Emergence of cell-type specific responses to expectation violations

Dimitra Maoutsa

Emergence of cell-type specific responses
to expectation violations

in mouse L2/3 V1

Dimitra Maoutsa

in mouse L2/3 V1

Dimitra Maoutsa

An unenthusiastic introduction to predictive coding and then some

in mouse L2/3 V1

Dimitra Maoutsa

[generated with DALL-E]

familiar/expected/predictable

[generated with DALL-E]

familiar/expected/predictable

novel/unexpected/unpredictable

[generated with DALL-E]

familiar/expected/predictable

novel/unexpected/unpredictable

[generated with DALL-E]

familiar/expected/predictable

novel/unexpected/unpredictable

[generated with DALL-E]

familiar/expected/predictable

novel/unexpected/unpredictable

[generated with DALL-E]

Novelty: response to information that is not expected or predicted in a given context on the basis of prior experience. Note that we distinguish here between two types of novelty (Box 2): unrelated information that does not strongly match any schema, and incongruent information that is inconsistent with a dominant schema. Within the present SLIMM framework, only the latter improves memory, and note that this type of novelty cannot exist without a schema (i.e. the two concepts are intimately related).

what are predictions and expectations?

Novelty: response to information that is not expected or predicted in a given context on the basis of prior experience. Note that we distinguish here between two types of novelty (Box 2): unrelated information that does not strongly match any schema, and incongruent information that is inconsistent with a dominant schema. Within the present SLIMM framework, only the latter improves memory, and note that this type of novelty cannot exist without a schema (i.e. the two concepts are intimately related).

what are predictions and expectations?

Novelty: response to information that is not expected or predicted in a given context on the basis of prior experience. Note that we distinguish here between two types of novelty (Box 2): unrelated information that does not strongly match any schema, and incongruent information that is inconsistent with a dominant schema. Within the present SLIMM framework, only the latter improves memory, and note that this type of novelty cannot exist without a schema (i.e. the two concepts are intimately related).

prediction

noun [C or U]

UK /prɪˈdɪk.ʃᵊn/ US /prɪˈdɪk.ʃᵊn/

B2

a statement about what you think will happen in the future:

what are predictions and expectations?

Novelty: response to information that is not expected or predicted in a given context on the basis of prior experience. Note that we distinguish here between two types of novelty (Box 2): unrelated information that does not strongly match any schema, and incongruent information that is inconsistent with a dominant schema. Within the present SLIMM framework, only the latter improves memory, and note that this type of novelty cannot exist without a schema (i.e. the two concepts are intimately related).

prediction

noun [C or U]

UK /prɪˈdɪk.ʃᵊn/ **US** /prɪˈdɪk.ʃᵊn/

B2

a statement about what you think will happen in the future:

what are predictions and expectations?

Novelty: response to information that is not expected or predicted in a given context on the basis of prior experience. Note that we distinguish here between two types of novelty (Box 2): unrelated information that does not strongly match any schema, and incongruent information that is inconsistent with a dominant schema. Within the present SLIMM framework, only the latter improves memory, and note that this type of novelty cannot exist without a schema (i.e. the two concepts are intimately related).

prediction

noun [C or U]

UK /prɪˈdɪk.ʃən/ US /prɪˈdɪk.ʃən/

B2

a statement about what you think will happen in the future:

predict verb

pre·dict pri-ˈdikt

predicted; predicting; predicts

Synonyms of *predict* >

transitive verb

: to declare or indicate in advance

especially **:** foretell on the basis of observation, experience, or scientific reason

what are predictions and expectations?

Novelty: response to information that is not expected or predicted in a given context on the basis of prior experience. Note that we distinguish here between two types of novelty (Box 2): unrelated information that does not strongly match any schema, and incongruent information that is inconsistent with a dominant schema. Within the present SLIMM framework, only the latter improves memory, and note that this type of novelty cannot exist without a schema (i.e. the two concepts are intimately related).

prediction

noun [C or U]

UK /prɪˈdɪk.ʃən/ US /prɪˈdɪk.ʃən/

B2

a statement about what you think will happen in the future:

predict verb

pre·dict pri-ˈdikt

predicted; predicting; predicts

Synonyms of *predict* >

transitive verb

: to declare or indicate in advance

especially **:** foretell on the basis of observation, experience, or scientific reason

what are predictions and expectations?

prediction

noun [C or U]

UK /prɪˈdɪk.ʃən/ US /prɪˈdɪk.ʃən/

B2

a statement about what you think will happen in the future:

predict verb

pre·dict pri-ˈdikt

predicted; predicting; predicts

Synonyms of *predict* >

transitive verb

: to declare or indicate in advance

especially : foretell on the basis of observation, experience, or scientific reason

expect verb

ex·pect ik-ˈspekt

expected; expecting; expects

Synonyms of *expect* >

transitive verb

1 a : to consider probable or certain

expect to be forgiven

expect that things will improve

b : to consider reasonable, due, or necessary

expected hard work from the students

c : to consider bound in duty or obligated

they *expect* you to pay your bills

2 : to anticipate or look forward to the coming or occurrence of

we *expect* them any minute now

expected a telephone call

Novelty: response to information that is not expected or predicted in a given context on the basis of prior experience. Note that we distinguish here between two types of novelty (Box 2): unrelated information that does not strongly match any schema, and incongruent information that is inconsistent with a dominant schema. Within the present SLIMM framework, only the latter improves memory, and note that this type of novelty cannot exist without a schema (i.e. the two concepts are intimately related).

what are predictions and expectations?

prediction

noun [C or U]

UK /prɪˈdɪk.ʃən/ US /prɪˈdɪk.ʃən/

B2

a statement about what you think will happen in the future:

predict *verb*

pre·dict pri-ˈdikt

predicted; predicting; predicts

Synonyms of *predict* >

transitive verb

: to declare or indicate in advance

especially : foretell on the basis of observation, experience, or scientific reason

expect *verb*

ex·pect ik-ˈspekt

expected; expecting; expects

Synonyms of *expect* >

transitive verb

1 **a** : to consider probable or certain

expect to be forgiven

expect that things will improve

b : to consider reasonable, due, or necessary

expected hard work from the students

c : to consider bound in duty or obligated

they *expect* you to pay your bills

2 : to anticipate or look forward to the coming or occurrence of

we *expect* them any minute now

expected a telephone call

Novelty: response to information that is not expected or predicted in a given context on the basis of prior experience. Note that we distinguish here between two types of novelty (Box 2): unrelated information that does not strongly match any schema, and incongruent information that is inconsistent with a dominant schema. Within the present SLIMM framework, only the latter improves memory, and note that this type of novelty cannot exist without a schema (i.e. the two concepts are intimately related).

what are predictions and expectations?

prediction

noun [C or U]

UK /prɪˈdɪk.ʃən/ US /prɪˈdɪk.ʃən/

B2

a statement about what you think will happen in the future:

predict verb

pre·dict pri-ˈdikt

predicted; predicting; predicts

Synonyms of *predict* >

transitive verb

: to declare or indicate in advance

especially : foretell on the basis of observation, experience, or scientific reason

expect verb

ex·pect ik-ˈspekt

expected; expecting; expects

Synonyms of *expect* >

transitive verb

1 a : to consider probable or certain

expect to be forgiven

expect that things will improve

b : to consider reasonable, due, or necessary

expected hard work from the students

c : to consider bound in duty or obligated

they *expect* you to pay your bills

2 : to anticipate or look forward to the coming or occurrence of

we *expect* them any minute now

expected a telephone call

Novelty: response to information that is not expected or predicted in a given context on the basis of prior experience. Note that we distinguish here between two types of novelty (Box 2): unrelated information that does not strongly match any schema, and incongruent information that is inconsistent with a dominant schema. Within the present SLIMM framework, only the latter improves memory, and note that this type of novelty cannot exist without a schema (i.e. the two concepts are intimately related).

what are predictions and expectations?

Novelty: response to information that is not expected or predicted in a given context on the basis of prior experience. Note that we distinguish here between two types of novelty (Box 2): unrelated information that does not strongly match any schema, and incongruent information that is inconsistent with a dominant schema. Within the present SLIMM framework, only the latter improves memory, and note that this type of novelty cannot exist without a schema (i.e. the two concepts are intimately related).

prediction

noun [C or U]

UK /prɪˈdɪk.ʃən/ US /prɪˈdɪk.ʃən/

B2

a statement about what you think will happen in the future:

predict verb

pre·dict pri-ˈdikt

predicted; predicting; predicts

Synonyms of *predict* >

transitive verb

: to declare or indicate in advance

especially : foretell on the basis of observation, experience, or scientific reason

expect verb

ex·pect ik-ˈspekt

expected; expecting; expects

Synonyms of *expect* >

transitive verb

1 a : to consider probable or certain

expect to be forgiven

expect that things will improve

b : to consider reasonable, due, or necessary

expected hard work from the students

c : to consider bound in duty or obligated

they *expect* you to pay your bills

2 : to anticipate or look forward to the coming or occurrence of

we *expect* them any minute now

expected a telephone call

what are predictions and expectations?

prediction

noun [C or U]

UK /prɪˈdɪk.ʃən/ US /prɪˈdɪk.ʃən/

B2

a statement about what you think will happen in the future:

predict verb

pre·dict pri-ˈdikt

predicted; predicting; predicts

Synonyms of *predict* >

transitive verb

: to declare or indicate in advance

especially : foretell on the basis of observation, experience, or scientific reason

Novelty: response to information that is not expected or predicted in a given context on the basis of prior experience. Note that we distinguish here between two types of novelty (Box 2): unrelated information that does not strongly match any schema, and incongruent information that is inconsistent with a dominant schema. Within the present SLIMM framework, only the latter improves memory, and note that this type of novelty cannot exist without a schema (i.e. the two concepts are intimately related).

expect verb

ex·pect ik-ˈspekt

expected; expecting; expects

Synonyms of *expect* >

transitive verb

1 a : to consider probable or certain

expect to be forgiven

expect that things will improve

b : to consider reasonable, due, or necessary

expected hard work from the students

c : to consider bound in duty or obligated

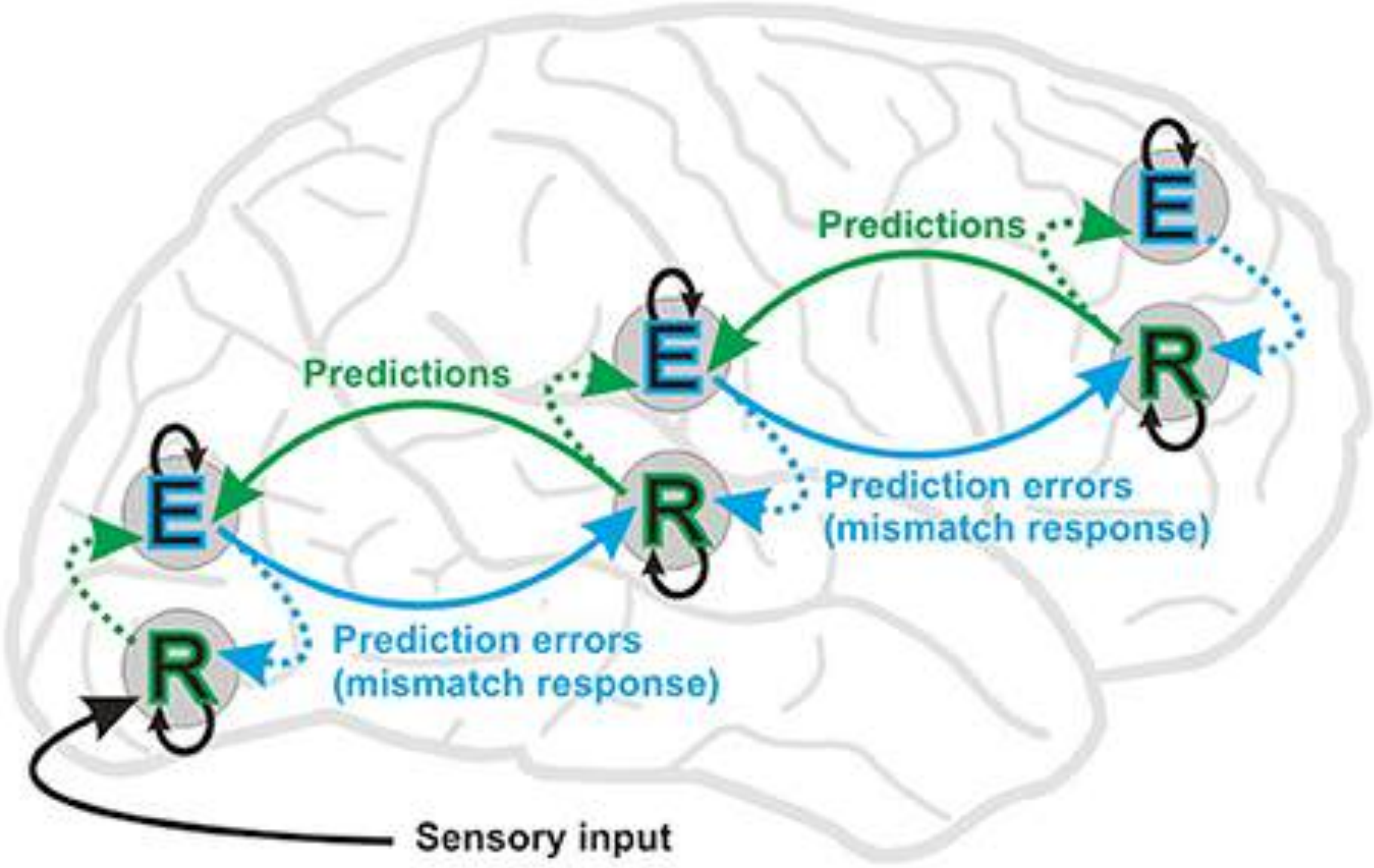
they *expect* you to pay your bills

2 : to anticipate or look forward to the coming or occurrence of

we *expect* them any minute now

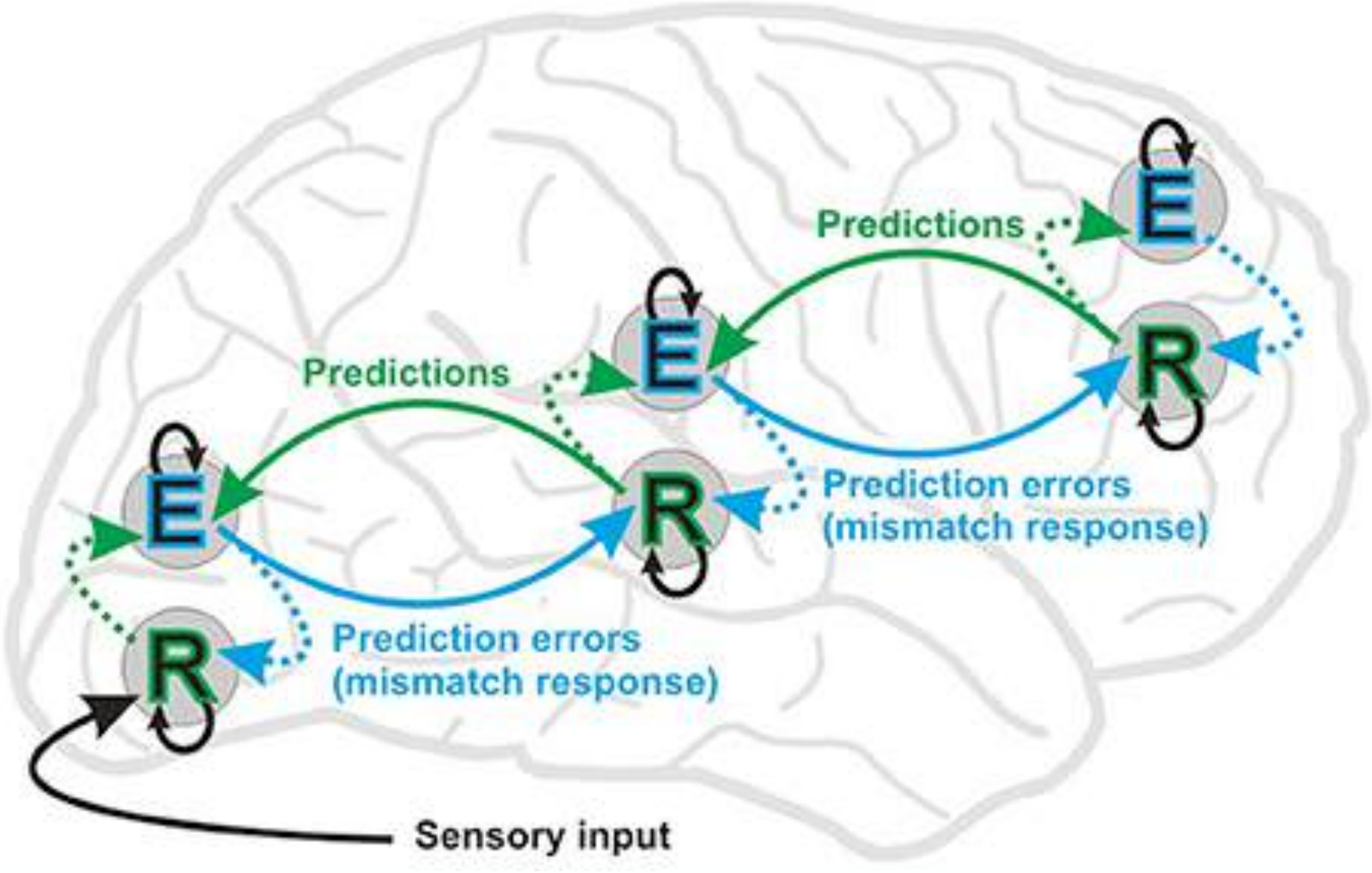
expected a telephone call

implies some sort of stored knowledge or memory that generates these predictions



- cortex maintains **internal representations** of the external world
- internal models generate **predictions of sensory inputs**
- internal models are **refined** based on **prediction errors** between predicted and actual input
- **predictions** generated in downstream cortical areas and **communicated through cortico-cortical projections**

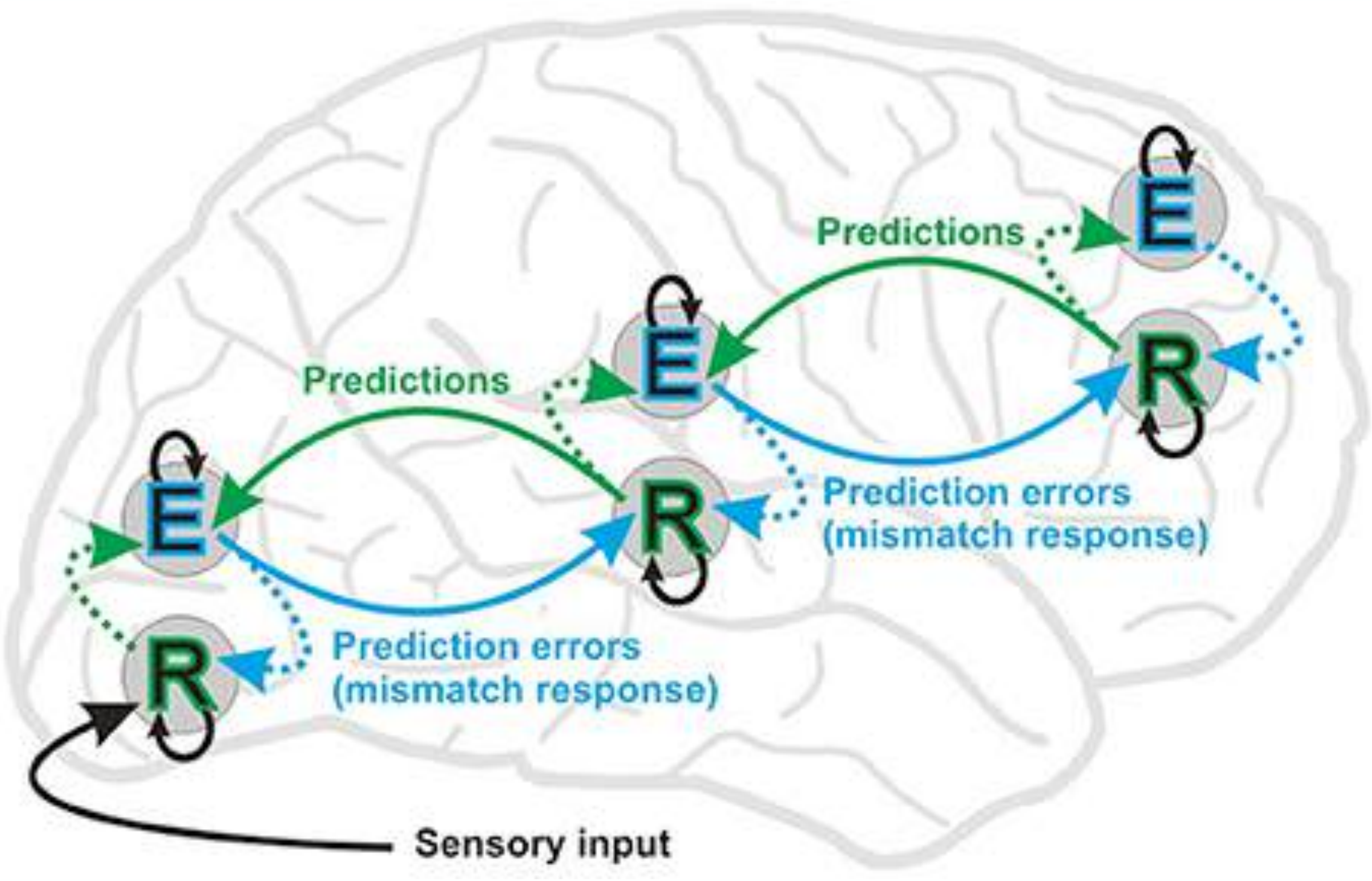
how are predictions generated?



- cortex maintains **internal representations** of the external world
- internal models generate **predictions of sensory inputs**
- internal models are **refined** based on **prediction errors** between predicted and actual input
- **predictions** generated in downstream cortical areas and **communicated through cortico-cortical projections**

how are predictions generated?

predictive coding theory



- cortex maintains **internal representations** of the external world
- internal models generate **predictions of sensory inputs**
- internal models are **refined** based on **prediction errors** between predicted and actual input
- **predictions** generated in downstream cortical areas and **communicated through cortico-cortical projections**

[Stefanics et al., 2014]

Sensory cortices under a predictive coding lens

Expectations and predictions

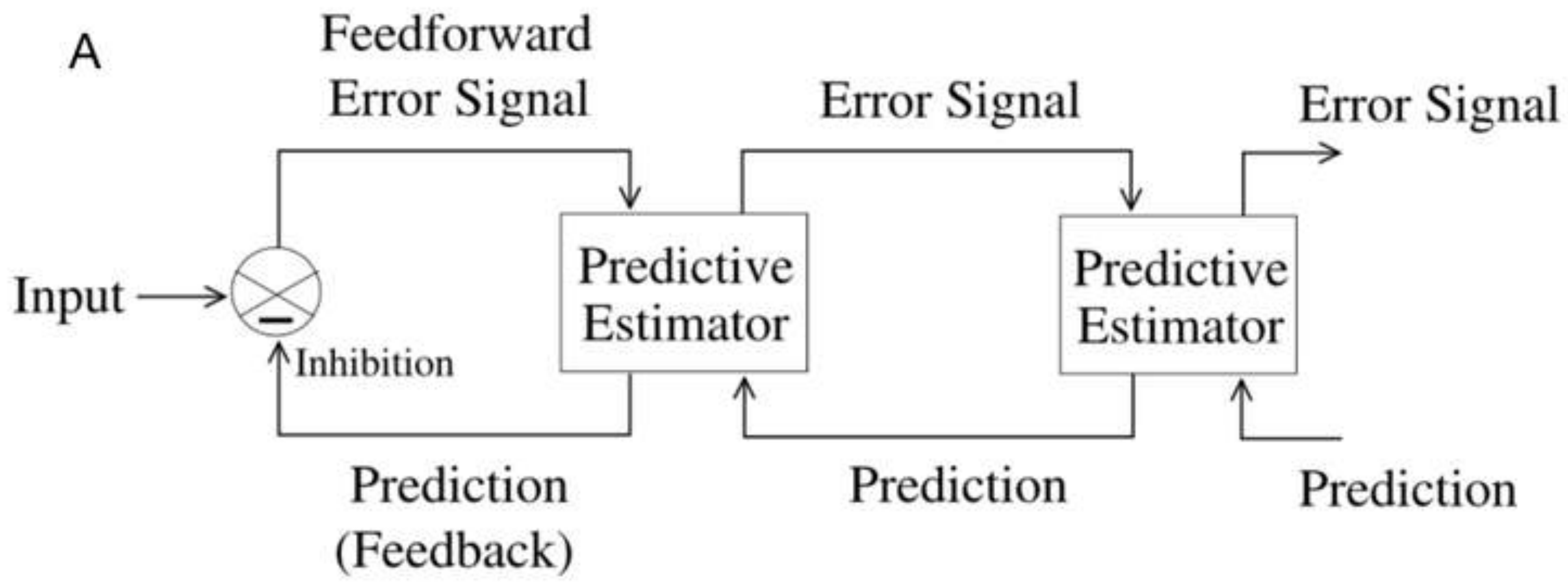
- **Internal models** used to **predict future behaviour**
- Prediction error = Sensory input - Prediction (from internal model)
- Prediction errors used to refine an internal model of the world

- the top-down predictions convey prior beliefs based on learned expectations, while the boom-up prediction errors carry evidence from the current input.

But who provides the internal model -> **hierarchical predictive coding**

Hierarchical predictive coding

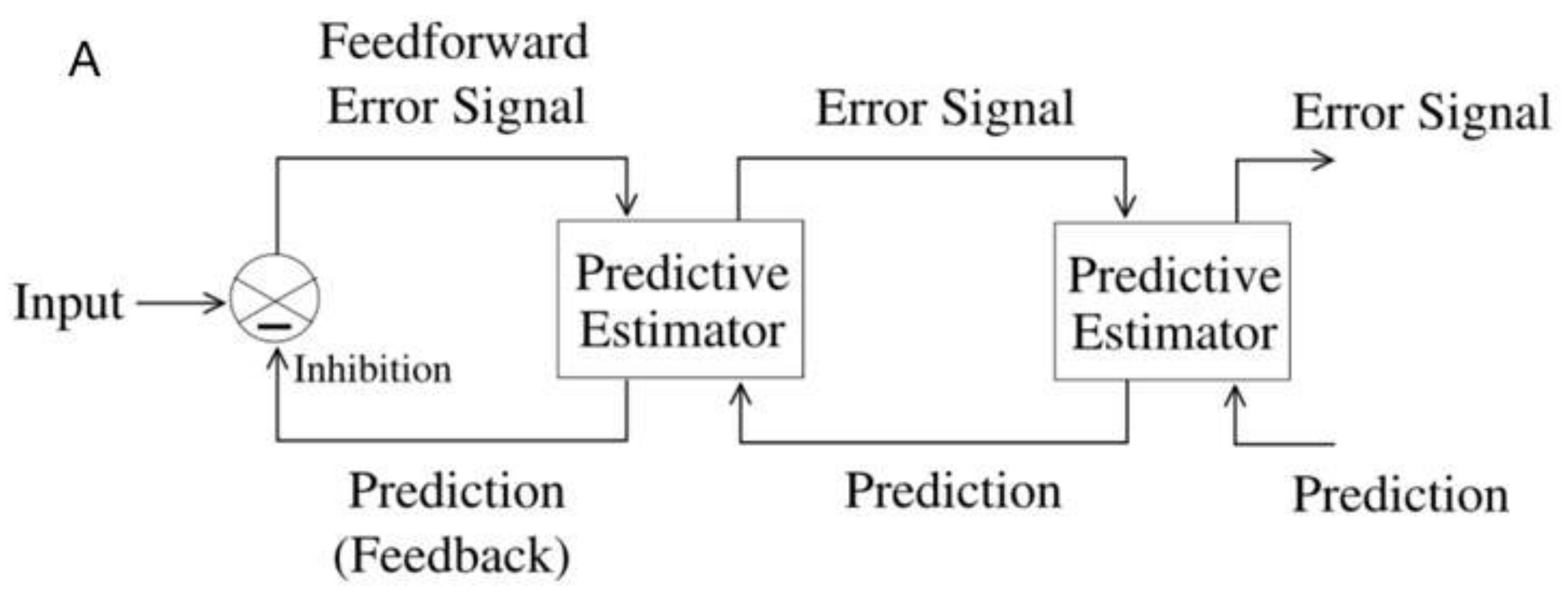
Prediction errors



- Are they transmitted through feedback
- Or are they represented within V1?
- Each level of visual hierarchy is engaged in a predictive loop.
- Higher order areas send predictions about expected sensory input to lower-level areas, which compare this prediction with actual sensory data
    - Prediction: higher visual areas generate predictions on what should be seen based on prior knowledge
    - Prediction error: lowest areas compute the difference between actual in[ut and prediction -> prediction error (prediction error as residual activity between prediction and sensory input)
- Rather than transmitting the full sensory input neuron in the vis cortex communicate mainly prediction errors up the hierarchy, allowing the brain to focus on unexpected or novel stimuli
- Rao: that layer 2/3 neurons compute and convey the prediction errors while neurons in the deeper layers 5/6 maintain the state estimate

Hierarchical predictive coding

Prediction errors



- Are they transmitted through feedback
- Or are they represented within V1?
- Each level of visual hierarchy is engaged in a predictive loop.
- Higher order areas send predictions about expected sensory input to lower-level areas, which compare this prediction with actual sensory data
  - Prediction: higher visual areas generate predictions on what should be seen based on prior knowledge
  - Prediction error: lowest areas compute the difference between actual in[ut and prediction -> prediction error (prediction error as residual activity between prediction and sensory input)
- Rather than transmitting the full sensory input neuron in the vis cortex communicate mainly prediction errors up the hierarchy, allowing the brain to focus on unexpected or novel stimuli
- Rao: that layer 2/3 neurons compute and convey the prediction errors while neurons in the deeper layers 5/6 maintain the state estimate

[Jiang & Rao, 2023]

Two theories for top-down feedback

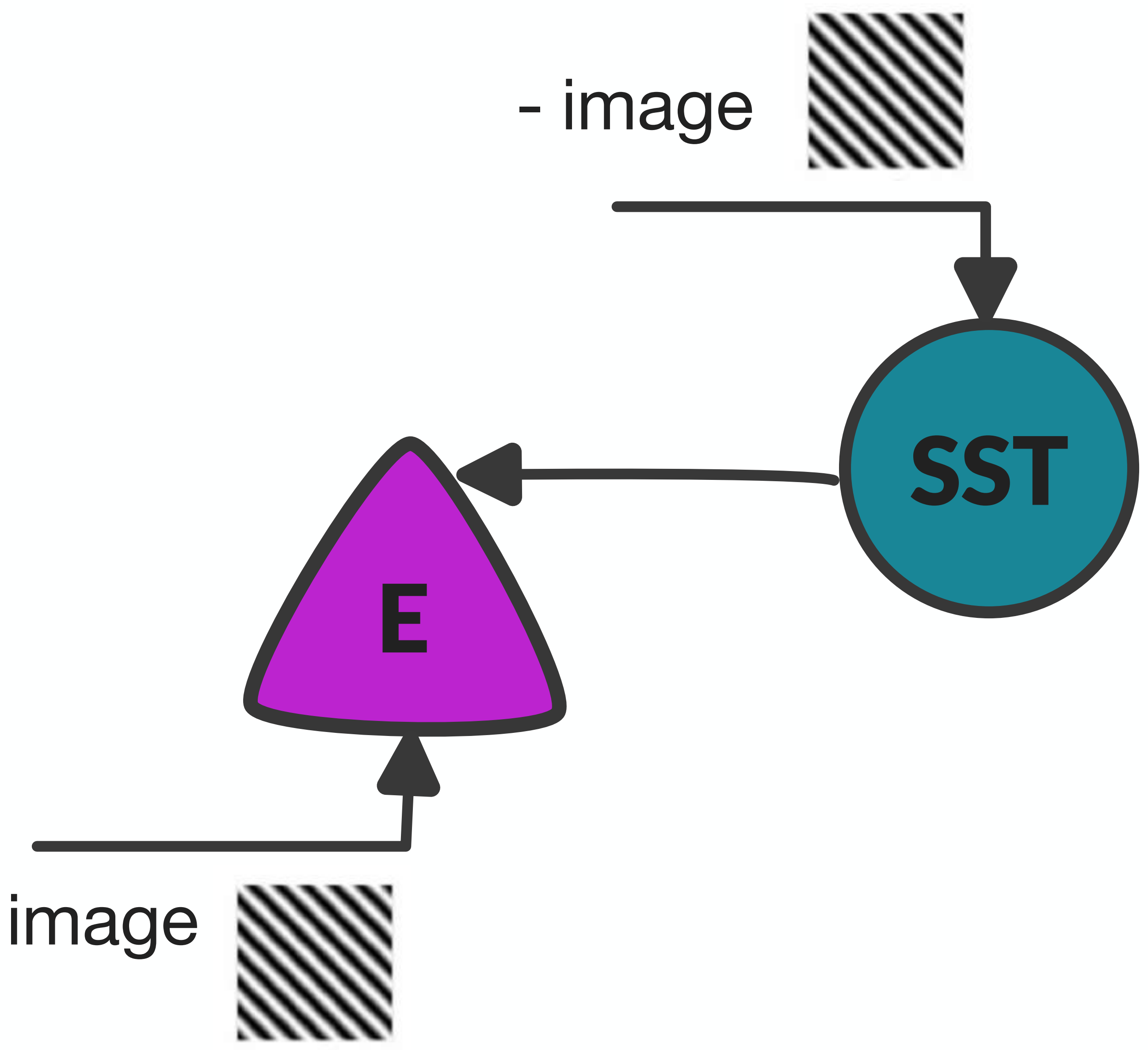
either predictive signal or prediction error

- Negative image hypothesis

Two theories for top-down feedback

either predictive signal or prediction error

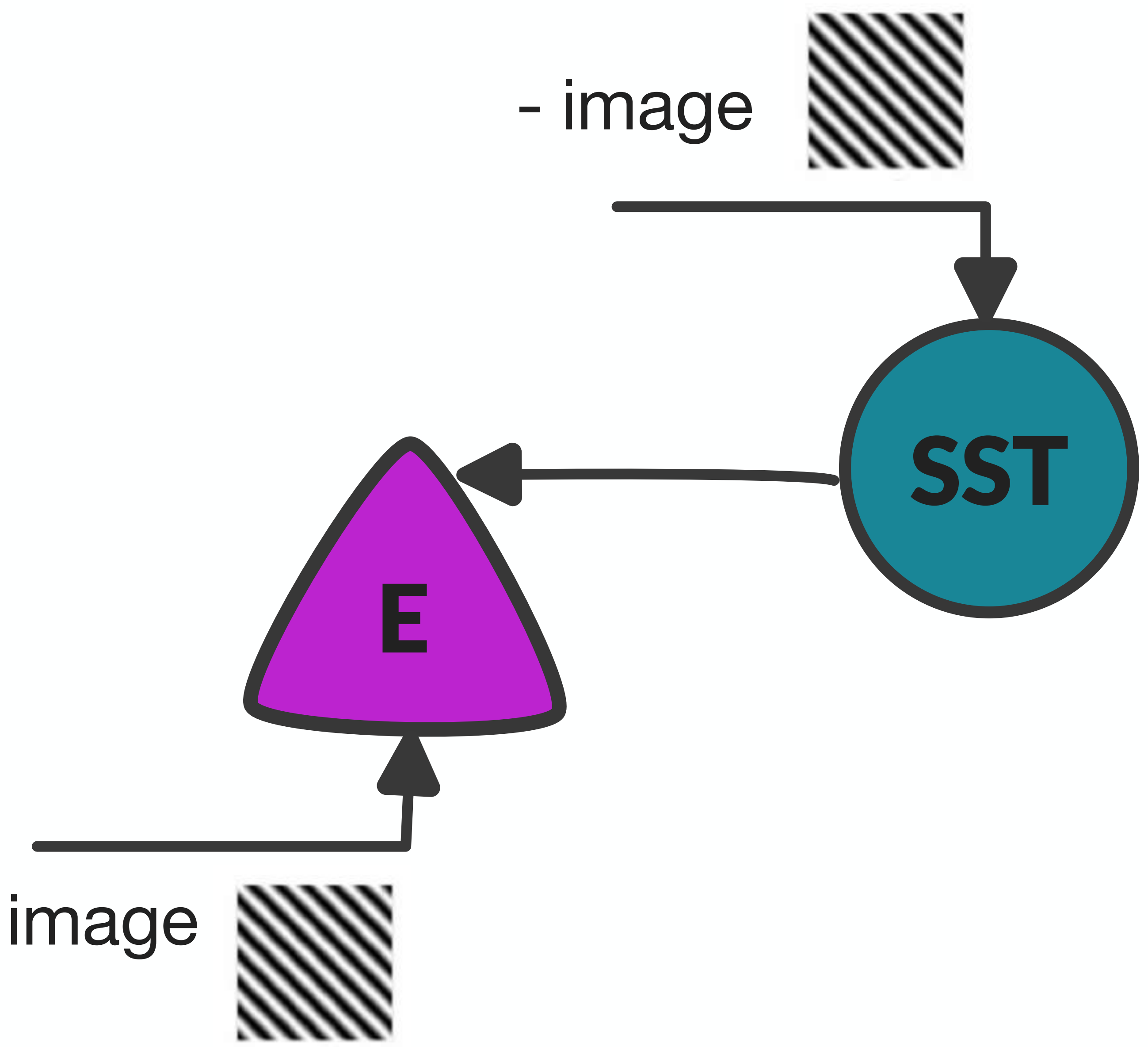
- Negative image hypothesis



Two theories for top-down feedback

either predictive signal or prediction error

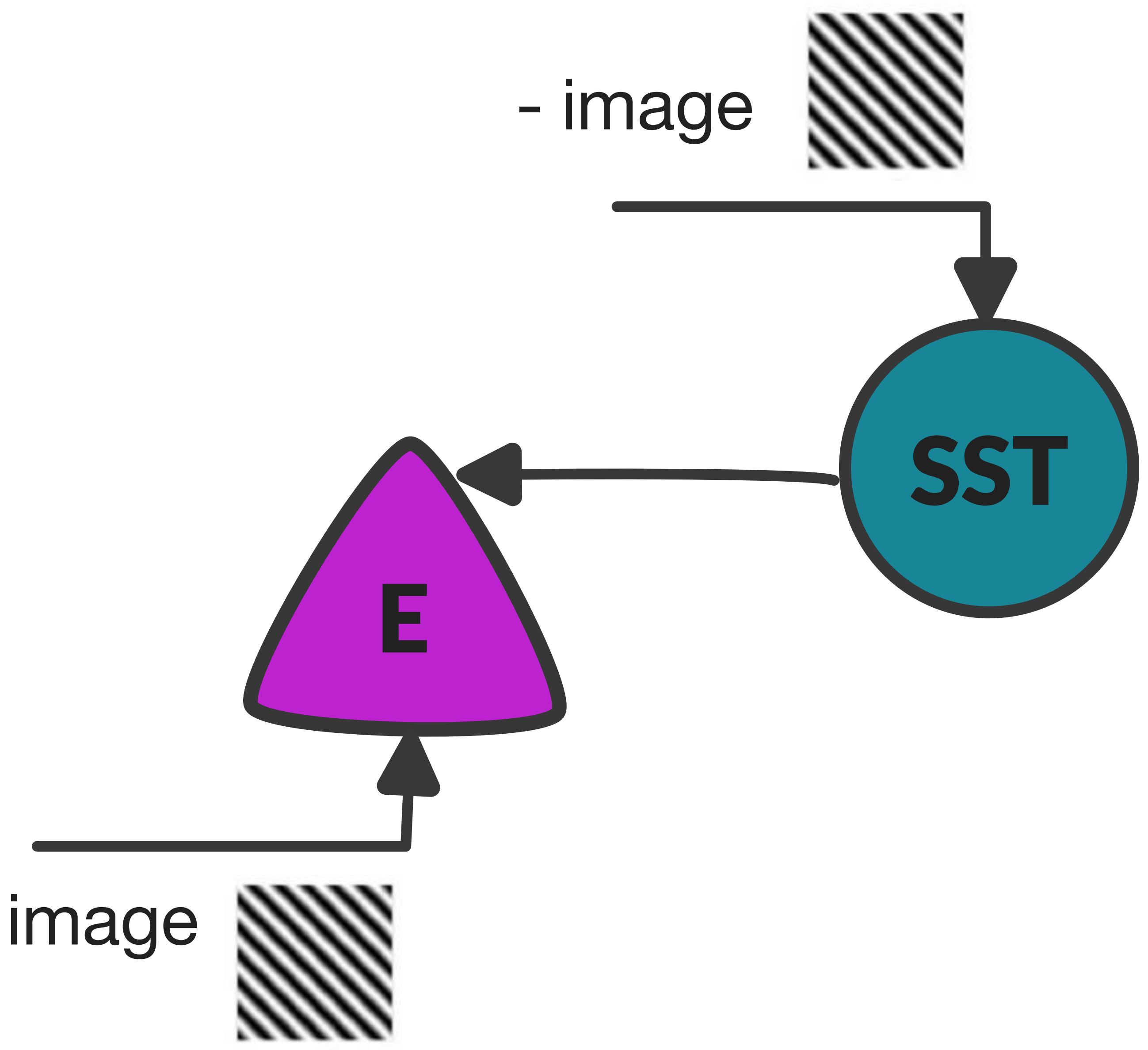
- Negative image hypothesis
- Prediction error hypothesis



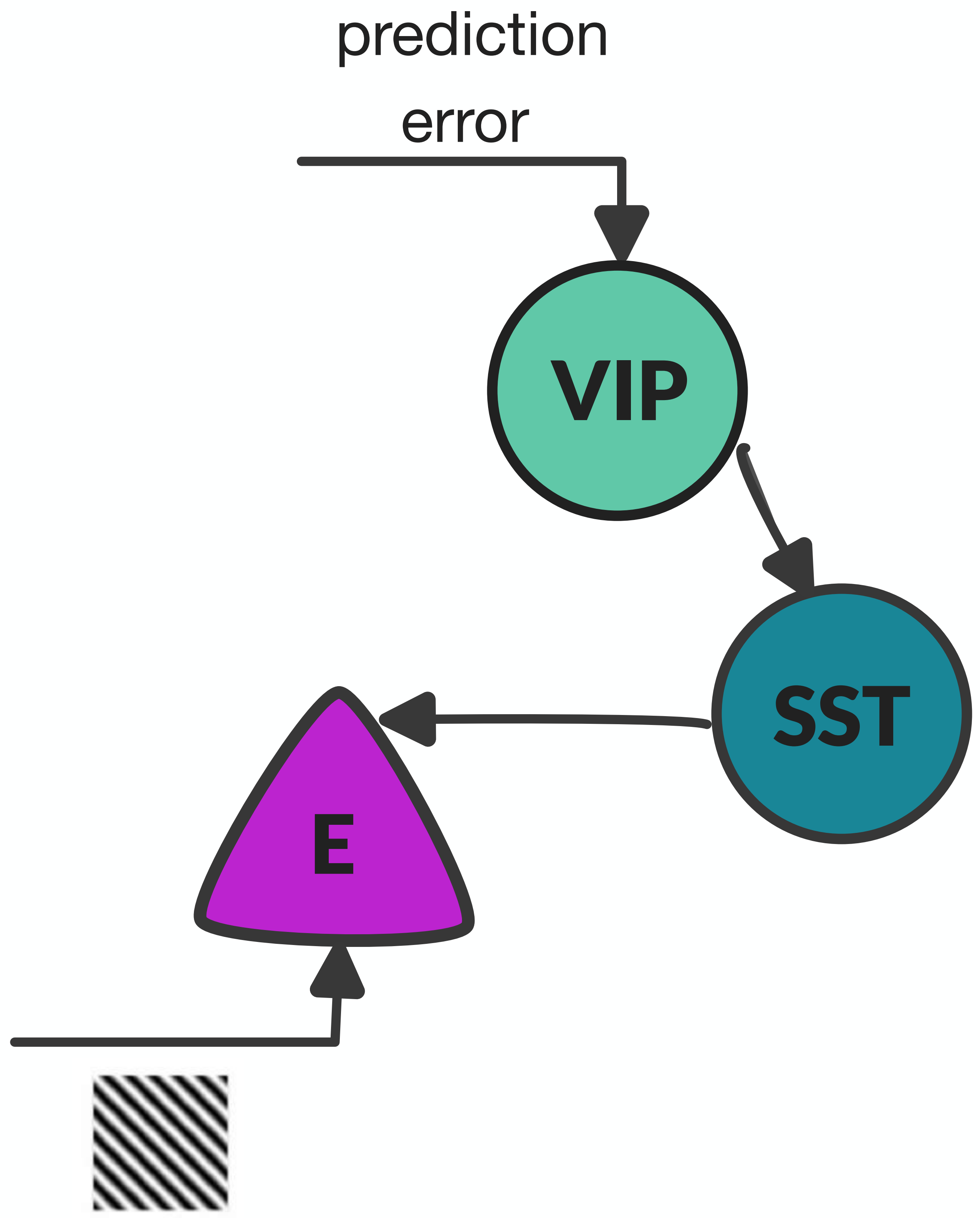
Two theories for top-down feedback

either predictive signal or prediction error

- Negative image hypothesis



- Prediction error hypothesis



What is nature of prediction error signals?

- Discrepancy between prediction and sensory input - [Rao Ballard]
- Pronounced unexpected input [Furutachi 2024]

- Prediction errors may guide learning [Gillon et al., 2021]

Perception as inference

binocular rivalry & prior bias

if perception is only bottom up, why would
perception of same stimulus
alternate between two different interpretations?

Feedback from higher order areas changes operating mode of cortical circuits

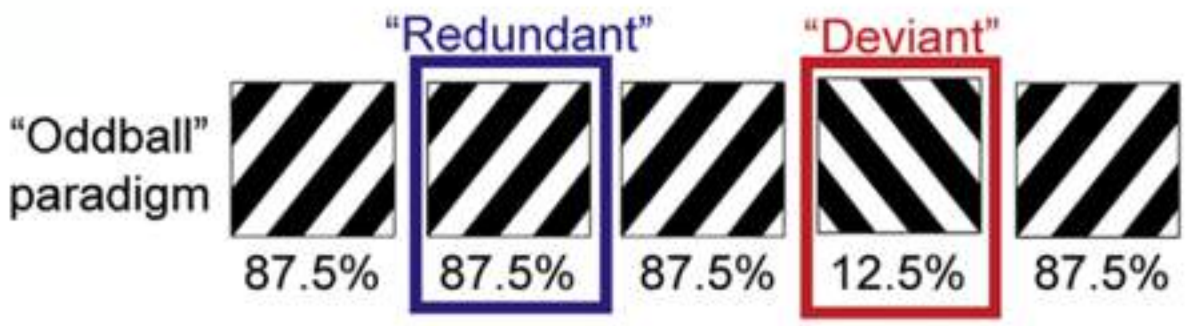
Perception as inference

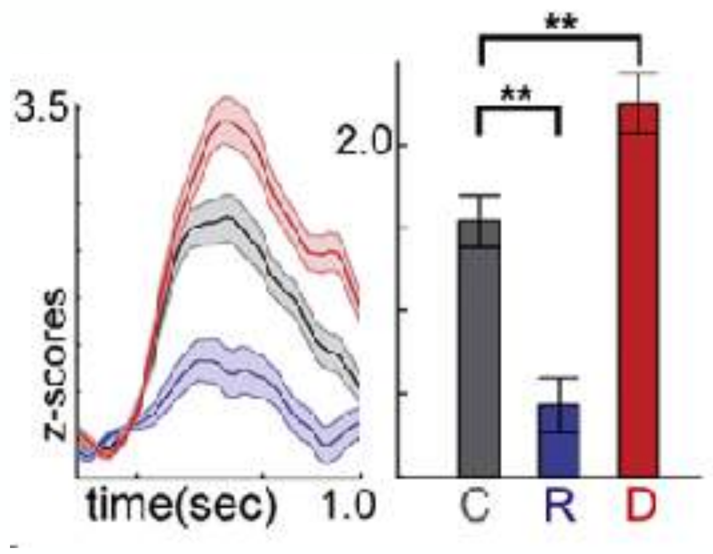
binocular rivalry & prior bias

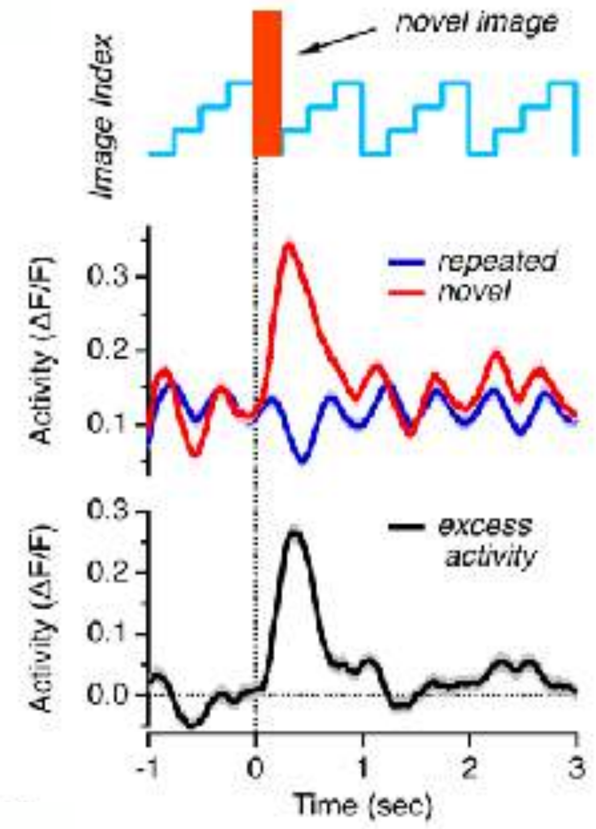
if perception is only bottom up, why would
perception of same stimulus
alternate between two different interpretations?

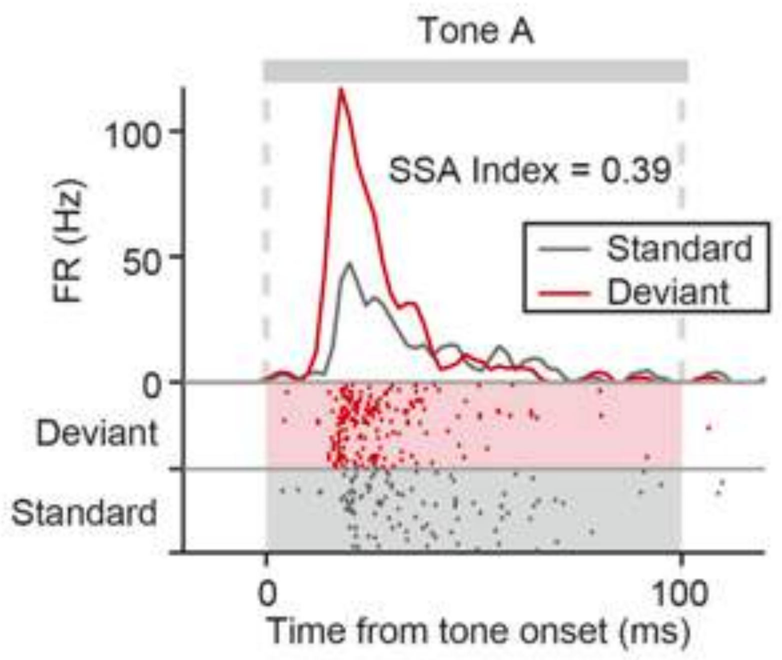
Feedback from higher order
areas changes operating mode
of cortical circuits

Representation suppression or stimulus-specific adaptation

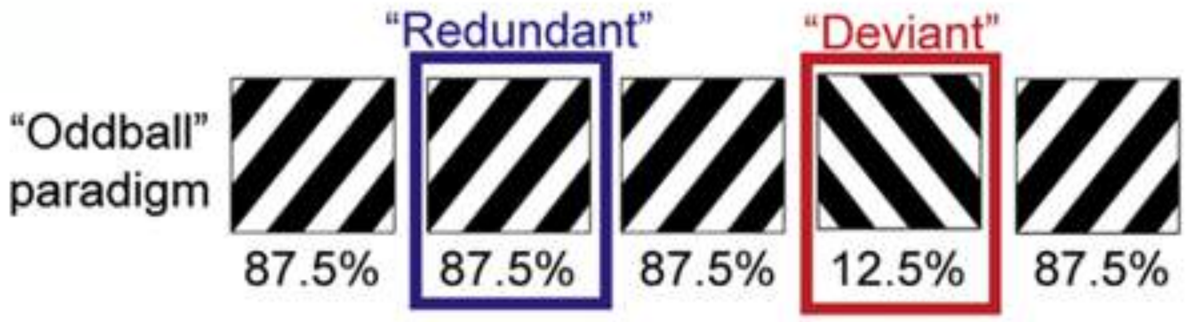


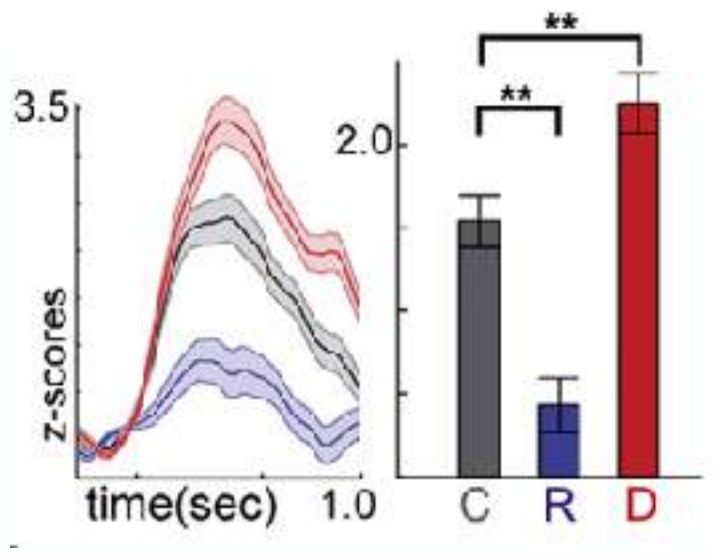


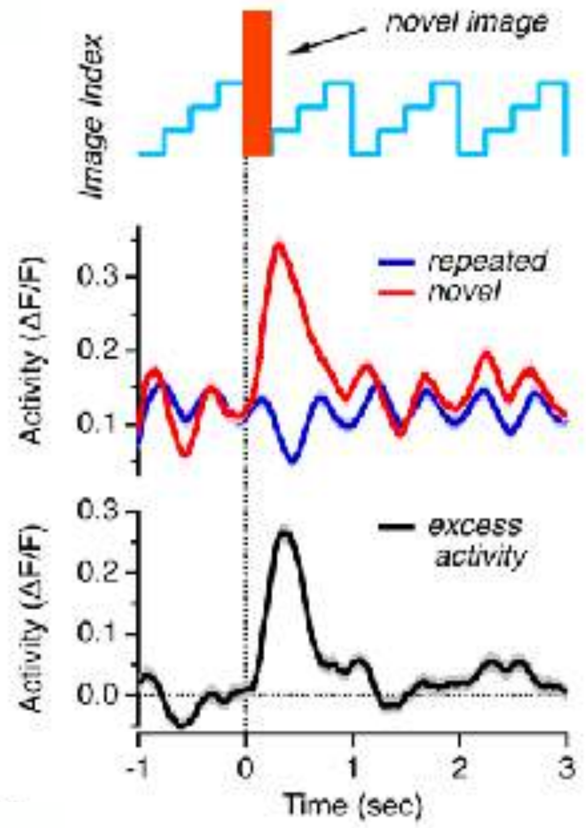


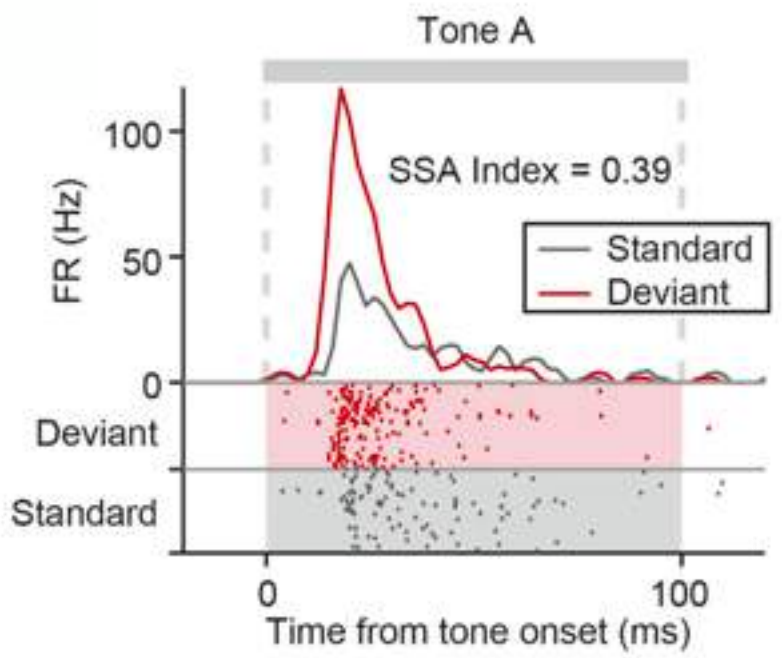


Representation suppression or stimulus-specific adaptation



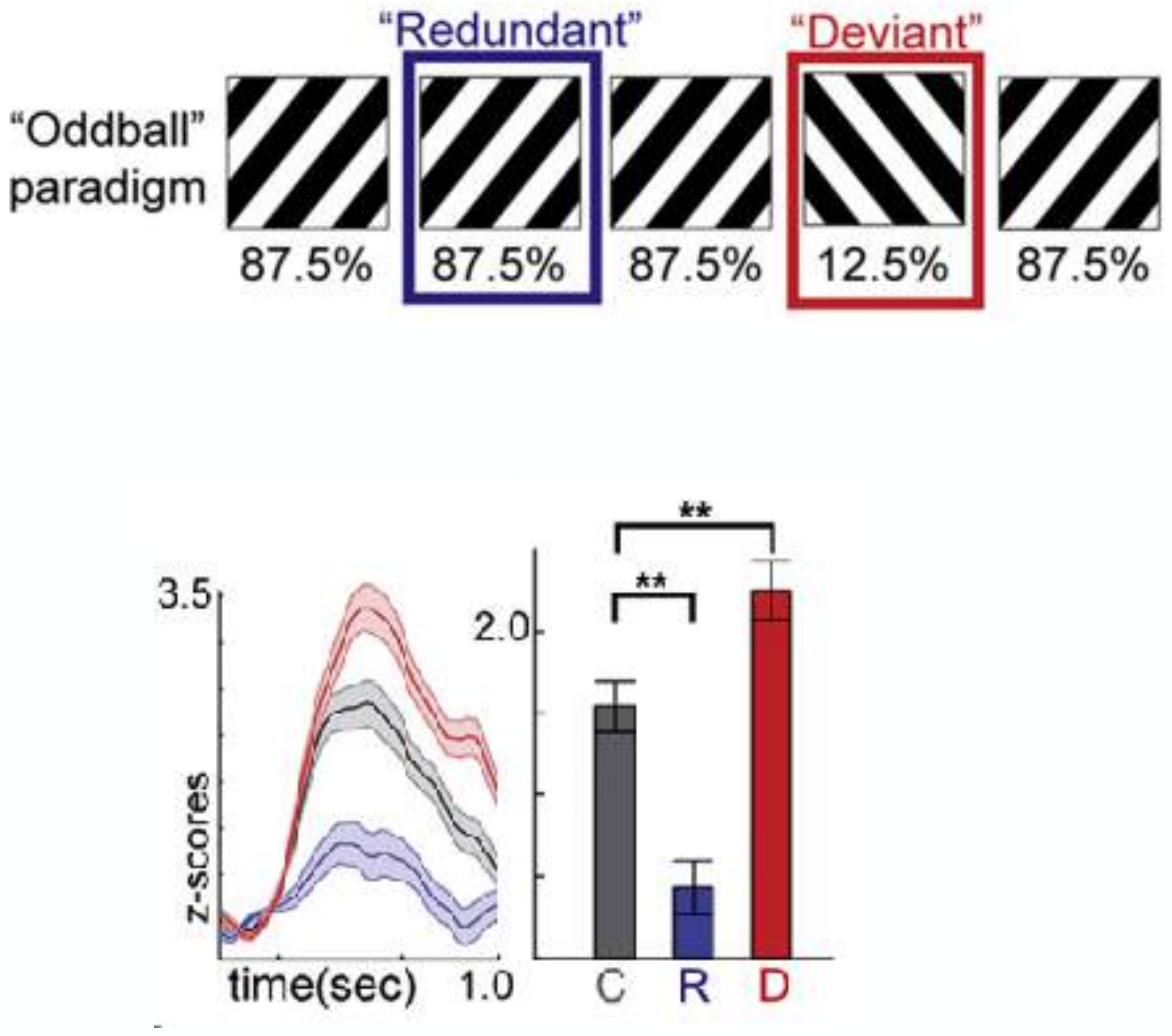


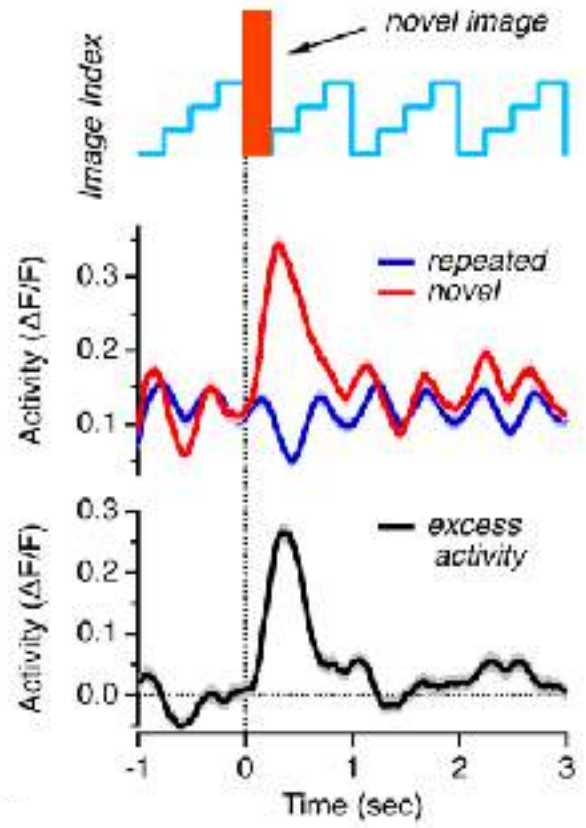


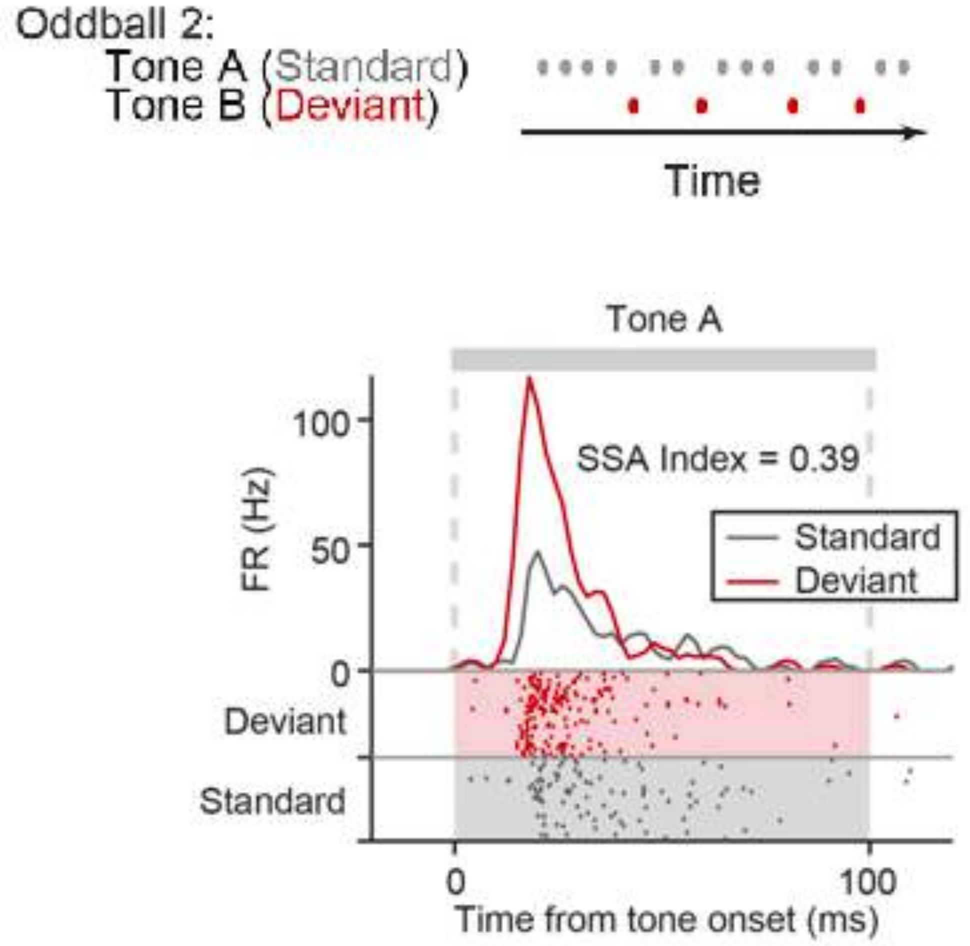


mouse - visual cortex
[Hamm & Yuste, 2016]

Representation suppression or stimulus-specific adaptation



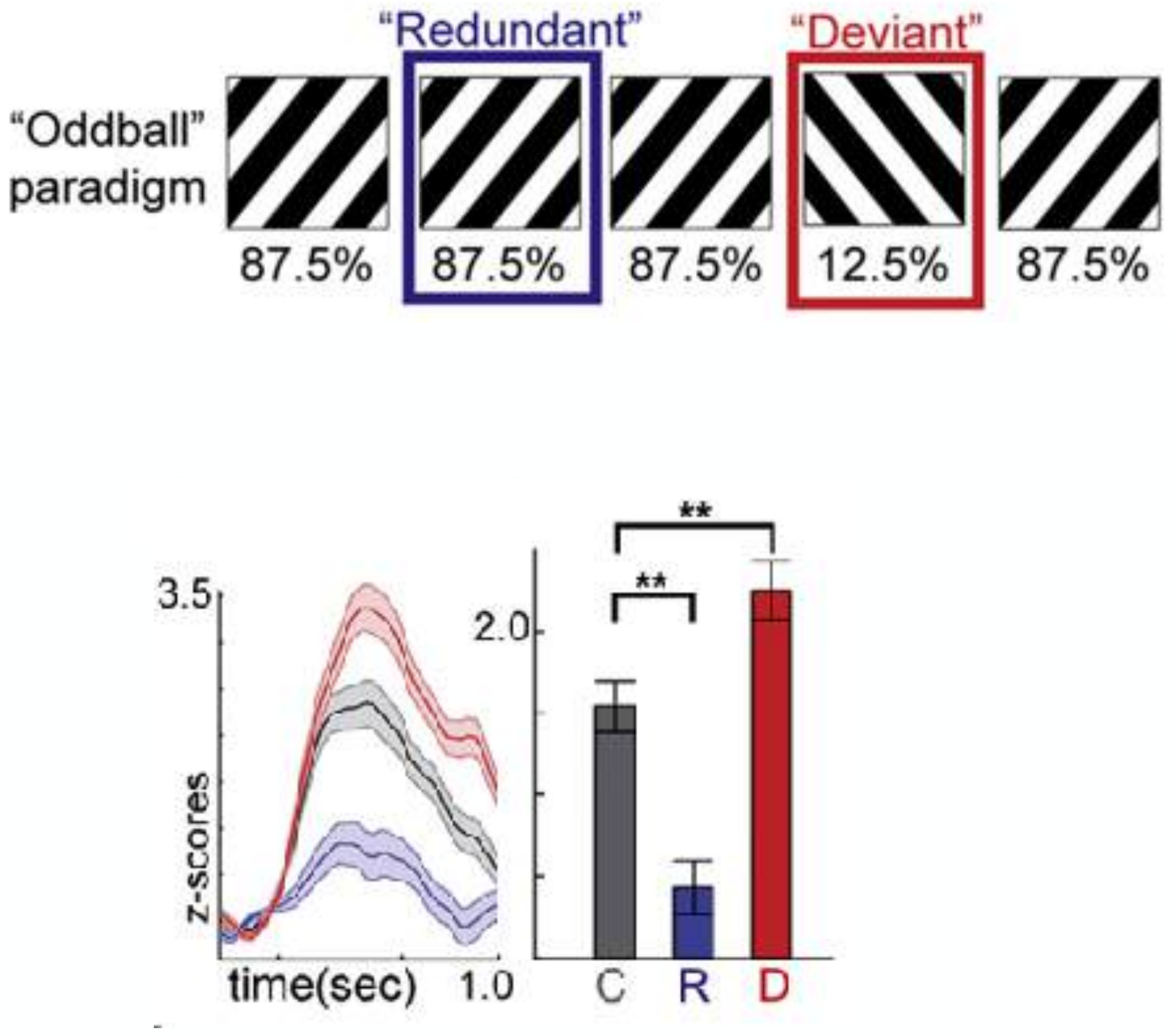




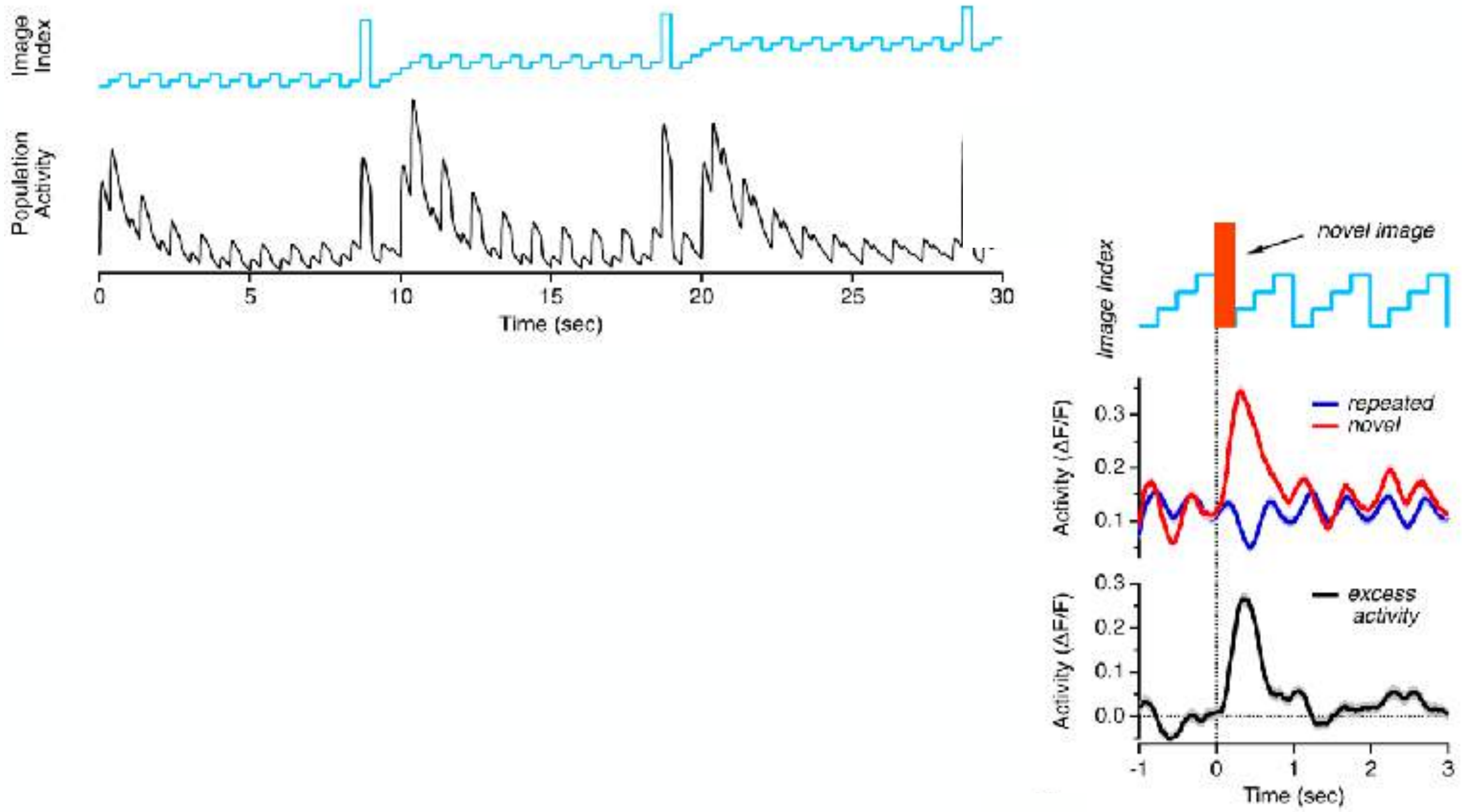
mouse - visual cortex
[Hamm & Yuste, 2016]

mouse - auditory cortex
[Natan et al., 2015]

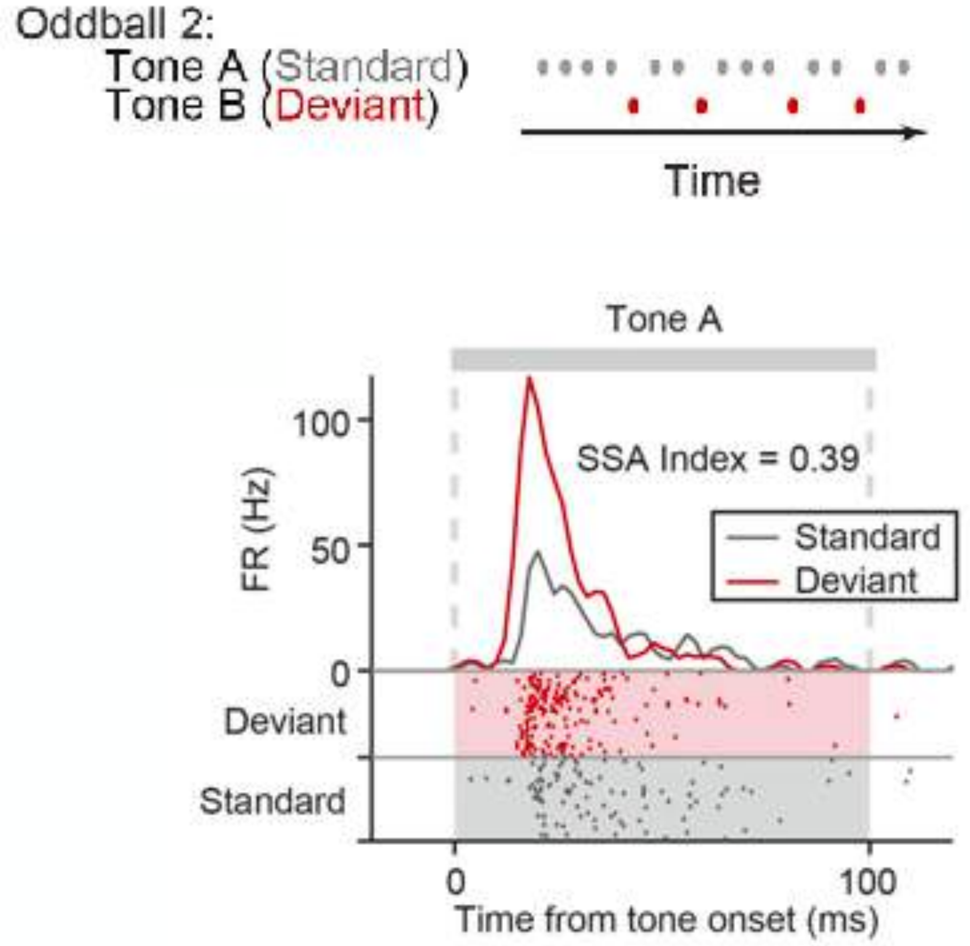
Representation suppression or stimulus-specific adaptation



mouse - visual cortex
[Hamm & Yuste, 2016]

