Comparing integration rules in visual search

Stefano Baldassi

Preeti Verghese

Search performance for a target tilted in a known direction among vertical distractors is well explained by signal detection theory models. Typically these models use a maximum-of-outputs rule (Max rule) to predict search performance. The Max rule bases its decision on the largest response from a set of independent noisy detectors. When the target is tilted in either direction from the reference orientation and the task is to identify the sign of tilt, the loss of performance with set size is much greater than predicted by the Max rule. Here we varied the target tilt and measured psychometric functions for identifying the direction of tilt from vertical. Measurements were made at different set sizes in the presence of various levels of orientation jitter. The orientation jitter was set at multiples of the estimated internal noise, which was invariant across set sizes and measurement techniques. We then compared the data to the predictions of two models: a Summation model that integrates both signal and noise from local detectors and a Signed-Max model that first picks the maxima on both sides of vertical and then chooses the value with the highest absolute deviation from the reference orientation. Although the function relating thresholds to set size had a slope consistent with both the Signed-Max and the Summation models, the shape of individual psychometric functions was in the most crucial conditions better predicted by the Signed-Max model, which chooses the largest tilt while keeping track of the direction of tilt.

Keywords: visual search, spatial vision, signal detection theory, identification

Introduction

When a visual target does not pop out, increasing the number of distractors makes this stimulus harder to find. Classical reaction time studies suggest that each item in the display is scanned sequentially (for a review, see Wolfe, 2000). However, more recent studies based on accuracy measures and signal detection theory (SDT) have disproved the necessity for serial processing, showing that parallel processing of the elements can explain the impairment in performance with increasing set size (for a review, see Verghese, 2001). Here we will refer to these as parallel models for the sake of clarity, even though we think they are “not-necessarily-serial” explanations of search behavior.

Parallel models share the idea that an integration rule combines the responses to individual inputs to yield a decision variable. However, they may differ substantially in the specific rule they implement. In particular, by using threshold measures across various visual dimensions, such as contrast, length, speed, and orientation discrimination, different studies have shown that the increase in thresholds with increasing set size is well fit by the prediction of a SDT model that chooses the strongest response among independent detectors (Palmer, 1994; Palmer, Ames, & Lindsey, 1993; Palmer, Verghese, & Pavel, 2000; Shaw, 1980; Shaw, 1982; Shiu & Pashler, 1995; Verghese & Stone, 1995; Solomon, Lavie, & Morgan, 1997). For this class of models, every new element contributes uncertainty. In other words, the magnitude of the observed threshold increase is consistent with a model that predicts only an increase in uncertainty with increasing set size. Similar models have also been successful in explaining the additional disruption of performance occurring in conjunction search tasks (Eckstein, 1998). All of these studies used paradigms well suited to this integration rule, commonly referred to as maximum-of-outputs rule or Max rule. However, as postulated formerly by Green and Swets (1966) and by Shaw (1982), information might be integrated in a different way via mechanisms that pool together individual responses by what is generally referred to as the summation rule or Sum rule (Graham, Kramer, & Yager, 1987).

Two studies using an identification task (Baldassi & Burr, 2000; Morgan, Ward, & Castet, 1998) have suggested that the information about a given visual feature (the target) may be diluted by adding neutral elements (distractors) through the action of a pooling mechanism. In particular, Baldassi and Burr (2000) observed that orientation thresholds varied with set size according to the square root of the number of elements, and that this relationship persisted across the whole range of noise levels and experimental conditions. In another task where subjects needed to locate the target, the rise in thresholds was substantially shallower, being about one half of that observed for identification. Due to the
difference in the behavior of thresholds in these two tasks, the authors suggested that these tasks dissociated Sum from Max behavior in visual search. The function relating thresholds to set size in the location task had a loglog slope of about 0.25, consistent with a standard Max rule. However, in the identification task, it is debatable whether the square-root relation (log-log slope of 0.5) reflects summation of target and distractor responses, or whether it is due to a variant of the Max rule suited to the particular nature of this task. In fact, both studies (Baldassi & Burr, 2000; Morgan, Ward, & Castet, 1998) used a type of 2-alternative task where the target was always present, but had a value that was randomly on one side of the distractor(s), or the other. In particular, they measured orientation discrimination with a task where the target was tilted clockwise (CW) or counterclockwise (CCW) from vertical. Following Baldassi and Burr (2000), we will call this an identification task because the observer has to identify the sign of the target orientation with respect to a mean, as opposed to the standard odd-man-out search tasks where the target varies along a single direction.

