A method for detection of inattentional feature blindness

Aire Raidvee¹ · Mai Toom¹ · Jüri Allik^{1,2}

Accepted: 22 December 2020 / Published online: 2 March 2021

© The Psychonomic Society, Inc. 2021

Abstract

In ensemble displays, two principal factors determine the precision with which the mean value of some perceptual attribute, such as size and orientation, can be discriminated: inefficiency and representational noise of each element. Inefficiency is mainly caused by biased inference, or by inattentional (feature) blindness (i.e., some elements or their features are not processed). Here, we define *inattentional feature blindness* as an inability to perceive the value(s) of certain feature(s) of an object while the presence of the object itself may be registered. Separation of the effects of inattentional (feature) blindness and perceptual noise has escaped traditional analytic methods because of their trade-off effects on the slope of the psychometric discrimination function. Here, we propose a method that can separate the effects of inattentional feature blindness from that of the representational noise. The basic idea is to display a set of elements from which only one contains information relevant for solving the task, while all other elements are “dummies” carrying no useful information because they do not differ from the reference. If the single informative element goes unprocessed, the correct answer can only be given by a random guess. The guess rate can be modeled similarly to the lapse rate, traditionally represented by λ . As an illustration, we present evidence that the presence versus lack of inattentional feature blindness in orientation pooling depends on the feature types present in the display.

Keywords Ensemble perception · Inattentional feature blindness · Representational noise · Lapsing rate

It is not easy to overlook inattentional blindness, especially after nearly 25 million people have watched the original invisible gorilla demonstration.¹ Because in the original setting, about a half of spectators do not notice a gorilla among ball-throwing humans (Simons & Chabris, 1999), it can be speculated that there are millions who have momentary attention lapses, due to which they miss an object which literally strikes the eye. Unsurprisingly, inattentional blindness can be studied in more controlled experimental conditions. For example, Mack and Rock developed a study protocol according to which observers were asked to judge which of the two arms of a briefly displayed cross was longer (Mack & Rock, 1998). On some trials, an unexpected object appears simultaneously with the cross at some place on the screen. About one quarter of observers are inattentionally blind when the judged cross is presented at fixation and the unexpected object is presented

peripherally. However, when positions were reversed, inattentional blindness reached 75% of all trials (Mack & Rock, 1998). Thus, inattentional blindness has come to stay and there are no excuses for ignoring this phenomenon.

Still, there are many perceptual tasks involving high numbers of visual elements that (at least at first glance) seem to escape inattentional blindness. One of such tasks is estimating the mean value of some stimulus attribute, typically size or orientation, of an object ensemble. Miller and Sheldon (1969) were the first who noticed that the average length and the average orientation of a set of lines were estimated approximately with the same accuracy as the length and orientation of a single element displayed alone. A number of studies that followed bolstered a conviction that the mechanism processing the mean value of an attribute pooled together across multiple elements—also called *ensemble perception*—is as accurate as inspection of a single element presented alone (Ariely, 2001; Baek & Chong, 2019; Chong & Treisman, 2003, 2005; Whitney & Leib, 2018). Because the discrimination accuracy was almost independent of the number of elements (Ariely, 2001, Fig. 4), it was regarded as evidence of a smart mechanism capable of evading severe capacity limits that attention and short-term memory are known to have (Cowan, 2001, 2015; Jevons, 1871). This was also the main reason why perception of ensemble properties was viewed as a solution to

¹ <https://www.youtube.com/watch?v=vJG698U2Mvo>

✉ Aire Raidvee
aire.raidvee@ut.ee

¹ Institute of Psychology, University of Tartu, Näituse 2,
50409 Tartu, Estonia

² Estonian Academy of Sciences, Tallinn, Estonia

coping with information overflow. Indeed, the replacement of a large number of objects by a single statistical summary representing them all looks like an ingenious solution (Alvarez, 2011; Alvarez & Oliva, 2008; Chong & Treisman, 2005; Kahneman, 2011; Whitney & Leib, 2018).

