

Poking Left To Be Right? A Model-Based Analysis of Temporal Order Judged by Mice

Jan Tünnermann and Ingrid Scharlau

Cognitive Psychology, Paderborn University, Germany

ABSTRACT

KEYWORDS

temporal order judgments,
theory of visual attention,
mice

The theory of visual attention (TVA) provides a formal framework for the assessment of visual attention and related processes. Its center is a mathematical model of visual encoding processes and discretely defined components of attention. Building on this model, TVA offers quantitative and process-related explanations for a variety of phenomena in the domain of visual attention. Because the theory relies on very general assumptions which might hold true for other domains of sensory processing, we tested its possible explanatory value for tactile processing in mice. Reanalyzing published data of temporal-order judgments by mice, we show how a TVA-based analysis identifies the processes which drive observable behavior and that it comes to conclusions quite different from those of conventional analyses of temporal-order judgments. According to this analysis, despite the same overall capacity dedicated to the task, some mice assume attentional biases toward one side, possibly to optimize their overall performance. We suggest that TVA's concepts provide a powerful point of vantage to find explanations for observable behavior where conventional analysis easily leads to dead ends.

INTRODUCTION

The theory of visual attention (TVA) was originally proposed by Bundesen (1990) and has by now been adopted by many theorists (e.g., Hung, Driver, & Walsh, 2005; Logan, 1996; Schneider, 2013). It provides a common, formal framework for the assessment of visual attention and related processes. In particular, TVA offers a mathematical model of visual encoding processes and discretely defined components of attention, most prominently stimulus processing rates, visual short-term memory capacity, and a threshold time of perception. TVA thereby can provide a means to precisely talk and think about concepts and explanations concerning visual attention. It thus may help to overcome difficulties in this field which arise from the dangers of loose terminological conventions, informal definitions, and weak theories (cf. Taagepera, 2008).

So far, however, a major difficulty hinders TVA's application to many domains of interest in attention research. This difficulty arises from the fact that TVA mainly builds upon a single experimental paradigm, the *letter report task*: Most of the data from which TVA parameters are inferred come from experiments in which participants identify letters that are briefly presented on a screen and masked af-

ter a variable interval (it is not mandatory to use letters; any highly learned and sufficiently large set of categorizable stimuli would suffice, but letters serve the purpose most conveniently). More specifically, the growth of identification accuracy across increasing masking intervals provides the data from which the attention-related parameters are estimated. To facilitate research, specific ready-to-use versions of the experimental paradigm have been published (see, e.g., Vangkilde, Coull, & Bundesen, 2012). However, these ready-to-use versions, as well as the general approach, require that researchers can pose their questions in a way compatible with the letter-report paradigm. Clearly, not all interesting perception experiments can be transformed into a letter—or, speaking more generally, multiple-stimulus report experiment. What if participants cannot read letters, for instance, because they are preliterate children, or even literate children for whom the task of naming letters is too challenging or too boring? Comparable dif-

Corresponding author: Jan Tünnermann, Cognitive Psychology, Paderborn University, Germany.

Email: jan.tuennermann@uni-paderborn.de

faculties hamper studies with cognitively impaired persons or animals, such as lab mice, who, outside the TVA world, contribute substantially to the study of attention.

It is possible to adapt TVA models for particular experimental paradigms, as A. Petersen, Kyllingsbæk, and Bundesen (2012) demonstrated for the attentional dwell-time paradigm. Not all researchers, however, have the possibility to do the extensive mathematical and modeling work which is necessary for this. Thus, it would be a considerable benefit if more generally applicable experimental paradigms would become available for them.

During the past years, we have been working on one such method: estimating TVA parameters from temporal-order judgments (TOJs). The method has first been proposed by Tünnermann, Petersen, and Scharlau (2015) and subsequently been applied to different problems. These include quantification of visual salience (Krüger, Tünnermann, & Scharlau, 2016, 2017), the influence of peripheral visual cues on processing speed (Tünnermann & Scharlau, 2016), and the estimation of processing speed in natural images (Tünnermann, Krüger, & Scharlau, 2017). What is new in our approach is that virtually any stimulus material can be used, as long as two asynchronous stimuli can be presented. In the present paper, we demonstrate that this TVA-based TOJ method can even go a step farther: It enables the estimation of TVA's processing rates from lab animals.

Reporting which of two events appeared first is one of the simplest tasks involving stimuli distributed in space and time. Unless only very small temporal intervals are used, humans do not need much training or instruction for deciding which of two stimuli—the left or the right, the green or the red, or whatever is convenient—has been shown first. This makes the task suitable for very different participants, including children or cognitively impaired people. Partly for that reason, TOJs are popular in measuring the timing of information processing and the distribution of processing resources. To give a few examples of their widespread and long-ranging use, TOJs have been applied to study visual (e.g., Rutschmann, 1966), auditory (e.g., Hirsh, 1959), and tactile (e.g., Hirsh & Sherrick, 1961; Miyazaki, Yamamoto, Uchida, & Kitazawa, 2006) processing, the integration of these modalities (e.g., Zampini, Shore, & Spence, 2003), the influence of attention within (e.g., Scharlau, Ansong, & Horstmann, 2006) and across modalities (e.g., Zampini, Shore, & Spence, 2005), the control of attention (e.g.,

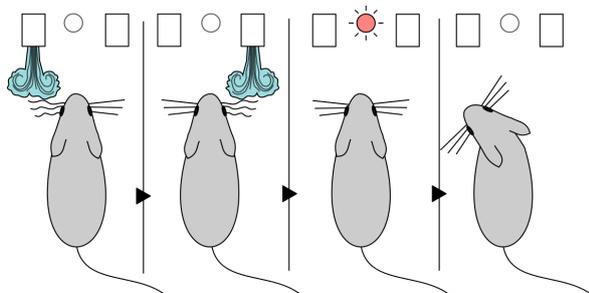


FIGURE 1.

Illustration of Wada, Moizumi, and Kitazawa's (2005) tactile mouse temporal-order judgment task.

Born, Kerzel, & Pratt, 2015), the processing of emotion (e.g., Lake, 2016; Schofield, Yousef, & Denson, 2017), pain (e.g., Vanden Bulcke, Crombez, Burnez, & Van Damme, 2015), causality (e.g., Bechlvianidis & Lagnado, 2016), and many more.

TOJs can also be gained from animals. Wada, Moizumi, and Kitazawa (2005) and Wada, Higo, Moizumi, and Kitazawa (2010) showed that mice can perform tactile TOJs, though after extended training. Short air puffs were delivered to the left and right whiskers of the mice (see Figure 1). After an additional go signal, the mice poked their heads into a response hole to indicate which air puff, left or right, had been delivered first or second, depending on the condition. Based on the idea that ordering of stimuli is vital for animals, the authors aimed at building a model for animal TOJs. Their data is the basis for the present study, and we will treat it in a TVA-based framework—that is, we will derive TVA parameters from their order responses. To motivate this approach, we will first show which problems conventional analysis runs into when building models of TOJs. We then explain how our TOJ-based extension of TVA solves these problems and why TVA's model might be a reasonable choice for modeling animal tactile data. Reanalysis of Wada et al.'s (2005) data finally shows which substantive conclusions can be drawn from this approach. We finally detail shortcomings and advantages of the approach. Our goal is to show that the TVA approach is promising not only within the field of visual attention, and that it is worth the effort to expand it beyond this field.

