Workshop on Visual Attention and Visual Short-Term Memory

Attention Cascade Model: Evolution and Application

Dr. Shui-I Shih
University of Southampton

George Box:

Essentially, all models are wrong, but some are useful. (p. 424)

…… the practical question is how wrong do they have to be to not be useful. (p. 74)

Box, G. E. P. & Norman R. D. (1987). *Empirical Model-Building and Response Surfaces*, Wiley.

Attention and Memory

- Memory in an attention model
  - Limited capacity storage
  - Decison algorithm

- Attention in memory research
  - manipulation

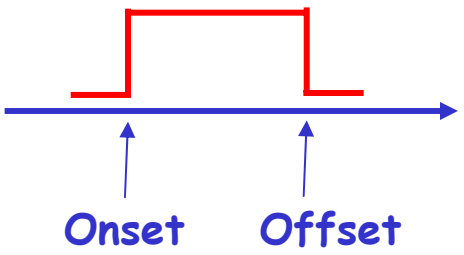
Attention Gating Models

- Reeves, A., & Sperling, G. (1986). Attention gating in short-term visual memory. *Psychological Review, 93 (2),* 180-206.
- Sperling, G., & Weichselgartner, E. (1995). Episodic theory of the dynamics of spatial attention. *Psychological Review,* 102, 503-532.
- Shih, S., & Sperling, G. (2002). Measuring and modeling the trajectory of visual spatial attention. *Psychological Review, 100 (2),* 260-305.

- Elaborate attention mechanism
- Account for a variety of experimental paradigms (e.g., whole report, partial report, 3x3 RSVP partial report, spatial cuing – simple or choice RT, and discrimination, etc.)
  - But ***not*** the attentional blink – limited working memory capacity

Theoretical Components

- Sensory/perceptual processor



- Long-term memory
- Attention control mechanism
- Working memory

Attention Gating Models

Cognitive accounts of the attentional blink

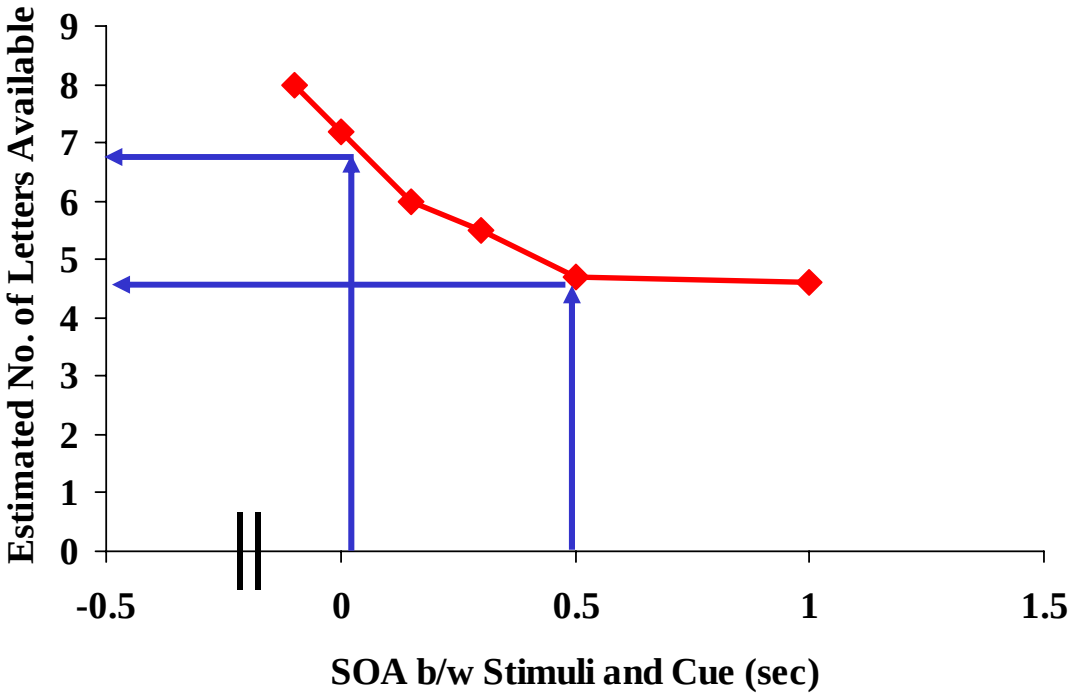
Outline of the Presentation

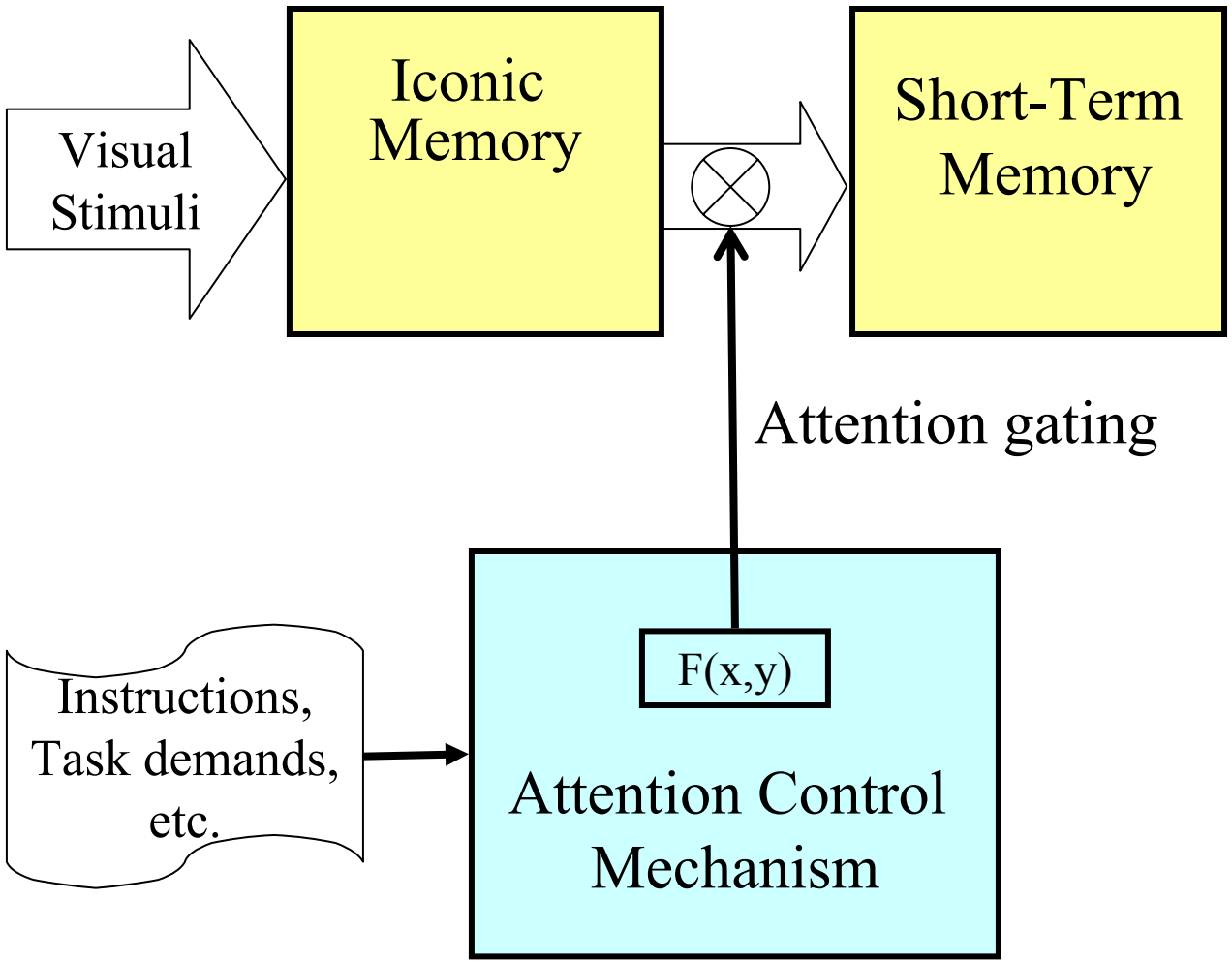
- Attention gating models
  - Evolution
  - Improvement
  - Working memory
- Cognitive accounts of the attentional blink
- Attention cascade model
  - The model
  - Some examples – computation, performance
- Cognitive aging

Sperling (1960)

D	Q	L	U
J	N	R	Y
S	B	P	Z

Partial-report




Visual
Stimuli
Iconic
Memory
Short-Term
Memory
Attention gating
F(x,y)
Instructions,
Task demands,
etc.
Attention Control
Mechanism

Reeves and Sperling (1986)

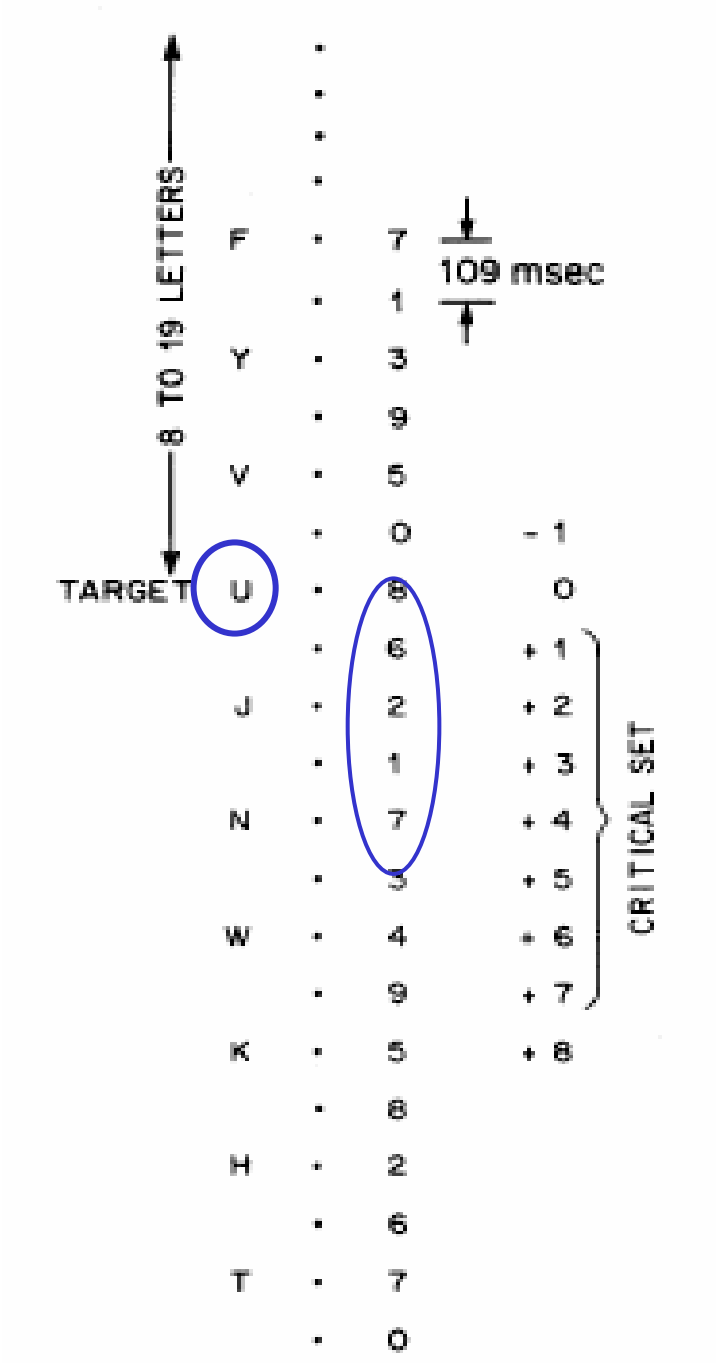
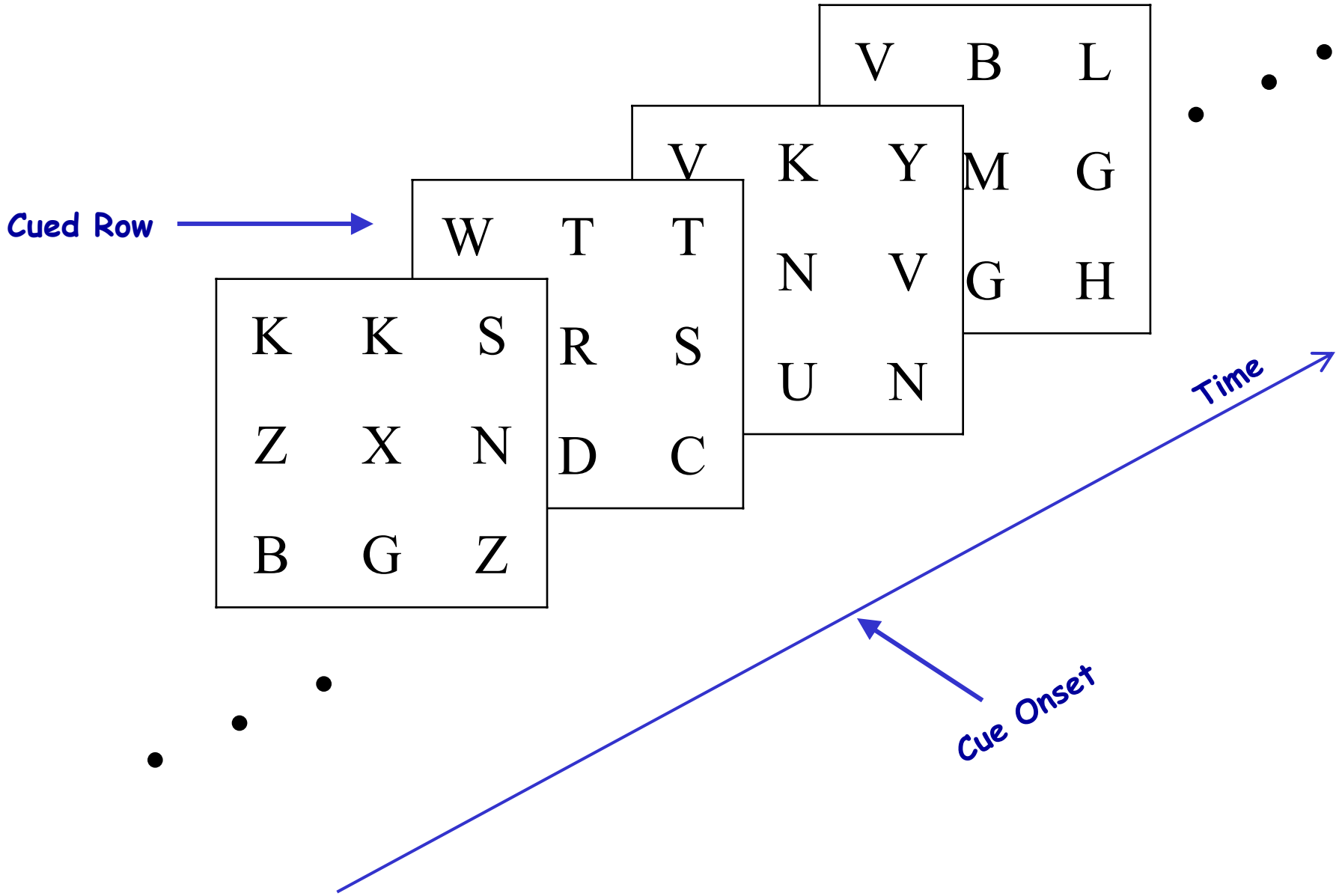
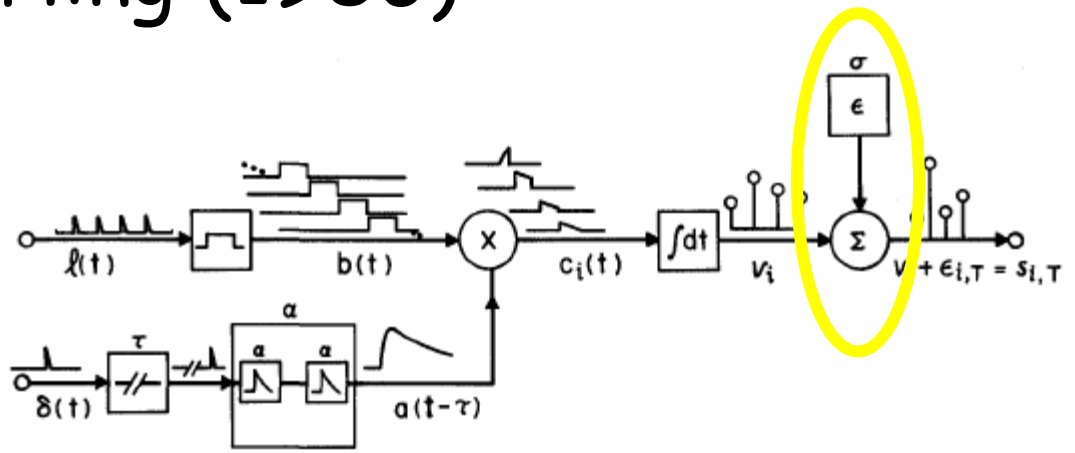


