# SUPPLEMENTARY MATERIALS
# Some Tasks Demands Induce Collapsing Bounds: Evidence from a Behavioral Analysis

James J. Palestro[a], Emily Weichart[a], Per B. Sederberg[a], Brandon M. Turner[a,*]

[a]*Department of Psychology, The Ohio State University*

**Abstract**

Traditional models of choice response time assume that sensory evidence accumulates for choice alternatives until a threshold amount of evidence has been obtained. Although some researchers have characterized the threshold as varying randomly from trial to trial, these investigations have all assumed that the threshold remains fixed across time within a trial. Despite decades of successful applications of these models to a variety of experimental manipulations, the time-invariance assumption has recently been called into question, and a time-variant alternative implementing collapsing decision thresholds has been proposed instead. Here, we investigated the fidelity of the collapsing threshold assumption by assessing relative model fit to data from a highly constrained experimental design that coupled a within-subject mixture of two classic response time paradigms – interrogation and free response – within a random dot motion (RDM) task. Overall, we identified strong evidence in

*Corresponding Author

*Email address:* turner.826@gmail.com (Brandon M. Turner)

This research was supported by Air Force Research Lab contract FA8650-16-1-6770.

favor of collapsing decision thresholds, suggesting that subjects may adopt a dynamic decision policy due to task characteristics, specifically to account for the mixture of response time paradigms and motion strengths across trials in the mixed response signal task. We conclude that time-variant mechanisms may serve as a viable explanation for the strategy used by human subjects in our task.

*Keywords:* fixed boundaries, collapsing boundaries, likelihood-free methods, random dot motion, interrogation

---

Here, we present additional details, tables, figures, and analyses not included in the main text. In the first section, we provide details about the evolution of evidence in the diffusion decision model (DDM; Ratcliff, 1978; Ratcliff and Rouder, 1998). Second, we present the a table that displays the arrangement of parameters featured in each of the 16 model variants under consideration in our comparison. Next, we present the unstandardized (i.e., raw) model fit statistics. In the fourth section, we present an additional model comparison analysis using approximated posterior model probabilities and an assessment of model fit to observed data using bar plots. The fifth section provides the numerical approximated Bayes factor scores for each subject. The sixth section provides additional analyses exploring the impact of practice on model fits. The final section provides the inferred task representation plots for the remaining 12 subjects not included in the main text.

## 1. The Diffusion Decision Model

The diffusion decision model (DDM; Ratcliff, 1978; Ratcliff and Rouder, 1998) assumes that, on the presentation of a stimulus $S$ (e.g., a random dot display), an observer accumulates sensory evidence as noisy samples and integrates these samples over the course of a trial. The integration process can be described in terms of the differential equation

$$de\left(t\right) = vdt + sdW, \tag{1}$$

where $e\left(t\right)$ is the value of the integrated evidence at time $t$, $v$ is the mean of the accumulated samples (i.e., drift rate), and $s$ represents the moment-to-moment variability in the drift rate (Ratcliff, 1978; Ratcliff and Rouder, 1998;

Turner et al., 2017). The parameter $v$ dictates the strength of evidence, so in the dot motion task, $s$ would indicate the strength of coherence toward one direction or another. However, due to some perceptual biases, the motion coherence may not map directly onto the diffusion process in Equation 1. Denoting the coherence variable as $C$, and $\zeta$ as the coherence scaling factor,

$$v = C\zeta \ .$$

The values of $C$ can be positive, arbitrarily indicating strength of motion to the right, or negative, indicating motion to the left. If we assume that no decision bound interferes with the integration process (e.g., in the interrogation paradigm), Equation 1 dictates how observers integrate evidence from the initial presentation of the stimulus until the time it is removed from the screen. If $W$ in Equation 1 denotes a Wiener process, as it does in the DDM (see Stone, 1960; Smith, 2000, for a detailed overview), the sensory evidence $e(t)$ is assumed to be sampled continuously over time from a normal distribution such that $e(t) \sim N\left(\mu\left(t\right), \sigma\left(t\right)\right)$ with mean $\mu(t)$ standard deviation $\sigma(t)$, such that

$$\begin{aligned} \mu(t) &= vt \\ \sigma(t) &= s\sqrt{t}. \end{aligned} \tag{2}$$

When modeling data from an interrogation paradigm, the typical approach is to compare the state of evidence at the interrogation time $t$ to some criterion $c(t)$. Given our choices about how positive evidence relates to rightward motion, a rightward response would be generated if $e(t) > c(t)$. Otherwise, a leftward response would be made.

The DDM includes several sources of variability beyond simple sensory integration that complicate Equation 1. These extra sources of variability have proven effective in accounting for a number of empirical benchmarks in decision-making tasks (Ratcliff and McKoon, 2008; Ratcliff and Rouder, 1998; Ratcliff and Tuerlinckx, 2002). The first source of variability is the within-trial variability term $\sigma_w$ already discussed in Equation 1. The second source of variability is trial-to-trial (i.e., between-trial) variability in the drift rate, which results from variability in the mean stimulus value of the stimuli used in the task. In line with Equation 1, this variability in implemented by perturbing the trial-specific value of $\mu$, typically from a normal distribution such that $\mu \sim \mathcal{N}(v, \eta)$. Here, the standard deviation parameter $\eta$ represents the between-trial variability in drift rate. The third source of variability is variability in the starting point of the diffusion process. This source is typically thought of as being independent of the sensory properties of the stimulus itself, and instead is internal to the observer. Starting point variability is implemented in a similar manner as the between-trial variability in drift: on each trial, the initial evidence value is perturbed by a normal distribution centered at $z_0$ with a standard deviation $S_0$, such that $e(t = 0) \sim \mathcal{N}(z_0, S_0)$. Between-trial variability in the starting point allowed the diffusion model to account for the relative speeds of errors and correct responses in the SAT manipulation (e.g., Ratcliff et al., 2006; Starns and Ratcliff, 2010; White et al., 2009). Considering these three sources of variability, the mean $\mu(t)$ and standard deviation $\sigma(t)$ from Equation 2 becomes

$$\begin{aligned} \mu(t) &= vt + z_0 \\ \sigma(t) &= \sqrt{S_0^2 + s^2 t + \eta^2 t^2}. \end{aligned} \tag{3}$$

The relative importance of each source of variability is investigated in the main text.

## 2. Model Variants

The arrangement of parameters featured in each of the 16 variants is displayed in Table 1. Model variants were determined by the type of boundary (time-invariant versus time-variant) and the combination of freely estimated and fixed parameters. If a parameter is freely estimated in the model, it is denoted as "FE" in the table. If the parameter is fixed, it is denoted as "F." If a parameter is absent from the model (e.g, the parameters associated with the collapsing boundary that are not included in the time-variant models), the space is left blank. This model order (1-16) will remain consistent in the following sections, and models will be referred to by their model number denoted in Table 1.

## 3. Raw Model Fit Statistics

Tables 1 and 2 provide the model fit statistics on their original scale for each of the 16 models and each subject. The best fitting model for each subject is in boldface type. For all subjects, the best fitting model is a variant of the time-variant diffusion model with collapsing boundaries, with Class 3 ($\lambda$ free) accounting for the most subjects overall. The second best class of models was Class 4, which freely estimated both $a'$ and $\lambda$, best accounting for 5 out of 14 subjects. The remaining two subjects were best accounted for by either Class 1 (i.e., the time-invariant class) or Class 2, which freely estimated $a'$.