To fully appreciate the advantages of this approach, we first turn to the traditional method of analyzing TOJ data, which Wada et al. (2005), and many other studies, used. Temporal-order judgment ex-

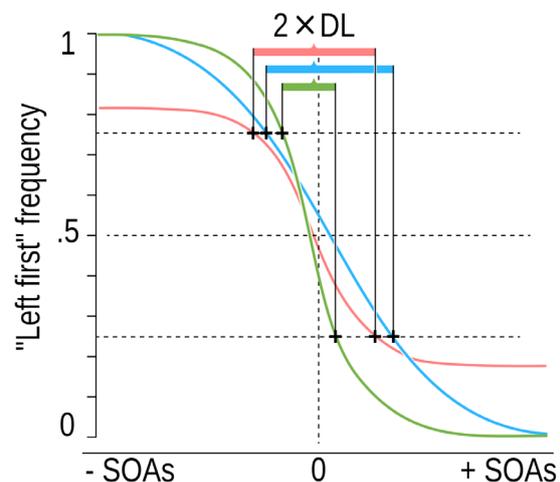


FIGURE 2.

Exemplary temporal-order judgment (TOJ) psychometric functions. Precision of the judgments is derived from their slope, indexed by the width of an interval between two percentiles. The difference limen (DL) marked in the illustration is half the distance between the .25-level and the .75-level intersection. The green and the blue curves, differing in precision, converge to zero and one, the red curve to different values, which happens in the presence of random errors. The point of subjective simultaneity (PSS) is the intersection with .5. SOA = Stimulus Onset Asynchrony.

periments typically yield *s*-shaped distributions of judgment frequency across temporal intervals (see Figure 2). Wada et al. fitted a psychometric function to the TOJ data and estimated its parameters, a common procedure both historically (Finney, 1971; Guilford, 1954; Woodworth & Schlosberg, 1961) and today (e.g., Vibell, Klinge, Zampini, Nobre, & Spence, 2017).

Old Methods Reaching Limits

Analyses of this type describe the *s*-shaped psychometric functions with two parameters, the point of subjective equality and a measure of temporal resolution (just notable difference, difference limen). In terms of the order judgment, the point of subjective equality is that objective interval at which the observers cannot decide which stimulus was first—that is, the interval at which the probability of each order judgment is .5. This is commonly interpreted as the point of subjective simultaneity (PSS, though see, for instance, Jaskowski, 1993; Weiß & Scharlau, 2011). In Figure 2, this point can be seen as the intersection of the psychometric functions with the horizontal line at .5. The difference limen is an index of the mean slope in the inner quartile of the psychometric function and is typically understood as a measure of judgment precision—if the slope is steep (the green function in Figure 2), the observer can discriminate well between the two events; if it is shallow, the observer's discrimination is poor (the blue function in Figure 2; markings indicate the difference limen).

In the mouse data, the asymptotes of the psychometric functions reached neither zero nor one (similar to the red function in Figure 2). That is, even with the largest temporal separations used in the experiment (750 ms), the mice still made a substantial amount of errors—around 30%. To account for this, Wada et al. (2005) included parameters for the upper and lower asymptote of performance in their model. This made it necessary to adapt the calculation of the two parameters that are of central interest: Their PSS lies halfway between the asymptotes, and the width parameter captures the temporal resolution at 84% of the asymptote, serving the same purpose as the difference limen explained above.

Using this conventional analysis of whiskers-based TOJs, Wada et al. (2005) found that, even after extensive training, mice performed at weaker temporal resolution than humans in comparable tasks with air puffs to fingers of the right and left hand (difference limen of 160 compared to 124 ms). Interindividual variability was, however, high, because at least one mouse seems to have performed at better resolution than the humans. More interestingly, with the mice, the PSS was strongly shifted by 133 ms, indicating that they had a tendency towards judging “left first” at short intervals or perceived tactile signals from the left faster than those from the right.

To explain this shift, Wada et al. (2005) suggested lateralization of TOJs to the right hemisphere, where the majority of left-whisker signals converge. Part of the shift could then be explained by interhemispheric conduction delays of a few, approximately 10 ms (Nicholls, 1996). Shuler, Krupa, and Nicolelis (2001) discuss interactions between signals in the primary sensory cortex as a further factor that could cause delays by suppressing responses to contralateral whiskers after

stimulation at ipsilateral ones. Wada et al. speculate that this suppression could be especially pronounced at small temporal intervals, but they leave its strength unquantified. Both explanations are tentative, and it is unlikely that they can explain the large, 133 ms displacement of the PSS, as the authors concede (p. 174).

With respect to the weaker performance reflected in the asymptotes, Wada et al. (2005) suggest that it might result from underdeveloped—compared to humans—cognitive processes involved in task execution or, alternatively, different signal convergence in the mouse primary sensory cortex. However, the data indicate that zero and one would not be reached if even longer intervals were included. Instead, it appears that the curves converge to values different from zero and one—values of .29 and .73. This is an intriguing finding: The mice seem to have a high base rate of errors. How is this to be explained?

The study of Wada et al. (2005) thus raises at least two questions: Does the PSS shift really indicate interhemispheric conduction delays and early interaction of sensory signals? What causes the high base rate of errors in the TOJ task? Conventional TOJ analysis cannot answer these questions because PSS and difference limen have restricted informative value or explanatory power. In the first place, they are summary parameters that describe the observer's performance but provide only a weak link to the internal processes that generate the judgment (see Tünnermann, 2016). Despite their ostensible connection to psychological processing (simultaneity perception, precision, bias), they are estimated by a model-free procedure—that is, without a formal model that would describe how cognitive processes are linked to judgment functions. Therefore, neither the functions (a cumulative Gaussian) nor the parameters themselves have a deeper meaning with respect to the processes which result in the temporal-order perception.

Furthermore, estimation of the descriptive parameters of conventional TOJ analysis presupposes a certain data structure. If the observed judgment distributions are symmetric and converge to zero and one, the shift of subjective simultaneity away from zero can be understood as a relative difference in the encoding latencies of the two targets. However, if the judgment distributions are not symmetrical, do not converge to zero or one (as in the mice data), are shifted vertically, contain central plateaus, or typically feature a combination of these distortions, the summary performance parameters lose even more of their sparse meaning. Such deviations from symmetry have been observed and are even expected on theoretical grounds (Allan, 1975; Sternberg & Knoll, 1973). They can originate from a variety of relevant and interesting causes: differential processing speeds, attentional lapses, response biases, or interactions with decision functions (see García-Pérez & Alcalá-Quintana, 2012). Importantly, the standard analysis pushes such influences into the latency difference (shifted PSS) and the judgment precision (difference limen), rendering a possibly distorted picture in these parameters' interpretations. Wada et al. (2005) evade some of these problems by amending a cumulative Gaussian model with error parameters. This deals with the asymmetry at a formal level. However, the parameters of this model describe the observed data, without describing the underlying processes, and other descriptions are well possible.

METHODS

The Theory of Visual Attention as a Model of Visual-Attentional Processing

The computational model we use to reanalyze Wada et al.'s (2005) data is based on Bundesen's (1990) TVA. The remainder of this section provides a brief introduction to the concepts of TVA, their mathematical formulation, and how they are applied to create a model of TOJs, followed by a justification of the transfer of TVA to domains outside human visual attention.

Theory of visual attention is a race model. This means that stimuli take part in a (metaphorical) race in which each stimulus progresses independently of the others. Stimuli race for a slot in visual short-term memory (VSTM), which only provides a limited number of (metaphorical) slots. Only stimuli which reach VSTM before it is filled up are represented at a level at which they become relevant for behavior. Arrival in VSTM is a precondition for reporting stimulus features (Bundesen, 1990). The order in which stimuli arrive in VSTM determines perceived order (Tünnermann, 2016; a similar idea is spelled out by Alcalá-Quintana & García-Pérez, 2013).

