Lecture 6: Attention

Samuel Gershman

Harvard University

Roadmap

- The world is filled with stimuli, but only some of them are important at any given time.

Roadmap

- The world is filled with stimuli, but only some of them are important at any given time.
- The brain focuses on important stimuli through attentional mechanisms.

Roadmap

- The world is filled with stimuli, but only some of them are important at any given time.
- The brain focuses on important stimuli through attentional mechanisms.
- We'll look at two mechanisms: (1) expected stimuli receive higher posterior probability, rendering them more detectable; (2) attentional cues direct limited resources to important stimuli, amplifying information.

What is attention?

- Attention is the enhanced processing of certain signals, typically (but not always) at the expense of other signals.

What is attention?

- Attention is the enhanced processing of certain signals, typically (but not always) at the expense of other signals.
- It can be induced "top-down" (e.g., by preparatory cues, volitional orienting) or "bottom-up" (e.g., capture of attention by salient or unexpected stimuli).

What is attention?

- Attention is the enhanced processing of certain signals, typically (but not always) at the expense of other signals.
- It can be induced "top-down" (e.g., by preparatory cues, volitional orienting) or "bottom-up" (e.g., capture of attention by salient or unexpected stimuli).
- Our goal is to formalize attention in the Bayesian decision theory framework.

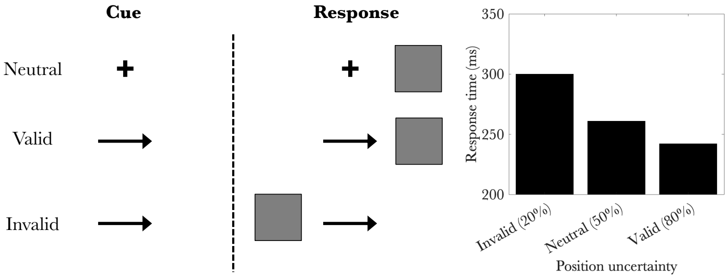
Attention as prior probability

- The most straightforward way to think about attention in a Bayesian framework is through the prior: if some events have higher prior probability, they will be more easily detected.

Attention as prior probability

- The most straightforward way to think about attention in a Bayesian framework is through the prior: if some events have higher prior probability, they will be more easily detected.
- To illustrate, we'll study the classic Posner cueing task.

The Posner cueing task



Validity effect: Invalid - Valid response time

The Posner cueing task

- Subjects maintain central fixation throughout the trial.

The Posner cueing task

- Subjects maintain central fixation throughout the trial.
- This means that the orienting of attention is "covert" in the sense that the sensory signals are largely unchanged. Only the internal direction of attention is altered by the cues.

Modeling the Posner cueing task

- Binary hidden state, $s \in \{\mathrm{Left}, \mathrm{Right}\}$, encoded by neural activity x at the time of the target appearance.

Modeling the Posner cueing task

- Binary hidden state, $s \in \{\text{Left}, \text{Right}\}$, encoded by neural activity x at the time of the target appearance.
- We will assume that the spike count of neuron d is Poisson-distributed with rate $f_d(s)$—i.e., $x_d|s \sim \text{Poisson}(f_d(s))$.

Modeling the Posner cueing task

- Recall how evidence accumulation can be implemented by a readout neuron whose input current $I(t)$ linearly weights spikes from the encoding population:

  $$I(t) = \sum_d w_d z_d(t)$$

  where $z_d(t) = 1$ if neuron d spikes at time t (0 otherwise), and the synaptic strength is given by:

  $$w_d = \log \frac{f_d(\text{Left})}{f_d(\text{Right})}$$

Modeling the Posner cueing task

- Recall how evidence accumulation can be implemented by a readout neuron whose input current $I(t)$ linearly weights spikes from the encoding population:

$$I(t) = \sum_d w_d z_d(t)$$

where $z_d(t) = 1$ if neuron d spikes at time t (0 otherwise), and the synaptic strength is given by:

$$w_d = \log \frac{f_d(\text{Left})}{f_d(\text{Right})}$$

- The input current is integrated by the membrane potential $\mu(t)$:

$$C\dot{\mu} = I(t)$$

where C is the membrane capacitance and $\mu(0)$ is the potential at the time of cue presentation ($t = 0$).