Table 1: **Model Variants.** Each row corresponds to a specific model we fit to data, where parameters (i.e., columns) were either freely estimated (FE) or fixed (F). If a parameter is not applicable, such as in the time-invariant models, the space is left blank.

| Model | Decision Boundaries | $a$ | $v$ | $z_0$ | $s$ | $\tau_{er}$ | $S_\tau$ | $S_0$ | $\eta$ | $a'$ | $\lambda$ | k | Free Parameters |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Fixed | FE | FE | FE | F | FE | FE | F | F | | | | 5 |
| 2 | Fixed | FE | FE | FE | F | FE | FE | F | FE | | | | 6 |
| 3 | Fixed | FE | FE | FE | F | FE | FE | FE | F | | | | 6 |
| 4 | Fixed | FE | FE | FE | F | FE | FE | FE | FE | | | | 7 |
| 5 | Collapsing | FE | FE | FE | F | FE | FE | F | F | FE | F | F | 6 |
| 6 | Collapsing | FE | FE | FE | F | FE | FE | F | FE | FE | F | F | 7 |
| 7 | Collapsing | FE | FE | FE | F | FE | FE | FE | F | FE | F | F | 7 |
| 8 | Collapsing | FE | FE | FE | F | FE | FE | FE | FE | FE | F | F | 8 |
| 9 | Collapsing | FE | FE | FE | F | FE | FE | F | F | F | FE | F | 6 |
| 10 | Collapsing | FE | FE | FE | F | FE | FE | F | FE | F | FE | F | 7 |
| 11 | Collapsing | FE | FE | FE | F | FE | FE | FE | F | F | FE | F | 7 |
| 12 | Collapsing | FE | FE | FE | F | FE | FE | FE | FE | F | FE | F | 8 |
| 13 | Collapsing | FE | FE | FE | F | FE | FE | F | F | FE | FE | F | 7 |
| 14 | Collapsing | FE | FE | FE | F | FE | FE | F | FE | FE | FE | F | 8 |
| 15 | Collapsing | FE | FE | FE | F | FE | FE | FE | F | FE | FE | F | 8 |
| 16 | Collapsing | FE | FE | FE | F | FE | FE | FE | FE | FE | FE | F | 9 |

Table 2: **Model fit statistics for Subjects 1-7.** boldface type represents the lowest BIC value for each subject.

| Model | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| 1 | 1049.27 | 1042.37 | 1093.19 | 1008.76 | 803.54 | 919.73 | 982.01 |
| 2 | 953.62 | 928.95 | 1134.93 | 863.97 | 784.63 | 807.44 | 986.10 |
| 3 | 885.27 | 932.26 | 781.04 | 824.48 | 801.99 | 839.36 | 910.01 |
| 4 | 1001.95 | 908.22 | 1046.31 | 891.85 | 775.69 | 791.2 | 945.51 |
| 5 | 926.87 | 920.86 | 991.25 | 874.36 | 798.97 | 867.23 | 913.18 |
| 6 | 889.56 | 1134.45 | 1040.52 | 761.77 | 762.4 | 732.06 | 832.11 |
| 7 | 833.55 | 930.09 | 780.56 | 821.79 | 767.78 | 833.68 | 892.91 |
| 8 | 890.09 | 925.9 | 786.16 | 894.18 | 781.46 | 844.88 | 912.16 |
| 9 | 1028.06 | 1028.89 | 882.41 | 786.15 | 772.47 | 775.89 | 739.83 |
| 10 | 783.61 | 902.80 | 1041.80 | **712.44** | 772.06 | **668.09** | 729.12 |
| 11 | 769.1 | 889.06 | 772.27 | 791.82 | 779.11 | 754.18 | **715.78** |
| 12 | **763.61** | 972.59 | 841.95 | 822.66 | 793.02 | 805.73 | 793.89 |
| 13 | 773.29 | 1055.65 | 930.19 | 842.7 | 780.21 | 808.74 | 971.92 |
| 14 | 788.95 | **864.94** | 900.79 | 731.19 | **759.90** | 725.36 | 806.56 |
| 15 | 784.72 | 922.83 | **769.63** | 805.99 | 789.51 | 783.16 | 827.23 |
| 16 | 801.43 | 970.06 | 808.69 | 849.1 | 788.87 | 850.85 | 801.99 |

Table 3: **Model fit statistics for Subjects 8-14.** boldface type represents the lowest BIC value for each subject.

| Model | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
|---|---|---|---|---|---|---|---|
| 1 | 775.71 | 991.92 | 724.68 | 992.44 | 501.22 | 915.63 | 769.59 |
| 2 | 634.18 | 883.91 | 687.46 | 989.68 | 529.26 | 814.55 | 680.70 |
| 3 | 767.44 | 959.03 | 706.02 | 947.96 | 499.26 | 730.39 | 758.59 |
| 4 | 616.46 | 916.72 | 696.79 | 916.11 | 503.74 | 720.85 | 705.31 |
| 5 | 652.17 | 938.47 | 687.78 | 970.18 | 491.42 | 777.46 | 757.84 |
| 6 | 591.90 | **844.78** | 674.59 | 857.79 | 496.09 | 730.25 | 655.59 |
| 7 | 658.51 | 919.95 | 677.78 | 937.88 | 496.51 | 700.61 | 734.95 |
| 8 | 627.44 | 908.64 | 672.07 | 829.66 | 497.42 | 774.14 | 718.03 |
| 9 | 607.86 | 845.82 | 660.56 | 817.11 | 484.81 | 700.03 | 676.48 |
| 10 | 631.55 | 893.08 | 652.55 | **702.74** | 485.65 | **671.05** | 677.13 |
| 11 | 607.87 | 877.35 | **648.10** | 741.94 | **476.54** | 718.87 | 692.81 |
| 12 | 637.92 | 891.72 | 666.32 | 759.94 | 501.03 | 775.38 | 639.13 |
| 13 | 638.97 | 875.00 | 694.13 | 846.25 | 485.23 | 907.59 | 691.22 |
| 14 | **546.29** | 968.94 | 657.48 | 743.69 | 532.9 | 684.89 | **610.12** |
| 15 | 629.47 | 878.16 | 666.24 | 791.53 | 484.00 | 714.01 | 718.07 |
| 16 | 570.64 | 870.55 | 677.58 | 772.04 | 495.96 | 705.65 | 678.64 |

## 4. Model Comparison

In addition to comparing subjects using standardized BIC scores, we also compared subjects by calculating the approximate posterior model probabilities as given by Wasserman (2000). Figure 1 shows the resulting posterior model probability values obtained. Each model performance statistic is color coded according to the legend on the right hand side. For the approximate posterior model probability, values closer to one (warmer colors) indicate the better fit, relative to the 15 other model variants for each subject. The left panel of Figure 1 provides a schematic of the various models investigated here. Parameters are represented as nodes in each column, where parameters that were freely estimated are empty, parameters that were fixed are solid, and parameters that are not applicable are shown as "x"s. From left to right, the columns are the model numbers, between-trial variability in starting point $S_0$, between-trial variability in drift $\eta$, the asymptotic boundary setting $a'$, and the scaling parameter of the collapsing decision boundary $\lambda$. For ease of discussion, the models were grouped into four classes: Models 1-4 (Class 1; red) are the time-invariant models, Models 5-8 (Class 2; blue) are the time-variant models with $a'$ free and $\lambda$ fixed, Models 9-12 (Class 3; green) are the time-variant models with $a'$ fixed and $\lambda$ free, and Models 13-16 (Class 4; purple) are the time-variant models with both $a'$ and $\lambda$ free.