However, these allures for automatic and high-capacity mechanisms were slightly calmed down by Myczek and Simons (2008) who calculated that a typical precision that was observed in the mean size perception task can be achieved by processing (noiselessly) only two to three elements and completely ignoring the rest. Thus, there is no need for a capacity-unlimited system if the same job can be done by a focused attention mechanism, capable of processing only a few elements at a time. In response to these critical comments, researchers who defended ensemble perception capacity correctly noted that simulation-based data cannot estimate the number of elements actually processed (Ariely, 2008). Even before Myczek and Simons's (2008) paper, researchers, and especially model-builders, routinely used an idea that observer's decisions are not based on all the available information but only a relatively small fraction of it (e.g., Burgess & Barlow, 1983). For example, in a study on pooling local orientation signals, Dakin (2001) proposed that in making decisions about the mean orientation, observers were ignoring a substantial proportion—a fixed power of the actual number—of elements (namely, the square root of the set size). In a similar manner, Solomon et al. (2011, Experiment 2) concluded that their observers used only $\sqrt{N}$ of all objects that were displayed if their task was to estimate the mean size of eight circles. This implies, for example, that out of eight displayed circles, only up to three would be attended and the rest ignored due to inattentional feature blindness towards their orientation. Yet another study concluded that for visual estimation of mean size a constant proportion of all elements was used (Legge, Gu, & Luebker, 1989). In our study on mean size discrimination (Allik et al., 2013, Fig. 2), we presumed that information of about half of the eight displayed elements was omitted because they were unattended in almost every trial. According to calculations, even a set of four elements was not protected from inattentional feature blindness because the best model fit was obtained when one element (or 25%) from the available four was systematically ignored. It is important to notice that all these estimates were obtained by fitting models to data, not as a result of any experimental procedure that could unconfound the effects of inattention and the perceptual noise with which each element was represented.

The main reason why modeling or simulation-based data cannot tell inattentional feature blindness apart from representational noise is that the indicator typically analyzed is the slope of discrimination function. The threshold defined as the slope of the discrimination function increases when the representational noise becomes larger. But it also increases when the fraction of used information becomes smaller, and there is no

way to tell which of these two possible factors caused a specific discrimination function slope (Allik et al., 2013). Thus, progress in understanding ensemble perception is in finding a unique indicator of inattentional feature blindness that cannot be confused with representational noise.

In a typical psychophysical task, the main interest is in establishing perceptual thresholds—stimulus intensities required to produce a given level of performance. Thresholds are derived from a psychometric function which describes the dependence of an observer's performance on some physical aspect of the stimulus (Wichmann & Hill, 2001). One factor that complicates accurate estimation of thresholds is the human observer's inclination to commit stimulus-independent errors or simply lapses. Because lapses can bias the threshold readings derived from the psychometric function, psychophysicists developed procedures for taking the lapsing rate into account (Klein, 2001; Prins, 2012; Strasburger, 2001; Treutwein & Strasburger, 1999; Wichmann & Hill, 2001). Typically, the lapsing rate λ is defined as a gap between the 100% perfect performance and the observed maximal rate of correct answers (e.g., Strasburger, 2001, Fig. 1). Although the lapse rate may be as low as 0.05% (Arditi, 2006), there is no guarantee that can protect us from much higher lapsing rates.