Modeling the Posner cueing task

- $\mu(0)$ is the critical variable, because that's what encodes the cue-dependent prior:

$$\mu(0) = \log \frac{p(\text{Left}|\text{Cue})}{p(\text{Right}|\text{Cue})}$$

Modeling the Posner cueing task

- $\mu(0)$ is the critical variable, because that's what encodes the cue-dependent prior:

$$\mu(0) = \log \frac{p(\text{Left}|\text{Cue})}{p(\text{Right}|\text{Cue})}$$

- The membrane potential $\mu(t)$ at time t represents the posterior log odds conditional on the cue and sensory evidence.

Modeling the Posner cueing task

- $\mu(0)$ is the critical variable, because that's what encodes the cue-dependent prior:

$$\mu(0) = \log \frac{p(\text{Left}|\text{Cue})}{p(\text{Right}|\text{Cue})}$$

- The membrane potential $\mu(t)$ at time t represents the posterior log odds conditional on the cue and sensory evidence.
- We can link the model to behavior by assuming that a detection response is made whenever the firing rate of the readout neuron crosses one of two thresholds representing the two spatial locations.

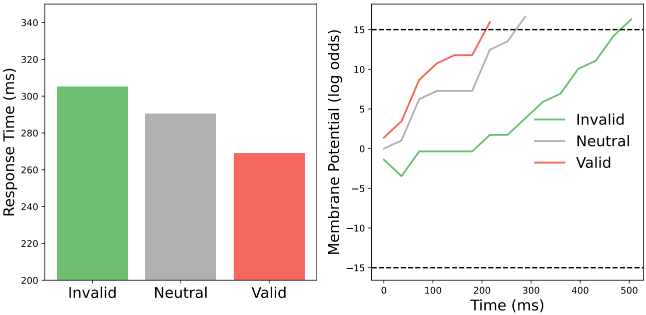
Modeling the Posner cueing task

- The model can explain the main behavioral finding: the arrow cue drives the membrane potential up or down depending on its direction, which in turn speeds the crossing of the decision threshold when the target appears in the expected location.

Modeling the Posner cueing task

- The model can explain the main behavioral finding: the arrow cue drives the membrane potential up or down depending on its direction, which in turn speeds the crossing of the decision threshold when the target appears in the expected location.
- When the cue appears in the unexpected location, the evidence needs to move in the opposite direction, making response time longer—this formalizes the concept of "reorienting" that Posner and others have invoked to explain responses on invalid trials.

Modeling the Posner cueing task



Acetylcholine and cued attention

- Yu & Dayan [2005] posited that acetylcholine reports *expected uncertainty* about target stimuli conditional on cues.

Acetylcholine and cued attention

- Yu & Dayan [2005] posited that acetylcholine reports *expected uncertainty* about target stimuli conditional on cues.
- Acetylcholine should control the size of the validity effect: greater uncertainty about the cued target location will increase response times on valid trials and decrease response times on invalid trials.

Acetylcholine and cued attention

- In support of the model, nicotine (an acetylcholine agonist) decreased the validity effect, while scopolamine (an acetylcholine antagonist) increased it [Phillips et al 2000].

Acetylcholine and cued attention

- In support of the model, nicotine (an acetylcholine agonist) decreased the validity effect, while scopolamine (an acetylcholine antagonist) increased it [Phillips et al 2000].
- However, this study appears to be an outlier, because several other studies found opposite effects.

Acetylcholine and cued attention

- In support of the model, nicotine (an acetylcholine agonist) decreased the validity effect, while scopolamine (an acetylcholine antagonist) increased it [Phillips et al 2000].
- However, this study appears to be an outlier, because several other studies found opposite effects.
- Acetylcholine is clearly involved in cued attention, but perhaps not in the specific way posited by Yu & Dayan. The bulk of data are compatible with a model in which acetylcholine reports *expected certainty*.

Modeling acetylcholine and cued attention

- One possible formalization: acetylcholine reports $\mu(0)$.

Modeling acetylcholine and cued attention

- One possible formalization: acetylcholine reports $\mu(0)$.
- Suppressing acetylcholine should have the effect of abolishing cue-dependent attentional modulation of the readout neuron.

