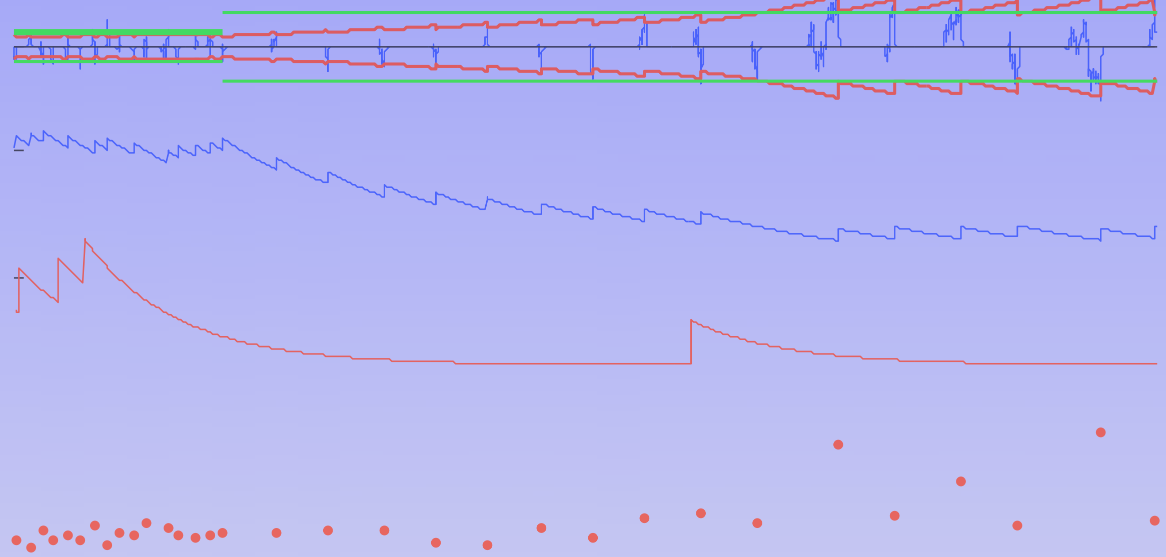
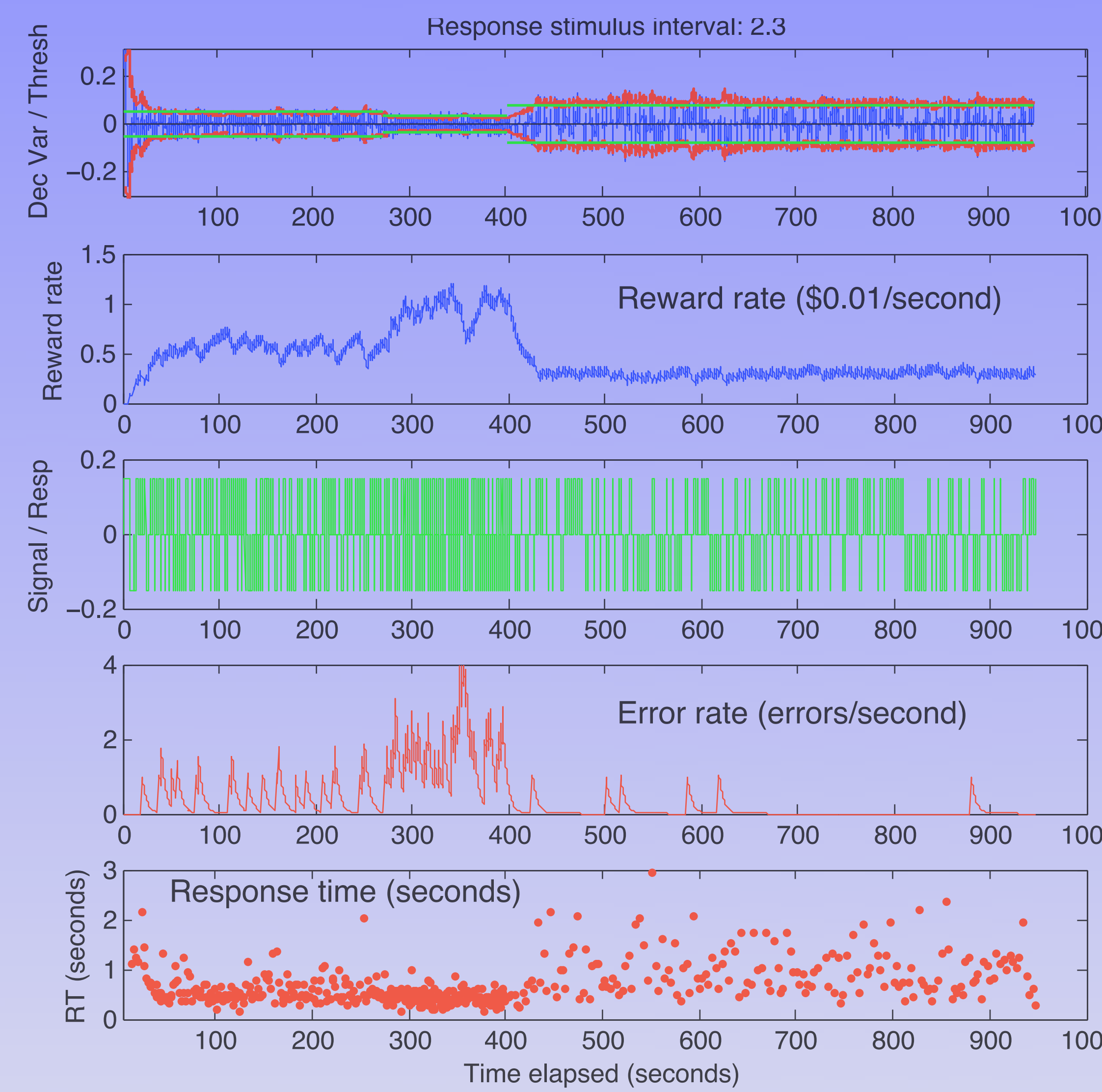
Adaptive performance in two-alternative decision making





Patrick Simen

Collaboration with Phil Holmes and Jonathan D. Cohen
Princeton University

Deciding, by drift and diffusion

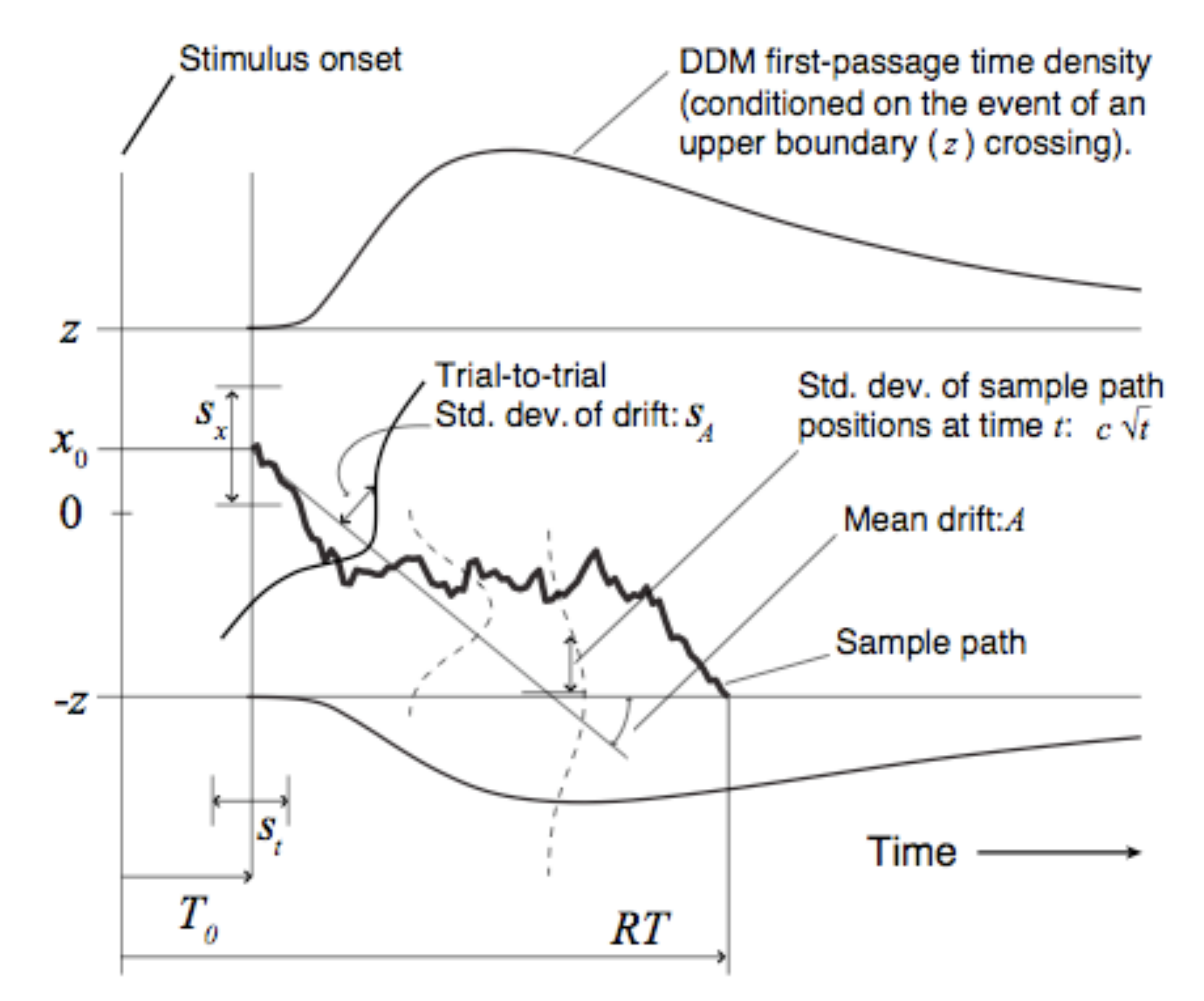
Continuous time:

$$dX = A\,dt + \sigma dW$$

Discrete time (a random walk):

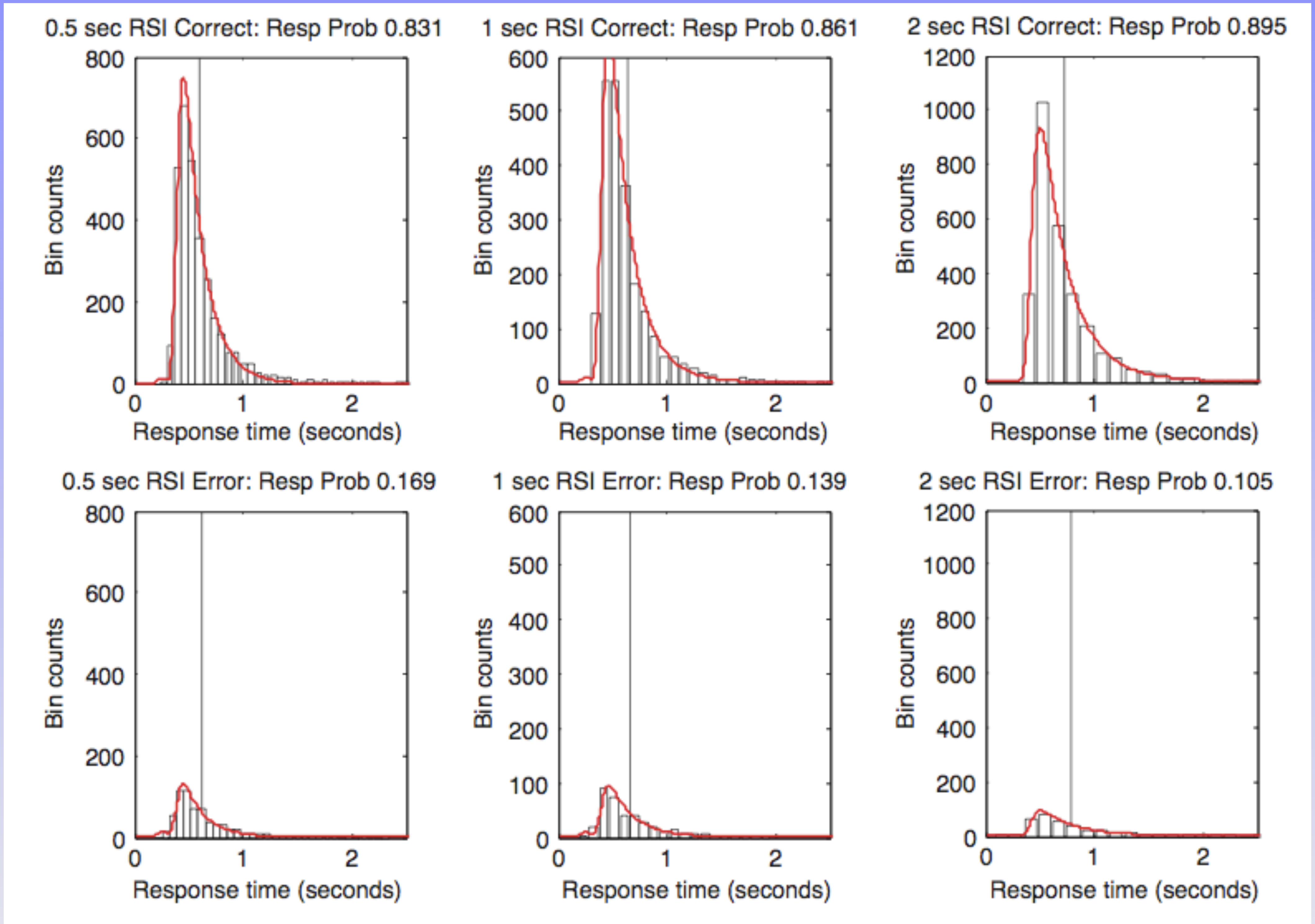
$$X_{new} = X_{old} + \ldots$$
$$A \cdot \Delta_t + \ldots$$
$$\sigma \cdot \sqrt{\Delta_t} \cdot \xi,$$

with ξ standard normal



cf. Ratcliff & Rouder (1998) *Psych Science*

Fits data well

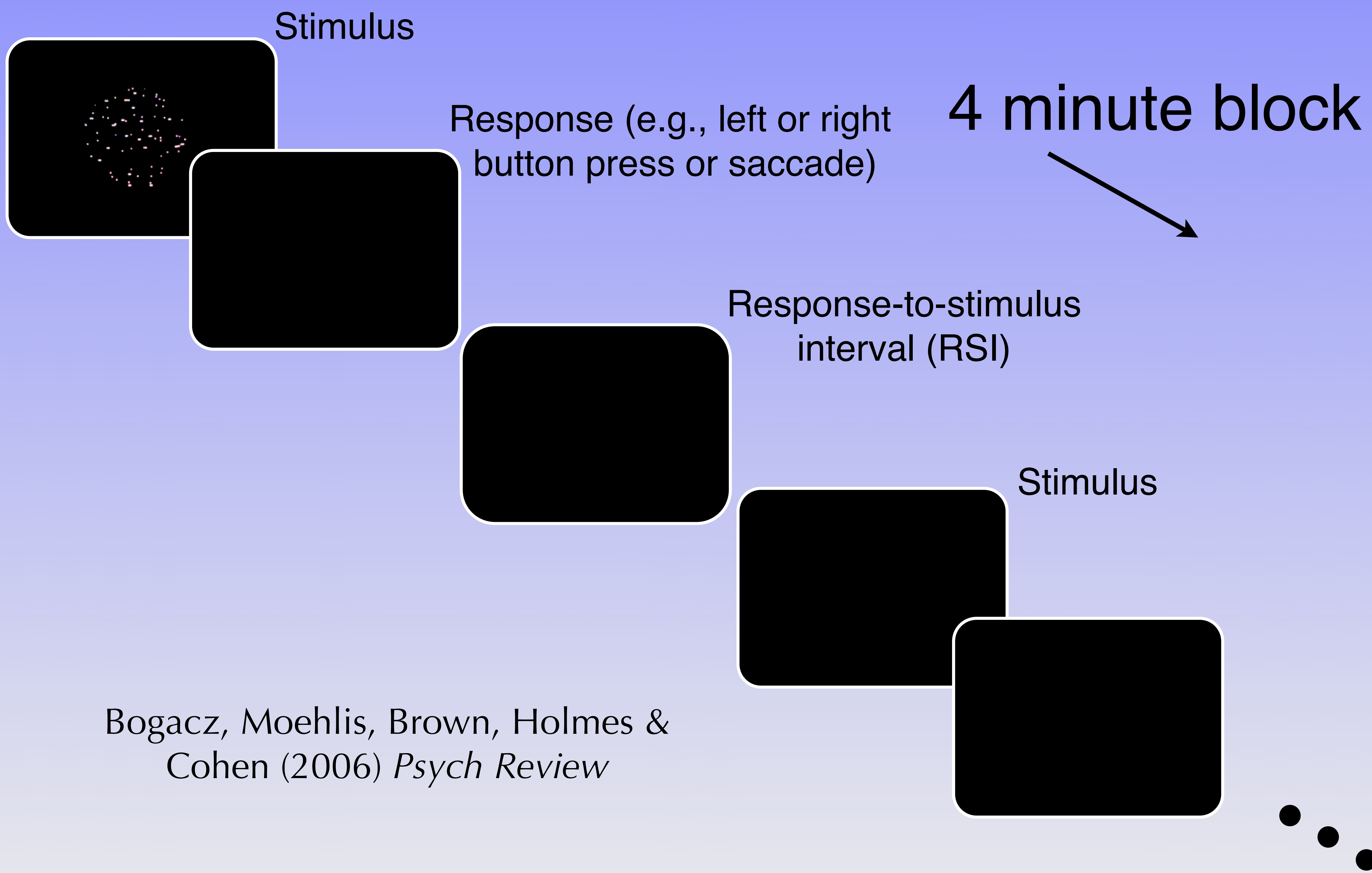


Simen, Contreras, Buck, Hu, Holmes & Cohen (in press), *J Exp Psych: Human Perception & Performance*

Outline

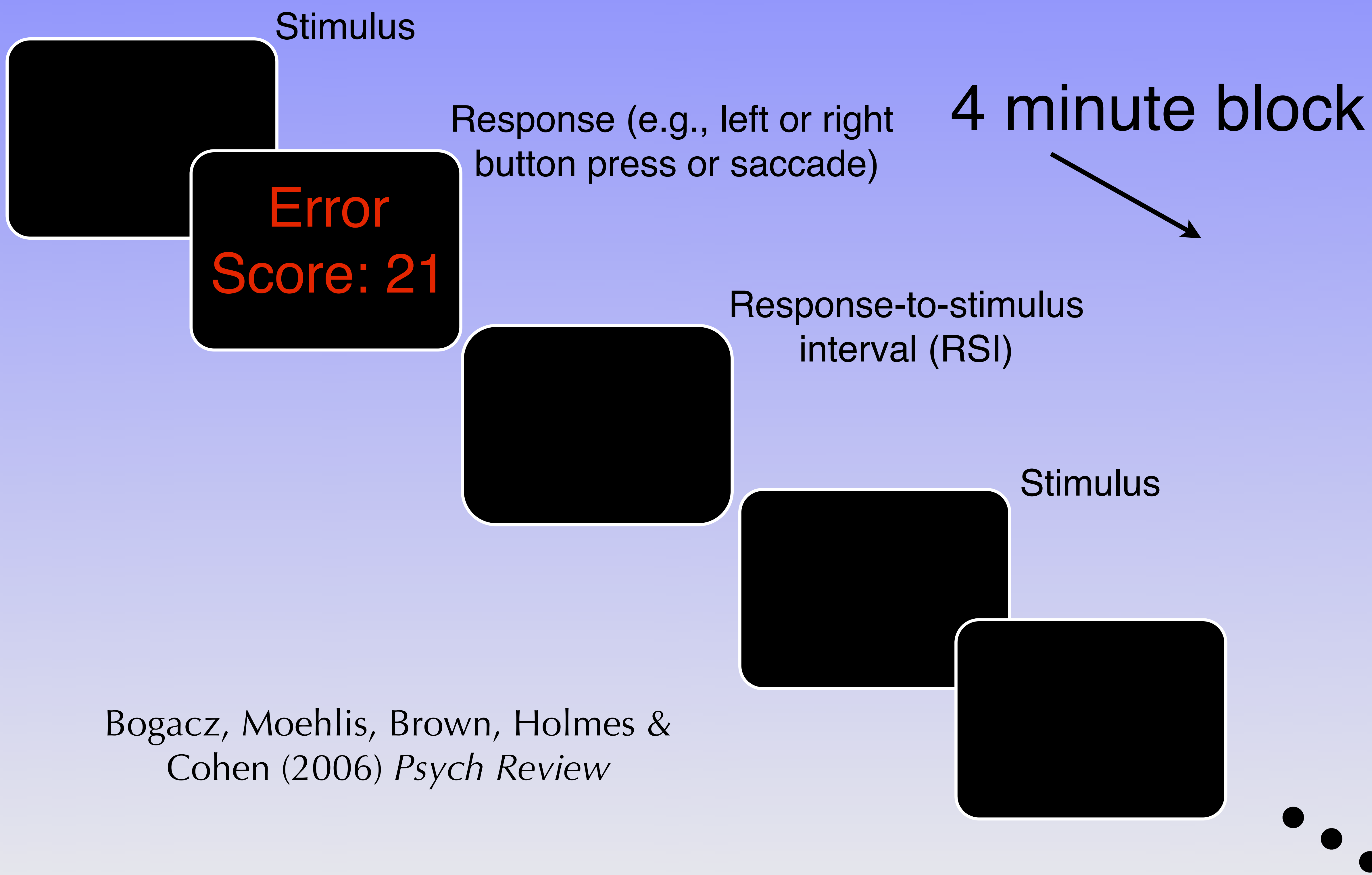
- Can simple conditioning/RL principles determine parameter specification in a "neurally plausible" implementation of this decision-making model?
- Optimal (reward maximizing) behavior statistics can be efficiently predicted for a simple, generic task that mimics life in a Skinner box.
- People seem to exhibit the predicted average behavior in this task; they also appear to implement a very simple learning algorithm for discovering nearly optimal model parameters.

The task: simple 2AFC, with overtones of operant conditioning



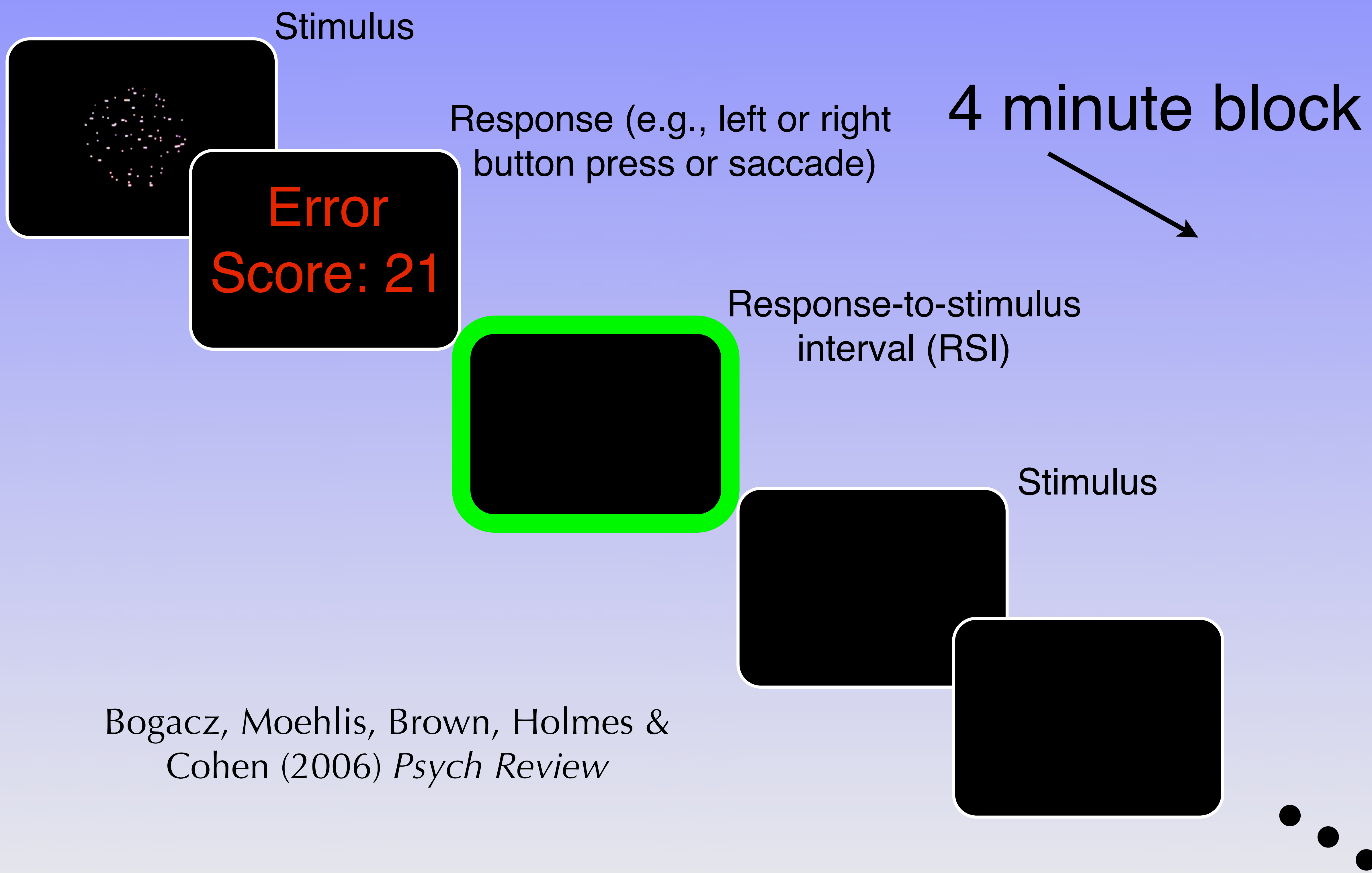
Bogacz, Moehlis, Brown, Holmes & Cohen (2006) *Psych Review*

The task: simple 2AFC, with overtones of operant conditioning



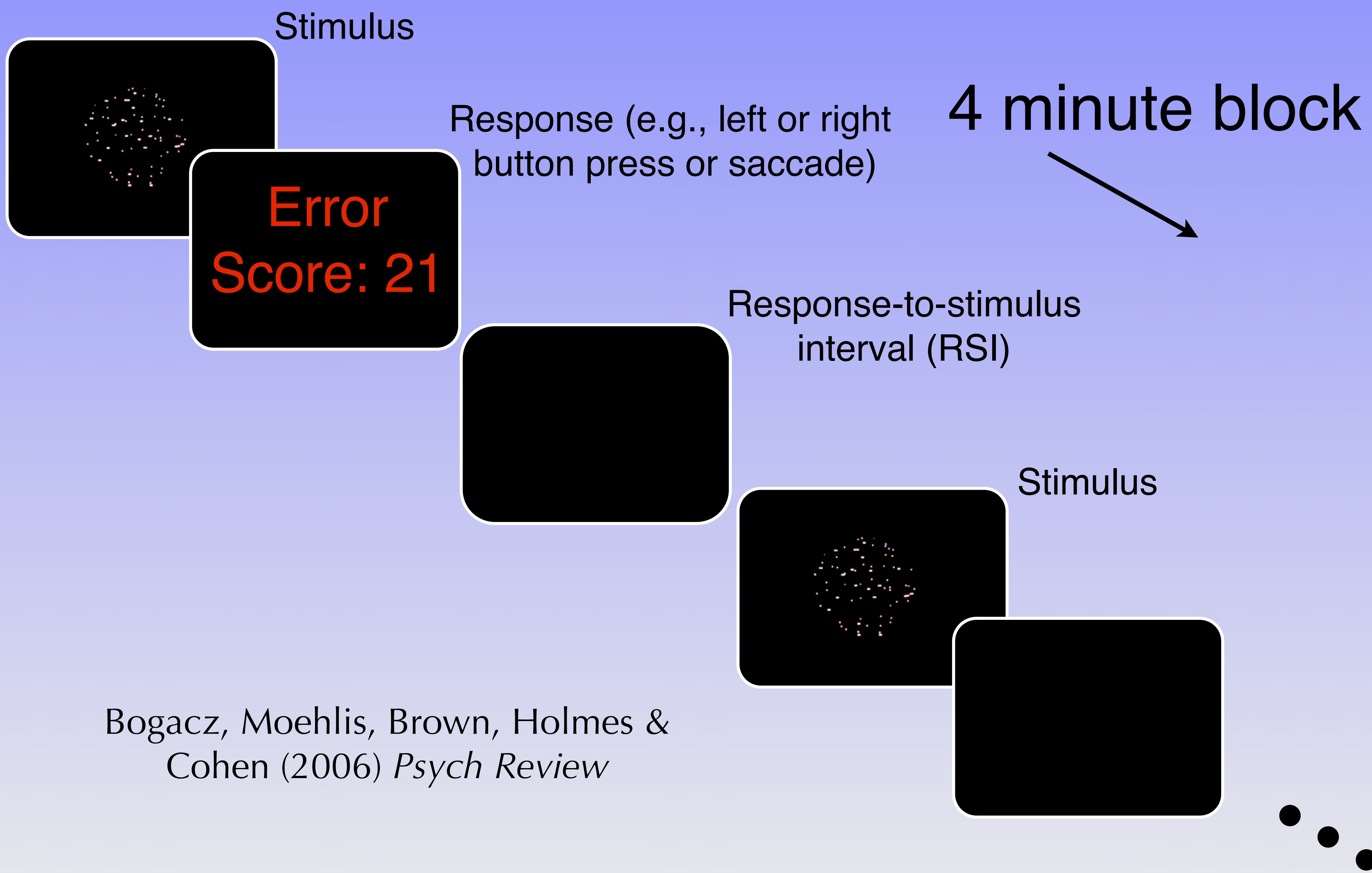
Bogacz, Moehlis, Brown, Holmes & Cohen (2006) *Psych Review*