Here, we use a modified form of the psychometric discrimination function (Thurstone, 1927) to account for the lapsing rate. The discrimination function describes the probability of choices when the observer was asked to tell which of the two presented stimuli is larger on some perceptual attribute. For instance, we can ask which of two displayed circles is larger or which oriented elements are tilted further from the vertical orientation. Because we have an excellent internal frame of reference, we can present a single line and ask whether its tilt is clockwise or counterclockwise relative to the imagined vertical as an implicit standard (cf. Nachmias, 2006). A typical empirical discrimination function shows how well we can discriminate the direction of tilt. In most cases, the choice probability *Prob* can be sufficiently adequately described by a normal cumulative distribution function (shortly *normcdf*) which relates the choice probability to the stimulus attribute based on which the decisions were made. In this study, this attribute was the relative orientation— $\Delta\phi$ is negative for counterclockwise, and positive for clockwise tilts. The probability that the tilt direction was identified as positive (clockwise) can be described by a function $Prob = normcdf(\Delta\phi, \mu, \sigma)$, where μ and σ are the mean and standard deviation of the normal distribution function, respectively. However, if in some proportion of trials the observer's answers are not stimulus related—that is, the answers result from inattention or motor error—we need a slightly modified function. By an analogy with the detection function, the discrimination function requires that the lapsing rate λ should be taken into account. If 2λ is the lapsing rate or proportion of trials in which the observer's answers were not based on the stimulus

information occurring stochastically, then a modified function could be written in the following form: $Prob = (1 - 2\lambda) \times \text{normcdf}(\Delta\phi, \mu, \sigma) + \lambda$, where 2λ is the proportion of trials in which the correct answer is determined randomly (see Fig. 1 for an illustration). Effectively the same formula was used previously to take into account lapsing rates (Solomon, 2010, Equation 2; Solomon et al., 2016, Equation 1). Compared with a customary two-parametric discrimination function (with parameters μ and σ), the amplitude of the modified function is compressed by a factor of $1 - 2\lambda$ and a gap λ is left below and above preventing the function tails from asymptotically approaching either 0% or 100% correct answers. Thus, lapse rate of 2λ alters the shape of the discrimination function.

As we said, the lapses could occur for various reasons, perhaps the two most common being motor errors and lack of attention. What is common to all these reasons, however, is that they are not related to the stimulus information and their occurrence is arbitrary. If in multielement displays one of the elements was unattended then this loss could be compensated for by the stimulus information carried by other elements. Yet this is true only if other elements indeed carry some information that is usable in answering the posed question. Thus, there seems to be a special case when only one element contains all the information relevant for answering the question correctly. If this element is lost, then the correct answer can only be guessed randomly because other elements cannot provide any useful information.

This leads us to a potential indicator that is capable of detecting inattentional feature blindness while not confusing it with representational noise. The principal idea is to present multiple elements from which only one has a tilt that differs

from the reference. More specifically, we present the observer with a set of Gabor patches from which only one has an orientation different from the vertical reference. All other elements are vertically aligned and serve as “dummies” because they serve no use in answering the posed question. The logic of the experimental design is that even if only one element out of the N displayed elements goes unattended, there is a probability of $1/N$ that the observer has no information to answer the question correctly. If this happens then the correct answer can be randomly guessed between two response alternatives. The trials when this happens are lapses because the given answers are not related to stimulus information. Technically, our task is to compare which form of the discrimination function—a normal cumulative distribution function or one that also takes into account the lapse rate—provides a better fit to the empirical discrimination function. Because the most conspicuous trace left by lapses, including those that were caused by inattentional feature blindness, is the λ gap between 100% and 0% correct answers and the actually reached maximum or minimum, the proposed test essentially lies in detecting the magnitude of the gaps between the discrimination function asymptotes and the theoretical minimum/maximum.

We also need to mention a relation to a well-known formula for the correction of random guessing (Blackwell, 1952). If a detection task is organized in a manner where trials in which the signal is present are alternating with trials containing no signal then we can correct the probability of the “Yes” answer by subtracting from it the false-alarm rate that was obtained from trials where the absence of the stimulus was wrongly identified as its presence. Because the comparative judgment does not contain any noise-only trials, we have no explicit measure of the possible bias in the random guessing rate. However, we can consider this as a reverse task for finding the best estimate for the random guessing rate.

Method

Apparatus

Stimuli were presented on a flat LCD monitor (60 Hz) at a distance ensuring that one pixel of the screen would subtend about 1 min of arc of the visual field. Stimuli were generated using MATLAB (The MathWorks, Inc.), the Psychophysics Toolbox (Brainard, 1997), and CircStat Toolbox (Berens, 2009).