Figure 1, Reeves and Sperling (1986).

Shih and Sperling (2002)

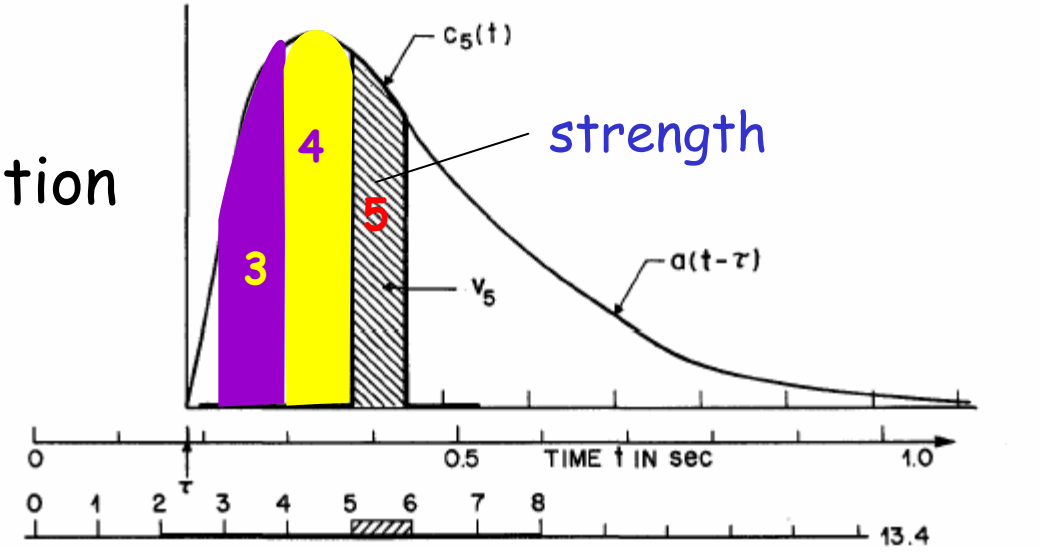


Reeves and Sperling (1986)



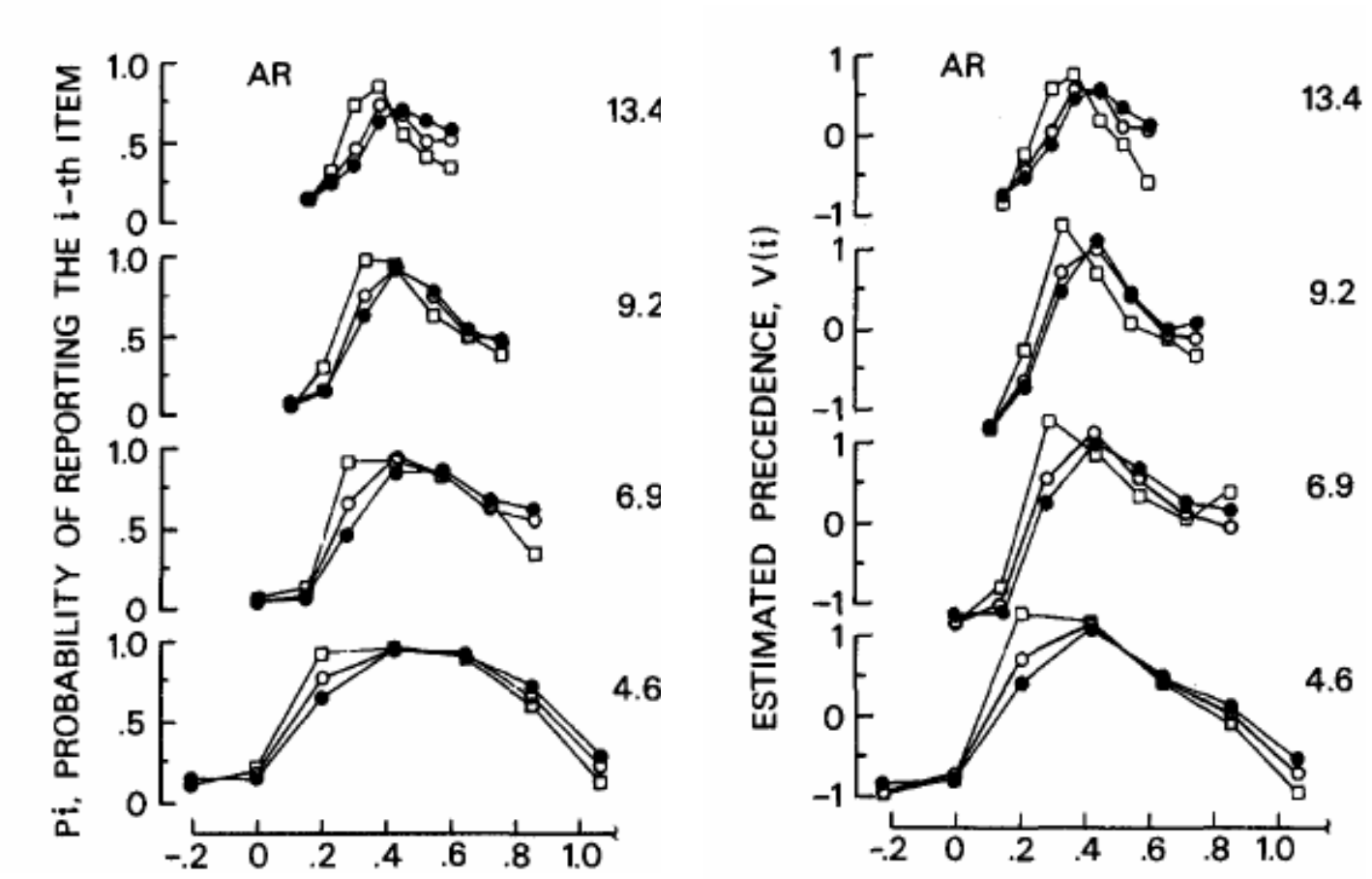
Attention gating function

(>= 2)



Data

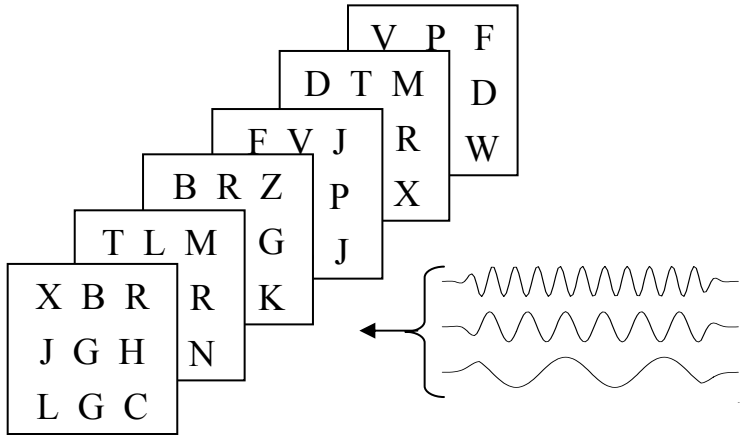
Model



Time after Target of the i-th Item is Seconds

Episodic theory of Attention

- Sperling and Weichselgartner (1995)
  - Spatial cuing paradigm (go/no go RT, choice RT, and discrimination, etc.), attention RT
- Shih and Sperling (2002)



Episode: Pre-Cued Cued Post-Cued

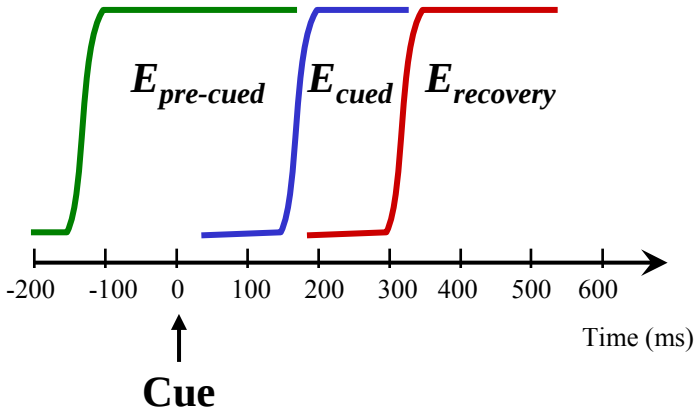
Spatial Attention, *F(x, y)*

0	0	0
0	0	0
0	0	0

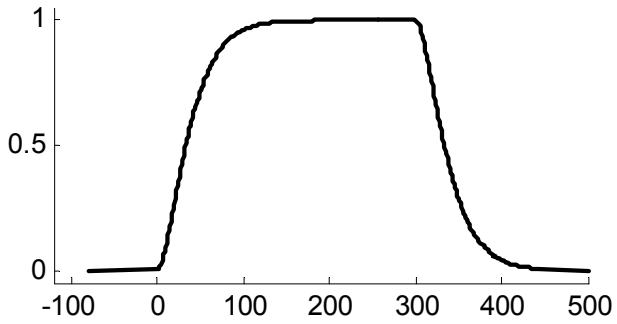
1	1	1
0	0	0
0	0	0

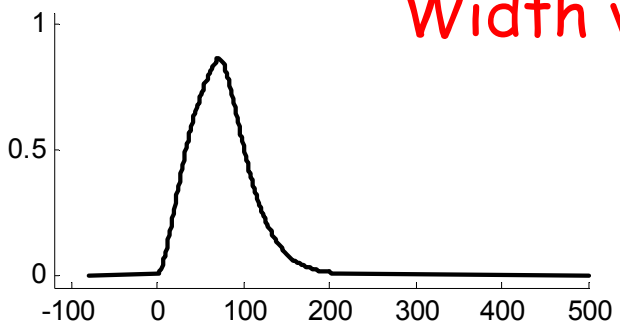
0	0	0
0	0	0
0	0	0

Temporal Transition, G(t):