The task: simple 2AFC, with overtones of operant conditioning



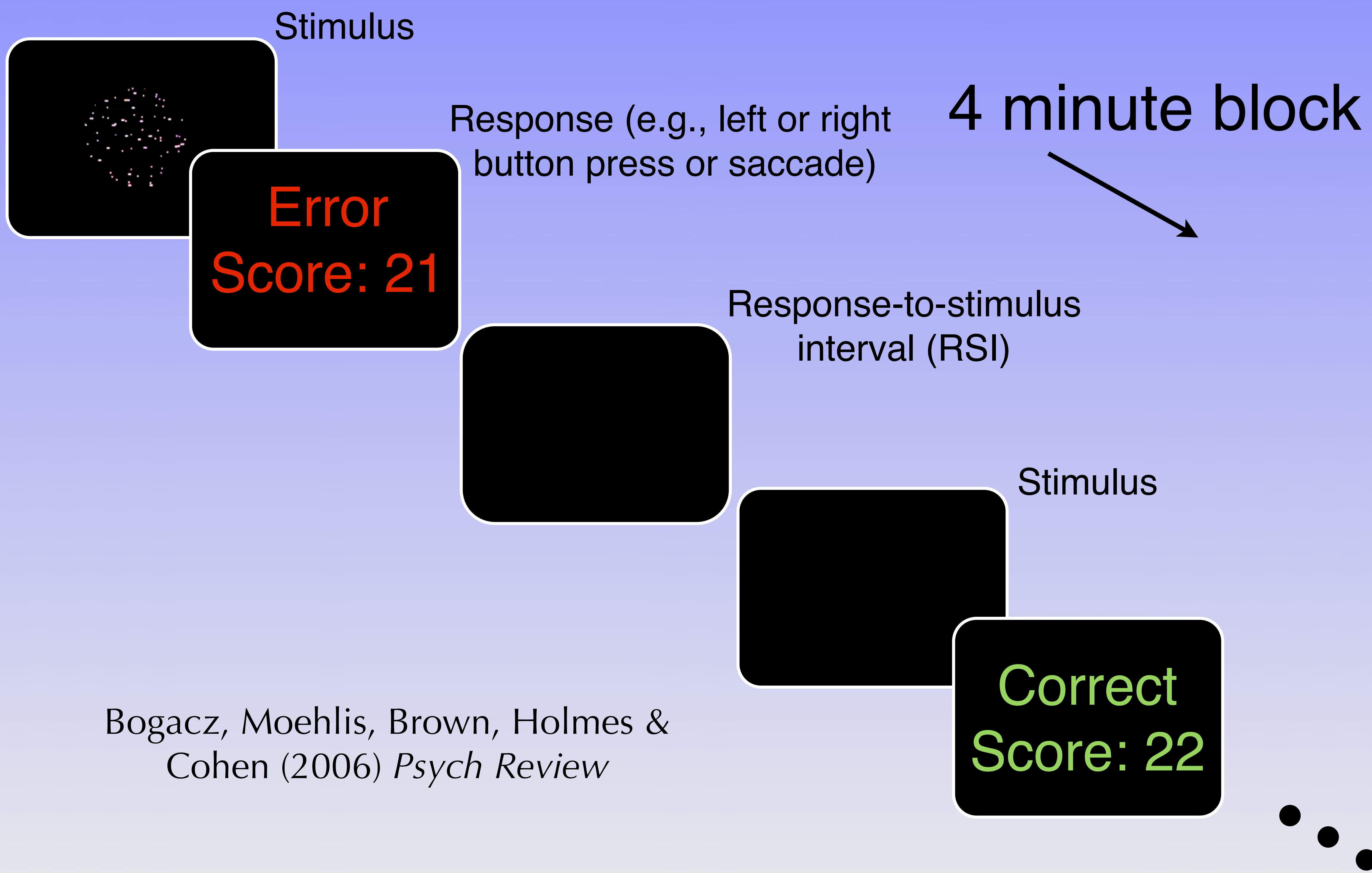
Bogacz, Moehlis, Brown, Holmes & Cohen (2006) *Psych Review*

The task: simple 2AFC, with overtones of operant conditioning



Bogacz, Moehlis, Brown, Holmes & Cohen (2006) *Psych Review*

The task: simple 2AFC, with overtones of operant conditioning



Bogacz, Moehlis, Brown, Holmes & Cohen (2006) *Psych Review*

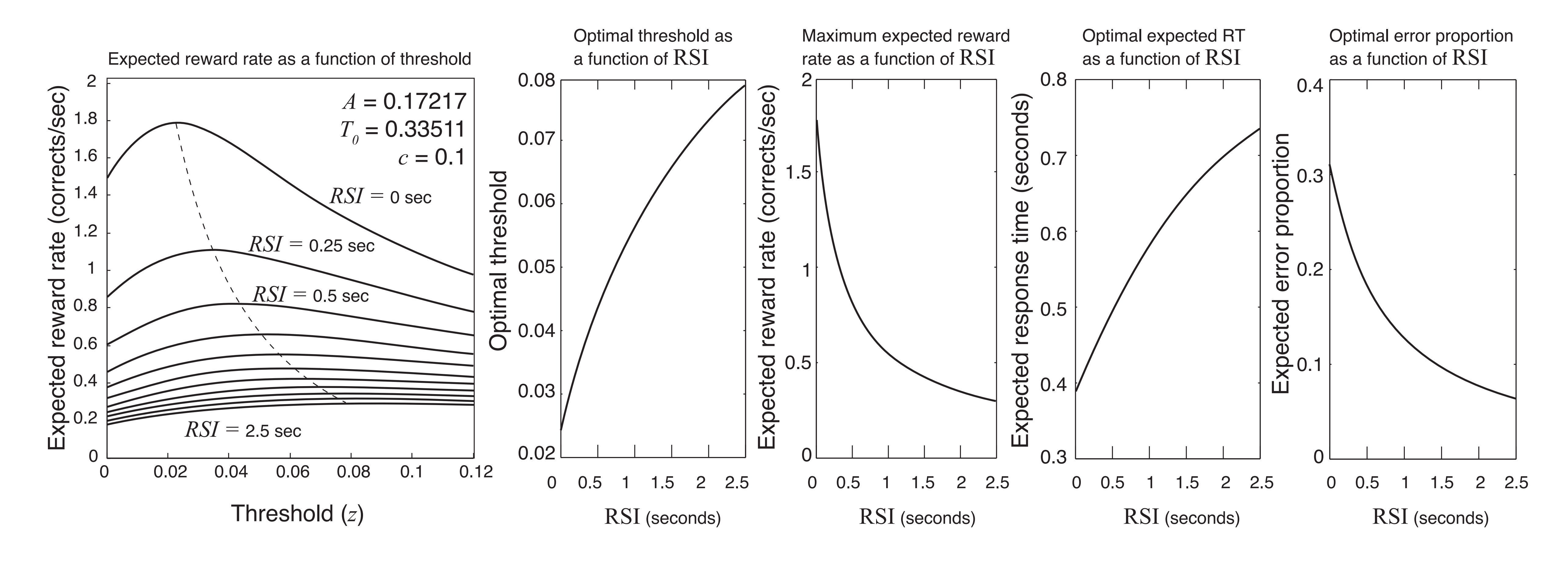
Reward-maximizing performance: predictions of Bogacz, Moehlis, Brown, Holmes & Cohen (2006) *Psych Review*

Q: What's the very best an ideal observer could do in a 2AFC task with rewards for corrects, no rewards for errors?

A: Reduce Ratcliff's diffusion model to Stone's sequential probability ratio test (SPRT) model by setting its extra parameters to 0, and feed it samples of log likelihood information.

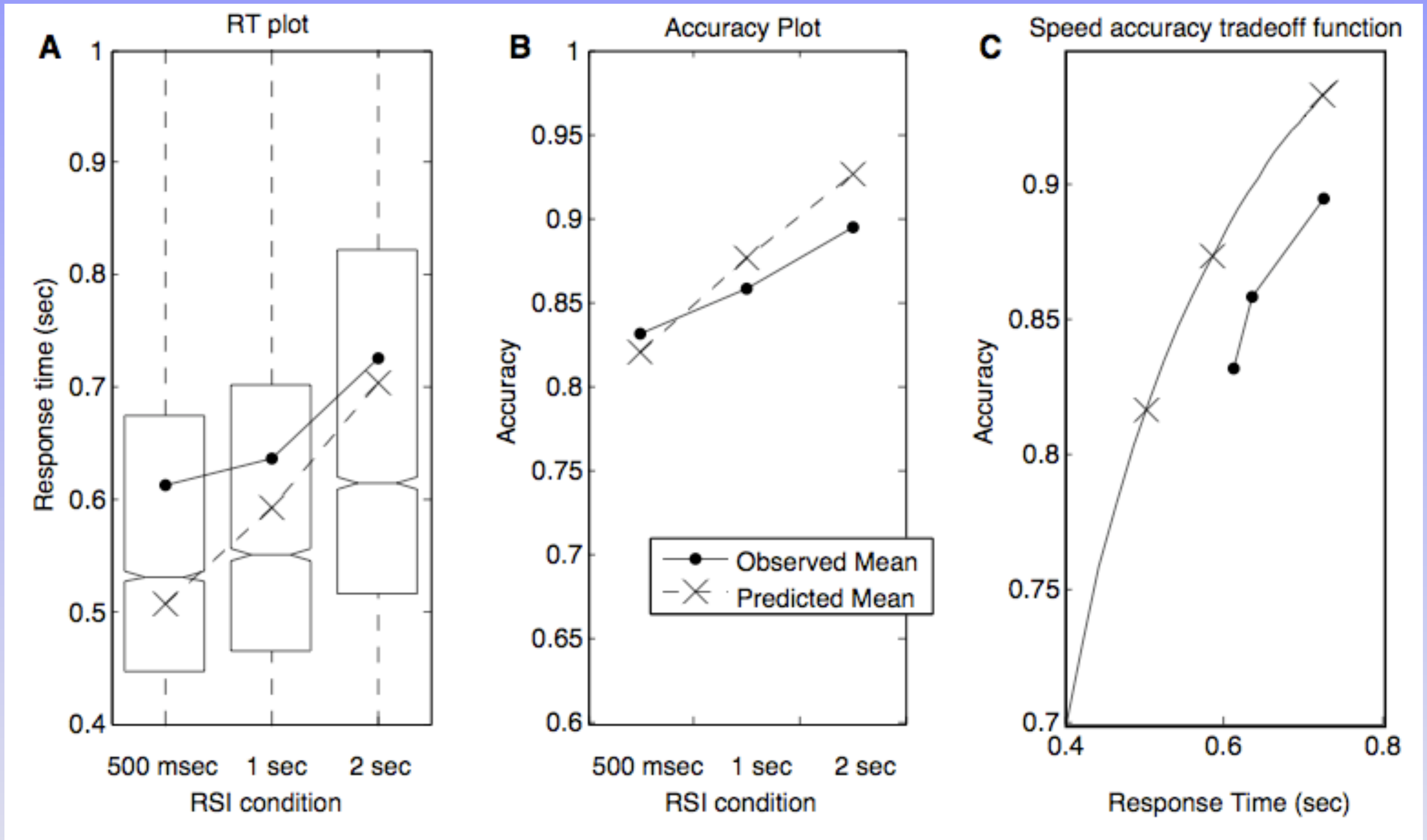
Set starting point and threshold so as to maximize expected reward (cf. Edwards, 1965 and Link, 1975)

Optimal response to changes in RSI: adapt threshold as follows



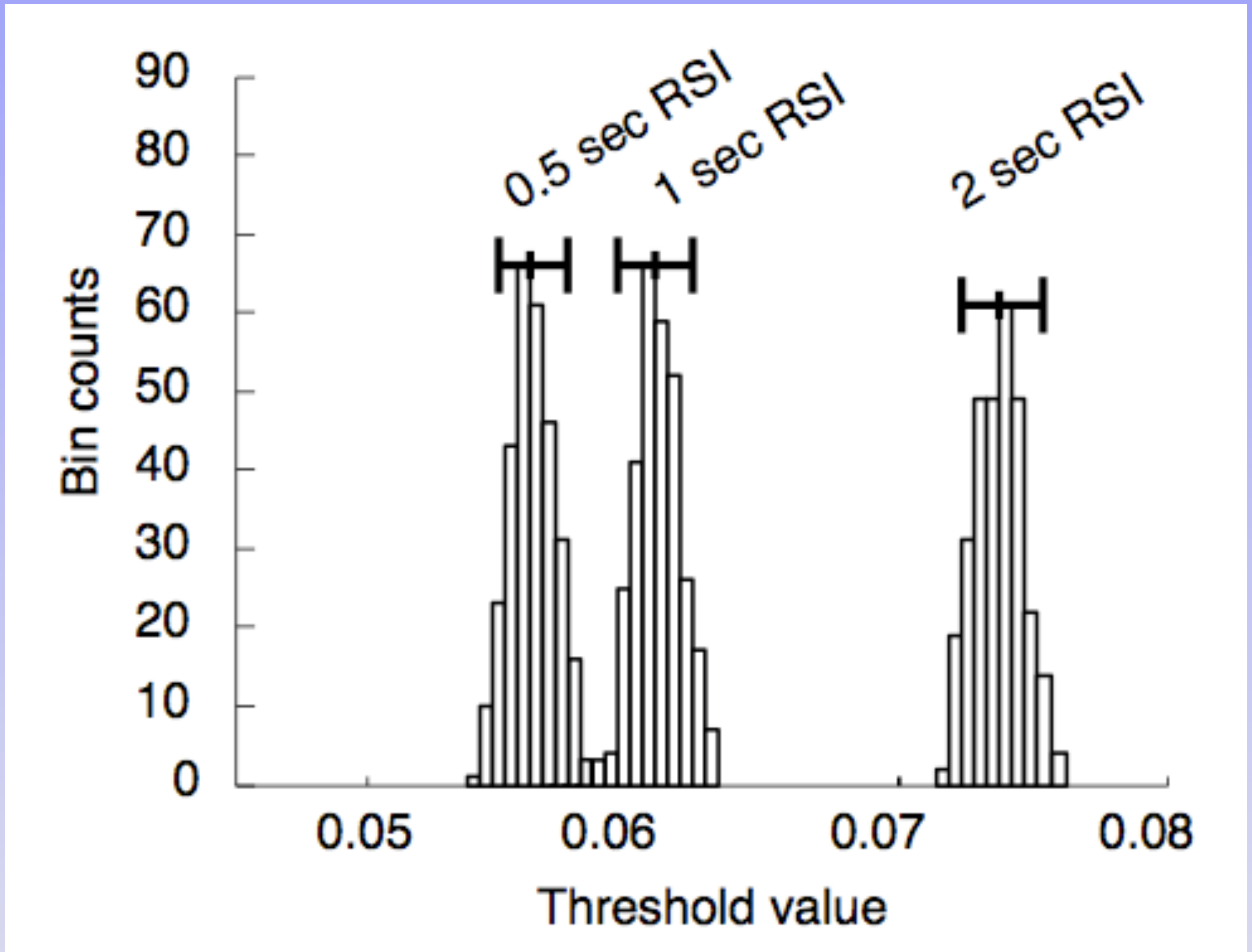
Simen, Contreras, Buck, Hu, Holmes & Cohen (in press), *Journal of Experimental Psychology: Human Perception & Performance*

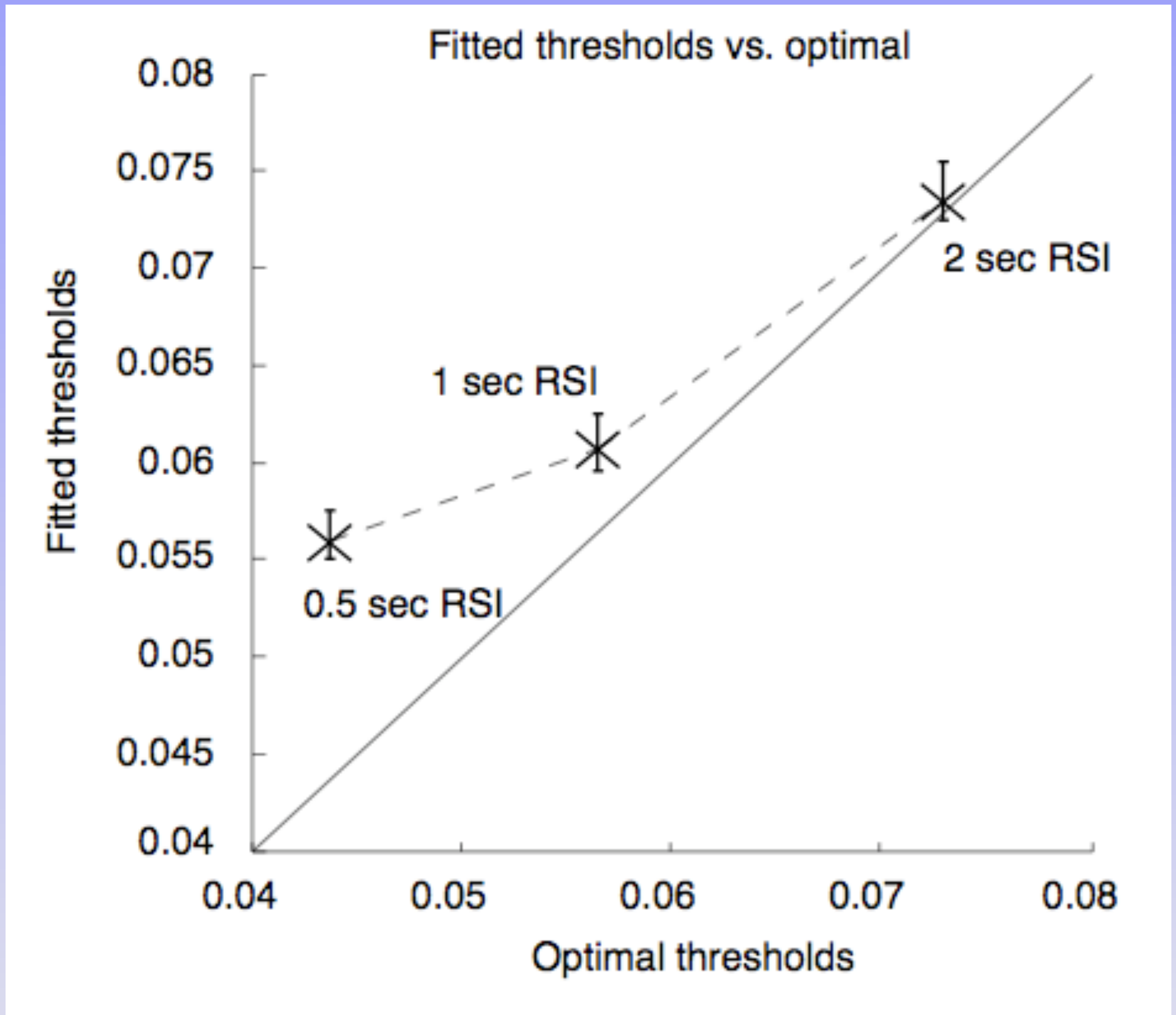
Human performance in **Expt 1**: moving dots, blocked by RSI = {0.5, 1, 2} sec, left and right button press responses



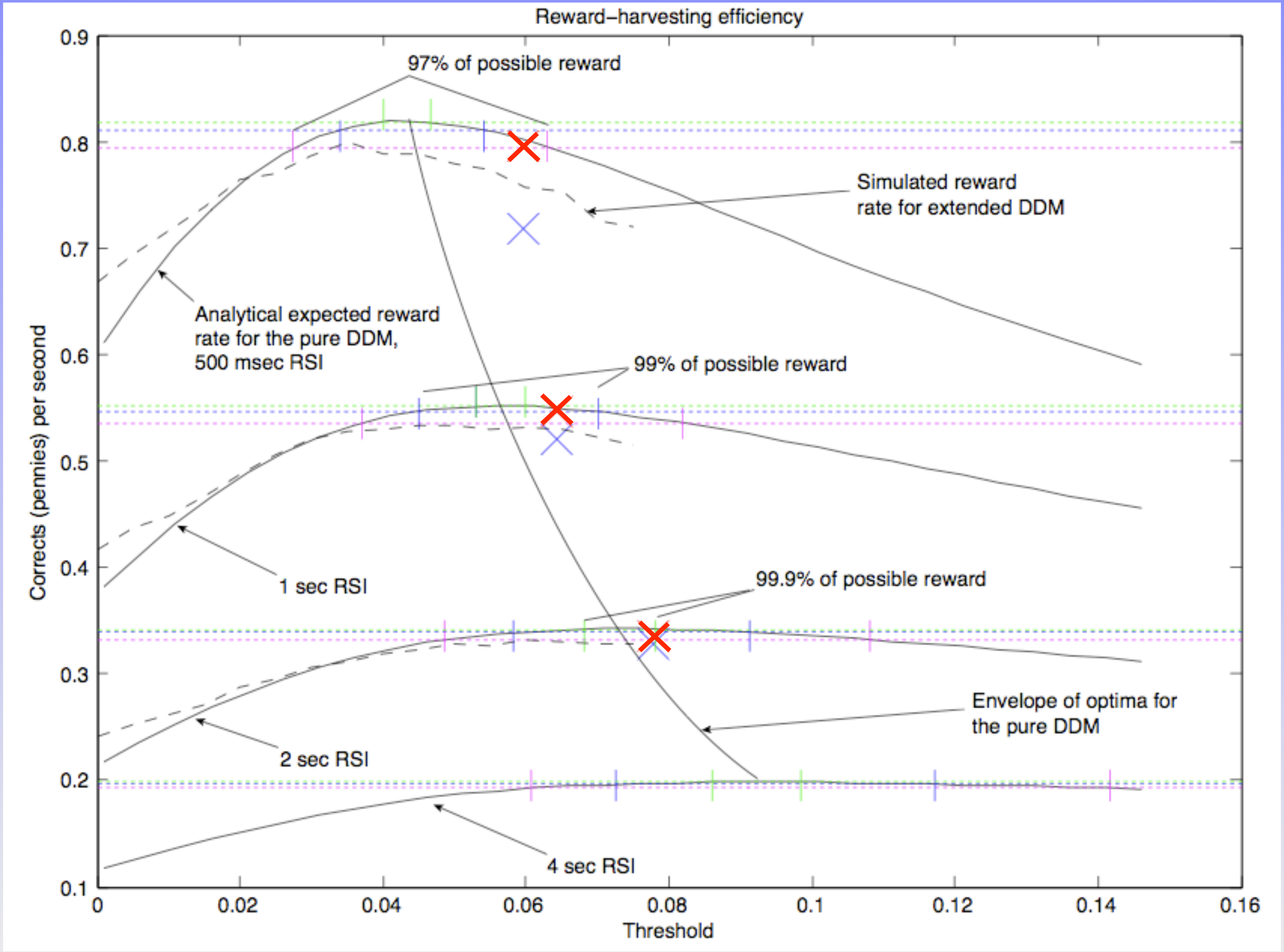
Simen, Contreras, Buck, Hu, Holmes & Cohen (in press), *Journal of Experimental Psychology: Human Perception & Performance*

Observed threshold changes in **Expt 1** vs. optimal changes

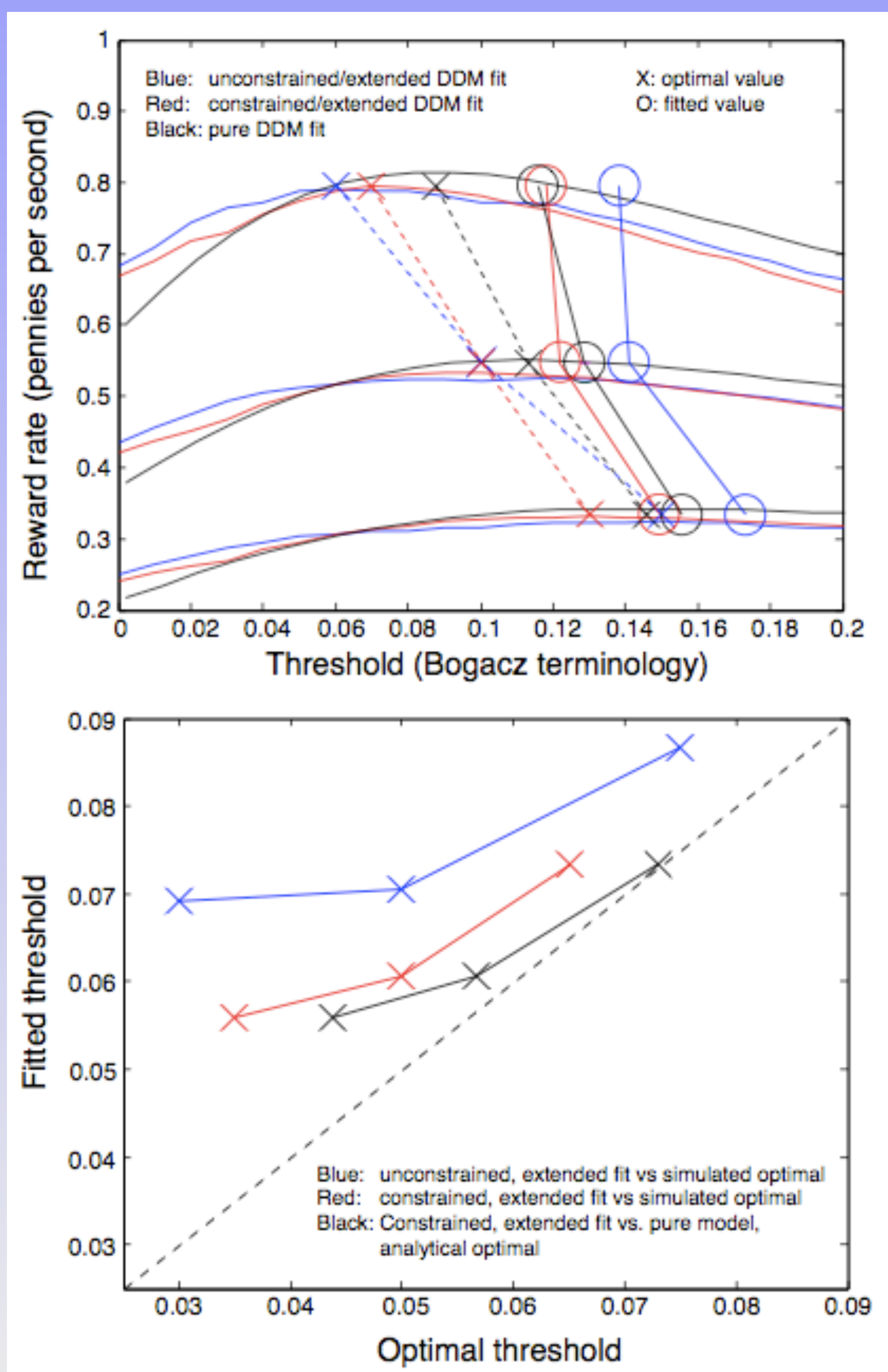


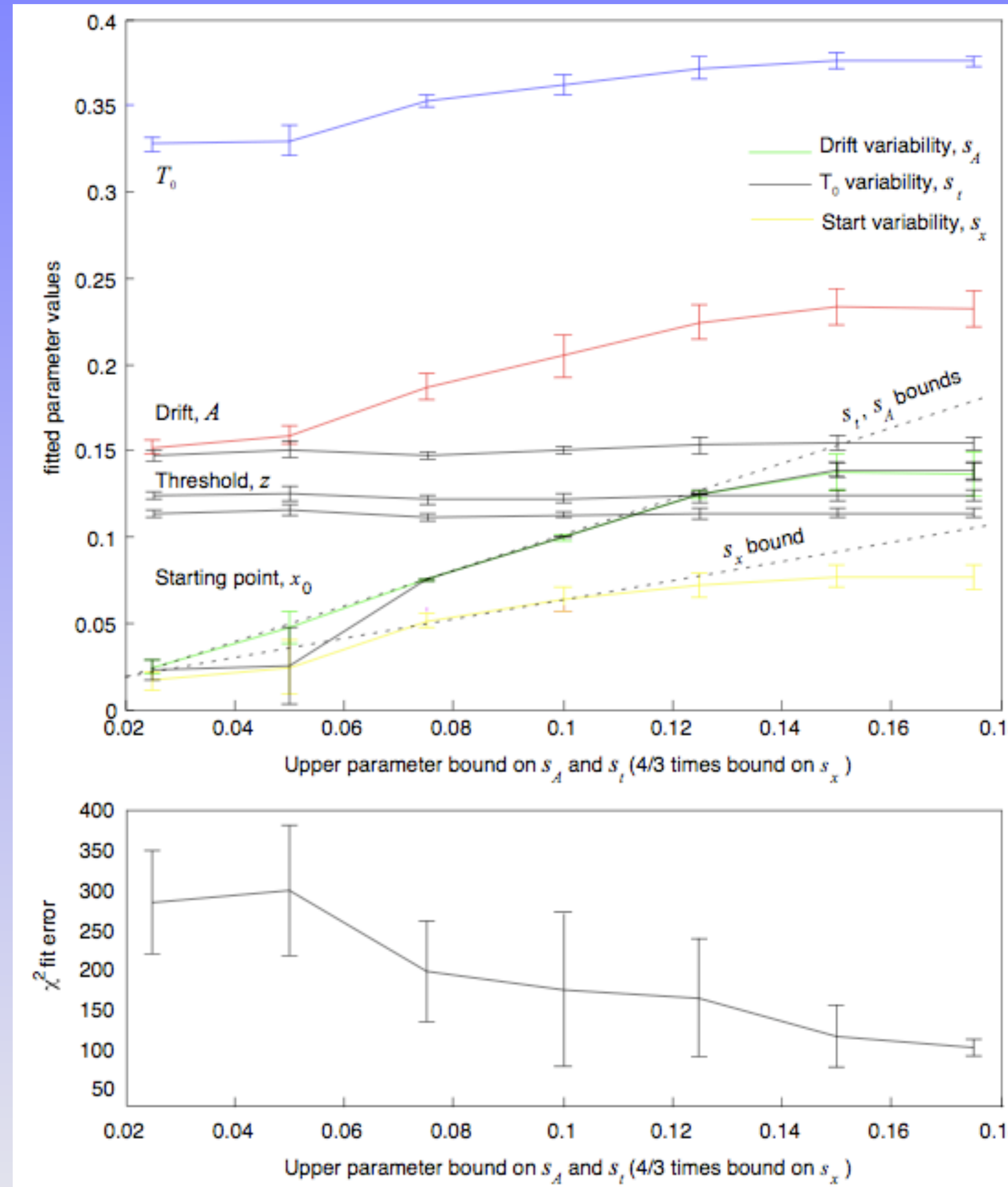


Expt 1: People are very good



Constrained fitting (bounding s_t, s_A, s_z) gives better account of data than unconstrained fits

