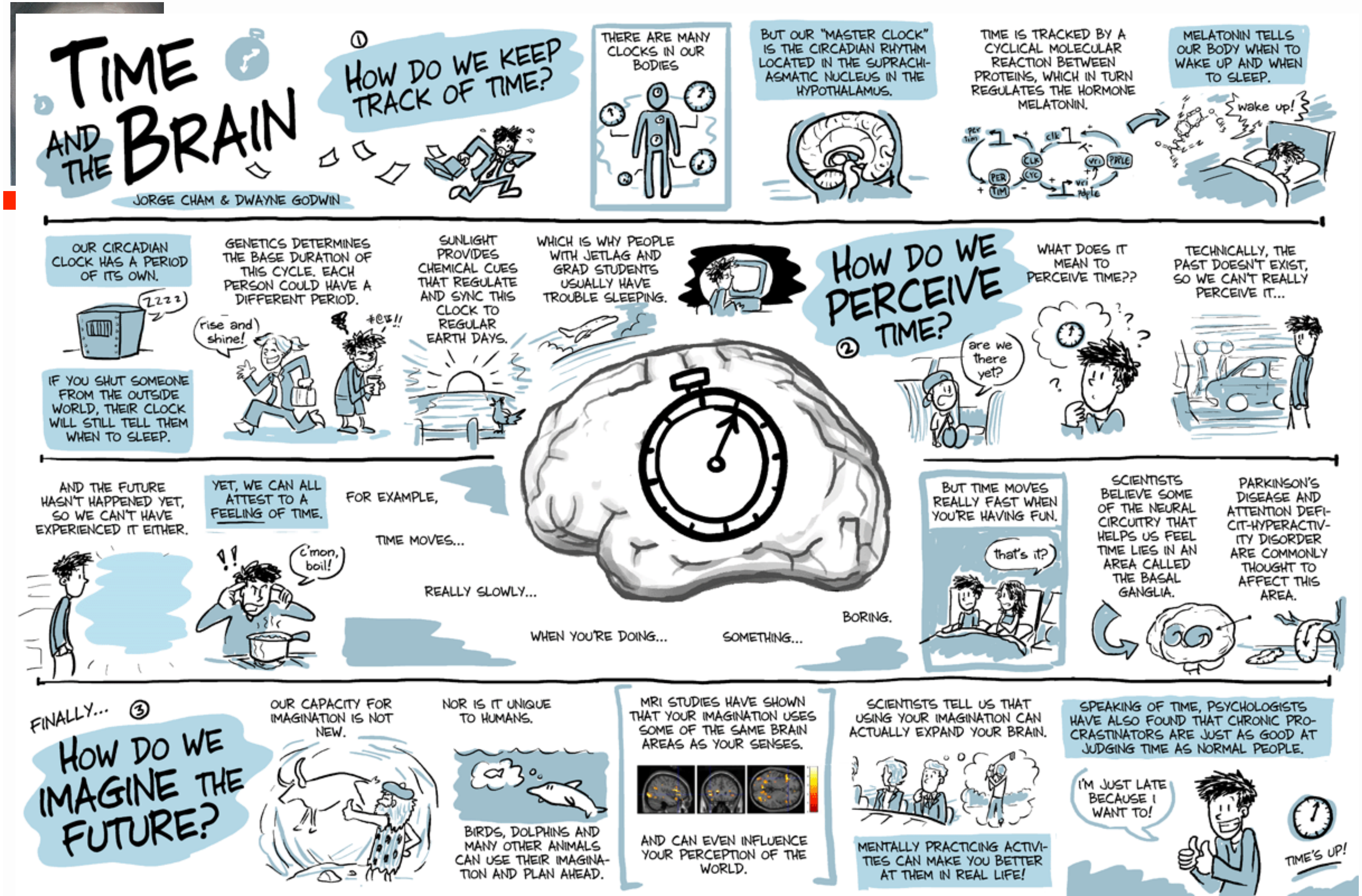
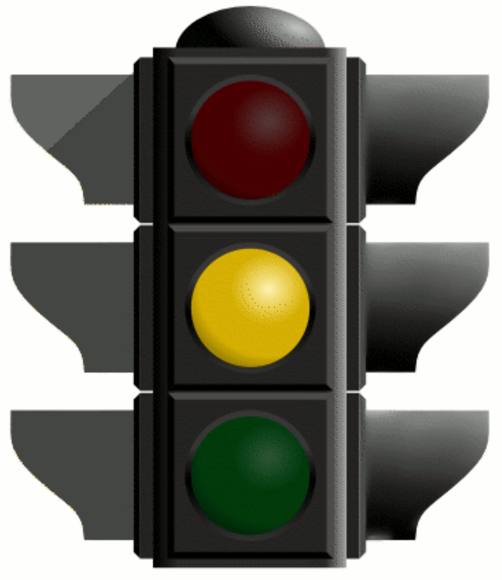
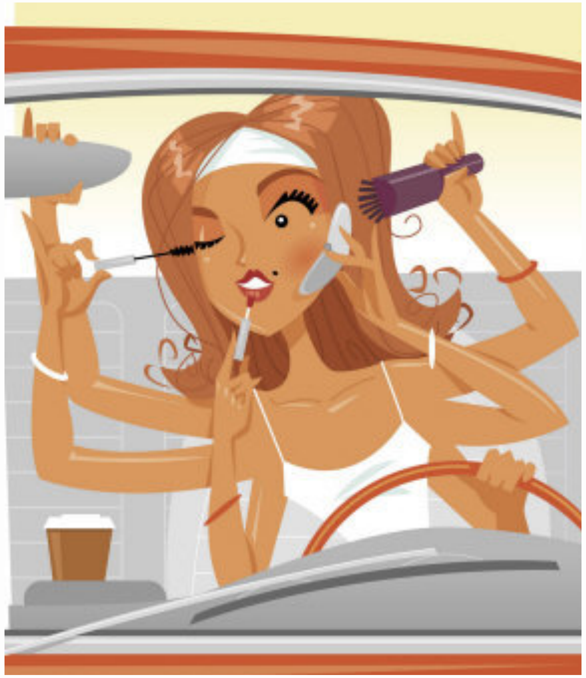


TIME AND THE BRAIN
JORGE CHAM & DWAYNE GODWIN
① HOW DO WE KEEP TRACK OF TIME?
THERE ARE MANY CLOCKS IN OUR BODIES
BUT OUR "MASTER CLOCK" IS THE CIRCADIAN RHYTHM LOCATED IN THE SUPRACHI-ASMATIC NUCLEUS IN THE HYPOTHALAMUS.
TIME IS TRACKED BY A CYCLICAL MOLECULAR REACTION BETWEEN PROTEINS, WHICH IN TURN REGULATES THE HORMONE MELATONIN.
MELATONIN TELLS OUR BODY WHEN TO WAKE UP AND WHEN TO SLEEP.
wake up!
OUR CIRCADIAN CLOCK HAS A PERIOD OF ITS OWN.
ZZZZ
IF YOU SHUT SOMEONE FROM THE OUTSIDE WORLD, THEIR CLOCK WILL STILL TELL THEM WHEN TO SLEEP.
GENETICS DETERMINES THE BASE DURATION OF THIS CYCLE. EACH PERSON COULD HAVE A DIFFERENT PERIOD.
rise and shine!
#@$!!
SUNLIGHT PROVIDES CHEMICAL CUES THAT REGULATE AND SYNC THIS CLOCK TO REGULAR EARTH DAYS.
WHICH IS WHY PEOPLE WITH JETLAG AND GRAD STUDENTS USUALLY HAVE TROUBLE SLEEPING.
② HOW DO WE PERCEIVE TIME?
are we there yet?
WHAT DOES IT MEAN TO PERCEIVE TIME??
TECHNICALLY, THE PAST DOESN'T EXIST, SO WE CAN'T REALLY PERCEIVE IT...
AND THE FUTURE HASN'T HAPPENED YET, SO WE CAN'T HAVE EXPERIENCED IT EITHER.
YET, WE CAN ALL ATTEST TO A FEELING OF TIME.
c'mon, boil!
FOR EXAMPLE,
TIME MOVES...
REALLY SLOWLY...
WHEN YOU'RE DOING...
SOMETHING...
BORING.
BUT TIME MOVES REALLY FAST WHEN YOU'RE HAVING FUN.
that's it?
SCIENTISTS BELIEVE SOME OF THE NEURAL CIRCUITRY THAT HELPS US FEEL TIME LIES IN AN AREA CALLED THE BASAL GANGLIA.
PARKINSON'S DISEASE AND ATTENTION DEFI-CIT-HYPERACTIV-ITY DISORDER ARE COMMONLY THOUGHT TO AFFECT THIS AREA.
FINALLY... ③ HOW DO WE IMAGINE THE FUTURE?
OUR CAPACITY FOR IMAGINATION IS NOT NEW.
NOR IS IT UNIQUE TO HUMANS.
BIRDS, DOLPHINS AND MANY OTHER ANIMALS CAN USE THEIR IMAGINA-TION AND PLAN AHEAD.
MRI STUDIES HAVE SHOWN THAT YOUR IMAGINATION USES SOME OF THE SAME BRAIN AREAS AS YOUR SENSES.
AND CAN EVEN INFLUENCE YOUR PERCEPTION OF THE WORLD.
SCIENTISTS TELL US THAT USING YOUR IMAGINATION CAN ACTUALLY EXPAND YOUR BRAIN.
MENTALLY PRACTICING ACTIVI-TIES CAN MAKE YOU BETTER AT THEM IN REAL LIFE!
SPEAKING OF TIME, PSYCHOLOGISTS HAVE ALSO FOUND THAT CHRONIC PRO-CRASTINATORS ARE JUST AS GOOD AT JUDGING TIME AS NORMAL PEOPLE.
I'M JUST LATE BECAUSE I WANT TO!
TIME'S UP!

Time Estimation embedded in a general cognitive architecture

Estimating time is a component of skilled performance



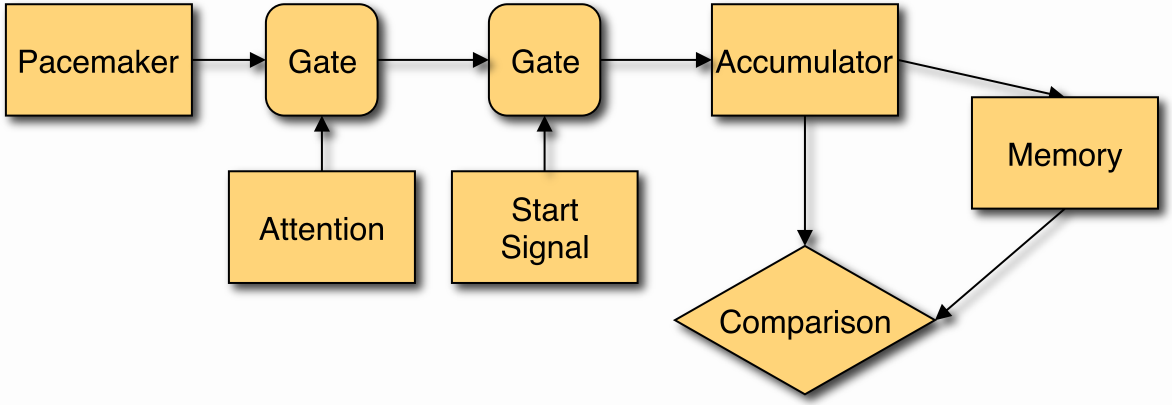


Overview

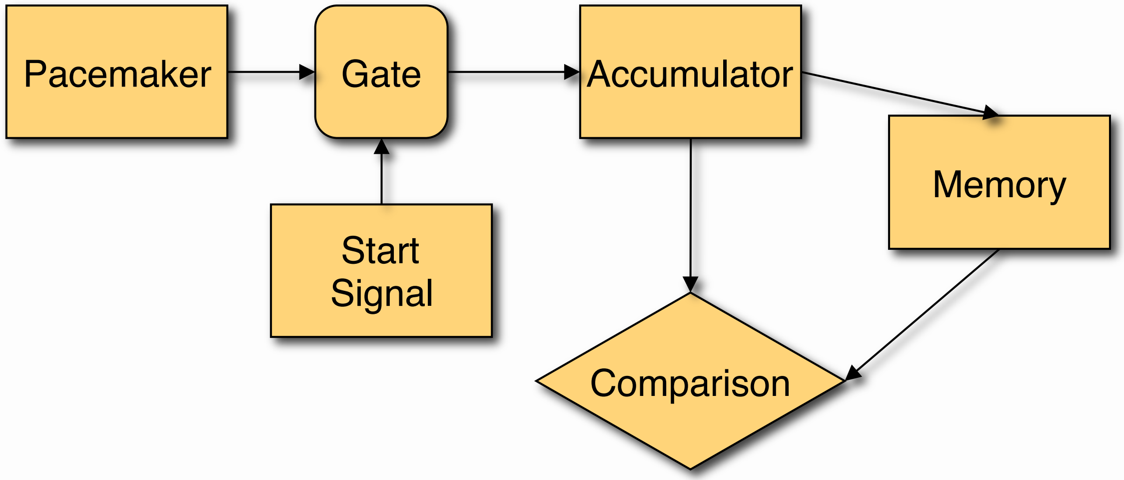
- The time estimation module: how does it work? (Relevant for the assignment)
  - Bisection
- The two papers: mental manipulation of interval representations
  - Memory for time intervals

