Learning and Prediction

A gentle Intro to AIS students

Uri Hasson

May 2022

AIS M.A Course

Table of contents

1. Background: Learning and anticipation

2. Gradual consolidation; slow forgetting

3. Error-driven-learning (Rescorla-Wagner) model of learning based on AFO data

4. Relationship between prediction and behavior (AFO and saccades)

Background: Learning and anticipation

Studying learning: What knowledge people have

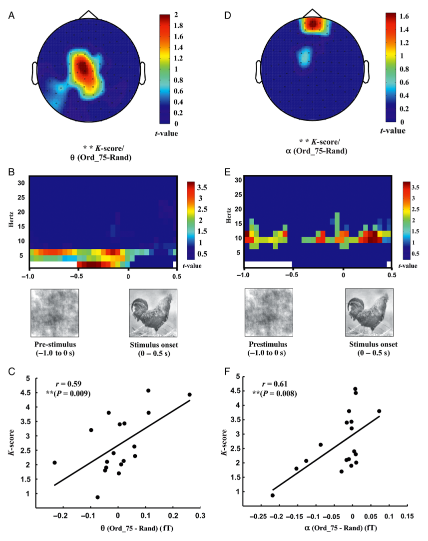
- Learning theory: advanced by quantifying neurobiological responses to stimuli with different levels of probability/stimulus-features.
- Learning optimizes perception and behavior, with **Surprise** (sometimes called prediction error) long-considered key component because surprise (Response) is a clue to what people know.
- But (I): S-R depends on more than what people know: mediated by low-level perception, accumulation of evidence, report biases, and response-initiation processes (noise)
- BUT (II) S-R also ignores the issue of whether people use learning to construct Expectations or Predictions..

Studying learning: What knowledge people have

- Learning theory: advanced by quantifying neurobiological responses to stimuli with different levels of probability/stimulus-features.
- Learning optimizes perception and behavior, with **Surprise** (sometimes called prediction error) long-considered key component because surprise (Response) is a clue to what people know.
- But (I): S-R depends on more than what people know: mediated by low-level perception, accumulation of evidence, report biases, and response-initiation processes (noise)
- BUT (II) S-R also ignores the issue of whether people use learning to construct Expectations or Predictions..
- Studies of Expectation probe the state of a cognitive system prior to stimulus presentation and independent of stimulus-guided responses.

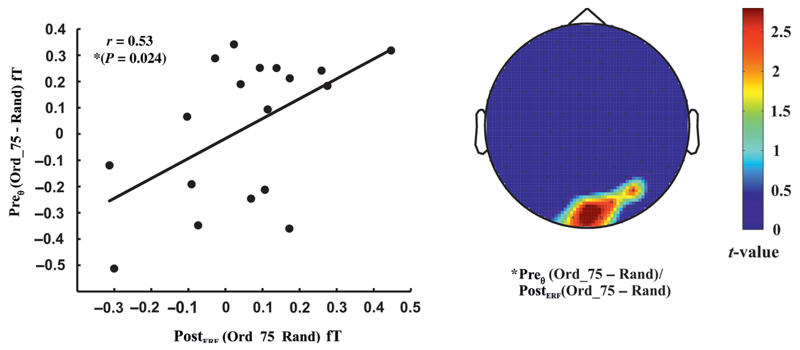
MEG: learning produces predictions

- MEG: Image categories presented randomly or predictably.
- Pre-stimulus activity correlates with WM.
- theta and alpha effects
- Cashdollar, Nathan, et al. doi: 10.1093/cercor/bhw138.



.. And predictions impact processing

Stronger activity for predictable stimulus ($Ord75 - Rand$) for people more sensitive to series order (greater prestimulus theta-band power for Ordered vs. Random series)



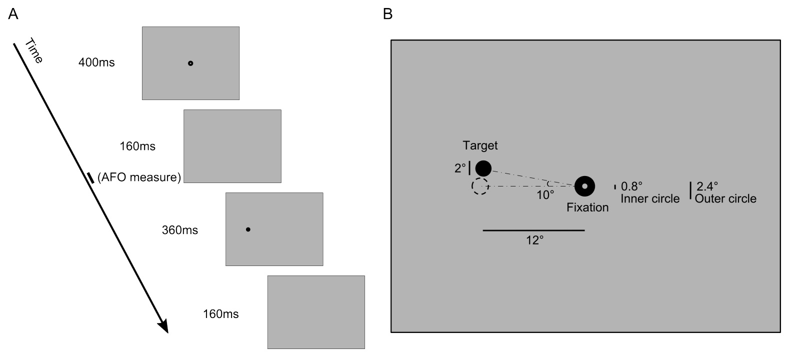
The logic of studying learning and prediction