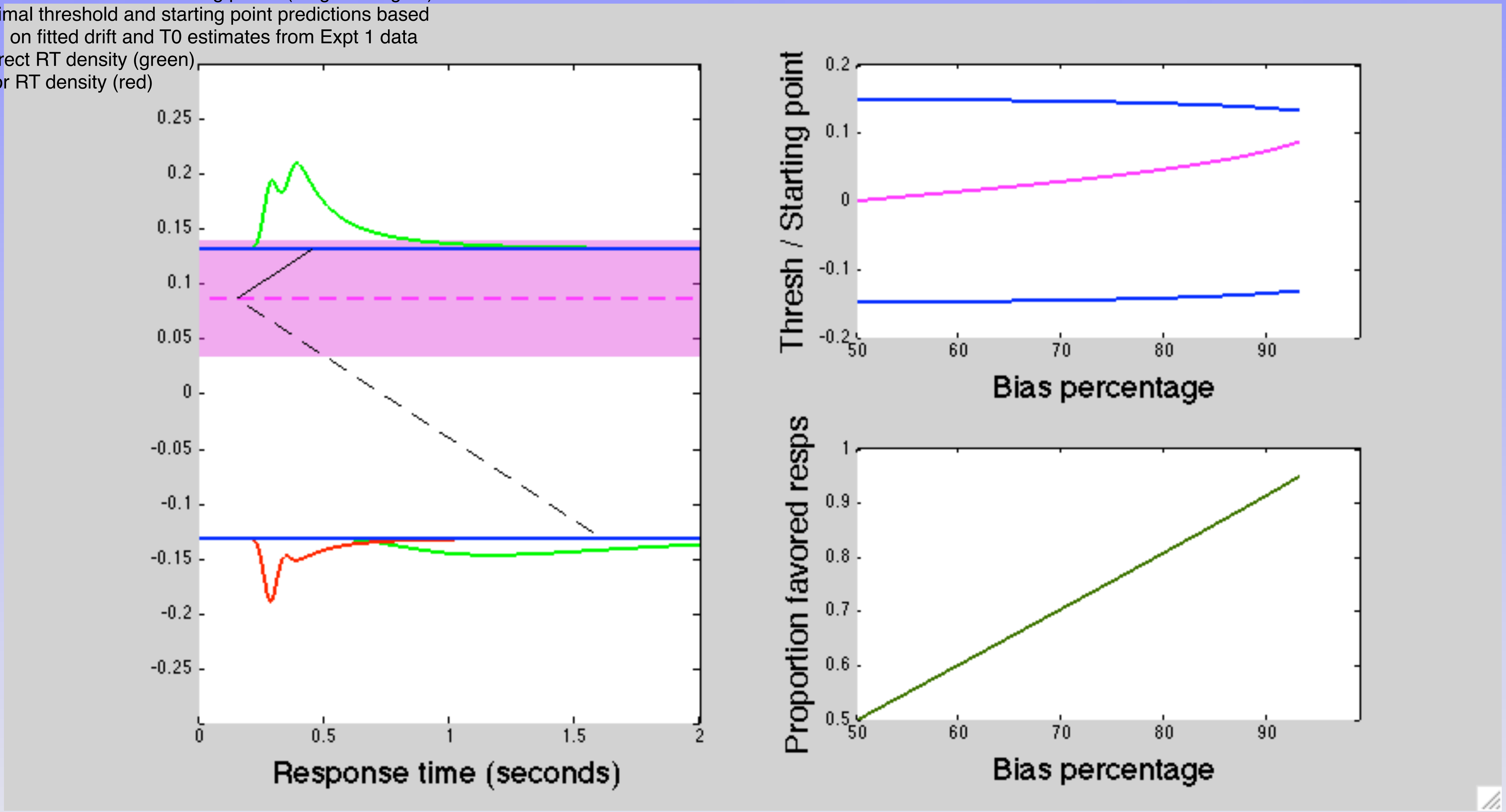




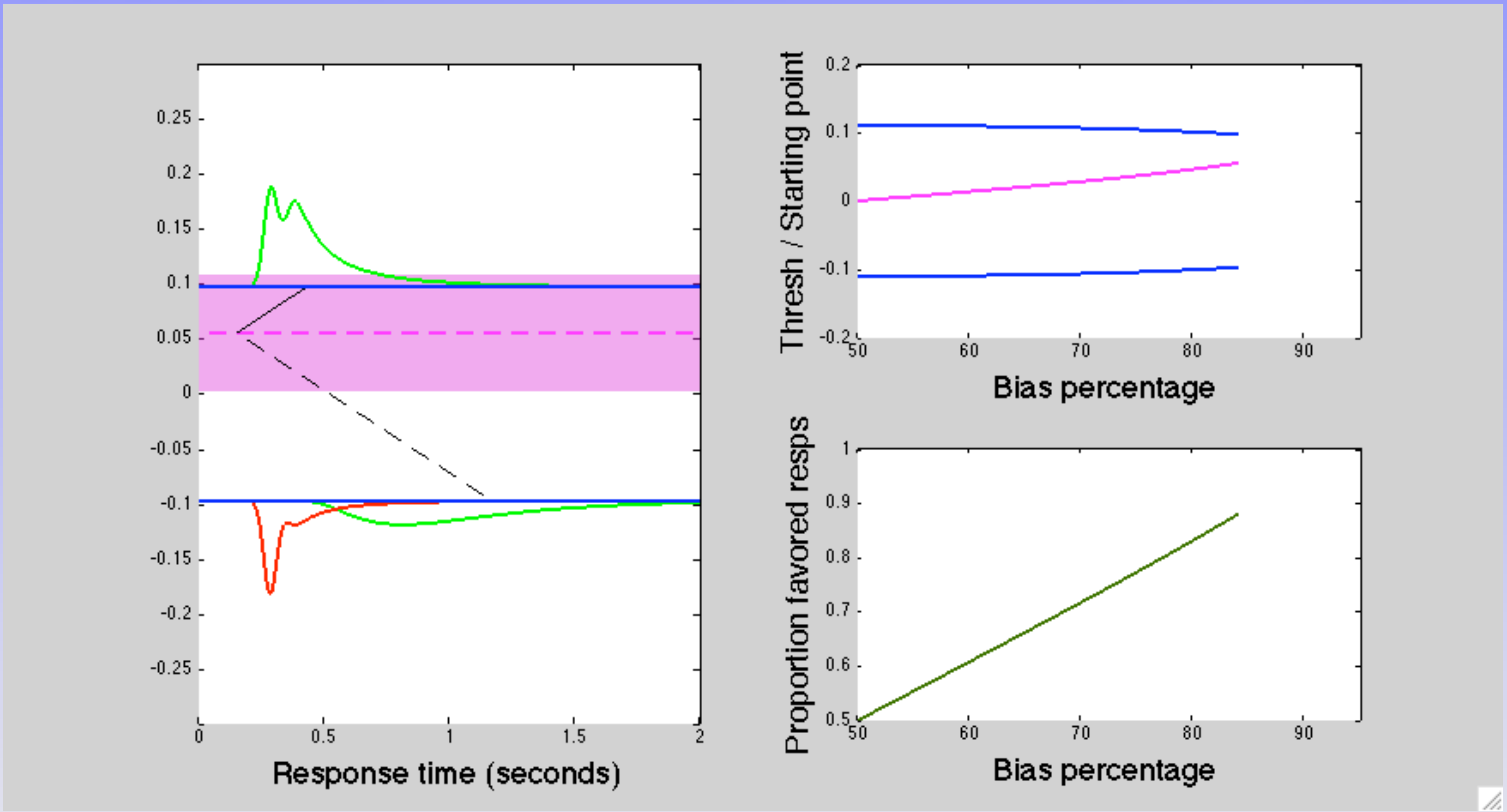
Optimal threshold and starting point parameterization and resulting behavior for unequal stimulus odds (50:50 --> 99:1)

RSI = 2 sec

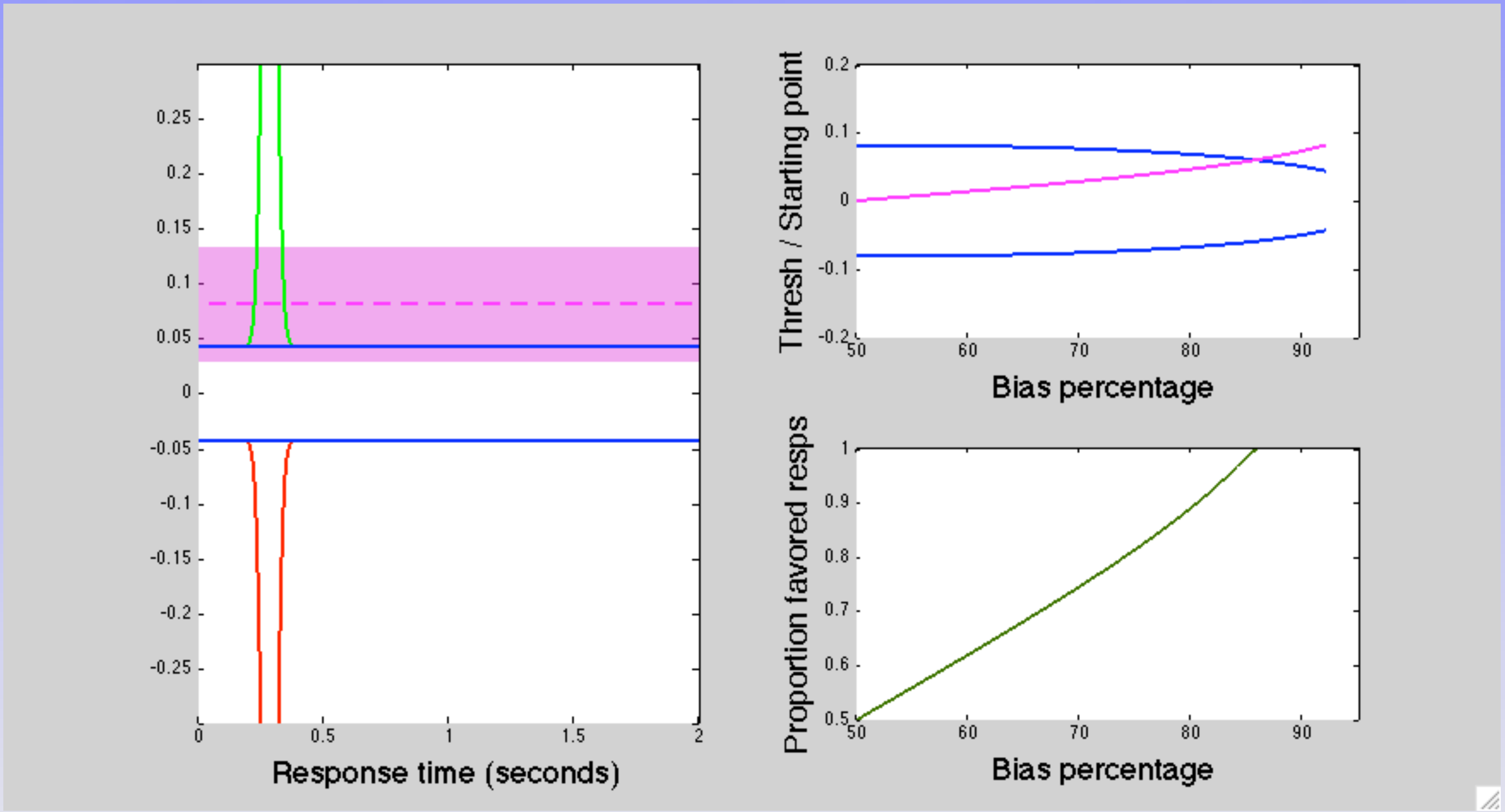
Optimal threshold (blue)
Optimal starting point (magenta dashed line)
Uniform distribution of starting points (magenta region)
Optimal threshold and starting point predictions based
on fitted drift and T0 estimates from Expt 1 data
Correct RT density (green)
Error RT density (red)



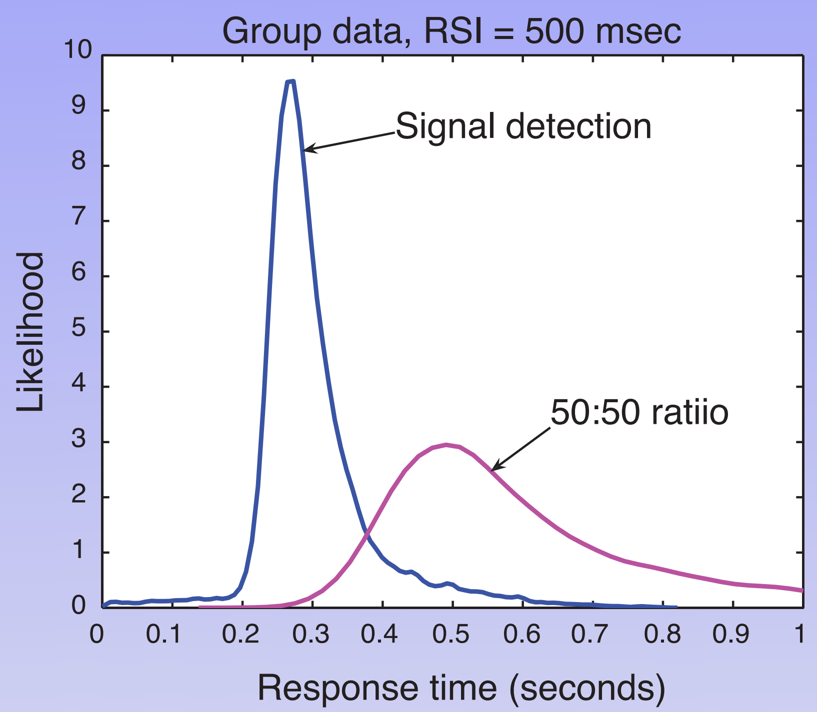
Unequal stimulus odds (50:50 --> 99:1)
RSI = 1 sec



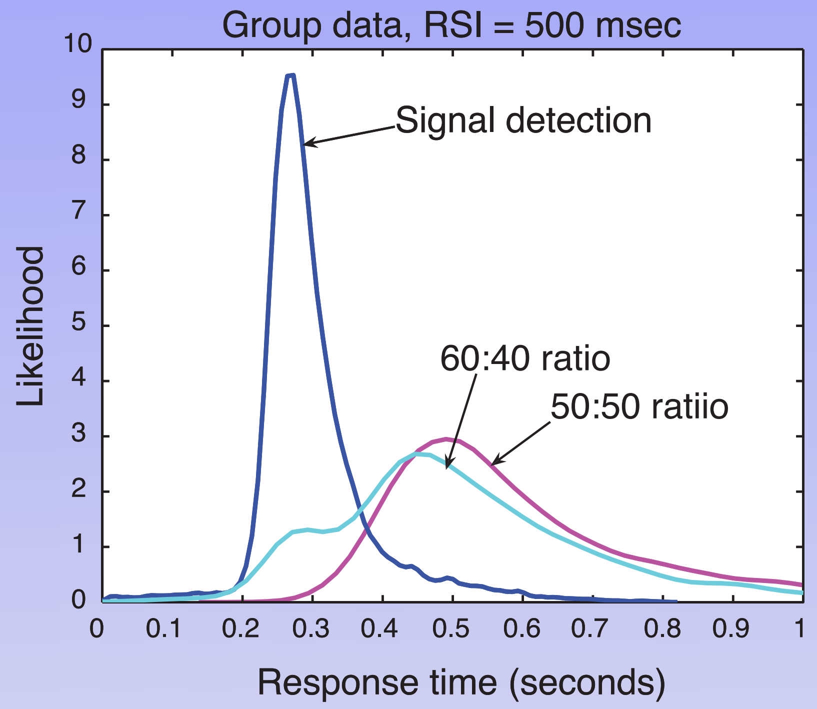
Unequal stimulus odds (50:50 --> 99:1) RSI = 0.5 sec



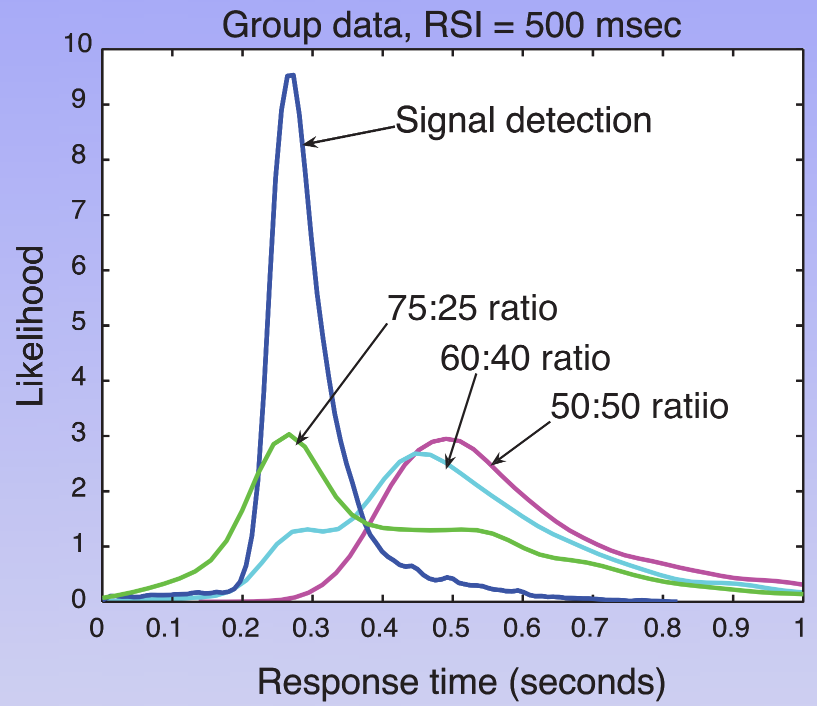
Empirical RTs for **Expt. 2**: 10 subjects, RSI = {0.5, 1, 2} sec x Bias = {60:40, 75:25, 90:10}, compared to signal detection RTs (blue)



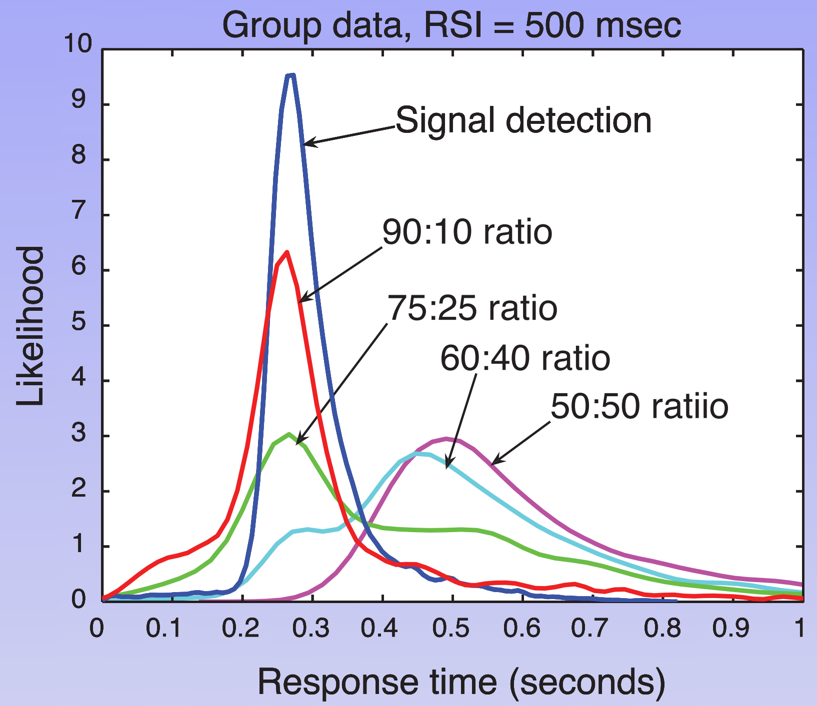
Empirical RTs for **Expt. 2**: 10 subjects, RSI = {0.5, 1, 2} sec x Bias = {60:40, 75:25, 90:10}, compared to signal detection RTs (blue)



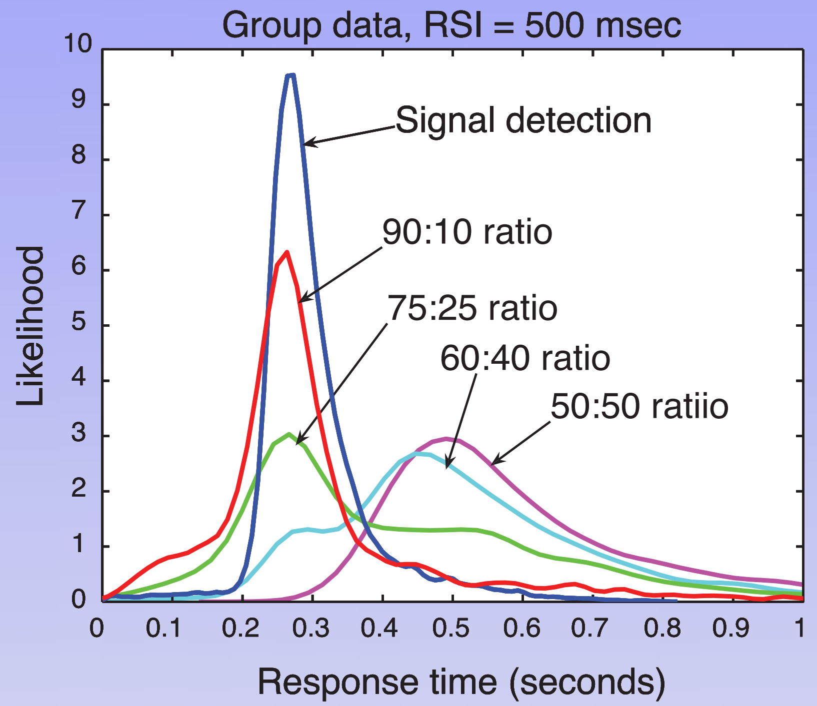
Empirical RTs for **Expt. 2**: 10 subjects, RSI = {0.5, 1, 2} sec x Bias = {60:40, 75:25, 90:10}, compared to signal detection RTs (blue)

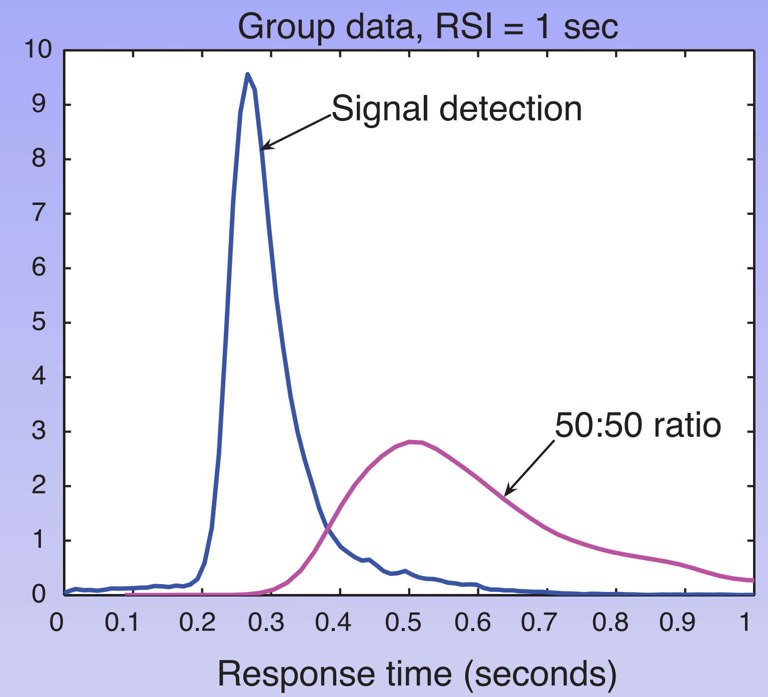


Empirical RTs for **Expt. 2**: 10 subjects, RSI = {0.5, 1, 2} sec x Bias = {60:40, 75:25, 90:10}, compared to signal detection RTs (blue)

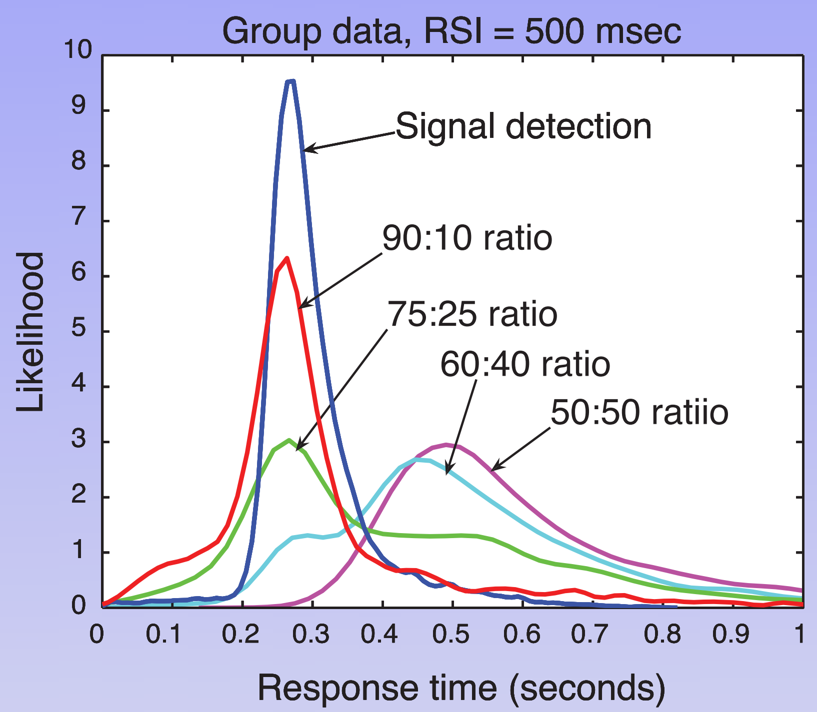


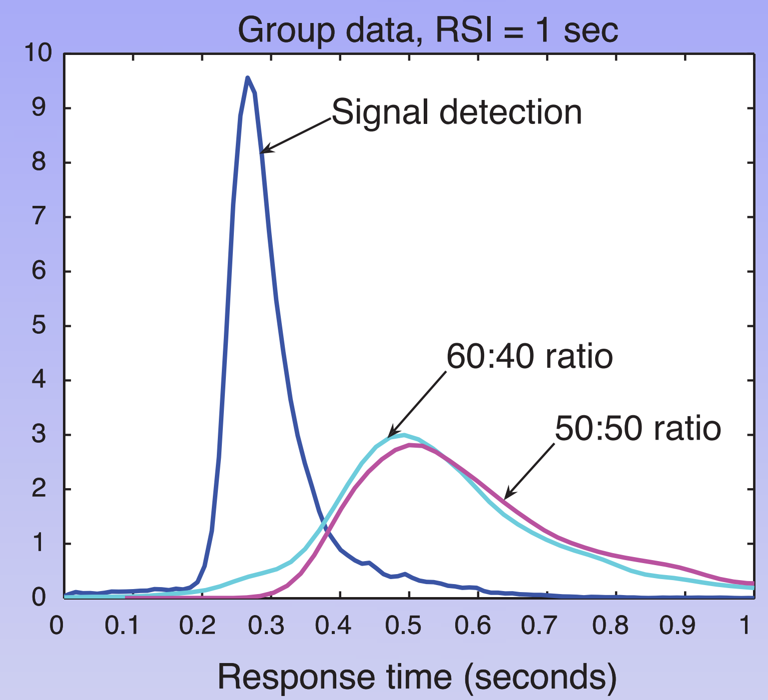
Empirical RTs for **Expt. 2**: 10 subjects, RSI = {0.5, 1, 2} sec x Bias = {60:40, 75:25, 90:10}, compared to signal detection RTs (blue)



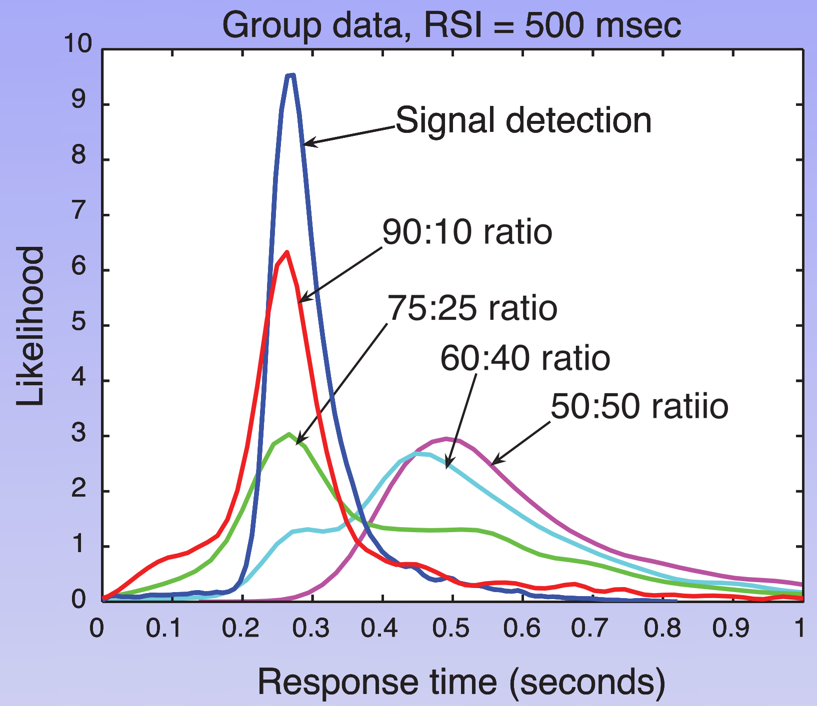


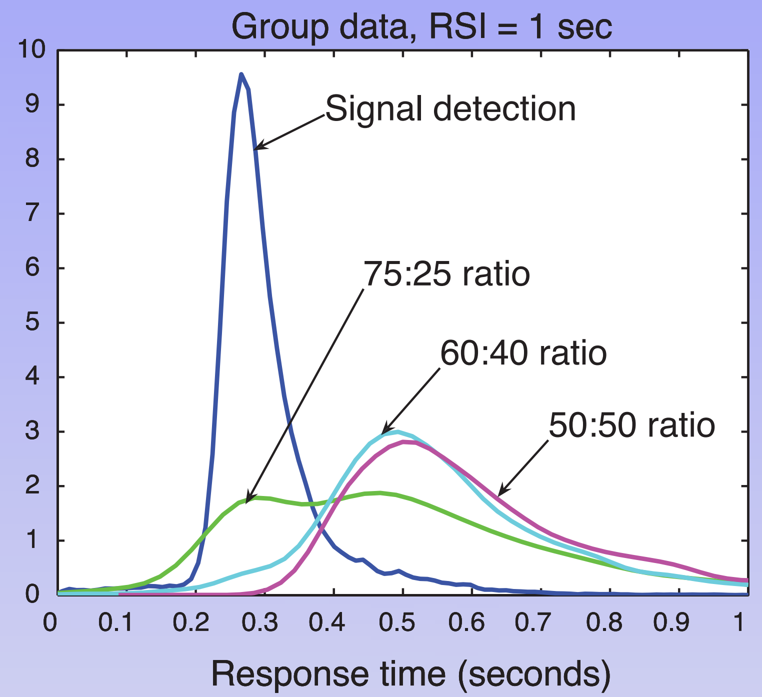
Empirical RTs for **Expt. 2**: 10 subjects, RSI = {0.5, 1, 2} sec x Bias = {60:40, 75:25, 90:10}, compared to signal detection RTs (blue)



