Data files

Overview

This directory contains raw experimental data, organized data, Wishart model fits and predictions, as well as calibration and transformation matrices that allow data to be transformed from the model space to other color spaces (e.g., RGB and DKL).

Broadly, the files fall into three categories:

1. **Organized data and model predictions**
    - `Noise_ellipses_sub#.csv`: Model-predicted internal noise ellipses.
    - `Thres_ellipses_sub#.csv`: Model-predicted threshold ellipses at 66.7% correct performance.
    - `Bestfit_W_sub#.csv`: Best-fit Wishart model weight matrices.
    - `trial_data_pooled_by_type_sub#.csv`: Trial-level data pooled across sessions and organized by trial type (AEPsych, MOCS, and pre-generated Sobol trials).
    - `thres_at_validation_conditions_sub#.csv`: Estimated thresholds using validation data.
    - `thres_at_validation_conditions_WPPM_sub#.csv`: Model-predicted thresholds at the same validation conditions, obtained from Wishart model fits.
2. **Raw experimental data**
    - `Unity_trial_data_sub#_##_session#__YYYY-MM-DD_HH-MM-SS.csv`: Trial-by-trial raw data recorded during the experiment, including stimulus values, responses, and timing. They were not sorted by trial type or conditions.
    - `sub#_##_session#_YYYY-MM-DD_HH-MM-SS.txt`: Dual-computer communication logs.
    - `trial_status_sub#_session#.txt`: Logs describing the execution history of each trial, including delays, reordering, and advancement due to fallback strategies.
3. **Calibration and transformation**
    - `M_2DWToDKLPlane_DELL_MM-DD-YYYY.csv`, `M_2DWToRGB_DELL_MM-DD-YYYY.csv`, `M_DKLPlaneTo2DW_DELL_MM-DD-YYYY.csv`, `M_RGBTo2DW_DELL_MM-DD-YYYY.csv`: Transformation matrices between the model space and other color representations.

Notably, all stimulus values are stored in the model space, bounded between −1 and 1. Conversion to other color spaces requires applying the corresponding transformation matrices.

Organized data and model predictions

▼ **Noise ellipses**

`Noise_ellipses_sub#.csv`

This file contains ellipses that characterize the internal noise limiting color perception, evaluated on a finely sampled grid of reference stimuli. We assume that internal noise follows a multivariate Gaussian distribution centered on the true color stimulus, with the noise structure specified by a 2 x 2 covariance matrix.
The file first includes model-estimated noise ellipses derived from the fit to the original dataset, followed by noise ellipses obtained from 120 bootstrapped datasets. The bootstrapped results are sorted in descending order of similarity to the original dataset, quantified by the Normalized Bures Similarity (NBS) score summed across a finely sampled grid.
For computation and visualization of 95% confidence intervals, bootstrapped datasets with NBS scores in the lowest 5% were excluded.

Column name	Description	Values / Notes
`grid_ref_fine`	Finely sampled reference coordinates in the 2D model space	Linearly spaced from −0.85 to 0.85 with 103 steps per dimension, chosen to exactly include the coarse grid `grid_ref` (−0.7 to 0.7, 7 steps per dimension) in `Thres_ellipses_sub#.csv`
`Sigmas_noise_grid_org`	Covariance matrices characterizing noise ellipses from the WPPM fit to the original dataset	$\left[[\sigma^2_{dim1}, \sigma_{dim1,dim2}], [\sigma_{dim1,dim2}, \sigma^2_{dim2}]\right]$; derived from the original dataset
`Sigmas_noise_grid_btst#_rank#`	Covariance matrices characterizing noise ellipses from WPPM fits to bootstrapped datasets	`btst#` indexes the bootstrap replicate; `rank#` indexes the descending NBS rank (0 = highest similarity to the original dataset; 119 = lowest). Indices `rank0` - `rank113` correspond to the 95% confidence interval

▼ Threshold ellipses

`Thres_ellipses_sub#.csv`

This file contains ellipses that characterize color discrimination threshold contours at 66.7% correct performance, evaluated on a coarse grid of reference stimuli. Each ellipse is represented by a 2 × 2 covariance matrix, which defines the corresponding threshold contour.
The file first includes model-predicted threshold ellipses derived from the fit to the original dataset, followed by ellipses obtained from 120 bootstrapped datasets. The bootstrapped results are sorted in descending order of similarity to the original dataset, using the same NBS-based ordering as in `Noise_ellipses_sub#.csv`.