Table 1 provides an overview of TVA's main parameters, their symbols, meanings, and units. It also indicates whether the parameters can be estimated with the TOJ task.

The rates at which stimuli are processed are represented by $\nu(x, i)$. These are the rates with which stimuli x are encoded as members of the category i . As will become clear below, we can often neglect the i in the term and call the rates ν_x , which can be thought of as the speed with which stimulus x is processed. This parameter is central for explaining TOJs: The faster a stimulus is processed, the earlier it enters VSTM and thus will be reportable. A comparable stimulus which is processed more slowly will be represented later and thus has a disadvantage in temporal-order perception.

The overall processing rate, or capacity, C represents the available processing resources for the task. It is the sum of the rates at which

individual stimuli are processed. In the TOJ task, C is the sum of the two ν values for the two stimuli.

Parameter w_x^* is the relative attentional weight of stimulus x . It determines how much of the available resources C are spent on stimulus x , in effect determining its rate ν_x . The leftward bias Wada et al. (2005) speculated about might be a higher relative w_x^* for a stimulus on the left compared to one on the right side, which would show up in a higher ν_x . Again, this is a parameter of high relevance for explaining temporal-order perception.

Parameter K is the VSTM limit, the number of elements that can be represented there. It is important for, and can be measured with, the whole report task. The VSTM capacity needs to be taken into account in situations with more elements than the typical VSTM limit. With two targets, as in the present study, the capacity likely plays no role.

The TVA parameter t_0 is the maximum ineffective exposure duration. If a stimulus is shown for a duration shorter than t_0 , it has no chance of being encoded at all. Typically, t_0 is considered equal for the equal stimuli in a TOJ, canceling out in the VSTM arrival times (cf. Tünnermann, 2016, pp. 75–77), and it can therefore be ignored in the present study.

With these parameters and their relations in mind, the reader can follow the analysis conducted in this study and may skip to the results and discussion without loss of continuity. For readers interested in TVA's encoding model and how it can explain TOJs, we present its mathematical details and further low-level parameters (listed in Table 2) below. Before that, we will turn to the question whether the TVA framework can be meaningfully applied to other modalities.

The Theory of Visual Attention as a Model for Tactile Processing

TVA is a theory of visual attention. For the time being, we ignore this fact and use it as TWA, a *theory of whiskers' attention*. This is maybe not too far-fetched. The only assumption strictly required for the application of TVA's mathematical modeling is that the categorization latencies are exponentially distributed. Such distributions were already used to model modalities other than vision, for example, auditory signals (Alcalá-Quintana & García-Pérez, 2013), and they are common as arrival time distributions (Colonius & Diederich, 2011; Heath, 1984). Alcalá-Quintana and García-Pérez (2013) point out that exponential distributions have the crucial features of being probabilistic, but de-

TABLE 1.
Important High-Level Parameters of the Theory of Visual Attention (TVA) Which Are Directly Estimated From Data

Parameter	Meaning	Unit	Task
$\nu(x, i)$	Rate at which stimulus x is encoded in category i	Hz	WR, TOJ
w_x^*	Relative attentional weight of stimulus x		WR, TOJ
C	Overall processing rate	Hz	WR, TOJ
K	VSTM storage capacity	Items	WR
t_0	Maximum ineffective exposure duration	ms	WR

Note: These are fundamental TVA parameters that can be measured with Whole Reports (WR) or Temporal-Order Judgment (TOJ) Tasks. VSTM = Visual Short-Term Memory.

TABLE 2.
Low-Level Parameters of the Theory of Visual Attention (TVA) Which Can Be Inferred From High-Level Parameters and the Experimental Design

Parameter	Meaning	Unit
$\eta(x, i)$	(Objective) sensory evidence for x belonging to category k	(Hz)
π_j	(Subjective) pertinence of sensory category j	
β_i	(Subjective) bias for making reports of category i	

spite being probabilistic, no arrivals can occur before the actual onset of the stimulus.

As has been shown above, TVA dissolves attention into several different parameters (e.g., ν , w , C , K , t_0). All these parameters contribute to attention, but none in itself is attention. From this follows another argument why it is justified to understand TVA as an example of a generic model which captures processing in several modalities: A model-based approach is instructive even if the details of the order judgment function in tactile mouse TOJs may differ from those in human visual attention. The way in which TVA derives an order judgment function from low-level visual processing and attention components can act as a template for how specialists in rodent tactile processing could derive a more specific model—that is, if they disagree with some detail of the model proposed here.

There are at least two more substantive reasons why the attempt to transfer or expand TVA analysis to tactile processing in mice might be useful in the long run. Firstly, tactile processing serves similar functions in rodents as vision does in humans. Rats and mice use their snout whiskers to build spatial representations of their environment, to locate objects, and to explore textures (C. C. Petersen, 2007). Rodents use whiskers actively to establish somatotopic maps in the primary somatosensory cortex (Mitchinson, 2016). Therefore there are, on an abstract level, parallels between the whisker system and the human visual system, which forms similar maps via similar active sensing. Secondly, synchrony perception in rats is consistent enough with those of humans and other species to lead Schormans et al. (2017) to suggest that rats provide a proper general model for audiovisual temporal synchrony at both the neuronal and perceptual level.

Therefore, we feel that it is justified to skip the visual attention-related interpretation and naming of the parameters and test the use of TVA as TWA.

Theory of Visual Attention Modeling in More Detail

According to TVA, the encoding duration of a single stimulus is described by a shifted exponential distribution. The probability that a stimulus x is encoded until time t (relative to the presentation) is

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

where ν_x is the rate at which stimuli of type x are encoded. Even a single stimulus competes for multiple possible categorizations of the form “stimulus x belongs to category i ”. Therefore, the rate ν_x is the sum of the individual categorization rates $\nu(x, i)$:

$$\nu_x = \sum_{i \in R} \nu(x, i). \quad (2)$$

These rates can be further dissected:

$$\nu(x, i) = \eta(x, i) \beta_i w_x^* \quad (3)$$

where $\eta(x, i)$ is the sensory evidence that stimulus x is a member of report category i . The component β_i is a bias for the tendency to report categorizations i independently of the stimulation. In the case of Wada et al.'s (2005) study, for instance, the training increased the mice's β for categorizing air puffs as targets. The β_i for other categories are effectively zero for the task. We thus can use ν_x to mean “the rate at which stimulus x is encoded in the relevant category”.

The attentional weights w_x^* reflect the distribution of attention and result from normalizing absolute weights by the sum of all stimulus weights in the visual field S , $w_x^* = \frac{w_x}{\sum_{z \in S} w_z}$. These weights consist of further low-level parameters:

$$w_y = \sum_{j \in R} \eta(y, j) \pi_j. \quad (4)$$

Above, in Equation 3, η reflected the evidence that a stimulus belongs to a report category. Here in Equation 4, η reflects a different kind of evidence: $\eta(y, j)$ is the sensory evidence that stimulus y is a member of filter category j . The term π_j reflects the pertinence of a filter category. In Wada et al.'s (2005) mouse TOJs, *left* and *right* can be understood as such filter categories, and one could be more relevant (higher π value) than the other. However, from the experimenter's point of view, the pertinence value π_{left} and π_{right} should be equal because the mice were not trained to prefer one of the categories.

