Video Article

Measuring Attention and Visual Processing Speed by Model-based Analysis of Temporal-order Judgments

Jan Tünnermann[1], Alexander Krüger[1], Ingrid Scharlau[1]

[1]Faculty of Arts and Humanities, Paderborn University

Correspondence to: Jan Tünnermann at jeti@mail.upb.de

Abstract

This protocol describes how to conduct temporal-order experiments to measure visual processing speed and the attentional resource distribution. The proposed method is based on a new and synergistic combination of three components: the temporal-order judgments (TOJ) paradigm, Bundesen's Theory of Visual Attention (TVA), and a hierarchical Bayesian estimation framework. The method provides readily interpretable parameters, which are supported by the theoretical and neurophysiological underpinnings of TVA. Using TOJs, TVA-based estimates can be obtained for a broad range of stimuli, whereas traditional paradigms used with TVA are mainly limited to letters and digits. Finally, the meaningful parameters of the proposed model allow for the establishment of a hierarchical Bayesian model. Such a statistical model allows assessing results in one coherent analysis both on the subject and the group level.

To demonstrate the feasibility and versatility of this new approach, three experiments are reported with attention manipulations in synthetic pop-out displays, natural images, and a cued letter-report paradigm.

Video Link

The video component of this article can be found at https://www.jove.com/video/54856/

Introduction

How attention is distributed in space and time is one of the most important factors in human visual perception. Objects that capture attention because of their conspicuity or importance are typically processed faster and with higher accuracy. In behavioral research, such performance benefits have been demonstrated in a variety of experimental paradigms. For instance, allocating attention to the target location speeds up the reaction in probe detection tasks[1]. Similarly, the accuracy of reporting letters is improved by attention[2]. Such findings prove that attention enhances processing, but they remain hopelessly mute about how this enhancement is established.

The present paper shows that low-level mechanisms behind attentional advantages can be assessed by measuring the processing speed of individual stimuli in a model-based framework that relates the measurements to fine-grained components of attention. With such a model, the overall processing capacity and its distribution among the stimuli can be inferred from processing speed measurements.

Bundesen's Theory of Visual Attention (TVA)[3] provides a suitable model for this endeavor. It is typically applied to data from letter report tasks. In the following, the fundamentals of TVA are explained and it is shown how they can be extended to model temporal-order judgment (TOJ) data obtained with (almost) arbitrary stimuli. This novel method provides estimates of processing speed and resource distribution which can be readily interpreted. The protocol in this article explains how to plan and conduct such experiments and details how the data can be analyzed.

As mentioned above, the usual paradigm in TVA-based modeling and estimation of attention parameters is the letter report task. Participants report the identities of a set of letters which is briefly flashed and typically masked after a varying delay. Among other parameters, the rate at which visual elements are encoded into visual short-term memory can be estimated. The method has been successfully applied to questions in fundamental and clinical research. For instance, Bublak and colleagues[4] assessed which attentional parameters are affected in different stages of age-related cognitive deficits. In fundamental attention research, Petersen, Kyllingsbæk, and Bundesen[5] used TVA to model the attentional dwell time effect, the observer's difficulty in perceiving the second of two targets at certain time intervals. A major drawback of the letter report paradigm is that it requires sufficiently overlearned and maskable stimuli. This requirement limits the method to letters and digits. Other stimuli would require heavy training of participants.

The TOJ paradigm requires neither specific stimuli nor masking. It can be used with any kind of stimuli for which the order of appearance can be judged. This extends the stimulus range to pretty much everything that could be of interest, including direct cross-modal comparisons[6].

Investigating attention with TOJs is based on the phenomenon of attentional prior entry which is a measure of how much earlier an attended stimulus is perceived compared to an unattended one. Unfortunately, the usual method for analyzing TOJ data, fitting observer performance

psychometric functions (such as cumulative Gaussian or logistic functions), cannot distinguish whether attention increases the processing rate of the attended stimulus or if it decreases the rate of the unattended stimulus[7]. This ambiguity is a major problem because the question whether the perception of a stimulus is truly enhanced or if it benefits because of the withdrawal of resources from a competing stimulus is a question of both fundamental and practical relevance. For instance, for the design of human-machine interfaces it is highly relevant to know if increasing the prominence of one element works at the expense of another one.

The TOJ task usually proceeds as follows: A fixation mark is presented for a brief delay, typically a randomly drawn interval shorter than a second. Then, the first target is presented, followed after a variable stimulus onset asynchrony (SOA) by the second target. At negative SOAs, the *probe*, the attended stimulus, is shown first. At positive SOAs, the *reference*, the unattended stimulus, leads. At an SOA of zero, both targets are shown simultaneously.

Typically, presenting the target refers to switching the stimulus on. Under certain conditions, however, other temporal events, such as a flicker of an already present target or offsets are used[8].

In TOJs, responses are collected in an unspeeded manner, usually by keys mapped to the stimulus identities and presentation orders (*e.g.*, if stimuli are squares and diamonds, one key indicates "square first" and another one "diamond first"). Importantly, for the evaluation, these judgments must be converted to "probe first" (or "reference first") judgments.

In the present work, a combination of the processing model of TVA and the TOJ experimental paradigm is used to eliminate the problems in either individual domain. With this method, readily interpretable speed parameters can be estimated for almost arbitrary visual stimuli, enabling to infer how the observer's attention is allocated to competing visual elements.

The model is based on TVA's equations for the processing of individual stimuli, which will be shortly explained in the following. The probability that one stimulus is encoded into visual short-term memory before the other is interpreted as the probability of judging this stimulus as appearing first. The individual encoding durations are exponentially distributed[9]:

$$F(t) = \begin{cases} 1 - e^{-v_{x,i}(t-t_0)}, & \text{if } t > t_0 \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

The maximum ineffective exposure duration t_0 is a threshold before which nothing is encoded at all. According to TVA, the rate $v_{x,i}$ at which object x is encoded as member of a perceptual category i (such as color or a shape) is given by the *rate equation*,

$$v_{x,i} = \eta_{x,i} \beta_i \frac{w_x}{\sum_{z \in S} w_z} . \quad (2)$$

The strength of the sensory evidence that x belongs to category i is expressed in $\eta_{x,i}$, and β_i is a decision bias for categorizing stimuli as members of category i. This is multiplied by attentional weights. Individual attentional weights w_x are divided by the attentional weights of all objects in the visual field. Hence, the relative attentional weight is calculated as

$$w_x = \sum_{j \in R} \eta_{x,i} \pi_{j}, \quad (3)$$

where R represents all categories and $\eta_{x,i}$ represents the sensory evidence that object x belongs to category j. The value π_j is called pertinence of category j and reflects a bias to make categorizations in j. The overall processing capacity C is the sum of all processing rates for all stimuli and categorizations. For a more detailed description of TVA, refer to Bundesen and Habekost's book[9].