Participants

Three adult participants—S1, S2, and S3—with normal or corrected-to-normal vision took part in the experiment. One participant was naïve to the principles underlying the stimulus generation.

Fig. 1 An illustration of how the shape of a habitual psychometric function ($\lambda = 0$) changes if the random guessing parameter λ takes values of 5%, 10%, or 20%

Procedure

Each trial started with the presentation of a fixation cross (1 s), after which a set of eight stimulus elements (either Gabor patches or colorful ellipses; see Fig. 2 for an illustration) were presented for 250 ms against a homogenous gray background in the central annulus of the screen (diameter 10.6° of visual angle). Each stimulus element was surrounded by an inhibitory radius to avoid spatial overlap between the elements. Participant's task was to indicate with the corresponding mouse button press whether the presented set had an average tilt to the left or right compared with the reference orientation, which was vertical. Thus, the reference was implicit without any specific reference stimulus. All stimulus elements except for one were always aligned vertically. One element was odd, with an orientation $\Delta\phi \neq 0$ tilted either clockwise or counterclockwise from the vertical ($\Delta\phi = 0$). Because the zero mean tilt was avoided, the mean (or summary) tilt value of all N elements was always smaller or larger than the vertical reference. In each trial, the choice between two buttons was recorded from which a psychometric function relating *Prob* (the right button meaning “tilted clockwise”) to the mean orientation $\Delta\phi$ of all eight elements was constructed. Auditory feedback on response correctness was provided in each trial (a beep of 400 Hz in case of correct, and 600 Hz in case of an incorrect response).

Experiment with colored ellipses

The ensemble display contained eight colorful ellipses positioned randomly. Long and short axes of each colorful ellipse subtended 0.9 and 0.65 degrees of visual angle, respectively. The mean color of the ten ellipses was randomized on a trial-by-trial basis along an isoluminant color wheel (CIE $L^*a^*b^*$ color space) generated by MemToolBox (Suchow et al., 2013), and the circular standard deviation of the color distribution in each trial was 60° .

Fig. 2 Examples of the two types of stimuli used in this study: (a) One of eight colored ellipses, or (b) one of eight Gabors was tilted away from the vertical

Experiment with Gabor patches

Stimulus display consisted of eight randomly positioned Gabor patches—sinusoidal gratings embedded in a Gaussian envelope, with phase of 0° , 2.3 cycles per degree, the value of the spatial constant of the Gaussian hull function was 11, Michelson contrast of 75%, 1 degree of visual angle per grating.

Results

Figure 3 demonstrates probabilities *Prob* related to the mean orientation $\Delta\phi$ relative to the internal vertical reference. Data points were approximated with two different discrimination functions, a traditional one and the modified version in which a random guessing, or lapsing rate λ was also estimated. The best approximations are shown for the habitual (dotted line) and for the modified (dashed line) psychometric functions. Parameter values for the modified psychometric functions are given, with the estimates of random guessing rates.

In Gabor stimuli, the differences between two discrimination functions are very small, demonstrating that it does not matter if the discrimination function is corrected for the lapsing rate λ or not. This can be verified by a visual inspection as well. If there is a detectable number of lapses—a critical element containing all usable information is unattended—there should be a gap between the empirical data points and 0% and 100% correct answers. Because there are no such gaps, it cannot be concluded that in case of Gabor stimuli, inattentional feature blindness happened on any number of trials.

In the experiment with colored ellipses, differences between the two psychometric functions are obvious, with the lapse rates of 0.1, 0.16, and 0.22 for the three subjects. This means that for the three subjects, in 20%, 32%, and 44% of the trials,

Fig. 3 The probability of choosing the clockwise tilt as a function of the summary orientation approximated by a customary (dotted lines) and a modified (dashed line) psychometric function for the observers S1, S2, and S3

respectively, the answer was given at random. Given the near zero lapse rates in the experiment with Gabor stimuli in the same set of subjects, it can be concluded with reasonable certainty that the majority of lapses were inattentional, not motor. Thus, we have shown that in case of colorful ellipses, in a fraction of trials, the orientation value of at least one stimulus element goes unprocessed—evidence of inattentional feature blindness.

