Psychophysics

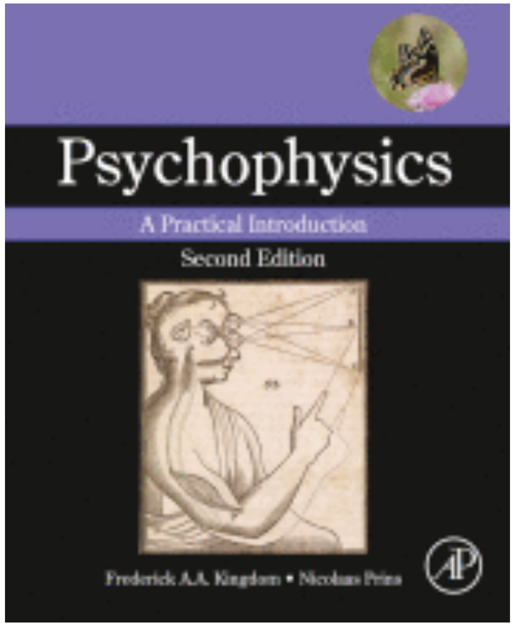
Adaptive Methods

Josef Mana

Goals

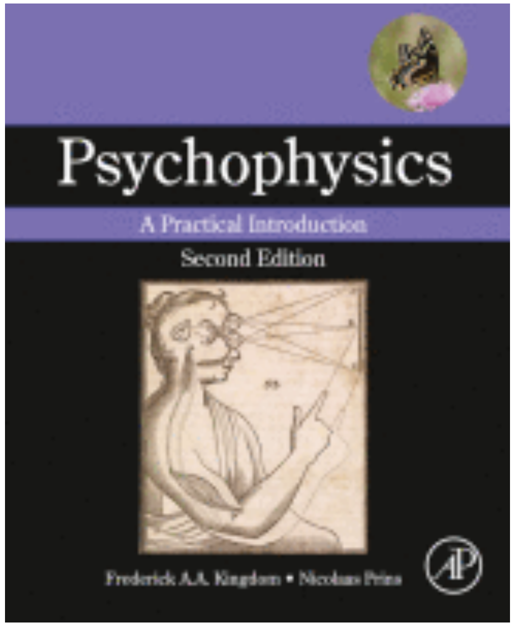
- explain what and why are adaptive methods
- name most commonly used adaptive methods
- be able to implement at least one adaptive method
- explain to a random person at least one adaptive method

Primary resource



Primary resource

- Kingdom and Prins (2016)



CHAPTER

5

Adaptive Methods*

Frederick A.A. Kingdom[1], *Nicolaas Prins*[2]
[1]McGill University, Montreal, Quebec, Canada; [2]University of Mississippi, Oxford, MS, USA

OUTLINE

Components of a psychophysical experiment

- Stimulus *(the patch)*
- Task *(e.g., 2IFC, adjustment)*
- Method *(e.g., adaptive)*
- Analysis *(calculations)*
- Measure *(contrast detection threshold)*

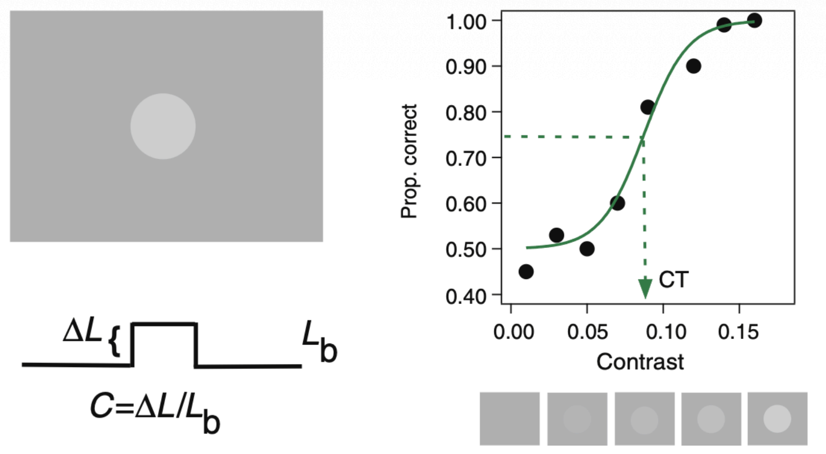


FIGURE 2.2 Top left: circular test patch on a uniform background. Bottom left: luminance profile of the patch and the definition of Weber contrast. Right: results of a standard two-interval-forced-choice (2IFC) experiment. The various stimulus contrasts are illustrated on the abscissa. Black circles are the proportion of correct responses for each contrast. The green curve is the best fit of a psychometric function, and the calculated contrast detection threshold (CT) is indicated by the arrow. See text for further details. L = luminance; L_b = luminance of background; ΔL = difference in luminance between patch and background; C = Weber contrast.

Components of a psychophysical experiment

- Stimulus *(the patch)*
- Task *(e.g., 2IFC, adjustment)*
- Method *(e.g., adaptive)* 👈
- Analysis *(calculations)*
- Measure *(contrast detection threshold)*

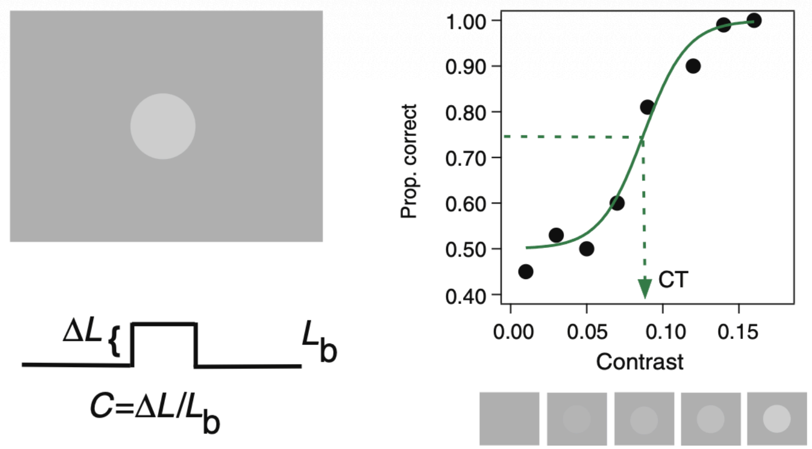


FIGURE 2.2 Top left: circular test patch on a uniform background. Bottom left: luminance profile of the patch and the definition of Weber contrast. Right: results of a standard two-interval-forced-choice (2IFC) experiment. The various stimulus contrasts are illustrated on the abscissa. Black circles are the proportion of correct responses for each contrast. The green curve is the best fit of a psychometric function, and the calculated contrast detection threshold (CT) is indicated by the arrow. See text for further details. L = luminance; L_b = luminance of background; ΔL = difference in luminance between patch and background; C = Weber contrast.

Methods

- Classical:
  - Method of limits
  - Method of constants
  - Method of adjustment
- Adaptive
  - Staircase
  - Bayesian and ML procedures

Why adaptive methods

- measurement is time-consuming
- potential measurement space might be vast
- *"sweat factor (K)"* (Taylor and Creelman 1967) 💦

Why adaptive methods

- measurement is time-consuming
- potential measurement space might be vast
- *“sweat factor (K)”* (Taylor and Creelman 1967) 💦

$$K = N\sigma^2$$

Why adaptive methods

Efficiency (Taylor and Creelman 1967)

$$K = N\sigma^2$$

Why adaptive methods

Efficiency (Taylor and Creelman 1967)[1]

$$K = N\sigma^2$$

$$dp/dL = V$$

1. p is "true" event probability, V is a slope of PF near p

Why adaptive methods

Efficiency (Taylor and Creelman 1967)[1]

$$K = N\sigma^2$$

$$dp/dL = V$$

$$K_{min} = \frac{p(1-p)}{V}$$

1. p is "true" event probability, V is a slope of PF near p

Why adaptive methods

Efficiency (Taylor and Creelman 1967)[1]

$$K = N\sigma^2$$

$$dp/dL = V$$

$$K_{min} = \frac{p(1-p)}{V}$$

$$Efficiency = \frac{K_{min}}{K_{empirical}}$$

1. p is "true" event probability, V is a slope of PF near p

General adaptive methods workflow

1. Decide upon a **measure** (e.g., threshold at 77.85% response probability)
2. *(optional)* Decide upon an **analysis** (i.e., select PF)
3. Find an **efficient** way of stimulus presentation (e.g., interleaved staircase)
4. Decide *where* to **start** the experiment (e.g., at expected threshold)
5. Decide *when* to **terminate** the experiment (e.g., after fixed number of trials)
6. **Compute**/estimate the measure (e.g., decision threshold)
7. …

General research workflow

1. Decide upon an **estimand**
2. *(optional)* Decide upon an **analysis**
3. Conduct the **experiment**
4. Compute statistical **estimate**
5. …

General research workflow

1. Decide upon an **estimand** (e.g., threshold at 77.85% response probability)
2. *(optional)* Decide upon an **analysis** (e.g., select PF)
3. Conduct the **experiment** (e.g., using interleaved staircase method)
4. Compute statistical **estimate** (e.g., decision threshold)
5. …

Adaptive methods cookbook

- Up/Down methods 🪜
  - vanilla (Dixon and Mood 1948)
  - transformed (Wetherill and Levitt 1965)
  - weighted (Kaernbach 1991)
  - transformed & weighted (Garcıá-Pérez 1998)
- “Running fit” methods 👟
  - the best PEST (Pentland 1980)
  - QUEST (Watson and Pelli 1983; Watson 2017)
  - the Psi method (Kontsevich and Tyler 1999)
  - Psi-method variants (Prins 2013)

Up/Down methods

Up/Down method (Dixon and Mood 1948)

- dropping weights from different heights on explosive mixtures
  - explosion occurs → increase height
  - explosion does not occur → decrease height

Up/Down methods

Up/Down method (Dixon and Mood 1948)

- dropping weights from different heights on explosive mixtures
  - explosion occurs → increase height
  - explosion does not occur → decrease height