In our novel method, Equation 1, which describes the encoding of individual stimuli, is transformed into a model of TOJs. Assuming that selection biases and report categories are constant within an experimental task, the processing rates v_p and v_r of the two target stimuli probe (p) and reference (r) depend on C and the attentional weights in the form $v_p = C \cdot w_p$ and $v_r = C \cdot w_r$, respectively. The new TOJ model expresses the success probability $P_{p^{1st}}$ that a participant judges the probe stimulus to be first as a function of the SOA and the processing rates. It can be formalized as follows:

$$P_{p^{1st}}(v_p, v_r, SOA) = \begin{cases} 1 - e^{-v_p|SOA|} + e^{-v_p|SOA|}\left(\frac{v_p}{v_p+v_r}\right) & \text{if } SOA < 0 \\ e^{-v_r|SOA|}\left(\frac{v_p}{v_p+v_r}\right) & \text{otherwise.} \end{cases} \quad (4)$$

A more detailed description of how this equation is derived from the basic TVA equations is described by Tünnermann, Petersen, and Scharlau[7].

For the sake of simplicity, the parameter t_0 is omitted in the model in Equation 1. According to the original TVA, t_0 should be identical for both targets in the TOJ task, and, therefore, it cancels out. However, this assumption may sometimes be violated (see section Discussion).

For fitting this equation to TOJ data, a hierarchical Bayesian estimation scheme[11] is suggested. This approach allows to estimate the attentional weights w_p and w_r of the probe and reference stimuli and the overall processing rate C. These parameters, the resulting uptake rates v_p and v_r, and attention-induced differences between them, can be assessed on the subject and group levels along with estimated uncertainties. The hierarchical model is illustrated in Figure 1. During the planning stage for an experiment, convenient Bayesian power analysis can be conducted.

The following protocol describes how to plan, execute and analyze TOJ experiments from which processing speed parameters and attentional weights for visual stimuli can be obtained. The protocol assumes that the researcher is interested in how an attentional manipulation influences the processing speeds of some targets of interest.
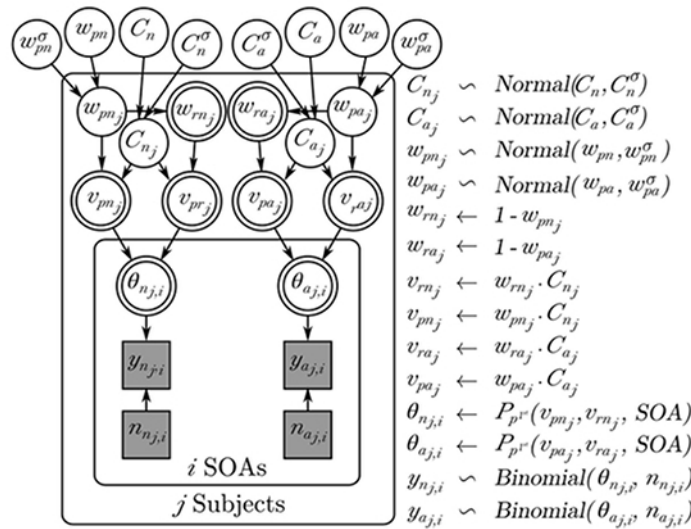
Figure 1: Graphical model used in the Bayesian estimation procedure. Circles indicate estimated distributions; double circles indicate deterministic nodes. Squares indicate data. The relations are given on the right side of the figure. The nodes outside the rounded frames ("plates") represent mean and dispersion estimates of TVA parameters (see Introduction) on the group level. In the "j Subjects" plate, it can be seen how attentional weights (w) are combined with the overall processing rates (C) to from stimulus processing rates (v) on the subject level. Plate "i SOAs" shows how these TVA parameters are then transformed (via the function $P_{p^{1st}}$ described in the Introduction) into the success probability (θ) for the binomially distributed responses at each SOA. Therefore, the θ together with the repetitions of the SOA (n) describe the data points (y). For more details on notation and interpretation of graphical models, refer to Lee and Wagenmakers[23]. Note that for the sake of clarity, the nodes that represent differences of parameters have been omitted. These deterministic parameters are indicated in the figures of the experimental results instead. Please click here to view a larger version of this figure.

Protocol

NOTE: Some steps in this protocol can be accomplished using custom software provided (along with installation instructions) at http://groups.upb.de/viat/TVATOJ. In the protocol, this collection of programs and scripts is referred to as "TVATOJ".

1. Selection of Stimulus Material

1. **Select stimuli according to the research question.**
   NOTE: In general, two targets are shown at different locations on the screen. Stimuli that have been used with the present method include, for example, shapes, digits, letters, singletons in pop-out displays, and objects in natural images. The latter three types were used in this protocol.
   NOTE: Several different stimulus types are included in the TOJ plugin ("psylab_toj_stimulus" provided with TVATOJ) for the experiment builder OpenSesame[12].
2. When creating new stimulus types, make sure that the properties of interest have to be encoded for the judgment by making them important for the task or select stimuli where the properties of interest are encoded automatically (*e.g.*, singletons in pop-out displays).

2. Power Estimation and Planning

1. **Perform a Bayesian power analysis by simulating data sets with the selected model, planned design (SOA distribution and repetitions), sample sizes, and hypothesized parameters. Estimate whether it is likely to reach the research goal (for instance, a certain difference in the parameters). If the power is not sufficient, alter the design by adding or shifting SOAs or repetitions and repeat the analysis.**
   1. To use the provided TVATOJ software, open and edit the script "exp1-power.R". Follow the comments in the file to adjust it for the specific analysis. For general information on Bayesian power estimation refer to Kruschke[13].

