Motion processing impaired by transient spatial attention: Potential implications for the magnocellular pathway

Motion processing impaired by transient spatial attention: potential implications for the Magnocellular pathway

Andrea Pavan[1*], Seyma Koc Yilmaz[2,3], Hulusi Kafaligonul[2,3], Luca Battaglini[4], and Steven P. Blurton[5]

[1] Department of Psychology, University of Bologna, Viale Berti Pichat, 5, 40127, Bologna, Italy

[2] National Magnetic Resonance Research Center (UMRAM), Bilkent University, 06800 Ankara, Turkey

[3] Interdisciplinary Neuroscience Program, Aysel Sabuncu Brain Research Center, Bilkent University, 06800, Ankara, Turkey

[4] Dipartimento di Psicologia Generale, University of Padova, Via Venezia 8, 35131 Padova, Italy

[5] Department of Psychology, University of Copenhagen, Øster Farimagsgade 2A, 1353 København, Denmark

***Corresponding Author**

Andrea Pavan

University of Bologna

Department of Psychology

Viale Berti Pichat, 5, 40127, Bologna, Italy

Email: andrea.pavan2@unibo.it

Abstract

Spatial cues presented prior to the presentation of a static stimulus usually improve its perception. However, previous research has also shown that transient exogenous cues to direct spatial attention to the location of a forthcoming stimulus can lead to reduced performance. In the present study, we investigated the effects of transient exogenous cues on the perception of briefly presented drifting Gabor patches. The spatial and temporal frequencies of the drifting Gabors were chosen to mainly engage the magnocellular pathway. We found better performance in the motion direction discrimination task when neutral cues were presented before the drifting target compared to a valid spatial cue. The behavioral results support the hypothesis that transient attention prolongs the internal response to the attended stimulus, thus reducing the temporal segregation of visual events. These results were complemented by applying a recently developed model for perceptual decisions to rule out a speed-accuracy trade-off and to further assess cueing effects on visual performance. In a model-based assessment, we found that valid cues initially enhanced processing but overall resulted in less efficient processing compared to neutral cues, possibly caused by reduced temporal segregation of visual events.

Keywords: Visuospatial attention, exogenous cue, temporal segregation, computational modelling

Introduction

Observers are typically faster and more accurate at detecting a target that appears in an attended location rather than in an unattended location (Posner, 1980). Furthermore, spatial attention attracted by a transient exogenous cue can enhance performance in tasks involving spatial resolution (Carrasco & Yeshurun, 2009) such as texture segmentation (Yeshurun & Carrasco, 2000), Vernier stimuli, and Landolt-squares (Yeshurun & Carrasco, 1999). However, in some cases and under specific stimulus conditions, a transient exogenous cue can negatively affect the performance. For example, Yeshurun and Levy (2003) reported that an exogenous cue impairs performance in a temporal gap detection task in which observers had to report whether they perceived two disks which appeared for 47 ms with a delay interval (ISI) between 11 and 34 ms, or whether they perceived only a single disk. Hein, Rolke and Ulrich (2006) presented a dot either in the left or right visual hemi-field. After a variable stimulus onset asynchrony (SOA; ranging from 17 to 83 ms) a second dot spatially adjacent to the first one was presented. The temporal order was randomly varied from trial to trial. Participants had to indicate which of the two dots in the temporal sequence was presented first. When the exogenous cue indicated the visual hemi-field where the two dots were going to appear (i.e., valid cue condition), participants' performance significantly dropped compared to an un-cued condition in which there was no spatial cue indicating the target's visual hemi-field. Nicol and colleagues (2009) replicated the findings of Hein et al. (2006) and showed that the temporal resolution was better in un-cued locations. Additionally, Yeshurun and Marom (2008) found that transient attention can prolong the apparent duration of an attended disk that appeared for a duration of 70 ms. Similarly, Yeshurun and Hein (2011) showed that when the location of a two-frame apparent motion sequence was cued by a transient exogenous cue, participants' performance on direction discrimination significantly dropped.

These results are consistent with the hypothesis that transient attention prolongs the internal response to the attended stimulus (Yeshurun & Marom, 2008). Such attentional prolongation of the response reduces temporal sensitivity and thus temporal segregation. Therefore, events occurring at different time points might be integrated into a single event (Yeshurun & Hein, 2011). Yeshurun and Hein (2011) suggested that transient attention directed to a spatial location facilitates the activity of the parvocellular system (i.e., visual cells with small receptive fields, with superior spatial acuity but poor temporal acuity and prolonged response

persistence; Denison, Vu, Yacoub, Feinberg, & Silver, 2014; Derrington & Lennie, 1984; Livingstone & Hubel, 1988), and suppresses the magnocellular system (Yeshurun, 2004).

Yeshurun (2004) used isoluminant stimuli to ensure that the performance in a temporal gap detection task was mainly supported by the parvocellular system. They found that, under these stimulus conditions, the deleterious effect of the exogenous cue was attenuated or eliminated. Megna, Rocchi, and Baldassi (2012) evaluated the influence of transient spatial cues on the spatial and temporal dimensions of perceptual filters. They presented a spatiotemporal noise matrix in each visual hemi-field. The matrix was composed of bars changing in luminance over time. One of the matrices contained the target, which was a luminance increment of the middle bar spatiotemporally centered in the noise matrix. In the cueing conditions, a small horizontal bar, which preceded the target by 120 or 250 ms, indicated the visual hemi-field where the target would appear. The observers had to report which matrix (left or right) included the target. When the matrices were presented at a small eccentricity (3 deg), transient attention caused the perceptual templates to be sharper in space, as reflected by high spatial frequency components and a large temporal integration window. These results further support the parvocellular explanation of transient exogenous attention by showing that in the presence of timed spatial exogenous cues observers rely on noisy evidence lasting longer and with finer spatial arrangement. Yeshurun and Sabo (2012) provided further evidence that transient attention facilitates parvocellular activity using two pedestal paradigms: a pulse-pedestal paradigm favoring parvocellular activity and a steady-pedestal paradigm favoring the magnocellular system. They found that an exogenous cue that attracts transient attention to the location of the upcoming target improved performance with a pulse-pedestal but had no effect with the steady-pedestal. Though they did not find a decrease in the performance with the steady-pedestal paradigm, most of the previous studies mentioned above showed inefficient temporal processing in the cued region, and in line with Yeshurun's (2004) hypothesis of magnocellular suppression.

Based on the previous findings, it is likely that transient exogenous attention impairs accuracy in a task mainly mediated by the magnocellular system and in which the temporal information is a crucial factor such as in motion perception. According to Adelson and Bergen (1985), motion processing is a combined response between spatial and temporal filters, therefore one would expect that if temporal resolution is reduced by presenting a transient cue, then even motion perception will be affected. Attentional modulation of motion perception was supported

by a reversed apparent motion paradigm (Yeshurun & Hein, 2011). In this paradigm (Experiment 4 in Yeshurun & Hein, 2011), observers were briefly presented with two consecutive sine-wave gratings with a 90-deg phase shift, which produces the perception of motion in the direction of the phase shift with short ISIs between frames and in the opposite direction with longer ISIs (Takeuchi & DeValois, 1997). When the location of the target (i.e., one of six possible locations around the fixation) was exogenously cued (i.e., valid cue condition), the observers were less likely to report reversed apparent motion compared to when spatial attention was not allocated to any specific location (i.e., neutral cue condition). Transient attention reduced the perception of reversed motion in a similar way to increasing the spatial frequency of the gratings (thus biasing parvo-dominated processing), implying that the temporal response was more sustained in this condition, allowing reversed motion to be perceived only with longer ISIs. On the other hand, Liu, Fuller and Carrasco (2006) showed that transient attention could increase performance in a motion task. They presented two random dot kinematograms (RDKs), and the observers had to judge which pattern was more coherent and report the motion direction. They found that an exogenous cue that preceded one of the two RDK increased subjective appearance of motion coherence and improved motion direction discrimination. However, the processing of global motion occurs at a late stage of visual processing and requires the integration of different local motion cues into a global motion percept (Newsome, Britten, & Movshon, 1989; Newsome & Pare, 1988). Therefore, one could argue that transient attention may alter temporal processing only at the local motion level.

In the present study, we investigated whether transient attention negatively affects motion perception from drifting Gabor patches. Building on the above-mentioned studies, we tested the hypothesis that exogenous attention prolongs the internal response to the stimuli presented in the cued location by reducing temporal sensitivity and thus impairing motion direction discrimination (Yeshurun & Hein, 2011). This is because the prolongation of the response to the cued stimulus could reduce the temporal segregation of visual events. Such modulation in the temporal domain is expected to mainly affect the processing of moving stimuli which have spatiotemporal characteristics that are appropriate and mainly engage the magnocellular system.

There has been extensive research on the anatomical projections and functional specialization of M and P cells (Denison et al., 2014; Tootell, Hamilton, & Switkes, 1988; Zhang, Zhou, Wen, & He, 2015) and debates on whether they form pathways that carry out

distinct processing of different stimulus characteristics (Kaplan, 2012). Despite the views challenging parallel processing and arguing for shared magnocellular and parvocellular contribution in the processing of stimulus characteristics (Sincich & Horton, 2005), studies on temporal and spatial tuning properties, contrast gain of P and M cells (Alitto, Moore, Rathbun, & Usrey, 2011; Kaplan & Benardete, 2001; Movshon, Kiorpes, Hawken, & Cavanaugh, 2005) and lesion studies (Merigan, Katz, & Maunsell, 1991; Schiller, Logothetis, & Charles, 1990) suggest that specific combinations of stimulus parameters can be used to index the activity of M/P system, or at least bias the response of either. In other words, the processing of certain stimuli can be not exclusively, but predominantly carried out by one system.

Therefore, we used drifting Gabor patches with a temporal frequency (i.e., 20 Hz; see Liu, Bryan, Miki, Woo, Liu, & Elliott, 2006) and a range of low spatial frequencies optimally engaging magnocellular processing (Derrington & Lennie, 1984). The rationale was that if transient attention prolongs the temporal response to stimuli presented in the cued location, thus favoring the parvocellular system over the magnocellular system, then we would expect decreased accuracy on a motion direction discrimination task when valid exogenous spatial cues are used.

Methods

Participants

Twenty-two observers with normal or corrected-to-normal vision, who were naïve to the purpose of the study, took part in the experiment. Two participants were excluded from data analysis, since their average accuracy scores were below the pre-determined 75% accuracy threshold (Aydin, Ogmen, & Kafaligonul, 2021; Hung & Carrasco, 2021). Accordingly, the data of the remaining 20 participants (12 females and 8 males, with a mean age of 25.8 ± 3.9 years) were included in the further analysis. Prior to their participation, all observers were informed about experimental procedures and gave written informed consent. All experimental procedures were in accordance with the World Declaration of Helsinki (2013) and approved by the local Ethics Committee at Bilkent University.

Apparatus

Stimuli were generated using MATLAB with the Psychophysics Toolbox (Brainard, 1997; Kleiner, Brainard, & Pelli, 2007; Pelli, 1997) and displayed on a NEC MultiSync LCD 2190UXP 21-inch monitor with a refresh rate of 60 Hz and a spatial resolution of 1600 x 1200 pixels. The screen luminance was measured and calibrated using a SpectroCAL (Cambridge Research Systems, Rochester, Kent, UK) photometer. The stimuli were presented on a gray background (30.7 cd/m^2). Gaze coordinates were recorded using a monocular eye-tracker (Eye Trac 6, Applied Science Laboratories, Bedford, MA, USA) at a sampling rate of 50 Hz. Participants sat in a dark room at a viewing distance of 57 cm from the monitor. Viewing was binocular. Head movements were constrained by a chin and forehead rest.