Time estimation: many models

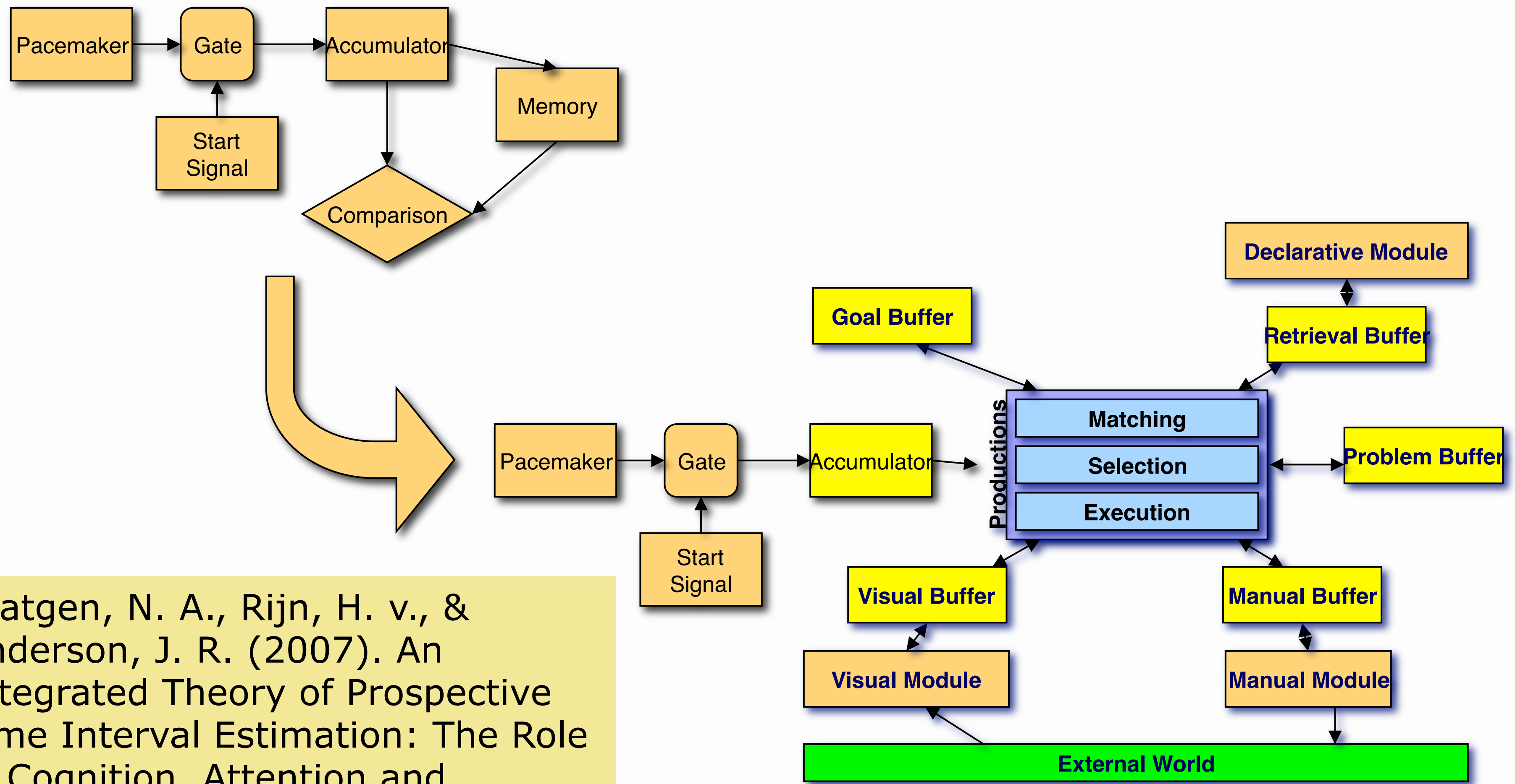
Attentional gating
(Zakay, Block)



Internal clock (Gibbon, Meck)



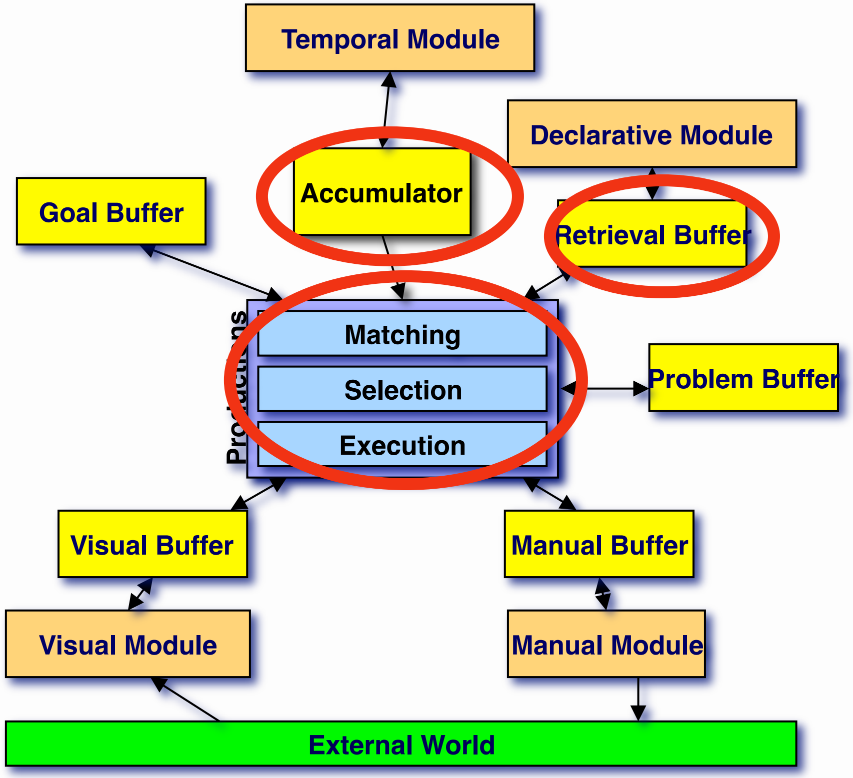
Embed time estimation in a general cognitive architecture



Taatgen, N. A., Rijn, H. v., & Anderson, J. R. (2007). An Integrated Theory of Prospective Time Interval Estimation: The Role of Cognition, Attention and Learning. *Psychological Review, 114*(3), 577-598.

Attention

- ACT-R basically conforms with central bottleneck theories (e.g., Pashler, see Salvucci & Taatgen, 2008)



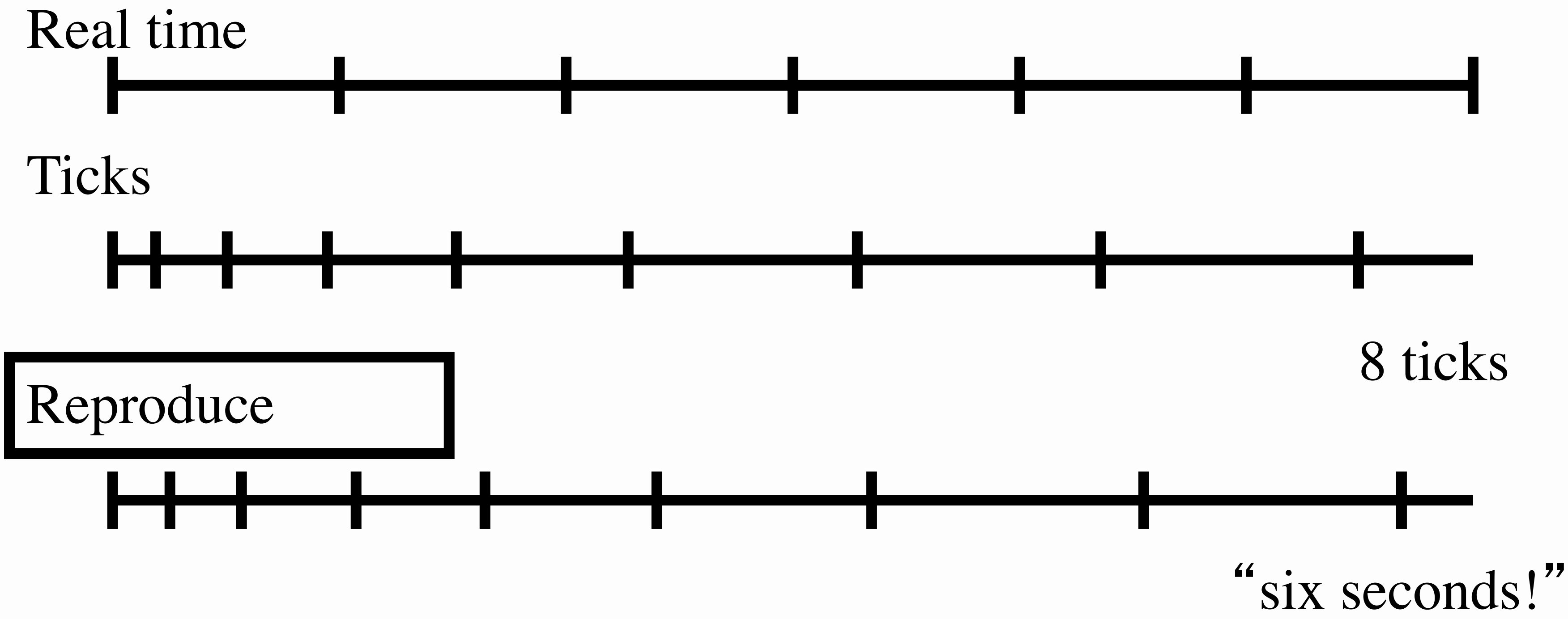
Learning

- ACT-R uses a form of instance learning (Logan)
  - Experiences with an interval are stored in memory
  - Memory is subjected to decay
- ACT-R uses rule learning: can explain how judging a particular interval can be automated

Implementation of time estimation

- Idea: metronome that starts ticking fast but gradually slows down
- The current value of the accumulator is available to the rest of cognition
- At any moment, the accumulator can be read or matched in the condition of a production

Example: estimate 6 seconds



How long are the pulses?