3. Specification or Programming of the Experiment

1. **Use an experiment builder or psychophysical presentation library to implement the experiment.**
   1. To use the OpenSesame TOJ plugin provided in TVATOJ, drag the "psylab_toj_stimulus" plugin into a trial presentation loop. Alternatively, open the "simple-toj.osexp" example experiment in OpenSesame.
   2. Select the desired stimulus type from the dropdown menu "Stimulus type" in the psylab_toj_stimulus configuration. Follow the instructions in TVATOJ for adding new stimulus types if required.
2. **Specify the trials as described in the following steps.**

1. For every experimental condition, create trials with the planned SOAs. When using the psylab_toj_stimulus plugin and OpenSesame, add all varied factors as variables to the trial loop (*e.g.* "SOA").
2. Add rows to the table to realize all factor combinations (**e.g.**, seven SOAs, from -100 to 100 msec, crossed with the experimental conditions "attention" and "neutral"). Adjust the loop's "Repeat" attribute to create sufficient repetitions (see protocol step 2 for determining the distribution and repetition of SOAs).
   NOTE: Typically, at most 800 trials can be presented within one hour. If more repetitions are needed, consider splitting the experiment into several sessions. Make sure that the "Order" attribute of the loop is set to "Random" before running the experiment.
3. In the psylab_toj_stimulus plugin configuration, add placeholders (*e.g.* "[SOA]") for the varied factors in the respective fields. Enter constant values in the fields of factors that are not varied.
   NOTE: Before running the experiment, make sure that accurate timing is guaranteed. If appropriate timing behavior of newer monitors was not verified, use CRT monitors and synchronize with the vertical retrace signal[12].

4. Experimental Procedure

1. **Welcoming and briefing of the participants**
   1. Welcome the participants and inform them about the general form of the experiment (computer-based perception experiment). Inform the participants about the prospective duration of the experiment. Obtain the participants' informed consent to participate in the experiment.
   2. Ensure that the participants show normal or corrected-to-normal vision (optimally by conducting short vision tests). Some deficits, such as color blindness, may be tolerable if they do not interfere with the research question for the particular type of stimulus material.
   3. Provide a quiet booth where the experiment is conducted. Adjust the chair, chin rest, keyboard position, and so on, to ensure optimal viewing and response conditions for the experiment.
   4. Make the participants aware that the experiment requires attention and mental focus and can be fatiguing. Ask them to take short breaks when required. It is, however, equally important not to perform these simple tasks under strong attentional strain. Tell the participants that it is okay to make some errors.

2. **Instruction and warm-up**
   1. Present onscreen instructions for the task, detailing the presentation sequence and response collection procedure. Inform the participants that the task is to report the order in which the targets arrived, and that this will be difficult in some trials. Ask the participants to report their first impression when they cannot tell the order for certain, and let them guess if they have no such impression at all.
      NOTE: In the binary TOJs used here, there is no option to indicate the perception of simultaneity. To avoid excessive guessing, do not point out the presence of trials with simultaneously presented targets explicitly. Let these simply be difficult trials with the instructions outlined above.
   2. To avoid eye movements during the trials, ask participants to fixate a mark which is shown in the center of the screen. Ask them to rest their head on a chin rest.
   3. Ask the participants to take short breaks if necessary. Let them know when breaks are allowed and when they must be avoided (*e.g.*, during the target presentation and before the response).
   4. Include a short training in which participants can get used to the task. To this end, present a random subset of the experimental trials (see protocol step 3.2).
      NOTE: Because the task itself is rather simple, ten to twenty trials are usually sufficient. It can be advantageous to increase the participants' confidence in their performance in this task. This can be done by slowing down the presentation and providing feedback.
   5. Obtain the participants' confirmation that they have understood the task (let them explain it) and that they have no further questions.

3. **Running the main experiment**
   1. Let the experimental software start with the presentation of the main trials. Leave the booth for the main experiment.

5. Model-based Analysis of the TOJ Data

1. Convert the raw data files into counts of "probe first" judgments for every SOA. For instance, run the script "os2toj.py" provided with TVATOJ.
2. Run the Bayesian estimation procedure to estimate the main parameters w_p and C, the derived ones v_p and v_r and the differences of the parameters. For this purpose, run the script "run-evaluation.R" after editing it according to the instructions provided in the file.
3. When the sampling has finished, the differences of interest for the research questions can be assessed. Examples can be found in the following section.

Representative Results

In the following, results obtained with the proposed method are reported. Three experiments measured the influence of different attentional manipulations with three highly different types of stimulus material. The stimuli are simple line segments in pop-out patterns, action space objects in natural images, and cued letter targets.

Experiment 1: Salience in pop-out displays

Experiment 1 aimed at measuring the influence of visual salience on the processing speed of line segments in a synthetic pattern. Subjects judged which of two target line segments (left or right) in a background pattern of oriented line segments flickered first. In half of the trials, the probe was a color pop-out (see **Figure 2a**). More background on TOJ-based assessment of pop-out can be found in a study by Krüger and colleagues[8], where local orientation salience was manipulated instead of color. The distribution and frequency of the SOAs are shown in **Figure 2b**.

A Bayesian power analysis was conducted as described in the protocol step 2. For typical group overall rates C (M = 70 Hz, SD = 20) and a hypothetical advantage from about 7 Hz for the salient target in the attention condition (resulting from an attentional weight of M = 0.55, SD = 0.02), 200 simulations were performed. The success rate for detecting the advantage was calculated for the lower boundary of the 95 % HDI (highest density interval) being above 4 Hz and fulfilling additional requirements regarding the difference to the control condition (see TVATOJ example "power-exp1.R" for all details). The success rate for reaching this goal under the hypothetical conditions with 25 participants turned out to be 0.88 with a 95 % HDI ranging from 0.82 to 0.92.

For the actual experiment, 30 participants were recruited. One participant was excluded from the analysis because he did not follow the instructions but always pressed the same key.
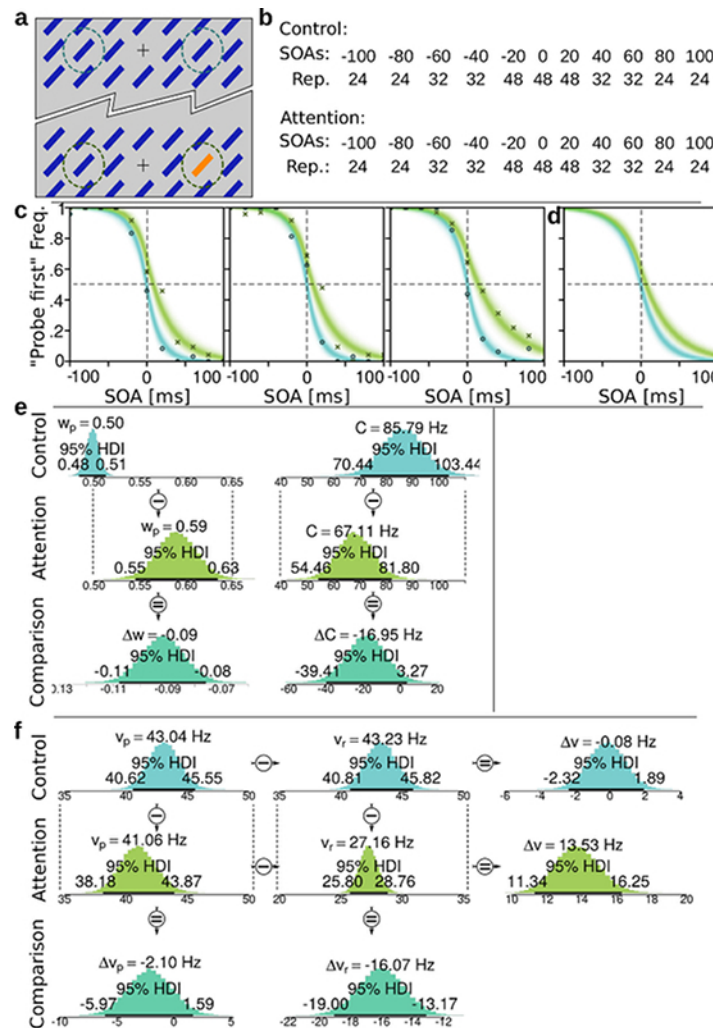