Modeling acetylcholine and cued attention

- One possible formalization: acetylcholine reports $\mu(0)$.
- Suppressing acetylcholine should have the effect of abolishing cue-dependent attentional modulation of the readout neuron.
- Consistent with this hypothesis, Davidson et al [2000] showed that attentional modulation of parietal cells during performance of the Posner task is suppressed in a dose-dependent manner by scopolamine.

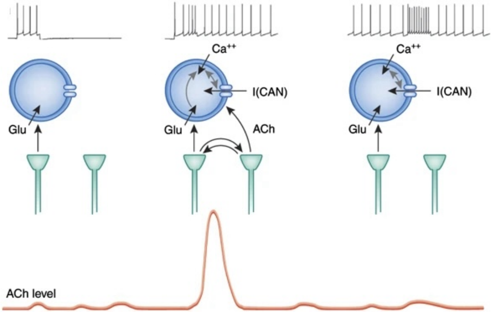
A possible biophysical mechanism for acetylcholine's attentional effects

- When a neuron is sufficiently depolarized, calcium channels are opened, resulting in the activation of a calcium-sensitive nonspecific cation current, which causes further depolarization, further calcium influx, and so on—a self-sustaining loop producing persistent spiking.

A possible biophysical mechanism for acetylcholine's attentional effects

- When a neuron is sufficiently depolarized, calcium channels are opened, resulting in the activation of a calcium-sensitive nonspecific cation current, which causes further depolarization, further calcium influx, and so on—a self-sustaining loop producing persistent spiking.
- Subsequent stimuli will be more effective at activating neurons with elevated activity.

A possible biophysical mechanism for acetylcholine's attentional effects



[Hasselmo & Sarter 2011]

Effects on tuning

- Spatial priors can also be used to orient attention towards particular features.

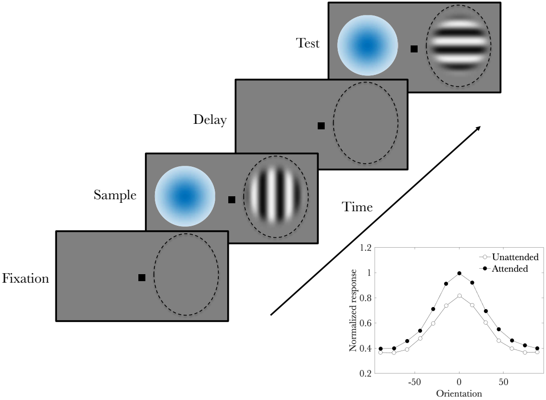
Effects on tuning

- Spatial priors can also be used to orient attention towards particular features.
- This can be realized through modulation of tuning functions.

Effects on tuning

- Spatial priors can also be used to orient attention towards particular features.
- This can be realized through modulation of tuning functions.
- Example [McAdams & Maunsell 1999]: attention to orientation multiplicatively enhances orientation tuning.

Effects on tuning in V4



[McAdams & Maunsell 1999]

Modeling effects on tuning

- Idea: prior is higher in the attended location. In the extreme case (0 probability outside the attended location), the posterior takes the form of a "spotlight" that selectively amplifies information in a specific location.

Modeling effects on tuning

- Idea: prior is higher in the attended location. In the extreme case (0 probability outside the attended location), the posterior takes the form of a "spotlight" that selectively amplifies information in a specific location.
- Rao [2005] developed a model of the delayed match-to-sample task in which V4 neurons reported the posterior probability over orientation (s_1) given lower-level visual inputs (x), marginalizing over location (s_2):

$$p(s_1|x) = \sum_{s_2} p(s_1|s_2, x)p(s_2|x)$$

where $s_2 \in \{\text{Left}, \text{Right}\}$.

Modeling effects on tuning

- Idea: prior is higher in the attended location. In the extreme case (0 probability outside the attended location), the posterior takes the form of a "spotlight" that selectively amplifies information in a specific location.
- Rao [2005] developed a model of the delayed match-to-sample task in which V4 neurons reported the posterior probability over orientation (s_1) given lower-level visual inputs (x), marginalizing over location (s_2):

$$p(s_1|x) = \sum_{s_2} p(s_1|s_2, x)p(s_2|x)$$

  where $s_2 \in \{\mathrm{Left}, \mathrm{Right}\}$.
- Because $p(s_2|x)$ increases with $p(s_2)$, the posterior over orientation is modulated multiplicatively by the location prior.