t_0 = the starting pulse

$$t_{n+1} = at_n + \text{noise}(M = 0,\ SD = b \cdot at_n)$$

Noise = logistic noise

Two parameter sets:

t_0 = 11 ms, a = 1.1, b = 0.015

t_0 = 100 ms, a = 1.02, b = 0.015

Assignment for the next two weeks

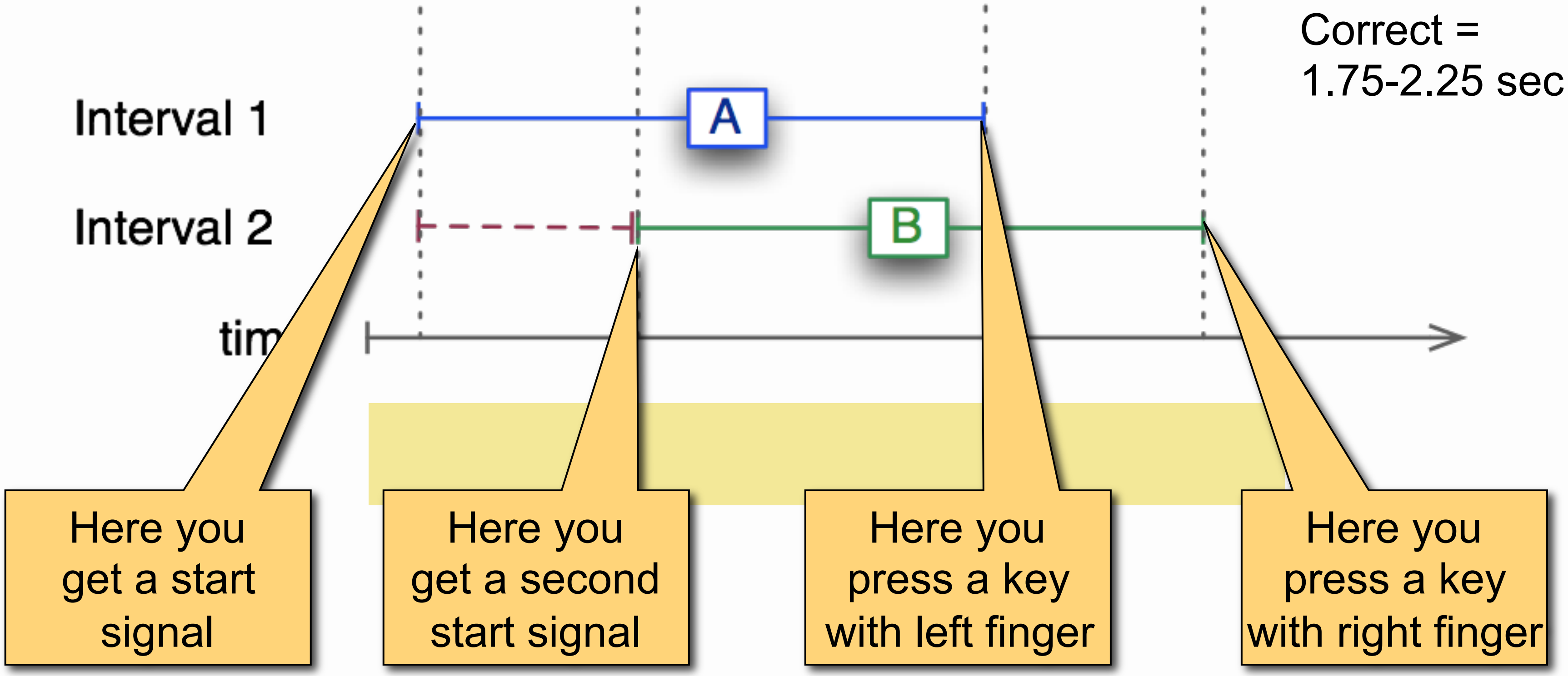
- Implement the ACT-R timing module
  - Specifically, you will need
    - A function that converts pulses into time
    - A function that converts time into pulses
- Model Bisection

Cognitive Architecture allows modeling time perception in real tasks

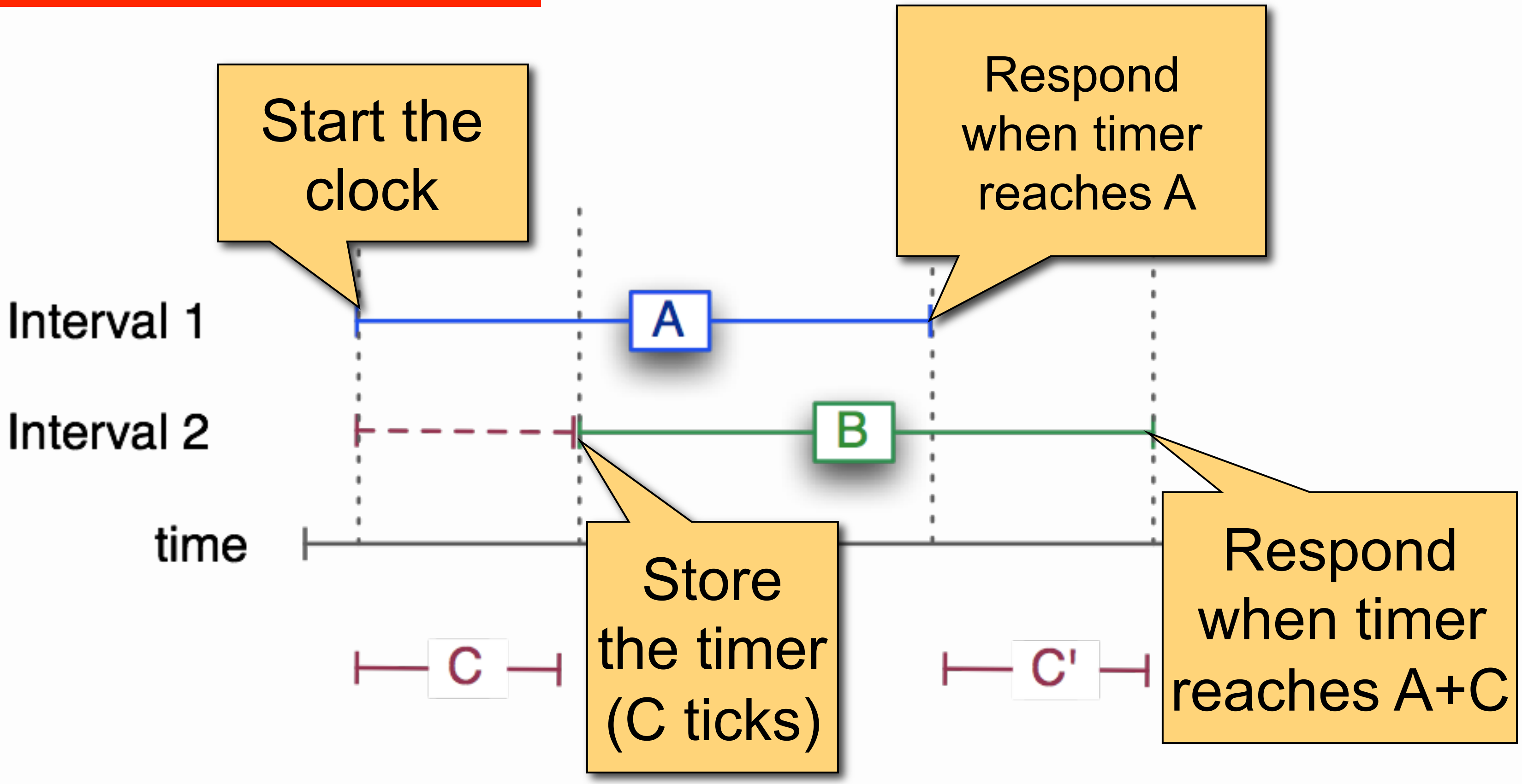
(Borst & Taatgen, 2007)