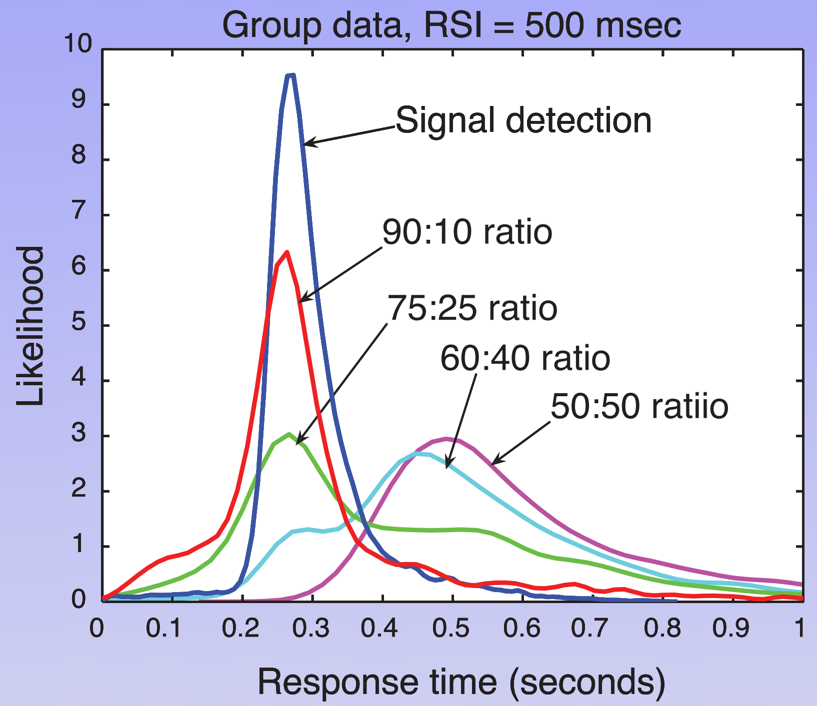


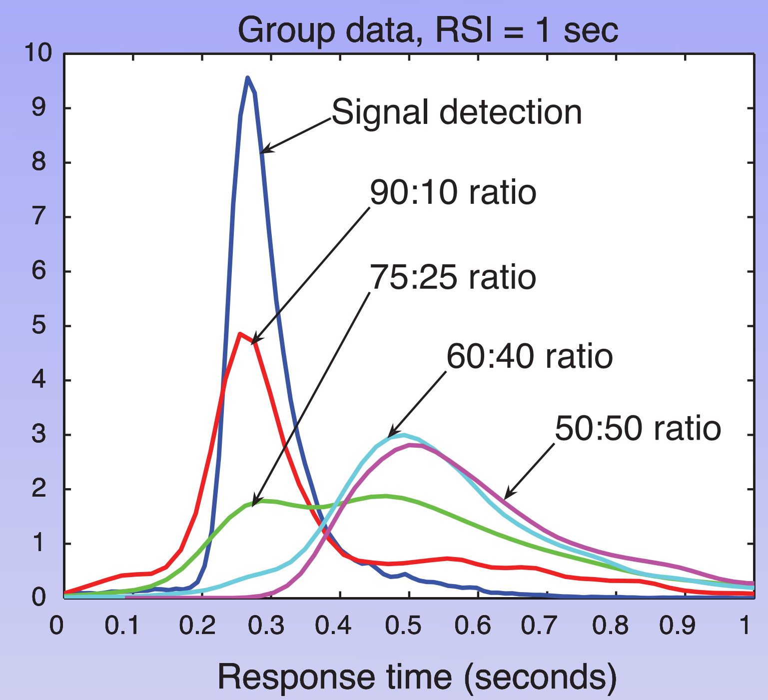
Empirical RTs for **Expt. 2**: 10 subjects, RSI = {0.5, 1, 2} sec x Bias = {60:40, 75:25, 90:10}, compared to signal detection RTs (blue)



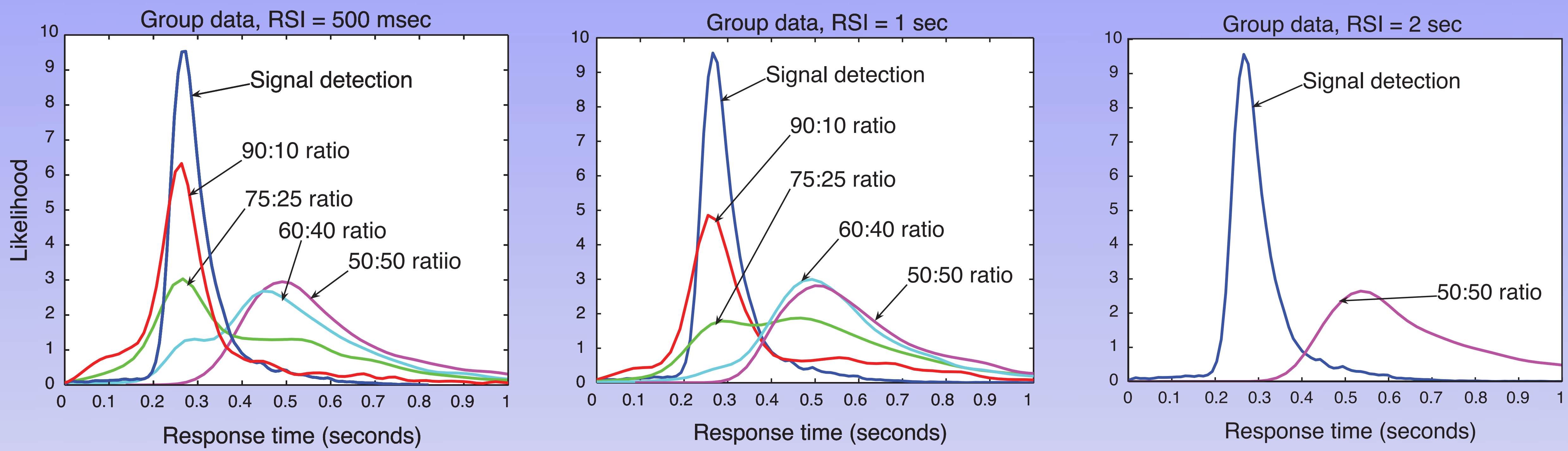


Empirical RTs for **Expt. 2**: 10 subjects, RSI = {0.5, 1, 2} sec x Bias = {60:40, 75:25, 90:10}, compared to signal detection RTs (blue)

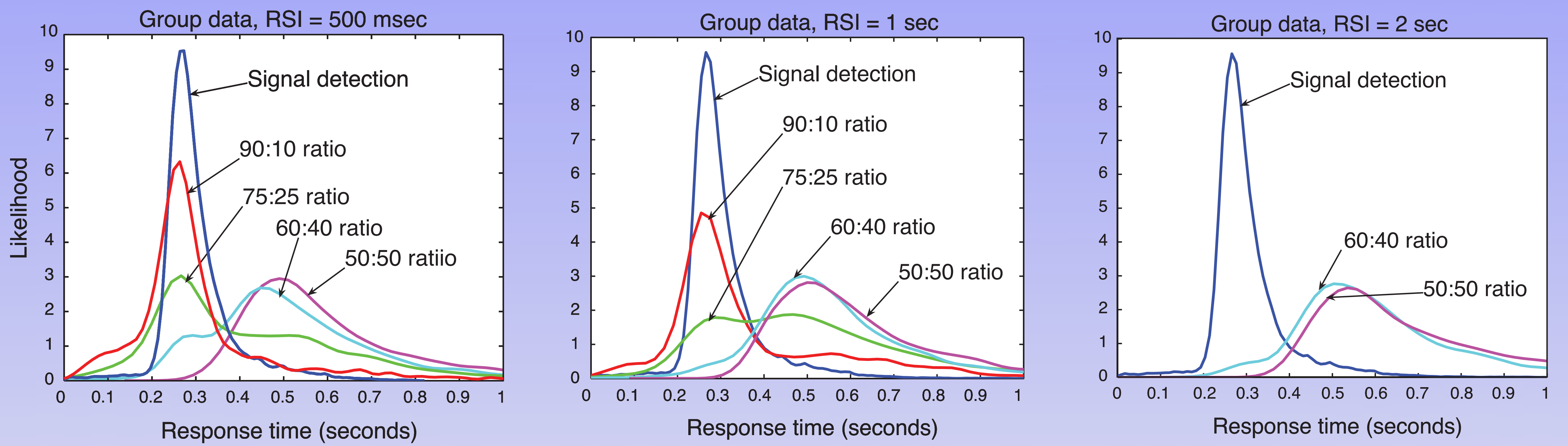




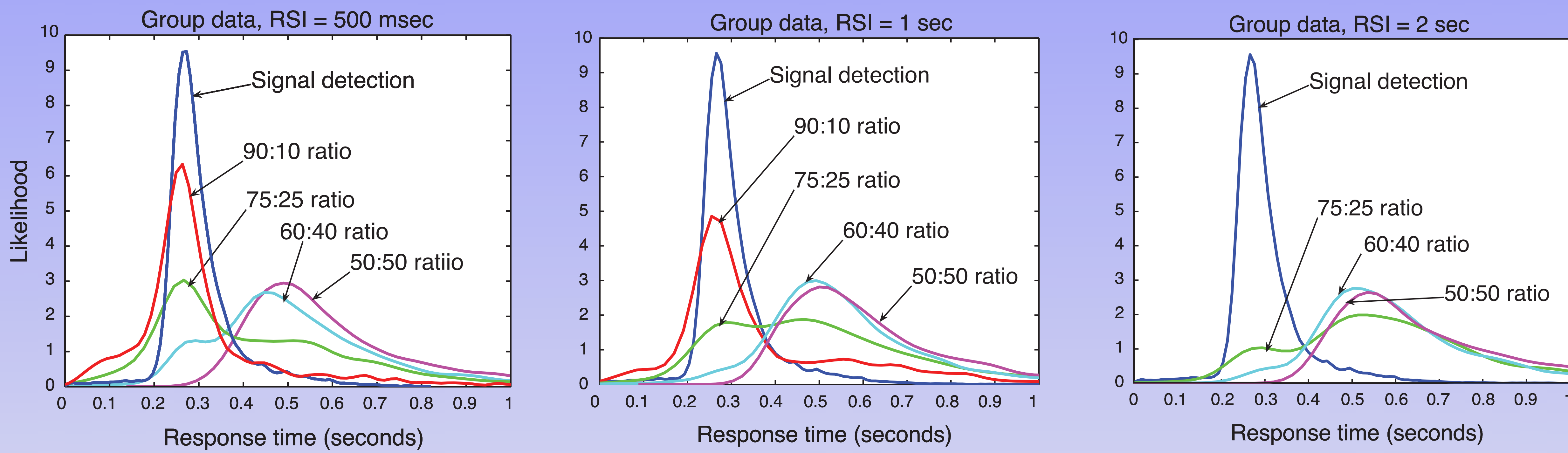
Empirical RTs for **Expt. 2**: 10 subjects, RSI = {0.5, 1, 2} sec x Bias = {60:40, 75:25, 90:10}, compared to signal detection RTs (blue)



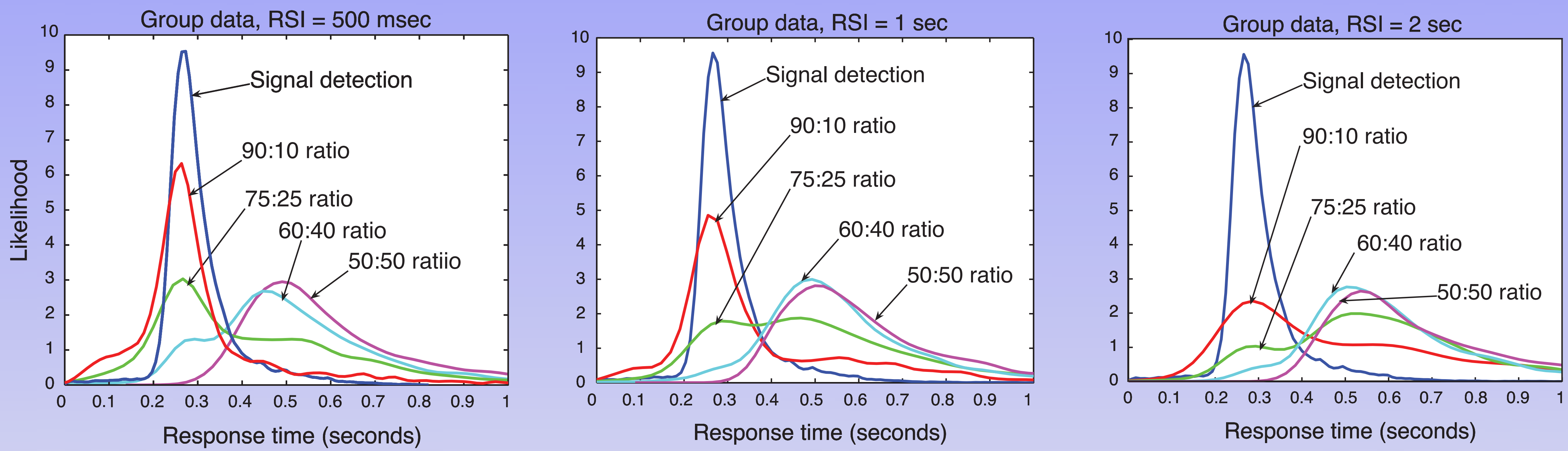
Empirical RTs for **Expt. 2**: 10 subjects, RSI = {0.5, 1, 2} sec x Bias = {60:40, 75:25, 90:10}, compared to signal detection RTs (blue)



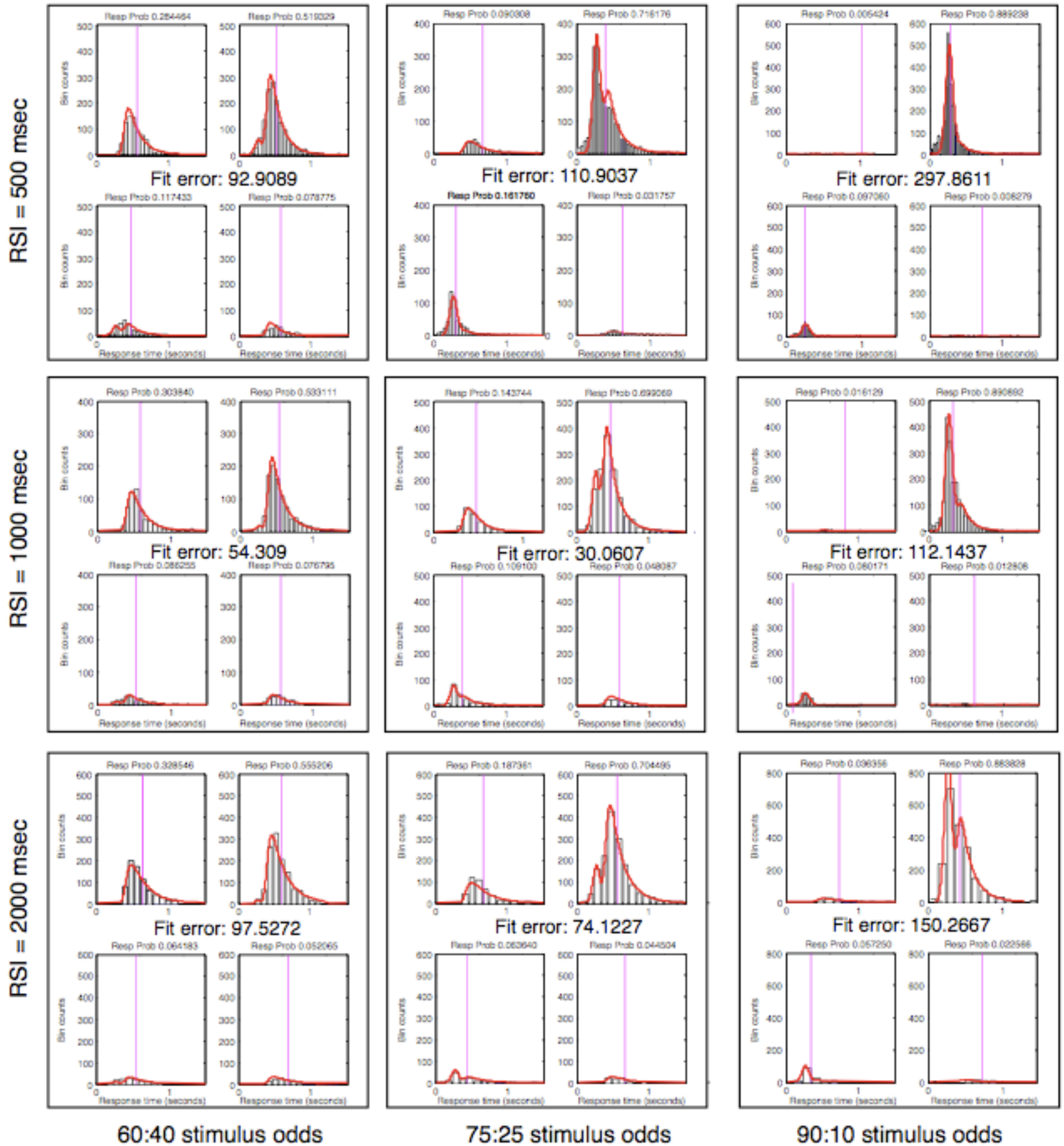
Empirical RTs for **Expt. 2**: 10 subjects, RSI = {0.5, 1, 2} sec x Bias = {60:40, 75:25, 90:10}, compared to signal detection RTs (blue)



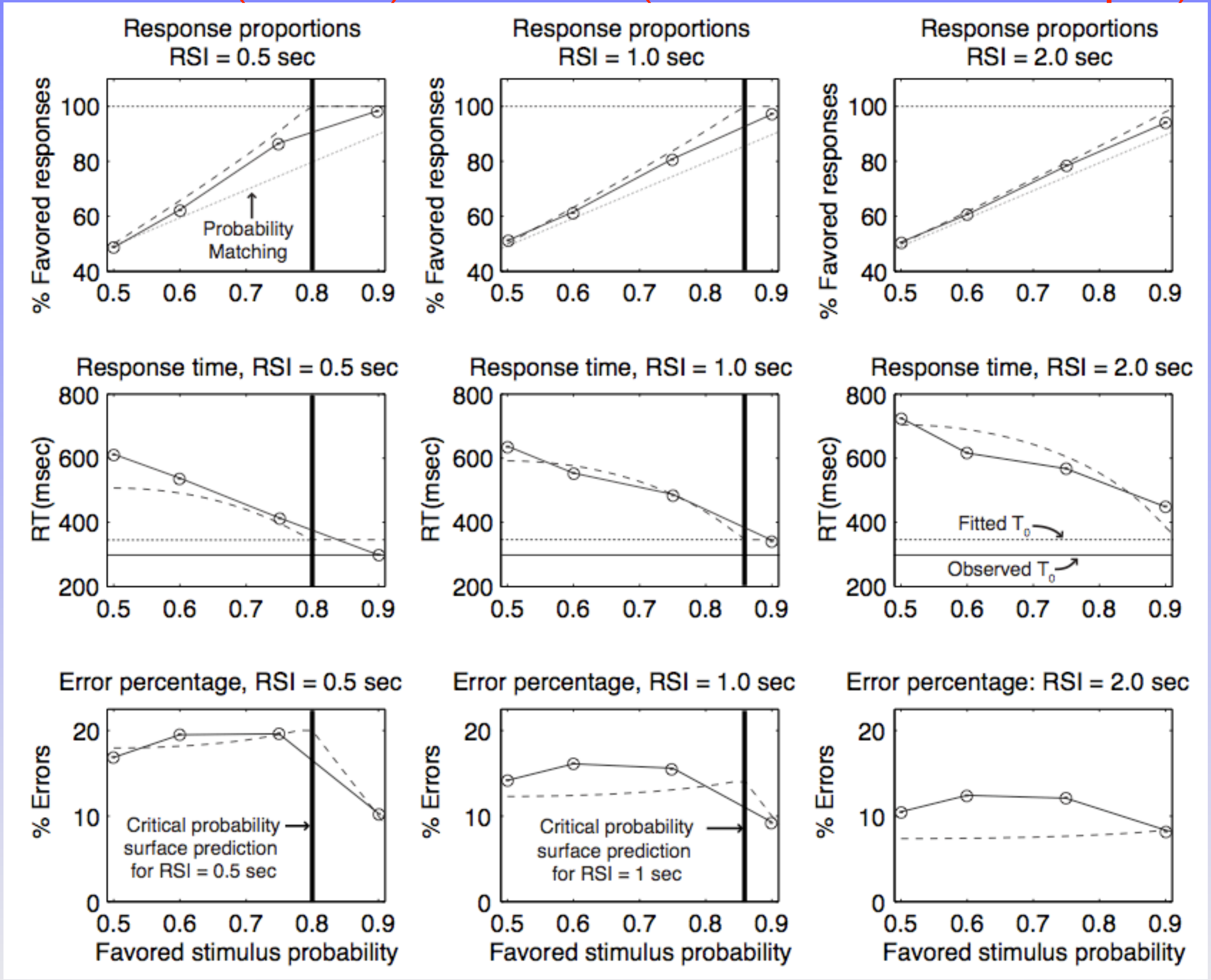
Empirical RTs for **Expt. 2**: 10 subjects, RSI = {0.5, 1, 2} sec x Bias = {60:40, 75:25, 90:10}, compared to signal detection RTs (blue)



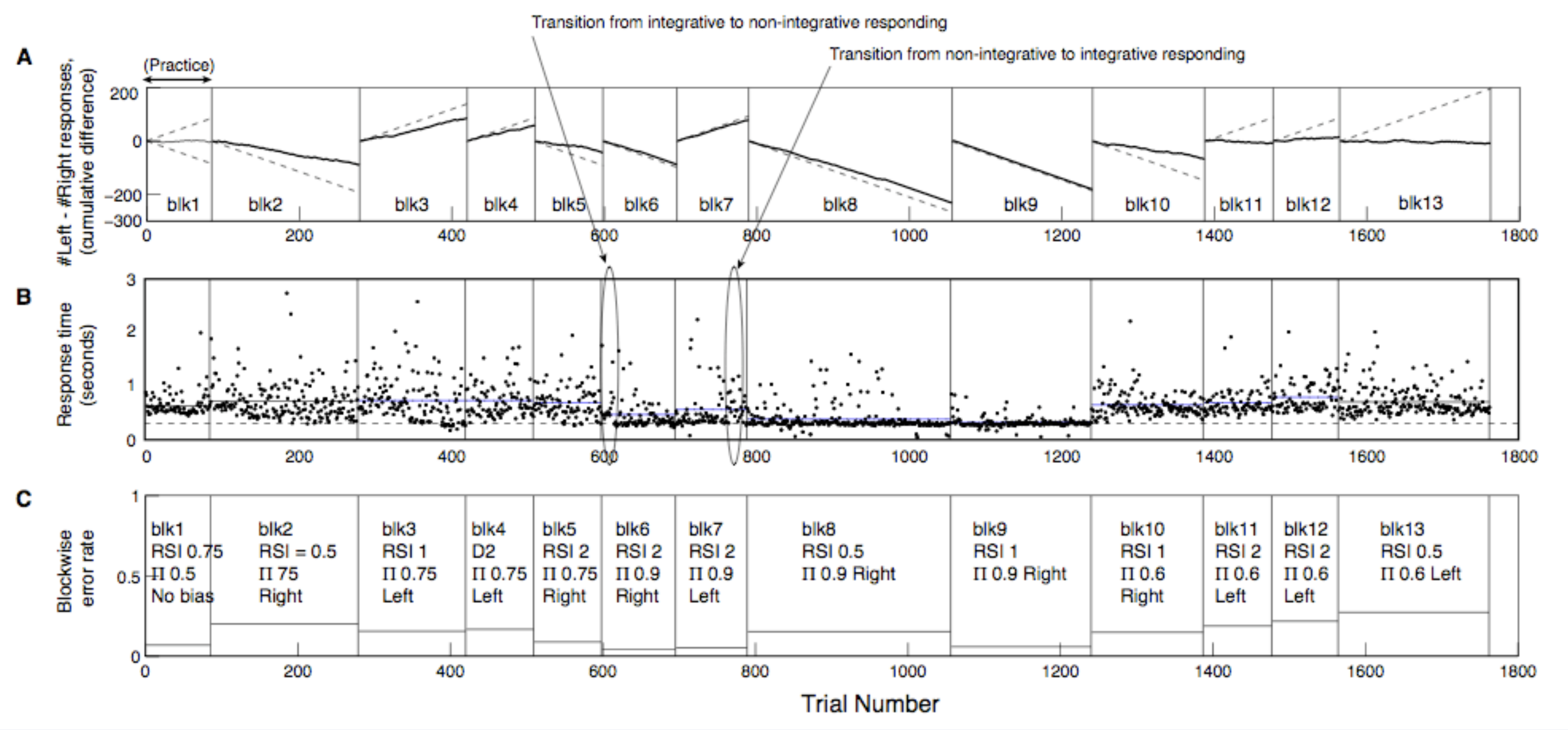
Mixture model fits (red) vs. empirical (black histograms)



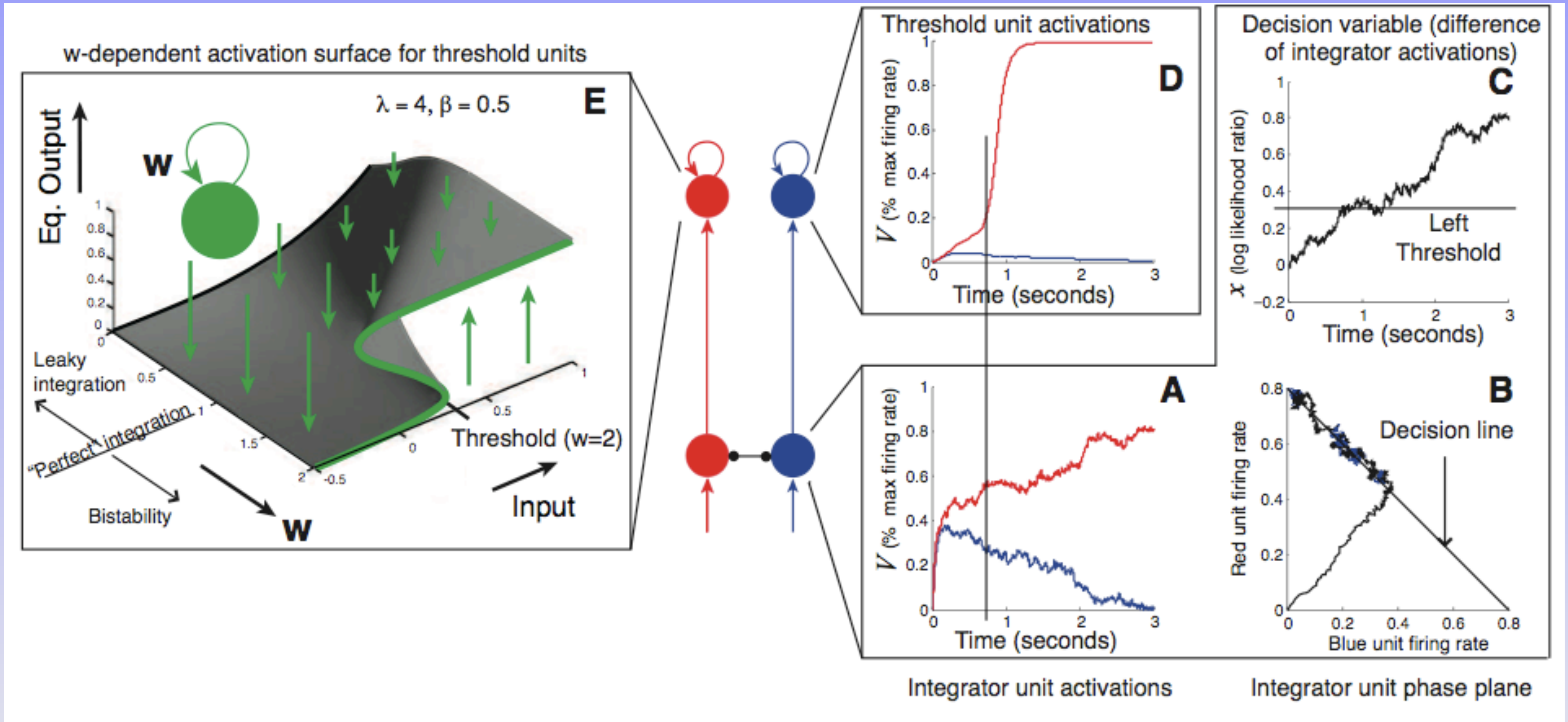
Predicted (dashed), observed (solid lines, circle data pts.)