- How do we quantify learning from observable behavior?
- How do we Separate predictive processes from responses
- Can we quantify the relation between predictions and responses

An experimental approach

- $N = 21$ Volunteers.
- Each Series $\leftarrow$ 100 trials with target side, but not precise location determined by Markov process (pret70, pret30).
- Proportion of presentations on the left and right screen sides set at 50% in both conditions (marginals).
- Twenty random-location trials appended to each series to study wash-out effects.
- 10 series presented in each condition.
- Tower mounted video based Eye-tracking at 1000Hz.

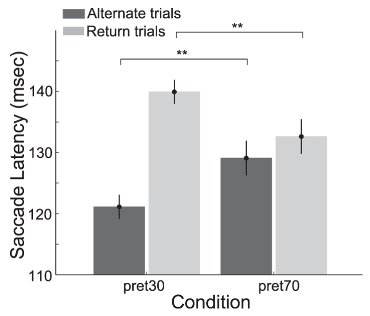
Basic paradigm



Definitions I

1. Return vs. Alternation trial: Screen-side of the last-presented target was the same / opposite as the one that preceded it.

Saccade latencies provide evidence for learning



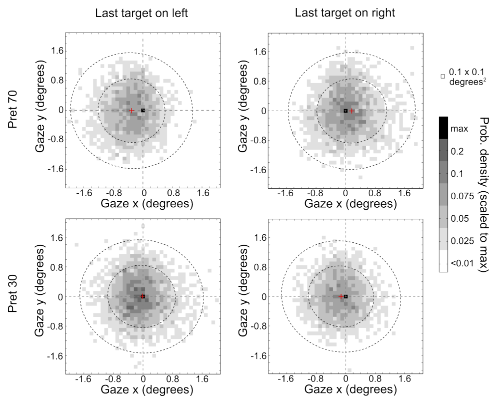
- Saccades to target presented on alternate sides faster when alternates are more predictable (pret30). $RT_{\text{alternations}}$: $pret30 < pret70$
- Saccades to targets presented on return sides faster when returns are more predictable (pret70). RT_{returns} : $pret70 < pret30$
- Also, in general, returns are slower than alternations because of Inhibition of Return (explore; not exploit)

Definitions II

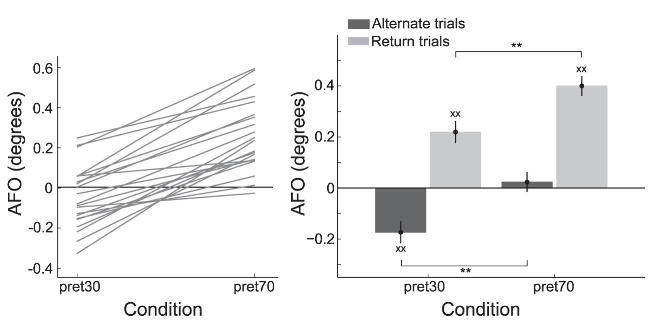
1. Anticipatory Gaze offset: absolute gaze location in relation to center.
2. Anticipatory Fixation Offset: Recoding of Gaze Offset in relation to prior target: horizontal deviation from center coded as positive if to the side of the last presented target, negative otherwise. Measured during final 10ms of pre-target blank screen.

Anticipatory Gaze offsets: descriptives

Density of fixation location during last 10 ms of pretarget blank screen. The single dark point: screen center. Red points: mean values for condition; inner/outer circles mark areas encompassing 50% and 90% of all fixations. When returns expected, anticipatory bias towards location of last target.



AFO (horizontal shift) indicates active prediction



- Left panel: AFO tracks predictive structure.
- BUT. Also independent impact of the very last trial. Example: Even when returns are frequent (pret70), if last trial was alternate, much reduced prediction.
- The behavior is subtle: AFO values $\sim 0.4^\circ$ from center. Within the spatial zone of the just-removed fixation symbol.

Validation

Internal (Split-half) reliability: deriving two separate ΔAFO values per participant: one from odd trials and one from even trials. $SH = 0.90$.

Studying the time scale of learning.

