesis of a periodicity in perception started to emerge in the middle of the twentieth century. For instance, it was shown that the visual threshold for detecting a flash of light varies periodically in the few milliseconds preceding the onset of an eye saccade (Latour, 1967); and that the visual threshold for detecting two flashes displayed successively varies periodically as a function of the time interval between them (Latour, 1967). These results were viewed as evidence for discrete information sampling based on the following logic: if, as postulated by the hypothesis of discrete information sampling, little or no information is being processed during some moments, the probability of detecting a brief stimulus should vary as a function of the state of this extraction process. The detection rate should decrease if the stimulus is presented during a "no extraction" state. Likewise, a stimulus with an onset occurring during a "no extraction" state should have to wait until the next extraction state before the processing begins, thus leading to a longer reaction time. As a result, another piece of evidence in support of discrete information processing is that periodicities can be observed in reaction time distributions

* Corresponding authors. Addresses: Département de psychoéducation et de psychologie, Université du Québec en Outaouais, CP 1250, succursale Hull, Gatineau (Québec), Canada J8X 3X7. Tel.: 819-595-3900 #2551; fax: 819-595-2250 (C. Blais), Département de psychologie, Université de Montréal, CP 6128, Succ. Centre-ville, Montréal, QC, Canada H3C 3J7. Tel.: +1 514 343 7550; fax: +1 514 343 5787 (F. Gosselin).

E-mail addresses: caroline.blais@uqo.ca (C. Blais), frederic.gosselin@umontreal.ca (F. Gosselin).

0010-0277/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.cognition.2013.04.009

(Dehaene, 1993; Latour, 1967; Venables, 1960; White & Harter, 1969).

A renewal of interest for the hypothesis of discrete perception came from the observation that the wagon-wheel illusion, which consists of perceiving a spoked wheel as rotating differently from its true rotation (i.e. more slowly, stationary, or with a reversed direction of rotation) can occur under continuous illumination (Purves, Paydarfar, & Andrews, 1996; Reddy, Remy, Vayssiere, & VanRullen, 2011; Simpson, Shahani, & Manahilov, 2004; VanRullen, 2006, 2007; VanRullen, Reddy, & Koch, 2006; VanRullen et al., 2005). The wagon-wheel illusion was first observed in movies (i.e. under stroboscopic presentation). The cinematic version of this illusion occurs because the speed at which the camera captures the information differs from the frequency of rotation of the spoked wheel, therefore resulting in temporal aliasing. However, the occurrence of this illusion under continuous illumination cannot be attributed to stimulus presentation constraints. Thus, some researchers have hypothesized that a discrete sampling of information by the visual system could be the basis for the illusion (Andrews & Purves, 2005; Andrews, Purves, Simpson, & VanRullen, 2005; Purves et al., 1996; Rojas, Carmona-Fontaine, López-Calderón, & Aboitiz, 2006; Simpson et al., 2004; VanRullen et al., 2005).

The wagon-wheel illusion under continuous illumination has been used as a tool to determine the properties of information sampling through time. For example, it was shown that the illusion is most prevalent when the wheel rotates at a temporal frequency of about 10 Hz, and it was proposed (based on a motion-energy model) that a system that samples information at a rate of 15 Hz could account for these data (VanRullen et al., 2005). It was also shown that attention was required for the illusion to occur (VanRullen et al., 2005) and that a decrease in power of the 13 Hz band of the EEG power spectrum was correlated with the occurrence of the illusion (VanRullen et al., 2006), suggesting that the discrete mechanism is attention-driven and that 13 Hz cortical oscillations are a potential candidate for this mechanism. However, other studies using different experimental settings have obtained results suggesting a sampling frequency other than 13 Hz (e.g. Busch & VanRullen, 2010; Busch et al., 2009; Dehaene, 1993; Landau & Fries, 2012; Latour, 1967; Mathewson et al., 2009, 2010; VanRullen, Carlson, & Cavanagh, 2007).

An alternative explanation for the wagon-wheel illusion under continuous illumination was proposed, which posits that the illusion is caused by perceptual rivalry (Holcombe, Clifford, Eagleman, & Pakarian, 2005; Kline & Eagleman, 2008; Kline, Holcombe, & Eagleman, 2004, 2006). Due to the debate regarding the origin—perceptual rivalry vs. discrete sampling—of the illusion, our understanding of the temporal properties of visual information sampling would likely benefit from using a method different from that of the wagon-wheel illusion paradigm. Furthermore, because the wagon-wheel illusion principally recruits motion perception mechanisms, the properties of the information sampling of other domains of vision remain to be examined. Hence, supporting evidence for perceptual oscillations came from studies demonstrating that the visual system can be trained, using a rhythmic stimulation, to become more sensitive to stimuli presented in phase with this rhythmic stimulation (Lakatos, Karmos, Mehta, Ulbert, & Scroeder, 2008; Mathewson et al., 2010, 2012; see also Jones, Moynihan, MacKenzie, & Puente, 2002 for a similar phenomenon in audition). Moreover, it was shown that the probability that a stimulus is detected or reaches consciousness is modulated by the phase of the ongoing brain oscillations in low frequency rhythms (Busch & VanRullen, 2010; Busch et al., 2009; Lakatos et al., 2008; Mathewson et al., 2009, 2012). Rohenkohl et al. (2012) have also shown that the improvement in detecting stimuli that are in phase with a rhythmic stimulation occurs via a contrast gain. These observations support the idea that perception is discrete and that the state of the extraction process may determine whether a stimulus is perceived or not.

If perception is indeed discrete, its properties remain to be clarified. For example, it was proposed that a sampling occurring at a rate of around 13–15 Hz (VanRullen et al., 2005, 2006) could be related to the occurrence of the wagon-wheel illusion under continuous light, and perceptual modulations at a rate of around 7–10 Hz have been observed in detection tasks (Busch & VanRullen, 2010; Busch et al., 2009; Landau & Fries, 2012; Mathewson et al., 2009). Does that mean that the sampling frequency varies as a function of the kind of visual processing required by the task? Relatedly, does this information sampling synchronize with the visual stimulation and if so, how? For example, is the sampling resetting its phase at the beginning of each trial in a visual perceptual experiment (i.e. with stimulus onset)? Or is the sampling evolving without synchronizing with the external world, as a passive ongoing oscillation (i.e. random)? Synchronization with the stimulation is more consistent with the periodicities observed in reaction time distributions as well as in the visual threshold preceding the onset of an eye saccade. Nevertheless, how the information sampling synchronizes with the external word remains unclear.