The observers also reported their impressions about displays they were shown and strategies they used to solve the task. Nobody said that they saw one pop-out and a bunch of identical elements. Because all orientation deviations from the reference were relatively small, the structure of stimulus values did not strike the observer's eye. Also, there was no preferred strategy reported—sometimes, the perceived tilt was reported based on some kind of average of all elements' orientations, and sometimes, it was based on a few elements.

Discussion and conclusions

Because inattentional (feature) blindness and representational noise have an identical effect on the slope of the discrimination function, the effects of these two factors have been inseparable from each other (cf. Allik et al., 2013, p. 38). This

uncertainty led to presentation of explanations in which representational noise played no significant role (Myczek & Simons, 2008) or in which a massive involvement of inattentional blindness was presumed—of course, in addition to noisy representations (Allik et al., 2013; Dakin, 2001; Myczek & Simons, 2008; Simons & Myczek, 2008; Solomon et al., 2016; Solomon et al., 2011). In extreme cases, the proposed models assumed that a substantial fraction of elements were missed because of attentional failure. In addition to suggestions that, for example, the square root of the total number of elements are not attended, there has been no evidence about a fixed upper capacity limit. Modeling experiments suggest that the human observer is subject to not noticing all elements, even if their number is four or fewer.

However, the main goal of this paper was to provide an experimental method for detection of inattention in a way that the effects of inattention would be clearly separable from the effects of other factors. Clearly, the task and results from the first experiment with Gabor patches are not novel (see, e.g., Solomon, 2010), because we found no signs of lapsing. This result, however, is valuable in combination with the result from the experiment with colorful ellipses, which detects inattention without confounding with other variables and shows that inattentional feature blindness in orientation pooling

depends on specific feature types. Further studies can show how easily the conclusions about pooling local orientation signals are generalizable, for example, to summarizing size or position information. Despite the evidence of inattentional feature blindness (or lack of it) when orientation information from eight randomly distributed elements was pooled, there would be no guarantees that information about sizes of multiple elements could be executed without losses. However, our previous study suggested that all elements are obligatorily analyzed when the task is to estimate the mean size, at least when there are only four elements to judge (Allik et al., 2014). It would be honest to admit that we have no idea from what number of elements upward does inattentional feature blindness start to operate either for pooling orientation or size information. For example, there is evidence that the perceptual system can attend to only a limited number of spatial positions when the task is to average across multiple positions (Hess et al., 2003). Furthermore, if we increase the number of elements to extremes, then at some point we inevitably pass a threshold up to which all elements are processed. Data presented in this paper do not claim that inattentional feature blindness in (orientation pooling of) Gabor patches cannot exist in principle. Rather, what we demonstrate is that in principle, feature types exist where inattentional feature blindness toward at least one local signal takes place. This means that some feature types are more vulnerable to inattention.

Here, we have used the term “inattentional feature blindness” to accommodate for the differences in the conspicuity of various features types. Since its introduction, the term “inattentional blindness,” probably coined by Arien Mack and Irvin Rock (Rock et al., 1992; Mack & Rock, 1998), has referred to a situation in which an object that would have been easily detected in isolation goes unnoticed. The difference from the works of Mack and Rock (1998) and Simons and Chabris (1999) is that we did not deliberately engage the observers’ attention away from the informative element, and yet (although its presence may have been noticed), at least one of its attributes was not always registered with sufficient accuracy. We show that the registration of this critical feature, here orientation, depends on other aspects of the objects that carry the orientation information.