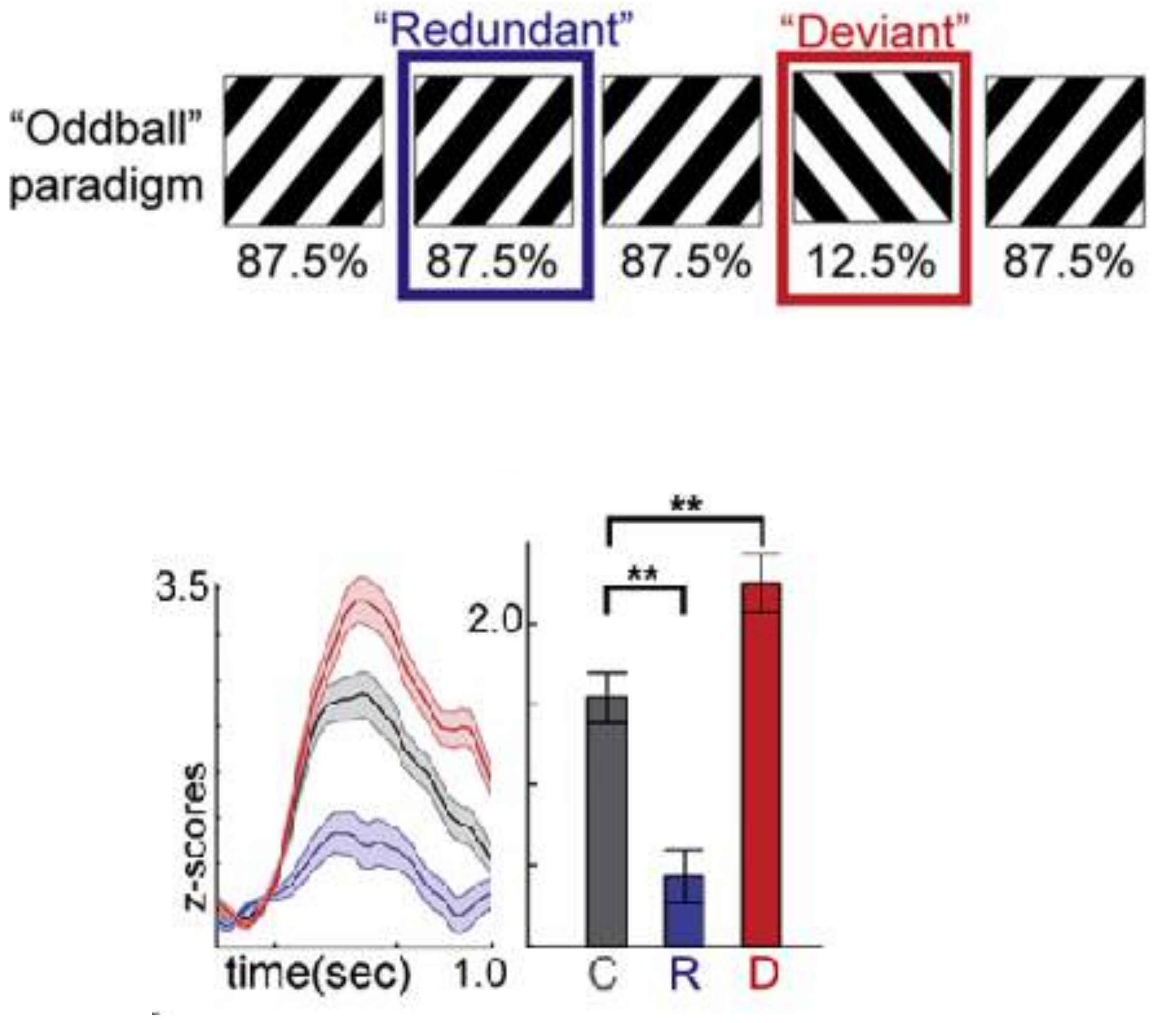


mouse - visual cortex
[Homann et al., 2021]



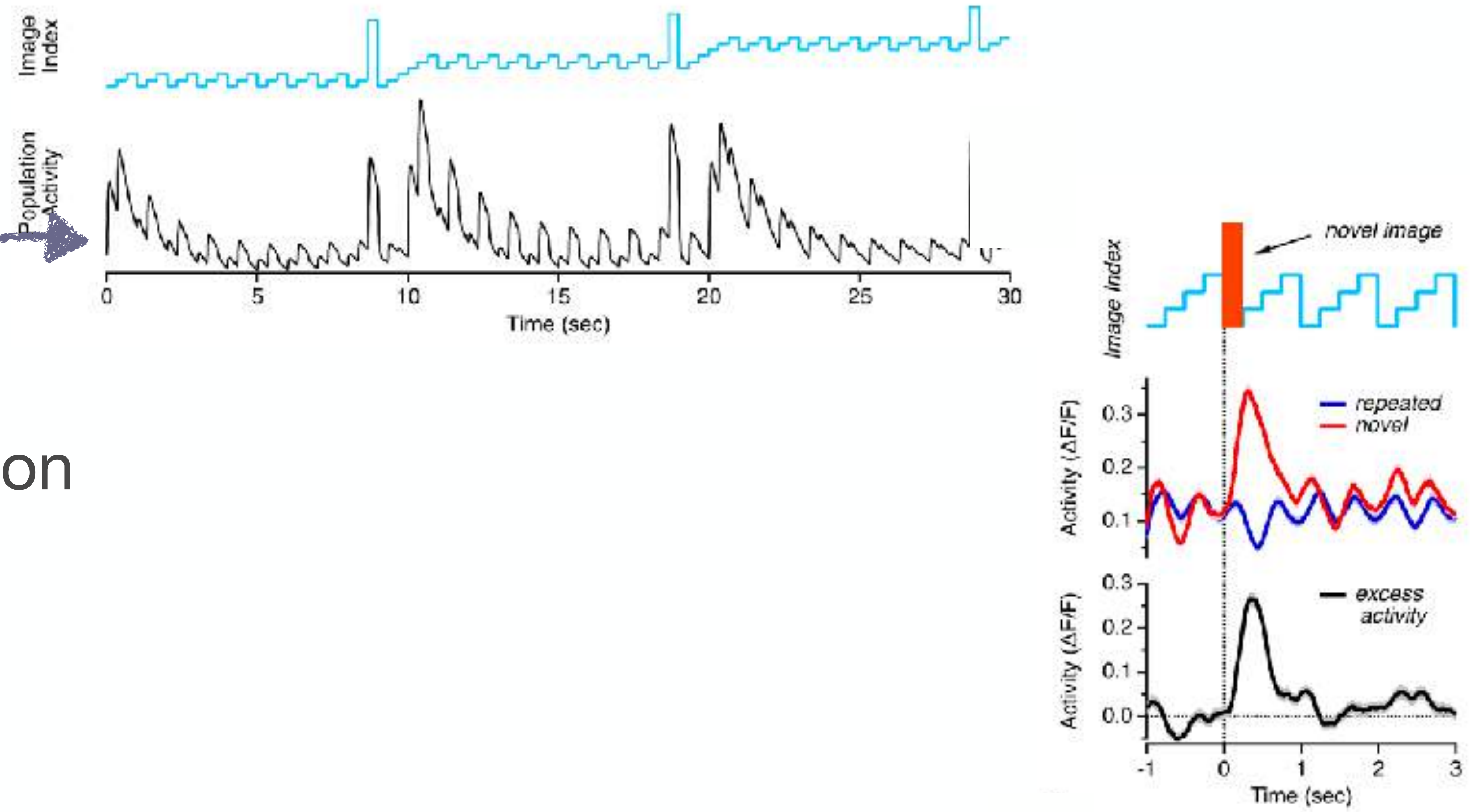
mouse - auditory cortex
[Natan et al., 2015]

Representation suppression or stimulus-specific adaptation

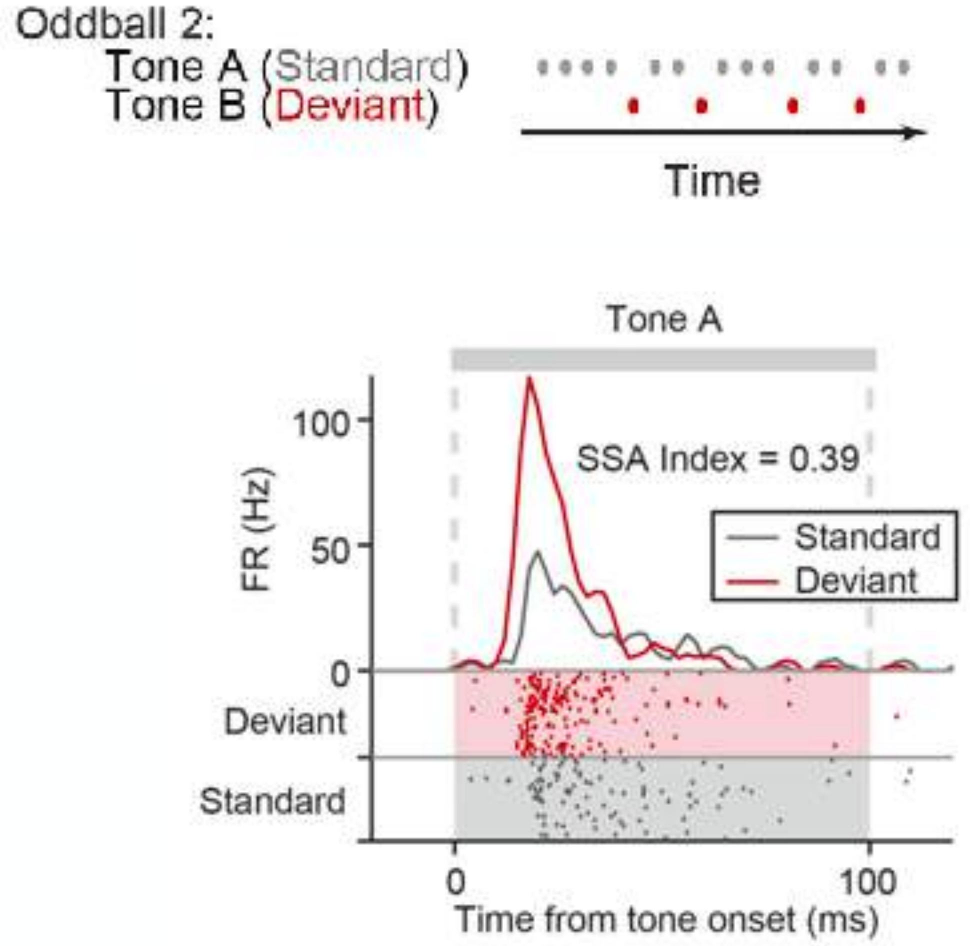


mouse - visual cortex
[Hamm & Yuste, 2016]

adaptation

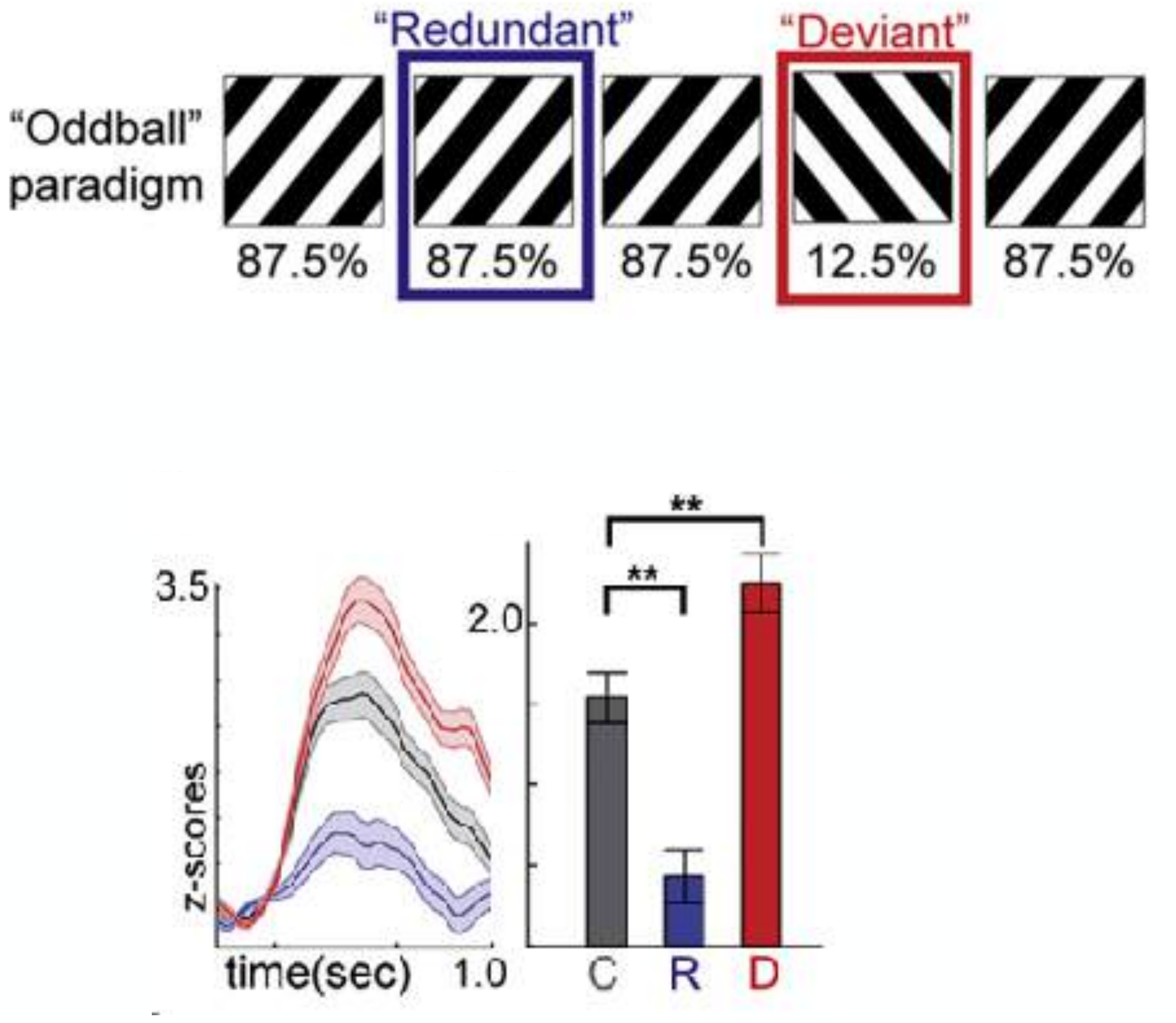


mouse - visual cortex
[Homann et al., 2021]



mouse - auditory cortex
[Natan et al., 2015]

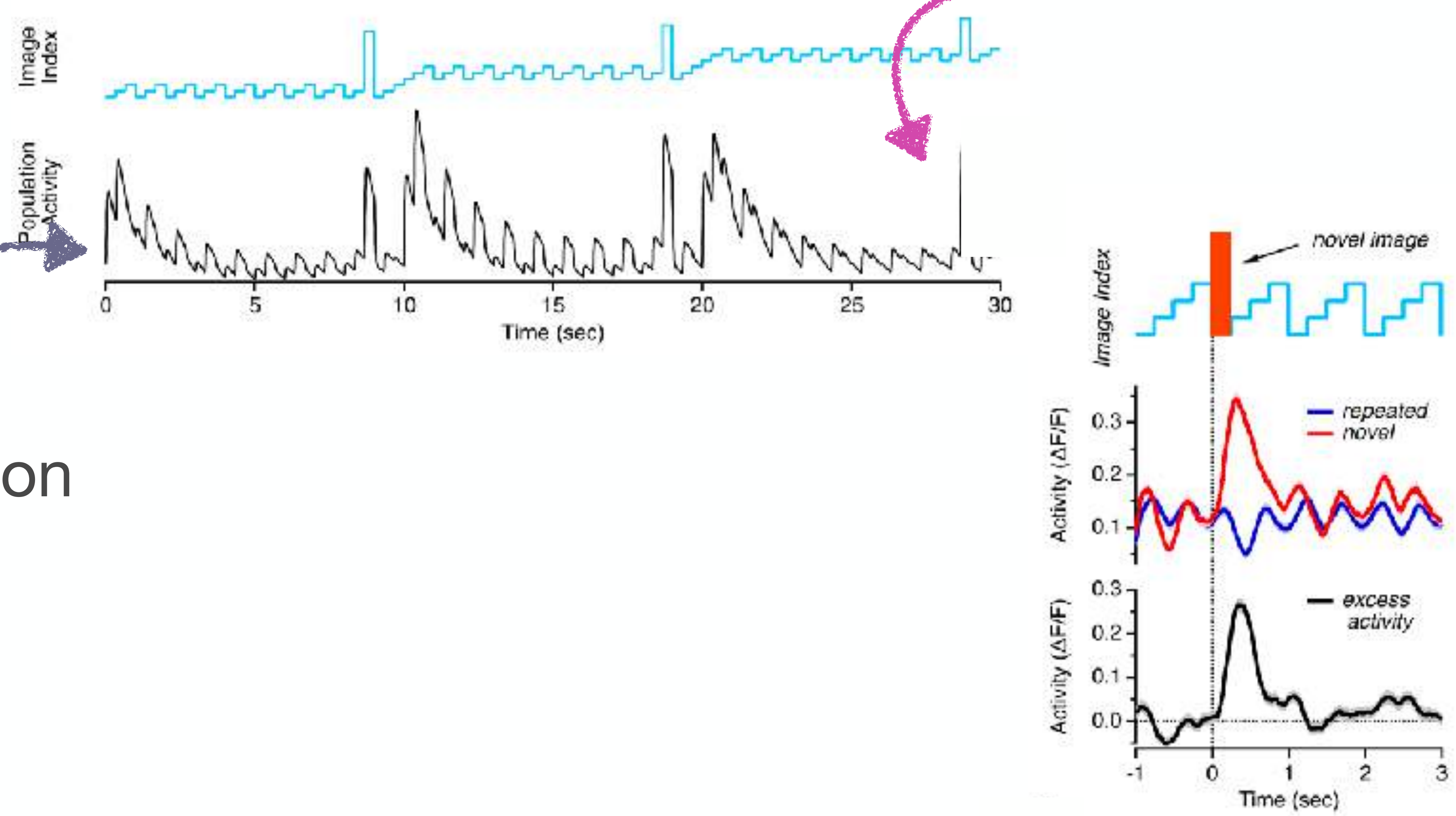
Representation suppression or stimulus-specific adaptation



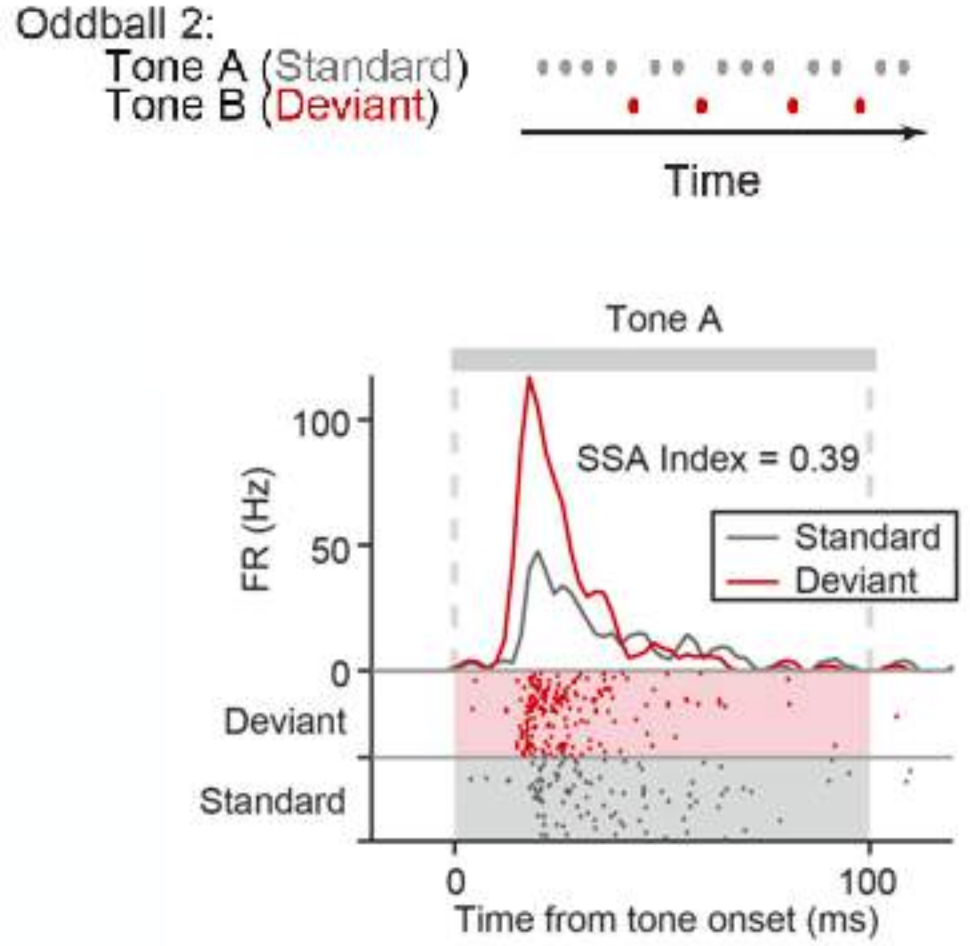
mouse - visual cortex
[Hamm & Yuste, 2016]

novelty response

adaptation



mouse - visual cortex
[Homann et al., 2021]



mouse - auditory cortex
[Natan et al., 2015]

What are the functional implications of this adaptation?

What are the functional implications of this adaptation?

I will not answer this question

~~What are the functional implications of this adaptation?~~

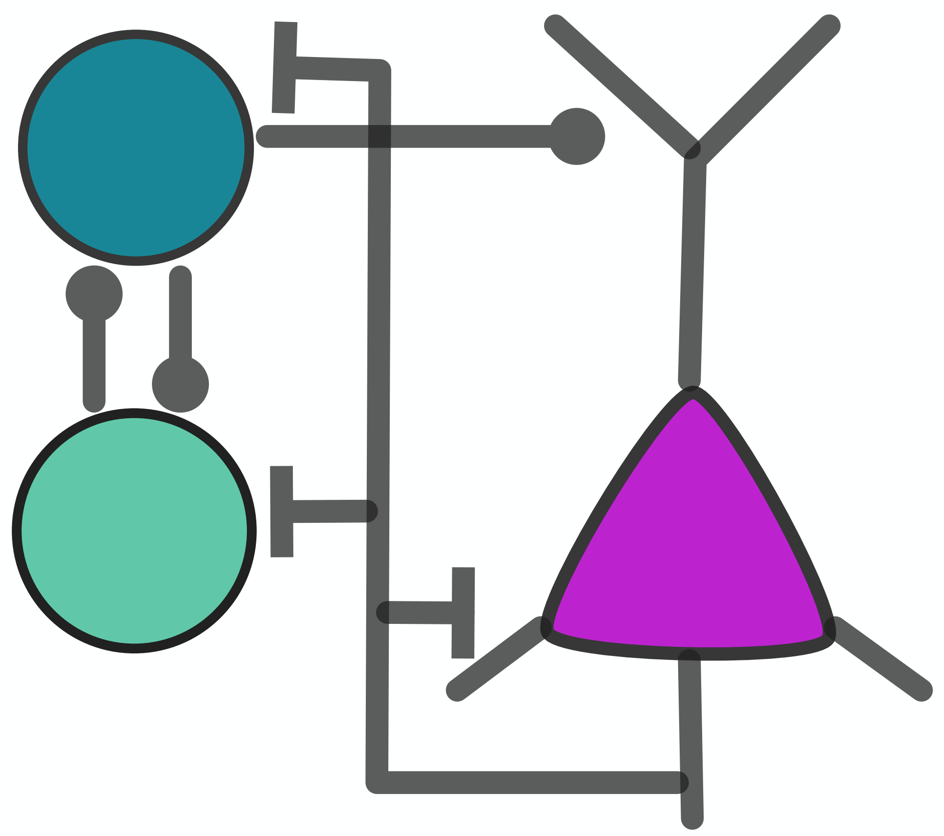
I will not answer this question

What is the mechanism that drives the adaptation?

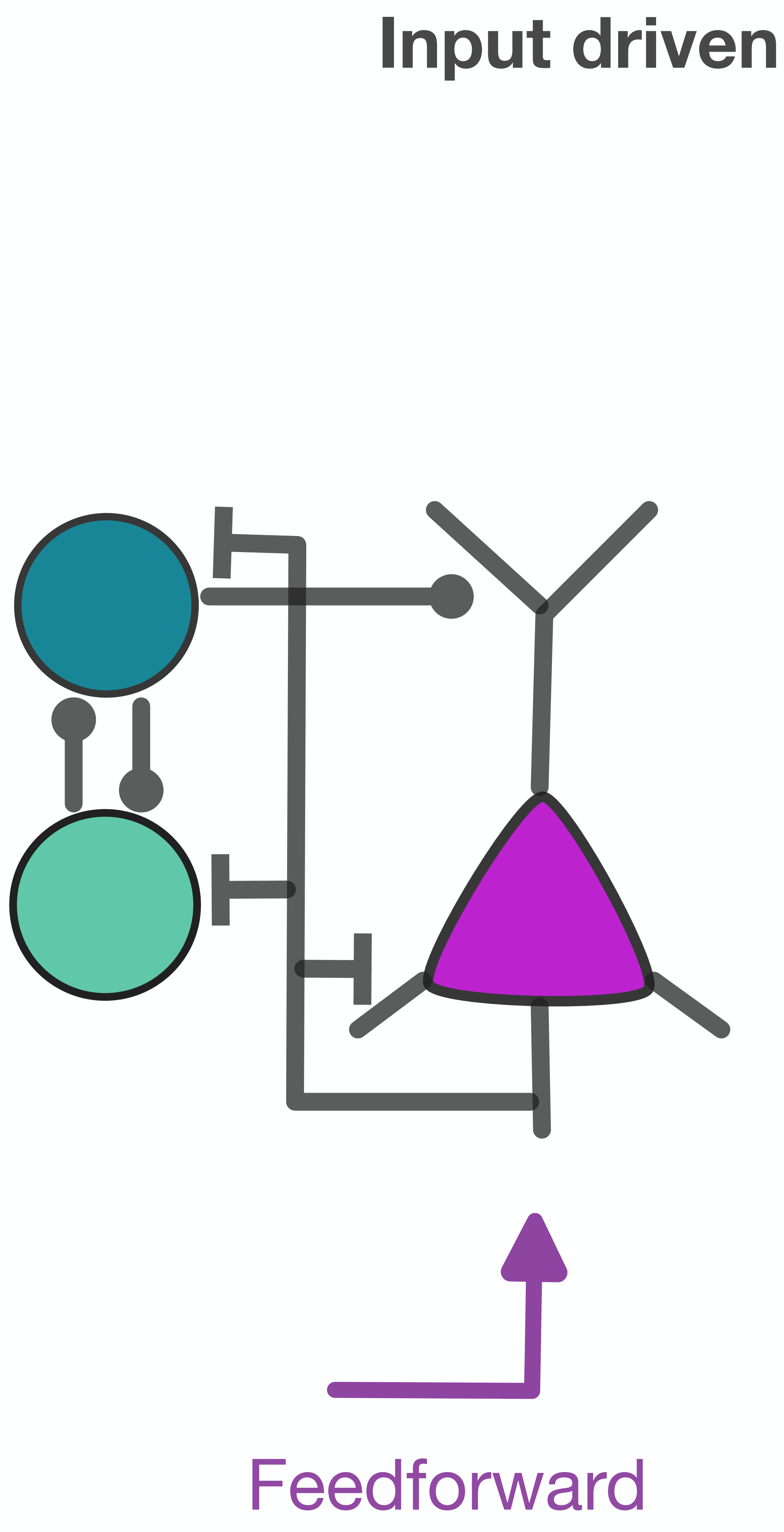
What are the mechanisms that drive these responses?

What are the mechanisms that drive these responses?

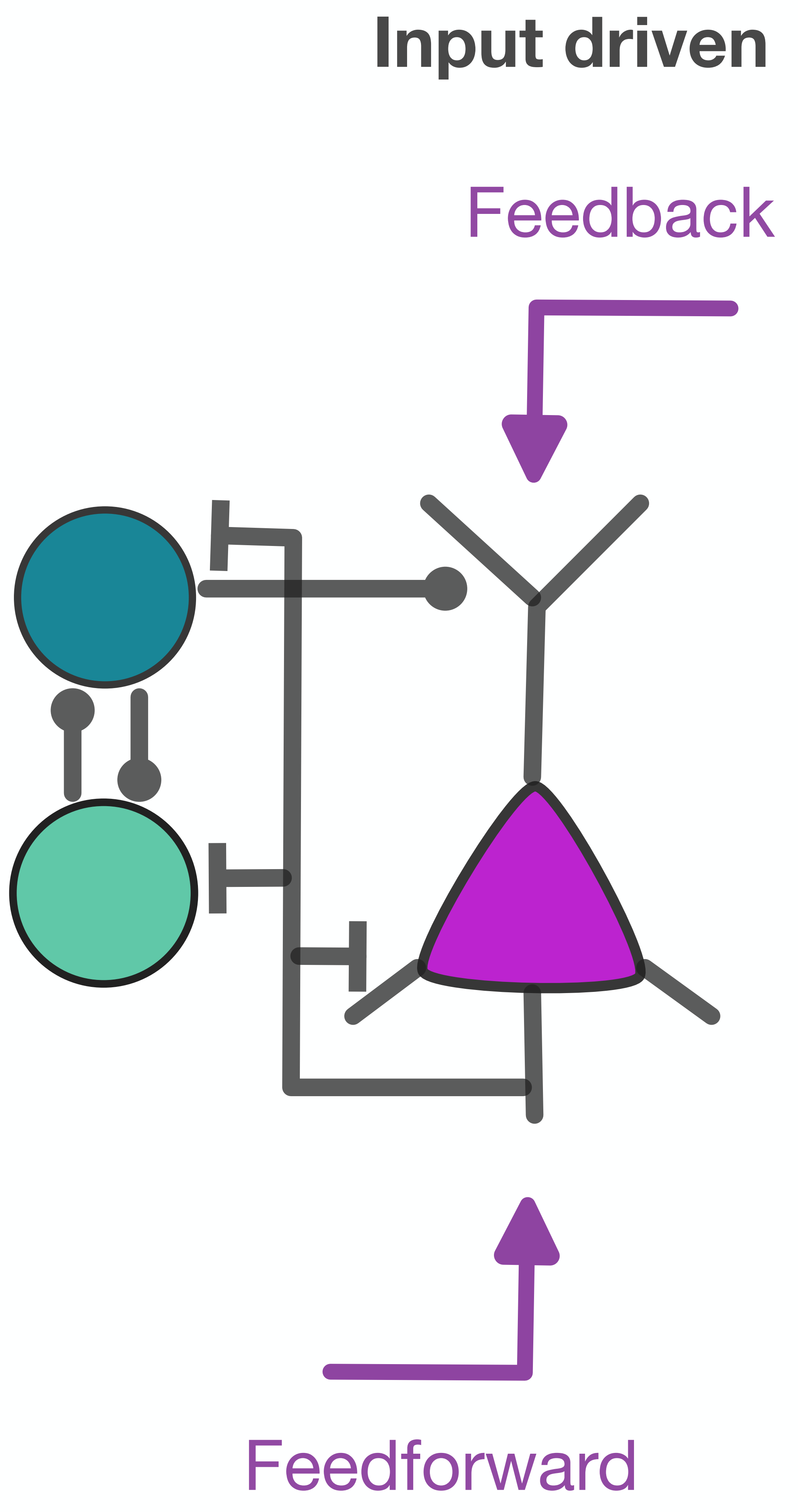
Input driven



What are the mechanisms that drive these responses?

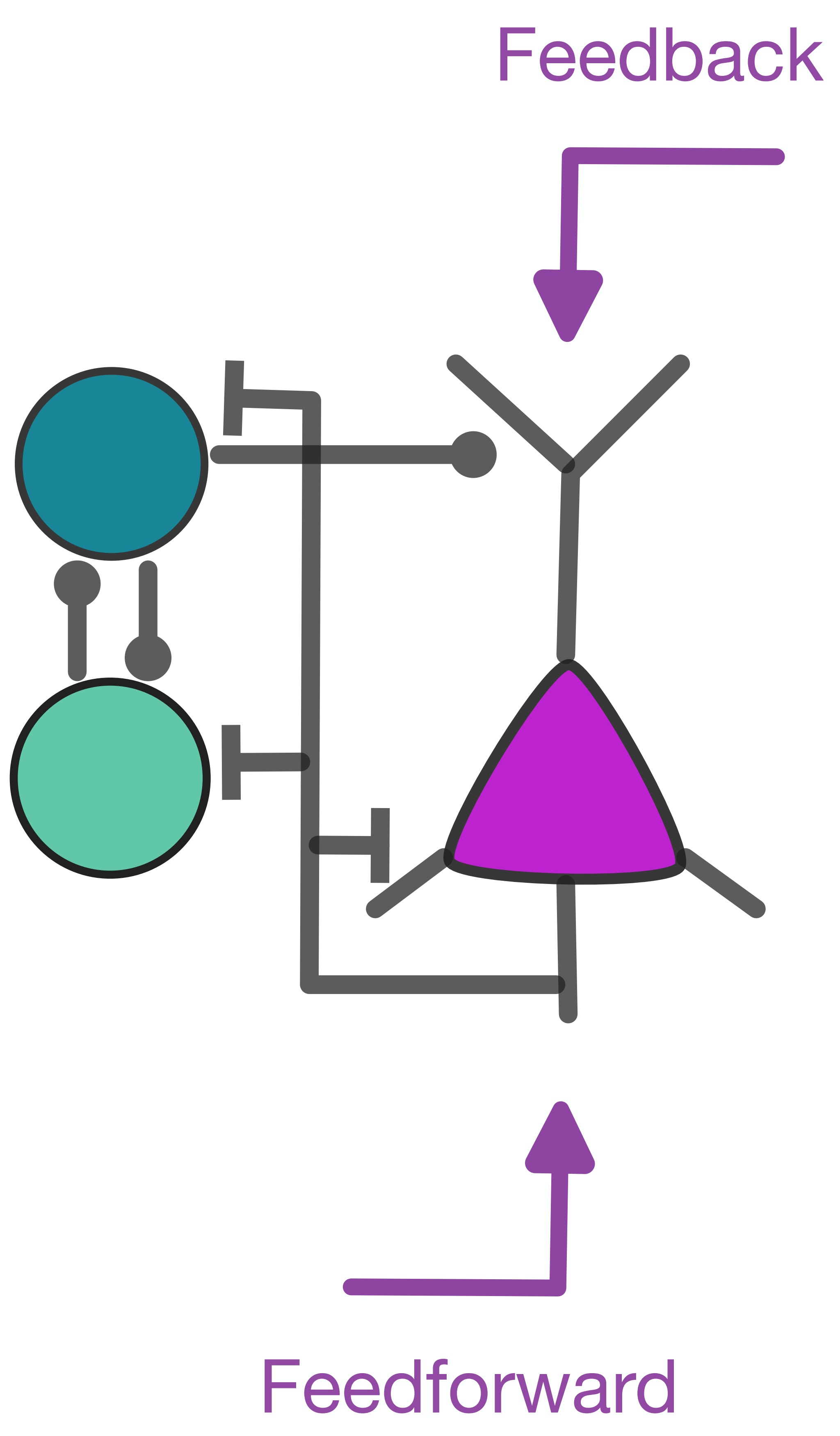


What are the mechanisms that drive these responses?



What are the mechanisms that drive these responses?

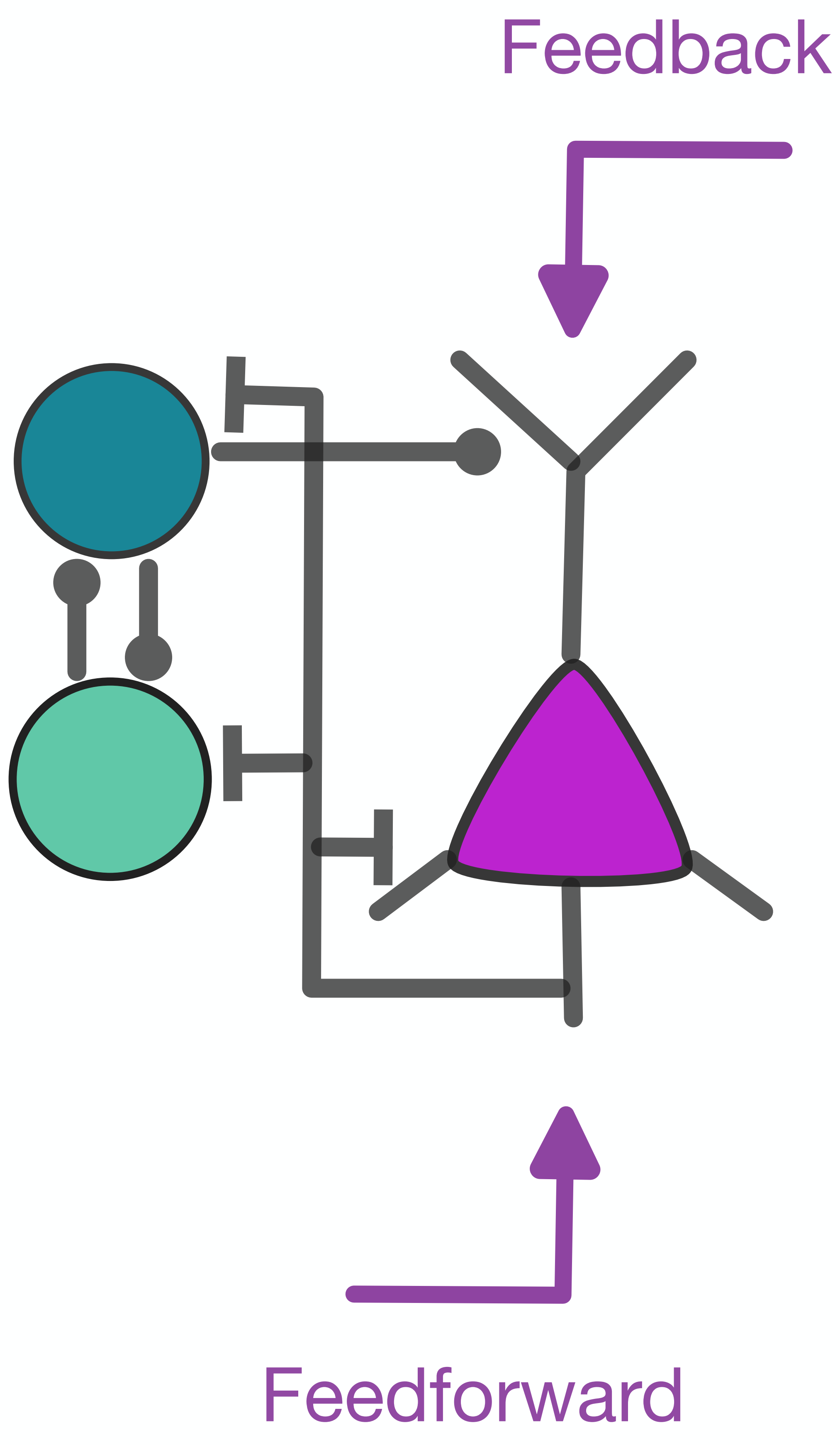
Input driven



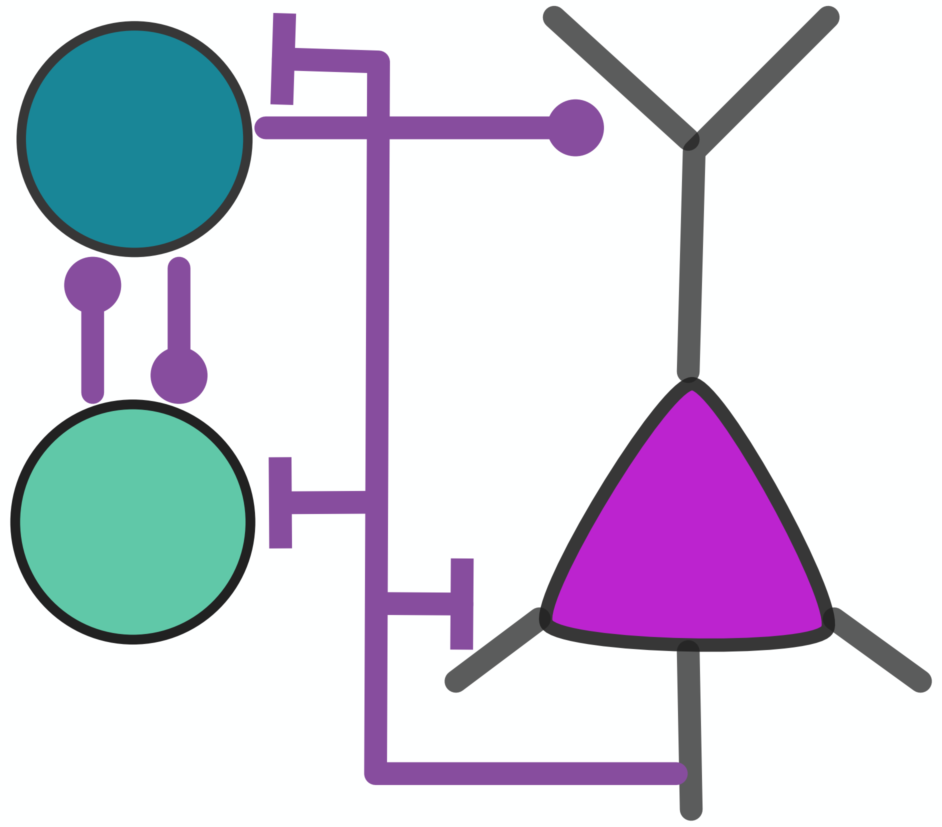
Recurrently driven

What are the mechanisms that drive these responses?

Input driven

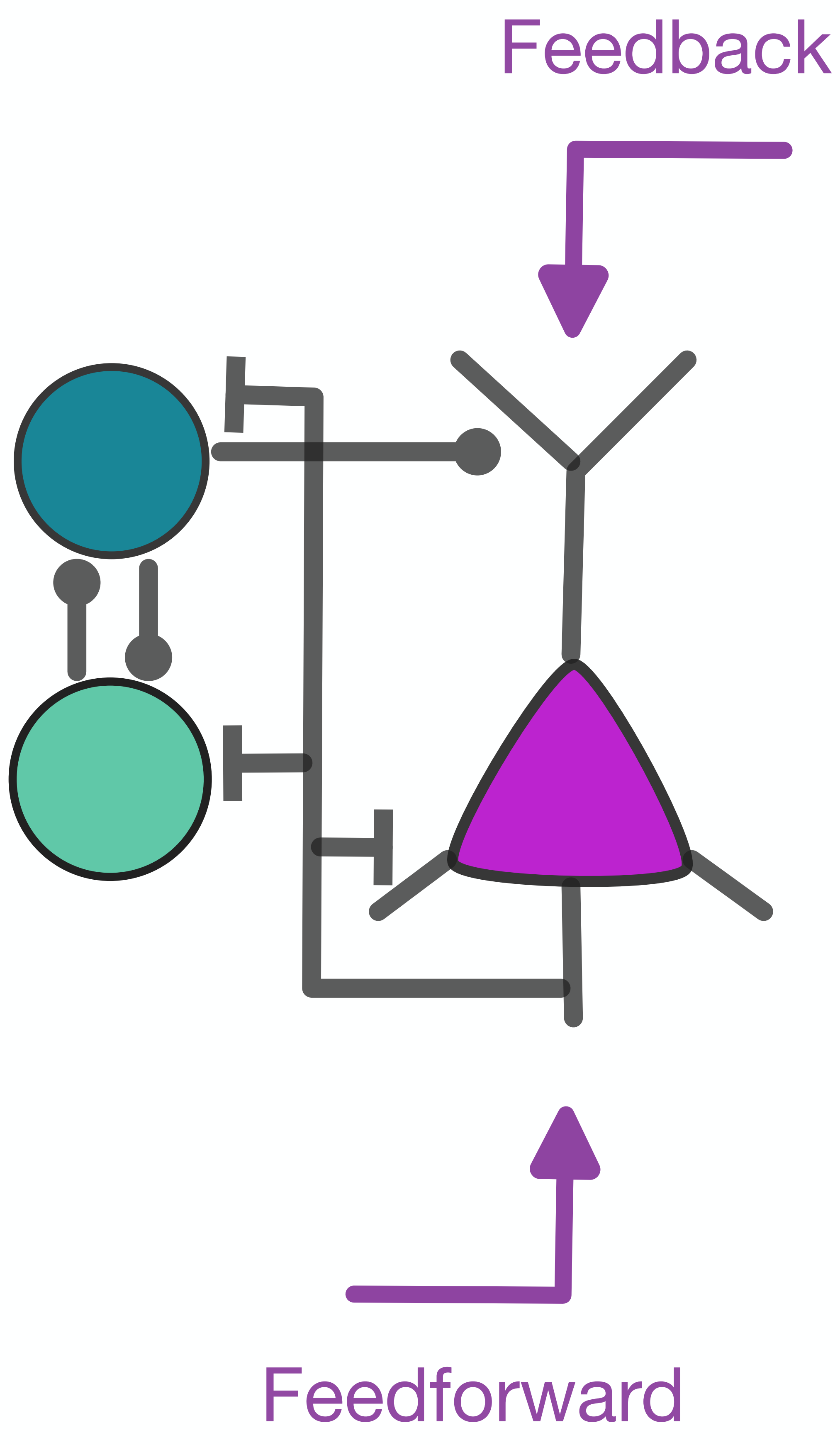


Recurrently driven

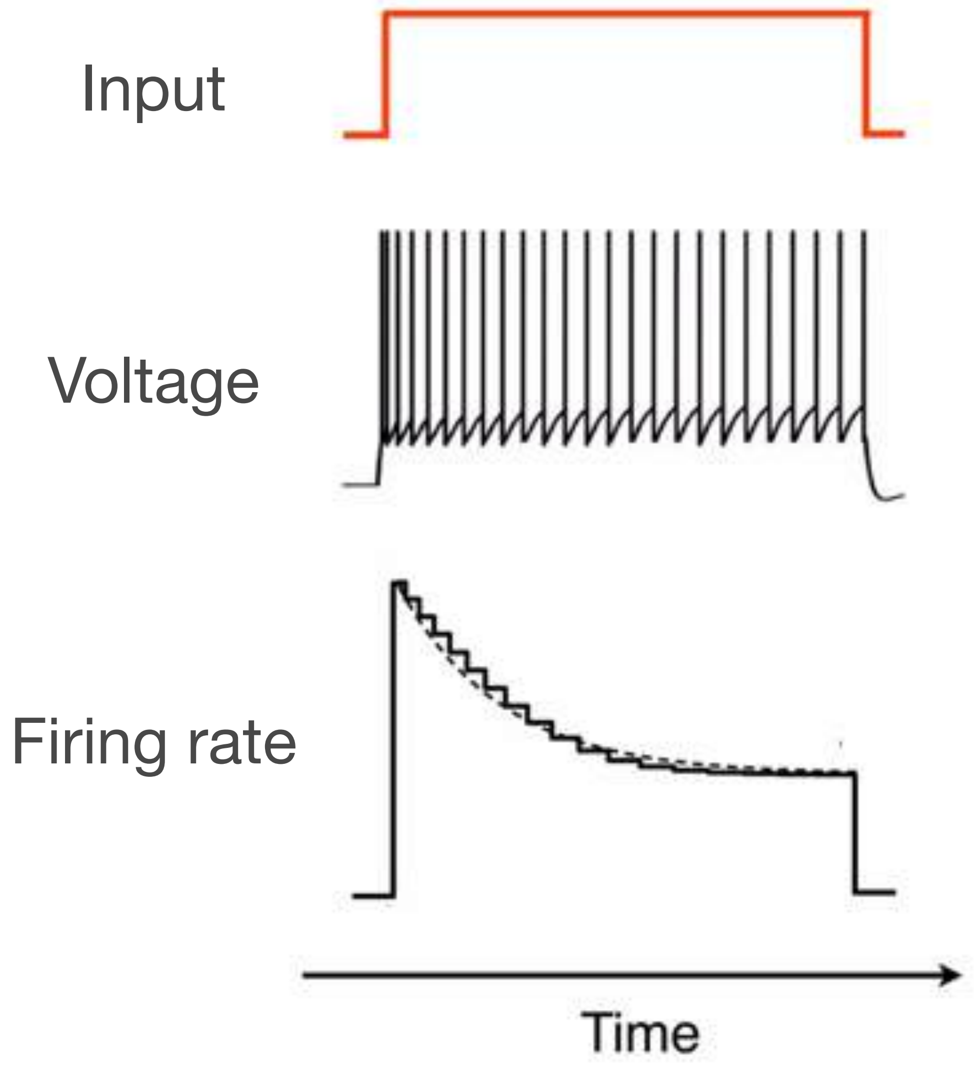


What are the mechanisms that drive these responses?

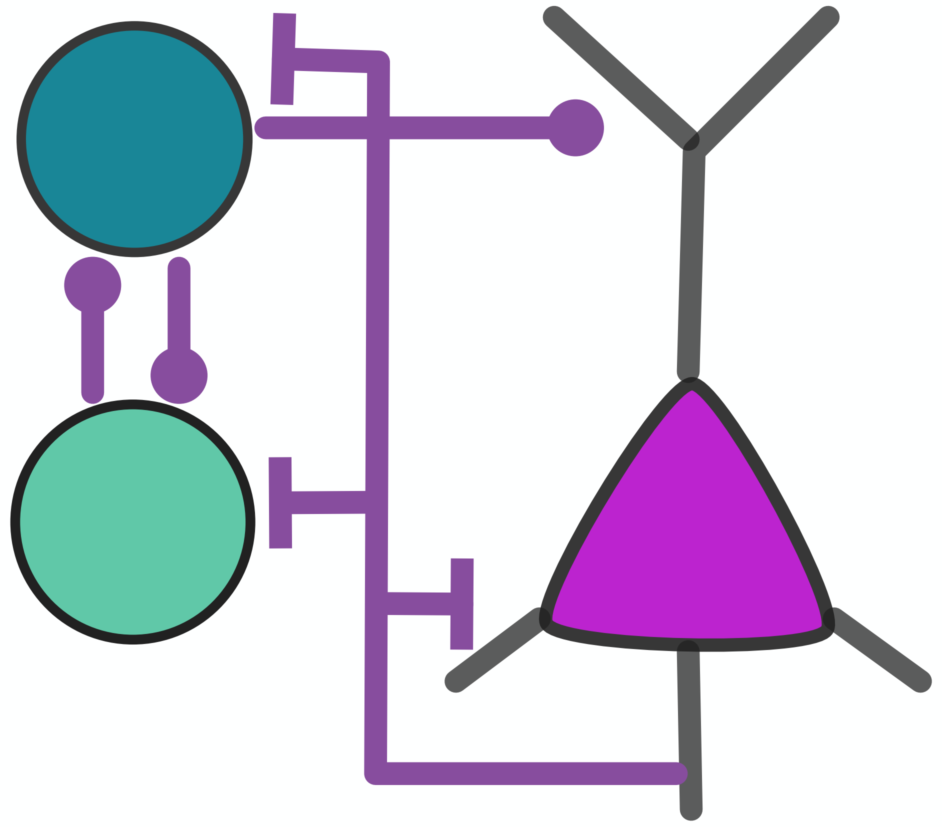
Input driven



Spike frequency adaptation

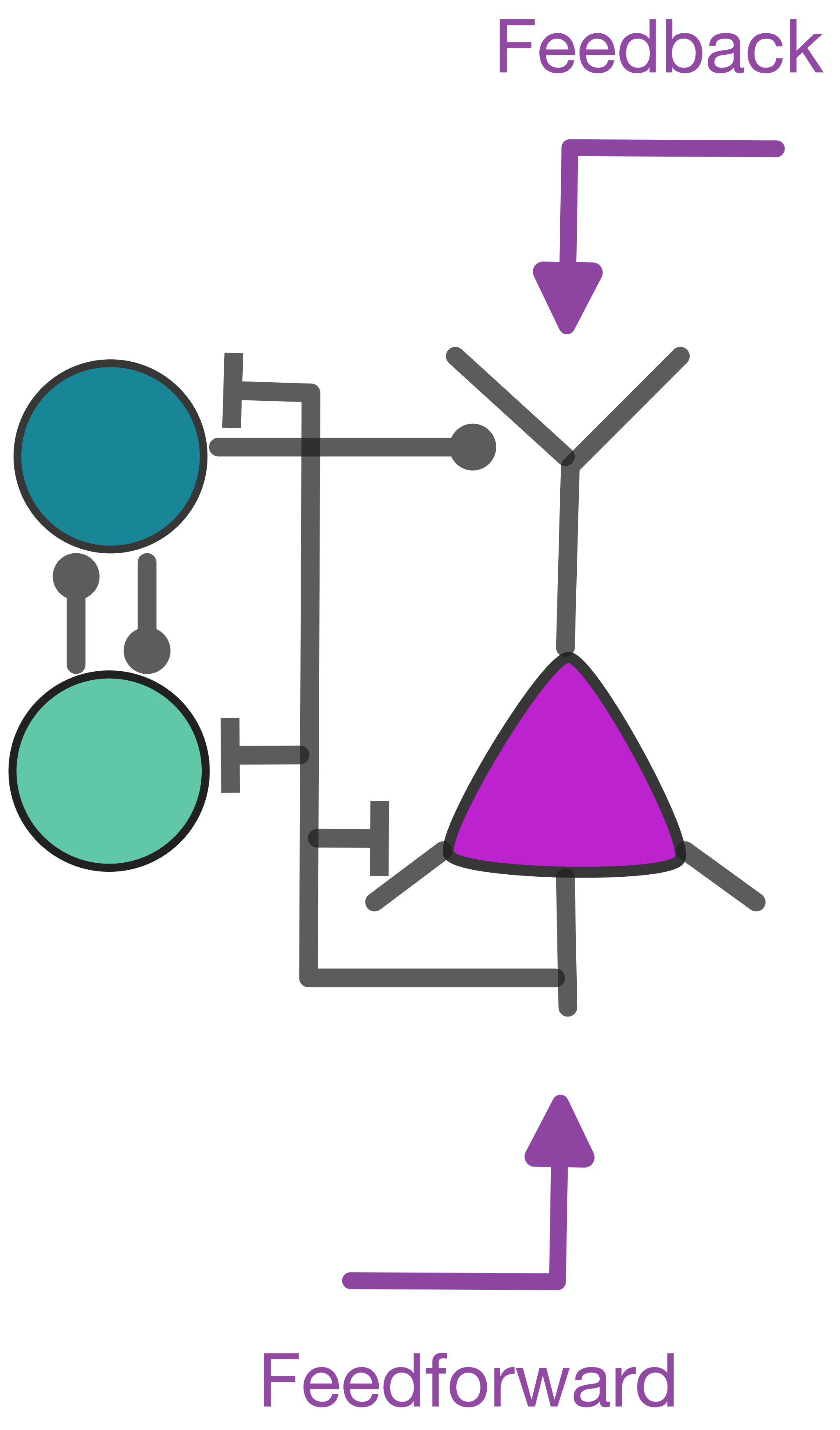


Recurrently driven

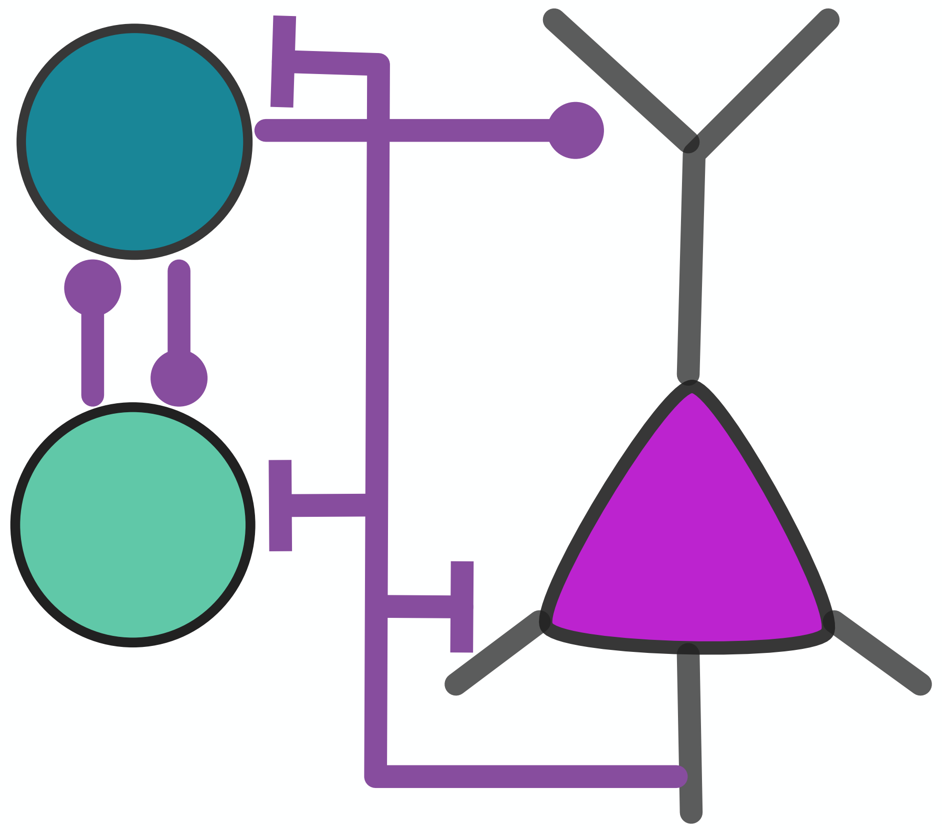


What are the mechanisms that drive these responses?

Input driven



Recurrently driven

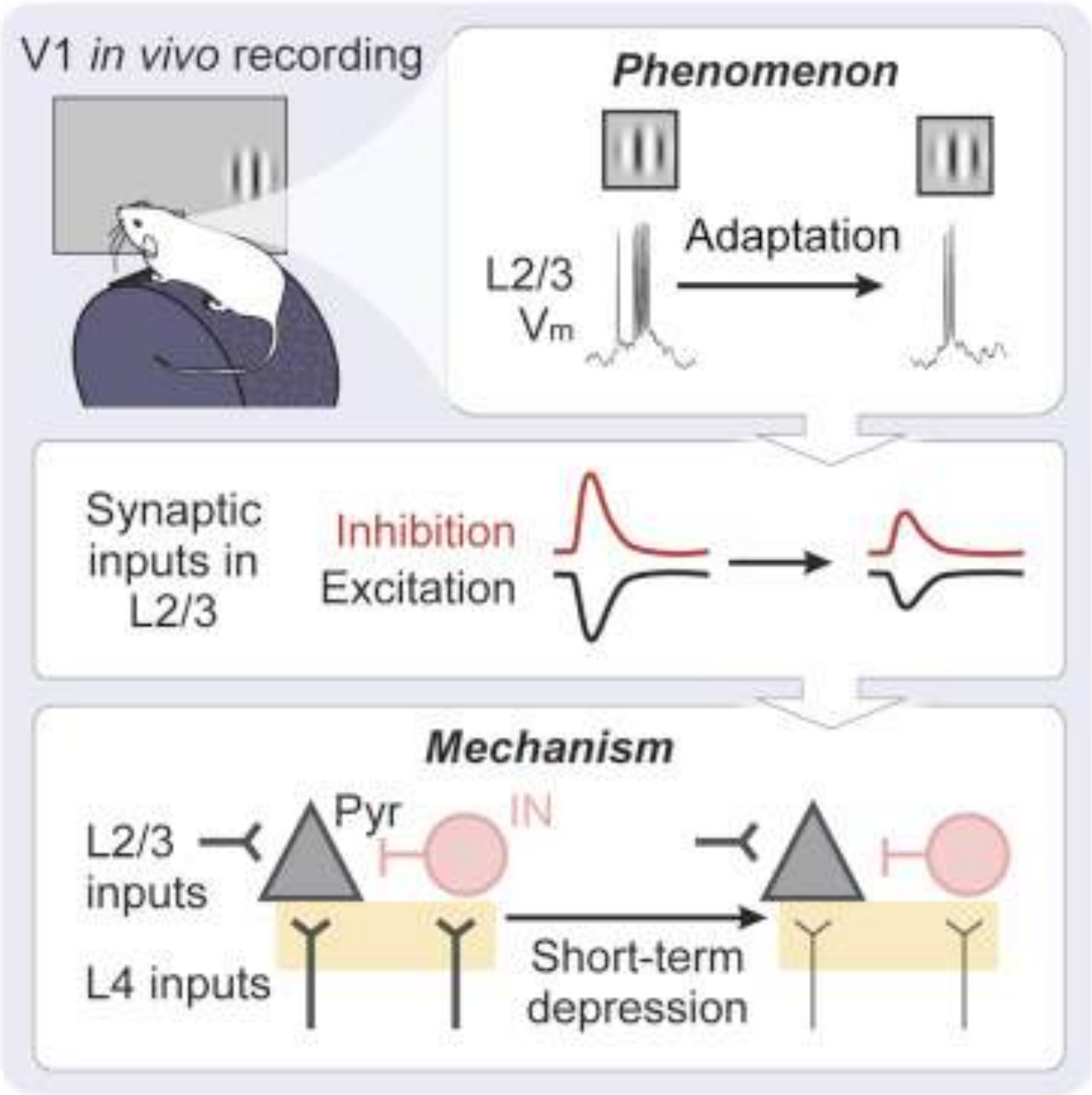


How do we think about this

- What is driving this adaptation?
- Reduced excitatory input:
  - Reduced input excitation
  - Reduced recurrent excitation
- Increased inhibitory input:
  - Recurrent
  - Feedback inhibition from higher order thalamus or other cortical areas

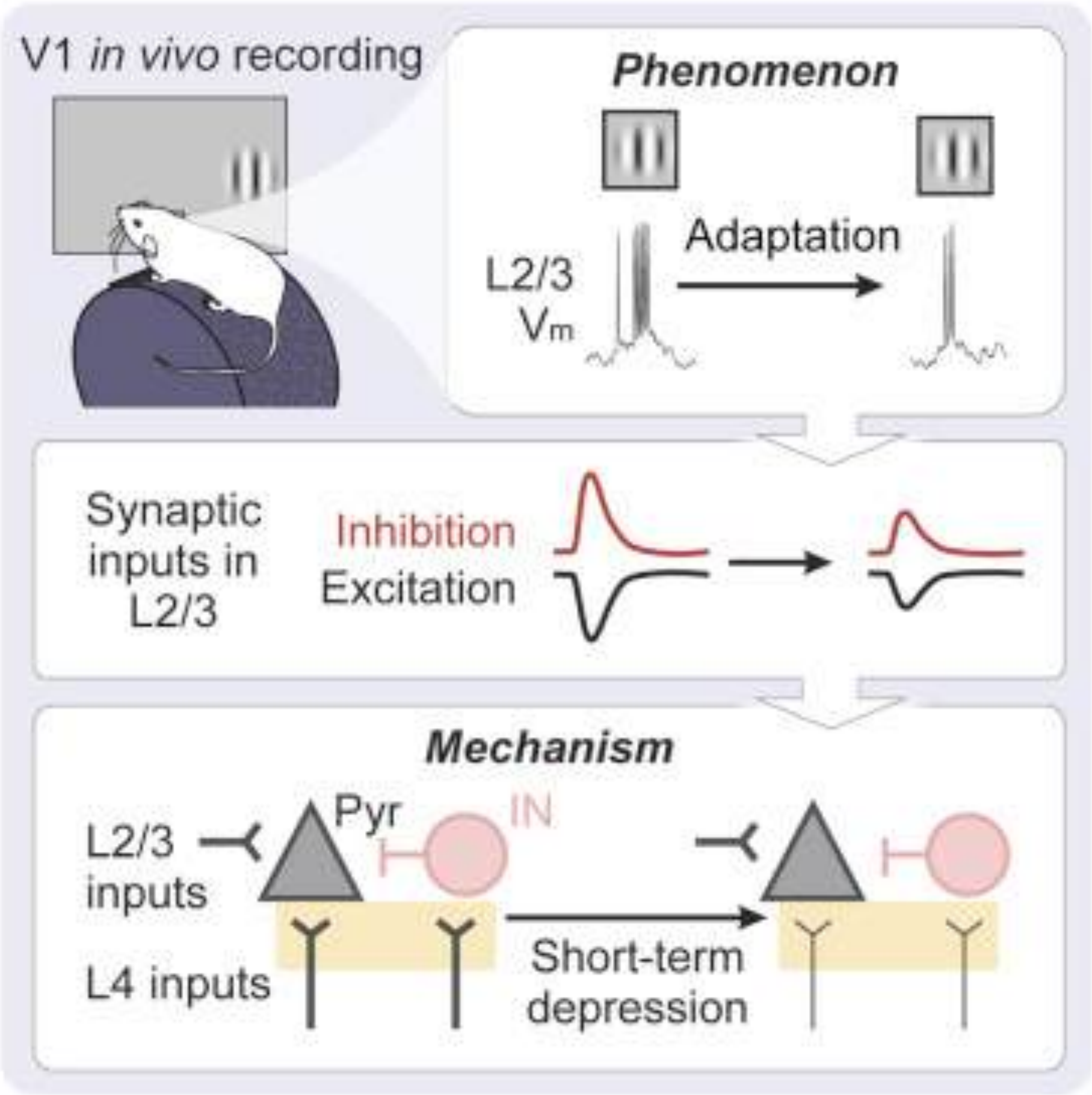
Short term depression of input synapses

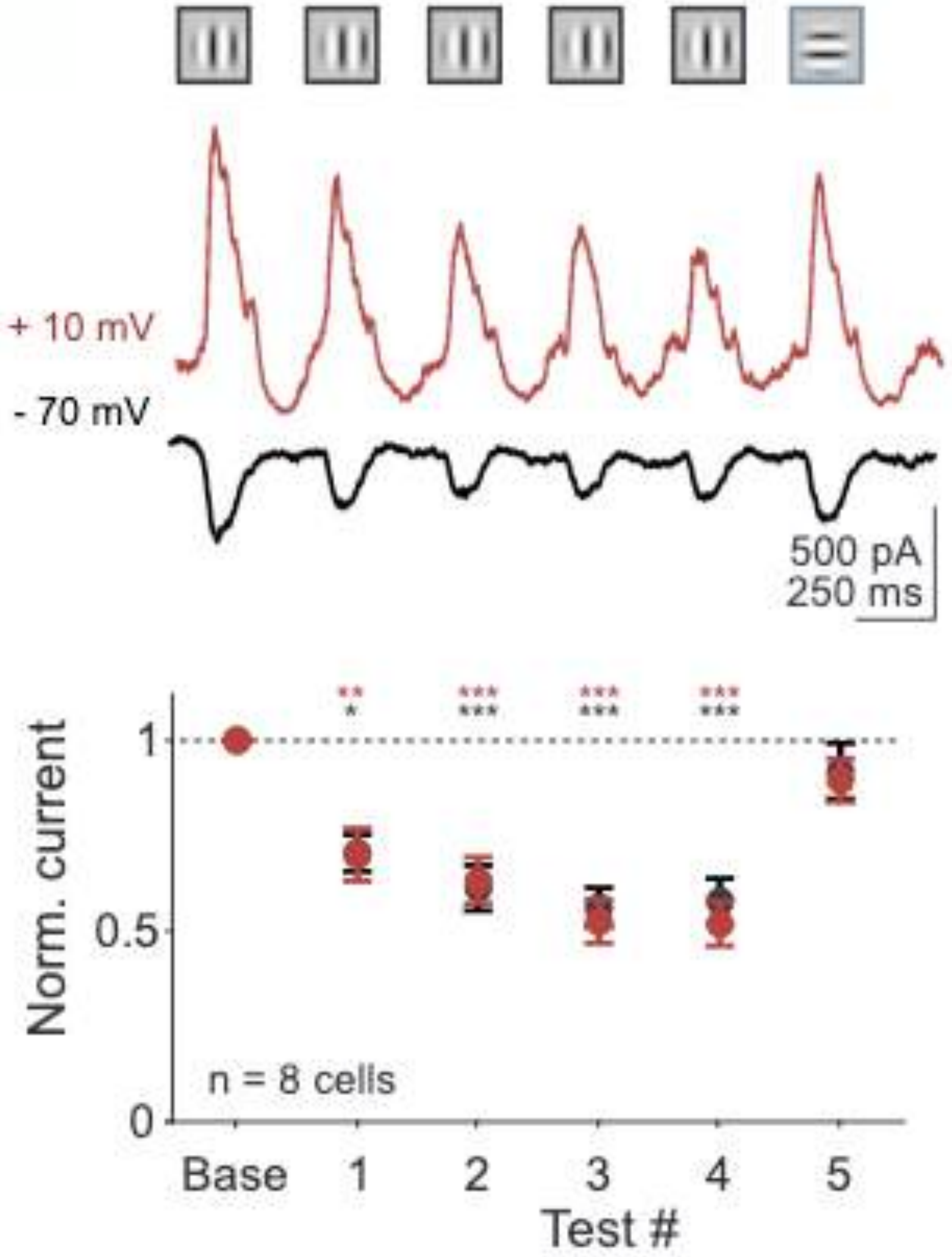
Short term depression of input synapses



[Li & Glickfeld, 2023]

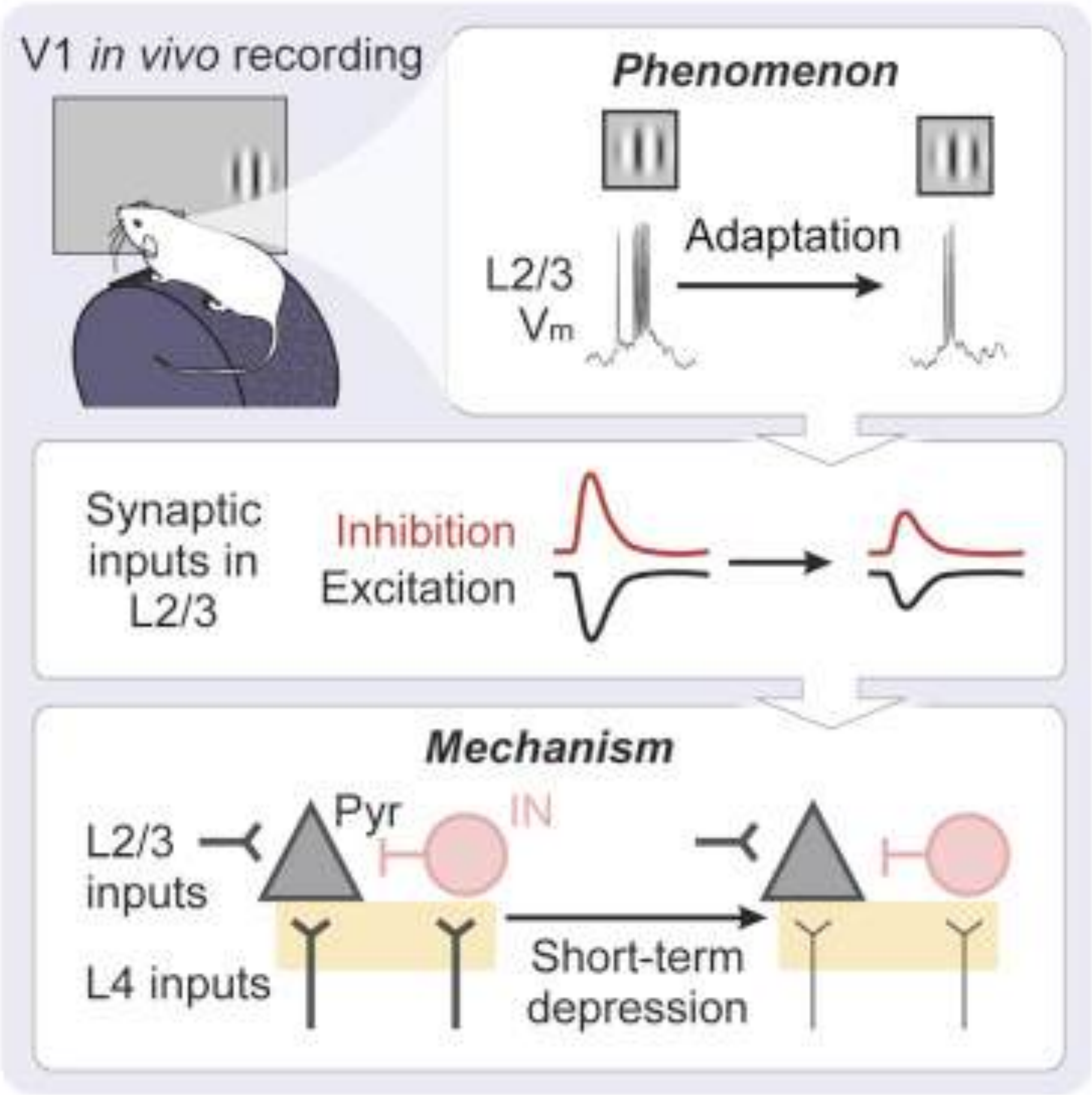
Short term depression of input synapses

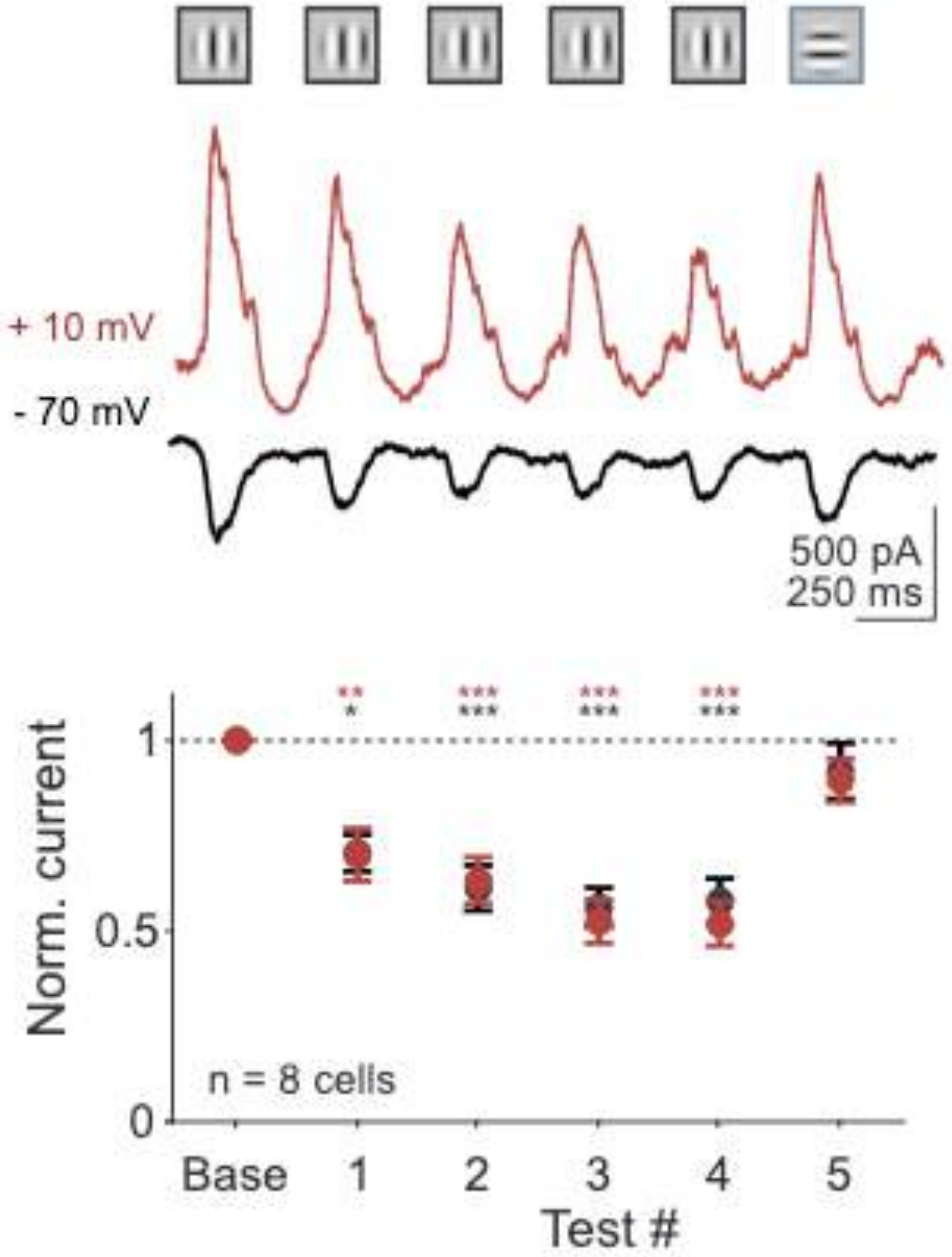




[Li & Glickfeld, 2023]

Short term depression of input synapses

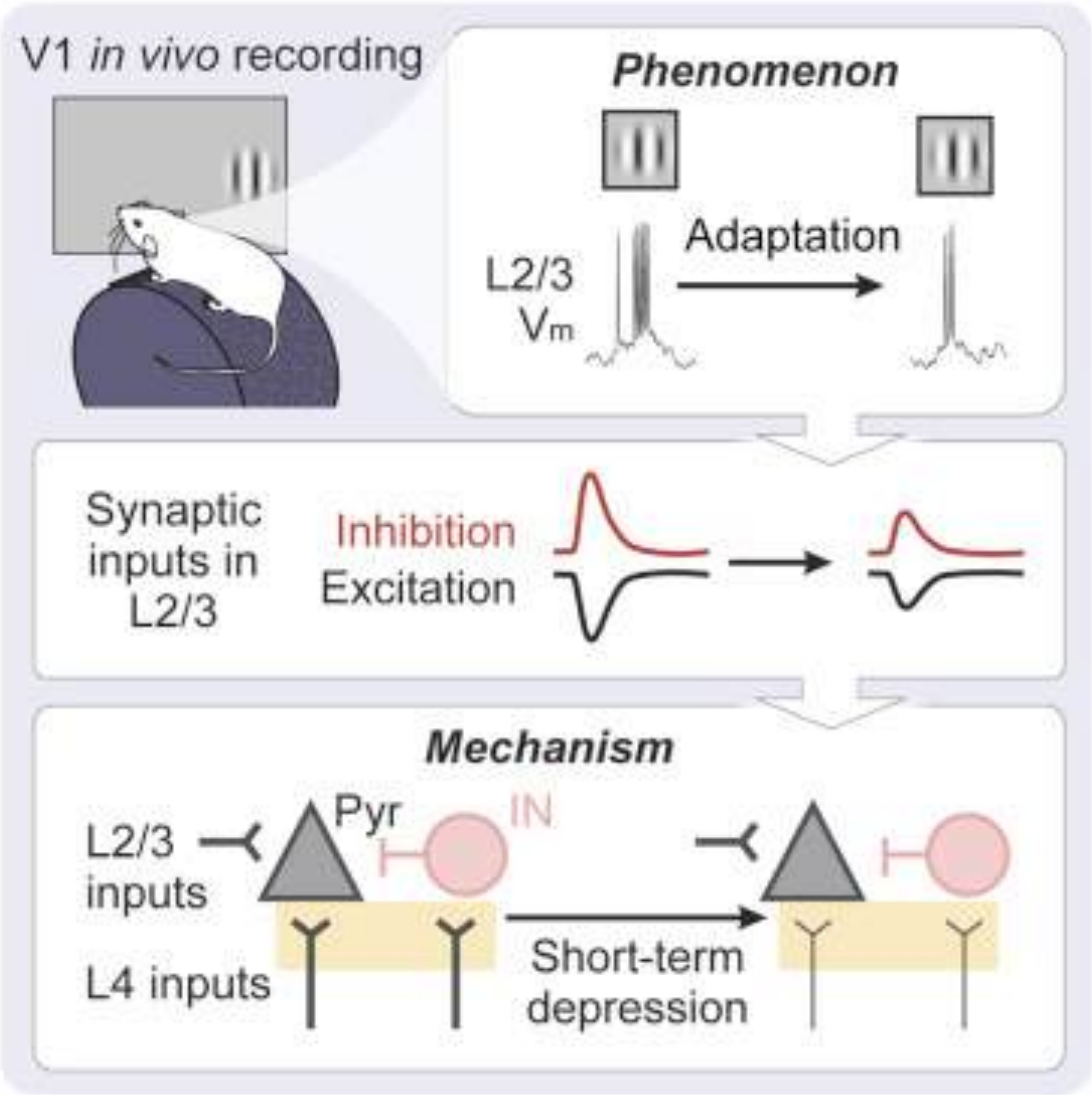


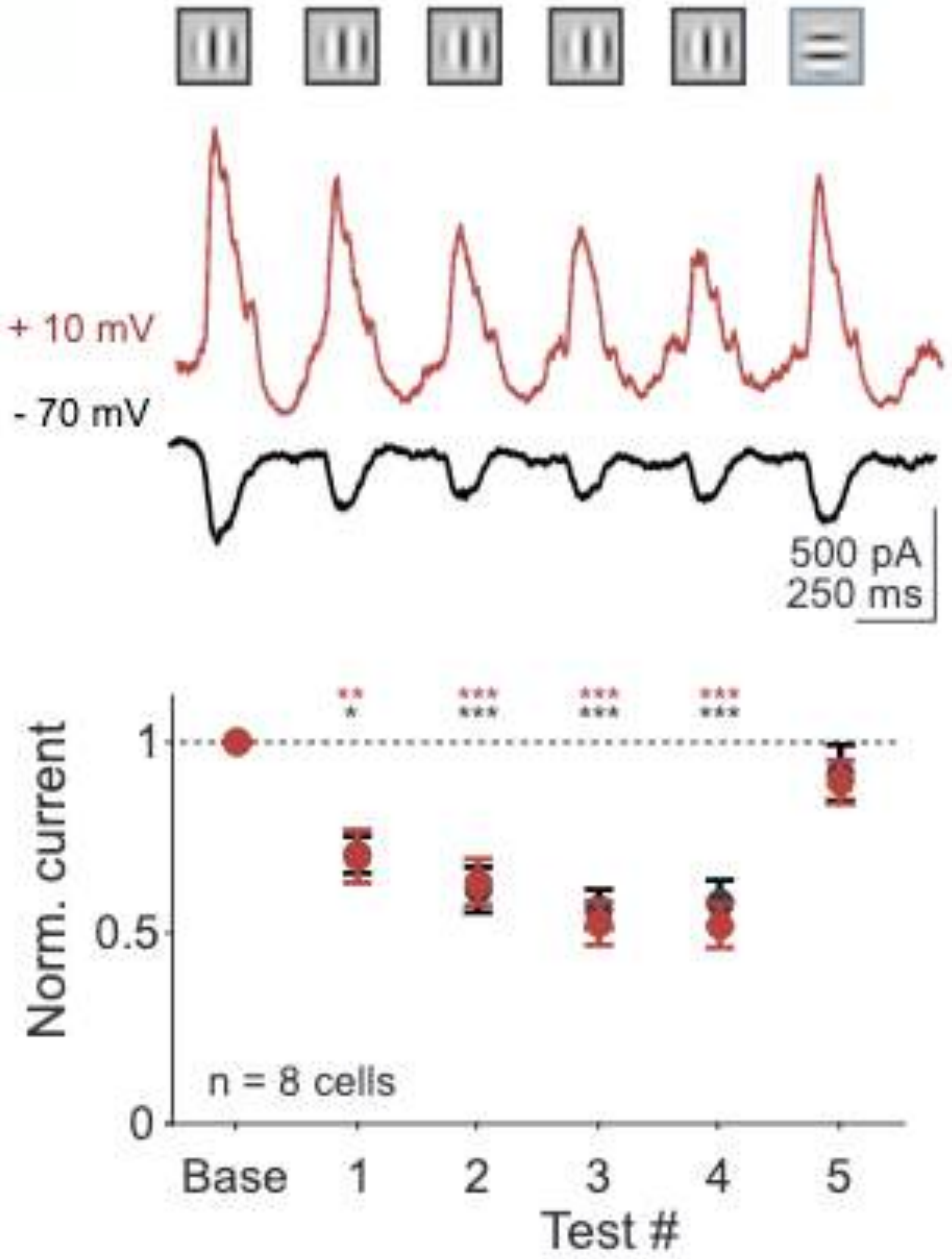


Input specific short-term depression of the **L4->L2/3** connections cause stimulus specific adaptation

[Li & Glickfeld, 2023]

Short term depression of input synapses





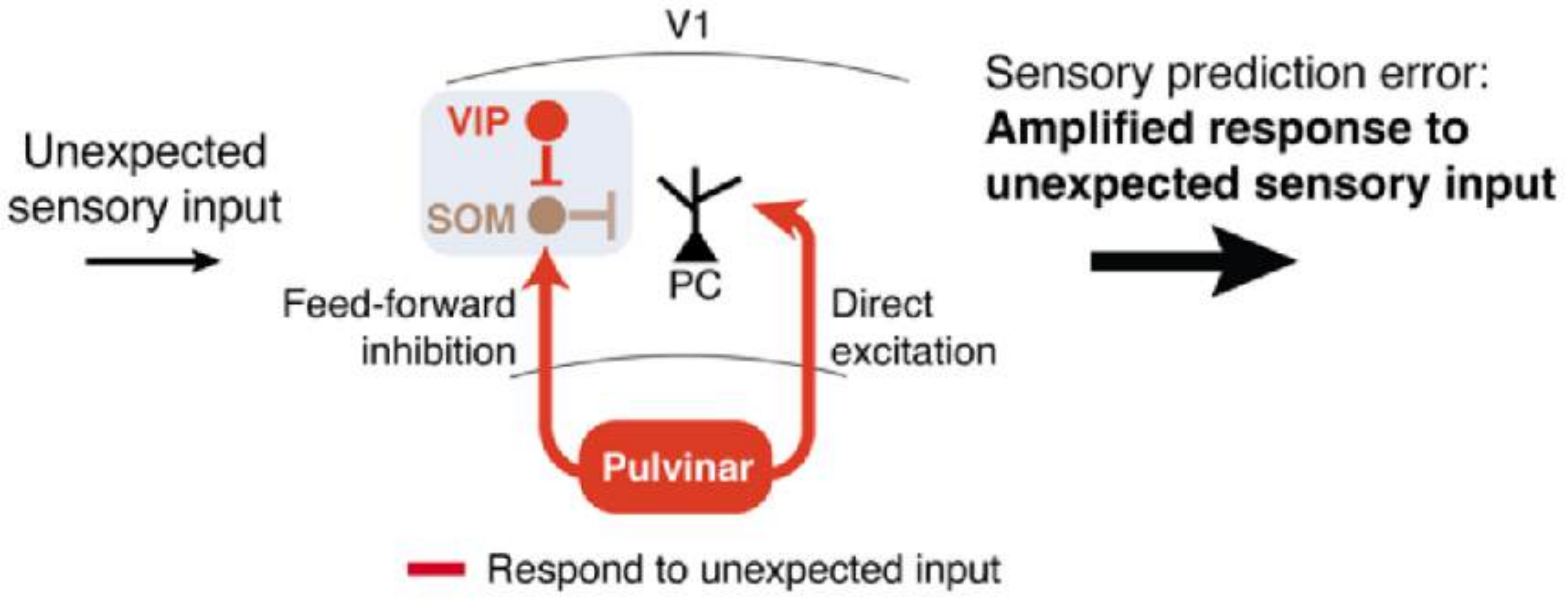
Input specific short-term depression of the **L4->L2/3** connections cause stimulus specific adaptation

Only relevant for **short** timescales!