The aim of this study is to compare the experimental data of an identification task with models that consider both Sum- and Max-of-outputs decision rules. Moreover, we will compare the outcome of different models by considering their effects on the whole psychometric function. Psychometric functions reveal information beyond a summary measure such as threshold at a criterion percent correct (e.g., Solomon & Morgan, 2001). Basic studies have shown that the slope of the psychometric function increases with uncertainty or the number of detectors that the observer monitors (Burgess & Oghadrehanian, 1984; Pelli, 1983; Swensson & Judy, 1981; Tanner, 1961; Tyler & Chen, 2000). Specific parameters associated with the steepness of the function and with its horizontal position allow quantifying, at least through relative estimates, the number of detectors an observer monitors for any given task. Some studies have used this as a tool to reveal the functioning of the visual system when a decision must be made among many competitors, as in the case of detecting motion trajectories in dynamic noise (Vergheese & McKee, 2002). Therefore, uncertainty analysis may be a useful tool to reveal mechanisms of visual search.

Models

In this section, we will examine in detail the theoretical background leading to quantitative predictions of the Max and the Sum models.

The two models have at least three stages between the input, consisting of target and distractors, and the output, which is the observer’s response. Both models share a common first stage, which assumes that the response to each stimulus element is an independent noisy variable. Each response is drawn from a distribution whose mean is linearly related to the stimulus value, and whose variance reflects the internal noise (plus added external noise; see “Methods” section). The second stage, which combines the noisy responses, is crucially different between the two. The Sum model postulates that these responses are added together resulting in a single variable whose mean and variance are the sum of the individual means and variances (Green & Swets, 1966). The Max model postulates that the decision depends on the variable producing the strongest response. As the task is to identify the target’s direction of tilt, the Max rule should take this into account. We will call the modified Max rule that we propose the Signed-Max rule. The output of the second stage is fed to a third and final stage, where a decision is taken.

The following paragraphs outline the principles of the Max model and develop the Signed-Max version as a variant that we propose for identification.

As reported in the “Introduction,” many studies have successfully explained search performance in tasks where the target was tilted in a predetermined direction (e.g., clockwise) with a model that bases its decision on the largest response in the relevant detectors. For example, in a 2-interval forced-choice experiment with vertical distractors and a CW tilted target, the observer monitors the activity of a CW tilted detector at each of the n locations in the 2 intervals and chooses the interval that evokes the greatest response in this class of detectors. Therefore, a correct decision will be made if the maximum output comes from any of the locations in the signal interval. If we assume that the response of the tilted detector is a sample from the probability density function, f(r) for a vertical stimulus and f(r-kθ) for a tilted stimulus, then the probability of a correct response is given by

\[ P_{\text{correct}}(\theta) = \int_{-\infty}^{\infty} \left[ f(r-k\theta) F(r)^{2n-1} - (n-1) f(r) f(r-k\theta) F(r)^{2n-2} \right] dr \] (1)

where r is the response from a CW-oriented detector, θ is the orientation of the target, k is a sensitivity parameter that scales the orientation, and F(θ) is the probability distribution

\[ \int_{-\infty}^{r} f(x) dx \]

We will refer to Equation 1 as the standard Max rule. Here we assume that f(x) has a Gaussian distribution. The standard Max rule assumes that at each location there is a single detector tuned to, say, a CW orientation away from vertical. In an identification task, we then need to consider a modified form of this rule because the target can be tilted in either direction from vertical.

We propose that at each location there is one detector tuned for CW and one for CCW orientation (Figure 1).
These detectors have preferred orientations centered beyond the largest tilt used in our experiment (28° ± jitter), say 30° and -30° from vertical. This allows the response of these detectors to increase with increasing tilt away from vertical. These assumptions are consistent with known physiology (e.g., Blakemore & Campbell, 1969; Hubel & Wiesel, 1968): the orientation tuning of simple cells in primate cortex is broad (full width at half height equal to 22.5°) and the orientation dimension is sparsely sampled (about 7 unique orientations). The width of this distribution does not represent the variability in response to a single tilt, but rather the bandwidth of these orientation-tuned detectors.

For both the target and the distractor locations, one response is taken from each of these two detectors. A correct choice is made when the largest response comes from the detector whose preferred direction of tilt matches the target tilt. Consider the CW tilted target (dark gray) shown along the abscissa of Figure 1. It produces a response in both the CW and the CCW detectors (shown by the blue and red horizontal lines, respectively). Due to noise, these responses vary around a mean response, as shown by the blue and red distributions on the vertical panel at the right for the CW and CCW detectors, respectively. A vertical distractor produces equal mean responses in the two detectors, as shown by the green distribution (there are two overlapping distributions of responses to the distractor from CW and CCW detectors).