Thus, what could be regarded as an original contribution of this paper is a method of detecting inattentional feature blindness. The previous methods are based on registering changes in the slope of the discrimination function. This method, as we pointed out in the Introduction, leads to indeterminate results because two different factors—representational noise and inattentional blindness—can cause these changes. Looking for a unique indicator of inattentional blindness, we turned

to the traditional method of taking the lapsing rate λ into account among the parameters of a psychometric function. We took this well-elaborated idea and applied the lapsing rate concept to the discrimination functions. We also used a stimulus set in which missing even a single element deprived the observers from all usable information based on which the posed question could be answered correctly. An experimental protocol that satisfied this requirement was assigning all relevant information to only one of the displayed elements, leaving the remaining elements as “dummies” that did not contribute any useful information. If the single critical element was not attended, the only way to give a correct answer was random guessing, which placed this trial into the category of lapses.

Although we hold in high regard the attempt by Myczek and Simons’s (2008) to explain mean size perception by picking out only two to three elements that are used for the answer, there is no conclusive experimental evidence yet that this is the method by which the mean orientation of a set of elements is actually established. For this explanation, inattentional blindness is presumed, which we found in determining orientation of ellipses, but were not able to detect in pooling the orientations of Gabor patches, even if there were up to eight elements. Based on the current results alone, it cannot be quantified with sufficient accuracy whether more than one element went unprocessed, as ignoring the “dummy” elements was inconsequential.

An indicator by which inattentional feature blindness can be detected without confusion with representational noise would also challenge model builders. It is no longer sufficient to merely demonstrate that a good model fit can be obtained based on the assumption that a considerable number of elements were neglected (e.g., Allik et al., 2013; Dakin, 2001; Solomon & Morgan, 2017; Solomon et al., 2011). More direct tests are needed to demonstrate the degree to which inattentional blindness plays any role in ensemble perception.

To conclude, the proposed method for testing operations of inattentional blindness opens several new perspectives. A general attitude shared by the ensemble perception community appears to be that the ability to compute mean values or averages is a highly generalizable capacity, which transfers from simple attributes such as orientation and size to more complex ones such as facial expressions (cf. Whitney & Leib, 2018). However, from the current results, it seems that different perceptual attributes are differently vulnerable to inattentional blindness. There may also be substantial differences in the number of elements that can be handled without losing any of them due to inattentional feature blindness. We also need to mention that inattentional blindness is not the only mechanism that requires more rigorous testing. The ability to infer mean values is often presumed rather than placed under rigorous testing. Although a

test showed that the sizes of four circles were averaged exactly according to the axioms of arithmetic (Allik et al., 2014), there is no guarantee that these axioms would still apply if the number of elements is different and/or some other perceptual attribute is tested. For instance, there is some evidence that orientation pooling may not be as accurate as was previously thought (Solomon et al., 2016). We hope that the proposed approach to detecting inattentional feature blindness helps to advance the knowledge about ensemble perception.

Acknowledgements Correspondence could be sent to Aire Raidvee, Institute of Psychology, University of Tartu, Näituse 2, Tartu 50409, Estonia, E-mail: aire.raidvee@ut.ee. Aire Raidvee was supported by the Mobilitas Pluss Returning Researcher Grant (MOBTP91) by Estonian Research Council. Mai Toom was supported by the Estonian Research Council grant PRG770.

We especially thank Shaul Hochstein for helpful suggestions and detailed feedback. For useful comments and discussion, we thank Joshua Solomon, Sang Chul Chong, and Andrey Chetverikov.

We thank Endel Pöder for participating in early stages of this project. We are grateful to Tiit Mogom for his enduring support and assistance.

Open practices statement The data and materials for the experiments reported here are available upon request. None of the experiments was preregistered.

