Supplementary Material

The model presented in this article is an extension of the model presented by Petersen et al. (2012). For simplicity, Petersen et al. assumed that the attentional weight of a mask, $w_M$, was equal to the attentional weight of a target, $w_T$, and that the processing rate of a mask, $v_M$, was equal to the processing rate of a target, $v_T$. Here, we extend the model by proposing that targets and masks may have different values of attentional weight and that the processing rate of a mask may deviate from the processing rate of a target. These changes allow the model to account for the data in the three conditions described in this article: one in which T1 and T2 are masked (Difficult condition), one in which T1 is unmasked and T2 is masked (Easy condition), and one in which the mask of T1 is presented without T1 and T2 is masked (Mask condition). Figure S1 and Figure 5 shows the behavior of the model in the Difficult and Easy conditions for a variety of parameter values. In the following, the assumptions and the mathematical details behind the model will be described. We have chosen to primarily explain how the performance on T2 is modeled since T2 is of most interest in relation to the arguments put forward in the article.

Redistribution of attentional resources

When an object (i.e., T1, T2, mask of T1) is presented on a previously empty location, we assume that a redistribution of the attentional resources (neurons) will occur. The times of the redistributions are referred to as $t_{01}$ (i.e., $t_0$ for T1), $t_{02}$ (i.e., $t_0$ for T2) and $t_{01M1}$ (i.e., $t_0$ for the mask of T1), respectively. However, if a target is postmasked (i.e., the mask is presented at the same location as the target) only the presentation of the target will initiate a redistribution of attentional resources. Thus, no redistribution occurs when the mask of T2 is presented since it is always preceded by T2. Furthermore, we assume that $t_{01}$, $t_{02}$, and $t_{01M1}$ are approximately normally distributed with a common standard deviation $\sigma_0$ and means $\mu_0$, $\text{SOA} + \mu_0$, and $\tau_1 + \mu_0$, respectively, where $\tau_1$ is the exposure duration of T1 and SOA is the stimulus onset asynchrony between T1 and T2.
Probability of encoding T2

The probability of encoding T2 \( (p_{T2}) \) depends on the order by which the redistributions (i.e., \( t_{01} \) and \( t_{02} \)) occur and if they occur before or after the presentation of the masks (i.e., at \( \tau_1 \) and SOA + \( \tau_2 \)). Five conditions exist: Both redistributions occur before \( \tau_1 \) (i.e., \( t_{01} \leq t_{02} \leq \tau_1 \)), both redistributions occur after \( \tau_1 \) (i.e., \( \tau_1 \leq t_{01} \leq t_{02} \)), the redistribution initiated by the presentation of T1 (Difficult and Easy conditions) or the mask of T1 (Mask condition) occurs before \( \tau_1 \) whereas the redistribution initiated by T2 occurs after \( \tau_1 \) (i.e., \( t_{01} \leq \tau_1 \leq t_{02} \)), and the redistribution initiated by the presentation of T1 (Difficult and Easy condition) or the mask of T1 (Mask condition) occurs after the redistribution initiated by T2 (i.e., \( t_{01} > t_{02} \)) and either before the presentation of the mask of T2 (i.e., \( t_{01} \leq \text{SOA} + \tau_2 \)) or after (i.e., \( t_{01} > \text{SOA} + \tau_2 \)). That is, \( p_{T2} \) is found as a sum of five probabilities:

\[
p_{T2} = p[T2&(t_{01} \leq t_{02} \leq \tau_1)] + p[T2&(\tau_1 \leq t_{01} \leq t_{02})] + p[T2&(t_{01} \leq \tau_1 \leq t_{02})] + p[T2&(t_{01} > t_{02})&(t_{01} \leq \text{SOA} + \tau_2)] + p[T2&(t_{01} > t_{02})&(t_{01} > \text{SOA} + \tau_2)]
\]

(S1)

