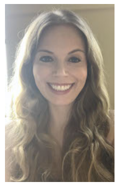
The importance of causal inference in multisensory learning

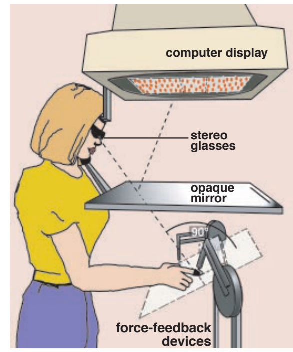
Michael S. Landy
NYU

Fangfang Hong



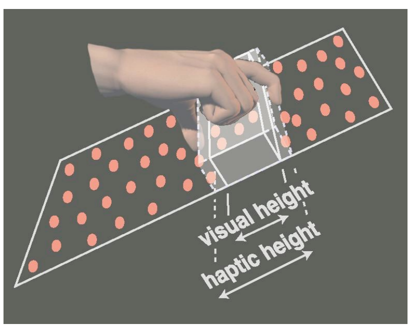
Stephanie Badde

“Traditional” cue integration



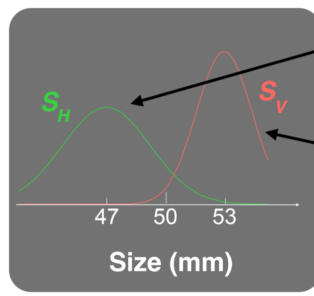
Maloney & Landy, 1989; Clark & Yuille, 1990; Landy et al., 1995; Ernst & Banks, 2002; etc.

“Traditional” cue integration



Maloney & Landy, 1989; Clark & Yuille, 1990; Landy et al., 1995; Ernst & Banks, 2002; etc.

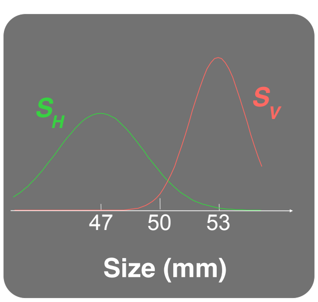
“Traditional” cue integration



$p(m_H | s_H)$

$p(m_V | s_V)$

“Traditional” cue integration



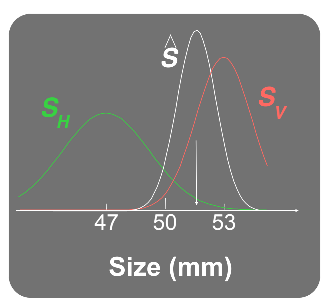
$$\hat{s} = w_H m_H + w_V m_V$$

$$w_H = \frac{r_H}{r_H + r_V}$$

$$w_V = \frac{r_V}{r_H + r_V}$$

$$\text{Reliability } r = \frac{1}{\sigma^2}$$

“Traditional” cue integration



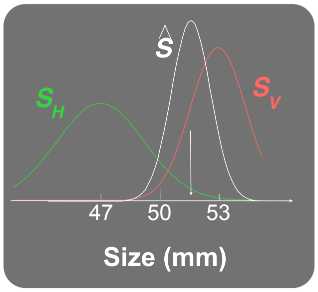
$$\hat{s} = w_H m_H + w_V m_V$$

$$w_H = \frac{r_H}{r_H + r_V}$$

$$w_V = \frac{r_V}{r_H + r_V}$$

$$\text{Reliability } r = \frac{1}{\sigma^2}$$

“Traditional” cue integration



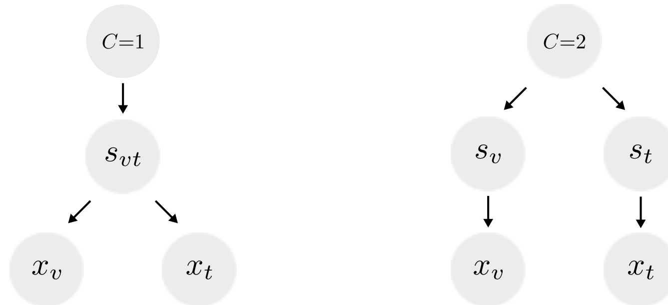
$$\hat{s} = w_H m_H + w_V m_V$$

$$\sigma^2_{HV} = \frac{\sigma^2_H \sigma^2_V}{\sigma^2_H + \sigma^2_V}$$

$$r_{HV} = r_H + r_V$$

Variance of combined estimate is lower than variance of either single-cue estimate

