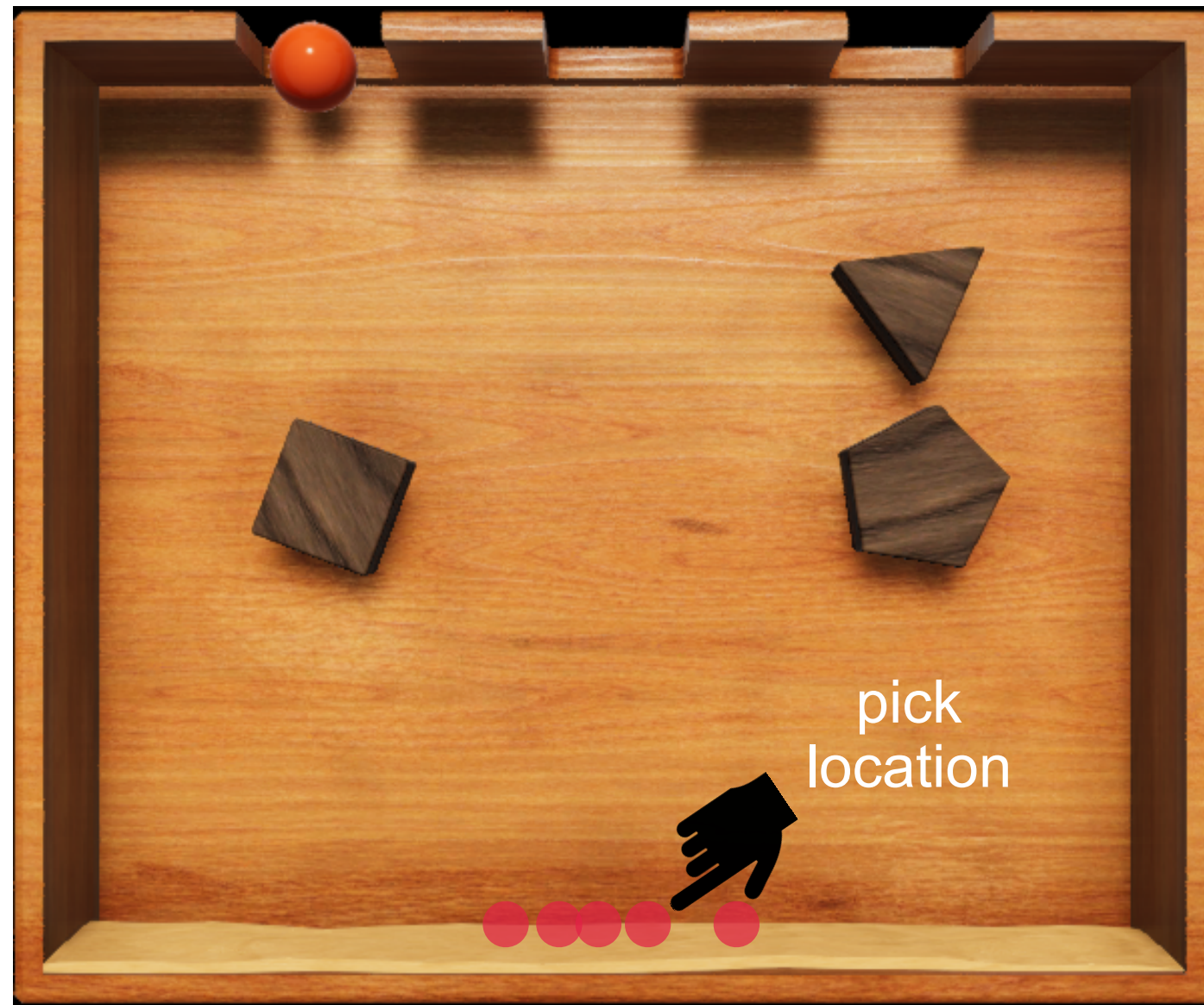


A Experiment 1: Prediction

Where will the ball land?



B

Experiment 2: Inference

Where was the ball dropped from?

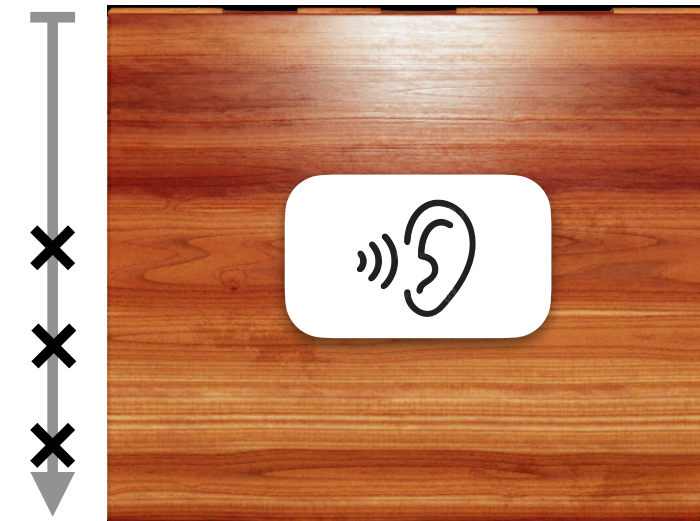
No Sound + Ball Visible



1 2 3



Sound + Ball Visible



1 2 3



Sound + Ball Occluded



1 2 3