In the first four conditions, the presentation of T1 affects the processing rate of T2. However, in the last condition, the presentation of T1 does not affect the processing rate of T2, resulting in a rate of \( C \) from \( t_{02} \) until SOA + \( \tau_2 \). Thus, the five probabilities are given by,

\[
p[T2&(t_{01} \leq t_{02} \leq \tau_1)] = \int_{-\infty}^{\tau_1} \frac{1}{\sigma_0} \phi\left(\frac{t_{01} - \mu_0}{\sigma_0}\right) \left( \int_{t_{01}}^{\tau_1} \frac{1}{\sigma_0} \phi\left(\frac{t_{02} - (\text{SOA} + \mu_0)}{\sigma_0}\right) \right)
\]

\[
(1 - e^{-(\tau_2 / (t_{01} \leq t_{02} \leq \tau_1)) / (\text{SOA} + \tau_2)} \text{dt}_{02}) \text{dt}_{01}
\]

(S2)

\[
p[T2&(\tau_1 \leq t_{01} \leq t_{02})] = \int_{\tau_1}^{\text{SOA} + \tau_2} \frac{1}{\sigma_0} \phi\left(\frac{t_{01} - \mu_0}{\sigma_0}\right) \left( \int_{t_{01}}^{\text{SOA} + \tau_2} \frac{1}{\sigma_0} \phi\left(\frac{t_{02} - (\text{SOA} + \mu_0)}{\sigma_0}\right) \right)
\]

\[
(1 - e^{-(\tau_2 / (\tau_1 \leq t_{01} \leq t_{02})) / (\text{SOA} + \tau_2)} \text{dt}_{02}) \text{dt}_{01}
\]

(S3)
\[ p[T2 \&(t_{01} \leq \tau_1 \leq t_{02})] = \int_{-\infty}^{\tau_1} \frac{1}{\sigma_0} \phi\left(\frac{t_{01} - \mu_0}{\sigma_0}\right) \left( \int_{\tau_1}^{\infty} \frac{1}{\sigma_0} \phi\left(\frac{t_{02} - (SOA + \mu_0)}{\sigma_0}\right) dt_{02} \right) dt_{01} \quad (S4) \]

\[ (1 - e^{-\nu_{T2}(t_{01} \leq \tau_1 \leq t_{02})(SOA + \tau_2 - t_{02})}) dt_{02} \]

\[ p[T2 \&(t_{01} > t_{02}) \&(t_{01} \leq SOA + \tau_2)] = \int_{-\infty}^{\infty} \frac{1}{\sigma_0} \phi\left(\frac{t_{02} - (SOA + \mu_0)}{\sigma_0}\right) \left( \int_{t_{02}}^{\infty} \frac{1}{\sigma_0} \phi\left(\frac{t_{01} - \mu_0}{\sigma_0}\right) dt_{01} \right) dt_{02} \quad (S5) \]

\[ (1 - e^{-C(t_{01} - t_{02}) - \nu_{T2}(t_{01} > t_{02})(SOA + \tau_2 - t_{01})}) dt_{01} \]

\[ p[T2 \&(t_{01} > t_{02}) \&(t_{01} > SOA + \tau_2)] \]

\[ = (1 - \Phi\left(\frac{SOA + \tau_2 - \mu_0}{\sigma_0}\right)) \int_{-\infty}^{\infty} \frac{1}{\sigma_0} \phi\left(\frac{t_{02} - (SOA + \mu_0)}{\sigma_0}\right) \left( \int_{t_{02}}^{\infty} \frac{1}{\sigma_0} \phi\left(\frac{t_{01} - \mu_0}{\sigma_0}\right) dt_{01} \right) dt_{02} \quad (S6) \]

where \( \nu_{T2} \) is the processing rate of T2. The next section explains how the processing rate of T2 is calculated in the first four conditions.

**Processing rate of T2**

**Condition 1.** If \( t_{01} \leq t_{02} \leq \tau_1 \) (i.e., both redistributions occur before \( \tau_1 \)), the processing rate of T2 is given by,

\[ \nu_{T2}(t_{01} \leq t_{02} \leq \tau_1) = C\left(\frac{w_{T2}}{\sum_{z \in S} w_z} x p_{\text{free}}(t_{01} \leq t_{02} \leq \tau_1)\right) \]

\[ + 0 \times p_{\text{locked}}(t_{01} \leq t_{02} \leq \tau_1) + 1 \times p_{\text{released}}(t_{01} \leq t_{02} \leq \tau_1) \quad (S7) \]