Column name	Description	Values / Notes
`grid_ref`	Reference coordinates in the 2D model space	Linearly spaced from −0.7 to 0.7 with 7 steps per dimension

`Sigmas_thres_grid_org`	Covariance matrices characterizing threshold ellipses from the WPPM fit to the original dataset	$\left[\left[\sigma^2_{dim1}, \sigma_{dim1,dim2}\right], \left[\sigma_{dim1,dim2}, \sigma^2_{dim2}\right]\right]$; derived from the original dataset
`Sigmas_thres_grid_btst#_rank#`	Covariance matrices characterizing threshold ellipses from WPPM fits to bootstrapped datasets	`btst#` indexes the bootstrap replicate; `rank#` indexes the descending NBS rank (0 = highest similarity to the original dataset; 119 = lowest)

▼ Best-fit weight matrix

`Bestfit_W_sub#.csv`

> 💡 This file contains the model-estimated weight matrices applied to the 5 × 5 basis functions. The original weight array has shape 5 × 5 × 2 × 3 and is unrolled into a tabular format while preserving the indices of all four dimensions.
> The file first includes the best-fit weight values from the original dataset, followed by the corresponding weight values from 120 bootstrapped datasets. The bootstrap ordering is consistent with `Noise_ellipses_sub#.csv` and `Thres_ellipses_sub#.csv`.

Column name	Description	Values / Notes
`i,j,k,l`	Indices for the four dimensions of the weight matrix	$i \in \{0,1,2,3,4\}$, $j \in \{0,1,2,3,4\}$, $k \in \{0,1\}$, $l \in \{0,1,2\}$
`W_org`	Weight value at index `(i, j, k, l)`	Estimated from the original dataset
`W_btst#_rank#`	Weight value at index `(i, j, k, l)` based on bootstrapped datasets	`btst#` indexes the bootstrap replicate; `rank#` indexes the descending NBS rank (0 = highest similarity to the original dataset; 119 = lowest)

▼ Trial-by-trial processed data

`trial_data_pooled_by_type_sub#.csv`

> 💡 This file contains AEPsych trials, validation trials (method of constant stimuli; MOCS), and pre-generated Sobol trials that were inserted as needed, pooled across sessions. The trial data are grouped by trial type. To retrieve trial data ordered by the actual execution sequence, see `Unity_trial_data_sub#_##_session#__YYYY-MM-DD_HH-MM-SS.csv`.

Column name	Description	Values / Notes
`TrialType`	Trial type and type-specific identifier	Format: `TYPE_i_COND_j`, where `TYPE ∈ {AEPsych, MOCS, PregenSobol}` and `i` is the type-specific index. For `AEPsych` and `MOCS`, `i ∈ {0, ..., 5999}`; for `PregenSobol`, `i` is subject-specific. `COND` denotes the trial condition. For `AEPsych`, `COND` can either be Sobol or adaptive; for `MOCS`, there were 25 different reference

		conditions. `PregenSobol` trials do not include a `COND` field.
`xref`	Values of the reference stimulus in 2D model space	The values were restricted to be in the range `[-0.75, 0.75]`
`x1`	Values of the comparison stimulus in 2D model space	The values were restricted to be in the range `[-1, 1]`
`y`	Response	`1`: correct identification of the odd stimulus `0`: incorrect response