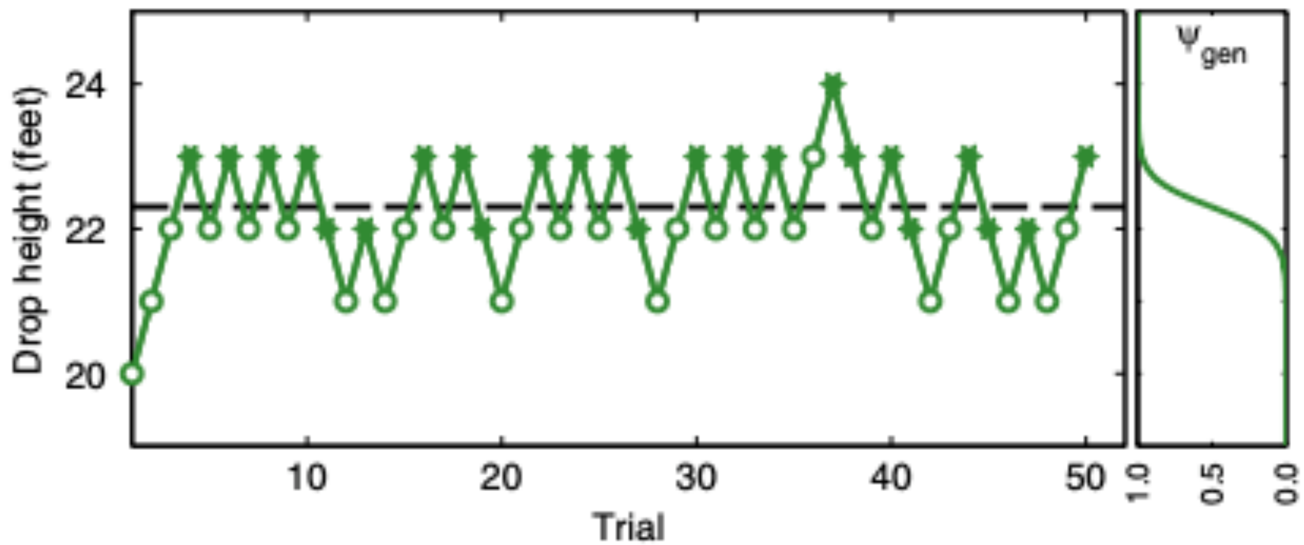


FIGURE 5.1 Simulated run of the up/down method of Dixon and Mood (1948). The figure follows the example described in the text. Weights are dropped from various heights (ordinate) on explosive mixtures. If an explosion occurs (filled symbols) the drop height is reduced by 1 foot on the next trial, and if no explosion occurs (open circular symbols), the drop height is increased by 1 foot. The targeted height (22.3 feet) is indicated by the broken line. Responses were generated by a Logistic function with $\alpha = 22.3$, $\beta = 5$, $\gamma = 0$, and $\lambda = 0$. The generating function (ψ_{gen}) is shown on the right.

Up/Down methods

Transformed up/down method (Wetherill and Levitt 1965)

- decision to decrease stimulus intensity is based on a few preceding trials
  - 1↑/1↓ => 50% correct
  - 1↑/2↓ => 70.71% correct
  - 1↑/3↓ => 79.37% correct

Up/Down methods

Transformed up/down method (Wetherill and Levitt 1965)

- decision to decrease stimulus intensity is based on a few preceding trials
  - 1↑/1↓ => 50% correct
  - 1↑/2↓ => 70.71% correct
  - 1↑/3↓ => 79.37% correct

- starts with 1↑/1↓ and switches after the first reversal (efficiency)
- steps of the same size

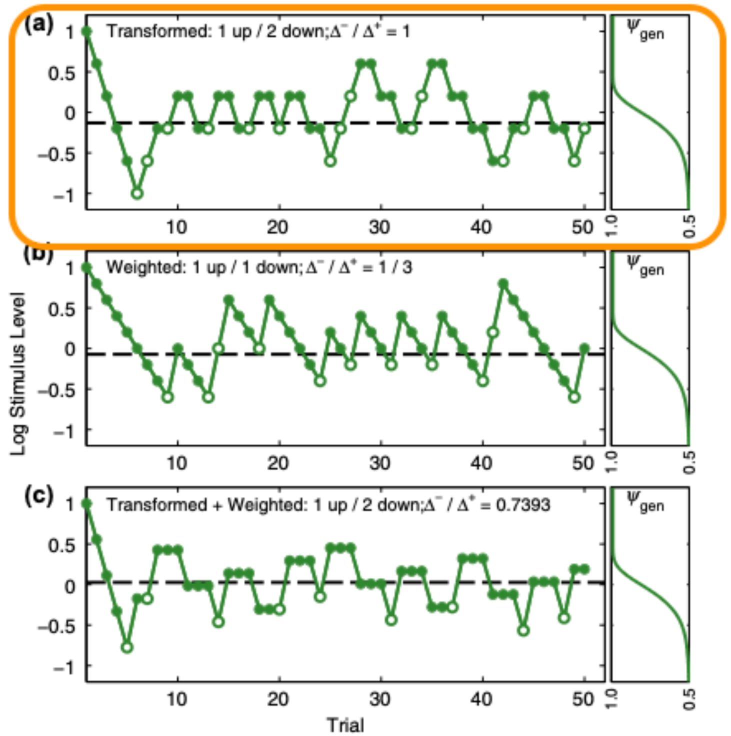


FIGURE 5.2 Examples of simulated staircases following a transformed up/down rule (a); a weighted up/down rule (b); and a transformed and weighted up/down rule (c). Correct responses are indicated by the filled symbols, and incorrect responses are indicated by open symbols. Stimulus levels corresponding to the targeted percent correct values are indicated by the broken lines (note that the different procedures target different performance levels). In all example runs the responses were generated by a Gumbel function with $\alpha = 0$, $\beta = 2$, $\gamma = 0.5$, and $\lambda = 0.01$. The generating PF (ψ_{gen}) is shown to the right of each graph. Δ^+: size of step up; Δ^-: size of step down (see Section 5.2.3).

Up/Down methods

Weighted up/down method (Kaernbach 1991)

- steps of unequal size ($\Delta^+ \uparrow$, $\Delta^- \downarrow$)

Up/Down methods

Weighted up/down method (Kaernbach 1991)

- steps of unequal size ($\Delta^+ \uparrow, \Delta^- \downarrow$)
- target probability correct:

$$\psi_{target} = \frac{\Delta^+}{\Delta^+ + \Delta^-}$$

Up/Down methods

Weighted up/down method (Kaernbach 1991)

- steps of unequal size ($\Delta^+ \uparrow, \Delta^- \downarrow$)
- target probability correct:

$$\psi_{target} = \frac{\Delta^+}{\Delta^+ + \Delta^-}$$

$$\frac{\Delta^-}{\Delta^+} = \frac{1 - \psi_{target}}{\psi_{target}}$$

Up/Down methods

Weighted up/down method (Kaernbach 1991)

- steps of unequal size ($\Delta^+ \uparrow, \Delta^- \downarrow$)
- target probability correct:

$$\psi_{target} = \frac{\Delta^+}{\Delta^+ + \Delta^-}$$

$$\frac{\Delta^-}{\Delta^+} = \frac{1 - \psi_{target}}{\psi_{target}}$$

- e.g., for $\psi_{target} = 0.75$:

$$\frac{\Delta^-}{\Delta^+} = \frac{1 - 0.75}{0.75} = \frac{1}{3}$$

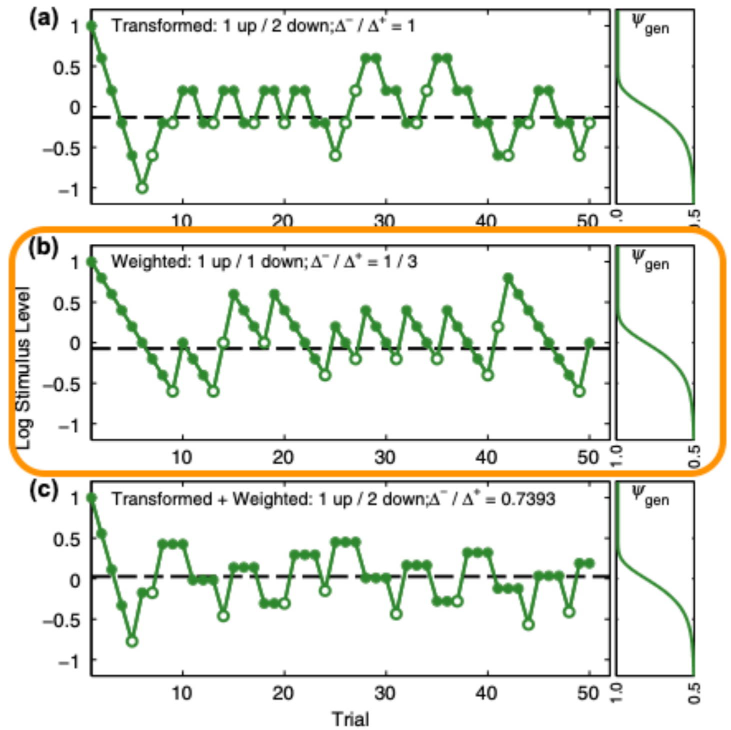


FIGURE 5.2 Examples of simulated staircases following a transformed up/down rule (a); a weighted up/down rule (b); and a transformed and weighted up/down rule (c). Correct responses are indicated by the filled symbols, and incorrect responses are indicated by open symbols. Stimulus levels corresponding to the targeted percent correct values are indicated by the broken lines (note that the different procedures target different performance levels). In all example runs the responses were generated by a Gumbel function with $\alpha = 0$, $\beta = 2$, $\gamma = 0.5$, and $\lambda = 0.01$. The generating PF (ψ_{gen}) is shown to the right of each graph. Δ^+: size of step up; Δ^-: size of step down (see Section 5.2.3).

Up/Down methods

Weighted up/down method (Kaernbach 1991)

- steps of unequal size ($\Delta^+ \uparrow, \Delta^- \downarrow$)
- target probability correct:

$$\psi_{target} = \frac{\Delta^+}{\Delta^+ + \Delta^-}$$

$$\frac{\Delta^-}{\Delta^+} = \frac{1 - \psi_{target}}{\psi_{target}}$$

- e.g., for $\psi_{target} = 0.75$:

$$\frac{\Delta^-}{\Delta^+} = \frac{1 - 0.75}{0.75} = \frac{1}{3}$$

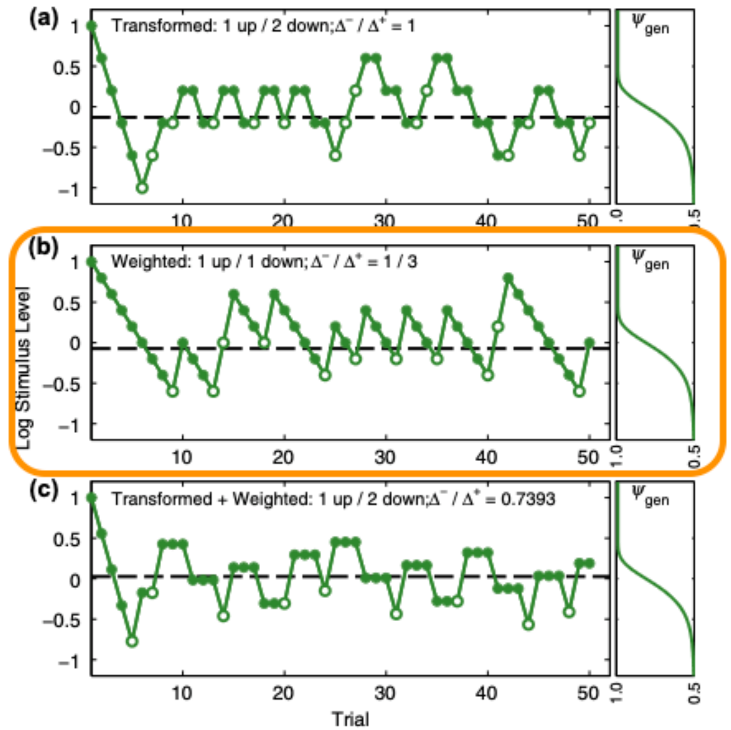


FIGURE 5.2 Examples of simulated staircases following a transformed up/down rule (a); a weighted up/down rule (b); and a transformed and weighted up/down rule (c). Correct responses are indicated by the filled symbols, and incorrect responses are indicated by open symbols. Stimulus levels corresponding to the targeted percent correct values are indicated by the broken lines (note that the different procedures target different performance levels). In all example runs the responses were generated by a Gumbel function with $\alpha = 0$, $\beta = 2$, $\gamma = 0.5$, and $\lambda = 0.01$. The generating PF (ψ_{gen}) is shown to the right of each graph. Δ^+: size of step up; Δ^-: size of step down (see Section 5.2.3).