We propose an alternative experimental approach in order to address these questions. The method consists of systematically varying the signal-to-noise ratio of stimuli through time while maintaining stimuli energy constant and examining how different temporal profiles of signal-to-noise ratio impact performance. If perception is indeed discrete, the greater the overlap between the SNR's profile and the observer's information processing profile, the better the performance should be.

The method that we propose is essentially a classification image technique (e.g. Eckstein & Ahumada, 2002; Gosselin & Schyns, 2004). Classification image techniques have been used to probe time before, but never to specifically verify the presence of periodicities in information processing (e.g. Blais et al., 2009; Fiset et al., 2009; Gold & Shubel, 2006; Neri & Heeger, 2002; Neri & Levi, 2008; Vinette, Gosselin, & Schyns, 2004). For instance, Vinette, Gosselin and Schyns (2004) used temporal classification images to reveal the time course of visual features utilization in a face identification task. On each trial, different facial areas were selected and rendered visible while the rest of the face was hidden behind a gray mask. The visibility of the facial areas also varied through time, so that different

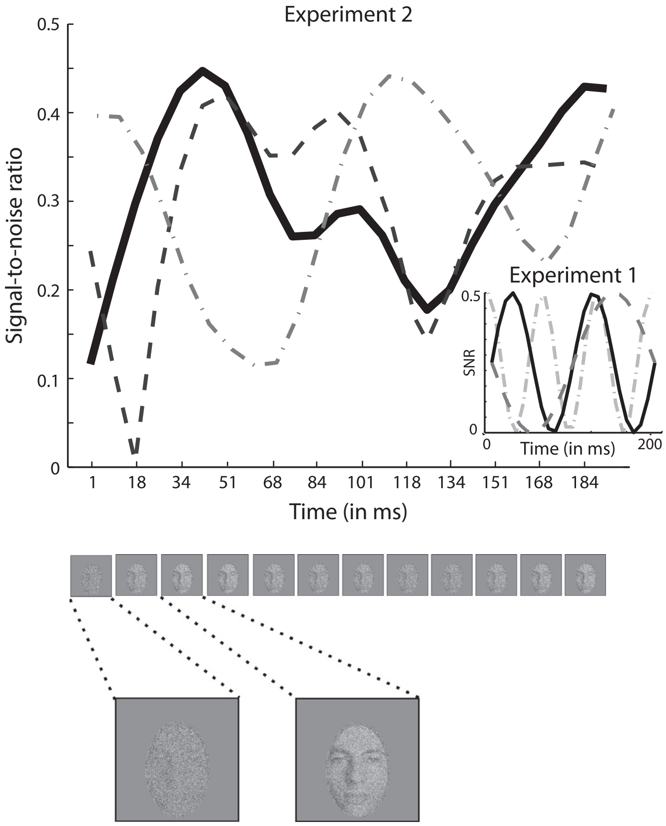

Fig. 1. Top: examples of signal-to-noise ratio profiles in Exp. 1 (inset) and 2. The different line styles represent different trials. Bottom: faces embedded in white noise with the signal-to-noise ratio for that trial represented by a solid black line in the Exp. 2 graph.

facial areas were visible at different moments within the trial. The underlying assumption was that if the information useful for the task is available at the right moment, the probability that the participant responds correctly will increase. In contrast, if the information useful for the task is not available at the right moment during the processing, then the probability that the participant responds correctly will decrease. They found that the left eye was the first visual area used in a face identification task, followed by the right eye. Interestingly, 10 Hz oscillations were observed in the time course of information use for a face identification task (Vinette et al., 2004; see Blais et al., 2009, for similar findings during word reading).

Moreover, Neri and Levi (2008) examined directional tuning in human observers using a classification image technique and found some oscillations—albeit slower than in other studies (i.e. about 3 Hz)—in their participant's sensitivity to target direction: it peaked between 30 and 60 ms, declined to 0 between 120 and 180 ms, and rose again between 240 and 300 ms. They proposed a self-normalization process with a delay of 90 ms to explain this oscillation. However, this oscillation in directional sensitivity also fits within a discrete processing framework: the frames on which the system was not processing (or processing less) were the ones on which directional sensitivity declined.

Here, two experiments were conducted using the method proposed above (see Fig. 1), i.e. modulating stimulus visibility through time. In Exp. 1, the signal-to-noise ratio of stimuli was modulated as a function of time with one sine function characterized by a particular frequency (i.e. 5, 10, 15 or 20 Hz) and phase (0, $\pi/3$, $2\pi/3$, π, $4\pi/3$ or $5\pi/3$) that varied randomly across trials. Exp. 2 modulated the signal-to-noise ratio of stimuli using a composite waveform function, to allow us to characterise the sampling properties more precisely.

2. Experiment 1

2.1. Methods

2.1.1. Participants

Six students from the Université de Montréal took part in this experiment. All had normal or corrected-to-normal visual acuity. All procedures were carried out with the ethics approval of the Université de Montréal.

2.1.2. Material and stimuli

The stimuli were displayed on a high-resolution Sony monitor at a refresh rate of 120 Hz. The experiment ran on a Macintosh G4 computer. The experimental program was written in Matlab, using functions from the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997). The monitor was calibrated to allow linear manipulations of luminance. The background luminance was of 38.9 cd/m^2.

Stimuli were ten grayscale pictures of faces subtending 7×7 degrees of visual angle—five female and five male faces with a neutral expression were used. The eyes, nose and mouth position on those pictures were aligned using rotation, translation and scaling. Faces were displayed through an oval aperture subtending 4.8×7 degrees of visual angle to reveal only their inner features. The luminance and spatial spectra were equated across stimuli. The viewing distance was maintained at 57 cm by using a chinrest.

