Appendix A

In this section, we detail the equations we used to carry out MLE of the nROUSE and diffusion race model parameters. For the nROUSE model, three parameters were allowed to vary: a temporal attention coefficient, the inhibition term, and the noise multiplier. All other values were fixed to the previously stated default values. To determine the likelihood of the data given the current parameters, the nROUSE model is first simulated. Let F_k and T_k represent the identification latencies generated by simulating the nROUSE for the foil and target, respectively, in condition k. Furthermore, let ν represent the noise multiplier. Then the probability of correctly picking the target is

$$\alpha_k = 1 - \Phi(0|F_k - T_k, \nu[F_k^2 + T_k^2]), \tag{A.1}$$

where $\Phi(x|\mu, \sigma^2)$ is the Gaussian cumulative distribution function given an independent variable x and the distribution mean μ and variance σ^2. Let $Y = \{y_1, ..., y_K\}$ be the number of correctly identified targets for each of the K conditions, and $N = \{n_1, ..., n_K\}$ represent the corresponding total number of trials. Finally, let Θ represent the set of parameters for the nROUSE model. The likelihood function for the nROUSE model is then

$$f_n(Y|N, \Theta) = \prod_{k=1}^{K} \left[\binom{n_k}{y_k} \alpha_k^{y_k} (1 - \alpha_k)^{n_k - y_k} \right], \tag{A.2}$$

where $\binom{n}{y}$ is the binomial coefficient.

The diffusion race model assumes that for each racer, evidence accumulates towards a threshold according to a one-boundary Wiener process. As noted before, this means the finishing times for the racer follow an inverse Gaussian distribution. For a finishing time $T = t$ the cumulative distribution function of the inverse Gaussian is

$$F(t|\kappa, \xi) = \Phi\left(\frac{\kappa}{\sqrt{t}} \left[\frac{\xi t}{\kappa} - 1 \right] \right) + \exp\{2\xi\kappa\} \Phi\left(-\frac{\kappa}{\sqrt{t}} \left[\frac{\xi t}{\kappa} + 1 \right] \right), \tag{A.3}$$

where Φ is the cumulative distribution function for the standard normal distribution, and κ is a threshold towards which evidence accumulates with an average rate of ξ. For identification purposes, we fix the variance of the evidence accumulation (σ^2; the coefficient of drift) to 1, allowing it to be factored out of the equation. Note that equation A.3 is defined only for positive, non-zero values of t and κ. For simplicity, we also assume ξ can only be positive and non-zero. The probability density function is

$$f(t|\kappa, \xi) = \frac{\kappa}{\sqrt{2\pi t^3}} \exp\left\{ -\frac{1}{2t} (\kappa - \xi t)^2 \right\}. \tag{A.4}$$

With equations A.3 and A.4, we can now define the joint density function for the diffusion race model. Let $Y = y$ be an observed choice, where $Y = \{0, 1\}$. Furthermore, let $T = t$ now represent an observed response time. This response

time is the sum of the decision time (i.e., the finishing time for the fastest racer), and a non-decision component τ encapsulating the duration of mechanisms such as the motor response and stimulus encoding. The joint likelihood of observing t and y is the likelihood of the finishing time $t - \tau$ weighted by the probability that the racer for the alternate choice has yet to finish. Therefore, the joint density is

$$\begin{aligned} f_d(y, t|\kappa_1, \kappa_0, \xi_1, \xi_0, \tau) = & \, y f(t - \tau|\kappa_1, \xi_1)[1 - F(t - \tau|\kappa_0, \xi_0)] + \\ & (1 - y) f(t - \tau|\kappa_0, \xi_0)[1 - F(t - \tau|\kappa_1, \xi_1)], \end{aligned} \tag{A.5}$$

where $\{\xi_0, \xi_1\}$ are the rates of evidence accumulation and $\{\kappa_0, \kappa_1\}$ are the thresholds corresponding to choices 0 and 1 respectively. The weights y and $1-y$ ensure that only the likelihood for the appropriate finishing time is included, thereby conditioning on choice.