Dual timing paradigm

Interval =
2 sec
Correct =
1.75-2.25 sec

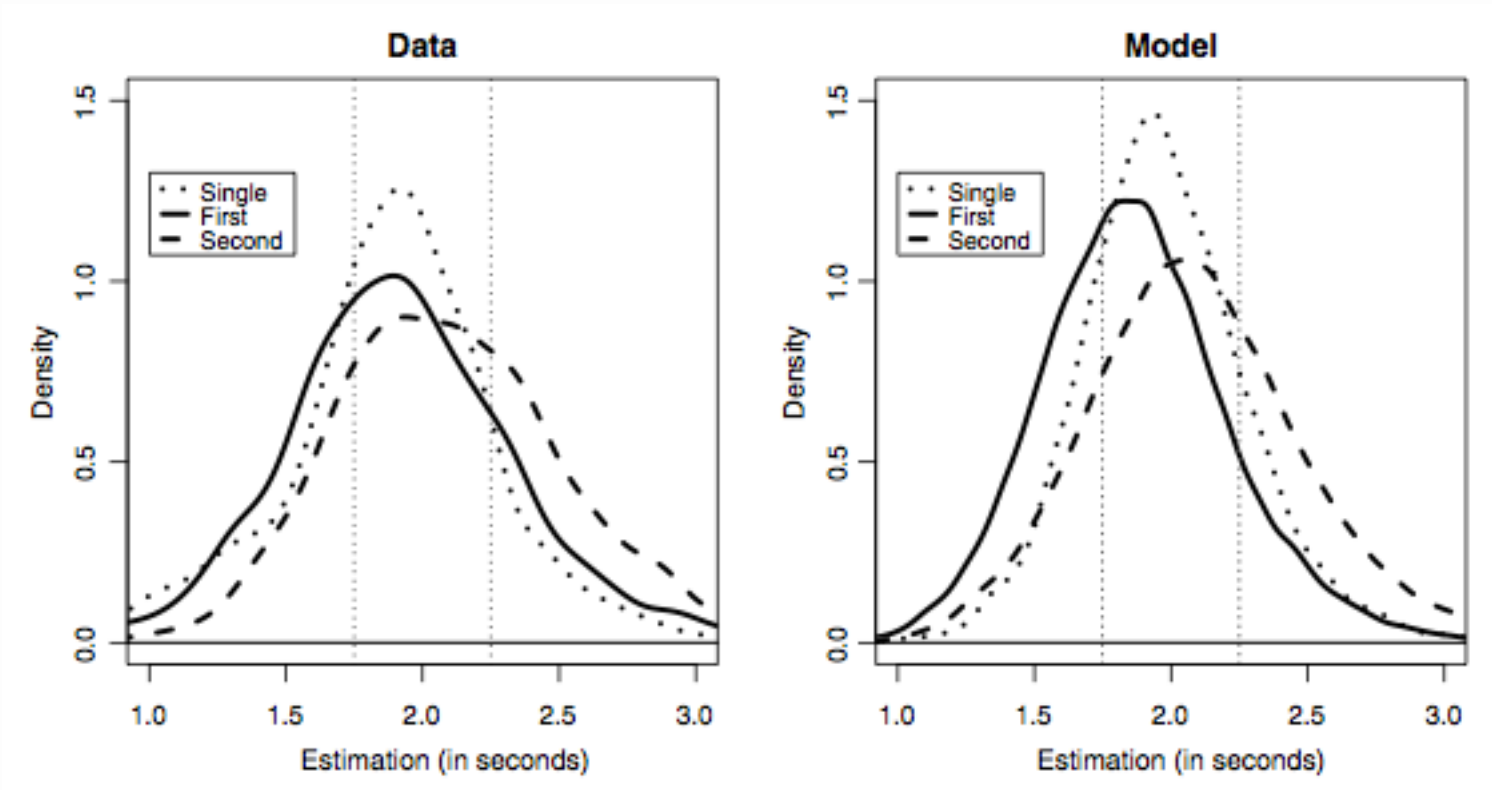


Model



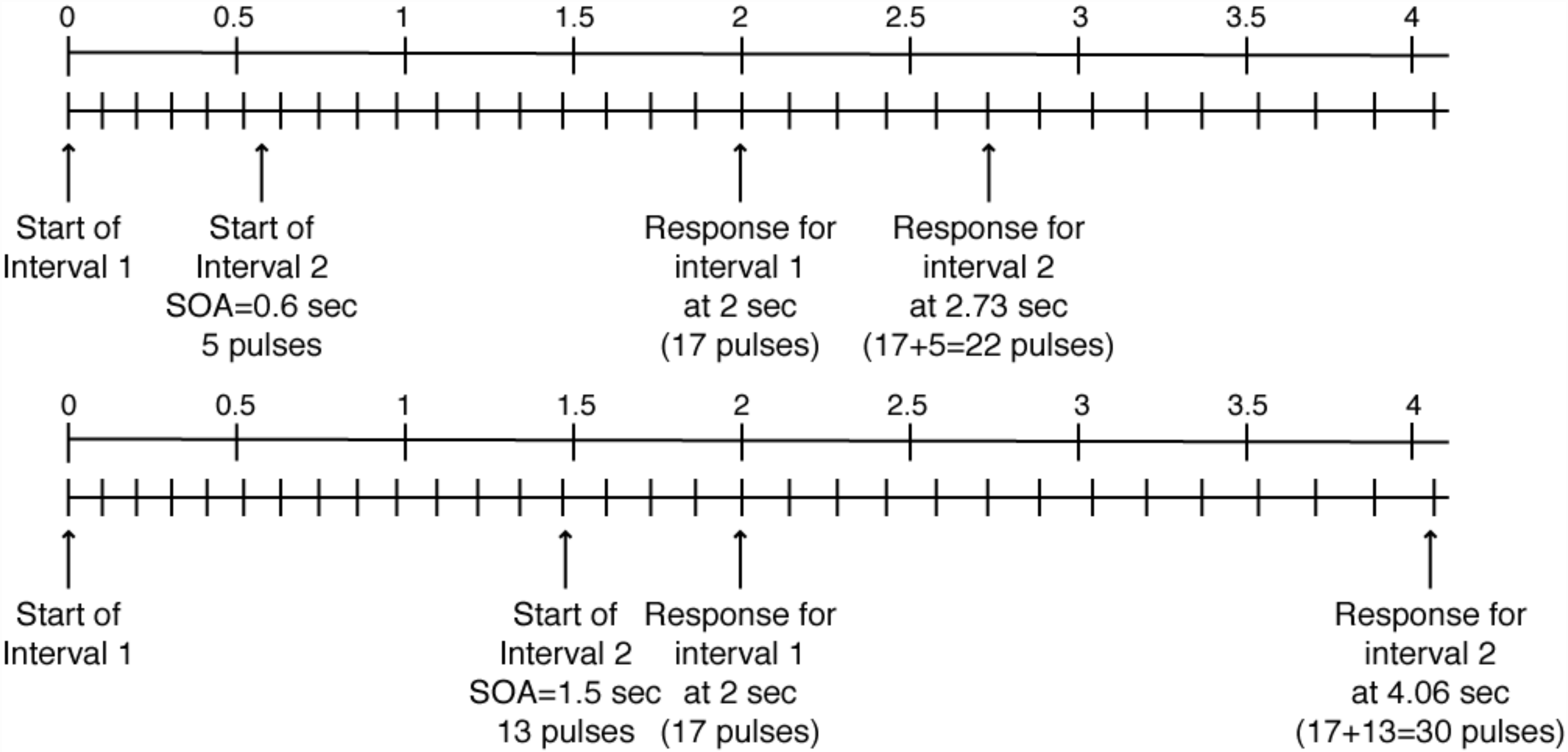
van Rijn & Taatgen (2008)

Results

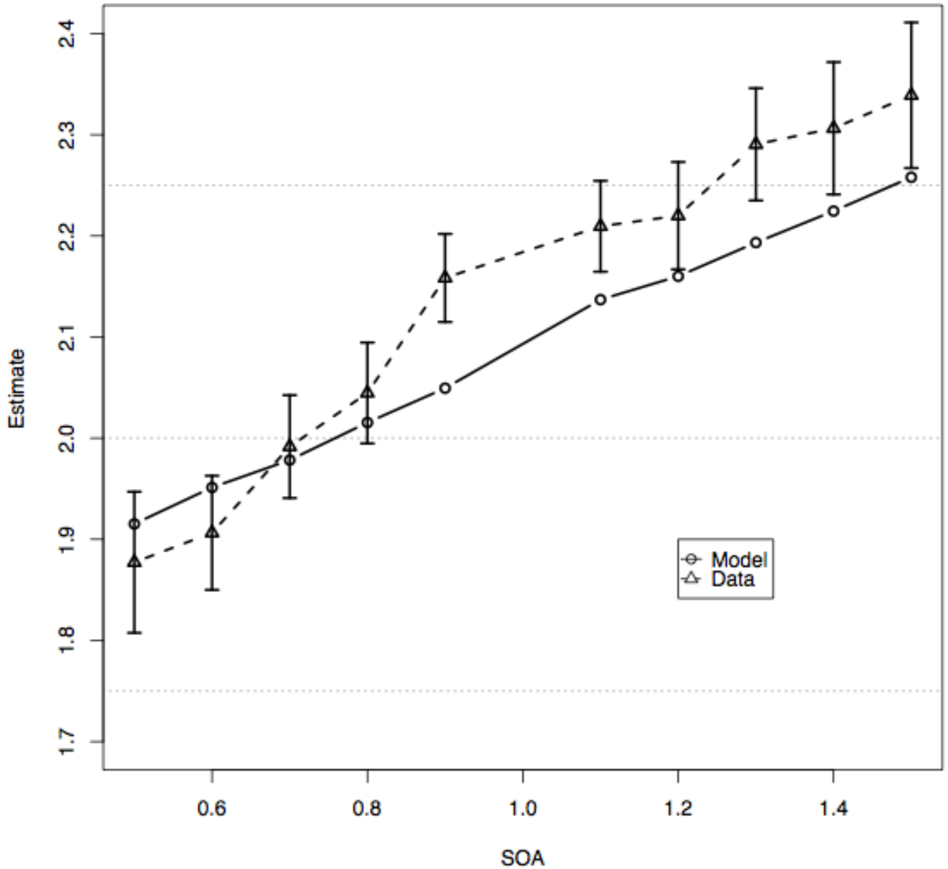


Data
Model
Single
First
Second
Density
Estimation (in seconds)
0.0
0.5
1.0
1.5
2.0
2.5
3.0

Non-linear scale biases second estimate



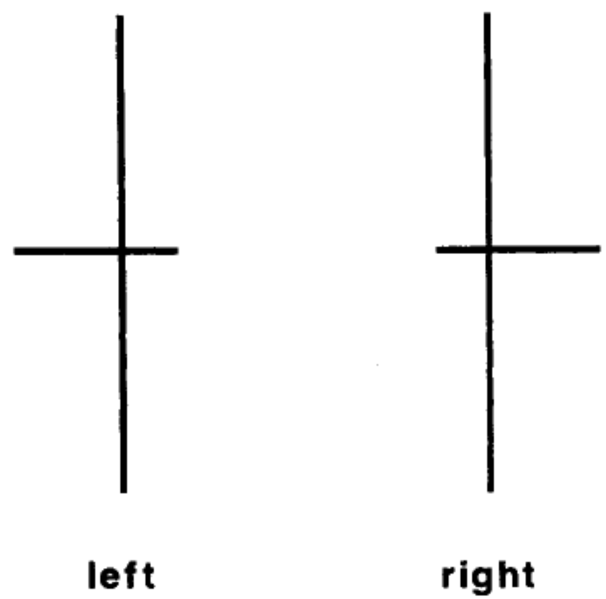
Result: Estimate of second interval as a function of SOA



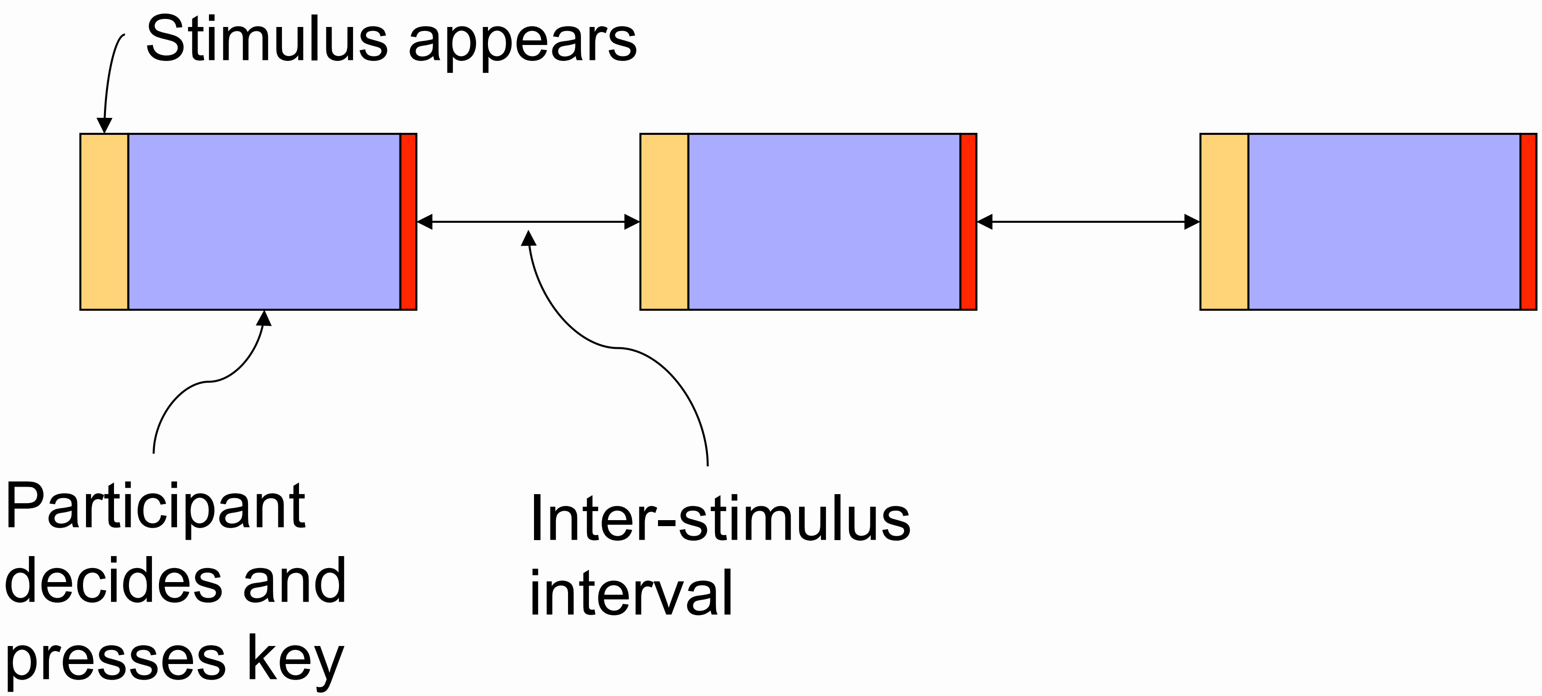
- Relationship between the ACT-R temporal module and the SBF model
- Staddon & Higa alternative: decaying memory
- Resetting the timer after each interval, can it explain the data?
- Will estimation become progressively worse if we have a long series of overlapping intervals?

