Supplementary Material

Curve fitting Analysis. To compare the dynamics of upright and inverted face encoding, a shifted exponential function was fitted to the averaged *d*' values at each time point (Wickelgren & Corbett, 1977):

$$\hat{d}' = A\left\{1 - \exp\left[-R(t - I)\right]\right\} \text{ for } t > I\text{, otherwise } \hat{d}' = 0. \qquad (1)$$

Here, A is the asymptote, R is the rate of approach to asymptote, I is the intercept, and t is the stimulus encoding duration. The inverse of the rate, $1/R$, is expressed in seconds.

The eight possible models derived from this exponential function, which differed in the number of free parameters, were fitted to the 16 (2 stimulus types x 8 presentation durations) data points, d_i. More specifically, the models differed in whether the intercepts, rates of approach to the asymptote, or the asymptotes themselves were the same or different for the upright and inverted curves fitted to the observed data. This was notated with a 1 or 2 respectively (e.g. 2I, 1R, 2A). The least constrained model (6 parameters; 2I, 2R, 2A), allowed different intercepts, rates of approach to the asymptote, and asymptotes for the performance functions for upright and inverted faces, while the most constrained model possible (3 parameters; 1I, 1R, 1A), assuming equivalent intercept, rate, and asymptotes for upright and inverted faces. Fits were calculated for the average sensitivity across participants.

Data fitting was implemented in Matlab (Mathworks) using a simplex hill-climbing function to iteratively adjust parameters to maximize the value of the variance accounted for (r^2). The range of values from which the starting point of the curve fitting functions could be selected were limited as follows: intercept, 0.01 - 0.5, growth rate 0.01 – 10.0, asymptote, 0.1 - 5.0). To increase the plausibility of the resulting best fit models, the fitting procedure was restricted so that the asymptote parameter for inverted faces could not exceed that for upright faces. Two thousand iterations were performed per model to fit the data. Goodness of fit was assessed by calculating r^2 values (equation 2).

$$r^2 = 1 - \frac{\sum_{i=1}^{N}\left(d_i - \hat{d}_i\right)^2}{\sum_{i=1}^{N}\left(d_i - \bar{d}\right)^2} \tag{2}$$

where (d_i) refers to the observed data point i, $\hat{d}_i$ is the value predicted by equation 1, and $\bar{d}$ is the overall mean.

Nested models (whose parameters are proper subsets or super-sets) were statistically compared with an *F*-test comparing a full (greater number free parameters) to a reduced model (fewer number of free parameters) as follows in equation 3 (Lu & Dosher, 1998):

$$F(df_1, df_2) = \frac{\left(r_{full}^2 - r_{reduced}^2\right)/df_1}{\left(1 - r_{full}^2\right)/df_2} \tag{3}$$

where $df_1 = k_{\text{full}} - k_{\text{reduced}}$, and $df_1 = N - k_{\text{full}}$. The *k* refers to the number of free parameters in each model, and *N* is the number of predicted data points. This F-test incorporates an adjustment for the number of free parameters. The resulting F-values were converted to a t-value to evaluate the nested models. This conversion was done so that a one-tailed test could be applied. A one-tailed test is most appropriate because, logically, the fit of the nested model (i.e. the model with less free parameters) must be lower than, or at best equal to, that for the less constrained model.

Goodness of fit was also evaluated by comparing—between models— the $r^2_{adjusted}$, which is the proportion of variance accounted for after it is adjusted for the number of free parameters, *k* (Reed, 1976).

$$r^2 = 1 - \frac{\sum_{i=1}^{N}\left(d_i - \hat{d}_i\right)^2 / (N-k)}{\sum_{i=1}^{N}\left(d_i - \bar{d}\right)^2 / (N-1)} \tag{4}$$

where *N* is the number of data points (d_i), $\hat{d}_i$ is the value predicted by equation 1, *k* is the number of free parameters, and $\bar{d}$ is the overall mean.

<u>Experiment 1: Curve-fitting Results</u>

Curve fitting. A curve was fitted to the average sensitivity (d') measures for each group for each of the eight possible models. The data were best described by a model in which the asymptote and intercept, but not the rate of approach to asymptote, differed for upright and inverted faces (Figure S1). More specifically, constraining the rate parameter so that it was the same for both upright and inverted faces did not result in a significant drop in the variance accounted for across the two conditions (2I, 1R, 2A, $r^2 = .9772$), relative to the full model (2I, 2R, 2A, $r^2 = .9813$), $t(10)=1.48$, $p=.084$. However, constraining the intercept parameter (1I, 2R, 2A, $r^2 = .9698$), $t(10)=2.48$, $p=.0164$, or the asymptote (1I, 2R, 2A, $r^2 = .8953$), $t(10)=6.78$, $p<.0001$, so they were the same for both upright and inverted faces did result in a significantly worse fit relative to the full model. The best-fitting model not only confirmed the considerable difference in the asymptote level of performance between the two groups (.92 d' difference), but it also suggests that the onset of recognition performance for upright faces occurs approximately 33 ms before that for inverted faces. However, we find no evidence for a difference in the rate of perceptual encoding for upright and inverted faces.

