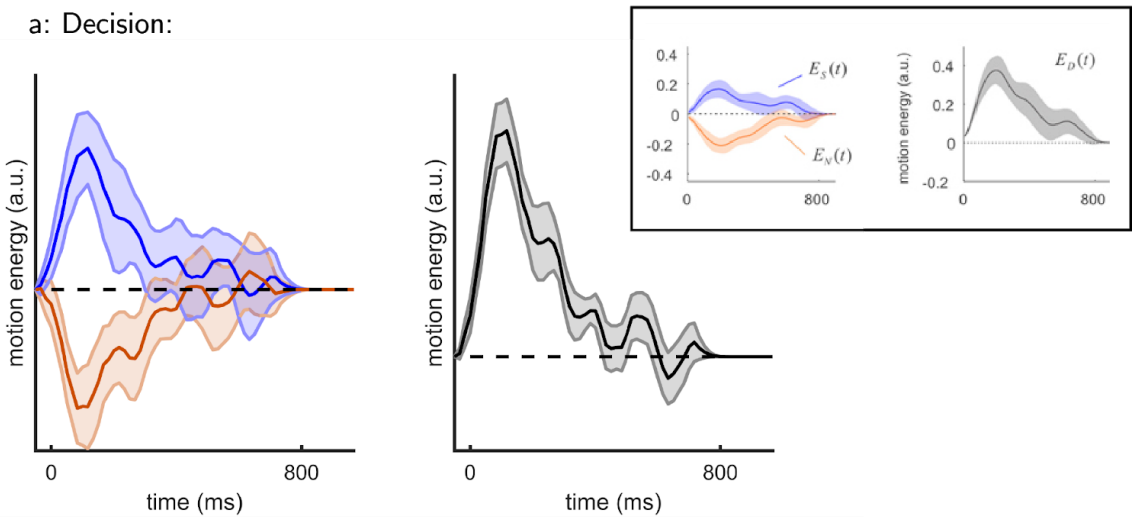
Experiment 2: pre-registration

Objectives

1. Replicate the positive evidence bias in discrimination: stronger effect of positive compared to negative evidence on decision confidence (Experiment 1, Zylberberg et al., 2012; Experiments 1 and 2).



2. Replicate the absence of a positive evidence bias in detection judgments: similar effects for negative and positive evidence on decision confidence, revealed by pseudo-discrimination analysis (Experiment 1).
3. Replicate the absence of an effect for positive or negative evidence on confidence in 'no' judgments, as revealed by pseudo-discrimination analysis and reverse correlation analysis of correct rejection trials (Experiment 1).

