Using the Attention Cascade Model to Computationally Account for the Age Differences in an Attentional Blink (AB) Task

Shui-I Shih, University of Southampton, United Kingdom (sis@soton.ac.uk)

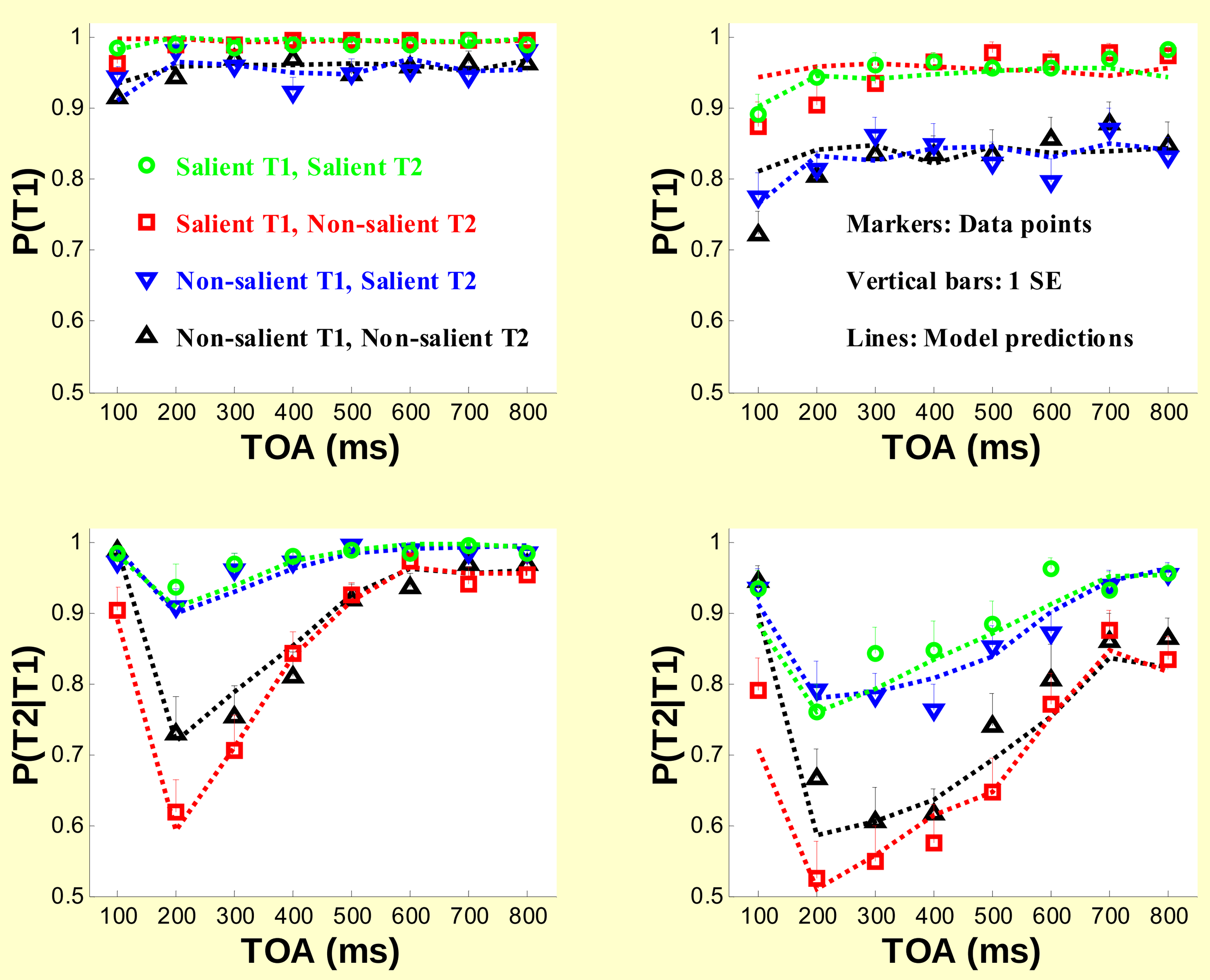
The Study

- **Introduction**
  - The **attention cascade model** (Shih, 2008) is a general, mathematical model of attention and working memory. It is applied here to characterize cognitive aging.

- **Method**
  - Task: search for two targets (T1 and T2) in rapid serial visual presentation (SOA = 100 ms)
    - ⇒ T1: 3, 5, 7, 9 ; T2: 2, 4, 6, 8 ; Distractor: black letters
  - Design: T1 Salience (red/black) x T2 Salience (green/black) x T1-T2 Lag

- **Results**

Young (n = 19, M = 22.4 yrs) **Older (n = 23, M = 72.6 yrs)**



- The older group performed worse.
- The older group exhibited greater and longer AB — a loss of performance on a later target, T2, when an earlier target, T1, is processed.
- Target salience improved accuracy.