We assume that the decision requires a comparison between the outputs of the two detectors. We represent that comparison as the difference between CW and CCW responses at each location, as subtraction has been proposed by a number of biologically plausible models of neural competition (e.g., Desimone & Duncan, 1995). In the small vertical panel on the right of Figure 1, the distribution of responses of the CCW (red dashed line) and CW (blue solid line) detectors are by their nature positive. These responses are shown in the top and middle panels of Figure 2, which plot the response distribution of the CW and CCW detectors, respectively, to 3 different stimulus orientations. The bottom panel of Figure 2 shows the distributions of the difference between the responses of the two detectors, what we can call a difference detector. This subtraction operation produces distractor responses that have a zero mean, CW responses that have a positive mean value, and CCW responses that have a negative mean. Note that while the responses of the two individual detectors have a standard deviation equal to 1, the difference distribution has a standard deviation that is larger by a factor of $\sqrt{2}$. The

---

**Figure 1.** Representation of the Signed-Max model. The big horizontally oriented panel shows the two detectors tuned to opposite orientations with respect to vertical, as a function of the angle of a stimulus. The detectors for CCW and CW orientations are represented by dashed red and continuous blue lines, respectively. The small oriented gratings along the abscissa show the orientation range spanned by these detectors, with the vertical orientation in green and a possible target orientation in dark gray. Because we do not have vertical targets, the two detectors produce responses of different strengths to a tilted target. The mean response strengths of the two detectors are marked by horizontal lines that project to the vertical panel on the right, where the response of the CCW detector to the target is represented by a dashed red line, that of the CW detector by a continuous blue line, and the overlapping mean response of the two detectors to the distractor(s) by a dotted green line. The small vertical panel on the right shows the variability of these mean responses. The response probability distributions are centered at the output response of the respective detectors and have noise (σ) from internal and, in our experiment, external sources. This figure shows the effect of a CW-oriented stimulus, but the same rationale holds for a CCW target.
distributions in the bottom panel of Figure 2 represent the output of this difference operation.

To evaluate which direction of tilt has the larger response, we determined whether the largest positive or the largest negative response has the greater absolute value. Any value generated by the target or the distractors would be the largest, in absolute magnitude, if all the other samples produced responses between −r and r.

According to the Signed-Max model, the probability of choosing the clockwise orientation is given by

\[
P_{\text{correct}}(\theta) = \frac{\int_{-\infty}^{\infty} f\left(\frac{r - k\theta}{\sqrt{2}}\right) \int_{-\infty}^{r} f\left(\frac{r'}{\sqrt{2}}\right) dr \, dr' - \int_{-\infty}^{-\infty} f\left(\frac{r - k\theta}{\sqrt{2}}\right) \int_{-\infty}^{-r} f\left(\frac{r'}{\sqrt{2}}\right) dr' \, dr}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f\left(\frac{r - k\theta}{\sqrt{2}}\right) \int_{-\infty}^{r} f\left(\frac{r'}{\sqrt{2}}\right) dr \, dr' - \int_{-\infty}^{\infty} \int_{-\infty}^{-\infty} f\left(\frac{r - k\theta}{\sqrt{2}}\right) \int_{-\infty}^{-r} f\left(\frac{r'}{\sqrt{2}}\right) dr' \, dr}
\]

(2)

where r is the response from an individual detector, θ is the orientation difference, fr(r) is the probability density function for responses, k a sensitivity parameter, and n is the uncertainty parameter. The Signed-Max model is a variant of the standard Max model that takes absolute values and keeps track of their sign. The equation is the sum of two terms. The first part of the equation calculates the probability that the response to a CW target (a sample from the distribution \( f\left(\frac{r - k\theta}{\sqrt{2}}\right) \)) is larger in magnitude than the response to vertical distractors (n-1 samples from the distribution \( f\left(\frac{r}{\sqrt{2}}\right) \)). Note that these responses are samples from the difference distribution of CW and CCW responses. The second part calculates the probability that one of the n-1 distractor responses has the largest positive value, and the response from the target and the other distractor elements is smaller in magnitude. The integration limits in the first and second parts of the equation go from −r to +r. This is because any value generated by the target or distractor will be the largest in absolute magnitude if all of the other responses lie between −r and r. The integration limits of the entire expression go from 0 to \( \infty \) because only positive responses are correct for a clockwise target.1