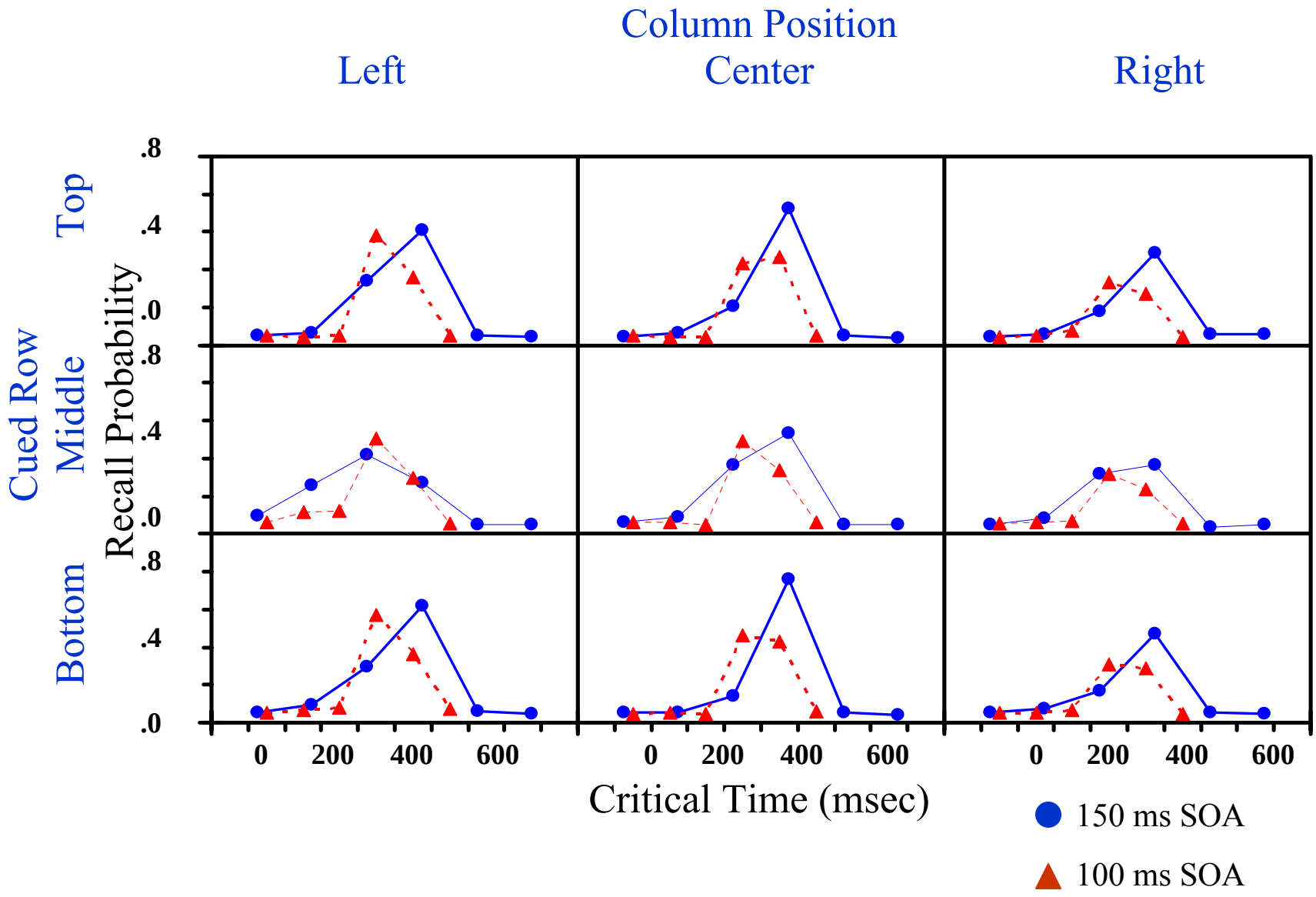


$$A_{cue}(r, t) = F_{cue}(r) * [G(t-t_{cue}) - G(t-t_{post})]$$





Width vs. Onset

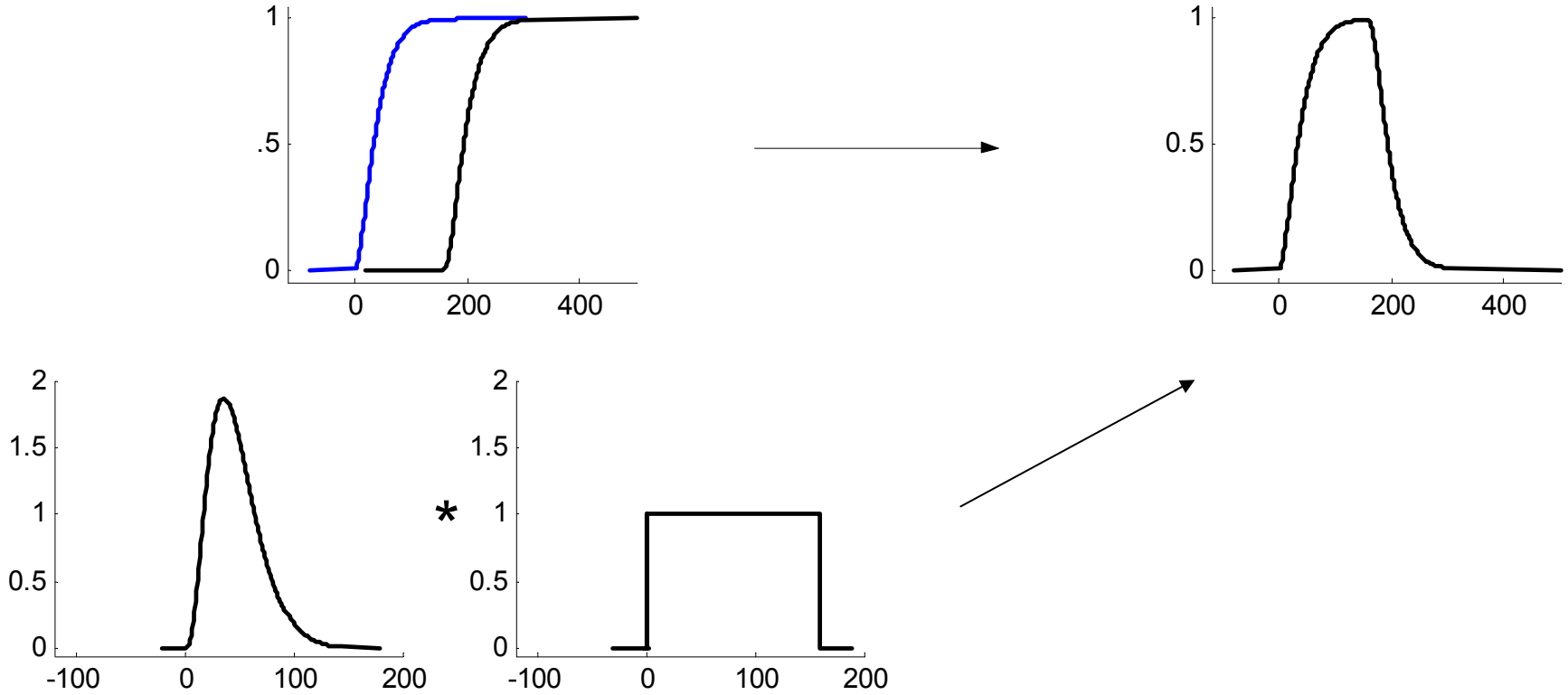

Column Position
Left
Center
Right
Cued Row
Top
Middle
Bottom
Recall Probability
.8
.4
.0
0
200
400
600
Critical Time (msec)
150 ms SOA
100 ms SOA

- Visual attention ≡ a sequence of episode
- Episode ≡ spatial distribution of attention
- Transition is not instantaneous

- Six primary parameters
- 1000+ data points per observer

Issues

- Cue interpretation vs. Attention transition
- Assumptions regarding the attention gating function



Onset of AW

Attention Window (AW)

-- Two stages

-- Processing time *iid* exponential pdf

Issues

- Cue interpretation vs. Attention transition
- Assumptions regarding the attention gating function
- Physical salience
  - automatic vs. controll attentional process (Weichselgartner & Sperling, 1987)
    - a constant number of stages, but varied rate
    - A constant rate, but varied number of stages

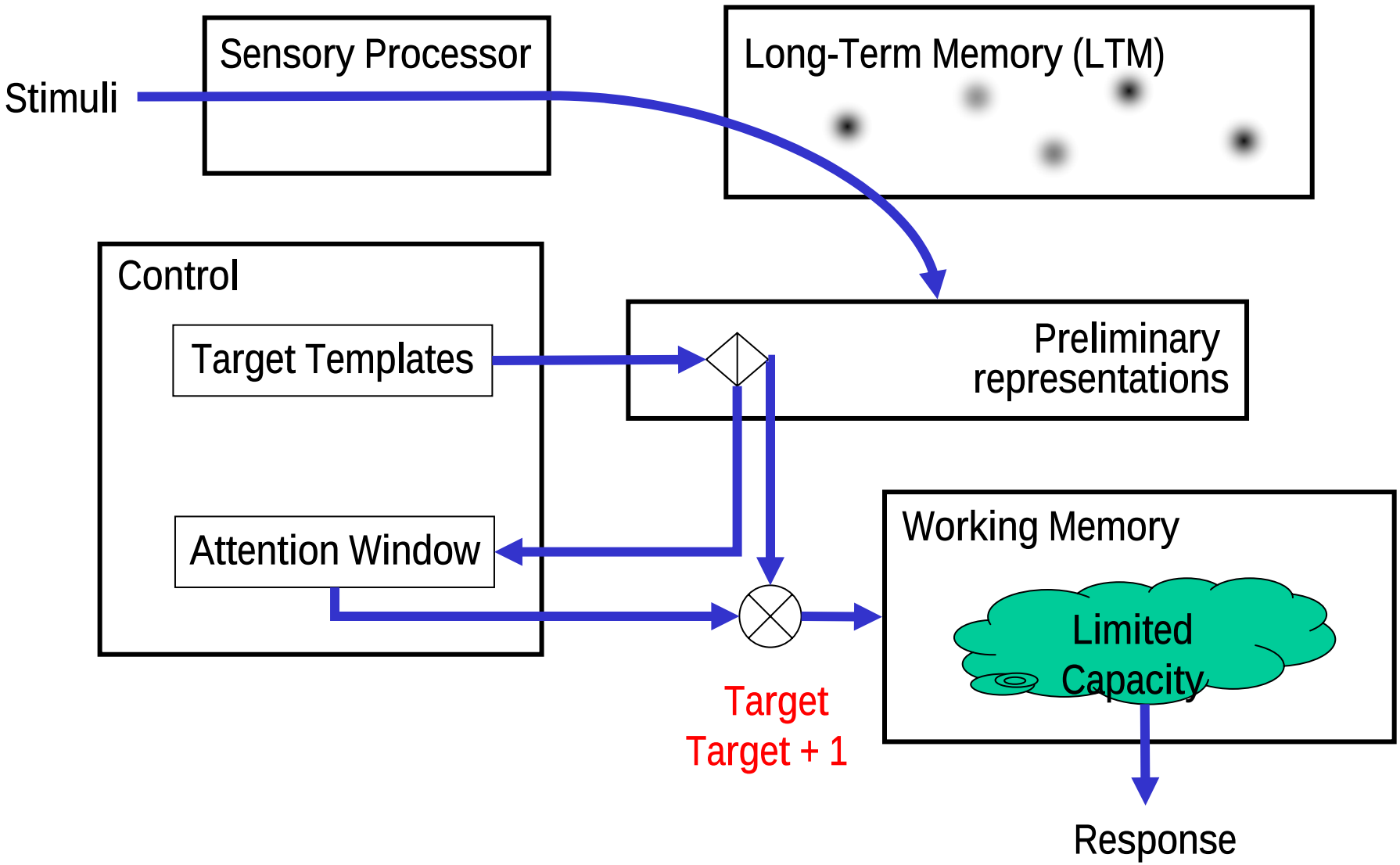
Plans

- Attentional blink experiments
  - Address the three issues
    - Attentional gating models
  - Account for the attentional blink
    - Working memory

Cognitive Accounts of the AB

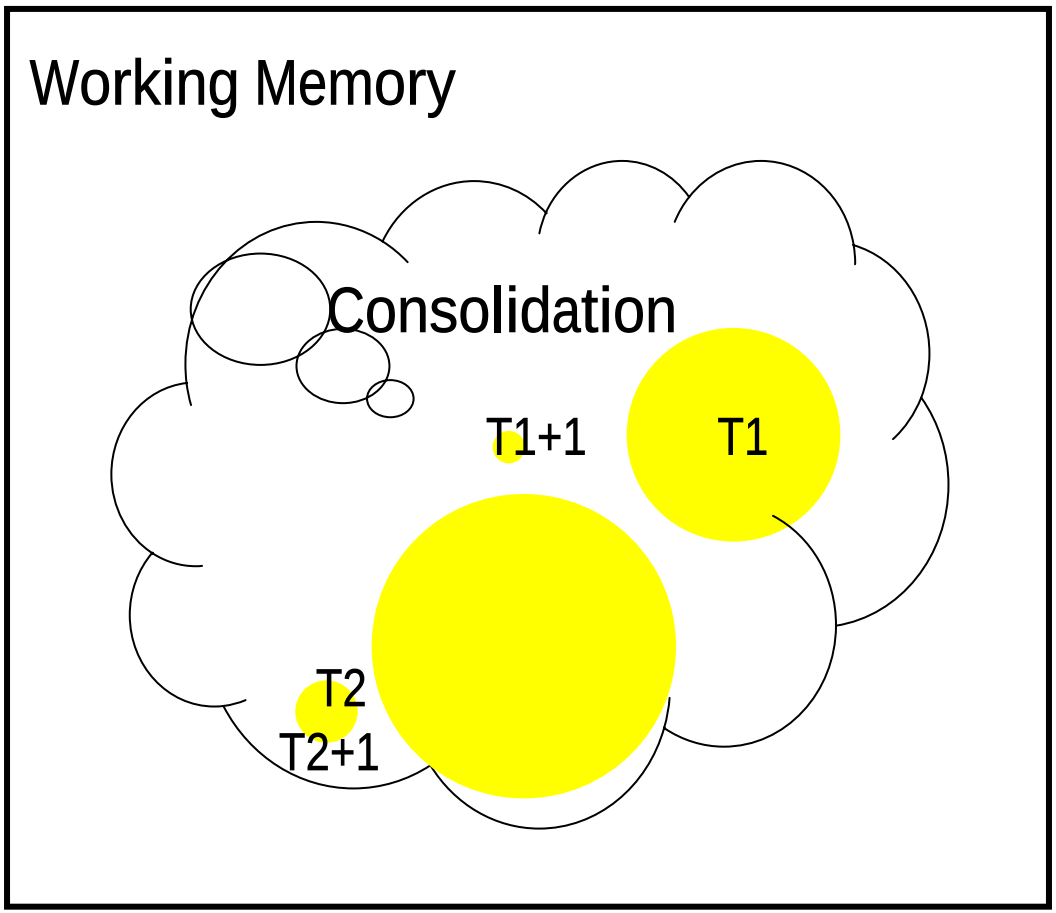
- Inhibition model (Raymond, Shapiro, & Arnell, 1992)
- Interference model (Shapiro, Raymond, & Arnell, 1994)
- Two-stage model (Chun & Potter, 1995)
- Two-stage competition model (Potter, Staub, & O'Connor, 2002)
- Central interference model (Jolicoeur, 1999)
- Hypothesis of attentional dwell time (Duncan, Ward, & Shapiro, 1994)
- Hypothesis of temporal loss of control (Di Lollo, Kawahara, Ghorashi, Enns, 2005)