[Li & Glickfeld, 2023]

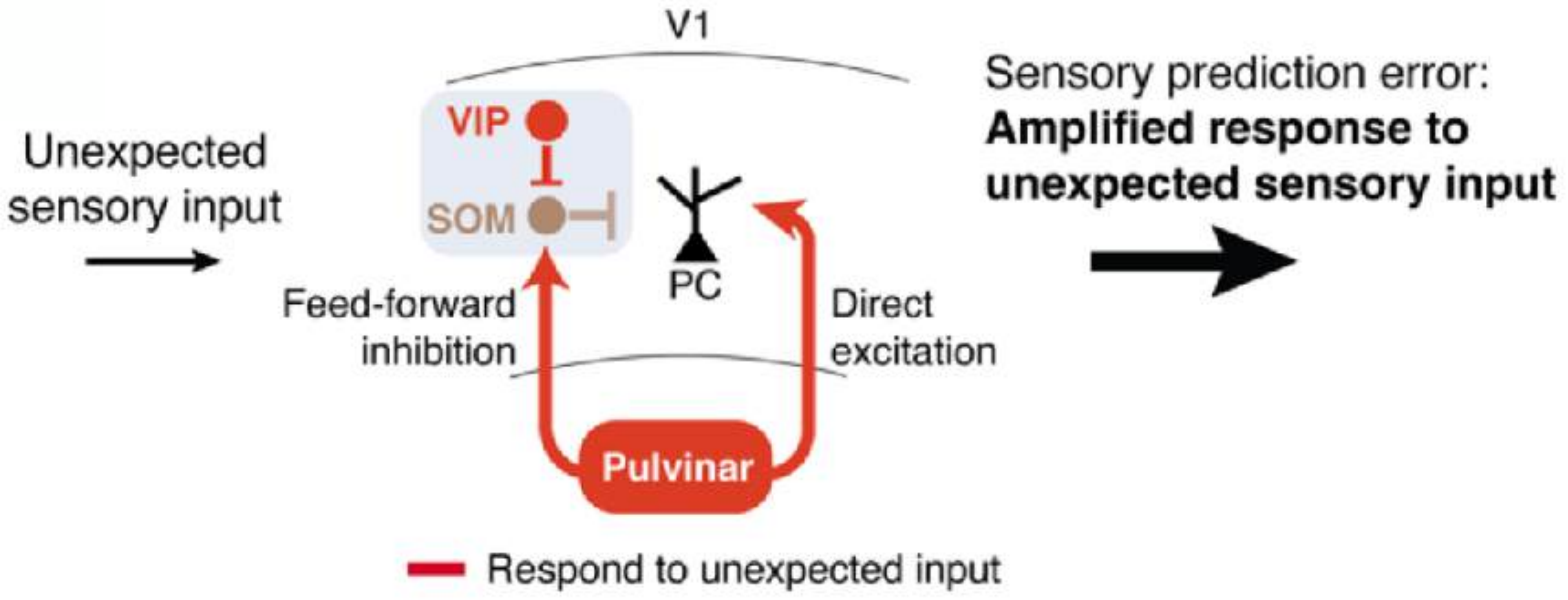
Pulvinar-driven response amplification

Pulvinar-driven response amplification



[Furutachi et al., 2023]

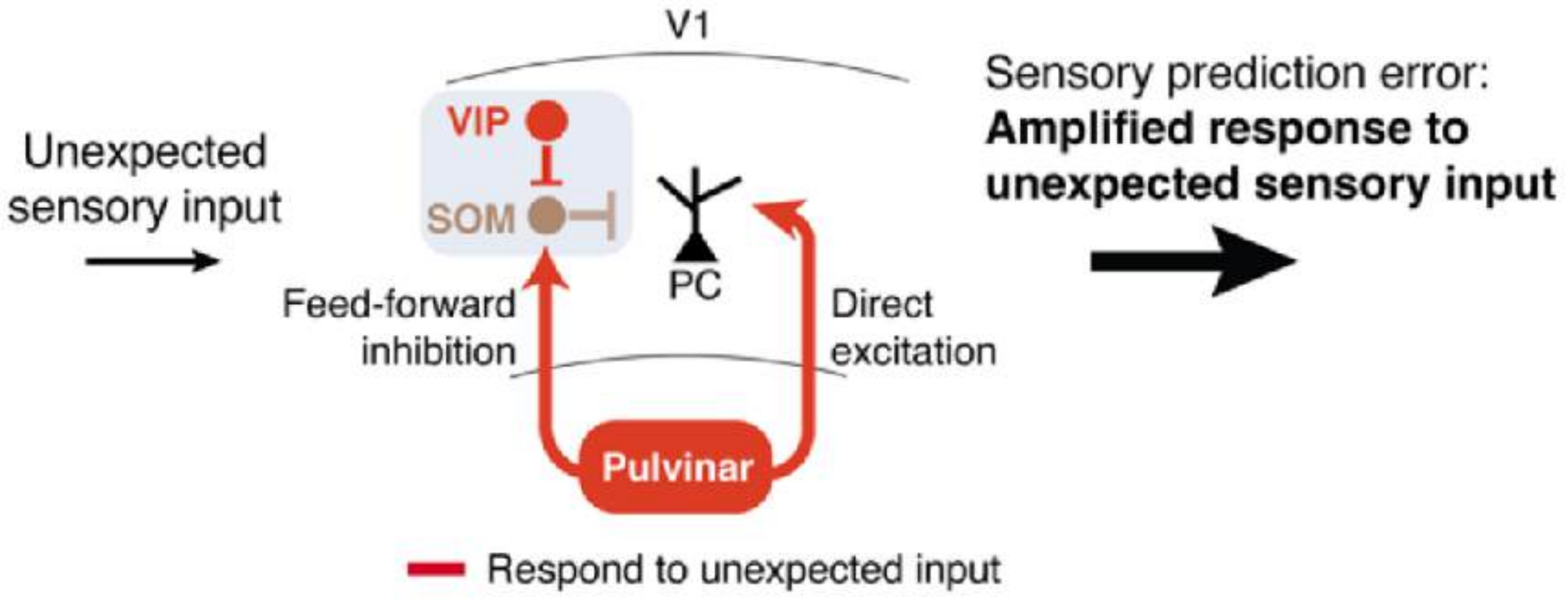
Pulvinar-driven response amplification



- **Pulvinar** drives **E cells** and a **subset of SST cells** (**SST1**)

[Furutachi et al., 2023]

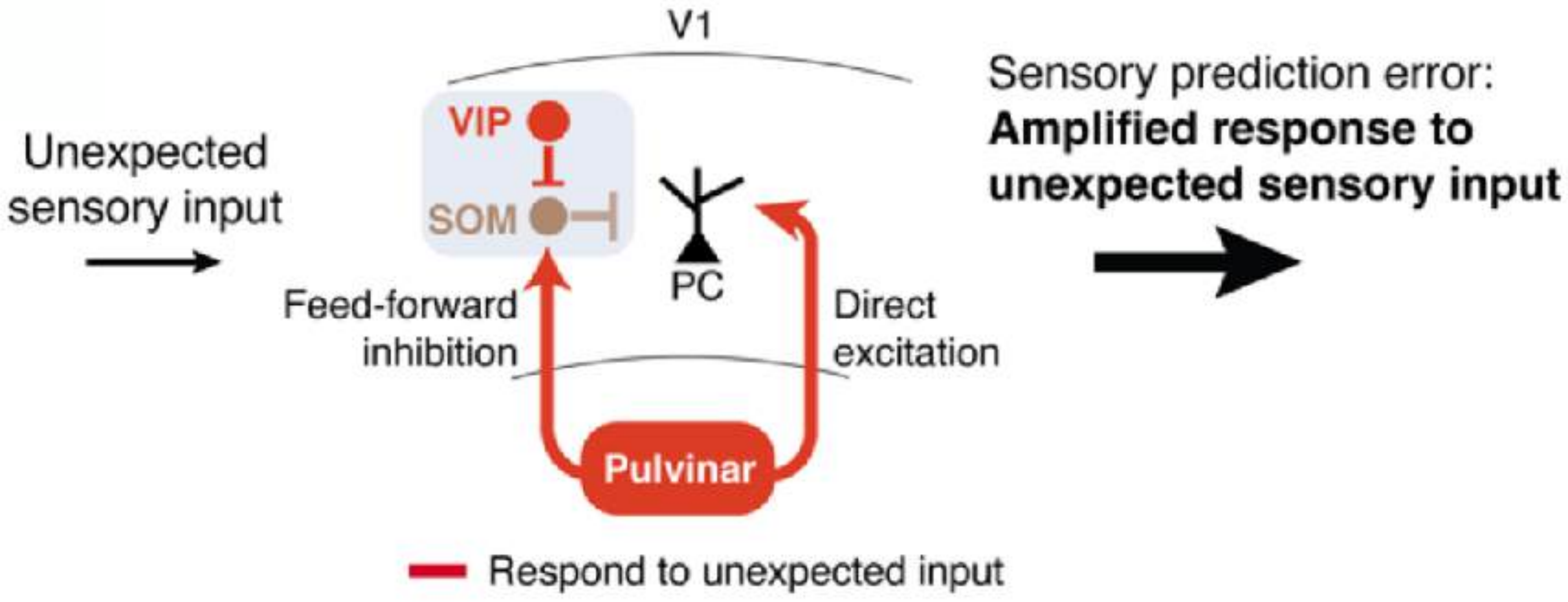
Pulvinar-driven response amplification



- **Pulvinar** drives **E cells** and a **subset of SST cells** (**SST1**)
  - SST1 get suppressed by unexpected stimuli due to inhibition from VIPs

[Furutachi et al., 2023]

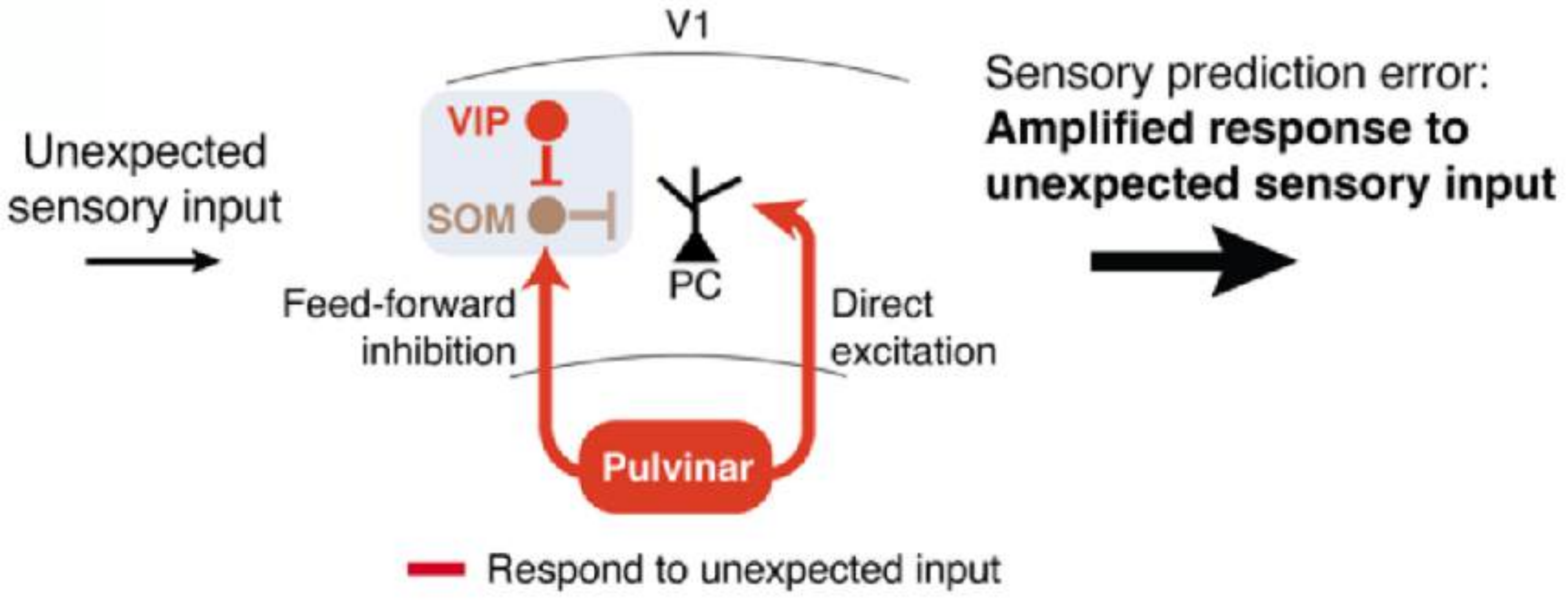
Pulvinar-driven response amplification



- **Pulvinar** drives **E cells** and a **subset of SST cells** (**SST1**)
  - SST1 get suppressed by unexpected stimuli due to inhibition from VIPs

[Furutachi et al., 2023]

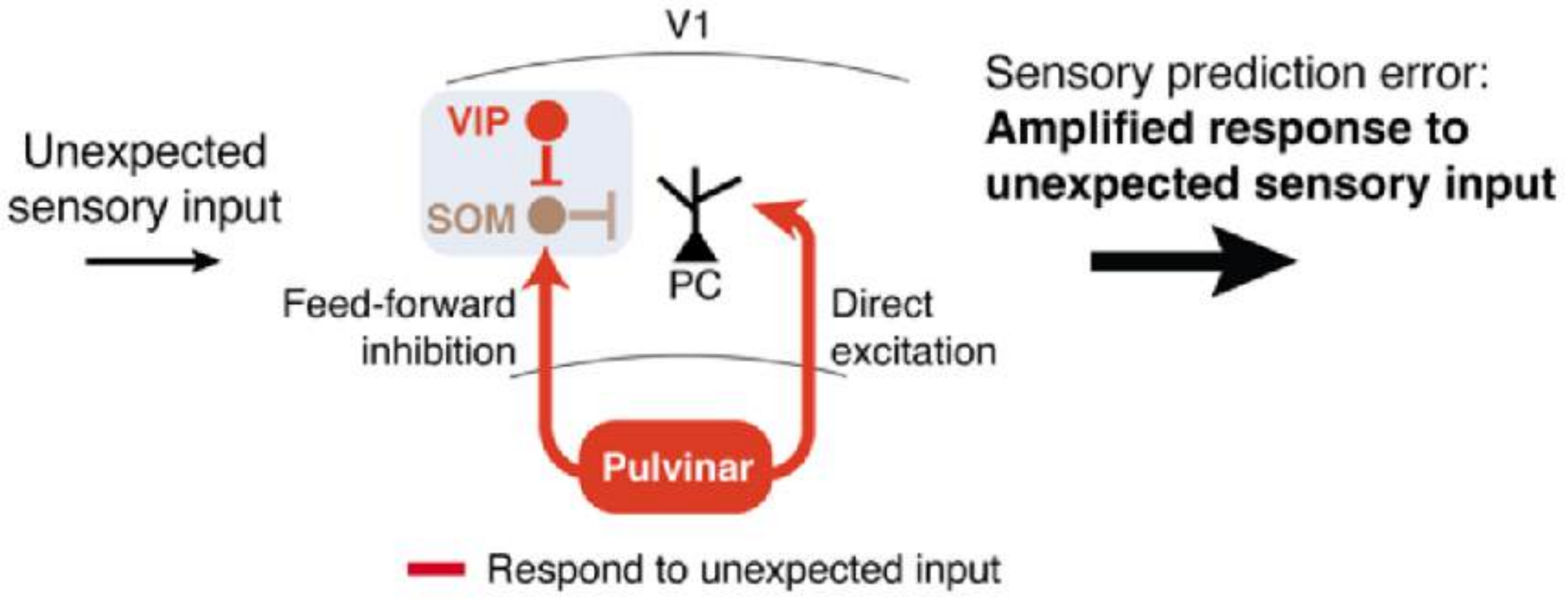
Pulvinar-driven response amplification



- **Pulvinar** drives **E cells** and a **subset of SST cells** (**SST1**)
  - SST1 get suppressed by unexpected stimuli due to inhibition from VIPs
- SST cells not driven by the pulvinar (SST2) were **activated** by **unexpected stimuli**

[Furutachi et al., 2023]

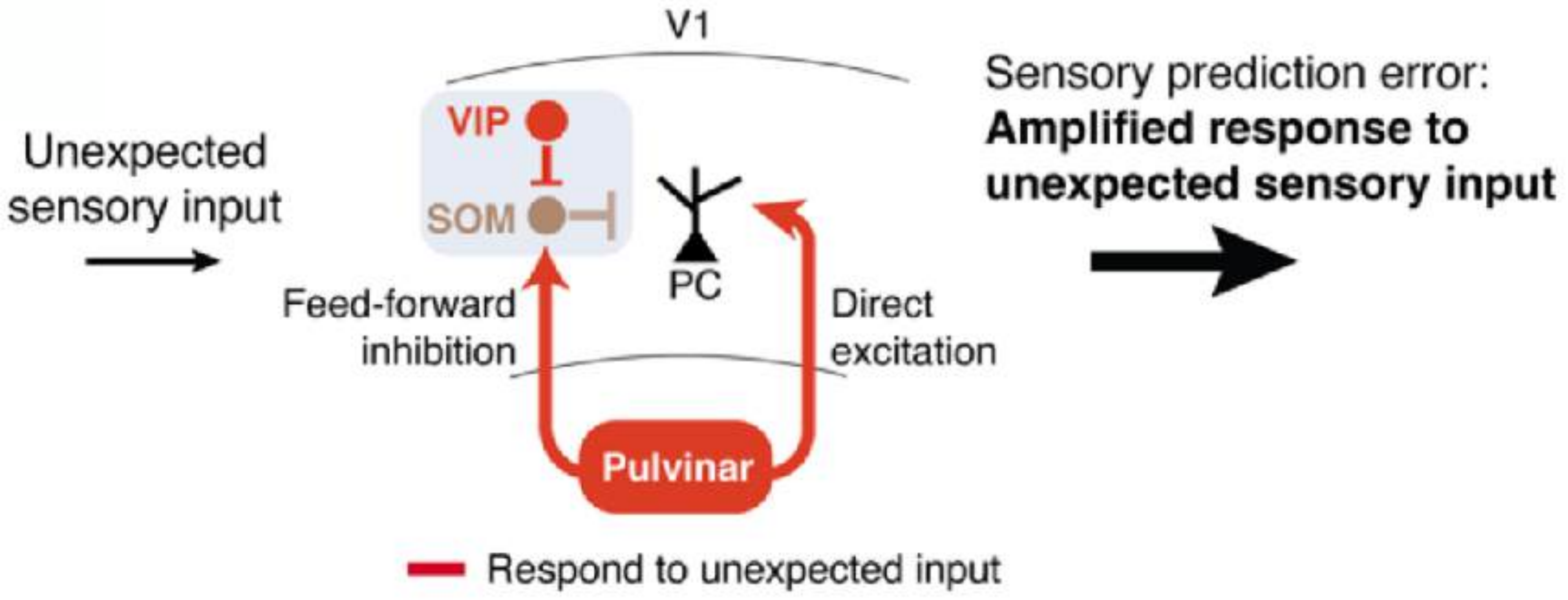
Pulvinar-driven response amplification



- **Pulvinar** drives **E cells** and a **subset of SST cells** (**SST1**)
  - SST1 get suppressed by unexpected stimuli due to inhibition from VIPs
- SST cells not driven by the pulvinar (SST2) were **activated** by **unexpected stimuli**
- VIPs simultaneous active with pulvinar input, **inhibit SST1** and **disinhibit E cells**

[Furutachi et al., 2023]

Pulvinar-driven response amplification

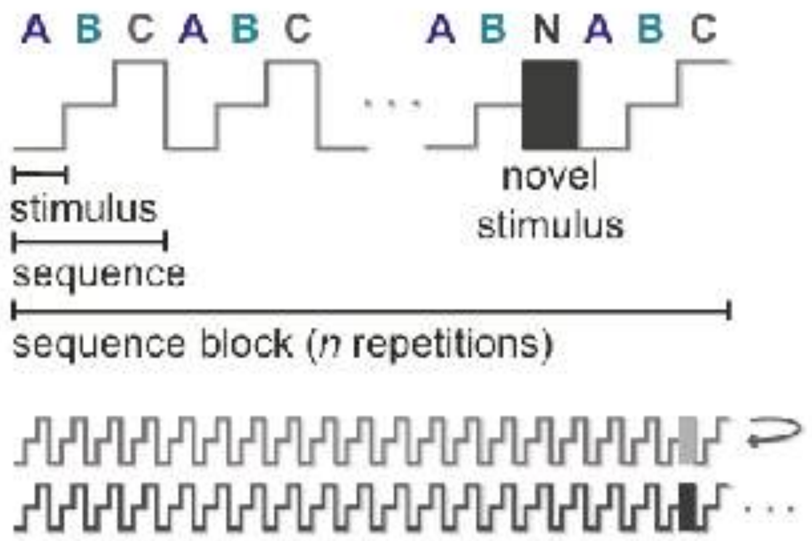


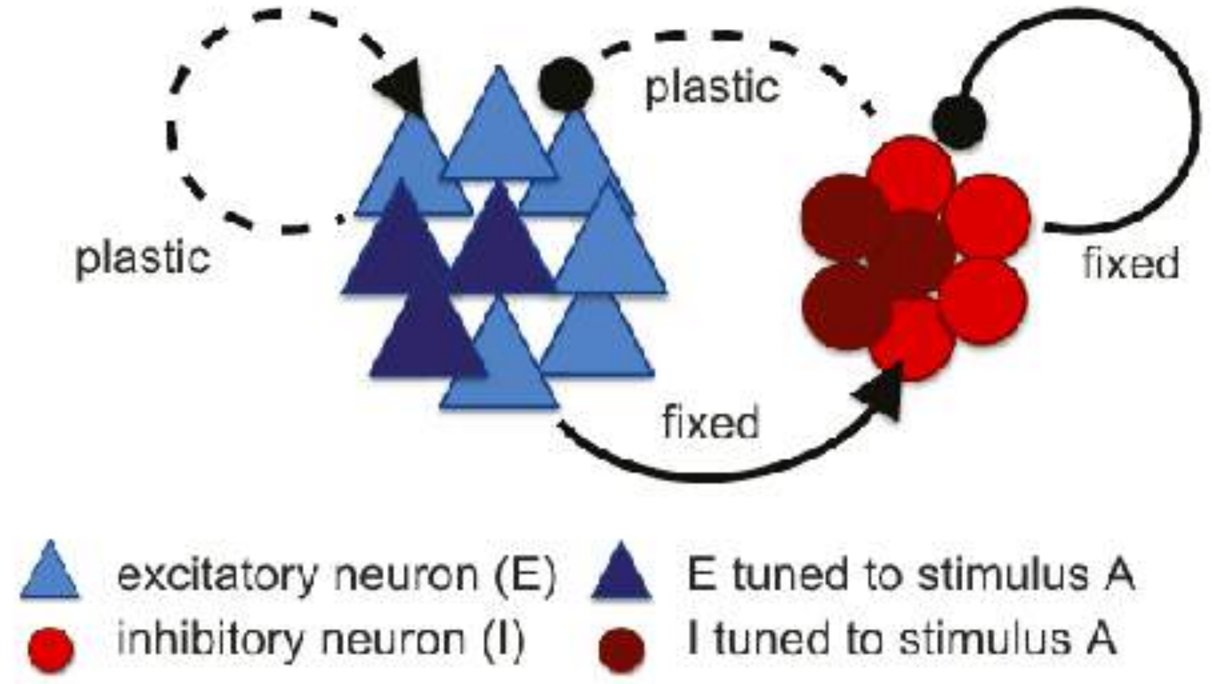
- **Pulvinar** drives **E cells** and a **subset of SST cells** (**SST1**)
  - SST1 get suppressed by unexpected stimuli due to inhibition from VIPs
- SST cells not driven by the pulvinar (SST2) were **activated** by **unexpected stimuli**
- VIPs simultaneous active with pulvinar input, **inhibit SST1** and **disinhibit E cells**

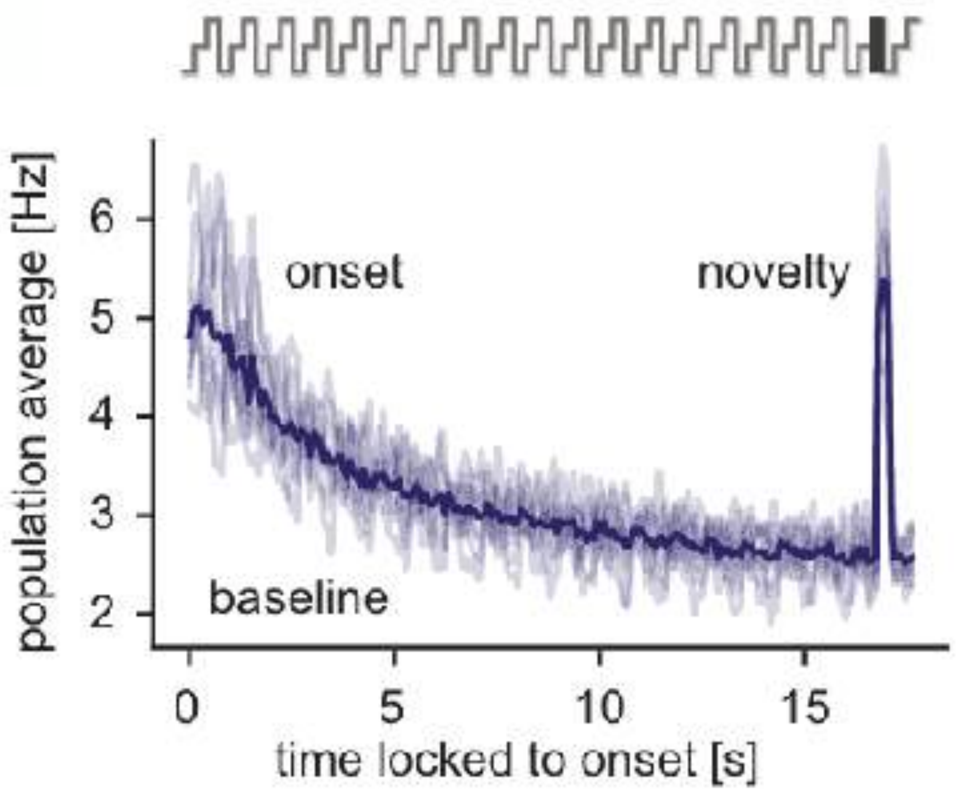
[Furutachi et al., 2023]

Adaptation through recurrent inhibitory plasticity

Adaptation through recurrent inhibitory plasticity

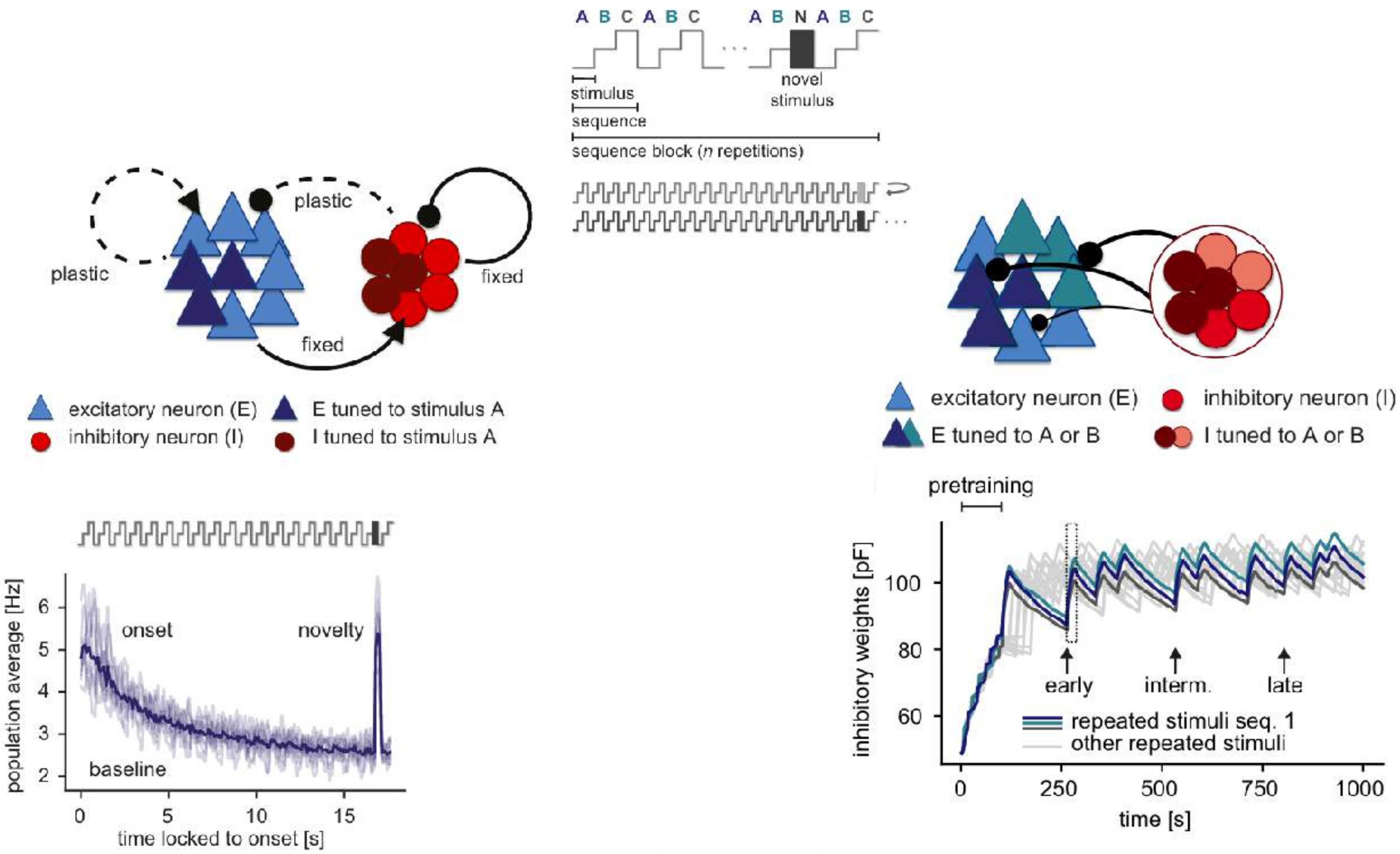






[Schulz, Miehl et al., 2021]

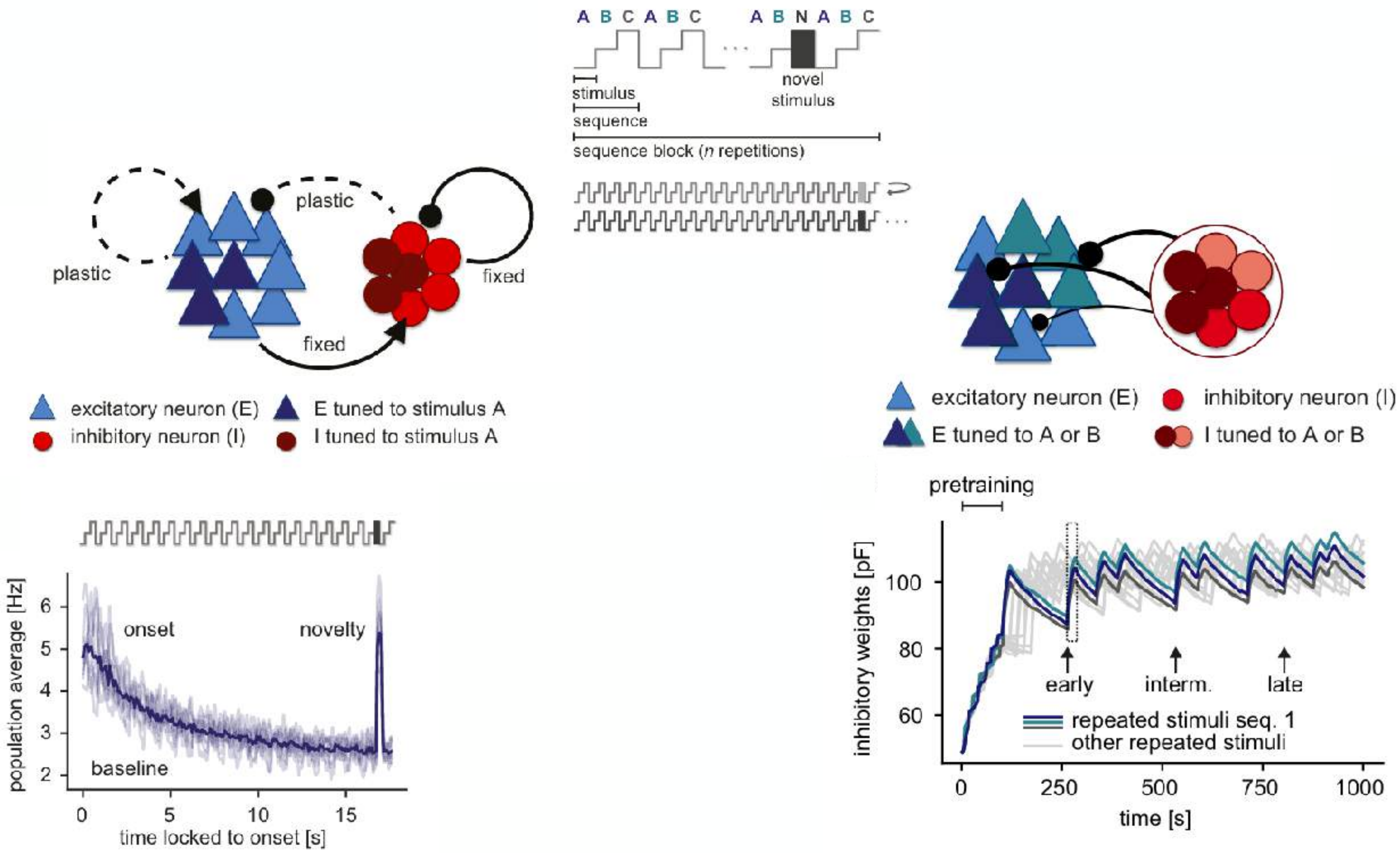
Adaptation through recurrent inhibitory plasticity



Increased recurrent inhibition on stimulus **tuned** excitatory **assemblies** reduces population responses

[Schulz, Miehl et al., 2021]

Adaptation through recurrent inhibitory plasticity



Increased recurrent inhibition on stimulus **tuned** excitatory **assemblies** reduces population responses

How are **different inhibitory cell-types involved** in the **adaptation process** and the resulting **novelty responses**?

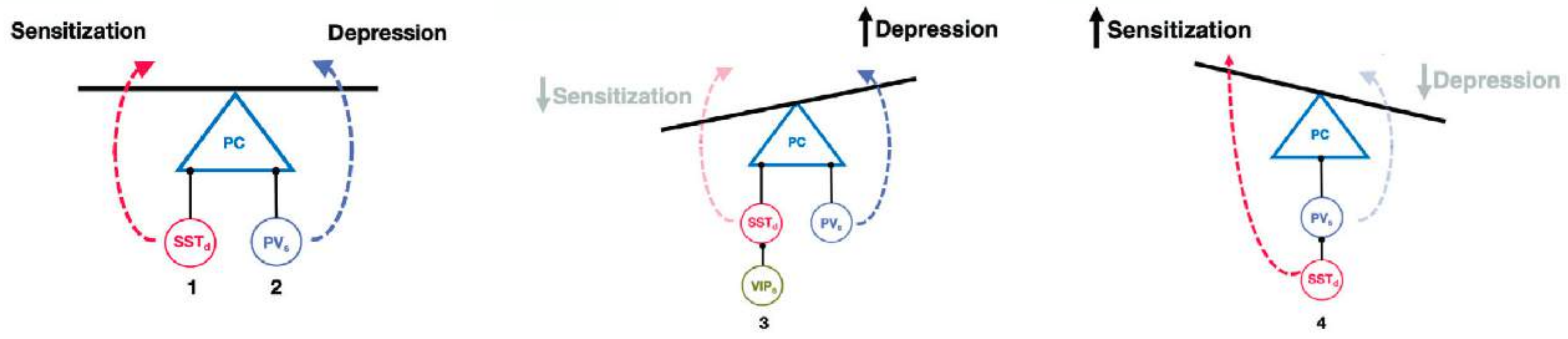
[Schulz, Miehl et al., 2021]

Differential involvement of interneuron subtypes in adaptation

- In mouse V1, activation of different disinhibition pathways results in opposing effects on the gain of E cells
    - Opposing contribution of activation of VIP -> SST and SST->PV pathway

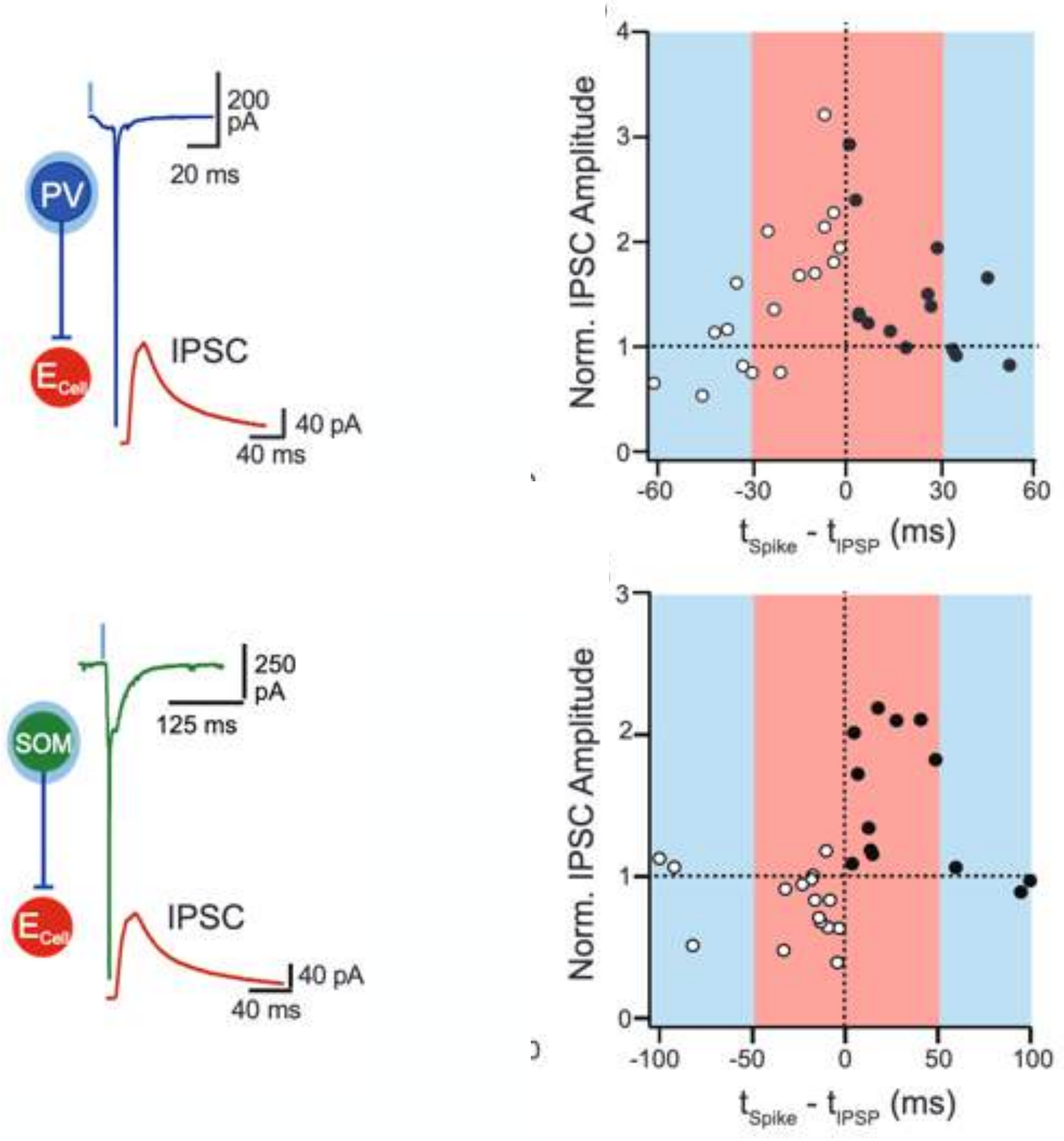
Differential involvement of interneuron subtypes in adaptation

- In mouse V1, activation of different disinhibition pathways results in opposing effects on the gain of E cells
    - Opposing contribution of activation of VIP -> SST and SST->PV pathway



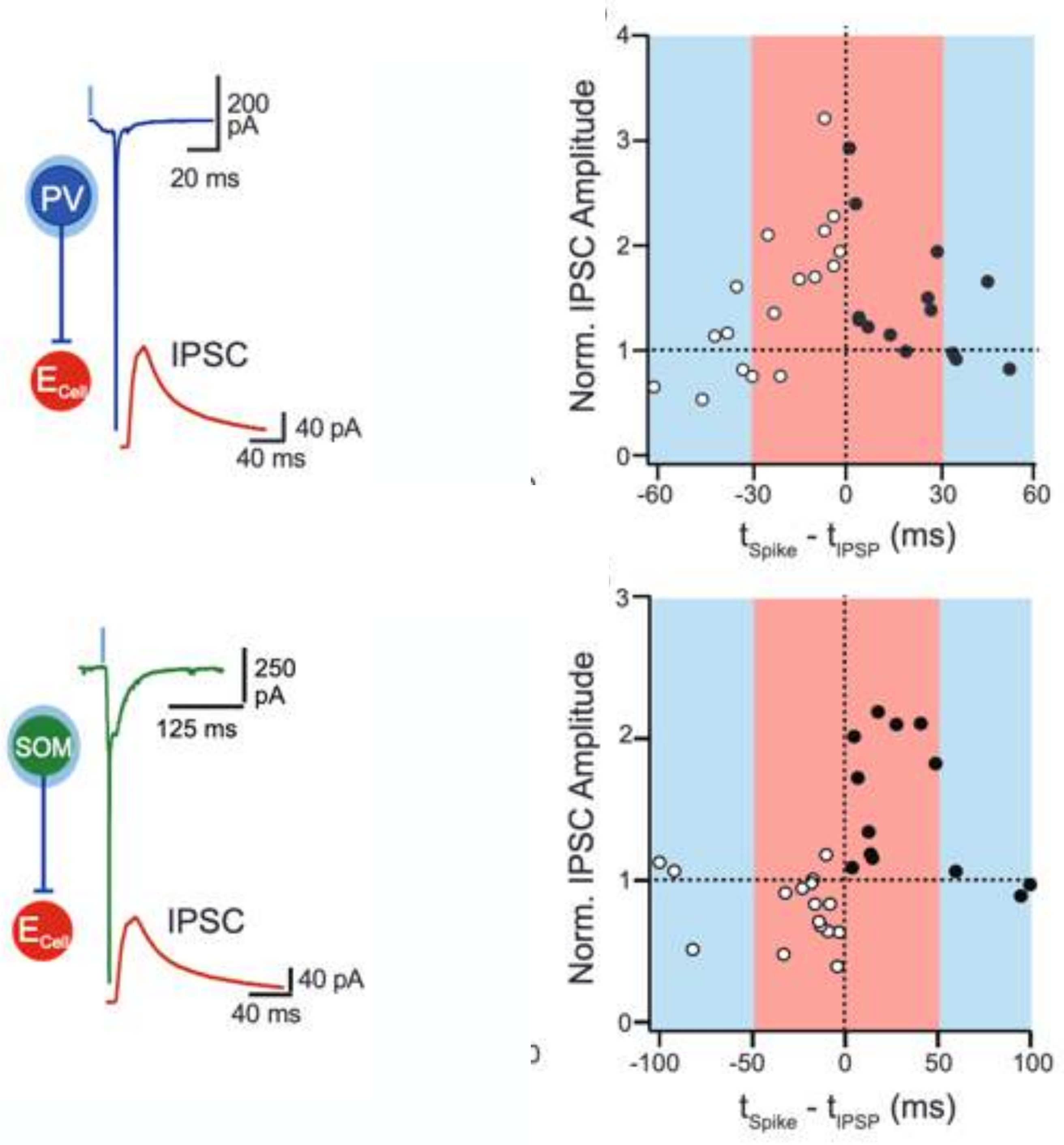
Differential involvement of interneuron subtypes in adaptation

- In mouse V1, activation of different disinhibition pathways results in opposing effects on the gain of E cells
    - Opposing contribution of activation of VIP -> SST and SST->PV pathway
- Different forms of plasticity for different interneuron type
    - PV stabilise assemblies
      SST accelerate their formation



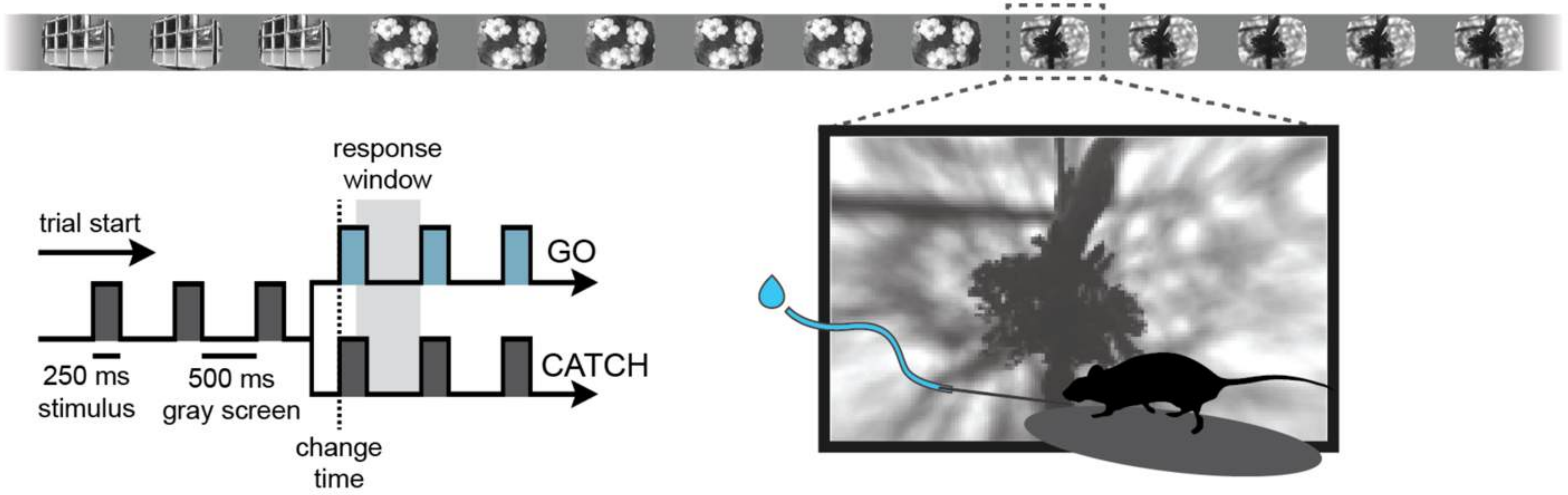
Differential involvement of interneuron subtypes in adaptation

- In mouse V1, activation of different disinhibition pathways results in opposing effects on the gain of E cells
    - Opposing contribution of activation of VIP -> SST and SST->PV pathway
- Different forms of plasticity for different interneuron type
    - PV stabilise assemblies SST accelerate their formation

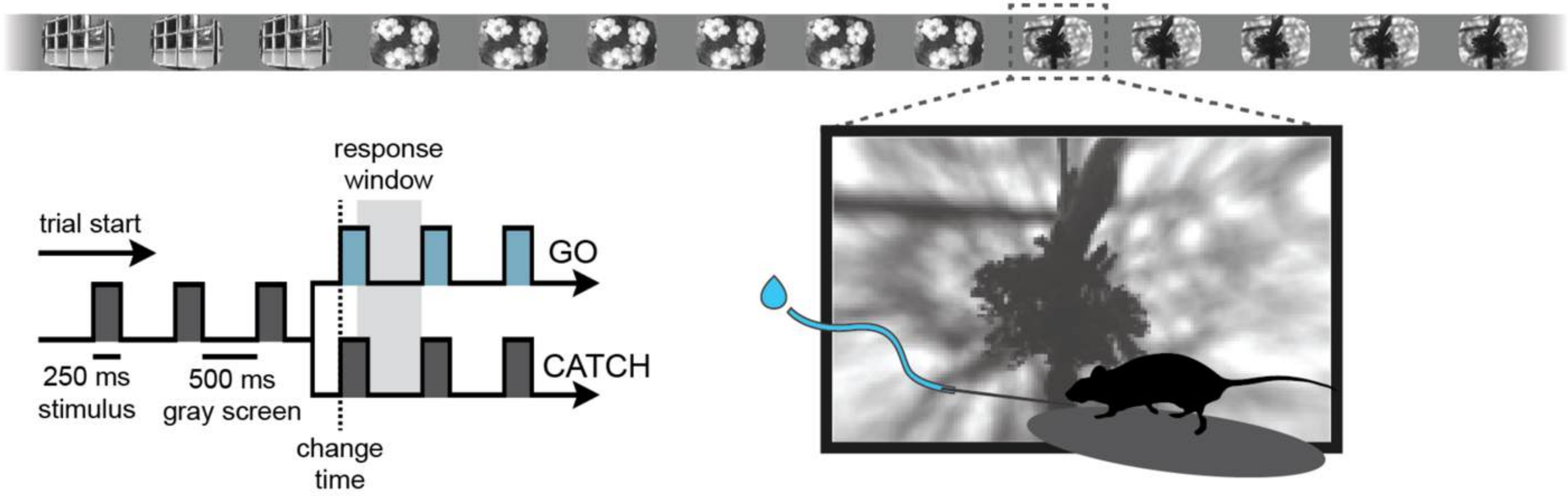


[Lagzi et al., 2021]

Sequential visual discrimination task



Sequential visual discrimination task



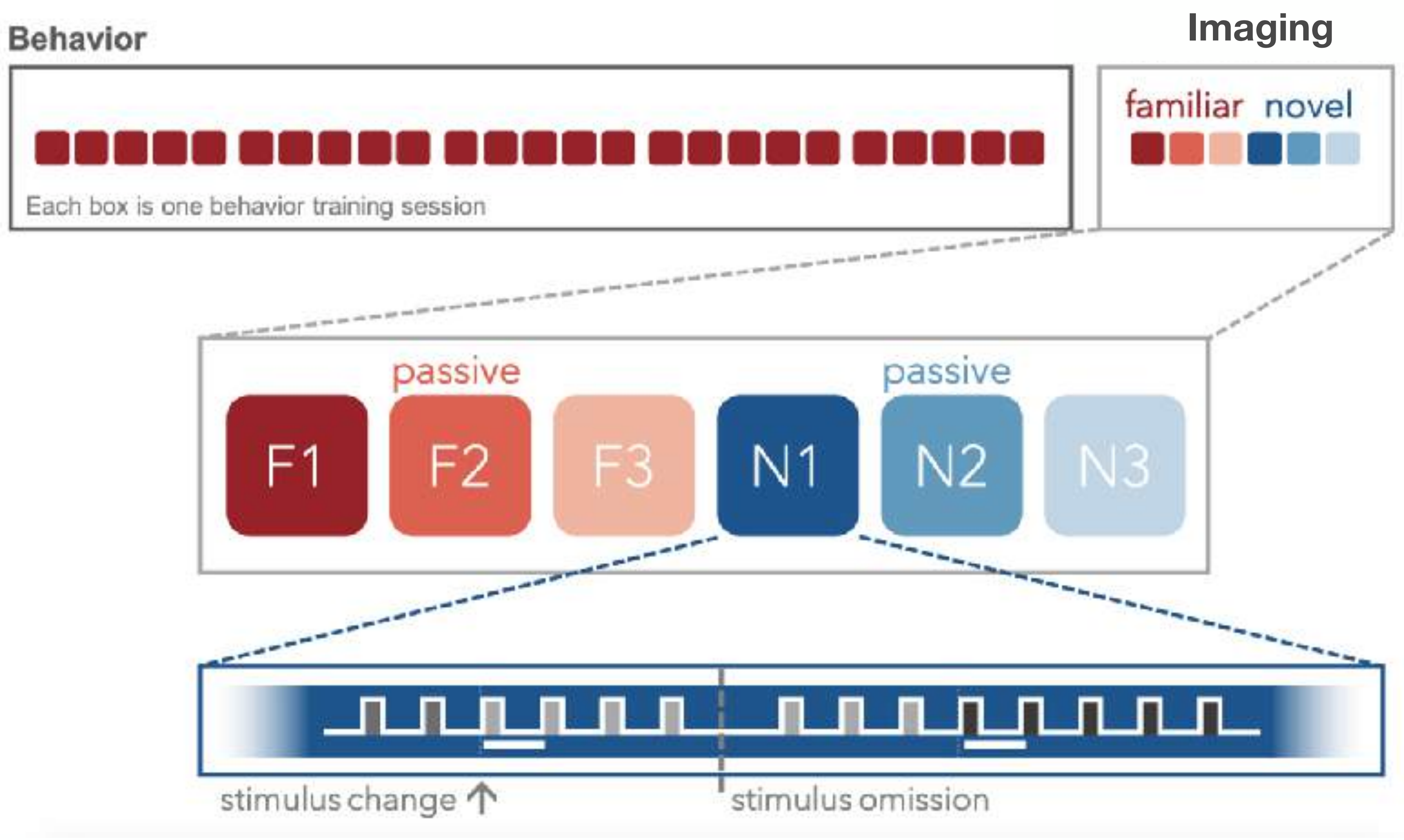
[Garrett et al., 2019]

Allen Institute visual behavior dataset

Calcium imaging

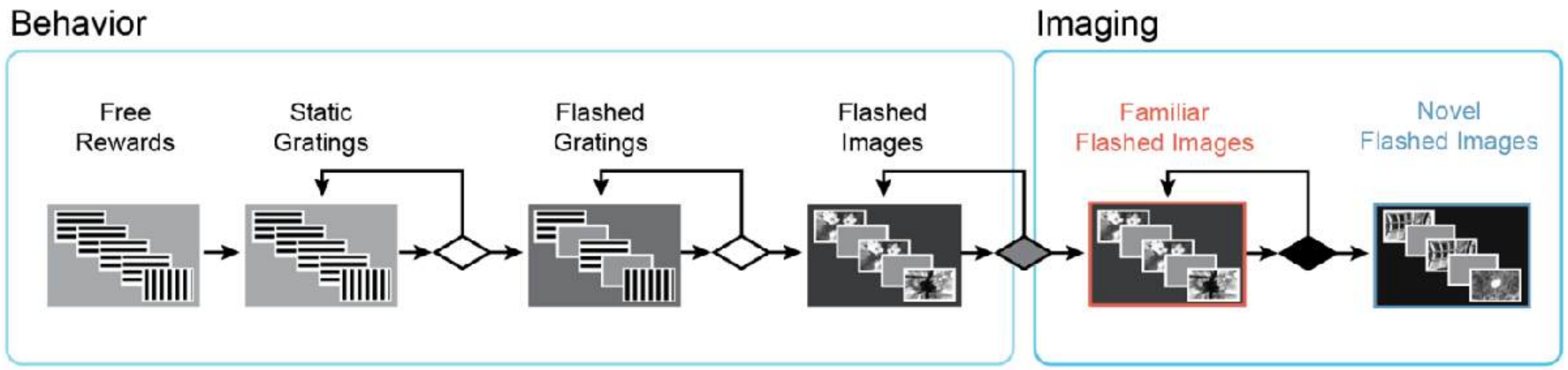
Allen Institute visual behavior dataset

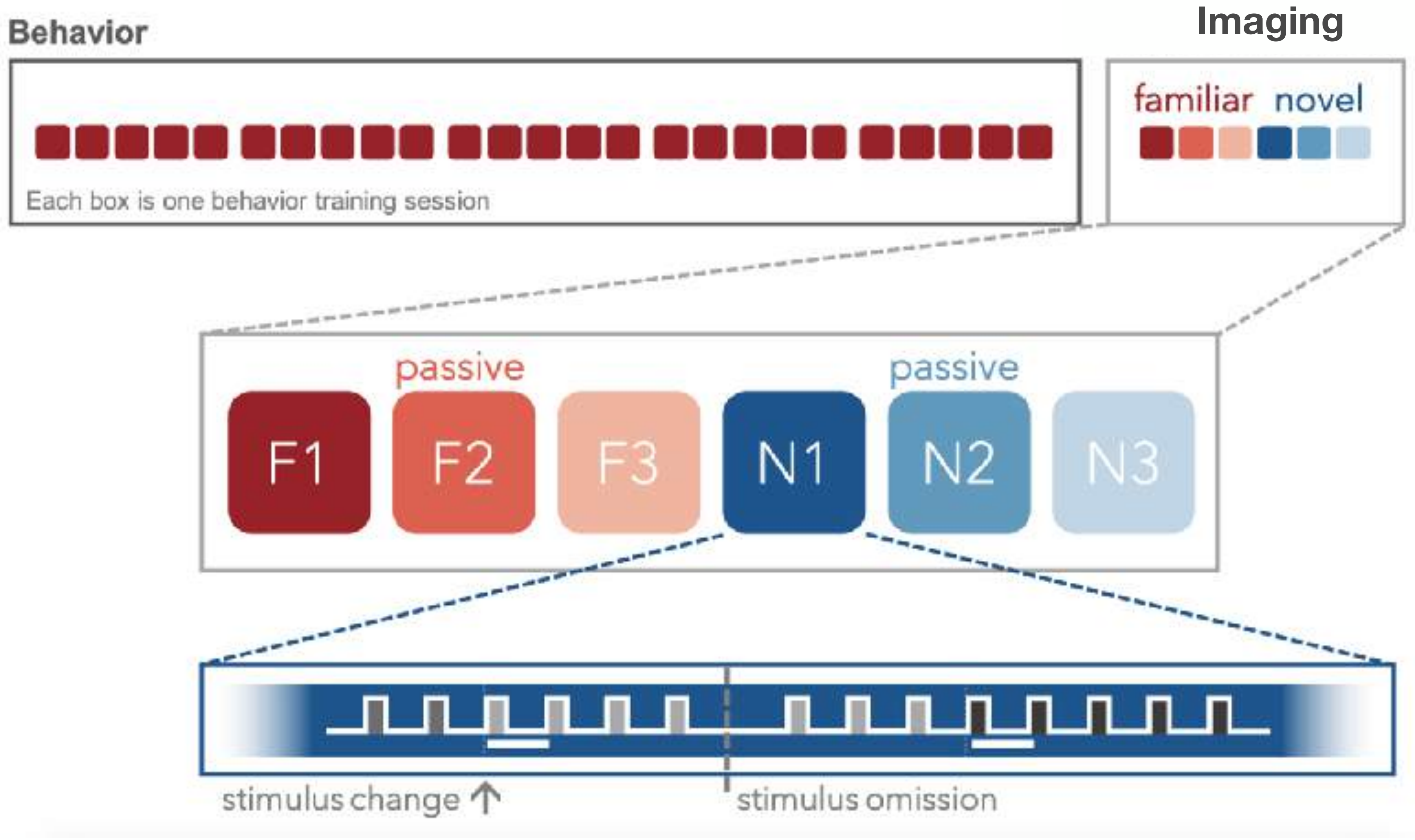
Calcium imaging



Allen Institute visual behavior dataset

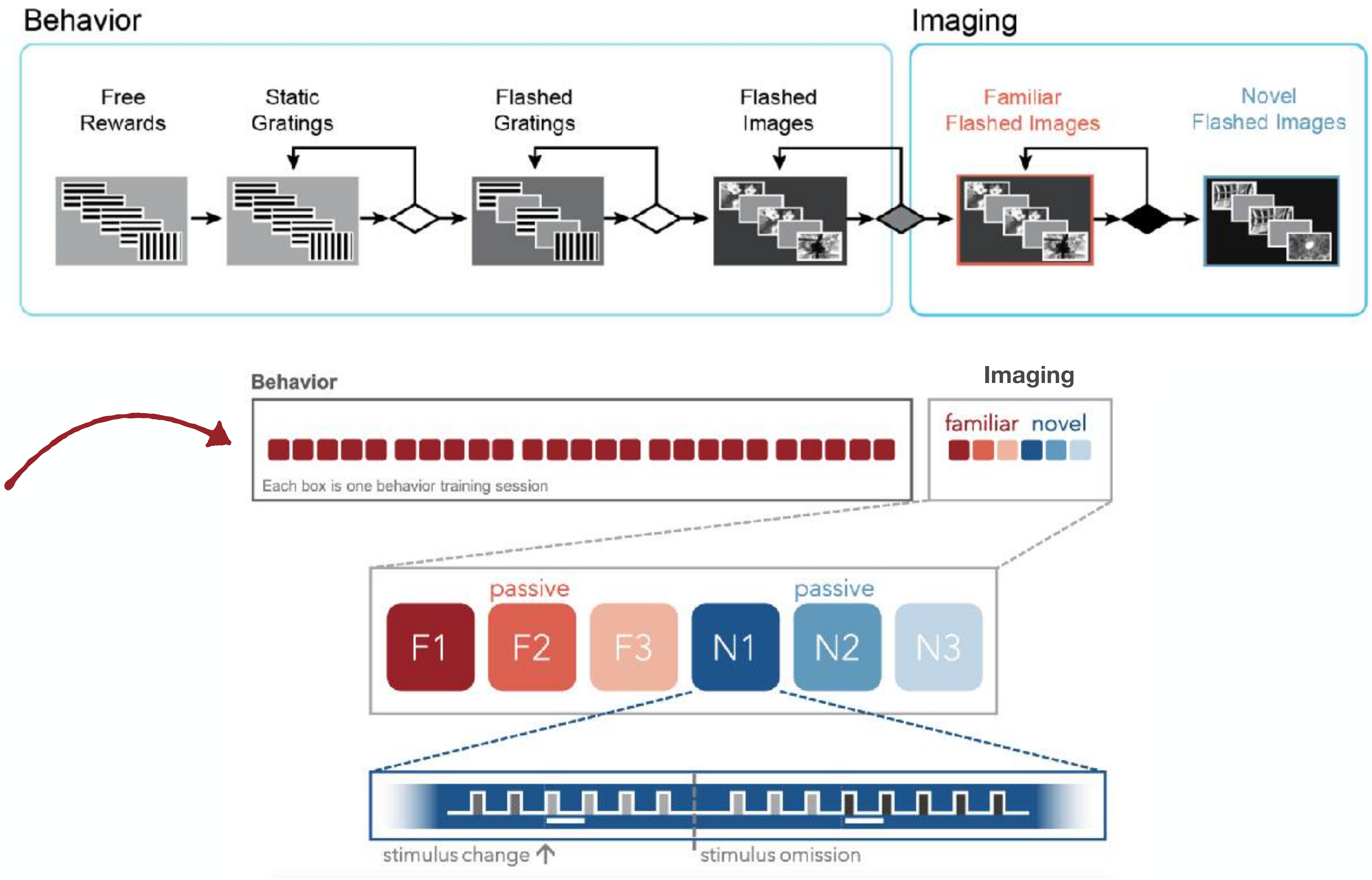
Calcium imaging





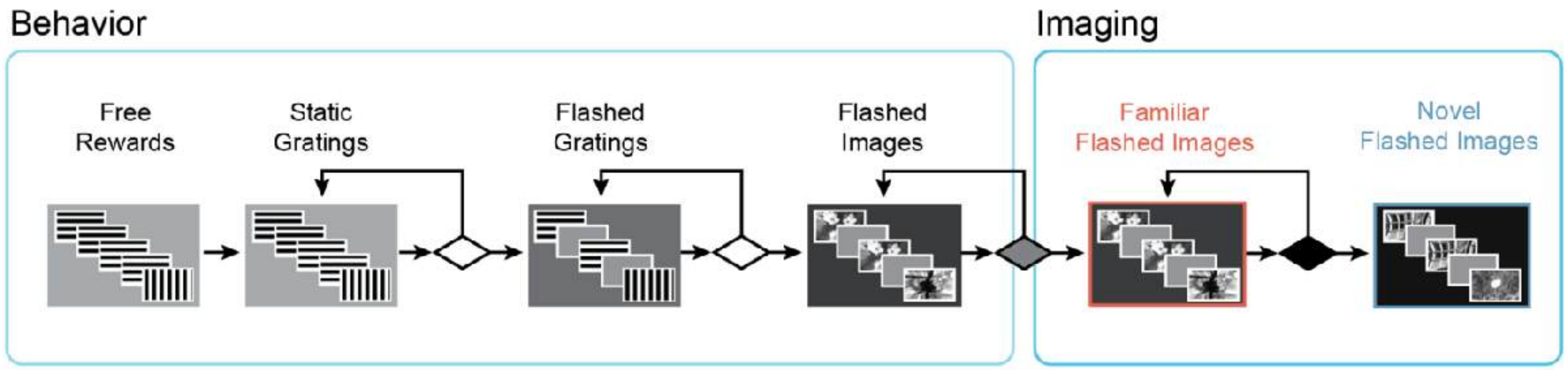
Allen Institute visual behavior dataset

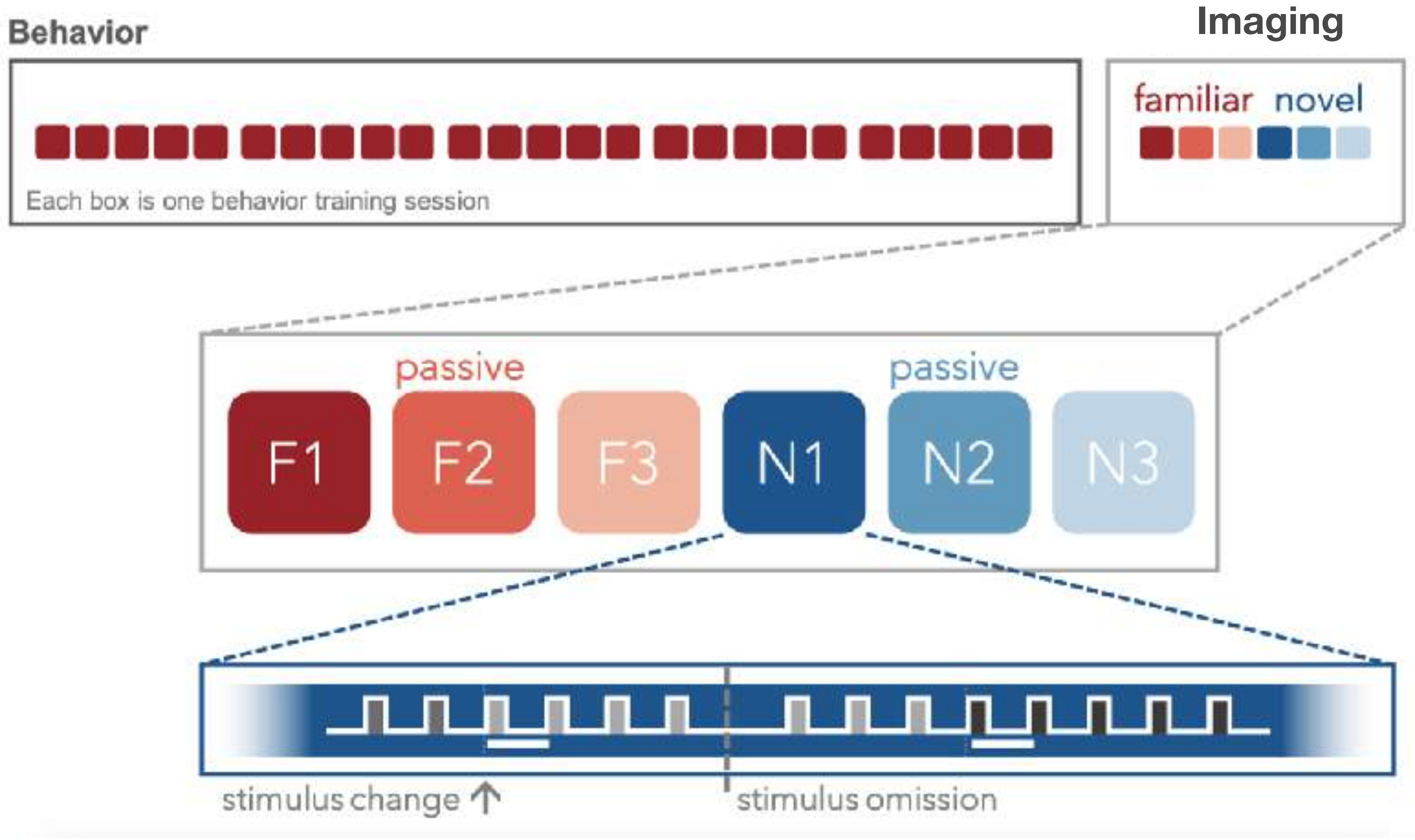
Calcium imaging



Allen Institute visual behavior dataset

Calcium imaging

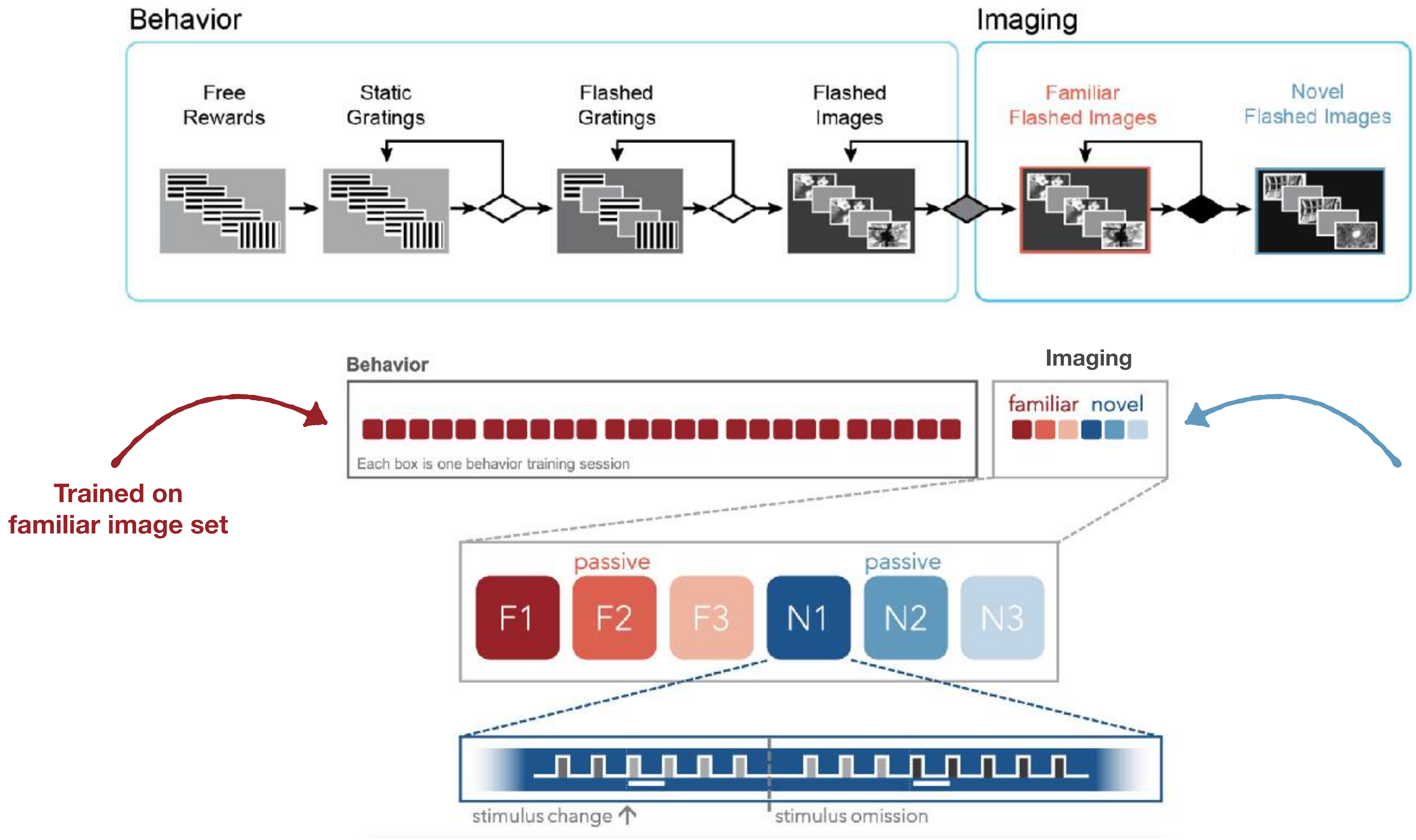




Trained on familiar image set

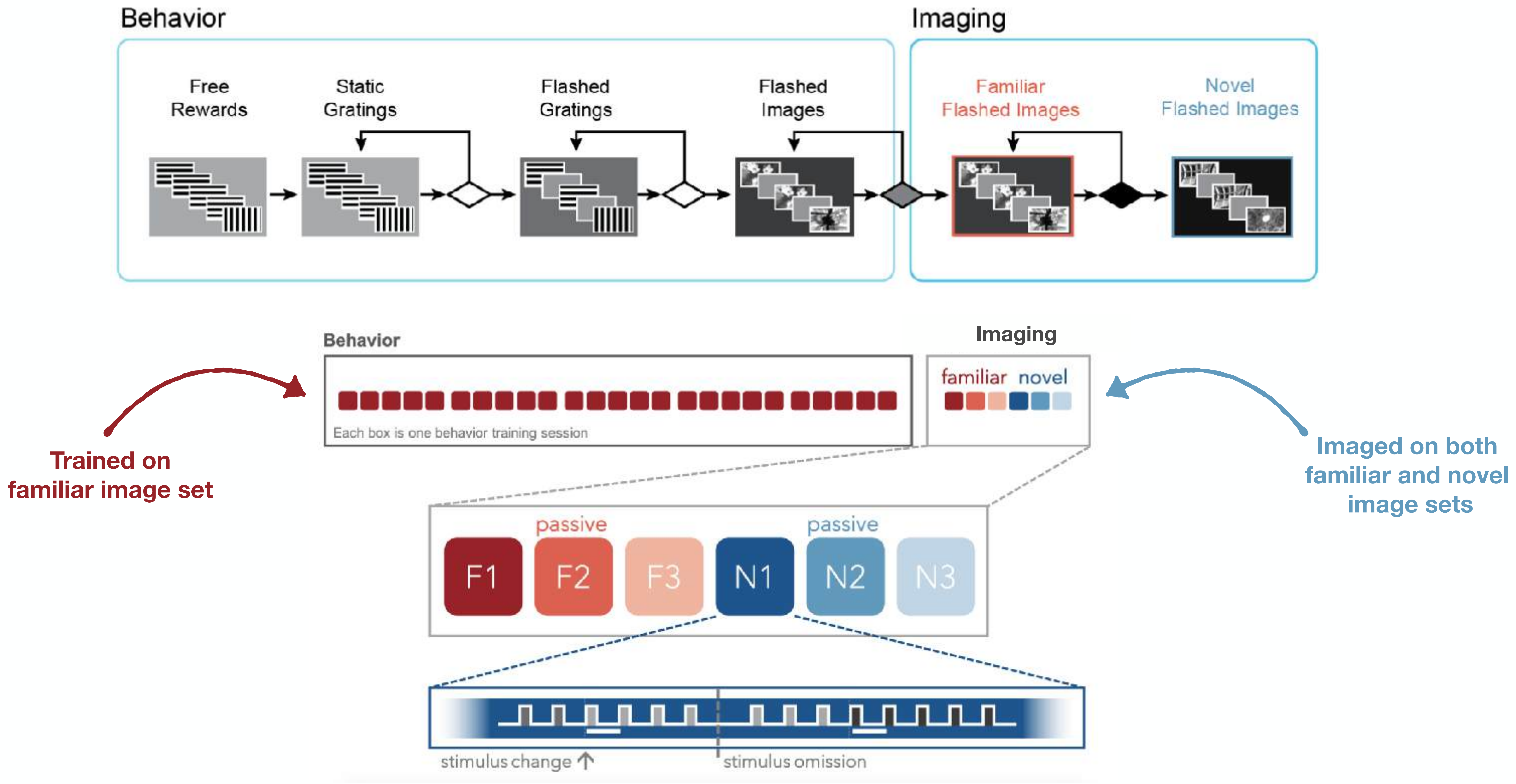
Allen Institute visual behavior dataset

Calcium imaging



Allen Institute visual behavior dataset

Calcium imaging



Prediction errors during training
(Pre-training - not recorded)