For the Sum model, if we assume Gaussian distributions of equal variance, then the sum of distributions resulting from 1 CW and n-1 vertical distractors has a mean equal to kθ, where θ is the CW target angle and the variance is nσ^2. The probability that a sample from such a distribution has a value greater than 0 (CW) is

\[
P_{\text{correct}}(\theta) = \int_{0}^{\infty} e^{-\frac{(r - k\theta)^2}{2n}} dr
\]

(3)

where \( k = \frac{1}{\sigma_{\text{tot}}} \) and \( \sigma_{\text{tot}}^2 = \sigma_{\text{int}}^2 + \sigma_{\text{ext}}^2 \).

In Equations 1 and 2, k is a parameter that defines signal-to-noise ratio. In Equation 3, the factor k/n modulates the signal-to-noise ratio. Because k and n cannot be independently evaluated in this formulation, we set k equal to the total variance and calculated the best-fitting value for the parameter n.

For both the Signed-Max and the Sum rules, the above equations predict the probability of responding correctly for a CW target. The same rationale applies to CCW targets as we assume no bias in the model for CW and CCW targets. In fact, we verified experimentally this assumption by comparing the psychometric functions for CW responses to that of CCW responses.

The two models have different signatures on the overall shape of the psychometric functions. For the Signed-Max model, increasing the set size changes the
schematic of the model prediction for set size 2 to 16. With increasing set size (Figure 3, blue curves). If instead target and distractors signals are summed together, then the distractors decrease the signal to noise ratio, with the net effect of a rightward shift of the psychometric function with increasing set size (Figure 3, red curves). Figure 3 is a schematic of the model prediction for set size 2 to 16.

![Signed-Max Model and Sum Model](image)

Figure 3. Schematic representation of the signature effects of the Signed-Max model (left panel) and the Sum model (right panel) on the psychometric functions generated at different set sizes (set size grows from left to right). The Signed-Max model predicts steeper functions with increasing set size, while the Sum model predicts a rightward shift of the whole function. The two sets of predictions diverge progressively with increasing set size. To compare the two models, each panel shows one model's predictions superimposed on the other (dashed lines). Note that these psychometric functions are plotted in log-linear coordinates.

To determine which one of the two models provides a better account of the data, we used Equations 2 and 3 to fit the curves relating the percentage of correct responses to the angle of the target in our task. If we are able to verify that the major source of noise limiting this task is local (at the level of the individual elements) and that central noise (the noise source at the level of the integration) is negligible, then we will be able to work out predictions for both the Sum and the Max models on a common and comparable ground. Furthermore, we have the advantage of using empirical estimates of internal noise. Both models have two variables. The sensitivity parameter $k$ represents a scalar that relates orientation to detector response. The parameter $n$ represents uncertainty, or the number of detectors (locations) monitored. Uncertainty will affect the Sum model by increasing the variance of the summed distribution, and will affect the Signed-Max model by increasing the units to be monitored. We used an iterative procedure to find the best-fitting values of $k$ and $n$ for a given set of data. In the Signed-Max model, we left both $k$ and $n$ free to vary, whereas we effectively had only one free parameter for the Sum model. Because both $k$ and $n$ modulate the signal-to-noise ratio in the Sum model, we set $k$ equal to the inverse of the internal noise plus the external noise, and allowed $n$ to vary. It is of interest to note that the values of $k$ returned by the Signed-Max fit, where it varied freely, closely matched the fixed value of $k$ in the Sum model.

**Methods**

Stimuli and procedure were designed to match the Baldassi and Burr (2000) task. The stimuli were generated in Matlab using the Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997), and displayed on a Sony Multiscan 210GS monitor at a refresh rate of 75 Hz. Two observers performed the experiment, one author (S.B.) and a naïve paid subject; both of them had normal or corrected vision.

Each individual stimulus was a Gabor patch of space constant equal to $0.5^\circ$ and spatial frequency of 2 cpd, displayed at 50% contrast (Figure 4). The patches were $5^\circ$ eccentric from fixation and their positions were distributed to maximize the angular separation at different set sizes. Set size was varied from 1 to 16 and exposure time was 8 frames (106.7 ms). The viewing distance was 57 cm and the mean luminance 19 cd/m$^2$. Pilot data indicated that crowding effects were under control. Thresholds obtained with two elements were similar whether the elements were on opposite sides of fixation or were separated by $22.5^\circ$ (the angular separation for 16 elements).