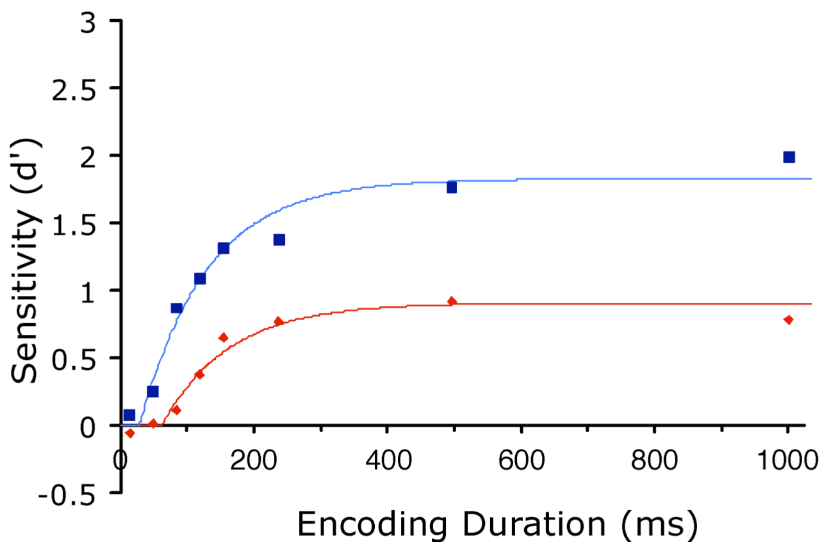

Figure S1. The performance (d') for upright and inverted faces in Experiment 1 was best described by a shifted exponential function that equated the rate of approach to asymptote (9.86), but had different performance onsets (intercepts; upright, 27 ms; inverted, 60 ms) and asymptote performance levels (upright, d'=1.83; inverted, .91) for upright and inverted faces (r^2 = .9772, $r^2_{adjusted}$ = .9688).

Experiment 2: Curve-fitting Results

Curve fitting. A curve was fit to the average sensitivity (d') measures from the expert and novice groups for each of the eight possible models. Constraining the asymptote (2I, 2R, 1A, r^2 = .9747), t(10)=4.99, p=.00025, resulted in a significantly worse fit relative to the full model (2I, 2R, 2A, r^2 = .9928). Constraining the intercept parameter (1I, 2R, 2A, r^2 = .9920), t(10)=1.02, p=.167, or the rate parameter (2I, 1R, 2A, r^2 = .9912), t(10)=1.46, p=.087, so that they were the same for both upright and inverted faces did not significantly affect the fit relative to the full model. Notably, the intercept and rate parameters appeared to trade off in their ability to account for the difference in performance between car experts and novices: constraining both the rate and intercept parameters within the same model (1I, 1R, 2A, $r^2 = .9830$), $t(11)=3.51$, $p=.0025$, resulted in a significant impact on the fit relative to the full model.

On closer inspection it appeared that the trade-off between the rate and intercept parameters may be occurring between different stages in the time-course of object processing. Specifically, the intercept-constrained model could better account for the longer encoding durations while the rate constrained model could better account for the shorter encoding durations. A similar reduction in fit to the early data points was observed for the intercept-constrained model (1I, 2R, 2A) in Experiment 1. Given that the focus of our study was on the time course of perceptual encoding, the curves were re-fitted only to the shorter encoding durations where performance was still rising sharply (12 ms, 48 ms, 83 ms, 118 ms, and 153 ms) to see if the rate- and intercept-constrained models could be distinguished on the basis of their ability to account for performance when encoding duration was more limited. The curve resulting from the model in which the rate was controlled, so that it could not vary across the expert and novice groups (2I, 1R, 2A, $r^2 = .9877$), fitted equally well as the curve resulting from the full model (2I, 2R, 2A, $r^2 = .9878$), $t(4)=.17$ $p=.436$. However, the reduction in the fit of the curve that assumed that the intercept was the same for car experts and novices (1I, 2R, 2A, $r^2 = .9783$), relative to the curves that assumed all three parameters differed across the group, approached significance, $t(4)=1.77$ $p=.076$. In addition, to achieve this fit, the intercept-constrained model equated the asymptote parameters for experts and novices, which is *not* consistent with the behavioral performance of the two groups[1]. Moreover, once the fits for the models were adjusted for the number of free parameters ($r^2_{adjusted}$), the model in which the intercept and asymptote varied across experts and novices, but not the rate, was associated with the highest $r^2_{adjusted}$ value, even exceeding that for the full model. This suggests that the onset of performance for car experts starts approximately 43 ms earlier than that among novices (Figure S2). It also reveals that asymptotic level of performance among experts is 1.18 d' units greater than that among novices.