Modeling effects on tuning

- Rao's model can also explain competitive attentional phenomena.

Modeling effects on tuning

- Rao's model can also explain competitive attentional phenomena.
- When multiple stimuli are presented in the receptive field of a V4 neuron, activity is suppressed [Reynolds et al 1999].

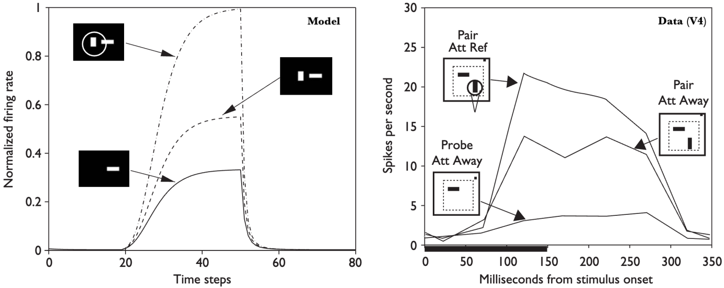
Modeling effects on tuning

- Rao's model can also explain competitive attentional phenomena.
- When multiple stimuli are presented in the receptive field of a V4 neuron, activity is suppressed [Reynolds et al 1999].
- The model explains this suppression due to uncertainty about orientation.

Modeling effects on tuning

- Rao's model can also explain competitive attentional phenomena.
- When multiple stimuli are presented in the receptive field of a V4 neuron, activity is suppressed [Reynolds et al 1999].
- The model explains this suppression due to uncertainty about orientation.
- When a monkey is trained to attend to the location where the horizontal bar is presented, activity is restored—a pattern recapitulated by the model due to the elevated prior probability assigned to the location.

Attentional competition in V4



[Rao 2005; Reynolds et al 1999]

Connection to normalization

- The normalization model [Reynolds & Heeger 2009] describes the firing rate of a neuron as a function of stimulus feature (s_1, here orientation) and location (s_2), which we write as $s = (s_1, s_2)$:
$$f(s) = \frac{E(s)A(s)}{S + \nu}$$
where ν is a gain control parameter, $E(s)$ is the stimulus drive, $A(s)$ is the attention field, and S is the suppressive drive derived from pooling the excitatory drive $E(s)A(s)$ over a wider range of s (the suppressive field).

Connection to normalization

- The normalization model [Reynolds & Heeger 2009] describes the firing rate of a neuron as a function of stimulus feature (s_1, here orientation) and location (s_2), which we write as $s = (s_1, s_2)$:

$$f(s) = \frac{E(s)A(s)}{S + \nu}$$

  where ν is a gain control parameter, $E(s)$ is the stimulus drive, $A(s)$ is the attention field, and S is the suppressive drive derived from pooling the excitatory drive $E(s)A(s)$ over a wider range of s (the suppressive field).
- This model can explain many of the attentional phenomena already discussed.

Connection to normalization

- We can map the normalization model onto the Bayesian model if we think of the likelihood $p(x|s)$ as encoding the stimulus drive $E(s)$, the prior $p(s)$ as encoding the attention field $A(s)$, and the marginal likelihood $p(x) = \sum_s p(x|s)p(s)$ as encoding the suppressive drive S (pooled excitatory drive).

Connection to normalization

- We can map the normalization model onto the Bayesian model if we think of the likelihood $p(x|s)$ as encoding the stimulus drive $E(s)$, the prior $p(s)$ as encoding the attention field $A(s)$, and the marginal likelihood $p(x) = \sum_s p(x|s)p(s)$ as encoding the suppressive drive S (pooled excitatory drive).
- Bayes' rule then implements a form of normalization with $\nu = 0$:
$$p(s|x) = \frac{p(x|s)p(s)}{p(x)} = \frac{E(s)A(s)}{S}$$

Normalization and the contrast response function

- For neurons in visual areas (e.g., V4), firing rate generally increases with stimulus contrast.

Normalization and the contrast response function

- For neurons in visual areas (e.g., V4), firing rate generally increases with stimulus contrast.
- Some studies observed a multiplicative modulation of the contrast response function by attention—a bigger boost in firing for higher-contrast attended stimuli.