Each trial began with the observer fixating a small $0.05^\circ$ fixation square that was always on, followed by the stimulus presentation. Subjects were asked to report the direction of tilt of the single target that appeared at a random location around the notional circle. In the set size 1 condition, the target was randomly displayed in 1

![Figure 4. Example of the stimuli used. Set size increases from left to right, going from 2 to 16.](image)
out of 4 predetermined locations. This avoids foveation of a fixed target location but does not add significant uncertainty as the target was supra-threshold. Acoustic feedback was provided and a response triggered the next trial. Sessions were blocked by set size and jitter level. In some conditions, we added noise (orientation jitter) to the stimuli. This dimensional noise (e.g., Verghese & Stone, 1995) is characterized by variability within the dimension of interest (orientation), rather than the standard pixel contrast-modulated noise used in similar studies. It has the advantage of directly affecting orientation detectors responsible for the task, and not requiring any assumption about the way contrast is related to orientation. Different amounts of jitter were used in separate sessions. We decided not to use the QUEST procedure (Watson & Pelli, 1983) as in the Baldassi and Burr (2000) study, but rather a fixed set of angles that spanned the whole range of performance from chance to perfect, interleaved within a session. While adaptive procedures concentrate trials around the inferred threshold value, we wanted the same number of trials over the entire psychometric function. Orientation jitter was introduced by setting the standard deviation of the orientation distribution of both target and distractors to a multiple of the internal orientation noise estimate (threshold for 1 element). The mean of the distractor distribution was 0 (vertical), and the mean of the target distribution was ± target angle, with 50% probability of being CW or CCW. Target and distractors were independently drawn from these noisy orientation distributions. In the no-noise condition, the added noise was 0, so target and distractors were displayed at the mean of their distributions.

All these conditions produced a set of 20 psychometric functions (4 set sizes x 5 noise levels) for each subject for the main experiment. Set sizes included 2, 4, 8, and 16 elements, and noise levels included no noise, 0.5, 1, 2, and 4 times the internal noise level estimated in the absence of external noise (see below). A block of trials had a fixed set size and noise level with 6 interleaved values of target orientation (in steps of 1 or 0.5 octaves), each presented 20 times. Each block was repeated 4 to 6 times to yield 80 to 120 trials per orientation. The equivalent internal orientation noise was estimated in two ways. For both observers, we obtained an additional psychometric function for set size 1 with 10 target orientations (100 trials each). The best-fitting k value was used to determine the equivalent internal orientation noise, a fixed parameter in the model fits to each observer’s data. As an alternative measure to estimate the equivalent internal orientation noise for that task, which is a key parameter of the study, observer S.B. measured additional orientation thresholds at set size 1 at different orientation noise levels.

Results

Internal Noise Estimate

As we stated previously, the two models have a common first stage that produces independent and noisy responses. To provide an empirical estimate of this noise, we measured psychometric functions for a set size of 1 with no external noise for both observers. We then calculated the standard deviation of the Gaussian distribution underlying these functions, which was 1.35 for S.B. and 1.38 for V.A. We also measured psychometric functions with various amounts of added orientation noise for S.B. for the set size 1 condition. We then used the thresholds obtained with added noise (82% criterion) to estimate the equivalent internal orientation noise by performing an analysis similar to Pelli’s (1985) equivalent noise measurement for contrast. Orientation thresholds are plotted as a function of external orientation noise in Figure 5, where the points represent the data and the line represents the fit through the data calculated using the following equation:

$$\theta_{\text{thr}} = \frac{1.29}{k} \sqrt{\sigma_{\text{int}}^2 + \sigma_{\text{ext}}^2}$$

where $\sigma_{\text{int}}$ is internal variability and $\sigma_{\text{ext}}$ the orientation jitter that we added. The numerator is the d’ value for 82% correct. The text box in Figure 5 reports the actual estimate of $\sigma_{\text{int}}$, or equivalent internal orientation noise, for one observer. This value is almost identical to the estimated standard deviation of the psychometric function for set size 1 with no noise. More importantly, it was similar to the equivalent noise values estimated by reanalyzing thresholds measured across set sizes ranging from 1 to 16 using pixel contrast-modulated noise (Baldassi & Burr, 2000). The latter set of data showed that the internal noise does not change with set size, suggesting that the dominating source of internal noise arises locally from each element, rather than globally.
following the integration process (as in Morgan et al., 1998). This provides empirical support for the common first stage of the two models stated in the Models section. The value of $k$ estimated from internal noise was consistent with the estimate of $k$ from the Signed-Max model fits to the data. Therefore, we have converging evidence from various sources to justify the use of this noise estimate in the fits of the Sum model. It is reasonable to use the standard deviation at the set size 1 with no noise as a fixed parameter because this value seems to be independent of both the set size and the added external noise.