Figure 1 shows that all subject's data are best explained by one or two model variants, and these model variants are some variant of the time-invariant diffusion model. Across subjects, most of the models with high posterior probabilities belong to Class 3 ($\lambda$ is free), but Class 4 also performs well. These results suggest that freeing $\lambda$ (Class 3) substantially improves

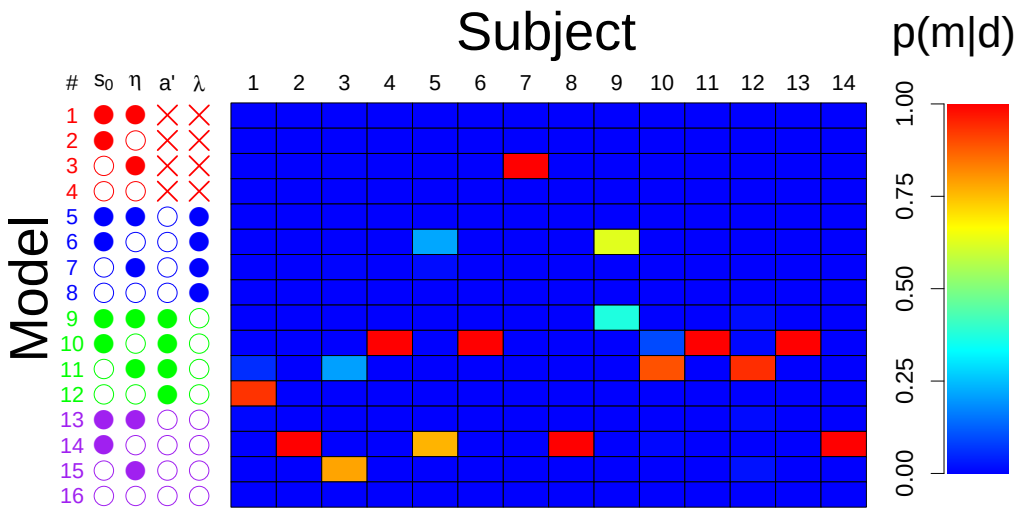


Figure 1: **Model Comparison.** Each box illustrates the posterior model probability obtained for each subject (columns) and model (rows) combination, color coded according to the legend on the right-hand side. Higher posterior probabilities (i.e., warmer colors) correspond to better model performance. The panel on the left summarizes the models investigated here, where the columns correspond to the model number and the specific model parameters that were either fixed (filled circle), free to vary (empty circle) or not applicable (an "x" symbol). For convenience, the models were also grouped into classes: time-invariant (Models 1-4; red), time-variant models with $a'$ free (Models 5-8; blue), time-variant models with $\lambda$ free (Models 9-12; green), and time-variant models with both $a'$ and $\lambda$ free (Models 13-16; purple).

model performance for the majority of subjects in our study. However, freeing $a'$ in addition to $\lambda$ (Class 4) improved model fit in such a way that the penalty induced by the additional parameter was not enough to damage the performance of the model. Again, while the posterior model probability lacks the finer gradient of the standardized BIC model comparison, the clear worst performer is Class 1, with all four models showing 0% probability for all subjects.

### *4.1. Fit to Observed Data*

Figures 2, 3, 4, and 5 display the model fits to observed data in another way. Instead of plotting each model's fit against the observed data as in Figure 5 of the main text, Figures 2, 3, 4, and 5 show the accuracy (y-axis) for both the data (grey) and the model predictions (colors) as a bar plots for each interrogation time (x-axis) and for each level of coherency (columns). These figures allow us to see in greater detail how closely the model predictions are to the observed datar and where the models have over- or under-predicted the accuracy relative to the observed data.

While the predictions across all four models are highly variable, Model 4 and Model 14 tend to predict a different pattern of accuracy for the 0.15, 0.25, and 0.35 coherency conditions. Whereas Model 7 and and Model 11 predict an increase in accuracy with an increase in interrogation time in these conditions, the accuracy as predicted by Models 4 and 14 tends to plateau for these conditions in the left panel (i.e.,"All Responses") and decreases in the bottom right panel (i.e., "Responses After the Cue").

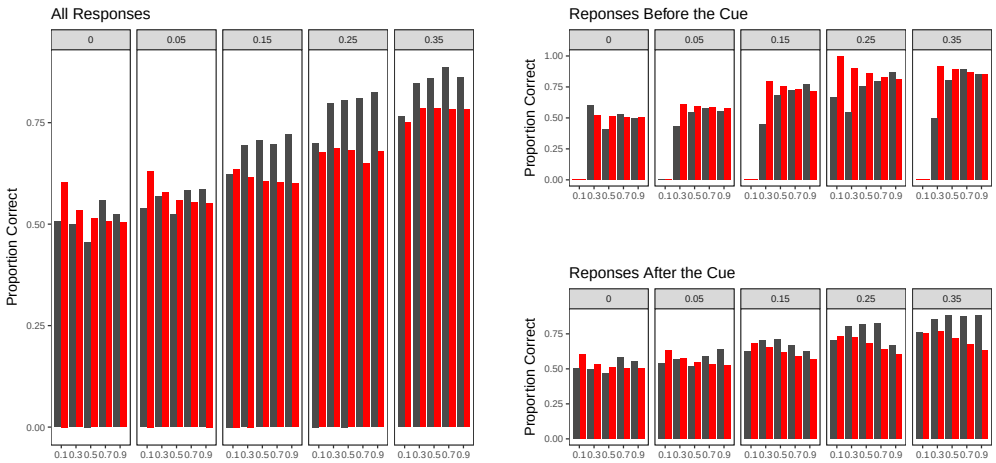


Figure 2: **Model Predictions of Accuracy Across Independent Variables versus Observed Data.** The left panel shows the observed accuracy (gray) and predicted accuracy from Model 4 (red) ($y$-axis) as bar plots for each interrogation time ($x$-axis) and for each coherency (panels), collapsed across response times. The right panel shows the same information as in the left panel, but separated based on whether the response was made prior to (top panel) or after (bottom panel) the stimulus disappeared.

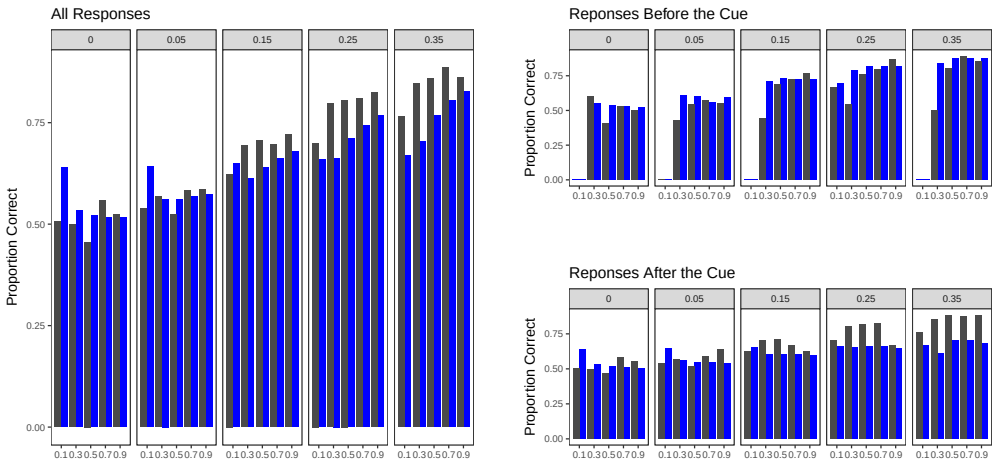


Figure 3: **Model Predictions of Accuracy Across Independent Variables versus Observed Data.** The left panel shows the observed accuracy (gray) and predicted accuracy from Model 7 (blue) ($y$-axis) as bar plots for each interrogation time ($x$-axis) and for each coherency (panels), collapsed across response times. The right panel shows the same information as in the left panel, but separated based on whether the response was made prior to (top panel) or after (bottom panel) the stimulus disappeared.