Common Assumptions



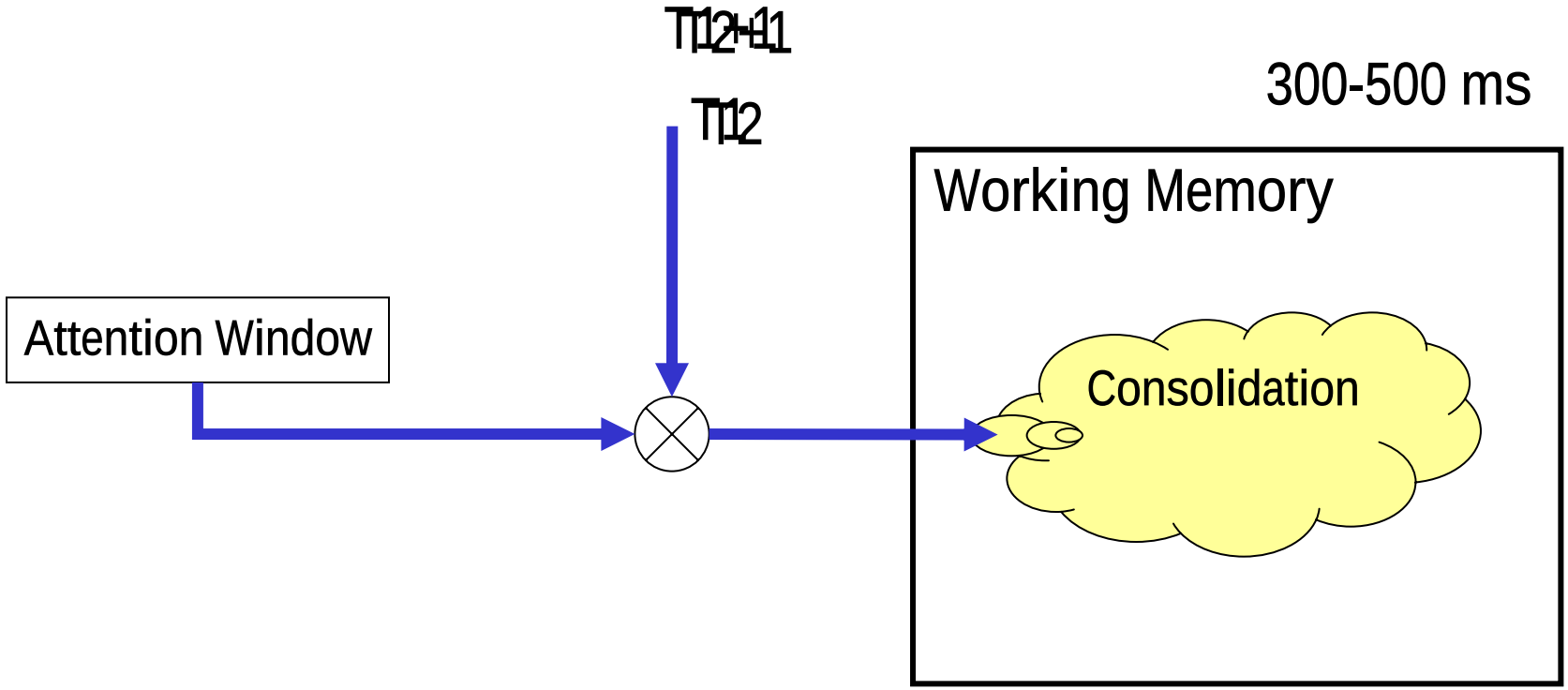
Interference Model

(Shapiro, Raymond, & Arnell, 1994)