**Psychometric Function Analysis**

Panels A and B in Figure 6 report the data for the two observers, S.B. and V.A., and the fit of the two models, Signed-Max (Equation 2) and Sum (Equation 3). The variance of each observer's internal representation was set to be equal to the estimated variance for the set size 1 condition measured with no noise; under added noise conditions, we summed the internal and external noise variances. In both figures, each horizontal line of subplots shows a different set size, from 2 to 16, while the columns mark the orientation noise levels used in the experiment. The column labels represent the orientation noise, which was a multiple of the threshold for one element (1.35 for S.B., 1.38 for V.A.). In the added noise condition, the orientation jitter represented the standard deviation of the Gaussian distribution from which the actual orientation of each element, target and distractors, was drawn. The black symbols within each subplot represent the experimental data, whereas the smooth curves are the model fits for the Signed-Max and the Sum models, in solid blue and dashed red, respectively. The subplots with a slightly darker background indicate the conditions where the Sum model had a statistical advantage over the Signed-Max model. Statistical comparisons between the two models have been performed by comparing the weighted $\chi^2$ values for each model fit for each separate condition (subplot in Figure 6). As previously stated in more detail, we expect the two models to differ from each other in the way they fit the whole psychometric function across the conditions we explored. In particular, a different signature should be represented by diverging slopes of the fits as the set size increases. Indeed, this is what we observed for both our subjects. When the set size is small, 2 and 4, the two models yield similar trends to the fits, with virtually overlapping functions. When instead the set size increases up to 16, the overall picture is that psychometric functions have increasing slope consistent with the Signed-Max model, whereas the Sum model predicts shifts along the abscissa without changing slope. This is true for both observers, although it is more evident in V.A.’s data.

<table>
<thead>
<tr>
<th>Set Size</th>
<th>Noise</th>
<th>S.B. SignedMax</th>
<th>S.B. Sum</th>
<th>Difference</th>
<th>V.A. SignedMax</th>
<th>V.A. Sum</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0.764</td>
<td>0.689</td>
<td>-0.075</td>
<td>0.933</td>
<td>0.816</td>
<td>-0.117</td>
</tr>
<tr>
<td>0.5</td>
<td>1</td>
<td>1.433</td>
<td>1.25</td>
<td>-0.183</td>
<td>4.869</td>
<td>6.313</td>
<td>1.444</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>3.136</td>
<td>2.617</td>
<td>-0.519</td>
<td>2.433</td>
<td>2.167</td>
<td>-0.266</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>4.104</td>
<td>3.644</td>
<td>-0.46</td>
<td>1.056</td>
<td>0.939</td>
<td>-0.117</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>2.328</td>
<td>2</td>
<td>-0.328</td>
<td>0.947</td>
<td>1.97</td>
<td>1.023</td>
</tr>
<tr>
<td>0.5</td>
<td>1</td>
<td>1.298</td>
<td>1.383</td>
<td>0.085</td>
<td>1.623</td>
<td>1.298</td>
<td>-0.325</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>1.528</td>
<td>1.273</td>
<td>-0.255</td>
<td>0.495</td>
<td>0.412</td>
<td>-0.083</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>0.51</td>
<td>0.566</td>
<td>0.056</td>
<td>0.569</td>
<td>3.9</td>
<td>3.331</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>3.533</td>
<td>3.029</td>
<td>-0.504</td>
<td>2.428</td>
<td>2.017</td>
<td>-0.411</td>
</tr>
<tr>
<td>0.5</td>
<td>0.5</td>
<td>0.983</td>
<td>1.98</td>
<td>0.997</td>
<td>1.371</td>
<td>1.16</td>
<td>14.629</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1.854</td>
<td>3.544</td>
<td>1.69</td>
<td>0.615</td>
<td>6.82</td>
<td>6.205</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>1.173</td>
<td>66.2</td>
<td>65.027</td>
<td>19.3</td>
<td>484.286</td>
<td>464.986</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>1.031</td>
<td>0.99</td>
<td>-0.041</td>
<td>1.611</td>
<td>1.818</td>
<td>0.207</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0.327</td>
<td>23.125</td>
<td>22.798</td>
<td>2.13</td>
<td>3.529</td>
<td>1.399</td>
</tr>
<tr>
<td>0.5</td>
<td>1</td>
<td>0.286</td>
<td>1.711</td>
<td>1.425</td>
<td>1.03</td>
<td>670</td>
<td>668.97</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>1.211</td>
<td>1.663</td>
<td>0.452</td>
<td>0.697</td>
<td>38.286</td>
<td>37.589</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>1.393</td>
<td>1.875</td>
<td>0.482</td>
<td>2.827</td>
<td>43.25</td>
<td>40.423</td>
</tr>
</tbody>
</table>