Figure 2: Experiment 1. (**a**) Targets (marked with circles for illustration) in the neutral (upper part) and attention (lower part) condition. (**b**) SOA distributions. (**c**) Three exemplary subject-level response counts (points) and posterior predictive curves (shaded area; intensity represents likelihood with regard to 100 simulated repetitions at fine-grained SOAs). Blue indicates control and green attention condition. (**d**) Group-level posterior predictive curves. (**e**) Posterior distributions of the overall rate C and attentional weights w_p and w_r. (**f**) Posterior distributions of v_p and v_p and their differences. Please click here to view a larger version of this figure.

The data was fitted with the TOJ model derived from TVA (as described in the Introduction) using a hierarchical Bayesian procedure implemented in JAGS[10]. **Figure 2c** shows three exemplary subject-level plots of the raw data and posterior predictive estimates obtained from subject-level samples of the fitted model. Group-level posterior predictive curves are shown in **Figure 2d** and parameter estimates in **Figure 2e** and **2f**. Depending on the research question, parameters C and the relative attentional weights w_p and w_r (**Figure 2e**) or the individual rate parameters v_p and v_r (**Figure 2f**) can be assessed. If the overall processing capacity C was changed by the manipulation, the latter parameters can show whether and how the individual stimulus processing rates have changed.

The proposed hierarchical Bayesian estimation procedure offers a wealth of results. For example, all parameters can be assessed for each participant on the subject level. Typically, there is interest in tendencies in the population. Hence, the results on the group level are discussed. The histograms show distributions over the parameter space. The modes of the distributions are stated to indicate the parameters' central tendencies. The 95 % HDIs mark the ranges in which the true values lie with a probability of 95% according to model and data (for further details on how to interpret the Bayesian statistics, refer to Kruschke[11,13]).

Figure 2e shows estimates of the means across subjects for the weight and overall rate parameters. An attentional benefit for the salient stimulus can be seen in the attention condition. The central tendency of parameter w_p is 0.59, and its 95 % HDI ranges from 0.55 to 0.63. Hence, salience shifted the attentional weight away from the neutral value of 0.5. In the control condition, where none of the targets was salient, a neutral weight of w_p= 0.5 was obtained (95 % HDI: 0.48 to 0.51). The corresponding row "Comparison" shows that the difference between the w_p weights across conditions is 0.09, and the 95 % HDI of this difference ranges from -0.11 to 0.07. Hence, there is a reliable difference between the two weights in favor of the salient stimulus.

However, does this mean that the salient target was processed faster? The difference in weights together with the shared overall rate C in the attention condition indicate that it was processed faster than the non-salient target in this condition. However, an important question is whether it was also processed faster than the targets of the control condition. Taking the estimates of the processing rates into account, the answer must be no. The C estimate shown in **Figure 2e** is lower in the attention condition by a difference of almost 17 Hz. In the corresponding "Comparison" plot, 0, no difference, is just at the fringe of the 95 % HDI; hence it is highly unlikely. Considering the individual rates of the probe (v_p) and reference (v_r) stimulus in both conditions (**Figure 2f**), it is clear that the advantage of the salient stimulus results from a 16 Hz reduction of the processing rate of the non-salient stimulus in the attention condition. A possible interpretation of these results is that the salient target leads to a suppression of the non-salient target in the attention condition and thereby benefits in relation.

Note that in this experiment, even though the appearances of probe and reference were identical in the neutral condition, the delay between trial start and probe event was constant. Therefore, participants could have directed attention toward this point in time, thus shifting the attentional weight away from the neutral value 0.5. Consequently, the actual attentional weight of the probe stimulus in the control condition must be estimated and fixed at 0.5. Fixing the parameter is possible when the participant cannot tell even in principle which is the probe and which the reference stimulus, as in the control condition in Experiment 3.

Experiment 2: Action space advantages in natural images
The second experiment measured attentional advantages for objects in the action space in natural images. From change blindness studies it is known that center-of-interest objects benefit from meaning-driven orienting in natural images[14]. This effect is absent when images are obscured by upside-down presentation. In unpublished change blindness experiments, we found an action space advantage with a set of images with changes in action-space and background objects (there is a published replication with similar stimulus material in reference[15]). We hypothesize that these action space objects, which are close to the observer and possibly graspable, exhibit a similar advantage in their processing rates.

Thus, the proposed TOJ-based method is tested with natural images. Action space (probe) and more distant (reference) objects, which appeared abruptly in natural images, constituted the targets for the TOJ procedure (see **Figure 3a**). In a between-subjects control condition, upside-down versions of the same images were used. These are known to have reduced context effects in change detection experiments[15,16]. The SOA between the onsets was varied according to the distribution shown in **Figure 3a**.

A power estimation was performed exactly as for Experiment 1, except that between-subject comparisons between experimental and control condition were conducted. The success rate for achieving the goal was estimated with 0.92 (95 % HDI: 0.88 to 0.96) with 35 simulated participants per condition (details can be found in the TVATOJ example "power-exp2.R").