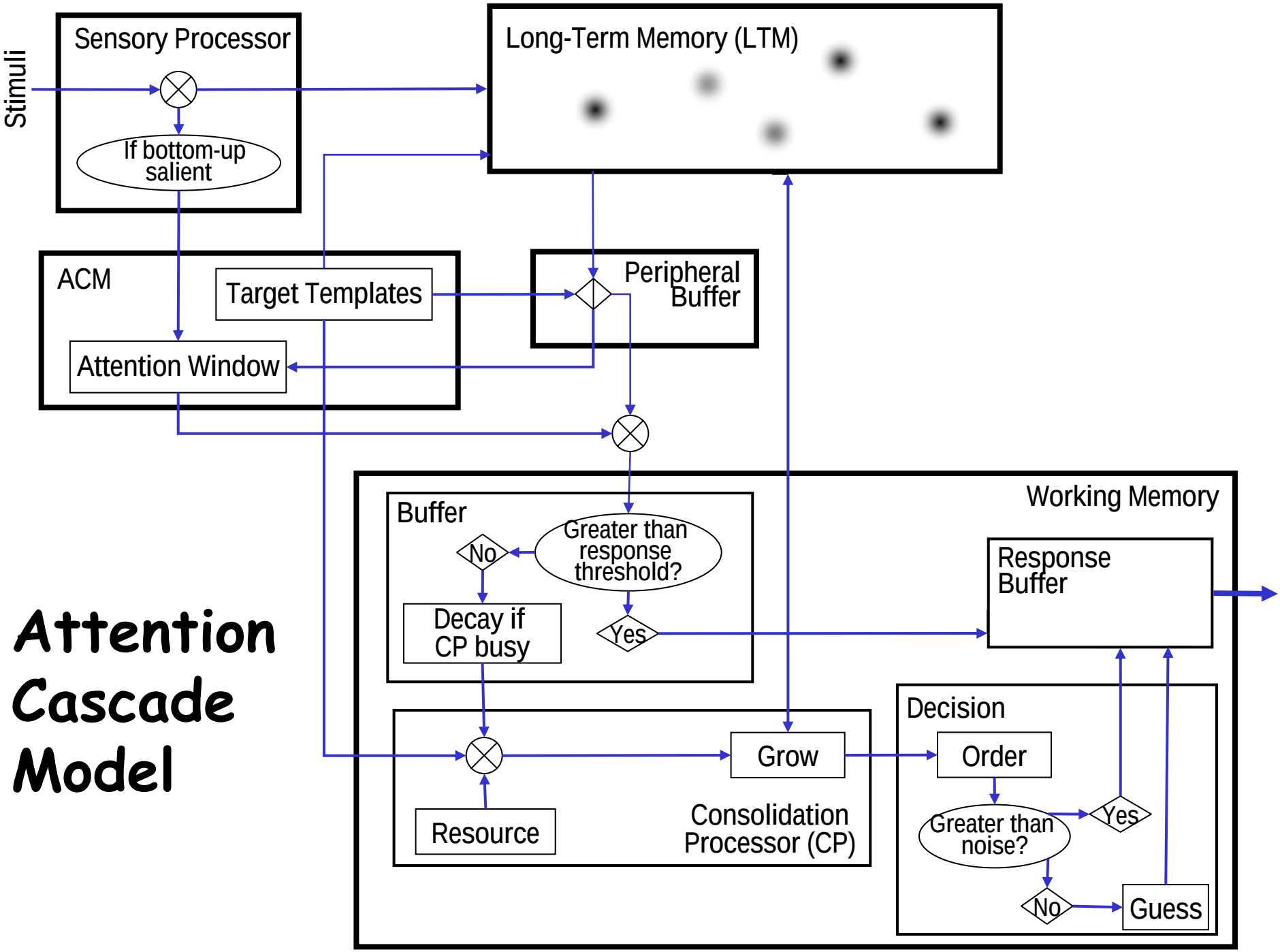


300-500 ms

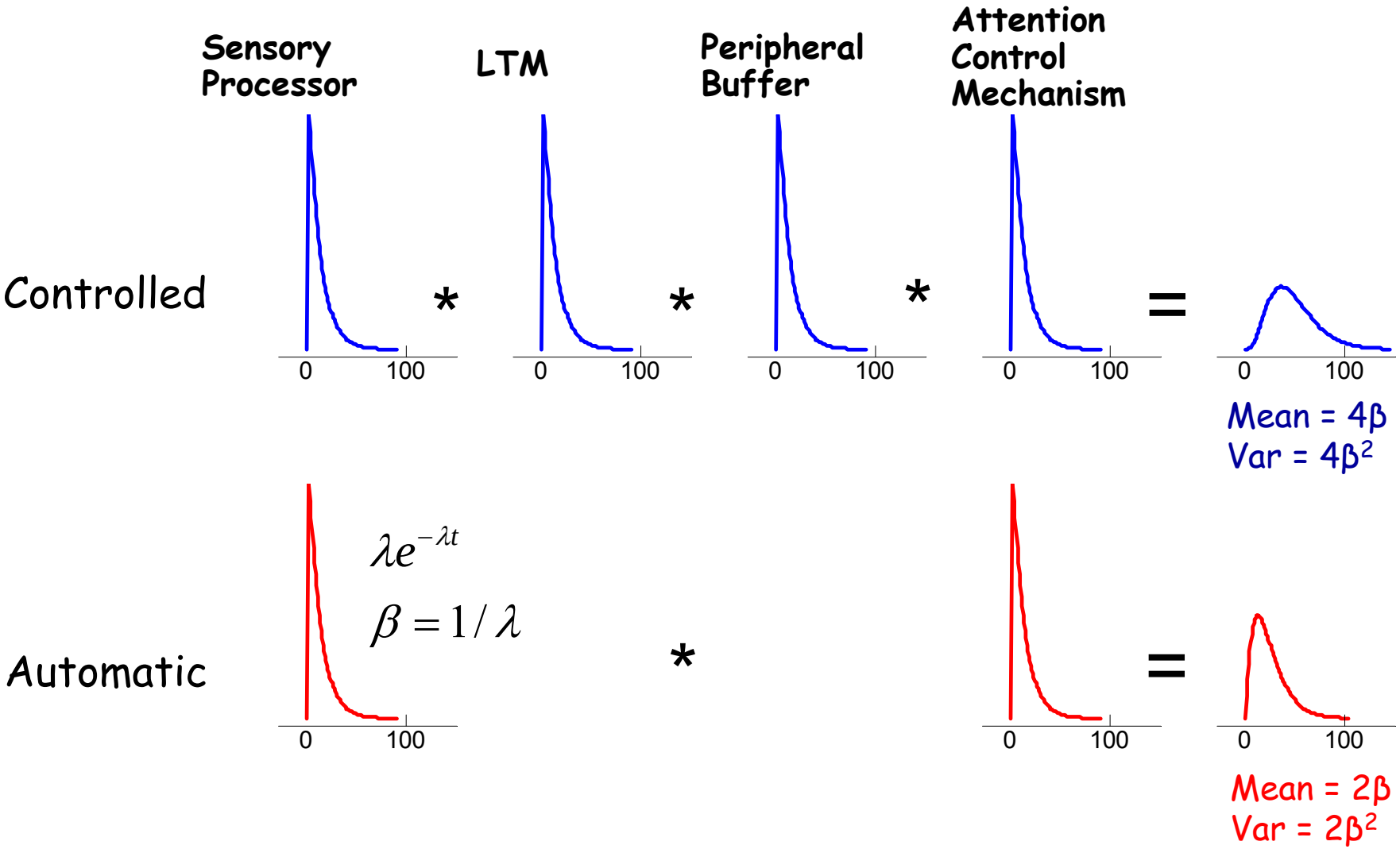
Two-Stage Model



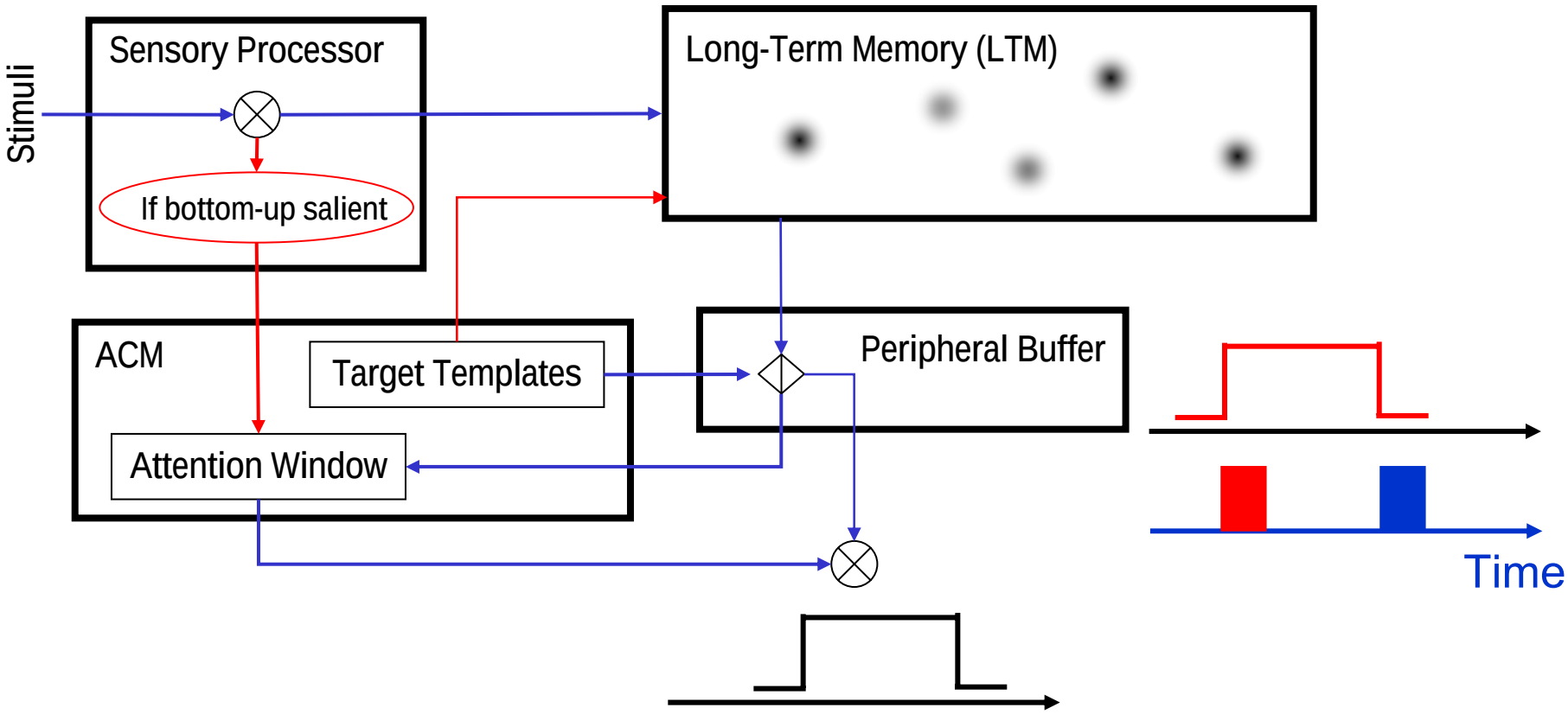
Attention Cascade Model

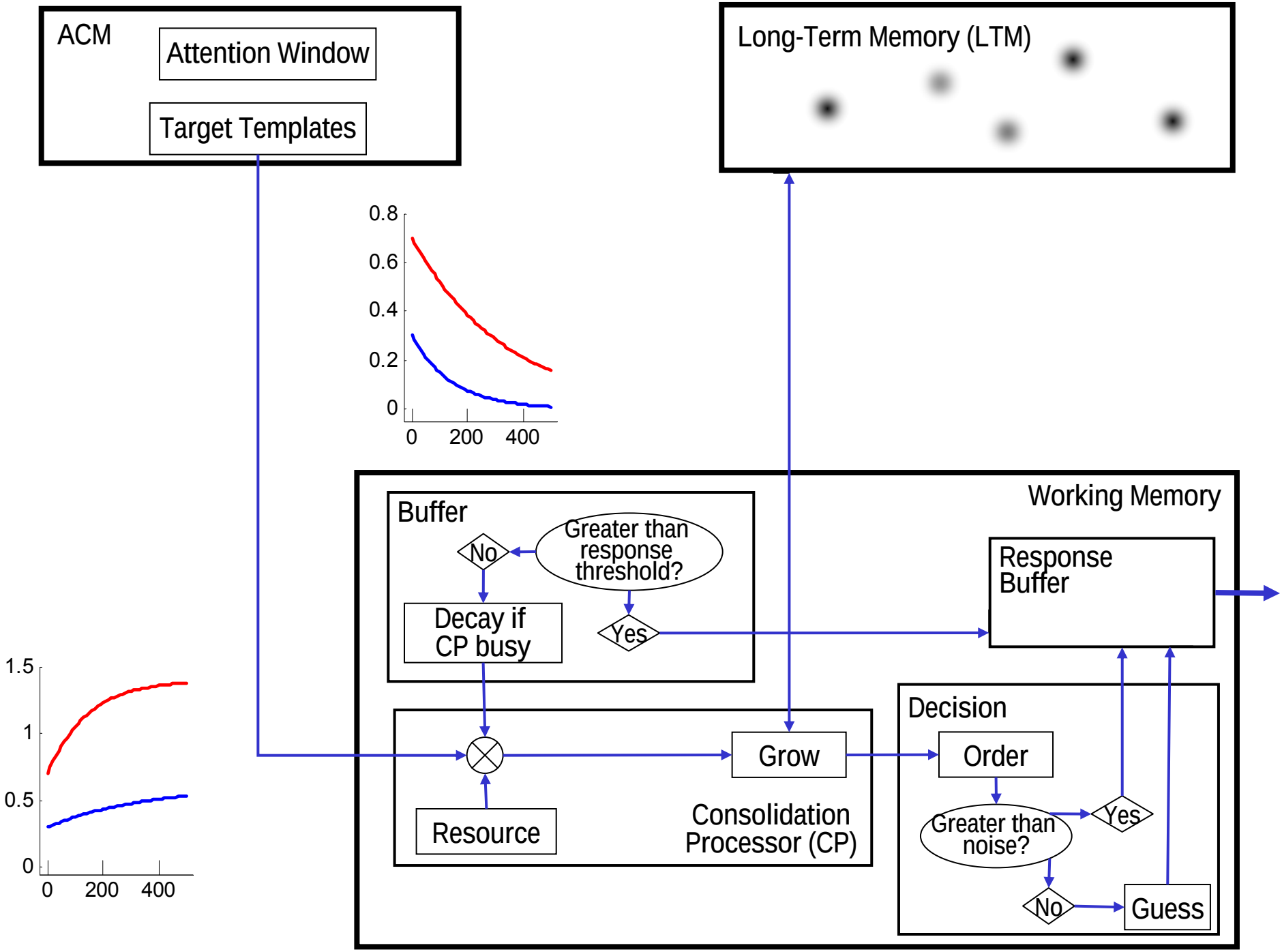


Triggering Time Distribution of the Attention Window



Attention Gating

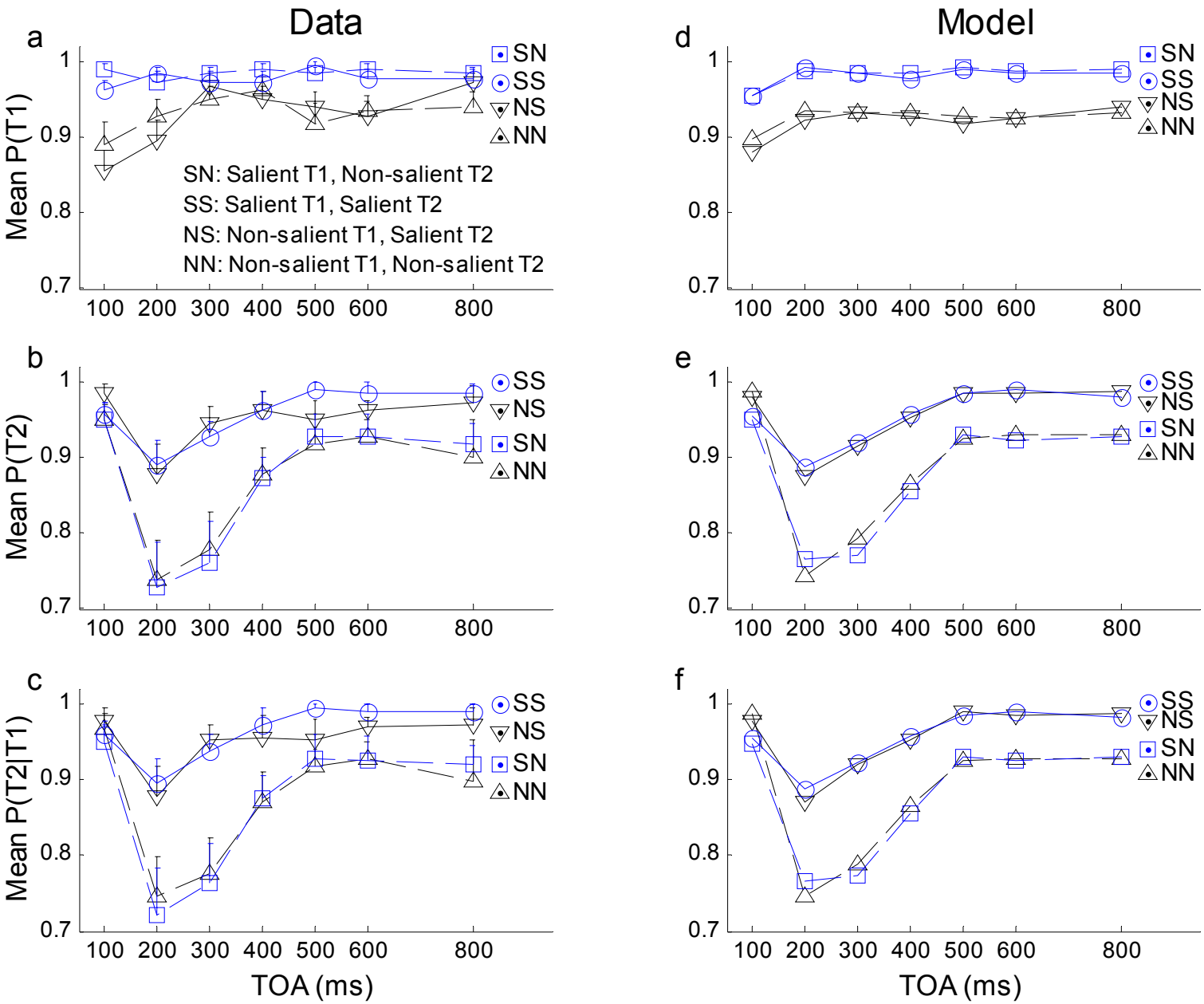



ACM
Attention Window
Target Templates
Long-Term Memory (LTM)
0.8
0.6
0.4
0.2
0
0
200
400
Working Memory
Buffer
Greater than response threshold?
No
Decay if CP busy
Yes
Response Buffer
Grow
Resource
Consolidation Processor (CP)
Decision
Order
Greater than noise?
Yes
No
Guess
1.5
1
0.5
0
0
200
400

Parameters

	Assumed	Estimated	Derived
Preliminary Rep (PR)			
Attention Window (AW)		*w*	
Pre-AW stage	2 or 4 stages	β	TTD of AW
Initial strength, s_i	PR-AW overlap		$0 <= s_i <= 1$
Consolidation duration		π	
Queuing time			Max(0, π – TOA)
Decay rate	Exponential pdf		$1 - s_i$
Strength after decay, q_i			$0 <= q_i <= 1$
Resource		*C*	
Assigned resource, r_i			$r_i = q_i \ \mathrm{Min}(1, C/\sum q_i)$
Growth rate	Exponential CDF		r_i
Mean of internal noise	Gaussian	μ	
SD of internal noise		σ	

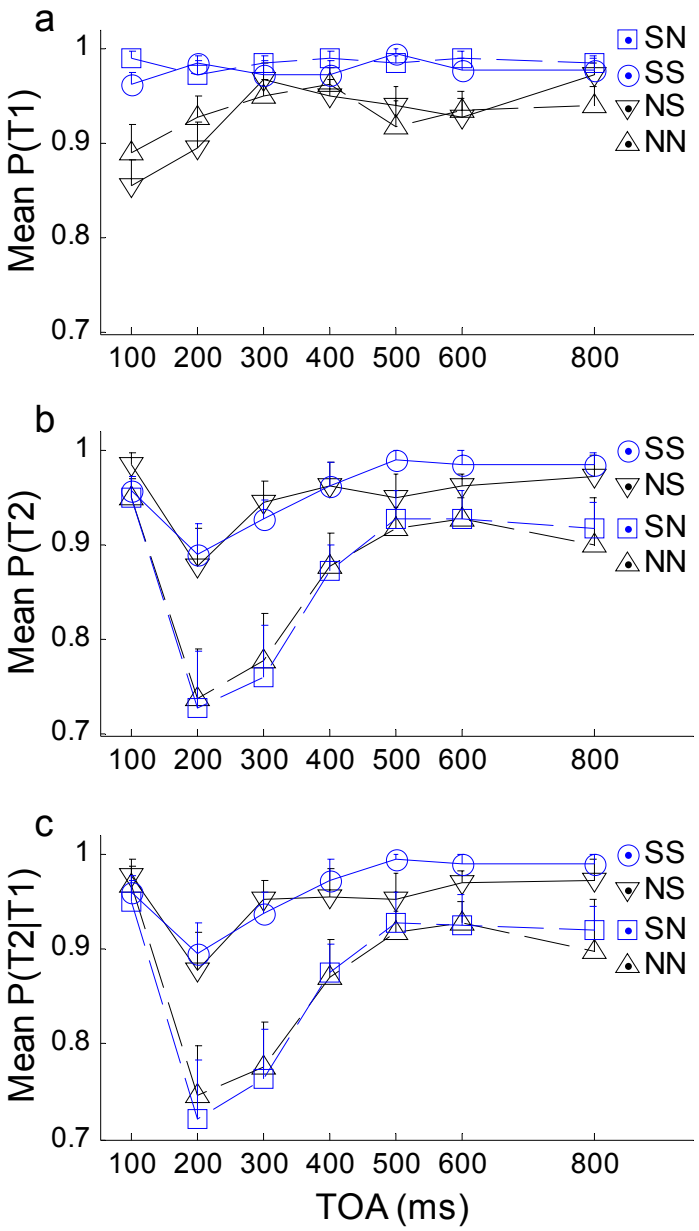
Shih and Reeves (2007)



Shih and Reeves (2007), Exp 1

- T1 Salience x T2 Salience x TOA
- Photometrically equiluminant red, green, and yellow characters
- *SOA* = 100 ms