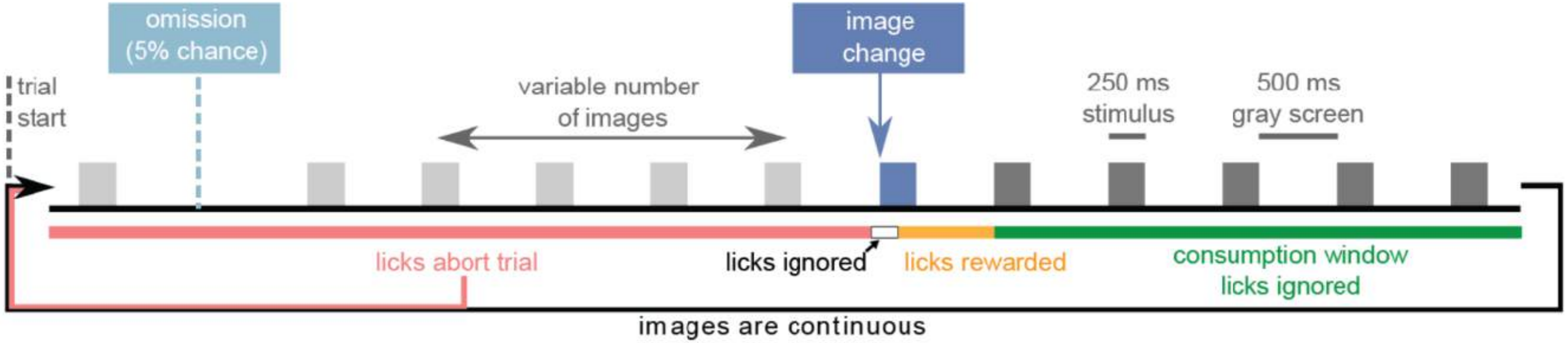
- Initial association between gratings and reward: positive prediction error after each stimulus presentation
- Association between stimulus change and reward:
  negative prediction error after each image presentation;
  positive prediction error after each stimulus change
- Trained animal:
  no prediction error after each stimulus change

Different forms of expectation violations

- **Stimulus expectation violations**
    - **Temporal expectation violations**/ absence of expected stimulus (omissions)
    - **Stimulus identity expectation violation**
        - **Contextual** expectation violation (*contextual stimulus novelty*)
        - **Familiarity** expectation violation (*absolute stimulus novelty*)
- **Reward expectation violations**

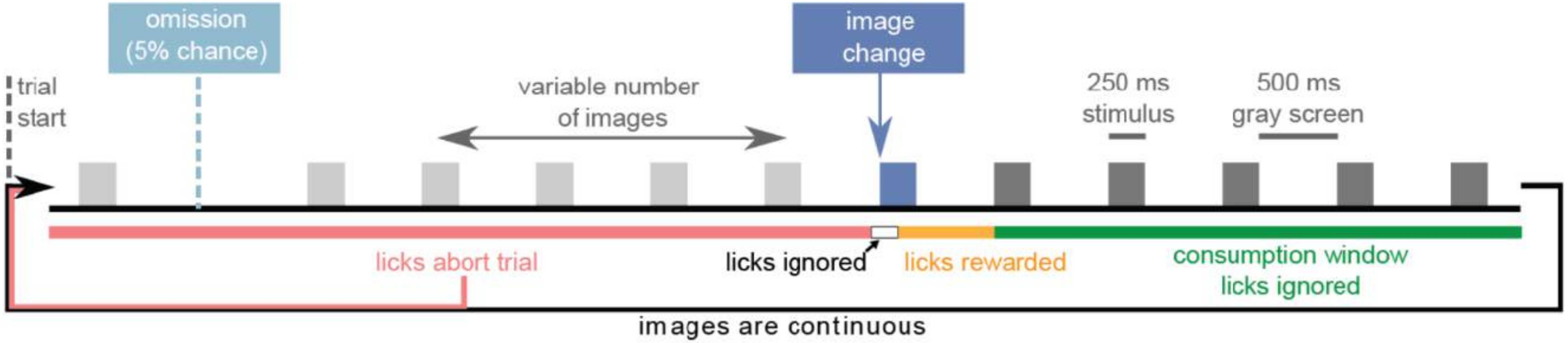
Allen Institute visual behavior dataset

Calcium imaging

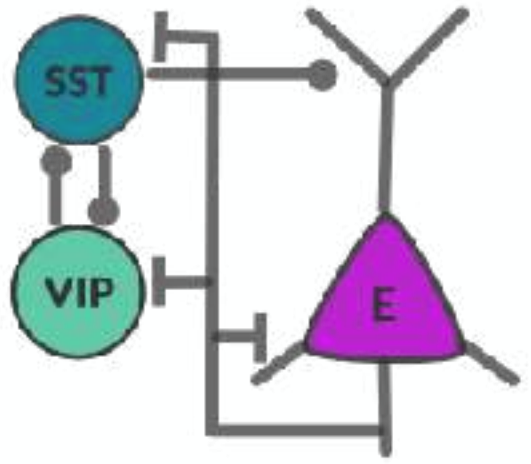


Allen Institute visual behavior dataset

Calcium imaging

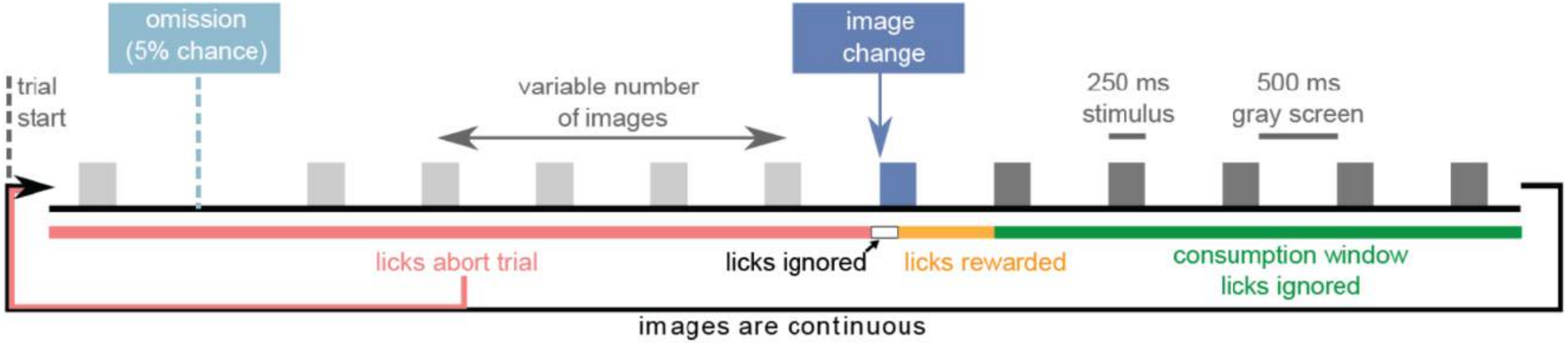


recorded cell types

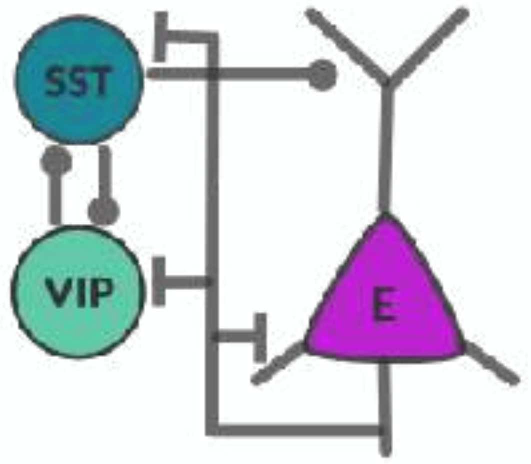


Allen Institute visual behavior dataset

Calcium imaging



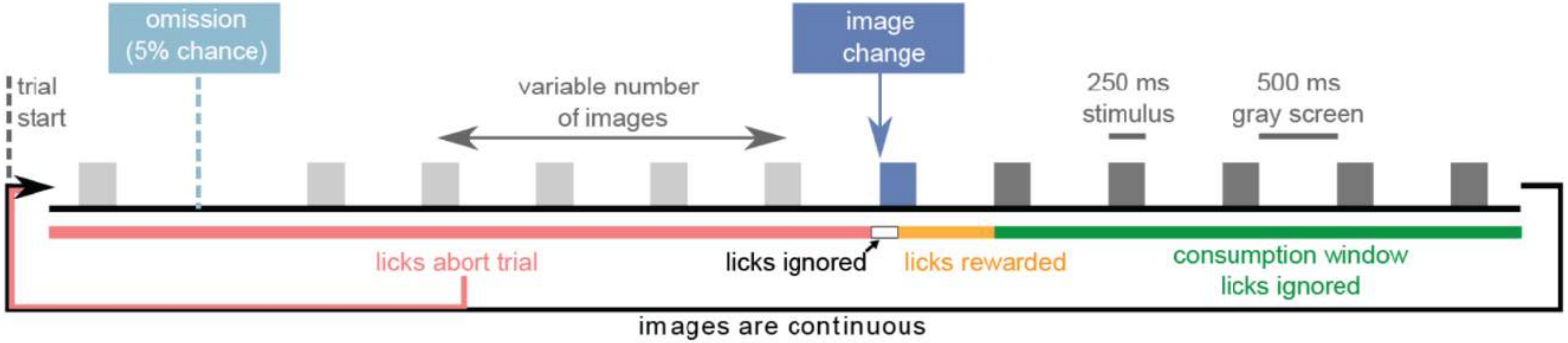
recorded cell types



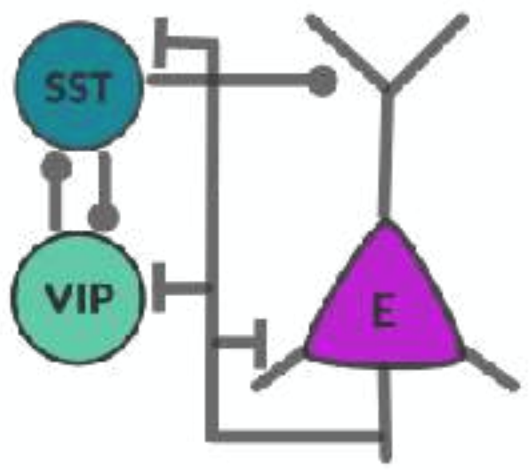
Recordings
for each cell type
come from different mice

Allen Institute visual behavior dataset

Calcium imaging



recorded cell types



Recordings
for each cell type
come from different mice

image change

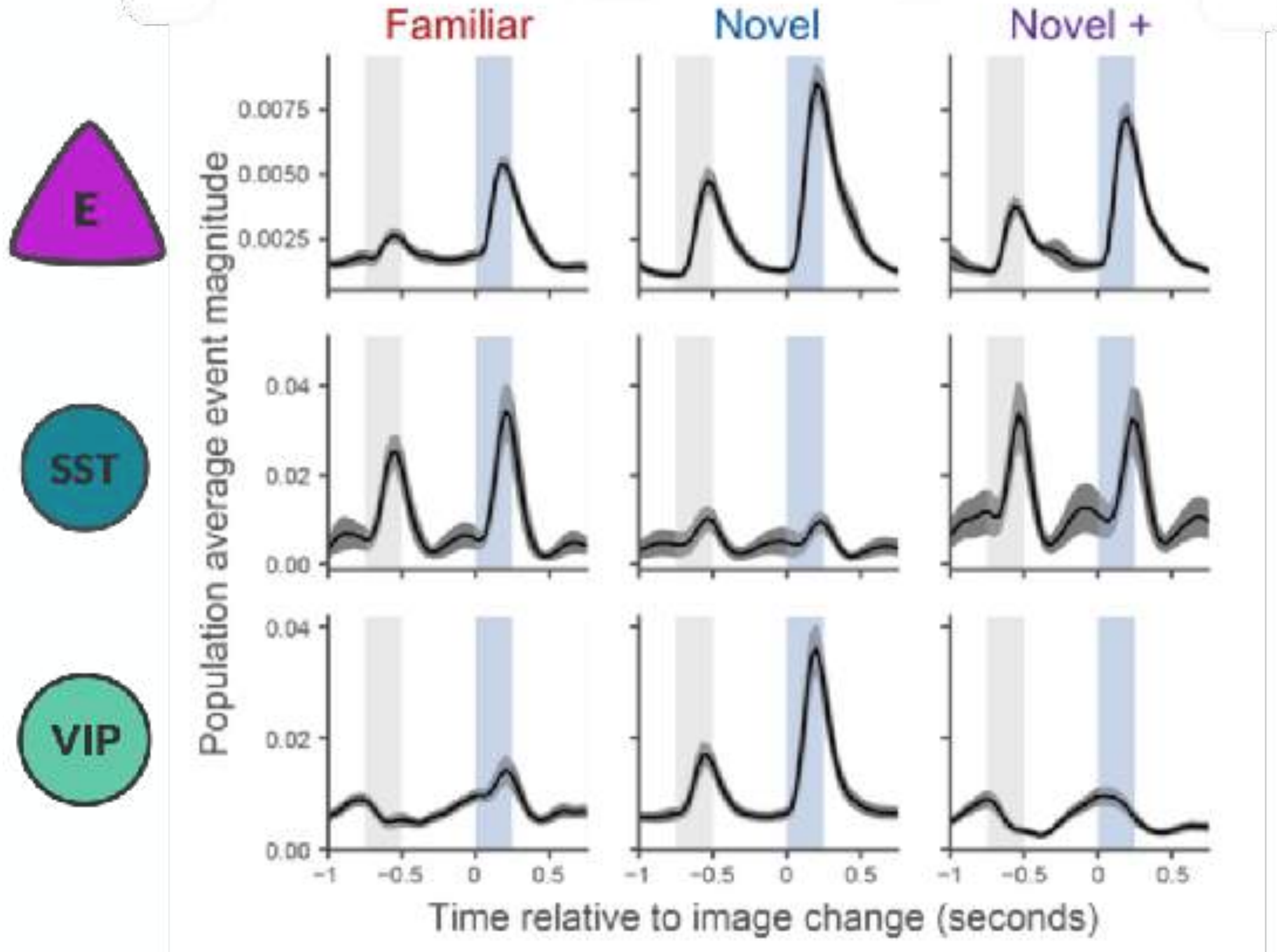


Image change responses

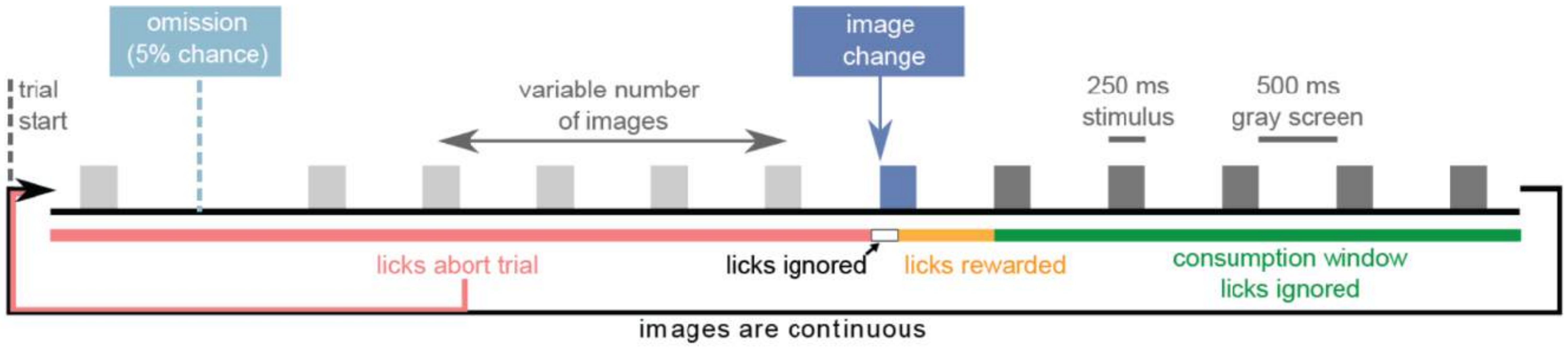


Image change responses

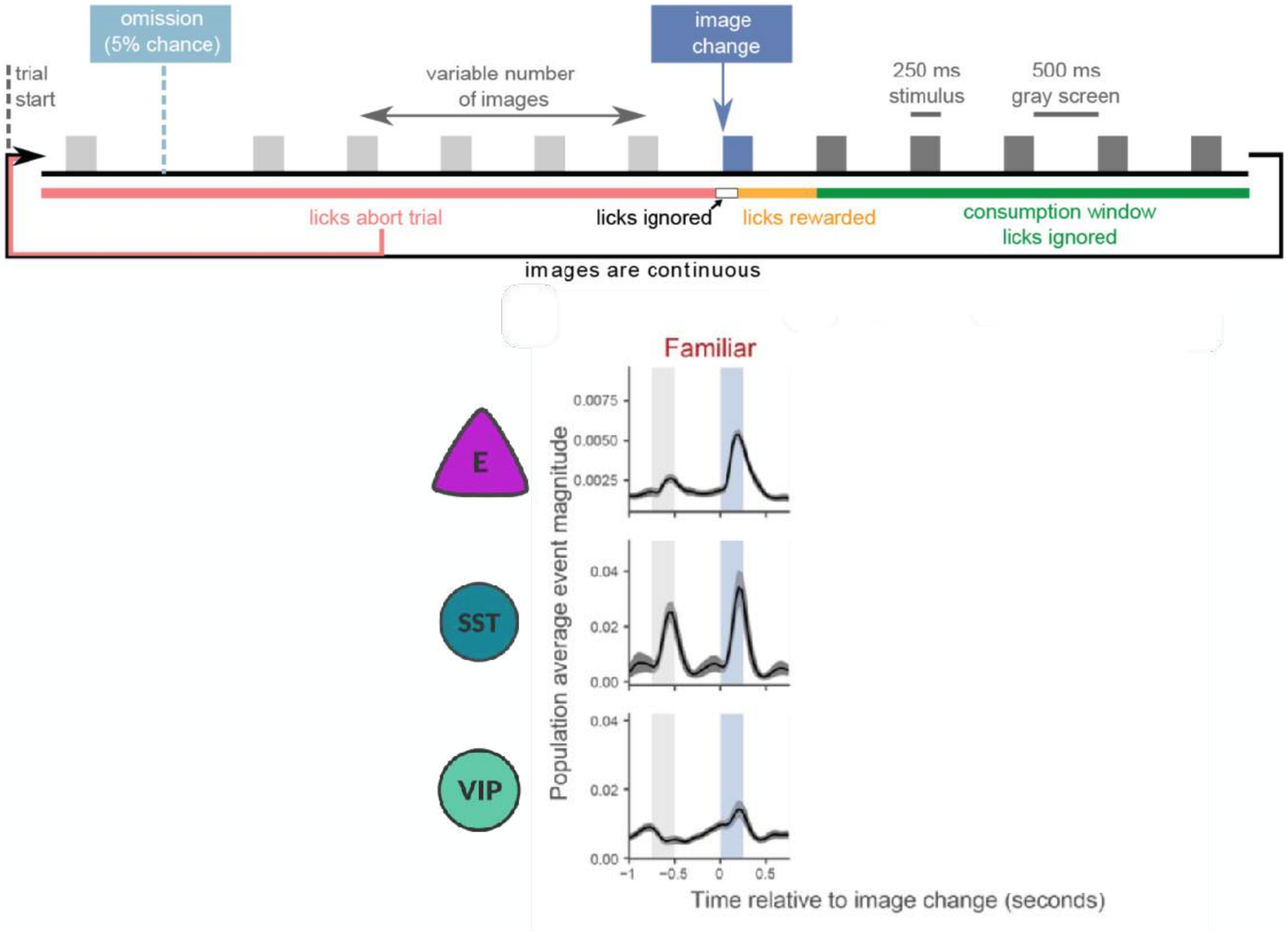


Image change responses

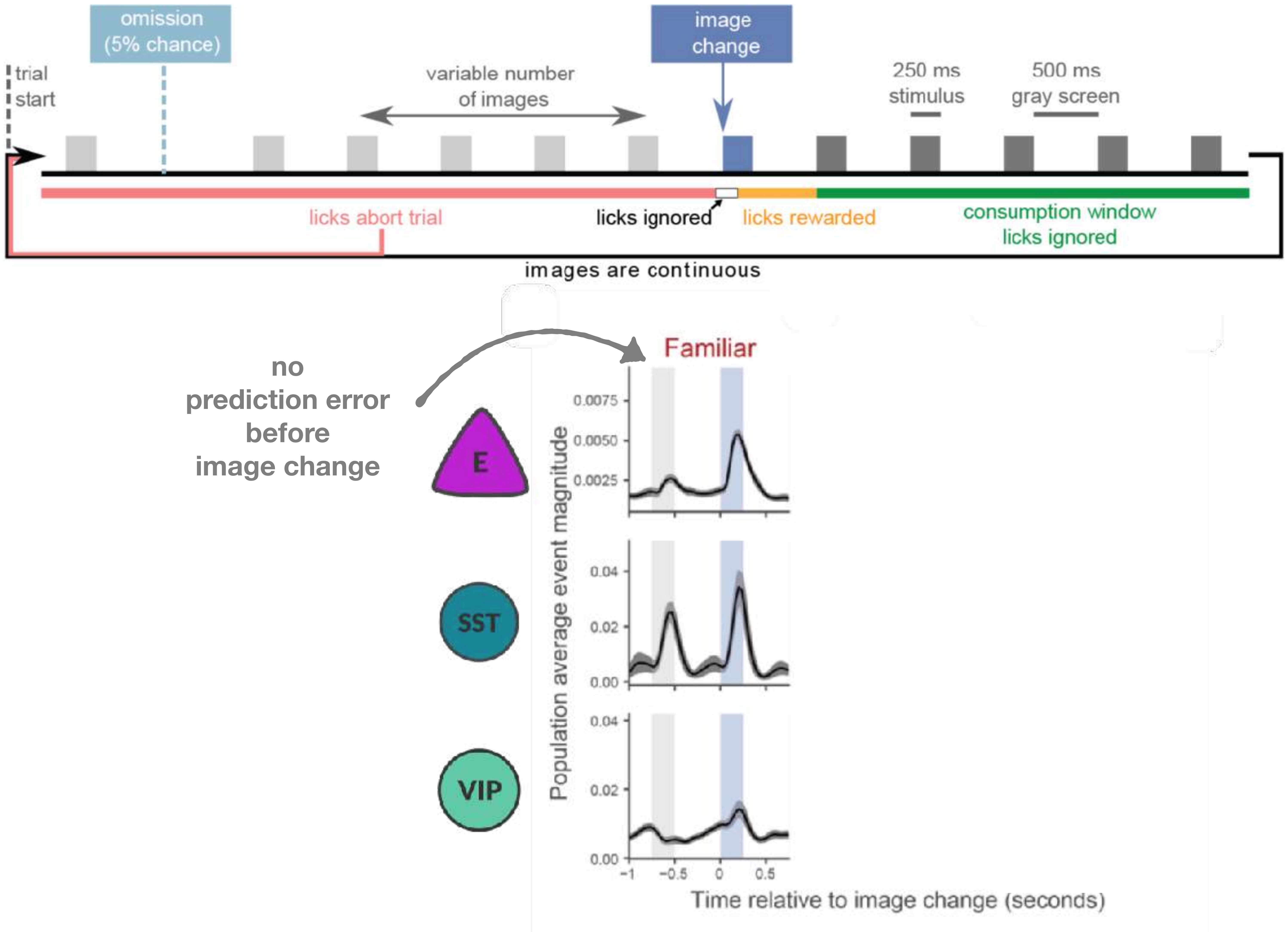


Image change responses

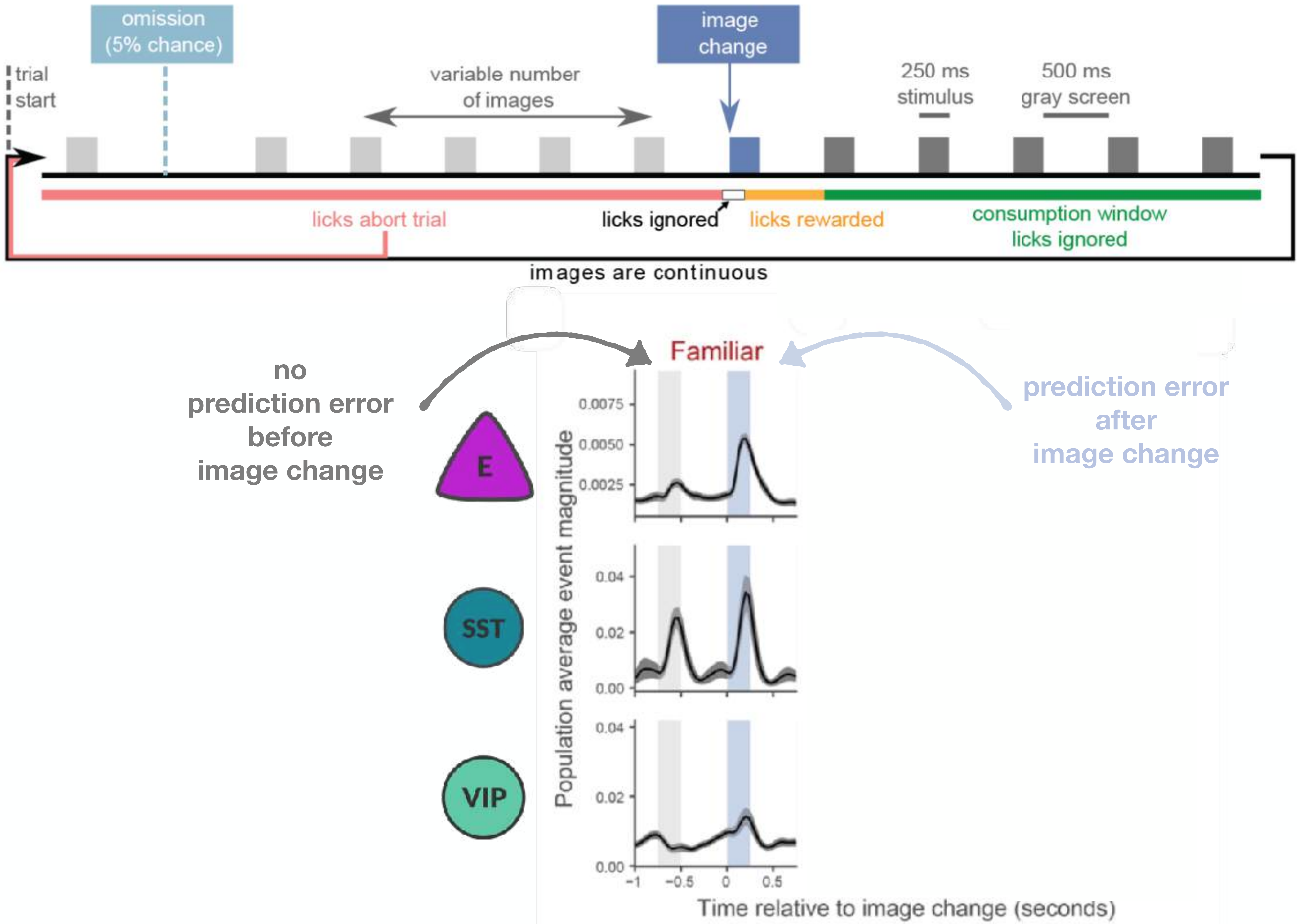


Image change responses

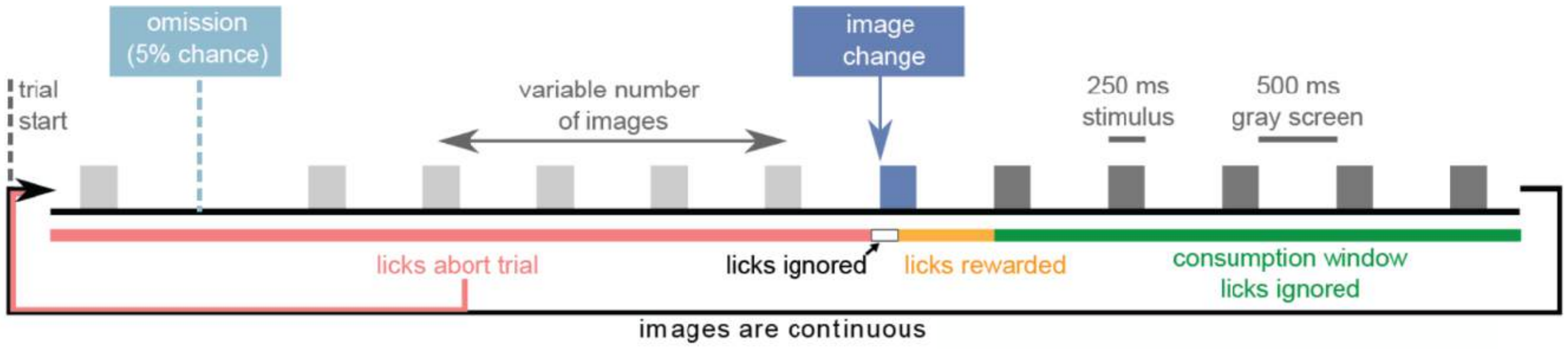
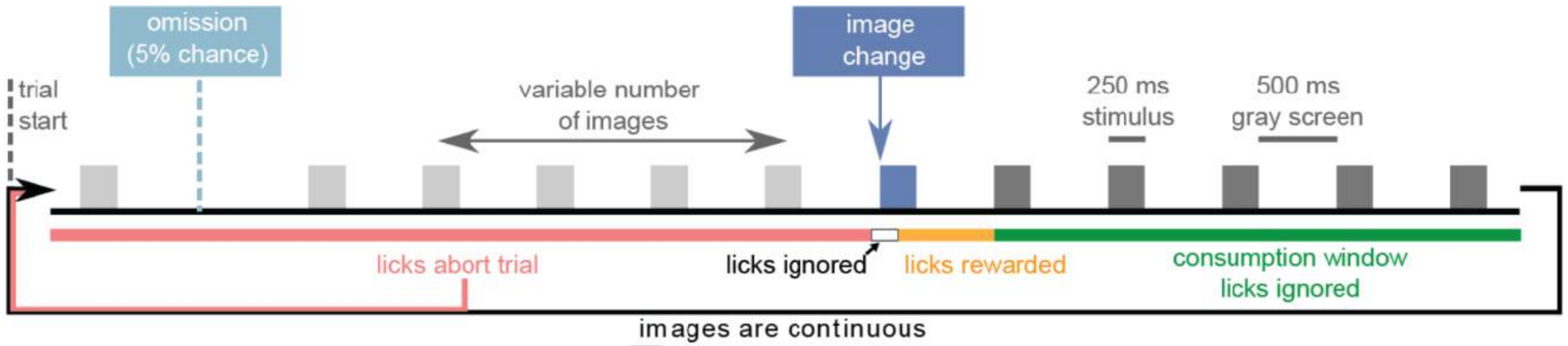


Image change responses



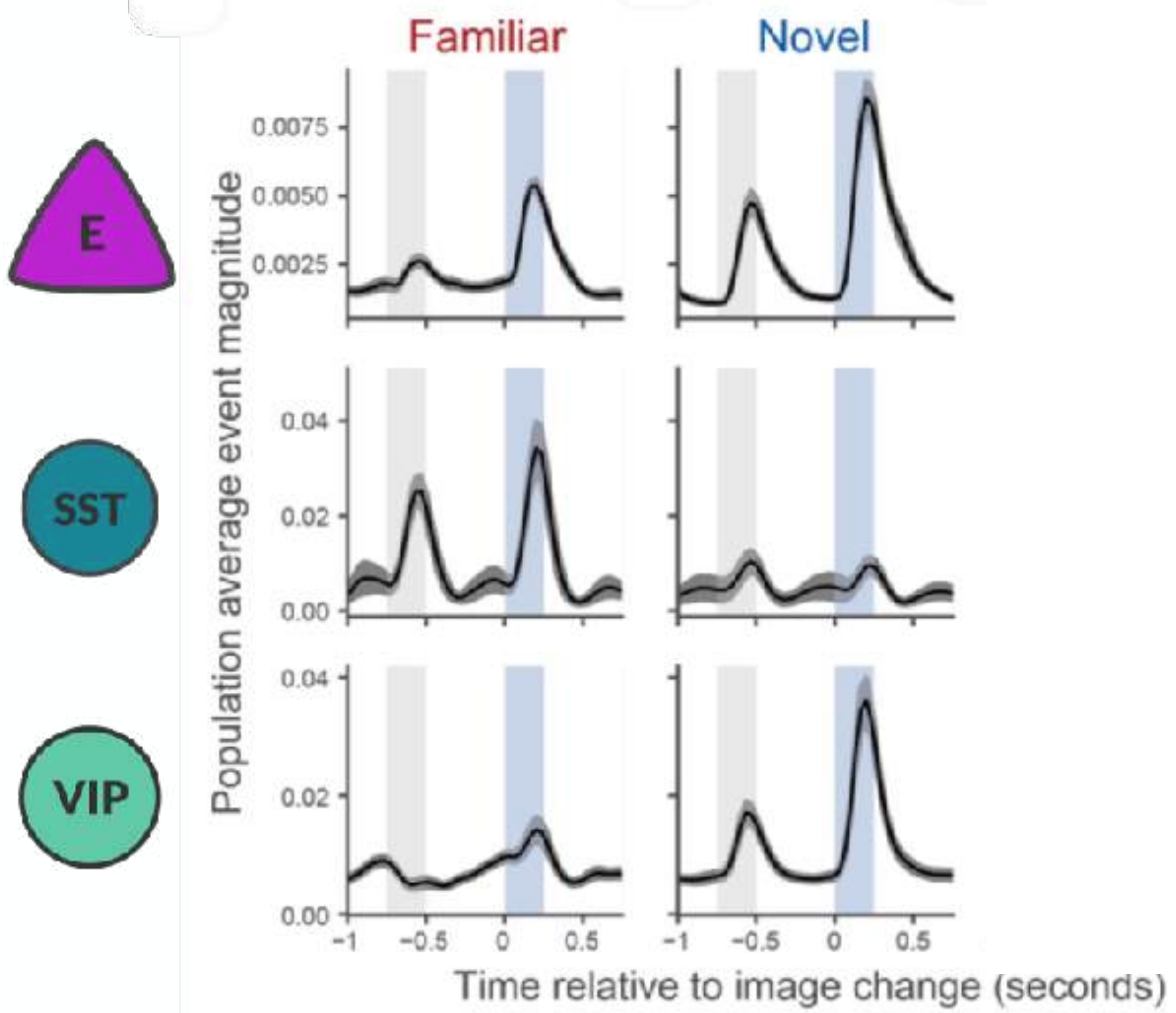
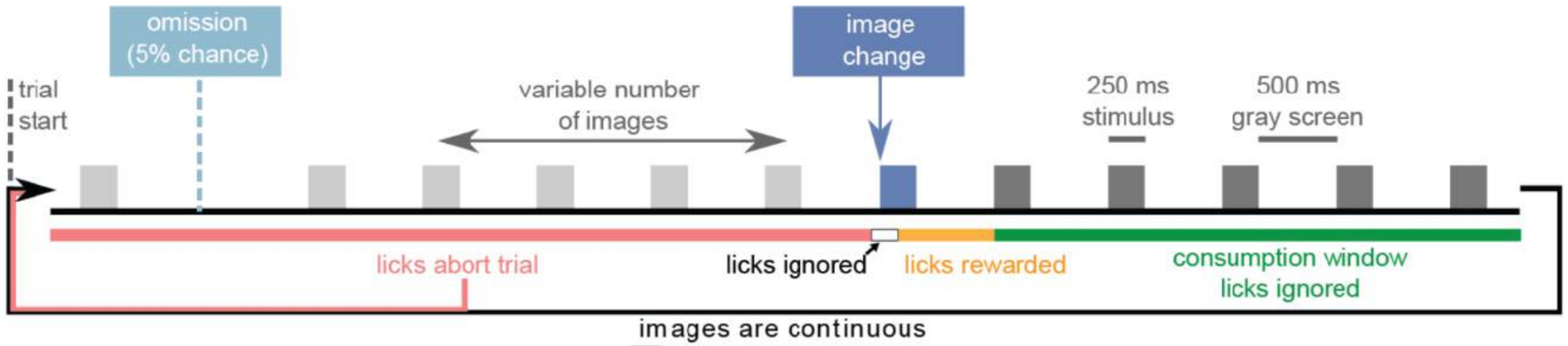


Image change responses



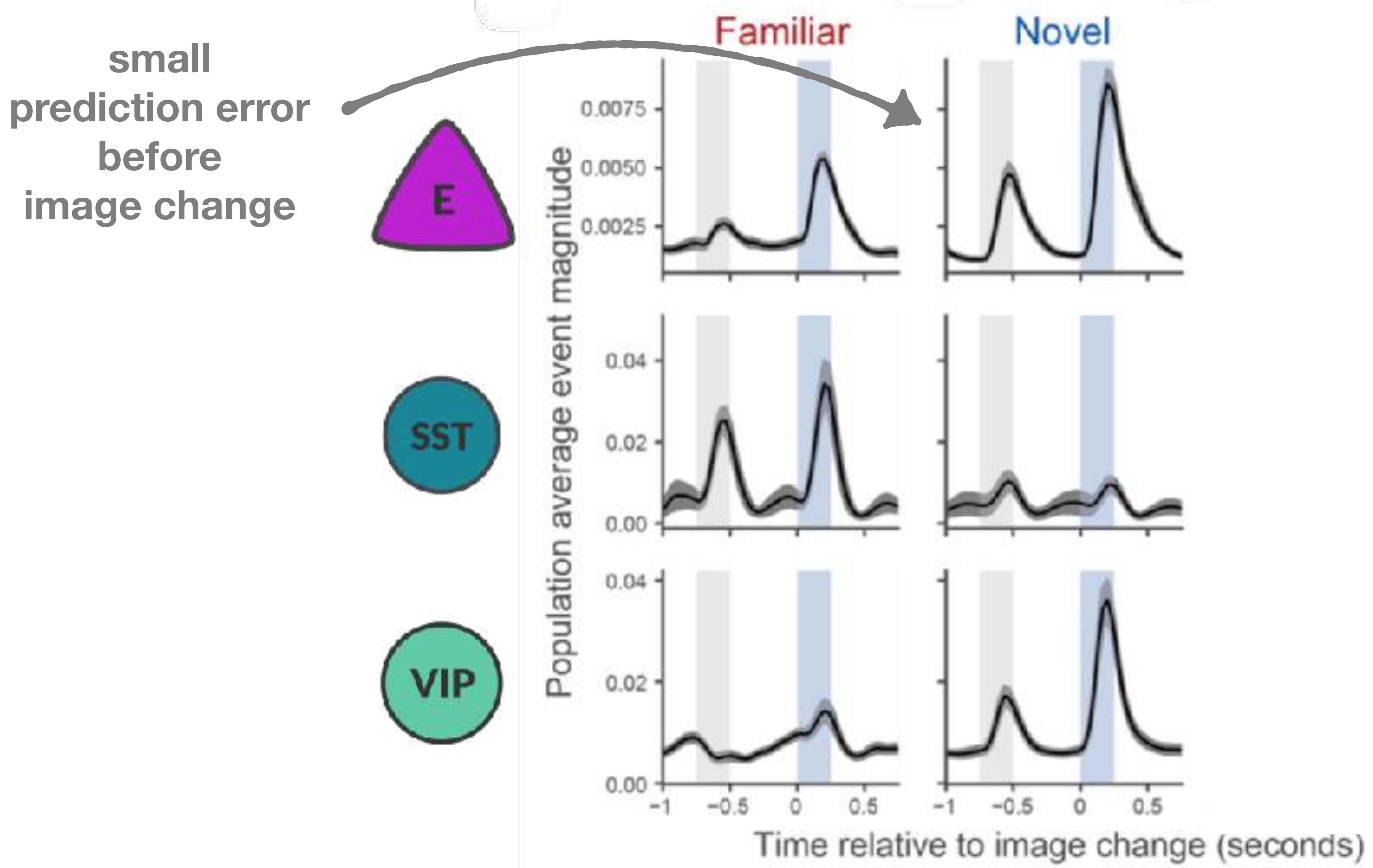


Image change responses

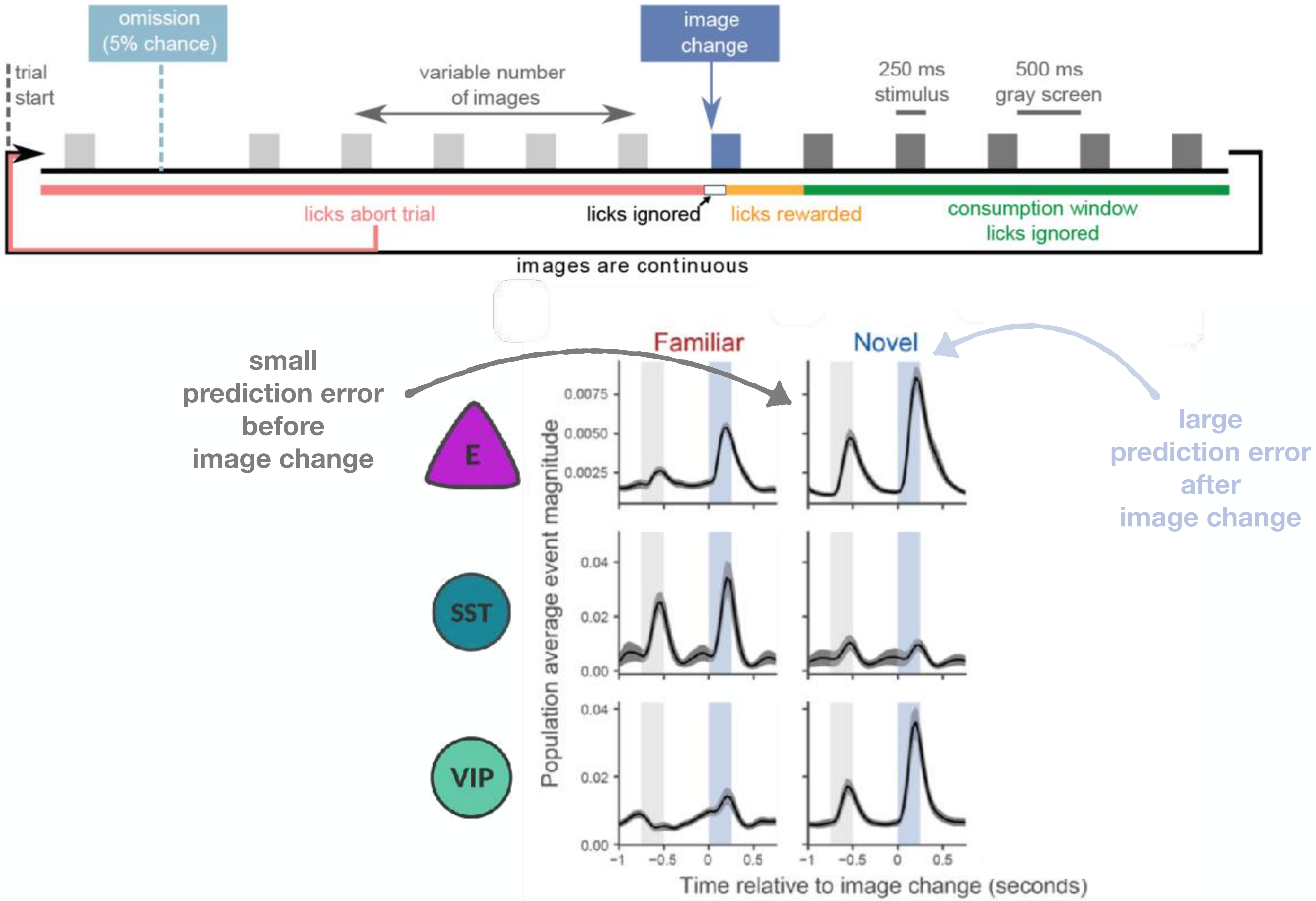
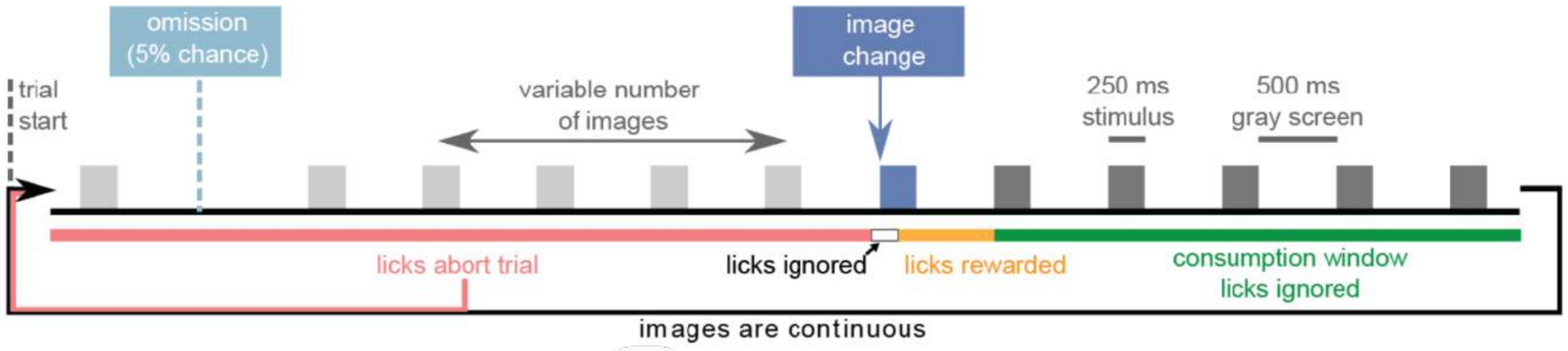


Image change responses



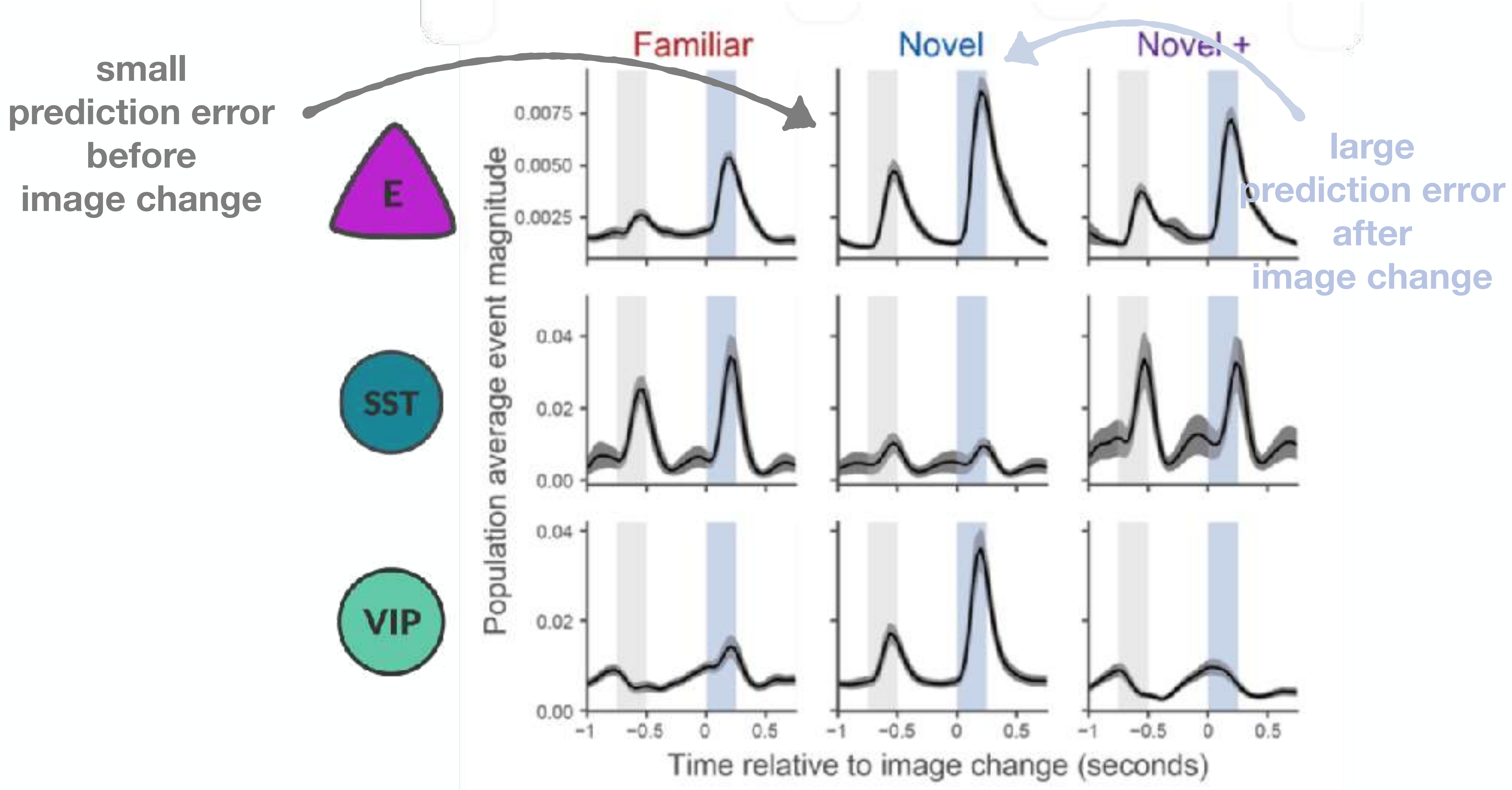


Image change responses

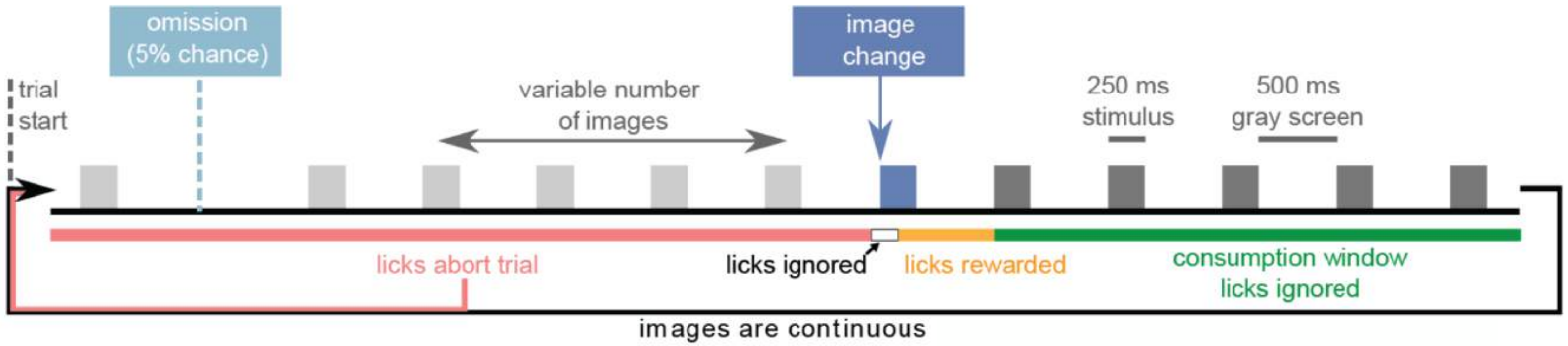


Image change responses

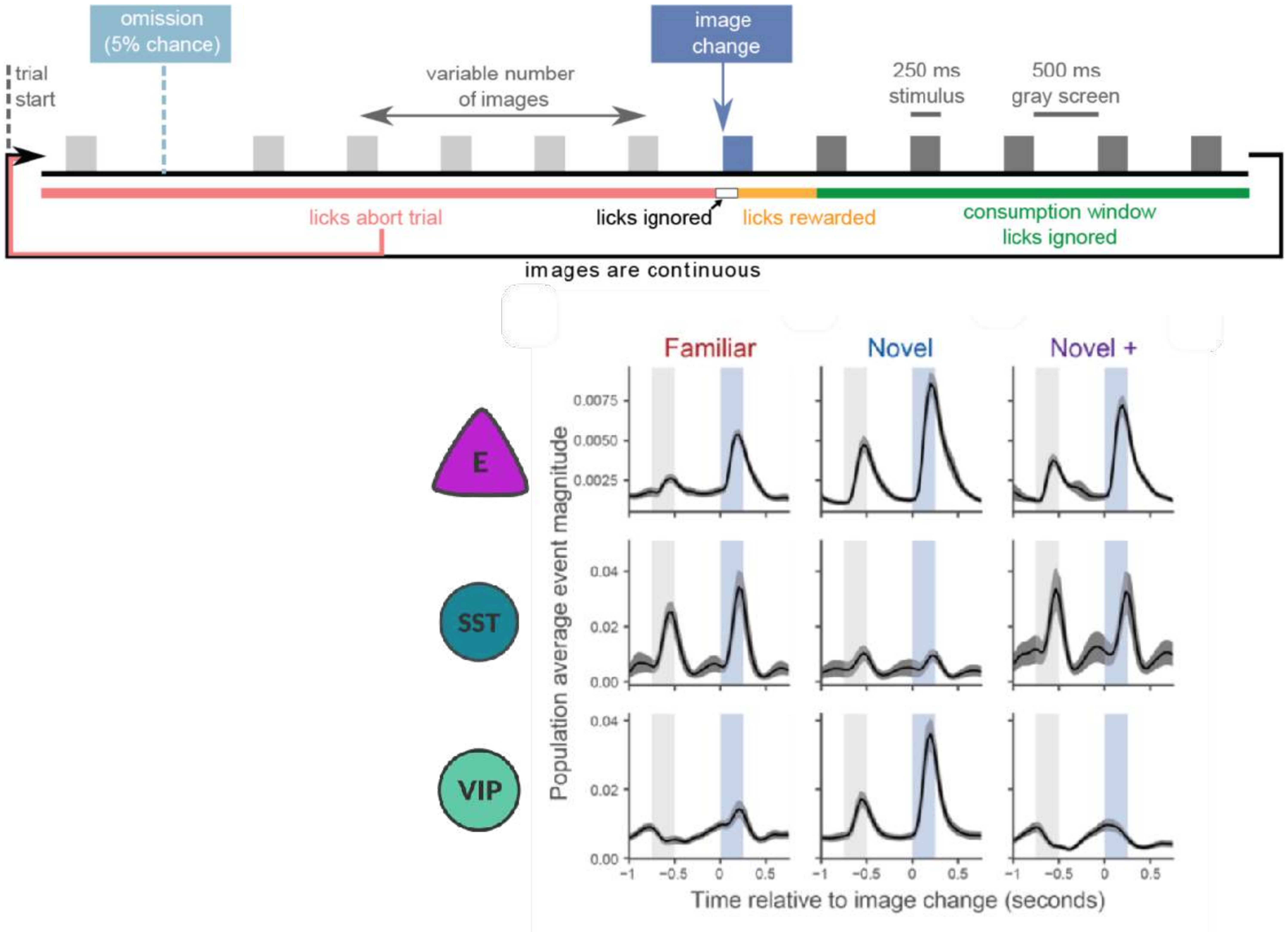
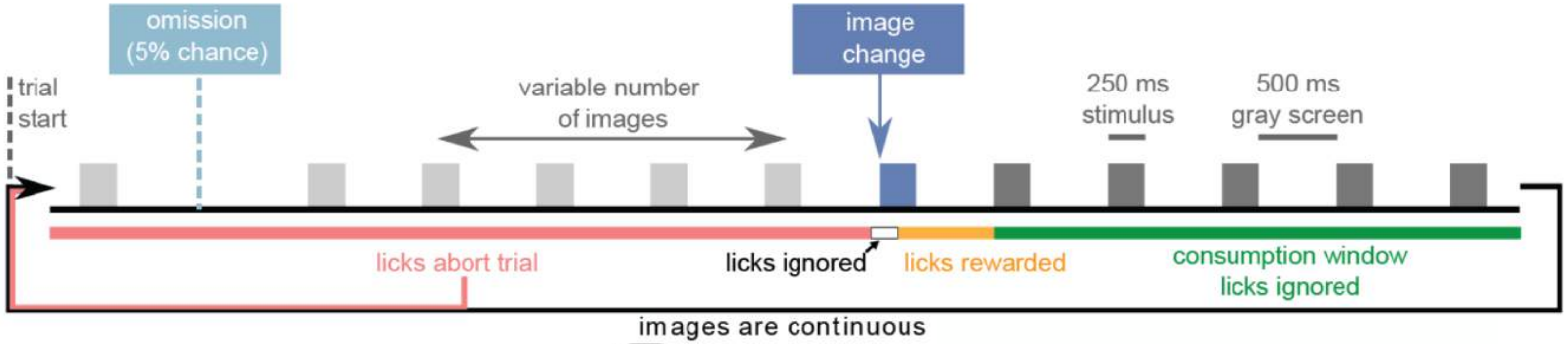


Image change responses



small
prediction error
before
image change

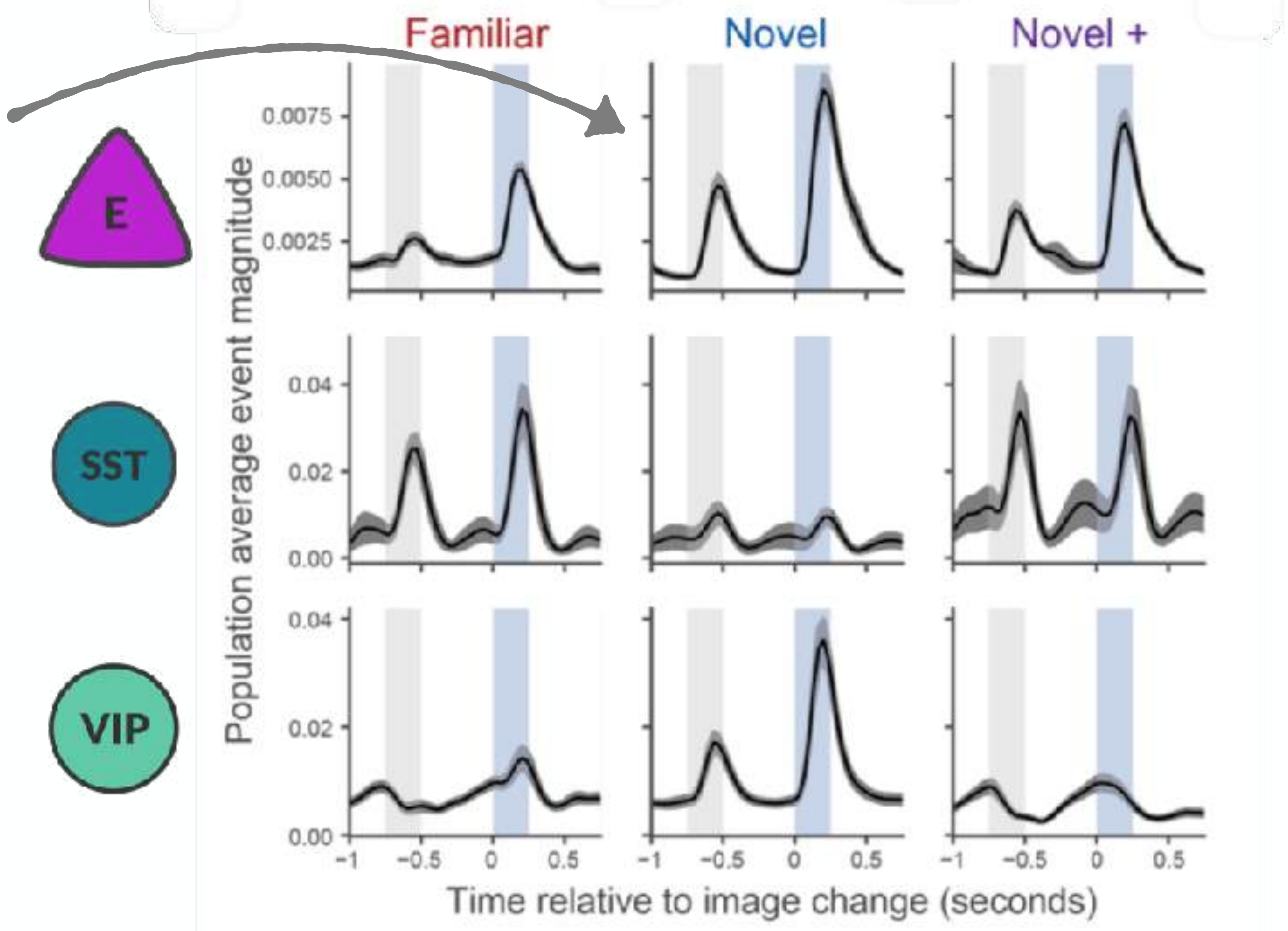


Image change responses

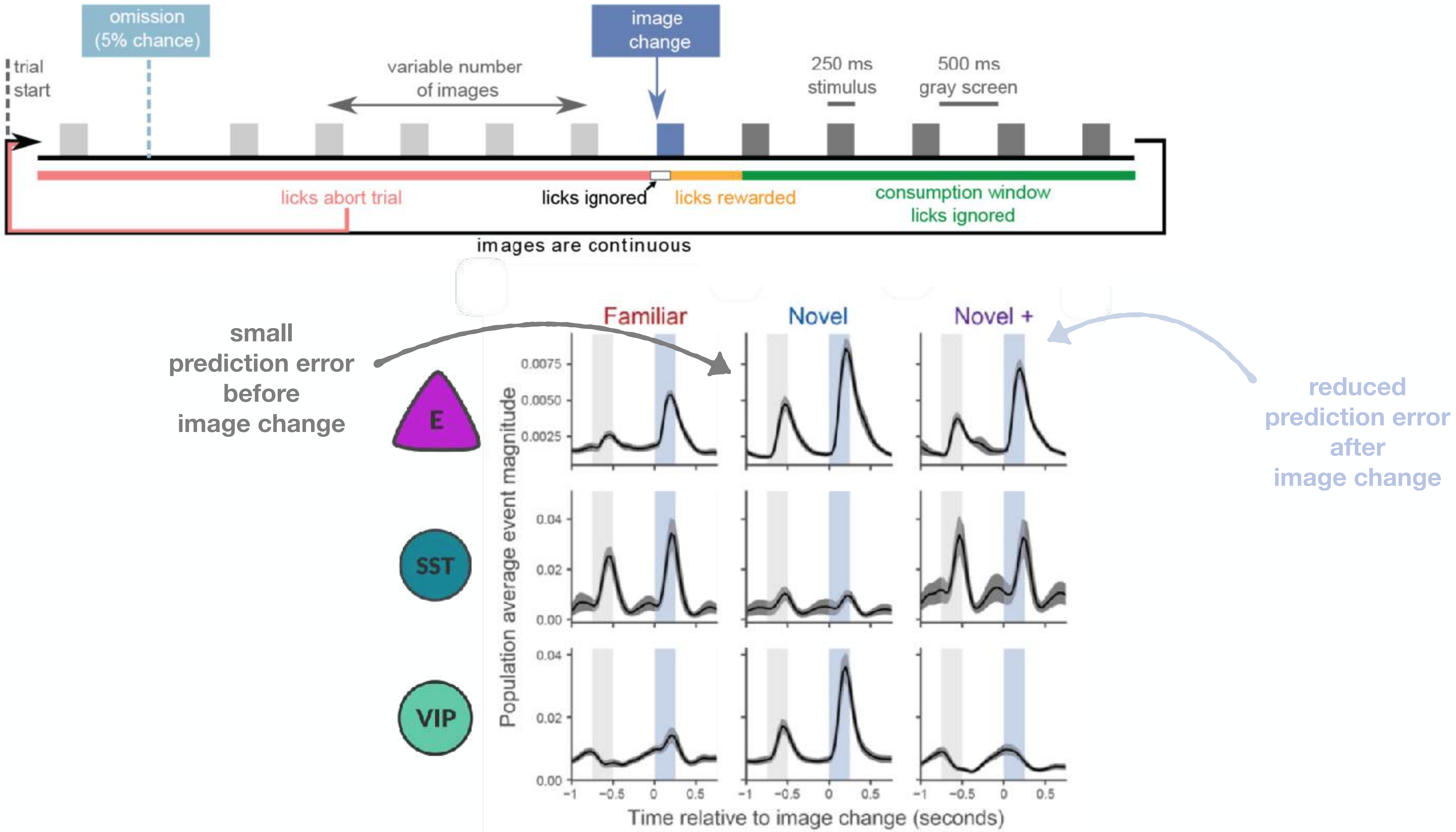


Image omission responses

omission would correspond
to a negative prediction error

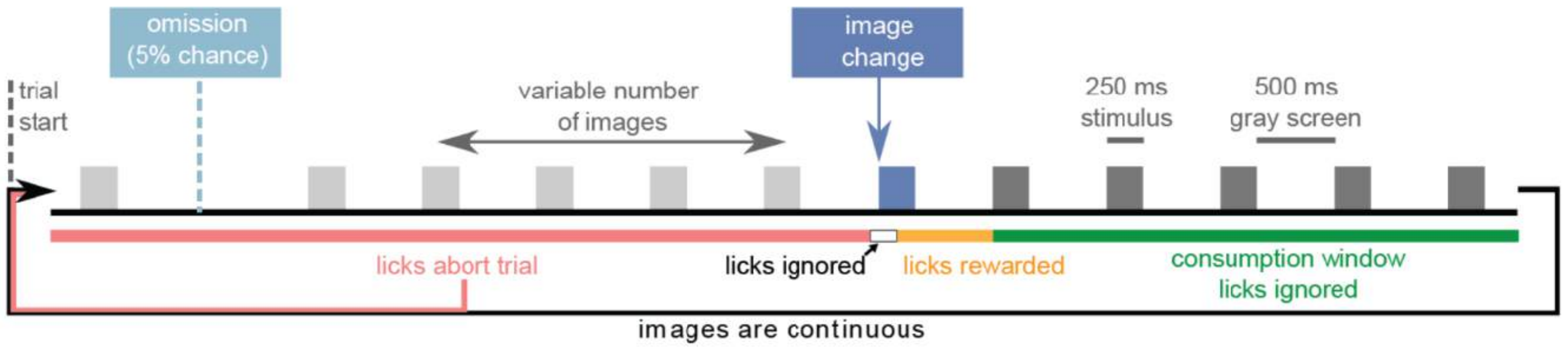
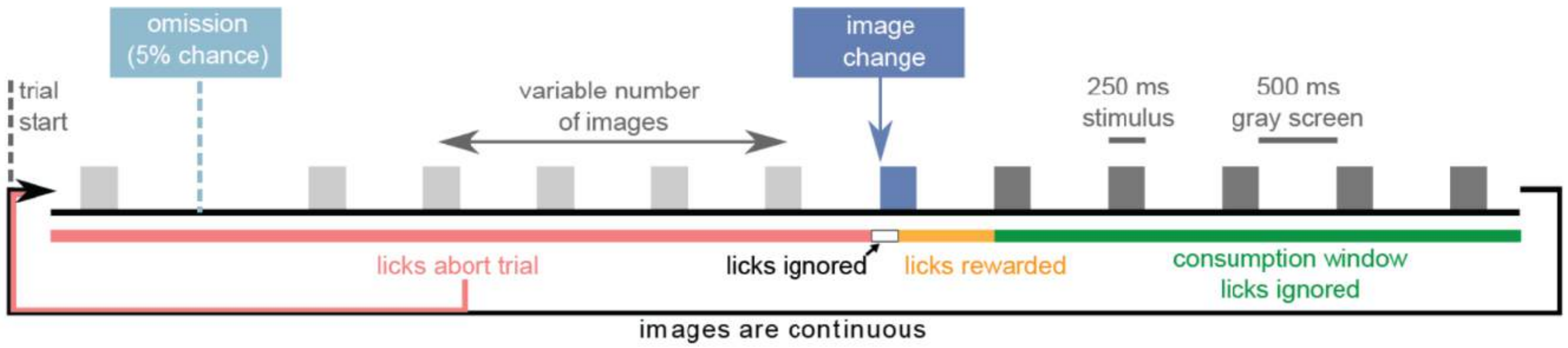