We varied the signal-to-noise ratio sinusoidally as a function of time while maintaining stimulus energy constant. On each frame (f) of every trial, a weighted sum of a face stimulus (set to unit energy by dividing by $\sqrt{E}$, E being its original energy) and a white Gaussian noise field of the same size (set to unit energy by dividing by $\sqrt{N}$, N being its original energy)[1] was performed. A different noise sample was generated on each trial, but this noise sample was used on every time frame within one trial. The weights (a and b, respectively, for the signal and the noise) were chosen so that the signal-to-noise ratio changed according to a sinus with a frequency of 5, 10, 15 or 20 Hz and a phase (i.e. relative to stimulus onset) of 0, $\pi/3$, $2\pi/3$, π, $4\pi/3$, or $5\pi/3$—resulting in a total of 24 possible signal-to-noise ratio functions (*signal-to-noise*). The sinusoidal variation was determined by the function $\textit{signal-to-noise}(f) = A \cdot \sin(wt + \Phi)/2 + 0.5$, where A is the maximum amplitude of the signal (controlled with QUEST, see below), and w and Φ are the frequency and the phase of the modulation, respectively. More specifically, $a = \sqrt{(\textit{signal-to-noise}(f) \cdot b^2)}$ and $b = \sqrt{(E/(\textit{signal-to-noise}(f) + 1))}$.[2]

2.1.3. Procedure

Participants had to perform a face identification task. This task was selected because it is relatively taxing for the visual system, and the maximum signal-to-noise ratio could therefore remain relatively high. Each face in our set of stimuli was associated with a keyboard key, and participants had to press on the appropriate key to identify the face during the experimental phase. Before the experimental phase, the participants had 10 min to familiarize themselves with the set of 10 faces and their associated keys. A practice phase then began: each practice block contained 75 trials. On each trial, a fixation cross was first displayed at the center of the screen for a duration of 500 ms. It was immediately replaced by the full contrast stimulus (i.e. without noise) for a duration of 200 ms. A homogeneous gray screen was displayed at the end of stimulus presentation and remained on screen until the participant responded. The participant's task was to identify the face as quickly and as accurately as possible by pressing on the appropriate key. The participants had the choice of looking at the keyboard to find the appropriate response key or to learn the position of the keys by heart (in which case they did not have to search the keyboard). The practice phase ended when the participant's performance reached 95% on two successive blocks.

Each subject completed 7 blocks of 160 trials in the experimental phase. The sequence of events on each trial of the experimental phase was the same as in the practice phase, with the exception that the signal-to-noise ratio of the stimulus presented was manipulated through time as described in the Material and stimuli section. Performance was maintained at 75% by adjusting the maximum signal-to-noise ratio on a trial-by-trial basis with QUEST, irrespective of the frequency and the phase of the temporal profile of the signal-to-noise ratios (Watson & Pelli, 1983).

2.2. Results and discussion

The maximum signal-to-noise ratio necessary to maintain performance at 75% correct had a median of 0.33 across participants (first quartile: 0.27, third quartile: 0.59). On average, the peak signal contrast was of 0.25 (std = 0.12), and the RMS noise contrast was of 0.27 (std = 0.01). On average, participants responded 1.3 s (std = 3.1 s) following stimulus onset.

We searched for the pattern of frequencies and phases that best predicted the subjects' performance using an analysis procedure (see Fig. 2a) that amounts to a multiple linear regression on the properties of the sampling functions that were used during the experiment (explanatory variables) and on the participant's response accuracy (predictor variables). The logic here is that when the profile of the signal-to-noise ratio on a trial match the properties of the participant's sampling profile, then the probability of a correct response should increase. In contrast, when the profile of the signal-to-noise ratio does not match the participant's sampling profile, then the probability of a correct response should decrease. Thus, by allocating positive weights to the frequencies and phases that characterized the signal-to-noise ratio profiles that led to a correct response and negative weights to the frequencies and phases that characterized the signal-to-noise ratio profiles leading to an incorrect response, we can determine which properties were correlated with accuracy. In other words, for each trial of the experiment, we created a 4×6 matrix (i.e. four frequencies, six phases) and we put a negative or a positive weight (i.e. depending on the accuracy for that trial) in the appropriate cell, i.e. the one corresponding to the frequency and phase of the temporal profile of the signal-to-noise ratio on that trial. The weights corresponded to the accuracies transformed into z-scores (i.e. using the mean accuracy across all trials and the standard deviation of the accuracies across all trials). All these weighted matrices were then summed in order to produce a 4×6 matrix of regression coefficients that, henceforth, we will call the classification image. The accuracies were transformed into z-scores to take into account the different number of correct and incorrect trials, so that even if there

[1] The energy of a stimulus can be defined by the sum of the squared contrast of the pixels of the image (or, in other words, by the sum of the squared differences between each pixel luminance value and the mean luminance of the image). Thus, by dividing an image by the square root of its original energy, the resultant energy of this image is 1. Here, we set the energy of the image and of the noise to a value of 1 in order to be able to entirely manipulate the total energy with the weights a and b. The value of these weights was adjusted so that the total energy of the stimulus (signal + noise) was constant across time, and equal to the original energy of the face image.

[2] This non-linear contrast manipulation introduces high-temporal frequency harmonics (i.e. above 20 Hz). These artefacts, however, accounted for only 6.04% of the total energy, which is negligible.