- Even when there is predictable regularity in the world, people do not always predict the most likely outcome (we saw: strong impact of last trial).
- This is irrational. People will be most successful if they always bet on the most likely outcome.
- Still, people cannot ignore the impact of the most recent past (perhaps useful for non-stationary environments).
- We need a Mathematical Modeling approach to see quantify how information is integrated over time to determine current behavior.
- The model can be applied to quantify Responses (Saccade Latency) or prediction (AFO)

AFO: Modeling impact of 6 prior transitions

$$AFO = \beta_1 S_1 + \beta_2 S_2 + \beta_3 S_3 + \beta_4 S_4 + \beta_5 S_5 + \beta_6 S_6 + c + \varepsilon$$

Where:

$S = 1$ last target is a return

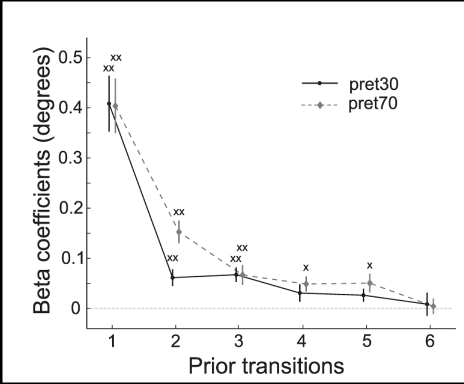
$S = 0$ last target is a alternation

$k = 1 \ldots 6$ last transitions

- Positive β: a return at lag k prior transitions is associated with larger AFO values (shift towards previous target).
- Regression model fit per participant per condition (pret70, pret30)

How the recent past impacts prediction (AFO)

- Recent 'return' trials bias prediction towards side of most recent target. Very strong effect of last trial.
- Decay is rapid
- Impact stronger for pret70



Saccade Latency: Impact of each of of 6 prior transitions

$$SL = \beta_1 S_1 + \beta_2 S_2 + \beta_3 S_3 + \beta_4 S_4 + \beta_5 S_5 + \beta_6 S_6 + c + \varepsilon$$

Where:

$S = 1$ current target is a return

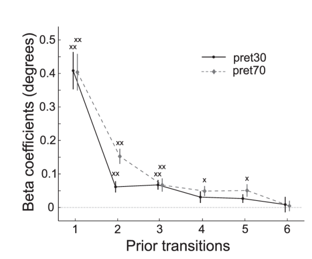
$S = 0$ current target is a alternation

$k = 1 \dots 6$ last transitions

- Positive β: a return at lag k prior transitions increaesd SLs.
- Regresion model fit per participant per condition (pret70, pret30) but seperately for return and alternation saccades (control: IOR, surprise)

Results: Six prior transitions (ii)

AFO results



Saccade Results

- No impact of recent past on response times for 'predicted' targets: Alternation saccades in pret30 Return saccades in pret70
- Return saccades (surprising targets) in pret30: impact limited to immediately preceding transition: return saccades faster when preceded by a return. Interim: 'minor' impact of past in pret30.
- Alternation saccades in pret70: $\beta_1 \ldots \beta_4$ significantly positive: recent return trials slow down an alternation saccade (vs. returns)

Interim Summary: basic descriptive statistics

1. Saccade latencies confirm statistical learning
2. AFOs were small, on average around 0.4°
3. AFO reflected learning of global statistics
4. AFO was strongly impacted by the immediately preceding transition, and (more weakly) by the preceding 3-5 transitions.
5. SL were more weakly impacted by recent trials and manifestation of this impact depended on the type of saccade made.
6. SL reflect IOR; AFOs do not.

Gradual consolidation; slow forgetting

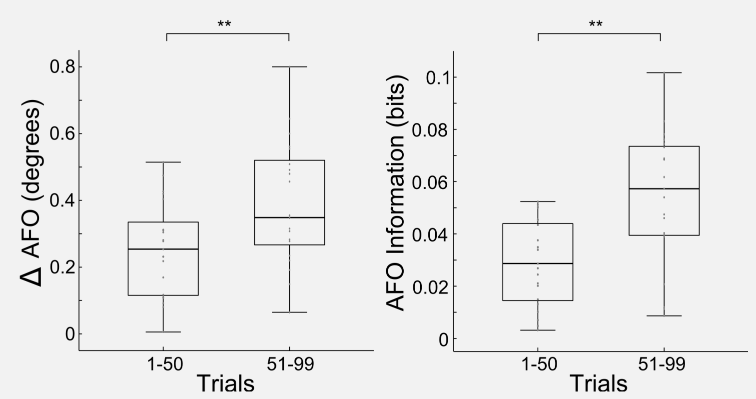
Studying time scales of knowledge consolidation and process of forgetting

Each series consisted of 100 trials, followed by 20 random trials

1. Is there consolidation of knowledge over the entire set of 100 trials?
2. Is there rapid forgetting during the washout trials?