Causal inference



Körding et al., 2007

Ventriloquism and its aftereffect

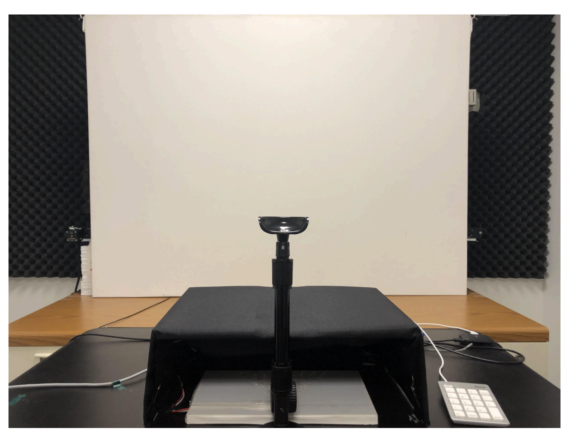
Ventriloquism and its aftereffect

- Multisensory integration
- Causal inference
- Bayesian modeling of integration (e.g., Körding et al., 2007)
- Recalibration from consistent discrepancy

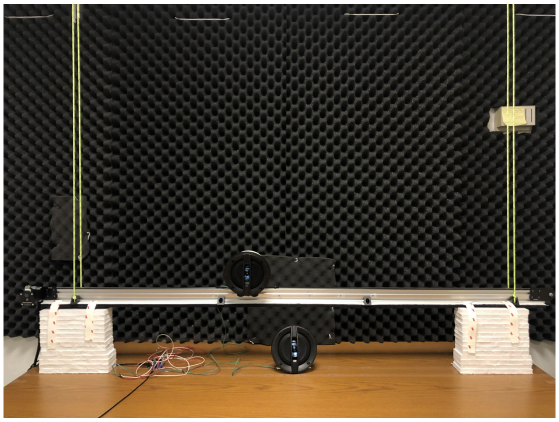
1. Models of multisensory recalibration

- Reliability-based: Several studies find that vision dominates for parameters for which it is more reliable, but not when it is the less-reliable cue. Burge, Girshick & Banks (2010) tested and confirmed a reliability-based model for vision and haptic cue integration
- Fixed-ratio: Zaidel, Turner & Angelaki (2011) found that a fixed-ratio model fit best for visual-vestibular integration, and argued that a statistical flaw led to Burge et al.’s conclusions
- Ernst & di Luca (2011) suggested a “coupling prior” based on the tendency for cues to become miscalibrated