There were 39 subjects in the attention condition and 38 in the control condition of the actual experiment. (Some subjects participated in both conditions. To our knowledge, this does not compromise the Bayesian statistical analysis. Treating the mixed data as between-subjects reduces the power compared to considering the within-subject differences.) Again, one participant (the same person in both conditions) was removed from the analysis of each condition, due to having given intentionally random responses throughout the experiment.
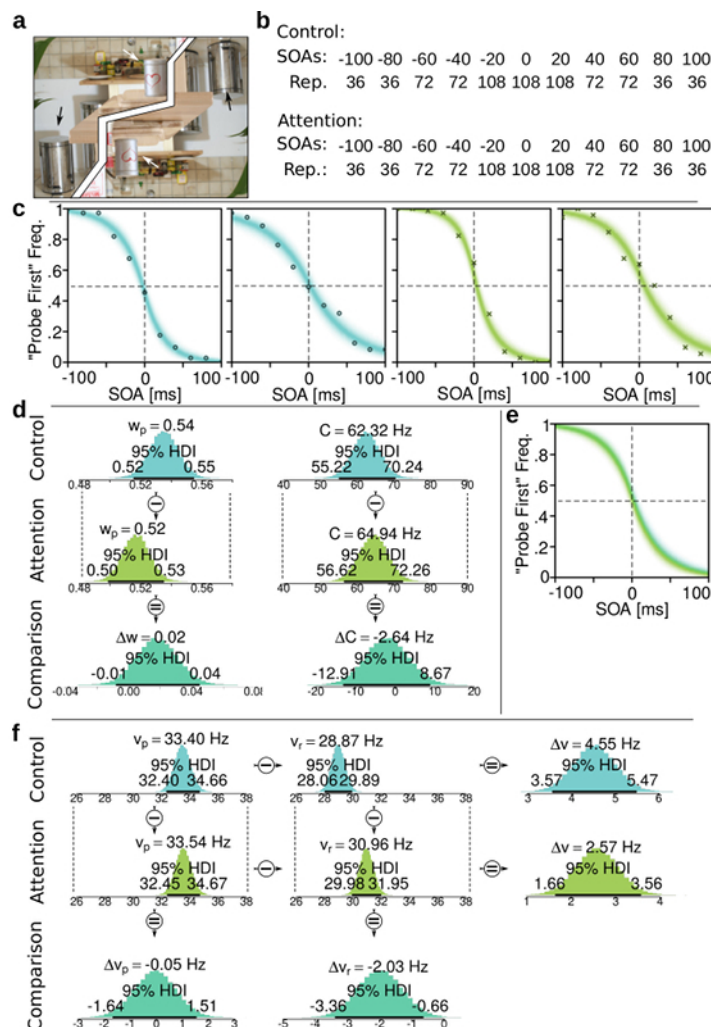
Figure 3: Experiment 2. (**a**) Action space (marked with white arrows) and background (black arrow) targets in the neutral (left) and attention (right) condition. (**b**) SOA distribution. (**c**) Two exemplary subject-level plots from the neutral (blue) condition and two plots from the attention condition (green) with response counts (points) and posterior predictive curves shaded area; intensity represents likelihood with regard to 100 simulated repetitions at fine-grained SOAs). (**d**) Posterior distributions of the overall rate C and attentional weights w_p and w_r. (**e**) Group-level posterior predictive curves. (**f**) Posterior distributions of v_p and v_r and their differences. Please click here to view a larger version of this figure.

This data was fitted in the same manner as in the first experiment. The only difference was that because of the between-subjects design, parameter differences between the two conditions could not be calculated during the sampling at the subject level. This reduces the power in comparison to within-subject differences.

The exemplary subject-level and the group-level posterior predictive distributions in **Figure 3c** show distributions that are barely shifted against each other. The two conditions overlap almost perfectly in the group-level posterior predictive plot (**Figure 3e**), so that it may seem as if the attention manipulation had not worked at all. Inspecting the posterior distributions of the parameters, however, reveals that there indeed is an advantage for the action space objects. The w_p estimate in the attention condition is shifted away from the neutral state of 0.5, which is only at the very left end of the 95 % HDI. Curiously, however, it is also shifted in the control condition, with the 95 % HDI even excluding 0.5, suggesting that the inversion of the images did not remove the potential action space advantage.

Considering the comparison of the individual stimulus processing rates (**Figure 3f**, "Comparison"), an effect of attention can be seen for the rate of the reference stimulus v_r. However, the difference points in the direction opposite to the hypothesis and is small, reflecting a rate change of only 2 Hz (95 % HDI: -3.36 to 0.66).

Therefore, it must be concluded that either (a) the attentional advantage of action space objects is due to a factor that is not affected by the scene inversion, such as salience or visibility. Alternatively, (b) the scene inversion does not reduce the action space effects as intended, or (c) the power of the present experiment was too small to detect the effect. Explanations (a) and (b), or a combination, are the likely ones. In our unpublished change-blindness experiments mentioned earlier, which were conducted with the same images, there was still an advantage (though reduced) for the action space objects in inverted scenes.

In the context of this method-centered paper, however, alternative (c) may be the most interesting one. Therefore, the magnitudes of possibly overlooked effects will be briefly discussed. Looking at the comparison of the attentional weights, the lower bound of the 95% HDI, which reflects

the hypothesized direction, is at -0.01. Hence, only weights larger by 0.01 in the attended compared to the control condition are probable. This difference is small compared to the other experiments, and the odds are against even such a small effect. This is reflected by the upper HDI bound reaching 0.04. Looking at the processing rates is helpful because the rates in Hz can be readily interpreted as processing speed.

The differences between the two conditions are shown in the "Comparison" row in **Figure 3f**. Difference Δv_r between the reference stimuli is negative, -2.03 Hz, and the 95 % HDI excludes 0. The negative difference reflects an increase in the processing rate of the reference targets, the background objects, which too is against the action space advantage hypothesis. A small attentional advantage is still possible in the processing rates of the probe targets, Their difference Δv_p is estimated close to zero, but the 95 % HDI ranges from -1.64 Hz to 1.51 Hz. Even though a value close to zero is most probable, rate effects up to 1.64 Hz in favor of the hypothesis, and up to 1.51 Hz against it, remain possible concerning the 95% HDI. Overall, these results are not favorable for the original hypothesis, but their discussion showed how meaningful sizes of possibly missed effects can be conveniently extracted from the results. Note that for accepting null results, such as the lacking reduction of the action space advantage by rotating the images, regions of practical equivalence can be defined and their overlap with the 95 % HDI can be tested[11] (see **Discussion Section**).

Experiment 3: Spatial cueing in letter recognition
The third experiment investigated the limits of the proposed TVA-based TOJ model and shows how the model can be extended to deal with these difficulties.