Up/Down methods

Transformed & weighted up/down method (García-Pérez 1998)

- number of consecutive incorrect responses (D) before step down & step size differ

Up/Down methods

Transformed & weighted up/down method (García-Pérez 1998)

- number of consecutive incorrect responses (D) before step down & step size differ

$$\psi_{target} = \left(\frac{\Delta^+}{\Delta^+ + \Delta^-}\right)^{\frac{1}{D}}$$

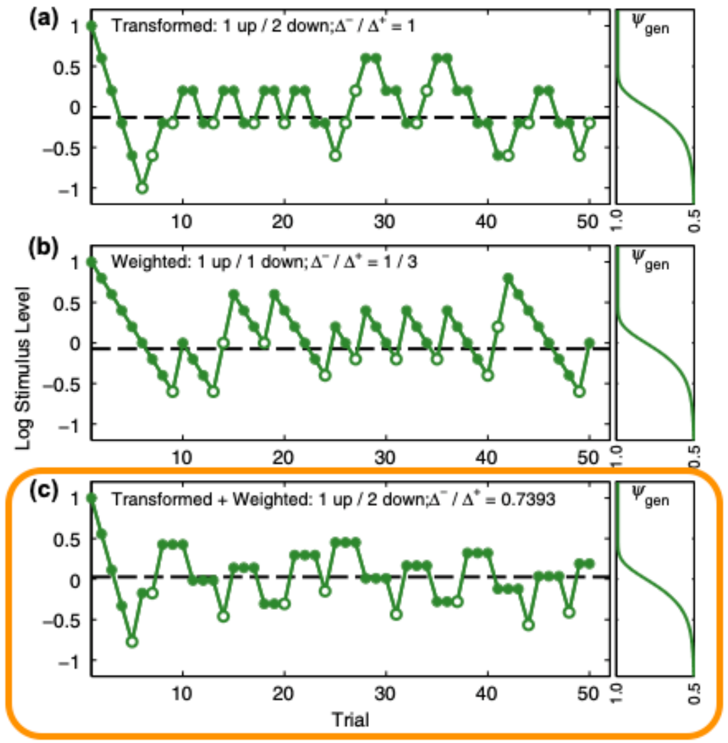


FIGURE 5.2 Examples of simulated staircases following a transformed up/down rule (a); a weighted up/down rule (b); and a transformed and weighted up/down rule (c). Correct responses are indicated by the filled symbols, and incorrect responses are indicated by open symbols. Stimulus levels corresponding to the targeted percent correct values are indicated by the broken lines (note that the different procedures target different performance levels). In all example runs the responses were generated by a Gumbel function with $\alpha = 0$, $\beta = 2$, $\gamma = 0.5$, and $\lambda = 0.01$. The generating PF (ψ_{gen}) is shown to the right of each graph. Δ^+: size of step up; Δ^-: size of step down (see Section 5.2.3).

Up/Down methods

Transformed & weighted up/down method (García-Pérez 1998)

- number of consecutive incorrect responses (D) before step down & step size differ

$$\psi_{target} = \left(\frac{\Delta^+}{\Delta^+ + \Delta^-}\right)^{\frac{1}{D}}$$

- $D = 1$ → weighted up/down
- $\Delta^+ = \Delta^-$ → transformed up/down

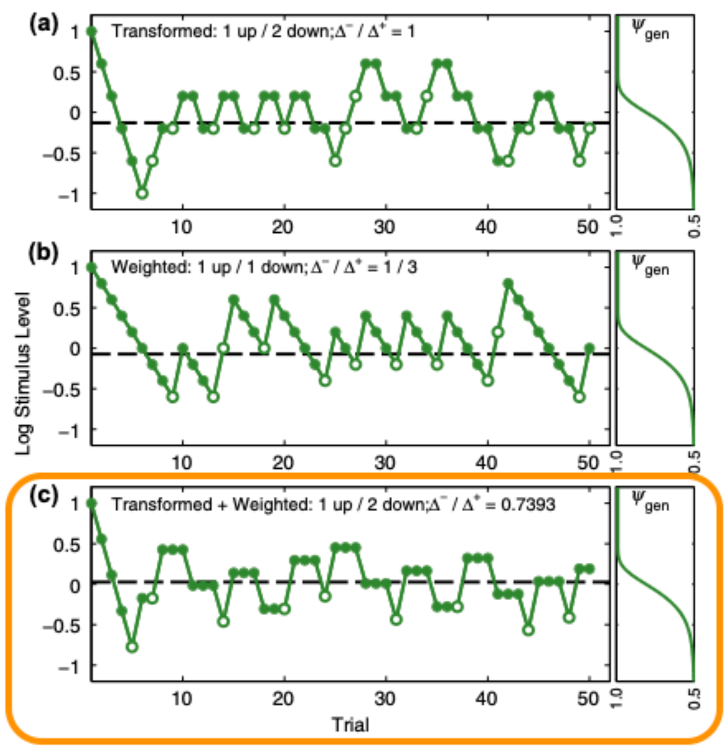


FIGURE 5.2 Examples of simulated staircases following a transformed up/down rule (a); a weighted up/down rule (b); and a transformed and weighted up/down rule (c). Correct responses are indicated by the filled symbols, and incorrect responses are indicated by open symbols. Stimulus levels corresponding to the targeted percent correct values are indicated by the broken lines (note that the different procedures target different performance levels). In all example runs the responses were generated by a Gumbel function with $\alpha = 0$, $\beta = 2$, $\gamma = 0.5$, and $\lambda = 0.01$. The generating PF (ψ_{gen}) is shown to the right of each graph. Δ^+: size of step up; Δ^-: size of step down (see Section 5.2.3).

Up/Down methods

Termination criteria & threshold estimate

- after a specific number of **reversals** of direction
  - average of last N trials
- after a specified number of **trials**
  - average of last N trials
- ***"hybrid adaptive procedure"***
  - up/down method to select stimulus intensities
  - PF to derive threshold using all the data
- up/down methods are *"non-parametric"*
  - assume only monotonic relationship, no particular PF shape
  - cannot estimate slope (data concentrated near threshold)

Up/Down methods

Practical tips

- random trials with **interleaved** up/down staircases to hide the mechanism
- **multiple runs** → between-run SD indicates *reliability* + pooled data for *PF estimation*
- **units** → corresponding to linear changes in *internal representation* (linear or log)
- **step size** → not as clear as might have seen

Up/Down methods

Practical tips

- random trials with **interleaved** up/down staircases to hide the mechanism
- **multiple runs** → between-run SD indicates *reliability* + pooled data for *PF estimation*
- **units** → corresponding to linear changes in *internal representation* (linear or log)
- **step size** → not as clear as might have seen

TABLE 5.1 Ratios of down stepsize and up stepsize Δ^{-}/Δ^{+} that will reliably converge on the targeted ψ values given by Eqn (5.2); these values are suggested by García-Pérez (1998) and are based on a large number of simulated runs.

Rule	Δ^{-}/Δ^{+}	Targeted ψ (%)
1 up/1 down	0.2845	77.85
1 up/2 down	0.5488	80.35
1 up/3 down	0.7393	83.15
1 up/4 down	0.8415	85.84

Up/Down methods

Practical tips

- random trials with **interleaved** up/down staircases to hide the mechanism
- **multiple runs** → between-run SD indicates *reliability* + pooled data for *PF estimation*
- **units** → corresponding to linear changes in *internal representation* (linear or log)
- **step size** → not as clear as might have seen
  - steps up to $\frac{\sigma}{2}$[1]

TABLE 5.1 Ratios of down stepsize and up stepsize Δ^-/Δ^+ that will reliably converge on the targeted ψ values given by Eqn (5.2); these values are suggested by García-Pérez (1998) and are based on a large number of simulated runs.

Rule	Δ^-/Δ^+	Targeted ψ (%)
1 up/1 down	0.2845	77.85
1 up/2 down	0.5488	80.35
1 up/3 down	0.7393	83.15
1 up/4 down	0.8415	85.84

1. σ is the spread of underlying PF

Up/Down methods

Practical tips

- random trials with **interleaved** up/down staircases to hide the mechanism
- **multiple runs** → between-run SD indicates *reliability* + pooled data for *PF estimation*
- **units** → corresponding to linear changes in *internal representation* (linear or log)
- **step size** → not as clear as might have seen
  - steps up to $\frac{\sigma}{2}$
- What to do at the fringes? (E.g., in weighted, what if previous stimulus was supposed to have had “negative intensity”?)
  - make a step relative to the negative value (García-Pérez 1998)
  - use a “running fit” adaptive method

Running fit

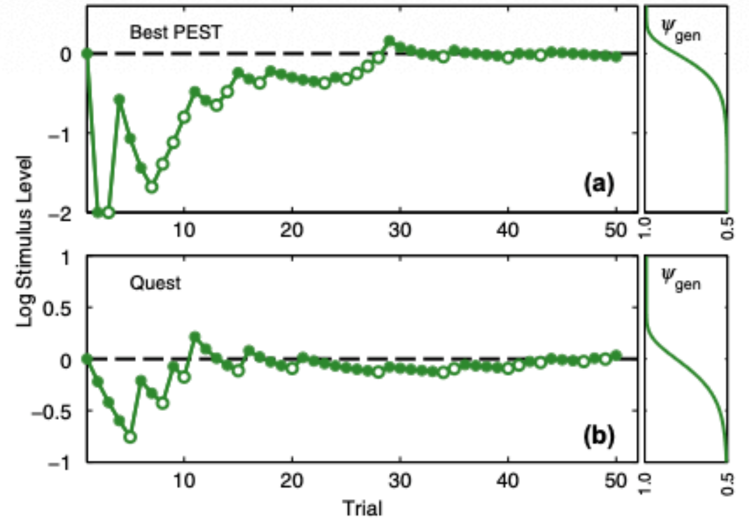


FIGURE 5.3 Examples of a simulated best PEST staircase (a) and a Quest staircase (b). Correct responses are indicated by the filled symbols, and incorrect responses are indicated by open symbols. Stimulus levels corresponding to the threshold of the generating PFs are indicated by the broken lines. In both example runs the responses were generated by a Gumbel function with $\alpha = 0$, $\beta = 2$, $\gamma = 0.5$, and $\lambda = 0.01$. The generating PF (ψ_{gen}) is shown to the right of each graph.