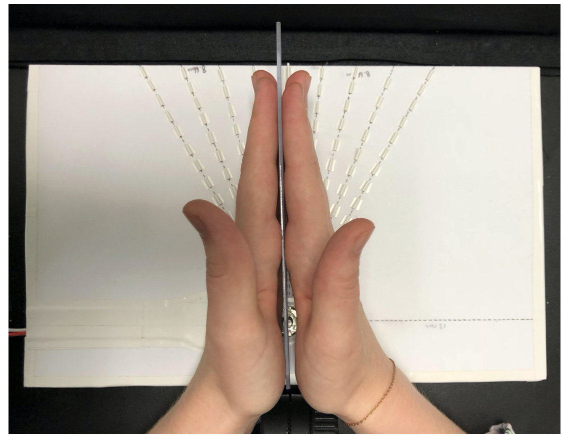
Apparatus: Projection screen



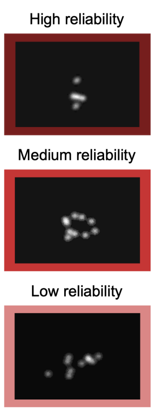
Apparatus: Movable speaker



Apparatus: Pointing device



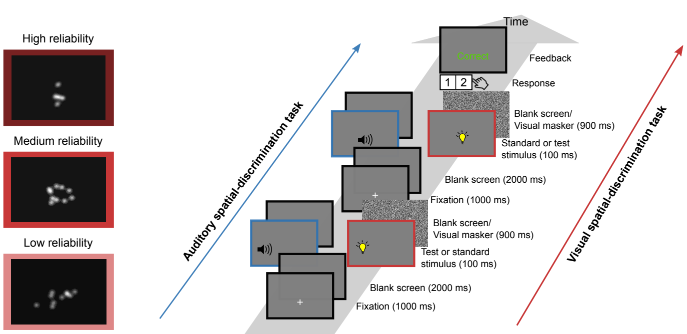
Visual stimuli with varied reliability



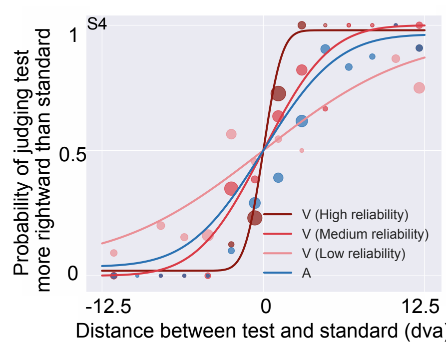
Experimental tasks/sessions

1. Unimodal spatial discrimination: To measure JNDs
2. Bimodal spatial discrimination: To measure bias
3. Pointing practice: For practice and to measure motor noise)
4. Recalibration. Sessions: 2 directions x 3 visual reliabilities
   - Pre-recalibration phase: Unimodal spatial localization
   - Recalibration phase: Bimodal localization task
   - Post-recalibration phase: Same as pre-recalibration

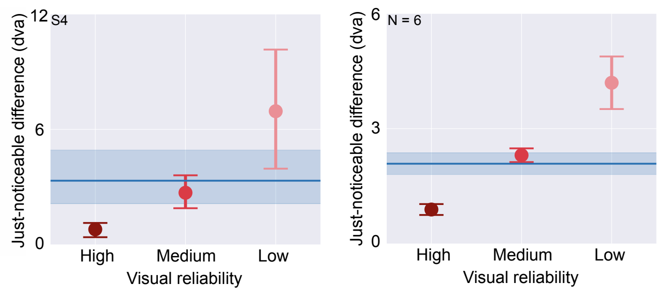
Phase 1: Unimodal spatial discrimination



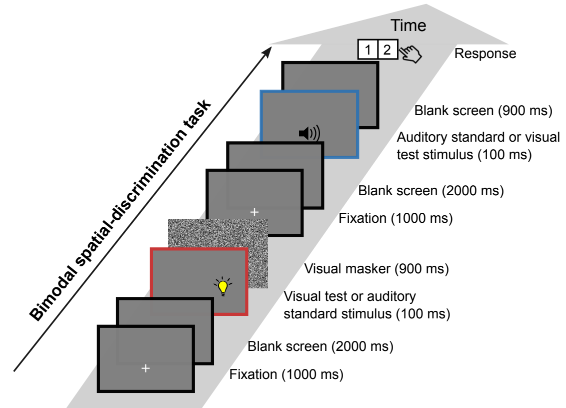
Phase 1: Unimodal spatial discrimination



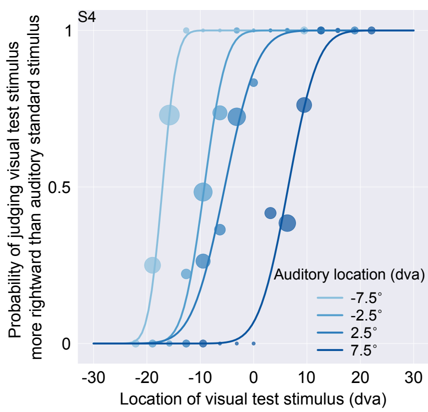
Phase 1: Unimodal spatial discrimination