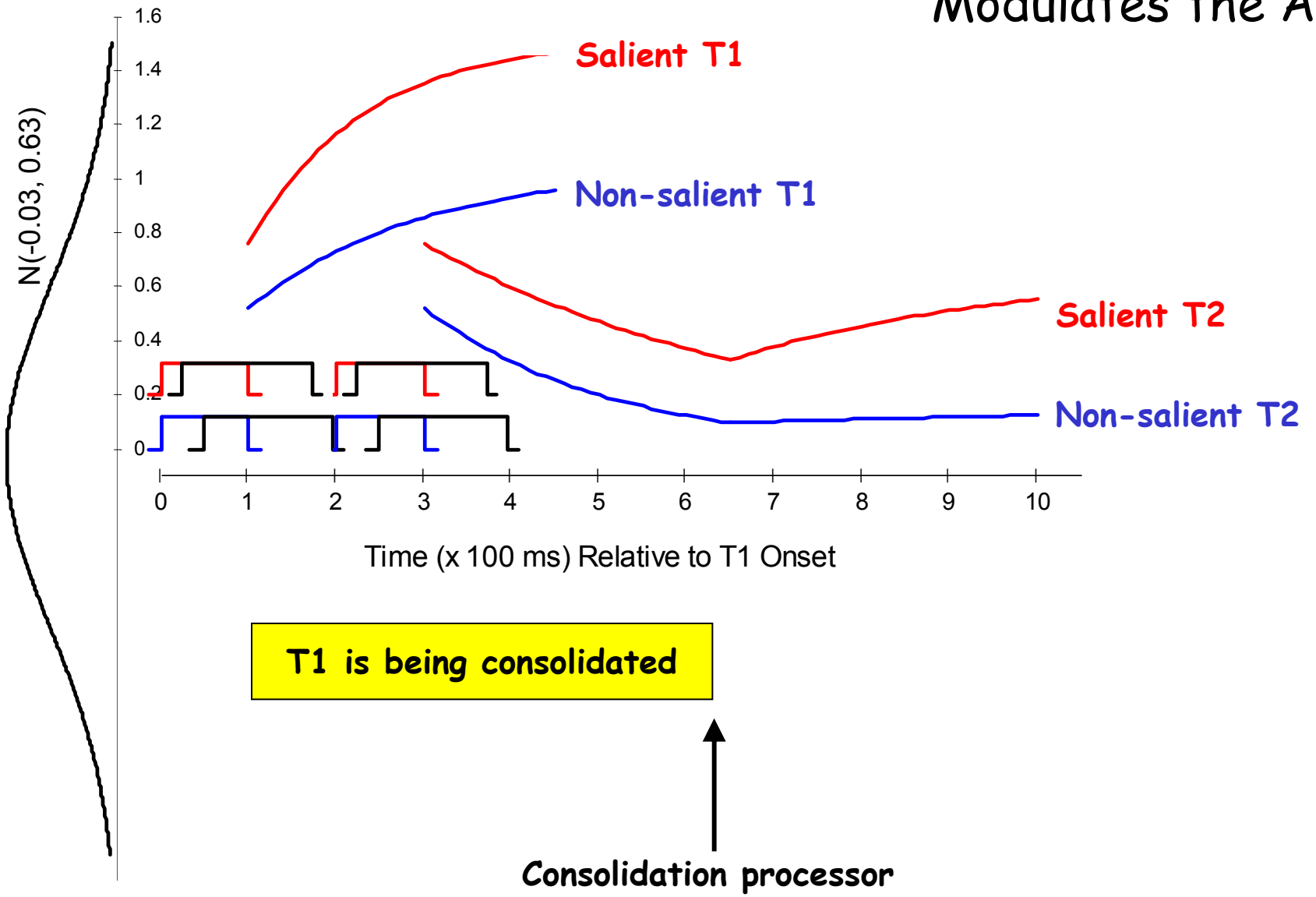
N-N-100	. . . U S 3 6 P T X . . .
N-S-200	. . . G X 5 L 8 C V . . .
S-N-300	. . . B D 7 R Y 2 K . . .
S-S-600	. . . H S 9 P X V C V 4 N . . .



	T1	T2
SS	**Salient**	**Salient**
SN	**Salient**	Non-salient
NS	Non-salient	**Salient**
NN	Non-salient	Non-salient