Trial-by-trial performance, for one subject, **Expt 2**

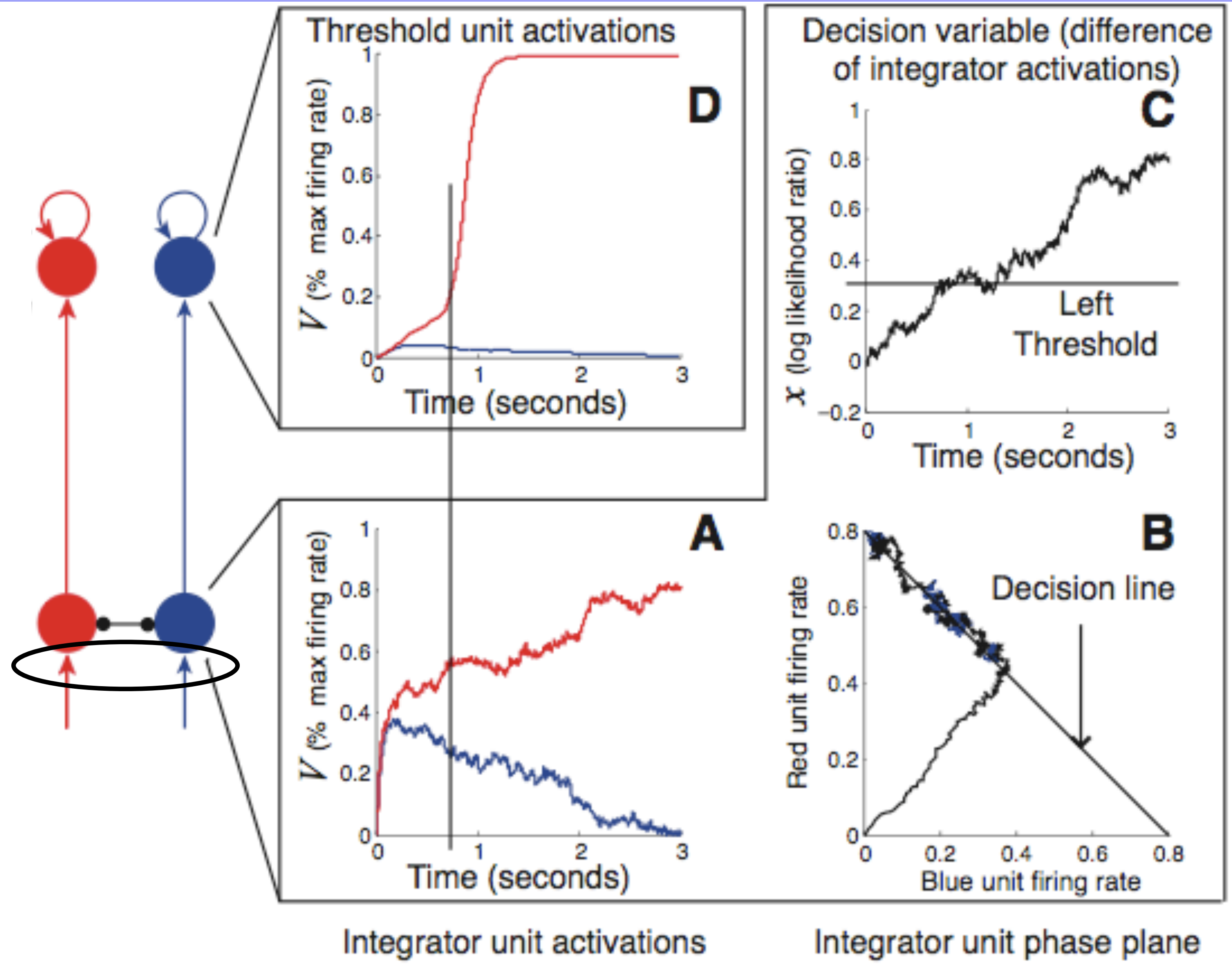


Neural circuit that also does "matching"



Simen, P. & Cohen, J. D. (in press), *Brain Research,*
"Explicit melioration by a neural diffusion model"

Neural circuit that also does “matching”



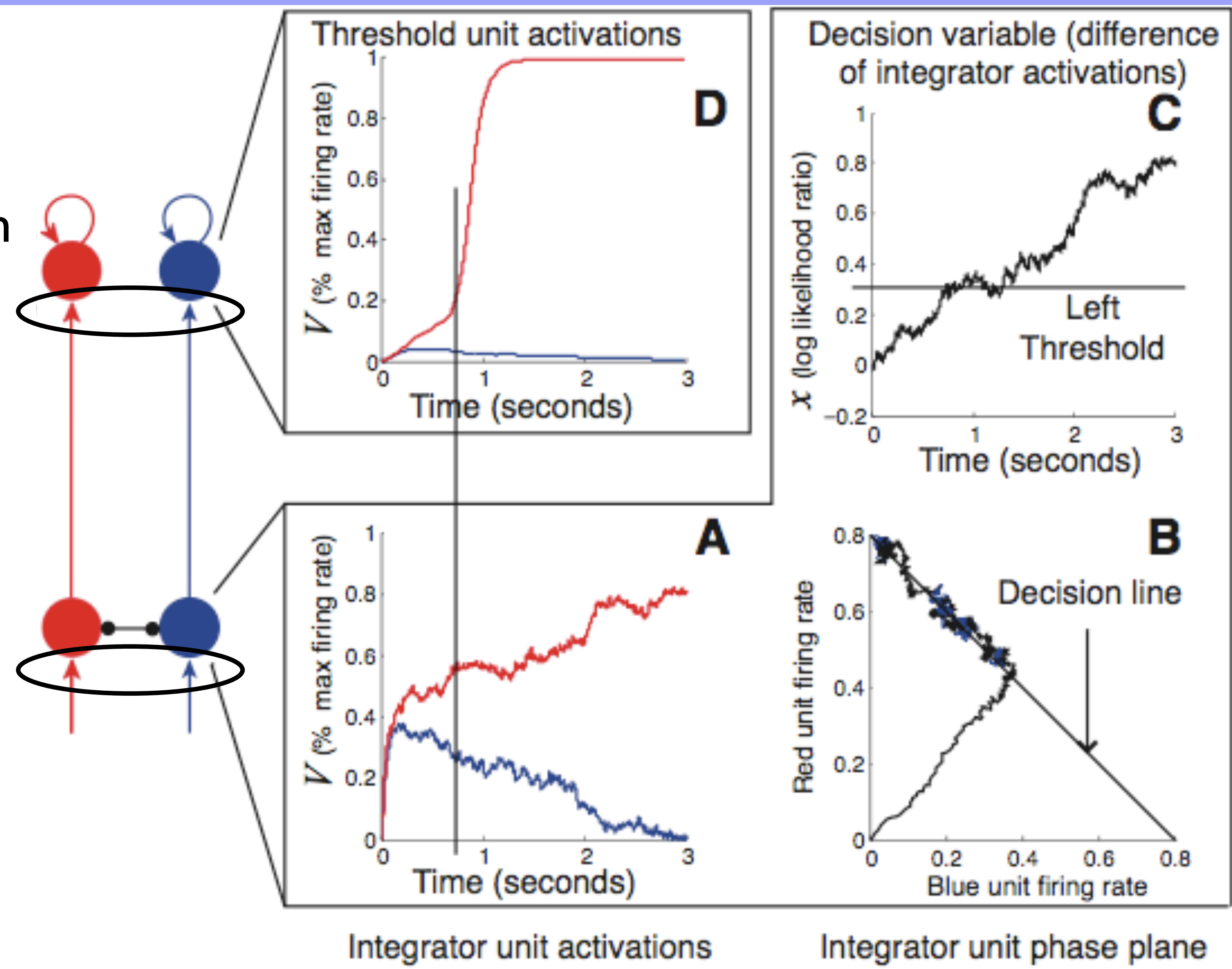
Adapting these weights modifies drift;
(with FIXED THRESHOLDS, results in
SOFTMAX choice rule;
cf. Montague & Berns, 2002)

Simen, P. & Cohen, J. D. (in press), *Brain Research,*
“Explicit melioration by a neural diffusion model”

Neural circuit that also does "matching"

Adapting these weights modifies thresholds; (with DRIFT = 0, results in MATCHING/MELIORATION)

Adapting these weights modifies drift; (with FIXED THRESHOLDS, results in SOFTMAX choice rule; cf. Montague & Berns, 2002)



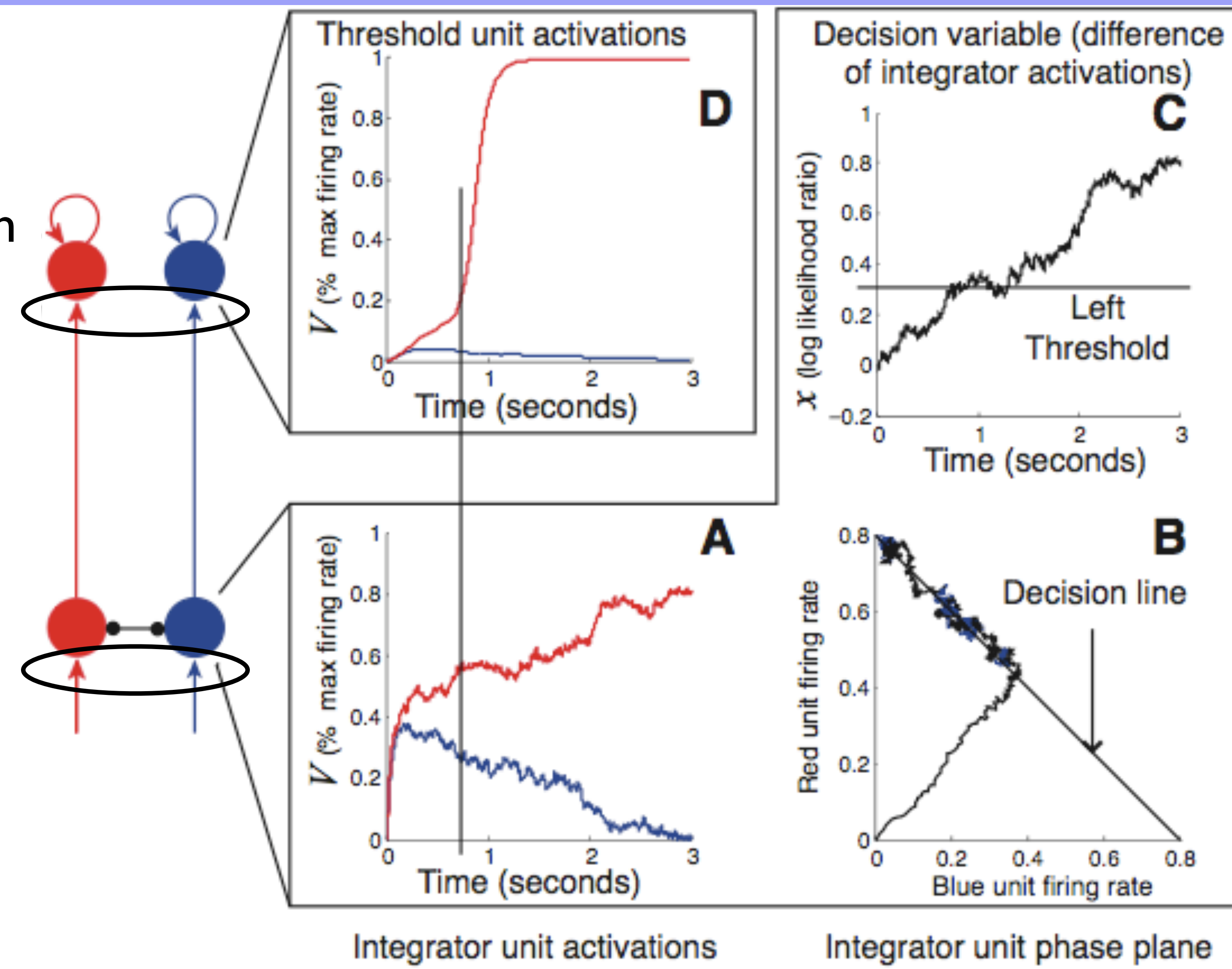
Simen, P. & Cohen, J. D. (in press), *Brain Research,* "Explicit melioration by a neural diffusion model"

Neural circuit that also does "matching"

Adapting these weights modifies thresholds; (with DRIFT = 0, results in MATCHING/MELIORATION)

$$\dot{w}_i = -w_i(t) + r_i(t)$$

Adapting these weights modifies drift; (with FIXED THRESHOLDS, results in SOFTMAX choice rule; cf. Montague & Berns, 2002)



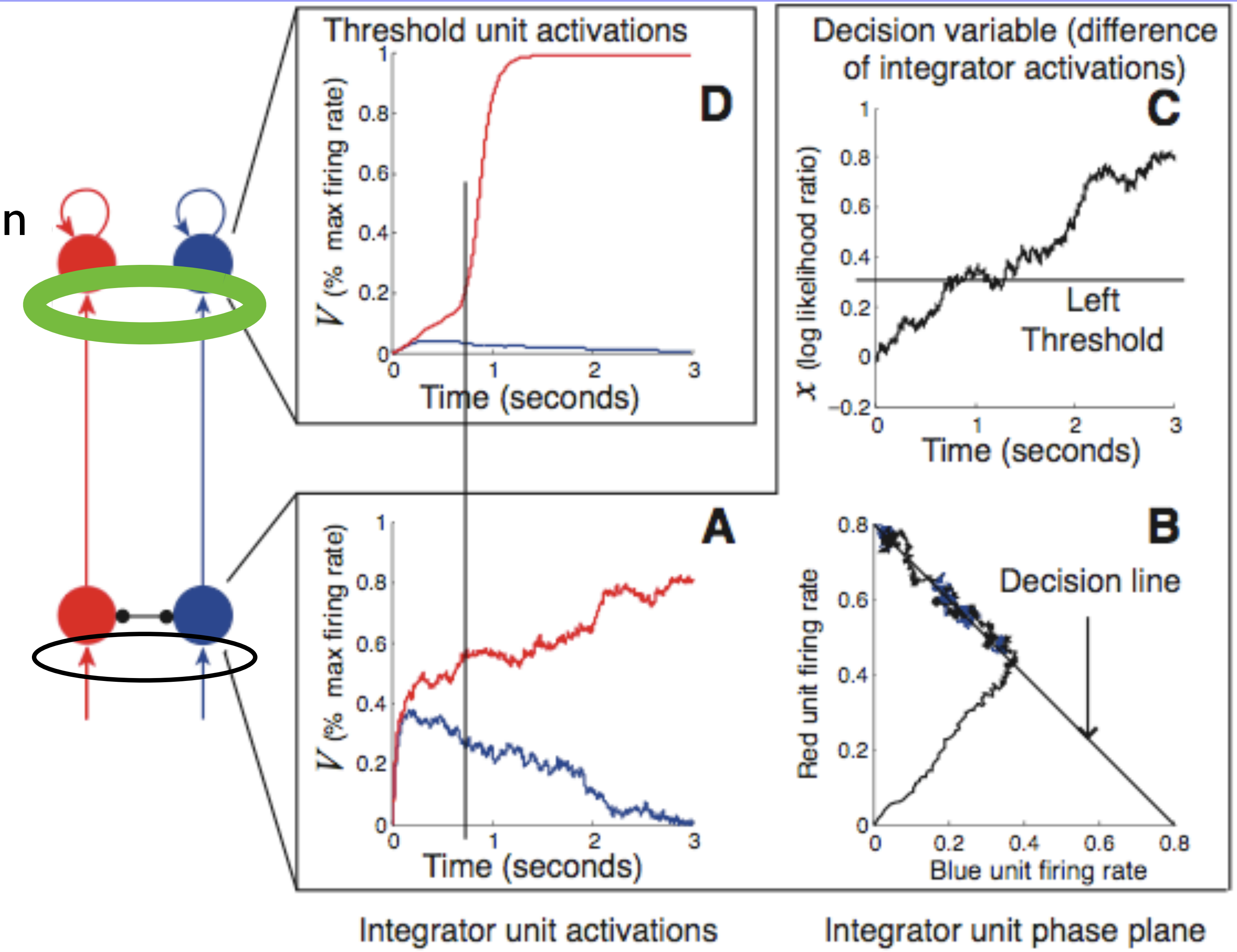
Simen, P. & Cohen, J. D. (in press), *Brain Research*, "Explicit melioration by a neural diffusion model"

Neural circuit that also does "matching"

Adapting these weights modifies thresholds; (with DRIFT = 0, results in MATCHING/MELIORATION)

$$\dot{w}_i = -w_i(t) + r_i(t)$$

Adapting these weights modifies drift; (with FIXED THRESHOLDS, results in SOFTMAX choice rule; cf. Montague & Berns, 2002)



Simen, P. & Cohen, J. D. (in press), *Brain Research,* "Explicit melioration by a neural diffusion model"

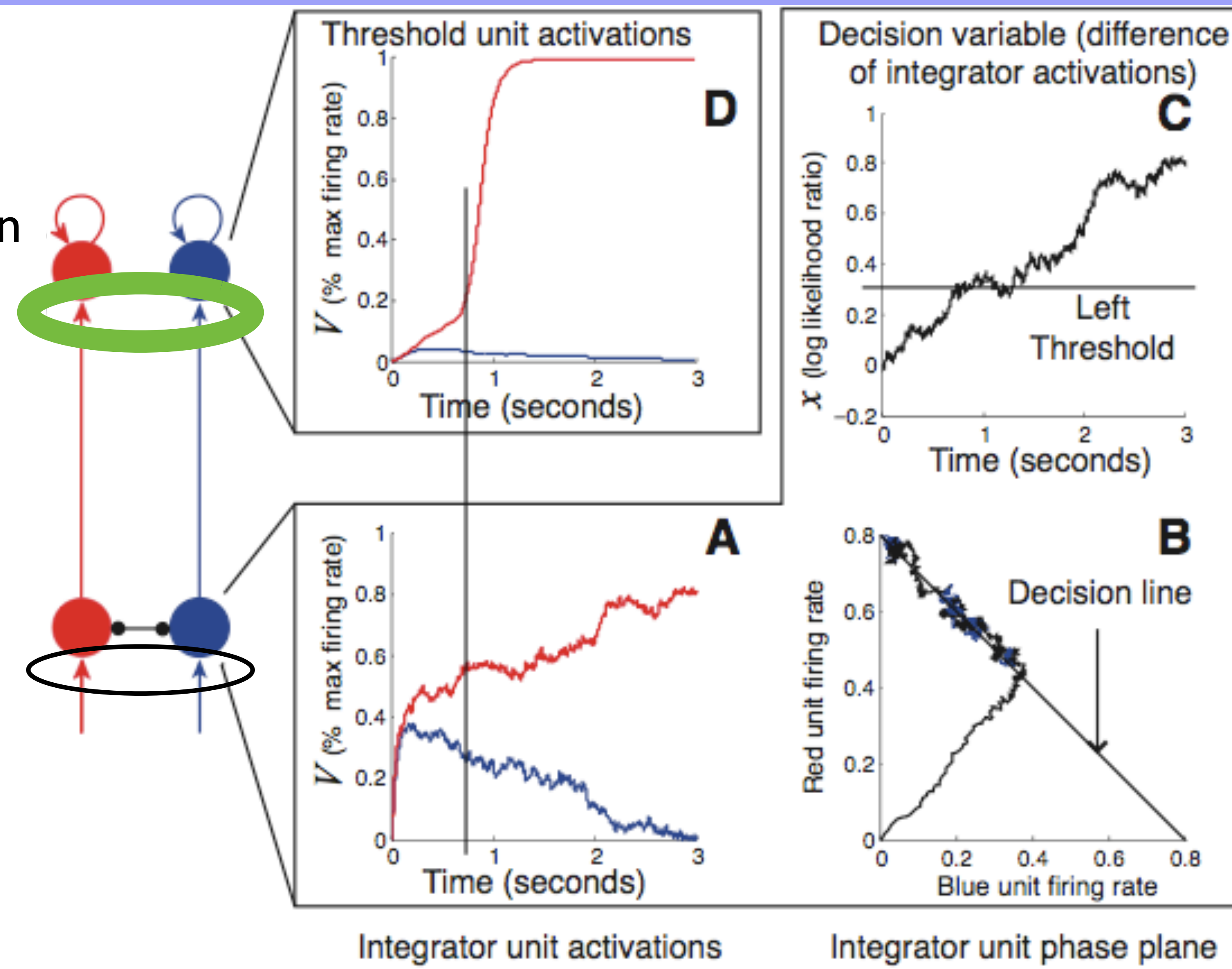
Neural circuit that also does “matching”

Adapting these weights modifies thresholds; (with DRIFT = 0, results in MATCHING/MELIORATION)

$$\dot{w}_i = -w_i(t) + r_i(t)$$

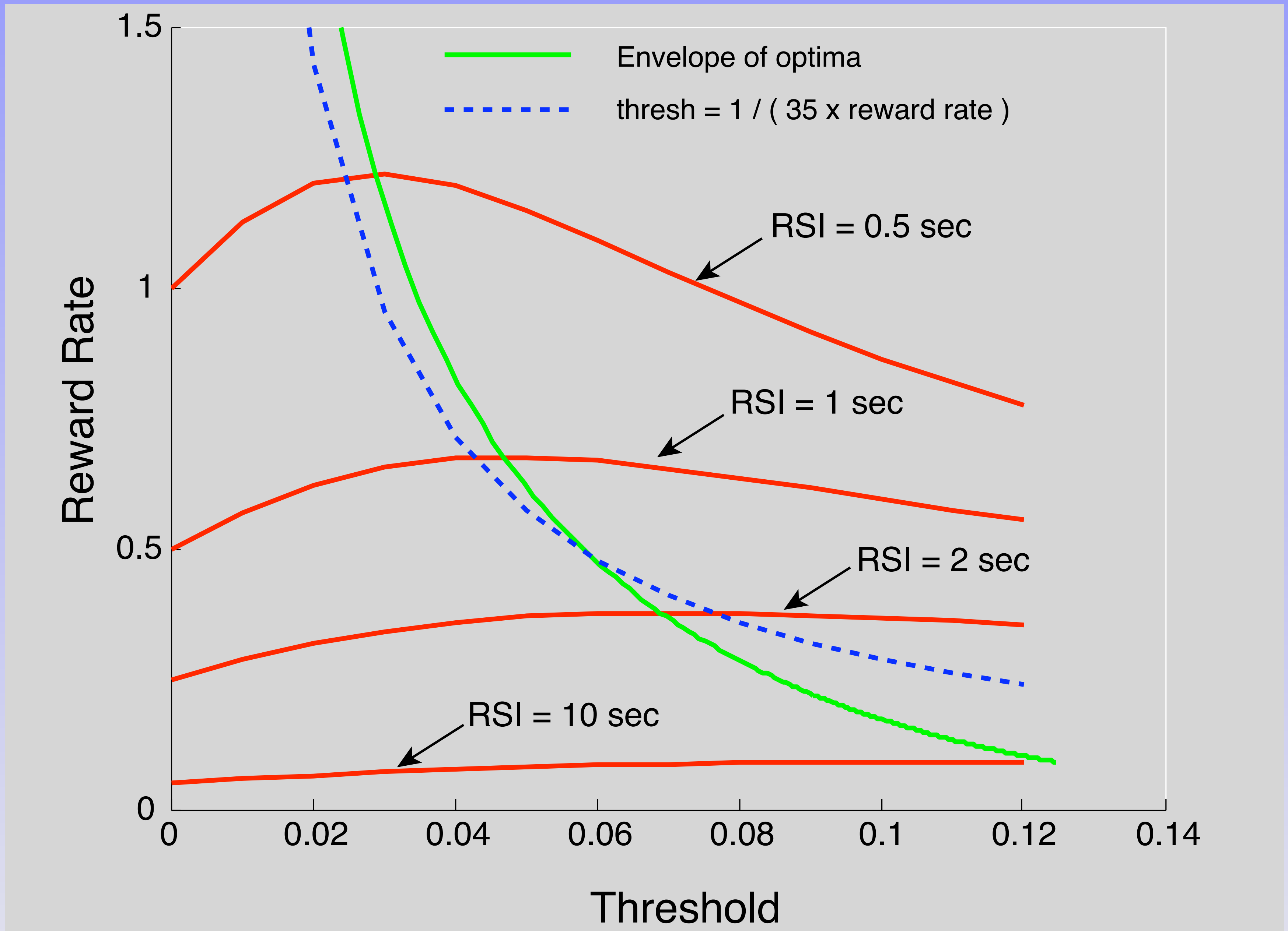
~~Adapting these weights modifies drift; (with FIXED THRESHOLDS, results in SOFTMAX choice rule; cf. Montague & Berns, 2002)~~

FOR OUR TASK

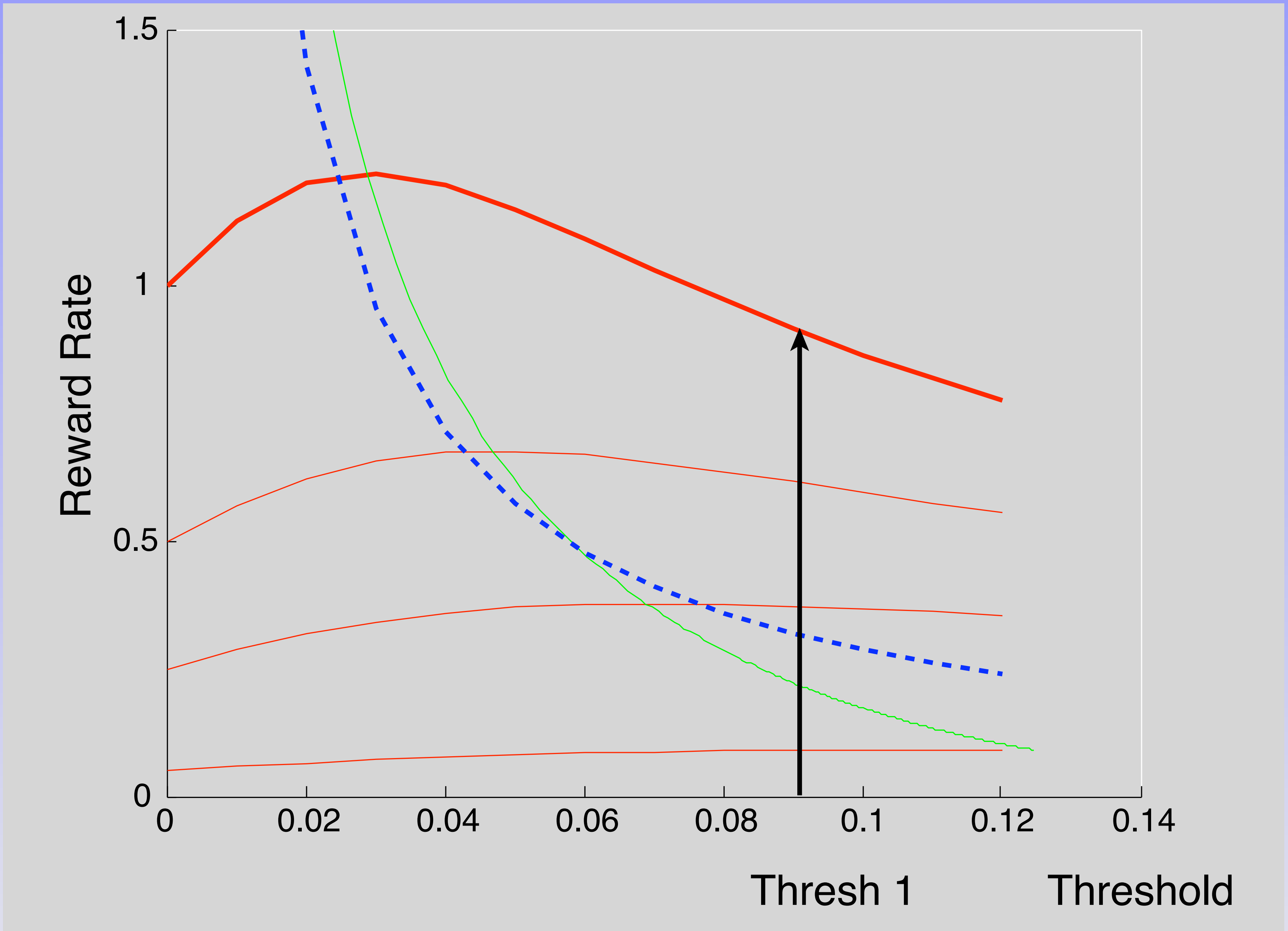


Simen, P. & Cohen, J. D. (in press), *Brain Research,* “Explicit melioration by a neural diffusion model”

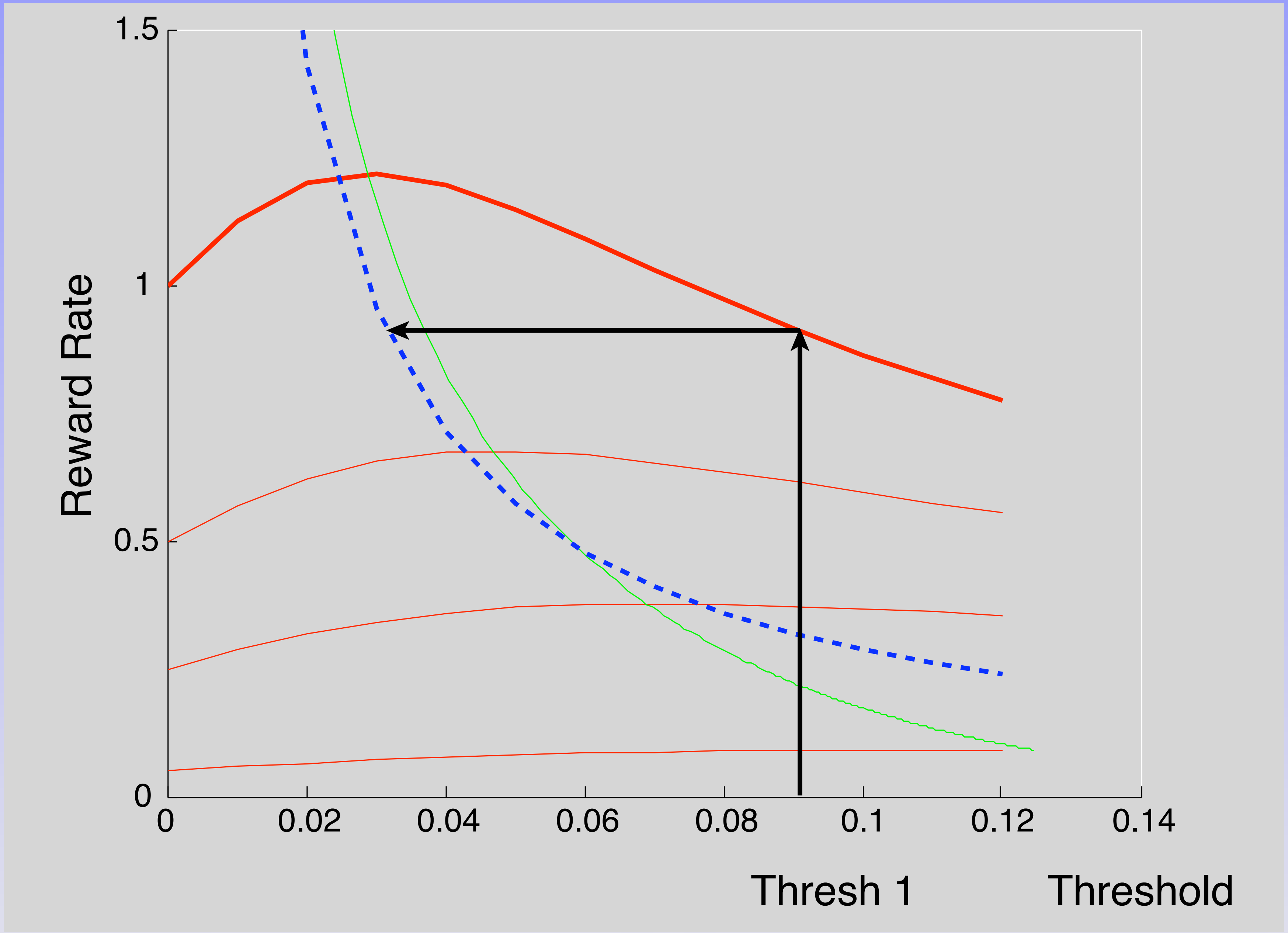
Hyperbolic function (blue) approximates envelope of optima (green) for 50:50 odds task (**Expt 1**)