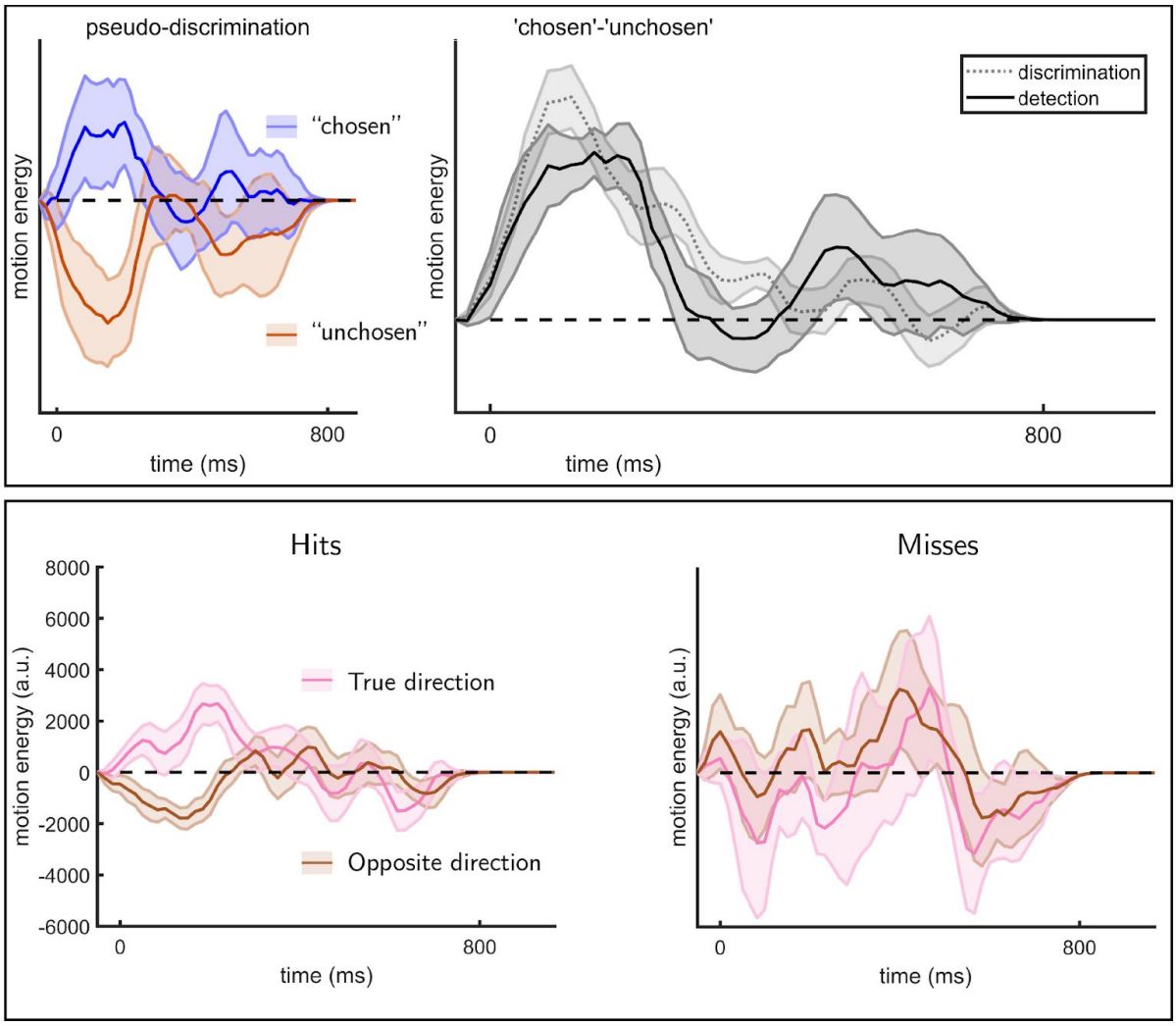


Figure 1: Decision (top panel) and confidence (lower panel) reverse correlation kernels for detection, using pseudo-discrimination analysis on detection judgments (Experiment 1).

Materials and Methods

Participants

Participants will be recruited through Prolific. We will aim for a net sample size of 100 participants who will meet our exclusion criteria for both tasks. Task-wise analysis will be applied to the data of all participants that met our task-specific inclusion criteria.

Experimental Paradigm

The experiment will consist of two tasks (Detection and Discrimination) presented in separate blocks. A total of 56 trials of each task will be delivered in 2 blocks of 28 trials each (~5 seconds per trial; ~2 minutes and 20 seconds per block). The order of experimental blocks will be interleaved, starting with discrimination.

The first discrimination block will start after an Introduction section. The introduction will include instructions about the stimuli and confidence scale, four practice trials and four confidence practice trials (~3 minutes). A second introduction section will be presented before the second block (first detection block; ~2 minutes). Introduction sections will be followed by multiple-choice comprehension questions, to monitor participants' understanding of the main task and confidence reporting interface. To encourage concentration, feedback will be given at the end of the second and fourth blocks about overall performance and mean confidence in the task. Overall, the total experiment is expected to take about 20 minutes:

	Duration
Introduction (1+2)	3 minutes
112 trials in 4 blocks	15 minutes
Inter-block feedback and rest	2 minutes
Total	**20 minutes**

Unlike the lab-based experiment, there will be no calibration of difficulty for the two tasks. The rationale for this is that in Experiment 1 participants' perceptual thresholds for motion discrimination were highly similar, and staircasing took a long time to converge. Furthermore, in

order to directly compare detection and discrimination and complement our findings from Experiment 1, luminance will be set to be equal for the two tasks, such that discrimination performance is expected to be higher overall.

The experiment will terminate after 2 blocks if accuracy for both of the tasks will fall below 55%, or after an incorrect response for one of the two multiple choice questions. In such an occasion, data will not be used for the main analysis.

Trial Structure

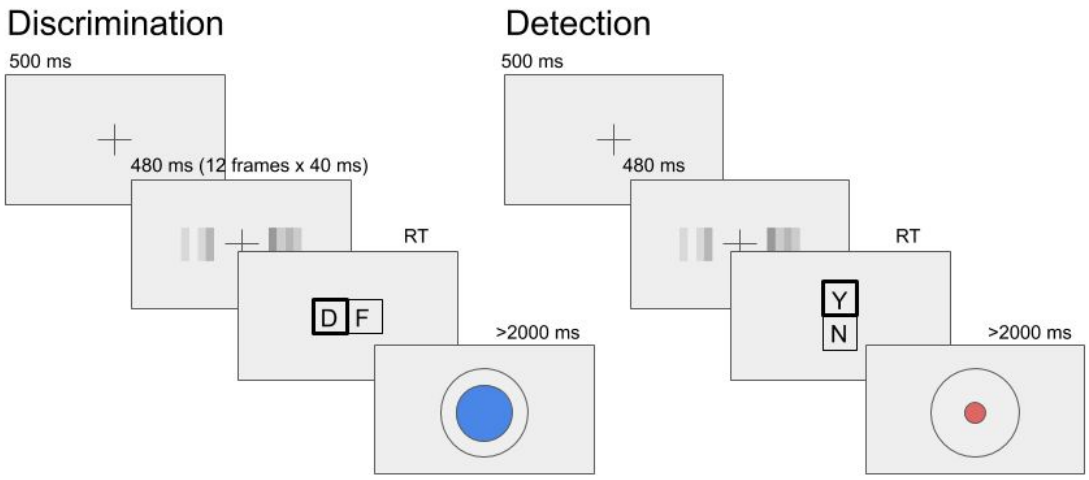


Figure 2: Trial strucure for the discrimination and detection tasks.

