Evidence weighting in confidence judgments for detection and discrimination

Matan Mazor1, Lucie Charles4, Roni Maimon-Mor4,5, & Stephen M. Fleming1,2,3

1 Wellcome Centre for Human Neuroimaging, UCL

2 Max Planck UCL Centre for Computational Psychiatry and Ageing Research

3 Department of Experimental Psychology, UCL

4 Institute of Cognitive Neuroscience, UCL

5 FMRIB Centre, Nuffield Department of Clinical Neuroscience, University of Oxford

Author note

The authors have no conflicting interests to declare.

Correspondence concerning this article should be addressed to Matan Mazor, 12 Queen Square, London WC1N 3BG. E-mail: [mtnmzor@gmail.com](mailto:mtnmzor@gmail.com)

Abstract

Confidence in perceptual decisions is more sensitive to evidence in support of the decision than to evidence against it. This positive evidence bias (PEB) has been demonstrated in confidence ratings in binary discrimination decisions between two stimulus categories. Recent theoretical proposals suggest that a PEB is due to observers adopting a detection-like strategy when rating their confidence, one that has functional benefits for metacognition in real-world settings where detectability and discriminability often go hand in hand. However, it is unknown whether, or how, a PEB is also in play for detection decisions about the presence or absence of a stimulus. In three experiments (one lab-based and two online) we first successfully replicate a PEB in discrimination confidence. We then show that a PEB is observed in detection decisions, where participants report the presence or absence of a stimulus, regardless of its identity. We discuss our findings in relation to models that account for a positive evidence bias as emerging from a confidence-specific heuristic, and alternative models where decision and confidence are generated by the same, Bayes-rational process.

*Keywords:* confidence, detection, metacognition

*Word count:* X

Evidence weighting in confidence judgments for detection and discrimination

# Compuatational models

## General framework

### Generative model.

Stimuli were represented as two vectors of 12 values each {\_l} and {\_r}, corresponding to the two sensory channels (e.g., the right and left stimuli in Exp. 2). In the discrimination task, one sensory channel transmitted pure noise (that is, samples were centered around zero), and one channel had additional signal in it (samples were centered around a non-zero value). The signal channel was chosen randomly for each trial with equal probability. In the detection task, both sensory channels transmitted pure noise.

On top of the presented noise, we added perceptual noise to the stimulus, resulting in a degraded representation of each sensory channel . Importantly, this additional noise affected the agent’s decisions and confidence ratings, but did not affect the stimulus itself such that trial-wise estimates of stimulus energy were unaffected for the reverse correlation analysis. The noise was channel and time specific.

### Belief update.

Agents kept track of three quantities, the log likelihood for signal in the right versus the left channels (), and the log likelihood for the presence of signal in one of the channels, versus noise in both channels (). Both values were set to 0 at the beginning of each trial. They were then updated from the second timepoint and on (we used the first timepoint as a control, to make sure reverse correlation analysis is not showing any effect of evidence at this timepoint). Log likelihood ratios were updated according to the following rule:

and

where is the true probability density function of values conditioned on signal being present in the channel, and is the true probability density function of values conditioned on signal being absent. is the probability that the signal is in the right channel (based on all previous samples) and is the probability that the signal is in the left channel (conditioned on signal presence). Note that subjects are rationally incorporating accurate beliefs about the effect of sensory noise on evidence strength in updating their beliefs about the world state.

### Decision.

In discrimination, agents decided ‘right’ when and ‘left’ otherwise. In detection, agents decided ‘present’ when and ‘absent’ otherwise.

### Confidence.

Confidence was the probability of being correct, given an equal prior over the two world states. This equals in discrimination and in detection.

## Vanilla

In the vanilla model, sensory noise was sampled from .

A python simulation is available [in the project’s GitHub](https://github.com/matanmazor/reverseCorrelation/blob/main/simulations/equal_variance.ipynb).

## Stimulus-dependent noise

In the stimulus-dependent noise model, sensory noise was sampled from .

A python simulation is available [in the project’s GitHub](https://github.com/matanmazor/reverseCorrelation/blob/main/simulations/unequal_variance.ipynb).

## Random attention

In the random attention noise model, sensory noise was sampled from . However, subjects had only access to one sensory channel per timepoint.

At the beginning of each trial, one of the two channels was chosen at random with equal probability to be the preferred channel. Then, at each timepoint, one of the two channels was selected to be attended, with a probability of being the preferred channel. Subjects then used information from this channel only to update their beliefs. For example, upon attending the left channel:

and

A python simulation is available [in the project’s GitHub](https://github.com/matanmazor/reverseCorrelation/blob/main/simulations/random_attention.ipynb).

## Goal-directed attention

In the goal-directed attention noise model, sensory noise was sampled from . However, subjects had only access to one sensory channel per timepoint.

The probability of attending the right channel was set to , where . This made agents heavily biased to attend the channel that is more likely to include signal.

A python simulation is available [in the project’s GitHub](https://github.com/matanmazor/reverseCorrelation/blob/main/simulations/goal_directed_attention.ipynb).