Stimuli

Stimuli were drifting Gabor patches and consisted of a sinusoidal carrier enveloped by a stationary Gaussian window with $\sigma = 0.68$ deg (approximately 3.0 deg in diameter). Narrow Gabor patches were used to limit spatial summation. In fact, Tadin, Lappin, Gilroy, and Blake (2003), using a motion direction discrimination task, found that at high contrast (i.e., 92%), the direction discrimination of a drifting Gabor patch was impaired (i.e., higher duration thresholds) with increasing the stimulus size (up to 5 deg), suggesting an effect of centre-surround antagonism at the level of individual neurons (i.e., surround suppression; Glasser & Tadin, 2011; Tadin & Lappin, 2005). However, when using low contrast Gabors (2%), the authors found lower duration thresholds with increasing stimulus size, suggesting a shift from surround suppression to spatial summation for low contrast and wide Gabor patches (Tadin et al., 2003).

We used five spatial frequencies (0.15, 0.25, 0.35, 0.45, and 0.55 c/deg). The target Gabor patch had a Michelson contrast of 0.22 (with a min luminance of 26.1 cd/m^2 and a max luminance of 40.8 cd/m^2) and drifted at 20 Hz. The temporal frequency and contrast were selected to mainly tap the magnocellular system (Liu et al., 2006; Merigan, Byrne, & Maunsell, 1991; Merigan & Maunsell, 1990). The target Gabor patch could appear in one of eight possible locations equidistant around the fixation point. The distance between the central fixation point and the center of the Gabor patch was 7.0 deg. A cue was presented before the presentation of the target Gabor patch which consisted of a circular frame (approximately 5.0 deg in diameter) and could be either informative of the location of the target Gabor patch (i.e., valid cue

condition) or uninformative (i.e., neutral cue condition) (Yeshurun & Hein, 2011). The circular cues appeared simultaneously at all the locations around the fixation point. The contrast polarity of the cues (i.e., white cue and black cue) was counterbalanced in two separate versions of the experiment to limit the contribution of polarity-specific paracontrast masking effect, which was shown to be minimal at the stimulus onset asynchrony (SOA) used in the current experiment (Kafaligonul, Breitmeyer, & Ogmen, 2009). In the white cue version of the experiment, neutral cues were all black whereas the valid cue was white. In the other version of the experiment neutral cues were white whereas the valid cue was black (see Figure 1). The two versions of the experiment were counterbalanced across participants. In the fixation training (see Figure S.1A in the Supplementary Material for details), a flickering random-dot pattern with half black and half white pixels was used (Guzman-Martinez, Leung, Franconeri, Grabowecky, & Suzuki, 2009) together with a fixation marker combining bull's eye and cross hair that was shown to minimize eye movements (Thaler, Schütz, Goodale, & Gegenfurtner, 2013).

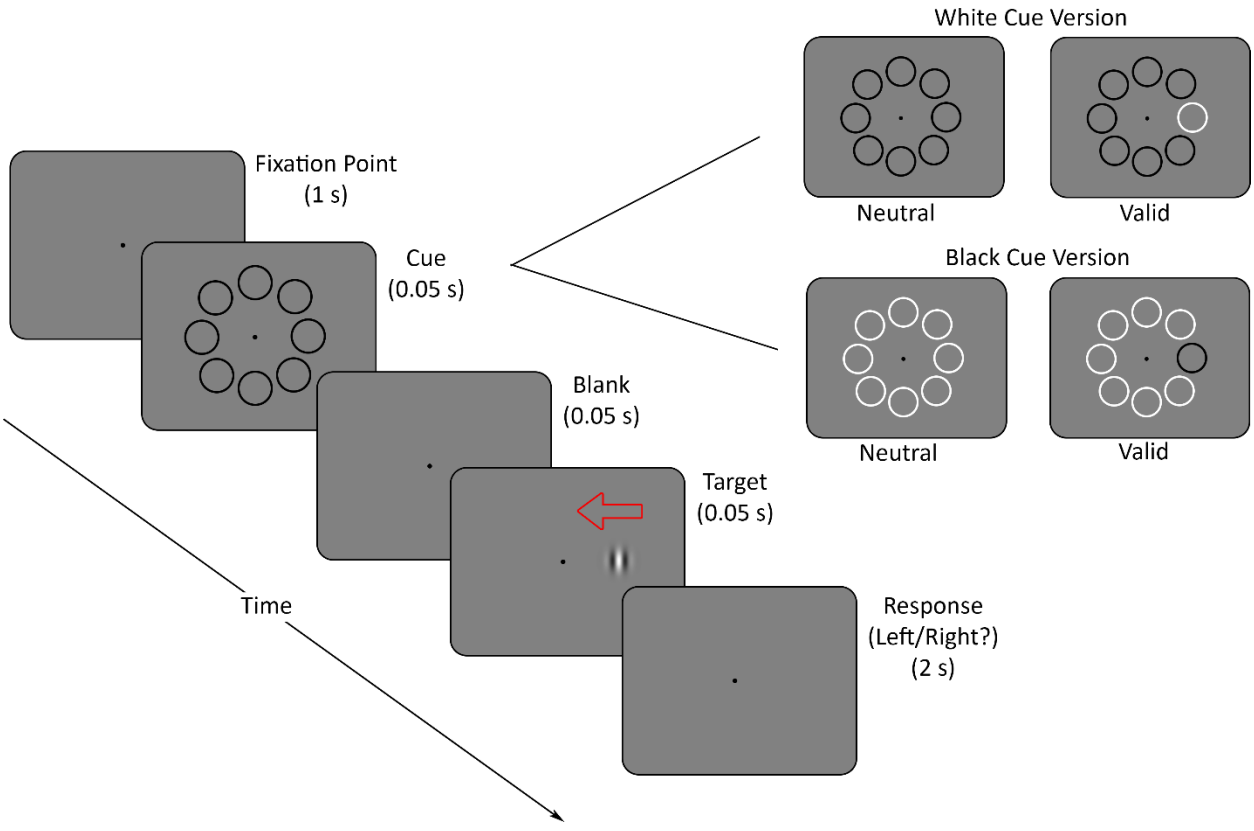

Figure 1. Schematic representation of the stimuli and procedure used in the experiment. The valid cue is represented by the single white/black circular frame with an opposite contrast

polarity. In the neutral cue condition, all the circular frames were either black or white and were non-informative of the forthcoming target's location. Observers had to discriminate the direction of the target Gabor patch. The red arrow indicates leftward motion direction and was not displayed during the experiment. The contrast of the Gabor patch was increased for illustrative purposes.

Procedure

At the beginning of each session, participants completed a brief training to maintain central fixation. The procedure of the fixation training is summarized in the Supplementary Material (Figure S.1B). The participants also performed an initial experimental training on motion direction discrimination and performed as many blocks as needed to get an accuracy ≥ 0.75 (Aydin, Ogmen, & Kafaligonul, 2021; Hung & Carrasco, 2021). On average, they completed 1.15 ± 0.37 training blocks, with the same length and procedure as the main experimental blocks. There were eight blocks of equal length in the main experiment. The eye-tracker was recalibrated using a nine-point standard procedure before each block.

In the main experiment, each trial started with a fixation point with a duration of 1.0 s and a pure tone with a duration of 0.05 s (frequency: 500 Hz). The fixation point was followed by a display including the circular cues for 0.05 s. After a 0.05 s blank screen (i.e., SOA of 0.1 s), the drifting Gabor patch appeared at one of the eight locations for 0.05 s. These short durations triggered a transient response and prevented eye movements towards the Gabor patch (Carpenter, 1988; Hein et al., 2006; Wright & Ward, 2008). On half of the trials the cue was valid and indicated the location of the forthcoming drifting Gabor patch. On the other half of the trials, the cue was neutral and uninformative of the Gabor patch location (see Figure 1). The motion direction of the Gabor patch (i.e., rightward, or leftward) was randomized across trials. Observers had to discriminate the motion direction of the Gabor patch (Method of Single Stimuli; Morgan, Dillenburger, Raphael, & Solomon, 2012), by pressing the 'A' key if it was perceived to drift leftwards or the 'L' key if it was perceived to drift rightwards. Observers were instructed to respond as accurately and quickly as possible. They had 2 s to respond. Feedback was provided on a trial basis after each response during the inter-trial interval, by displaying at the center of the screen the response time (in milliseconds) in green for correct responses and in red for incorrect responses. The trials, in which the observer did not respond within the specified

time window, were recorded as unanswered. In addition, instead of the response time, a warning (i.e., "Please, respond faster") was displayed at the center of the screen during the feedback period. Accuracy and reaction times were measured. Gaze coordinates were recorded simultaneously with the behavioral responses.

There were 2 (valid vs. neutral cue) x 5 (spatial frequencies) conditions. Observers performed 192 trials per spatial frequency for a total of 960 trials (i.e., 120 trials per block), which were presented in a random order. For each spatial frequency, there were 96 presentations of the valid cue and 96 presentations of the neutral cue. Gabor patches with each spatial frequency were displayed 24 times at each of the eight spatial locations.

Analysis of Eye Movements, Accuracy Scores, and Reaction Times

To ensure that the stimuli were processed at the periphery and covert attention was effectively manipulated, eye movements during the experiment were analyzed in *ASL Results* software (Applied Science Laboratories, Bedford, MA, USA). First, the stimulus presentation period (i.e., the first 1.15 s) in each trial was extracted by discarding the response period. Second, the fixations during this period were detected in each trial. If the gaze was stable for a minimum of 0.1 s within a maximum of 1.0 deg, then it was regarded as the start of a fixation. If three consecutive samples of gaze coordinates deviated from this initial position or pupil data was lost for longer than 0.2 s (typical length of a blink), then it was regarded as the end of the fixation. Following the detection of fixations, the fixation coordinates in each trial were compared to a 2 x 2 deg central square window around the fixation point. Finally, trials in which fixations were not detected and trials with fixation coordinates outside this window (a total of 15.1 ± 9.0%) were excluded from further analyses. After the exclusion of trials with no fixations and trials with excessive eye movements (i.e., trials in which participants failed to fixate at the central window), it was checked whether a similar number of trials were analyzed in each condition. A repeated measures ANOVA with the mean number of removed trials as dependent variable showed that the number of removed trials did not significantly differ between conditions ($M_{removed}$ = 14.5 ± 0.5 trials per condition, Cue type: $F_{1, 19} = 0.154$, $p = .70$, $\eta^2_p = .01$; Spatial frequency: $F_{4, 76} = 0.600$, $p = .66$, $\eta^2_p = .03$; Cue type x Spatial frequency interaction: $F_{4, 76} = 0.147$, $p = .96$, $\eta^2_p = .01$). In addition, the two subgroups of participants, who performed different versions of the experiment (i.e., white cue and black cue), were compared using a Mann-Whitney *U* test due to a

violation of the normality of residuals for accuracy scores. Their accuracy scores were comparable ($U = 41.0$, $p = .52$).

Following trial exclusion based on eye movements, accuracy was calculated as the proportion of correct responses in each condition. Median and mean reaction times were calculated in each condition for correct responses only. Reaction times data for correct responses only were checked for outliers using the method of the median absolute deviation (MAD) for asymmetric distributions using a cutoff of three (Leys, Ley, Klein, Bernard, & Licata, 2013; Rousseeuw & Croux, 1993). On average we found 6.32% (SD = 1.55%) of outlier RTs. Outlier RTs were included in the raw RT analysis and analysis of RT distribution. Additionally, there were six participants who missed only a few trials. Missed trials were 0.04% of the total trials (SD = 0.06%) and were excluded from subsequent analyses.

According to Shapiro-Wilk tests, residuals of accuracy scores and reaction times (RTs) were not normally distributed ($W = 0.903$, $p < .001$ and $W = 0.944$, $p < .001$, for accuracy scores and reaction times, respectively). Since the common nonparametric tests such as Friedman are inadequate to evaluate interactions, we used the Aligned Rank Transform (ART), a procedure for the nonparametric analysis of variance in multifactor designs (Higgins, Blair, & Tashtoush, 1990; Higgins & Tashtoush, 1994; Salter & Fawcett, 1993; Wobbrock, Findlater, Gergle, & Higgins, 2011). We applied the alignment and rank transformation to the mean proportion of correct responses and the mean reaction times. With this technique, parametric tests (typically ANOVA) can be implemented once the data is aligned and ranked for each main and interaction effect. The pairwise comparisons for the significant main and interaction effects were conducted using the ART-C procedure, which was developed for reliable contrast tests (Elkin, Kay, Higgins, & Wobbrock, 2021).