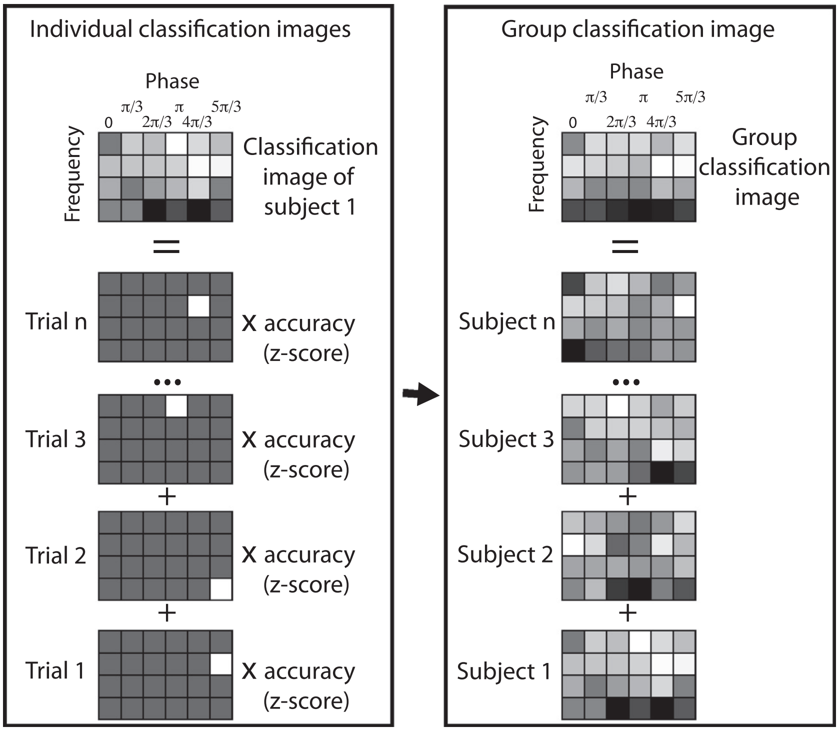

Fig. 2a. Illustration of the analysis procedure in the Fourier domain. Left panel: each sinusoidal function used during the experiment was Fourier transformed, and the magnitude values for each frequency were placed into the corresponding cell of a 4×6 matrix (i.e. cells colored in white indicate the frequency and phase of the signal on that trial). Note that in Exp. 1 we did not need to Fourier transform the signal function: it was a pure sine wave, so we knew exactly the frequency and the phase of the signal presented. Therefore, we only had to put a negative or a positive weight (i.e. depending on the accuracy on that trial) in the appropriate cell, i.e. the one that corresponded to the frequency and phase of the temporal profile of the signal-to-noise ratios on that trial. Individual classification images were generated by calculating a weighted sum of all these matrices. The weights consisted of the participant's accuracy, transformed into *z*-scores, on each trial. Right panel: the group classification image was calculated by summing the individual classification images.

were more correct than incorrect trials, the two kinds of trials would have an equal importance in the classification image.

The same procedure was then repeated one thousand times on shuffled accuracies to determine statistical significance. In other words, by using the same SNR profiles as those that were created during the experiment and by shuffling the accuracies of all the participants, we randomly generated 1000 matrices of regression coefficients, thus producing 24,000 regression coefficients (since there were 24 cells—4 frequencies $\times$ 6 phases—in each matrix). We then used the 50 highest (i.e. 99.9 percentile) values as our statistical threshold (i.e. 100-5/24, since a Bonferroni correction was applied to adjust for the multiple comparisons across frequencies and phases). The results of this analysis, displayed in Fig. 3, show that one frequency—10 Hz with a phase between $4\pi/3$ and 2π—was positively correlated with accuracy. There was a difference of 7% between the mean accuracy associated with the SNR profile that led to the worst performance and the mean accuracy associated with the SNR profile that led to the best performance.

The observation that some frequencies are correlated with accuracy is consistent with the idea that visual information is processed in a discrete manner. In fact, as explained in the Introduction, this hypothesis predicts that the performance will be influenced by the amount of overlap between the SNR's profile and the observer's information processing profile. Here, we find that the subjects' performance is indeed influenced by the SNR's profile; the performance is best when the SNR varies at a frequency of 10 Hz with a phase between $4\pi/3$ and 2π.

Exp. 1 also provides some information regarding how the information processing mechanism synchronizes with the environment. Indeed, the fact that only a restricted range of phases attained statistical significance for each frequency suggests that the information sampling is in phase with the beginning of a trial.

In Exp. 2, we will use composite waveforms instead of pure sine waves to modulate the SNR through time. This will allow us to verify if our findings can be replicated using a different stimulus generation procedure. Moreover, composite waveforms will allow us to characterize the sampling properties more precisely.

3. Experiment 2

3.1. Methods

3.1.1. Participants

Six students from the Université de Montréal who had not participated in the previous experiment took part in

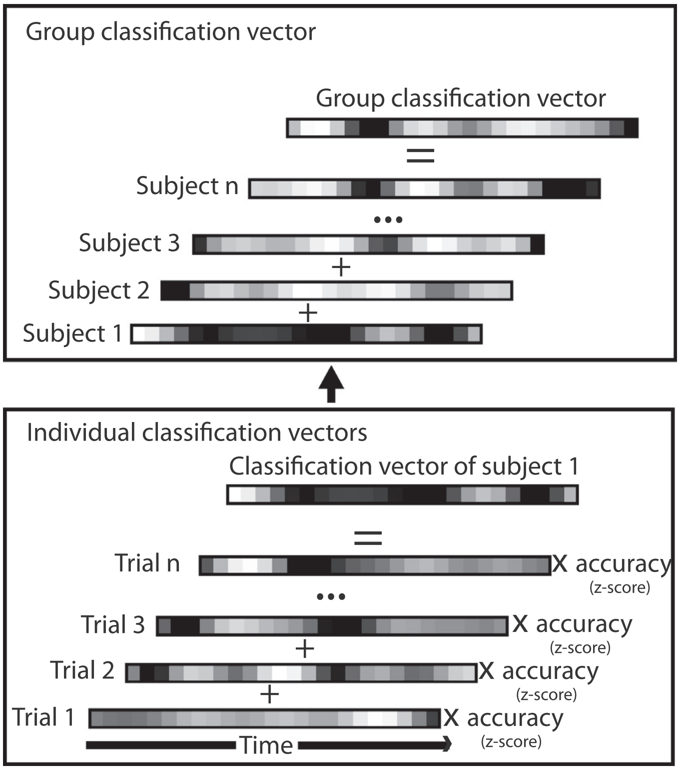

Fig. 2b. Illustration of the analysis procedure in the temporal domain. Bottom panel: individual classification vectors were generated by calculating a weighted sum of all the signal-to-noise ratio functions used during the experiment. The weights consisted of the participant's accuracy, transformed into *z*-scores, on each trial. The different shades of gray represent the different signal-to-noise ratios at each temporal frame. Top panel: the group classification vector was calculated by summing the individual classification vectors.

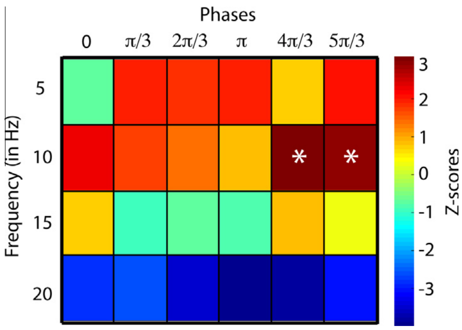

Fig. 3. Group classification image obtained in Exp. 1, computed in the Fourier domain. Red and blue correspond to positive and negative correlations, respectively. White star indicates significant combinations of frequency and phase. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

this experiment. All had normal or corrected-to-normal visual acuity. All procedures were carried out with the ethics approval of the Université de Montréal.