What is it that gets the proposed method into trouble? In the two previous experiments, the participants had to judge two temporal events. Now we add a third temporal event, a peripheral cue that is shown 110 msec before the probe stimulus to direct attention toward it. This third event presents difficulties to simple the TVA-based TOJ model, for which only two stimuli are explicitly modeled.
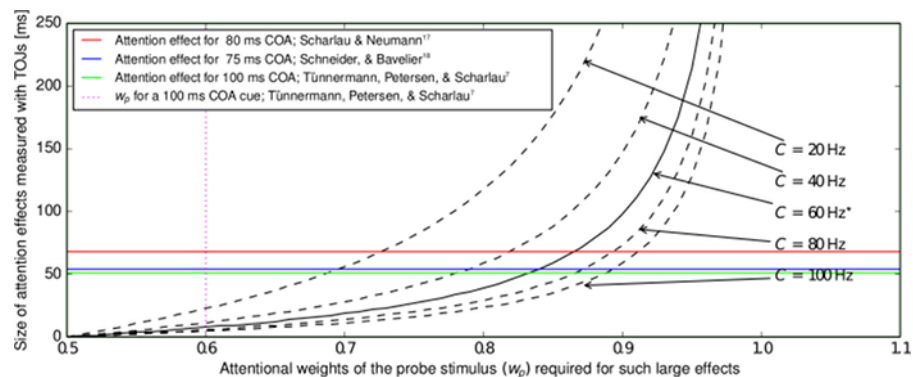

Figure 4: Typical effects of cues on perceptual latencies. Magnitudes of attention effects typically found in TOJs with peripheral cues (horizontal lines). Magnitudes predicted by the TVA-based TOJ model for increasing attentional weights of the probe stimulus (curves). The solid curve corresponds to typically observed parameters. Please click here to view a larger version of this figure.

Magnitudes of attentional benefits reported in the TOJ literature already hint at these difficulties. The latency differences induced by peripheral cues are often as high as 50% and sometimes even as high as 80% of the cueing interval[19]. As illustrated in **Figure 4**, such a large shift would require unlikely attentional weights close to $w_p = 0.9$ at typical processing rates. Furthermore, such extreme weights lead to highly skewed psychometric distributions. These would have a steep slope at one end and a shallow slope at the other end. In a weaker manifestation, this can be seen in the posterior predictive plots of the first experiment (Figure 2c and d). Such strongly distorted curves are rarely reported. When the data from the present peripheral cueing experiment is fitted with the TVA-based TOJ model, the posterior predicted curves strongly deviate from the actual data pattern.

Importantly, however, peripheral cues produce the strongest and most reliable effects of attention in psychophysical TOJ[20]. Therefore, it is worthwhile to apply a model-based assessment with an extended version of the proposed model. Alcalá-Quintana and García-Pérez[21] proposed a TOJ model based on general assumptions of exponential stimulus encoding. This model contains an additional parameter that allows for large shifts without altering the slopes of the psychometric curves. Alcalá-Quintana and García-Pérez used it for crossmodal TOJs, where such shifts originate from delays between modalities. Hence, to model data from cued TOJs, we include their parameter τ. A delay between the start of the encoding processes could account for the expected large lateral shifts. The parameter τ may even have a TVA-compatible interpretation. However, this is not entirely unproblematic and will be discussed later. To keep the model parsimonious, other parameters suggested by Alcalá-Quintana and García-Pérez (response biases, lapses, and a minimum possible temporal resolution) were not included.

Formally, the original psychometric model in Equation 4 is modified by replacing the term SOA by an adjusted term $SOA_{adj} = SOA + τ$. This adjustment also reflects the interpretation of τ: the starts of exponential encoding processes are now not only separated by the SOA, but an additional constant delay is added. In the hierarchical Bayesian model, subject level τ is sampled from a group-level normal distribution.

An explicit power analysis was not performed for this experiment. Because the within-subjects design is similar to the one from Expriment 1, a similar power is expected for effects in the rates and attentional weights. The expected large lateral shift to be captured by the τ parameter is much larger and more stable than the rate and weight effects typically are, so that no power problem can be expected for detecting it either.

Data was collected for 32 participants (among them the three authors) according to the experimental procedure described in the protocol step 4. Participants had to report the order of two letters. In half of the trials the probe stimulus was preceded (110 msec) by a peripheral four-dot cue (see **Figure 5a**). A detailed description of the stimulus material can be found in Tünnermann, Petersen, and Scharlau's study[7]. The SOAs and their frequencies are shown in **Figure 5b**. Each participant performed one or two sessions.
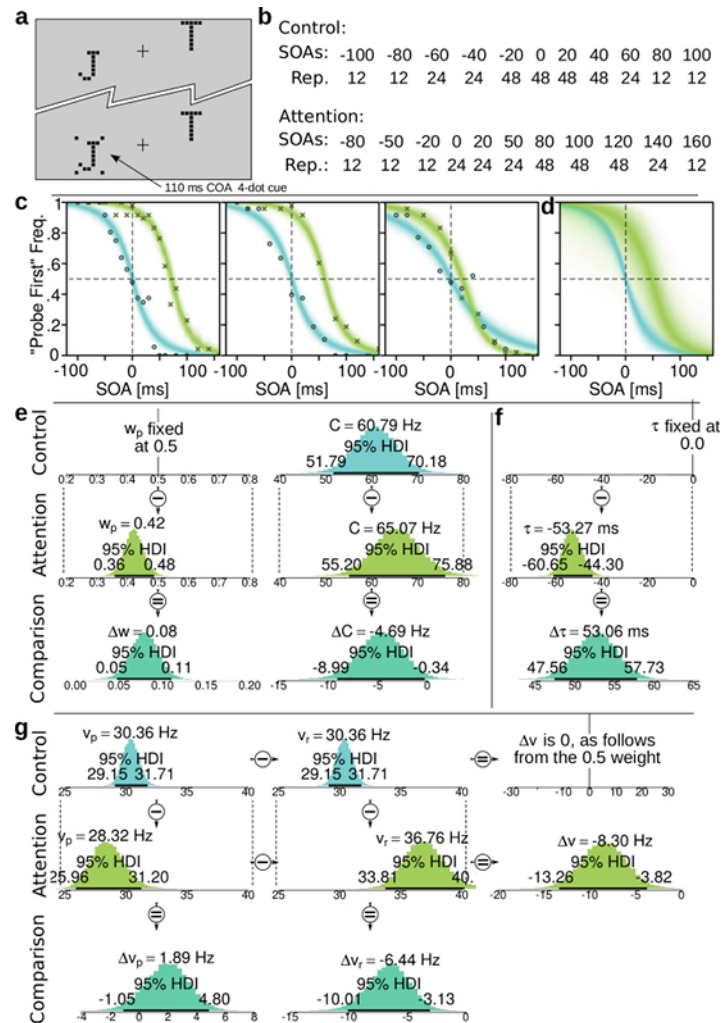
Figure 5: Experiment 3. (**a**) Targets, arbitrarily designated as probe and reference in the the neutral condition (upper part). In the attention condition (lower part) the probe stimulus was preceded (110 msec) by a four-dot cue. (**c**) Three exemplary subject-level response counts (points) and posterior predictive distributions (shaded area; intensity represents likelihood with regard to 100 simulated repetitions at fine-grained SOAs). Blue represents the neutral and green the attention condition. (**d**) Group-level posterior predictive curves. (**e**) Posterior distributions of the overall rate C and attentional weights w_p and w_r. (**f**) Posterior distributions of τ and its difference from zero. (**g**) Posterior distributions of v_p and v_r and their differences. Please click here to view a larger version of this figure.