where

\[ \frac{w_{T2}}{\sum_{z \in S} w_z} = \begin{cases} \frac{w_{T2}}{w_{T1} + w_{T2}} = \frac{1}{2} & \text{Difficult and Easy conditions} \\ \frac{w_{T2}}{w_{M1} + w_{T2}} = \frac{1}{\alpha + 1} & \text{Mask condition} \end{cases} \quad (S8) \]

and \( p_{\text{free}} \) is the proportion of neurons that become distributed according to the attentional weights, \( p_{\text{locked}} \) is the proportion of neurons that remain locked to T1, and \( p_{\text{released}} \) is the proportion of neurons that have been
released from T1 and are exclusively available for T2. The relative attentional weight of a mask is denoted \( \alpha = w_{T1}/w_{T2} \).

\( p_{\text{free}}, p_{\text{locked}}, \) and \( p_{\text{released}} \) may also represent the probabilities with which a single neuron is found in the three stages. Thus, the above proportions can be derived by defining the behavior of a single neuron. The times it takes to lock and release a single neuron are assumed to be exponentially distributed with rate parameters \( \lambda_1 \) and \( \lambda_d \), respectively. If we furthermore assume that the locking process starts immediately after an object has been encoded into VSTM and the subsequent release process starts directly after a neuron has been locked in a feedback loop, the three proportions (probabilities) are given by

\[
p_{\text{free}}(t_{01} \leq t_{02} \leq \tau_1) = 1 - \int_{t_{01}}^{t_{02}} v_{T1} e^{-v_{T1}(t-t_{01})} \left( \int_t^{t_{02}} \lambda_1 e^{-\lambda_1 (t'-t)} \, dt' \right) \, dt
\]

\[
= e^{v_{T1}(t_{01}-t_{02})} + \frac{v_{T1}}{\lambda_1 - v_{T1}} \left( e^{v_{T1}(t_{01}-t_{02})} - e^{\lambda_1(t_{01}-t_{02})} \right)
\]

\[
p_{\text{locked}}(t_{01} \leq t_{02} \leq \tau_1) = \int_{t_{01}}^{t_{02}} v_{T1} e^{-v_{T1}(t-t_{01})} \left( \int_t^{t_{02}} \lambda_1 e^{-\lambda_1 (t'-t)} (1 - \int_t^{t_{02}} \lambda_d e^{-\lambda_d (t''-t')} \, dt'') \, dt' \right) \, dt
\]

\[
= \frac{\lambda_1}{\lambda_d - \lambda_1} \left( e^{v_{T1}(t_{01}-t_{02})} - e^{\lambda_1(t_{01}-t_{02})} \right)
\]

\[
- \frac{v_{T1}}{\lambda_d - v_{T1}} \left( e^{v_{T1}(t_{01}-t_{02})} - e^{\lambda_d(t_{01}-t_{02})} \right)
\]

\[
p_{\text{released}}(t_{01} \leq t_{02} \leq \tau_1) = \int_{t_{01}}^{t_{02}} v_{T1} e^{-v_{T1}(t-t_{01})} \left( \int_t^{t_{02}} \lambda_1 e^{-\lambda_1 (t'-t)} \left( \int_t^{t_{02}} \lambda_d e^{-\lambda_d (t''-t')} \, dt'' \right) \, dt' \right) \, dt
\]

\[
= 1 - (p_{\text{free}}(t_{01} \leq t_{02} \leq \tau_1) + p_{\text{locked}}(t_{01} \leq t_{02} \leq \tau_1))
\]

where \( v_{T1} = C \) in all three test conditions (i.e., Difficult, Easy and Mask conditions), \( t \) is the time when the letter identity of T1 (or the mask for T1) is encoded into VSTM, \( t' \) is the time when a neuron representing the encoded identity of T1 is locked by becoming embedded in a positive feedback loop gated by the unit representing T1 in the VSTM map, and \( t'' \) is the time when the feedback loop is broken such that the neuron is released. At any given time, a neuron is in just one of the three states, so \( p_{\text{free}} + p_{\text{locked}} + p_{\text{released}} = 1 \).
Note, that in the Mask condition \( t_{01} (t_{01M1}) \) is normally distributed with mean \( \tau_1 + \mu_0 \) and \( \nu_{T1} = \nu_{M1} = C \).