Analysis of response time distributions

In addition to the analysis of mean RT and response accuracy mentioned above, we further investigated the observed data based on the level of RT distributions. With this approach, we sought to gain a more detailed insight into how both response accuracy and response speed were affected by the different cues. First, we present a description of the data, followed by a model-based assessment. The aim of the descriptive approach is to illustrate the patterns observed in the data which the model needs to explain. Statistical testing was then performed in

the model-based assessment. For the model-based approach, we fitted a computational RT model to the data, which can simultaneously account for response speed and response accuracy. For the distributional analyses, we pooled responses across the different spatial frequencies to get less noisy estimates, whereas the model was fitted to all conditions separately, as is common in RT modeling.

A common approach to compare RT distributions is the quantile-quantile (Q-Q) plot. In this representation, quantiles of one distribution are plotted against those of another distribution. If the two distributions are similar, the Q-Q plot shows a straight line along the main diagonal. If one distribution dominates the other, the line will lie below or above the main diagonal. For the Q-Q plot, we estimated 20 quantiles (2.5%, 7.5%, …, 97.5%) of the valid and neutral cue RT distribution. The quantiles were obtained from the nonparametric Kaplan-Meier estimate (Kaplan & Meier, 1958), treating incorrect responses as censored observations in this analysis. That is, in trial with incorrect responses, the correct response was not observed but assumed to have occurred any time after the incorrect response. Further, we investigated response accuracy over different response latencies with a conditional accuracy function (CAF). The CAF plots response accuracy as a function of response latency, indicating the presence (or absence) of fast and/or slow errors. To estimate the CAF, we created five RT bins (0–20%, …, 80–100%) from the RT distribution of valid and neutral cues, and estimated response accuracy separately for each bin. Confidence intervals for accuracy were estimated by a bootstrap procedure using a mixed-effects design (cue type as fixed effect, participants as random effect).

Response time model analysis

To further investigate cueing effects on RT distributions and to analyze response times and response accuracy in a single analysis, we fitted the Poisson random walk model (Blurton, Kyllingsbæk, Nielsen, & Bundesen, 2020) to the observed data. The Poisson random walk model is an RT model for speeded responses in an n-alternative forced-choice task. It is an extension of the Theory of Visual Attention (TVA, Bundesen, 1990) to model perceptual decisions in tasks with mutually confusable stimuli. It is assumed that, upon stimulus presentation, tentative categorizations of the form "stimulus x belongs to category i" are made continuously over time (Figure 2A), until one perceptual category has accumulated a critical number of tentative categorizations more than all other categories (sequential sampling). With sequential sampling

models for RT, the entire observed dataset can be analyzed in a single analysis because the model predicts the whole RT distribution of correct and incorrect responses (including response accuracy).

In the special case of two perceptual categories as in the present study, the Poisson random walk model is very similar to the diffusion decision model (DDM, Ratcliff & McKoon, 2008). Like the DDM, the Poisson random walk model is based on a noisy stochastic process between absorbing barriers, representing response criteria for each of the two alternatives (Figure 2B). In the current application of the model, the two perceptual categories may be represented with *L* (Gabor patch appears moving leftward) and *R* (Gabor patch appears moving rightward). Each perceptual category is associated with a response criterion (k_L and k_R). These criteria can be set farther apart, resulting in slower but more accurate decisions or closer, resulting in faster but less accurate decisions (speed-accuracy trade-off).

The benefit of applying computational RT models is that they allow estimation of cognitively meaningful parameters. As already mentioned, the Poisson random walk contains a response criterion (k) that determines how many categorizations are needed to conclusively select a perceptual category. Further, categorizations for each category i are made with a rate v_i, representing the processing speed for that category. The higher the processing rate, the faster tentative categorizations are collected. In a task with two categories and mutually confusable stimuli, there are two possible categorizations with different rates, one categorization is correct, the other incorrect. Clearly, the rate for the correct categorization (v_c) will usually be greater than that for the incorrect categorization (v_e). It is useful to define $p = v_c/(v_c + v_e)$ as the probability that a given categorization is for the correct perceptual category. This probability mainly depends on stimulus confusability: with higher confusability, p will be closer to 0.5, whereas with highly discriminable stimuli p will be close to unity. This parameter therefore represents the quality of information sampled from a stimulus. Due to the normalization, we call this parameter normalized processing rate. It is further useful to define $C = v_c + v_e$, that is, the sum of processing rates, which can also be conceived as visual processing capacity (Bundesen, 1990; Townsend & Ashby, 1983). C defines the processing limit (parallel processing with fixed capacity; Shibuya & Bundesen, 1988) and therefore reflects the efficiency with which a stimulus is processed. Finally, there is a non-decision time (T_0), representing the latency of all processes that are otherwise not described by the model, such as motor execution or response selection.

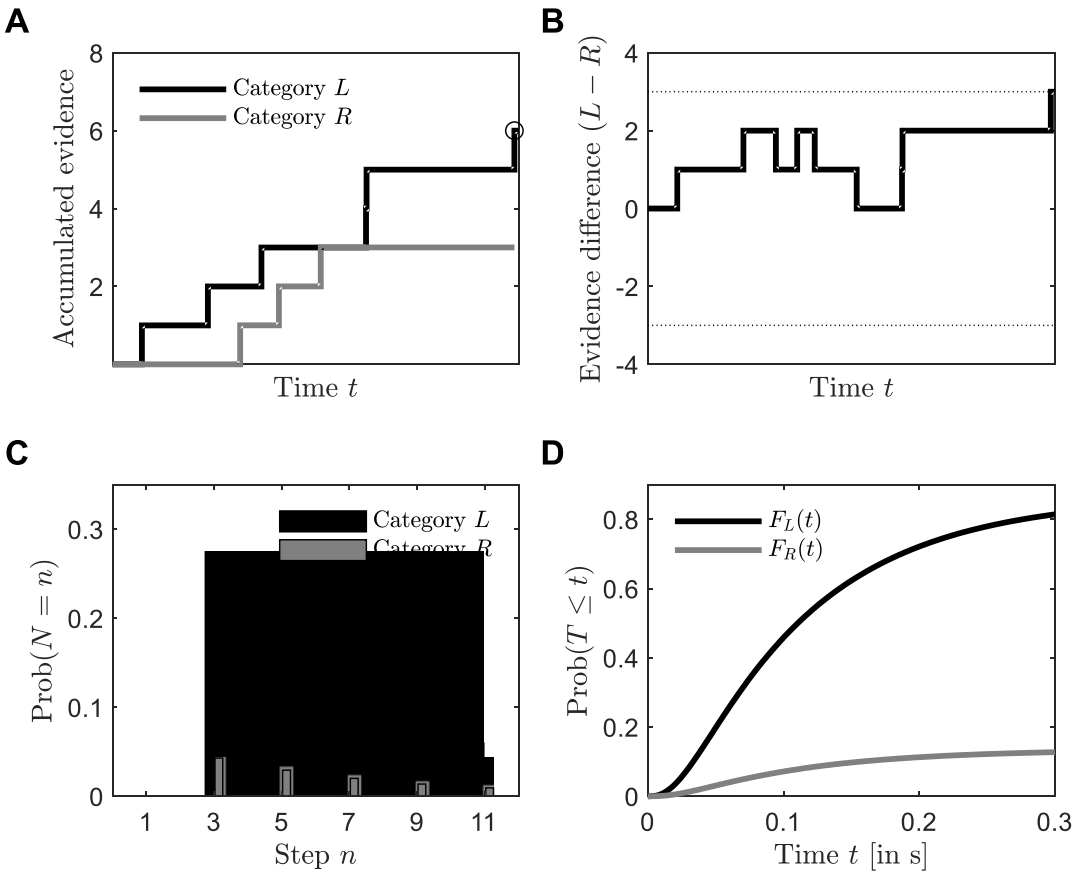

Figure 2. Overview of the Poisson random walk model for two perceptual categories *L* and *R*. (**A, B**) Random realization of a single trial with a leftward moving grating. Evidence for either category is accumulated over time. A category is conclusively selected and reported as soon as it has accumulated *k* more categorizations than the other (here: $k = 3$). The categorization that leads to reaching the criterion *L* is marked with a circle in (**A**). (**C**) Theoretical probability distribution for reaching either criterion after exactly *n* steps in the random walk. Reaching the criterion for *L* is more likely overall, however, there is a probability that the criterion for *R* is reached first, leading to an incorrect response. (**D**) Theoretical cumulative distribution functions for reaching criterion *L* or *R* at or before time *t*. This is the model prediction for RT, it is based on the probability distribution shown in (**C**) and the distribution of waiting time until *n* categorizations have been made (not shown).

In an extension to the model outlined so far, there are three additional sources of RT variability. The first source of additional variability comes from the assumption of trial-by-trial

variation in the processing rates. Instead of assuming the same processing rates throughout the experiment, the processing rates are assumed to vary across trials. The amount of variation is denoted by parameter ξ. The trial-by-trial parameter ξ is best interpreted as a scaling factor $1/\xi$, indicating how large the processing rate variance is compared to the mean processing rate (small values represent large variance compared to the mean). The second source of additional variability is a variable starting point Z_0. With a variable starting point, the decision process does not start at the neutral position at stimulus onset but may be biased towards either criterion across trials. The amount of variation is denoted by the variance of Z_0, that is, $\sigma_z^2 = Var(Z_0)$. As in the DDM, this variability reflects effects of premature sampling (Ratcliff & Rouder, 1996), that is, sampling (noise) before stimulus onset (Laming, 1968). With this assumption it is possible to explain fast errors, that is, errors that are faster than predicted by the base model described above. Conversely, the assumption of trial-by-trial variation in processing rates allows the model to predict slow errors, that is, errors that are slower than predicted by the base model.

The third source of additional variation in the model is the assumption of temporally inhomogeneous processing. Under this assumption, processing rates v_i are no longer constant over time (Figure 3A). Instead, the processing rates $v_i(t)$ vary over time; they are zero at stimulus onset and upon stimulus presentation, gradually increase up to a maximum. This response function is modeled as a scaled gamma distribution function with variable rate μ and fixed shape n. The effect of μ is that high values lead to a faster increase to the asymptotic level, whereas lower values lead to a slower increase. Apart from being physiologically more plausible than a constant processing rate (Christensen, Markussen, Bundesen, & Kyllingsbæk, 2018), this functional form has been suggested by Smith and Van Zandt (2000) to reflect the time course of early visual processing. Early visual processing can be conceived as filters, some of which exhibit a sustained response to a stimulus; their output shall be approximated by this response function (Smith & Van Zandt, 2000). This functional form provides additional flexibility in the leading edge of the predicted RT distribution but has otherwise negligible effects. This is because the speed of perceptual decisions depends on the integrated processing rates rather than the processing rates $v_i(t)$. Since the time-varying rate will asymptote to the constant rate v_i sooner or later, only the fastest responses will be affected by the assumption of temporal inhomogeneity (Figure 3B).