Effect of Stimulus Salience

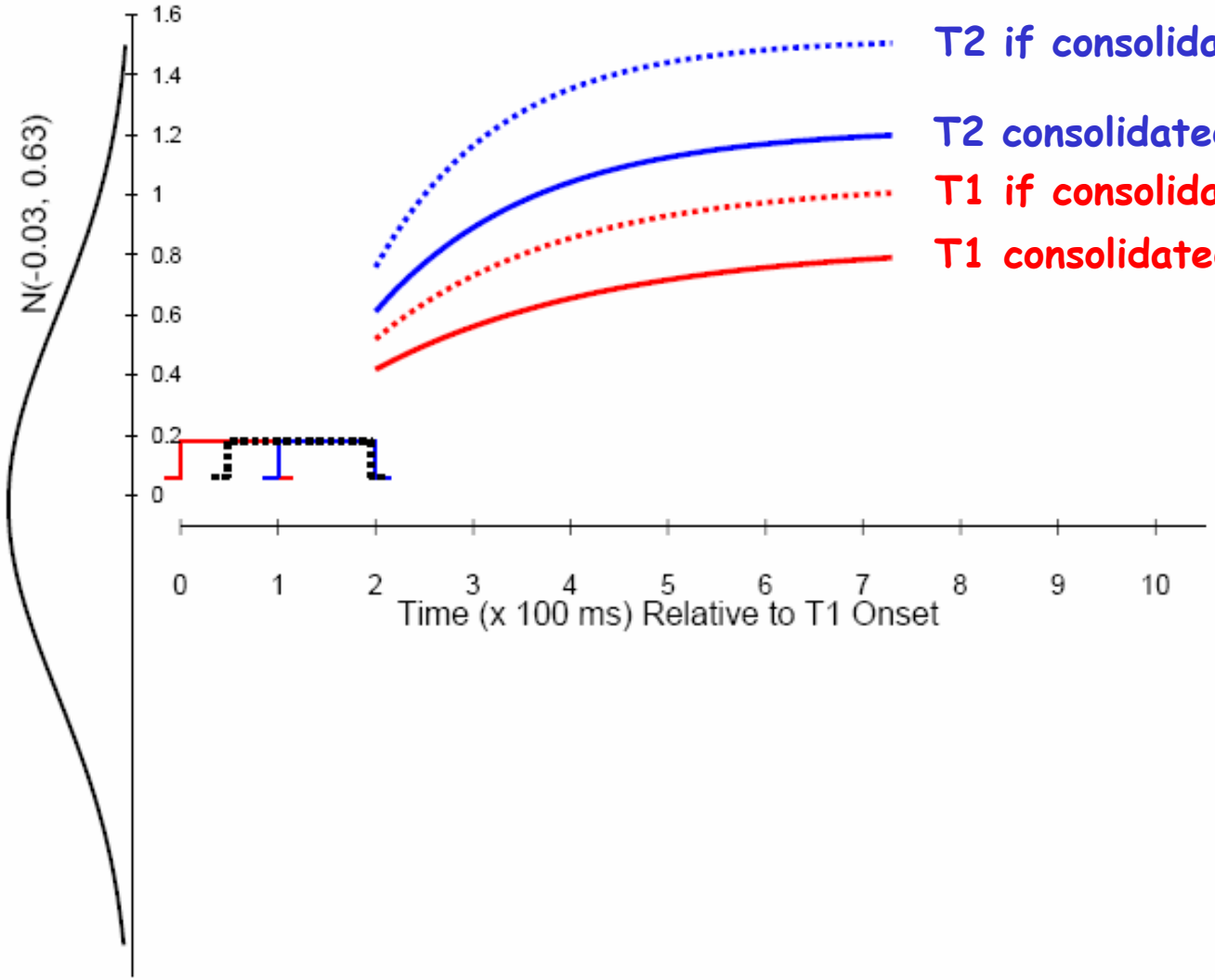
T2 Salience Modulates the AB



T1 is being consolidated

Consolidation processor becomes available

T2 Superiority or Lag 1 Sparing at a "Short" TOA



T2 if consolidated without competition

T2 consolidated with T1

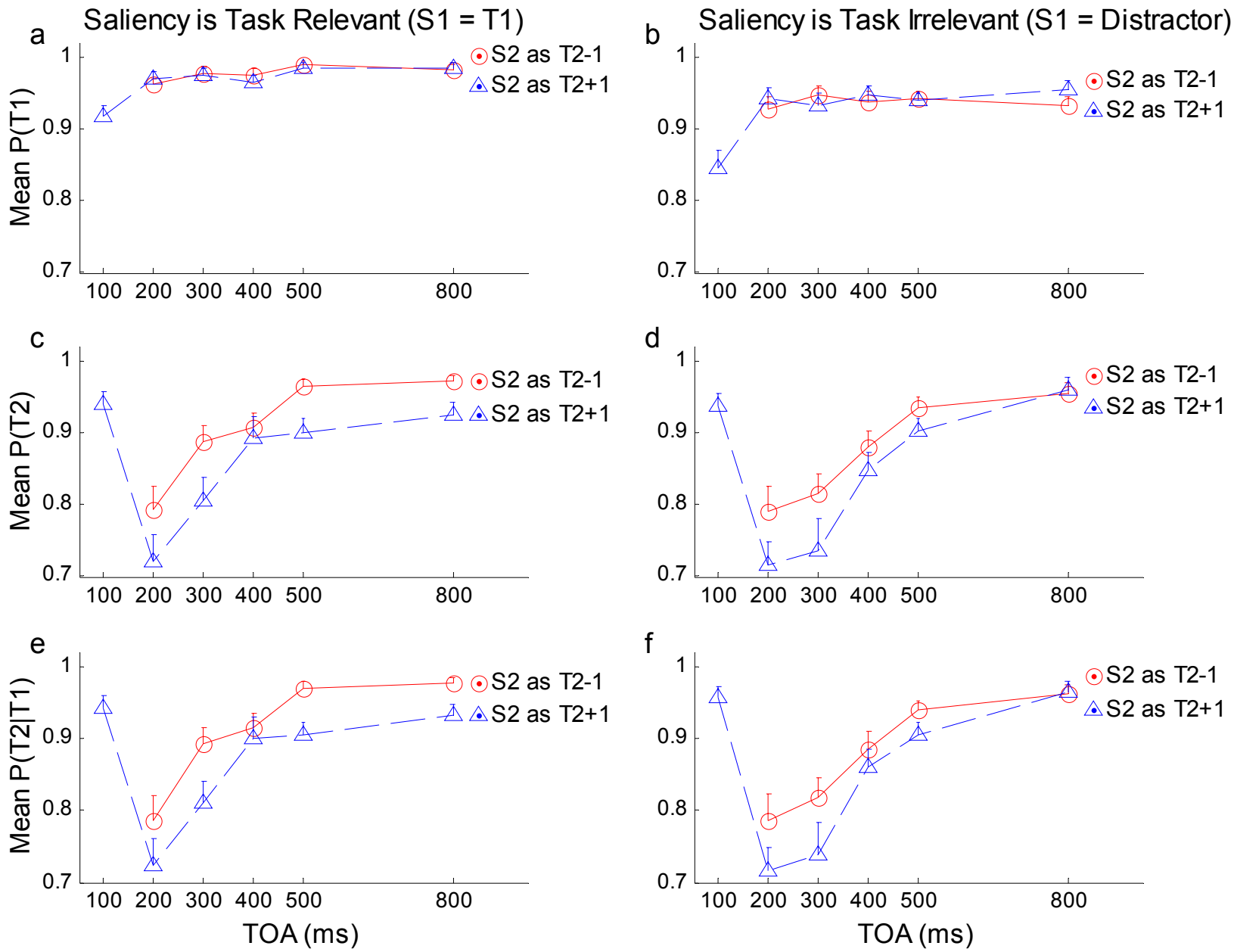
T1 if consolidated without competition

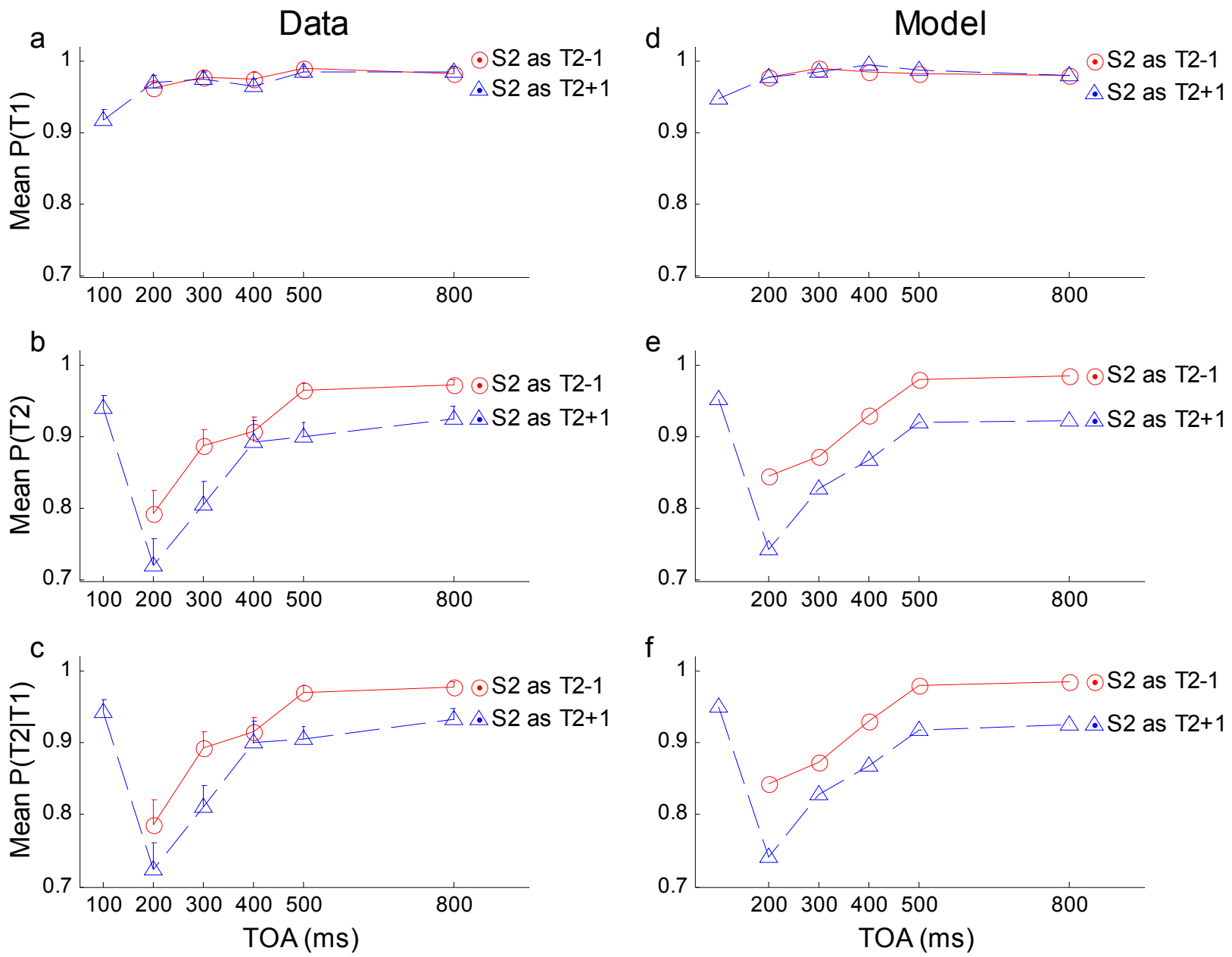
T1 consolidated with T2

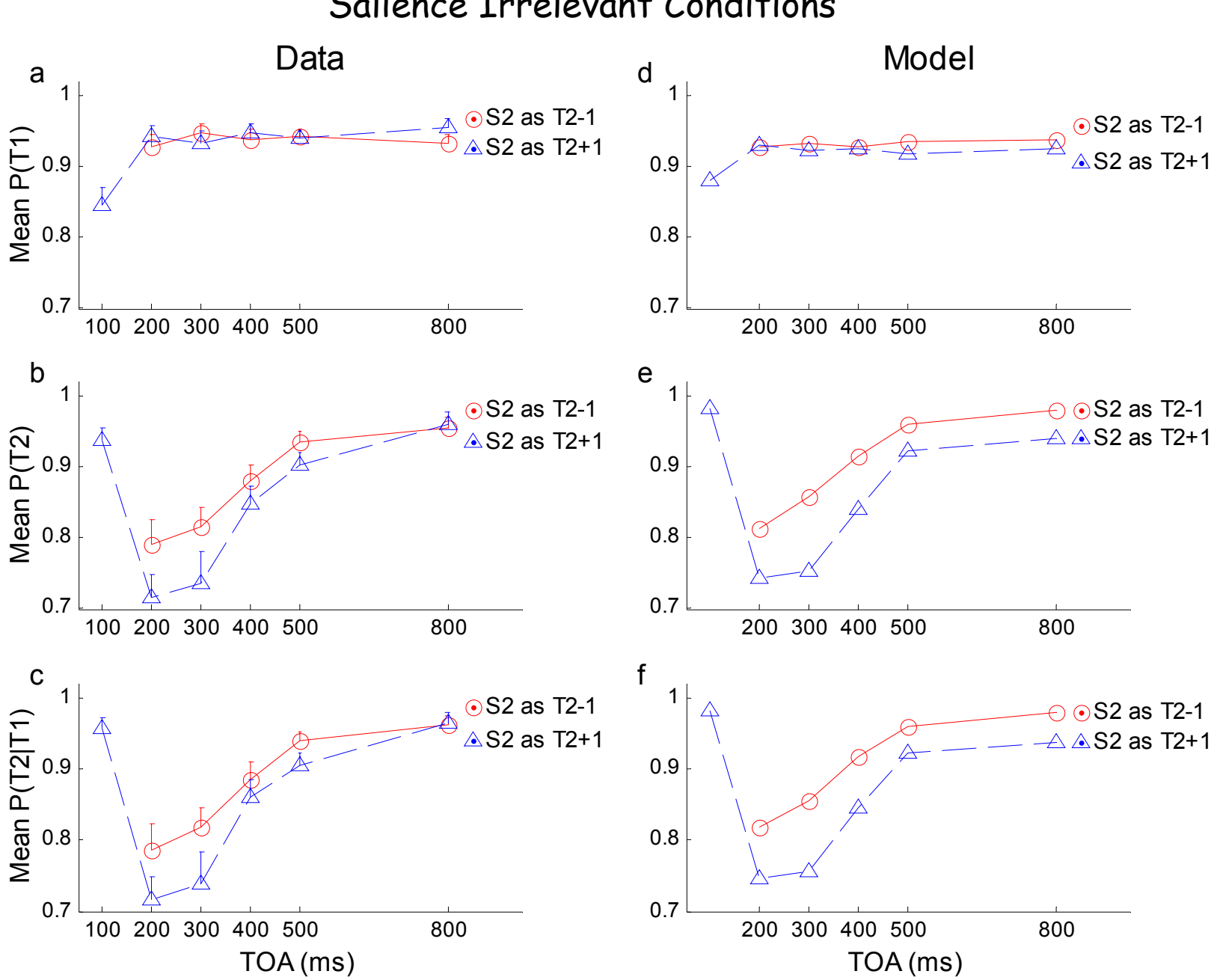
Shih and Reeves (2007), Exp 2

- Two salient items (S1 and S2)
- Non-salient T2
- Task relevance of saliency
  - T1 = S1 OR Non-salient T1 and S1 occurred several items before T1
- S2-T2 Lag
  - S2 as T2-1 OR T2+1

R-100-T2+1	**. . . U M 3 C T 6 P T X . . .**
I-500-T2-1	**. . . P S G X 5 L 8 C V . . .**


a
Saliency is Task Relevant (S1 = T1)
Mean P(T1)
b
Saliency is Task Irrelevant (S1 = Distractor)
c
Mean P(T2)
d
e
Mean P(T2|T1)
TOA (ms)
f
TOA (ms)
S2 as T2-1
S2 as T2+1
1
0.9
0.8
0.7
100
200
300
400
500
800


Salience Relevant Conditions
Data
Model
a
b
c
d
e
f
Mean P(T1)
Mean P(T2)
Mean P(T2|T1)
TOA (ms)
S2 as T2-1
S2 as T2+1
1
0.9
0.8
0.7
100
200
300
400
500
800


Salience Irrelevant Conditions
Data
Model
a
b
c
d
e
f
Mean P(T1)
Mean P(T2)
Mean P(T2|T1)
TOA (ms)
S2 as T2-1
S2 as T2+1
1
0.9
0.8
0.7
100
200
300
400
500
800

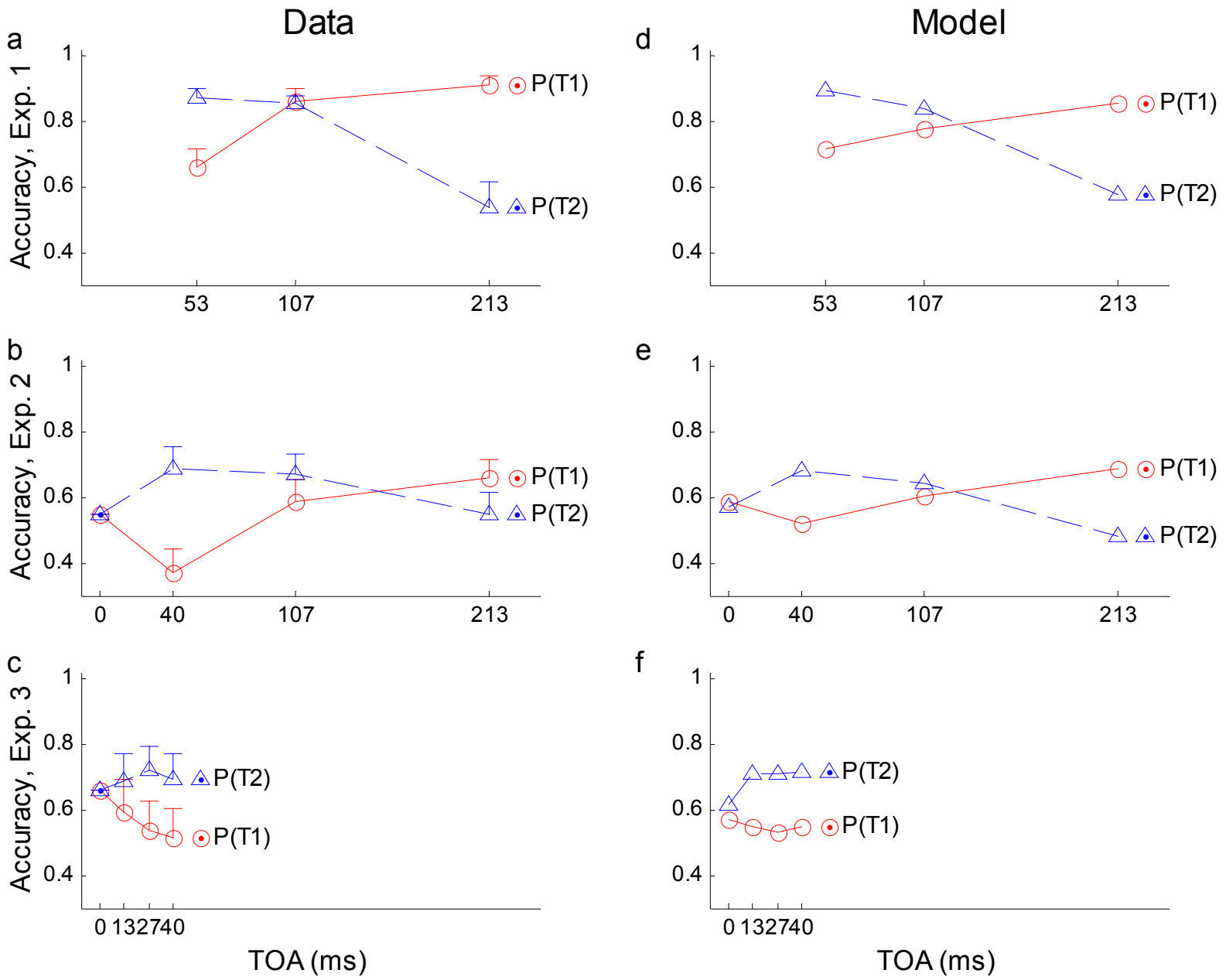
Potter, Staub, and O'Connor (2002)

- SOA = 53 ms
- Shorter TOAs
- T1 and T2: words
- Distractors: ##### or %%%%

	Exp 1	Exp 2	Exp 3
RSVP	Single	Unsynchronized dual	
TOA (ms)	53, 107, 213	0, 40, 107, 213	0, 13, 27, 40

- Exp 2&3 vs. 1: acuity, items entering WM, noise, etc.
- (μ, σ)

Potter et al. (2002)