**Condition 2.** If, however, \( \tau_1 \leq t_{01} \leq t_{02} \) (i.e., both redistributions occur after \( \tau_1 \)) the processing rate of T2 is given by,

\[
\nu_{T2}|(\tau_1 \leq t_{01} \leq t_{02}) = C(W_{T2}^{*} \times p_{\text{free}})|(\tau_1 \leq t_{01} \leq t_{02}) + 0 \times p_{\text{locked}}|(\tau_1 \leq t_{01} \leq t_{02}) + 1 \times p_{\text{released}}|(\tau_1 \leq t_{01} \leq t_{02})
\]

(S12)

where

\[
\frac{W_{T2}^{*}}{\sum_{x \in S} W_x} = \begin{cases} \frac{W_{T2}}{W_{T1} + W_{T2}} = \frac{1}{2} & \text{, Difficult condition} \\ \frac{W_{T2}}{W_{T2}} = 1 & \text{, Easy condition} \\ \frac{W_{T2}}{W_{M1} + W_{T2}} = \frac{1}{\alpha + 1} & \text{, Mask condition} \end{cases}
\]

(S13)

and

\[
p_{\text{free}}|(\tau_1 \leq t_{01} \leq t_{02}) = 1 - \int_{t_{01}}^{t_{02}} \nu_{M1} e^{-\nu_{M1}(t-t_{01})} \left( \int_{t}^{t_{02}} \lambda_1 e^{-\lambda_1(t-t')} dt' \right) dt
\]

\[
= e^{\nu_{M1}(t_{01}-t_{02})} + \frac{\nu_{M1}}{\lambda_1 - \nu_{M1}} \left( e^{\nu_{M1}(t_{01}-t_{02})} - e^{\lambda_1(t_{01}-t_{02})} \right)
\]

(S14)

\[
p_{\text{locked}}|(\tau_1 \leq t_{01} \leq t_{02})
\]

\[
= \int_{t_{01}}^{t_{02}} \nu_{M1} e^{-\nu_{M1}(t-t_{01})} \left( \int_{t}^{t_{02}} \lambda_1 e^{-\lambda_1(t-t')} \left( 1 - \int_{t'}^{t_{02}} \lambda_4 e^{-\lambda_4(t-t'')} dt'' \right) dt' \right) dt
\]

\[
= \frac{\lambda_1}{\lambda_1 - \nu_{M1}} \left( e^{\nu_{M1}(t_{01}-t_{02})} - e^{\lambda_1(t_{01}-t_{02})} \right)
\]

\[
- \frac{\nu_{M1}}{\lambda_1 - \nu_{M1}} \left( e^{\nu_{M1}(t_{01}-t_{02})} - e^{\lambda_4(t_{01}-t_{02})} \right)
\]

(S15)

\[
p_{\text{released}}|(\tau_1 \leq t_{01} \leq t_{02}) = \int_{t_{01}}^{t_{02}} \nu_{M1} e^{-\nu_{M1}(t-t_{01})} \left( \int_{t}^{t_{02}} \lambda_1 e^{-\lambda_1(t-t')} \left( \int_{t'}^{t_{02}} \lambda_4 e^{-\lambda_4(t-t'')} dt'' \right) dt' \right) dt
\]

\[
= 1 - (p_{\text{free}}|(\tau_1 \leq t_{01} \leq t_{02}) + p_{\text{locked}}|(\tau_1 \leq t_{01} \leq t_{02}))
\]

(S16)

where the processing rate of the mask is given by,
\[ v_{M1} = \begin{cases} 0, & \text{Easy condition} \\ C, & \text{Difficult and Mask conditions} \end{cases} \] (S17)