Table 1. Weighted $\chi^2$ Values

Weighted $\chi^2$ values obtained by fitting Equations 2 and 3 (Signed-Max and Sum models, respectively) to the psychometric functions obtained by the two observers, S.B. and V.A. The difference column reports the difference between the Sum and the Signed-Max models; negative values indicate the advantage of the Sum (italic) and positive values of the Signed-Max (bold) model.
Figure 6. Individual psychometric functions plotting percent correct (on linear axes) versus angle (on log axes) for observers S.B. (A) and V.A. (B). Noise level increases along the rows from left to right, and set size increases along the columns from top to bottom. Each plot shows the model fits for the Sum model (red dashed curves) and for the Signed-Max model (blue continuous curves). Plots drawn on a gray background indicate conditions where the statistical comparison was in favor of the Sum model. For all the others, the Signed-Max model dominate.
Interestingly, the statistical comparison between the models shows a very slight advantage for the Sum model at low set sizes, where the two models do not effectively differ from each other. In this case, the advantage is numerically subtle and substantially null. When instead the two fits appear different, especially at high set sizes, then the advantage of the Signed-Max over the Sum model becomes much bigger, and even visual scrutiny shows that it clearly fits the data better.

This is evident in Table 1, which shows weighted $\chi^2$ values for the Signed-Max and the Sum models for the two observers. The estimate of $n$ averaged across noise conditions corresponds to the actual set size for observer S.B. for both the Signed-Max and the Sum models. Observer V.A.’s average estimate of $n$ reflects the number of elements in the display at low set sizes, but exceeds that number for the two larger set sizes used. While the uncertainty increases with set size, this increase is not proportional to the displayed set size. So it cannot simply be explained by an intrinsic uncertainty factor. This naïve observer exhibits non-optimal behavior only at larger set sizes.

**Set Size Functions**

The analysis of the entire psychometric function in the previous section shows an advantage for the Signed-Max model. How do the predictions of the model compare to the function relating thresholds to set size? Figure 7 plots the thresholds (just noticeable differences) as a function of the set size on loglog coordinates (dark blue circles) along with the predictions of the Sum and the Signed-Max models (red dotted and blue continuous line, respectively) and of the standard Max rule for tasks with target varying in a single direction (green dashed line) (e.g., Palmer, 1994). If we compare the Sum with the standard Max, it is clear that the former does a better job, even though some points appear to lie in between. However, the modified Max proposed in this study (i.e., the Signed-Max) is closer to the data than the other two models, especially at higher set sizes, where they generate clearly different predictions.

Even though the log-log slope of the threshold versus set size function is a standard measure for comparing models of visual search, plotting thresholds at a single criterion limits the scope of the possible conclusions. Indeed, the predictions of the two models for the psychometric functions sketched in Figure 4 show different behaviors at different criteria for thresholds. As the Sum model predicts a parallel shift of the whole function with increasing set size, the slope of the threshold versus set size functions should be independent of the criterion for threshold. However, the Signed-Max model predicts far apart psychometric functions at low criteria and converging functions at higher criteria, implying a change in the slope of the threshold versus set size function with changing criterion. Figure 8 plots the set size dependency, that is the log-log slope of the threshold versus set size function, at five different criteria for the two observers and compares them to the predictions of two models. Whereas the prediction for the Sum model is flat at a slope of 0.5, the prediction for Signed-Max ranges from about 0.65 to 0.2 with increasing criterion for threshold. The subjects’ data show a trend very similar to the predictions of the latter, with slopes ranging from 0.9 to 0.2 in the extreme cases. The systematic shift toward steeper slopes probably reflects the higher uncertainty associated with the two larger set sizes for observer V.A.