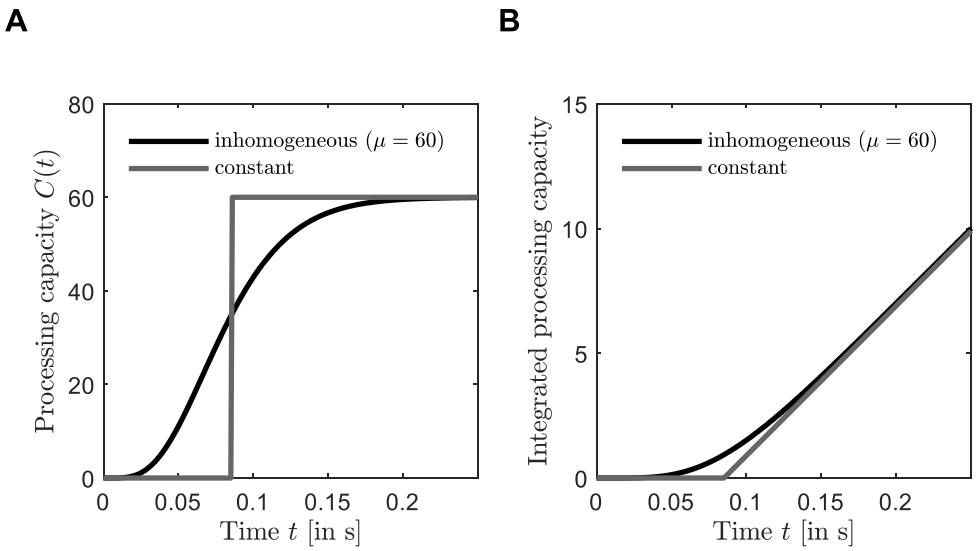

Figure 3. Inhomogeneous processing in the Poisson random walk. (**A)** Time dependent processing rates, represented by overall processing capacity C. In the temporally homogeneous case, tentative categorizations are made with constant speed (here: $C = 60$ Hz). In the inhomogeneous case, rates increase continuously until reaching an asymptote (here: $C(\infty) = 60$ Hz). (**B)** Integrated processing capacity of the two cases shown in (A). In a Poisson process, the number of expected events depends on the integrated event rate, rather than the event rate itself. This assumption on inhomogeneous processing affects the leading edge of predicted RT distributions only, as its effect becomes negligible with increasing time.

In the DDM, variability in the leading edge of RT distributions is achieved by introducing trial-by-trial variation in non-decision time (Ratcliff & Tuerlinckx, 2002) while leaving the drift rates constant within a trial. A notable exception is the time-dependent diffusion model proposed by Smith and Ratcliff (2009, see also Smith, Ratcliff, & Wolfgang, 2004). This extended version of the original DDM is relevant in this context, as it has been suggested as a RT model for spatial cueing effects. The extended DDM captures those cueing effects with different time courses of drift rates. Decisions based on stimuli presented at an attended location are based

on a drift rate that more quickly reaches its asymptote than those based on stimuli presented at an unattended location (Smith & Ratcliff, 2009). To model the data in this study, we largely followed Smith and Ratcliff's (2009) work in specifying the model and selecting model parameters for testing. Nonetheless, there are differences in our approach, as the extended diffusion model is a multi-stage processing model that includes encoding in visual short-term memory as an explicit processing stage. However, the Poisson random walk model enabled us to take up the most relevant aspects of the diffusion model using a computationally simpler model. We exploited this advantage to perform nested model tests of model fits based on individual datasets instead of group averaged data. Also, the Poisson random walk model includes visual processing capacity C, a parameter that has no direct counterpart in the DDM but has been central in modelling visual attention effects within the TVA (Bundesen, 1990; Bundesen, Vangkilde, & Petersen, 2014; Shibuya & Bundesen, 1988).

Altogether, the null model, that is, the model without parameters for cueing effects had eleven free parameters ($p_{.15}$, $p_{.25}$, $p_{.35}$, $p_{.45}$, $p_{.55}$, C, k, T_0, ξ, σ_z^2, μ) that were estimated from data. Five of these are normalized processing rates p_x for the five different spatial frequencies, which, together with processing capacity C, yield the (mean) processing rates for correct and incorrect categorizations. The remaining parameters were the response criterion k, the latency of all non-decision processes T_0, and three parameters (ξ, σ_z^2, and μ) for the three additional sources of RT variation. For more details on the model and model specification, we refer the interested reader to the Appendix.

Regarding cueing effects on the different model parameters, we were particularly interested in two different models. The first model under consideration is similar to the time-dependent diffusion model (Smith & Ratcliff, 2009): we allowed for different time courses of processing (μ_{valid}, $\mu_{neutral}$) expecting that valid cues will lead to a faster response in early visual processing. Additionally, visual processing capacity is different between cueing conditions (C_{valid}, $C_{neutral}$) in this model. The latter prediction can be derived from the TVA by noting that spatial cues will alter attention weights assigned to the different locations. The attention weights determine processing rates and thus processing capacity in a multiplicative fashion (Bundesen, 1990), so that higher weights would lead to higher processing speed. On the other hand, if transient attention increases the temporal integration time window (Hein et al., 2006; Yeshurun, 2004; Yeshurun & Hein, 2011; Yeshurun & Marom, 2008; Yeshurun & Sabo, 2012), this should

degrade the information that is sampled from the moving stimuli. In this case, we would expect a lower processing capacity if attention was directed at the location of the motion stimulus.

The second model that we fitted to the data was a model with different response criteria (k_{valid}, $k_{neutral}$) to test if cueing effects could be explained by a mere speed-accuracy trade-off rather than attentional effects due to the cue (e.g., Hein et al., 2006). In addition, we tested whether the starting point distributions ($\sigma^2_{z,valid}$, $\sigma^2_{z,neutral}$) differed between the cueing conditions. As stated above, this parameter reflects premature sampling (Laming, 1968; Rouder, 1996), that is, sampling noise before the onset of the imperative stimulus. Judging from the form of the estimated conditional accuracy functions (Figure 5B), we suspected starting point variation to be different for valid and neutral cueing conditions.

The remaining parameters were held fixed across cueing conditions in all models we investigated. The decision to fix non-decision time (T_0) and trial-by-trial variation of processing rates (ξ) was based on the results reported by Smith, Ratcliff, and Wolfgang (2004). Similarly, we restricted the normalized processing rates p_x to be the same across cueing conditions. This assumption is also based on the TVA—while attention weights alter the overall processing capacity, they should not differentially affect the relative strength of processing rates. The relative strengths of processing rates are rather determined by categorical bias (in addition to stimulus confusability). According to the TVA, categorical bias works independent from attentional weights and thus should not be affected by spatial cueing. After identifying the best fitting model, we tested post hoc the veracity of our assumptions.

Estimation of model parameters was performed by maximizing the likelihood function of the model. For the maximum likelihood estimation of model parameters, we removed all responses faster than 200 ms to avoid distortion of model parameters. This cut off criterion was applied to all datasets and led to the removal of four responses in one participant. To determine which model is most suitable given the data, we followed a two-fold approach. First, we fitted a model that included all parameters of interest to individual datasets and used *t*-tests to determine which parameters (if any) were significantly different between cueing conditions. In a second approach, we further tested the parameter of interest in a model comparison approach. Starting with the most restricted model with all parameters being equal between cueing conditions, we added parameters to the model and assessed the increased goodness-of-fit by means of the Bayes Information Criterion (BIC). The BIC introduces a sample-size dependent penalty for additional

parameters, so that models with lower BIC are preferred. Rather than performing a forward procedure, we repeatedly checked whether parameters that were excluded in previous steps would lead to a lower BIC in later steps. This complementary model evaluation can provide further insight because averaging parameter estimates across participants can mask effects that are indeed present (Donkin, Brown, & Heathcote, 2011).

Results

Analysis of Accuracy Scores

Figure 4A shows the median and mean proportion of correct responses as a function of the target's spatial frequency. An aligned rank transform ANOVA including Cue type (neutral vs. valid) and Spatial frequency (0.15, 0.25, 0.35, 0.45, and 0.55 c/deg) as within-subjects factors revealed a significant effect of the Cue type ($F_{1, 171} = 30.72$, $p < .001$, $\eta^2_p = .15$), a significant effect of the Spatial frequency ($F_{4, 171} = 4.91$, $p < .001$, $\eta^2_p = .10$), and a significant interaction between Cue type and Spatial frequency ($F_{4, 171} = 2.73$, $p = .031$, $\eta^2_p = .06$).

For the main effect of spatial frequency, post-hoc comparisons corrected with Holm's method ($\alpha = 0.05$) (Holm, 1979) yielded a significant difference between the spatial frequency at 0.15 c/deg and 0.55 c/deg ($p_{adj} = .04$), but not between 0.15 c/deg and 0.25 c/deg ($p_{adj} = .66$), 0.35 c/deg ($p_{adj} = 1.0$), and 0.45 c/deg ($p_{adj} = .95$). For the spatial frequency at 0.25 c/deg, Holm corrected post-hoc comparisons revealed a significant difference with respect to 0.55 c/deg ($p_{adj} < .001$), but not with respect to 0.35 c/deg ($p_{adj} = .36$) and 0.45 c/deg ($p_{adj} = .13$). For the spatial frequency at 0.35 c/deg, Holm corrected post-hoc comparisons did not reveal a significant difference with respect to 0.45 c/deg ($p_{adj} = 1.0$) and 0.55 c/deg ($p_{adj} = .13$). Additionally, 0.45 c/deg did not significantly differ from 0.55 c/deg ($p_{adj} = .36$).

For the interaction between Cue type and Spatial frequency, Holm corrected post-hoc comparisons revealed a significant difference between valid and neutral cue conditions only for the spatial frequency of 0.35 c/deg, in which the valid cue reduced the accuracy in motion direction discrimination ($p_{adj} < .001$). The difference between valid and neutral cues was not significant for the spatial frequencies of either 0.15 c/deg ($p_{adj} = .99$), 0.25 c/deg ($p_{adj} = .99$), 0.45 c/deg ($p_{adj} = .99$) and 0.55 c/deg ($p_{adj} = .25$).

Analysis of Reaction Times

Figure 4B shows the median and mean reaction times (RTs, in seconds) for correct responses only, calculated for each cue condition as a function of the spatial frequency. An aligned rank transform ANOVA was performed on the mean RTs including as within-subjects factors the Cue type and Spatial frequency revealed a significant effect of the Cue type ($F_{1, 171}$ = 148.07, $p < .0001$, $\eta^2_p = .46$), but no significant effect of the Spatial frequency ($F_{4, 171} = 2.39$, $p = .053$, $\eta^2_p = .05$) and interaction between Cue type and Spatial frequency ($F_{4, 171} = 1.45$, $p = .22$, $\eta^2_p = .03$). In general, median and mean RTs in the valid cue conditions were significantly lower than in the neutral cue condition. For spatial frequency, Holm corrected pairwise comparisons did not reveal any significant difference between spatial frequencies (all $p_{adj} > .05$).