▼ Thresholds at validation conditions

`thres_at_validation_conditions_sub#.csv`

This file contains estimated discrimination thresholds at the validation conditions. There were a total of 25 validation conditions, each paired with 12 comparison stimuli selected using the method of constant stimuli (MOCS). Each reference–comparison stimulus pair was tested 20 times, resulting in 240 trials per validation condition.
For each validation condition, we fitted a Weibull psychometric function with the guess rate fixed at 1/3. The threshold was defined as the stimulus level yielding 66.7% correct performance.
To estimate uncertainty, we performed 120 bootstrap resamples. For each bootstrapped dataset, the number of trials at each stimulus level was held constant at 20, matching the original experimental design, and thresholds were re-estimated accordingly.
Confidence intervals were computed by sorting the bootstrapped threshold estimates in ascending order and taking the central 95% of the resulting distribution.

Column name	Description	Values / Notes
`xref`	Reference stimulus in the 2D model space	Reference conditions were generated using a Sobol sequence with a subject-specific seed to ensure reproducibility
`x1_thres_org`	Comparison stimulus value at the 66.7% correct threshold	Estimated from the fitted psychometric function
`L2norm_thres_org`	Euclidean distance between the reference and comparison stimuli at threshold (original dataset)	Computed as $\|\|x_1 - x_{ref}\|\|_2$ using the threshold estimated from the original dataset
`L2norm_thres_btst#`	Euclidean distance between the reference and comparison stimuli at threshold (bootstrapped datasets)	`btst#` indexes the bootstrapped dataset

`thres_at_validation_conditions_WPPM_sub#.csv`

💡 This file contains the **same columns** as `thres_at_validation_conditions_sub#.csv`. The key differences are:

1. Thresholds were predicted using the fitted Wishart model
2. Confidence intervals were determined using the same NBS-based ordering as in `Noise_ellipses_sub#.csv`

Accordingly, instead of columns named `L2norm_thres_btst#`, this file uses `L2norm_thres_btst#_rank#`, where `rank#` indexes bootstrap replicates ordered by descending NBS similarity to the original dataset. Indices `rank0` – `rank113` correspond to the 95% confidence interval.

Raw experimental data

▼ **Trial-by-trial raw data**

`Unity_trial_data_sub#_##_session#__YYYY-MM-DD_HH-MM-SS.csv`

💡 This file records the surface colors of the reference and comparison stimuli presented during the experiment, along with the location of the odd stimulus (generated using participant- and session-specific seeds), the participant's response, and response time.

Column name	Description	Values / Notes
`TrialCounter`	Global trial index across the entire session	Integer. Starts at `0`. Ranges from `0–1199` for `sub1`, and `0–999` for other subjects (per session).
`TrialInfo`	Predetermined trial sequence number and trial-type–specific sequence number	Format: `Trial_i_TYPE_j` Scheduled trial with global index `i`, trial type (`AEPsych`, `MOCS`, or `Sobol`), and type-specific index `j`. For `AEPsych` and `MOCS`, `i` ranges from `0–999` and `j` ranges from `0–499`. For pre-generated `Sobol`, `i` starts at `1000` and `j` starts at `0`.
`OddLocation`	Location of the odd (comparison) stimulus	`1`: top `2`: bottom left `3`: bottom right
`PressedKey`	Participant's selected odd stimulus	`1_0_0`: top `0_1_0`: bottom left `0_0_1`: bottom right
`Ref`	Linear RGB values of the reference stimulus	`[R, G, B]`, normalized to the range `[0, 1]`
`Comp`	Linear RGB values of the odd comparison stimulus	`[R, G, B]`, normalized to the range `[0, 1]`
`Ref_Corrected`	Gamma-corrected RGB values of the reference stimulus	`[R, G, B]` after LUT-based gamma correction
`Comp_Corrected`	Gamma-corrected RGB values of the comparison stimulus	`[R, G, B]` after LUT-based gamma correction

`binaryResp`	Response	`1`: correct identification of the odd stimulus `0`: incorrect response
`exptITI`	Inter-trial interval	Duration in seconds
`RT`	Response time	Duration in seconds

▼ **Dual-computer communication log**

`sub#_##_session#_YYYY-MM-DD_HH-MM-SS.txt`

> 💡 The experiment was run using a dual-computer setup.
> This file logs all communication events between a high-performance PC (name: sas-mantaray-md.psych.upenn.edu), which handled trial placement computations, and a second computer responsible for stimulus presentation in Unity. The two computers communicated via a shared text file accessible to both machines throughout the experiment.