Appendix B

In this section, we present additional figures with the distribution of performance measures for subjects over conditions. Figure B.1 presents boxplots for the proportion correct across subject, shown for each condition. The endpoints of the boxplots were constructed by taking the 5% and 95% quantiles across subjects. The inner boxes were then based on the first and third quartiles, and the center line represents the median. For completeness, the observations that lay outside the endpoints of the boxplots are shown as points. For ease of comparison, the order of the conditions for the same-different task are mirrored relative to the forced-choice two-alternative task. When the correct choice was on the left or was 'same', the boxes are colored as dark gray. When the correct choice was on the right or was 'different', the boxes are colored as light gray.

Figure B.2 and B.3 present boxplots for the median response times across subject, shown for each condition and state of accuracy. Figure B.1 is specific to the two-alternative forced-choice task, while Figure B.2 is specific to the same-different task. Again, endpoints are the 5% and 95% quantiles, and individual points are median times that fell outside of the endpoints. The inner boxes still represent the first and third quartiles, with the center line representing the median of the distribution. The distributions for median times of errors are marked in light gray, and distribution for correct responses are marked in dark gray. Target primed conditions are shown on the left, while foil primed conditions are shown on the right, mirrored relative to the target primed conditions. The duration of the prime is given at the bottom of the plot. The type of correct response for each condition is given at the top of the plot, where for Figure B.1 'Lt' refers to left and 'Rt' refers to right, and for Figure B.2 'S' refers to a 'same' response and 'D' refers to a different response.

Figure B.4 presents the individual scatterplots per subjects showing the relationship between the perceptual identification speeds predicted by the nROUSE model (x-axis) and the drift rates for picking 'same' (y-axis) over the combination of each prime duration, type, and racer type (target versus foil). Each panel also included the best-fitting simple linear regression line and the associated Pearson's R.