Time estimation: anticipation

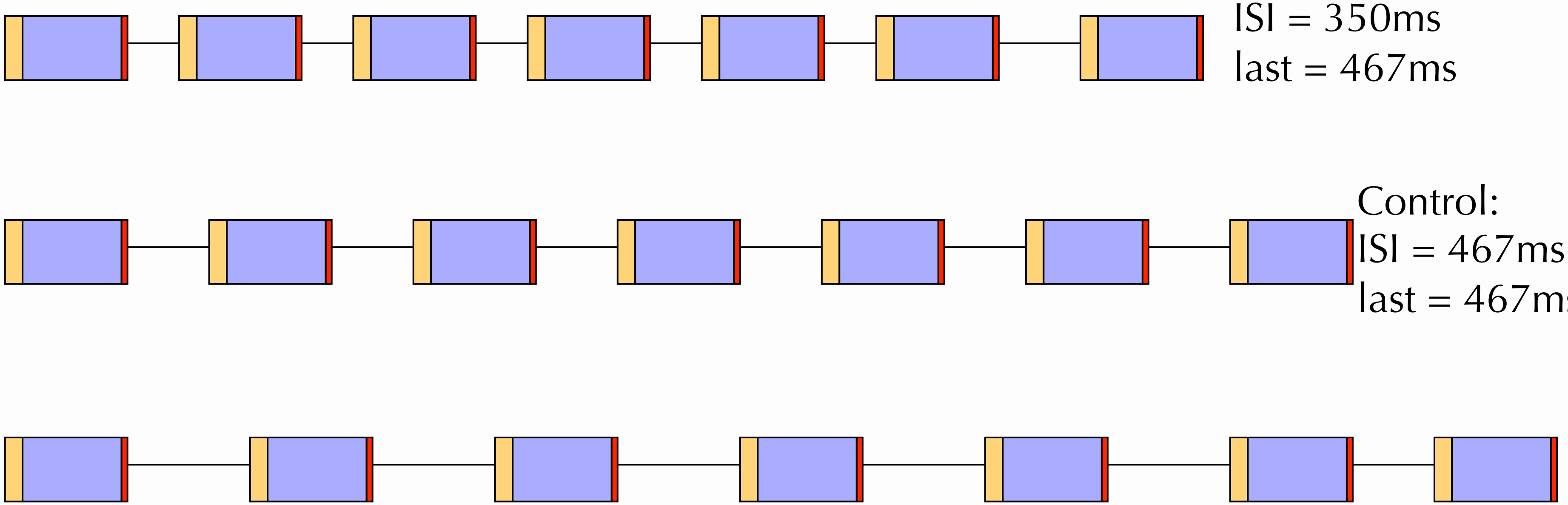
- Choice-reaction time experiment
- The inter-stimulus interval was varied on the last trial of each block



Time-line of a block



Design

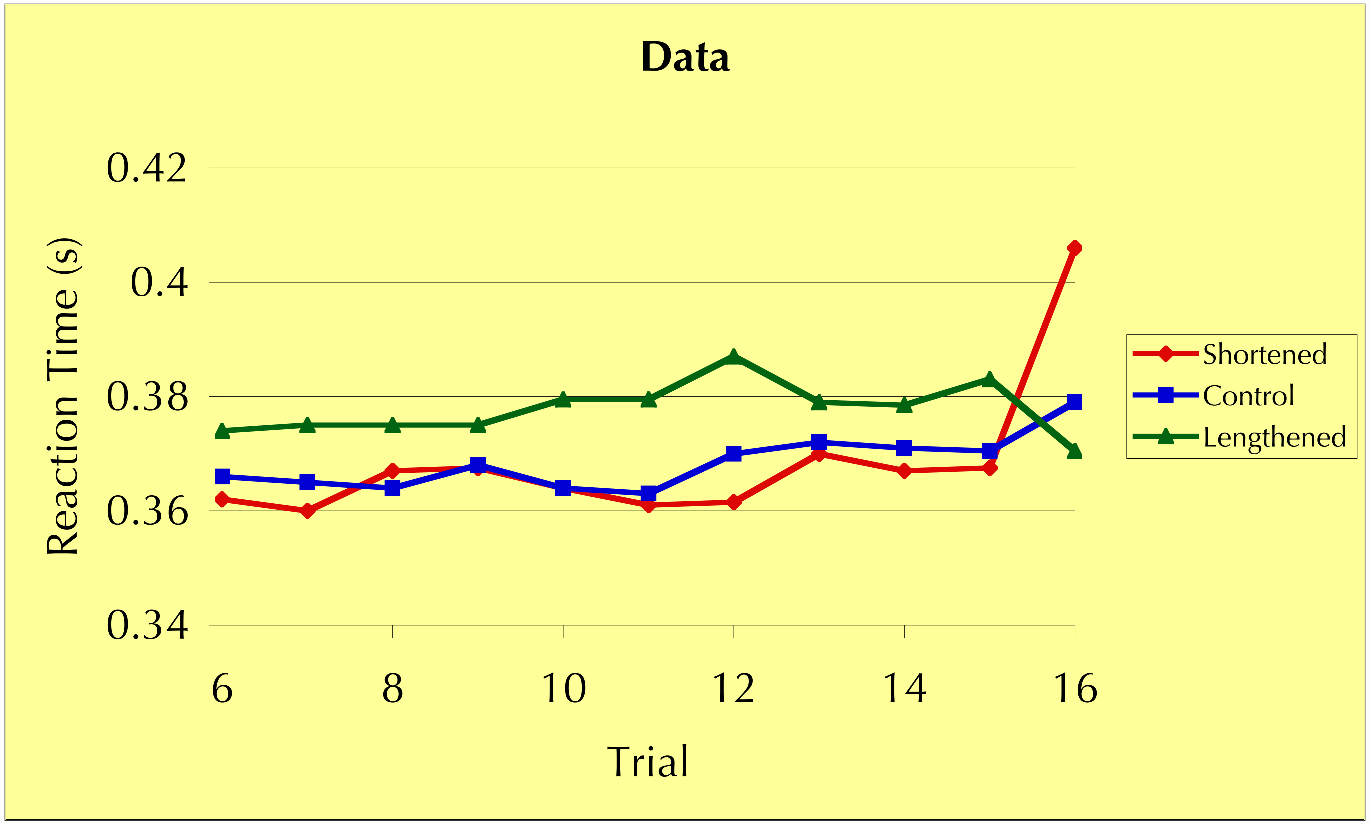


Lengthening:
ISI = 350ms
last = 467ms

Control:
ISI = 467ms
last = 467ms

Shortening:
ISI = 700ms
last = 467ms

Data


Data
0.42
0.4
0.38
0.36
0.34
Reaction Time (s)
6
8
10
12
14
16
Trial
Shortened
Control
Lengthened

Model

- What could be the profit of timing?
- If we know when and where the information will be, we can immediately do a +visual> on the location without waiting for the =visual-location>, saving up to 50ms
- But if the stimulus is early, we cannot use this advantage, because it will “surprise” us

At the start of the trial we start the timer

```
(p start-timer
   =goal>
     isa crt
     status nil
==>
   +temporal>
      isa time
   =goal>
     status waiting)
```

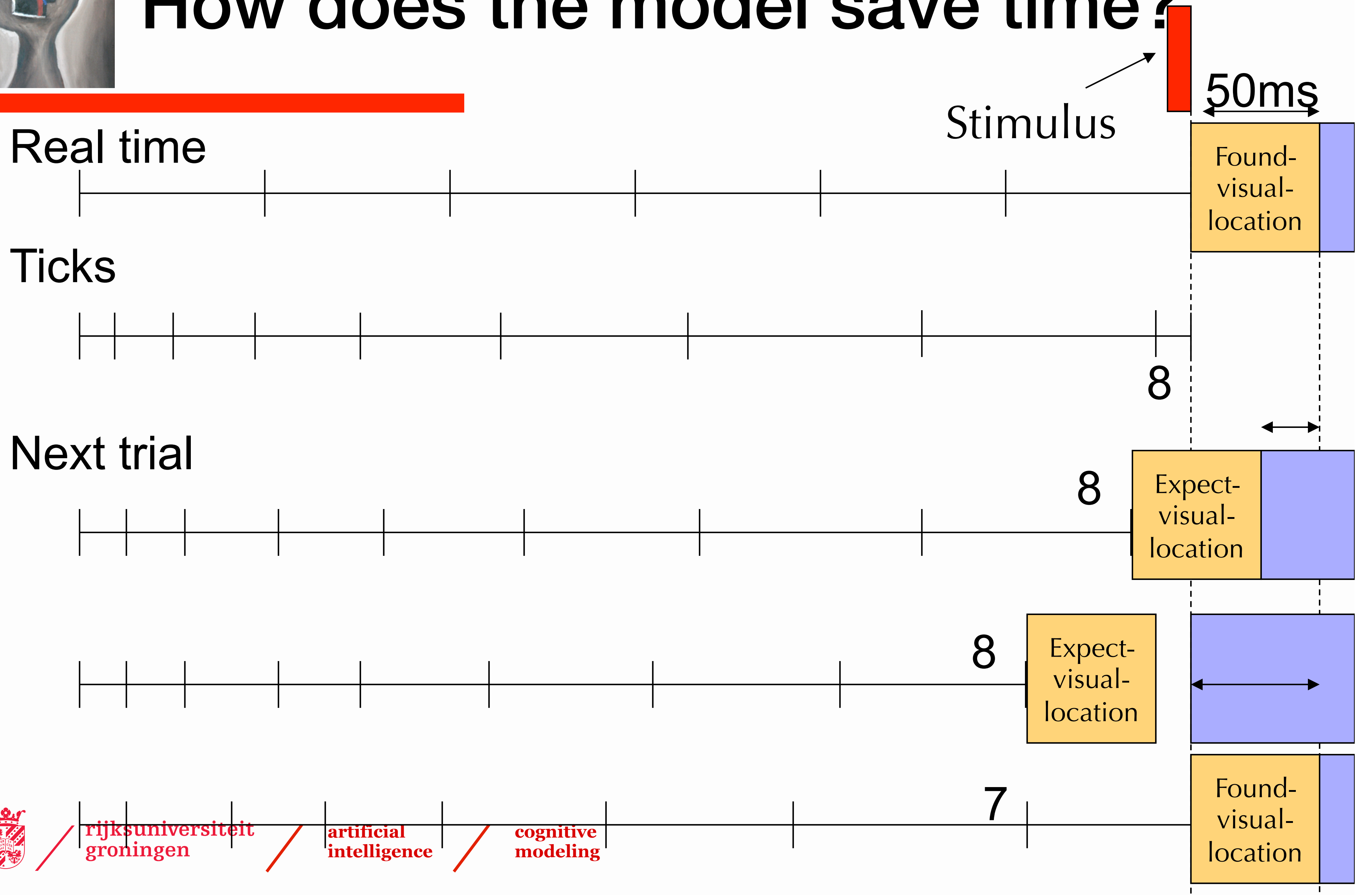
When the stimulus comes up, we process it, but also store the time

```
(p found-visual-location
   =goal>
     isa crt
     status waiting
   =visual-location>
     isa visual-location
   =visual-state>
     isa module-state
     modality free
   =temporal>
     isa time
     ticks =ticks
==>
   +visual>
     isa visual-object
     screen-pos =visual-location
   =goal>
     time =ticks
     loc =visual-location
     status wait-visual
```

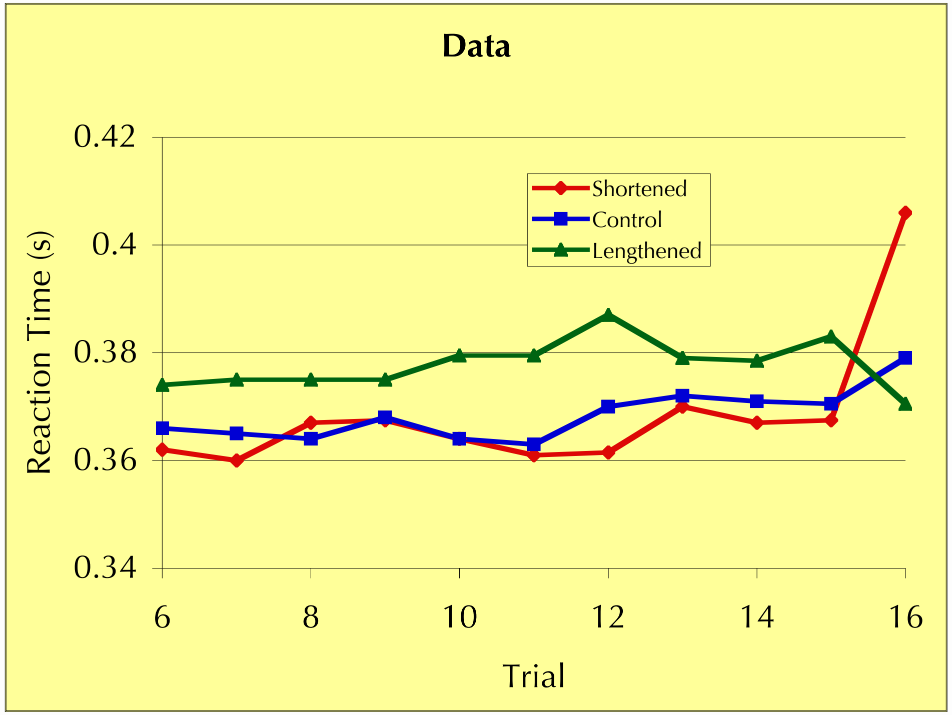
For the next trial, we can try to anticipate the stimulus, and save up to 50ms