Example lines:

```
Subject ID: 1
Subject initial: CH
Session: 10
Date and Time: 2025-02-27_18-32-00
```

Subject ID, subject initials, session number, and session start timestamp

```
2025-02-27 18:32:00 - sas-mantaray-md.psych.upenn.edu: Set_Up_to_Communicate
2025-02-27 18:32:22 - Unity: Ready_To_Communicate
```

Initial handshake between the two computers

```
2025-02-27 18:32:42 - sas-mantaray-md.psych.upenn.edu: Trial_0_MOCS_0 Ref_R0.79644970_G0.36802668_B0.75094980 Comp_R0.91060940_G0.33655663_B0.71020740 Image_Display
```

Computer A sent the trial information and the linear RGB values of the stimuli

```
2025-02-27 18:32:50 - Unity: Trial_0_MOCS_0 Ref_R0.79644970_G0.36802670_B0.75094980 Comp_R0.91060940_G0.33655660_B0.71020740 Resp1 Image_Confirmed
```

Computer B confirmed stimulus presentation and returned the registered participant response

▼ **Trial status log**

`trial_status_sub#_session#.txt`

This file logs the status of each trial. Within each session, an equal number of adaptive (AEPsych) trials and validation trials (method of constant stimuli; MOCS) were scheduled. MOCS trials were first pseudo-randomized, then paired one-to-one with AEPsych trials, and finally shuffled within each pair. This procedure produced a predetermined trial sequence.
Because AEPsych trial placement is computed online, delays occasionally occurred. To keep the experiment running, the next available MOCS trial was advanced when an AEPsych trial was not yet ready. If the delay persisted, additional MOCS trials were successively advanced. To limit imbalance between trial types, the number of MOCS trials allowed to precede AEPsych trials was capped at four. If this cap was reached and AEPsych computation was still incomplete, pre-generated Sobol trials were temporarily inserted.
Each log entry records the execution history of a scheduled trial. Tokens following the trial identifier describe how the trial was executed or deferred due to AEPsych computation delays, as well as the total elapsed waiting time before completion.

Log token definitions:

- `Trial_i_TYPE_j`: Scheduled trial with global index `i`, trial type (`AEPsych`, `MOCS`, or `Sobol`), and type-specific index `j`. For `AEPsych` and `MOCS`, `i` ranges from `0–999` and `j` ranges from `0–499`. For pre-generated `Sobol`, `i` starts at `1000` and `j` starts at `0`.
- `Completed_in_time`: Trial executed as scheduled, without delay or reordering.
- `Stepped_down_to_MOCS_k`: A scheduled `AEPsych` trial was deferred, and the `MOCS` trial with index `k` was executed instead.
- `Stepped_up_to_AEPsych_k`: An `AEPsych` trial with index `k` was executed during a slot originally scheduled for a `MOCS` trial.
- `Insert_Sobol_k`: A pre-generated `Sobol` trial with index `k` was executed because AEPsych computation was still incomplete and the maximum lead of four `MOCS` trials ahead of `AEPsych` trials had been reached.
- `Elapsed_time_t`: Time (in seconds) spent waiting for AEPsych trial placement to complete.
- `Completed`: Trial execution finished; the trial may have been executed earlier or later than scheduled.
- `Skipped`: Scheduled trial was skipped because it had already been completed earlier in the session.