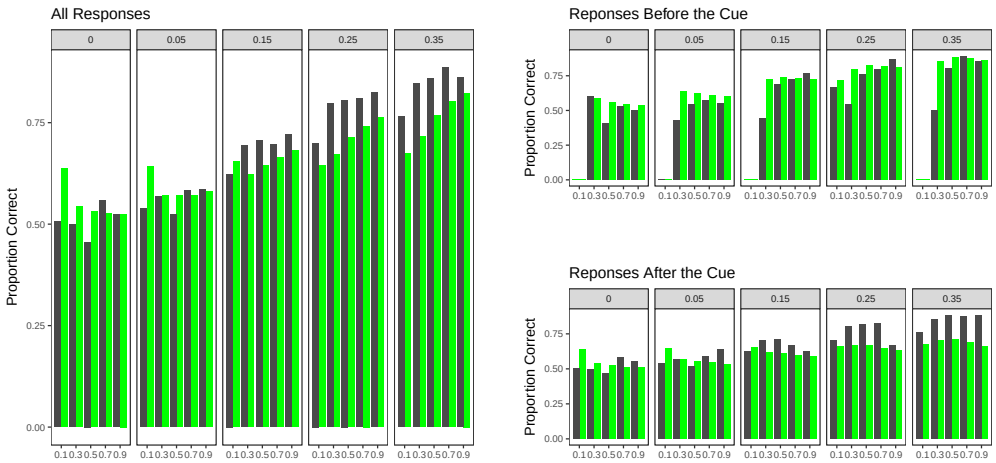


Figure 4: **Model Predictions of Accuracy Across Independent Variables versus Observed Data.** The left panel shows the observed accuracy (gray) and predicted accuracy from Model 11 (green) ($y$-axis) as bar plots for each interrogation time ($x$-axis) and for each coherency (panels), collapsed across response times. The right panel shows the same information as in the left panel, but separated based on whether the response was made prior to (top panel) or after (bottom panel) the stimulus disappeared.

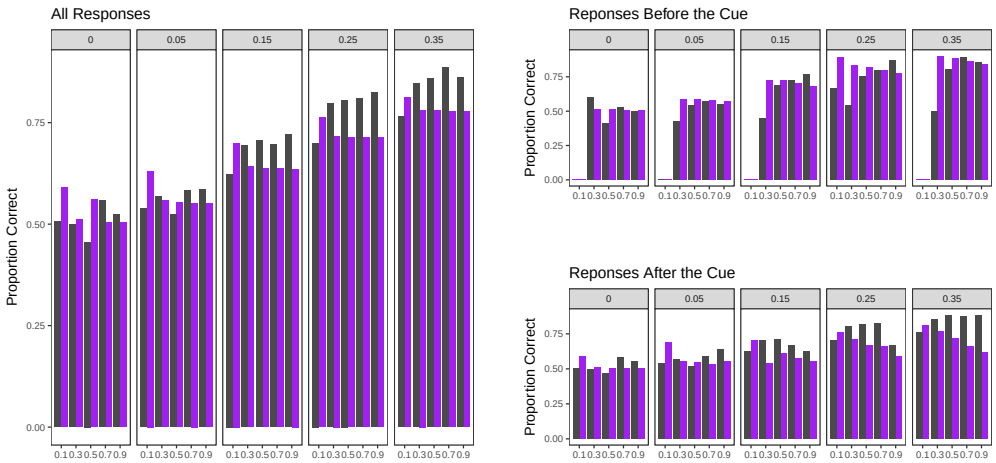


Figure 5: **Model Predictions of Accuracy Across Independent Variables versus Observed Data.** The left panel shows the observed accuracy (gray) and predicted accuracy from Model 14 (purple) ($y$-axis) as bar plots for each interrogation time ($x$-axis) and for each coherency (panels), collapsed across response times. The right panel shows the same information as in the left panel, but separated based on whether the response was made prior to (top panel) or after (bottom panel) the stimulus disappeared.

## 5. Bayes Factor Approximation

Table 3 displays the numerical approximated Bayes factor scores for each subject. Again, 13 of 14 subjects showed a preference for the time-invariant model, with the majority showing "strong" evidence for the time-variant model.

## 6. Practice

### *6.1. Practice Analysis*

Figure 6 illustrates the differences in fit between a linear function and a power function for the mean (left panel) and the standard deviation (right panel) of response times across each of the 13 experimental blocks. Due to the large difference in response times between the first and second block, both the mean and standard deviations appear to be best fit by a power function ($BIC_{mean} = 132.70$; $BIC_{std} = 122.33$) versus a linear function ($BIC_{mean} = 142.66$; $BIC_{std} = 131.10$).

### *6.2. Model Comparison with Trimmed Data*

Given the presence of practice effects in our data and the potential that "split" in the split half analysis performed in Section 5.4 of the main text is an arbitrary distinction, we refit Models 4, 7, 11 and 14 (i.e., the best fitting model from each class) to each subject's data after removing the first four blocks (i.e., 200 trials), which appeared to have the largest influence of practice. This model fit would allow us to see if removing the influence of practice impacted each model's ability to account for the data. Due to the

Table 4: **Bayes factor comparison**. Approximated Bayes factor values after comparing the best fitting time-variant model (iModel 11, which includes between-trial variability in starting point and freely estimates $\lambda$) to the best fitting time-invariant model (i.e., Model 4, where both sources of trial-to-trial variability are included) for each of the subjects in our data.

| Subject | Log(Bayes Factor) | Strength of Evidence | Model |
|---|---|---|---|
| 1 | 116.43 | Strong | Time-variant |
| 2 | 9.58 | Strong | Time-variant |
| 3 | 137.02 | Strong | Time-variant |
| 4 | 50.02 | Strong | Time-variant |
| 5 | -1.71 | Strong | Time-invariant |
| 6 | 18.51 | Strong | Time-variant |
| 7 | 114.87 | Strong | Time-variant |
| 8 | 4.3 | Strong | Time-variant |
| 9 | 19.69 | Strong | Time-variant |
| 10 | 24.35 | Strong | Time-variant |
| 11 | 57.09 | Strong | Time-variant |
| 12 | 13.6 | Strong | Time-variant |
| 13 | 0.99 | Moderate | Time-variant |
| 14 | 6.25 | Strong | Time-variant |