References

- Allik, J., Toom, M., Raidvee, A., Averin, K., & Kreegipuu, K. (2013). An almost general theory of mean size perception. *Vision Research*, 83, 25–39. <https://doi.org/10.1016/j.visres.2013.02.018>
- Allik, J., Toom, M., Raidvee, A., Averin, K., & Kreegipuu, K. (2014). Obligatory averaging in mean size perception. *Vision Research*, 101, 34–40. <https://doi.org/10.1016/j.visres.2014.05.003>
- Alvarez, G. A. (2011). Representing multiple objects as an ensemble enhances visual cognition. *Trends in Cognitive Sciences*, 15(3), 122–131. <https://doi.org/10.1016/j.tics.2011.01.003>
- Alvarez, G. A., & Oliva, A. (2008). The representation of simple ensemble visual features outside the focus of attention. *Psychological Science*, 19(4), 392–398. <https://doi.org/10.1111/j.1467-9280.2008.02098.x>
- Arditi, A. (2006). Lapse resistance in the verbal letter reporting task. *Vision Research*, 46(8), 1327–1330. <https://doi.org/10.1016/j.visres.2005.06.014>
- Ariely, D. (2001). Seeing sets: Representation by statistical properties. *Psychological Science*, 12(2), 157–162. <https://doi.org/10.1111/1467-9280.00327>
- Ariely, D. (2008). Better than average? When can we say that subsampling of items is better than statistical summary representations? *Perception & Psychophysics*, 70(7), 1325–1326. <https://doi.org/10.3758/pp.70.7.1325>
- Baek, J., & Chong, S. C. (2019). Distributed attention model of perceptual averaging. *Attention, Perception, & Psychophysics*. <https://doi.org/10.3758/s13414-019-01827-z>
- Berens, P. (2009). CircStat: A MATLAB toolbox for circular statistics. *Journal of Statistical Software*, 31(10), 1–21. <https://doi.org/10.18637/jss.v031.i10>
- Blackwell, H. R. (1952). Studies of psychophysical methods for measuring visual thresholds. *Journal of the Optical Society of America*, 42(9), 606–614. <https://doi.org/10.1364/JOSA.42.000606>
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10, 437–442. <https://doi.org/10.1163/156856897X00357>
- Burgess, A. E., & Barlow, H. B. (1983). The precision of numerosity discrimination in arrays of random dots. *Vision Research*, 23(8), 811–820. [https://doi.org/10.1016/0042-6989\(83\)90204-3](https://doi.org/10.1016/0042-6989(83)90204-3)
- Chong, S. C., & Treisman, A. (2003). Representation of statistical properties. *Vision Research*, 43(4), 393–404. [https://doi.org/10.1016/S0042-6989\(02\)00596-5](https://doi.org/10.1016/S0042-6989(02)00596-5)
- Chong, S. C., & Treisman, A. (2005). Statistical processing: computing the average size in perceptual groups. *Vision Research*, 45(7), 891–900. <https://doi.org/10.1016/j.visres.2004.10.004>
- Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24(1), 87–114. <https://doi.org/10.1017/S0140525X01003922>
- Cowan, N. (2015). George Miller's magical number of immediate memory in retrospect: Observations on the faltering progression of science. *Psychological Review*, 122(3), 536–541. <https://doi.org/10.1037/a0039035>
- Dakin, S. C. (2001). Information limit on the spatial integration of local orientation signals. *Journal of the Optical Society of America A: Optics Image Science and Vision*, 18(5), 1016–1026. <https://doi.org/10.1364/JOSAA.20.000974>
- Hess, R. F., Barnes, G., Dumoulin, S. O., & Dakin, S. C. (2003). How many positions can we perceptually encode, one or many? *Vision Research*, 43(14), 1575–1587. [https://doi.org/10.1016/s0042-6989\(03\)00122-6](https://doi.org/10.1016/s0042-6989(03)00122-6)
- Jevons, W. S. (1871). The power of numerical discrimination. *Nature*, 3, 281–282. <https://doi.org/10.1038/003281a0>
- Kahneman, D. (2011). *Thinking fast and slow*. Farrar, Straus & Giroux.
- Klein, S. A. (2001). Measuring, estimating, and understanding the psychometric function: A commentary. *Perception & Psychophysics*, 63(8), 1421–1455. <https://doi.org/10.3758/bf03194552>
- Legge, G. E., Gu, Y., & Luebker, A. (1989). Efficiency of graphical perception. *Perception & Psychophysics*, 46(4), 365–374. <https://doi.org/10.3758/bf03204990>
- Mack, A., & Rock, I. (1998). *Inattention blindness*. Cambridge, Mass.: MIT Press.
- Miller, A. L., & Sheldon, R. (1969). Magnitude estimation of average length and average inclination. *Journal of Experimental Psychology*, 81, 16–21. <https://doi.org/10.1037/h0027430>
- Myczek, K., & Simons, D. J. (2008). Better than average: Alternatives to statistical summary representations for rapid judgments of average size. *Perception & Psychophysics*, 70(5), 772–788. <https://doi.org/10.3758/pp.70.5.772>
- Nachmias, J. (2006). The role of virtual standards in visual discrimination. *Vision Research*, 46(15), 2456–2464. <https://doi.org/10.1016/j.visres.2006.01.029>
- Prins, N. (2012). The psychometric function: The lapse rate revisited. *Journal of Vision*, 12(6), 25–25. <https://doi.org/10.1167/12.6.25>
- Rock, I., Linnett, C. M., Grant, P., & Mack, A. (1992). Perception without attention: Results of a new method. *Cognitive Psychology*, 24(4), 502–534. [https://doi.org/10.1016/0010-0285\(92\)90017-V](https://doi.org/10.1016/0010-0285(92)90017-V)
- Simons, D. J., & Chabris, C. F. (1999). Gorillas in our midst: sustained inattention blindness for dynamic events. *Perception*, 28(9), 1059–1074. <https://doi.org/10.1068/p2952>
- Simons, D. J., & Myczek, K. (2008). Average size perception and the allure of a new mechanism. *Perception & Psychophysics*, 70(7), 1335–1336. <https://doi.org/10.3758/pp.70.7.1335>