**Condition 3.** Further, if \( t_{01} \leq \tau_1 \leq t_{02} \) (i.e., the redistribution initiated by the presentation of T1 (Difficult and Easy conditions) or the mask of T1 (Mask condition) occurs before \( \tau_1 \) whereas the redistribution initiated by T2 occurs after \( \tau_1 \)), the processing rate of T2 is given by,

\[ v_{T2}|(t_{01} \leq \tau_1 \leq t_{02}) = C\left(\frac{w_{T2}}{\sum w_z} \times p_{\text{free}}|(t_{01} \leq \tau_1 \leq t_{02})\right) + 0 \times p_{\text{locked}}|(t_{01} \leq \tau_1 \leq t_{02}) + 1 \times p_{\text{released}}|(t_{01} \leq \tau_1 \leq t_{02}) \] (S18)

where \( \frac{w_{T2}}{\sum w_z} \) is given by Equation S13 and

\[ p_{\text{free}}|(t_{01} \leq \tau_1 \leq t_{02}) = \int_{t_{01}}^{\tau_1} v_{T1} e^{-v_{T1}(t-t_{01})} \left(1 - \int_t^{t_{02}} \lambda_1 e^{-\lambda_1(t'-t)} \, dt'\right) \, dt + e^{-v_{T1}(\tau_1-t_{01})} \left(\int_{t_{01}}^{\tau_1} v_{M1} e^{-v_{M1}(t-\tau_1)} \left(\int_t^{t_{02}} \lambda_1 e^{-\lambda_1(t'-t)} \, dt'\right) \, dt \right) \]

\[ = e^{v_{T1}(t_{01}-\tau_1)+v_{M1}(\tau_1-t_{02})} + \frac{v_{T1}}{\lambda_1 - v_{M1}} \left(e^{v_{T1}(t_{01}-\tau_1)+\lambda_1(t_{01}-t_{02})} - e^{\lambda_1(t_{01}-t_{02})}\right) + \frac{v_{M1}}{\lambda_1 - v_{M1}} \left(e^{v_{T1}(t_{01}-\tau_1)+v_{M1}(\tau_1-t_{02})} - e^{v_{T1}(\tau_1-t_{01})-\lambda_1(\tau_1-t_{02})}\right) \] (S19)
\[ p_{\text{locked}}(t_{01} \leq \tau_1 \leq t_{02}) \]

\[ \begin{aligned}
&= \int_{t_{01}}^{t_1} v_{T1} e^{-v_{T1}(t-t_{01})} \left( \int_t^{t_{02}} \lambda_1 e^{-\lambda_1(t'-t)} \left( 1 - \int_{t'}^{t_{02}} \lambda_d e^{-\lambda_d(t''-t')} dt'' \right) dt' \right) dt \\
&\quad + e^{-v_{T1}(t_1-t_{01})} \int_{t_1}^{t_{02}} v_{M1} e^{-v_{M1}(t-t_1)} \left( \int_t^{t_{02}} \lambda_1 e^{-\lambda_1(t'-t)} \left( 1 - \int_{t'}^{t_{02}} \lambda_d e^{-\lambda_d(t''-t')} dt'' \right) dt' \right) dt \\
&\quad - \int_{t'}^{t_{02}} \lambda_d e^{-\lambda_d(t''-t')} dt'' dt' \\
&= \frac{\lambda_1}{\lambda_d - \lambda_1} \left( \frac{v_{T1}}{\lambda_1 - v_{T1}} \left( e^{v_{T1}(t_{01}-\tau_1)} + \lambda_1(t_{01}-t_{02}) - e^{\lambda_1(t_{01}-t_{02})} \right) \right. \\
&\quad \left. + \frac{v_{M1}}{\lambda_1 - v_{M1}} \left( e^{v_{T1}(t_{01}-\tau_1)} + \lambda_1(t_{01}-t_{02}) - e^{\lambda_1(t_{01}-t_{02})} \right) \right) \\
&\quad \left. - \frac{v_{T1}}{\lambda_d - v_{T1}} \left( e^{v_{T1}(t_{01}-\tau_1)} + \lambda_d(t_{01}-t_{02}) - e^{\lambda_d(t_{01}-t_{02})} \right) \right) \\
&\quad \left. - \frac{v_{M1}}{\lambda_d - v_{M1}} \left( e^{v_{T1}(t_{01}-\tau_1)} + \lambda_1(t_{01}-t_{02}) - e^{\lambda_d(t_{01}-t_{02})} \right) \right)
\end{aligned} \tag{S20}

\[ p_{\text{released}}(t_{01} \leq \tau_1 \leq t_{02}) \]