The extended model as described above was applied in the hierarchical Bayesian estimation procedure. (For participants who produced very steep psychometric curves in the first session, smaller SOAs were used in the second session. This can be seen, for example, in the leftmost plot **Figure 5c**, which contains additional data points at small SOAs.) Because of the more complex model, the powerful NUTS sampler from the Stan software package was employed in this analysis[22].

In the other experiments, the probe stimulus could have a different processing rate than the reference stimulus, even in the control condition. In Experiment 1, this was because the participants could have allocated attention to its predictable point in time relative to the trial onset. In Experiment 2, the inversion of images was not expected to entirely remove the advantage of action space objects. In this letter-based cueing experiment, however, participants would not even in principle be able to identify which is the probe and which the reference stimulus, because the same random letters were used and the time between trial and target onset did not allow to conclude the target type. Therefore, a truly neutral control condition is expected and w_p fixed at 0.5 and τ at 0 in the neutral condition.

As can be seen in **Figure 5c** and **5d**, the cue leads to a substantial shift of the psychometric function compared to the other experiments. Furthermore, the posterior plots in **Figure 5f** show that τ is estimated as a 53.27 ms benefit for the cued target. The 95% HDI on the difference ("Comparison" row) excludes all differences smaller than 47.56 (or larger than 57.73), rendering them highly improbable.

Interestingly, there is a change in attentional weights in favor of the uncued target (**Figure 5e**). The posterior distribution of w_p has its mode at 0.42. The neutral weight of 0.5 is not included in the 95 % HDI. For the C parameter, there is an increase by 4.69 Hz for the attention condition. Expressed in v-parameters (**Figure 5f**), it is most notable that the rate v_r of the reference stimulus in the attention condition increases.

In the previous two experiments, it was observed that the attention manipulations increased the attentional weight of the probe stimulus. In the present experiment, however, the pattern could reflect an interference of the cue with the target, thereby reducing its rate in the race for

encoding. At the same time, the cued target benefits from the faster processing due to the τ parameter. The latter can be linked to the reduction of the cued target's delays before or after the exponential races. Note, however, that concerning τ, an prolongation of a delay associated with the uncued stimulus explains the relative difference equally well.

Discussion

The protocol in this article describes how to conduct simple TOJs and fit the data with models based on fundamental stimulus encoding. Three experiments demonstrated how the results can be evaluated in a hierarchical Bayesian estimation framework to assess the influence of attention in highly different stimulus material. Salience in pop-out displays led to increased attentional weights. Also, increased weights were estimated for action space objects in natural images. However, due to the persisting advantage when spatial relations were disturbed by showing such images upside down, it is likely that another local attentional benefit leads to the weight increase. A peripheral cue, as used in Experiment 3, exhibits a negative influence on the attentional weight. However, it leads to a large effect in the τ parameter, which models a delay between the starting times of the encoding processes.

Most of the protocol follows common steps in conducting TOJs and perception experiments in general. Note, however, that the interpretation of the results in terms of TVA is tied to encoding the stimuli into visual short-term memory. The possibility of performing the TOJ by pure onset detection should be reduced as much as possible. Therefore, as mentioned in protocol step 1.2, it is crucial that the attributes of interest are either automatically encoded (which can be assumed for certain stimuli, *e.g.*, salience pop-outs) or encoding must be facilitated via the task (*e.g.*, reporting the stimulus identity).

It is advisable to inspect the summarized raw data ("probe first" counts across SOAs divided by the number of repetitions) before running the final analysis. This data should follow an S-shaped curve as shown in the psychometric functions in the figures of the Representative Results section. Note that due to the binomially distributed responses, the data points randomly deviate from the ideal path. The deviations increase with a decreasing number of repetitions. With few repetitions, the deviations are often relatively large, obscuring the ideal S shape. However, if the pattern clearly deviates from the usual curve, the mathematical model may need to be adjusted. For instance, when large lateral shifts are observed (as in Experiment 3 of this article), Alcalá-Quintana and García-Pérez's τ parameter can be included. If the curve does not converge to one and zero at its ends, additional lapse parameters[21] can be added.

It is possible to perform a formal model comparison as suggested in Alcalá-Quintana and García-Pérez[21] to decide between different models. Using models different from the one outlined in the Introduction, however, may affect whether results can be interpreted in terms of TVA.

In the results reported in this article, we stated the central tendencies of estimated differences along with their 95% HDIs. However, in the Bayesian framework, it is possible to accept or reject that there is no difference between two estimates. For this purpose, a ROPE (region of practical relevance) must be specified[11,13]. The ROPE indicates a small range around zero. Values within this range are considered practically equal to zero. If the ROPE does not overlap with the HDI, the null hypothesis is rejected. Meaningful ROPE limits depend on the research question or application. In contrast to TOJ analysis with traditional means, the TVA-based approach can guide the establishment of meaningful ROPE limits: Due to their meaningful units, the parameters can be related to estimates from other TVA paradigms (*e.g.*, whole reports, see reference[3]). Furthermore, processing rates can be converted into encoding durations (the expected value of the encoding duration E_x of stimulus x is $1/v_x$, see reference[7]) to inform ROPE limits. For example, if researchers are interested in whether an attention manipulation contributes to a reaction time reduction for a participant in a driving simulation, they could reason as follows: Reaction times (including motor components) are in the range of a few hundred milliseconds, therefore, if the attention manipulation changes the overall reaction only a few milliseconds, the change would be practically zero. Hence a ROPE from -2 to +2 ms could be applied to the difference of reference and probe encoding duration (E_r-E_p). If the ROPE of this difference completely includes the HDI, the result that there is no difference can be accepted. If HDI and ROPE do not overlap, the null hypothesis can be rejected. If neither is the case, no such point decision can be made. Further details concerning the Bayesian evaluation approach in general can be found, for instance, in Kruschke's book[13].