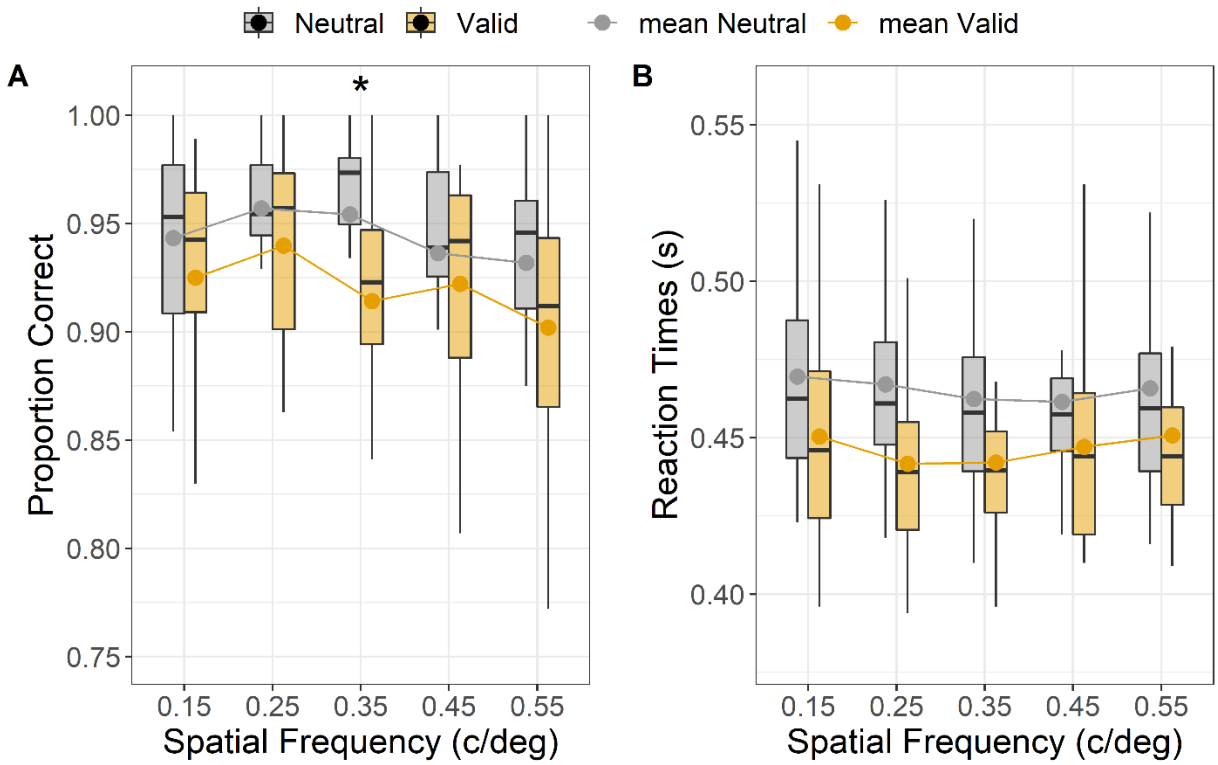

Figure 4. Boxplots of the proportion of correct responses (**A**), and reaction times (in seconds) based on correct trials (**B**) for the valid and neutral cue conditions as a function of the target's spatial frequency (c/deg). For each boxplot, the horizontal black line indicates the median, the boxes in the boxplots extend from the lower quartile to the upper quartile (i.e., the 25th and 75th percentiles). The difference between the upper quartile and the lower quartile is called interquartile range. The vertical lines extend to one and a half times the interquartile range, and they are limited to reaching actual data points. The dot within each boxplot represents the mean

accuracy (**A**) and mean reaction time (**B**) values. The asterisk in panel (**A**) indicates the significant difference ($p < .05$) between neutral and valid cue conditions for the spatial frequencies at 0.35 c/deg.

Descriptive analysis of RT distribution and conditional accuracy

In line with the results in the previous section, we found faster response times to validly cued targets across a wide range of the entire distribution (Figure 5). This difference was not constant, though, it was larger for the lowest quantiles and decreased with increasing the RT. That is, responses in the valid cue condition were initially faster than those in the neutral cue condition, but response speed to targets in the neutral cue condition caught up with increasing time so that late responses were about equally fast.

The analysis of response accuracy across the RT distribution showed that these fast responses in the valid cue condition were less accurate than the fastest responses to neutrally cued targets. With increasing response times, the observed accuracy was approximately the same between the two cueing conditions. That is, while the valid cues lead to faster responses, these fast responses were less accurate than those in the neutral cue condition. This analysis supplements the analysis of mean RT and accuracy, in that it describes more precisely, where the observed cueing effects are located. The response time model needs to account for these effects; and in testing for cueing effects its various parameters, we can describe those in terms of cognitive parameters.

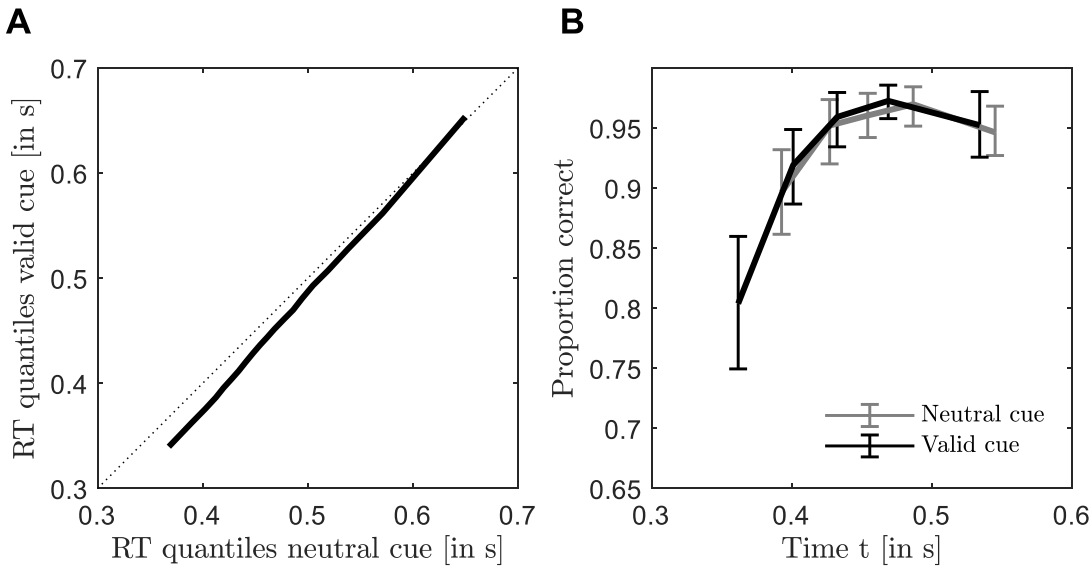

Figure 5. Analysis of RT distributions. (**A)** Quantile-Quantile (Q-Q) plot of empirical RT distributions in the two cueing conditions. Early responses are faster in the valid cue condition, but this difference diminishes as RTs increase, so slow responses are comparable in valid and neutral cue trials. (**B**) Conditional accuracy functions (CAFs) for valid and neutral cue conditions. Error bars denote 95% confidence intervals.

Fit of the Poisson random walk model

As described above, we fitted the Poisson random walk model to the observed RT distributions. The goal of this analysis was twofold: (i) to investigate the possibility that a speed-accuracy trade-off caused the patterns of RT and accuracy (Figure 4), (ii) to analyze the full RT distributions rather than mean RT only, as the full distribution can contain relevant information that is not captured by mean RT (cf. Figure 5).

We first fitted the Poisson random walk model to the individual data obtained in the valid cue condition and in the neutral cue condition. This model was specified so that separate

parameters were included for all parameters that we expected might contain cueing effects (i.e., μ, C, k, σ_z^2). The choice of which parameters to test was based on a priori considerations. In particular, we wanted to include all relevant parameters while excluding the least interesting to reduce the computational burden in the second analysis. The estimated parameters of this model are summarized in Table 1. The processing rates (p_x) for the correct categorization were similar across the spatial frequencies used in the experiment. The response criteria (k) were also rather similar for the two cueing conditions, but the criterion for validly cued targets was significantly higher than that for neutrally cued targets ($t = -2.343$, $df = 19$, $p = .030$). That is, in both cueing conditions our participants performed on average about three categorizations more for the correct than the incorrect perceptual category. Processing capacity (C) was significantly higher in neutrally cued targets than validly cued targets ($t = 2.328$, $df = 19$, $p = .031$), whereas the speed of early perceptual processing (μ) was significantly higher in validly cued targets than neutrally cued targets ($t = -4.526$, $df = 19$, $p < .001$). Valid cues lead to significantly higher starting point variation than neutral cues ($t = -4.491$, $df = 19$, $p < .001$). Trial-by-trial variation (ξ) in processing rates was low in both cueing conditions. This is also reflected in the data, in which we observed only few slow errors (error RT quantiles greater than correct RT quantiles). This is evident in Figure 7, where a wide range of quantiles of incorrect RT distributions are lower than corresponding quantiles of correct RT distributions, especially in validly cued trials. Individual parameter estimates of the model are available as supplemental material (Table S.1).

Table 1. Group average (± SD) parameter estimates of the Poisson random walk model.

	Cue type	
Parameter	Valid cues	Neutral cues
$p_{0.15}$	0.792 (± 0.059)	
$p_{0.25}$	0.816 (± 0.061)	
$p_{0.35}$	0.816 (± 0.069)	
$p_{0.45}$	0.806 (± 0.062)	

$p_{0.55}$	0.784 (± 0.056)	
ξ [$E(v_x)/Var(v_x)$]	20.6 (± 40.7)	
k [categorizations]	3.16 (± 0.64)	2.98 (± 0.61)
C [in Hz]	85.9 (± 22.6)	101.3 (± 30.0)
μ	37.6 (± 15.2)	28.6 (± 11.2)
σ_z^2	2.20 (± 1.27)	0.83 (± 1.16)
T_0 [in s]	0.251 (± 0.032)	

Note—Parameters: p_x denote (mean) normalized processing rates for the correct categorizations across the different spatial frequencies (x = 0.15, 0.25, 0.35, 0.45, and 0.55 c/deg), ξ represents trial-by-trial variation in processing rates, k is criterion separation, C is the visual processing capacity, μ determines the time course of processing after stimulus presentation (time inhomogeneous processing, higher means faster onset), σ_z^2 is the variance of the starting point, and T_0 is the latency of all non-decision processes. Standard deviation (SD) is calculated across participants.

As suggested by the conditional accuracy function (Figure 5B), starting point variation was higher after valid cues than neutral cues. Allowing for this difference, however, was not sufficient for the model to predict the difference in fast responses between valid and neutral cue conditions (Figure 5A). In addition to starting point variation, early visual processing was also faster in the valid cue condition. Although there was a significant criterion shift, the shift was in a direction opposite to that suggested by the analysis of mean RT and accuracy (Figure 4). Thus, the result speaks against a speed-accuracy trade-off as a simple explanation for the observed patterns; the parameter analysis based on model comparisons (see below) suggested that this difference is likely a statistical artefact caused by including a cueing effect in a parameter that is not affected by cueing validity.

Taken together, the results of the first model-based analysis suggest that the presentation of a valid cue led to significantly faster increase in processing rates but to a significantly lower efficiency in overall processing, that is, more efficient processing of the moving Gabor stimulus after neutral cues. Thus, processing after stimulus onset was initially faster, however, over time, neutrally cued targets were processed more efficiently, leading to late responses being about equally fast in neutrally cued targets and validly cued targets (Figures 5A and 6).

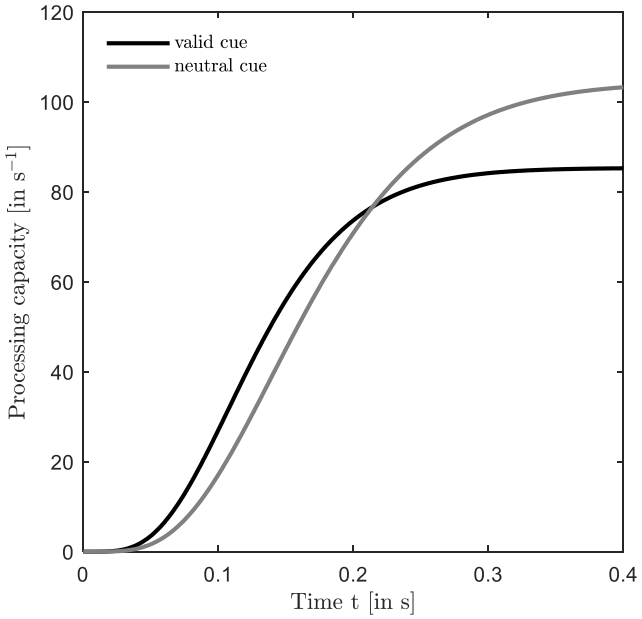

Figure 6. Estimated time-course of visual processing capacity (C) in the Poisson random walk model. Valid cues led to faster processing initially, which provides an early advantage. Later, neutral cues take over as they are processed more efficiently. The functional form is given by Eq. A3 in the Appendix, the relevant parameter estimates are taken from Table 1.

In the second approach, we fitted different models to the individual datasets and compared their goodness-of-fit by means of their Bayes information criterion (BIC), summed across all participants. Unlike the first analysis, we directly tested cueing effects on single parameters with this stepwise approach. Starting with the most restricted model in which all parameters were equal across cueing conditions, we added parameters to the model and compared their BIC to test if the added parameters substantially improve the model predictions

and should be kept in the model. Rather than a simple forward selection, we checked whether previously rejected parameters improved the fit in later iterations.