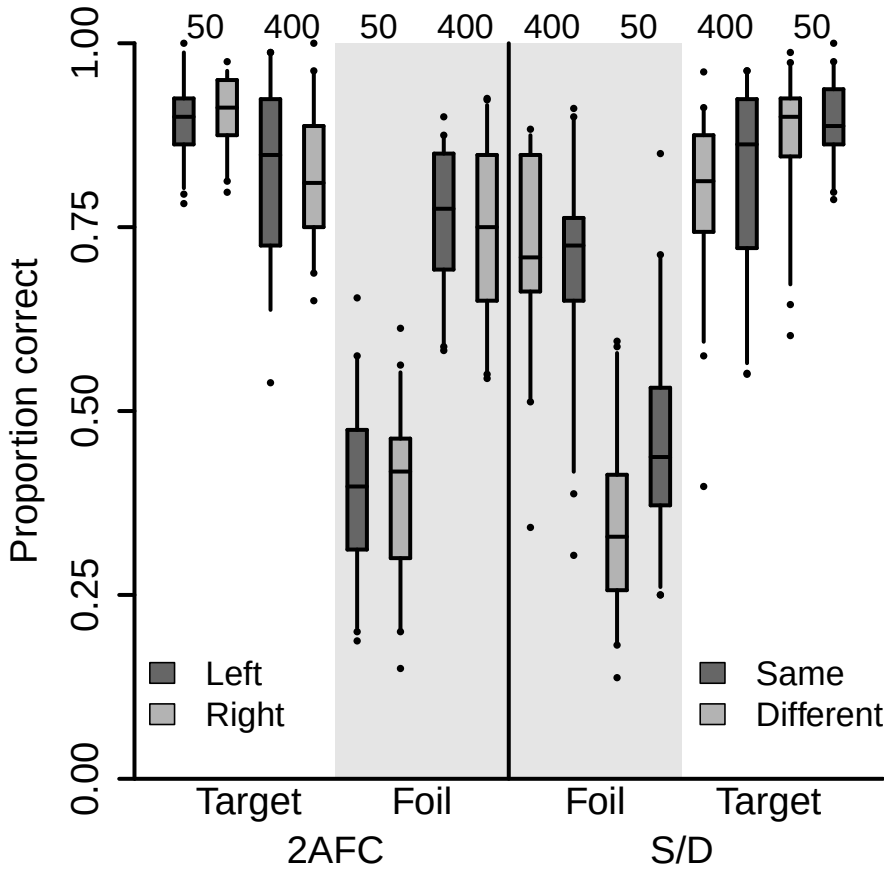

Figure B.1: Distributions across subjects for the proportion correct by conditions. Endpoints are the 5% and 95% quantiles, and the points are observations that fell outside of these endpoints. Inner boxes are the 1^{st} and 3^{rd} quartiles, and the center line is the median. The label '2AFC' refers to the two-alternative forced-choice task, while the label 'S/D' refers to the same-different task. The type of prime is given on the bottom of the plot, while the duration of the prime is given at the top. The color denotes which choice was correct in the given condition.

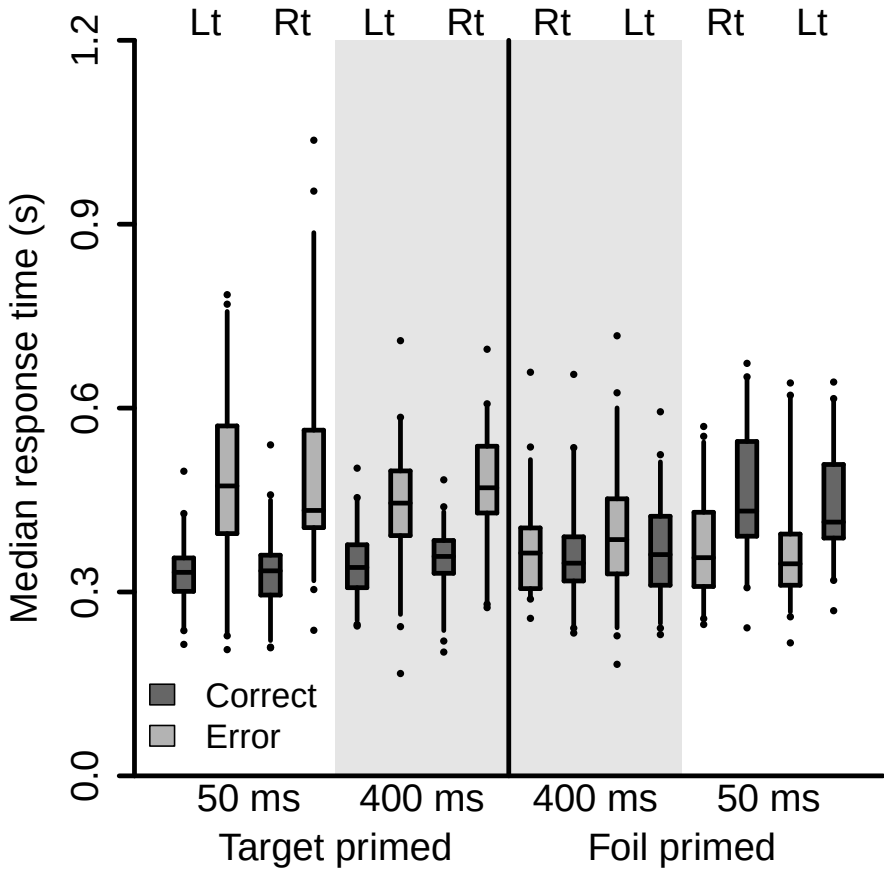

Figure B.2: Distributions across subjects for the median response by forced-choice two-alternative task conditions and state of accuracy. Endpoints are the 5% and 95% quantiles, and the points are observations that fell outside of these endpoints. Inner boxes are the 1^{st} and 3^{rd} quartiles, and the center line is the median. Target primed conditions are shown on the left, foil primed conditions are shown on the right, mirrored relative to the target primed conditions. The duration of the prime is given on the bottom of the plot, while the position of the correct response is given at the top ('Lt' is left and 'Rt' is right). The color denotes whether the median times are for errors or correct responses.