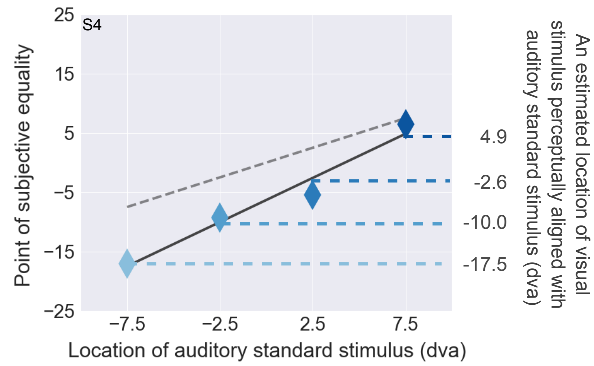
Phase 2: Bimodal spatial discrimination



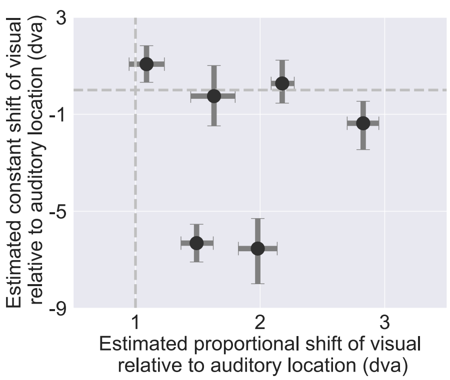
Phase 2: Bimodal spatial discrimination



Phase 2: Bimodal spatial discrimination



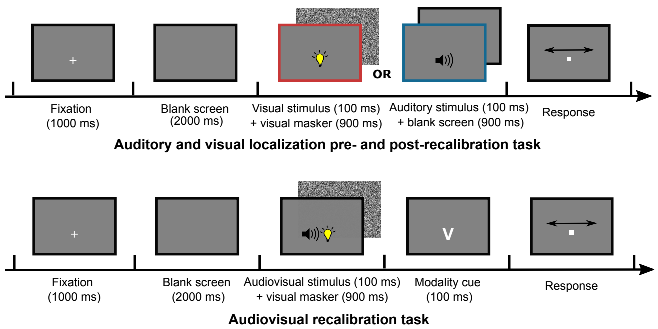
Phase 2: Bimodal spatial discrimination



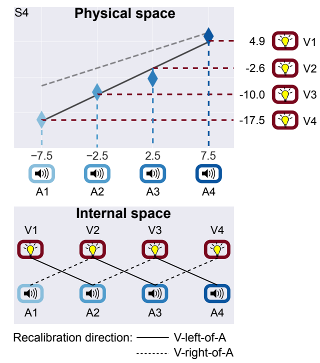
Phase 3: Pointing practice

- Provides practice with the pointing device
- Visual feedback was provided (else very noisy and likely biased)
- Provides estimates of motor noise for subsequent localization tasks: $\bar{\sigma}_r = 1.85$ deg

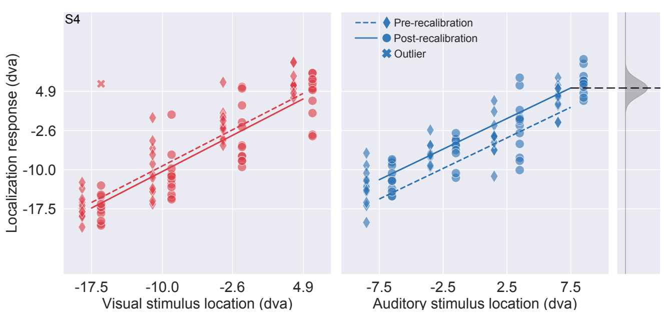
Main experiment: Recalibration



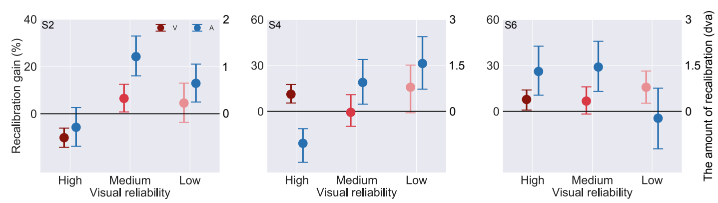
Main Experiment: Recalibration



Main Experiment: Recalibration



Main Experiment: Recalibration



The story so far

- No or tiny amount of visual recalibration. This is consistent with previous studies, although for them, visual reliability was typically higher than auditory
- Significant auditory recalibration
- Amount of auditory recalibration depends on relative reliability
- That dependence varies across participants; it can rise, fall, or rise then fall with increasing visual reliability
- What can account for these varying patterns of recalibration?

