Supporting Appendix

Let $h_s(t)$ and f_s denote the hit rate and false alarm rate (respectively) for detection of the contrast decrement probe as a function of its duration t, when s items have been pre-cued (s is the set size; by convention, we use $s=0$ for minimal attention). f_s is the probability of responding "probe present" in the absence of a probe (which is independent of t).

This section describes how, using the psychometric functions for full attention ($h_1(t)$ and f_1) and minimal attention ($h_0(t)$ and f_0) as given, we compute predicted psychometric functions ($h_s(t)$ and f_s) for divided attention (set size s=2, 3, or 4).

The observed hit rate is itself subject to contamination by false alarms (i.e. correct detections are at least as likely as false alarms, even if detection always fails), so there must exist a "true" detection rate $h_s^*(t)$ such that

$$h_s(t) = f_s + (1-f_s) \,.\, h_s^*(t) \qquad \forall\, s \in \mathrm{N} \qquad (1)$$

Thus:

$$h_s^*(t) = (h_s(t)-f_s) \,/\, (1-f_s) \qquad \forall\, s \in \mathrm{N}. \qquad (2)$$

Parallel Model. The parameter of the parallel model is the division cost c ($c \in [0\,;\,1]$) incurred for each additional item to be attended. As the set size s increases, the true hit rate and the false alarm rate are assumed to converge geometrically (with a convergence factor (1-c)) from the "full attention" to the "minimal attention" performance values, at the same probe duration t. That is:

$$h_s^*(\mathrm{t}) = (1-c)^{s-1} \,.\, h_1^*(t) + (1-(1-c)^{s-1}) \,.\, h_0^*(t) \qquad \forall\, s \in \mathrm{N}^* \qquad (3)$$

$$f_s = (1-c)^{s-1} \,.\, f_1 + (1-(1-c)^{s-1}) \,.\, f_0 \qquad \forall\, s \in \mathrm{N}^*. \qquad (4)$$

Once $h_s^*(t)$ is derived by using Eq. **3**, the predicted hit rate $h_s(t)$ can be computed by using Eq. **1**, and compared to the experimentally observed hit rate.

Sample-When-Divided Model. The parameter of the sampling models is the sampling duration τ ($\tau \in \mathrm{R}^+$; however, we will only explore a "plausible" range of parameter values, i.e. between 50 and 1,000 ms). The underlying assumption behind this model is that each of the s potential targets will be sampled in turn by attention for a duration τ, and will then be sampled again after a period of $s\tau$. (We do not assume a particular sampling order in each trial, but only that this order remains constant throughout any given trial).

In each trial, the first "attentional sample" of the element containing the probe may start with a different delay δ ($\delta \in [0\,;\,s\tau]$) with respect to trial onset. (Because there is no transient at trial onset, it is reasonable to assume that the ongoing sampling would not be systematically reset at the beginning of each trial.)

We describe a probe stimulus of duration t as a "boxcar" function σ of trial time x ($x \in \mathrm{R}^+$):

$$\sigma(d,t,x) = 1 \quad \text{if } d < x \leq d+t, \qquad (5)$$
$$= 0 \quad \text{otherwise}$$

where d is the delay before the probe onset. (The outcome of our calculations is not affected by whether d is assumed to be constant across trials, or to vary randomly. In the actual experiment, it did vary randomly from trial to trial).

We describe the i^{th} attentional sample ($i \in \mathrm{N}^*$) of the element containing the probe as a "boxcar" function α_s of trial time x:

$$\alpha_s(i,\delta,x) = 1 \quad \text{if } \delta + (i\text{-}1).\, s\tau < x \leq \delta + (i\text{-}1).\, s\tau + \tau \qquad \forall\, s \in N^*. \tag{6}$$
$$= 0 \quad \text{otherwise}$$

Finally, we also describe the i^{th} period during which the element containing the probe is outside of the main focus of attention (i.e., attention is sampling the other elements) as a "boxcar" function β_s of trial time x:

$$\beta_s(i,\delta,x) = 1 \quad \text{if } \delta + (i\text{-}1).\, s\tau + \tau < x \leq \delta + is\tau \qquad \forall\, s \in N^*. \tag{7}$$
$$= 0 \quad \text{otherwise}$$

We derive the predicted hit rate $h_s(t)$ as:

$$h_s(t) = (1/s\tau).\int_0^{s\tau} 1\text{- } (1\text{-}f_s).\prod_{i=1}^{\infty} \{(1\text{-}h_1^*(\int_0^{\infty}\sigma(d,t,x)\ .\ \alpha_s(i,\delta,x)\ dx))\ .\ (1\text{-}h_0^*(\int_0^{\infty}\sigma(d,t,x)\ .\ \beta_s(i,\delta,x)\ dx))\}\ d\delta \qquad \forall\, s \in N^*. \tag{8}$$

In the previous equation, all periods of attentional sampling α_s are checked against the probe stimulus function σ to determine how much of the probe was sampled by attention (if any). The probability of detecting the probe at this duration is given by the function h_1^*. By construction, this duration cannot be longer than τ. Similarly, the periods β_s are used to determine for how long at a time the probe was shown outside of the main focus of attention. The probability of detecting the probe at this duration is given by the function h_0^*. By construction, this duration cannot be longer than $(s\text{-}1).\tau$. All these probabilities (as well as the probability of making a false alarm) are combined independently, yielding the final probe detection probability for a given trial. Lastly, these are integrated over all possible delays δ of the ongoing sampling, to predict the average probe detection probability $h_s(t)$ across trials.