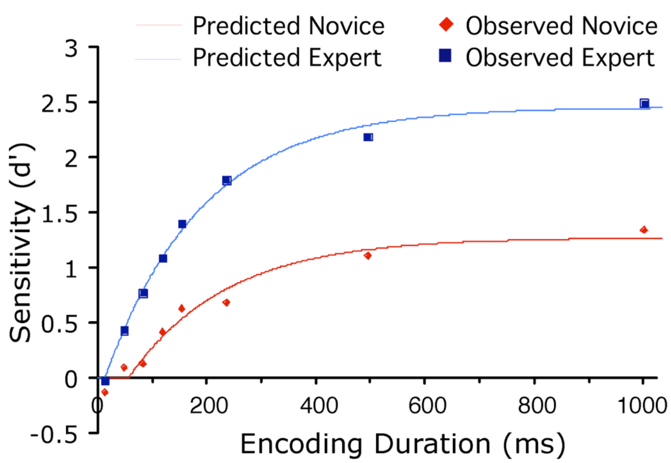

Figure S2. The performance (d') for cars among car experts and car novices in Experiment 2 was best described by a shifted exponential function that equated the rate of approach to asymptote (5.59), but had different performance onsets (intercepts; expert, 12 ms; novice, 55 ms) and asymptote performance levels (expert, d'=2.46; novice, 1.28) for car expert and car novice performance ($r^2 = .9912$, $r^2_{adjusted} = .9880$).

---

[1] Notably, the equivalence of the asymptote parameters for the functions describing expert and novice performance was a consequence of the restriction placed on the fitting procedure that the novice asymptote parameter could not exceed that for experts. Therefore, if the fitting procedure were restricted so that the asymptote parameter for experts actually *exceeded* that for novices, as is clearly the case in the observed data, the fit would presumably be reduced further.

Experiment 3: Curve-fitting Results

Curve fitting. A curve was fit to the average sensitivity (d') measures from the upright (Expriment 2) and inverted conditions for each of the eight possible models. Constraining the asymptote (2I, 2R, 1A, $r^2 = .8743$), $t(10)=13.172$, $p=.036$, resulted in a significantly worse fit relative to the full model (2I, 2R, 2A, $r^2 = .9490$). Constraining the intercept parameter (1I, 2R, 2A, $r^2 = .9375$), $t(10)=1.42$, $p=.199$, or the rate parameter (2I, 1R, 2A, $r^2 = .9360$), $t(10)=1.51$, $p=.186$, so that they were the same for both upright and inverted cars did not affect the fit relative to the full model. Constraining both the rate and intercept parameters within the same model (1I, 1R, 2A, $r^2 = .9347$), $t(11)=1.123$, $p=.248$, also failed to significantly impact the fit relative to the full model. This suggests that the onset of performance for upright and inverted cars is not significantly different, both starting at approximately 57-ms (Figure S3). However, the asymptotic level of performance for upright cars is .58 d' units greater than that for inverted cars.

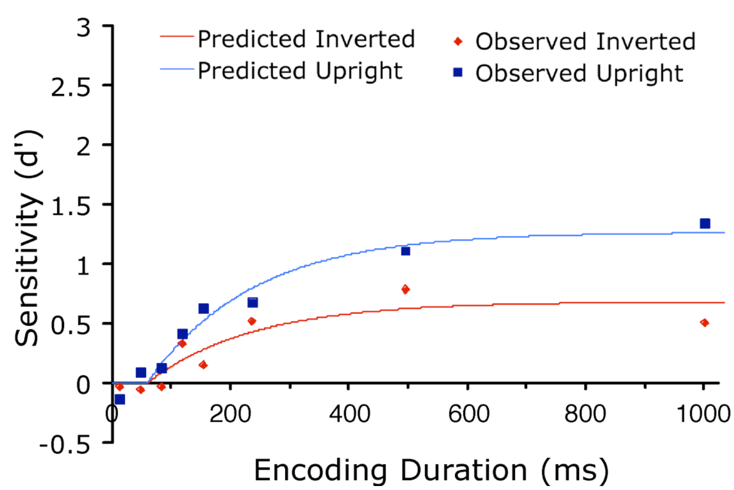

Figure S3. The performance (d') for upright and inverted cars among car novices in Experiment 3 was best described by a shifted exponential function that equated the rate of approach to asymptote (5.59) performance onsets (intercept; 57 ms), but had different asymptote performance levels (upright, d'=1.27; inverted, .67) for upright and inverted car performance ($r^2 = .9347$, $r^2_{adjusted} = .9020$).

References

Lu, Z-L., & Dosher, B. A. (1998). External Noise Distinguished Attention Mechanisms. *Vision Research*, 38(9), 1183-1198.

Reed, A. V. (1976). List length and the time course of recognition in immediate memory. *Memory and Cognition, 4*(1), 16-30.

Wickelgren, W. A., & Corbett, A. T. (1977). Associative interference and retrieval dynamics in yes-no recall and recognition. *Journal of Experimental Psychology: Human Learning and Memory, 3*(2), 189-202.