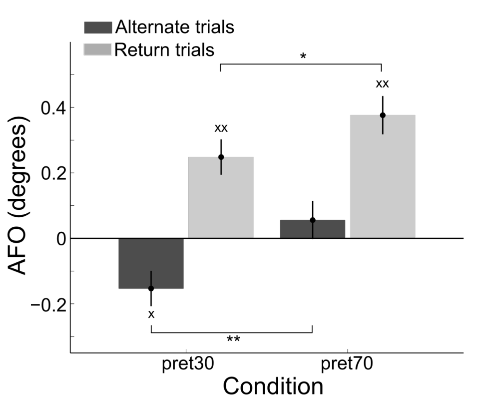
During learning: ΔAFO increases over trials

ΔAFO indicates difference in AFO position between pret70 and pret30.
'info' is M.I between AFO values and condition.



During washout: previous statistics impact AFO during random trials

Washout: During 'forgetting' people still show effects of Prior statistical Condition (and the immediately prior trial).



Error-driven-learning (Rescorla-Wagner) model of learning based on AFO data

Rescorla-Wagner learning model based on AFO

- Implementation: Response model maps beliefs about transitions to observed behavior (AFO).
- α: learning rate; K scaling factor transforming internal probability to overt behavior.

$$\begin{cases} P_{ret}(t+1) = P_{ret}(t) + \alpha(1 - P_{ret}(t)) & \text{after a return} \\ P_{ret}(t+1) = P_{ret}(t) - \alpha P_{ret}(t) & \text{after an alternation} \\ AFO(t+1) = K(P_{ret}(t+1)) \end{cases} \quad (1)$$

Example: current $P_{ret}(t) = 0.6$. Learning rate $\alpha = 0.1$

- if there is another return, $P_{ret}(t+1) = 0.64$
- if there is an alternation, $P_{ret}(t+1) = 0.54$

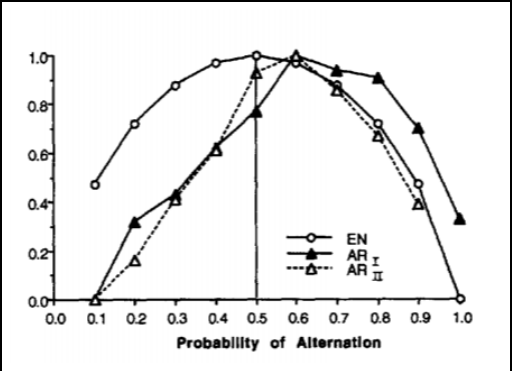
Rescorla-Wagner learning model based on AFO

An advantage of an error-driven learning model is that it allows modeling learning rate: the relative importance of a new observation in relation to prior knowledge. This is directly related to the drop-off curve of the Betas in the regression model (impact of past on present).

The scaling factor K further deals with the possibility that in different situations, the exact same knowledge can translate to behaviors of different magnitude.

Small complication: An intuitive notion is that the internal probability for returns > 0.5 people will start predicting returns because probability exceeds chance. The problem is that people mis-estimate chance in a systematic way. Their point of equilibrium is thought to be around $P(ret) = 0.6$

in binary series, subjective points of equilibrium strongly deviate from 50%; random binary series are subjectively perceived as containing too many streaks (Falk & Konold, 1997).



Rescorla-Wagner learning model based on AFO

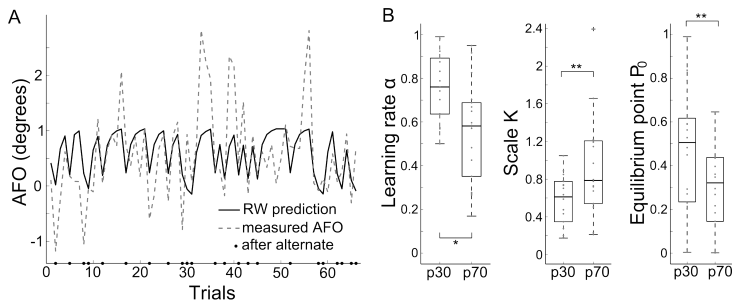
- So slightly revised equations
- α: learning rate; K scaling factor transforming internal probability to overt behavior; P_0 is a *probability equilibrium point* reflecting an internal estimate of probability of return above which a participant shows a gaze bias towards the return side.