Turning to more general issues, for the success of this protocol, it is crucial that there are only two stimuli that generate temporal signals at the target location. For example, a peripheral cue (as in Experiment 3) or masks[7] lead to large lateral shifts that cannot be accounted for by the current TVA-based model. Such situations are not uncommon and they were modeled by incorporating a parameter suggested by Alcalá-Quintana and García-Pérez[21]. In this extended model, the τ component cannot be clearly linked to a TVA mechanism. There is a tentative connection between τ and TVA, but there are some unsolved problems. Indeed, TVA assumes a short delay before encoding starts. Parameter t_0, which was discussed in the Introduction, is the maximum ineffective exposure duration before which nothing is encoded at all. The difference t_{0r}-t_{0p} could be understood as τ. However, t_0 is typically small, around 10 to 20 msec. Furthermore, the theory does not assume that it is influenced by attention. Nevertheless, t_0 reductions have been observed in letter recognition[7,24]. If one accepts this possibility, a further commitment must be made. Parameter τ was measured around 50 ms. Given the fact that t_{0p} of the cued stimulus can maximally be decreased by 10 to 20 ms because it is not larger in the first place, most of τ would come from increasing t_{0r} of the uncued to 50 to 60 msec. This magnitude is way beyond what sometimes is observed (around 10 msec). As a consequence of τ's unclear relation to TVA, some important questions cannot be answered. For example, it cannot be decided whether the delays of attended stimuli are reduced or if those of the unattended stimuli are prolonged (resulting in the observed τ difference).

The limitations of the technique mentioned above arise from the fact that only two stimuli are explicitly modeled with TVA. To improve on this, future research aims at extending the TVA-based model to more than two stimuli. In particular, explicitly modeling the cue in cued TOJ with TVA is an important objective of subsequent research[25].

Advantages of the protocol are the simplicity of the TOJ task which can utilize almost arbitrary stimuli, the thorough theoretical underpinning by TVA, and the Bayesian evaluation scheme. The TVA-based model is a large step forward from traditional model-free approaches. In the past, mostly generic psychometric functions have been fitted to TOJ data. Changes in their summary parameters PSS (point of subjective simultaneity) and DL (difference limen; a measure of discrimination performance) have been linked to attentional manipulations. At times, these parameters are over-interpreted. For example, it is frequently claimed that attention accelerates processing of the attended stimulus, whereas it could also be the case that the unattended stimulus is slowed[7]. In addition to this weakness, these parameters are rather indirect. They describe the

performance in the task and do not characterize the processes that produce it. The model-based analysis of TOJs improves on these drawbacks by providing meaningful parameters based on TVA.

Disclosures

The authors have nothing to disclose.

Acknowledgements

References

1. Posner, M. I. Orienting of attention. *Quarterly Journal of Experimental Psychology.*, **32** (1), 3-25 (1980).
2. Van der Heijden, A., Wolters, G., Groep, J., & Hagenaar, R. Single-letter recognition accuracy benefits from advance cuing of location. *Perception & Psychophysics.* **42** (5), 503-509 (1987).
3. Bundesen, C. A theory of visual attention. *Psychological Review.* **97** (4), 523-547 (1990).
4. Bublak, P., *et al.* Staged decline of visual processing capacity in mild cognitive impairment and Alzheimer's disease. *Neurobiology of Aging.* **32** (7), 1219-1230 (2011).
5. Petersen, A., Kyllingsbæk, S., & Bundesen, C. Measuring and modeling attentional dwell time. *Psychonomic Bulletin & Review.*, **19** (6), 1029-1046 (2012).
6. Vroomen, J., & Keetels, M. Perception of intersensory synchrony: A tutorial review. *Attention, Perception, & Psychophysics.*, **72** (4), 871-884 (2010).
7. Tünnermann, J., Petersen, A., & Scharlau, I. Does attention speed up processing? Decreases and increases of processing rates in visual prior entry. *Journal of Vision.*, **15** (3), 1-27 (2015).
8. Krüger, A., Tünnermann, J., & Scharlau, I. Fast and conspicuous? Quantifying salience with the Theory of Visual Attention. *Advances in Cognitive Psychology.*, **12** (1), 20-38 (2016).
9. Bundesen, C., & Habekost, T. *Principles of visual Attention: linking mind and brain.* Oxford University Press. Oxford, UK. (2008).
10. Plummer, M. JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. *In Proceedings of the 3rd international workshop on distributed statistical computing,.* **124**, 125 (2003).
11. Kruschke, J. K., & Vanpaemel, W. Bayesian estimation in hierarchical models. In Busemeyer, J., Townsend, J., Wang, Z. J., & Eidels A. (Eds.), *The Oxford handbook of computational and mathematical psychology.*, 279-299 (2015).
12. Mathôt, S., Schreij, D., & Theeuwes, J. OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods.*, **44** (2), 314-324 (2012).
13. Kruschke, J. K. *Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan.* 2nd ed. Academic Press. Boston, MA. (2015).
14. Rensink, R. A., O'Regan, J. K., & Clark, J. J. To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science.*, **8** (5), 368-373 (1997).
15. Tünnermann, J., Krüger, N., Mertsching, B., & Mustafa, W. Affordance estimation enhances artificial visual attention: Evidence from a change-blindness study. *Cognitive Computation.*, **7**, 525-538 (2015).
16. Shore, D. I., & Klein, R. M. The effects of scene inversion on change blindness. *The Journal of General Psychology.*, **127** (1), 27-43 (2000).
17. Scharlau, I., & Neumann, O. Temporal parameters and time course of perceptual latency priming. *Acta Psychologica.*, **113** (2), 185-203 (2003).
18. Schneider, K. A., & Bavelier, D. Components of visual prior entry. *Cognitive Psychology.*, **47** (4), 333-366 (2003).
19. Scharlau, I., & Neumann, O. Perceptual latency priming by masked and unmasked stimuli: Evidence for an attentional interpretation. *Psychological Research.*, **67** (3), 184-196 (2003).
20. Shore, D. I., Spence, C., & Klein, R. M. Visual prior entry. *Psychological Science.*, **12** (3), 205-212 (2001).
21. Alcalá-Quintana, R., & García-Pérez, M. A. Fitting model-based psychometric functions to simultaneity and temporal-order judgment data: MATLAB and R routines. *Behavior Research Methods.* **45** (4), 972-998 (2013).
22. Hoffman, M. D., & Gelman, A. The No-U-turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research.*, **15** (1), 1593-1623 (2014).
23. Lee, M. D., & Wagenmakers, E.-J. *Bayesian cognitive modeling: A practical course.* Cambridge University Press. Cambridge, UK. (2014).
24. Vangkilde, S., Bundesen, C., & Coull, J. T. Prompt but inefficient: Nicotine differentially modulates discrete components of attention. *Psychopharmacology.*, **218** (4), 667-680 (2011).
25. Tünnermann, J., & Scharlau, I., Peripheral visual cues: Their fate in processing and effects on attention and temporal-order perception. *Front. Psychol.* **7** (1442) (2016).