Running fit

The best PEST (Pentland 1980)

- estimates **threshold**
- assumes a specific form of PF
- intensity on next trial ← ML estimate now
- step sizes and directions differ
- needs bounds for the first trial

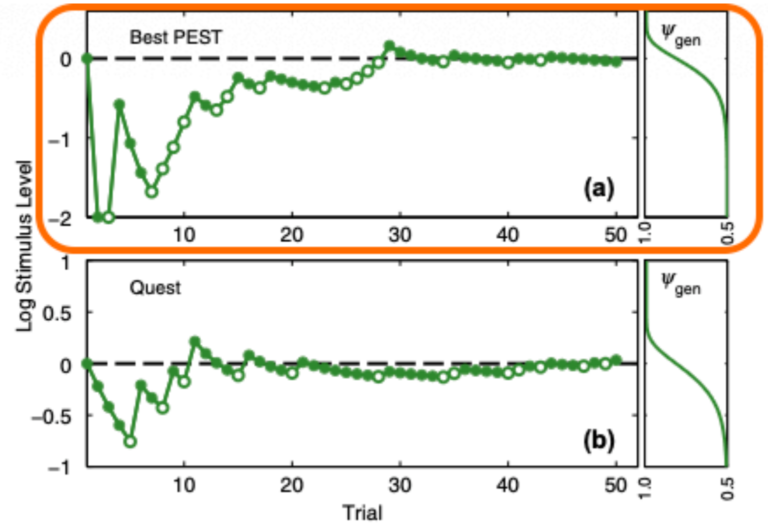


FIGURE 5.3 Examples of a simulated best PEST staircase (a) and a Quest staircase (b). Correct responses are indicated by the filled symbols, and incorrect responses are indicated by open symbols. Stimulus levels corresponding to the threshold of the generating PFs are indicated by the broken lines. In both example runs the responses were generated by a Gumbel function with $\alpha = 0$, $\beta = 2$, $\gamma = 0.5$, and $\lambda = 0.01$. The generating PF (ψ_{gen}) is shown to the right of each graph.

Running fit

Quest (Watson and Pelli 1983)

- *"Bayesian PEST"*
- Priors assist with first few trials
- Posterior mode or mean as the **threshold** estimate
- Use uniform prior? (equivalent to PEST) 🤔
- Divide by prior? (equivalent to ML) 🤔

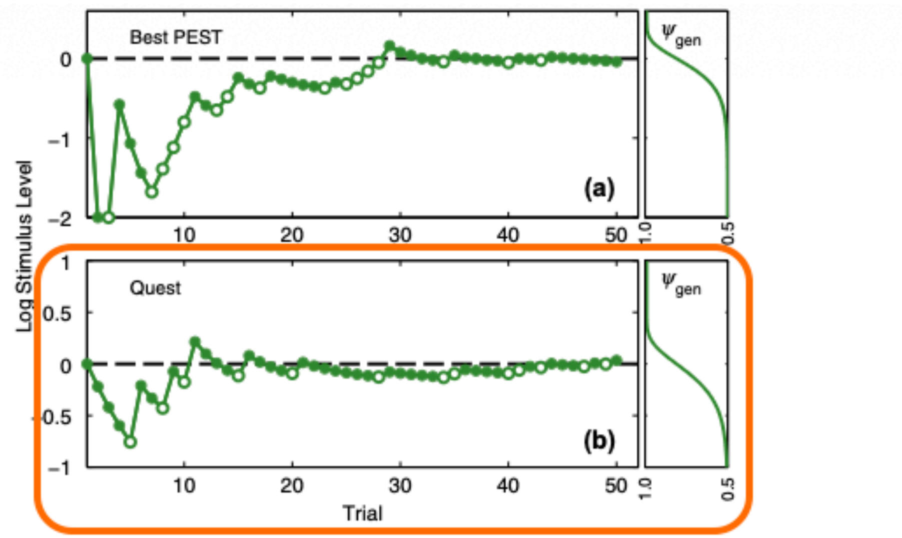


FIGURE 5.3 Examples of a simulated best PEST staircase (a) and a Quest staircase (b). Correct responses are indicated by the filled symbols, and incorrect responses are indicated by open symbols. Stimulus levels corresponding to the threshold of the generating PFs are indicated by the broken lines. In both example runs the responses were generated by a Gumbel function with $\alpha = 0$, $\beta = 2$, $\gamma = 0.5$, and $\lambda = 0.01$. The generating PF (ψ_{gen}) is shown to the right of each graph.

Running fit

Termination criteria & practical considerations

- after a specified number of **trials**
- after a specific number of **reversals** of direction
- *parameteric* → assumptions about PF shape required
  - mostly robust to violated assumptions
  - **slope** affects step size
  - allowing **lapses** (e.g., rate = 0.02)
- multiple runs/sessions

Running fit

The Psi method (Kontsevich and Tyler 1999)

- estimates both the **threshold** and the **slope**
- minimises the expected *entropy* in the posterior distribution
- may be hard on RAM
- **psi-marginal method** (Prins 2013)
  - estimates lapse rate as well
  - minimises entropy for threshold and slope only

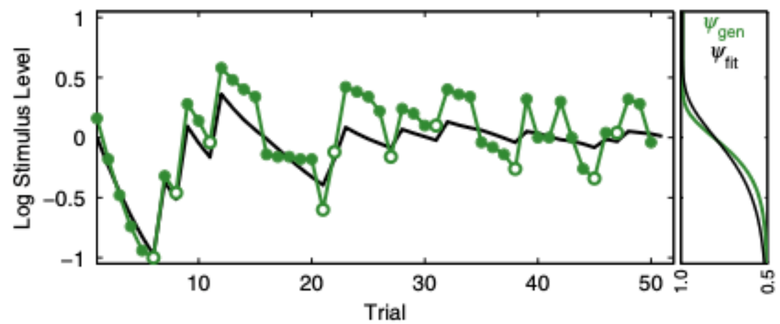
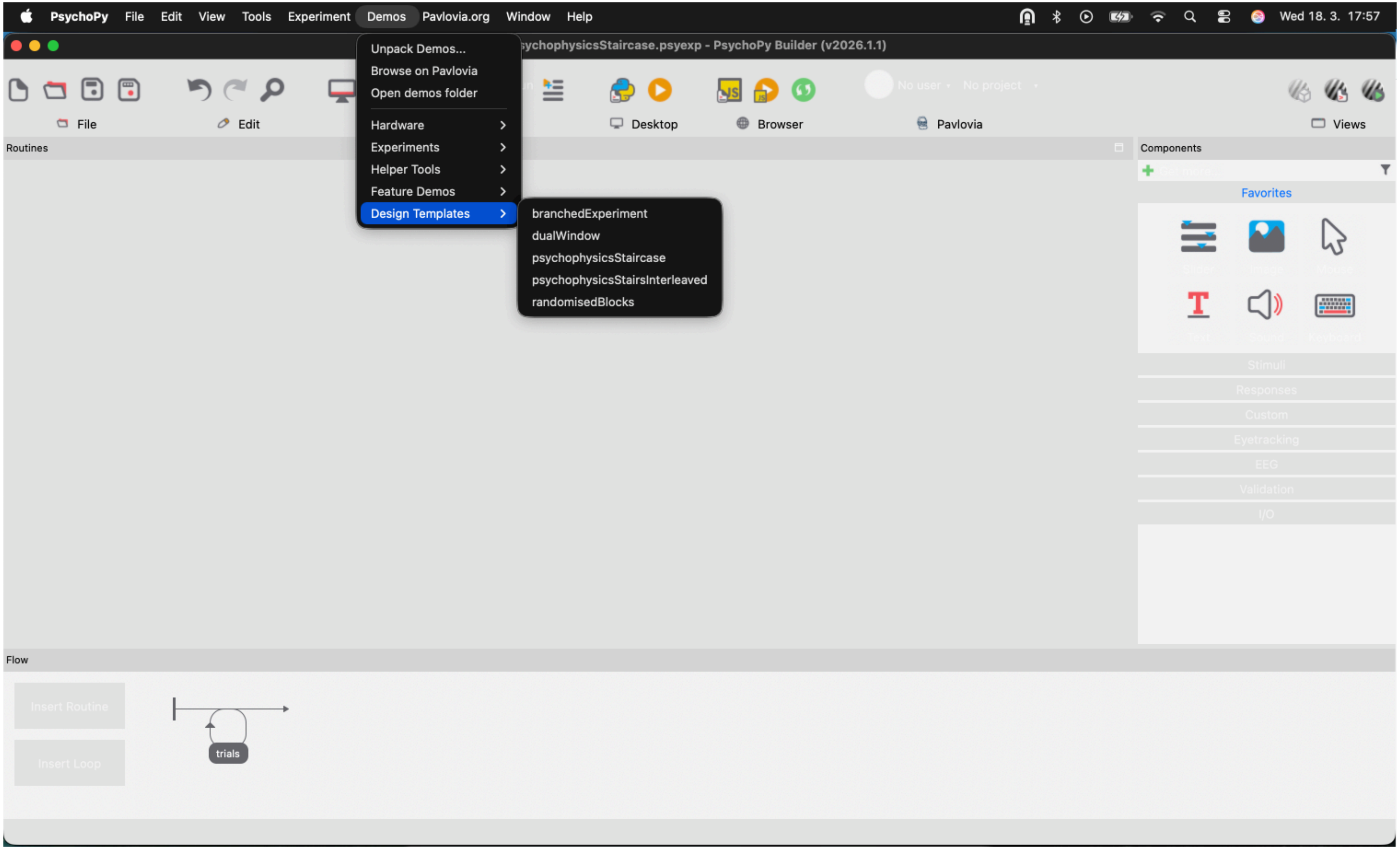
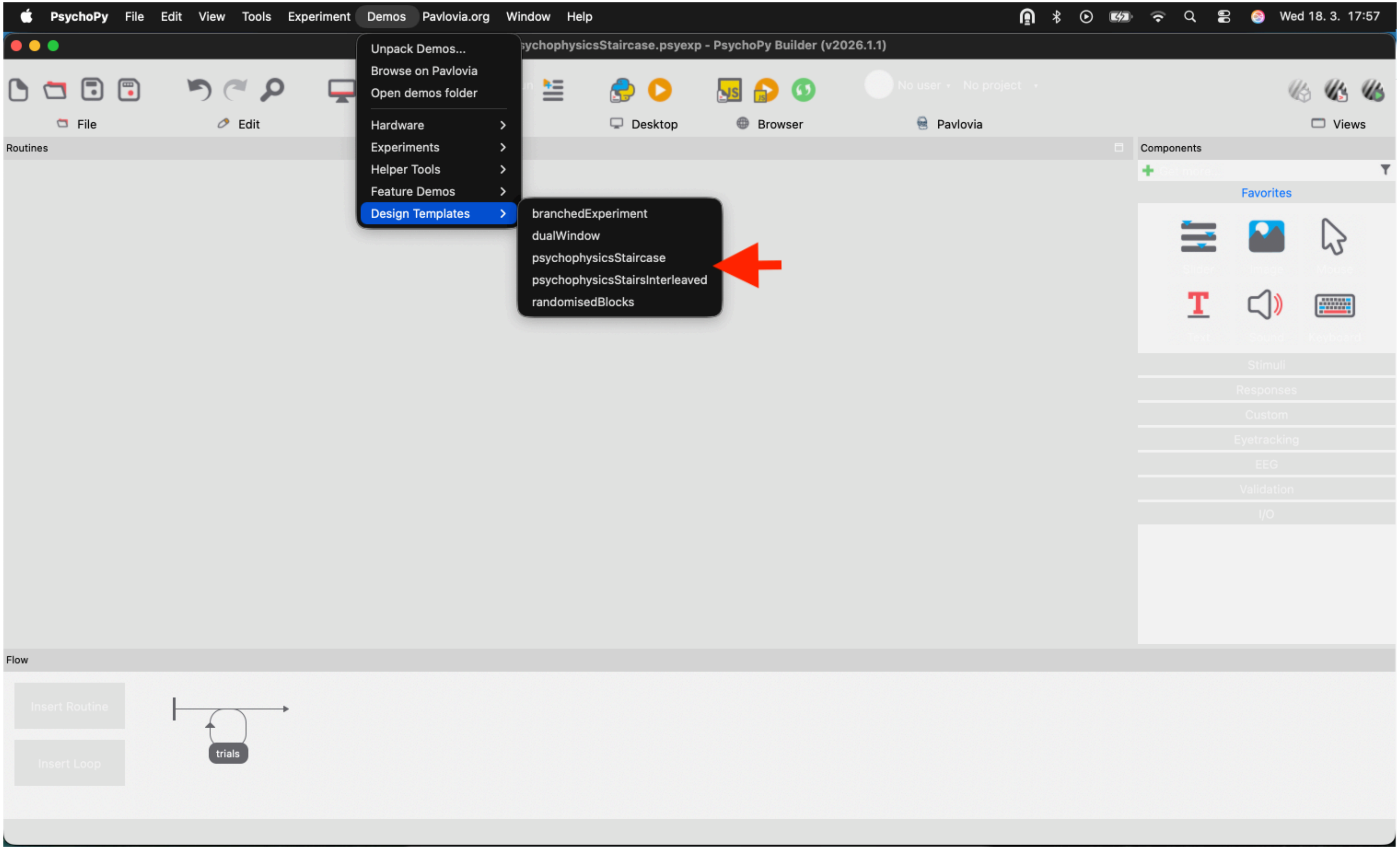