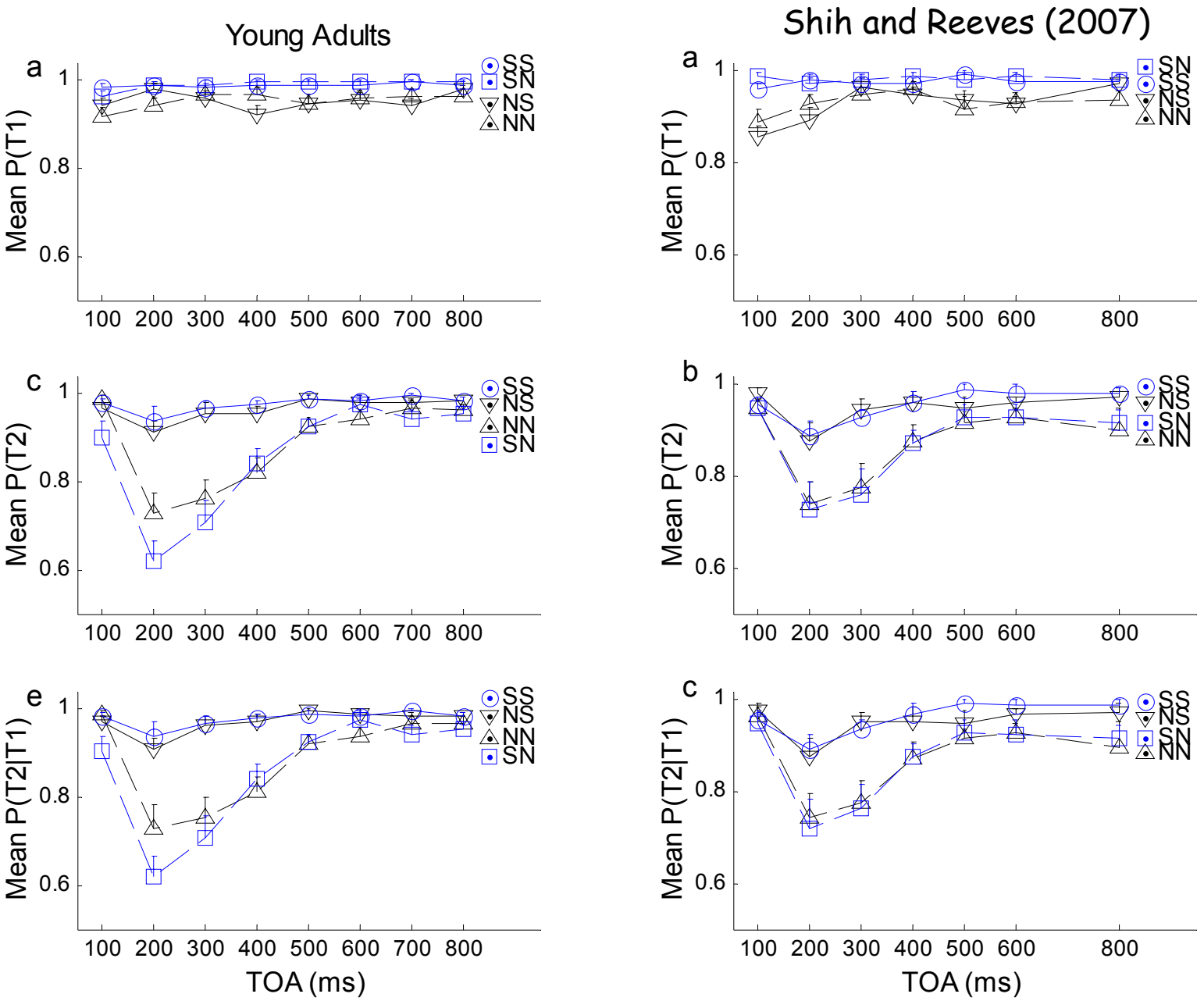


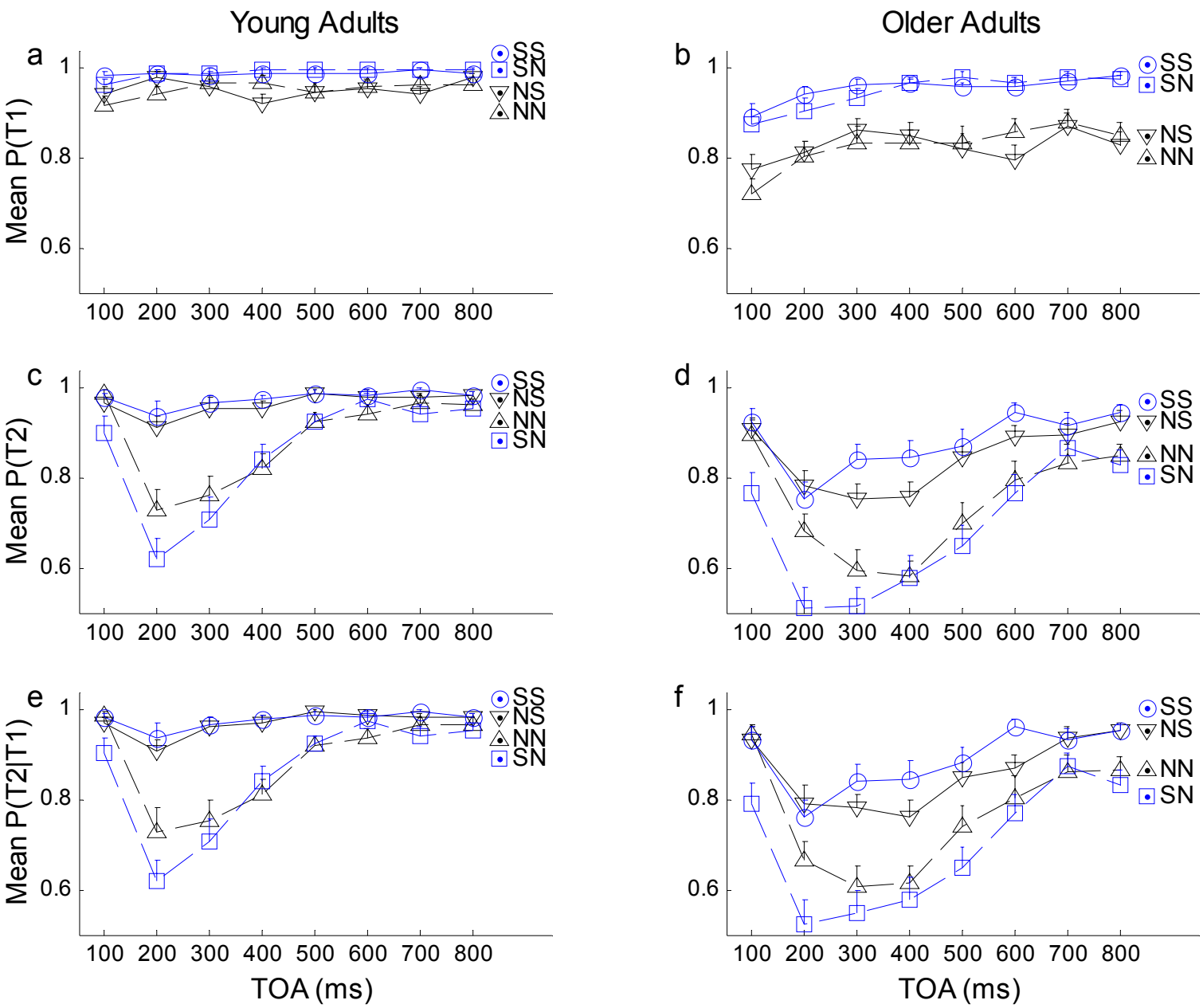
Cognitive Aging

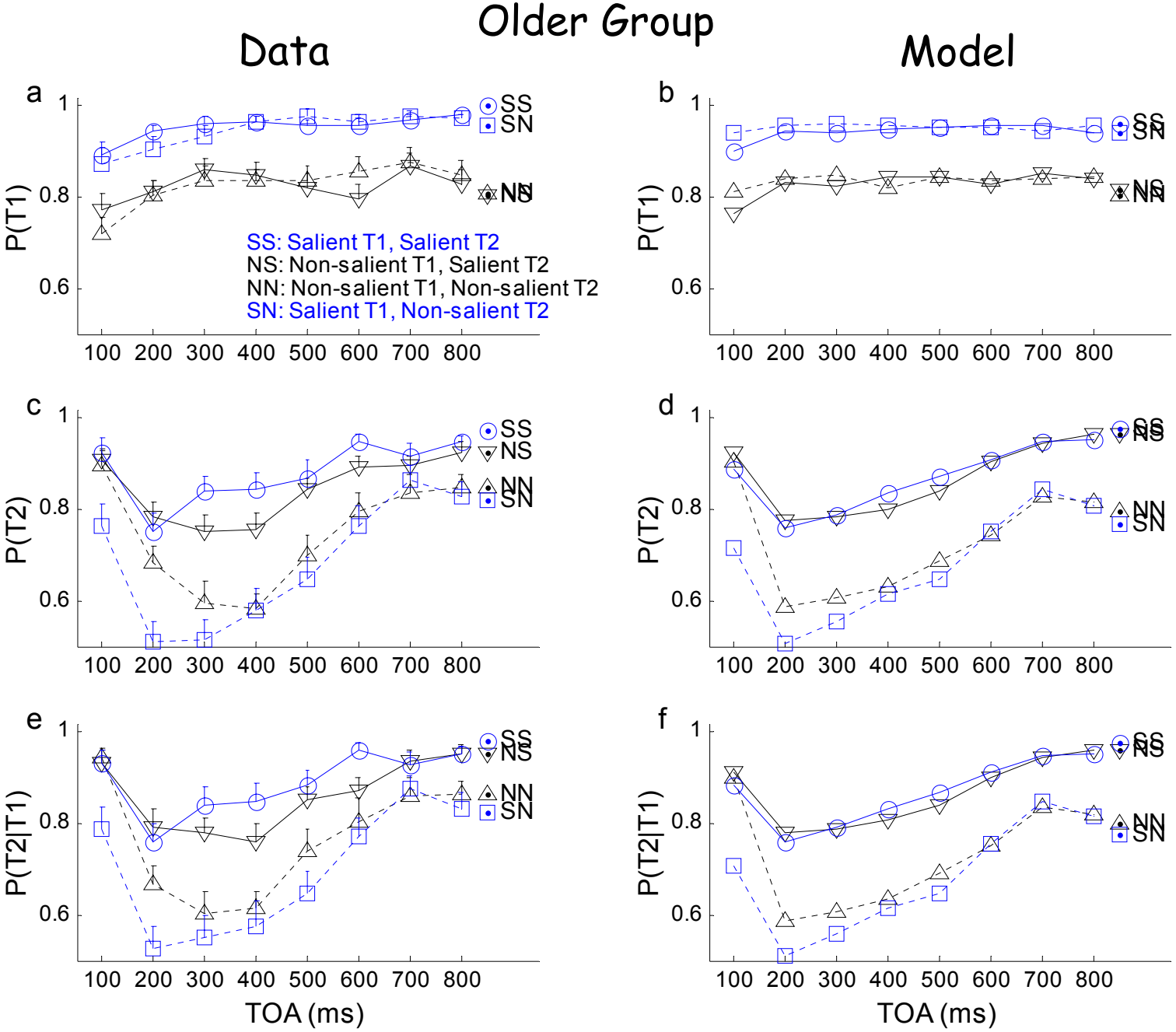
- Reduced
  - processing speed (e.g., Salthouse, 1996)
  - ability in inhibiting irrelevant information (e.g., Hasher & Zacks, 1988)
  - attentional resources or working memory capacity (e.g., Craik & Byrd, 1982)
  - sensory/perceptual processing efficiency

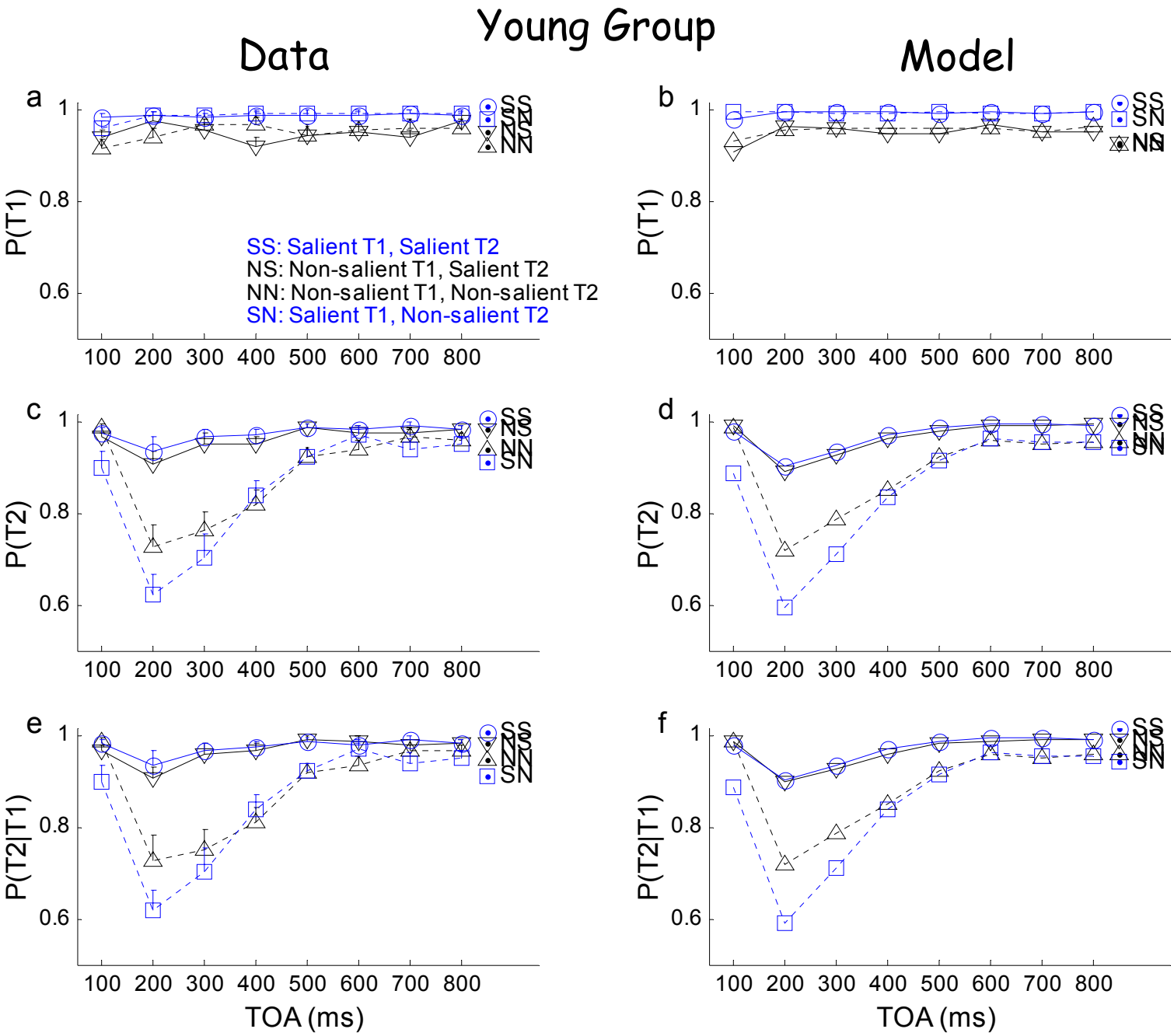
- T1 Salience x T2 Salience x TOA
- Bright red, bright green, and black characters
- Big font size
- SOA = 100 ms

N-N-100	. . . U S 3 6 P T X . . .
N-S-200	. . . G X 5 L 8 C V . . .
S-N-300	. . . B D 7 R Y 2 K . . .
S-S-600	. . . H S 9 P X V C V 4 N . . .


Young Adults
Shih and Reeves (2007)
a
Mean P(T1)
SS
SN
NS
NN
c
Mean P(T2)
b
e
Mean P(T2|T1)
c
TOA (ms)
1
0.8
0.6
100 200 300 400 500 600 700 800


Young Adults
Older Adults
a
b
c
d
e
f
Mean P(T1)
Mean P(T2)
Mean P(T2|T1)
SS
SN
NS
NN
1
0.8
0.6
100 200 300 400 500 600 700 800
TOA (ms)


Older Group
Data
Model
a
b
c
d
e
f
P(T1)
P(T2)
P(T2|T1)
TOA (ms)
SS
SN
NS
NN
SS: Salient T1, Salient T2
NS: Non-salient T1, Salient T2
NN: Non-salient T1, Non-salient T2
SN: Salient T1, Non-salient T2


Young Group
Data
Model
a
b
c
d
e
f
P(T1)
P(T2)
P(T2|T1)
1
0.8
0.6
100 200 300 400 500 600 700 800
TOA (ms)
SS
NS
SN
NN
SS: Salient T1, Salient T2
NS: Non-salient T1, Salient T2
NN: Non-salient T1, Non-salient T2
SN: Salient T1, Non-salient T2

Parameter	Young Group		Older Group	
	Optimum	Range	Optimum	Range
θ, initial masking factor (dimensionless)	0.52	(0.48, 0.58)	0.44	(0.36, 0.52)
β, Time constant (ms)	11	(10.5, 11.6)	14	(12.8, 14.3)
w, Width of attention window (ms)	140	(135, 150)	160	(> 120)
C, Capacity of ICP (Item per SOA unit)	0.98	(0.94, 1.01)	0.95	(0.90, 0.98)
π, Duration of ICP (ms)	590	(570, 610)	700	(650, 730)
μ_n, Mean of ICP noise (ms)	10	(7, 13)	25	(19, 28)
σ_n, SD of ICP noise (ms)	40	(35, 44)	65	(54, 73)
Mean R^2	0.958		0.920	
SD R^2	0.006		0.008	

Conclusions

- Useful
- Automatic vs. controlled
- Attention gating
- Elaborate working memory
- General theory of attention