Now we have reduced TVA's encoding equations to components that describe the stimulation and task set. How are these components experimentally assessed? The low-level components, such as sensory evidence and pertinence, cannot be directly estimated and are thus kept constant in the experiments. What is estimated are the processing rates ν of different stimuli. These are derived from TOJ data that are present as the frequency of judgments, in the mouse experiments the frequency of “left air puff occurred first” responses over various stimulus onset asynchronies (SOAs; *Left* is an arbitrary choice so that the following could be equally put in terms of “right first” judgments). The SOAs are the intervals between the left air puff L and the right air puff R . At negative SOAs, L is presented first, at positive SOAs, R leads (again an arbitrary choice). The TVA processing rates can be estimated from such data with the following equations (cf. Tünnermann et al., 2017):

The probability $P_{L^{1st}}$ of encoding the left air puff L before right air puff R is a function of their processing rates, ν_L and ν_R and the SOA. For negative SOAs:

$$P_{L^{1st}}(\nu_L, \nu_R, SOA) = 1 - e^{-\nu_L|SOA|} + e^{-\nu_L|SOA|} \left(\frac{\nu_L}{\nu_L + \nu_R} \right) \text{ for } SOA < 0, \quad (5)$$

where (according to Equation 1) $1 - e^{-\nu_L|SOA|}$ is the probability that the air puff L is encoded before R 's encoding even starts. With probability $e^{-\nu_L|SOA|}$, L was not yet encoded when R is presented. Therefore, they race against each other with this probability. Following Luce's choice axiom (e.g., see Luce, 1977), L then wins the race with probability $\nu_L/(\nu_L + \nu_R)$.

If the SOA is positive instead, $P_{L^{1st}}$ is calculated as

$$P_{L^{1st}}(\nu_L, \nu_R, SOA) = e^{-\nu_R |SOA|} \left(\frac{\nu_L}{\nu_L + \nu_R} \right) \text{ for } SOA \geq 0, \quad (6)$$

Probability $e^{-\nu_R |SOA|}$ models that air puff R was not encoded before L is presented. With this probability, again both stimuli race each other, leading to “left first” judgments with $\frac{\nu_L}{\nu_L + \nu_R}$.

The model described so far is based only on TVA’s equations (and concepts) of encoding the stimuli into an early short-term storage at which the arrival order judgment is assumed to take place. For the present use in mice, the model was extended by a parameter ϵ for response errors, for instance, errors caused by lapses or in response execution. This was done because, as already mentioned, Wada et al. (2005) observed that the psychometric functions in their analysis did not reach their usual asymptotes. Alcalá-Quintana and García-Pérez (2013) suggested including error parameters in such cases, which we implemented analogously to their description. The parameter ϵ captures a possible mismatch between the internal order perceptions and the observed judgments caused, for instance, by response execution errors:

$$P_{L^{1st}}^{lapses}(\nu_L, \nu_R, \epsilon, SOA) = (1 - \epsilon) \cdot P_{L^{1st}} + \epsilon \cdot (1 - P_{L^{1st}}). \quad (7)$$

The Bayesian Estimation Procedure

To fit the experimental data with the model described above, we employed a Bayesian hierarchical approach (see Kruschke & Vanpaemel, 2015), which is highly suitable for inferences based on a formal model. It yields subject-level and group-level estimates of the individual processing rates ν_L and ν_R and the error parameter ϵ . At its core, the model is based on the equations described above that link the observations to the parameters.

In the Bayesian statistical framework, prior beliefs about the parameter values are updated using the data and the model. Because we do not have specific prior information about the performance of mice in TOJs, we decided to use prior distributions which are only weakly informed. The group-level priors were parametrized as overall processing rate C , attentional weight w_L^* , and error parameter ϵ . Note that $C \cdot w_L^* = \nu_L$ and $C \cdot (1 - w_L^*) = \nu_R$. The following group-level distributions were used: The prior for the group mean of C , μ_C , is uniformly distributed within $[0 \dots 100 \text{ Hz}]$ and its standard deviation, σ_C , uniformly in $[0 \dots 50 \text{ Hz}]$. This means individual rates could be anywhere from almost zero up to the higher range of human visual processing rates, which typically are around 60 Hz (Tünnermann, 2016, p. 154). From Wada et al.’s (2005) PSS-based results we know that the performance of mice is much weaker, hence, these priors are quite conservative choices. The attentional weight priors were $\mu_{w_L^*}$ (mean) uniformly within $[0.01 \dots 0.99]$ and $\sigma_{w_L^*}$ (standard deviation) in $[0 \dots 0.2]$, which again allows for a large range of values with up to 99% of processing resources focused on one side. Even strong attention manipulations typically do not exceed 80%, and, given the fact that with no attention manipulation in

the present experiment, a weight of $w_L^* = 0.5$ was expected, these priors again reflect a very conservative choice. Finally, the priors on lapse error parameters μ_ϵ (mean) and σ_ϵ (standard deviation) were set to uniform distributions over $[0 \dots 0.5]$, which includes everything from no response errors to only response errors. All subject-level parameters come from normal distributions with the abovementioned means and SDs from the group level (cf. Tünnermann et al., 2017). The following provides an overview of the probabilistic model.

PRIORS

The priors on the subject level were defined as follows:

$$\begin{aligned} C &\sim \text{Normal}(\mu_C, \sigma_C) \\ w_L^* &\sim \text{Normal}(\mu_{w_L^*}, \sigma_{w_L^*})T(0, 1) \\ \epsilon &\sim \text{Normal}(\mu_\epsilon, \sigma_\epsilon)T(0, 1) \end{aligned}$$

GROUP-LEVEL PRIORS

The parameters of the priors above were themselves drawn from “hyper-prior” distributions, which represent the group level. That is, the prior decisions described in the text above were implemented here as group-level distributions.

$$\begin{aligned} \mu_C &\sim \text{Uniform}(0, 100), & \sigma_C &\sim \text{Uniform}(0, 50) \\ \mu_{w_L^*} &\sim \text{Uniform}(0.01, 0.99) & \sigma_{w_L^*} &\sim \text{Uniform}(0, 0.2), \\ \mu_\epsilon &\sim \text{Uniform}(0, 0.5), & \sigma_\epsilon &\sim \text{Uniform}(0, 0.5) \end{aligned}$$

DETERMINISTIC TRANSFORMATIONS

The following deterministic transformations were implemented to generate posterior distributions for the parameters of interest. How ν_L and ν_R were calculated from the probabilistic variables follows from TVA.

$$\begin{aligned} \nu_L &= C \cdot w_L^* \\ \nu_R &= C \cdot (1 - w_L^*) \\ \mu_{\nu_L} &= \text{mean}(\nu_L) \\ \mu_{\nu_R} &= \text{mean}(\nu_R) \\ \mu_\epsilon &= \text{mean}(\epsilon) \\ \nu_\Delta &= \nu_L - \nu_R \\ \mu_{\nu_\Delta} &= \text{mean}(\nu_\Delta) \end{aligned}$$

LIKELIHOOD

The likelihood which links the model to the observed “left first” counts y is based on binomial distributions at each SOA. The number of repetitions N is part of the data and the success probability p [first parameter of $\text{Binomial}(p, N)$] was calculated via the TVA-based TOJ model as, $P_{L^{1st}}^{lapses}(\nu_L, \nu_R, \epsilon, SOA)$ (see Equations 1 to 7):

$$y \sim \text{Binomial}(P_{L^{1st}}^{lapses}(\nu_L, \nu_R, \epsilon, SOA), N)$$

A TVA-Based Analysis of Mouse Data TOJs

In the following, we will test whether a TOJ analysis based on TVA which considers the encoding processes can provide insights into the causes of the mouse TOJ data pattern. Wada et al. (2005) kindly provided their raw data so that we could fit it with the TVA-based model.