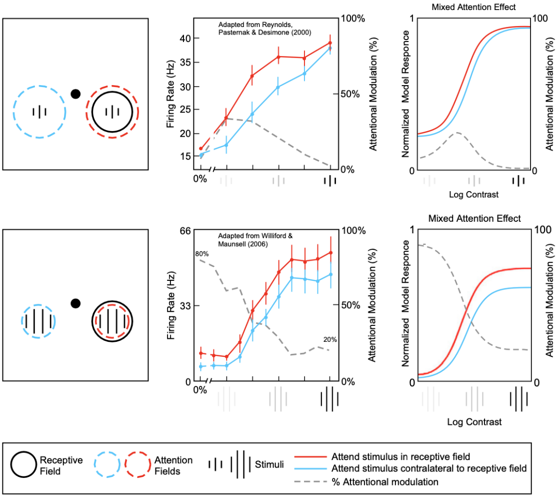
Normalization and the contrast response function

- For neurons in visual areas (e.g., V4), firing rate generally increases with stimulus contrast.
- Some studies observed a multiplicative modulation of the contrast response function by attention—a bigger boost in firing for higher-contrast attended stimuli.
- Other studies observed a shift in the contrast response function for attended stimuli—roughly the same boost in firing for all contrast levels.

Normalization and the contrast response function

- For neurons in visual areas (e.g., V4), firing rate generally increases with stimulus contrast.
- Some studies observed a multiplicative modulation of the contrast response function by attention—a bigger boost in firing for higher-contrast attended stimuli.
- Other studies observed a shift in the contrast response function for attended stimuli—roughly the same boost in firing for all contrast levels.
- To explain this discrepancy, Reynolds & Heeger proposed that multiplicative modulation arises when the attention field is small relative to the stimulus, whereas a shift arises when the attention field is large relative to the stimulus.

Attentional modulation of contrast response in V4



[Reynolds et al 2000; Williford & Maunsell 2006; Reynolds & Heeger 2009]

Normalization and the contrast response function

- Reynolds and Heeger assumed that stimulus drive is a linear function of contrast $g(s)$: $E(s) = \alpha g(s)$. They also assumed that the attentional field is approximately uniform with strength γ.

Normalization and the contrast response function

- Reynolds and Heeger assumed that stimulus drive is a linear function of contrast $g(s)$: $E(s) = \alpha g(s)$. They also assumed that the attentional field is approximately uniform with strength γ.
- In the case where the attention field is large relative to the stimulus, the attention field is approximately constant as a function of s and the contrast response function takes the following form:

  $$f(s) \approx \frac{\alpha\gamma g(s)}{\gamma g(s) + \nu}$$

  On a log scale, this shifts the entire contrast response function, as in Reynold et al [2000].

Normalization and the contrast response function

- In the case where the attention field is small relative to the stimulus, the attention field is approximately a focal point at the stimulus location, and the contrast response function takes the following form:

  $$f(s) \approx \frac{\alpha\gamma g(s)}{\gamma g(s) + \omega g(s) + \nu}$$

  where $\omega \in [0,1]$ scales the suppressive drive from the region surrounding the focal point.

Normalization and the contrast response function

- In the case where the attention field is small relative to the stimulus, the attention field is approximately a focal point at the stimulus location, and the contrast response function takes the following form:

  $$f(s) \approx \frac{\alpha\gamma g(s)}{\gamma g(s) + \omega g(s) + \nu}$$

  where $\omega \in [0, 1]$ scales the suppressive drive from the region surrounding the focal point.
- At low contrasts, this predicts a shift in the contrast response function, whereas at high contrast it predicts elevated attention that does not vary with contrast, broadly consistent with Williford & Maunsell [2006].

Bayesian interpretation of normalization

- Posterior probability increases with contrast due to its action on the likelihood (expressing the stimulus drive).