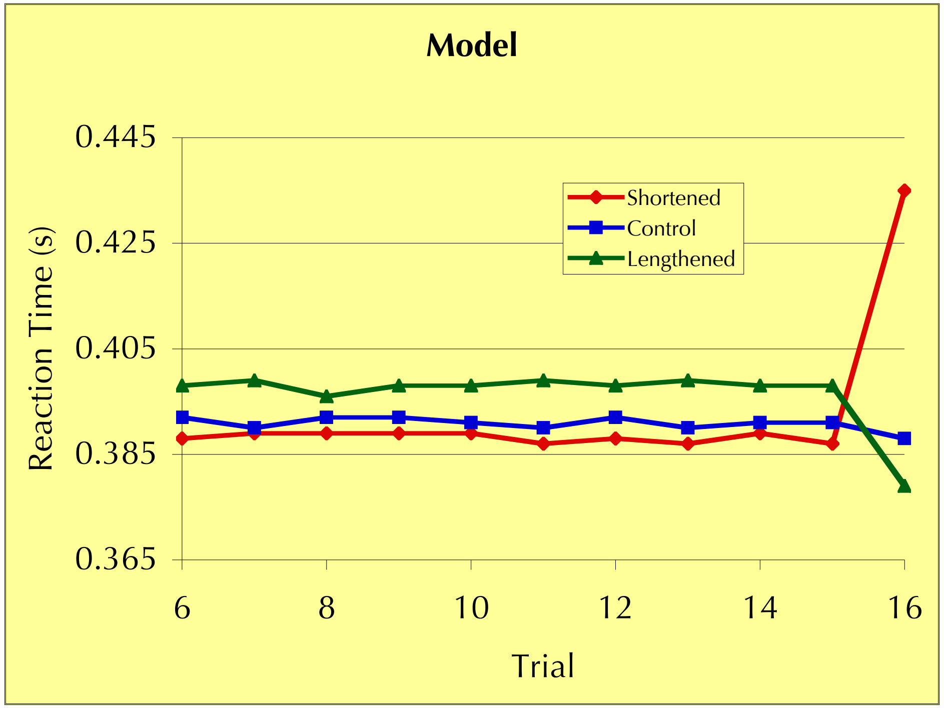
```
(p expect-visual-location
   =goal>
     isa crt
     status waiting
     time =ticks
     loc =loc
   =temporal>
     isa time
     ticks =ticks
==>
   +visual>
     isa visual-object
     screen-pos =loc
   =goal>
     status wait-visual)
```

How does the model save time?



Results of the model with no parameter fitting





What is the goal of this article? Because if you want to look at whether there is a temporal module in the brain or not, is it not easier to use falsification based on simple experiments? Such as, breathing and heart rate should be independent of the interval estimation task. Thus, given that the existence of the temporal module is your hypothesis, what is the advantage of using a cognitive model to test this as compared to simple experiments based on predictions?

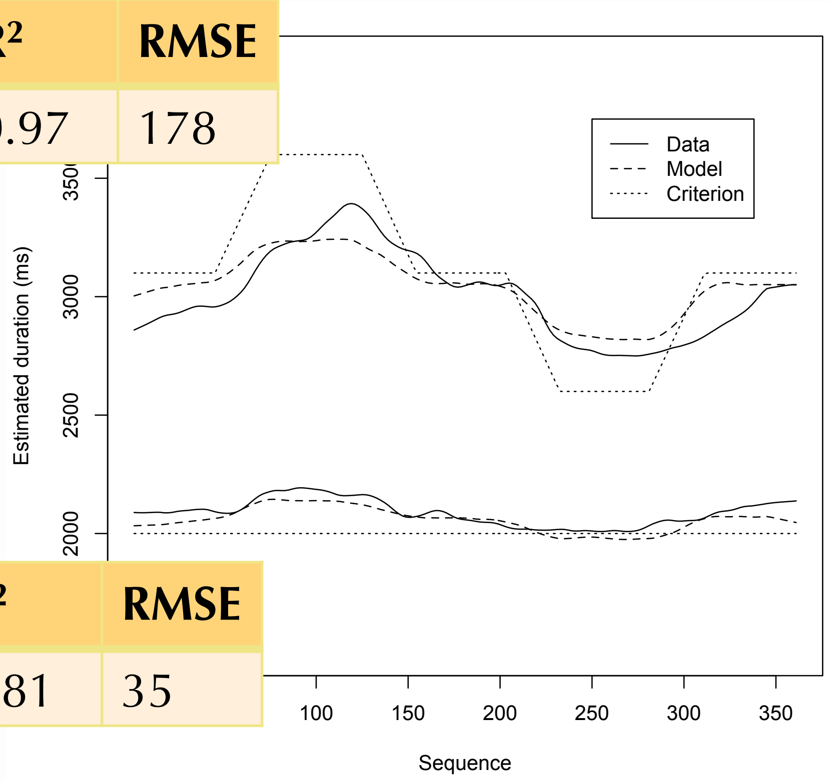
In the description of the interval estimation experiments it is mentioned that the participants were forbidden from counting. In general telling people not to think of something immediately makes them think of that particular thing. Would telling them not to count not have a similar effect? If not, would it have any other effect on the outcome of the experiment results, making it more plausible that the model resembles the experiment results, since the model uses some form of counting as well? In other words: would the model still correctly approximate the test results if the participants would not have counted because they did not know they had to keep track of time instead of just being told not to count?

- Does the proposed model also explain variances in rhythmic processes that humans produce? (E.g. in drumming, see this paper (http://bit.ly/DrumPaper) discussed in this Sixty Symbols video (https://youtu.be/GyLeBMdl_HU).)

Dikes and Rivers experiment

Goal: investigate the role of memory

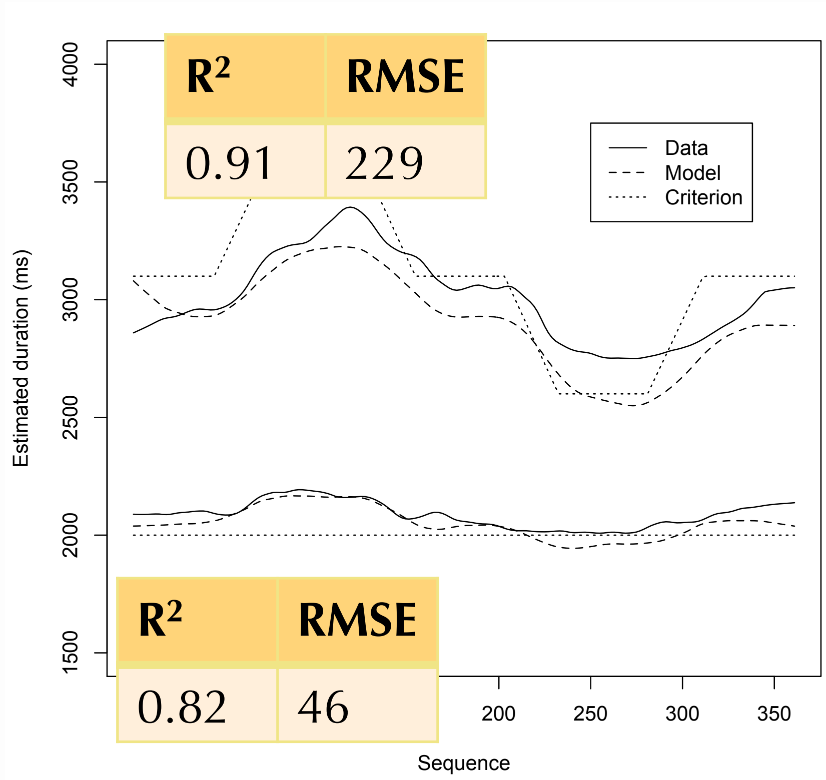
What is a better fit of the data?



R^2	RMSE
0.97	178

R^2	RMSE
0.81	35

Model A



R^2	RMSE
0.91	229

R^2	RMSE
0.82	46

Model B

Contents

- What is the experiment?
- What is Model A and Model B?
- Why is Model B a better after all?
- What lessons can we learn from this example?