RESULTS

The result of the estimation (using MCMC sampling via JAGS, see Plummer, 2003) is a set of posterior probability distributions of parameter values, as shown in Figure 4. Their peaks can be taken as the best guesses for a single parameter value. Furthermore, the shape of the distribution is informative: The 95% highest density interval (HDI) is the parameter value range with the 95% most likely values. The smaller the interval, the more certain is the estimate, which is often important information. This interval furthermore allows making a decision comparable to conventional statistical testing. For instance, if the posterior distribution of the processing rate difference ν_{Δ} ($= \nu_L - \nu_R$) does not include zero, the rates can be considered different. A more in-depth explanation of Bayesian statistics, especially of the hierarchical version which we apply here, can be found in Kruschke and Vanpaemel (2015).

On the group level (see Figure 3), the TVA-based estimates showed a processing speed advantage for the left stimulus ($\nu_L = 7.15$ Hz) compared to the right stimulus ($\nu_R = 4.47$ Hz). The lower boundary of the 95% HDI on $\mu_{\nu_{\Delta}}$ (group mean of $\nu_L - \nu_R$) is at 0.35 Hz—that is, the HDI excluded zero. This agrees with Wada et al.'s (2005) interpretation that mice have a leftward bias when performing TOJs. In TVA terms, they process tactile stimuli on the left faster than those on the right. The difference is approximately 3 Hz. In absolute values, the mice process the left-side tactile stimuli almost at double speed compared to the right-side ones.

These values can be converted into estimates of the signal arrival times. The latter are calculated as the expected values of the exponential distribution: $E_x = 1/\nu_x$. The difference in arrival times turned out to be $E_R - E_L = (1/4.47 \text{ Hz}) - (1/7.14 \text{ Hz}) = 84$ ms. This difference is smaller than Wada et al.'s (2005) PSS-based estimate of 133 ms, but still large compared to the few milliseconds expected on the basis of an interhemispheric conduction delay or sensory interaction. On the subject level, it would be even larger than 400 ms for those mice that showed a strong leftward bias (see below and Figure 4).

The low discrimination performance was reflected in the TVA estimates, too (see Figure 3). The overall capacity of the system C was estimated as the sum of the two ν values. The group-level estimate of $\mu_C = 11.16$ Hz corresponded to a performance roughly five to eight times

lower than human performance in simple visual TOJs (e.g., see Krüger et al., 2016; Tünnermann et al., 2017). The group-level error parameter μ_{ϵ} was estimated at 0.18.

The subject-level analysis showed two interesting features (see Figure 4). Strikingly, the overall processing rate C of approximately 12 Hz is highly similar in all mice (except for Mouse 7, which shows poor performance indicated by peaks at very low processing rates in the posterior density). This is true even when comparing mice that showed a strong leftward bias of the processing rate to mice with no bias (e.g., Mouse 2 to Mouse 5). This similarity in overall resources would not have shown up in conventional TOJ analysis.

The second interesting feature was that the strong leftward bias, indicated by a much higher estimate for the processing rate of the left stimulus, was only present in Mice 2, 3, and 6. According to our analysis, these three mice drive the bias observed on the group-level. By contrast, Wada et al. (2005) reported a leftward bias varying between PSS values of 112 to 200 ms.

We already mentioned that most of these biases are too large to be accounted for by the interhemispheric conduction delay and sensory interactions which Wada et al. (2005) referred to. How, then, can the difference between the results be explained, and can the TVA parameters provide a clue to an explanation?

Taken together, the TVA-based observations indeed provided a basis for a possible explanation of the leftward bias. The overall amount of available processing resources, captured in TVA parameter C , was similar in all mice. Some mice, however, may have learned a strategy in which they devoted as many resources as possible to the left stimulus. In TVA terms, they adjusted their pertinence for stimuli on the left side π_{left} in favor of that location. Normally, for instance, in typical TOJs in humans, such a strategy is disadvantageous, and in principle, it is disadvantageous for the mice too. However, they already performed very poorly, which rendered the disadvantage rather unimportant. Consider Mouse 3 (see Figure 4) that employed the leftward-bias strategy.

It responds correctly in only 57% of the trials. If the mouse would have distributed its processing resources equally (the optimal strategy), the performance would have only increased to 61%. If now the biasing strategy led to a reduction of lapses or response errors, it would have even been beneficial. Reducing ϵ to 0.05 for Mouse 3 would have increased its performance to 77% correct. Are the mice poking left to

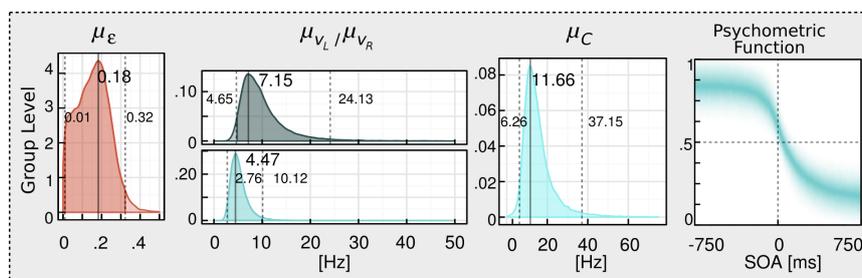


FIGURE 3.

Group-level attention parameter estimates for mouse temporal-order judgments, data taken from Wada et al. (2005). μ_{ϵ} = lapse-error parameter; μ_{ν_L} = left-side processing rate (dark shaded curves); μ_{ν_R} = right-side processing rate (bright shaded curves); μ_C = overall processing rate. The psychometric function is a posterior-predictive simulation at fine-grained stimulus onset.

be right? According to this explanation, mice reduce their lapse errors by focusing on the left side, which effectively maximizes their correct responses. At least two of the three biased mice (2 and 6) showed comparably small error parameters, suggesting that this strategy may indeed have decreased lapses or response errors.

This attentional bias strategy may explain the large size of the PSS shift, but not why the bias is toward the left side. It must be kept in mind, however, that, as our analysis revealed, only three mice in the seven mice sample drive this effect. They could have acquired the left-

ward bias by chance. The above explanation, for one, works equally well with a rightward bias.

Alternatively, the small interhemispheric conduction delay discussed by Wada et al. (2005), which leads to a small leftward bias, could have indeed been the reason for the biased performance. This small bias could have tipped the system away from the neutral resource distribution, a local maximum, toward the leftward bias strategy, which is another local maximum because of the reduction of response errors. This theory could be tested by experimentally inducing an opposite