**Discussion**

We used psychometric function analysis to discriminate between different strategies of visual search in an identification task. Previous speculations showed the Sum and the Max rule to be the appropriate combination rules for multidimensional stimuli, one better than the other under different conditions (Graham et al., 1987). Indeed, the different integration rules we considered have proven to be optimal (or close to optimal) in different visual tasks with compound displays. The Sum model is the optimal rule, producing the best predicted performance, for the so-called summation tasks where all the elements are equally informative (Graham et al., 1987; Green & Swets, 1966; Verghese & Stone, 1995). It also seems to be implemented in crowded displays (Parkes, Lund, Angelucci, Solomon, & Morgan, 2001) and in averaging local signals for the extraction of orientation-defined textures (Dakin & Watt, 1997). But the Max rule works more efficiently when the task is to
detect a single target among distractors and when there is little or no interaction between local elements.

Our results (Figure 7) showed a strong set size effect on thresholds at all noise levels (greater than predicted by the standard Max rule), similar to the data of Baldassi and Burr (2000). This allowed us to compare the Baldassi and Burr (2000) model with a version of the Max model modified for this type of task and procedure, that is the Signed-Max model. Fitting individual psychometric functions by allowing the uncertainty \((n)\) and the gain \((k)\) parameters to vary allowed us to quantitatively compare the outcomes of the two models (the gain parameter \(k\) was free only in the Signed-Max model).

Figure 7 shows the advantage of the Max rule for a location task in a similar display (Solomon & Morgan, 2001).

The success of the Signed-Max model in the present form suggests that the outputs of independent detectors (at least two for each stimulus location) and a comparison of the respective outputs constitute the decision variable. Moreover, we propose a plausible basis for the mechanisms that regulate the local decision related to any single stimulus. In fact, even though the identification task has been used extensively in the orientation (e.g., Morgan & Baldassi, 1997) and in other domains (Solomon et al., 1997), we think previous accounts either neglected or underestimated its singular nature. For example, in the Monte-Carlo simulation of Morgan et al. (1998), the authors used the absolute values generated from \(n = \text{set size independent random variables}\) to generate predictions of the Max model. Even though it is computationally equivalent to our account, an absolute Max is not satisfying conceptually. In our account, each direction of tilt away from vertical activates two mirrored detectors whose activity is monitored by the observer and labeled. Actually, a standard Max rule takes place on either side and then the two maxima are compared, similar to the Max rule applied to 2IFC tasks, where a max is assumed to be extracted in each interval, and the two maxima are compared. We think that this labeling takes place in tasks such as location search and \(m\)-alternative forced choice (e.g., 10AFC, Solomon & Morgan, 2001). Here the outputs are monitored in the location rather than the orientation domain, and the location producing the highest response is eventually chosen. The equivalent of Morgan and colleagues’ (1998) absolute Max in the positional context would be to
collapse all the positional information onto a single abstract space and take the Max. If such a strategy were used in an mAFC task, then the Max response would have to be remapped back onto the original space. We think this is not biologically plausible, nor economical, and assume that such tasks are instead accomplished by labeled detectors. The model sketched in Figure 1 is biologically plausible and can be extended to other tasks, once the nature of the task and the behavior of front-end filters are taken into account.

Moreover, as our model takes into account the actual behavior of physiologically plausible orientation detectors, it calls into play an extra stage in between the detectors response and the decision variable, as sketched in Figure 1, that is not considered in different accounts of the same task (e.g., Carrasco, Penpeci-Talgar, & Eckstein, 2000). We started from orientation detectors whose response as a function of orientation is non-monotonic (as opposed to the monotonic contrast response function). By considering the response of far-apart detectors (±30º), we obtain responses that grow monotonically with tilt away from vertical over our stimulus range (±28º). Therefore, rather than modeling the response of matched filters with preferred orientation peaking at the physical orientation at any given trial and location, we compute the probability that the detector that matches the direction of the target stimulus tilt produces the greater response. This choice of signed detectors makes sense given the fact that 6 different orientations were equally possible on any trial: the respective responses of two broadly tuned detectors differ by increasing amounts for larger deviations from vertical.

Conclusions

Although both the Sum and Max rules are plausible integration rules for visual search, we have shown that a version of the Max rule better describes search performance in an identification task. A detailed comparison of these rules was achieved by fitting the model to entire psychometric functions; analysis of thresholds versus set size functions alone would have obscured such differences.

Acknowledgments

We thank Miguel Eckstein for an extensive discussion of the manuscript. This work was made possible by National Eye Institute Grant R01EY12038 to P.V. Commercial Relationships: None.

Footnote

1The Signed-Max model defined above produces the same fit to the data as Equation A3 in Carrasco et al., 2000. We chose a different exposition than Carrasco et al. to reflect biological plausibility, that is, cortical detectors do not usually produce modulations below baseline.

References


Baldassi & Verghese 570