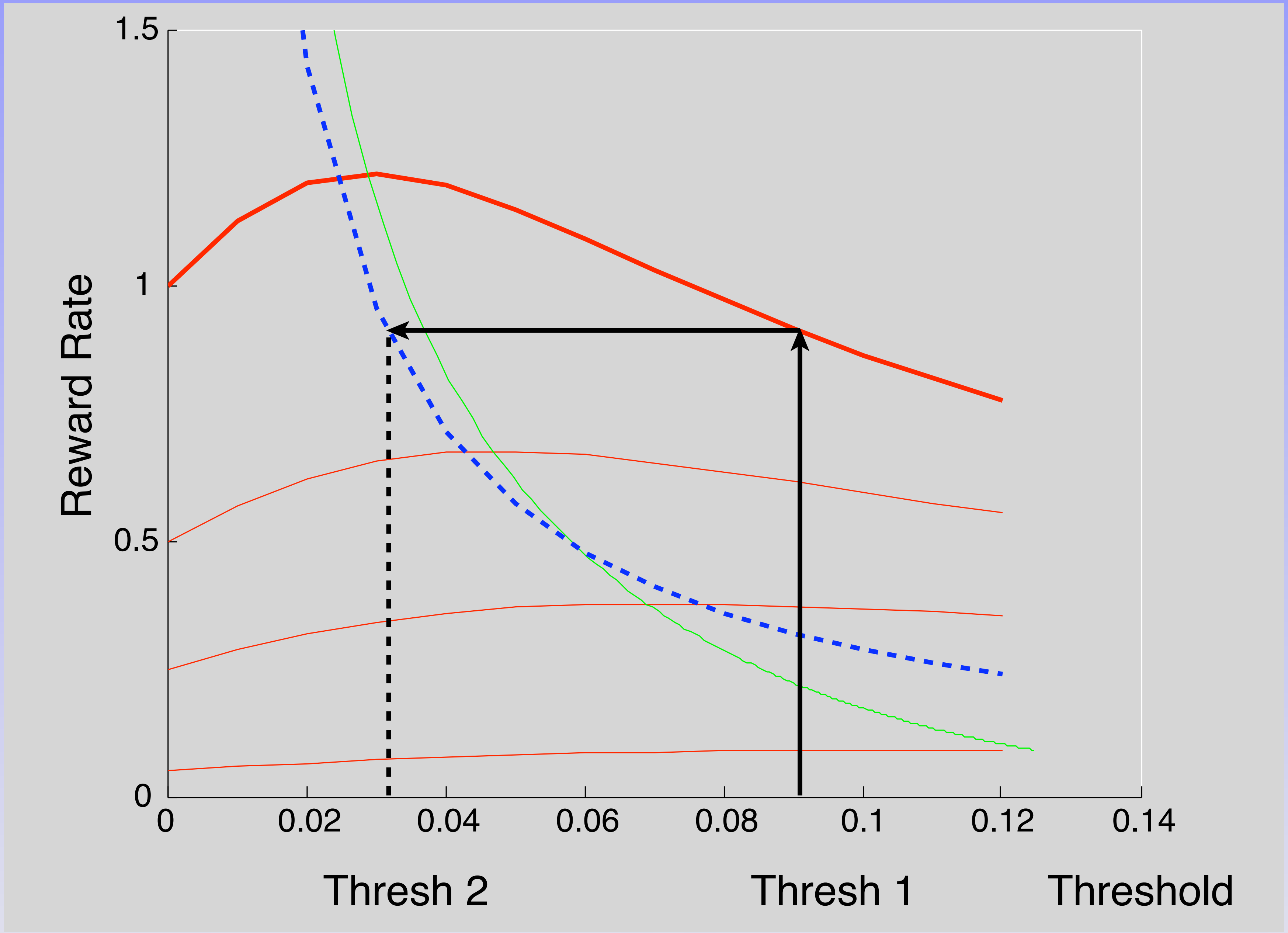
Trial-by-trial threshold adaptation



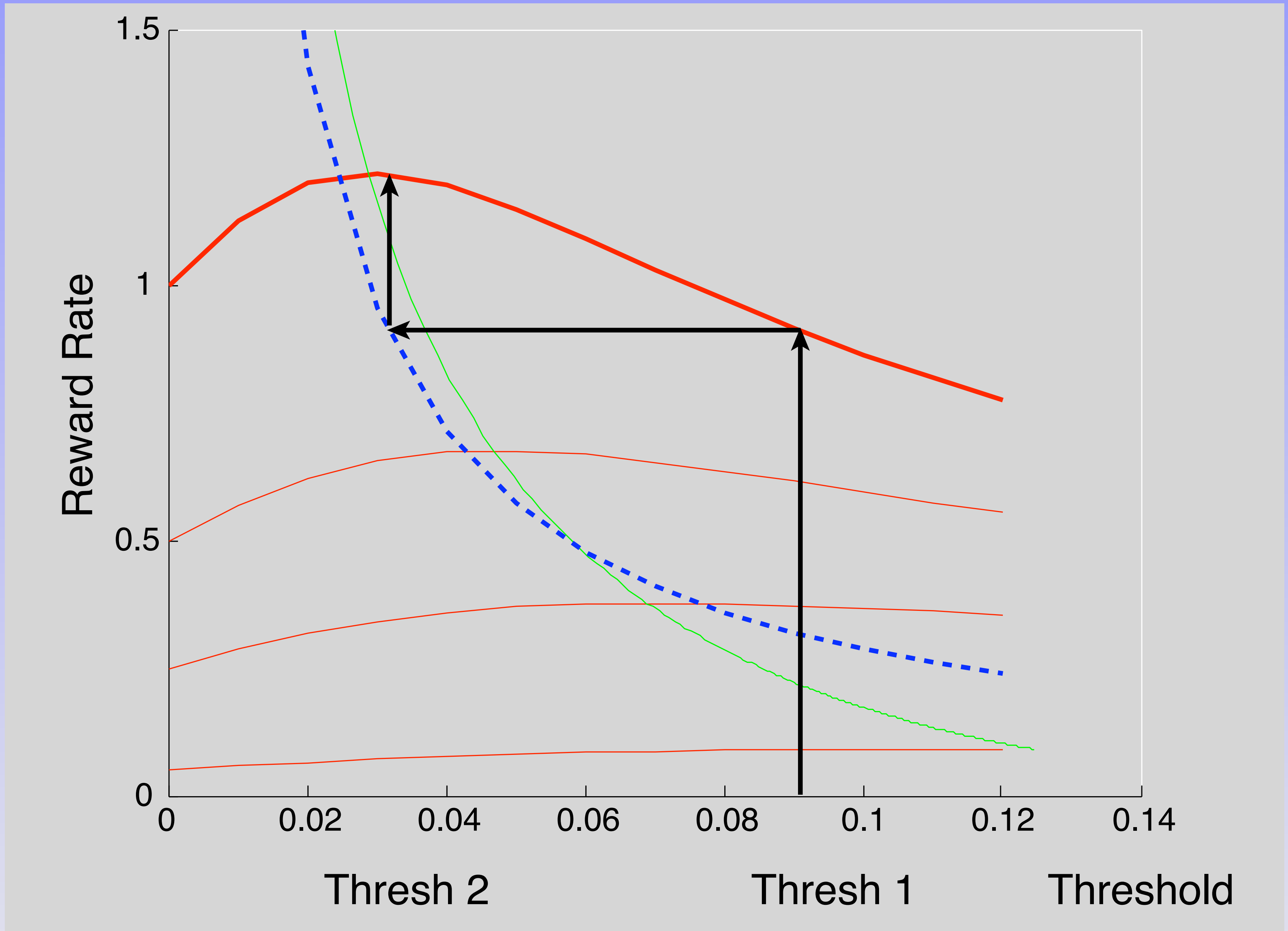
Trial-by-trial threshold adaptation



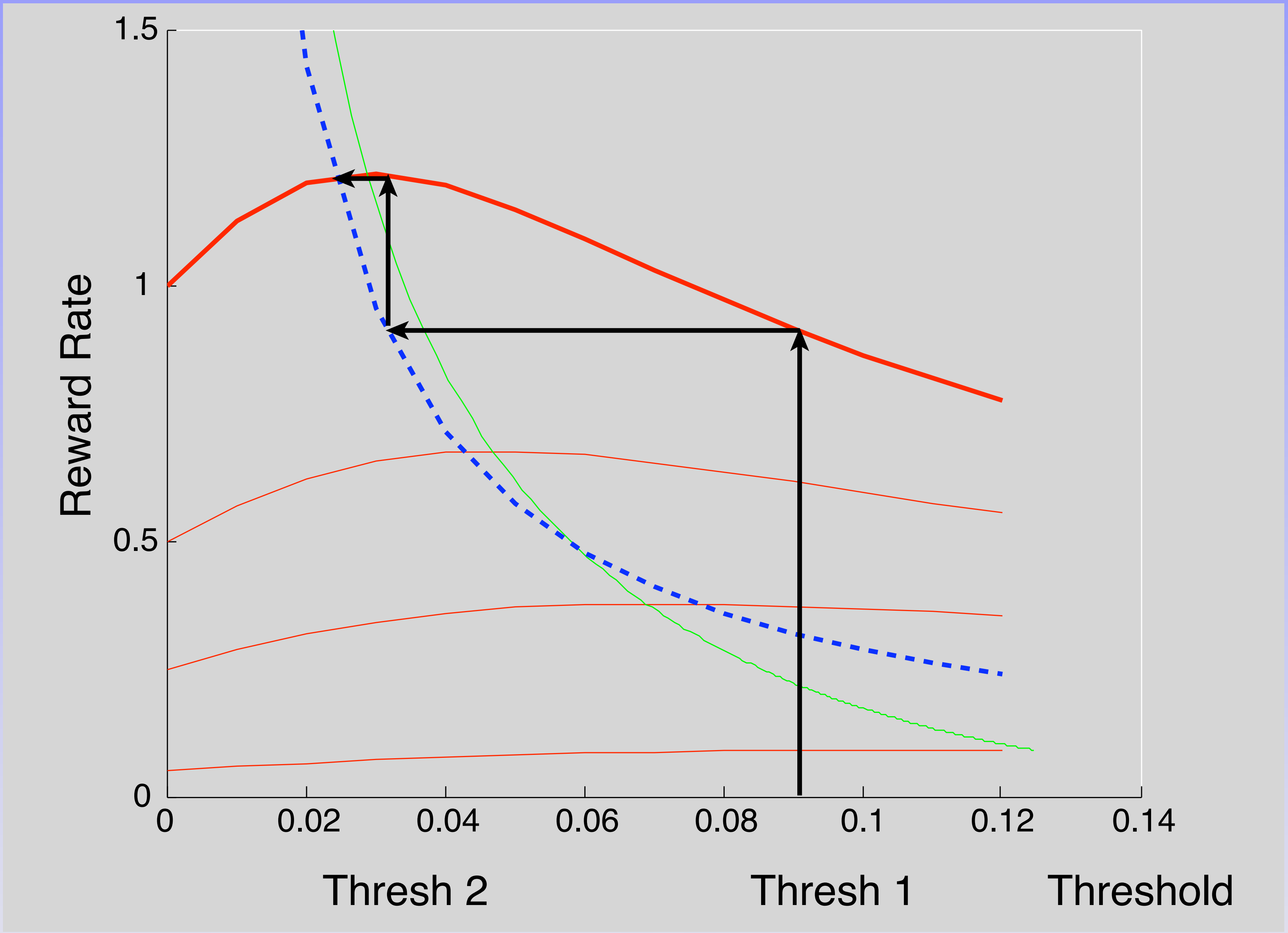
Trial-by-trial threshold adaptation



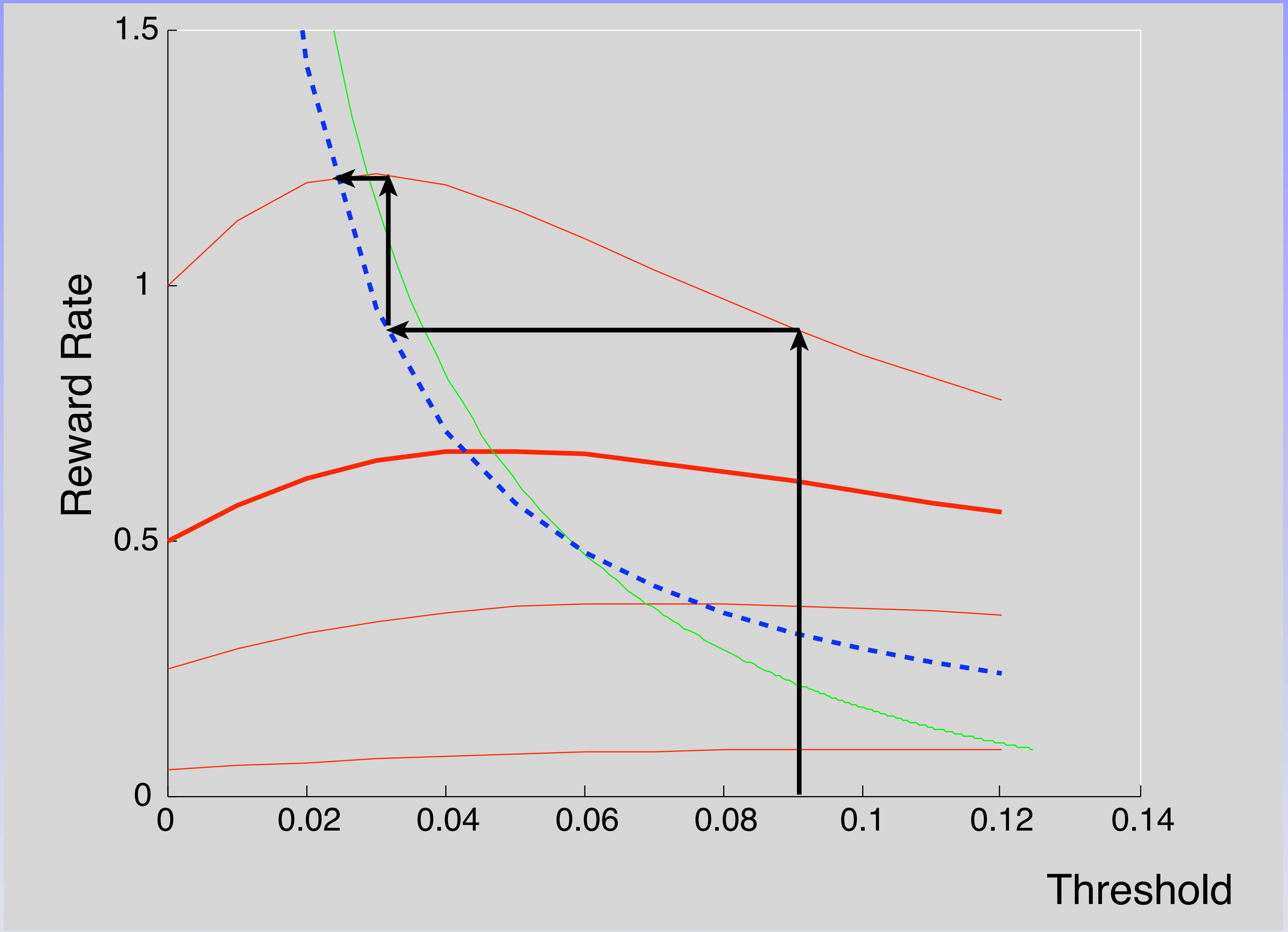
Trial-by-trial threshold adaptation



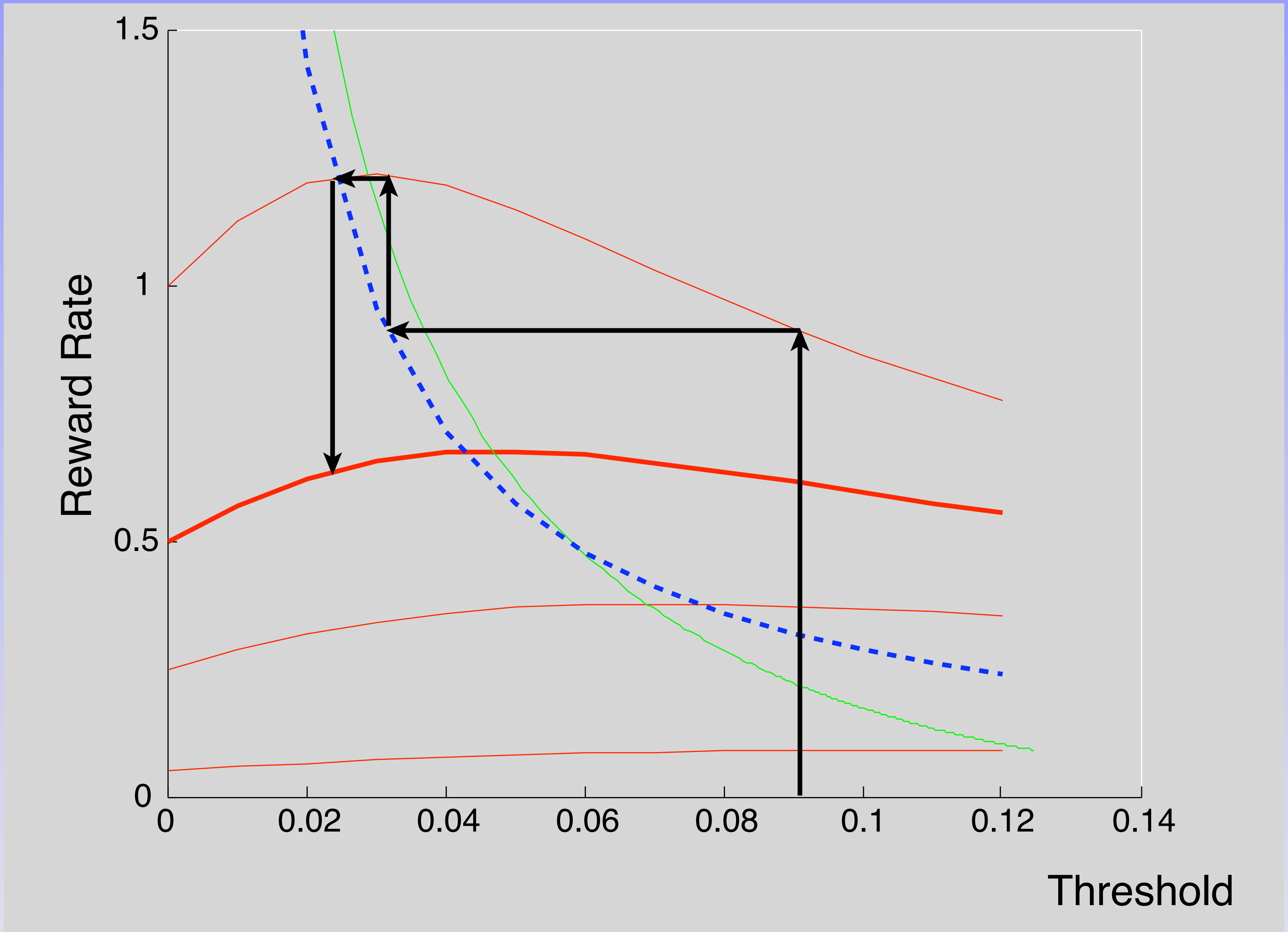
Trial-by-trial threshold adaptation



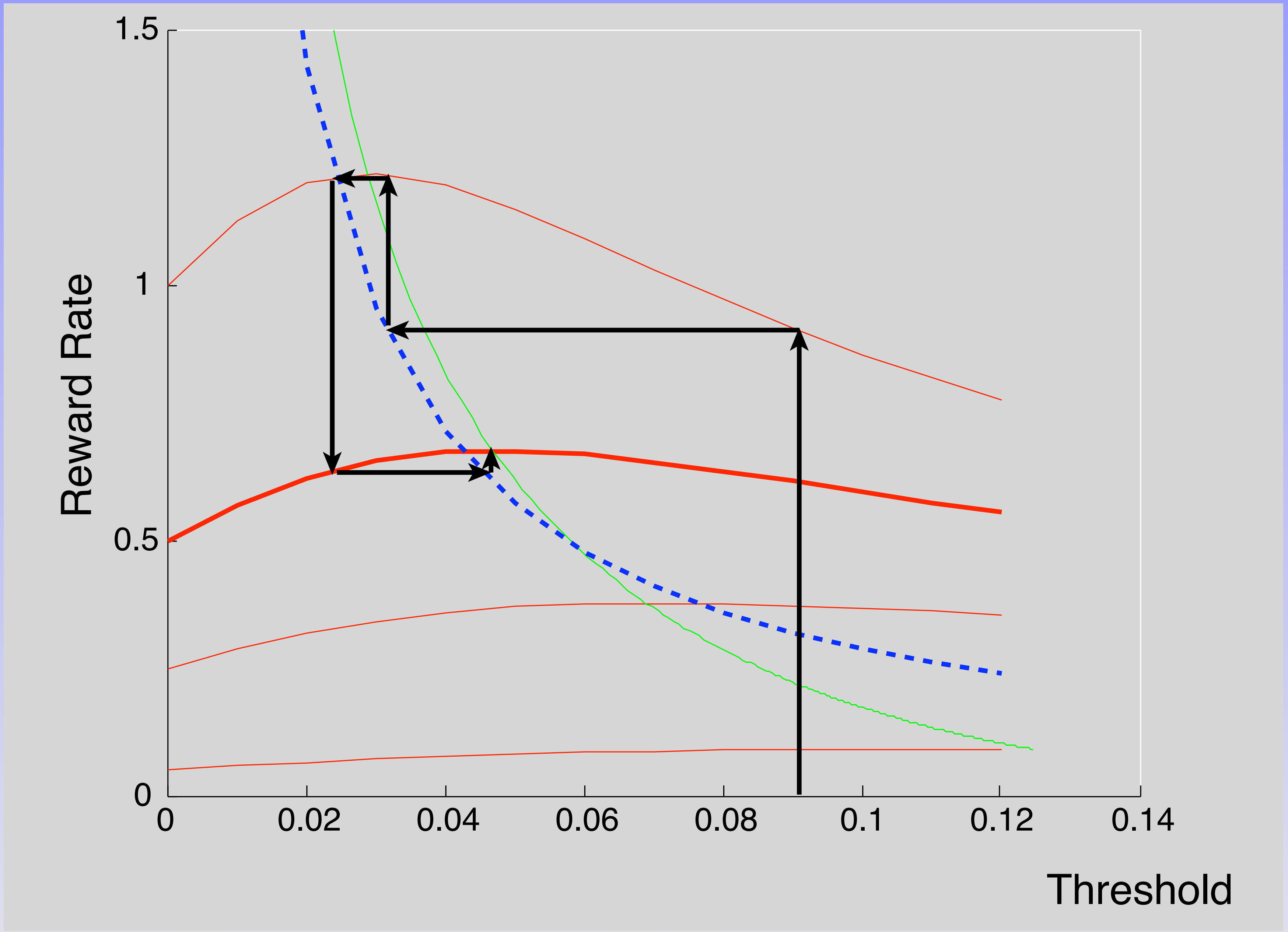
Trial-by-trial threshold adaptation



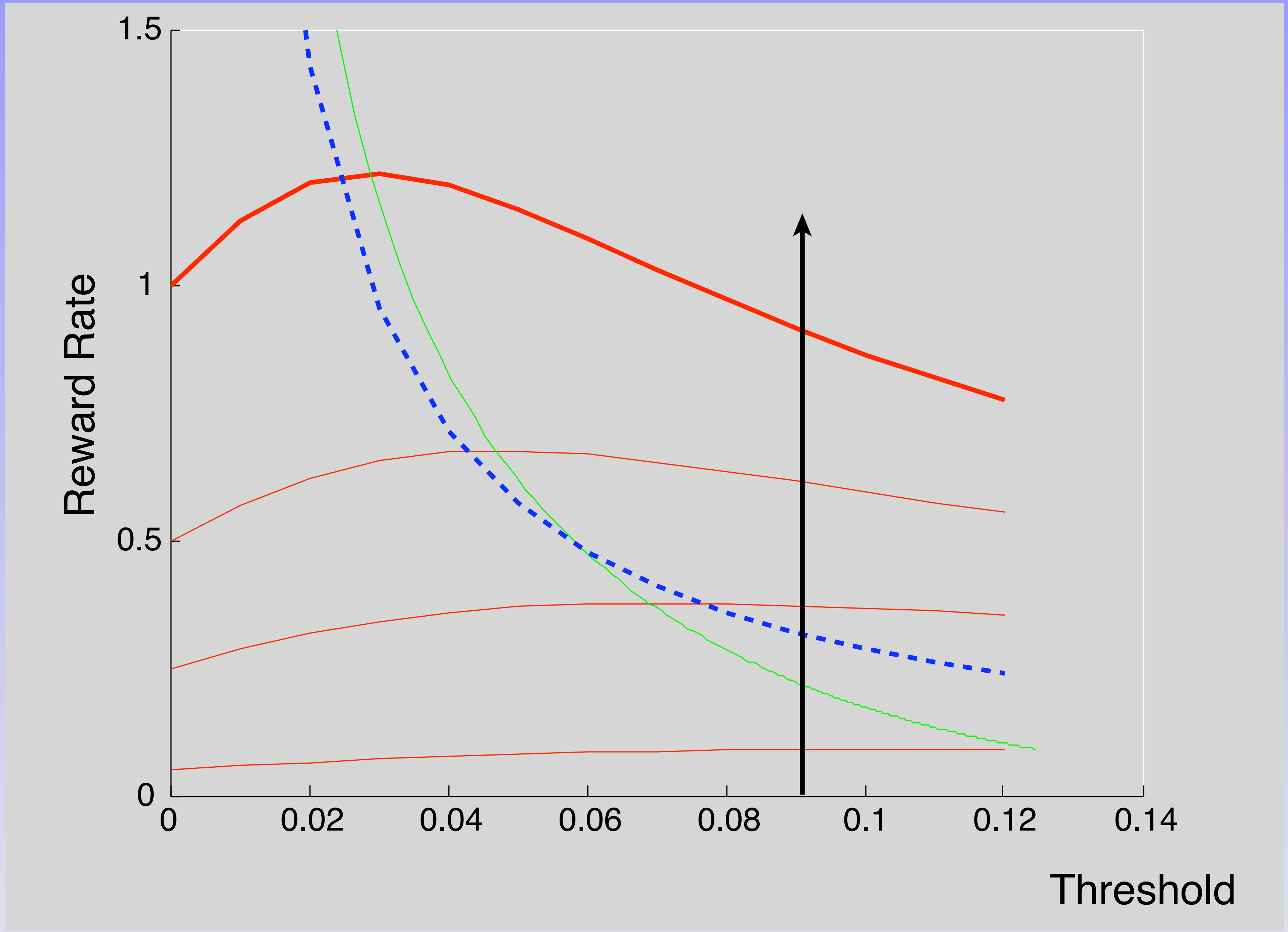
Trial-by-trial threshold adaptation



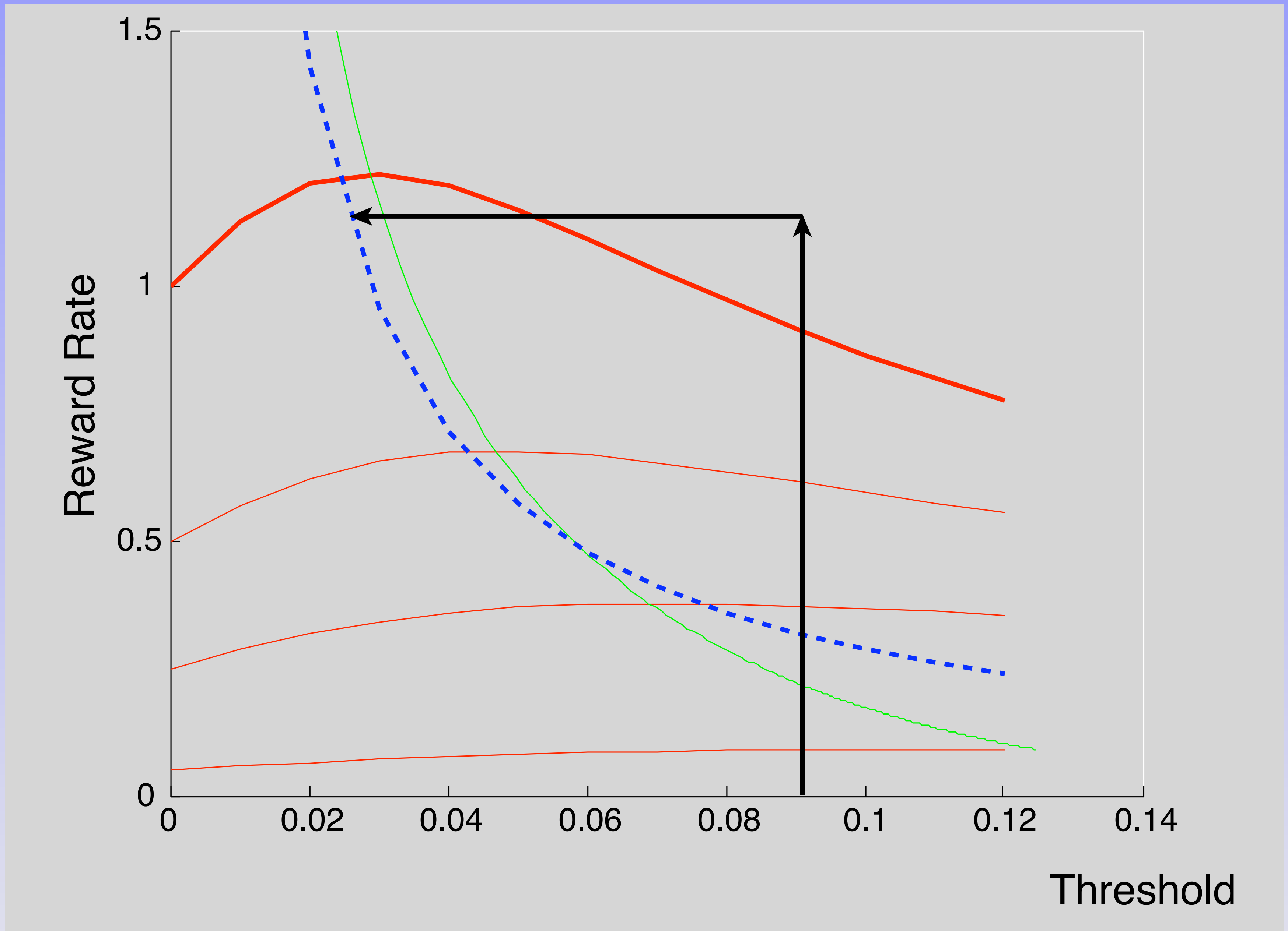
Trial-by-trial threshold adaptation



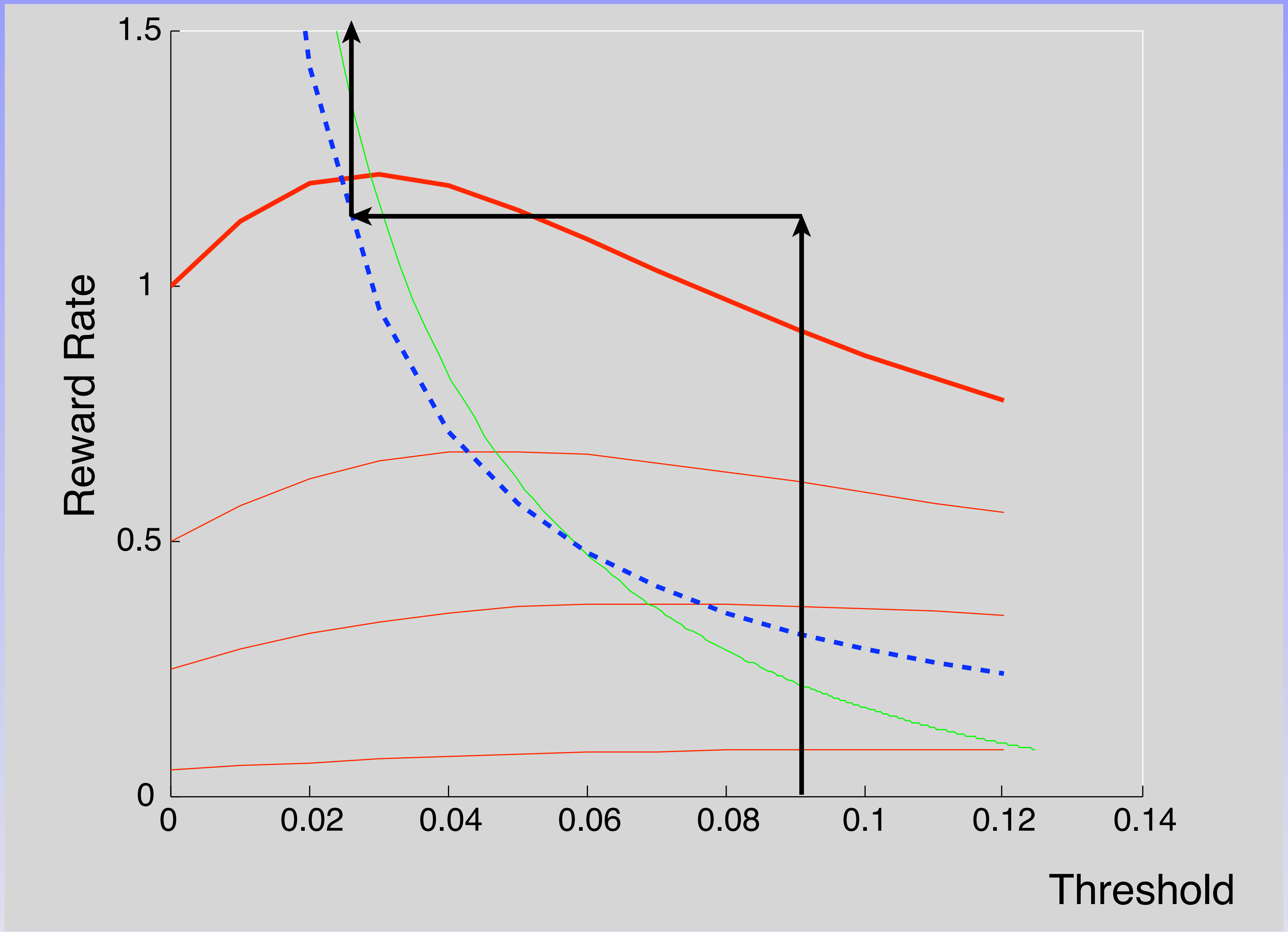
With noise



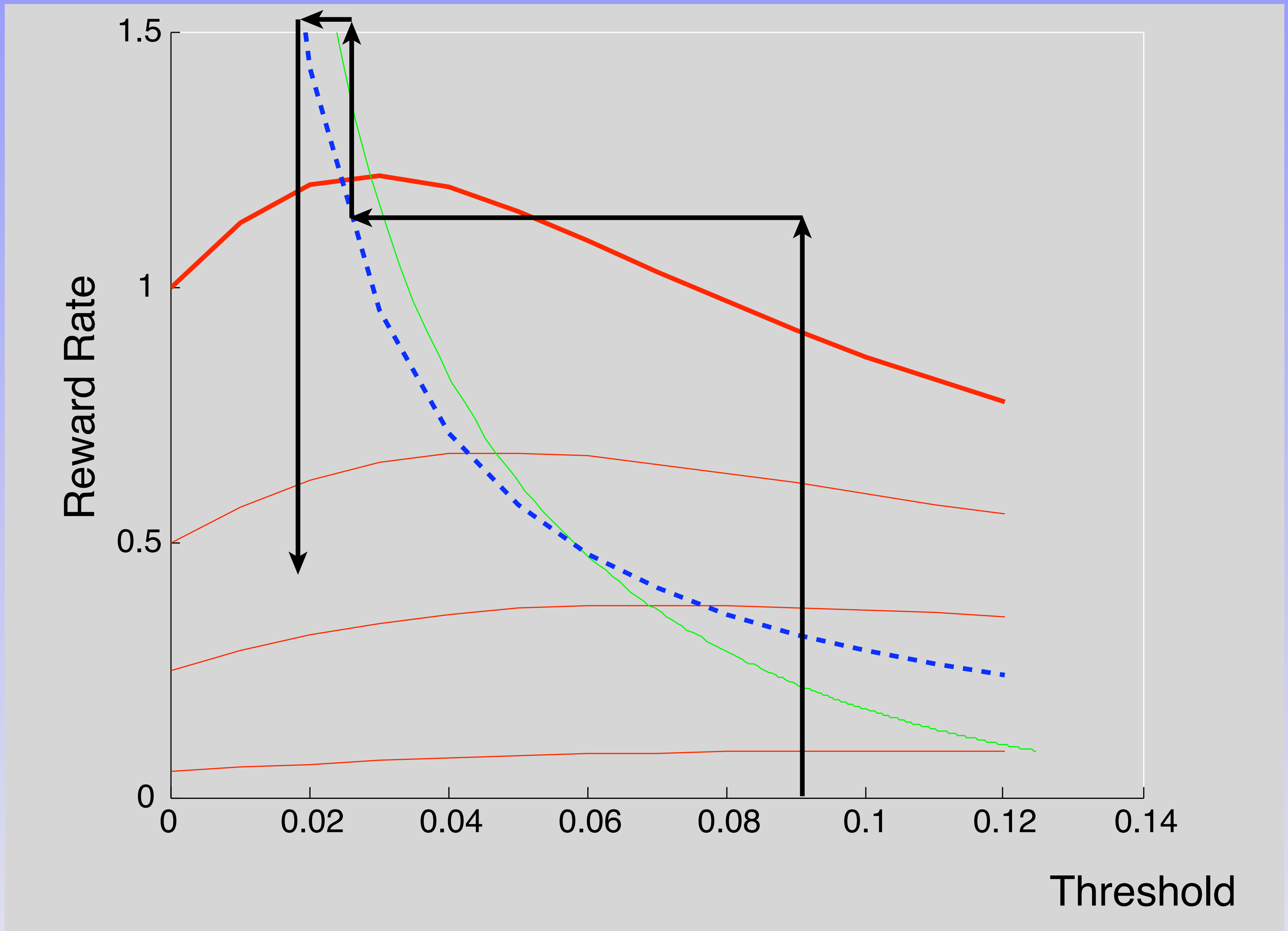
With noise



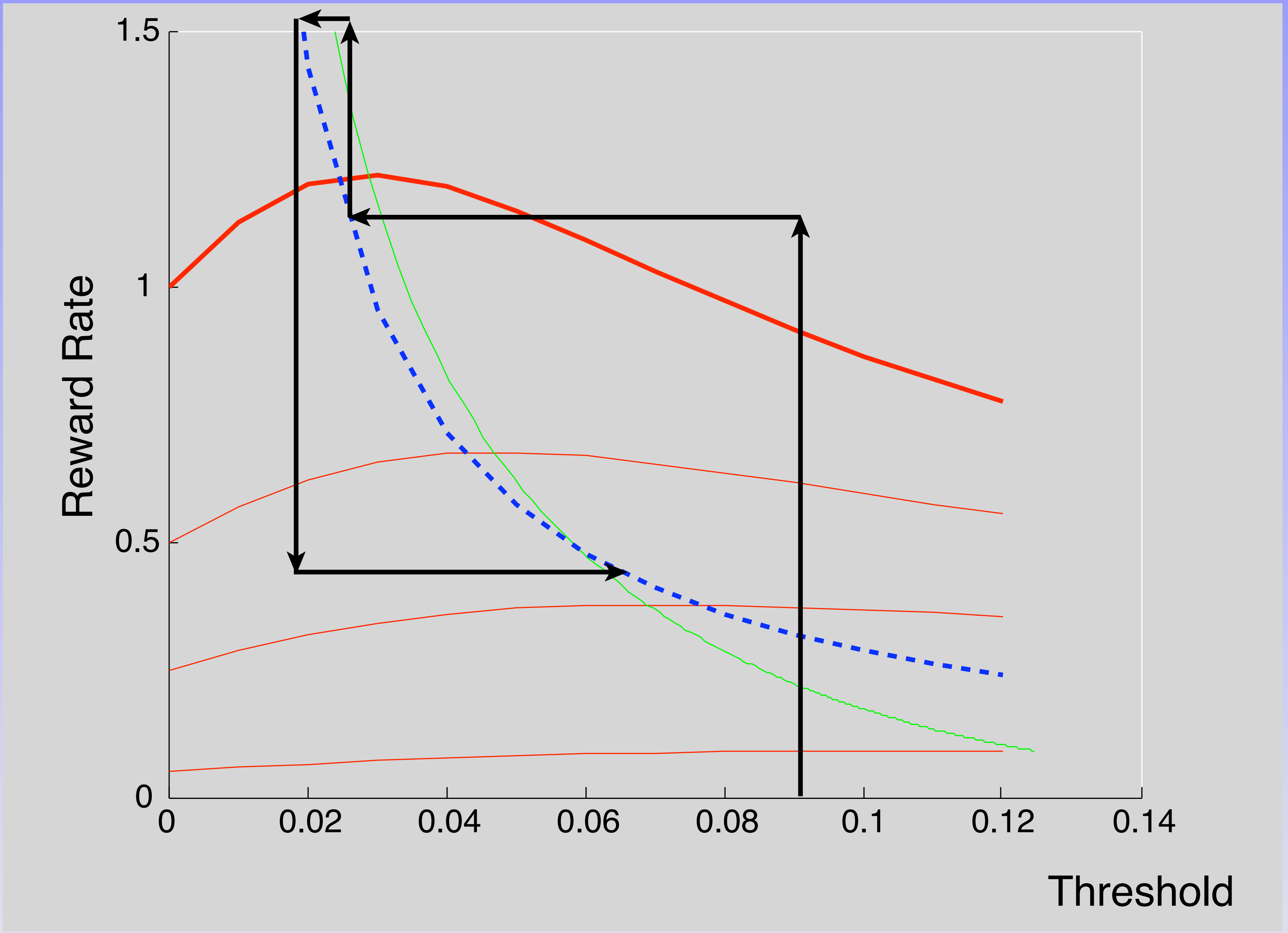
With noise



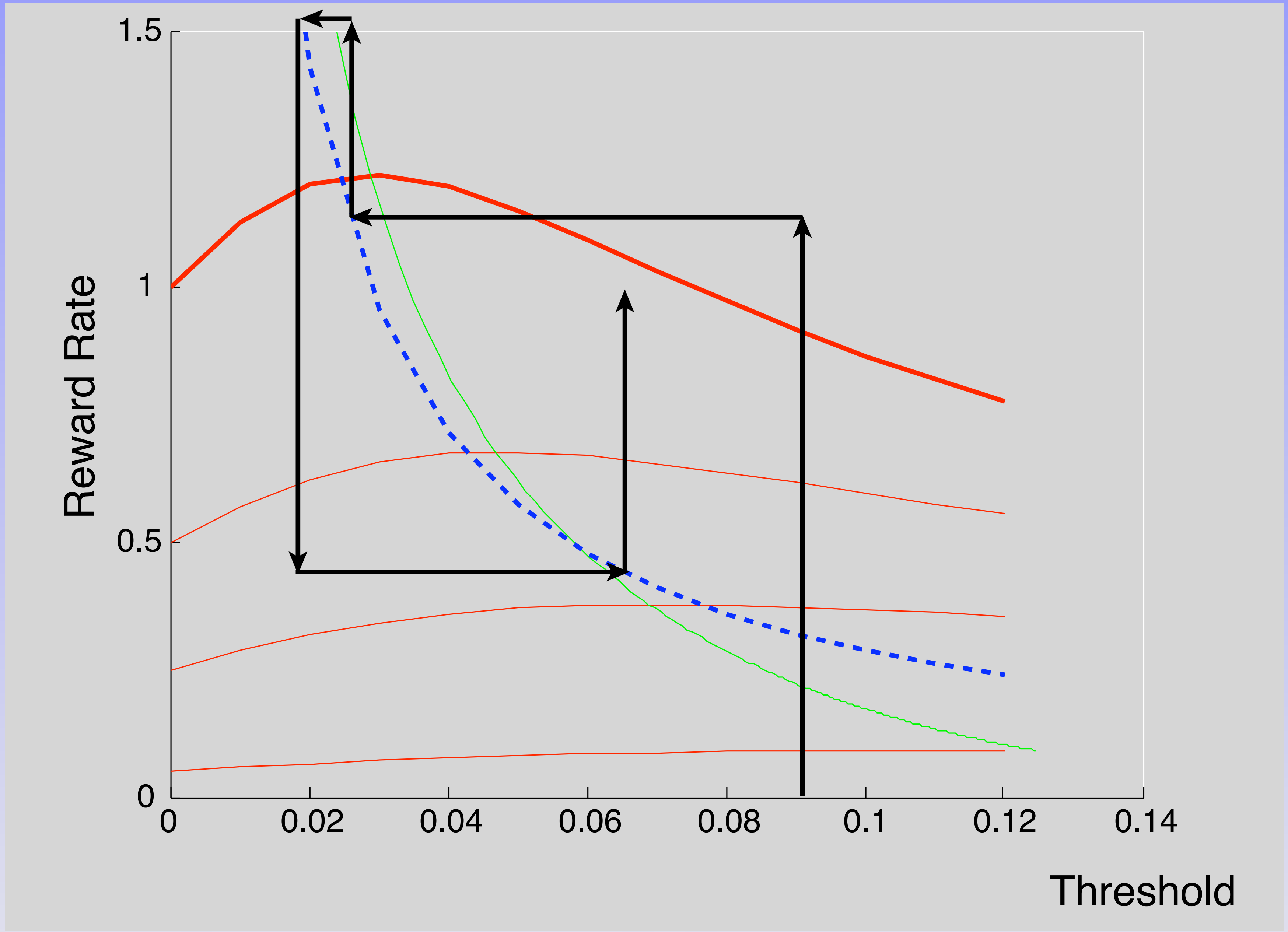
With noise



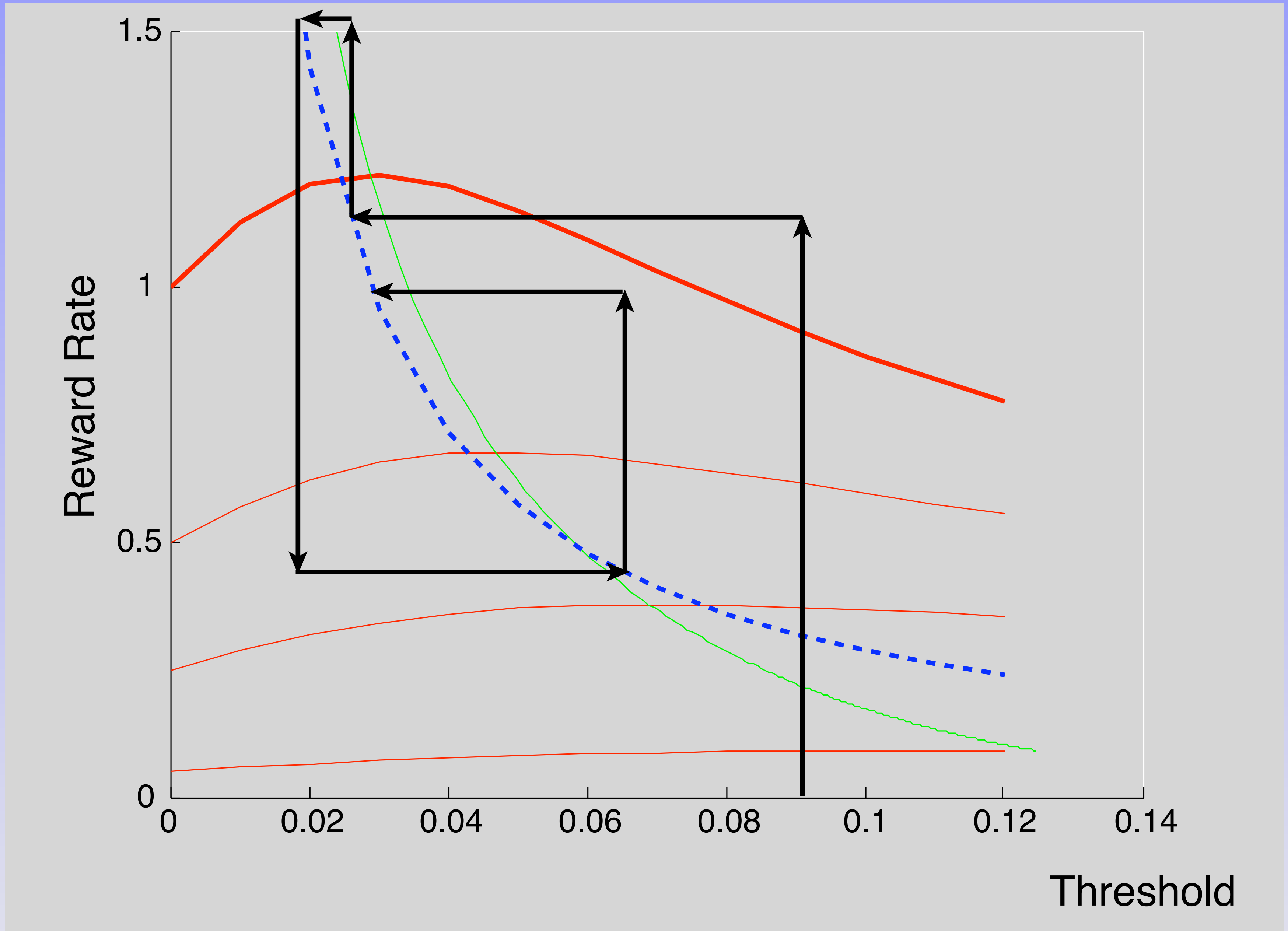
With noise



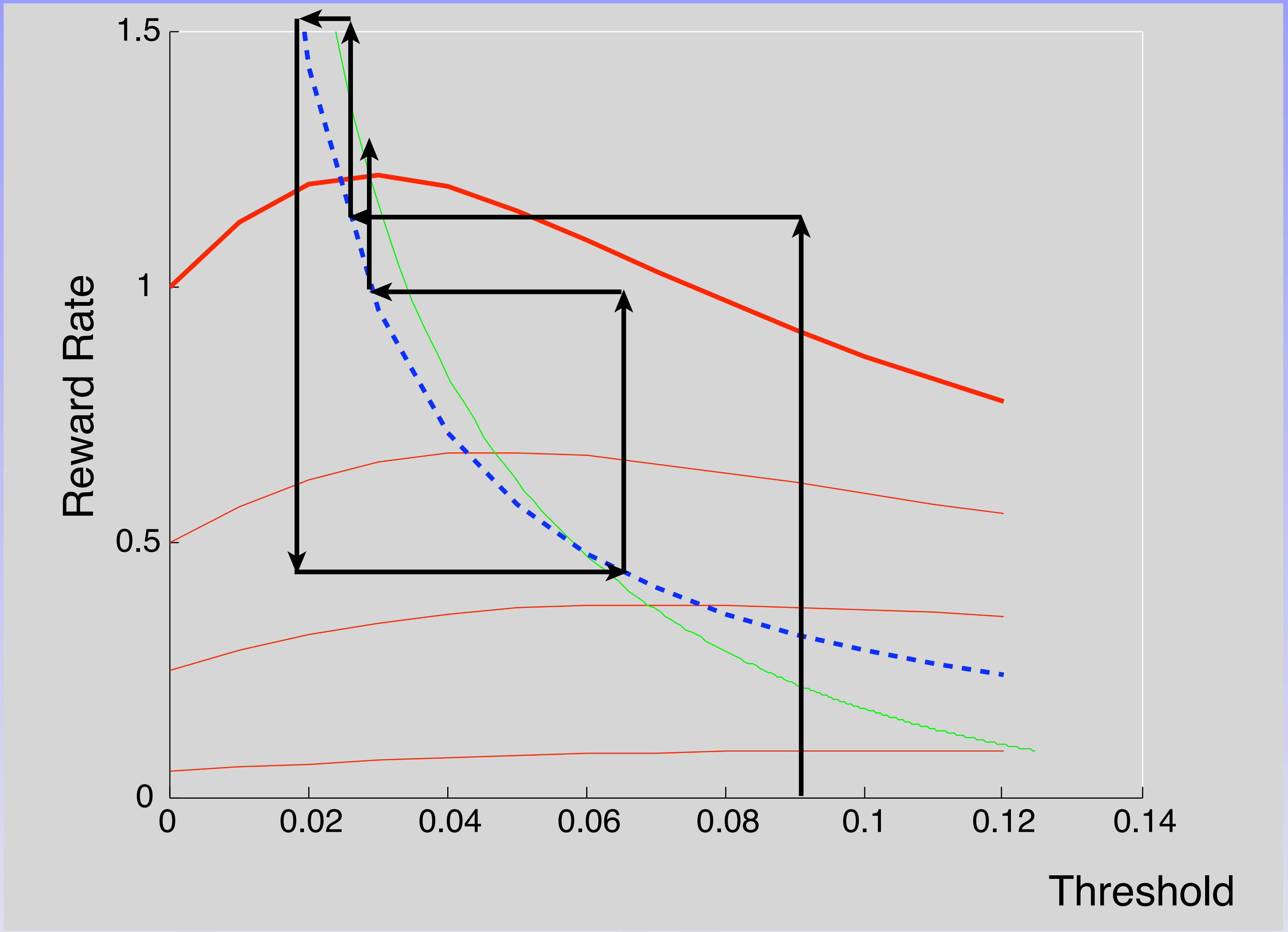
With noise



With noise



With noise



Moment-by-moment threshold adaptation and model performance in real time

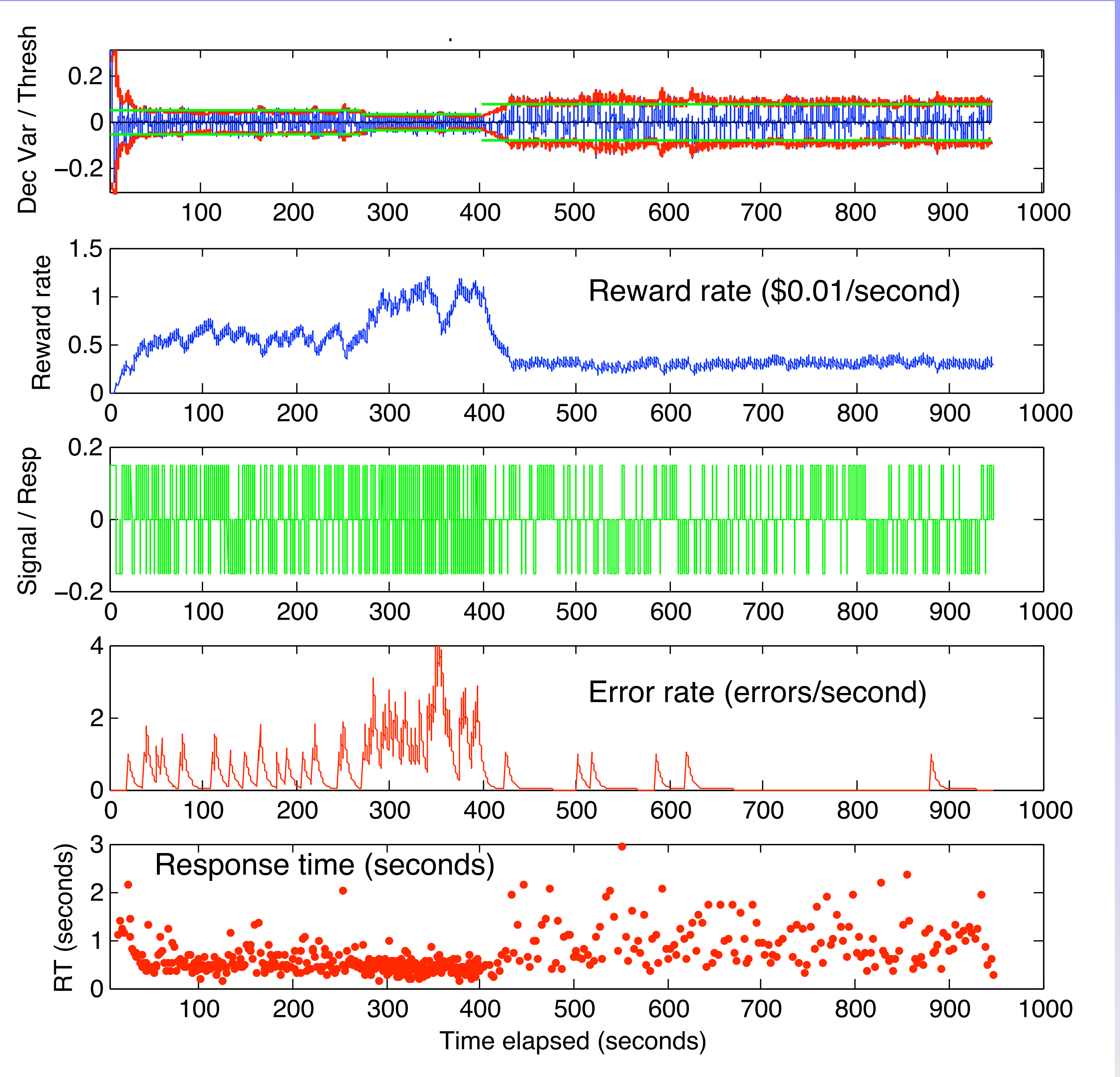