\[ \begin{aligned}
&= \int_{t_{01}}^{t_1} v_{T1} e^{-v_{T1}(t-t_{01})} \left( \int_t^{t_{02}} \lambda_1 e^{-\lambda_1(t'-t)} \left( 1 - \int_{t'}^{t_{02}} \lambda_d e^{-\lambda_d(t''-t')} dt'' \right) dt' \right) dt \\
&\quad + e^{-v_{T1}(t_1-t_{01})} \int_{t_1}^{t_{02}} v_{M1} e^{-v_{M1}(t-t_1)} \left( \int_t^{t_{02}} \lambda_1 e^{-\lambda_1(t'-t)} \left( 1 - \int_{t'}^{t_{02}} \lambda_d e^{-\lambda_d(t''-t')} dt'' \right) dt' \right) dt \\
&\quad - 1 - (p_{\text{free}}(t_{01} \leq \tau_1 \leq t_{02}) + p_{\text{locked}}(t_{01} \leq \tau_1 \leq t_{02}))
\end{aligned} \tag{S21}

In Condition 3, \( v_{T1} = C \) in all three test conditions (i.e., Difficult, Easy and Mask conditions) and \( v_{M1} \) is given by Equation S17.

**Condition 4.** In Conditions 1-3, we assumed that \( t_{01} \leq t_{02} \) which is most likely to happen when the SOA is long. If however the SOA is short, the distribution of \( t_{02} \) overlaps the distribution of \( t_{01} \), and it becomes likely that \( t_{01} > t_{02} \). In this case, T2 will have a head start and be assigned a processing rate of \( C \) in the interval between \( t_{02} \) and \( t_{01} \). After \( t_{01} \), the processing rate of T2 will be
\[ \nu_{T2}(t_{01} > t_{02}) = C \frac{w_{T2}}{w_{T1} + w_{T2}} \]  

(S22)

**Probability of encoding T1**

The probability of encoding T1 \((p_{T1})\) is found by a similar calculation to that of T2 and given by the following sum,

\[
p_{T1} = p[T1&(t_{02} \leq t_{01})&(t_{02} \leq \tau_1)] + p[T1&(t_{02} > t_{01})&(t_{02} \leq \tau_1)] + p[T1&(t_{02} > t_{01})&(t_{02} > \tau_1)]
\]

(S23)

**Guessing**

In both the current experiment and the experiment by Moore et al. (1996), participants were forced to respond, if necessary by guessing. To account for the guessing, we used a high-threshold guessing model, which assumes that participants report the identity of a target correctly if the target becomes encoded into VSTM, but guess at random among the \(N\) alternatives if the target fails to become encoded into VSTM. Formally, the adjusted probability of correctly reporting a target \(T\) using this guessing model can be defined as

\[
p_T^{adj}(SOA) = p_T(SOA) + (1 - p_T(SOA)) \frac{1}{N}
\]

(S24)

In the experiment by Moore et al. (1996), only two letter types for T1 and two digit types for T2 were used. Therefore, we used \(N = 2\) when the model was fitted to these data. In our own experiment, 26 different letter types were used for both T1 and T2, and thus \(N = 26\) was used when the model was fitted to the data.
Figure S1. Predicted probabilities of correctly reporting T1 ($p_{T1}$, dashed lines) and T2 ($p_{T2}$, solid lines) as functions of SOA in the Difficult (gray) and Easy (black) conditions for three different values of $C$: 20 Hz (thin lines), 40 Hz (medium size lines), and 80 Hz (thick lines) and three different values of $\mu$ (Panel A): 100 ms (thin lines), 180 ms (medium size lines), and 260 ms (thick lines); $\mu_d$ (Panel B): 500 ms (thin lines), 300 ms (medium size lines); $\mu_0$ (Panel C): 10 ms (thin lines), 20 ms (medium size lines), and 30 ms (thick lines); and $\sigma_0$ (Panel D): 10 ms (thin lines), 20 ms (medium size lines), and 30 ms (thick lines). In each panel, the parameters not varied were fixed at $\mu_l = 180$ ms, $\mu_d = 300$ ms, $\mu_0 = 20$ ms, $\sigma_0 = 20$ ms, and $\mu_{\text{decay}} = 200$ ms. For comparison, Figure 5 illustrates the effect of changing $C$ keeping all other parameters fixed.