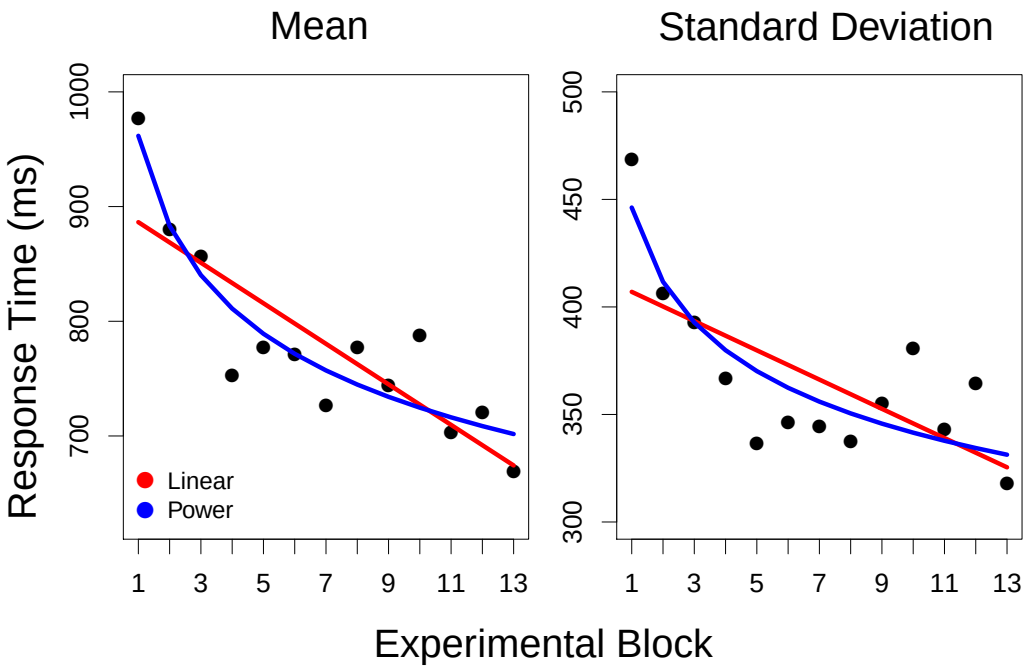


Figure 6: **Analysis of Practice Effects.** The left panel shows the best fitting linear (red) and power (blue) function to the mean response time across blocks. Similarly, the right panel shows the best fitting linear (red) and power (blue) function to the standard deviation of the response times across blocks.

removal of these four blocks, Subject 12 was dropped from the analysis due as only 9% of their data were remaining.

The results of these fits are in Figure 7, which shows the zBIC values, color coded according to the legend on the right side. Cooler colors represent lower BIC values and better model performance. While these fits are highly variable, the clear best fitting model across subjects is Model 11, with Models 7, 14, and 4 following in that order. Overall, the majority of subjects are best fit by collapsing bound models, suggesting that time-variant mechanisms serve as a more viable explanation for how subjects complete the task even after removing the influence of practice.

## 7. Inferred Task Representations

Inferred task representation plots for the remaining 12 subjects are displayed below. In these plots, the best fitting model is indicated by the color of the response time distribution in the top row of each figure. If the distribution is red, then the subject's data are better fit by the time-invariant model (Model 4). If the distribution is blue, then the subject's data are better fit by the time-variant model (Model 16).

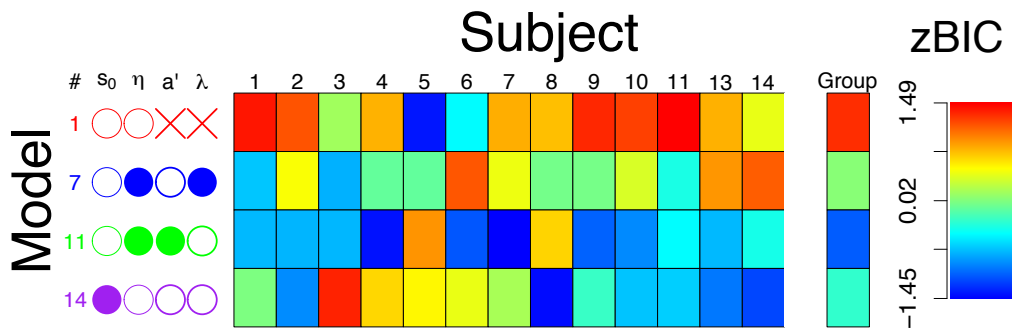


Figure 7: **Model Comparison of Trimmed Data.** Each box illustrates the z-scored BIC value obtained for each subject (columns) and model (rows) combination, color coded according to the legend on the right-hand side. Lower BIC values (i.e., cooler colors) correspond to better model performance. The panel on the left summarizes the models investigated here, where the columns correspond to the model number and the specific model parameters that were either fixed (filled circle), free to vary (empty circle) or not applicable (an "x" symbol). The models are color coded according to class: time-invariant (Model 4; red); time-variant with $a'$ free (Model 7 ; blue); time-variant with $\lambda$ free (Model 11; green); and time-variant with both $a'$ and $\lambda$ free (Models 13-16; purple).

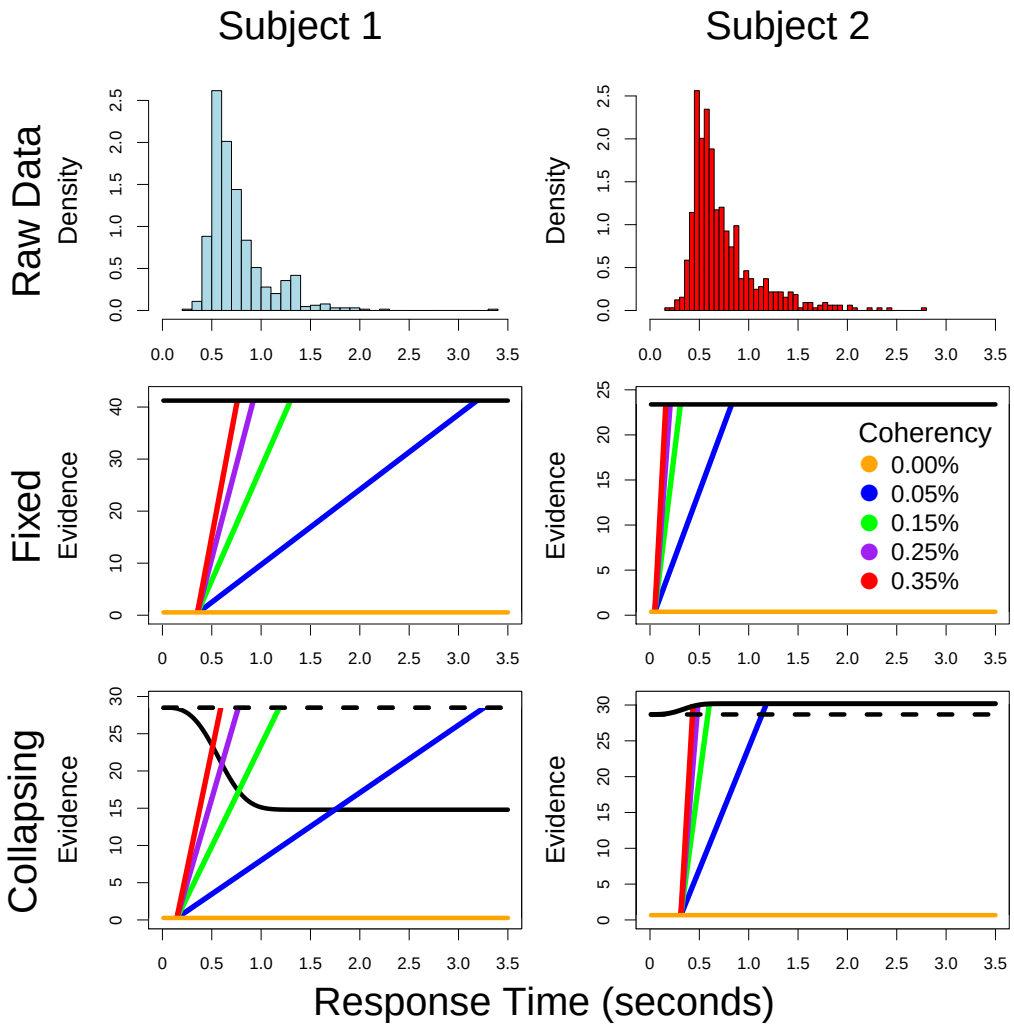


Figure 8: **Inferred Task Representations for subjects 1 and 2.** The columns correspond to two subjects' data. The top row shows the empirical response time distribution collapsed across choice and stimulus conditions. If the response time distribution is red, the subjects' data was best fit by the time-invariant model (Model 4). If the response time distribution is light blue, the subjects' data was best fit by the time-variant model (Model 16). The middle row shows the task representation inferred by a fixed boundary model, whereas the bottom row shows the task representation inferred by a collapsing boundary model. The drift rates for each coherency condition are shown, color coded according to the key in the middle right panel.