$$\begin{cases} P_{ret}(t+1) = P_{ret}(t) + \alpha(1 - P_{ret}(t)) & \text{after a return} \\ P_{ret}(t+1) = P_{ret}(t) - \alpha P_{ret}(t) & \text{after an alternation} \\ AFO(t+1) = K(P_{ret}(t+1) - P_0) \end{cases} \tag{2}$$

RW: model details II

- Model successfully cross-validated on left-out data for 19/20 participants. Parameters estimated from 9 series applied to 10th. We evaluate whether the predicted time series matches the true time series.
- It is also possible to construct a RW model that consists of two processes, each with its own learning rate (Bornstein & Daw, 2012). In this extended model, two estimations of the transition probability are updated independently based on two learning rates, α_1, α_2 and an overall summary statistic $P_{ret}(t)$ was their weighted average: $P_{ret}(t) = w\, P^{(1)}_{ret}(t, \alpha_1) + (1 - w)\, P^{(2)}_{ret}(t, \alpha_2)$. No advantage.

RW: Results



- A: Sample AFO series in pret70 and RW prediction based on independent parameter estimates from the 9 other series.
- B: For pret70: larger integration windows α, and internal probabilities translate into stronger behavioral impacts (K).
- C: In all, pret30 seems 'more random' psychologically.

RW: Summary

- Successful validation on left-out data per participant.
- Different learning rates for two processes with identical conditional entropy: supports the idea of a subjective perception of randomness.
- Modulations of K consistent with greater confidence in executing behavior in pret70.

Relationship between prediction and behavior (AFO and saccades)

Questions

1. Do AFOs predict subsequent saccade latencies in manner consistent with prior prediction?
2. Is there a difference in the amount of information AFOs and SLs carry about the experimental condition?

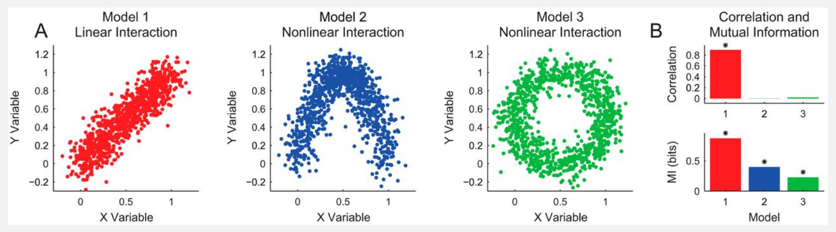
How does pre-saccade AFO constrain subsequent saccade latency?

If AFO measures prediction, the following relation should hold between measured AFO and following SL:

- Larger AFO values (putative anticipation of target on prior side) should be followed by faster return saccades, but slower alternation saccades.
- That is: negative AFO/SL correlation for returns, positive AFO/SL correlation for alternations.
- This is what is found. A larger anticipatory bias towards return predicts a faster return saccade, and also predicts a slower alternation saccade.

Which measure carries more information about pret? AFO or SL?

Mutual Information (MI) captures the amount of knowledge one variable provides about another (uncertainty about one variable that is reduced by knowing another). Does not assume any particular relationship between two variables.



Timme, N. M., & Lapish, C. (2018). A Tutorial for Information Theory in Neuroscience. eNeuro, 5(3).

Mutual information methods

$$I(x;w) = H(x) - H(x|w) = \sum_{x \in X} \sum_{w \in W} p(x,w)\, log\left(\frac{p(x,w)}{p(x)p(w)}\right) \qquad (3)$$

$$\begin{aligned} X &= \text{experimental condition (pret??); 1bit} \\ W &= (1)SL, or \\ &= (2)(SL|trial - type) \quad \text{factoring differential surprise}, or \\ &= (3)AFO \end{aligned}$$

None binary variable discretisized into six equally populated bins.

Mutual information results

AFO conveyed around twice as much information about the statistical process compared to SL: 0.0527 ± 0.0063 vs. 0.0245 ± 0.0050, $p < .001$

Summary: relation between AFO and SL

AFOs prior to a saccade contains information about saccade latencies in a manner consistent with anticipatory predictions.

AFOs contain more information about statistical context than that contained in SL distribution or SL distribution conditioned on trial-type.

Conclusions

- Learning can be studied by quantifying observable behavior over time.
- Macro-scale properties of the environment such as association-strength can be separated from micro-scale occurrences in recent past. Both impact behavior
- Learning translates into surprise signals: people react more slowly to surprising events
- Learning produces anticipation/prediction, which can be modeled with error-driven learning.
- Anticipatory signals can contain more information about the environment than stimulus-related responses