Both the model with different processing capacity (ΔBIC = −148.9) and the model with different response criteria (ΔBIC = −592.7) produced better fits than the null model. Allowing for a different time course of processing validly and neutrally cued targets in addition to processing capacity further improved the model fit of the processing capacity model (ΔBIC = −465.1). Similarly, different starting point variation improved both the criterion shift model (ΔBIC = −70.6) as well as the processing capacity and time course model (ΔBIC = −157.2). Eventually, the model with cueing effects in processing capacity (C), time course of processing (μ), and starting point variation (σ_z^2) was the best fitting model, despite having one parameter more than the criterion shift model with different starting point variation (ΔBIC = 108.0). The BIC values of all these model comparisons are summarized in Table 2. The fit of the winning model, averaged across all participants, is shown in Figure 7.

Table 2. Results of model comparisons.

Restricted parameters	Cueing parameters	Number of parameters	Σ BIC
C, k, T_0, μ, ξ, σ_z^2, $p_{0.15}$, …, $p_{0.55}$	None	11	–34.832
C, T_0, μ, ξ, σ_z^2, $p_{0.15}$, …, $p_{0.55}$	k	12	–35.425
k, T_0, μ, ξ, σ_z^2, $p_{0.15}$, …, $p_{0.55}$	C	12	–34.980
C, k, T_0, ξ, σ_z^2, $p_{0.15}$, …, $p_{0.55}$	μ	12	–35.350
k, T_0, ξ, σ_z^2, $p_{0.15}$, …, $p_{0.55}$	C, μ	13	–35.446
C, T_0, μ, ξ, $p_{0.15}$, …, $p_{0.55}$	k, σ_z^2	13	–35.495
k, T_0, ξ, $p_{0.15}$, …, $p_{0.55}$	C, μ, σ_z^2	14	–35.603

Note—BIC: Bayes Information Criterion. Lower values indicate better models.

Post-hoc we tested the correctness of our a-priori assumptions. In comparison with the winning model described above, the addition of a cueing effect on non-decision time (T_0) did not improve the fit (BIC = −35,571; ΔBIC = 32.7), nor did the addition of a cueing effect on the normalized processing rates (BIC = −35,534; ΔBIC = 68.6) or on the trial-by-trial variation of processing rates (ξ; BIC = −35,507; ΔBIC = 95.9).

Both the fits to individual data and the model comparison approach support the interpretation that cueing effects were not due to criterion shifts but rather reflect changes in the starting point, the time course of processing and the efficiency associated with processing the motion stimuli. This is also in line with earlier results that were not based on RT models but on median split of RT data (see Hein et al., 2006). The best fitting model had a total of 14 free parameters to explain 20 RT distributions (ten correct and ten incorrect RT distributions). Hence, the model is reasonably restricted, although the number of free parameters certainly contributed to the high consistency between the predictions of the model and the observed data (Figure 7).

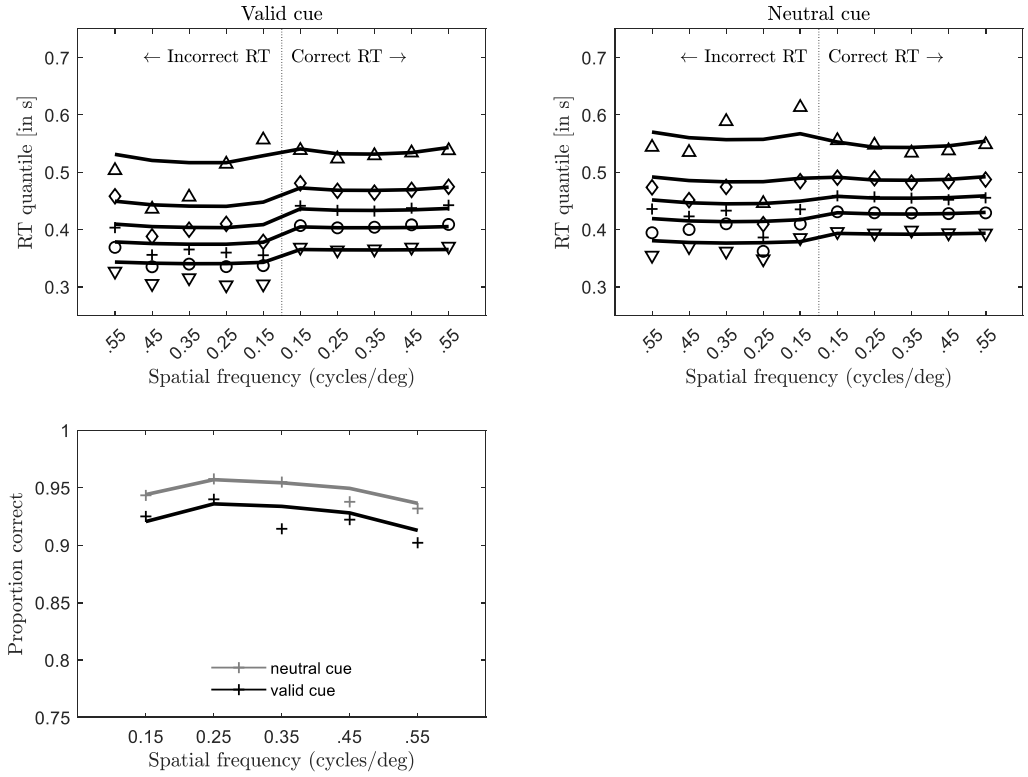

Figure 7. Empirical RT distributions and accuracy compared with the fit of the Poisson random walk model. Upper panels: Quantile plots of RT data (10%: downward pointing triangle, 30%: circles, 50% (median): crosses, 70%: diamonds, 90%: upward pointing triangles) and model predictions (unmarked points connected with lines). In each panel, the five leftmost distributions are incorrect RT distributions, the five rightmost distributions are correct RT distributions. Lower panel: observed frequency and predicted probability of correct

responses. Model predictions are obtained from the winning model of the model comparison. Both data and model predictions are averaged across participants.

Discussion

In the current study, we investigated the effects of transient spatial attention on motion processing. A target Gabor patch could appear in one of eight possible locations around the fixation point. A cue was presented shortly before the target Gabor patch and could be informative about the target's location (i.e., valid cue) or uninformative (i.e., neutral cue). Participants were asked to perform a motion direction discrimination task on the Gabor patch. Although the accuracy scores were close to ceiling across the spatial frequencies used, the results showed a robust main effect of the cue type, with neutral cues producing higher accuracy than the valid cues. The cue type and spatial frequency interaction was also significant, and we found a small (4.28%) but significant decrease in accuracy when the target Gabor patch had a spatial frequency of 0.35 c/deg only in the valid cue condition. This accuracy drop is in line with the previous findings indicating that exogenous cueing impairs temporal segregation processing and supports Yeshurun's (2004) hypothesis that transient attention facilitates parvocellular but inhibits the magnocellular activity. In a series of experiments, Yeshurun and Levy (2003) investigated the effects of transient attention on temporal and spatial resolution, and the underlying mechanisms. Their results showed that exogenous spatial cues impair temporal resolution while enhancing spatial resolution. This contrasting influence of transient attention on temporal and spatial processing was not dependent on the decrease in spatial summation, which is a consequence of the reduction of the receptive field size at the attended location. Instead, an additional outcome of the prioritized parvocellular processing over magnocellular processing, namely the longer response durations of spatiotemporal units, might be responsible for the decrease in temporal resolution by boosting temporal integration and hindering temporal segregation.

The neural correlates of spatial-attention-related enhancement in temporal integration were investigated in an ERP study by Akyurek and van Asselt (2015). The authors found that exogenous cues enhance color fusion (i.e., perception of two colors as a single color when they were displayed briefly and rapidly), and attentional modulation of temporal processing occurs at early stages of visual processing. Along with the absence of cueing effects in later stages, these

results suggest that the increased temporal integration may have an adaptive role in decreasing competition between successive stimuli. However, this would impair motion perception which depends on temporal segregation. Baruch and Yeshurun (2014) used the attentional attraction field model, which shifts receptive fields of neurons to the attended location in a hierarchical manner. To assess the temporal modulations induced by spatial attention, they simulated the neuronal activation in response to two events separated by a brief interval. When attention was directed to the location of these events, the likelihood for the events to merge increased, which in turn decreased the ability to segregate the events in the temporal domain. The present results are compatible with this account; transient cues may have increased the likelihood that luminance changes in successive frames are integrated in a single event, thus decreasing the accuracy on the motion direction discrimination task.

However, of these differences, only that in accuracy between valid and neutral cues in the 0.35 c/deg condition reached statistical significance. One could argue that the low accuracy in the valid cue condition at 0.35 c/deg depends on the transient cue-induced speed-accuracy trade-off. This hypothesis seems unlikely as for reaction times the interaction between spatial frequency and cue type did not reach significance and the results of the RT model also provide converging evidence against this possibility. It is conceivable that valid cues reduced response accuracy across a wider range of low spatial frequencies, but these failed to reach statistical significance after correction for multiple comparisons. This potential issue of statistical power is avoided in the modelling approach, in which the parameters relevant to our investigation are the same across all stimulus conditions and only vary between cueing conditions, thereby avoiding the problem of multiple comparisons.

To analyze response latency and accuracy simultaneously and check for a possible speed-accuracy trade-off, we fitted the Poisson random walk model (Blurton et al., 2020) which provides predictions of the RT distribution of both correct and incorrect responses. All our analyses support the conclusion that cueing affects processing rates and their time course, rather than a change in criterion due to cueing. In addition to these quantitative analyses, it is also worth noting that the temporal profile of stimulation (i.e., brief cue duration and cue-target interval) restrict the potential contribution of criterion shifts to the observed effects. It was extremely difficult to discern valid cues from neutral cues and, even if the cue type became apparent from trial to trial, this was way after the target had been presented. This does not

preclude implicit forms of learning, but a genuine speed-accuracy trade-off in which participants set the response criteria to trade speed for accuracy is unlikely to explain our results.

Rather, an intricate relationship between an initial speed-up due to valid cues combined with less efficient processing is a better description of the observed data. These two effects reflect different aspects of processing: the first has been reported previously in an experiment with cueing and limited exposure time (Smith & Ratcliff, 2009). In that study, the diffusion decision model (DDM) was extended to a time-inhomogeneous version in which drift rates increased gradually after stimulus onset. In applying this model, Smith and Ratcliff (2009) showed that drift rates increased more quickly after validly cued targets compared to those presented after invalid cues, like we found in the present study. The cueing effect on processing capacity, i.e., more efficient processing of neutrally cued targets compared to validly cued targets, has not been previously reported since this parameter is not included in the original DDM. A reduced processing capacity might reflect an attenuation in temporal segregation of a stimulus (e.g., Hein et al., 2006), possibly caused by inhibition of the magnocellular pathway (Yeshurun, 2004). If directing spatial attention indeed has a detrimental effect on perceiving temporally extended information, we would expect similar effects as reducing stimulus contrast in a static visual stimulus. Stimulus contrast, on the other hand, has been shown to have a clear and systematic relationship with processing speed in perceptual decisions with above threshold stimuli (Christensen et al., 2018).

Most importantly, the processing rates did not show a cueing effect, suggesting that the quality of information sampled from the stimuli was similar between the valid cue condition and the neutral cue condition. In other words, while the processing capacity was higher for neutral cues than valid cues, the probability to make a correct or incorrect categorization (i.e., the signal-to-noise ratio) was independent of the cue type. This strong assumption was derived from the Theory of Visual Attention (TVA; Bundesen, 1990). In TVA, processing rates are determined by perceptual evidence and two distinct attention mechanisms: filtering and pigeonholing. In short, filtering describes an attentional weight that is given to an object, whereas pigeonholing refers to a categorical bias mechanism. The Poisson random walk model is an extension of TVA, applicable to response times in perceptual decisions. Being founded in TVA, the RT model retains the general TVA assumptions, including filtering and pigeonholing. Put simply, filtering refers to processing capacity, whereas pigeonholing refers to the normalized processing rates in

the RT model. In line with the model specification, we assumed no perceptual bias and describe cueing effects as a change in attention weights. Contrary to TVA, however, is that although the validly cued targets received higher attention weights than the neutrally cued ones, processing speed was higher for the latter. As described above, we interpret this finding as a result of increased temporal integration due to transient attention, leading to a poorer stimulus representation from which the observers could not sample information as efficiently as from the stimulus representations of neutrally cued targets.