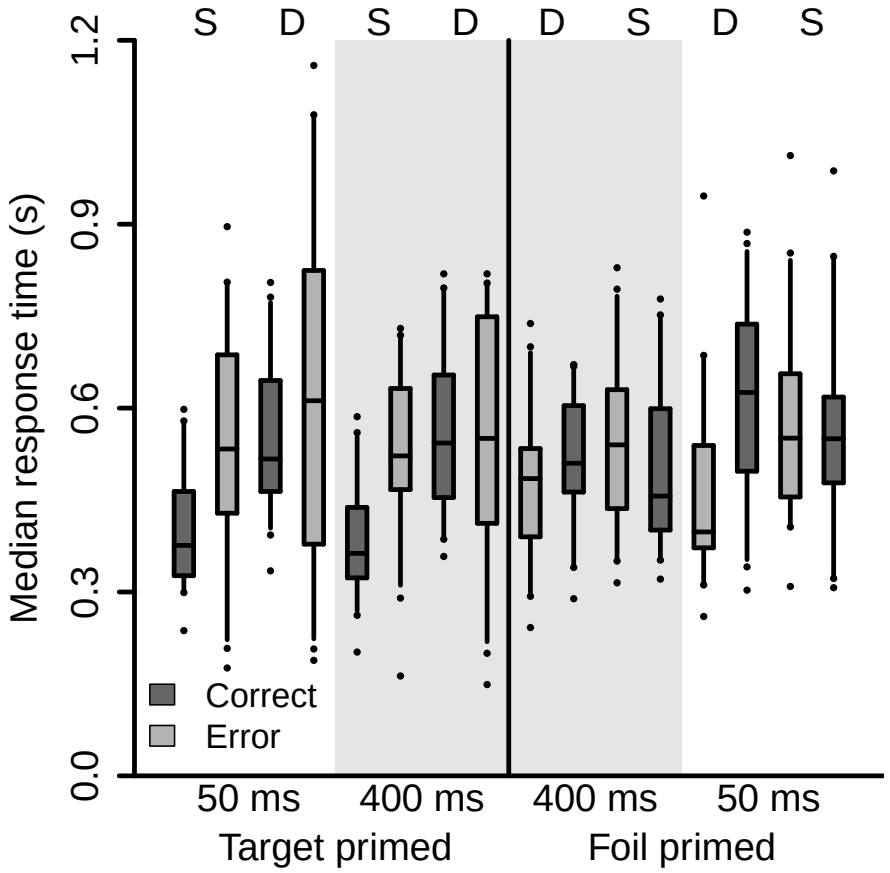

Figure B.3: Distributions across subjects for the median response by same-different task conditions and state of accuracy. Endpoints are the 5% and 95% quantiles, and the points are observations that fell outside of these endpoints. Inner boxes are the 1^{st} and 3^{rd} quartiles, and the center line is the median. Target primed conditions are shown on the left, foil primed conditions are shown on the right, mirrored relative to the target primed conditions. The duration of the prime is given on the bottom of the plot, while the type of correct response is given at the top ('S' is a 'same' response and 'Rt' is a 'different' response). The color denotes whether the median times are for errors or correct responses.