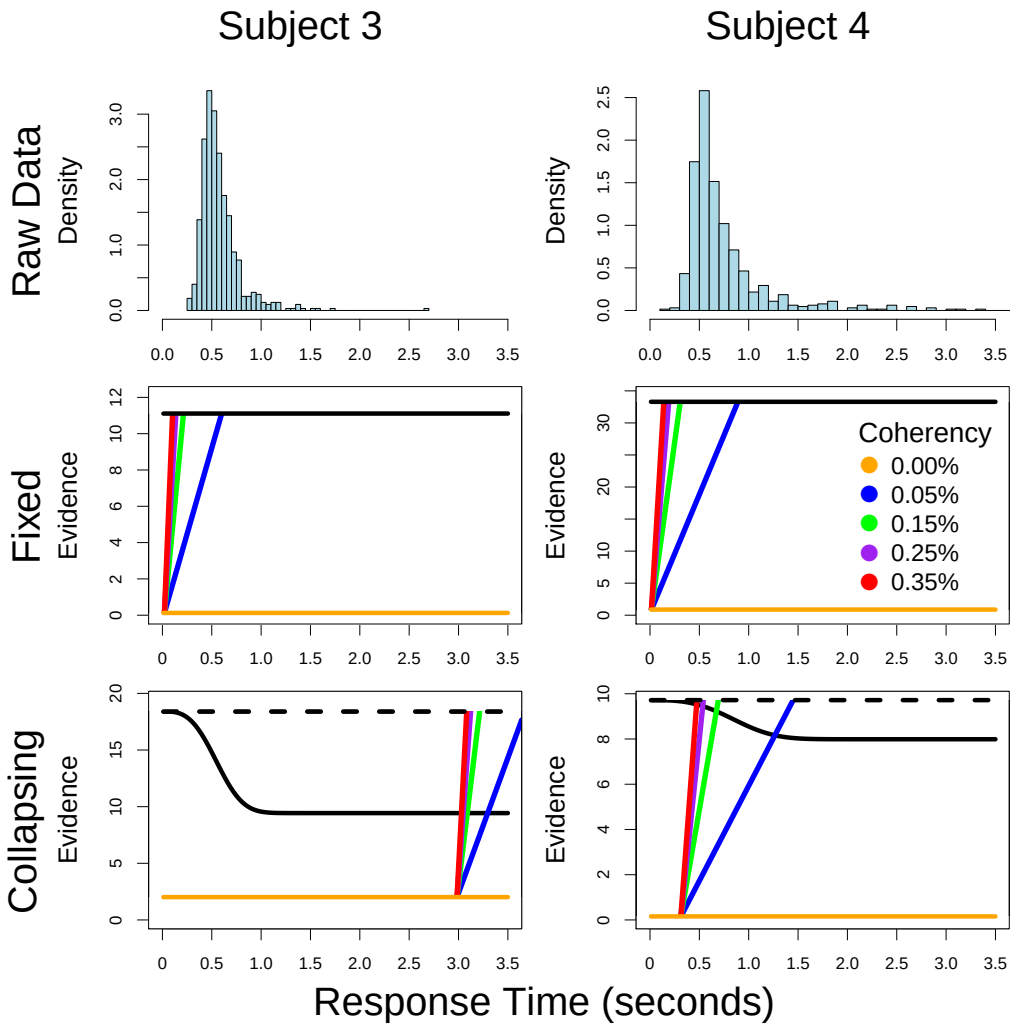


Figure 9: **Inferred Task Representations for Subjects 3 and 4.** The columns correspond to two subjects' data. The top row shows the empirical response time distribution collapsed across choice and stimulus conditions. If the response time distribution is red, the subjects' data was best fit by the time-invariant model (Model 4). If the response time distribution is light blue, the subjects' data was best fit by the time-variant model (Model 16). The middle row shows the task representation inferred by a fixed boundary model, whereas the bottom row shows the task representation inferred by a collapsing boundary model. The drift rates for each coherency condition are shown, color coded according to the key in the middle right panel.