Alternatively, it is possible that inhibition during the processing of a validly cued target decreased the processing speed after the initial advantage. We did not include inhibition in the RT model since the effect explained by inhibition would mimic the results from the model with different processing capacities, making it impossible to distinguish between these two possibilities. However, this certainly does not rule out the possibility that inhibition occurred, so this point remains open for further investigation. Concerning the initial facilitation/late inhibition hypothesis, a recent study of Zhang, Shelchkova, Ezzo, and Poletti (2021) investigating the time course of attentional effects in the fovea showed that exogenous attention was associated with an early enhancement in orientation discrimination. When target stimuli were preceded by a valid transient spatial cue within a short interval (i.e., 100 ms), sensitivity in orientation discrimination improved compared to neutral and invalid conditions. With longer intervals (i.e., hundreds of milliseconds) sensitivity was greater for the invalid compared to valid cues, which resembles an inhibition of return effect. A closer inspection of the results revealed that while sensitivity only slightly and non-significantly dropped for valid cues at longer intervals when compared to shorter intervals, there was a significant increase in sensitivity for neutral and invalid cues. These results indicated that fine detail is resolved better at the cued locations with short delays and acuity increases at the uncued locations with longer delays. The findings of Zhang et al. (2021) suggest that exogenous valid cues led to initial facilitation, which was subsequently replaced by facilitation for invalid and neutral cues in a spatial task (i.e., orientation discrimination). However, the time course of attentional modulation in terms of facilitation and possible inhibition in a spatiotemporal task with stimuli presented in the near periphery remains an open question and more evidence is necessary.

In the present study we extended the results of Yeshurun and Hein (2011) showing that an exogenous cue drawing attention towards a specific location diminishes the ability to

discriminate the target's direction of motion, notably when motion processing is likely to be mediated by the magnocellular system (Yeshurun & Levi, 2003; Hein et al., 2006; Nicol et al., 2009; Yeshurun & Hein, 2011). These results emerge presumably as a consequence of attention-induced facilitation of the parvocellular processing, leading to a relatively larger temporal integration window, which in turn impairs magnocellular-dominated motion processing on briefly presented targets. This explanation is consistent with the retino-cortical dynamics model of visual processing (Ogmen, 1993; Ogmen, Breitmeyer, & Melvin, 2003), postulating a mutual inhibition at the post-retinal level between neural populations that receive input from the sustained and transient channels, which are primarily driven by parvocellular and magnocellular cells, respectively. The current study provides the most compelling evidence that an exogenous cue reduces the accuracy in a behavioral task when temporal segregation is necessary.

Acknowledgments

We would like to thank Hakki Yilmaz for technical assistance in data collection. SPB was supported by The Independent Research Fund Denmark (9037-00169B).

Conflict of interest statement

On behalf of all authors, the corresponding author states that there is no conflict of interest.

Appendix

Details on the RT model

In the Poisson random walk model for two perceptual categories L (apparent leftward motion) and R (apparent rightward motion), let $v_L = v(x, L)$ and $v_R = v(x, R)$ be the processing rates for the perceptual categorizations "stimulus x belongs to category L" and "stimulus x belongs to category R", respectively. Evidence for each category is stored in a Poisson counter with their respective rate. Further, evidence in favor of category L is taken as evidence against R, and vice versa, leading to a random walk. That is, each time a categorization is made for L, the counter for that category is increased by one and the counter for R is decreased by one, and vice versa. A perceptual category i is conclusively selected as soon as it has accumulated k_i more categorizations than the other. Using the superposition principle of the Poisson processes (Townsend & Ashby, 1983), $p_L = v_L/(v_L + v_R)$ is the conditional probability to make a categorization "stimulus x belongs to category L" and $p_R = 1 - p_L = v_R/(v_L + v_R)$ is the conditional

probability to make a categorization "stimulus x belongs to category R"—both given that there is a categorization at some time t. The increments in the random walk process are independent and distributed as a (modified) binomial distribution: the increment is +1 with probability p_L and –1 with probability p_R. This definition corresponds to a random walk in which the upper criterion is associated with Category L and the lower with Category R.

Evidence is accumulated until one category, say Category L, has k_L more tentative categorizations than Category R. The total number of categorizations made until this happens is a random variable. Its distribution is obtained from solving the first passage problem for a simple random walk between two absorbing barriers (e.g., Feller, 1968, p. 354). This yields two distributions, $P_L(n \mid p_L, k_L, k_R)$ and $P_R(n \mid p_R, k_L, k_R) = P_L(n \mid 1–p_L, k_R, k_L)$ for the two possible outcomes. To obtain a prediction for the time needed to obtain n tentative categories, we use the Markov property of the Poisson process and obtain the density $f(t \mid n, C)$ of time T given n categorizations from the Erlang distribution (e.g., Forbes, Evans, Hastings, & Peacock, 2011, Ch. 15) with rate $C = v_R + v_R$, that is, the sum of all processing rates. In this context, C is the visual processing capacity, indicating the speed or effectivity with which a perceptual decision is made. Since the total number of categorizations is a latent variable, the model prediction for the probability density of RT is obtained by marginalizing:

$$f_L(t \mid p_L, k_L, k_R, C) = \sum_{n=k_L}^{\infty} P_L(n \mid p_L, k_L, k_R)\, f(t \mid n, C) \qquad \text{Eq. A1a}$$

$$f_R(t \mid p_L, k_L, k_R, C) = \sum_{n=k_R}^{\infty} P_R(n \mid p_L, k_R, k_L)\, f(t \mid n, C) \qquad \text{Eq. A1b}$$

For the cumulative distribution functions as displayed in Figure 2, one needs to substitute the density $f(t \mid n, C)$ of the Erlang distribution for the distribution function $F(t \mid n, C)$.

In addition to the variability in the simple random walk and Erlang distribution described above, the full model includes several sources of additional variability. The first is trial-by-trial variation in processing rates: The processing rates V_i are assumed to be gamma distributed across trials with mean v_i and variance v_i/ξ. Since a random variable defined as the sum of gamma distributed random variables is also gamma distributed, the visual processing capacity is gamma distributed with mean $C = v_A + v_B$ and variance $(v_A + v_B)/\xi$. In other words, all trial-by-trial variability is described by a single parameter (ξ), which is most easily interpreted as a scaling factor $1/\xi$, describing how large the variance of processing rates is in relation to their mean.

The second source of additional variability is a variable starting point Z_0. A suitable distribution for Z_0 is the beta-binomial distribution with parameters n, z_A, and z_B. This distribution is suitable because the increments in the full model with trial-by-trial variation in processing rates also follow this distribution. The beta-binomial distribution is a generalization of a binomial distribution with beta-distributed probability π. The parameter n is fixed ($n = k_A + k_B - 2$), whereas z_A, and z_B are free parameters. In the symmetric model for unbiased decisions, $z_A = z_B = z$, reducing the number of free parameters by one. As stated in the main text, we fitted the moments of the starting distribution instead of its parameter z for easier interpretability. Under the symmetry assumption, the expected value $E(Z_0)$ of the starting point does not depend on z, so we specified the entire distribution by its variance $\sigma_z^2 = Var(Z_0)$ and derived the distribution parameter as

$$z = \frac{n^2 - 4\,\sigma_Z^2}{2\,[4\,\sigma_Z^2 - n]} \qquad \text{Eq. A2}$$

From Eq. A2 it is easily seen that for a fixed n this parameter grows to infinity when $\sigma_z^2 = Var(Z_0)$ approaches $n / 4$. This is the limiting case in which a symmetric beta-binomial distribution becomes a binomial distribution with $p = .5$ and same n. To include the case $Var(Z_0) \leq n/4$, we used a binomial distribution with $p = 0.5$ and $n = \sigma_z^2/4$ as starting point distribution. This way, the starting point distribution is increasingly centered around the neutral starting point ($Z_0 = k$) with decreasing $Var(Z_0)$, so that for $Var(Z_0) = 0$ the process starts at the neutral starting point with probability 1 (i.e., no starting point variation).

The third source of additional variability in the model is temporally inhomogeneous processing. Rather than assuming a constant processing rate (or processing capacity), the time course of processing speed is modeled as a scaled gamma distribution function with variable rate μ and fixed shape n:

$$v(x, i, t) = v(x, i)\,[1 - \exp(-\mu t) \sum_{j=0}^{n-1} \frac{(\mu t)^j}{j!}] \qquad \text{Eq. A3}$$

This functional form has been suggested to approximate the time-dependent output from early perceptual filters (Smith & Van Zandt, 2000). This assumption can also be viewed to yield a

physiologically more plausible model than one with constant processing rates (Christensen et al., 2018). Technically, it introduces additional variation at the leading edge of the RT distribution, as the assumption only affects the speed of early, but not late responses. Analogous to Smith and Van Zandt (2000) and Smith and Ratcliff (2009), we fixed the shape parameter ($n = 5$) and left the rate parameter μ as a free parameter to be estimated from data. The effect of parameter μ is that a high value leads to a quick increase to the asymptotical level, approximating the constant processing model with increasing values of μ, whereas low values lead to a slowly increasing leading edge of the predicted RT distribution of all categorizations. It is assumed that both rates follow the same time course. A graphical illustration of this function (with $\mu = 60$) is displayed in Figure 3. This form of inhomogeneous processing rates influences early finishing processes only; with increasing time the effect becomes negligible, since there will always be a constant offset (included in the non-decision time T_0) so that constant processing will approximate inhomogeneous processing in the long run (cf. Figure 3B).

The full model entails an amount of flexibility that is usually not needed in empirical applications. For the present study, we introduced two restrictions to obtain a strong and testable version of the model. First, we did not distinguish between leftward moving Gabor patches and rightward moving Gabor patches but analyzed and fitted the responses as correct responses and errors. Thus, we introduced processing rates for correct perceptual decisions that were independent of motion direction (i.e., a leftward moving Gabor patch had the same processing rate for the perceptual category "moving leftward" as a rightward moving Gabor patch had for "moving rightward"). In line with this, we set $k_L = k_R = k$, that is, we specified a symmetric model for unbiased decisions. Second, we assume that the processing capacity C is constant across all stimuli used in the experiment; so, we introduced processing capacity C as a model parameter and rescaled the processing rates $p(x, i) = v(x, i)/C$, so that only a single processing rate was needed for each spatial frequency condition (since $p(x, L) + p(x, R) = 1$). Together with the first restriction, $p_x = p(x, correct)$ was defined as the probability of making a correct categorization, so that these rescaled (normalized) processing rates can be interpreted as the quality of information sampled from stimulus x.

References

Adelson, E. H., & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America. A, Optics and image science*, *2*(2), 284-299.

Akyurek, E. G., & van Asselt, E. M. (2015). Spatial attention facilitates assembly of the briefest percepts: Electrophysiological evidence from color fusion. *Psychophysiology*, *52*(12), 1646-1663.

Alitto, H. J., Moore, B. D., 4th, Rathbun, D. L., & Usrey, W. M. (2011). A comparison of visual responses in the lateral geniculate nucleus of alert and anaesthetized macaque monkeys. *The Journal of Physiology*, *589*(Pt 1), 87-99.

Aydin, A., Ogmen, H. & Kafaligonul, H. (2021). Neural correlates of metacontrast masking across different contrast polarities. *Brain Structure and Function, 266*, 3067-3081.

Baruch, O., & Yeshurun, Y. (2014). Attentional attraction of receptive fields can explain spatial and temporal effects of attention. *Visual Cognition, 22*(5), 704-736.