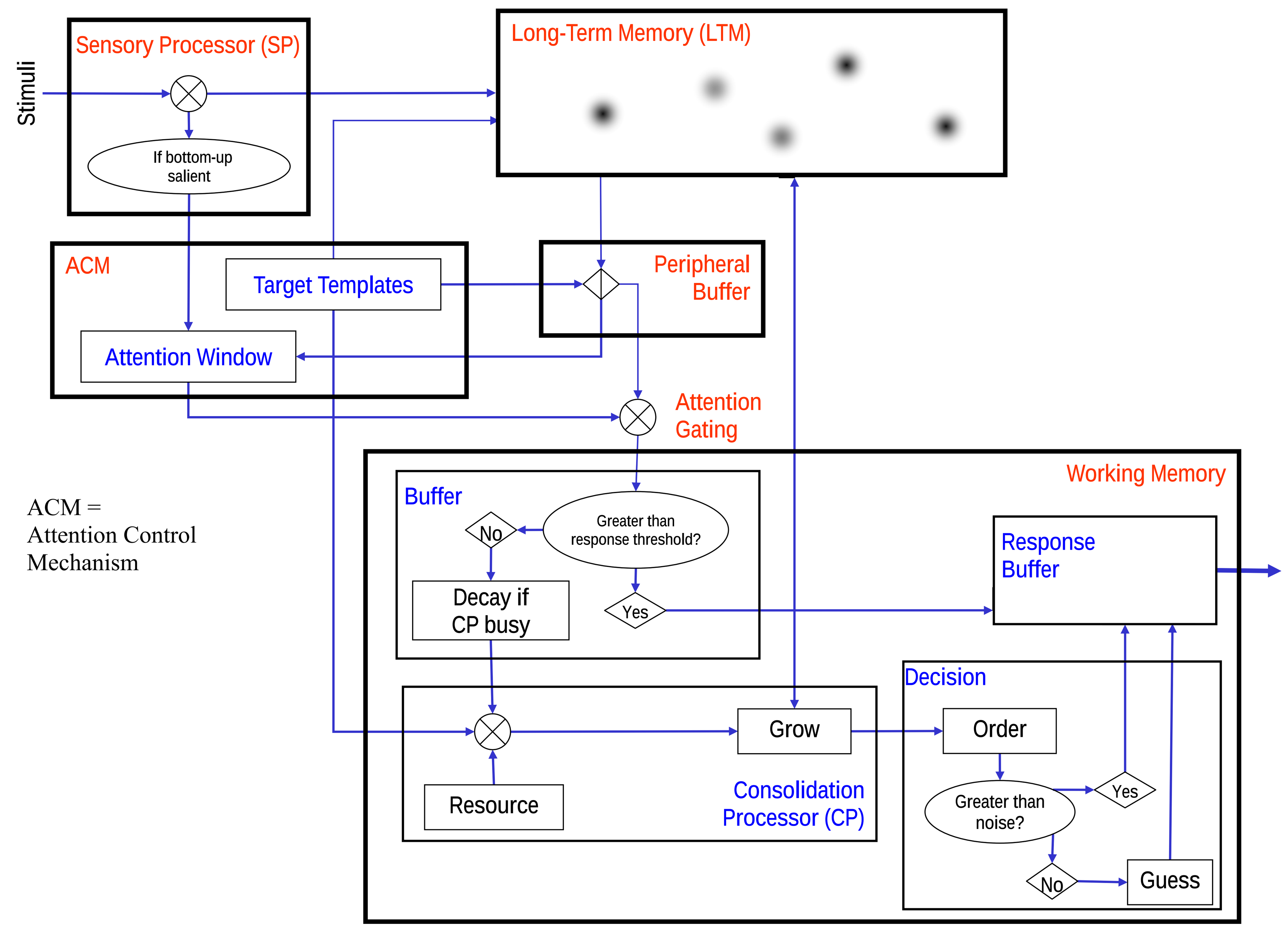
- **Computational accounts (Table 1)**
  - ◊ 96 data points [3 measures by 32 conditions] for each group; Measures: P(T1), P(T2) and P(T2|T1)
  - ◊ 10,000 bootstrap samples are used to estimate the optimum values and 95% confidence interval.
  - Using α = .05, the two groups did not differ in
    - ◊ The processing rate (1/β) prior to the WM stages
    - ◊ The width of the attention window
    - ◊ The capacity C of the consolidation processor
  - However, relative to the young, the older adults
    - ◊ Suffer more masking effect of the salient (and brighter) stimulus
    - ◊ Require a longer consolidation duration
    - ◊ Have greater mean and variance of the internal noise (assuming a Gaussian distribution)

- **Conclusion**
  - The attention cascade model relates the age differences in the AB task to the sensory and working memory components. The model may be a useful tool in comparative studies.