Image omission responses

omission would correspond
to a negative prediction error



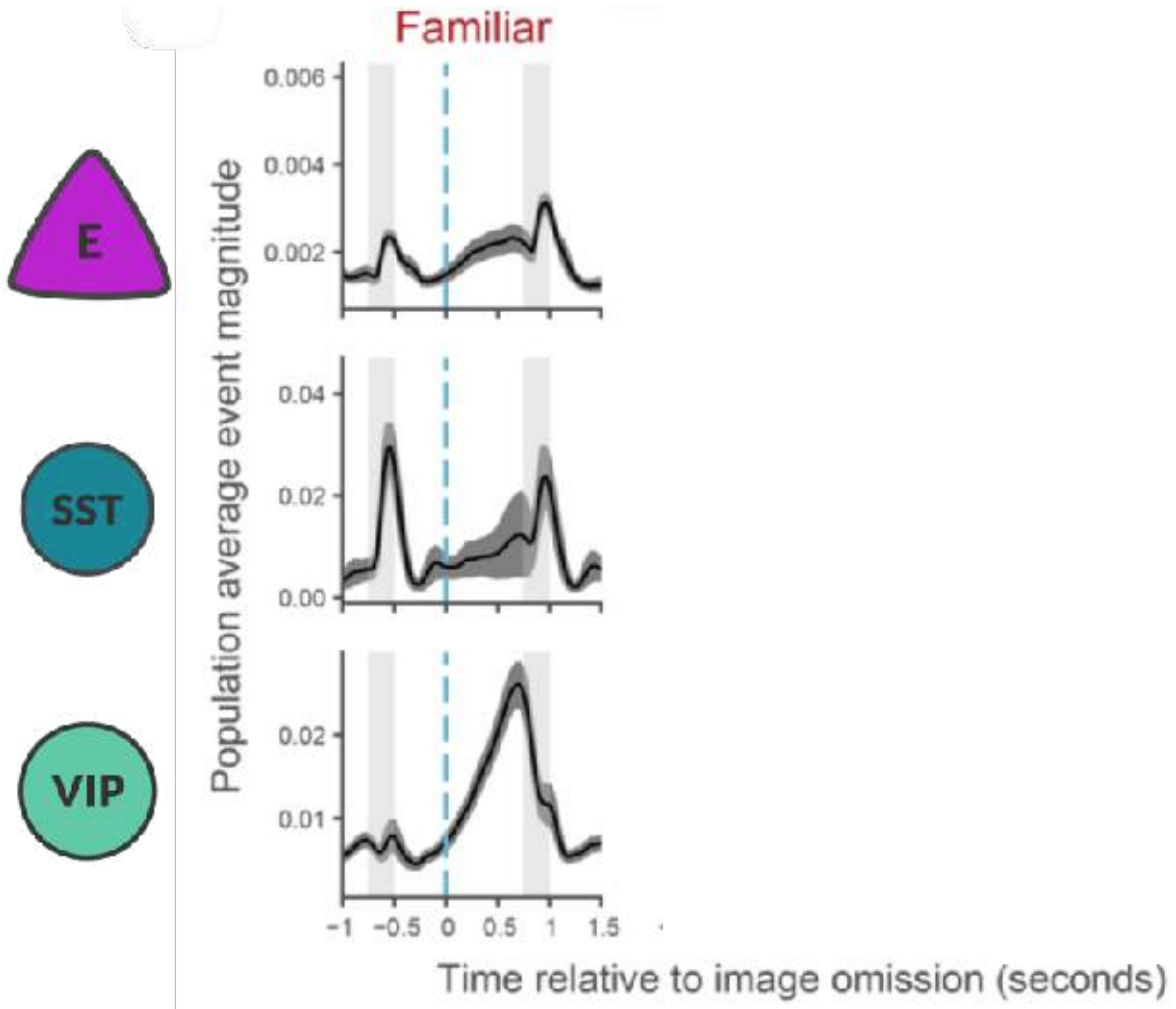
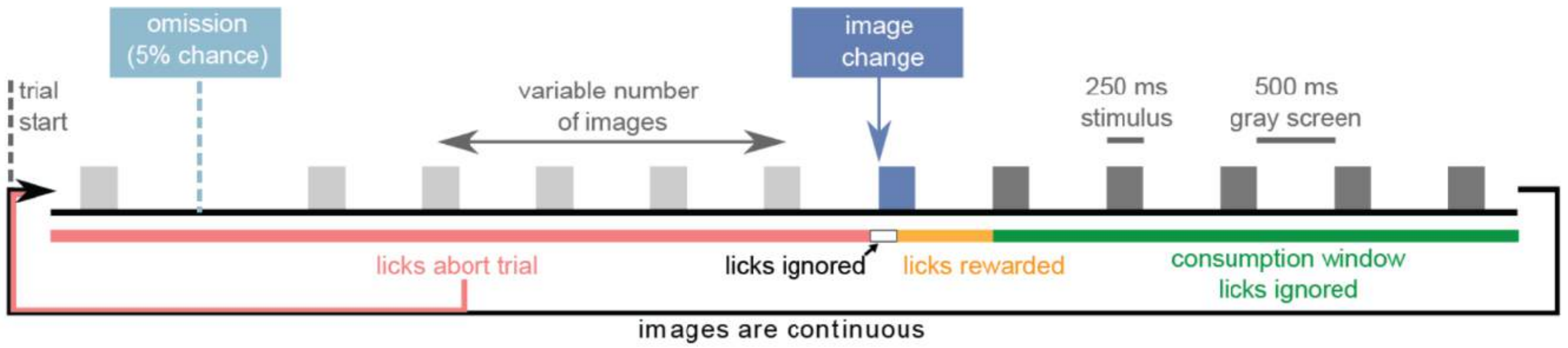


Image omission responses

omission would correspond
to a negative prediction error



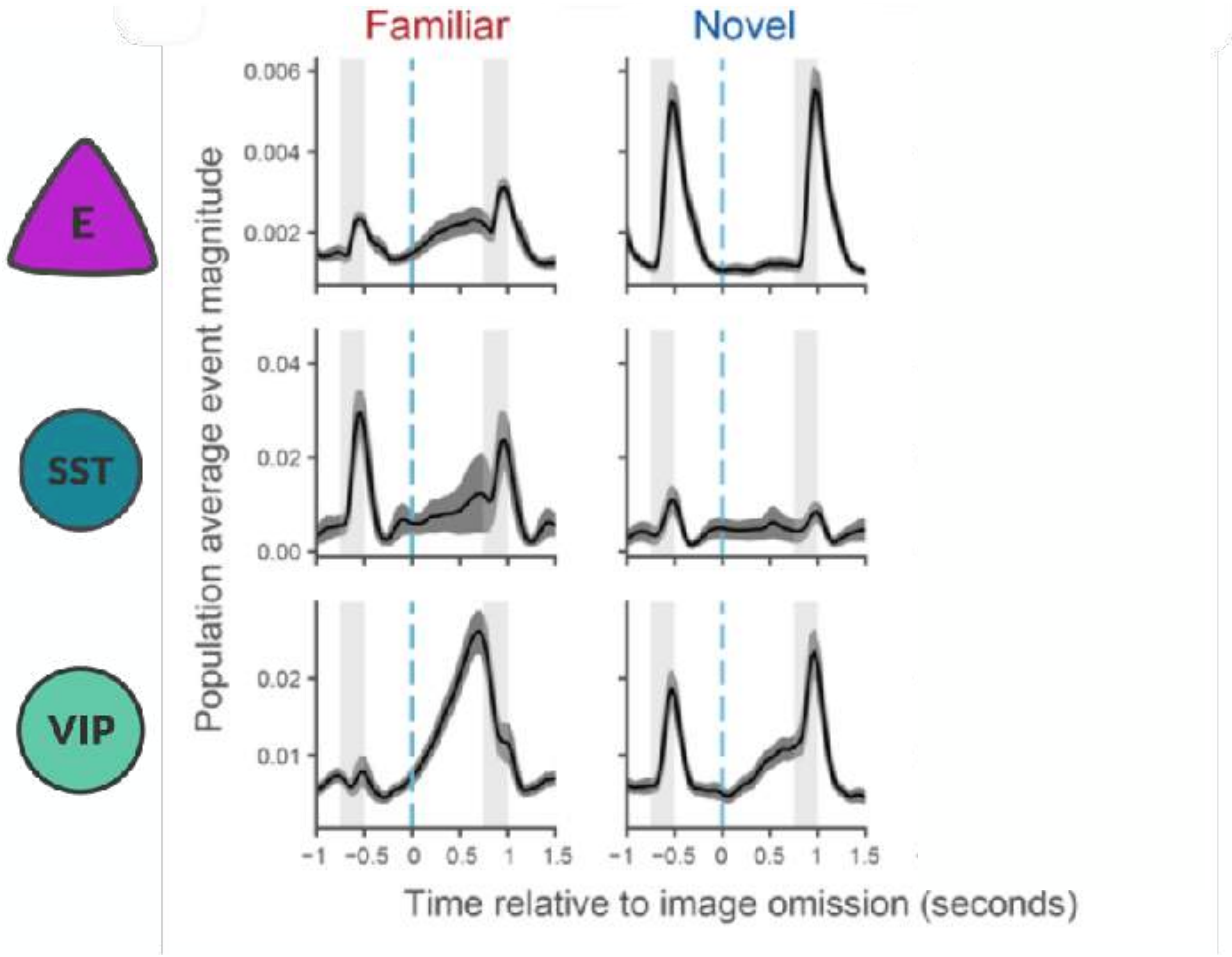
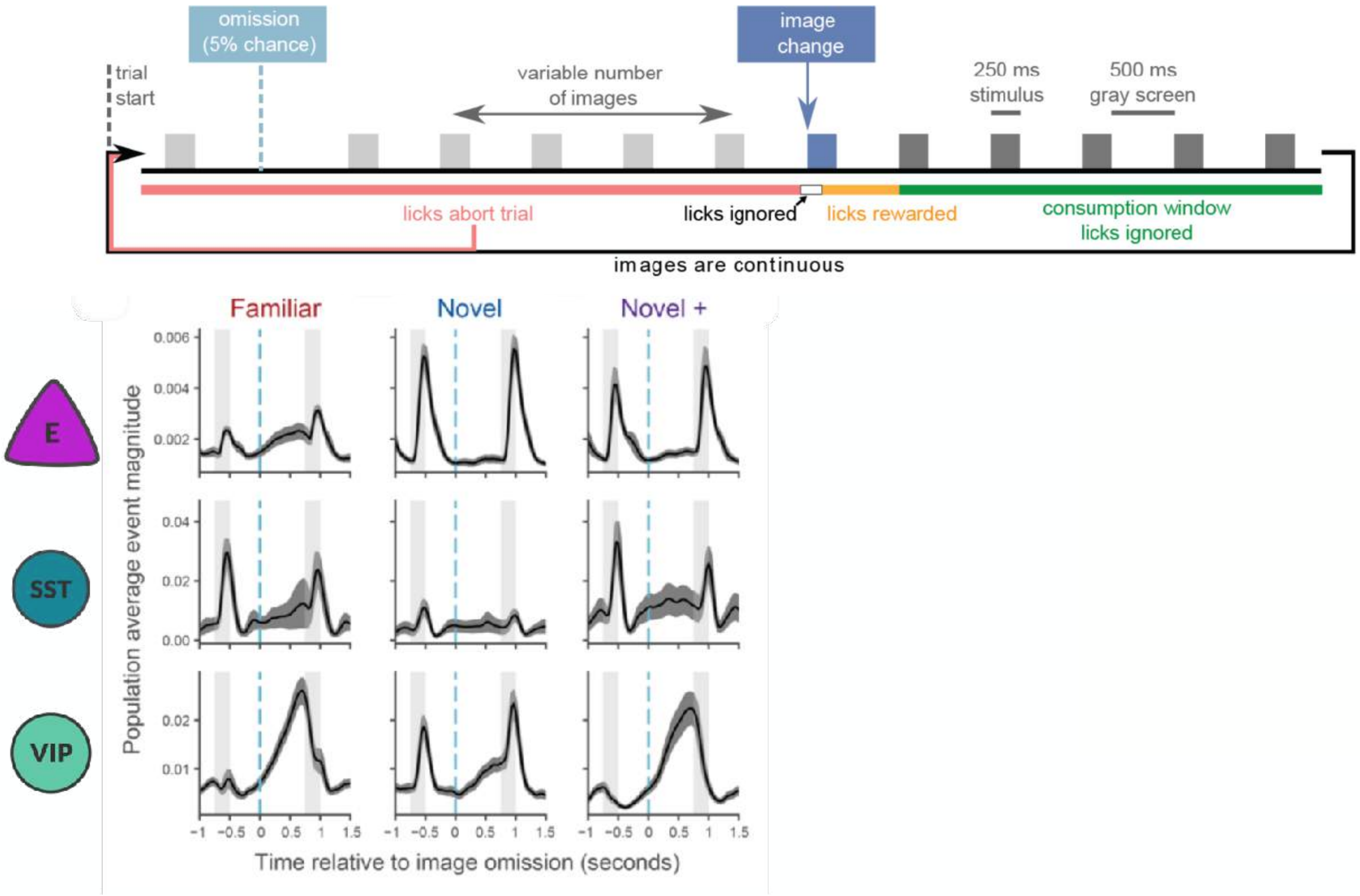


Image omission responses

omission would correspond
to a negative prediction error



Expectation violations

image change

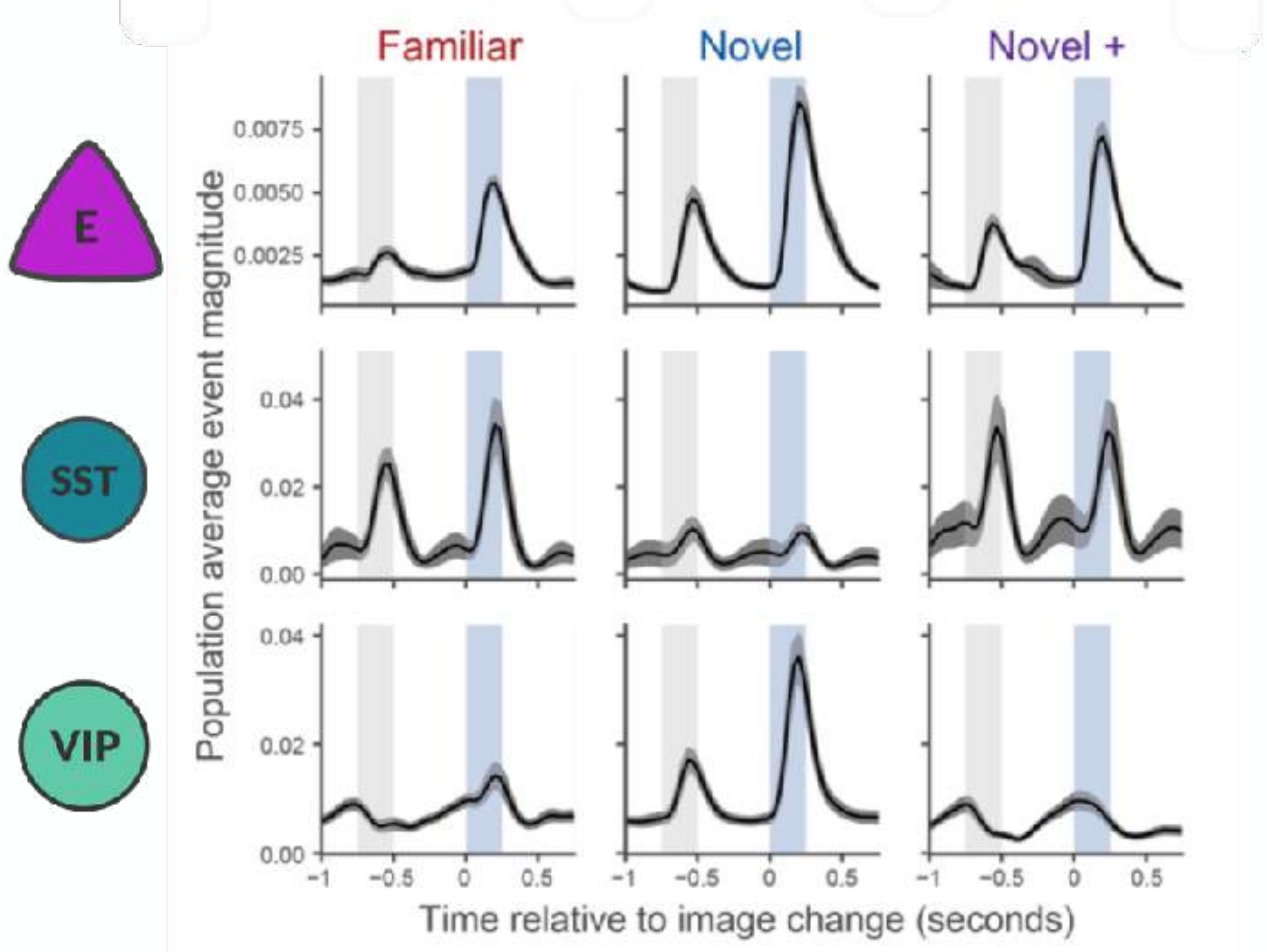
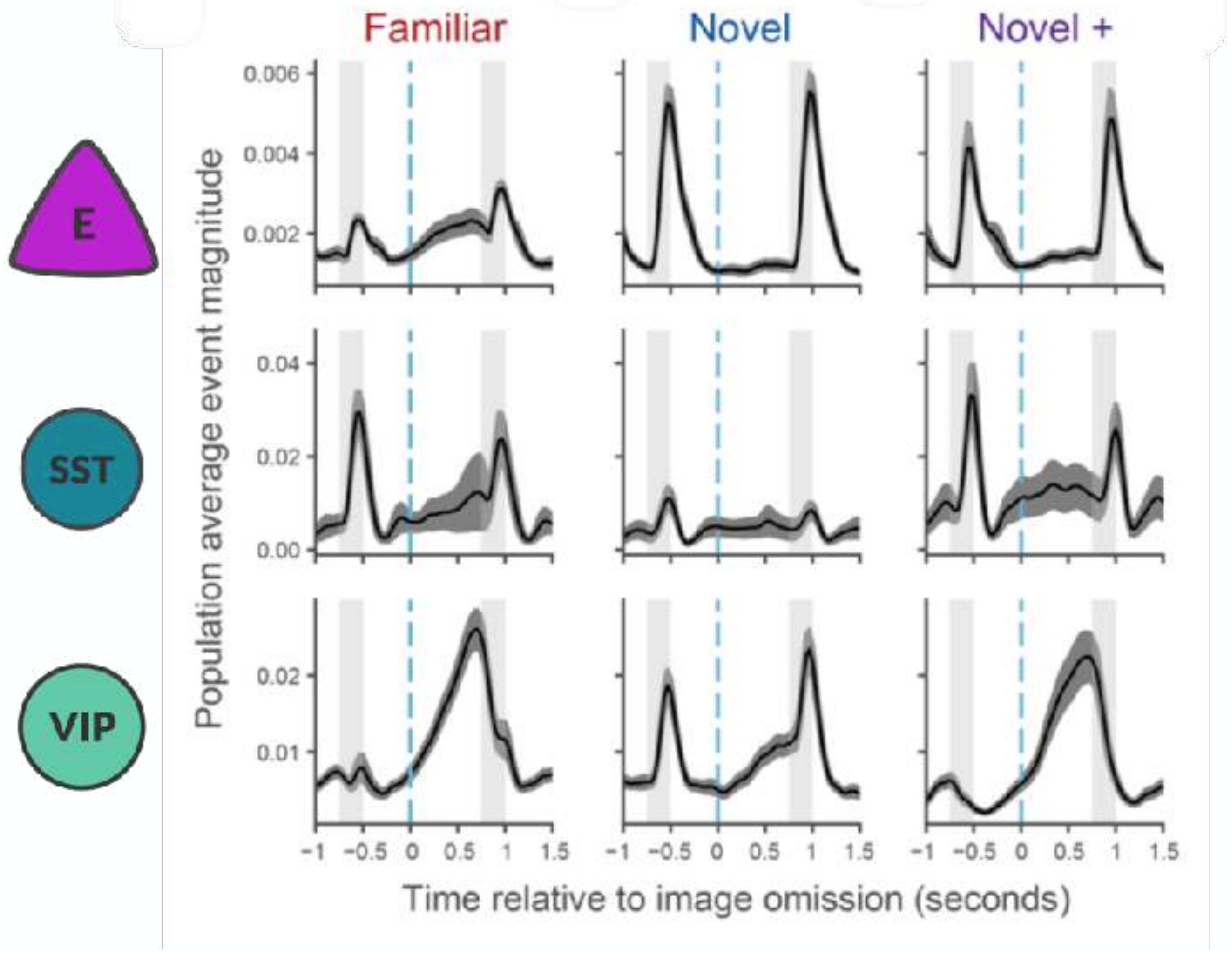


image omission



Expectation violations

Image identity

image change

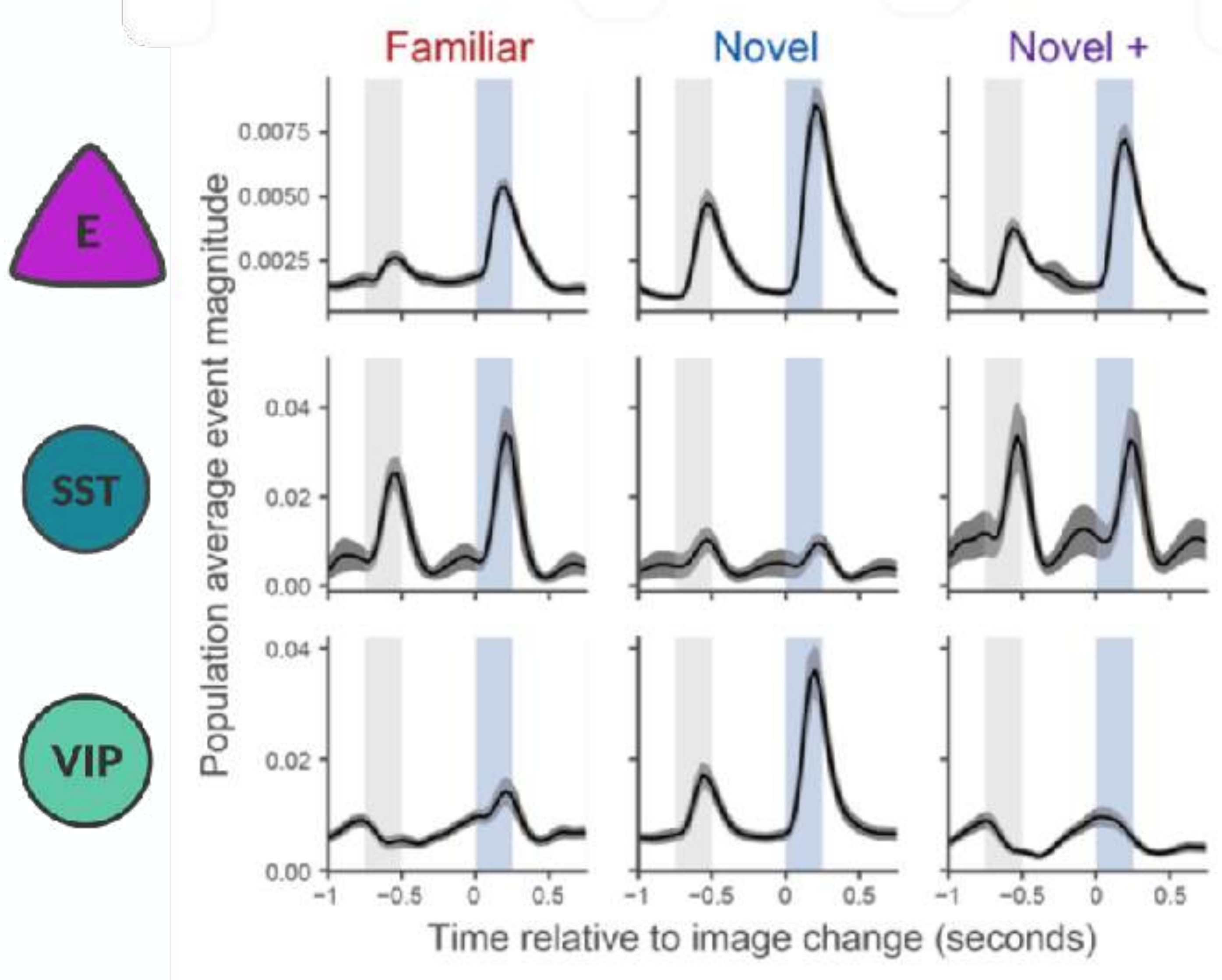
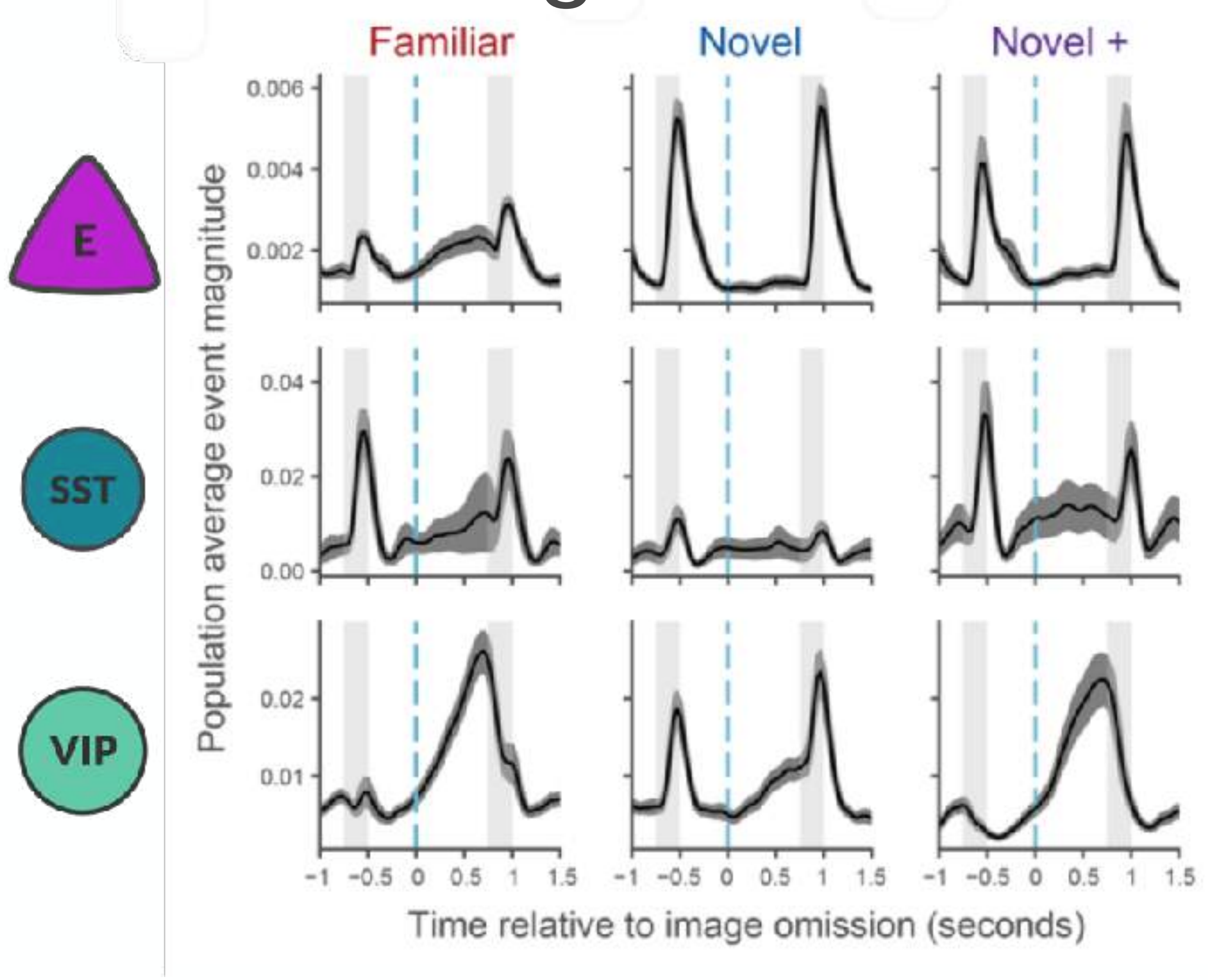


image omission



Expectation violations

- **Contextual** novelty *(familiar change)*

→ **Image identity**

image change

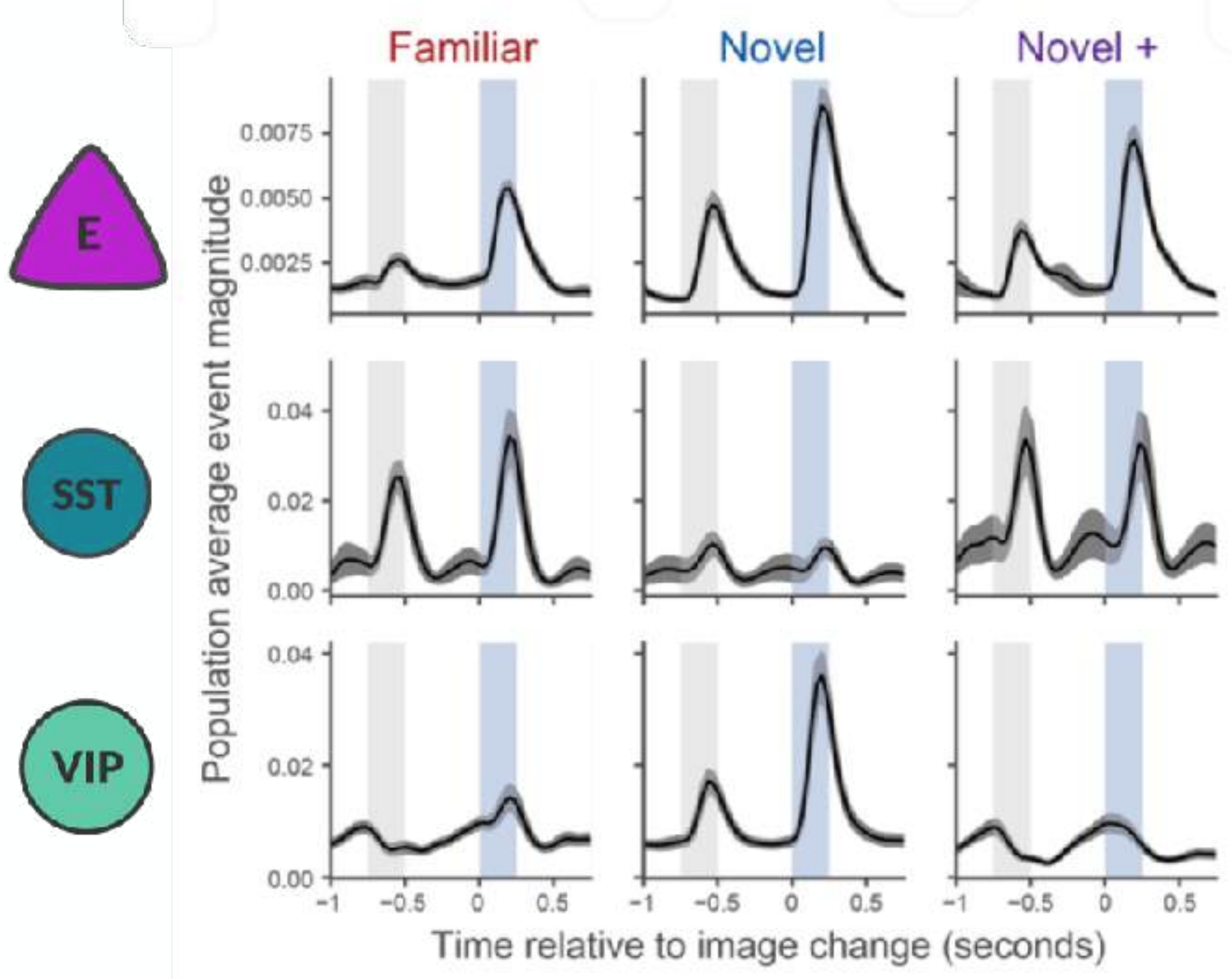
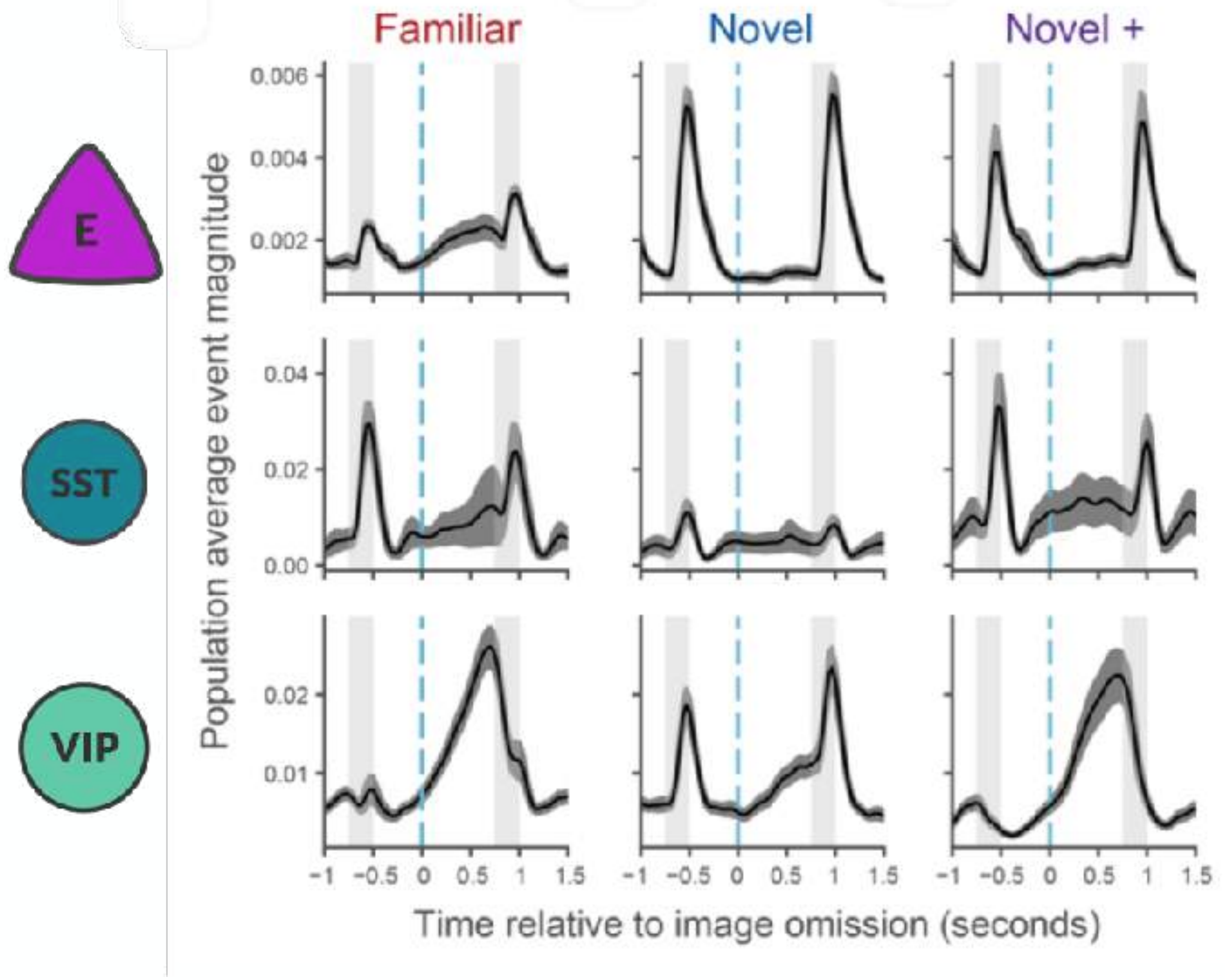


image omission



Expectation violations

- **Contextual** novelty *(familiar change)*
- **Absolute** novelty *(novel change)* → **Image identity**

image change

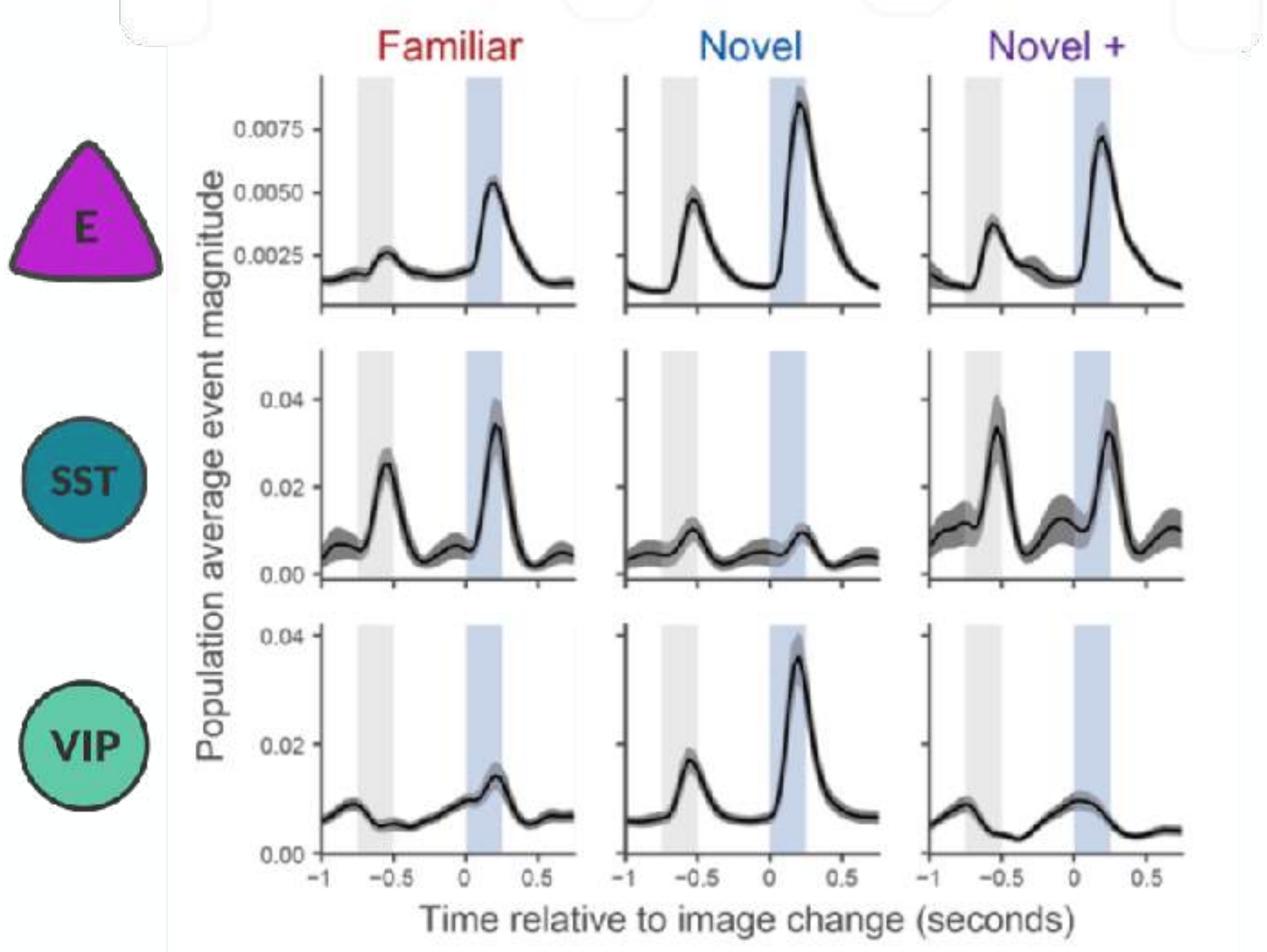
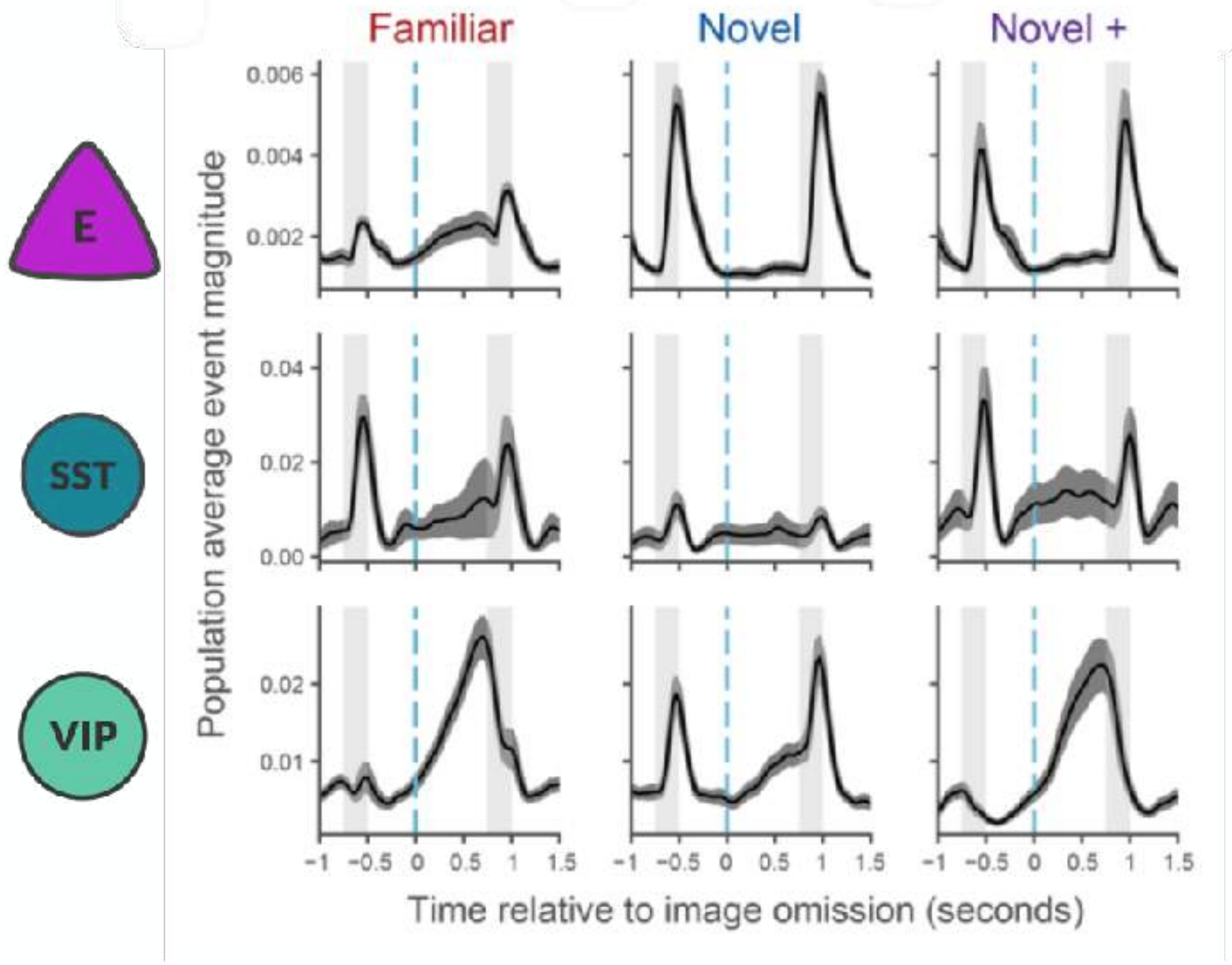


image omission

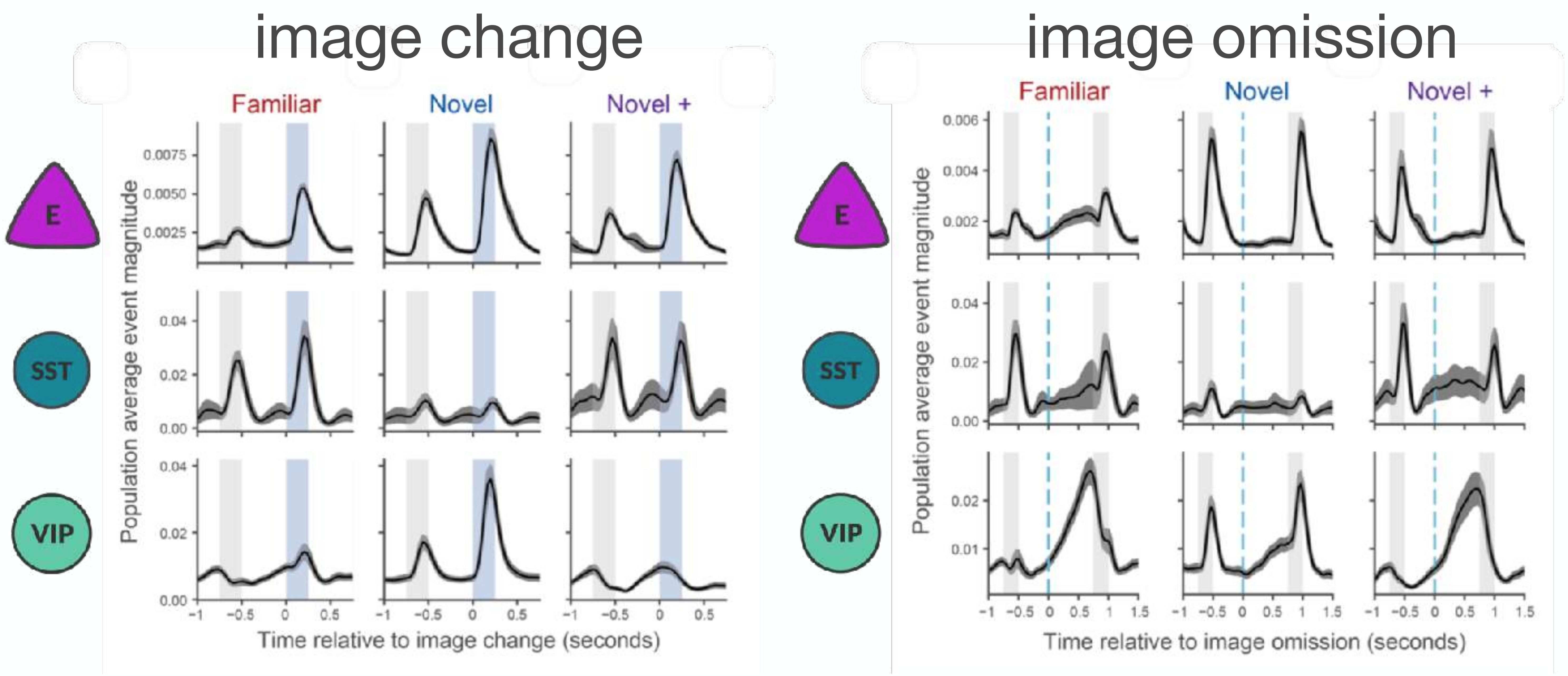


Expectation violations

- **Contextual** novelty *(familiar change)*
- **Absolute** novelty *(novel change)*

→ **Image identity**

Event timing



Expectation violations

- **Contextual** novelty *(familiar change)*
- **Absolute** novelty *(novel change)*

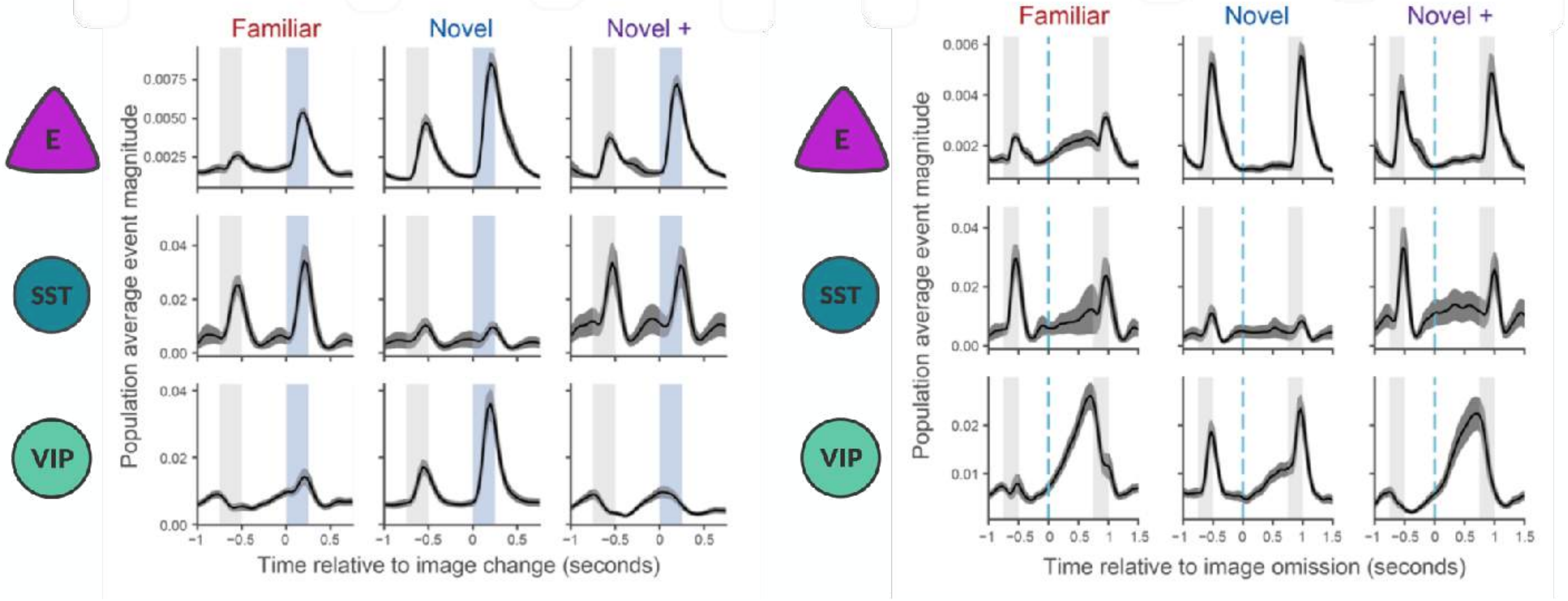
→ **Image identity**

- **Stimulus omissions**

→ **Event timing**

image change

image omission

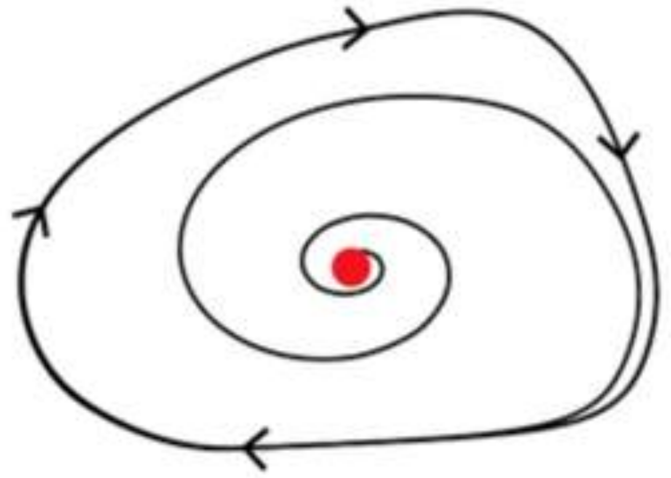


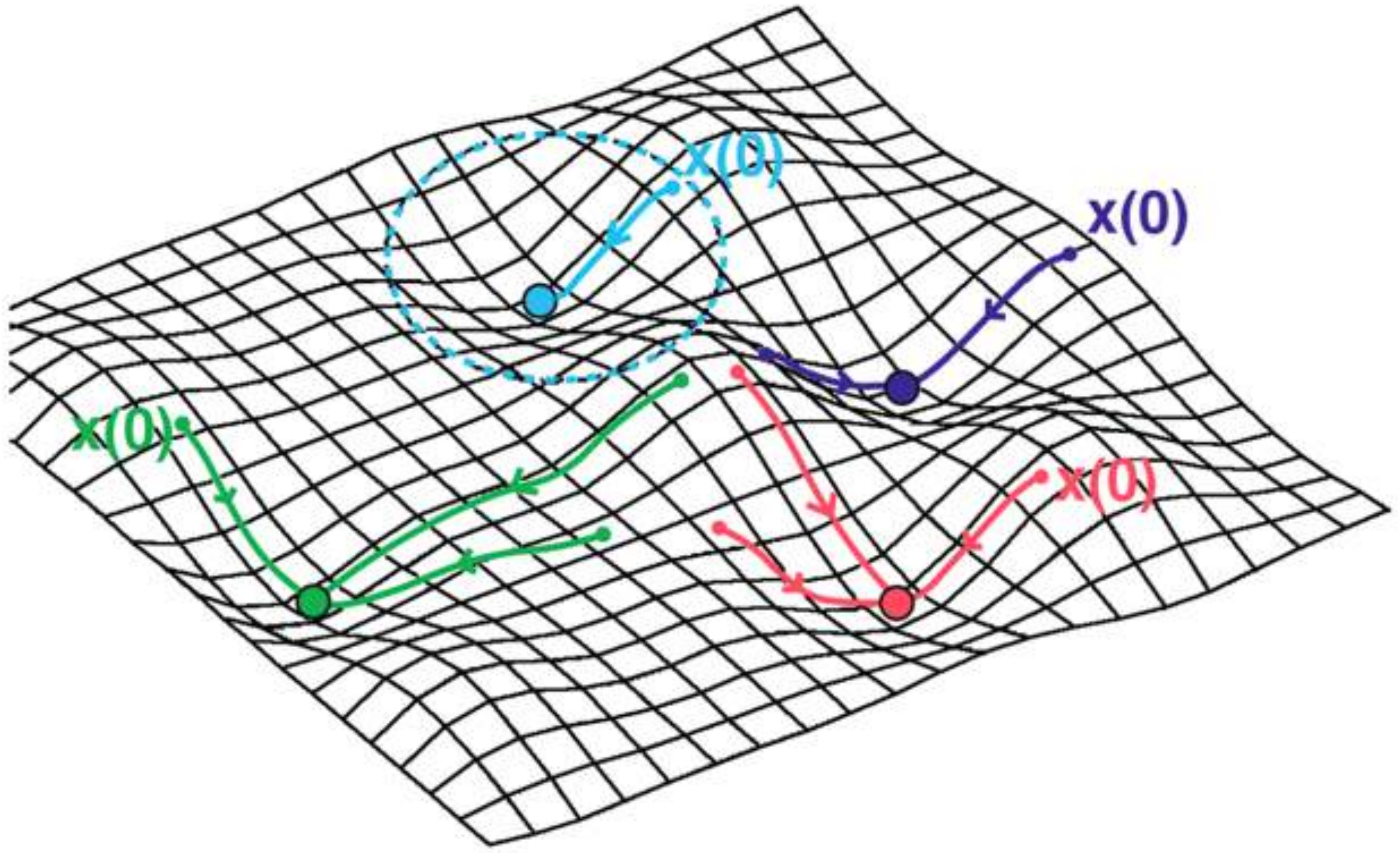
How are temporal expectations related to stimulus identity expectations?

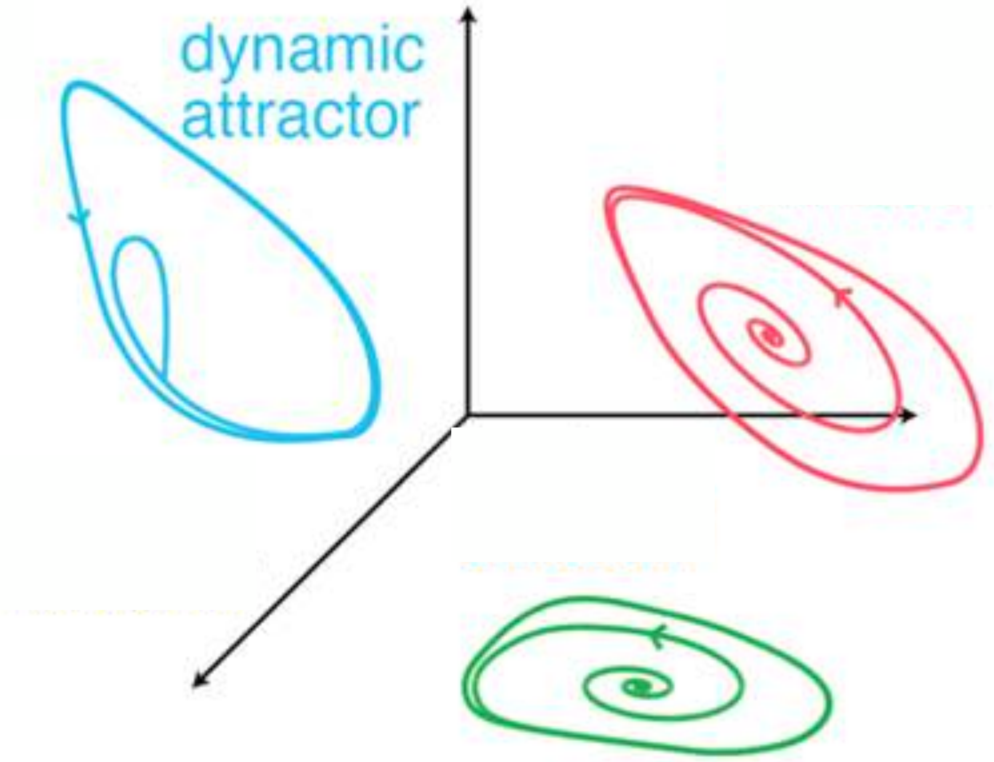
2 potential configurations

- subject learns independently frequency of image presentation from image identity
- temporal expectation independent of stimulus identity
- internal model comprises an independent associative memory (Hopfield network) an dynamic attractor (limit cycle)

- subject learns frequency of presentation together with image identity
- for each image the animal has an internal model with the dynamics of dynamic attractor







omission responses make the second option more probable

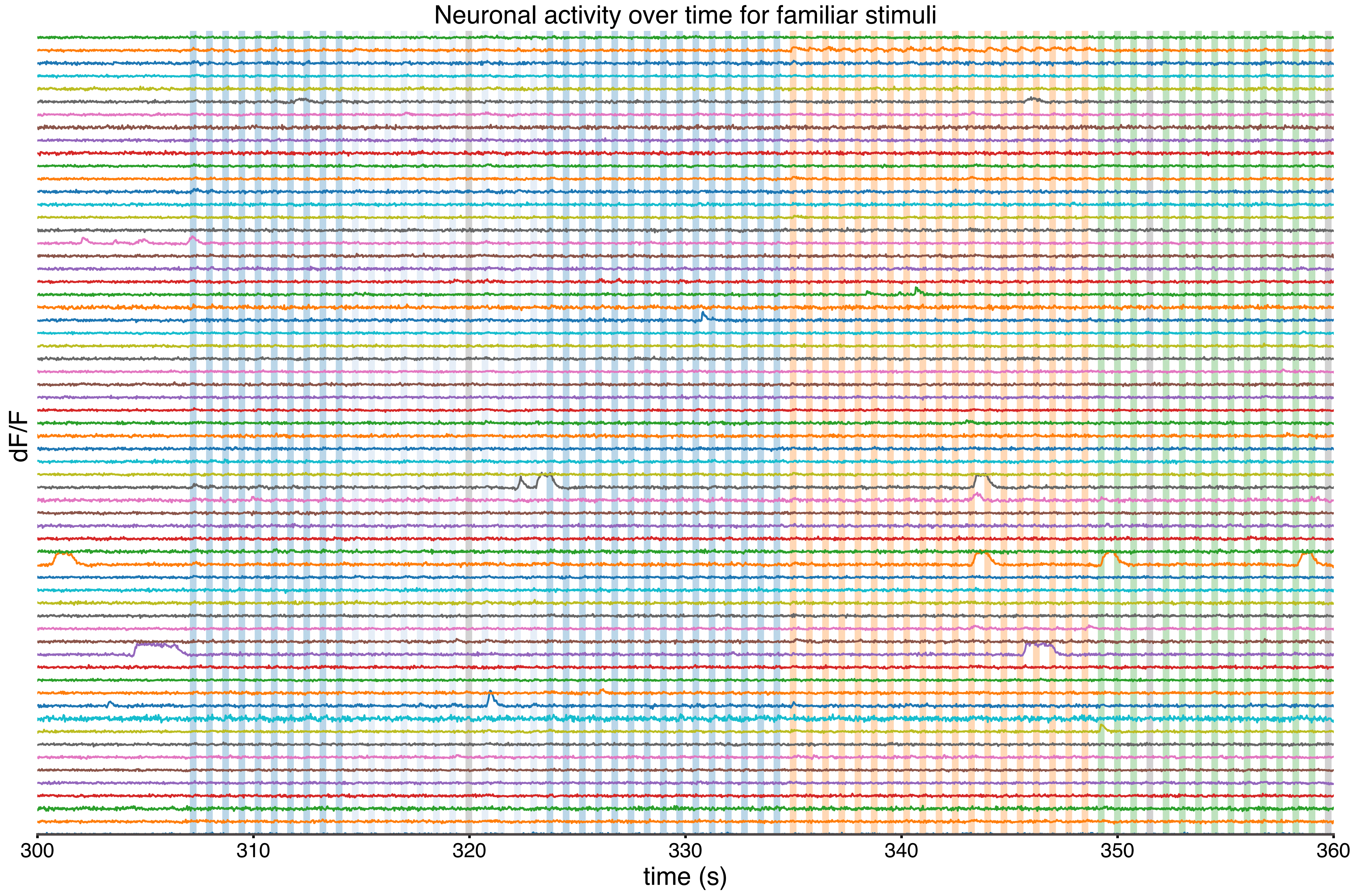
Questions

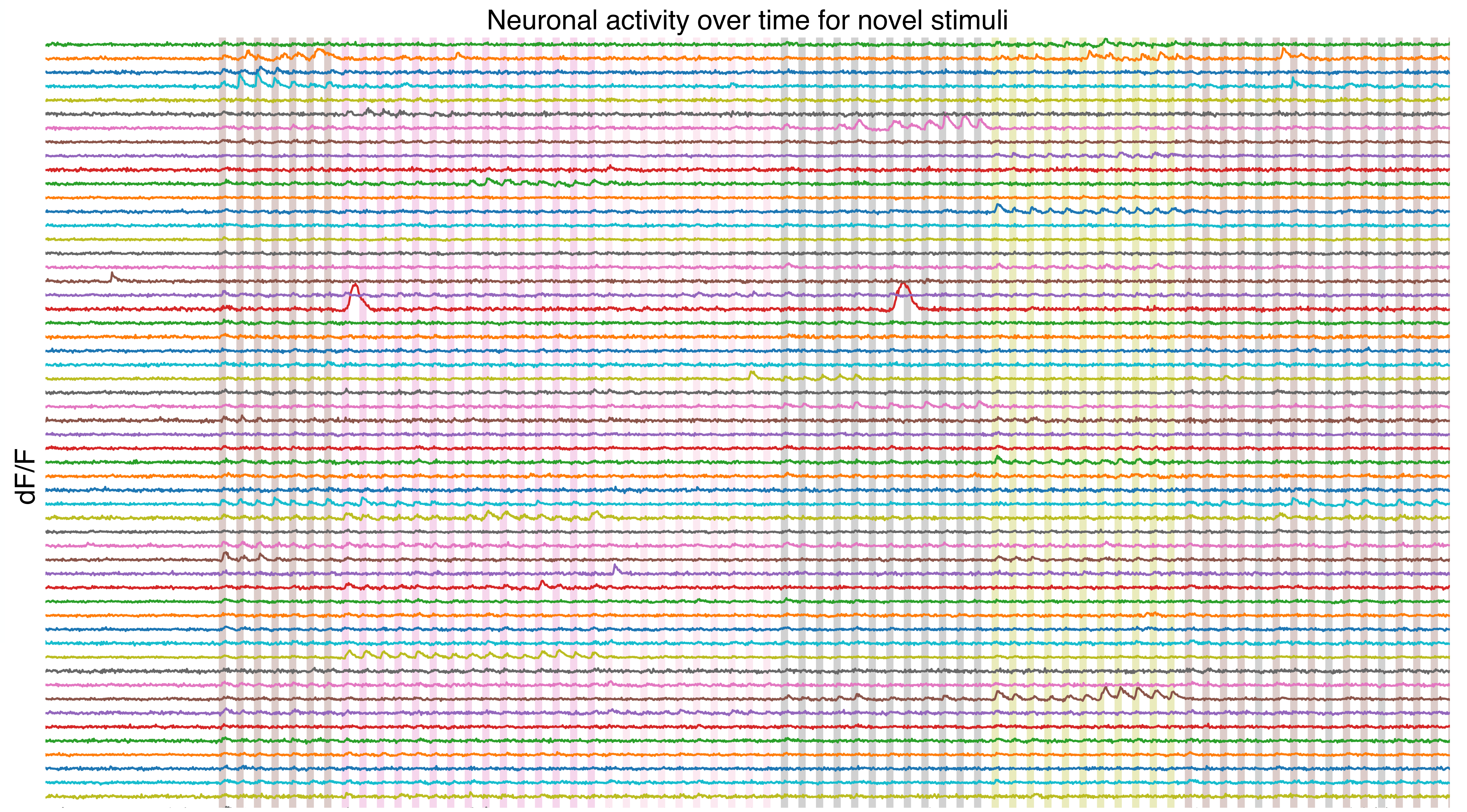
- How do **different cell-types** participate/respond to the **emergence of prediction errors**?
- What drives these responses?
- How are expectations formed in the first place?
- What is the **content of the feedback** from Higher Order Areas?
    - Furutachi et al 2024: how early visual processing reflects the sensory aspect of visual information, rather than the deviation from the expected input.
    - bastos Hamm, Cell reports 2023, ACC top down input to layer 1 encodes the expected stimulus, here the previous image

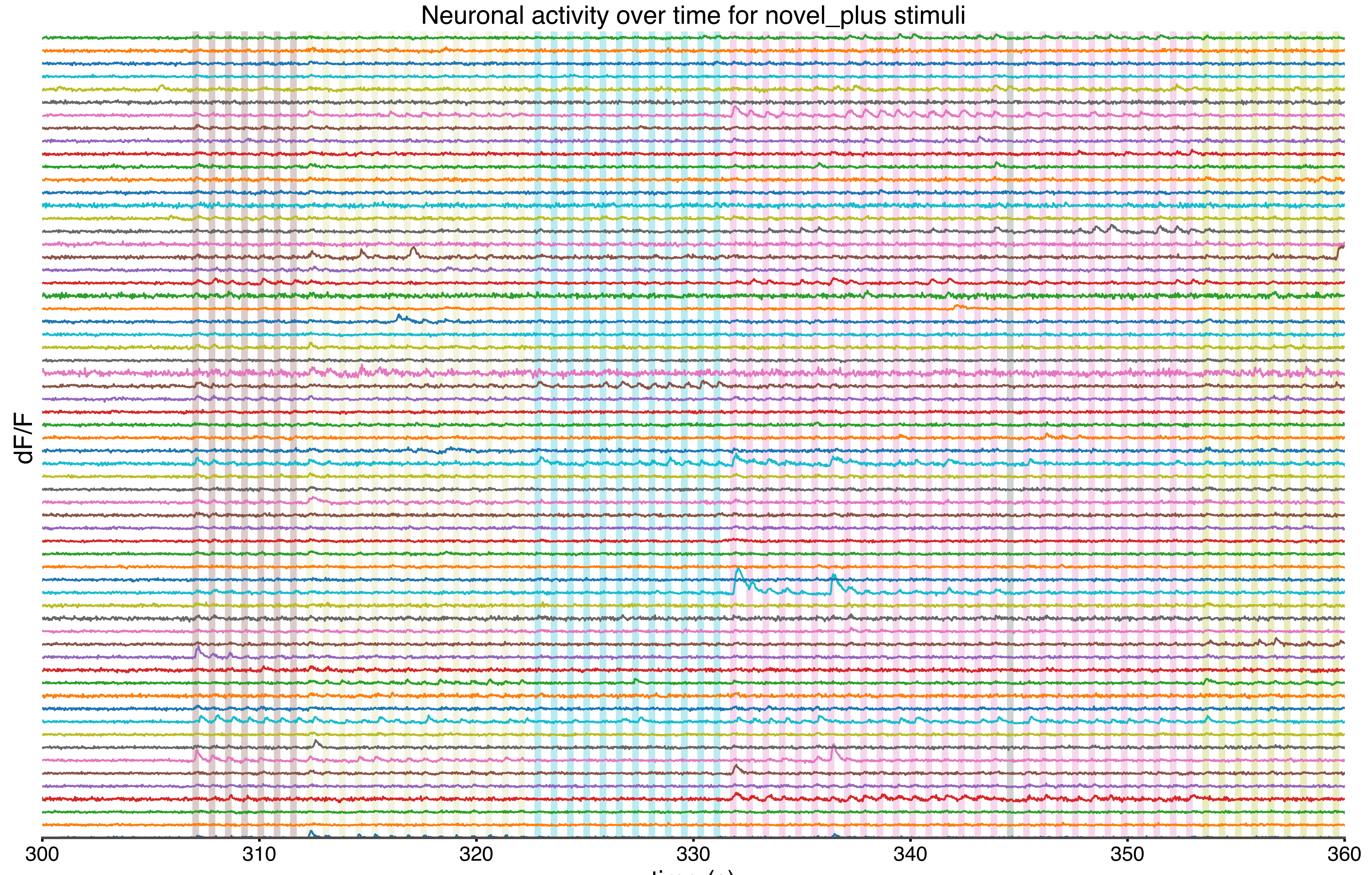
What is the nature of prediction errors in V1?