Modeling

We compare three models

- Reliability-based
- Fixed-ratio
- Causal-inference

We fit the models to the data from

- Unimodal discrimination
- Bimodal discrimination
- Pre- and post-recalibration bimodal localization

Modeling

Common model aspects

- Parameters for initial bias, assumed (without loss of generality) to be bias in the auditory signals
- Cumulative shifts from recalibration
- Thus, $s'_A = a_A s_A + b_A + \Delta_A$ and $s'_V = s_V + \Delta_V$
- These are corrupted by internal noise, e.g., $m'_A = s'_A + \varepsilon_A$
- Measurement reliability may differ between uni- and bimodal conditions
- Single lapse rate for discrimination, late noise for localization estimated from pointing practice
- Single supra-modal prior (very wide)

Reliability-based model

- Updates based on measurement discrepancy
- Updates proportional to the other modality's reliability
- Thus,

$$\Delta_A(t+1) = \Delta_A(t) + \alpha w_A \left(m'_V - m'_A\right)$$

and

$$\Delta_V(t+1) = \Delta_V(t) + \alpha w_V \left(m'_A - m'_V\right),$$

where

$$w_A = \frac{{\sigma'_{AV,V}}^{-2}}{{\sigma'_{AV,V}}^{-2} + {\sigma'_{AV,A}}^{-2}} = \frac{r_{AV,V}}{r_{AV,V} + r_{AV,A}}$$

$$w_V = 1 - w_A$$

Fixed-ratio model

- Updates in fixed proportion based on modality-specific learning rate
- Thus,

$$\Delta_A(t+1) = \Delta_A(t) + \alpha_A \left(m'_V - m'_A\right)$$

and

$$\Delta_V(t+1) = \Delta_V(t) + \alpha_V \left(m'_A - m'_V\right)$$

Causal-inference model

- Updates based on difference between measurement and estimate
- Thus,

$$\Delta_A(t+1) = \Delta_A(t) + \alpha_A \left(\hat{s}'_A - m'_A\right)$$

and

$$\Delta_V(t+1) = \Delta_V(t) + \alpha_V \left(\hat{s}'_V - m'_V\right)$$

- Also tested single-rate model, where $\alpha_A = \alpha_V$

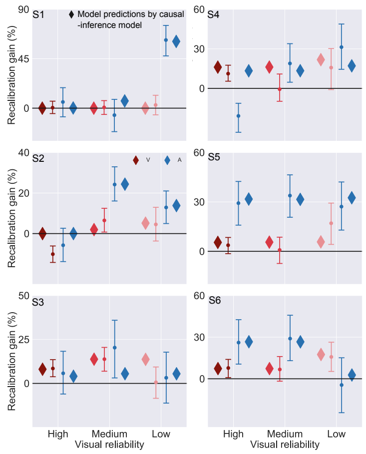
Causal-inference model

- Estimates are based on standard causal-inference integration model using model averaging
- Two scenarios: $C = 1$ and $C = 2$
- $\hat{s}'_{A,C=1} = \hat{s}'_{V,C=1}$ based on reliability-based integration of both measurements and the supra-modal prior
- $\hat{s}'_{A,C=2}$ based on reliability-based integration of auditory measurement and the supra-modal prior, and similarly for $\hat{s}'_{V,C=2}$
- $\hat{s}'_A$ and $\hat{s}'_V$ average the estimates for the two scenarios, weighted by their respective probabilities given the measurements

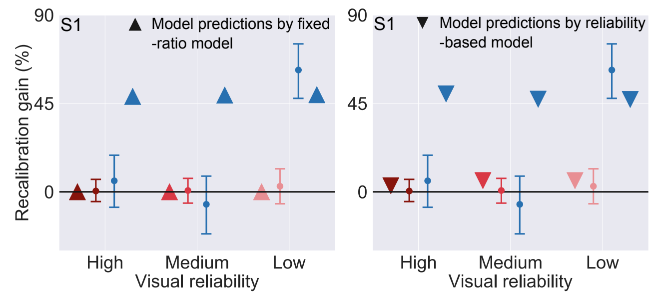
Free parameters for model fits

Θ	**Meaning**	**RB**	**FR**	**CI**	$\mathbf{CI}_{\alpha_V=\alpha_A}$
$\sigma_{AV,A}$	Measurement noise variance for the auditory stimulus in bimodal trials	✓	✓	✓	✓
σ_{AV,V_i}	Measurement noise variance for the visual stimulus in bimodal trials	✓	✓	✓	✓
μ'_P	The mean of the supra-modal prior distribution	-	-	-	-
σ'_P	The standard deviation of the supra-modal prior	-	-	-	-
a_A, b_A	The slope and the intercept of the linear function that captures biases in auditory measurements	✓	✓	✓	✓
$P(C=1)$	Prior probability of a common cause	-	-	✓	✓
α_A, α_V	Modality specific learning rate	-	✓	✓	-
α	Common learning rate	✓	-	-	✓
λ_{AV}	Lapse rate for the bimodal spatial-discrimination task	✓	✓	✓	✓
	Number of total parameters:	8	9	10	9