FIGURE 5.4 Example of a simulated psi method staircase. Correct responses are indicated by the filled symbols, and incorrect responses are indicated by open symbols. The black line displays the running estimate of the threshold based on the posterior distribution. The responses were generated by a Gumbel function with $\alpha = 0$, $\beta = 2$, $\gamma = 0.5$, and $\lambda = 0.01$. The generating function (ψ_{gen}) is shown on the right in green, and the function fitted by the psi method (ψ_{fit}) after completion of the 50 trials is shown on the right in black.

Adaptive methods in PsychoPy

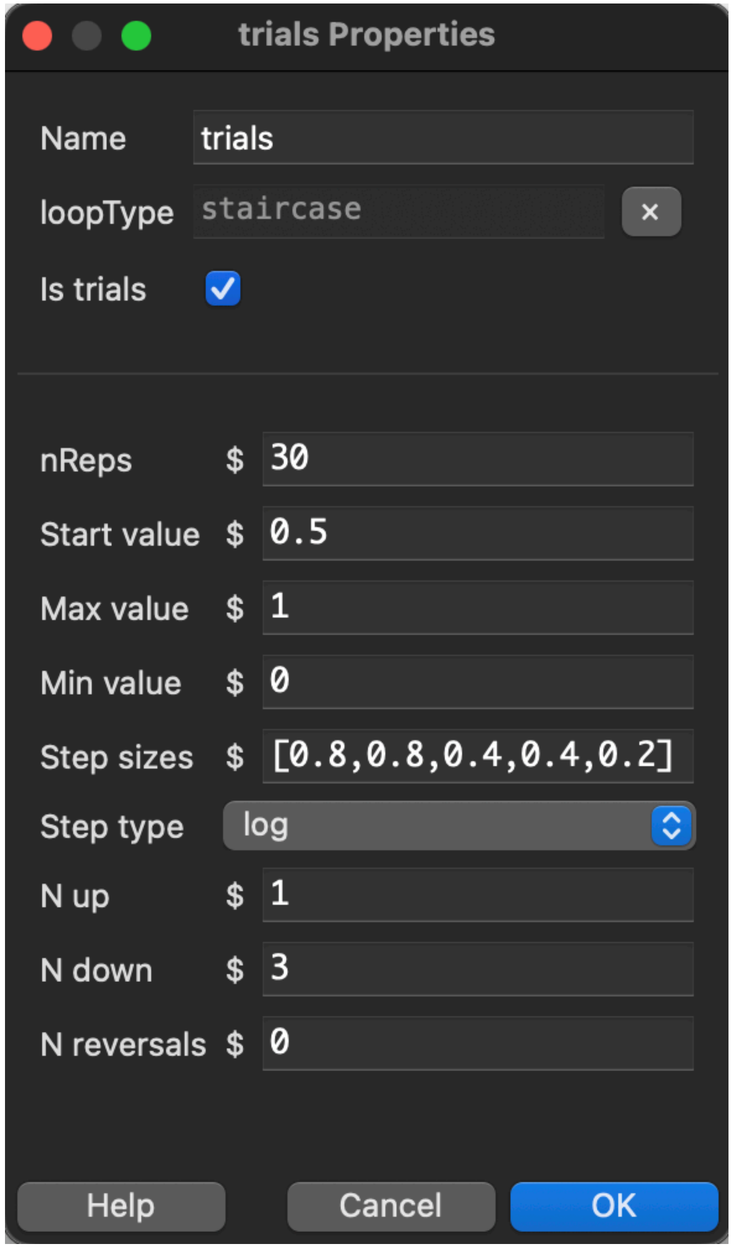


Adaptive methods in PsychoPy



Adaptive methods in PsychoPy

Staircase



readme (psychophysicsStaircase)

Detection threshold with a staircase

The experiment:

This experiment presents a Gabor (a sinusoid in a Gaussian window) with an orientation of your choosing. On each trial the subject reports whether or not they saw the stimulus with the up/down cursor keys. If they did see it 3 times in a row then the stimulus reduces in contrast, if they fail to see it, the stimulus immediately increases in contrast.

Analysing your data:

After the experiment a data/ folder will have been created for you, and it will contain an excel file with your data.

For a staircase the data file contains a number of things that could be used in the analysis; the intensity of the stimulus at each of the reversal points, the intensity on every trial and the response (correct or incorrect) on every trial too.

Commonly people analyse staircase data by taking the average of the last few (e.g. 4) reversal intensities. A more advanced analysis is provided as a Coder demo called JND_staircase analysis. This collapses data across all the levels presented during the run(s) and fits a psychometric curve to this averaged data.

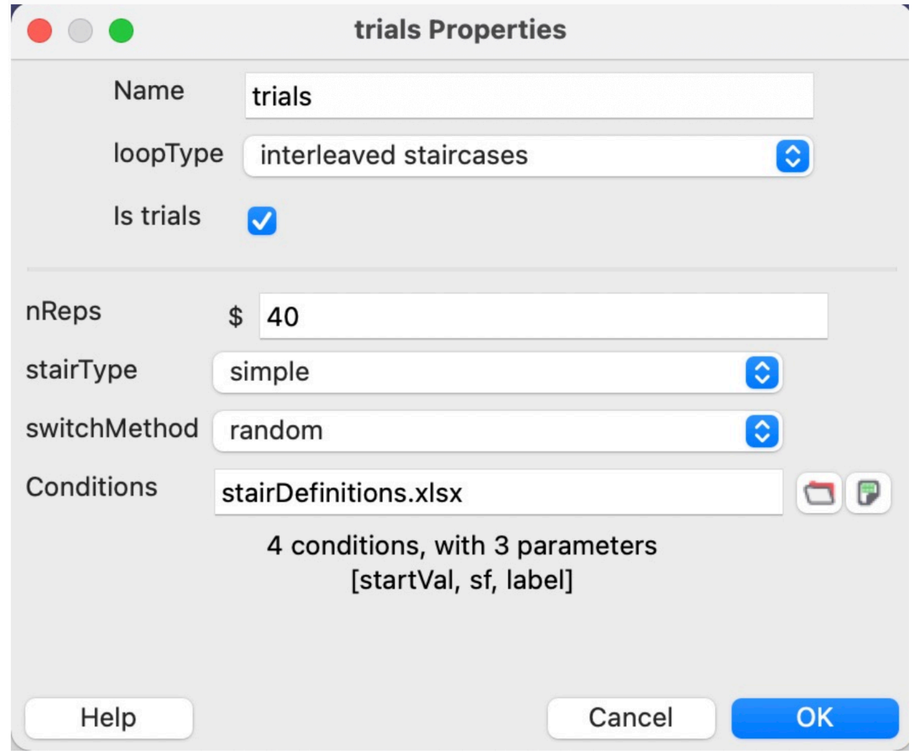
Notes:

The current value of the staircase can be found using the variable name 'level' or, equivalently, 'thisXXXX' where XXXX is the singular name of the loop (e.g. if you have a staircase loop called trials you could access thisTrial).

For this demo you can see the use of 'level' by looking at the advanced properties>color of the gabor stimulus.