3.1.2. Material and stimuli

The material and stimuli were the same as in Exp. 1, with the exception that the sampling function used to manipulate the signal-to-noise ratio changed according to a composite waveform made of four frequencies (5, 10, 15 or 20 Hz), the relative amplitude and the phase of each frequency varying randomly from one trial to the other. In order to do so, we created 24-cell vectors of white Gaussian noise that we then low-passed to keep the temporal frequencies at 20 Hz and under.

3.1.3. Procedure

The procedure was the same as the one used in Exp. 1.

3.2. Results and discussion

The maximum signal-to-noise ratio necessary to maintain performance at 75% correct had a median of 0.47 across participants (first quartile: 0.40, third quartile: 0.57). On average, the peak signal contrast was of 0.28 (std = 0.08), and the RMS noise contrast was of 0.26 (std = 0.01). On average, participants responded 1.2 s (std = 0.97 s) following stimulus onset.

We searched for the pattern of frequencies and phases that best predicted the subjects' performance using an analysis procedure (see Fig. 2a) that amounts to a multiple linear regression on the properties of the sampling functions that were used during the experiment (explanatory variables) and on the participant's response accuracy (predictor variables). That is, a Fourier transform was applied to each signal-to-noise function used during the experiment, and the magnitude values for each frequency were

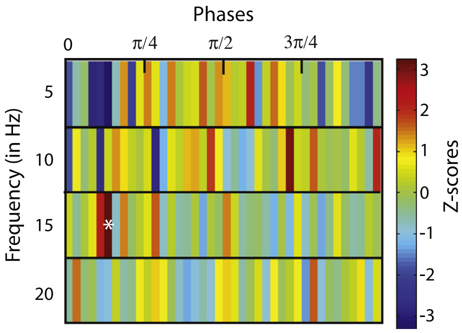

Fig. 4a. Group classification image obtained in Exp. 2, computed in the Fourier domain. Red and blue correspond to positive and negative correlations, respectively. White star indicates significant combinations of frequency and phase. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

placed into a 4 × 40 matrix (i.e. four frequencies, 40 linearly spaced phase bins from $[-\pi, \pi[$). A weighted sum of all those matrices was calculated, using the accuracies transformed into *z*-score values as weights. The same procedure was then repeated one thousand times on shuffled accuracies to determine statistical significance. In other words, by using the same SNR profiles as those created during the experiment and by shuffling the accuracies of all the participants, we randomly generated 1000 matrices of regression coefficients, thus producing 160,000 regression coefficients (since there were 160 cells—4 frequencies × 40 phases—in each matrix). We then used the 50 highest values (i.e. 99.9 percentile) as our statistical threshold (i.e. 100-5/160, since a Bonferroni correction was applied to adjust for the multiple comparisons across frequencies and phases).

The results of this analysis (see Fig. 4a) show that the 15 Hz frequency was positively correlated with accuracy. The performance was best when the signal-to-noise ratio started with a peak (phase of 0.15π). This finding in the Fourier classification image again shows that a phase was significantly correlated with correct responses. This corroborates the hypothesis that the information sampling was in phase with the stimulus (i.e. at least for the 15 Hz frequency).

We then searched for the temporal pattern of signal-to-noise ratios that best predicted the subjects' performance using an analysis procedure (see Fig. 2b) that amounts to a multiple linear regression on the sampling functions that were used during the experiment (explanatory variables) and on the participant's response accuracy (predictor variables). We called the result of this analysis a temporal classification vector (i.e. the signal-to-noise ratio function that best predicted subjects' performance). The temporal classification vector was computed by summing all the signal-to-noise functions used during the experiment weighted by the corresponding accuracy transformed into *z*-scores. The individual vectors of regression coefficients were then summed into a group vector, which we call a classification vector. This classification vector is presented in Fig. 4b. A Pixel test was applied in order to determine statistical significance ($Z_{crit} = 2.84$, $p < 0.05$). The time frames occurring between the stimulus onset and 25 ms after the stimulus onset and between 133 and 150 ms after the stimulus onset reached statistical significance, suggesting that when the stimulus was visible at those moments, the participants' performance increased. A glance at the temporal classification vector reveals that another group of frames, around 75 ms following stimulus onset, were very close to reaching statistical significance. The Pixel test allows us to reveal whether each piece of information (i.e. each pixel in an image, each frame in our SNR temporal profiles, etc.) is correlated with performance, but another test—the Cluster test—permits us to reveal the clusters of contiguous

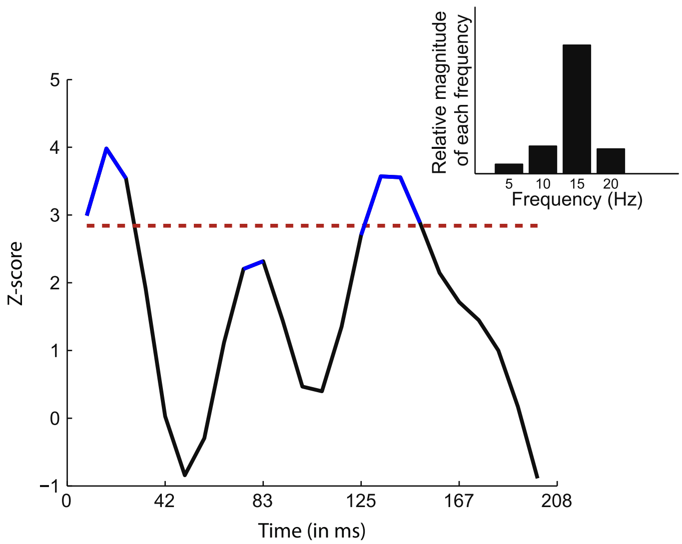

Fig. 4b. Group classification vector obtained in Exp. 2, computed in the temporal domain. The blue lines indicate the significant clusters of connected pixels above an arbitrary *z*-threshold of 2.38 ($p < 0.01$). The dashed red line indicates the significance threshold based on the pixel test. The inset shows the relative magnitude of the different frequencies composing the sinus discrete model described in footnote 4—that is, the one for which the temporal classification vector had the best fit with the temporal classification vector of the humans.