Blurton, S. P., Kyllingsbæk, S., Nielsen, C. S., & Bundesen, C. (2020). A Poisson random walk model of response times. *Psychological Review, 127*, 362-411.

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

Bundesen, C. (1990). A theory of visual attention. *Psychological Review, 97*, 523-547.

Bundesen, C., Vangkilde, S., & Petersen, A. (2014). Recent developments in a computational theory of visual attention. *Vision Research*, *116*, 210–218.

Carpenter, R.H.S. (1988). *Movements of the eyes* (2nd ed.). London: Pion.

Carrasco, M., & Yeshurun, Y. (2009). Covert attention effects on spatial resolution. *Progress in Brain Research, 176*, 65-86.

Christensen, J. H., Markussen, B., Bundesen, C., & Kyllingsbæk, S. (2018). A physiologically based nonhomogeneous Poisson counter model of visual identification. *Journal of Experimental Psychology: Human Perception and Performance, 44,* 1383-1398.

Denison, R. N., Vu, A. T., Yacoub, E., Feinberg, D. A., & Silver, M. A. (2014). Functional mapping of the magnocellular and parvocellular subdivisions of human LGN. *Neuroimage, 102*(2), 358-369.

Derrington, A. M., & Lennie, P. (1984). Spatial and temporal contrast sensitivities of neurones in lateral geniculate nucleus of macaque. *Journal of Physiology, 357*(1), 219-240.

Donkin, C. Brown, S., & Heathcote, A. (2011). Drawing conclusions from choice response time models: A tutorial using the linear ballistic accumulator. *Journal of Mathematical Psychology, 55*, 140-151.

Elkin, L. A., Kay, M., Higgins, J. J., & Wobbrock, J. O. (2021). An Aligned Rank Transform Procedure for Multifactor Contrast Tests. *arXiv preprint arXiv:2102.11824*.

Feller, W. (1968). *An introduction to probability theory and its applications*. *Vol. I (3rd ed.)*. New York: John Wiley & Sons.

Forbes, C., Evans, M., Hastings, N., & Peacock, B. (2011). *Statistical Distributions (4th ed.).* Hoboken: John Wiley & Sons.

Glasser, D. M., & Tadin, D. (2011). Increasing stimulus size impairs first- but not second-order motion perception. *Journal of Vision, 11*, 1–8.

Guzman-Martinez, E., Leung, P., Franconeri, S., Grabowecky, M., & Suzuki, S. (2009). Rapid eye-fixation training without eyetracking. *Psychonomic Bulletin & Review*, *16*(3), 491–496.

Hein, E., Rolke, B., & Ulrich, R. (2006). Visual attention and temporal discrimination: Differential effects of automatic and voluntary cueing. *Visual Cognition, 13*(1), 29-50.

Higgins, J. J., Blair, R. C., & Tashtoush, S. (1990). The aligned rank transform procedure. *Proceedings of the Conference on Applied Statistics in Agriculture*. Kansas State, 185-195.

Higgins, J. J., & Tashtoush, S. (1994). An aligned rank transform test for interaction. *Nonlinear World, 1*(2), 201-211.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics, 6*(2), 65-70.

Hung, S. C., & Carrasco, M. (2021). Feature-based attention enables robust, long-lasting location transfer in human perceptual learning. *Scientific Reports*, *11*(1), 13914.

Kafaligonul, H., Breitmeyer, B. G., & Ogmen, H. (2009). Effects of contrast polarity in paracontrast masking. *Attention, Perception, & Psychophysics, 71*(7), 1576–1587.

Kaplan, E. (2012). The M, P and K pathways of the primate visual system revisited. In J. Werner & L. Chalupa (Eds.), *The New Visual Neuroscience* (pp. 481-494). MIT Press.

Kaplan, E., & Benardete, E. (2001). The dynamics of primate retinal ganglion cells. *Progress in Brain Research*, *134*, 17-34.

Kaplan, E. L. & Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American statistical association*, *53*(282), 457–481.

Kleiner, M., Brainard, D., & Pelli, D. (2007). What's new in Psychtoolbox-3? *Perception, 36*, ECVP Abstract Supplement.

Laming, D. R. J. (1968). *Information theory of choice-reaction times*. London: Academic Press.

Leys, C., Ley, C. Klein, O., Bernarnd, P., Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology, 49*(4), 764-766.

Liu, C. S., Bryan, R. N., Miki, A., Woo, J. H., Liu, G. T., & Elliott, M. A. (2006). Magnocellular and parvocellular visual pathways have different blood oxygen level-dependent signal time courses in human primary visual cortex. *American Journal of Neuroradiology, 27*(8), 1628-1634.

Liu, T., Fuller, S., & Carrasco, M. (2006). Attention alters the appearance of motion coherence. *Psychonomic Bulletin & Review, 13*(6), 1091-1096.

Livingstone, M. S., & Hubel, D. H. (1988). Do the relative mapping densities of the magno- and parvocellular systems vary with eccentricity? *Journal of Neuroscience, 8*(11), 4334-4339.

Megna, N., Rocchi, F., & Baldassi, S. (2012). Spatio-temporal templates of transient attention revealed by classification images. *Vision Research*, *54*, 39-48.

Merigan, W. H., Byrne, C. E., & Maunsell, J. H. (1991). Does primate motion perception depend on the magnocellular pathway? *Journal of Neuroscience, 11*(11), 3422-3429.

Merigan, W. H., Katz, L. M., & Maunsell, J. H. (1991). The effects of parvocellular lateral geniculate lesions on the acuity and contrast sensitivity of macaque monkeys. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *11*(4), 994-1001.

Merigan, W. H., & Maunsell, J. H. (1990). Macaque vision after magnocellular lateral geniculate lesions. *Visual Neuroscience, 5*(4), 347-52.

Morgan, M., Dillenburger, B., Raphael, S., & Solomon, J. A. (2012). Observers can voluntarily shift their psychometric functions without losing sensitivity. *Attention, Perception, & Psychophysics, 74*, 185–193.

Movshon, J. A., Kiorpes, L., Hawken, M. J., & Cavanaugh, J. R. (2005). Functional maturation of the macaque's lateral geniculate nucleus. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *25*(10), 2712-2722.

Newsome, W. T., Britten, K. H., & Movshon, J. A. (1989). Neuronal correlates of a perceptual decision. *Nature, 341*, 52-54.

Newsome, W. T., & Paré, E. B. (1988). A selective impairment of motion perception following lesions of the middle temporal visual area (MT). *Journal of Neuroscience, 8*, 2201-2211.

Nicol, J. R., Watter, S., Gray, K., & Shore, D. I. (2009). Object-based perception mediates the effect of exogenous attention on temporal resolution. *Visual Cognition, 17*(4), 555-573.

Ogmen, H. (1993). A neural theory of retino-cortical dynamics. *Neural Networks, 6*(2), 245-273.

Ogmen, H., Breitmeyer, B. G., & Melvin, R. (2003). The what and where in visual masking. *Vision Research, 43*(12), 1337-1350.

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

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32*(1), 3–25.

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: theory and data for two-choice decision tasks. *Neural Computation, 20*, 873–922.

Ratcliff, R., & Tuerlinckx, F. (2002). Estimating parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameter variability. *Psychonomic Bulletin & Review*, *9*(3), 438-481.

Rouder, J. N. (1996). Premature sampling in random walks. *Journal of Mathematical Psychology*, *40*, 287-296.

Rousseeuw, P. J., & Croux, C. (1993). Alternatives to the median absolute deviation. *Journal of the American Statistical Association, 88*(424), 1273-1283.

Salter, K. C., & Fawcett, R. F. (1993). The art test of interaction: A robust and powerful rank test of interaction in factorial models. *Communications in Statistics: Simulation and Computation 22*(1), 137-153.

Schiller, P. H., Logothetis, N. K., & Charles, E. R. (1990). Role of the color-opponent and broad-band channels in vision. *Visual Neuroscience*, *5*(4), 321-346.

Shibuya, H. & Bundesen, C. (1988). Visual selection from multielement displays: measuring and modeling effects of exposure duration. *Journal of Experimental Psychology: Human Perception and Performance*, *14*, 591–600.

Sincich, L. C., & Horton, J. C. (2005). The circuitry of V1 and V2: integration of color, form, and motion. *Annual Review of Neuroscience*, *28*, 303-326.

Smith, P. L., & Ratcliff, R. (2009). An integrated theory of attention and decision making in visual signal detection. *Psychological Review, 116*, 283-317.

Smith, P. L., Ratcliff, R., & Wolfgang, B. J. (2004). Attention orienting and the time course of perceptual decisions: response time distributions with masked and unmasked displays. *Vision Research*, *44*, 1297–1320.

Smith, P. L., & Van Zandt, T. (2000). Time- dependent Poisson counter models of response latency in simple judgment. *British Journal of Mathematical and Statistical Psychology*, *53*(2), 293-315.

Tadin, D., & Lappin, J. S. (2005). Optimal size for perceiving motion decreases with contrast. *Vision Research, 45*, 2059–2064.

Tadin, D., Lappin, J. S., Gilroy, L. A., & Blake, R. (2003). Perceptual consequences of centre-surround antagonism in visual motion processing. *Nature*, *424*(6946), 312-315.

Takeuchi, T., & De Valois, K. K. (1997). Motion-reversal reveals two motion mechanisms functioning in scotopic vision. *Vision Research*, *37*(6), 745-755.

Thaler, L., Schütz, A. C., Goodale, M. A., & Gegenfurtner, K. R. (2013). What is the best fixation target? The effect of target shape on stability of fixational eye movements. *Vision Research*, *76*, 31-42.

Tootell, R. B., Hamilton, S. L., & Switkes, E. (1988). Functional anatomy of macaque striate cortex. IV. Contrast and magno-parvo streams. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *8*(5), 1594-1609.

Townsend, J. T., & Ashby, F. G. (1983). *Stochastic modeling of elementary psychological processes*. Cambridge University Press.

Wobbrock, J. O., Findlater, L., Gergle, D. and Higgins, J. J. (2011). The aligned rank transform for nonparametric factorial analyses using only ANOVA procedures. *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI '11).* New York: ACM Press, pp. 143-146.

World Medical Association. (2013). World Medical Association Declaration of Helsinki ethical principles for medical research involving human subjects. *JAMA: Journal of the American Medical Association, 310*(20), 2191–2194.

Wright, R. D., & Ward, L. M. (2008). Orienting of attention. New York: Oxford University Press.

Yeshurun, Y. (2004). Isoluminant stimuli and red background attenuate the effects of transient spatial attention on temporal resolution. *Vision Research, 44*(12), 1375-1387.

Yeshurun, Y., & Carrasco, M. (1999). Spatial attention improves performance in spatial resolution tasks1. *Vision Research, 39*(2), 293-306.

Yeshurun, Y., & Carrasco, M. (2000). The locus of attentional effects in texture segmentation. *Nature Neuroscience, 3*(6), 622.

Yeshurun, Y., & Hein, E. (2011). Transient attention degrades perceived apparent motion. *Perception, 40*(8), 905-918.

Yeshurun, Y., & Levy, L. (2003). Transient spatial attention degrades temporal resolution. *Psychological Science, 14*(3), 225-231.

Yeshurun, Y., & Marom, G. (2008). Transient spatial attention and the perceived duration of brief visual events. *Visual Cognition, 16*(6), 826-848.

Yeshurun, Y., & Sabo, G. (2012). Differential effects of transient attention on inferred parvocellular and magnocellular processing. *Vision Research, 74*, 21-29.

Zhang, Y., Shelchkova, N., Ezzo, R., & Poletti, M. (2021). Transient perceptual enhancements resulting from selective shifts of exogenous attention in the central fovea. *Current Biology, 31*(12), 2698-2703.

Zhang, P., Zhou, H., Wen, W., & He, S. (2015). Layer-specific response properties of the human lateral geniculate nucleus and superior colliculus. *NeuroImage*, *111*, 159-166.