Adaptive methods in PsychoPy

Interleaved Staircase



	A	B	C	D	E	F
1	label	startVal	sf	stepSizes	maxVal	minVal
2	low_2	0,001	2	[4,2,2,1]	1	0
3	high_2	0,1	2	[4,2,2,1]	1	0
4	low_8	0,001	8	[4,2,2,1]	1	0
5	high_8	0,1	8	[4,2,2,1]	1	0
6						

readme (interleaved_SF_contrast)

Measuring detection threshold using 2AFC task with interleaved adaptive staircases

The experiment

Measure your detection thresholds for two different spatial frequencies of grating. See notes about the staircases below. It's 160 trials long (4 staircases of 40 reps each) so it will take. Respond as quickly as you can and it will be over in about 5 mins!

Notes

1. This has been set up to use the 'testMonitor' with units of degrees. To make this work YOU NEED TO GO TO MONITOR CENTRE and set up the dimension and distance for testMonitor. You can see in the Experiment Settings that the default units for the experiment are then set to be 'deg' (which means spatial frequency is going to be in c/deg).
2. To decide whether the stimulus appears on the left or right a Code Component called `setSide`has been used to determine which is the correct key to press and the position of the grating.
3. The fixation point goes off between trials to alert the subject that the next trial is occurring
4. Right now the experiment uses seconds for timing (for simplicity) but to be precise N frames would be better, particularly for the grating stimulus duration.
5. This experiment obviously isn't using any precise way to control contrast (like a CRS Bits# box with a CRT screen), so the estimates of your

The staircases

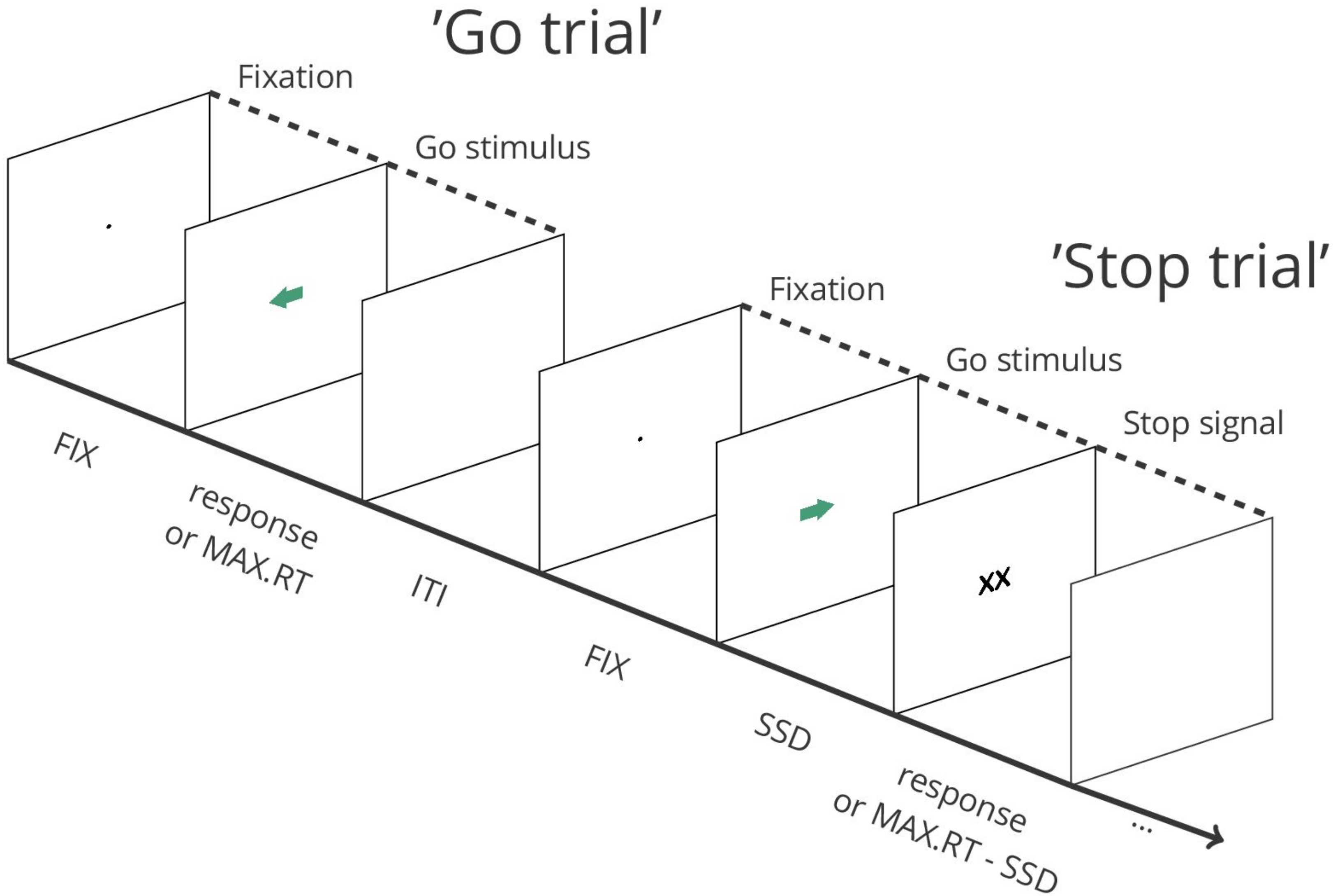
The experiment uses 4 adaptive staircases with interleaved trials. The set up of the stairs can be seen in the excel file in this folder: there a 2 spatial frequencies and for each there are two staircases, one starting at a very low contrast and another starting at a high contrast. The step size for the staircase is changing each time the staircase reverses in its direction, as determined by the stepSize column in the conditions file.

Case study

Stop Signal Task

Stop signal task

Task (Verbruggen et al. 2019)



Stop signal task

Measurement model (Matzke et al. 2013) 🏇🏇

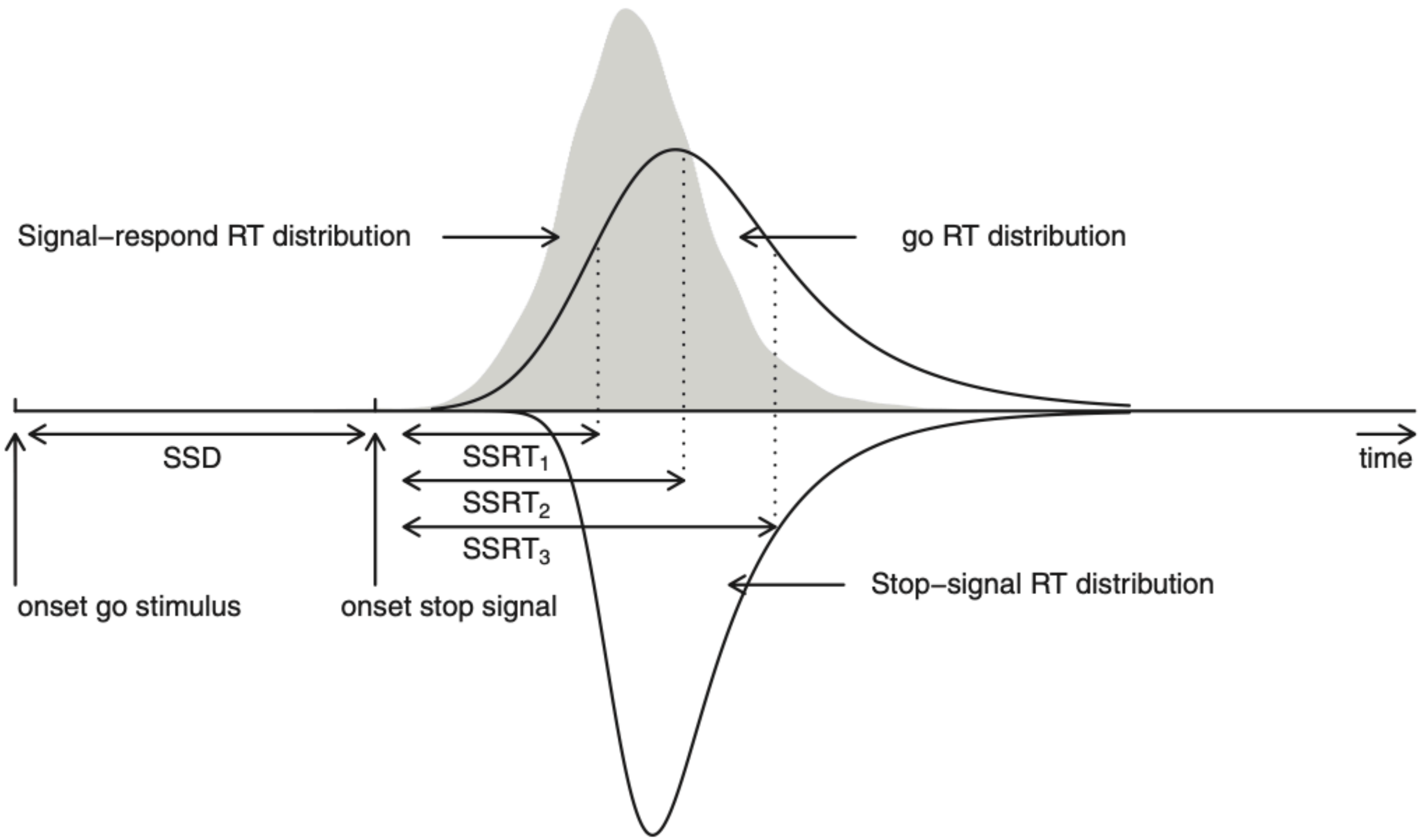
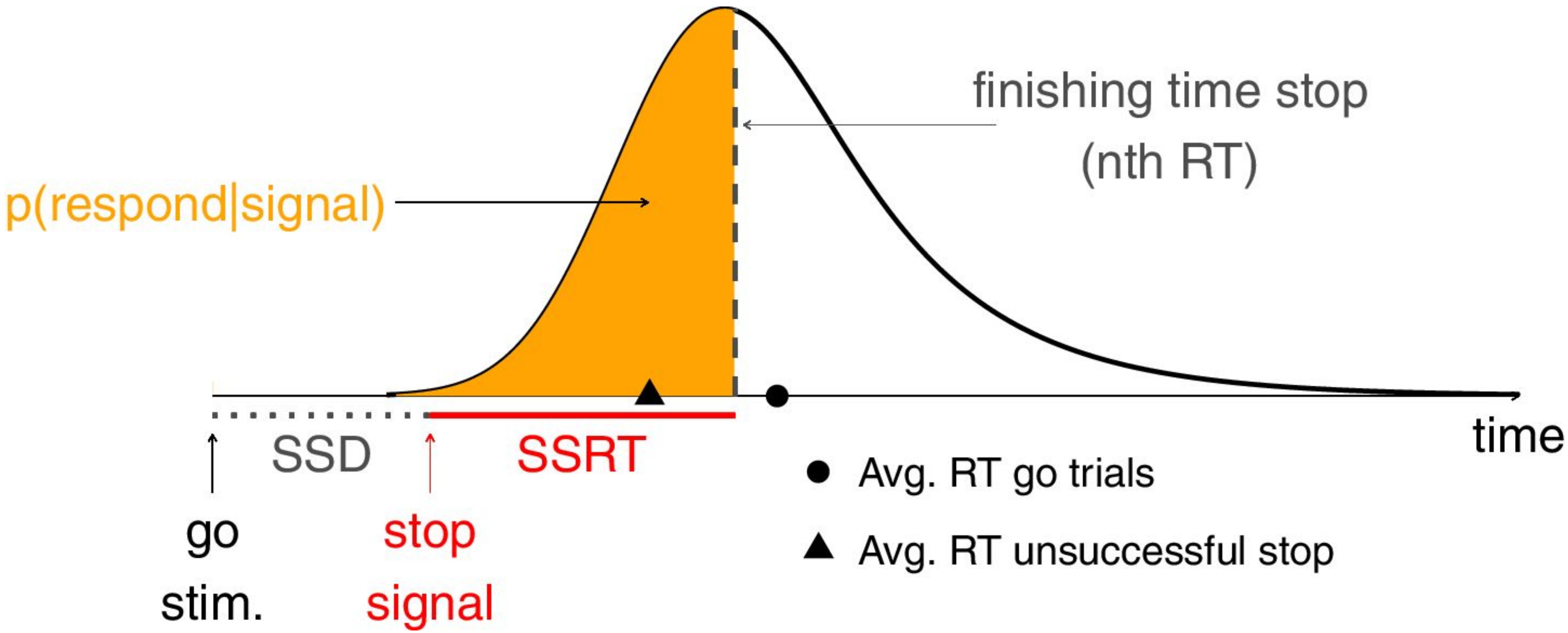