Optimal threshold – green
Decision process – blue
Model threshold – red

Reward rate estimate – blue
Exponentially weighted average *w(t)*, computed in continuous time by leaky integration of reward impulses *r* (sum of Dirac delta functions):

$$\dot{w}_i = -w_i(t) + r_i(t)$$

Signal – green (0.15: leftward moving dots; -0.15: rightward moving dot; 0: no signal present)

Error rate



Moment-by-moment threshold adaptation and model performance in real time

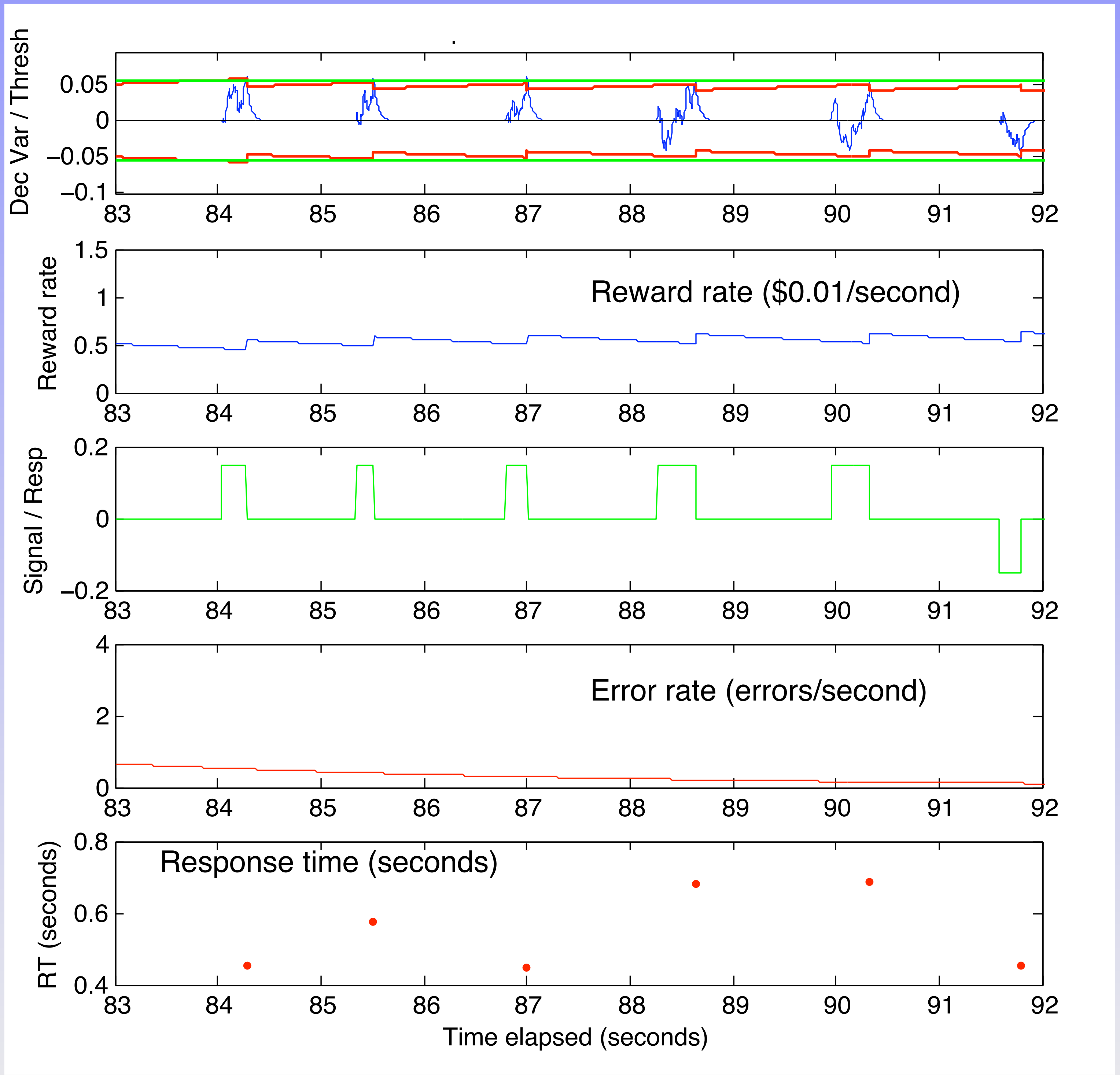
Optimal threshold – green
Decision process – blue
Model threshold – red

Reward rate estimate – blue
Exponentially weighted average $w(t)$, computed in continuous time by leaky integration of reward impulses r (sum of Dirac delta functions):

$$\dot{w}_i = -w_i(t) + r_i(t)$$

Signal – green (0.15: leftward moving dots; -0.15: rightward moving dot; 0: no signal present)

Error rate



Moment-by-moment threshold adaptation and model performance in real time

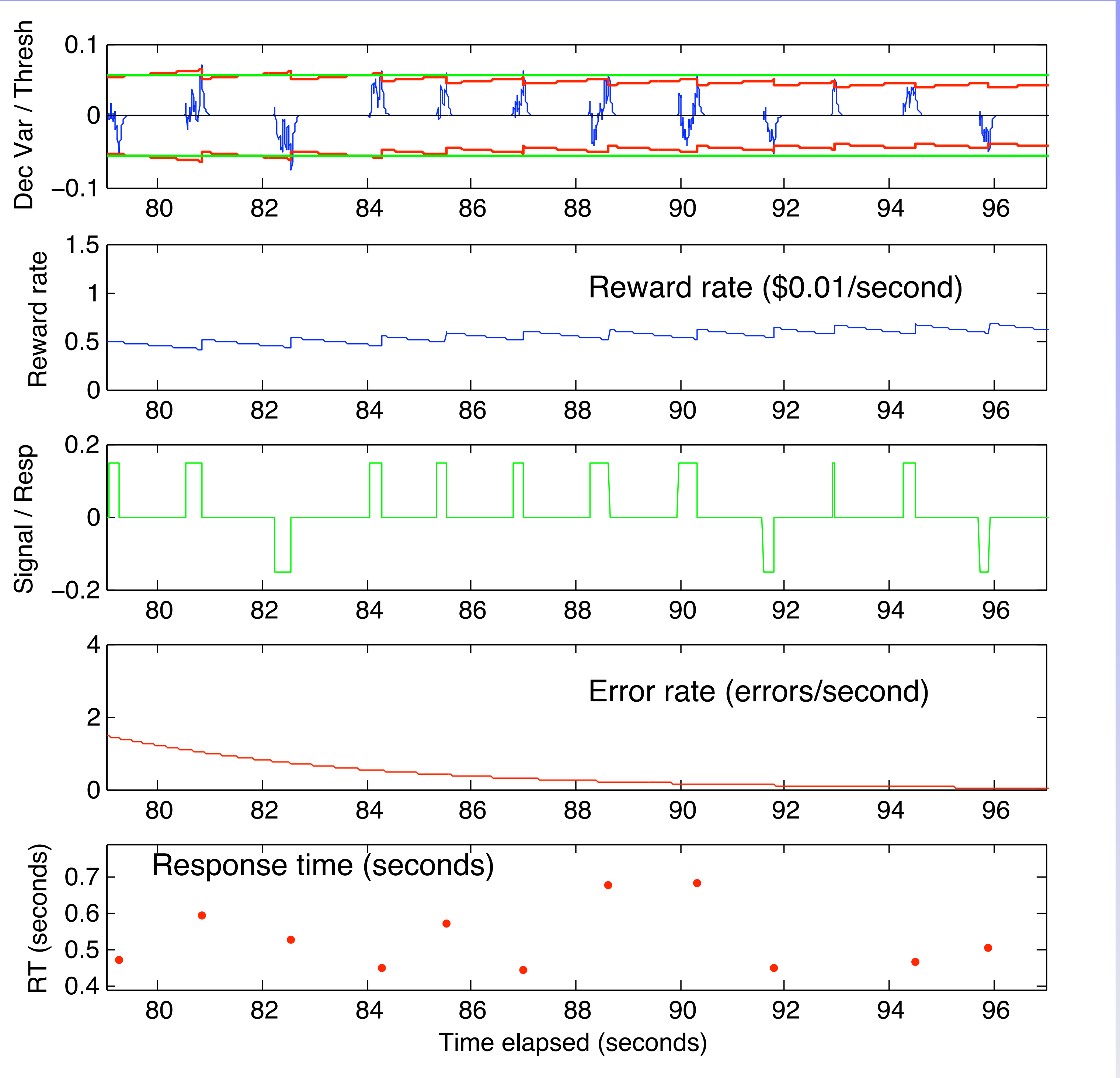
Optimal threshold – green
Decision process – blue
Model threshold – red