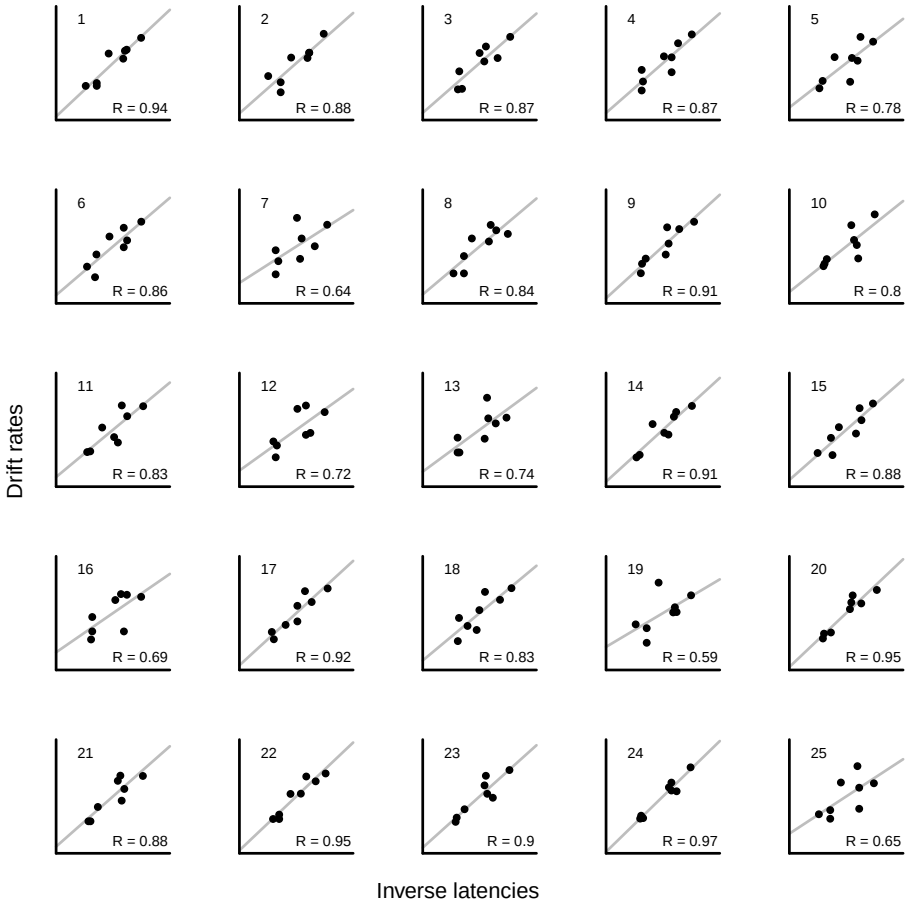

Figure B.4: Scatterplots of the relationship between the 8 inverse perceptual identification latencies (on the x-axis) and the 8 drift rates (on the y-axis) for picking 'same', shown separately for each subject. For each plot, the best fitting simple regression line is included, along with the associated value for Pearson's R. The subject ID number is shown at the top left.

Appendix C

In this section we report additional details regarding our data trimming approach. Trimming outliers can be potentially problematic, as researchers can (either intentionally or unintentionally) select exclusion criteria that strengthen evidence for their hypothesized findings (an issue typically labeled as 'researcher degrees of freedom').

We explored three different trimming approaches for our data. Initially, we simply used a set of global cut-offs. However, as noted in our main text, subjects varied greatly in terms of their overall speed on the task. The global cut-offs were only effective in excluding outliers for the slowest and fastest of subjects. The remaining outliers resulted in poor model convergence. Models either would not converge, or would get stuck in local maxima. Our second approach involved incorporating a mixture with a uniform density into the sequential sampling models we explored. This approach worked much better than the global cut-offs, but we still had convergence issues with estimates of the mixture probability. For convenience, we therefore used the two-step approach discussed in our main text. The mixture probabilities were easier to estimate due to the simpler, descriptive response time model we used. However, note that in in an ideal situation, our second approach would be preferable. Fortunately, regardless of the data trimming approach we used, our main findings never changed.

As noted in our main text, we first removed any responses faster than 200 ms, which excluded 0.27% of our data. We also excluded any responses slower than 2362 ms, which excluded an additional 0.11% of our data. Finally, the descriptive mixture model identified an additional 0.41% of the data as being more likely under the uniform density. Therefore, in total we excluded only 0.79% of our data (i.e., less than 1%).