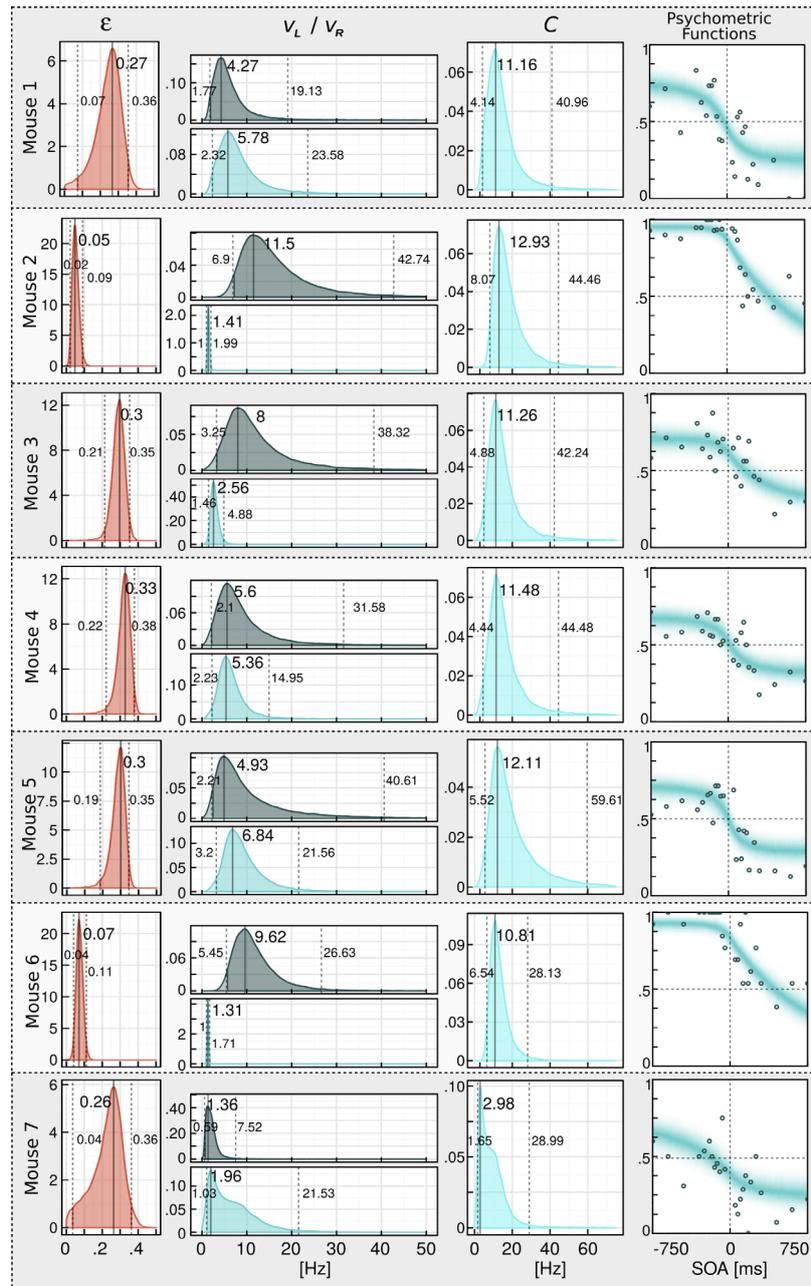


FIGURE 4.

Subject-level attention parameter estimates for mice temporal-order judgments (TOJs) from Wada et al. (2005). ϵ = lapse-error parameter; ν_L = left-side processing rate (dark shaded curves); ν_R = right-side processing rate (bright shaded curves); C = overall processing rate (sum of ν_L and ν_R). Psychometric functions are posterior predictive simulations at fine-grained stimulus onset asynchronies (SOAs) with 100 repetitions.

PSS shift that would be strong enough to counteract the interhemispheric delay and tip the system to the local maximum associated with a strong rightward bias strategy. In general, according to this explanation, gradual attention manipulations of different strengths will not lead to gradual changes in processing rates (or the PSS). Instead, the system would be pushed into one of the strategic local maxima.

Of course, this explanation is post hoc and would have to be tested on new data. Note that the point we want to make is not that this explanation is the correct one; what we want to show is that the TVA-based TOJ paradigm provides a means to shed light on the processing basis of perception and attention phenomena which may not be uncovered using more widespread analyses.

CONCLUSION OF THE TVA-BASED ANALYSIS

To sum up, the above reanalysis of data by Wada et al. (2005) shows that TVA or, more precisely, the TVA-based TOJ model can be applied to data gained from mice, data, moreover, that were difficult to understand by conventional TOJ analysis. Most importantly, the TVA-based explanation shows that all mice process the stimuli with the same overall capacity. Contrary to what Wada et al. report, we found the leftward bias only in some mice, possibly because in our analysis, the process-based parameters, whose uncertainties are respected in the Bayesian framework, better capture the individual differences in processing speed, lateral biases, and response execution errors. We therefore speculated that it is the result of a learned strategy in which the mice, despite their overall weak temporal discrimination, tried to avoid mistakes and were tipped away from a neutral distribution of their processing resources to both stimuli by the interhemispheric conduction delay. Whether this explanation is indeed correct cannot be decided on the basis of the present data, but we sketched how it might be tested.

Is the processing speed difference an attentional effect, or could a response bias have caused it? Importantly, TVA is a theory that somewhat explains away the traditional unitary concept of attention and decomposes it into discrete components of visual processing. Some of them are perceptual and control biases. According to the model we present, the difference in processing speeds originated from TVA's pertinence parameter π . If one wants to put a label on the effect, "attentional" would be appropriate because TVA's π is the component observers adjust to guide their attention to elements with a particular feature or to a certain location (Nordfang, Staugaard, & Bundesen, 2017).

While labeling the effect is a matter of terminology, it remains possible that the mouse behavior originated from a decision-level response bias that is not included in the model and would have remained undetected. The mice could have processed the stimuli at equal (or at any) rates and decided which side to poke independently of the stimulation. Note that Wada et al. (2005) carefully designed their training so that the mice did not acquire a response bias, so we consider this explanation very unlikely. However, process-based TOJ models can be extended to cover such biases. To this end, Alcalá-Quintana and García-Pérez

(2013) added two further parameters to their model. They lead to central plateaus in the psychometric functions at elevations different from 0.5. However, distinguishing such decision-level effects from alternative processing speed effects requires vastly more data than available from the mouse experiment (see Tünnermann, 2016, pp. 78–80). Hence, we prefer the simpler model and follow Wada et al.'s (2005) assumption that the effect originates from latency differences.

A full explanation of the mice data is not the focus of the present paper. What we wanted to show is that the TVA-based TOJ method can be applied to the study of animal attention and allows conclusions or at least questions that tap into the very processes which lead to the observed behavior. That is, although the above interpretations are speculative, as a proof of concept, the analysis shows that TVA-based TOJ analysis can be applied to data from animal experiments. Because mice can perform other simple tasks that assess visual attention (e.g., see Bushnell & Strupp, 2009), it should be possible to train them to perform visual TOJs, similarly to how Wada et al. (2005) trained mice for whisker-based tactile TOJs. Then, their visual attention parameters could be directly compared to those obtained for humans or other animals.

Aside from this central goal, the present paper demonstrates how substantive process-based models can be combined with hierarchical Bayesian analysis. This type of analysis allows for group-level and subject-level inferences in one coherent framework. Subject-level estimates inform each other via shrinkage toward the group-level estimates (e.g., Kruschke & Vanpaemel, 2015). This enables consistent subject-level fits even for such highly noisy data. In the present case, the subject level suggested that some animals used special response strategies. Such patterns would have been obscured when fitting data averaged over subjects, as is often done.

We hope the approach presented here motivates researchers of animal visual behavior to consider TOJs as an informative task. We believe they are powerful tools, especially when the data is analyzed with substantive process-based models—that is, models which describe the processes that lead to the TOJ curves instead of those which merely describe the shape of the curves or the judgment performance. Whether TVA provides a tenable basis for this, as assumed in the present study, remains to be seen. If not, it may still serve as a blueprint for the general characteristics of such models. There is some recent effort to model a typical animal paradigm of rodent attention with TVA. Fitzpatrick et al. (2017) modeled the five-choice continuous performance task with TVA and presented estimates of visual processing parameters in mice. In the future, it may be fruitful to compare results from this domain with TOJ-based analysis. Comparing parameters estimated from TOJs and the five-choice continuous performance task will help to understand how well current TVA captures rodent sensory processing.