information in a signal that are correlated with performance (Chauvin, Worsley, Schyns, Arguin, & Gosselin, 2005). The Pixel test and the Cluster test can both be conducted without the need to apply a *p*-value correction (Poline, Worsley, Evans, & Friston, 1997). We therefore performed a Cluster test (Chauvin et al., 2005) to verify if the middle peak in the temporal classification vector is correlated with the participants accuracy. The Cluster test was implemented as a bootstrap (10,000 vectors of regression coefficients were randomly generated) to determine statistical significance. The Cluster test gives a statistical threshold for the size of clusters of connected elements above an arbitrary *z*-score threshold. With the arbitrary *z*-score threshold set to 2.38, the statistical threshold corresponding to a family-wise error of $p = .01$ was a cluster size of 2. Thus, any cluster of two time frames or more reaching the significance threshold indicates that when some visual information was available at that specific moment, the probability that the participants answered correctly significantly increased. The temporal classification vector presented in Fig. 4b demonstrates how the different time frames were correlated with accuracy.[3]

The results of Exp. 2 show that within the 200 ms stimulus duration, there were three moments during which a high signal-to-noise ratio was particularly useful. There was a difference of 7% between the mean accuracy associated with the SNR profiles that were most correlated with the temporal classification vector (top 1%) and the mean accuracy associated with the SNR profiles that were least correlated with the temporal classification vector (bottom 1%). These oscillations in the usefulness of visual information through time are consistent with a discrete processing of information at ~15 Hz[4] (see the inset in Fig. 4b).

Moreover, the results of Exp. 2 offer some information about how the system synchronizes with the environment. The presence of three moments during which a high signal-to-noise ratio was particularly useful in the classification vector suggests that the sampling phase was not random: had this been the case, the moments at which information was sampled would have varied from trial to trial and would therefore have cancelled each other out in the classification vector across trials. Thus, Exp. 2 replicated the findings of Exp. 1, suggesting that information sampling synchronizes at least partly with the stimulation.

4. Discussion

In two experiments, the signal-to-noise ratio of faces was modulated through time such that at some moments, visual information was available whereas at other moments, no information was available. The primary aim of these experiments was to test the hypothesis that visual information would be sampled periodically. Both experiments indeed corroborate the existence of a periodical mechanism implicated in visual processing. This finding is in agreement with previous reports of periodicities in reaction time distributions (Dehaene, 1993; Latour, 1967; White & Harter, 1969) and in the visual threshold (Busch et al., 2009; Busch & VanRullen, 2010; Landau & Fries, 2012; Latour, 1967; Mathewson et al., 2009, 2010, 2012; Rohenkohl et al., 2012). It is also congruent with the hypothesis that the wagon-wheel illusion under continuous illumination may be caused by a discrete sampling of information (Andrews & Purves, 2005; Andrews et al., 2005; Purves et al., 1996; Rojas et al., 2006; Simpson et al., 2004; VanRullen et al., 2005).

Furthermore, our results revealed the 10 Hz (i.e. Exp. 1) and the 15 Hz (i.e. Exp. 2) frequencies as being correlated with performance. Perhaps the difference found across experiments in the frequency most correlated with accuracy is due to individual differences regarding the rate at which the discrete mechanism samples information, as different individuals participated in Exps. 1 and 2. If individual differences exist in the properties of the sampling, it could be indicated in future studies to use few observers who would each be submitted to many trials, rather than many observers submitted to few trials. Nevertheless, there was enough consistency between subjects to highlight significant combinations of frequency/phase in the classification images of Exps. 1 and 2 as well as in the classification vector of Exp. 2.

Another potential explanation for the different frequencies that correlated with accuracy in Exps. 1 and 2 is that the sampling frequency of most individuals is neither at 10 nor at 15 but somewhere between 10 Hz and 15 Hz. Consistently with this hypothesis, it was shown that the occurrence of the wagon-wheel illusion was correlated with magnitude changes in 13 Hz brain oscillations (VanRullen et al., 2006). It may be that the mechanism revealed in our study also underlies the wagon-wheel illusion under continuous light. This would also mean that sampling information in discrete snapshots is a mechanism that generalizes across (at least a subset of) domains of visual processing.

If a discrete mechanism is implicated in visual information processing, one question is whether this mechanism synchronizes with the external world. Reaction time studies and studies that have shown periodicities in the visual threshold preceding the onset of an eye saccade suggest that sampling is synchronized with the stimulation. Moreover, Landau and Fries (2012) have shown a periodic pattern in detection performance, and they demonstrated that a luminance stimulus that exogenously attracts attention could reset the phase of that periodic sampling. Mathewson et al. (2012) have also shown that a rhythmic visual stimulus can entrain an ongoing neural oscillation, such that the excitability cycles of that oscillation are phase-locked with the (predictable) appearance of the stimuli. Such a synchronization of the sampling with the onset of a stimulus is very similar to what was proposed by the triggered-moment models (Oatley, Robertson, & Scanlan,

[3] Note that we also constructed a temporal classification vector In Exp. 1, but nothing reached statistical significance. We think that this is due in part by the stimuli used in Exp. 1. In fact, only 24 SNR temporal profiles were tested; this is very few to construct a temporal classification vector (normally a classification image or a classification vector is computed by correlating random variations in a search space with performance).

[4] We have used a simple sinus discrete model with two free parameters and tried to minimize the square of the sum of the difference between the model and the human temporal classification vectors using the Nelder-Mead simplex method. The best fit had a frequency of 14.67 Hz and a phase of 0.74 rad ($r^2 = 0.7$).

1969; Sternberg & Knoll, 1973; Ulrich, 1987; Venables, 1960). This class of models proposes that the arrival of a stimulus in the system triggers a moment of processing. Our results support the idea of synchronization between the periodic sampling function and the external environment. Had the phase substantially varied from trial to trial in our experiments, no combination of frequency/phase would have been significantly correlated with accuracy (i.e. entire frequencies, irrespective of phase, would have correlated). It is possible, however, that other discrete mechanisms that are not in phase with stimulus onset exist—for instance, one operating at a lower or higher level of processing and at a different temporal frequency.