Figure 4. Graphical representation of the complete horse race model. RT = response time; SSD = stop-signal delay; SSRT = stop-signal reaction time.

Stop signal Task

Measure (Verbruggen et al. 2019)



Stop signal task

Non-parametric, mean method

- SSRT = mean of the inhibition function
- assumes that:

$$E(RT) = SSRT + E(SSD)$$

- then:

$$SSRT = E(RT) - E(SSD)$$

- valid only if staircase procedure is used
- even then likely biased

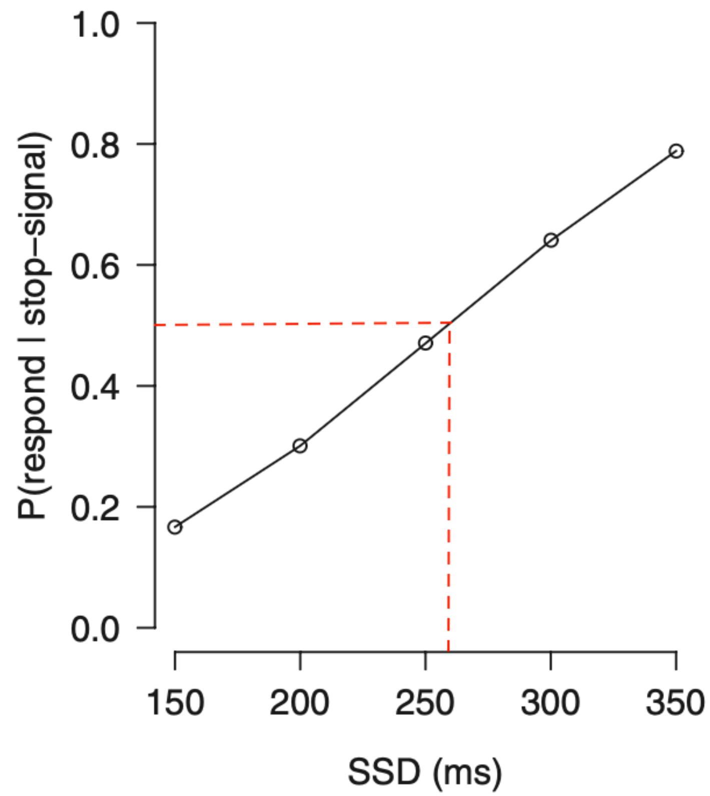


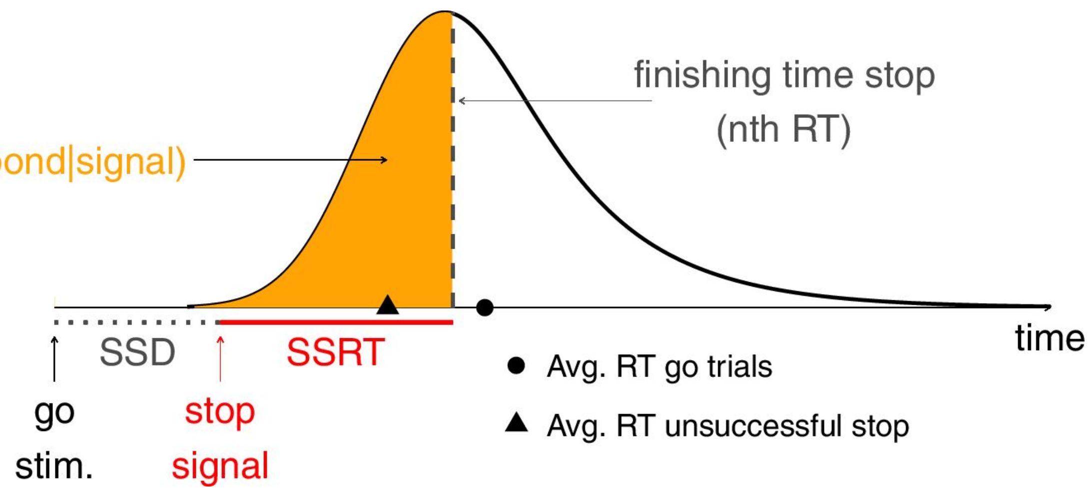
Figure 3. Example of an inhibition function based on synthetic data from the recovery study for the individual Bayesian parametric approach. The figure shows how the probability of responding on a stop-signal trial increases with increasing stop-signal delay (SSD).

Stop signal task

Non-parametric, integration method

- the point at which STOP processes finish estimating by “integrating” RT distribution at the point where $p(respond|signal) = 0.5$

1. sort RTs from fastest to slowest
2. *(optional)* assign $max(RT)$ to GO ommissions
3. find $p(respond|signal)$
4. calculate $n = p(respond|signal) \times n_{GO}$
5. $SSRT = n^{th} RT$



Stop signal task

Parametric, BEEST (Matzke, Love, and Heathcote 2016)

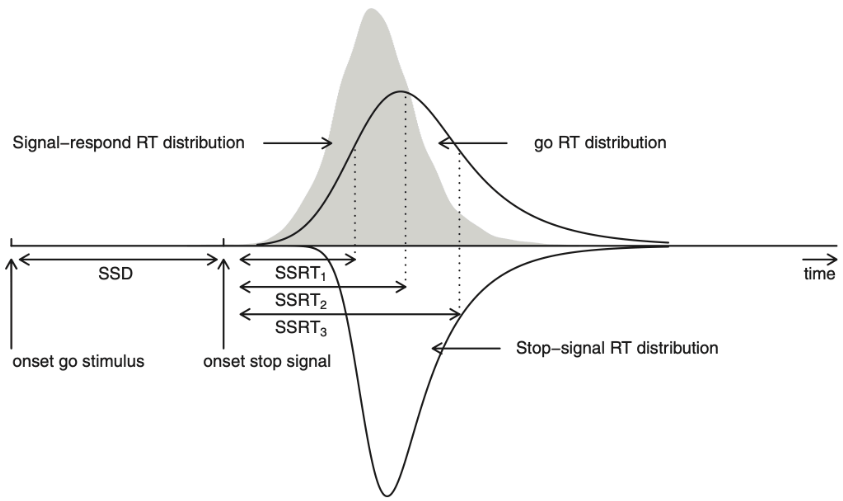


Figure 4. Graphical representation of the complete horse race model. RT = response time; SSD = stop-signal delay; SSRT = stop-signal reaction time.

$$
\begin{aligned}
&\text{Ł}_{SR}(\mu_{go}, \sigma_{go}, \tau_{go}, \mu_{stop}, \sigma_{stop}, \tau_{stop}, P(TF), SSD) \\
&= \prod_{r=1}^{R} \{P(TF) \times f_{go}(t_r; \mu_{go}, \sigma_{go}, \tau_{go}) \\
&\quad + [1 - P(TF)] \times [1 - F_{stop}(t_r; \mu_{stop}, \sigma_{stop}, \tau_{stop}, SSD)] \\
&\quad \times f_{go}(t_r; \mu_{go}, \sigma_{go}, \tau_{go})\}
\end{aligned} \tag{1}
$$

$$
\begin{aligned}
&\text{Ł}_{I}(\mu_{go}, \sigma_{go}, \tau_{go}, \mu_{stop}, \sigma_{stop}, \tau_{stop}, P(TF), SSD) \\
&= \prod_{i=1}^{I} [1 - P(TF)] \times \int_{-\infty}^{\infty} [1 - F_{go}(t_i; \mu_{go}, \sigma_{go}, \tau_{go})] \\
&\quad \times f_{stop}(t_i; \mu_{stop}, \sigma_{stop}, \tau_{stop}, SSD) dt_i,
\end{aligned} \tag{2}
$$

Stop signal task

Bonus, non-parametric, non-behavioural (Thunberg et al. 2024)