Finally, we want to draw attention to the fact that TVA is not the only model-based approach in the study of temporal-order perception or attention. Alternatives are independent-channels models with specified encoding processes or drift diffusion models. These models, too, infer process-related parameters from behavior, and thus can be directly meaningful for psychological understanding.

The independent-channels model was originally proposed by Sternberg and Knoll (1973) and was used to test statistical assumptions about processing and the decision mechanism. In its original form, it lacks a mechanistic model of the underlying processes. Some researchers extended this, as for instance, García-Pérez and Alcalá-Quintana (2012). The TVA-based model is such an extension, too. It should be noted that in the TOJ context, García-Pérez and Alcalá-Quintana's model and the TVA-based model we use are, at their cores, mathematically equal. This is the case because some components of the former were not relevant here and lapse parameters were added to the latter. Consequently, there is no point in comparing the two versions (in other contexts, they can be distinguished, but very precise data is required, see, e.g., Tünnermann, 2016, pp. 78–91; Tünnermann & Scharlau, 2018).

Drift diffusion models can describe TOJs, including the encoding processes, on a similar level of abstraction as TVA, but with different underlying concepts (e.g., Miller & Schwarz, 2006). Intriguingly, again, if applied to TOJs, their parameters can be interpreted in a very similar manner to those of the TVA-based TOJ model. For instance, if the model is set up in a certain way, the drift rate is similar to TVA's processing rate, not in terms of the absolute values, but in terms of parameter meaning. Therefore, predictions of the drift diffusion model and the TVA-based approach would be very similar. Again, highly accurate data would be required to distinguish the models—not only more accurate than the mouse data at hand, but also more accurate than data from typical experiments with humans (see, e.g., Tünnermann, 2016, pp. 125–129). While drift diffusion and independent-channels models may be able to fit the mouse data similarly well, they lack concepts deeper than the processing rates. By contrast, TVA offers a very fine-grained description of what makes up the processing rates on a conceptual level (which we have explained in the Methods section) and how they are implemented on a neural level (Bundesen, Habekost, & Kyllingsbæk, 2005). Moreover, as we explained in the Introduction, TVA parameters and mechanisms carry meaning beyond a specific experimental paradigm, beyond specific data models, beyond specific subfields of Cognitive Psychology, beyond a specific discipline, and perhaps—this is the idea behind this study—beyond specific modalities and species.

We cannot provide anything close to final conclusions on the value and appropriateness of these models, let alone a formal comparison. Neither was this the goal of the present paper. Differentiating and deciding between these models would have required both very precise data and a breadth of different findings related to the processes specified in the models. This is what further studies could aim at.

ACKNOWLEDGMENTS

We are very grateful to Makato Wada for sharing the data of the mouse TOJ study.

AUTHORNOTES

This paper is based in part upon an analysis conducted for Jan Tünnermann's doctoral dissertation at Paderborn University, Germany.

REFERENCES

- Alcalá-Quintana, R., & García-Pérez, M. A. (2013). Fitting model-based psychometric functions to simultaneity and temporal-order judgment data: MATLAB and R routines. *Behavior Research Methods, 45*, 972–998. doi: 10.3758/s13428-013-0325-2
- Allan, L. G. (1975). Temporal order psychometric functions based on confidence-rating data. *Perception & Psychophysics, 18*, 369–372. doi: 10.3758/BF03211214
- Bechlvianidis, C., & Lagnado, D. A. (2016). Time reordered: Causal perception guides the interpretation of temporal order. *Cognition, 146*, 58–66. doi: 10.1016/j.cognition.2015.09.001
- Born, S., Kerzel, D., & Pratt, J. (2015). Contingent capture effects in temporal order judgments. *Journal of Experimental Psychology: Human Perception and Performance, 41*, 995–1006. doi: 10.1037/xhp0000058
- Bundesen, C. (1990). A theory of visual attention. *Psychological Review, 97*, 523–547. doi: 10.1037/0033-295X.97.4.523
- Bundesen, C., Habekost, T., & Kyllingsbæk, S. (2005). A neural theory of visual attention: Bridging cognition and neurophysiology. *Psychological Review, 112*, 291–328. doi: 10.1037/0033-295X.112.2.291
- Bushnell, P. J., & Strupp, B. J. (2009). Assessing attention in rodents. In J. J. Buccafusco (Ed.), *Methods of behavior analysis in neuroscience* (2nd ed., pp. 119–145). Boca Raton, FL: CRC Press.
- Colonius, H., & Diederich, A. (2011). Computing an optimal time window of audiovisual integration in focused attention tasks: illustrated by studies on effect of age and prior knowledge. *Experimental Brain Research, 212*, 327–337. doi: 10.1007/s00221-011-2732-x
- Finney, D. J. (1971). *Probit analysis* (3rd ed.). Cambridge, England: Cambridge University Press.
- Fitzpatrick, C., Caballero-Puntiverio, M., Gether, U., Habekost, T., Bundesen, C., Vangkilde, S., . . . Petersen, A. (2017). Theory of visual attention (TVA) applied to mice in the 5-choice serial reaction time task. *Psychopharmacology, 234*, 845–855. doi: 10.1007/s00213-016-4520-6
- García-Pérez, M. A., & Alcalá-Quintana, R. (2012). On the discrepant results in synchrony judgment and temporal-order judgment tasks: A quantitative model. *Psychonomic Bulletin & Review, 19*, 820–846. doi: 10.3758/s13423-012-0278-y
- Guilford, J. P. (1954). *Psychometric methods* (2nd ed.). New York, NJ: McGraw-Hill.
- Heath, R. A. (1984). Response time and temporal order judgment in vision. *Australian Journal of Psychology, 36*, 21–34. doi: 10.1080/00049538408255075
- Hirsh, I. J. (1959). Auditory perception of temporal order. *The Journal of the Acoustical Society of America, 31*, 759–767. doi: 10.1121/1.1907782