Discrimination

Trial structure will closely follow Experiment 2 from Zylberberg and colleagues (2012), with a few adaptations. Following a fixation cross (500 ms), a rapid serial visual presentation (RSVP) will be presented (12 frames, presented at 25Hz), consisting of two sets of four adjacent vertical gray bars, displayed to the left and right of the fixation cross. On each frame, the luminance of the bars will be randomly sampled from a Gaussian distribution with a standard deviation of 10/255 units in the standard RGB 0-255 coordinate system. The average luminance of one set of bars will be that of the background (128/255). The average luminance of the other set will be 133/255, making this patch brighter on average. Participants will then report which of the two sets was brighter on average than the gray background using the ‘D’ and ‘F’ keys on the keyboard. After their response, they will rate their confidence on a continuous scale, by controlling the size of a colored circle with their mouse. High confidence will be indicated by

making the circle bigger, and low confidence will be indicated by making it smaller. To discourage hasty confidence ratings, the confidence rating scale will stay on the screen for at least 2000 milliseconds. Feedback about response accuracy will be delivered after the confidence rating phase.

Detection

Detection trials will be similar to discrimination trials, except that decisions will be made about whether the average luminance of either of the two sets was brighter than the gray backgroud, or not. In ‘different’ trials, luminance of the four bars in one of the sets will be sampled from a Gaussian distribution with mean 133/255, and luminance of the other set will be sampled from a Gaussian distribution with mean 128/255. In ‘same’ trials, luminance of both sets will be sampled from a distribution centered at 128/255. Decisions in Detection trials will be reported using the ‘y’ and ‘n keys (‘y’ for ‘yes’ and ‘n’ for ‘no’). Confidence ratings and feedback will be as in the discrimination task.

Analysis

1. Replicate the positive evidence bias in discrimination. This will be achieved by performing reverse correlation analysis on discrimination confidence ratings and luminance of the chosen and unchosen sets, following the procedure described in Zylberberg and colleagues (2012).
   Power: using a similar paradigm on 5 participants with 960 trials each, Zylberberg obtained a large effect size of d=1.4 (t(4)=2.99). Our within-participant measures will be much noisier with 56 instead of 960 trials per participant.
2. Replicate the absence of a positive evidence bias in detection judgments. This will be achieved by performing reverse correlation analysis on detection confidence ratings and luminance of the chosen and unchosen sets, using pseudo-discrimination analysis. In this analysis, we will assume that in the majority of ‘different’ trials, when participants responded ‘yes’ they correctly identified the brighter set. For example, a detection trial in which the brighter set was presented on the right and in which the participant responded ‘yes’ will be treated as a discrimination trial in which the participant responded ‘right’. Conversely, a trial in which the brighter set was presented on the right and in which the participant responded ‘no’ will be treated as a discrimination trial in which the participant responded ‘left’. These hypothetical responses will then be submitted to the same reverse correlation analysis described in the previous section confidence kernels.
   Power: We are predicting a null effect for this analysis. As a result, our ability to draw conclusions based on this contrast will depend on the obtained positive evidence bias effect size for discrimination.

3. Replicate the absence of an effect for positive or negative evidence on ‘no’ judgments.
   This will be achieved by performing reverse correlation analysis on detection confidence ratings and luminance of the chosen and unchosen sets, using pseudo-discrimination analysis on miss trials, and by applying reverse correlation analysis to correct rejection trials.
   Power: Similarly here we are interested in finding a null effect. We will have approximately 28 ‘no’ trials for each participant, so we want to be able to show that with a similar number ‘yes’ trials responses we were able to find an effect.

Participant Exclusion

Participant exclusion will be decided separately for the discrimination and detection tasks. Task-wise exclusion will be applied to participants that:

1. Perform with accuracy of below 55%.
2. Use the maximum confidence rating in more than 70% of the trials.

In addition, we will exclude participants from both tasks if they fail one or more of our multichoice task comprehension checks. Data collection will terminate once we obtain full datasets (detection and discrimination) from 100 participants.

Zylberberg, A., Barttfeld, P., & Sigman, M. (2012). The construction of confidence in a perceptual decision. *Frontiers in integrative neuroscience*, *6*, 79.