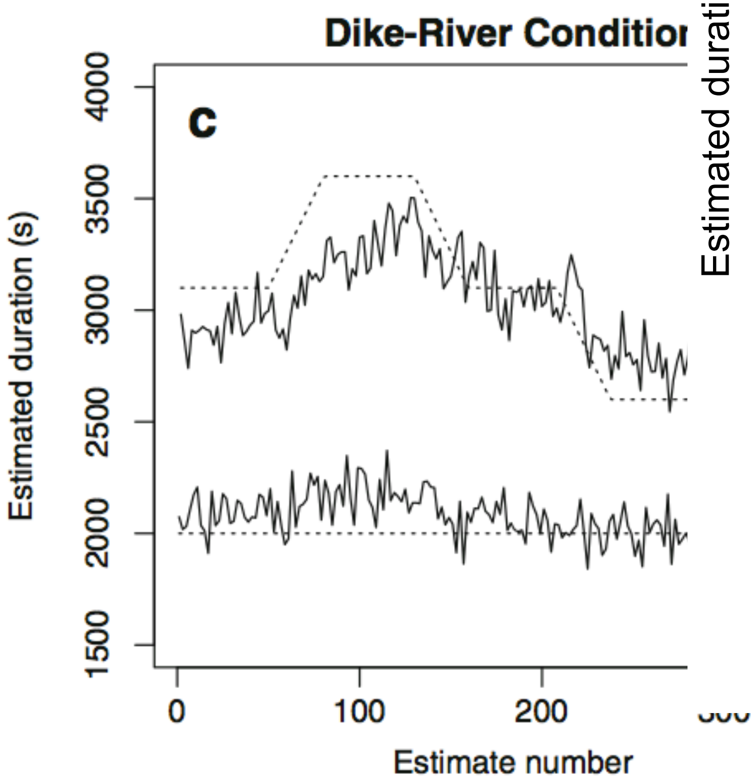
Dikes-and-Rivers experiment

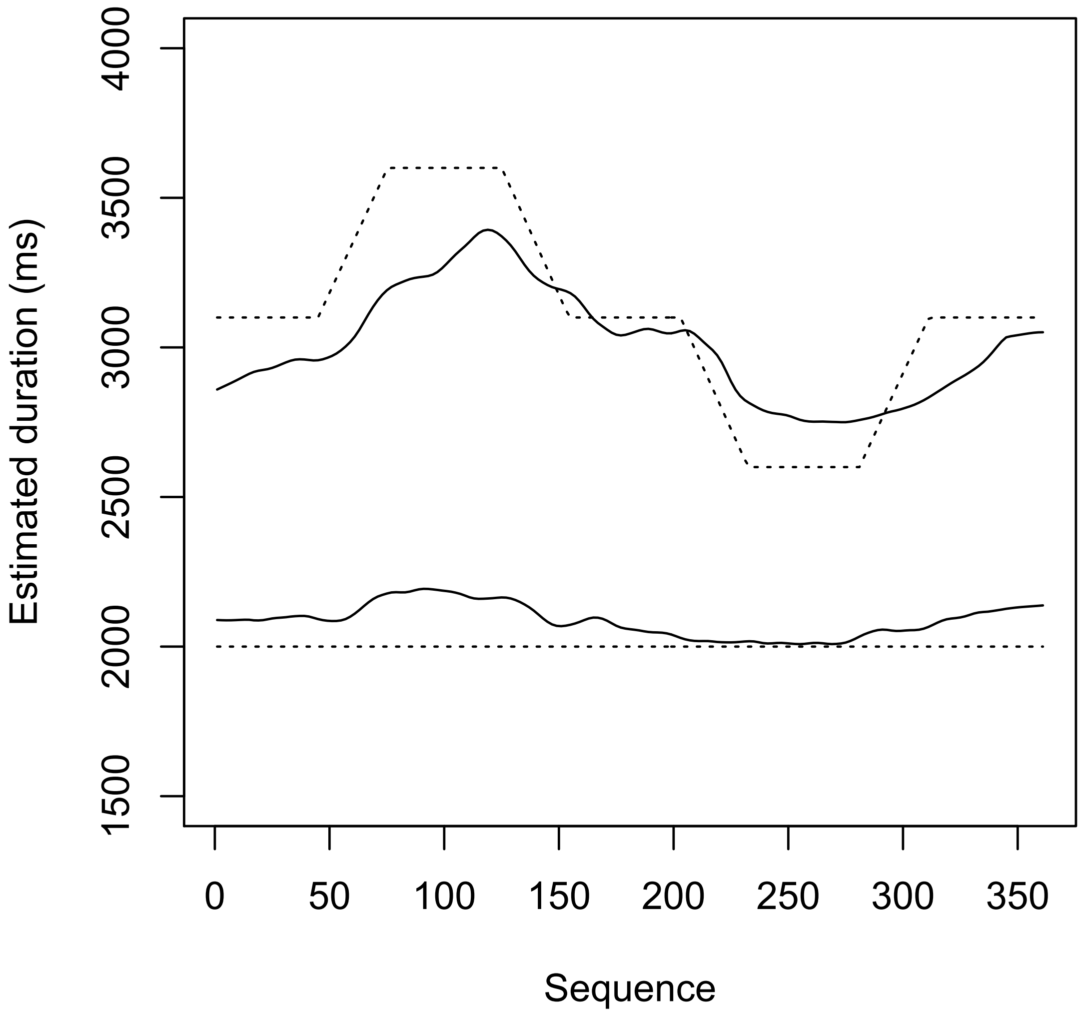
- Time estimation
- In earlier work, we found that if subjects need to memorize two time intervals, these intervals start to ”contaminate” each other
- The goal of this experiment is to investigate this contamination in more detail

Experiment

- Subjects alternate producing time intervals of a short and a long duration.
- Initially, short is 2 seconds and long is 3.1 seconds
- They receive feedback on whether their estimate is within +/- 12.5% of the target, and receive “too short” or “too long” as feedback otherwise.
- After a number of trials, the criterion for the long interval starts to change.

Results


Dike-River Condition
c
Estimated duration (s)
4000
3500
3000
2500
2000
1500
0
100
200
Estimate number


Estimated duration (ms)
4000
3500
3000
2500
2000
1500
0
50
100
150
200
250
300
350
Sequence

Analysis: Mixed-Effect models

Start with a simple regression equation:

$$\text{short}_{n,s} = \beta_0 + r_s + \varepsilon_{n,s}$$

Then add factors as long as the more complex equation fits the data significantly better than the previous model

$$\text{short}_{n,s} = \beta_0 + \beta_1\text{short}_{\text{n-1,s}} + \beta_2\text{short}_{\text{n-2,s}} + r_s + \varepsilon_{n,s}$$

Final set of factors for short interval

Fixed Effect	**Value of β**	***t* value**
Intercept	657 ms	4.6
short$_{n-1}$	0.385	8.3
short$_{n-2}$	0.085	3.3
short-fb-S$_{n-1}$	110 ms	3.1
short-fb-L$_{n-1}$	-208 ms	-6.5
long$_{n-1}$	0.16	5.1
long-fb-S$_{n-1}$	92.6 ms	3.2
long-fb-L$_{n-1}$	-163 ms	-4.2

For long interval

Fixed Effect	Value of β	*t* value
Intercept	695 ms	3.8
$long_{n-1}$	0.34	8.5
$long_{n-2}$	0.16	4.0
$long_{n-3}$	0.12	4.6
$long_{n-4}$	0.05	1.9
long-fb-S_{n-1}	159 ms	4.8
long-fb-L_{n-1}	-118 ms	-2.5
long-fb-S_{n-2}	82.9 ms	2.5
long-fb-L_{n-2}	3.8 ms	0.1
$short_{n-1}$	0.15	2.9
short-fb-S_{n-1}	85 ms	2.1
short-fb-L_{n-1}	-107 ms	-6.5

Model

- Based on ACT-R's declarative memory and temporal module, implemented in R
- Declarative memory using blending (Lebiere et al., 2007)

Declarative memory with blending

- Model uses base-level leaning and a mismatch penalty when on a short/long mismatch:

$$A(t) = \log(t - t_{creation})^{-d} + mismatchpenalty$$

Blending

- Each chunk has a probability of being retrieved:

$$P_i = \frac{e^{A_i/t}}{\sum_j e^{A_j/t}}$$

- The retrieved interval is a blend of all chunks

$$\text{Result value} = \sum_j P_j V_j$$

How to calculate the blend?

- Calculate the activations of all candidate chunks
- Apply mismatch penalties to chunks that do not match completely

- Example: retrieve a blend for long:
  - B(short, 9)=2.0
  - B(short,10)=1.0
  - B(long,16)=2.0
  - B(long,17)=1.5
- Apply penalty (of -2):
  - A(short, 9)=0.0
  - A(short,10)=-1.0
  - A(long,16)=2.0
  - A(long,17)=1.5

How to calculate the blend?

- Now calculate the probability of recall for each of the candidates, using:

$$P_i = \frac{e^{A_i/t}}{\sum_j e^{A_j/t}}$$

- Apply penalty:
  - A(short, 9)=0.0
  - A(short,10)=-1.0
  - A(long,16)=2.0
  - A(long,17)=1.5
- Results in (t=1):
  - p(short, 9)=0.076
  - p(short, 10)=0.0278
  - p(long, 16)=0.558
  - p(long, 17)=0.339

How to calculate the blend?

- Multiply each probability with the slot value, and add it all up to get the blended value

$$\text{Result value} = \sum_j P_j V_j$$

- Probabilities
  - p(short, 9)=0.076
  - p(short, 10)=0.0278
  - p(long, 16)=0.558
  - p(long, 17)=0.339
- 0.076*9 + 0.0278*10 + 0.558* 16 + 0.339 * 17 =
  15.65

Contents

- Model of Grosjean et al. data
- Dikes and Rivers paper
- How to model an experiment

Handling feedback

- In order to handle feedback, the model also stores the feedback on the previous trial in declarative memory
- "Too short" is stored as a positive number, "Too long" as a negative number, and "Correct" as 0.
- To determine the number of pulses to wait, the model retrieves the duration and the feedback from memory, and adds these together

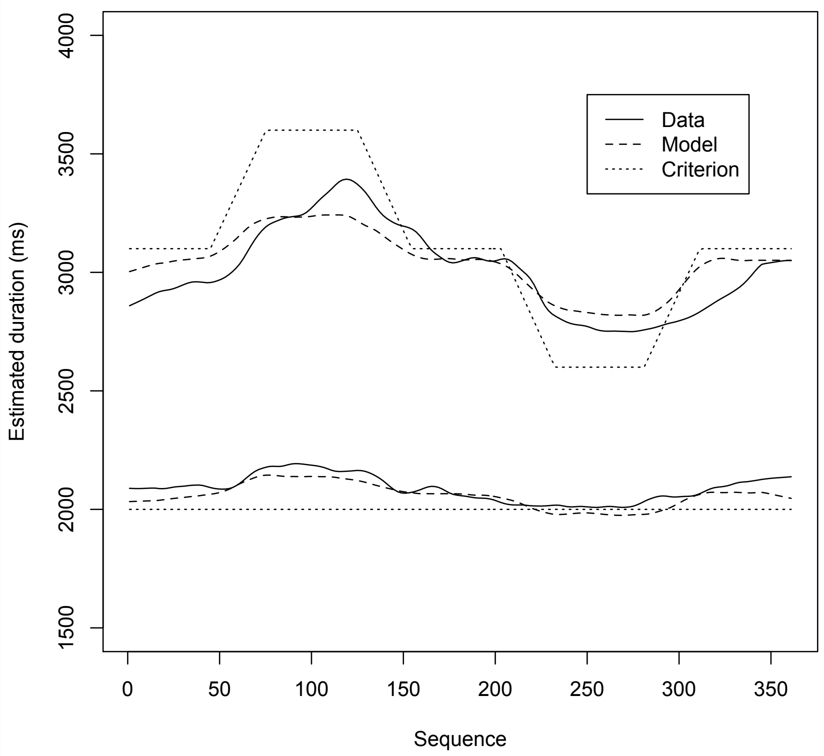
Table 3 Example of how the model calculates the number of pulses

Experience	Pulses	t (s)	A_j	P_j	P_jV_j	Feedback	t FB	A_j FB	P_j FB	P_jV_j FB	Grand Sum
long_{n-1}	28	4.5	–1.67	0.029	0.8	–2 (late)	0.7	–0.74	0.37	–0.75	
short_{n-1}	22	7.5	–1.01	0.797	17.5	0 (correct)	3.7	–0.65	0.58	0	
long_{n-2}	27	11.1	–2.12	0.003	0.08	+2 (early)	7.3	–1.91	0.001	0.002	
short_{n-2}	20	13.8	–1.31	0.171	3.43	+2 (early)	10	–1.15	0.05	0.10	
Sum					21.8					–0.65	21.15

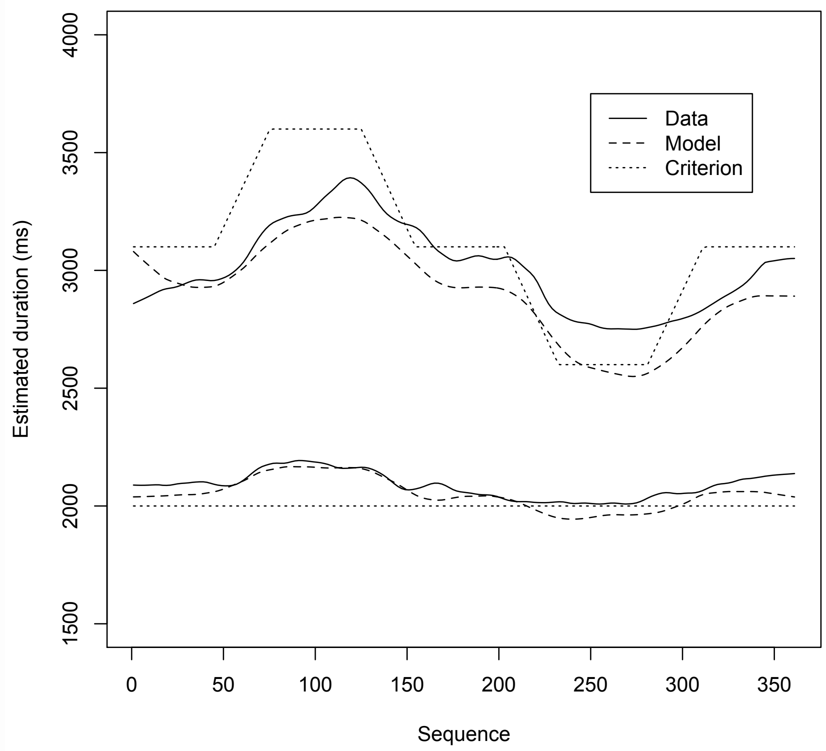
Model A and B are identical, except for parameters