- Furutachi et al 2024: Early visual processing reflects the sensory aspect of visual information, rather than the deviation from the expected input
  - Responses of E neurons to unexpected events over-represented these events - amplified stimulus specific responses
- Najafi et al 2024: VIPs function as non-specific error encoders, driving circuits to attend to context.
  VIPs encode a general surprise signal enriched with state-dependent information(state here is state of animal (arousal, locomotion etc), and not contextual state)

Calcium traces for E neurons

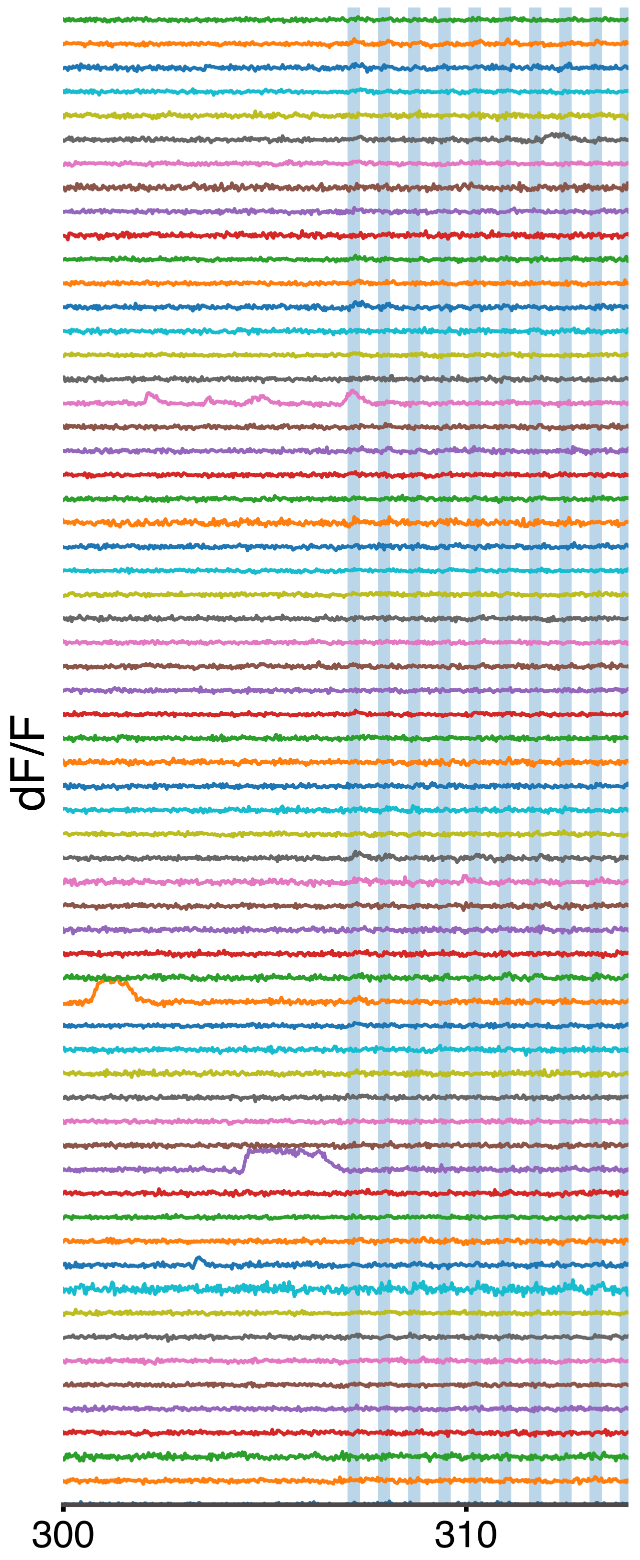




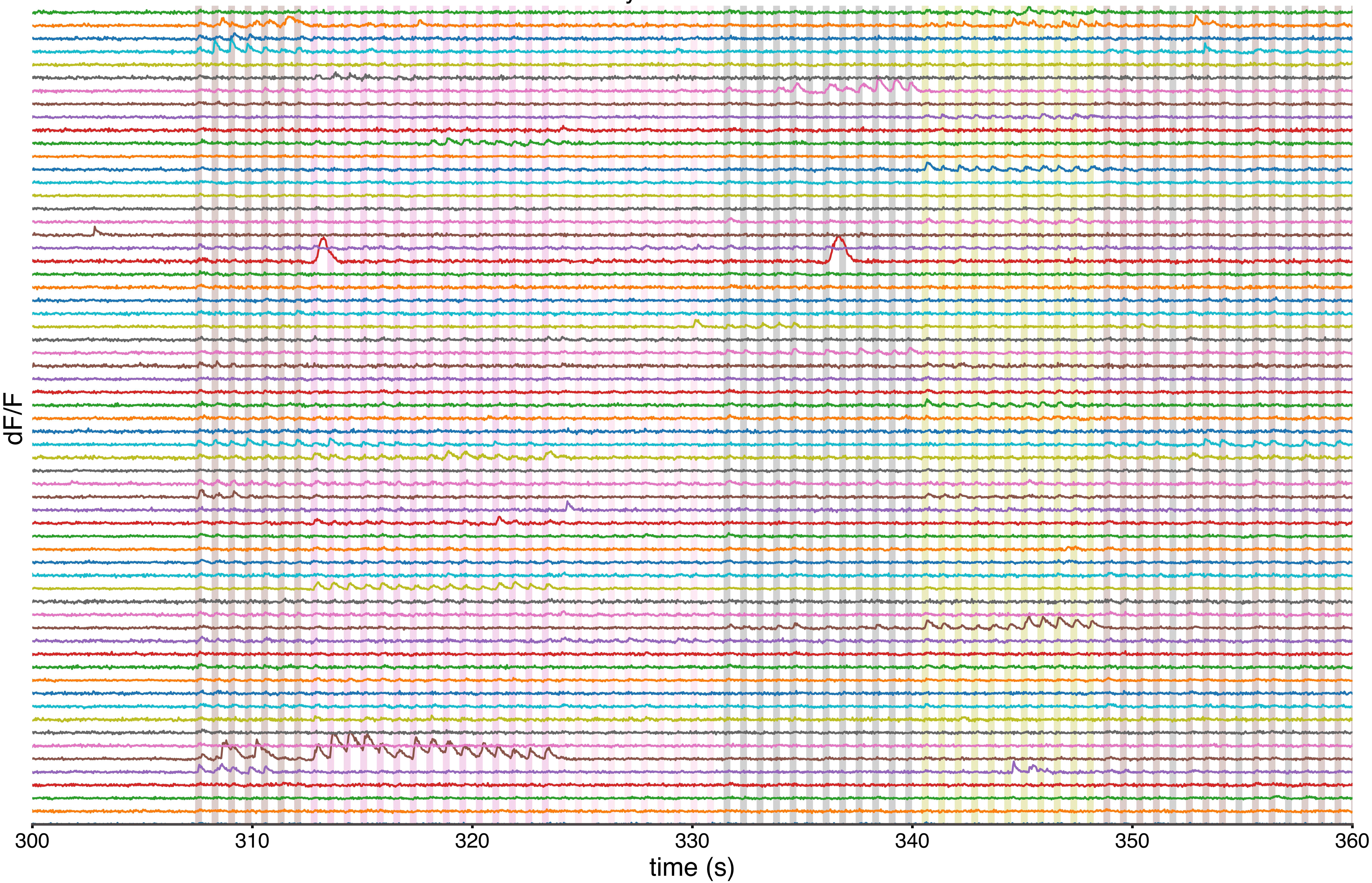


Neuronal activity over time for novel stimuli
Neuronal activity over time for familiar stimuli
dF/F
300
310
320
330
340
350
360
time (s)
l_plus stimuli
300
310
320
330
340
350
360

Calc
neu



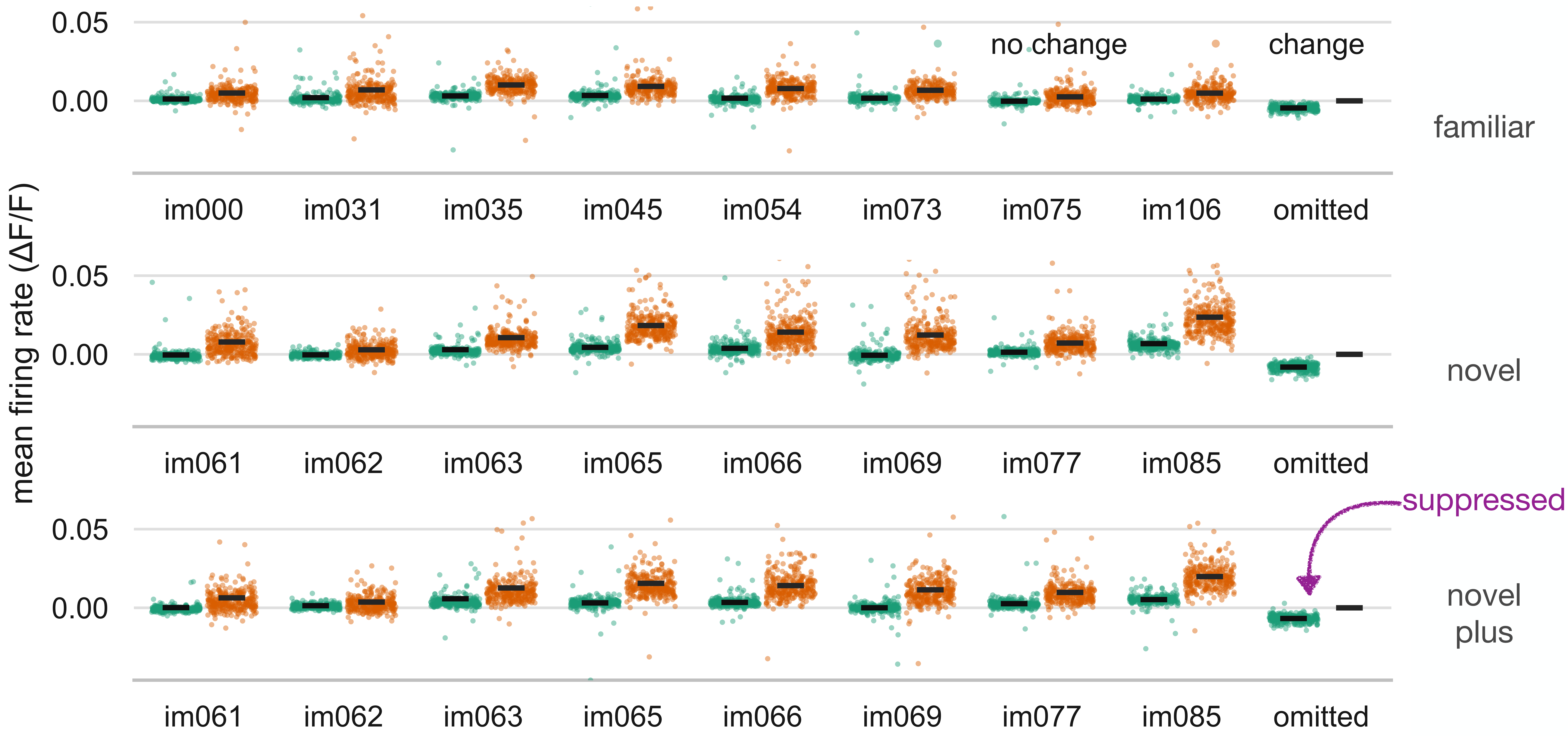
Neuronal activity over time for novel stimuli



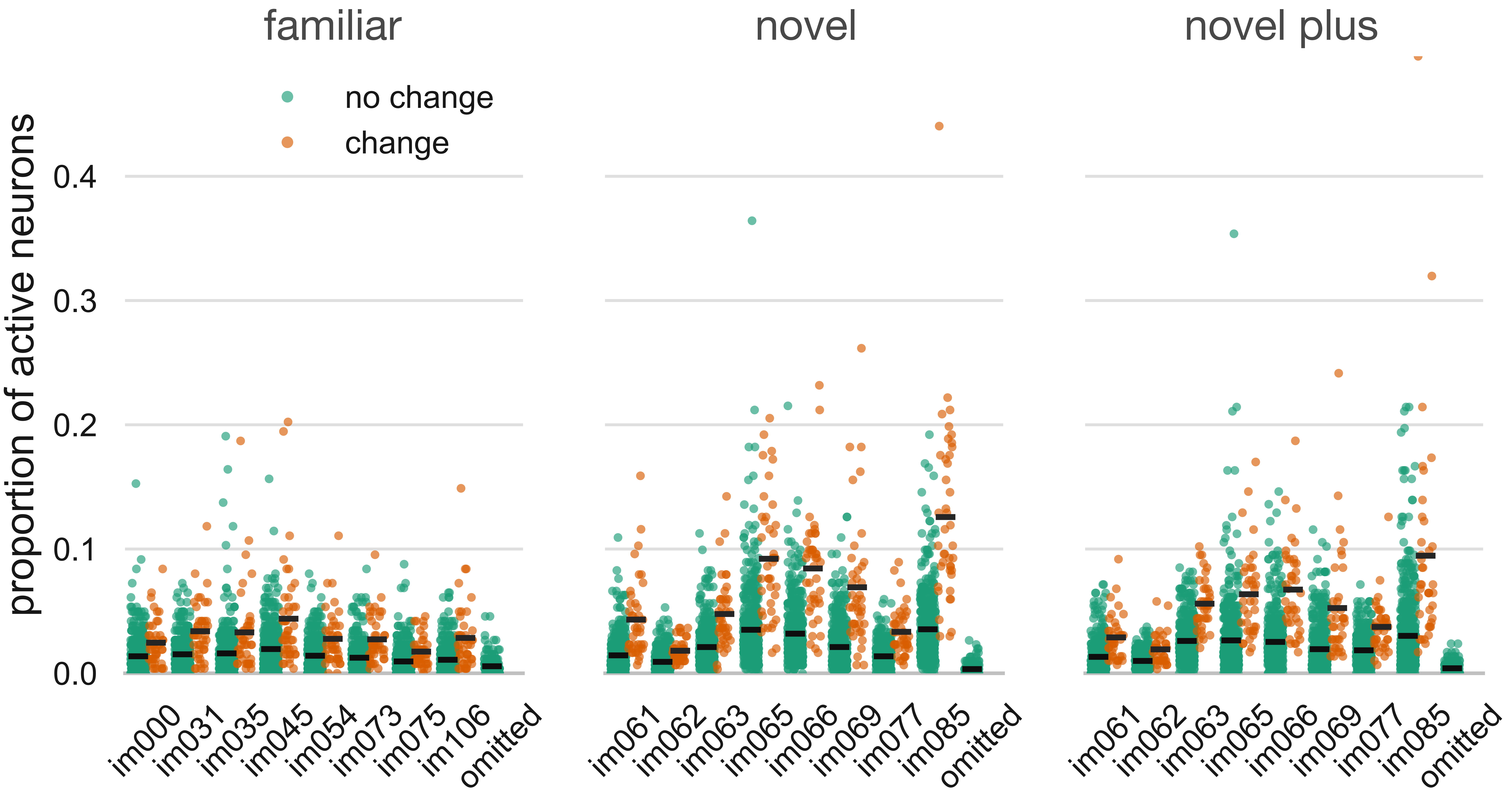
Mean E firing rates during image presentation

Mean E firing rates during image presentation

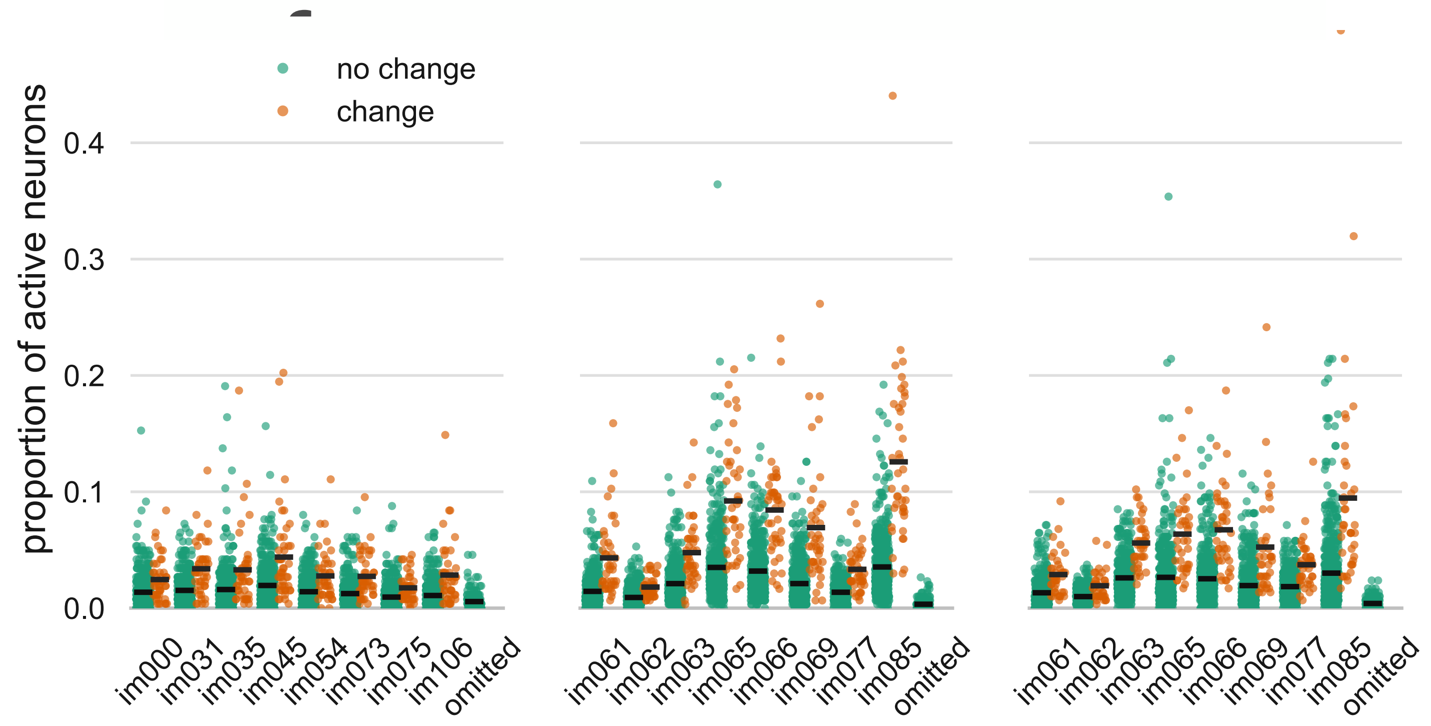
Mean E firing rates during image presentation

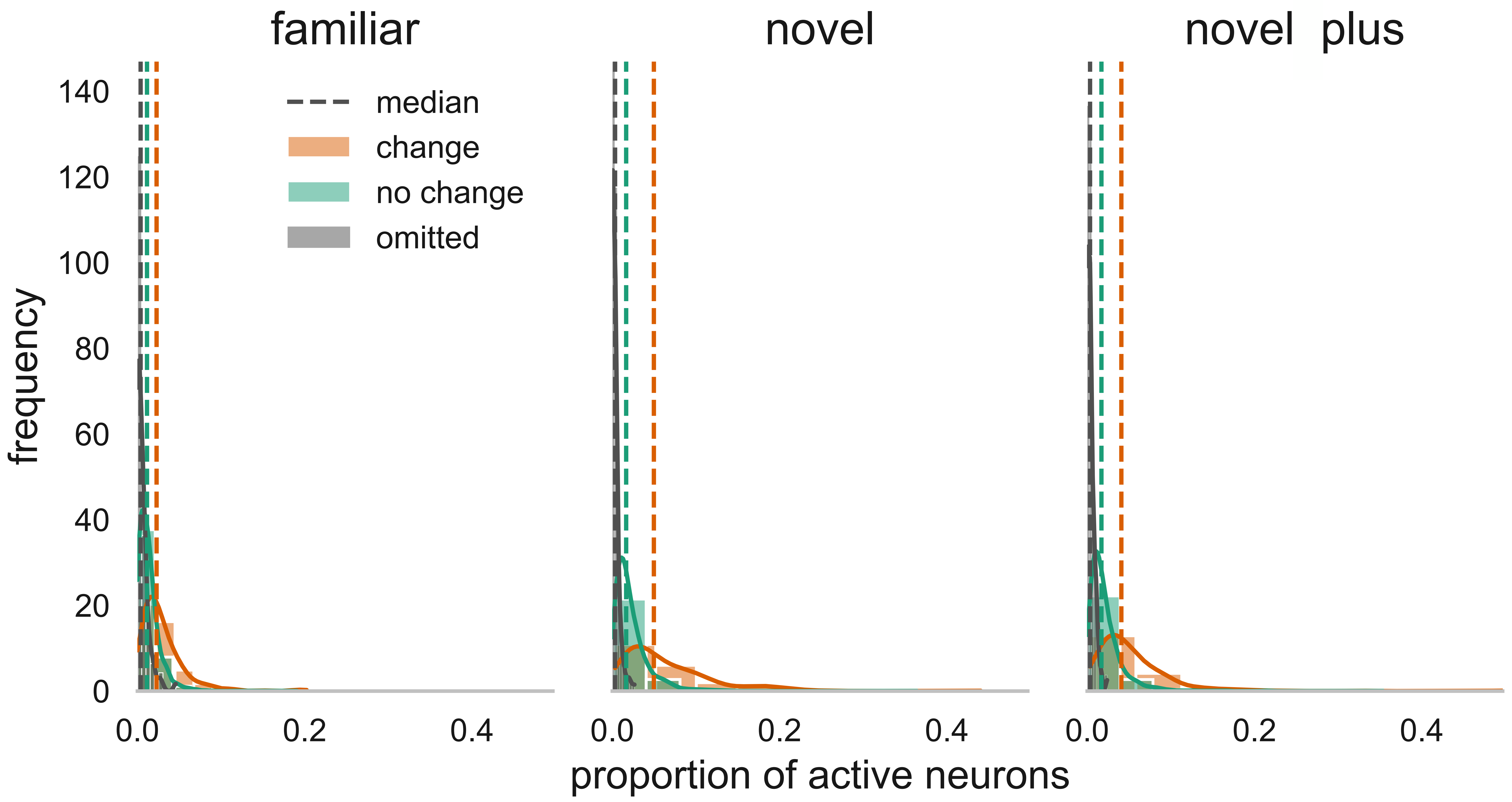


Responsive neurons per expected/unexpected stimulus

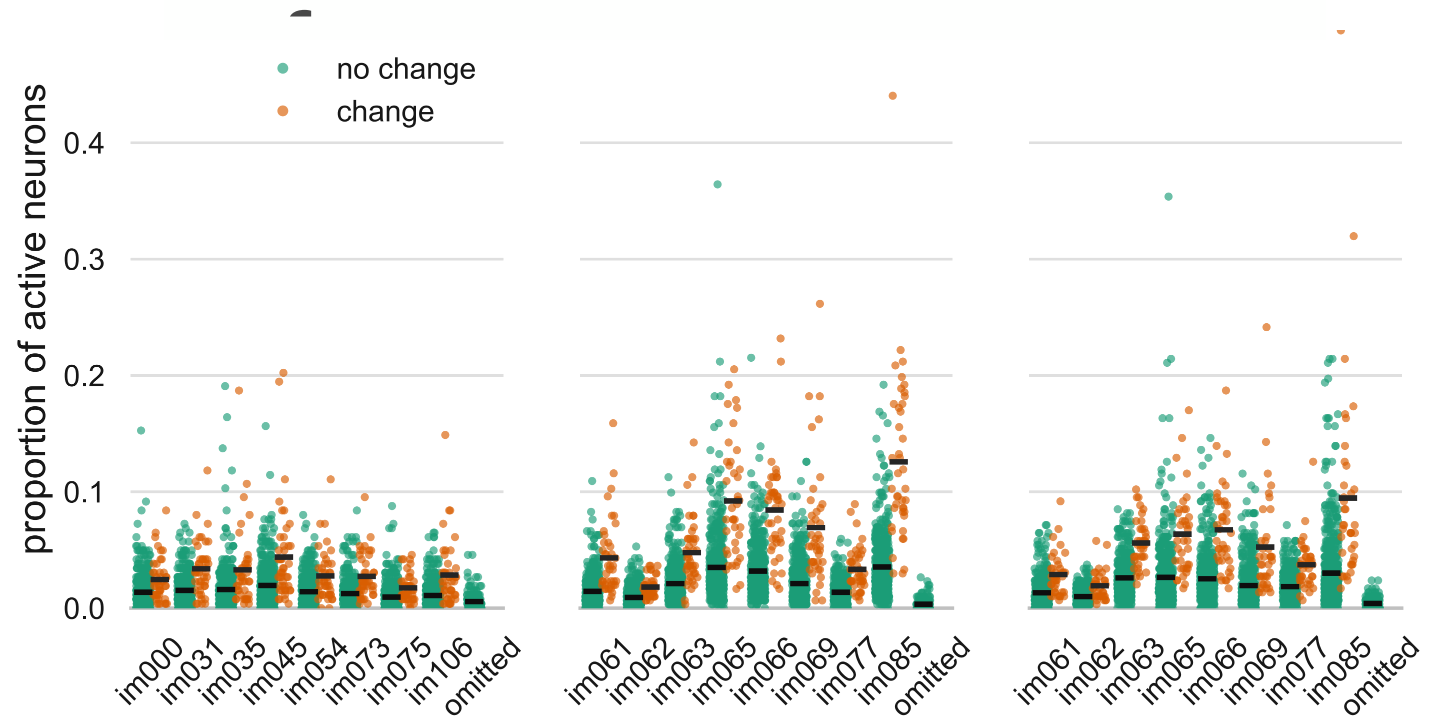


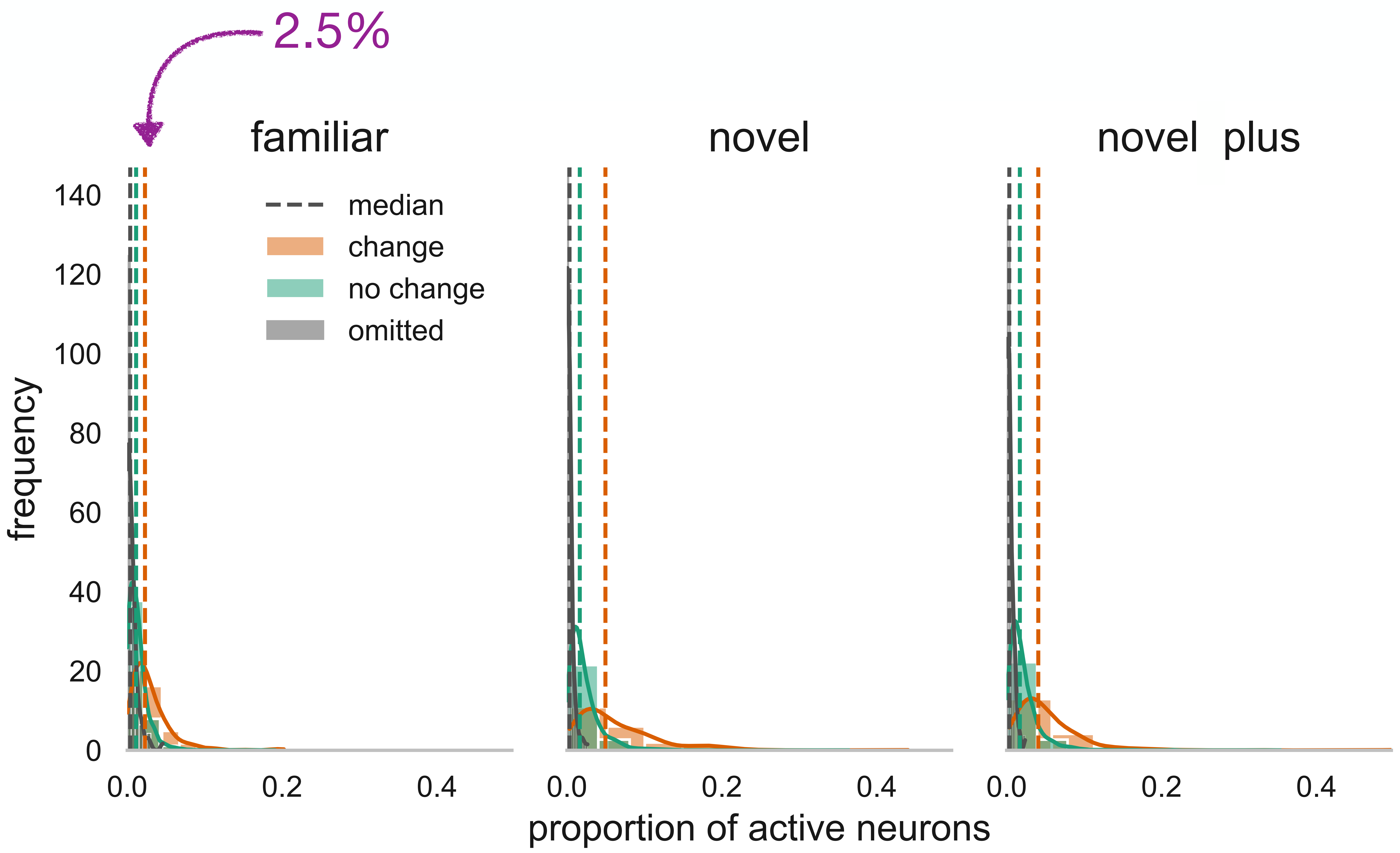
Responsive neurons per expected/unexpected stimulus



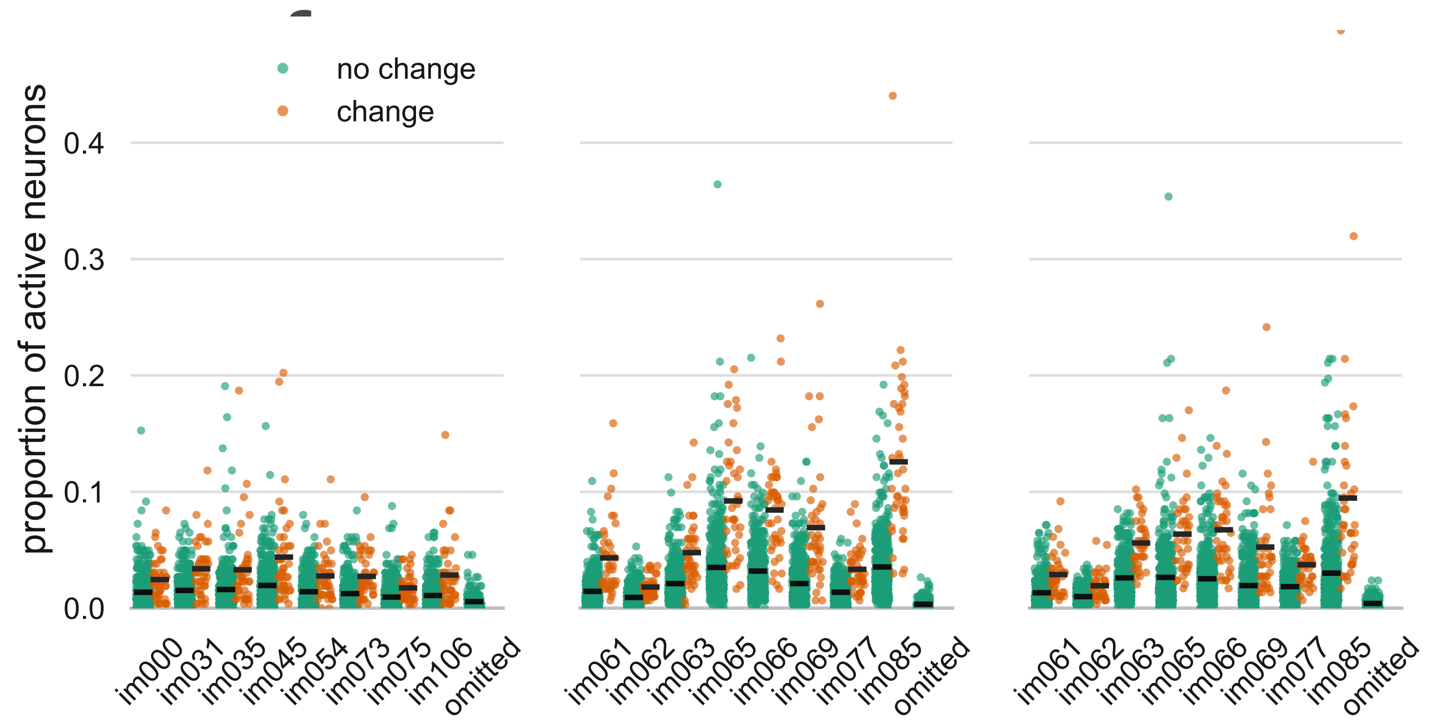


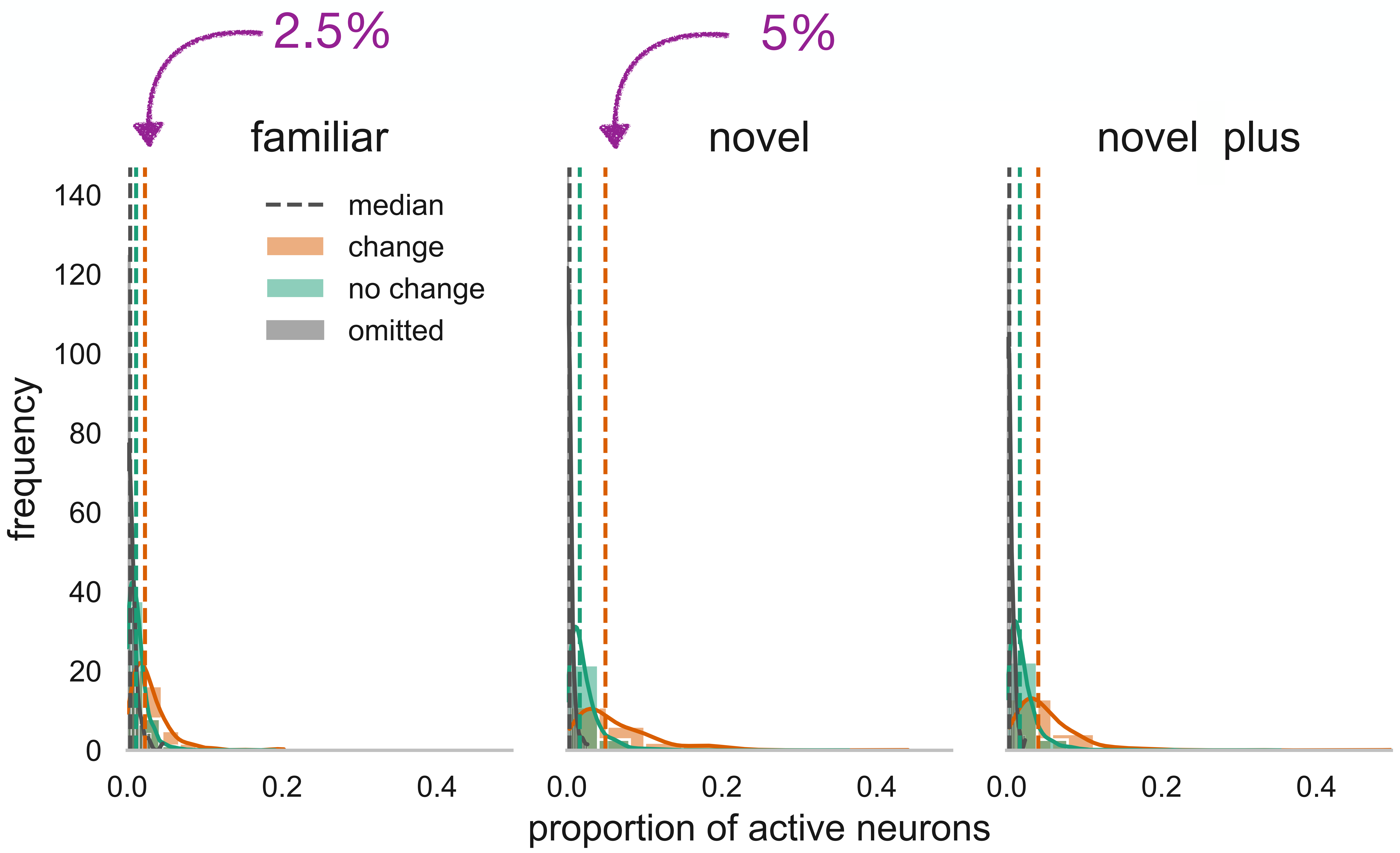
Responsive neurons per expected/unexpected stimulus



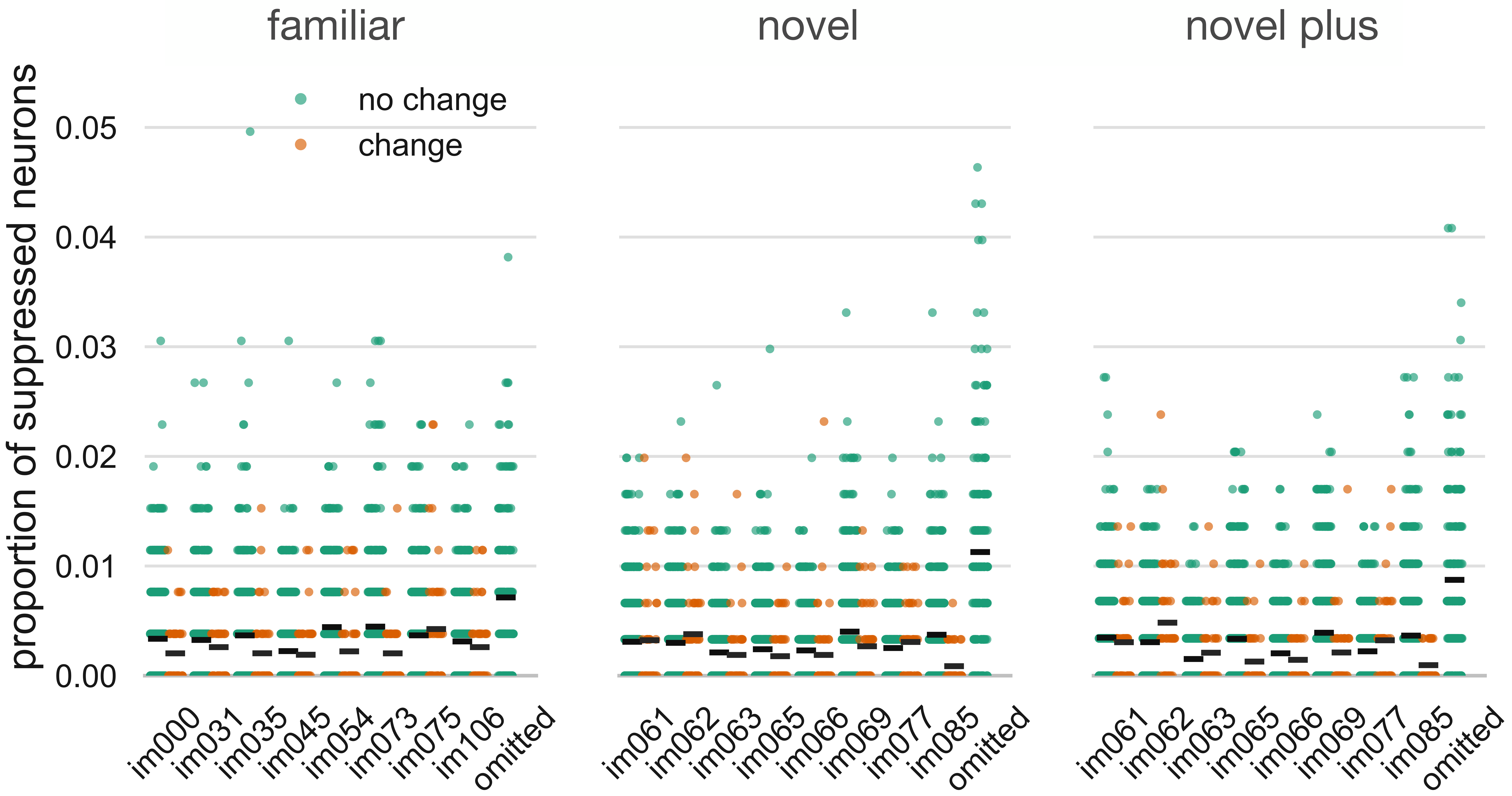


Responsive neurons per expected/unexpected stimulus

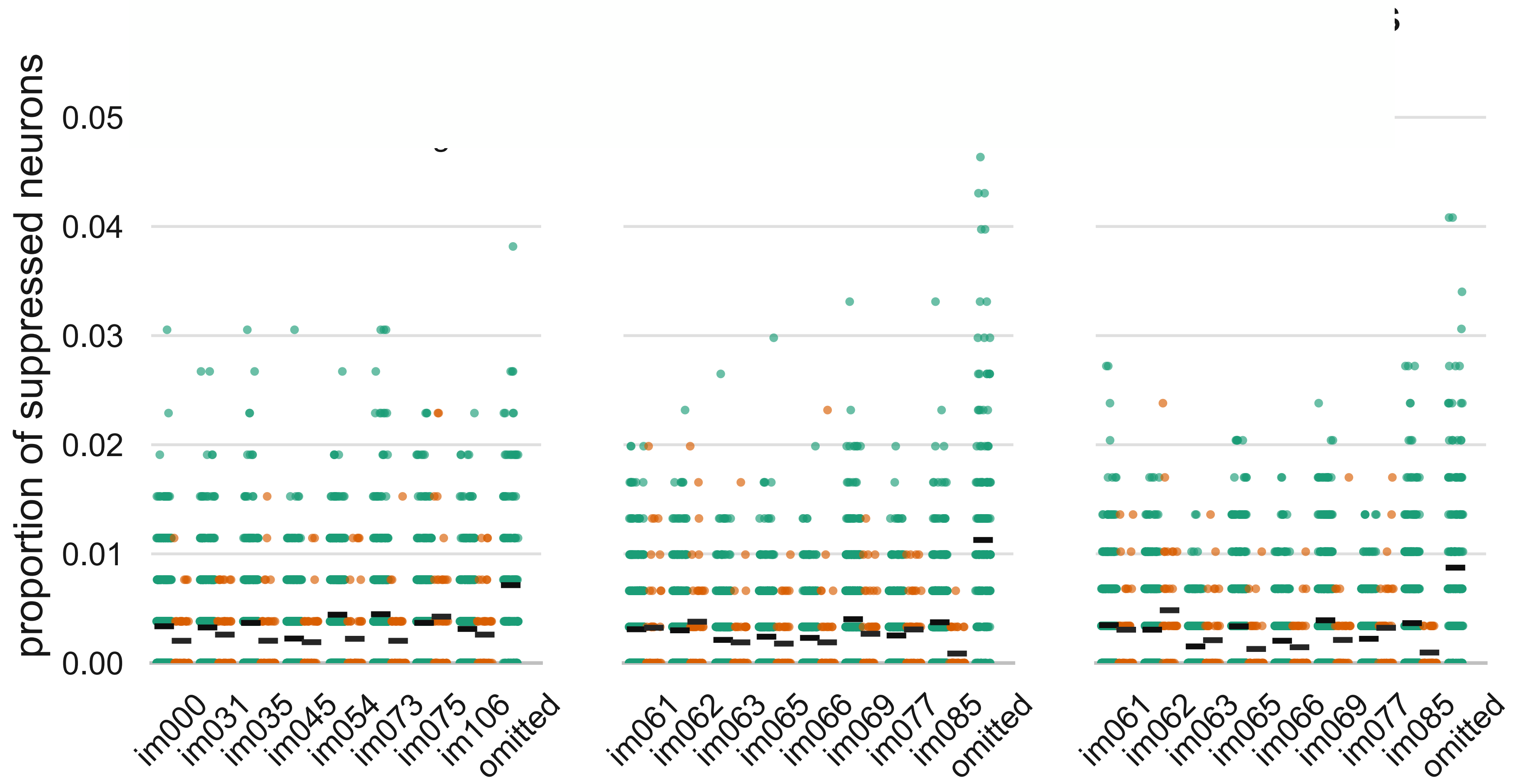


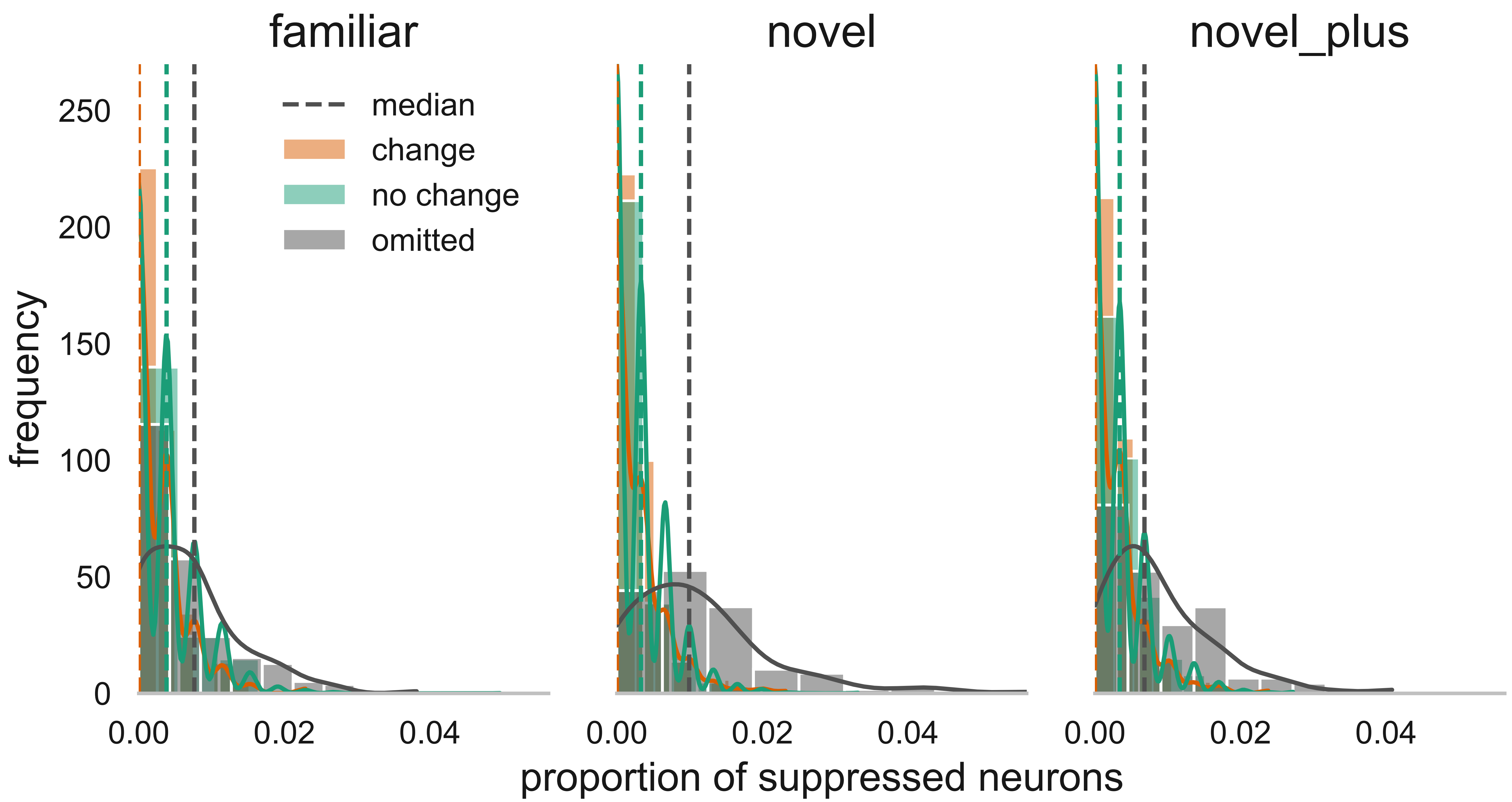


Suppressed neurons per expected/unexpected stimulus

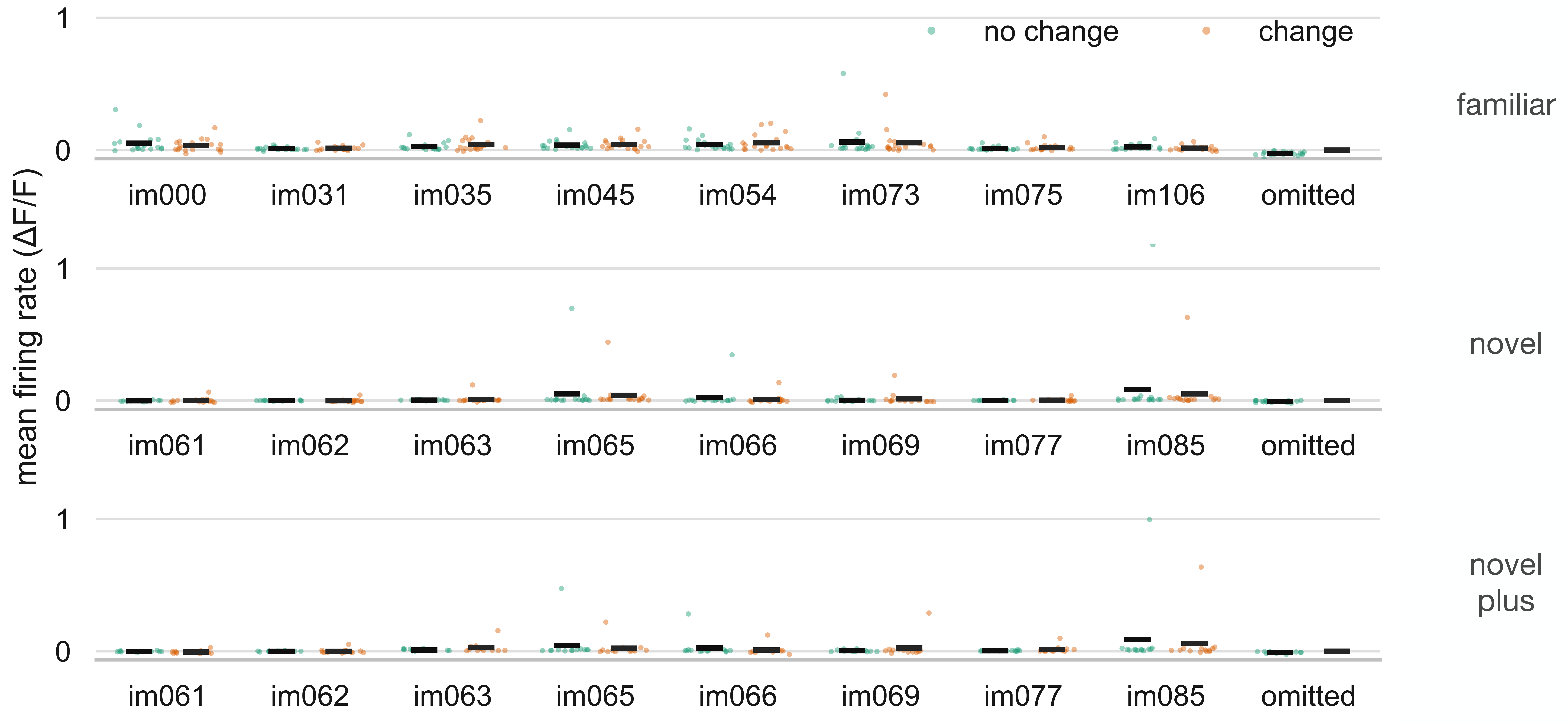


Suppressed neurons per expected/unexpected stimulus

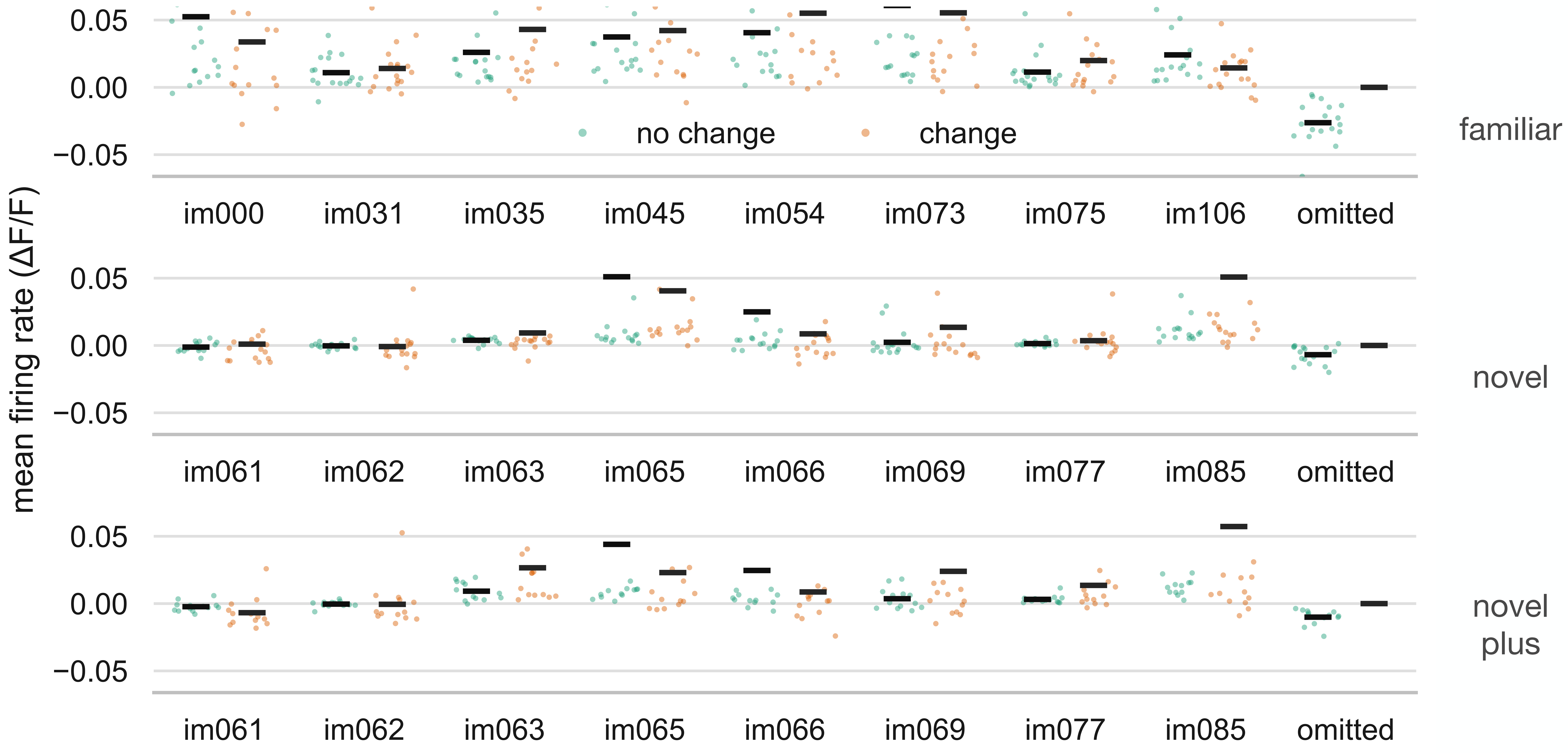




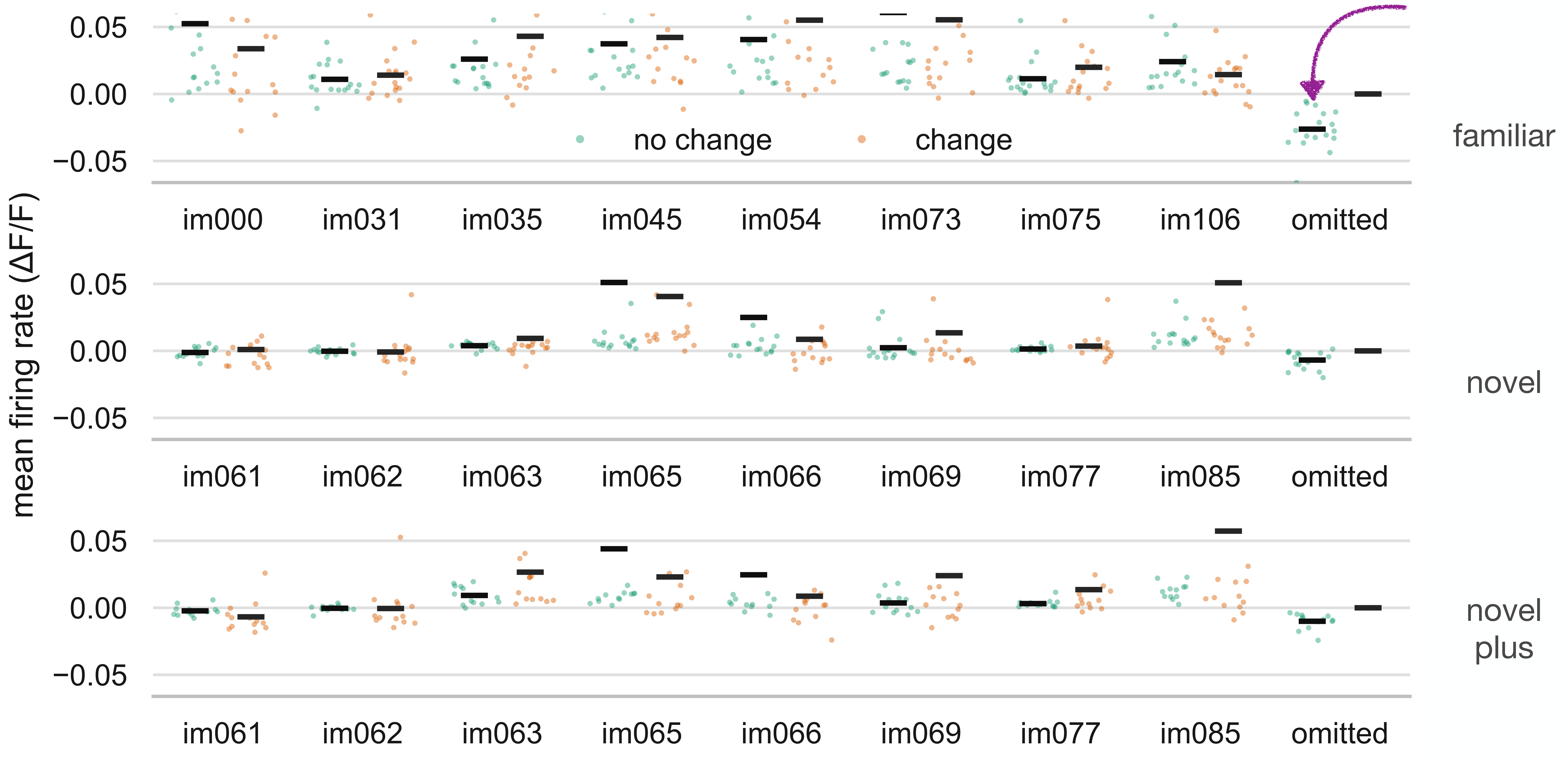
Mean SST firing rates during image presentation



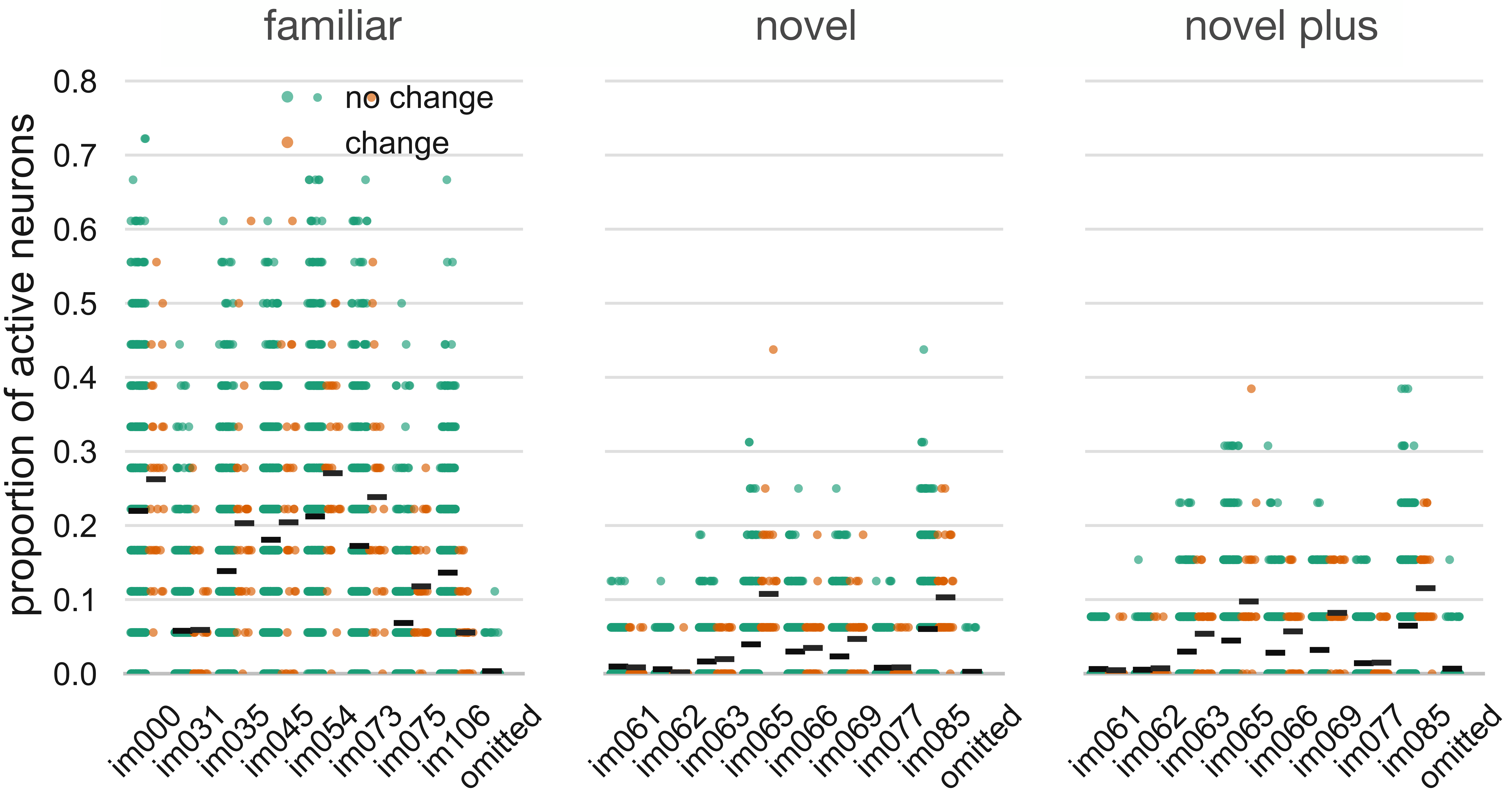
Mean SST firing rates during image presentation



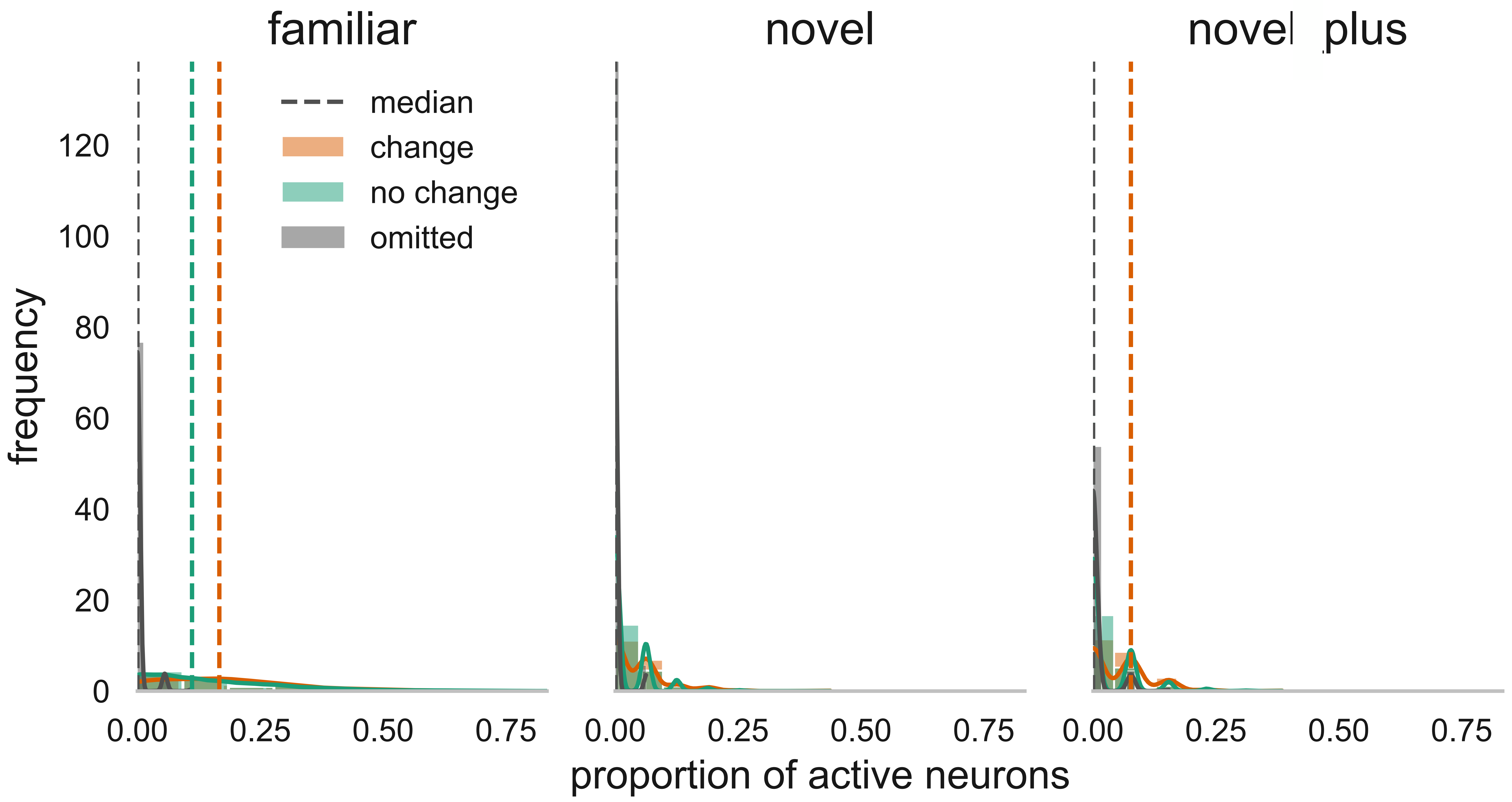
Mean SST firing rates during image presentation



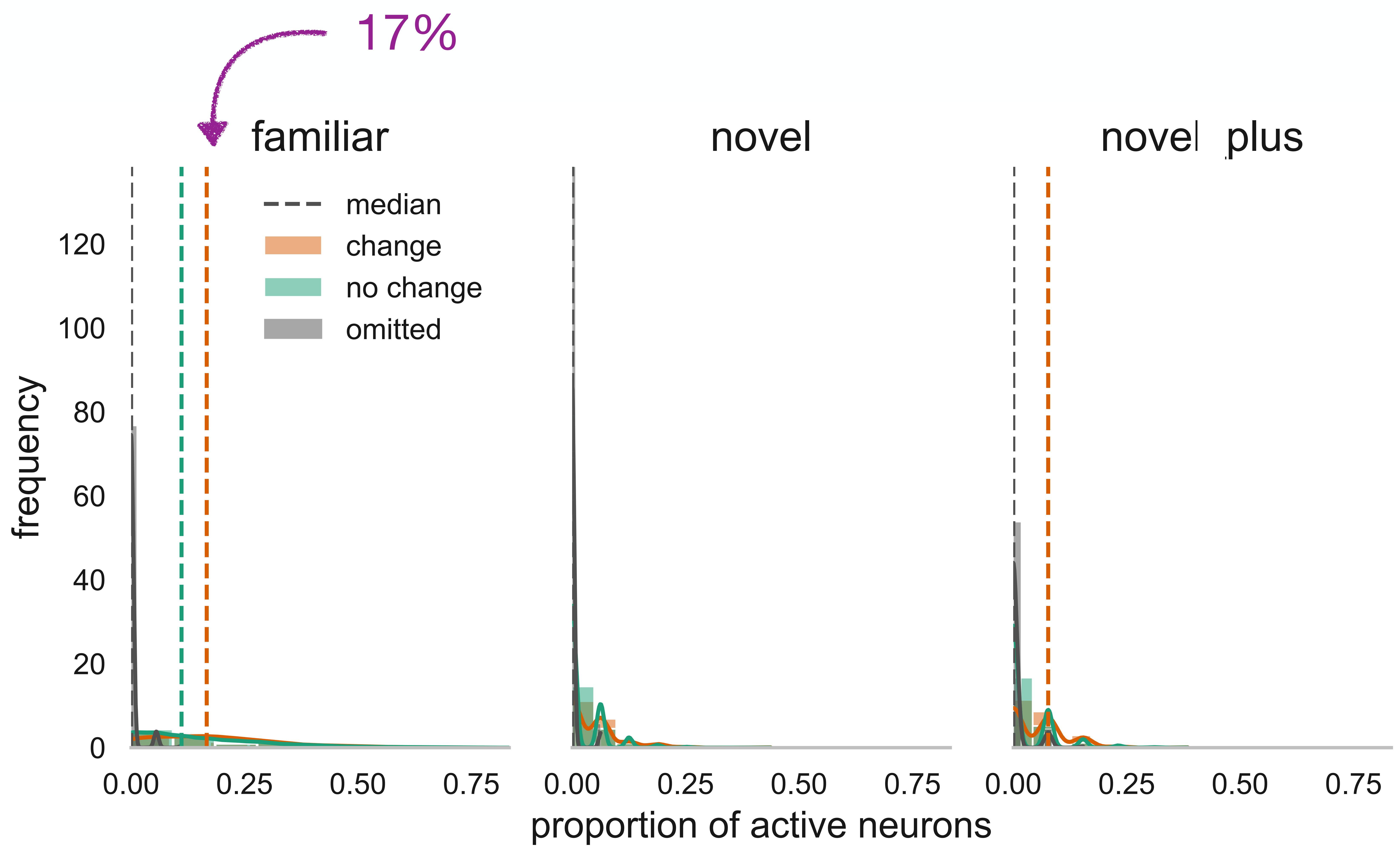
Responsive neurons per expected/unexpected stimulus



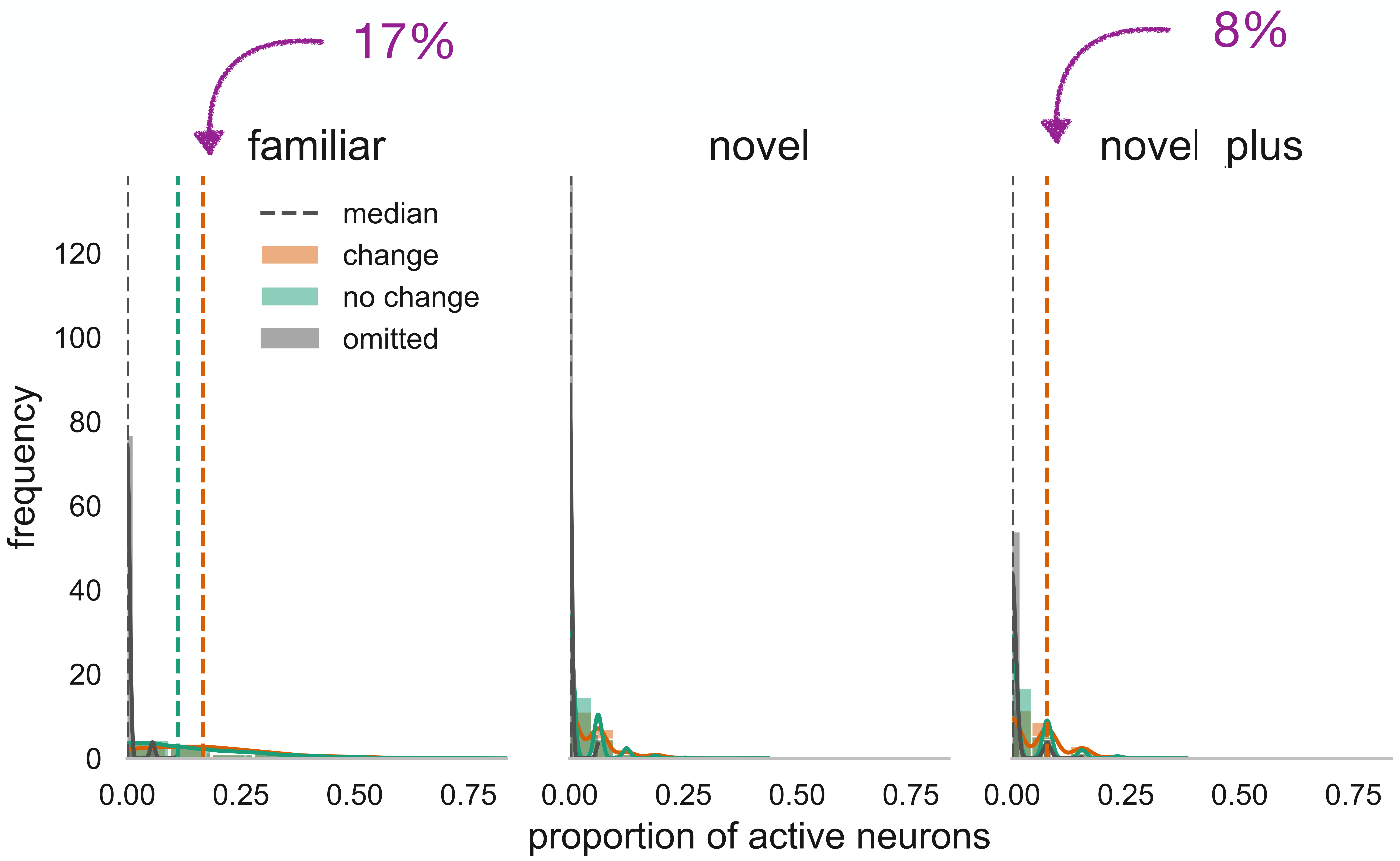
Responsive SST neurons per expected/unexpected stimulus

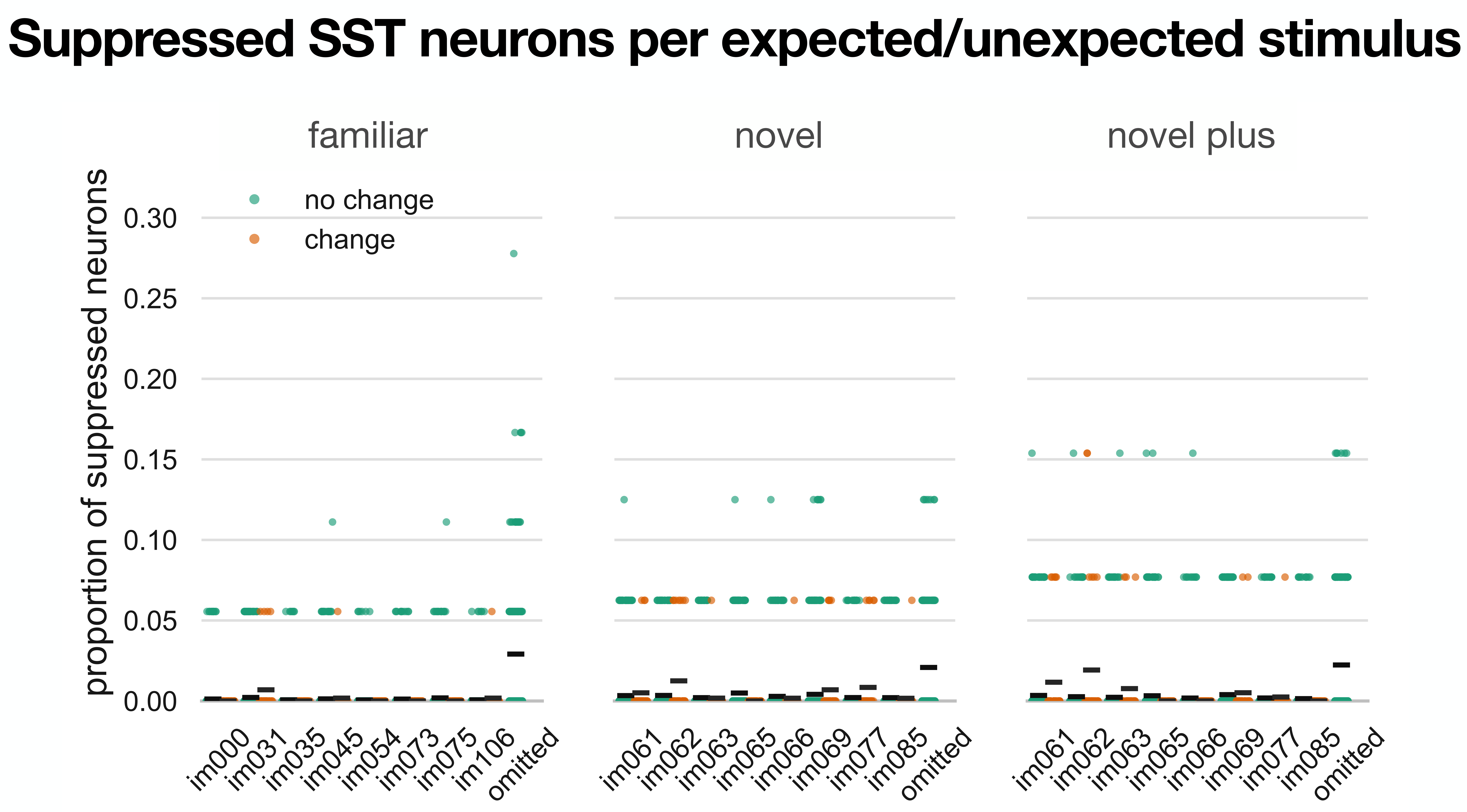


Responsive SST neurons per expected/unexpected stimulus

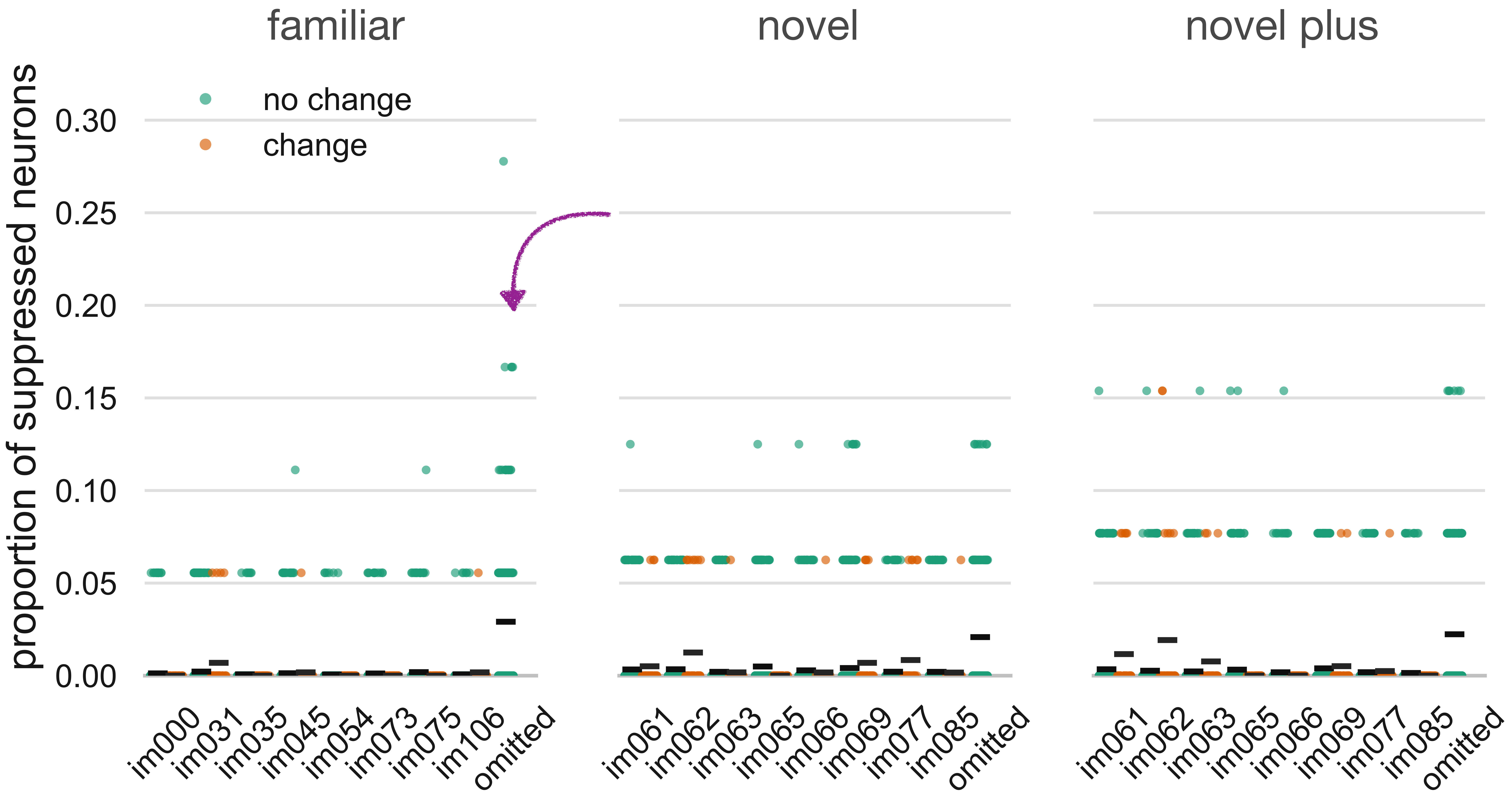


Responsive SST neurons per expected/unexpected stimulus

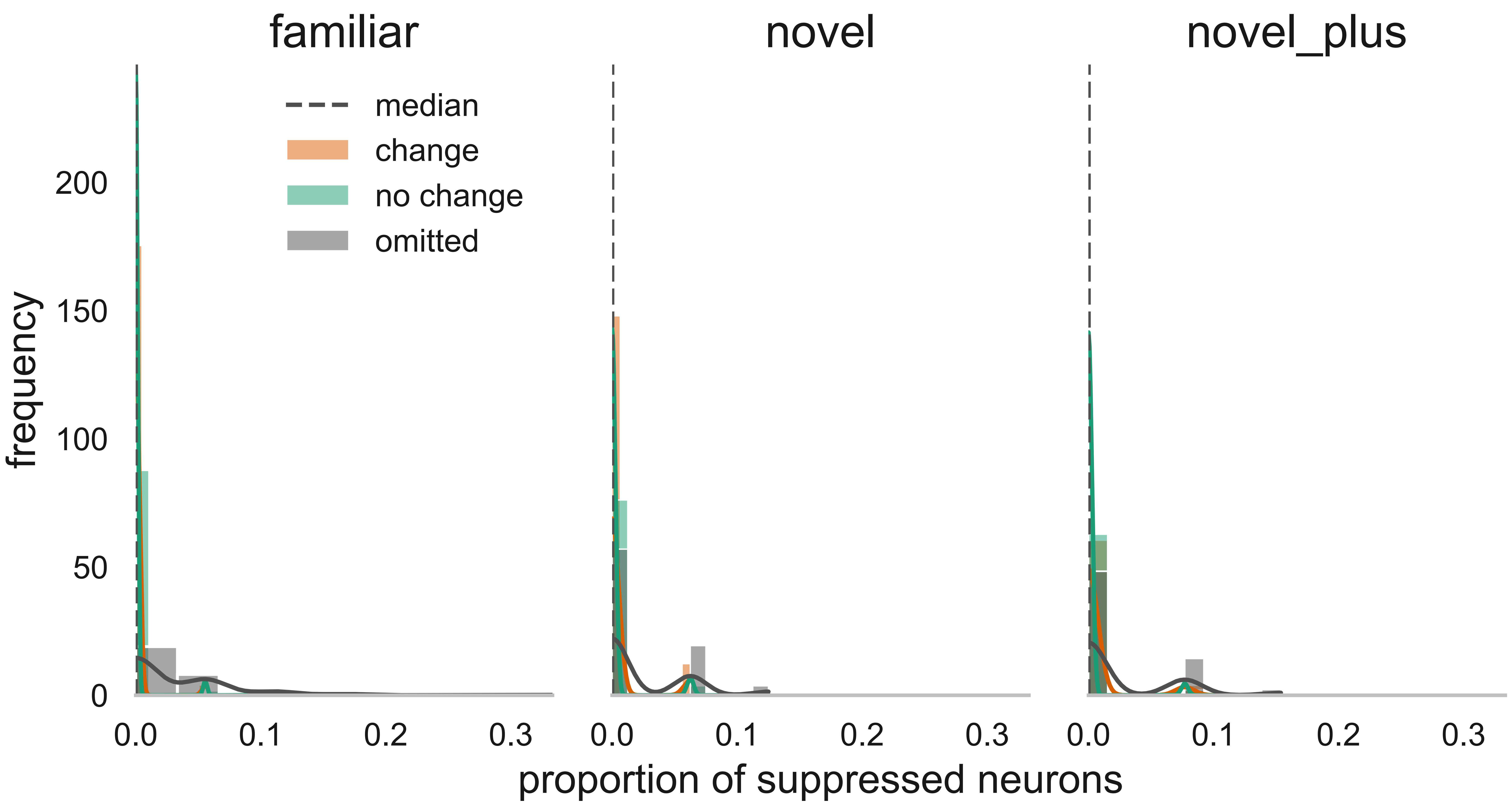



Suppressed SST neurons per expected/unexpected stimulus
familiar
novel
novel plus
no change
change
proportion of suppressed neurons
0.30
0.25
0.20
0.15
0.10
0.05
0.00
im000
im031
im035
im045
im054
im073
im075
im106
omitted
im061
im062
im063
im065
im066
im069
im077
im085
omitted
im061
im062
im063
im065
im066
im069
im077
im085
omitted

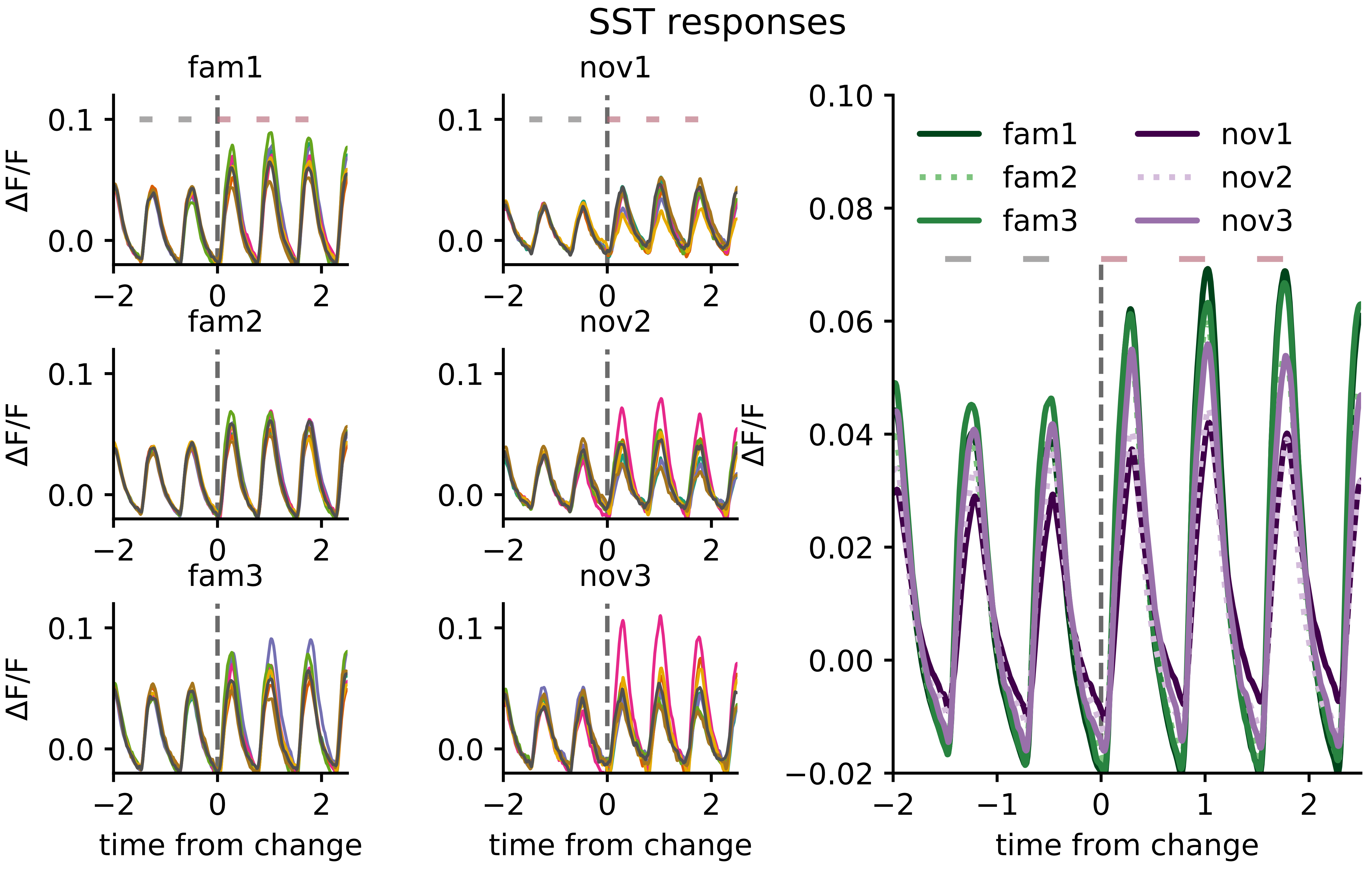
Suppressed SST neurons per expected/unexpected stimulus