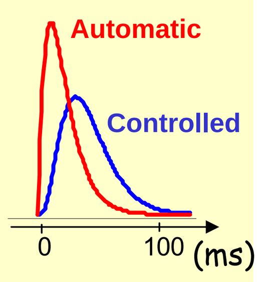
Reference:

Shih, S. (2008). The attention cascade model and attentional blink. *Cognitive Psychology, 56, 210-236.*

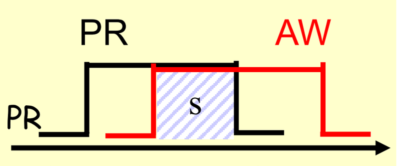
Descriptions of the Attention Cascade Model



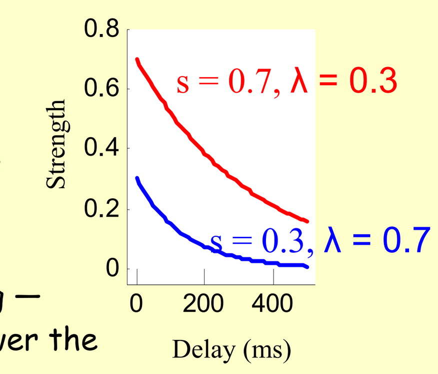
- **Sensory processor (SP)**
  - Interference may scale down the stimulus strength
  - Mandatory output to LTM; output to the ACM only if bottom-up salient (e.g., with a distinct color).
- **Long-term memory (LTM) traces**
  - Same activation level assumed for well-learned items
- **Preliminary representations (PRs) in the peripheral buffer**
  - Each PR is described by a rectangular function
    - ◊ Width = SOA (i.e., perceptually available)
    - ◊ Scalable height (e.g., due to masking)
- **Attention Control Mechanism (ACM)**
  - **Attention window** (AW) transfers PRs into the WM buffer
    - ◊ The AW is described by a rectangular function
      - ⇒ width (interval) modulated by the task demand, etc.
  - **Target templates** (TTs) ≡ task demands, behavioral goals
  - Two modes of triggering the AW
    - ◊ **Controlled**: by a top-down salient stimulus via TTs
    - ◊ **Automatic**: by a bottom-up salient stimulus via the SP.
  - The AW triggering time distribution is a 2nd- (automatic) or 4th-order (controlled) gamma function
    - ◊ Assuming the processing time in each pre-WM stage is *iid* as an exponential pdf with the time constant β



- **Attention gating**
  - The overlap between the AW and PR
  - The area under PR of the overlap defines the initial strength s of the input to the WM buffer



- **Working memory (WM)**
  - WM buffer
    - ◊ If s > response threshold, output to the response buffer
    - ◊ Otherwise, hold the inputs if the CP is engaged
    - ◊ Strength s **decays** exponentially while queuing — the greater the s, the slower the decay



  - Consolidation processor (CP)
    - ◊ Requires π ms. Once engaged, it takes no more inputs.
    - ◊ **Resource**: s is weighted according to its top-down salience subject to resource availability
    - ◊ **Grow**: the weighted s grows exponentially during consolidation
  - Decision
    - ◊ Order: an input with greater s is perceived to have occurred earlier
    - ◊ If the final s > noise, produce a correct response. Otherwise, make a guess.

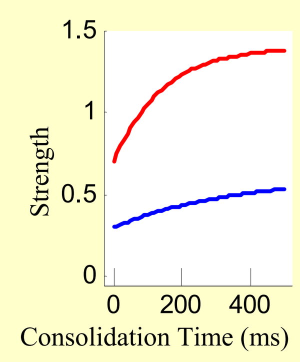


Table 1. Parameter Estimates of 2.5, 50, and 97.5 Percentiles for Each Group

Parameter	Young				Older		
	2.5	50	97.5		2.5	50	97.5
θ, initial masking factor (--)	0.49	**0.52**	0.55	*	0.41	**0.44**	0.48
β, time constant (ms) of pre-WM stages	9.3	**11.0**	12.6		11.5	**13.8**	15.8
w, width of attention window (ms)	135	**140**	164		141	**145**	194
C, CP capacity (item per SOA unit)	0.95	**0.98**	1.01		0.92	**0.95**	0.98
π, CP duration (ms)	580	**590**	635	*	684	**700**	747
μ_n, mean of CP noise (ms)	7	**10**	14	*	21	**25**	29
σ_n, SD of CP noise (ms)	36	**40**	43	*	61	**65**	70
R^2	0.83	0.92	0.95		0.79	0.87	0.91
Mean (SD) R^2		0.96 (0.006)				0.92 (0.008)	

Note. CP = consolidation processor. The values for the 2.5 and 97.5 percentile respectively provide the lower and upper bounds of the 95% confidence interval for the distribution of 10,000 bootstrap samples. The value for the 50 percentile coincides with the optimum estimate. R^2 denotes the amount of variance in the data that is accounted for by the model corrected for the number of free parameters. The value of R^2 is between 0 and 1, with 1 denoting a perfect fit. In each round of simulations, there are 1000 Monte Carlo trials for each of the 32 conditions; each condition provides three dependent measures: P(T1), P(T2), and P(T2|T1). The Mean (SD) R^2 are based on 100 rounds of Monte Carlo simulations using the optimum estimates.