Bayesian interpretation of normalization

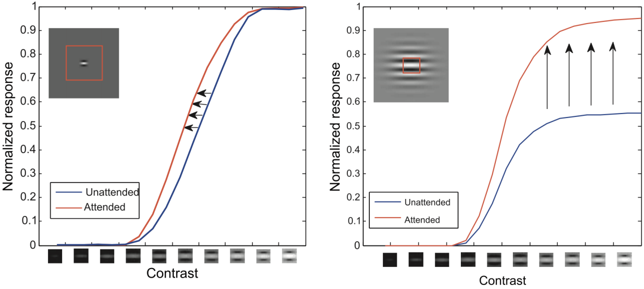
- Posterior probability increases with contrast due to its action on the likelihood (expressing the stimulus drive).
- Posterior probability also increase with attention due to its action on the prior.

Bayesian interpretation of normalization

- Posterior probability increases with contrast due to its action on the likelihood (expressing the stimulus drive).
- Posterior probability also increase with attention due to its action on the prior.
- The marginal likelihood expresses the suppressive drive, increasing with stimulus size due to the fact that more stimulus locations enter into the normalizing constant.

Bayesian interpretation of normalization

When the stimulus is small relative to the attention field (left), the contrast response function is shifted by attention. When the stimulus is large relative to the attention field (right), the contrast response function is scaled multiplicatively by attention.



[Chikkerur et al 2010]

Sharpened tuning

- The normalization model can also explain why attention sometimes sharpens stimulus tuning [e.g., Martinez-Trujillo and Treue 2004]. This contrasts with McAdams & Maunsell [1999], where the entire tuning curve is shifted upward.

Sharpened tuning

- The normalization model can also explain why attention sometimes sharpens stimulus tuning [e.g., Martinez-Trujillo and Treue 2004]. This contrasts with McAdams & Maunsell [1999], where the entire tuning curve is shifted upward.
- The variable over which attention is specified is different across studies. In McAdams & Maunsell, attention was specified spatially (i.e., whether the stimulus was inside or outside the neuron's spatial receptive field). In Martinez-Trujillo & Treue, attention was specified based on the stimulus feature (motion direction).

Sharpened tuning

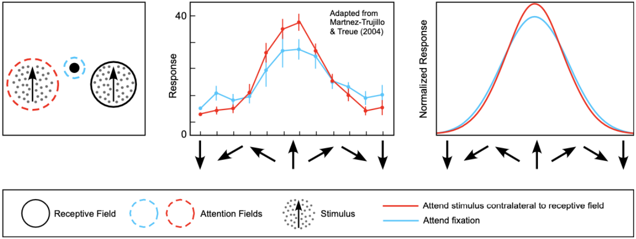
- The normalization model can also explain why attention sometimes sharpens stimulus tuning [e.g., Martinez-Trujillo and Treue 2004]. This contrasts with McAdams & Maunsell [1999], where the entire tuning curve is shifted upward.
- The variable over which attention is specified is different across studies. In McAdams & Maunsell, attention was specified spatially (i.e., whether the stimulus was inside or outside the neuron's spatial receptive field). In Martinez-Trujillo & Treue, attention was specified based on the stimulus feature (motion direction).
- In the normalization model, the attention field near the attended feature value is amplified, sharpening tuning.

Sharpened tuning

- The normalization model can also explain why attention sometimes sharpens stimulus tuning [e.g., Martinez-Trujillo and Treue 2004]. This contrasts with McAdams & Maunsell [1999], where the entire tuning curve is shifted upward.
- The variable over which attention is specified is different across studies. In McAdams & Maunsell, attention was specified spatially (i.e., whether the stimulus was inside or outside the neuron's spatial receptive field). In Martinez-Trujillo & Treue, attention was specified based on the stimulus feature (motion direction).
- In the normalization model, the attention field near the attended feature value is amplified, sharpening tuning.
- Bayesian interpretation: increase in prior probability for particular feature values. Different effects on tuning functions across experiments reflect spatial vs. feature-based priors.

Sharpened tuning

One stimulus was always shown in the receptive field of the recorded neuron (black circle), while attention was directed either to the fixation point (blue dashed circle) or to the contralateral stimulus (red dashed circle).



[Martinez-Trujillo & Treue 2004; Reynolds & Heeger 2009]

Study question

Why does attention sometimes multiplicatively enhance tuning curves and sometimes sharpen them?

Attention as noise reduction

- While prior probability can explain some aspects of attention, it's unlikely to be a complete account; not all attentional phenomena make sense as changes in the prior.