Suppressed SST neurons per expected/unexpected stimulus



Average response profiles to expected and unexpected per stimulus identity



Average response profiles to expected and unexpected per stimulus identity

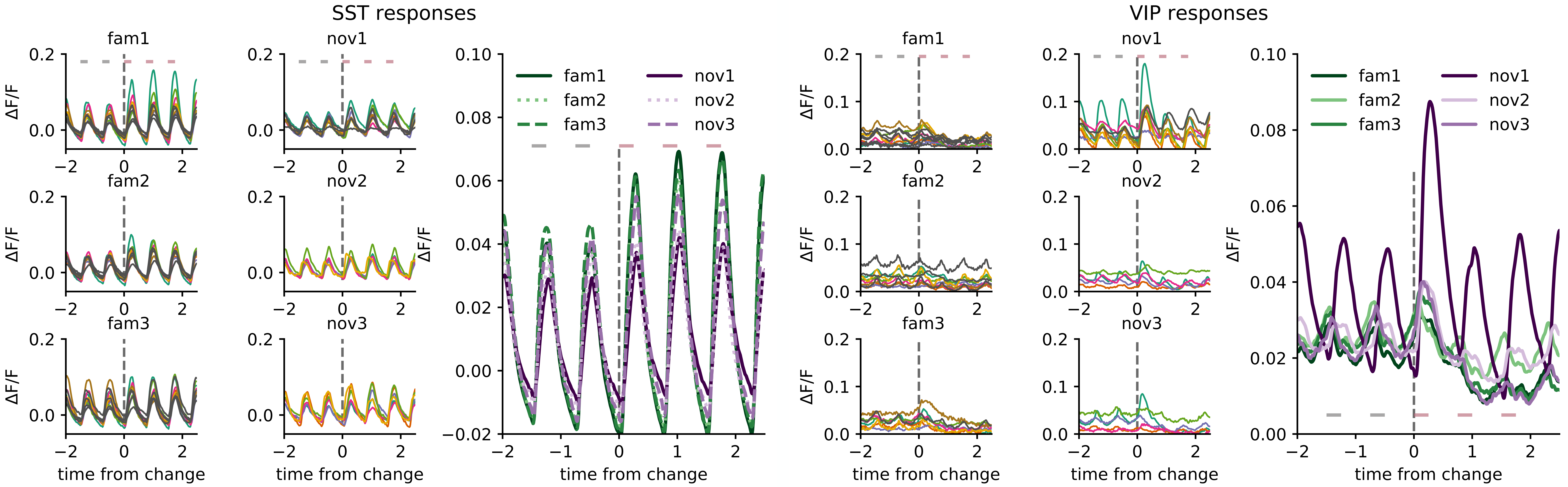
Average response profiles to expected and unexpected per stimulus identity

Average response profiles to expected and unexpected per stimulus identity

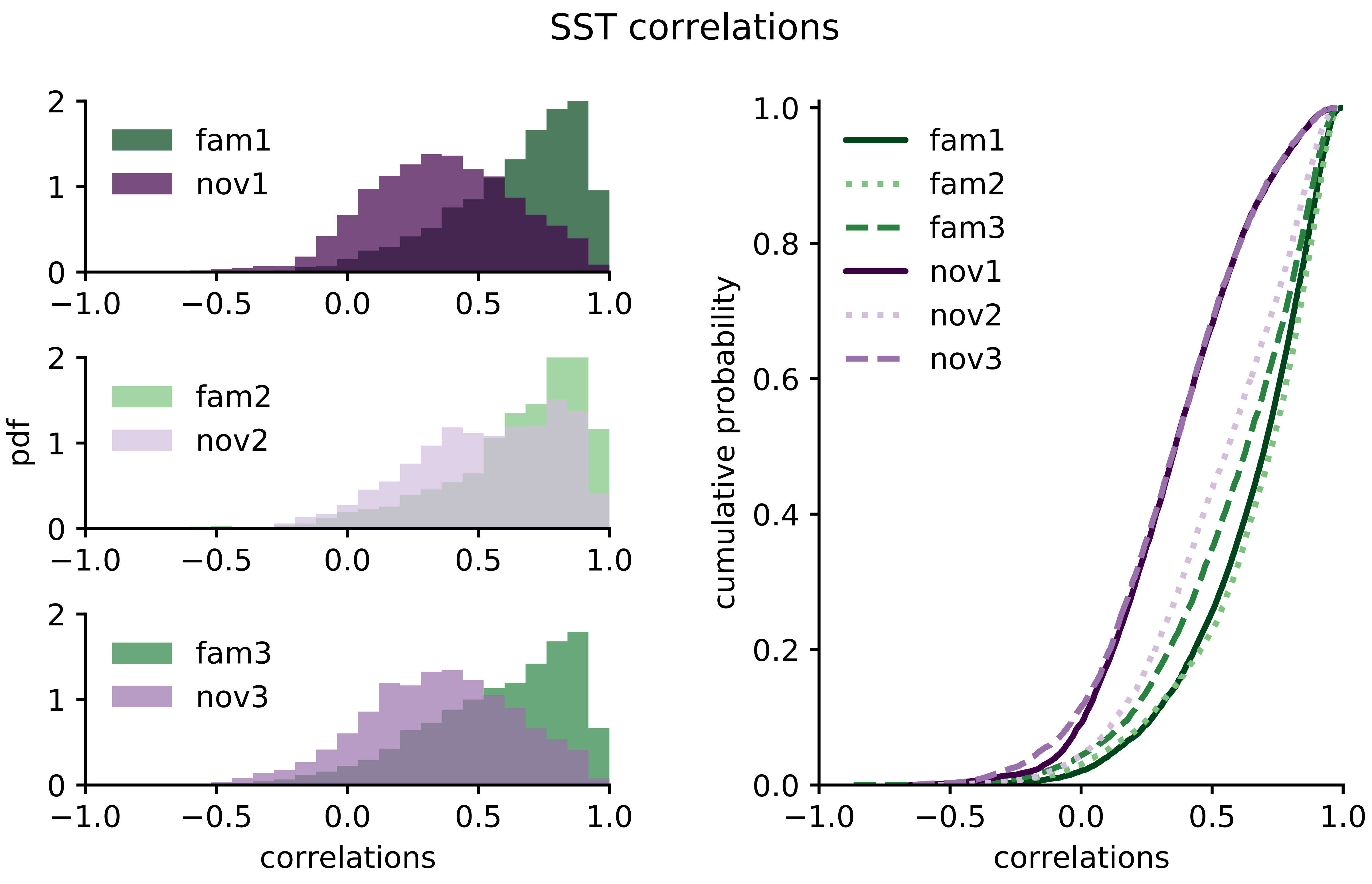
Average response profiles to expected and unexpected per stimulus identity

do all animals respond the same?

most mice respond similarly

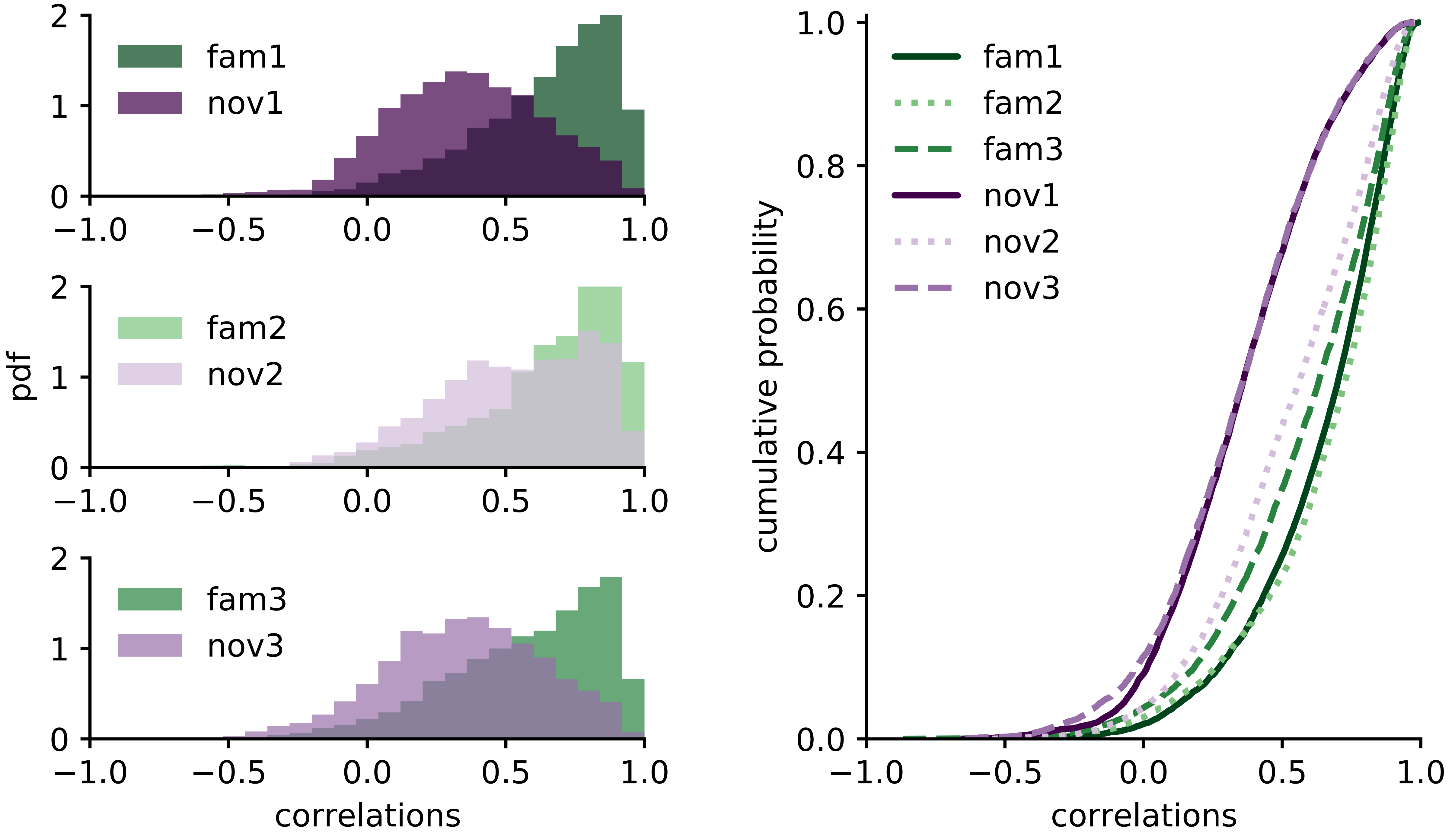


Different correlation profiles for each interneuron type

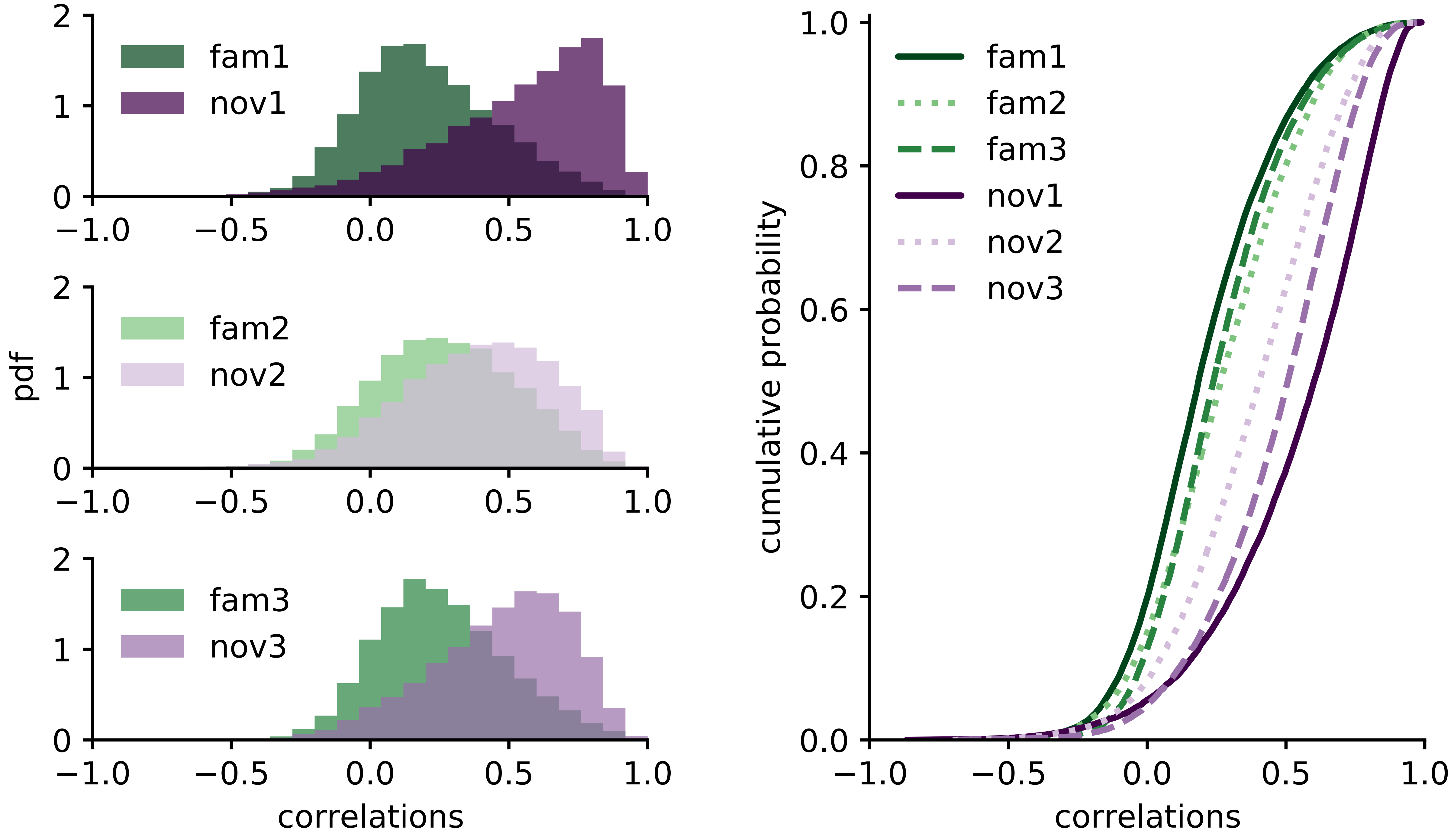


Different correlation profiles for each interneuron type

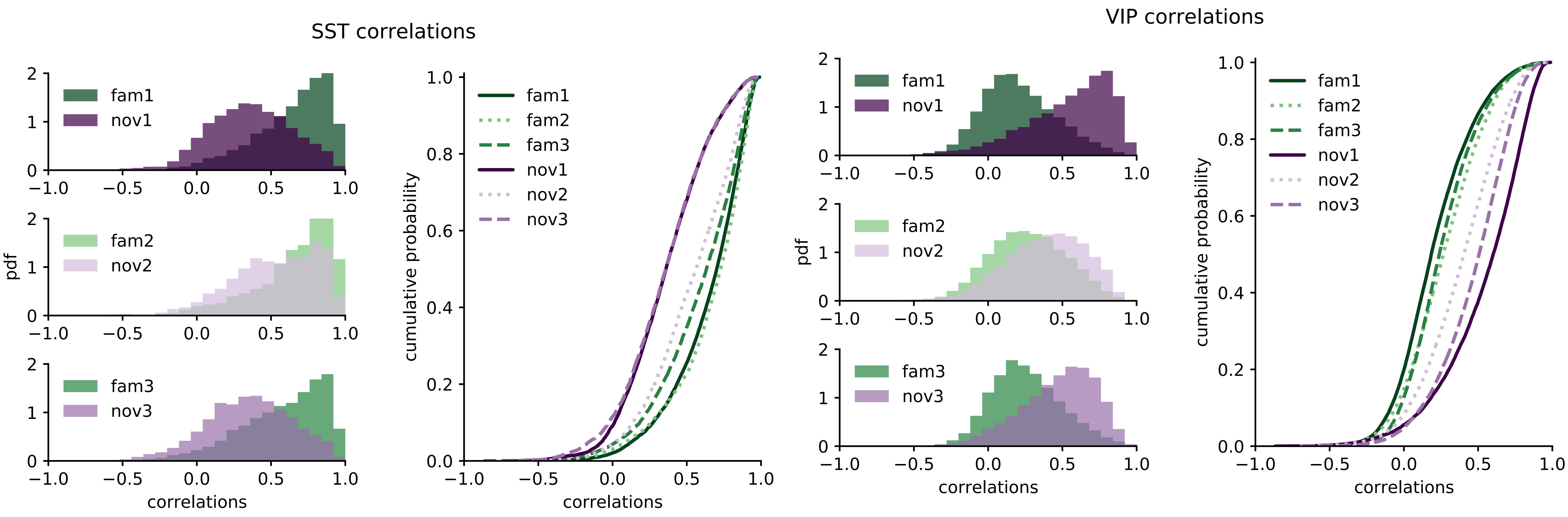
SST correlations



VIP correlations

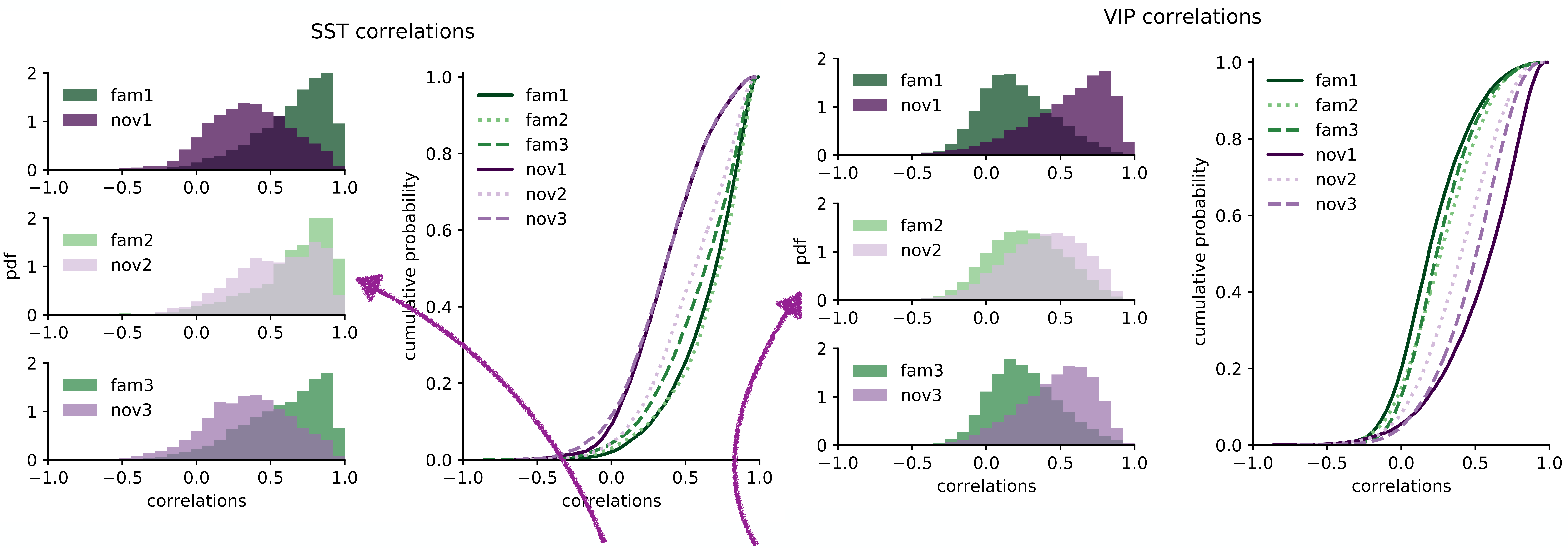


Different correlation profiles for each interneuron type



correlations become more similar
for passive/unrewarded sessions

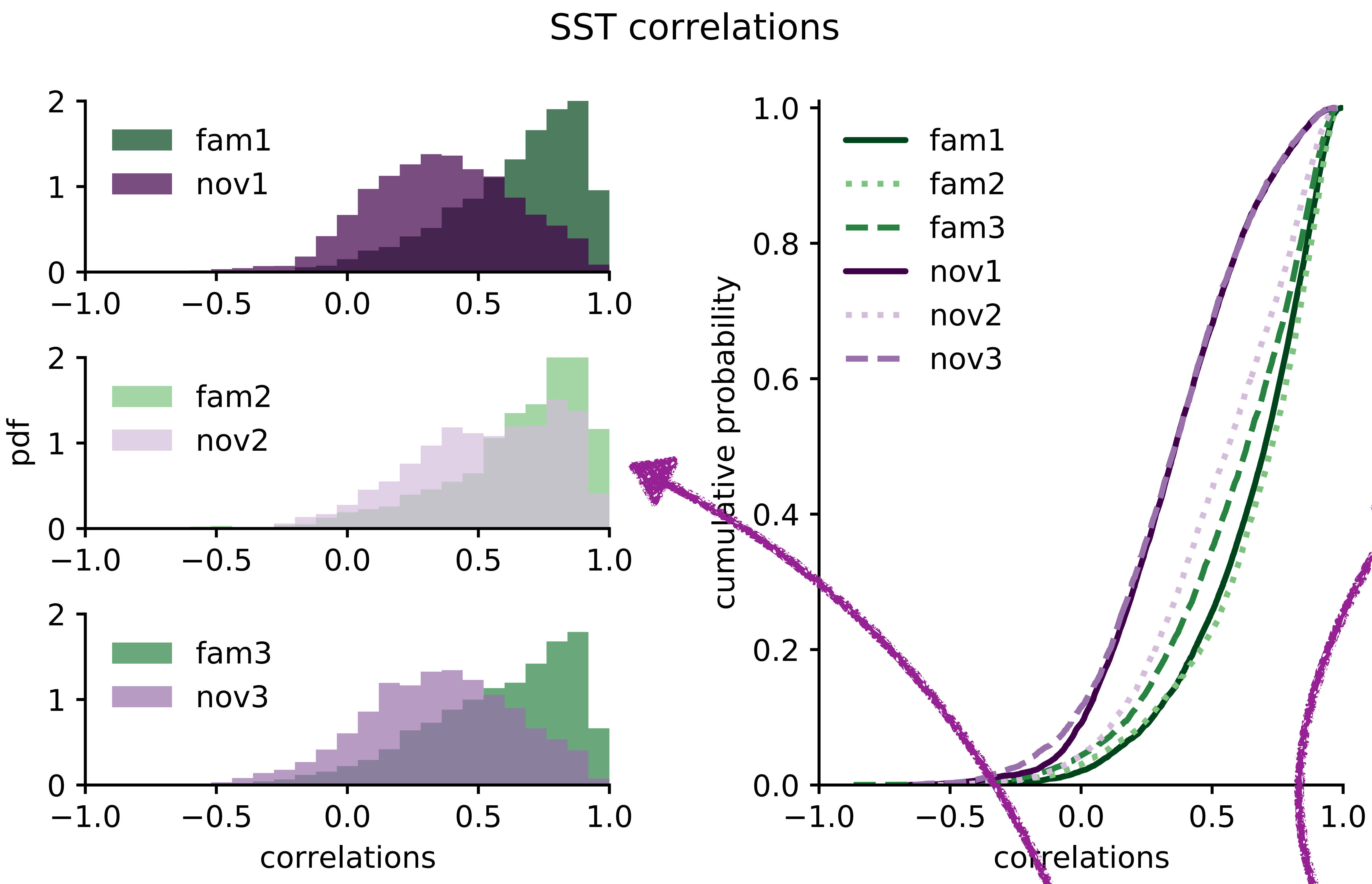
Different correlation profiles for each interneuron type

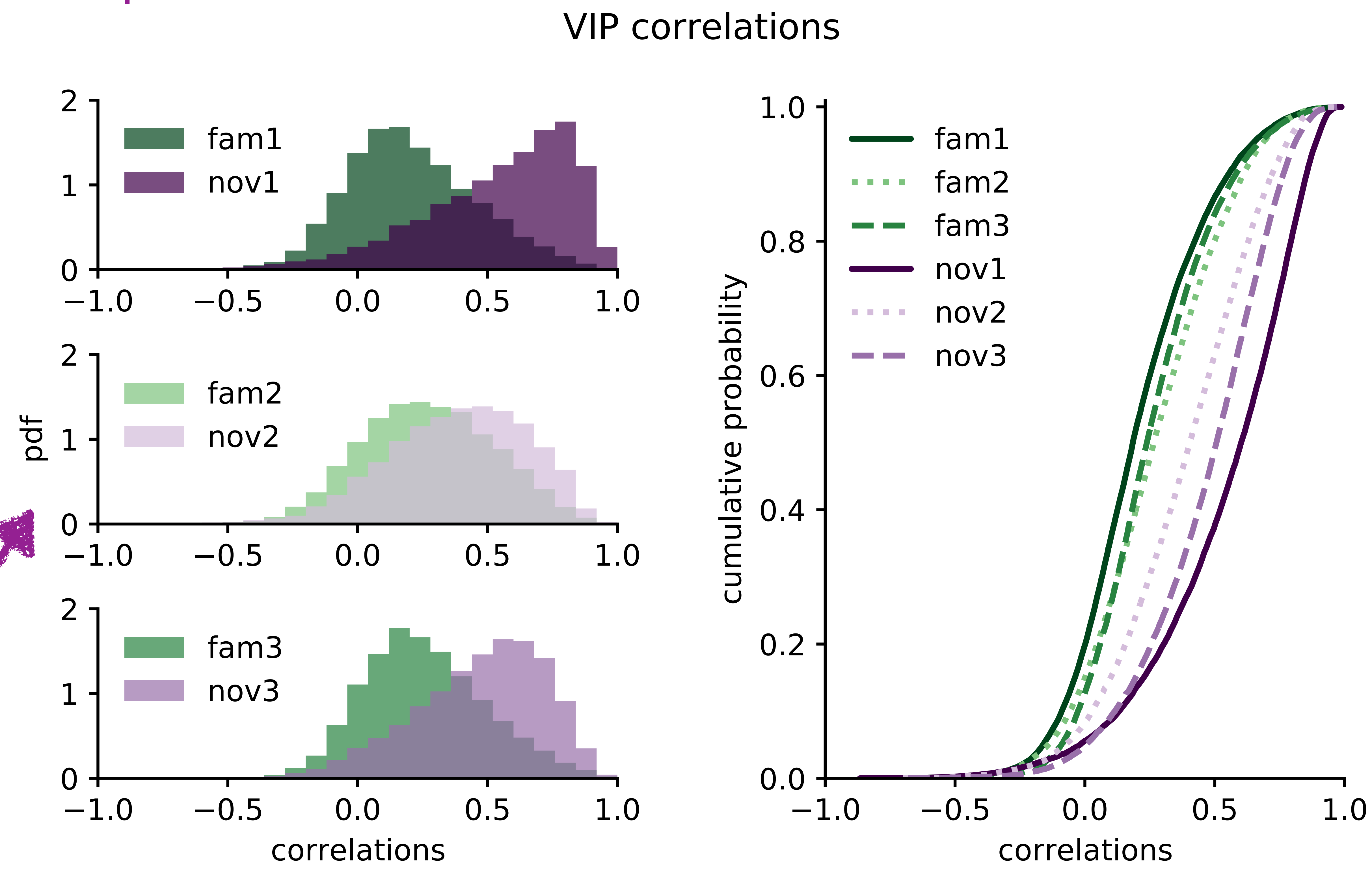


correlations become more similar
for passive/unrewarded sessions

Different correlation profiles for each interneuron type

different stimulus encoding based
on reward contingencies,
partly due to reduced top-down feedback

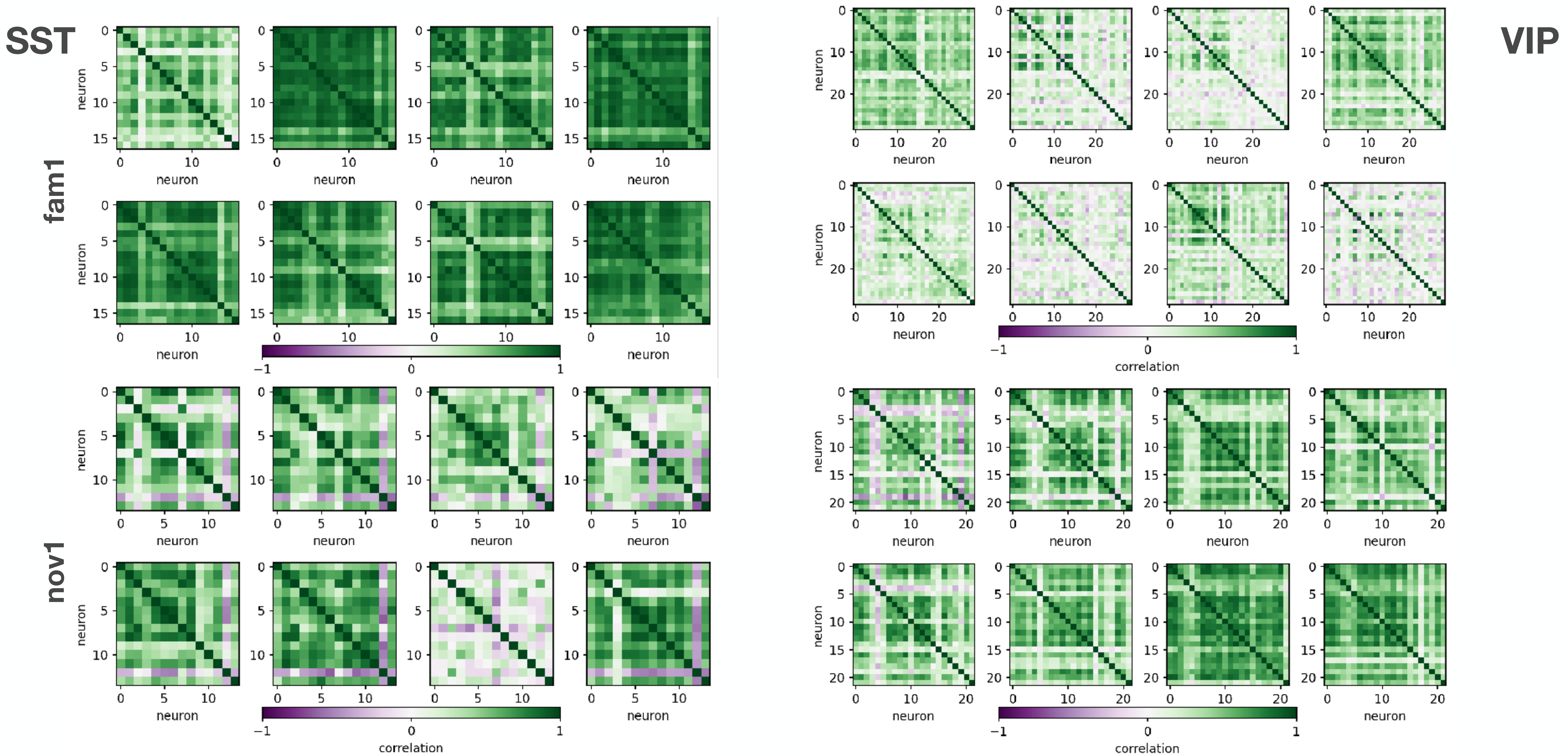




correlations become more similar
for passive/unrewarded sessions

how do correlations look for individual images?

Diverse SST responses on novel conditions



Model

familiar no change

- SSTs suppress PV & E
- top-down feedback to E - previous buffer
- top-down feedback is inhibitory on E to cancel out feedforward excitation

familiar change

- SSTs suppress PV & E
- top-down feedback to E - previous buffer
- error in top-down feedback suppresses some E and the E that correspond to the changed stim show pronounced responses

kick on VIP and SST1

novel no change

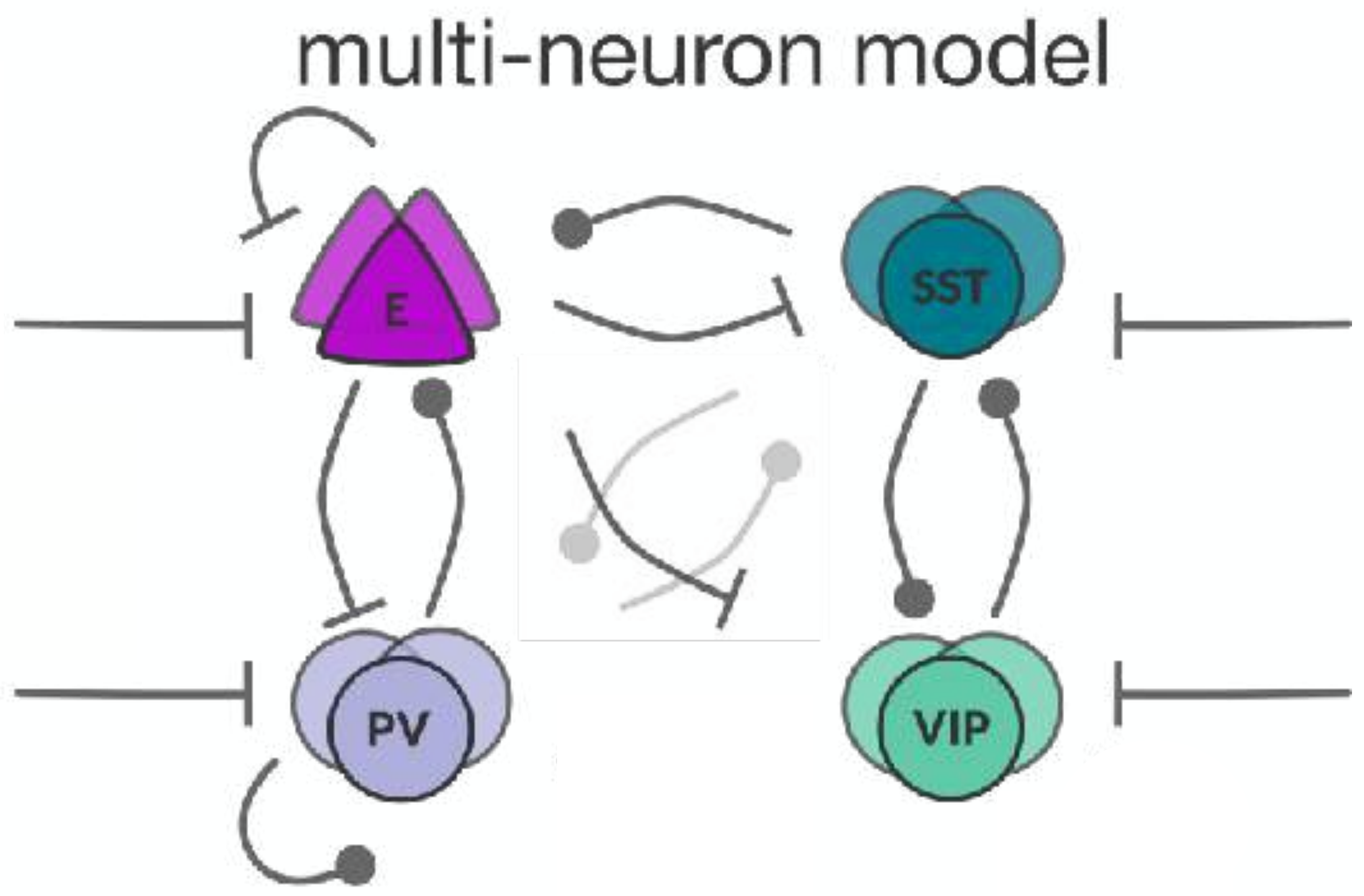
- PV and E are active - not yet suppressed by SSTs
- top-down feedback to E - previous buffer - but with less amplitude
- top-down feedback is inhibitory on E to cancel out part of feedforward excitation
- the small peak on VIPs may be due recurrent excitation from E and not due to top down

novel change

- PV and E are active - not yet suppressed by SSTs
- top-down feedback to E - previous buffer - now does not match input
- kick on VIP and SST1

Our strategy

Circuit parameters to be adjusted

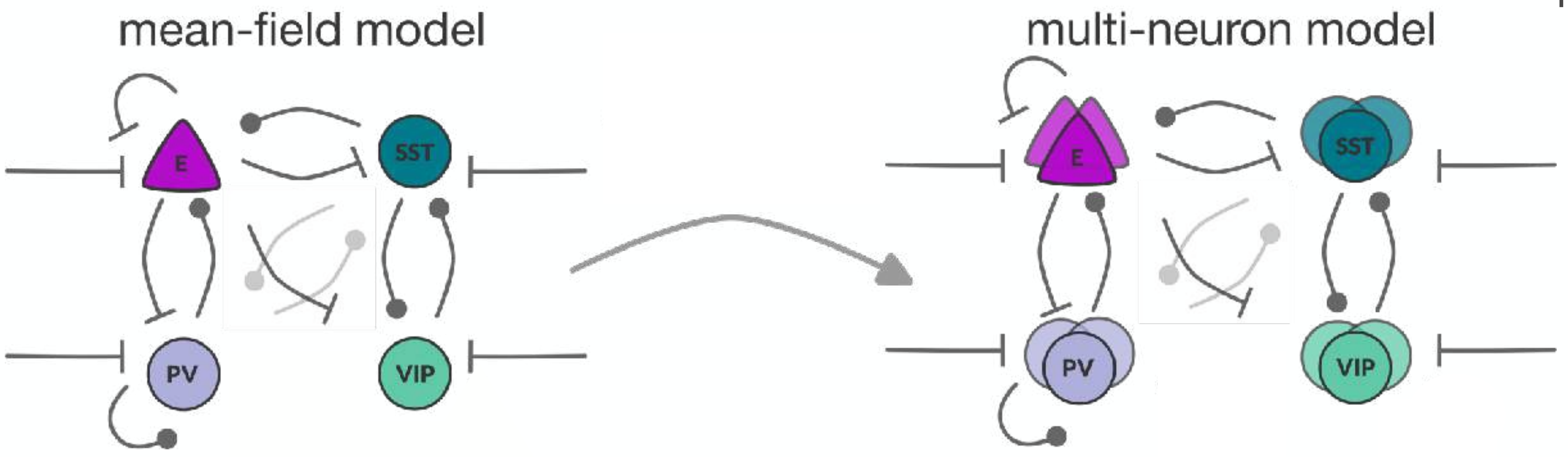


weights	W_{AB}
background inputs	I_A
stimulus inputs	I_A^{stim}
feedback inputs	I_A^{HOA}

Our strategy

Circuit parameters to be adjusted

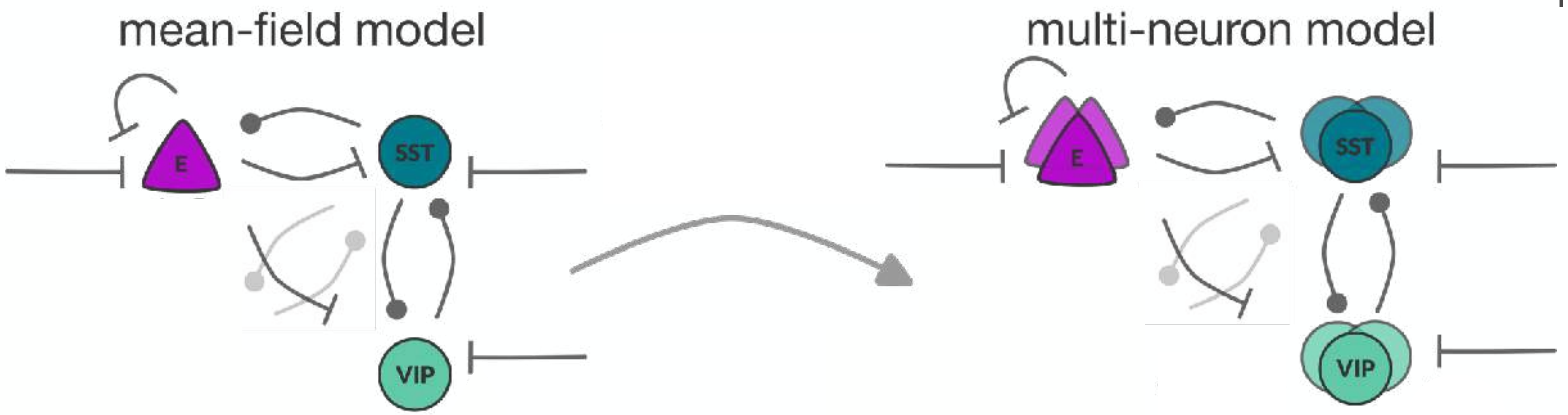
weights	W_{AB}
background inputs	I_A
stimulus inputs	I_A^{stim}
feedback inputs	I_A^{HOA}



Our strategy

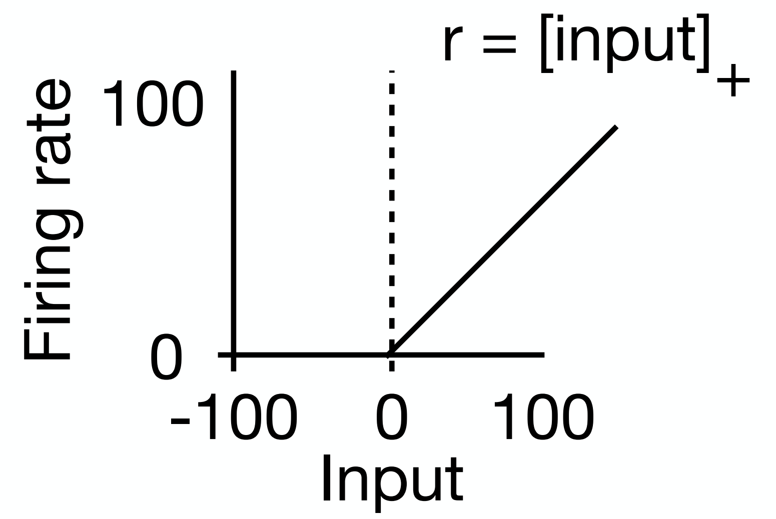
Circuit parameters to be adjusted

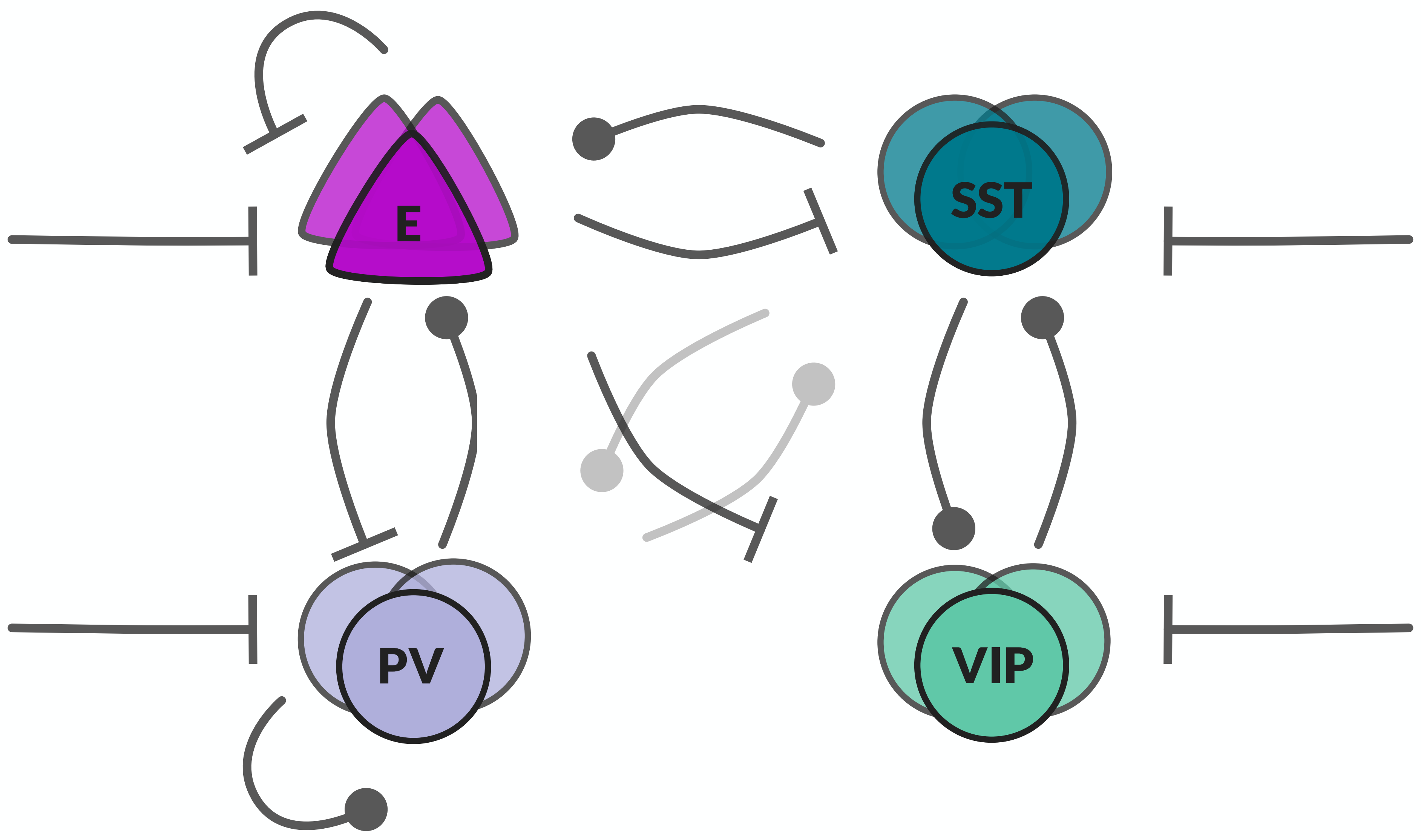
weights	W_{AB}
background inputs	I_A
stimulus inputs	I_A^{stim}
feedback inputs	I_A^{HOA}



Microcircuit model

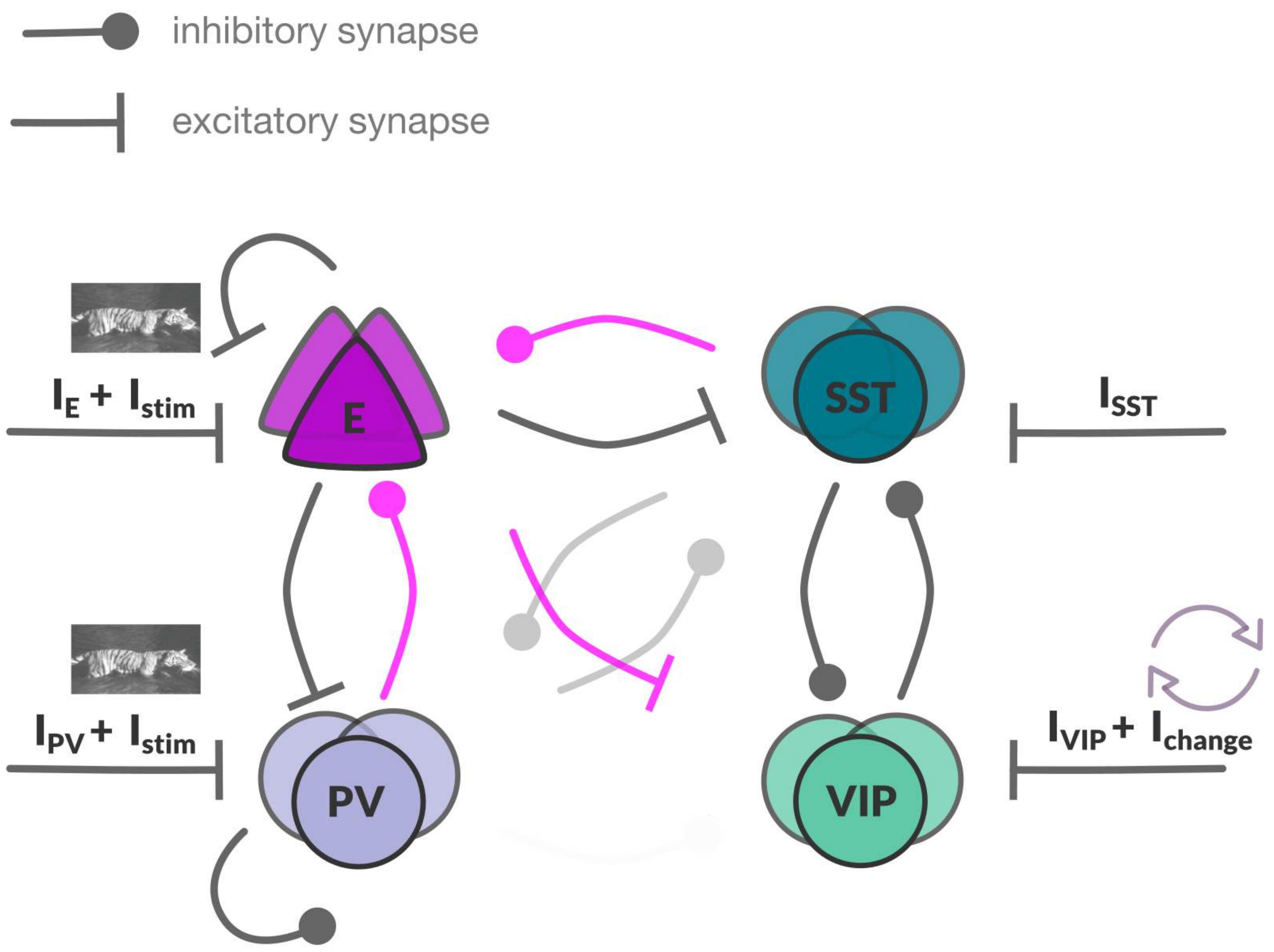
- multi-neuron **rate** model with experimentally observed:
  - ✦ relative counts [Billeh et al., 2020]
  - ✦ connectivity statistics
    & short-term plasticity parameters
    [Campagnola et al., 2020]



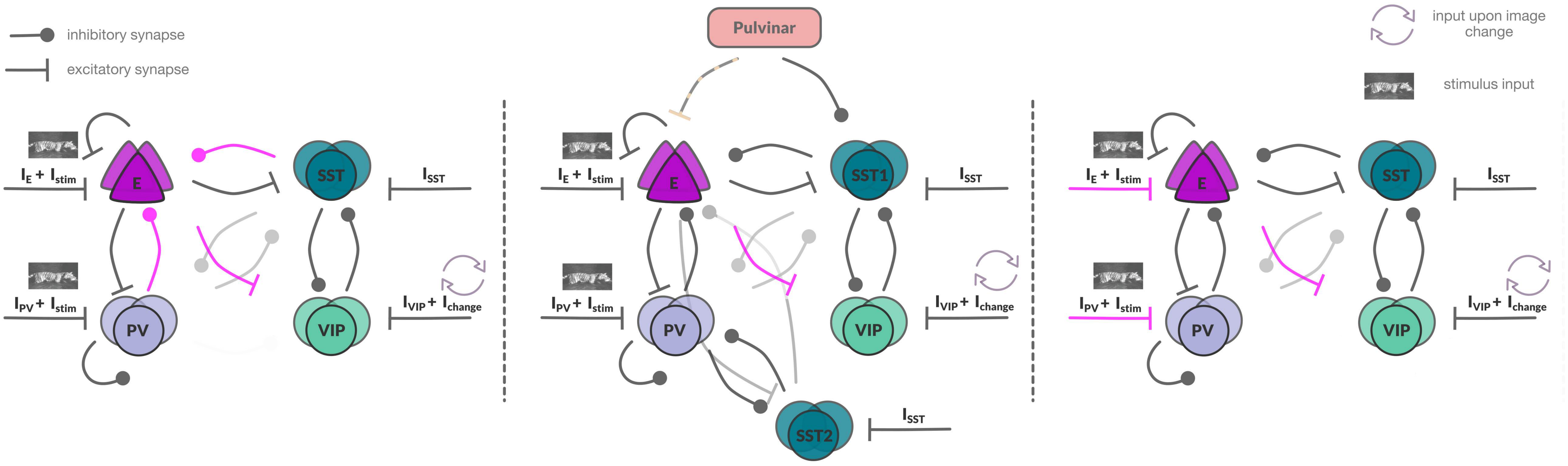


Microcircuit model

- multi-neuron **rate** model with experimentally observed
  - ✦ cell counts [Billeh et al., 2020]
  - ✦ connectivity statistics & short-term plasticity parameters [Campagnola et al., 2020]
- Inputs
  - ✦ Stimulus to E & PV
  - ✦ Background to set exper. baselines
  - ✦ VIP thalamic input on image change

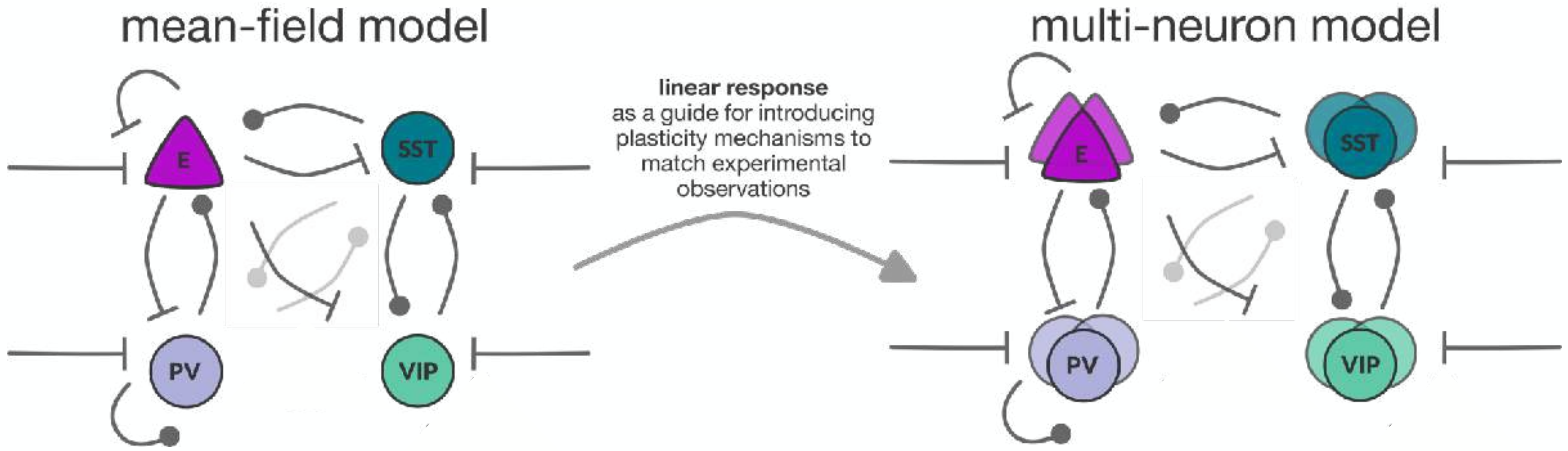


Potential sources of adaptation



Where should the adaptation happen?

- **Linear response** to
  - set strength of **external inputs**
  - reproduce **snapshots of responses**, i.e. before or after image change
- Start with observable network (3 cell types) and include PVs later



Linear response predicts system response to small perturbations

Linear response predicts system response to small perturbations

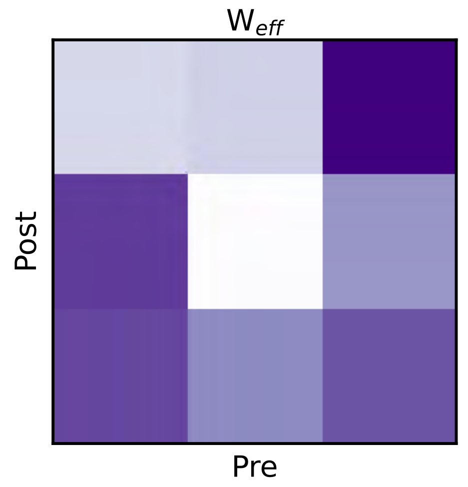
response $=$ sensitivity $\times$ input perturbation

Linear response predicts system response to small perturbations

response = sensitivity × input perturbation

Linear response predicts system response to small perturbations

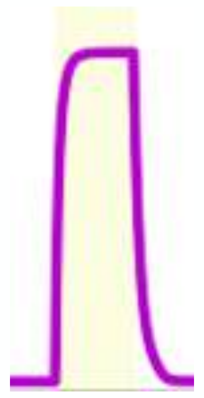
response = **sensitivity** × **input perturbation**

$\Big[\; I \;-$ 

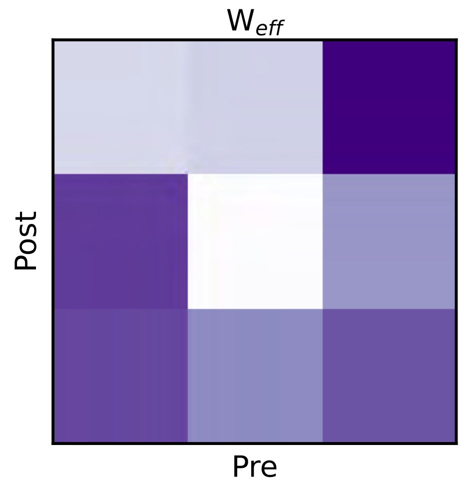
$\Big]^{-1}_{i,j}$

Linear response predicts system response to small perturbations

response = **sensitivity** × **input perturbation**



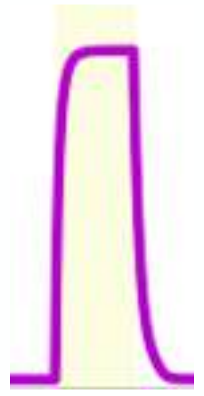
$[\, I -$



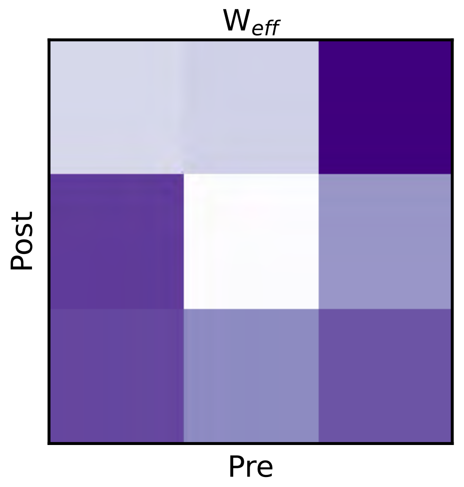
$]^{-1}_{i,j}$

Linear response predicts system response to small perturbations

response = **sensitivity** × **input perturbation**



[I -

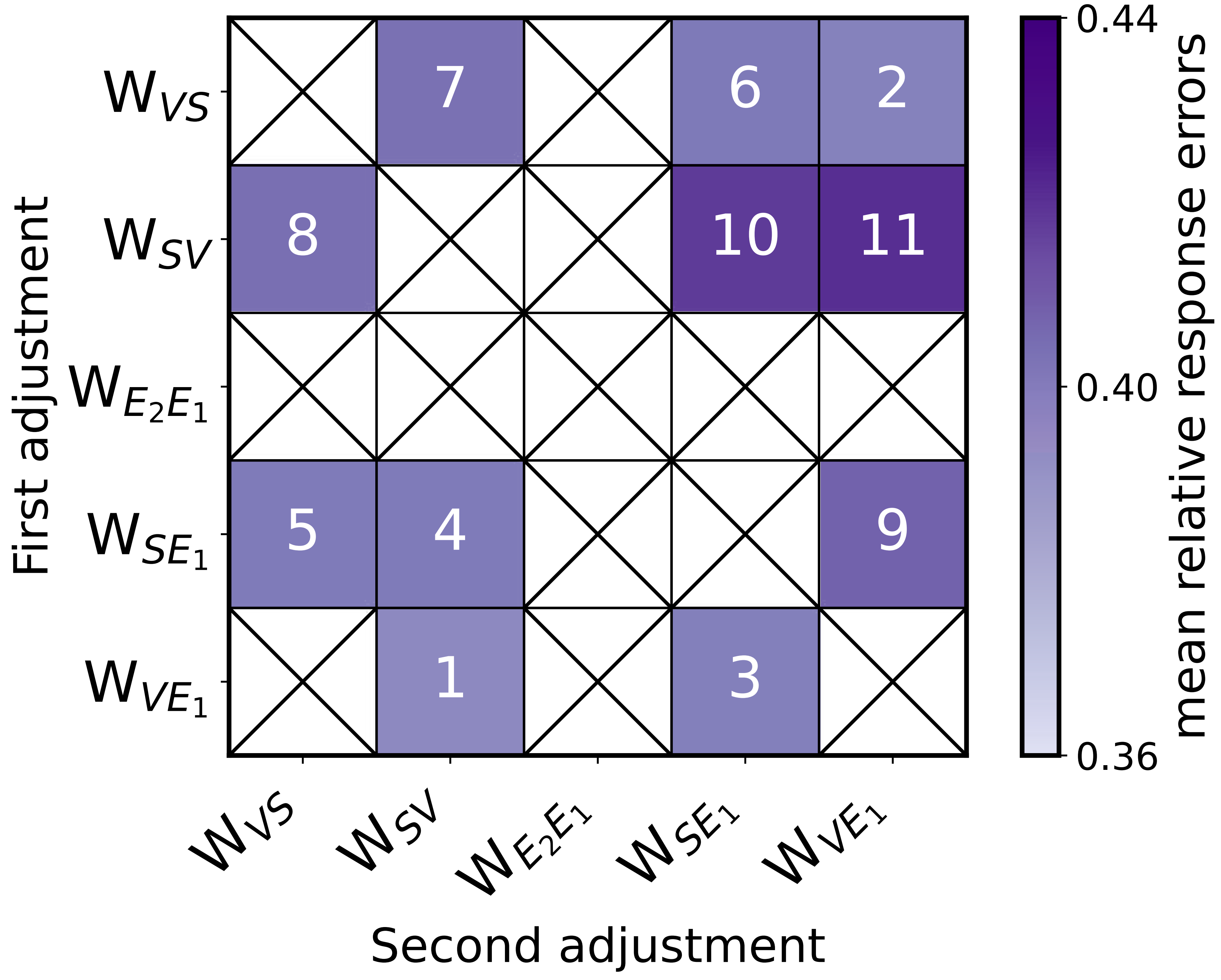


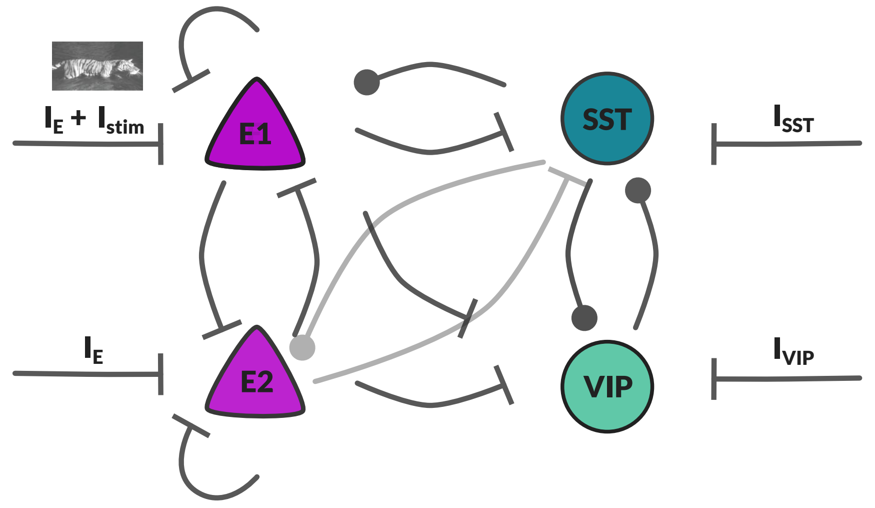
$]^{-1}_{i,j}$

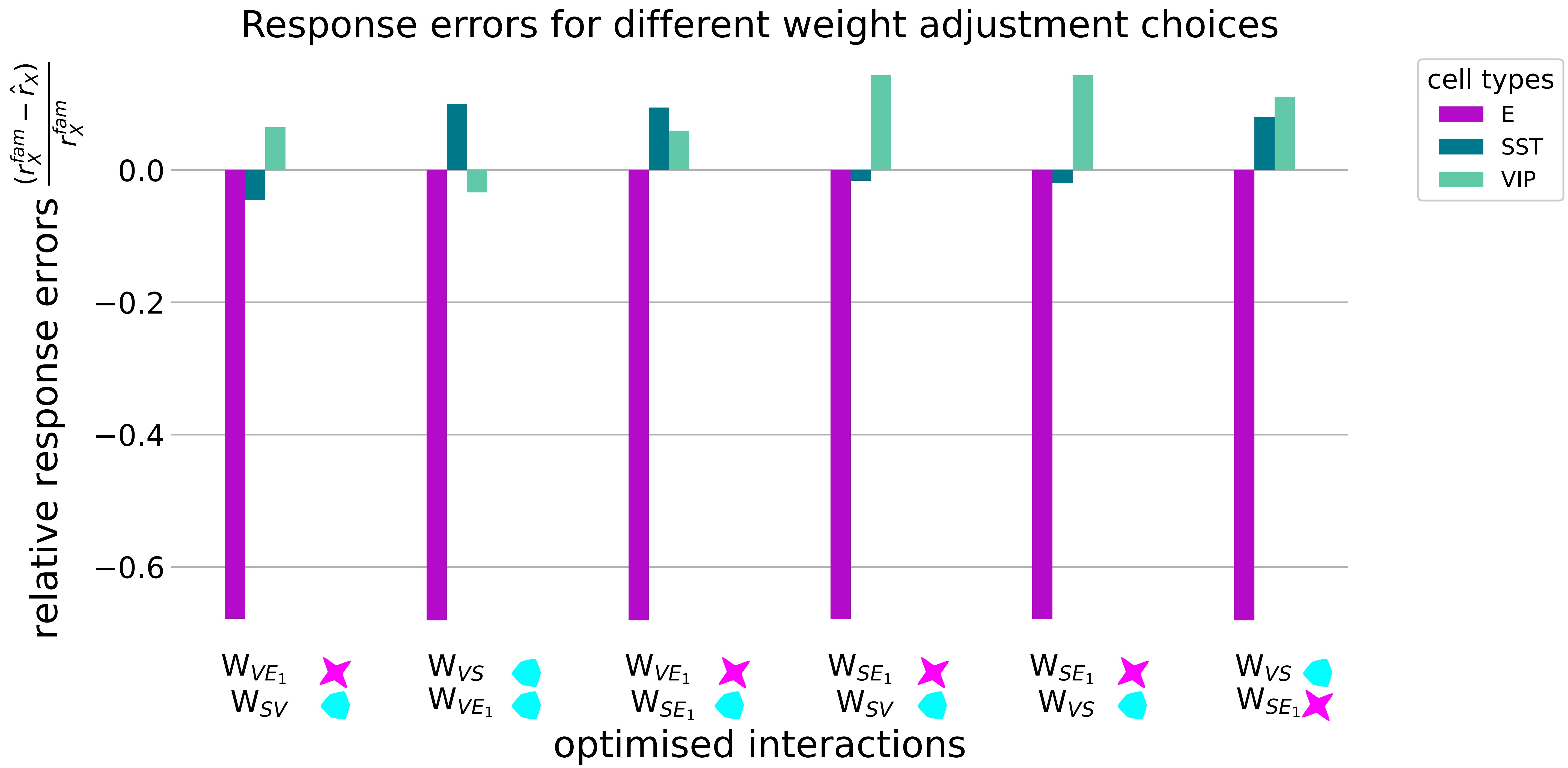
total response = Sum of (**sensitivity** × **single perturbations**)

Identify adjustments for adaptation

Modifying multiple interactions

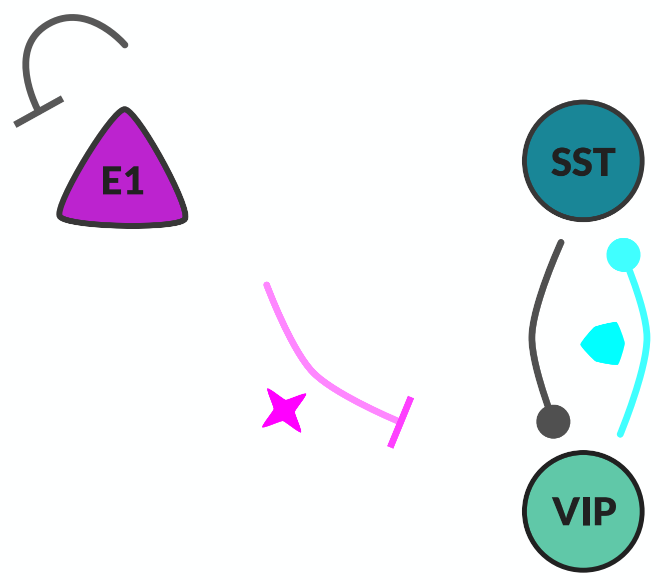




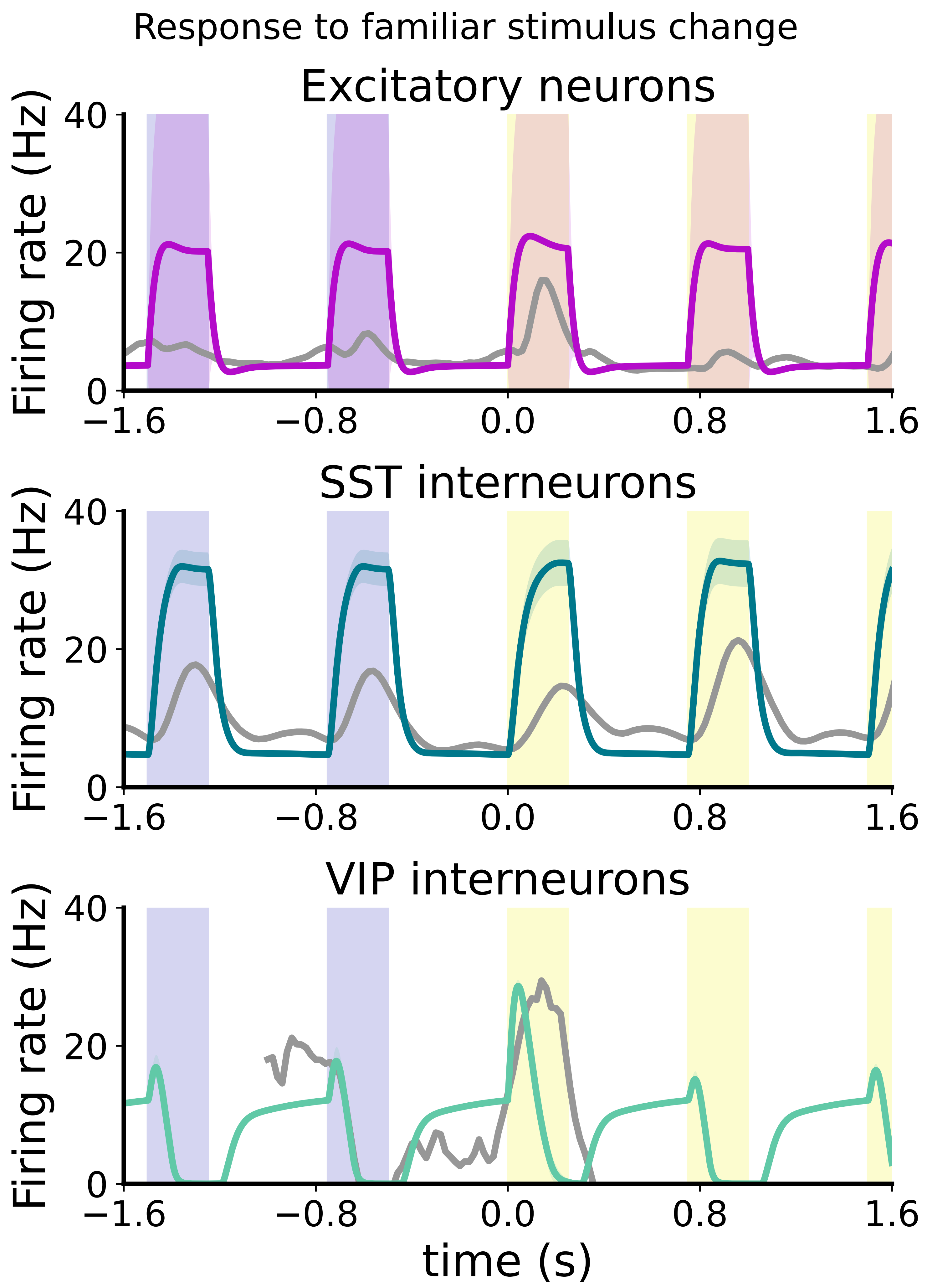


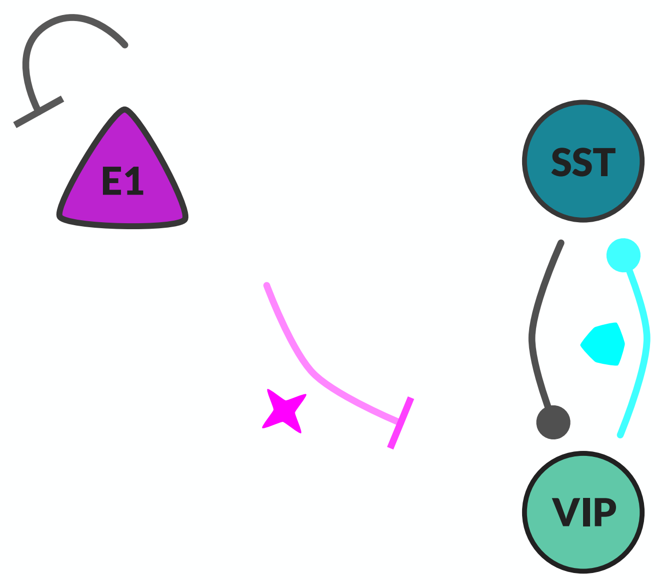
Anti-Hebbian plasticity on E-to-VIP synapses causes suppression of VIPs on stimulus onset

Anti-Hebbian plasticity on E-to-VIP synapses causes suppression of VIPs on stimulus onset