For the calculation of the false alarm rate within this model, we reason that the average number of attentional samples in a given trial is a constant, and so an observer employing this strategy would be as likely to make a false alarm on any given "probe-absent" trial, independent of set size. In other words:

$$f_s = \text{constant} = f_1 \qquad \forall\, s \in N^* \tag{9}$$

Sample-Always Model. This model is identical to the previous one, except for the fact that the hit rate $h_1(t)$ measured at set size 1 is already the result of a sampling process. In other words, there must exist a function $\theta(t)$ (with $t \in [0\ ;\ \tau]$), representing the probability of detecting the probe when attention actually samples it for a duration of t. From that function, we will be able to derive the predicted hit rates $h_s(t)$ as:

$$h_s(t) = (1/s\tau).\int_0^{s\tau} 1\text{- } (1\text{-}f_s).\prod_{i=1}^{\infty} \{(1\text{-}\theta(\int_0^{\infty}\sigma(d,t,x)\ .\ \alpha_s(i,\delta,x)\ dx))\ .\ (1\text{-}h_0^*(\int_0^{\infty}\sigma(d,t,x)\ .\ \beta_s(i,\delta,x)\ dx))\}\ d\delta \qquad \forall\, s \in N. \tag{10}$$

In other words, $\theta(t)$ plays the role that $h_1^*(t)$ was playing in the previous model.

We derive θ from observing $h_1(t)$ in the following way. For s=1, all attentional samples are directed to the probe location (so $\beta_1 = 0$, and the corresponding term in Eq. **10** disappears).

- For $t \in [0\ ;\ \tau]$, the probe will either be sampled by one or two consecutive attentional samples, depending on the sampling delay δ. Thus, we can write:

$$h_1(t) = 1\text{-}(1\text{-}f_1).(1/\tau).\{\int_0^{t}(1\text{-}\theta(x))\ .\ (1\text{-}\theta(t\text{-}x))\ dx + \int_t^{\tau} 1\text{-}\theta(t)\ dx\}$$

$$= 1-(1-f_1).(1/\tau).\{\int_0^t (1-\theta(x)) . (1-\theta(t-x))\, dx + (\tau-t).(1-\theta(t))\} \qquad \forall\, t \in [0;\tau]. \qquad (11)$$

The first and second terms in the curly braces correspond, respectively, to trials in which 2 or 1 attentional sample(s) will contain the probe.

- For probe durations $t+n\tau$ ($t \in [0;\tau]$, $n \in$ N*) longer than one sampling period, the probe will either be sampled by $(n+1)$ or $(n+2)$ samples (all but 2 of which, the first and last, will be "full"), depending on the sampling delay δ. Thus, we can write:

$$h_1(t+n\tau) = 1-(1-f_1).(1/\tau).\{(1-\theta(\tau))^n.\int_0^t (1-\theta(x)) . (1-\theta(t-x))\, dx + (1-\theta(\tau))^{n-1}.\int_t^\tau (1-\theta(x)) . (1-\theta(\tau+t-x))\, dx \}$$
$$\forall\, t \in [0;\tau],\ \forall\, n \in \text{N*}. \qquad (12)$$

We determined θ numerically by increments, using Eq. **11**. Starting up with $\theta(0)=0$, we considered the first data point t_1 of the experimentally obtained function $h_1(t)$. Assuming a linear progression of θ between 0 and t_1, we derived $\theta(t_1)$ from $h_1(t_1)$ and Eq. **11** (because of the linear assumption, this amounts to solving a 2nd degree equation in $\theta(t_1)$). Then, assuming a linear progression of θ between t_1 and the second data point t_2, we derived $\theta(t_2)$ from $h_1(t_2)$ and Eq. **11** in a similar way, and so on until the derivation of $\theta(\tau)$.

Additionally, from Eq. **12** we can get:

$$1-h_1(t+n\tau) = (1-\theta(\tau))^{n-1} . (1-h_1(t+\tau))$$
$$\Leftrightarrow \quad \theta(\tau) = 1-\{(1- h_1(t+n\tau)) / (1- h_1(t+\tau))\}^{1/(n-1)} \qquad \forall\, t \in [0;\tau],\ \forall\, n > 1 \qquad (13)$$

This provides another way of computing $\theta(\tau)$, using data points of h_1 higher than 2τ (if any are available, i.e. for $\tau<0.5$s). When possible, we applied both methods and took $\theta(\tau)$ to be the mean of the two values hence obtained.

We then used $\theta(t)$ ($t \in [0;\tau]$) to compute, according to Eq. **10**, the predicted $h_s(t)$ under conditions of divided attention ($s\geq2$), which we finally compared with the experimentally observed data, as done for the other two models.