<table>
<tr><th>Parameter</th><th>Model A</th><th>Model B</th></tr>
<tr><td>Noise parameter t</td><td>0.25</td><td>0.2</td></tr>
<tr><td>Mismatch penalty between short and long for interval retrieval</td><td>1.3</td><td rowspan="2">both 0.92</td></tr>
<tr><td>Mismatch penalty between short and long for feedback retrieval</td><td>0.8</td></tr>
<tr><td>Feedbackshift: how many pulses to add or subtract on the basis of feedback</td><td>8</td><td>1.8</td></tr>
</table>

Once more time the fit



Model A



Model B

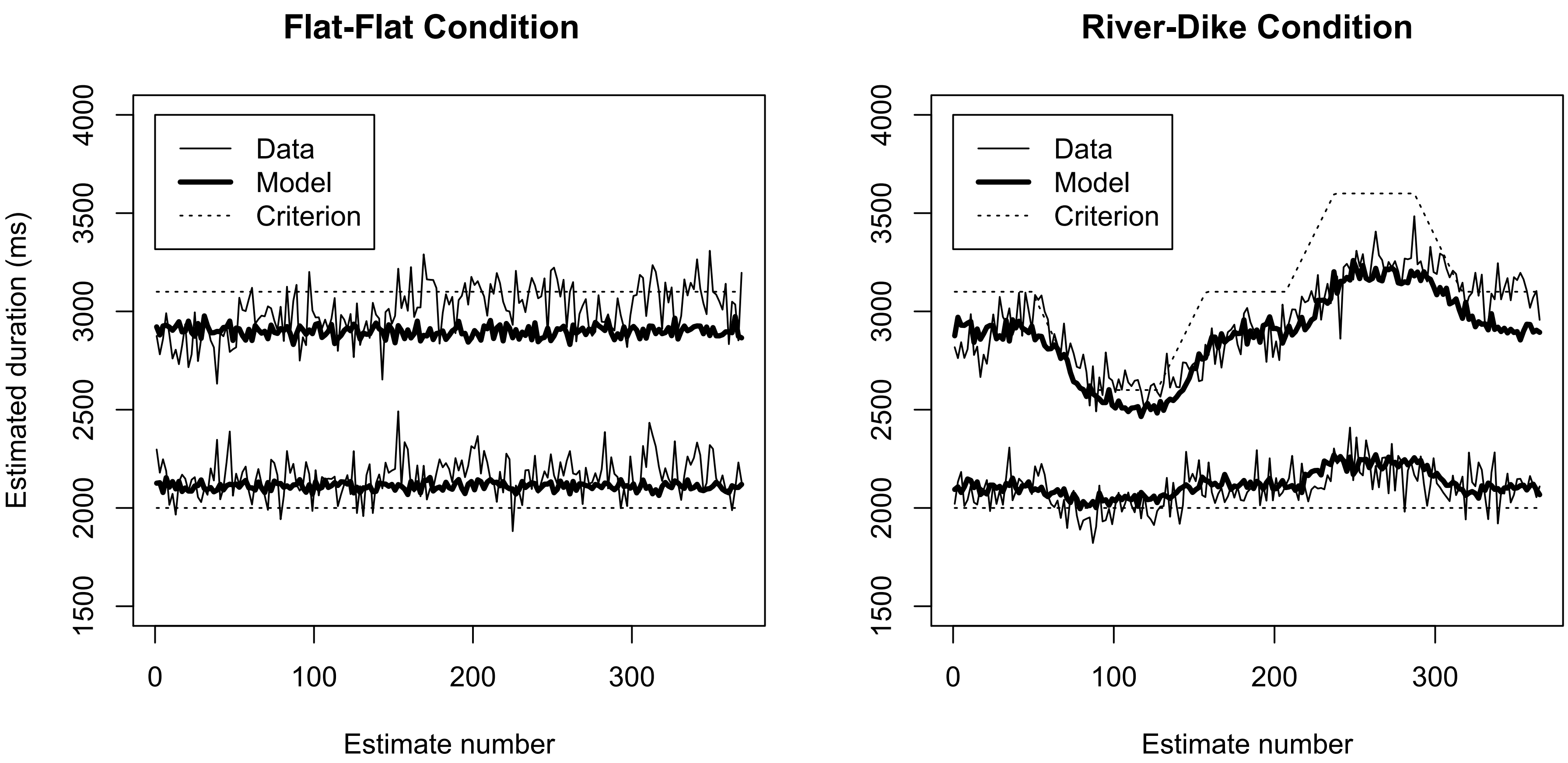
Which is better: run the model through the same regression as the data: Short interval

Fixed Effect	β data	β Model A	β Model B
Intercept	657 ms	2157 ms	789 ms
short_{n-1}	0.385	0.08	0.356
short_{n-2}	0.085	-0.03	0.048
short-fb-S_{n-1}	110 ms	487 ms	170 ms
short-fb-L_{n-1}	-208 ms	-521 ms	-153 ms
long_{n-1}	0.16	-0.06	0.15
long-fb-S_{n-1}	92.6 ms	432 ms	125 ms
long-fb-L_{n-1}	-163 ms	-534 ms	-211 ms

And the long interval…

Fixed Effect	β data	β model A	β model B
Intercept	695 ms	3162 ms	493 ms
$long_{n-1}$	0.34	0.011	0.22
$long_{n-2}$	0.16	0.012	0.25
$long_{n-3}$	0.12	0.003	0.12
$long_{n-4}$	0.05	0.001	0.09
long-fb-S_{n-1}	159 ms	626 ms	198 ms
long-fb-L_{n-1}	-118 ms	-744 ms	-251 ms
long-fb-S_{n-2}	82.9 ms	60 ms	90 ms
long-fb-L_{n-2}	3.8 ms	-142 ms	-57 ms
$short_{n-1}$	0.15	-0.07	0.18
short-fb-S_{n-1}	85 ms	326 ms	20 ms
short-fb-L_{n-1}	-107 ms	-492 ms	-35 ms

Model predictions



Questions

Regarding the rivers and dikes paper: If time to pulses is logarithmic, but the blending of pulses happens linear, are the changes, the tries a subject needs to find the correct long interval also much further from the target than at the short interval? And can you see the same (larger) slope towards longer durations in these tries? And if trying to compensate for a "too long" has a different effect (because of the larger ticks) than for a "too short", isn't it very hard for subjects to estimate the target without falling into the "too long" region again?

- You have shown that the cognitive model results in a better fit, and these are (only for the River-Dike condition?!) compared to alternative models. To help assess the validity of such a model, what is the difference in complexity between these models?

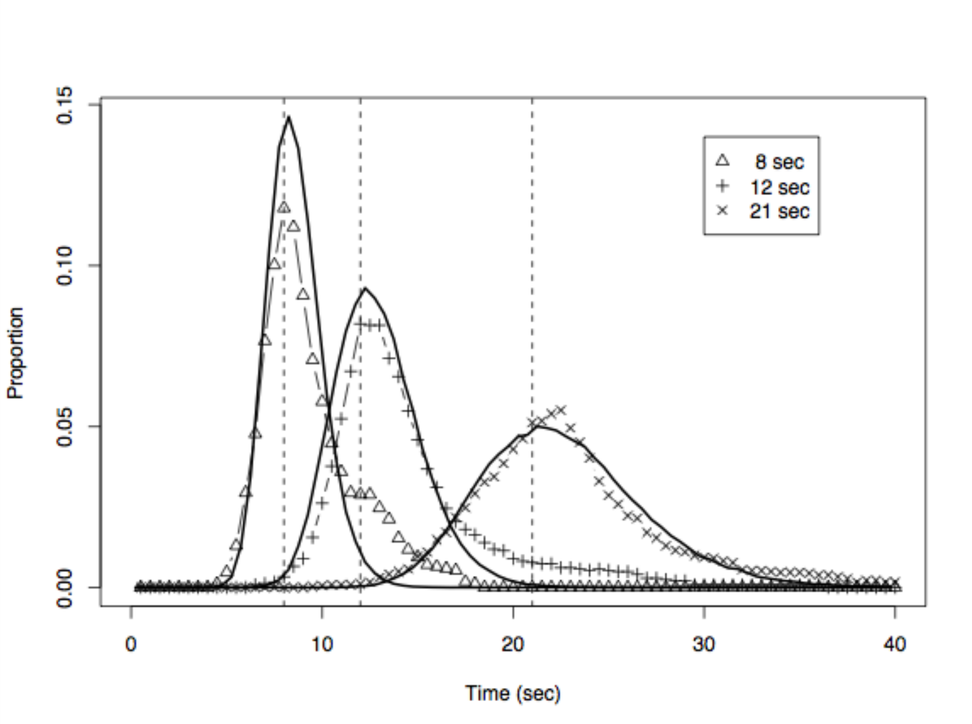
If you would change the +- 12.5% margin of the feedback on the subject's estimate, do you think the results (in particular the RD and DR manifestations) would change, and if so, in what manner? Would the proposed model show similar changes?

- In the experiment conducted by the authors, two circles and two colors are used to represent each interval. This allows for spatial location to potentially be added to the memory trace for the presented interval.
- That is, the interval could be encoded as a property of the circle, (as would be suggested by a binding pool model (Bowman and Wyble, 2007)). If that is true, than we would expect a greater influence if the intervals are distinguished only by color or only by circle.

Assignment this week

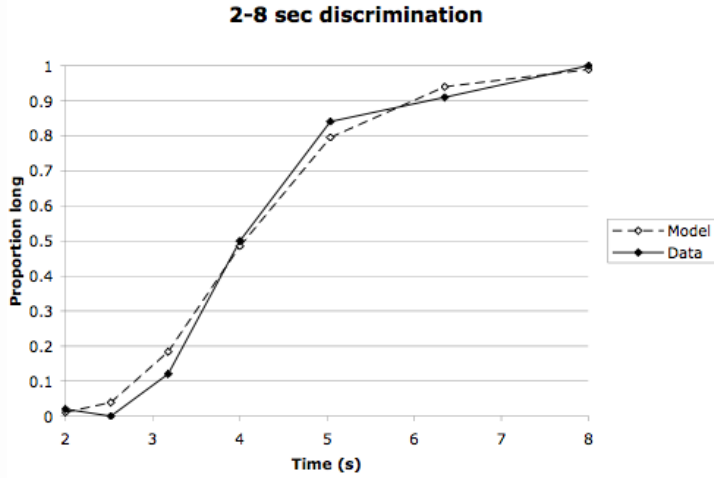
- Bisection experiments

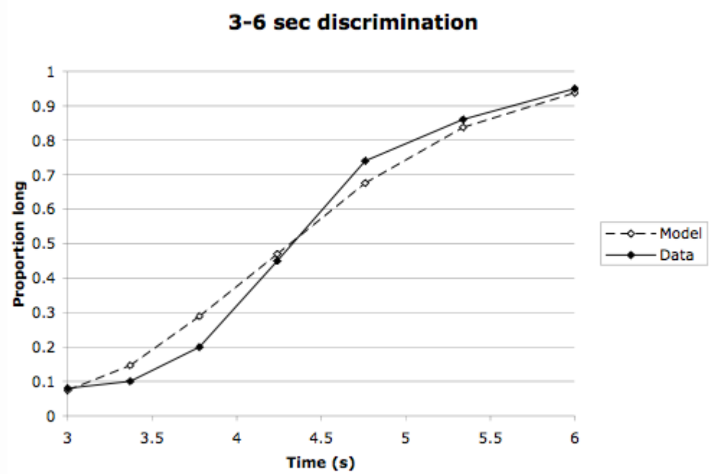
The module can fit basic time interval estimation properties



Scalar property

Data are from Rakitin et al. (1998)





Bisection

Data are from Penney, Gibbon & Meck (2000)

Build a proper experiment

- Your model should mimic running an actual experiment
- Outer loop for number of subjects
- Inner loop in which learn the short and long interval, and then proceed going through the the stimuli
- At the end, average the results an plot the graph