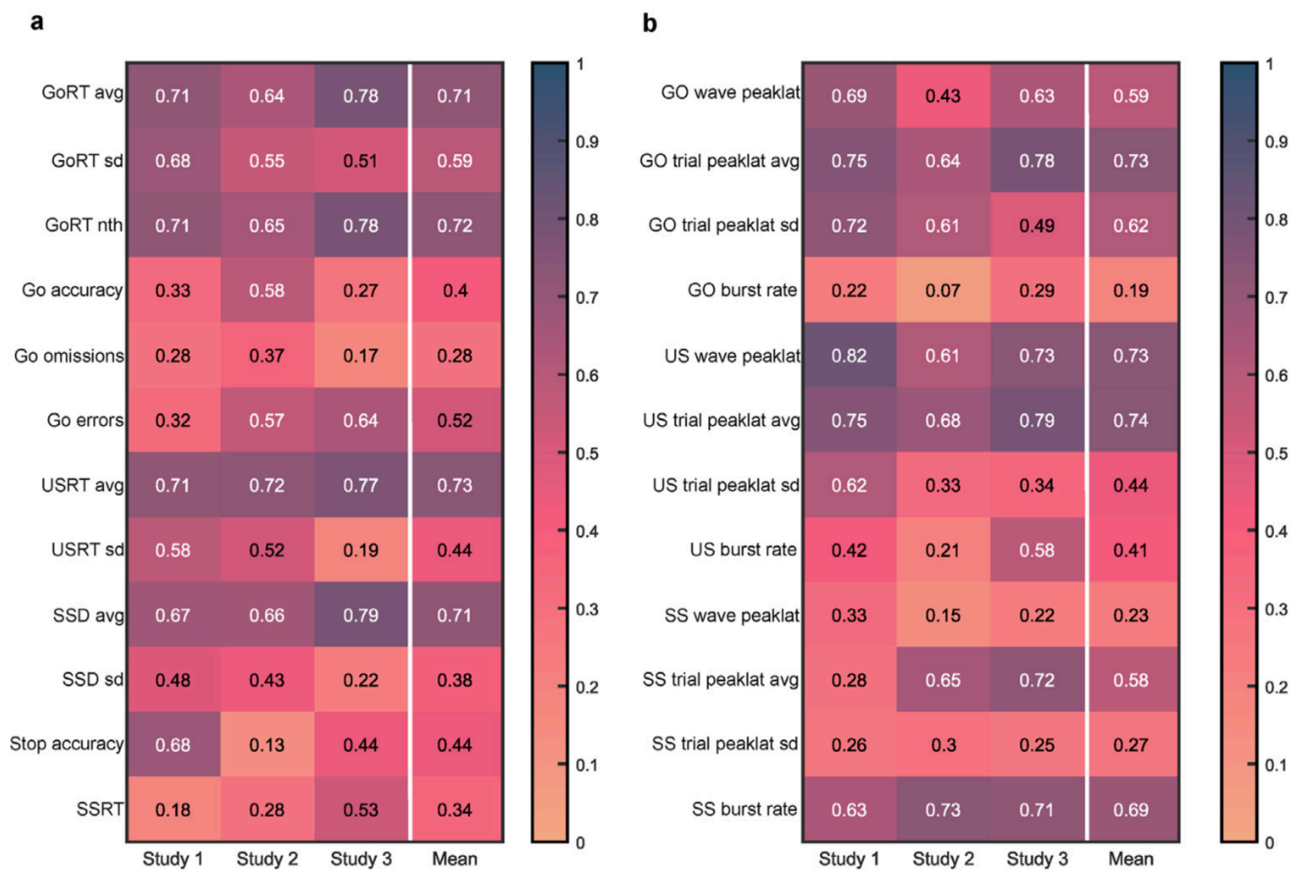
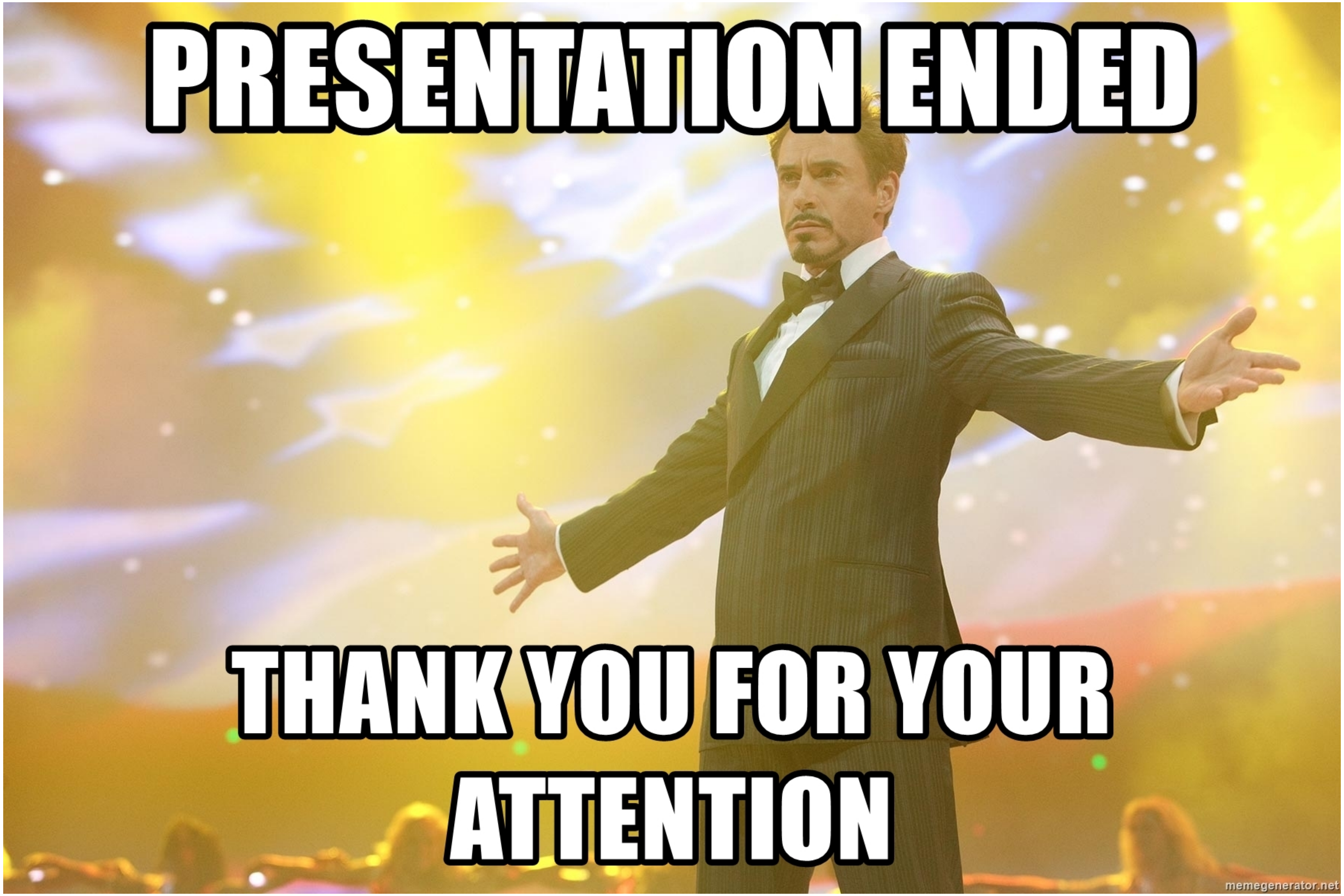


Fig. 3 – Test-retest reliability for task performance and electromyographic variables. Figures show the test-retest estimates for each study, as well as the mean estimate across studies. a. Test-retest reliability for task performance variables. b. Test-retest reliability for SST-derived EMG variables. Parameters are estimated from both participant-level and trial-level waveforms, denoted by wave and trial, respectively. Abbreviations: GoRT – go reaction times, USRT – unsuccessful stop reaction times, SSD – stop signal delay, SSRT – stop signal reaction time, avg – average, SD – standard deviation, peaklat – peak latency, GO – go trials, US – unsuccessful stop trials, SS - successful stop trials.


PRESENTATION ENDED
THANK YOU FOR YOUR ATTENTION
memegenerator.net

References

Dixon, W. J., and A. M. Mood. 1948. “A Method for Obtaining and Analyzing Sensitivity Data.” *Journal of the American Statistical Association* 43 (241): 109–26. https://doi.org/10.1080/01621459.1948.10483254.

Garcıá-Pérez, Miguel A. 1998. “Forced-Choice Staircases with Fixed Step Sizes: Asymptotic and Small-Sample Properties.” *Vision Research* 38 (12): 1861–81. https://doi.org/10.1016/s0042-6989(97)00340-4.

Kaernbach, Christian. 1991. “Simple Adaptive Testing with the Weighted up-down Method.” *Perception & Psychophysics* 49 (3): 227–29. https://doi.org/10.3758/bf03214307.

Kingdom, Frederick A. A., and Nicolaas Prins. 2016. “Chapter 5 - Adaptive Methods∗.” In *Psychophysics (Second Edition)*, edited by Frederick A. A. Kingdom and Nicolaas Prins, Second Edition, 119–48. San Diego: Academic Press. https://doi.org/10.1016/B978-0-12-407156-8.00005-0.

Kontsevich, Leonid L., and Christopher W. Tyler. 1999. “Bayesian Adaptive Estimation of Psychometric Slope and Threshold.” *Vision Research* 39 (16): 2729–37. https://doi.org/10.1016/s0042-6989(98)00285-5.

Matzke, Dora, Conor V. Dolan, Gordon D. Logan, Scott D. Brown, and Eric-Jan Wagenmakers. 2013. “Bayesian Parametric Estimation of Stop-Signal Reaction

Time Distributions." *Journal of Experimental Psychology: General* 142 (4): 1047–73. https://doi.org/10.1037/a0030543.

Matzke, Dora, Jonathon Love, and Andrew Heathcote. 2016. "A Bayesian Approach for Estimating the Probability of Trigger Failures in the Stop-Signal Paradigm." *Behavior Research Methods* 49 (1): 267–81. https://doi.org/10.3758/s13428-015-0695-8.

Pentland, Alex. 1980. "Maximum Likelihood Estimation: The Best PEST." *Perception & Psychophysics* 28 (4): 377–79. https://doi.org/10.3758/bf03204398.

Prins, N. 2013. "The Psi-Marginal Adaptive Method: How to Give Nuisance Parameters the Attention They Deserve (No More, No Less)." *Journal of Vision* 13 (7): 3–3. https://doi.org/10.1167/13.7.3.

Taylor, M. M., and C. Douglas Creelman. 1967. "PEST: Efficient Estimates on Probability Functions." *The Journal of the Acoustical Society of America* 41 (4A): 782–87. https://doi.org/10.1121/1.1910407.

Thunberg, Christina, Thea Wiker, Carsten Bundt, and René J. Huster. 2024. "On the (Un)reliability of Common Behavioral and Electrophysiological Measures from the Stop Signal Task: Measures of Inhibition Lack Stability over Time." *Cortex* 175 (June): 81–105. https://doi.org/10.1016/j.cortex.2024.02.008.

Verbruggen, Frederick, Adam R Aron, Guido PH Band, Christian Beste, Patrick G Bissett, Adam T Brockett, Joshua W Brown, et al. 2019. "A Consensus Guide to Capturing the Ability to Inhibit Actions and Impulsive Behaviors in the Stop-Signal Task." *eLife*

8 (April). https://doi.org/10.7554/elife.46323.

Watson, Andrew B. 2017. “QUEST+: A General Multidimensional Bayesian Adaptive Psychometric Method.” *Journal of Vision* 17 (3): 10. https://doi.org/10.1167/17.3.10.

Watson, Andrew B., and Denis G. Pelli. 1983. “Quest: A Bayesian Adaptive Psychometric Method.” *Perception & Psychophysics* 33 (2): 113–20. https://doi.org/10.3758/bf03202828.

Wetherill, G. B., and H. Levitt. 1965. “Sequential Estimation of Points on a Psychometric Function.” *British Journal of Mathematical and Statistical Psychology* 18 (1): 1–10. https://doi.org/10.1111/j.2044-8317.1965.tb00689.x.