Reward rate estimate – blue
Exponentially weighted average *w(t)*, computed in continuous time by leaky integration of reward impulses *r* (sum of Dirac delta functions):

$$\dot{w}_i = -w_i(t) + r_i(t)$$

Signal – green (0.15: leftward moving dots; -0.15: rightward moving dot; 0: no signal present)

Error rate



Moment-by-moment threshold adaptation and model performance in real time

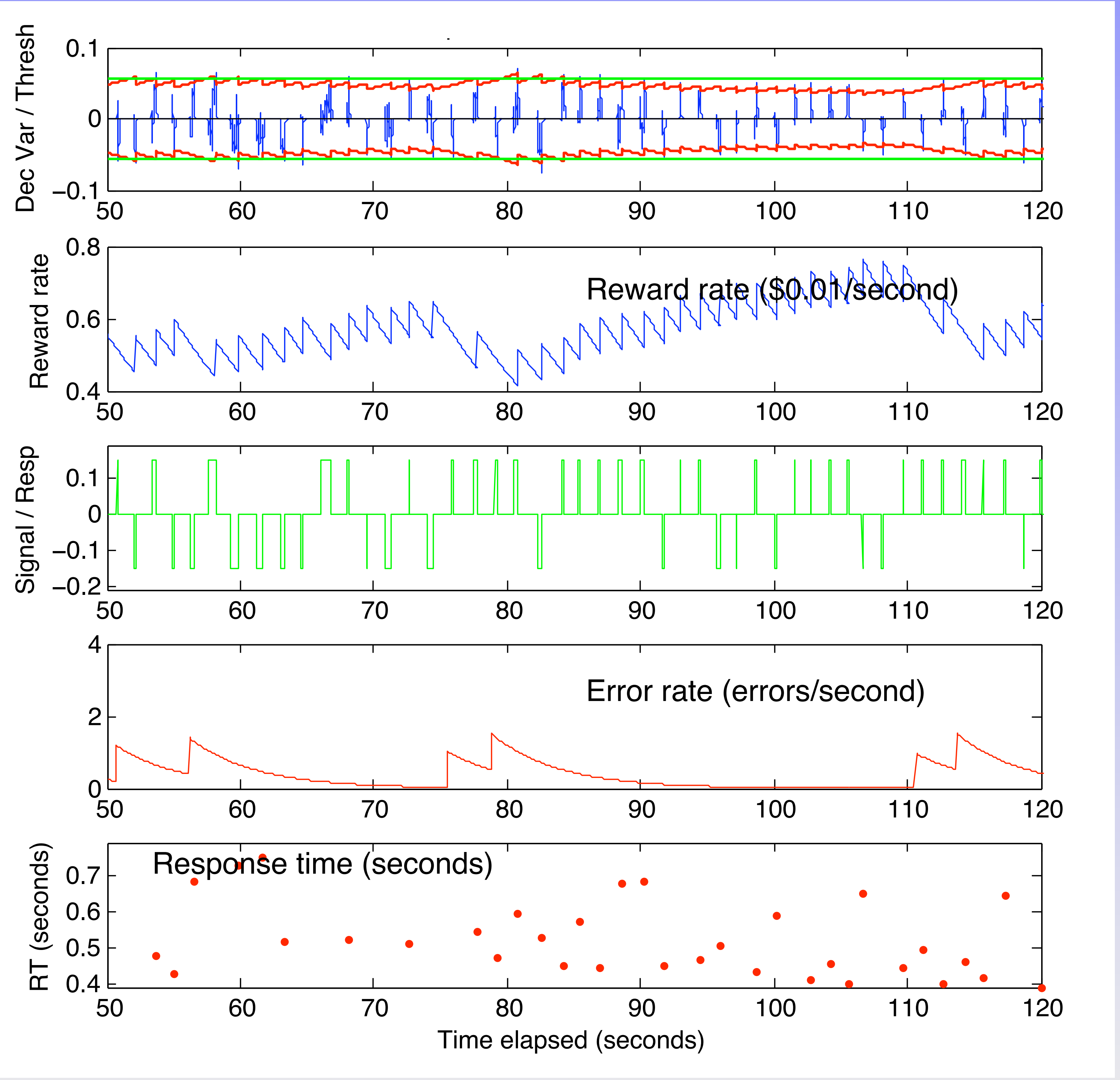
Optimal threshold – green
Decision process – blue
Model threshold – red

Reward rate estimate – blue
Exponentially weighted average *w(t)*, computed in continuous time by leaky integration of reward impulses *r* (sum of Dirac delta functions):

$$\dot{w}_i = -w_i(t) + r_i(t)$$

Signal – green (0.15: leftward moving dots; -0.15: rightward moving dot; 0: no signal present)

Error rate



Moment-by-moment threshold adaptation and model performance in real time

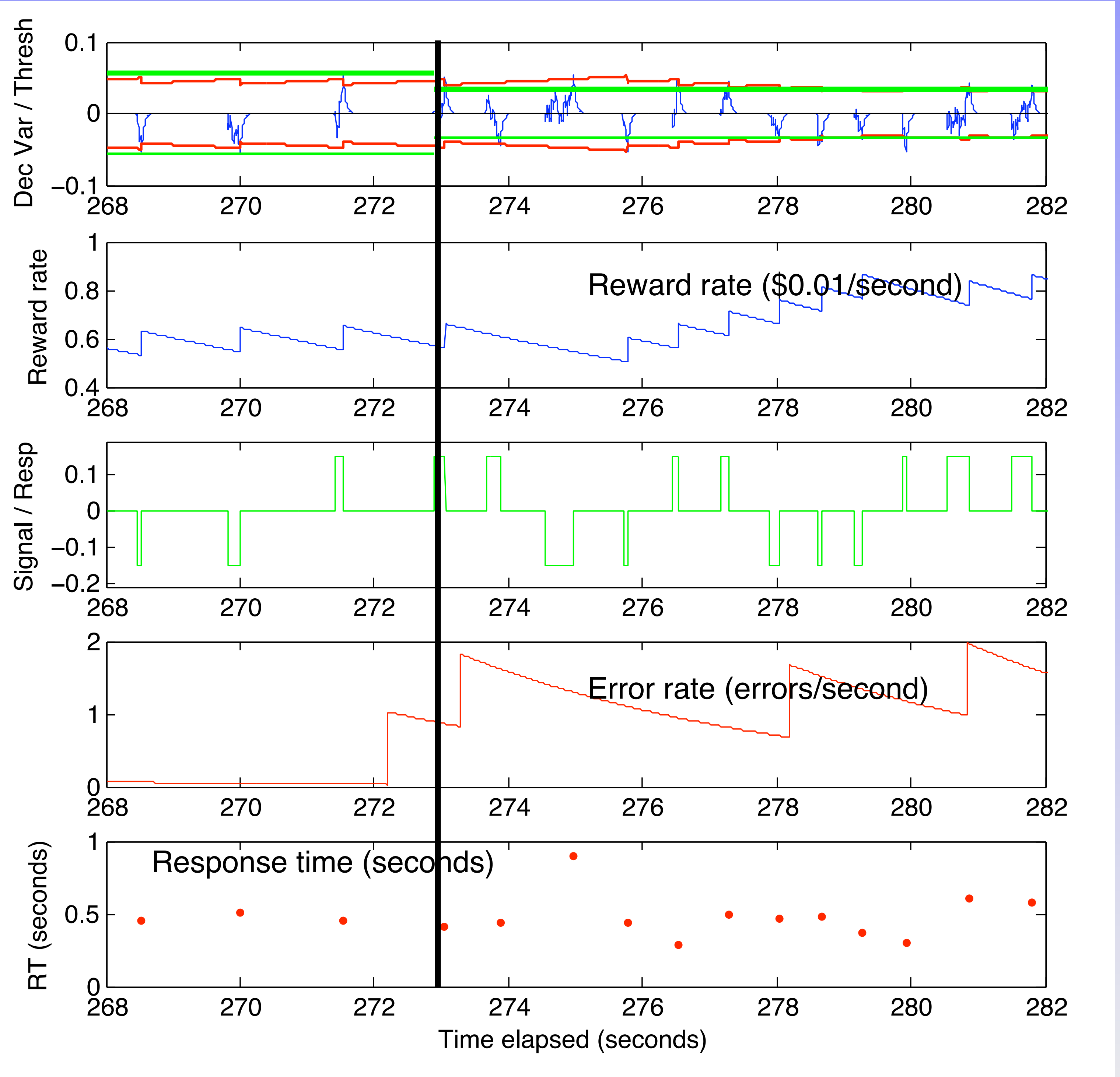
Optimal threshold – green
Decision process – blue
Model threshold – red

Reward rate estimate – blue
Exponentially weighted average *w(t)*, computed in continuous time by leaky integration of reward impulses *r* (sum of Dirac delta functions):

$$\dot{w}_i = -w_i(t) + r_i(t)$$

Signal – green (0.15: leftward moving dots; -0.15: rightward moving dot; 0: no signal present)

Error rate



RSI change from 1 sec to 0.5 sec

Moment-by-moment threshold adaptation and model performance in real time

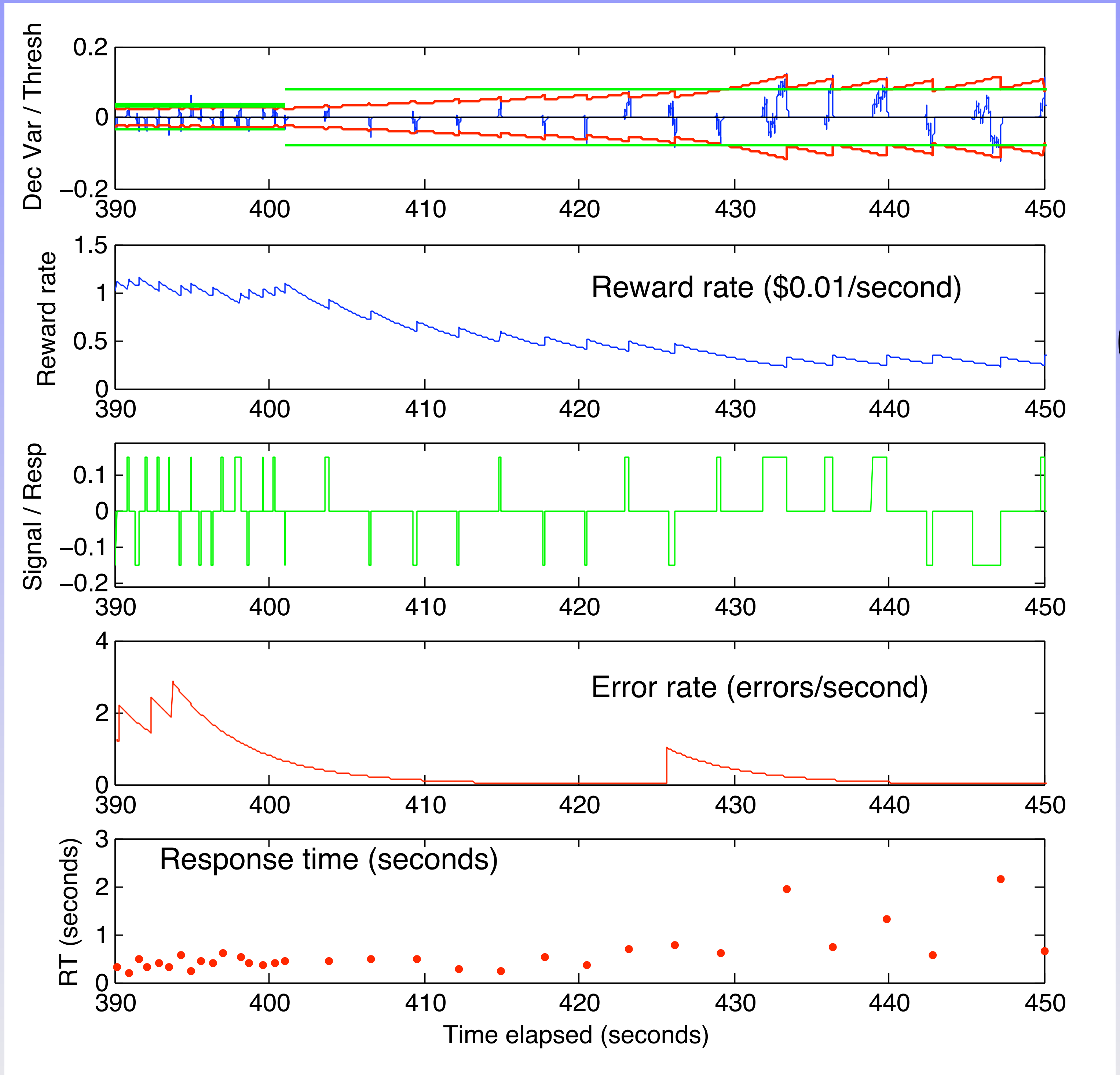
Optimal threshold – green
Decision process – blue
Model threshold – red

Reward rate estimate – blue
Exponentially weighted average *w(t)*, computed in continuous time by leaky integration of reward impulses *r* (sum of Dirac delta functions):

$$\dot{w}_i = -w_i(t) + r_i(t)$$

Signal – green (0.15: leftward moving dots; -0.15: rightward moving dot; 0: no signal present)

Error rate



RSI change from 0.5 sec to 2 sec

Moment-by-moment threshold adaptation and model performance in real time

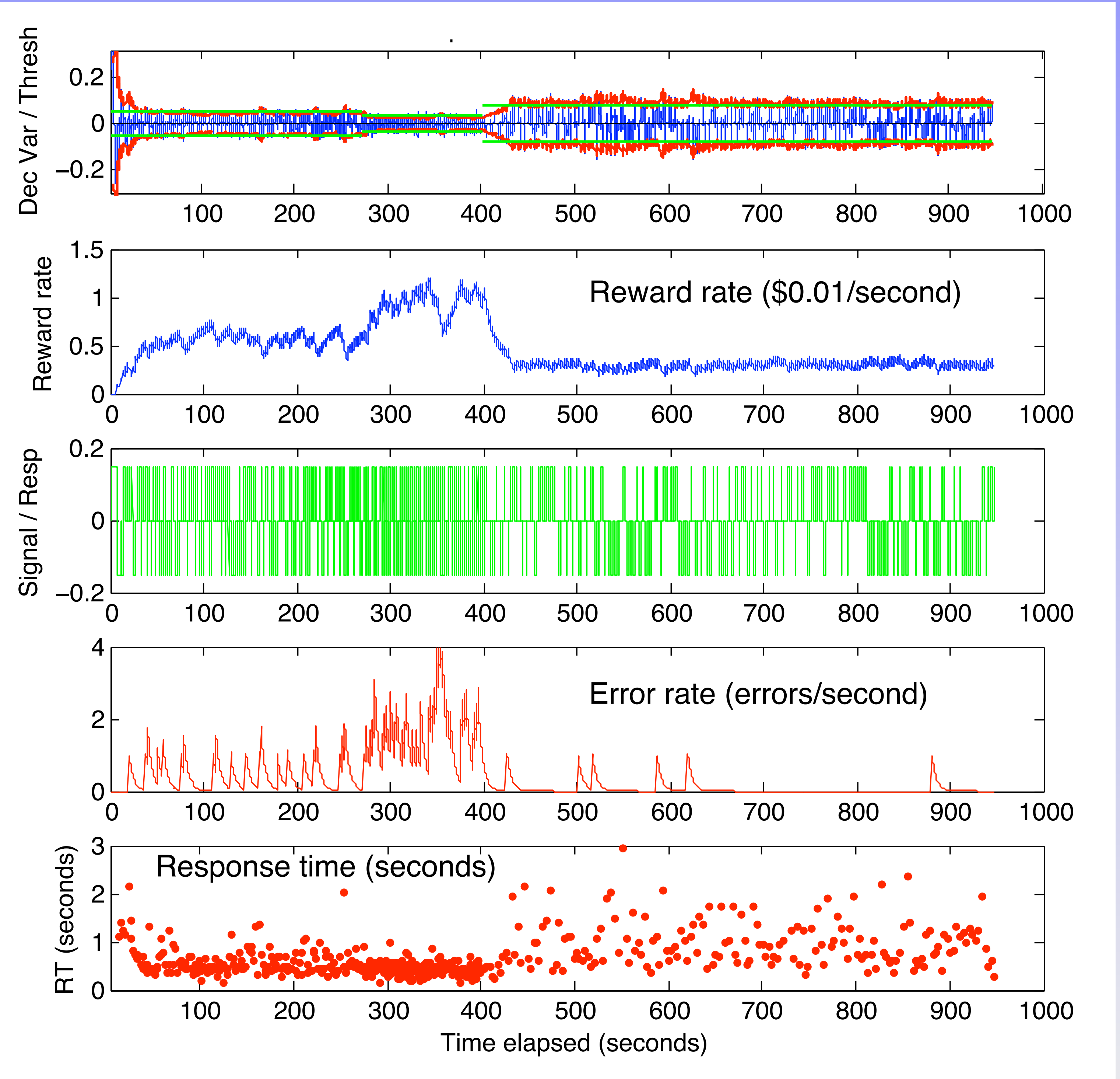
Optimal threshold – green
Decision process – blue
Model threshold – red

Reward rate estimate – blue
Exponentially weighted average *w(t)*, computed in continuous time by leaky integration of reward impulses *r* (sum of Dirac delta functions):

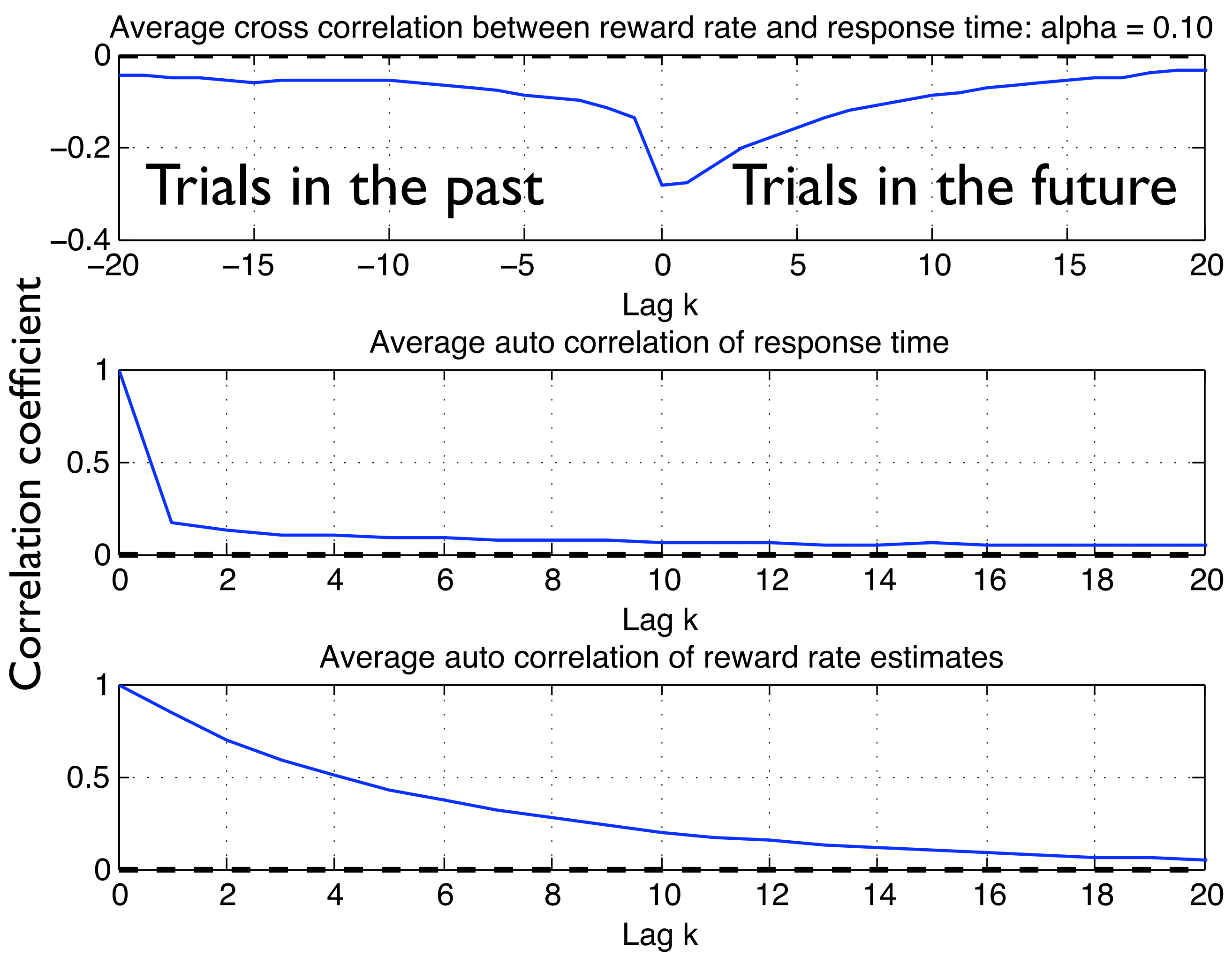
$$\dot{w}_i = -w_i(t) + r_i(t)$$

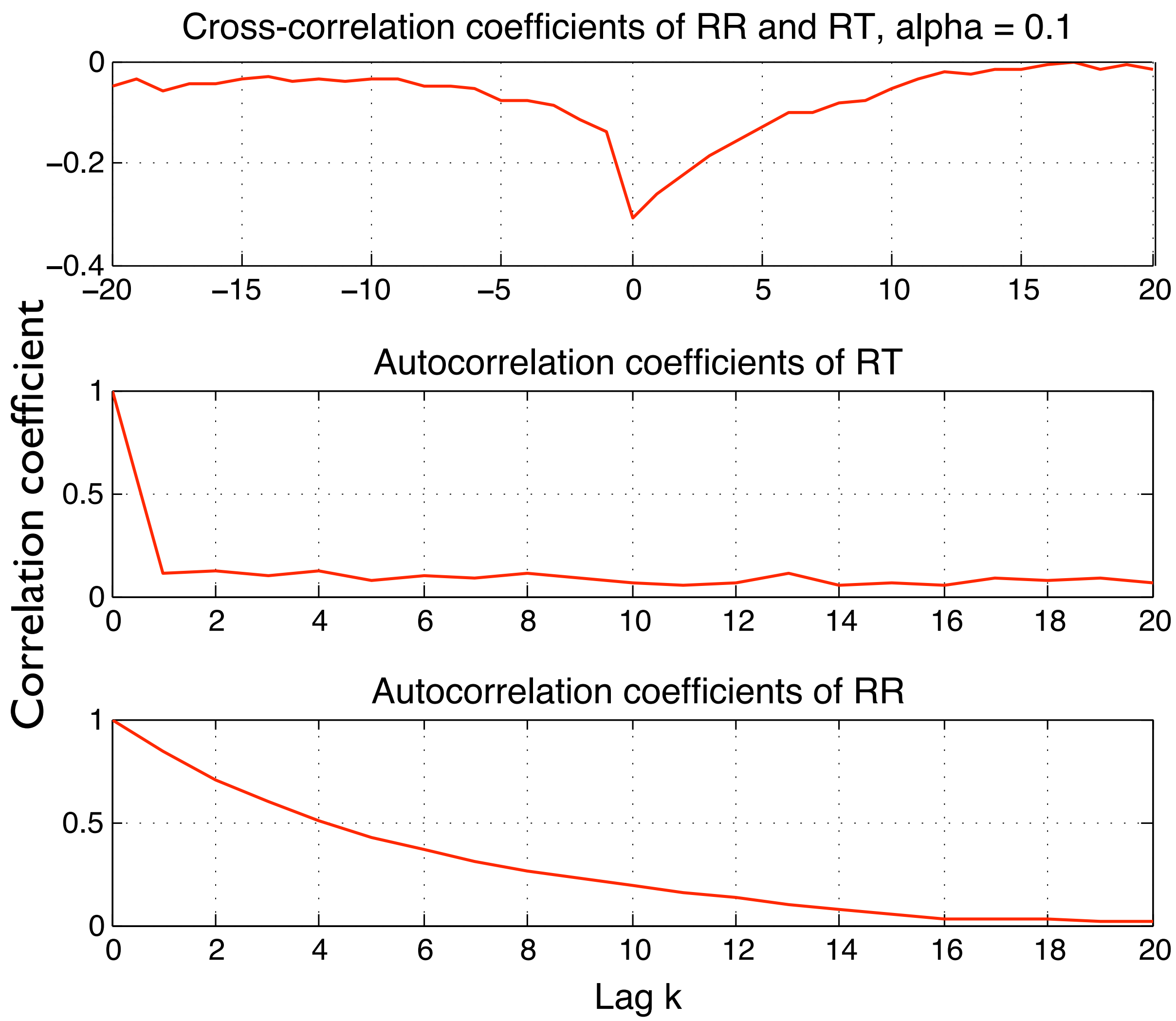
Signal – green (0.15: leftward moving dots; -0.15: rightward moving dot; 0: no signal present)

Error rate



Reward/RT cross-correlation and RT autocorrelation across trials





Empirical data

Model

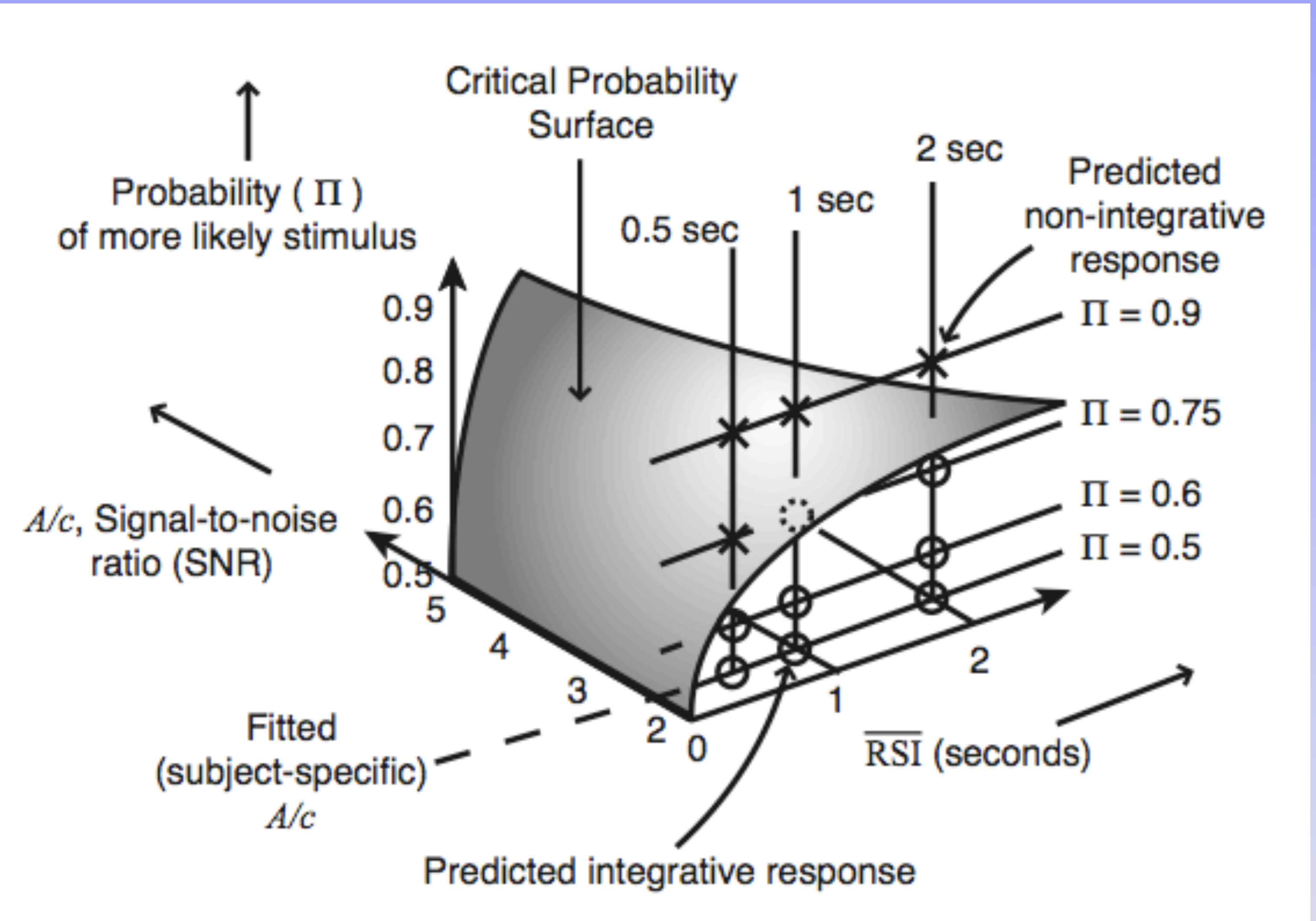
Conclusions

- Behavioral predictions of Bogacz et al. (2006) were borne out (see Simen et al., in press, JEP:HPP; also Bogacz et al., in press, Quarterly J of Euro Psych)
- Fast, nearly optimal adaptation to RSI changes, and autocorrelation in RT, can be explained by exponentially averaging rewards, then setting thresholds inversely proportional to reward rate
- Stochastic gradient ascent can be used over a longer time scale to learn the proportionality constant

Thanks to

- Jonathan D. Cohen
- Phil Holmes
- National Institute of Mental Health

Biased responding (unequal stimulus odds) creates a surface dividing integrative from non-integrative (fast-guess) responding



Bogacz, Moehlis, Brown, Holmes, Cohen (2006) *Psych Review*