On the other hand, it is also possible that oscillations in information intake may result from the continuous impulse response function (IRF) of the visual system (Bowen, 1989; Manahilov, 1995; Rashbass, 1976; Tyler, 1992; Watson, 1986). In that case, the oscillations that we have revealed in our data would not reflect a discrete sampling of information, but would instead be an epiphenomenon of the visual system's reaction to a visual stimulation. The energy in our stimuli was constant across time so, technically, these IRF models predict a constant sensitivity threshold. However, for the sake of the argument, we assumed here that the input of the models is the signal-to-noise variations in our experiments rather than energy per se. For instance, we have tested how Watson's 7-free-parameter working model (Watson, 1982) fitted our data. We minimized the sum of the square of the difference between that model and the human temporal classification vectors obtained in Exp. 2 using the Nelder-Mead simplex method. We found that 35% of the variance in our data could be explained by this continuous model. This compares rather poorly to the fit of a simple sinus discrete model with two free parameters, which explained 70% of the variance in our data. A likelihood ratio, corrected for the number of free parameters (Glover & Dixon, 2004), indicates that our participants' data in Exp. 2 are about 8,125 times more likely to have occurred if this discrete model were true than if the IRF model of Watson (1982) were true. We also tested a few other kinds of IRF models[5], but we found no evidence whatsoever that an IRF model could fit our data as well as a discrete sampling model. However, since we did not test for all the logically possible IRFs, it remains possible that this class of model could predict the oscillations we have observed in our participants' information utilization.

As mentioned above, the energy in our stimuli was constant through time. One consequence of this is a perfect negative correlation between the amount of noise and the amount of signal on each frame: when the signal level was high, the noise level was low and vice versa. Thus, our study does not allow us to tease apart three possibilities: the perceptual oscillation revealed in our classification images reflect an oscillating sensitivity (1) to the signal-to-noise ratio, (2) to the signal only, or (3) to the noise only.

To summarize, our results are in agreement with the hypothesis of a discrete sampling of visual information that synchronizes with the visual stimulation, a hypothesis that has gained popularity recently. Indeed, we showed oscillations in the time course of information utilization. We also demonstrated that the frequency and phase with which the visibility of a stimulus is varied influences performance: the information sampling function of our participants was synchronized with the beginning of the trial, and they performed better when the frequency of the signal-to-noise ratio varied at a rate of about 10–15 Hz. Yet, few investigations have been conducted on this topic so far, and a definitive conclusion regarding the discrete nature of visual sampling seems premature. For instance, an IRF model, which is not discrete by nature, could also predict oscillations in perception. More research will be needed to clarify both the perceptual and brain mechanisms underlying our results.

Acknowledgments

This work was supported by grants from the Natural Sciences and Engineering Research Council of Canada (CRSNG) to Martin Arguin and Frédéric Gosselin and by a scholarship from the Fonds Québécois de Recherche sur la Nature et les Technologies (FQRNT) to Caroline Blais.

References

Andrews, T., & Purves, D. (2005). The wagon-wheel illusion in continuous light. *Trends in Cognitive Sciences, 9*, 261–263.

Andrews, T., Purves, D., Simpson, W. A., & VanRullen, R. (2005). The wheels keep turning. *Trends in Cognitive Sciences, 9*, 560–561.

Blais, C., Fiset, D., Jolicoeur, P., Arguin, M., Bub, D., & Gosselin, F. (2009). Reading between eye saccades. *PLoS ONE, 4*(7), e6448. http://dx.doi.org/10.1371/journal.pone.0006448.

Bowen, R. W. (1989). Two pulses seen as three flashes: A superposition analysis. *Vision Research, 29*, 409–417.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*, 433–436.

Busch, N. A., Dubois, J., & VanRullen, R. (2009). The phase of ongoing EEG oscillations predicts visual perception. *Journal of Neuroscience, 29*, 7869–7876.

Busch, N. A., & VanRullen, R. (2010). Spontaneous EEG oscillations reveal periodic sampling of visual attention. *Proceedings of the National Academy of Science USA, 107*, 16048–16053.

Chauvin, A., Worsley, K. J., Schyns, P. G., Arguin, M., & Gosselin, F. (2005). Accurate statistical tests for smooth classification images. *Journal of Vision, 5*, 659–666.

Dehaene, S. (1993). Temporal oscillations in human perception. *Psychological Science, 4*, 264–270.

[5] We have also tried to define the impulse response function with two, three or four Gaussian radial-basis functions—Gaussian radial-basis functions are a popular choice for approximating arbitrary functions. The goodness of fit between the 6-parameter model (i.e. the sum of 2 Gaussians) and our data is about $r^2 = 0.58$; and the goodness of fit between the 9-parameter model (i.e. the sum of 3 Gaussians) and our data is about $r^2 = 0.68$ (the goodness of fit between the 12-parameter model (i.e. the sum of 4 Gaussians) and our data is again about $r^2 = 0.68$—so there is no need to go any further). Likelihood ratios corrected for the number of free parameters indicate that the human data is about 679 and 5444 times more likely to have occurred if a simple sinus discrete model were true than if the 6-parameter and the 9-parameters Gaussian radial-basis-function continuous model were true, respectively. Finally, for the sake of completeness, we have modeled *discrete* Gaussian radial-basis-function models and tried to fit them to the human data: the goodness of fit between the 6-parameter model and our data is about $r^2 = 0.68$ (against $r^2 = 0.58$ for the 6-parameter continuous model; likelihood ratio = 1.31); the goodness of fit between the 9-parameter model and our data is about $r^2 = 0.89$ (against $r^2 = 0.68$ for the 9-parameter continuous model; likelihood ratio = 2.91); and the goodness of fit between the 12-parameter model and our data is almost $r^2 = 1$ (against $r^2 = 0.68$ for the 12-parameter continuous model; likelihood ratio = 32).

Eckstein, M. P., & Ahumada, A. J. (2002). Classification images: A tool to analyze visual strategies. *Journal of Vision, 2* (i-i).

Fiset, D., Blais, C., Arguin, M., Tadros, K., Éthier-Majcher, C., Bub, D., et al. (2009). The spatio-temporal dynamics of visual letter recognition. *Cognitive Neuropsychology, 26*, 23–35.

Glover, S. R., & Dixon, P. (2004). Likelihood ratios: A simple and flexible statistic for empirical psychologists. *Psychonomic Bulletin & Review, 11*, 791–806.

Gold, J. M., & Shubel, E. (2006). The spatiotemporal properties of visual completion measured by response classification. *Journal of Vision, 6*, 356–365.

Gosselin, F., & Schyns, P. G. (2004). An image is worth thousands of trials: Rendering the use of visual information from spiking neurons to recognition. *Cognitive Science, 28*, 141–146.

Holcombe, A. O., Clifford, C. W. G., Eagleman, D. M., & Pakarian, P. (2005). Illusory motion reversal in tune with motion detectors. *Trends in Cognitive Sciences, 9*, 559–560.

Jones, M. R., Moynihan, H., MacKenzie, N., & Puente, J. (2002). Temporal aspects of stimulus-driven attending in dynamic arrays. *Psychological Science, 13*, 313–319.