Attention as noise reduction

- While prior probability can explain some aspects of attention, it's unlikely to be a complete account; not all attentional phenomena make sense as changes in the prior.
- When multiple stimuli are presented simultaneously and the subject has been trained to respond to only one of them, the response-relevant stimulus does not necessarily have higher probability in terms of its location or feature value.

Attention as noise reduction

- While prior probability can explain some aspects of attention, it's unlikely to be a complete account; not all attentional phenomena make sense as changes in the prior.
- When multiple stimuli are presented simultaneously and the subject has been trained to respond to only one of them, the response-relevant stimulus does not necessarily have higher probability in terms of its location or feature value.
- What's needed is a different conceptualization of attentional effects that captures allocation of limited/costly cognitive resources.

Optimizing signal precision: an elementary model

- Agent receives a set of N independent Gaussian signals with mean s and precision (inverse variance) λ, summarized by an aggregate signal $x \sim \mathcal{N}(s, 1/(N\lambda))$.

Optimizing signal precision: an elementary model

- Agent receives a set of N independent Gaussian signals with mean s and precision (inverse variance) λ, summarized by an aggregate signal $x \sim \mathcal{N}(s, 1/(N\lambda))$.
- Stimulus is drawn from a prior $s \sim \mathcal{N}(\bar{s}, 1/\lambda_0)$.

Optimizing signal precision: an elementary model

- Agent receives a set of N independent Gaussian signals with mean s and precision (inverse variance) λ, summarized by an aggregate signal $x \sim \mathcal{N}(s, 1/(N\lambda))$.
- Stimulus is drawn from a prior $s \sim \mathcal{N}(\bar{s}, 1/\lambda_0)$.
- Posterior mean $\hat{s}$ is a convex combination of the prior mean and the signal:

$$\hat{s} = wx + (1 - w)\bar{s}$$

  where

$$w = \frac{N\lambda}{\lambda_0 + N\lambda}$$

  is the *signal sensitivity*.

Optimizing signal precision: an elementary model

- Agent receives a set of N independent Gaussian signals with mean s and precision (inverse variance) λ, summarized by an aggregate signal $x \sim \mathcal{N}(s, 1/(N\lambda))$.
- Stimulus is drawn from a prior $s \sim \mathcal{N}(\bar{s}, 1/\lambda_0)$.
- Posterior mean $\hat{s}$ is a convex combination of the prior mean and the signal:

  $$\hat{s} = wx + (1 - w)\bar{s}$$

  where

  $$w = \frac{N\lambda}{\lambda_0 + N\lambda}$$

  is the *signal sensitivity*.
- New twist: the agent can adjust the signal precision—the allocation of attention is a kind of cognitive action which the agent can optimize.

Optimizing signal precision: an elementary model

- Utility is a monotonically decreasing and differentiable function of the squared error $\epsilon = (s - \hat{s})^2$.

Optimizing signal precision: an elementary model

- Utility is a monotonically decreasing and differentiable function of the squared error $\epsilon = (s - \hat{s})^2$.
- Taylor series approximation:

$$u(\epsilon) \approx u(0) - \beta(s - \hat{s})^2,$$

where $u(0)$ is the maximum achievable utility. and $\beta = -u''(0)$ is the *attentional incentive*, which determines the degree to which utility is contingent on error.

Optimizing signal precision: an elementary model

- Utility is a monotonically decreasing and differentiable function of the squared error $\epsilon = (s - \hat{s})^2$.
- Taylor series approximation:

$$u(\epsilon) \approx u(0) - \beta(s - \hat{s})^2,$$

  where $u(0)$ is the maximum achievable utility. and $\beta = -u''(0)$ is the *attentional incentive*, which determines the degree to which utility is contingent on error.
- Intuitively, agents should be more motivated to pay attention when this contingency is stronger.

Optimizing signal precision: an elementary model

- Expected utility:

$$\bar{u}(\lambda) = \mathbb{E}[u(\epsilon)|\lambda] \approx u(0) - \frac{\beta}{N\lambda + \lambda_0}$$

Optimizing signal precision: an elementary model

- Expected utility:

$$\bar{u}(\lambda) = \mathbb{E}[u(\epsilon)|\lambda] \approx u(0) - \frac{\beta}{N\lambda + \lambda_0}$$

- The agent achieves higher utility when: (i) the attentional incentive β is smaller (i.e., the agent doesn't need to pay attention to earn utility); (ii) the signal precision λ is larger; (iii) the prior precision λ_0 is larger; and (iv) the sample size N is larger.