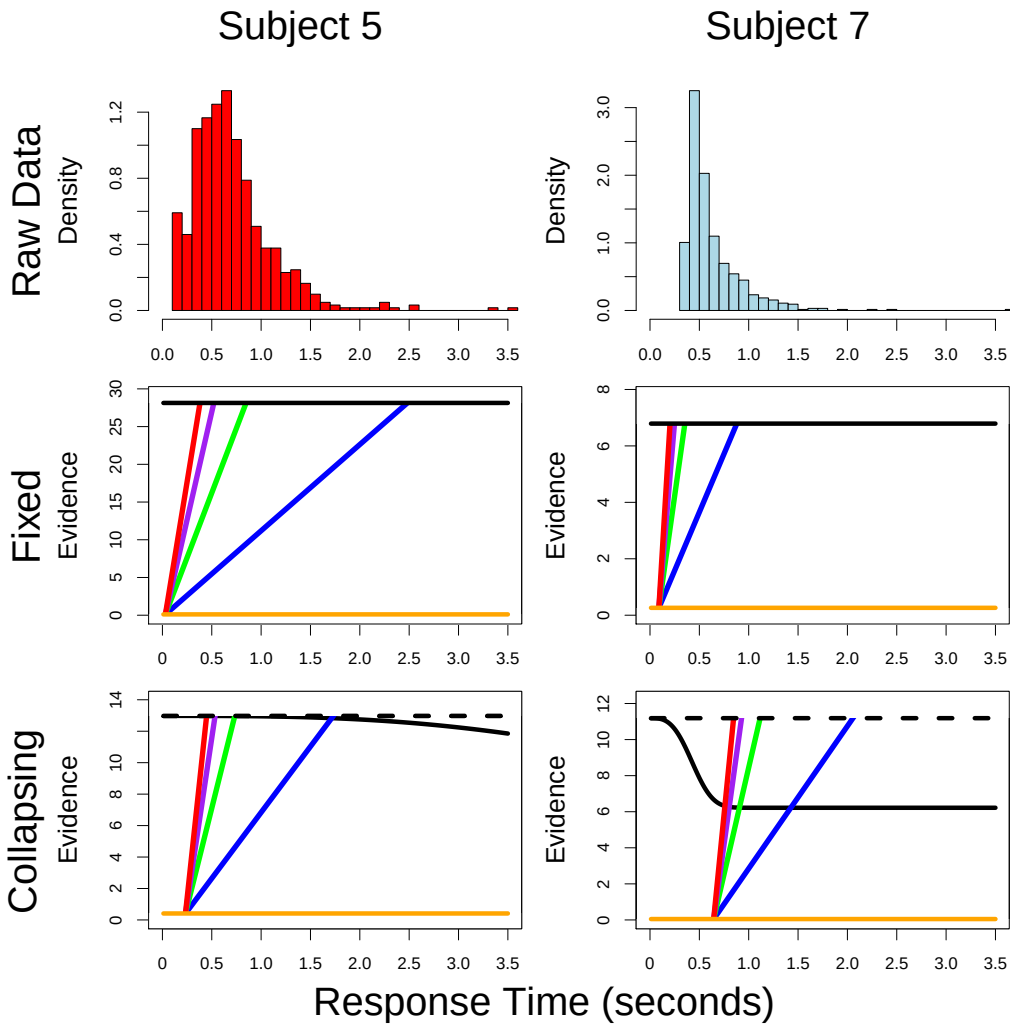


Figure 10: **Inferred Task Representations for Subjects 5 and 7.** The columns correspond to two subjects' data. The top row shows the empirical response time distribution collapsed across choice and stimulus conditions. If the response time distribution is red, the subjects' data was best fit by the time-invariant model (Model 4). If the response time distribution is light blue, the subjects' data was best fit by the time-variant model (Model 16). The middle row shows the task representation inferred by a fixed boundary model, whereas the bottom row shows the task representation inferred by a collapsing boundary model. The drift rates for each coherency condition are shown, color coded according to the key in the middle right panel.

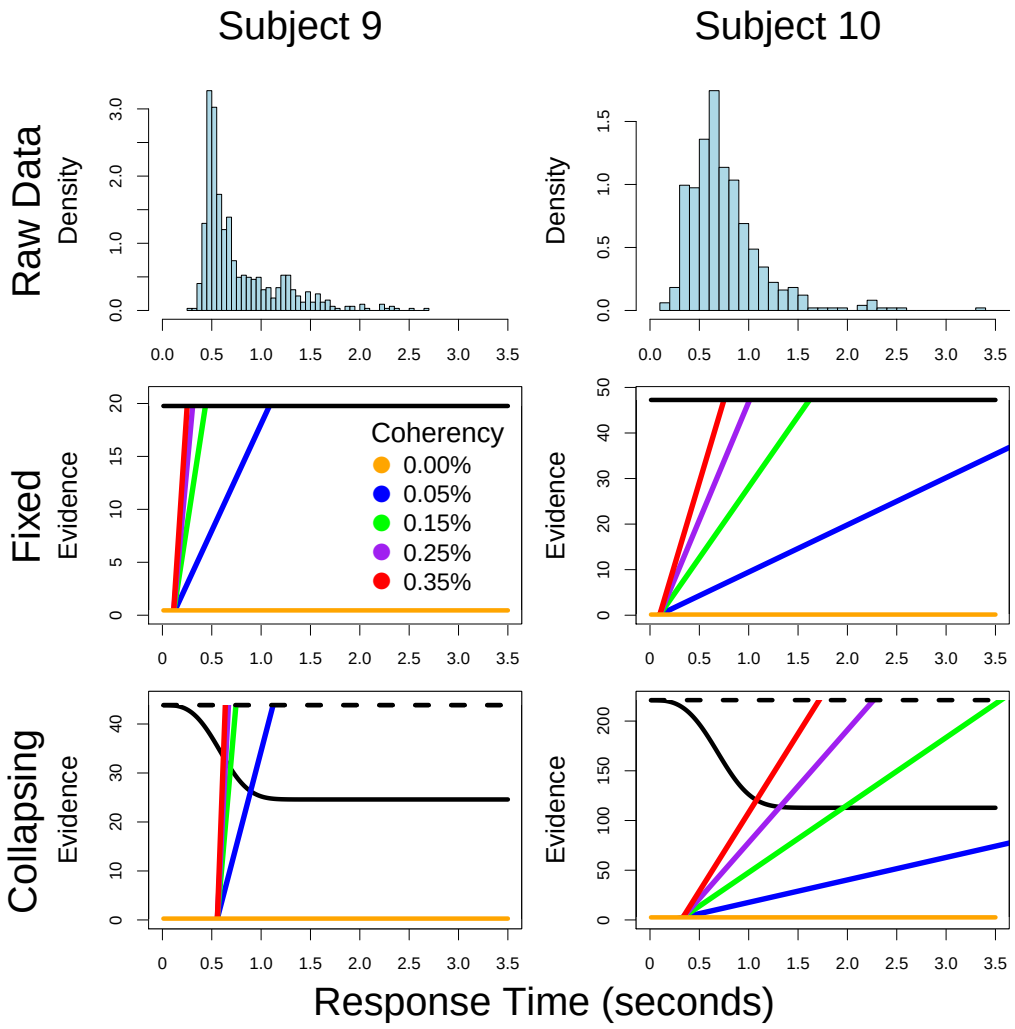


Figure 11: **Inferred Task Representations for Subjects 9 and 10.** The columns correspond to two subjects' data. The top row shows the empirical response time distribution collapsed across choice and stimulus conditions. If the response time distribution is red, the subjects' data was best fit by the time-invariant model (Model 4). If the response time distribution is light blue, the subjects' data was best fit by the time-variant model (Model 16). The middle row shows the task representation inferred by a fixed boundary model, whereas the bottom row shows the task representation inferred by a collapsing boundary model. The drift rates for each coherency condition are shown, color coded according to the key in the middle right panel.