- Hirsh, I. J., & Sherrick, C. E. (1961). Perceived order in different sense modalities. *Journal of Experimental Psychology*, *62*, 423–432. doi: 10.1037/h0045283
- Hung, J., Driver, J., & Walsh, V. (2005). Visual selection and posterior parietal cortex: Effects of repetitive transcranial magnetic stimulation on partial report analyzed by Bundesen's theory of visual attention. *Journal of Neuroscience*, *25*, 9602–9612. doi: 10.1523/JNEUROSCI.0879-05.2005
- Jaśkowski, P. (1993). Selective attention and temporal-order judgment. *Perception*, *22*, 681–689. doi: 10.1068/p220681
- Krüger, A., Tünnermann, J., & Scharlau, I. (2016). Fast and conspicuous? Quantifying salience with the theory of visual attention. *Advances in Cognitive Psychology*, *12*, 20–36. doi: 10.5709/acp-0184-1
- Krüger, A., Tünnermann, J., & Scharlau, I. (2017). Measuring and modeling salience with the theory of visual attention. *Attention, Perception, & Psychophysics*, *79*, 1593–1614. doi: 10.3758/s13414-017-1325-6
- Kruschke, J. K., & Vanpaemel, W. (2015). Bayesian estimation in hierarchical models. In J. R. Busemeyer, Z. Wang, J. T. Townsend, & A. Eidels (Eds.), *The Oxford handbook of computational and mathematical psychology* (pp. 279–298). Oxford, England: Oxford University Press.
- Lake, J. (2016). Recent advances in understanding emotion-driven temporal distortions. *Current Opinion in Behavioral Sciences*, *8*, 214–219. doi: 10.1016/j.cobeha.2016.02.009
- Logan, G. D. (1996). The CODE theory of visual attention: An integration of space-based and object-based attention. *Psychological Review*, *103*, 603–649. doi: 10.1037//0033-295X.103.4.603
- Luce, R. D. (1977). The choice axiom after twenty years. *Journal of Mathematical Psychology*, *15*, 215–233. doi: 10.1016/0022-2496(77)90032-3
- Miller, J., & Schwarz, W. (2006). Dissociations between reaction times and temporal order judgments: A diffusion model approach. *Journal of Experimental Psychology: Human Perception and Performance*, *32*, 394–412. doi: 10.1037/0096-1523.32.2.394
- Mitchinson, B. (2016). Tactile attention in the vibrissal system. In T. Prescott, E. Ahissar, & E. Izhikevich (Eds.), *Scholarpedia of touch* (pp. 771–779). Paris, France: Atlantis Press.
- Miyazaki, M., Yamamoto, S., Uchida, S., & Kitazawa, S. (2006). Bayesian calibration of simultaneity in tactile temporal order judgment. *Nature Neuroscience*, *9*, 875–877. doi: 10.1038/nn1712
- Nicholls, M. E. R. (1996). Temporal processing asymmetries between the cerebral hemispheres: Evidence and implications. *Laterality: Asymmetries of Body, Brain and Cognition*, *1*, 97–137. doi: 10.1080/713754234
- Nordfang, M., Staugaard, C., & Bundesen, C. (2017). Attentional weights in vision as products of spatial and nonspatial components. *Psychonomic Bulletin & Review*. Advance online publication. doi: 10.3758/s13423-017-1337-1
- Petersen, A., Kyllingsbæk, S., & Bundesen, C. (2012). Measuring and modeling attentional dwell time. *Psychonomic Bulletin & Review*, *19*, 1029–1046. doi: 10.3758/s13423-012-0286-y
- Petersen, C. C. (2007). The functional organization of the barrel cortex. *Neuron*, *56*, 339–355. doi: 10.1016/j.neuron.2007.09.017
- Plummer, M. (2003). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. In K. Hornik, F. Leisch, & A. Zeileis (Eds.), *Proceedings of the 3rd international workshop on distributed statistical computing* (pp. 1–10). Vienna, Austria: Technische Universität Wien.
- Rutschmann, R. (1966, May). Perception of temporal order and relative visual latency. *Science*, *152*(3725), 1099–1101. doi: 10.1126/science.152.3725.1099
- Scharlau, I., Ansorge, U., & Horstmann, G. (2006). Latency facilitation in temporal-order judgments: Time course of facilitation as a function of judgment type. *Acta Psychologica*, *122*, 129–159. doi: 10.1016/j.actpsy.2005.10.006
- Schneider, W. X. (2013). Selective visual processing across competition episodes: A theory of task-driven visual attention and working memory. *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences*, *368*, 1–13. doi: 10.1098/rstb.2013.0060
- Schofield, T. P., Yousef, H., & Denson, T. F. (2017). No experimental evidence for visual prior entry of angry faces, even when feeling afraid. *Emotion*, *17*, 78–87. doi: 10.1037/emo0000205
- Schormans, A. L., Scott, K. L., Vo, A. M. Q., Tyker, A., Tytlt, M., Stolzberg, D., & Allman, B. L. (2017). Audiovisual temporal processing and synchrony perception in the rat. *Frontiers in Behavioral Neuroscience*, *10*:246. doi: 10.3389/fnbeh.2016.00246.
- Shuler, M. G., Krupa, D. J., & Nicolelis, M. A. L. (2001). Bilateral integration of whisker information in the primary somatosensory cortex of rats. *Journal of Neuroscience*, *21*, 5251–5261.
- Sternberg, S., & Knoll, R. L. (1973). The perception of temporal order: Fundamental issues and a general model. In S. Kornblum (Ed.), *Attention and performance IV* (pp. 629–685). New York, NY: Academic Press.
- Taagepera, R. (2008). *Making social sciences more scientific: The need for predictive models*. Oxford, England: Oxford University Press.
- Tünnermann, J. (2016). *On the origin of visual temporal-order perception by means of attentional selection* (Doctoral dissertation, Paderborn University, Paderborn, Germany).
- Tünnermann, J., Krüger, A., & Scharlau, I. (2017). Measuring attention and visual processing speed by model-based analysis of temporal-order judgments. *Journal of Visualized Experiments*, *119*, e54856. doi: 10.3791/54856
- Tünnermann, J., Petersen, A., & Scharlau, I. (2015). Does attention speed up processing? Decreases and increases in processing rate in visual prior entry. *Journal of Vision*, *15*(3): 1, 1–27. doi: 10.1167/15.3.1

- Tünnermann, J., & Scharlau, I. (2016). Peripheral visual cues: Their fate in processing and effects on attention and temporal-order perception. *Frontiers in Psychology, 7*, 1442. doi: 10.3389/fpsyg.2016.01442
- Tünnermann, J., & Scharlau, I. (2018). *Stuck on a plateau: A model-based approach to fundamental issues in visual temporal order judgments*. Manuscript submitted for publication. Preprint available at www.preprints.org/manuscript/201804.0213/v1
- Vanden Bulcke, C., Crombez, G., Burnez, W., & Van Damme, S. (2015). Is attentional prioritization on a location where pain is expected modality-specific or multisensory? *Consciousness and Cognition, 36*, 246–255. doi: 10.1016/j.concog.2015.07.003
- Vangkilde, S., Coull, J. T., & Bundesen, C. (2012). Great expectations: Temporal expectation modulates perceptual processing speed. *Journal of Experimental Psychology: Human Perception and Performance, 38*, 1183–1191. doi: 10.1037/a0026343
- Vibell, J., Klinge, C., Zampini, M., Nobre, A., & Spence, C. (2017). Differences between endogenous attention to spatial locations and sensory modalities. *Experimental Brain Research, 235*, 2983–2996. doi: 10.1007/s00221-017-5030-4
- Wada, M., Higo, N., Moizumi, S., & Kitazawa, S. (2010). C-Fos expression during temporal order judgment in mice. *PLOS ONE, 5*, e10483. doi: 10.1371/journal.pone.0010483
- Wada, M., Moizumi, S., & Kitazawa, S. (2005). Temporal order judgment in mice. *Behavioural Brain Research, 157*, 167–175. doi: 10.1016/j.bbr.2004.06.026
- Weiß, K., & Scharlau, I. (2011). Simultaneity and temporal order perception: Different sides of the same coin? Evidence from a visual prior-entry study. *Quarterly Journal of Experimental Psychology, 64*, 394–416. doi: 10.1080/17470218.2010.495783
- Woodworth, R. S., & Schlosberg, H. (1961). *Experimental psychology*. New York, NY: Holt, Rinehart, & Winston.
- Zampini, M., Shore, D. I., & Spence, C. (2003). Multisensory temporal order judgments: The role of hemispheric redundancy. *International Journal of Psychophysiology, 50*, 165–180. doi: 10.1016/S0167-8760(03)00132-6
- Zampini, M., Shore, D. I., & Spence, C. (2005). Audiovisual prior entry. *Neuroscience Letters, 381*, 217–222. doi: 10.1016/j.neulet.2005.01.085

RECEIVED 08.06.2017 | ACCEPTED 21.01.2018