Model fit results: Causal-inference model



Failure to fit: Other models



Model comparison

Δ_{AIC}

Participant	RB	FR	CI	$CI_{\alpha_A=\alpha_V}$
S1	60	48.8	0	9
S2	15.3	16.8	0	20.1
S3	13.5	14.8	2.4	0
S4	13.6	5.2	1.6	0
S5	26.3	31	1.9	0
S6	0	6	2.8	1.7

Model

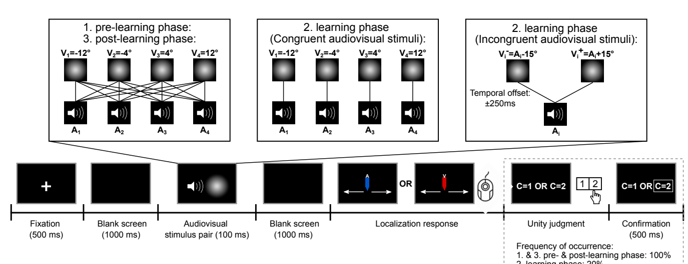
2. Is the common-cause prior flexible?

- The causal-inference model estimates location by first making a common-cause inference, which in turn depends on the prior probability of a common cause.
- Is that common-cause prior fixed, or can it adapt to current conditions?
- We test this by a similar experimental design in which the learning phase involves either a set of fully congruent or fully discrepant audiovisual stimuli.

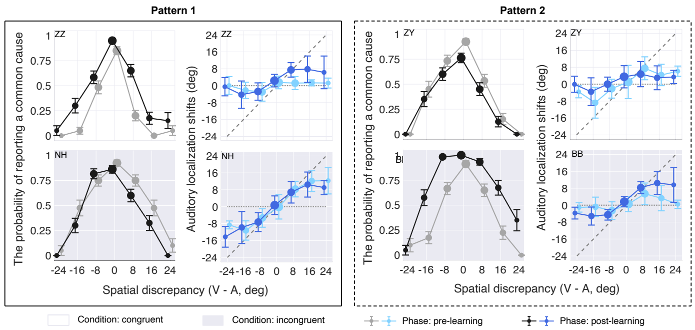
Experimental tasks/sessions

1. Unimodal spatial discrimination: To measure JNDs
2. Bimodal spatial discrimination: To measure bias
3. Pointing practice: For practice and to measure motor noise)
4. Learning. Two sessions: congruent vs. incongruent stimuli
   - Pre-recalibration phase: Unimodal spatial localization
   - Learning phase: Bimodal localization task
   - Post-recalibration phase: Same as pre-recalibration

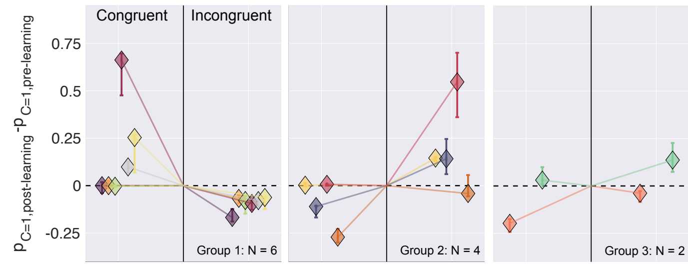
Main experiment: Prior Learning



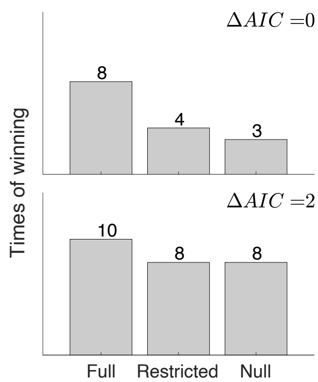
Main experiment: Sample results



Main experiment: Data patterns



Main experiment: Model comparison



Conclusions

- A causal-inference-based model of cue recalibration can explain the diverse patterns of recalibration with varying cue reliability across participants
- The prior for a common cause is flexible, adapting to recent conditions
- Given the presence of biases across cues, biases must be measured and taken into account in understanding multisensory integration
- Predictions are more clear-cut if stimuli across modalities are equated for these biases