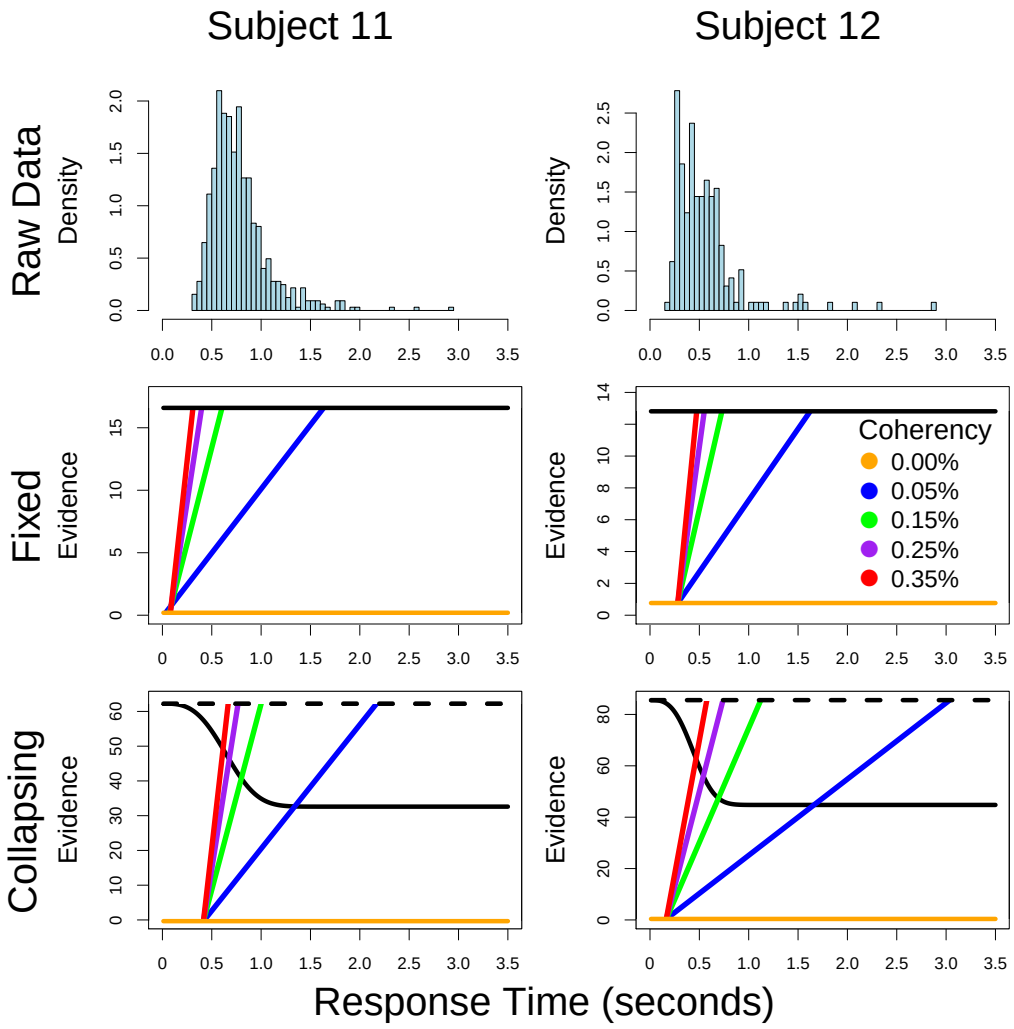


Figure 12: **Inferred Task Representations for Subjects 11 and 12.** The columns correspond to two subjects' data. The top row shows the empirical response time distribution collapsed across choice and stimulus conditions. If the response time distribution is red, the subjects' data was best fit by the time-invariant model (Model 4). If the response time distribution is light blue, the subjects' data was best fit by the time-variant model (Model 16). The middle row shows the task representation inferred by a fixed boundary model, whereas the bottom row shows the task representation inferred by a collapsing boundary model. The drift rates for each coherency condition are shown, color coded according to the key in the middle right panel.

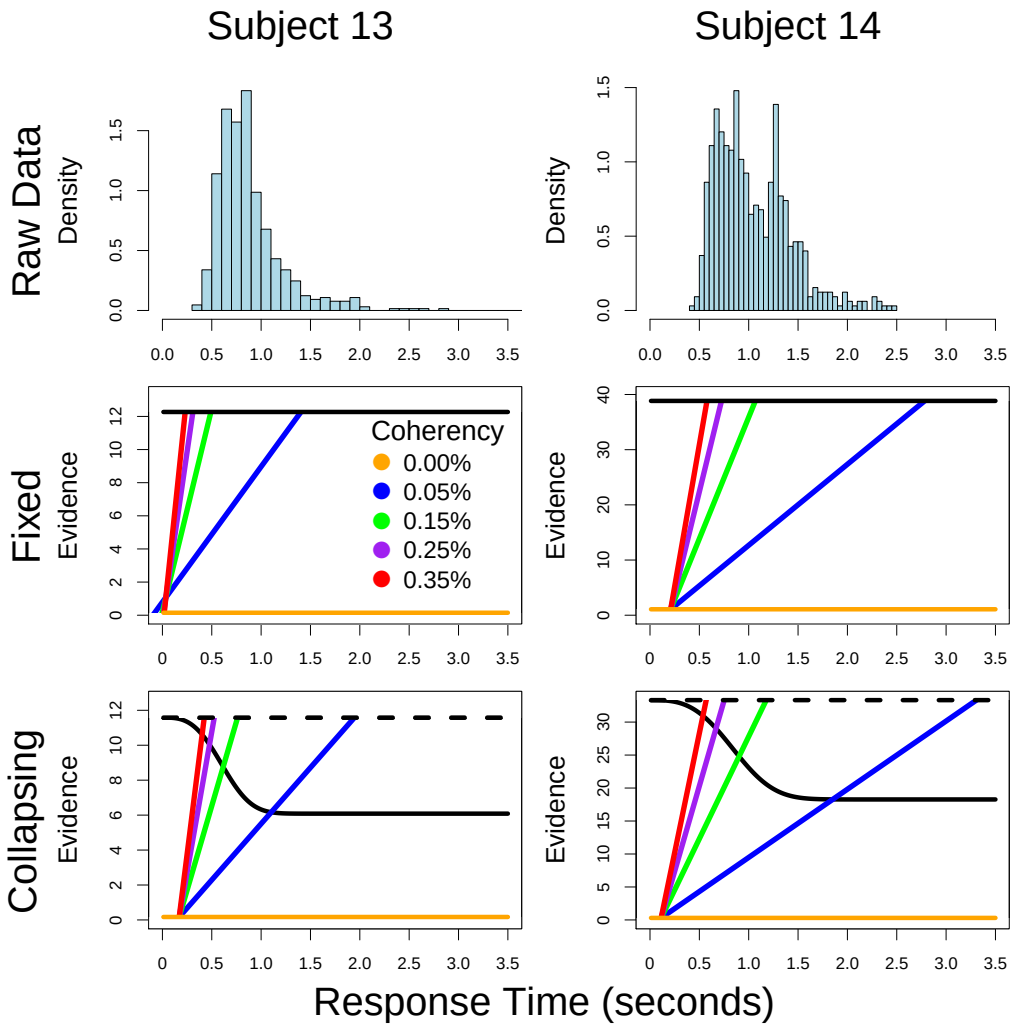


Figure 13: **Inferred Task Representations for Subjects 13 and 14.** The columns correspond to two subjects' data. The top row shows the empirical response time distribution collapsed across choice and stimulus conditions. If the response time distribution is red, the subjects' data was best fit by the time-invariant model (Model 4). If the response time distribution is light blue, the subjects' data was best fit by the time-variant model (Model 16). The middle row shows the task representation inferred by a fixed boundary model, whereas the bottom row shows the task representation inferred by a collapsing boundary model. The drift rates for each coherency condition are shown, color coded according to the key in the middle right panel.

## 8. References


Ratcliff, R., 1978. A theory of memory retrieval. Psychological Review 85, 59–108.

Ratcliff, R., McKoon, G., 2008. The diffusion decision model: theory and data for two-choice decision tasks. Neural Computation 20, 873–922.

Ratcliff, R., Rouder, J. N., 1998. Modeling response times for two-choice decisions. Psychological Science 9, 347–356.

Ratcliff, R., Thapar, A., McKoon, G., 2006. Aging, practice, and perceptual tasks: a diffusion model analysis. Psychological and Aging 21, 353–371.

Ratcliff, R., Tuerlinckx, F., 2002. Estimating parameters of the diffusion model: Approaches to dealing with contaminant reaction time and parameter variability. Psychonomic Bulletin and Review 9, 438–481.

Smith, P. L., 2000. Stochastic dynamic models of response time and accuracy: A foundational primer. Journal of Mathematical Psychology 44, 408–463.

Starns, J. J., Ratcliff, R., 2010. The effects of aging on the speed-accuracy compromise: boundary optimality in the diffusion model. Psychological Aging 25, 377–390.

Stone, M., 1960. Models for choice reaction time. Psychometrika 25, 251–260.

Turner, B. M., Gao, J., Koenig, S., Palfy, D., McClelland, J. L., 2017. The dynamics of multimodal integration: The averaging diffusion model, in Press.

198 Wasserman, L., 2000. Bayesian model selection and model averaging. Journal
199 of Mathematical Psychology 44, 92–107.

200 White, C., Ratcliff, R., Vasey, M., McKoon, G., 2009. Dysphoria and memory
201 for emotional material: A diffusion model analysis. Cognition and Emotion
202 23, 181–205.