Optimizing signal precision: an elementary model

- Cost function (KL divergence between posterior and prior):

$$c(\lambda) = \frac{1}{2} \log \left(1 + \frac{N\lambda}{\lambda_0} \right)$$

Optimizing signal precision: an elementary model

- Cost function (KL divergence between posterior and prior):

$$c(\lambda) = \frac{1}{2} \log \left(1 + \frac{N\lambda}{\lambda_0} \right)$$

- Precision optimization:

$$\lambda^* = \underset{\lambda}{\operatorname{argmax}}\, \bar{u}(\lambda) - \kappa c(\lambda)$$

where $\kappa > 0$ is an *attentional cost* parameter that captures the agent's limited capacity.

Optimizing signal precision: an elementary model

- Optimal precision:

$$\lambda^* = \max(0, (2\beta/\kappa - \lambda_0)/N)$$

Optimizing signal precision: an elementary model

- Optimal precision:

$$\lambda^* = \max(0, (2\beta/\kappa - \lambda_0)/N)$$

- Grows with the attentional incentive and shrinks with the attentional cost.

Optimizing signal precision: an elementary model

- Optimal precision:

$$\lambda^* = \max(0, (2\beta/\kappa - \lambda_0)/N)$$

- Grows with the attentional incentive and shrinks with the attentional cost.
- Smaller for higher prior precision: the agent doesn't need to pay attention to the signal as much when more confident prior to observing the signal.

Optimizing signal precision: an elementary model

- Optimal precision:

$$\lambda^* = \max(0, (2\beta/\kappa - \lambda_0)/N)$$

- Grows with the attentional incentive and shrinks with the attentional cost.
- Smaller for higher prior precision: the agent doesn't need to pay attention to the signal as much when more confident prior to observing the signal.
- All of these factors are attenuated with a larger sample size: the same amount of information can be obtained with less attention by observing for longer.

Dopamine and signal precision

- Several lines of evidence suggest that signal precision is controlled by the neuromodulator dopamine.

Dopamine and signal precision

- Several lines of evidence suggest that signal precision is controlled by the neuromodulator dopamine.
- Pharmacologically elevating dopamine increases sensitivity to stimuli.

Dopamine and signal precision

- Several lines of evidence suggest that signal precision is controlled by the neuromodulator dopamine.
- Pharmacologically elevating dopamine increases sensitivity to stimuli.
- For example, elevating dopamine using methylphenidate increases sensitivity in a motion discrimination task [Beste et al 2018].

Dopamine and signal precision

- Several lines of evidence suggest that signal precision is controlled by the neuromodulator dopamine.
- Pharmacologically elevating dopamine increases sensitivity to stimuli.
- For example, elevating dopamine using methylphenidate increases sensitivity in a motion discrimination task [Beste et al 2018].
- Patients with Parkinson's disease exhibit attenuated sensitivity to mechanical stimulation that was ameliorated in proportion to their dopamine medication dose [Wolpe et al 2018].

Study question

Which empirical phenomena are better explained in terms of prior probability vs. noise reduction?

Summary

- Attention is not a unitary construct—it consists of several components that affect behavior and neural activity in different ways. Nonetheless, we can explain many aspects of attention modulation within a unified framework.

Summary

- Attention is not a unitary construct—it consists of several components that affect behavior and neural activity in different ways. Nonetheless, we can explain many aspects of attention modulation within a unified framework.
- Key concepts: shifts in prior probabilities (e.g., due to spatial cues) and shifts in discriminability (e.g., due to optimization of signal precision).

Summary

- Attention is not a unitary construct—it consists of several components that affect behavior and neural activity in different ways. Nonetheless, we can explain many aspects of attention modulation within a unified framework.
- Key concepts: shifts in prior probabilities (e.g., due to spatial cues) and shifts in discriminability (e.g., due to optimization of signal precision).
- These attentional shifts are linked to distinct neural correlates. It remains unclear how they interact to produce the relevant psychophysical changes.