Kline, K. A., & Eagleman, D. M. (2008). Evidence against the snapshot hypothesis of illusory motion reversal. *Journal of Vision, 8*, 1–5.

Kline, K. A., Holcombe, A. O., & Eagleman, D. M. (2004). Illusory motion reversal is caused by rivalry, not by perceptual snapshots of the visual field. *Vision Research, 44*, 2653–2658.

Kline, K. A., Holcombe, A. O., & Eagleman, D. M. (2006). Illusory motion reversal does not imply discrete processing: Reply to Rojas et al.. *Vision Research, 46*, 1158–1159.

Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I., & Scroeder, C. E. (2008). Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science, 320*, 110–113.

Landau, A. N., & Fries, P. (2012). Attention samples stimuli rhythmically. *Current Biology, 22*, 1000–1004.

Latour, P. L. (1967). Evidence of internal clocks in the human operator. *Acta Psychologica, 27*, 341–348.

Manahilov, V. (1995). Spatiotemporal visual response to suprathreshold stimuli. *Vision Research, 35*, 227–237.

Mathewson, K. E., Fabiani, M., Gratton, G., Beck, D. M., & Lleras, A. (2010). Rescuing stimuli from invisibility: Inducing a momentary release from visual masking with pre-target entrainment. *Cognition, 115*, 186–191.

Mathewson, K. E., Gratton, G., Fabiani, M., Beck, D. M., & Ro, A. (2009). To see or not to see: Pre-stimulus alpha phase predicts visual awareness. *The Journal of Neuroscience, 29*, 2725–2732.

Mathewson, K. E., Prudhomme, C., Fabiani, M., Beck, D. M., Lleras, A., & Gratton, G. (2012). Making waves in the stream of consciousness: Entraining oscillations in EEG alpha and fluctuations in visual awareness with rhythmic visual stimulation. *Journal of Cognitive Neuroscience, 24*, 2321–2333.

Neri, P., & Heeger, D. J. (2002). Spatiotemporal mechanisms for detecting and identifying image features in human vision. *Nature Neuroscience, 5*, 812–816.

Neri, P., & Levi, D. M. (2008). Temporal dynamics of directional selectivity in human vision. *Journal of Vision, 8*, 1–11.

Oatley, K., Robertson, A., & Scanlan, M. P. (1969). Judging the order of visual stimuli. *Quaterly Journal of Experimental Psychology, 21*, 172–179.

Pelli, D. G. (1997). The Video toolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*, 437–442.

Poline, J.-B., Worsley, K. J., Evans, A. C., & Friston, K. J. (1997). Combining spatial extent and peak intensity to test for activations in functional imaging. *Neuroimage, 5*, 83–96.

Purves, D., Paydarfar, J. A., & Andrews, T. (1996). The wagon-wheel illusion in movies and reality. *Proceedings of the National Academy of Science, 93*, 3693–3697.

Rashbass, C. (1976). Unification of two contrasting models of the visual increment threshold. *Vision Research, 16*, 1281–1283.

Reddy, L., Remy, F., Vayssiere, N., & VanRullen, R. (2011). Neural correlates of the continuous Wagon-Wheel Illusion: A functional MRI study. *Human Brain Mapping, 32*, 163–170.

Rohenkohl, G., Cravo, A., Wyart, V., & Nobre, A. C. (2012). Temporal expectation improves the quality of sensory information. *The Journal of Neuroscience, 32*, 8424–8428.

Rojas, D., Carmona-Fontaine, C., López-Calderón, J., & Aboitiz, F. (2006). Do discreteness and rivalry coexist in illusory motion reversals? *Vision Research, 46*, 1155–1157.

Simpson, W. A., Shahani, U., & Manahilov, V. (2004). Illusory percepts of moving patterns due to discrete temporal sampling. *Neuroscience Letters, 375*, 23–27.

Sternberg, S., & Knoll, R. L. (1973). The perception of temporal order: Fundamental issues and a general model. In S. Kornblum (Ed.), *Attention and Performance IV* (pp. 629–685). New York: Academic Press.

Tyler, C. W. (1992). Psychophysical derivation of the impulse response through generation of ultrabrief responses: Complex inverse estimation without minimum-phase assumptions. *Journal of the Optical Society of America A, 9*, 1025–1040.

Ulrich, R. (1987). Threshold models of temporal-order judgments evaluated by a ternary response task. *Perception & Psychophysics, 42*, 224–239.

VanRullen, R. (2006). The continuous Wagon wheel illusion is object-based. *Vision Research, 46*, 4091–4095.

VanRullen, R. (2007). The continuous Wagon wheel illusion depends on, but is not identical to neuronal adaptation. *Vision Research, 47*, 2143–2149.

VanRullen, R., Carlson, T., & Cavanagh, P. (2007). The blinking spotlight of attention. *Proceedings of the National Academy of Science, 104*, 19204–19209.

VanRullen, R., & Koch, C. (2003). Is perception discrete or continuous? *TRENDS in Cognitive Sciences, 7*, 207–213.

VanRullen, R., Reddy, L., & Koch, C. (2005). Attention-driven discrete sampling of motion perception. *Proceedings of the National Academy of Science, 102*, 5291–5296.

VanRullen, R., Reddy, L., & Koch, C. (2006). The continuous Wagon wheel illusion is associated with changes in EEG power around 13 Hz. *Journal of Neuroscience, 26*, 502–507.

Venables, P. H. (1960). Periodicity in reaction time. *British Journal of Psychology, 51*, 37–43.

Vinette, C., Gosselin, F., & Schyns, P. G. (2004). Spatio-temporal dynamics of face recognition in a flash: It's in the eyes! *Cognitive Science, 28*, 289–301.

Watson, A. B. (1982). Derivation of the impulse response: Comments on the method of Roufs and Blommaert. *Vision Research, 22*, 1335–1337.

Watson, A. B. (1986). Temporal Sensitivity. In Boff, K., Kaufman, L., & Thomas, J. (Eds.), Handbook of Perception and Human Performance New York: Wiley.

Watson, A. B., & Pelli, D. G. (1983). QUEST: A Bayesian adaptive psychometric method. *Perception and Psychophysics, 33*, 113–120.

White, C. T., & Harter, M. R. (1969). Intermittency in reaction time and perception, and evoked response correlates of image quality. *Acta Psychologica, 30*, 368–377.