- Solomon, J. A. (2010). Visual discrimination of orientation statistics in crowded and uncrowded arrays. *Journal of Vision*, 10(14), 1–16. <https://doi.org/10.1167/10.14.19>
- Solomon, J. A., May, K. A., & Tyler, C. W. (2016). Inefficiency of orientation averaging: Evidence for hybrid serial/parallel temporal integration. *Journal of Vision*, 16(1). <https://doi.org/10.1167/16.1.13>
- Solomon, J. A., Morgan, M., & Chubb, C. (2011). Efficiencies for the statistics of size discrimination. *Journal of Vision*, 11(12):13, 1–11. <https://doi.org/10.1167/11.12.13>
- Solomon, J. A., & Morgan, M. J. (2017). Orientation-defined boundaries are detected with low efficiency. *Vision Research*, 138, 66–70. <https://doi.org/10.1016/j.visres.2017.06.009>
- Strasburger, H. (2001). Converting between measures of slope of the psychometric function. *Perception & Psychophysics*, 63(8), 1348–1355. <https://doi.org/10.3758/bf03194547>
- Suchow, J. W., Brady, T. F., Fournie, D., & Alvarez, G. A. (2013). Modeling visual working memory with the MemToolbox. *Journal of Vision*, 13(10). <https://doi.org/10.1167/13.10.9>
- Thurstone, L. L. (1927). A law of comparative judgments. *Psychological Review*, 34, 273–286. <https://doi.org/10.1037/h0070288>
- Treutwein, B., & Strasburger, H. (1999). Fitting the psychometric function. *Perception & Psychophysics*, 61(1), 87–106. <https://doi.org/10.3758/bf03211951>
- Whitney, D., & Leib, A. Y. (2018). Ensemble perception. *Annual Review of Psychology*, 69, 105–129. <https://doi.org/10.1146/annurev-psych-010416-044232>
- Wichmann, F. A., & Hill, N. J. (2001). The psychometric function: I. Fitting, sampling, and goodness of fit. *Perception & Psychophysics*, 63(8), 1293–1313. <https://doi.org/10.3758/bf03194544>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.