Example lines:

Trial_0_MOCS_0	Completed_in_time			
Trial_1_AEPsych_0	Stepped_down_to_MOCS_1	Stepped_down_to_MOCS_2	Elapsed_time_9.7788	Completed
Trial_2_MOCS_1	Stepped_up_to_AEPsych_0	Completed	Skipped	
Trial_3_AEPsych_1	Stepped_down_to_MOCS_3	Elapsed_time_5.8998	Completed	
Trial_4_MOCS_2	Stepped_up_to_AEPsych_0	Completed	Skipped	
Trial_5_AEPsych_2	Stepped_down_to_MOCS_4	Stepped_down_to_MOCS_5	Elapsed_time_10.6559	Completed
Trial_6_MOCS_3	Stepped_up_to_AEPsych_1	Completed	Skipped	

Trial_7_AEPsych_3	Stepped_down_to_MOCS_6	Elapsed_time_6.0782	Completed	
Trial_8_MOCS_4	Stepped_up_to_AEPsych_2	Completed	Skipped	
Trial_9_AEPsych_4	Stepped_down_to_MOCS_7	Stepped_down_to_MOCS_8	Elapsed_time_9.5915	Completed
Trial_10_AEPsych_5	Insert_Sobol_0	Elapsed_time_4.2021	Completed	
Trial_11_MOCS_5	Stepped_up_to_AEPsych_2	Completed	Skipped	
Trial_12_MOCS_6	Stepped_up_to_AEPsych_3	Completed	Skipped	
Trial_13_AEPsych_6	Insert_Sobol_1	Insert_Sobol_2	Elapsed_time_10.4792	Completed
Trial_14_MOCS_7	Stepped_up_to_AEPsych_4	Completed	Skipped	
Trial_15_AEPsych_7	Insert_Sobol_3	Elapsed_time_5.9229	Completed	
Trial_16_AEPsych_8	Insert_Sobol_4	Elapsed_time_5.8139	Completed	
Trial_17_MOCS_8	Stepped_up_to_AEPsych_4	Completed	Skipped	
Trial_18_MOCS_9	Completed_in_time			

The trials were executed in the following order:

Trial_0_MOCS_0 → Trial_2_MOCS_1 → Trial_4_MOCS_2 → Trial_1_AEPsych_0 → Trial_6_MOCS_3 → Trial_3_AEPsych_1 → Trial_8_MOCS_4 → Trial_11_MOCS_5 → Trial_5_AEPsych_2 → Trial_12_MOCS_6 → Trial_7_AEPsych_3 → Trial_14_MOCS_7 → Trial_9_AEPsych_4 → Trial_1000_Sobol_0 → Trial_10_AEPsych_5 → Trial_1001_Sobol_1 → Trial_1002_Sobol_2 → Trial_13_AEPsych_6 → Trial_1003_Sobol_3 → Trial_15_AEPsych_7 → Trial_1004_Sobol_4 → Trial_16_AEPsych_8 → Trial_18_MOCS_9

The actual execution order was also recorded in the `TrialInfo` column in `Unity_trial_data_sub#_##_session#__YYYY-MM-DD_HH-MM-SS.csv` from the same session.

Calibration and transformation

▼ **Transformation matrices**

`M_2DWToRGB_DELL_02242025_copy.csv`

> 💡 This file contains the transformation matrix that converts stimulus values from the 2D model space (bounded between −1 and 1) to RGB space (bounded between 0 and 1). The transformation is represented by a 3 × 3 matrix. To apply this transformation, stimulus vectors in the 2D model space are expanded with a constant `1` in the third dimension before matrix multiplication.
> The date string in the filename indicates the calibration date on which the transformation matrix was computed.

`M_RGBTo2DW_DELL_02242025_copy.csv`

> 💡 This file contains the transformation matrix that converts RGB values to model-space values. The transformed output has three columns, with the last column filled with `1`'s; this column can be removed to recover the corresponding 2D model-space values.

`M_2DWToDKLPlane_DELL_02242025_copy.csv`

> 💡 This file contains the transformation matrix that converts stimulus values from the 2D model space to DKL space. The

`M_DKLPlaneTo2DW_DELL_02242025_copy.csv`

> 💡 The file contains the transformation matrix that converts values from DKL space to the 2D model space.

▼ Monitor primaries