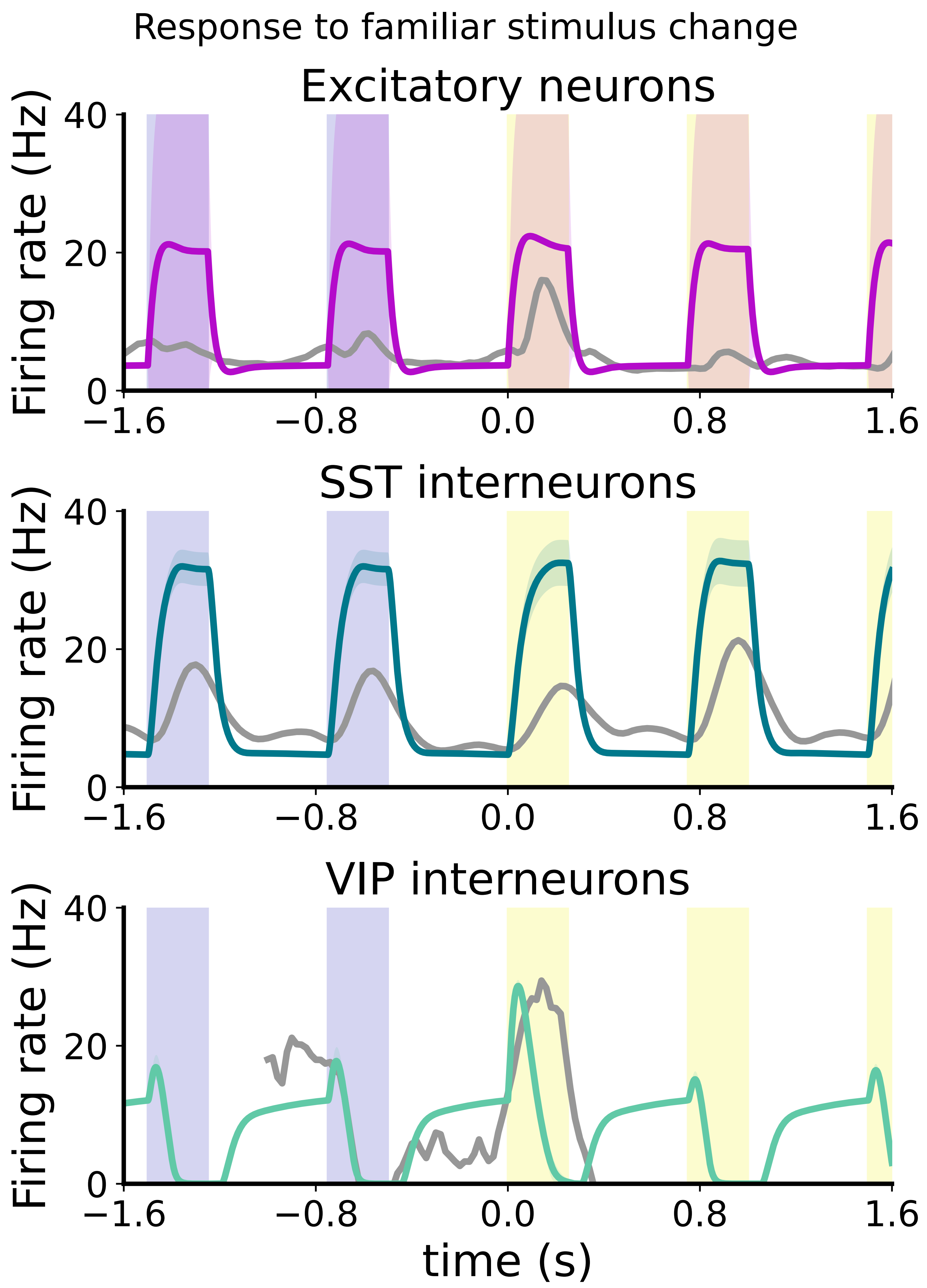


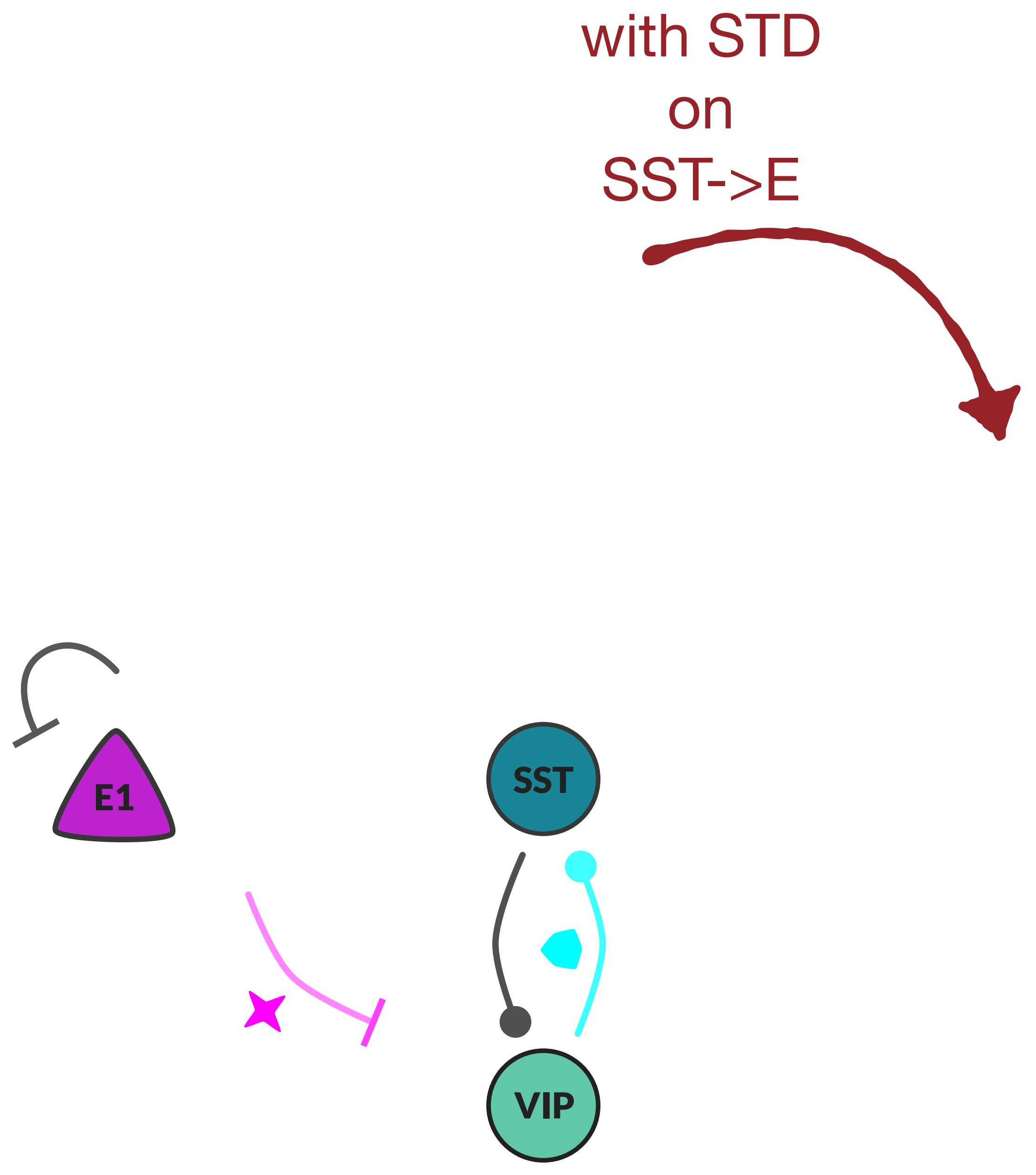
Anti-Hebbian plasticity on E-to-VIP synapses causes suppression of VIPs on stimulus onset



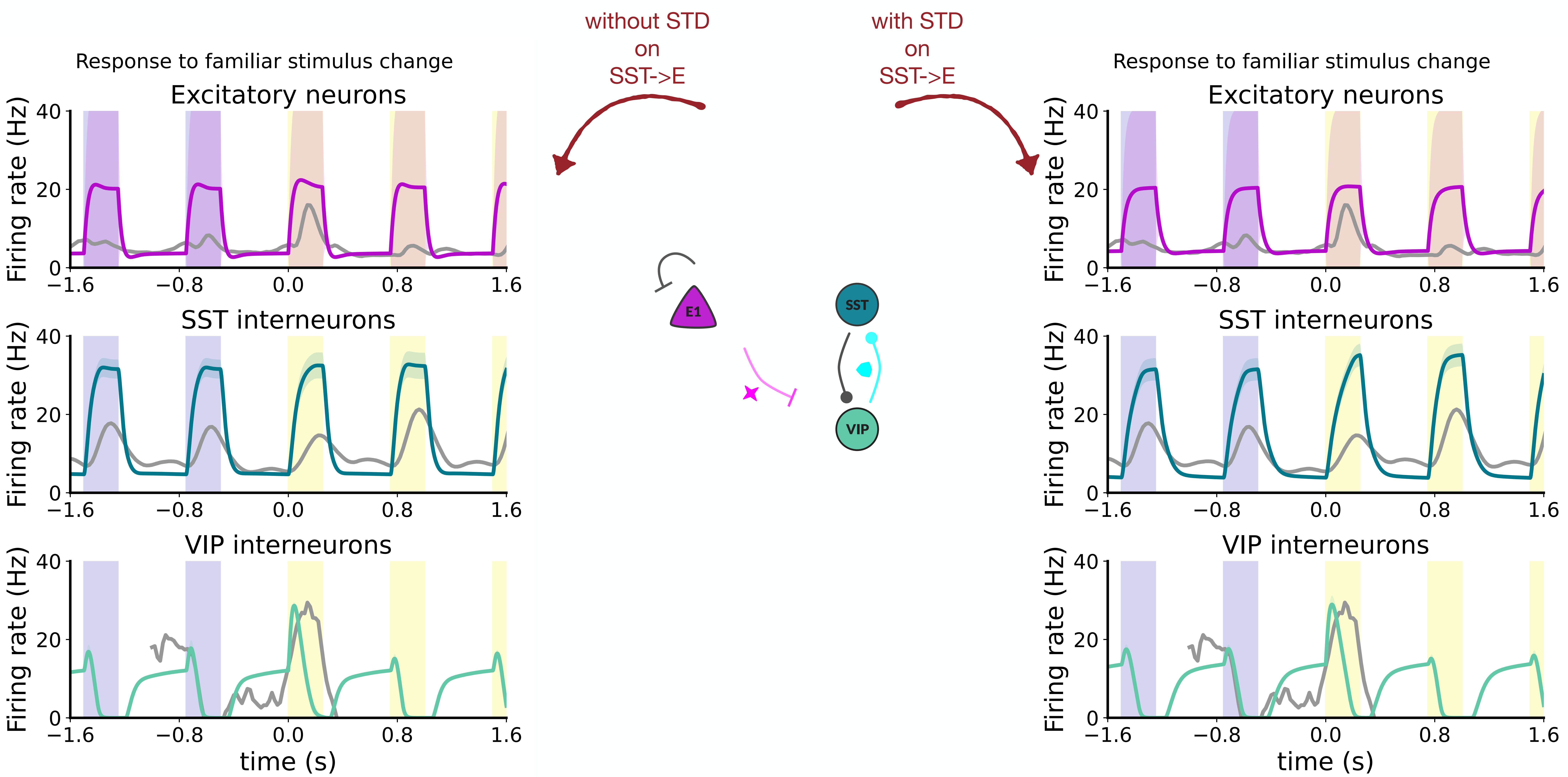


Anti-Hebbian plasticity on E-to-VIP synapses causes suppression of VIPs on stimulus onset

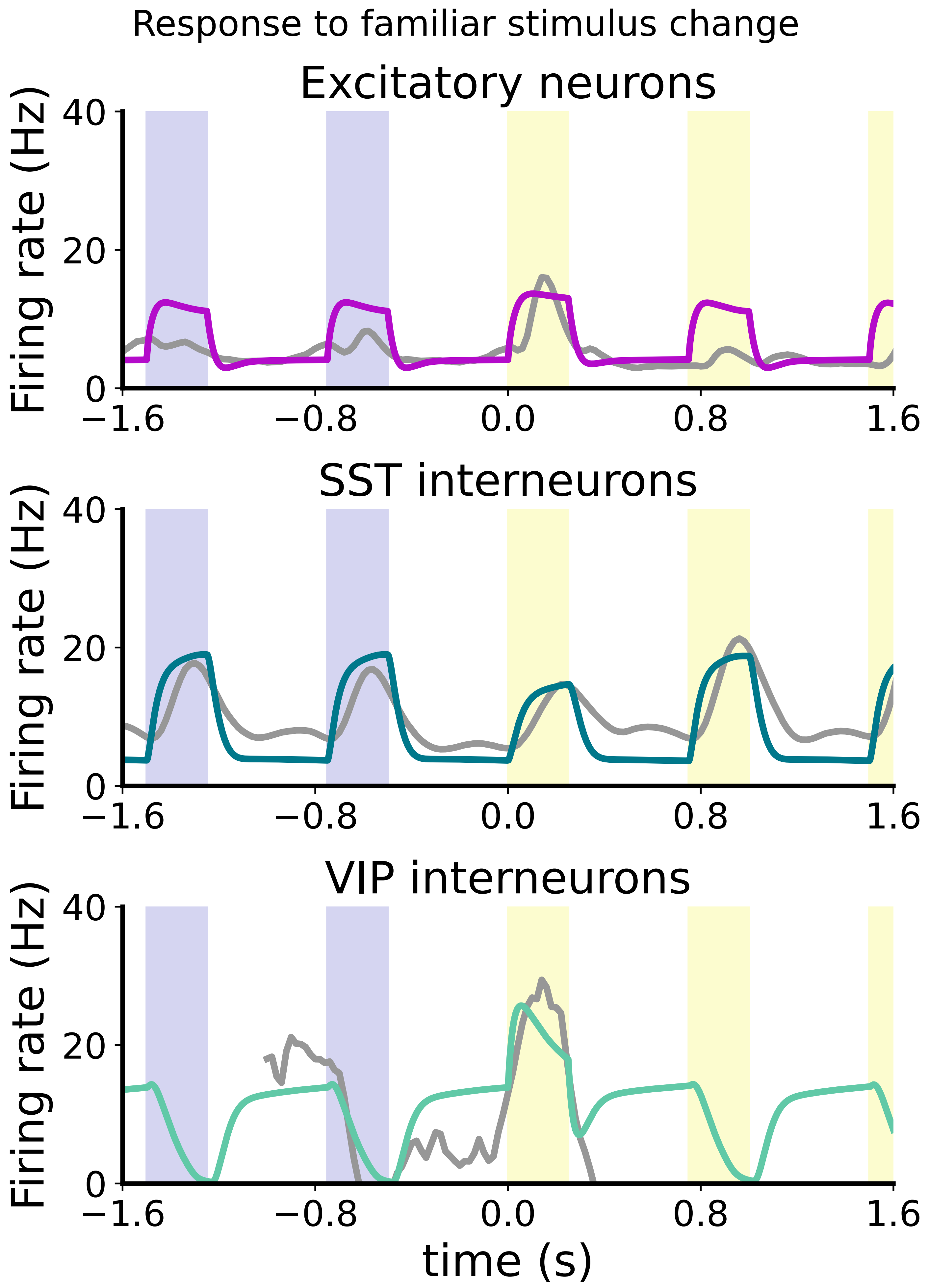


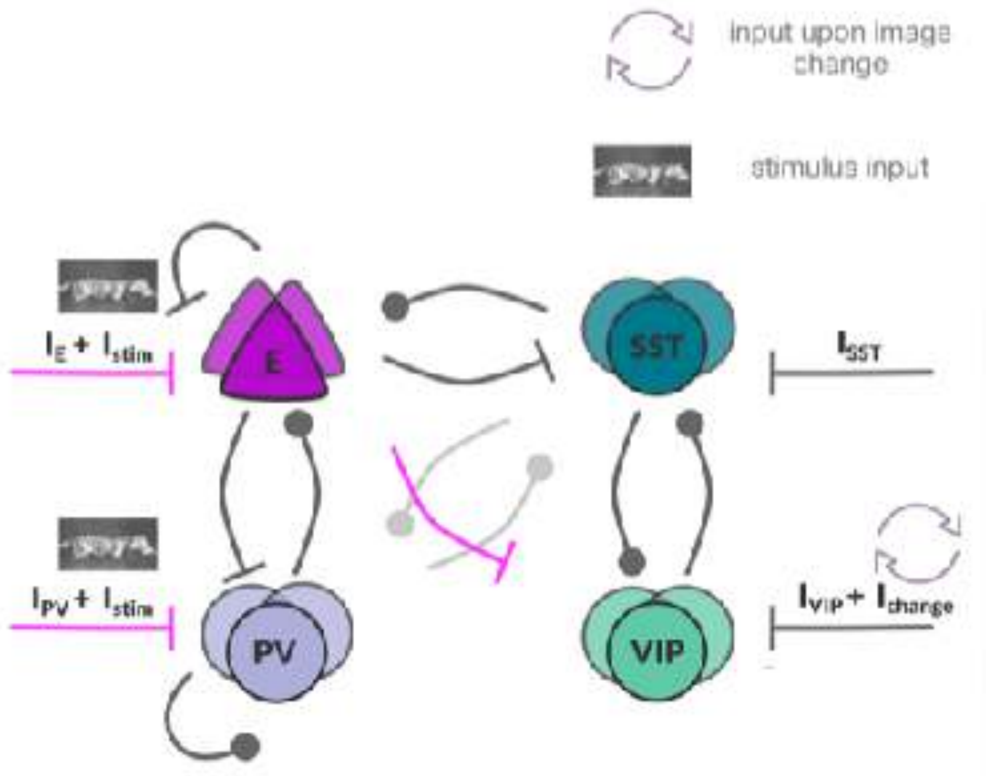


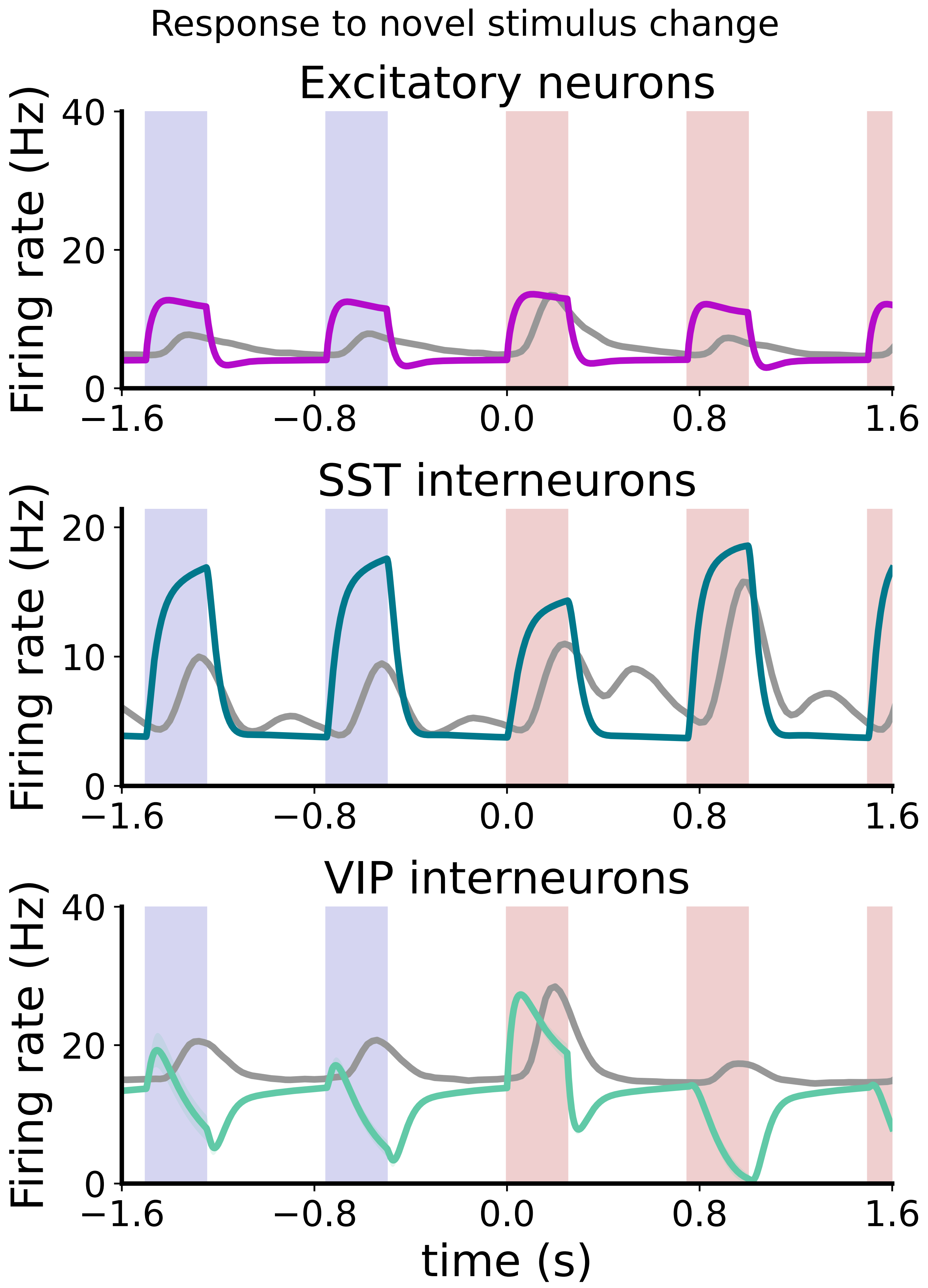
Anti-Hebbian plasticity on E-to-VIP synapses causes suppression of VIPs on stimulus onset



Short term depression on L4-to-L2/3 synapses reduces the excitatory drive on E cells

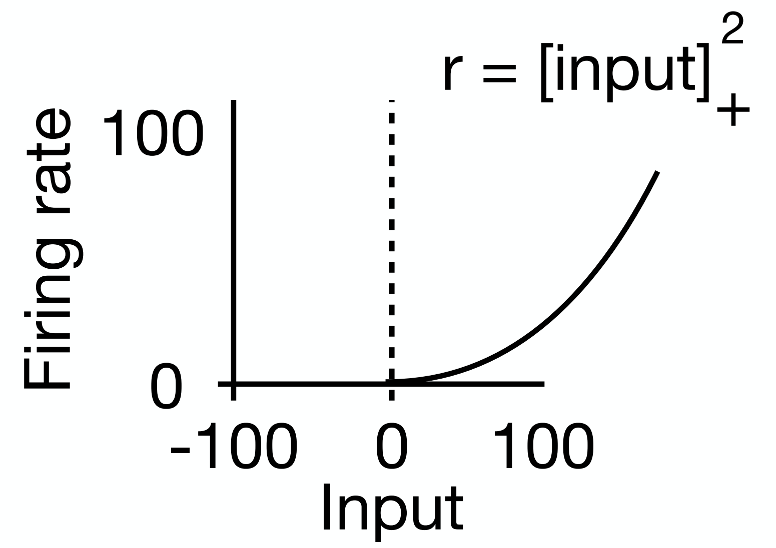


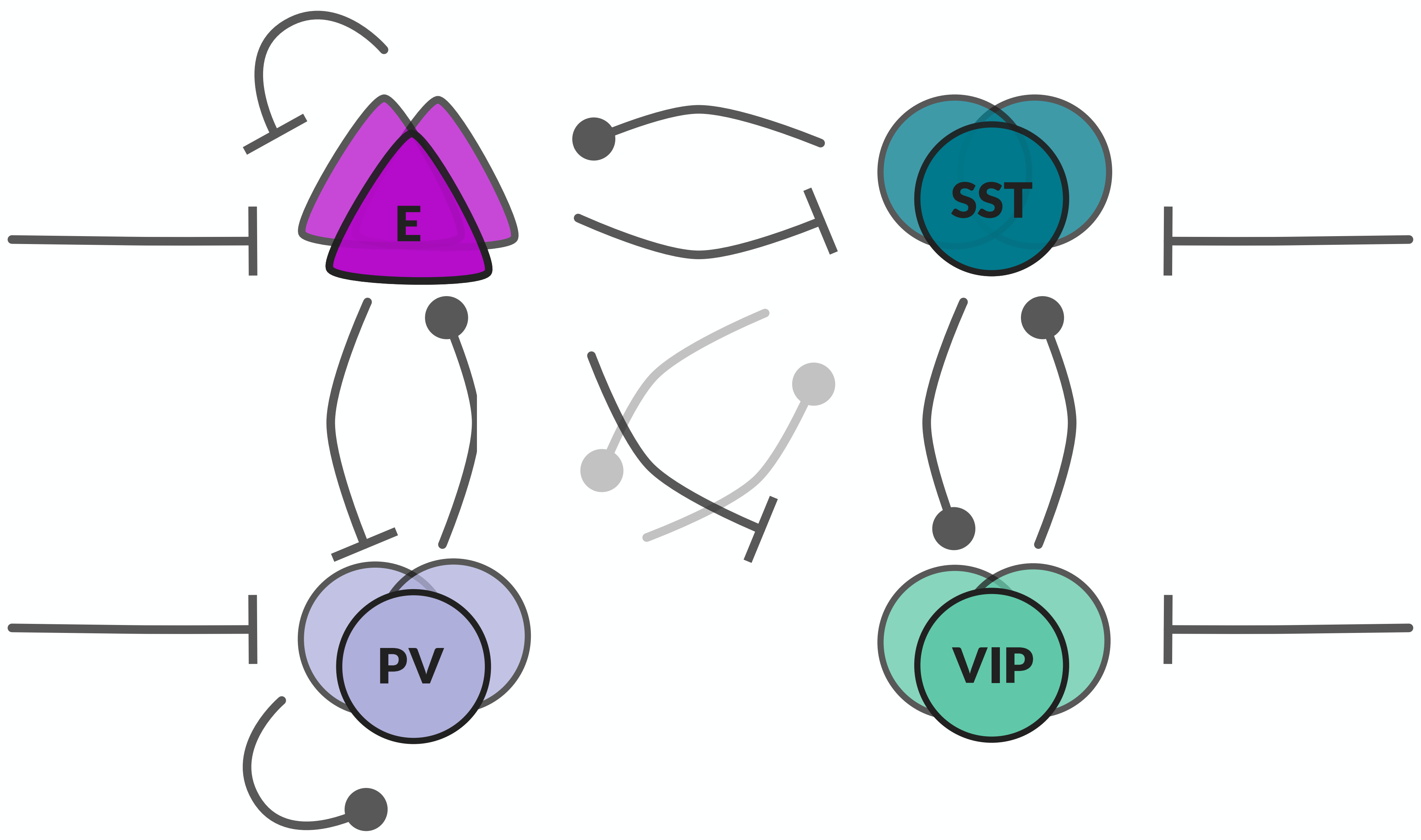




Microcircuit model

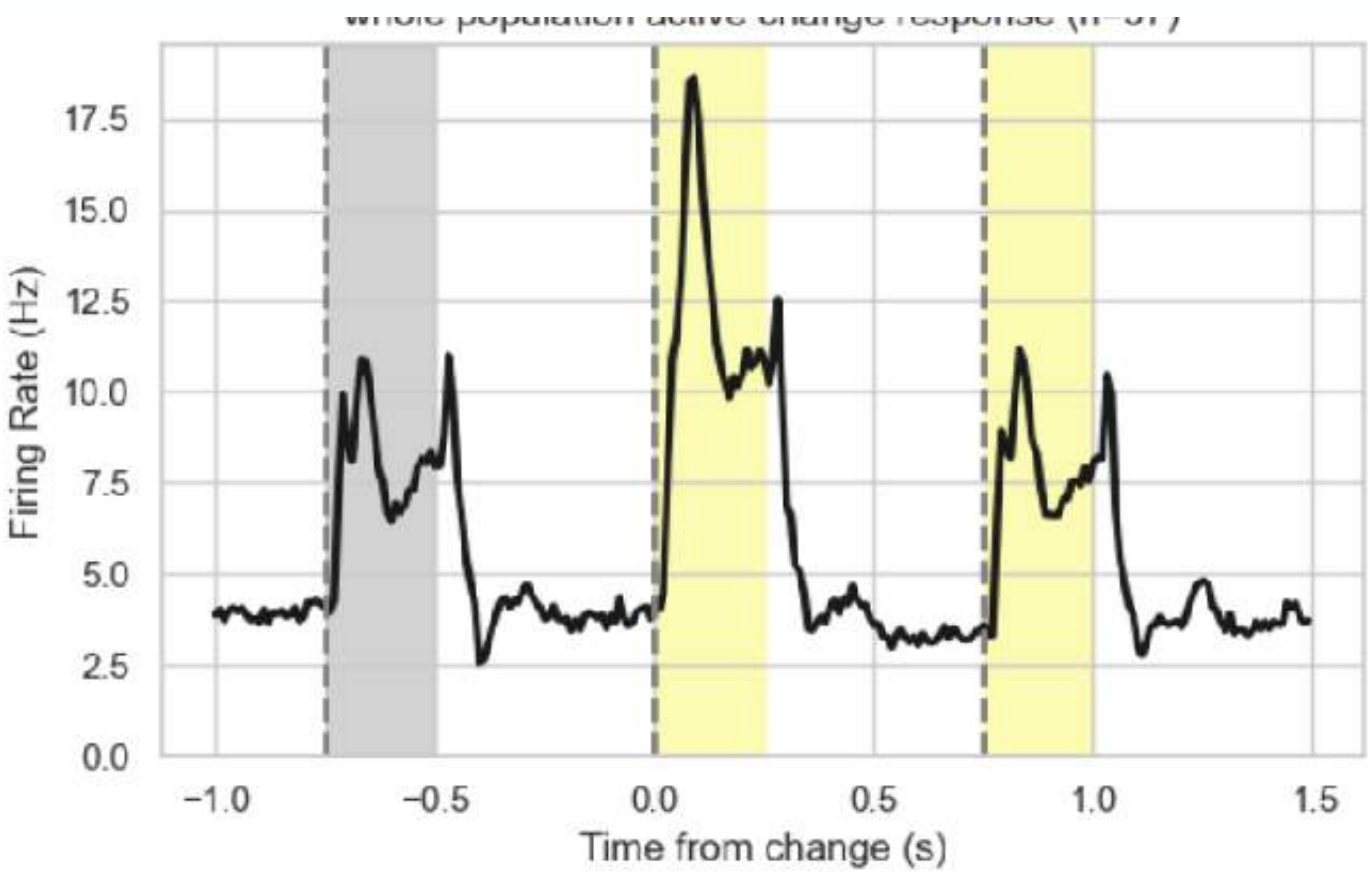
- multi-neuron **rate** model with experimentally observed:
  - ✦ relative counts [Billeh et al., 2020]
  - ✦ connectivity statistics
    & short-term plasticity parameters
    [Campagnola et al., 2020]
  - ✦ expansive nonlinearity



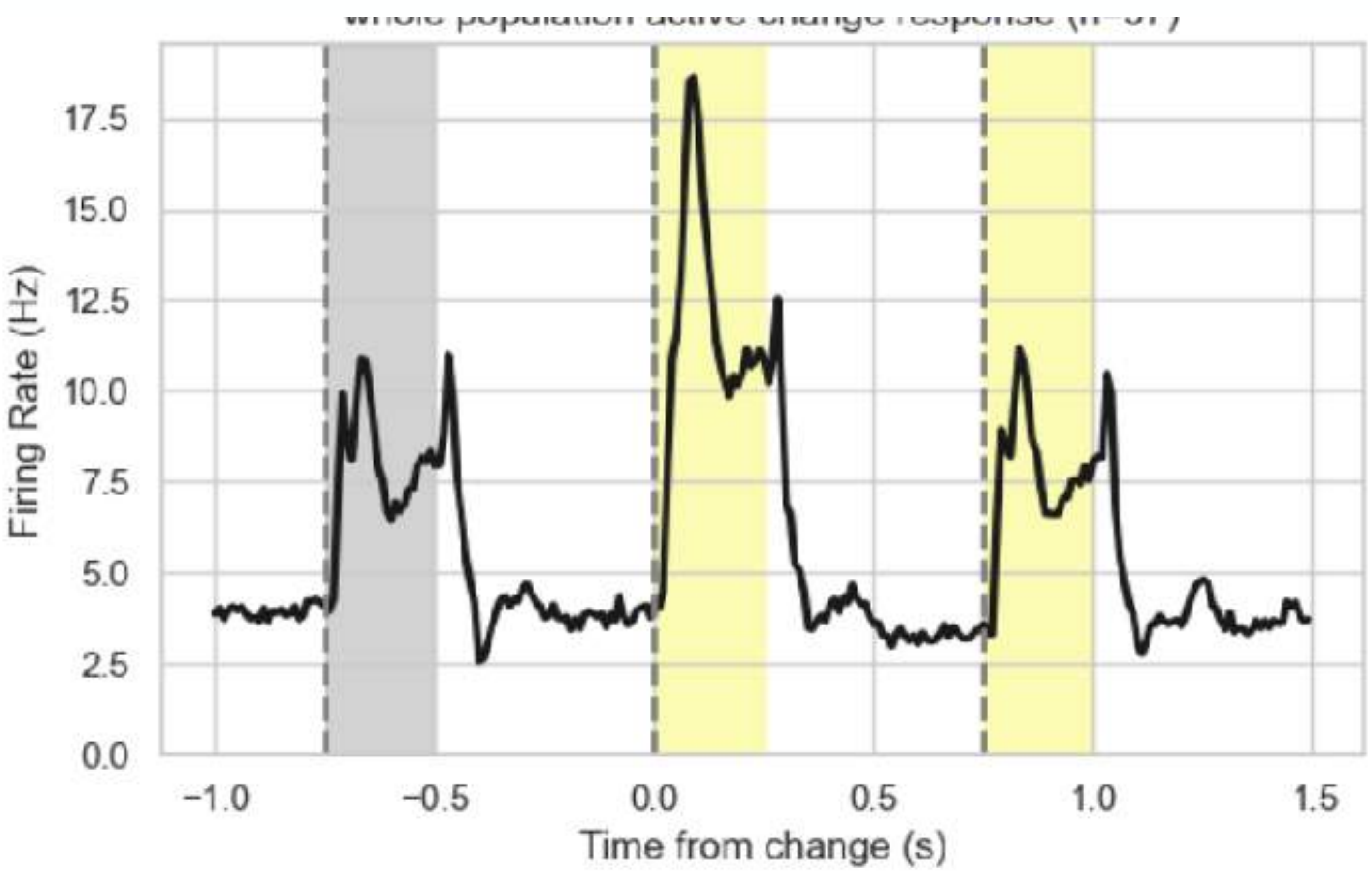


Response profiles before and after change

Familiar

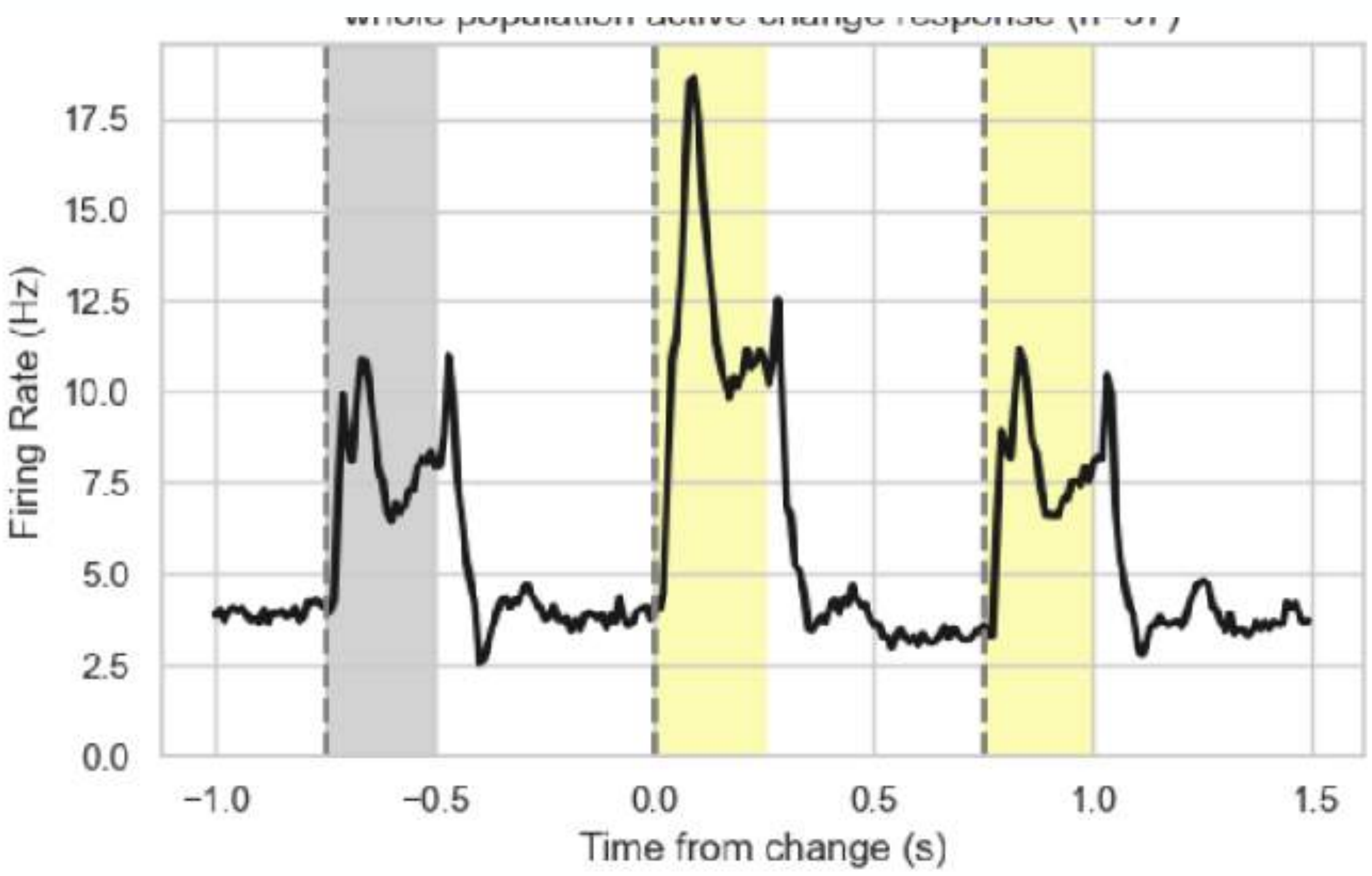


Novel



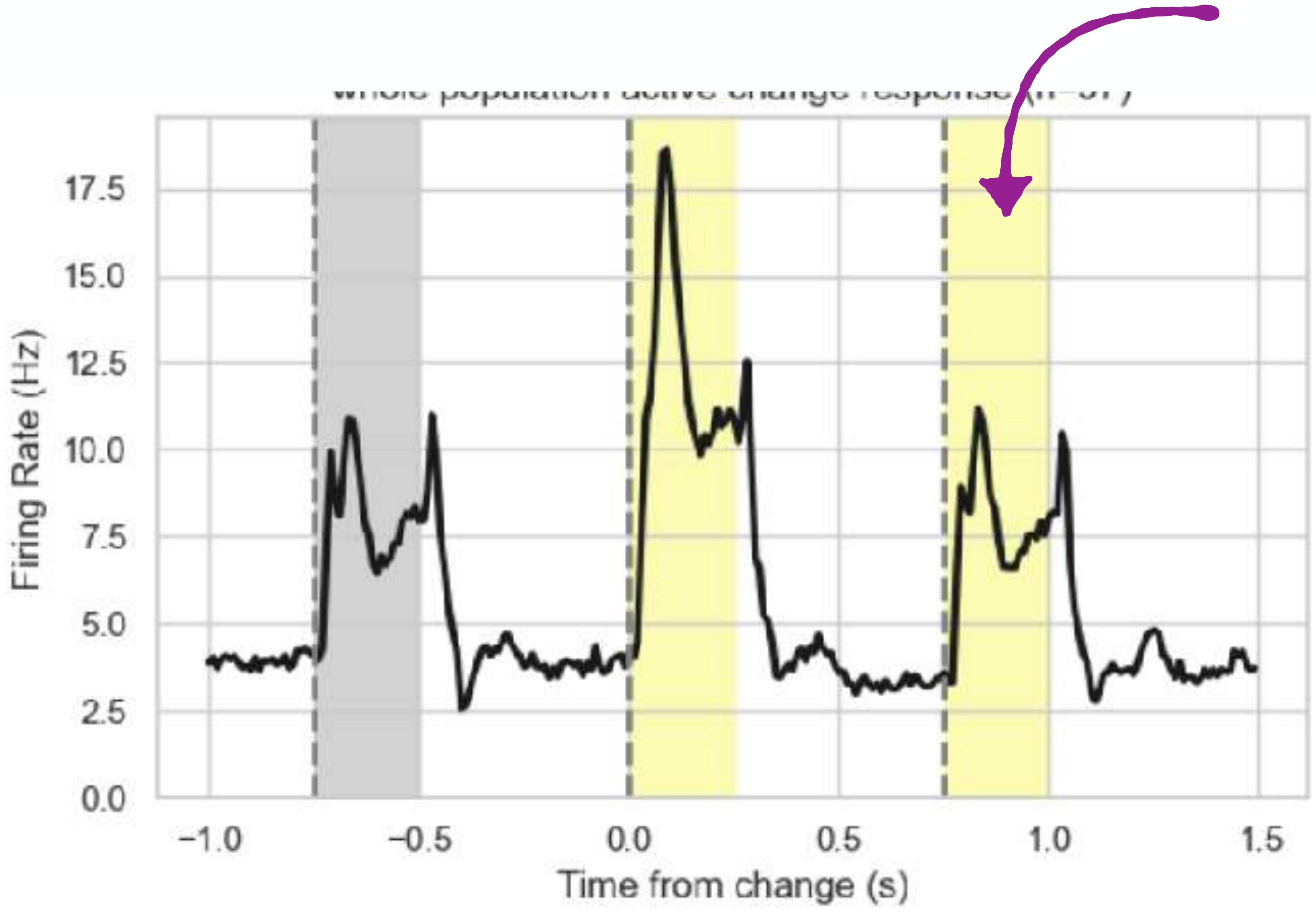
Response profiles before and after change

Familiar



Novel

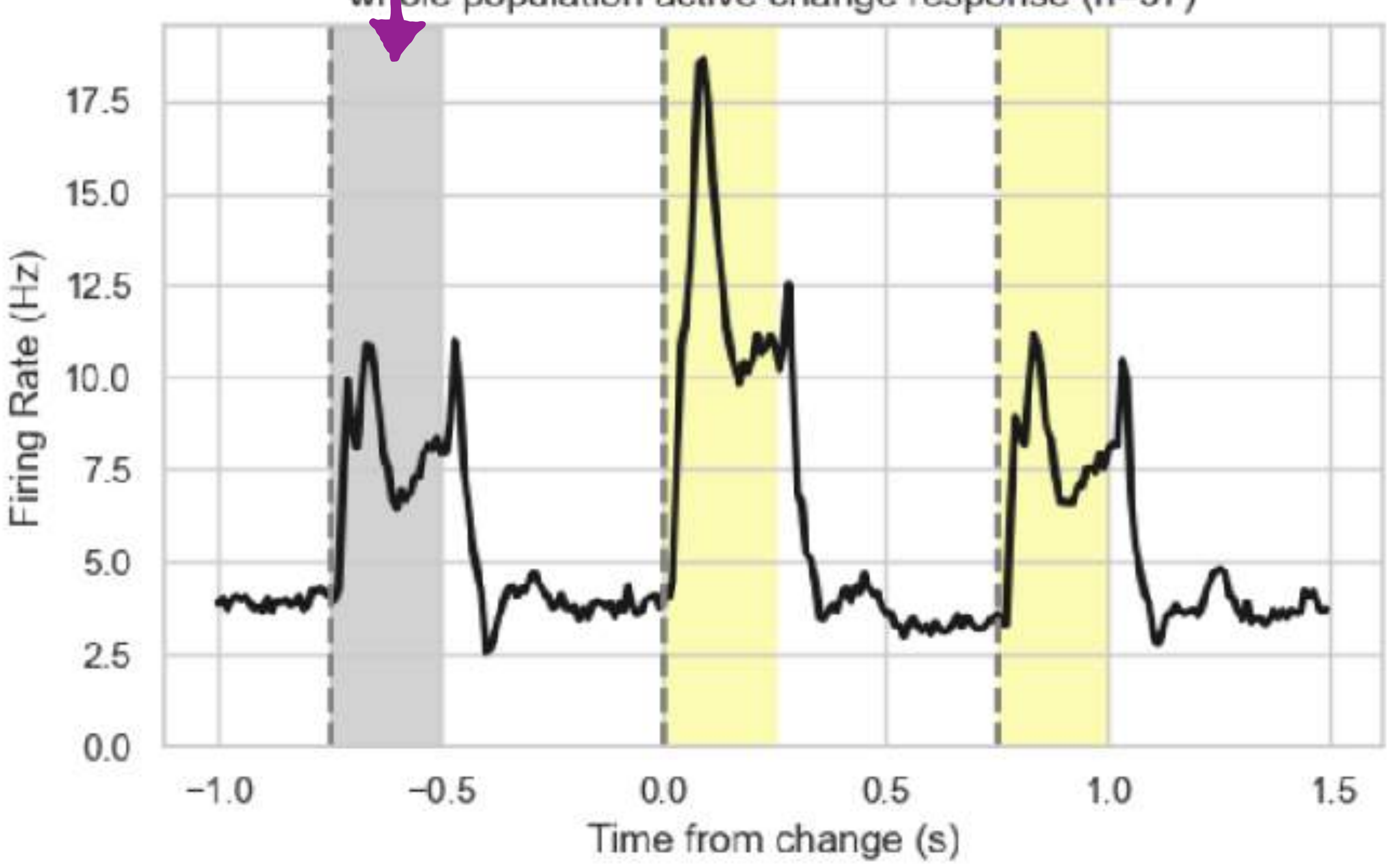
Response of naive network



Response profiles before and after change

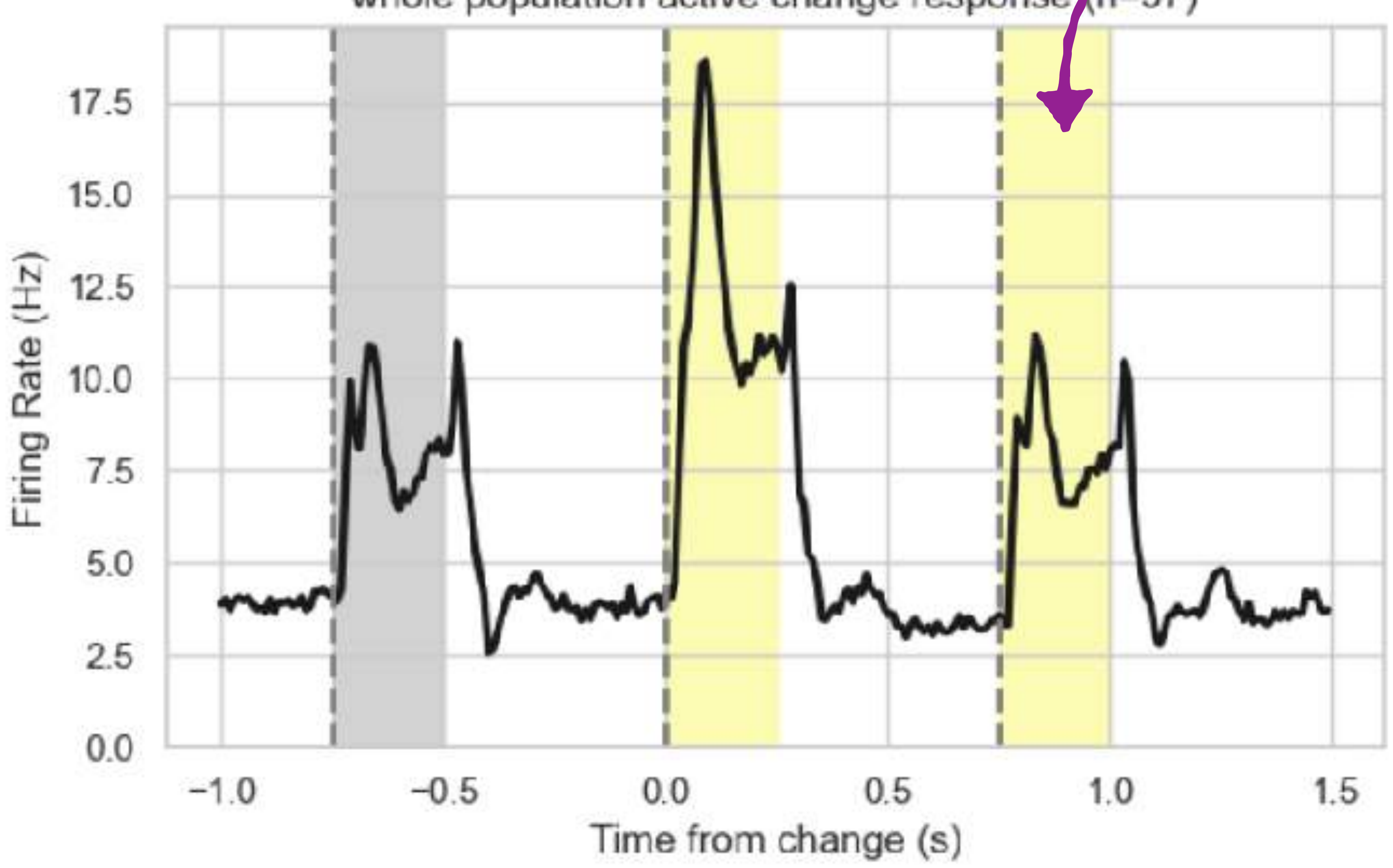
Familiar

Adapted response



Novel

Response of naive network

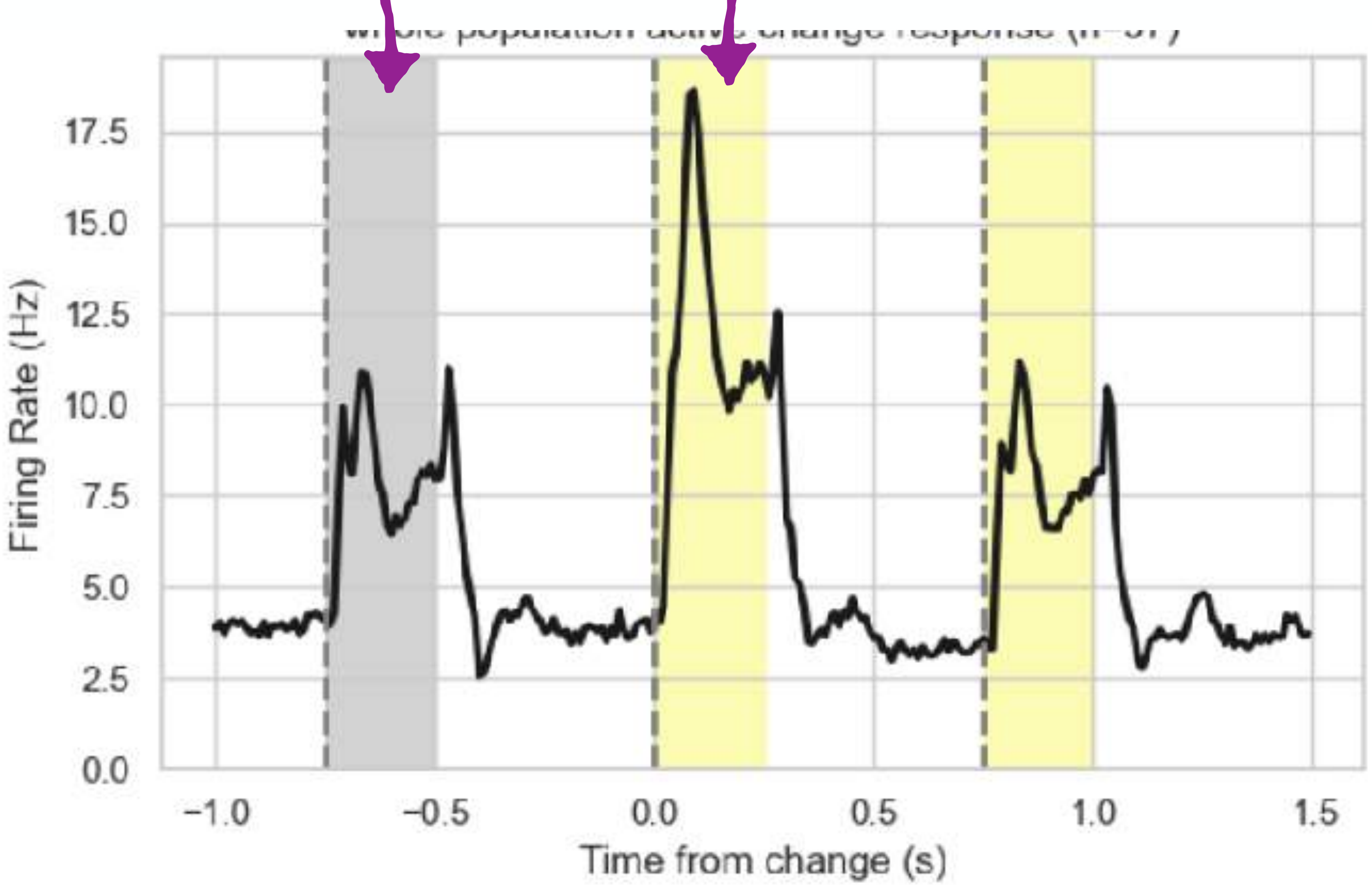


Response profiles before and after change

Familiar

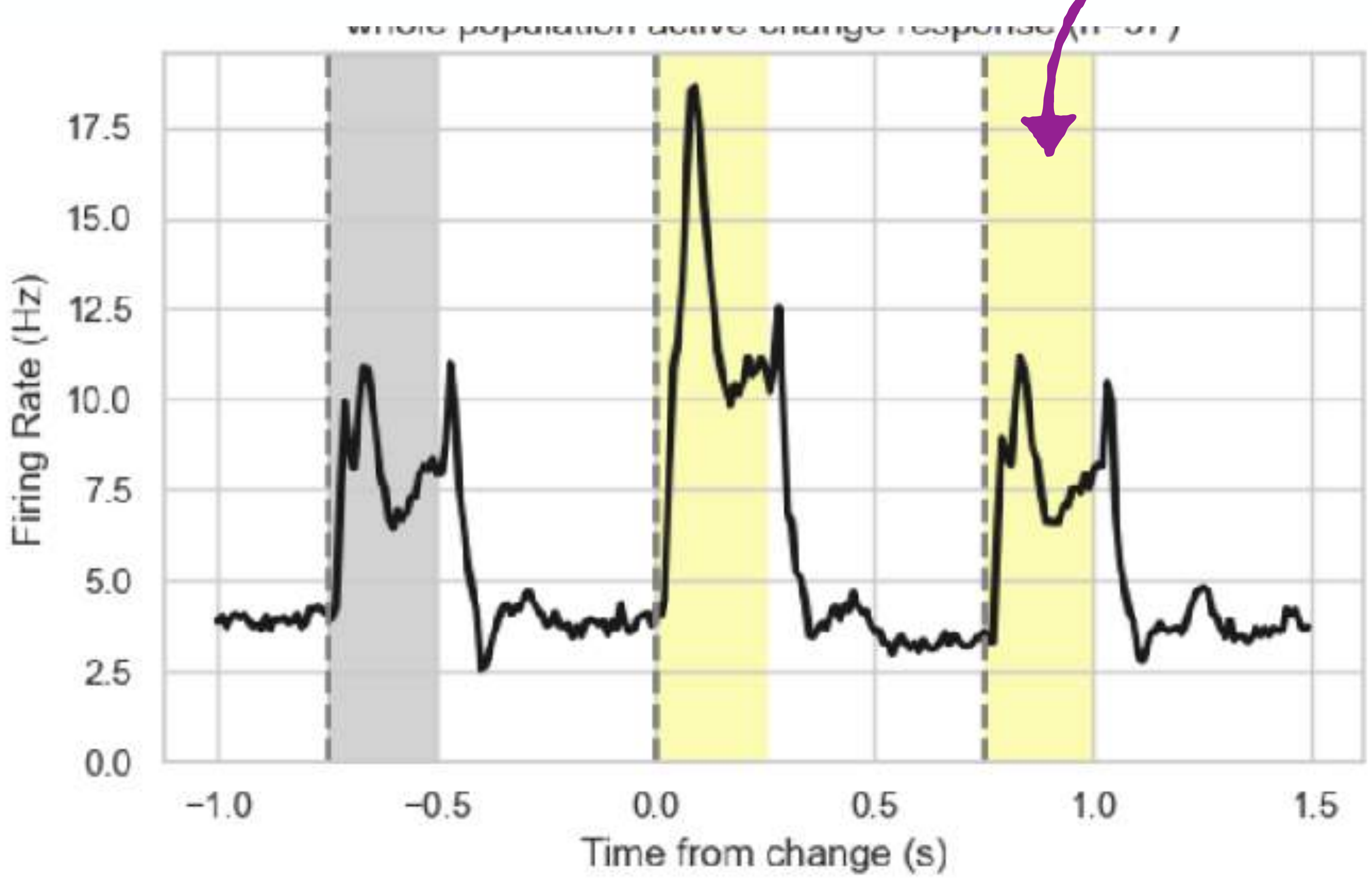
Adapted response

For VIP input estimation



Novel

Response of naive network

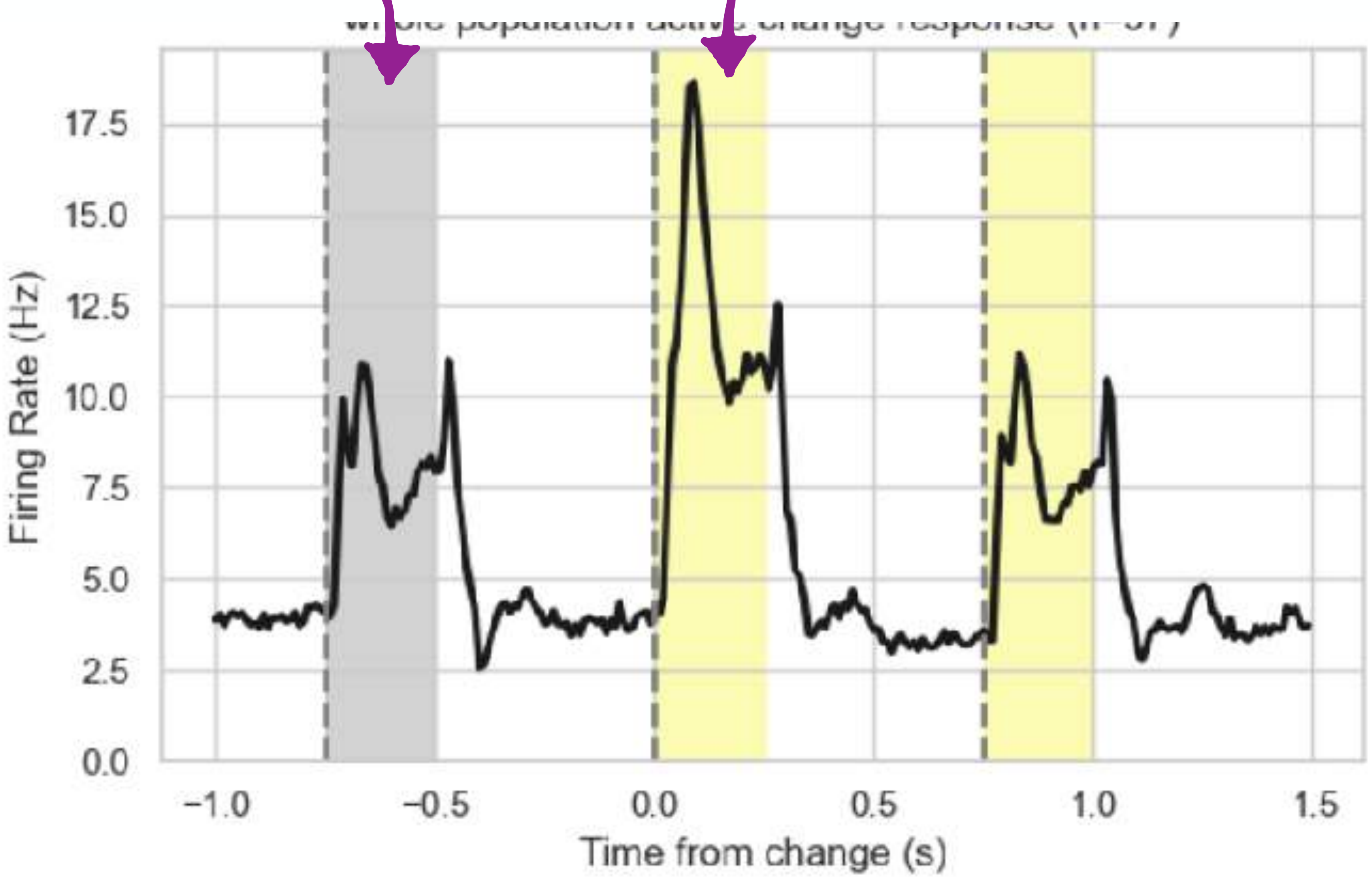


Response profiles before and after change

Familiar

Adapted response

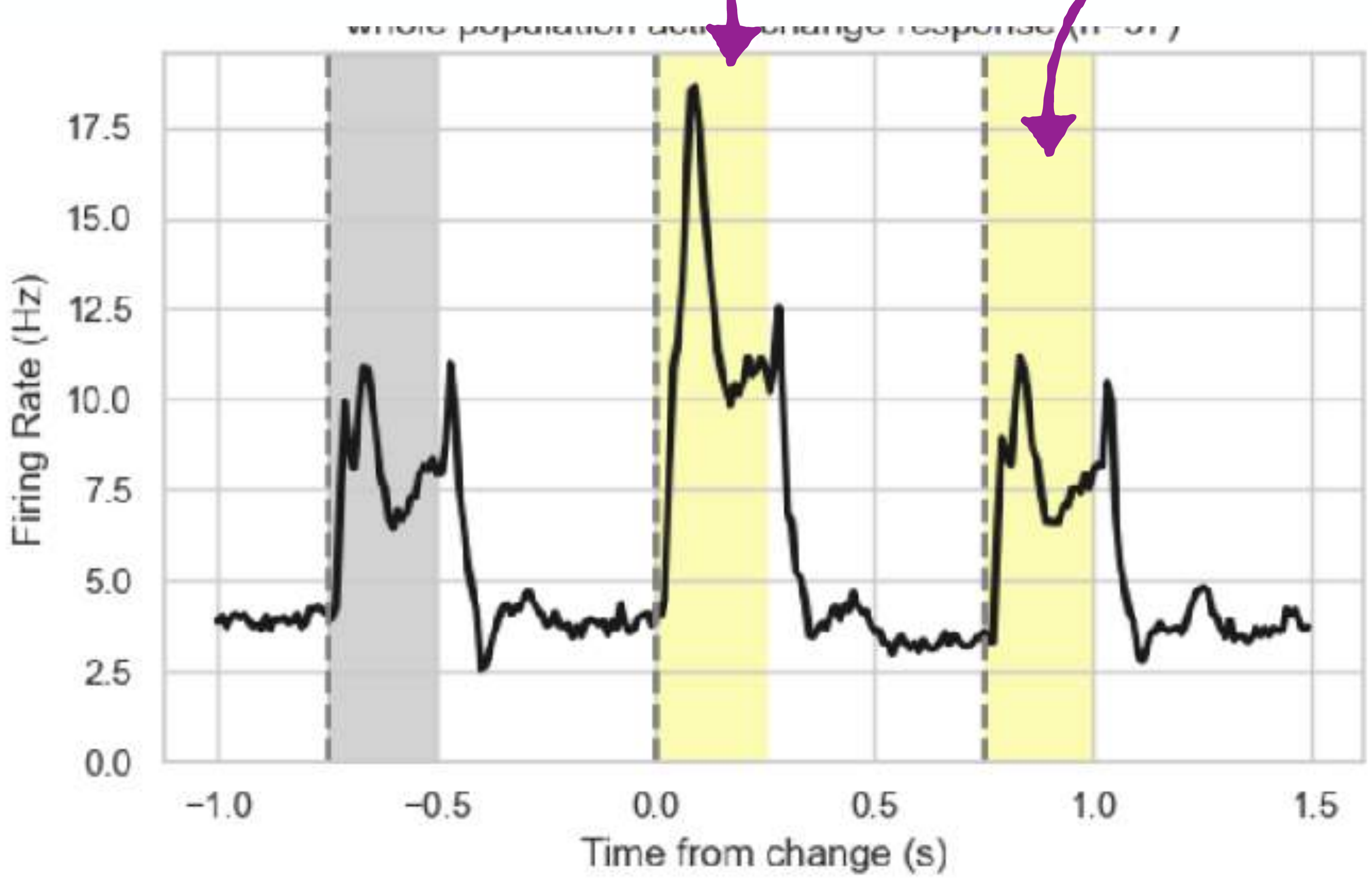
For VIP input estimation



Novel

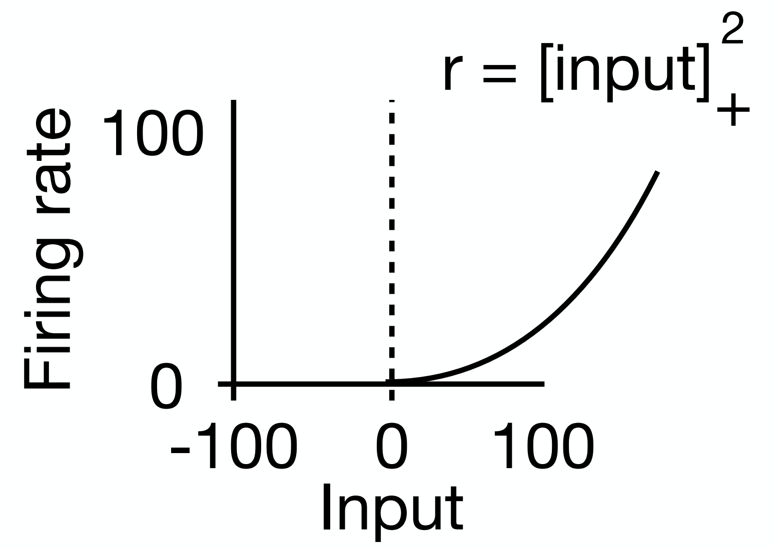
For validation

Response of naive network



Microcircuit model

- multi-neuron **rate** model with experimentally observed:
  - relative counts [Billeh et al., 2020]
  - connectivity statistics
    & short-term plasticity parameters
    [Campagnola et al., 2020]
  - expansive nonlinearity



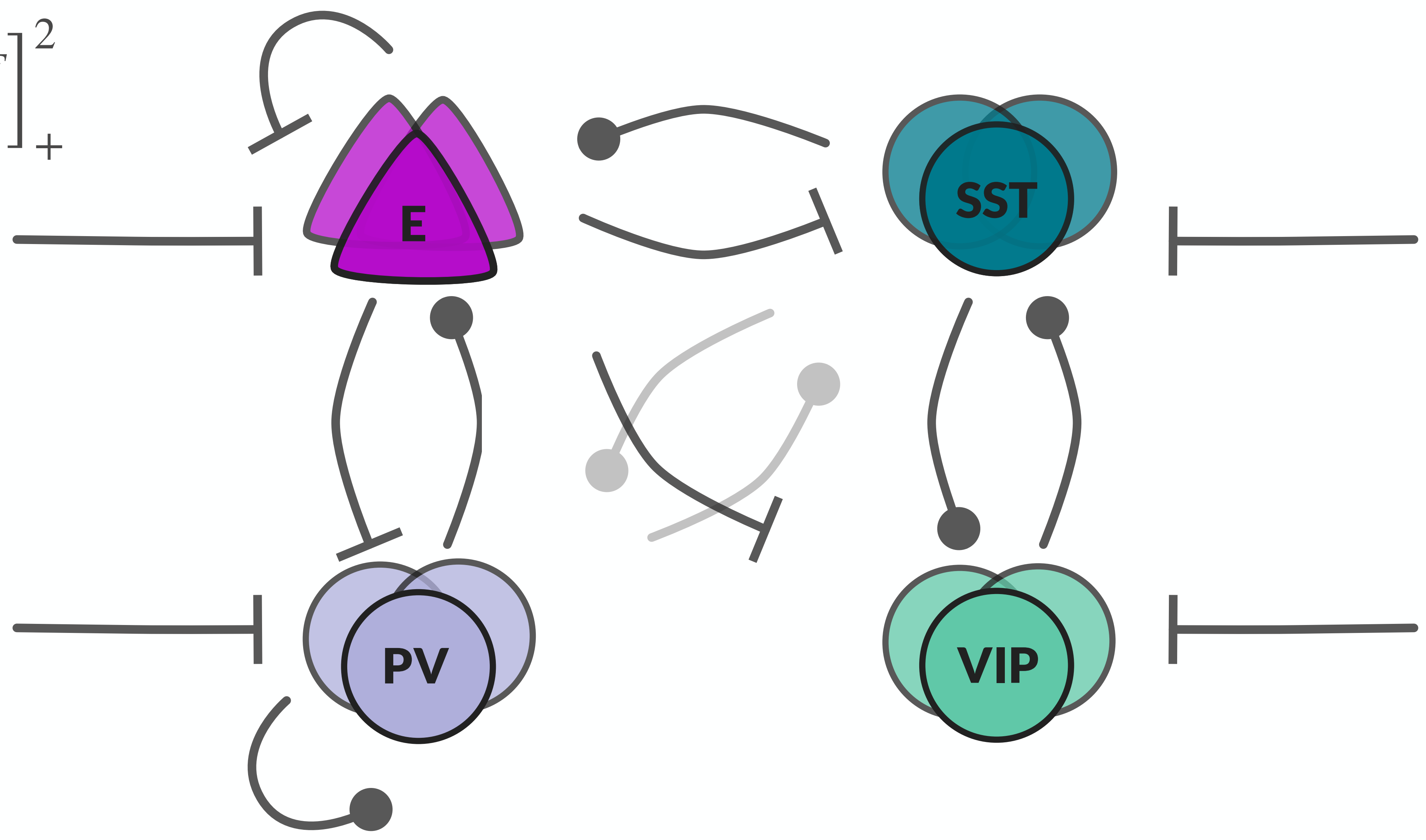
$$\tau_E \frac{\mathrm{d}r_E}{\mathrm{d}t} = -r_E + \left[w_{EE}r_E - w_{EP}r_P - w_{ES1}r_{S1} - w_{ES2}r_{S2} + I_E^B + I_E^{\text{stim}} + H^F\right]_+^2$$

$$\tau_P \frac{\mathrm{d}r_P}{\mathrm{d}t} = -r_P + \left[w_{PE}r_E - w_{PP}r_P - w_{PS2}r_{S2} + I_P^B + I_P^{\text{stim}}\right]_+^2$$

$$\tau_{S1} \frac{\mathrm{d}r_{S1}}{\mathrm{d}t} = -r_{S1} + \left[w_{S1E}r_E - w_{S1V}r_V + I_{S1}^B - I^{\text{pulvinar}}\right]_+^2$$

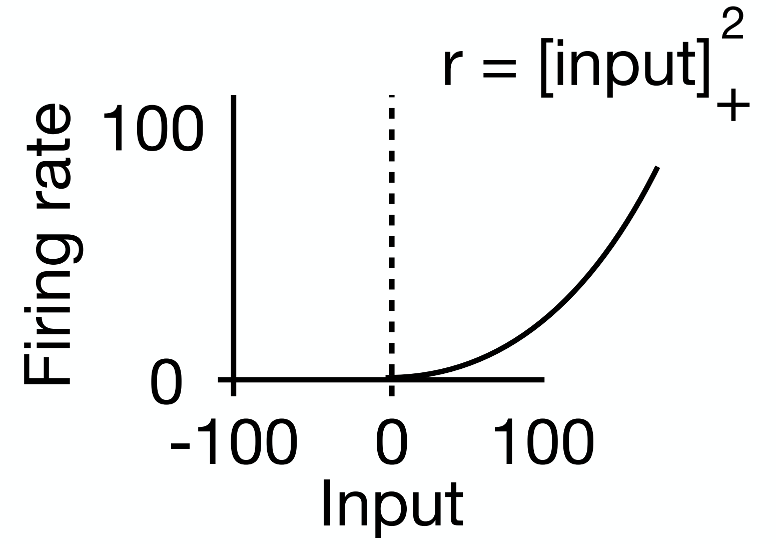
$$\tau_{S2} \frac{\mathrm{d}r_{S2}}{\mathrm{d}t} = -r_{S2} + \left[w_{S2E}r_E - w_{S2P}r_P + I_{S2}^B\right]_+^2$$

$$\tau_V \frac{\mathrm{d}r_V}{\mathrm{d}t} = -r_V + \left[w_{VE}r_E - w_{VS1}r_{S1} + I_V^B + I^{\text{kick}}\right]_+^2,$$



Microcircuit model

- multi-neuron **rate** model with experimentally observed:
  - ✦ relative counts [Billeh et al., 2020]
  - ✦ connectivity statistics
    & short-term plasticity parameters
    [Campagnola et al., 2020]
  - ✦ expansive nonlinearity



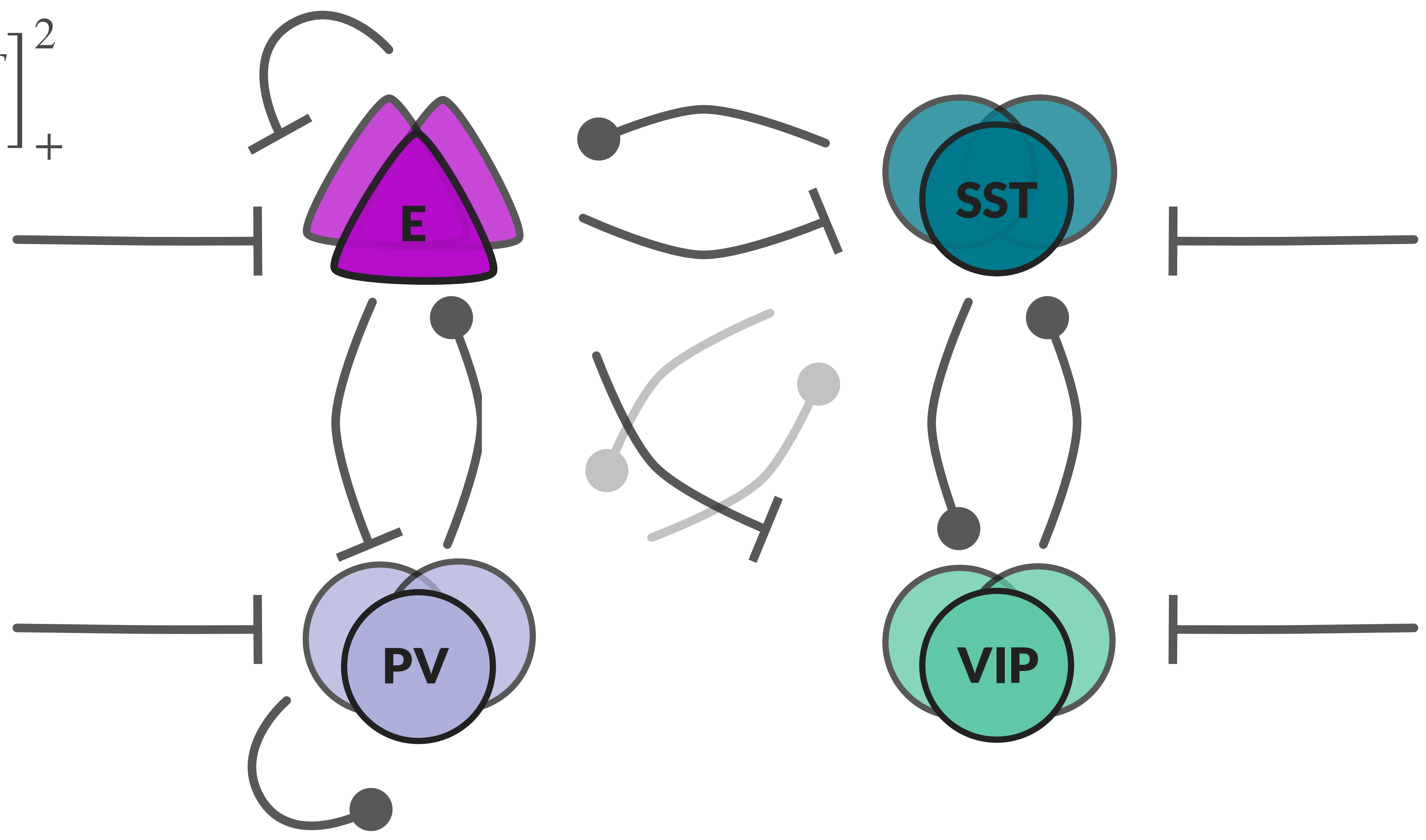
$$\tau_E \frac{\mathrm{d}r_E}{\mathrm{d}t} = -r_E + \left[w_{EE}r_E - w_{EP}r_P - w_{ES1}r_{S1} - w_{ES2}r_{S2} + I_E^B + I_E^{\text{stim}} + H^F\right]_+^2$$

$$\tau_P \frac{\mathrm{d}r_P}{\mathrm{d}t} = -r_P + \left[w_{PE}r_E - w_{PP}r_P - w_{PS2}r_{S2} + I_P^B + I_P^{\text{stim}}\right]_+^2$$

$$\tau_{S1} \frac{\mathrm{d}r_{S1}}{\mathrm{d}t} = -r_{S1} + \left[w_{S1E}r_E - w_{S1V}r_V + I_{S1}^B - I^{\text{pulvinar}}\right]_+^2$$

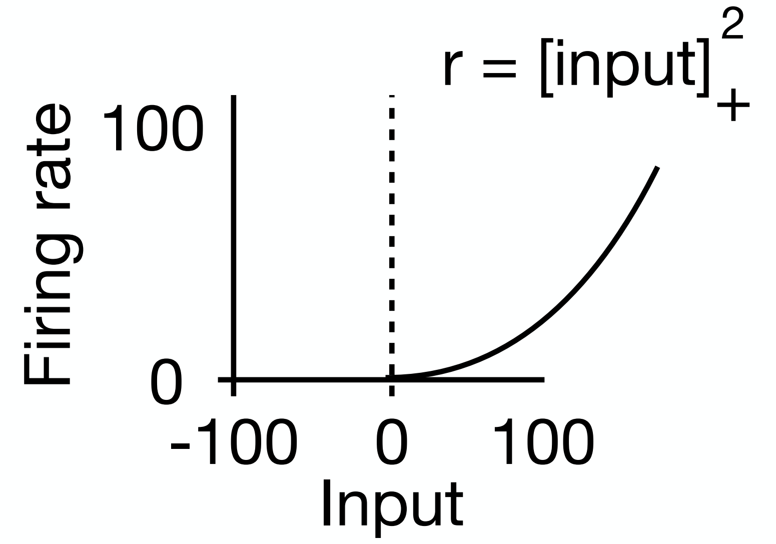
$$\tau_{S2} \frac{\mathrm{d}r_{S2}}{\mathrm{d}t} = -r_{S2} + \left[w_{S2E}r_E - w_{S2P}r_P + I_{S2}^B\right]_+^2$$

$$\tau_V \frac{\mathrm{d}r_V}{\mathrm{d}t} = -r_V + \left[w_{VE}r_E - w_{VS1}r_{S1} + I_V^B + I^{\text{kick}}\right]_+^2,$$



Microcircuit model

- multi-neuron **rate** model with experimentally observed:
  - ✦ relative counts [Billeh et al., 2020]
  - ✦ connectivity statistics
    & short-term plasticity parameters
    [Campagnola et al., 2020]
  - ✦ expansive nonlinearity



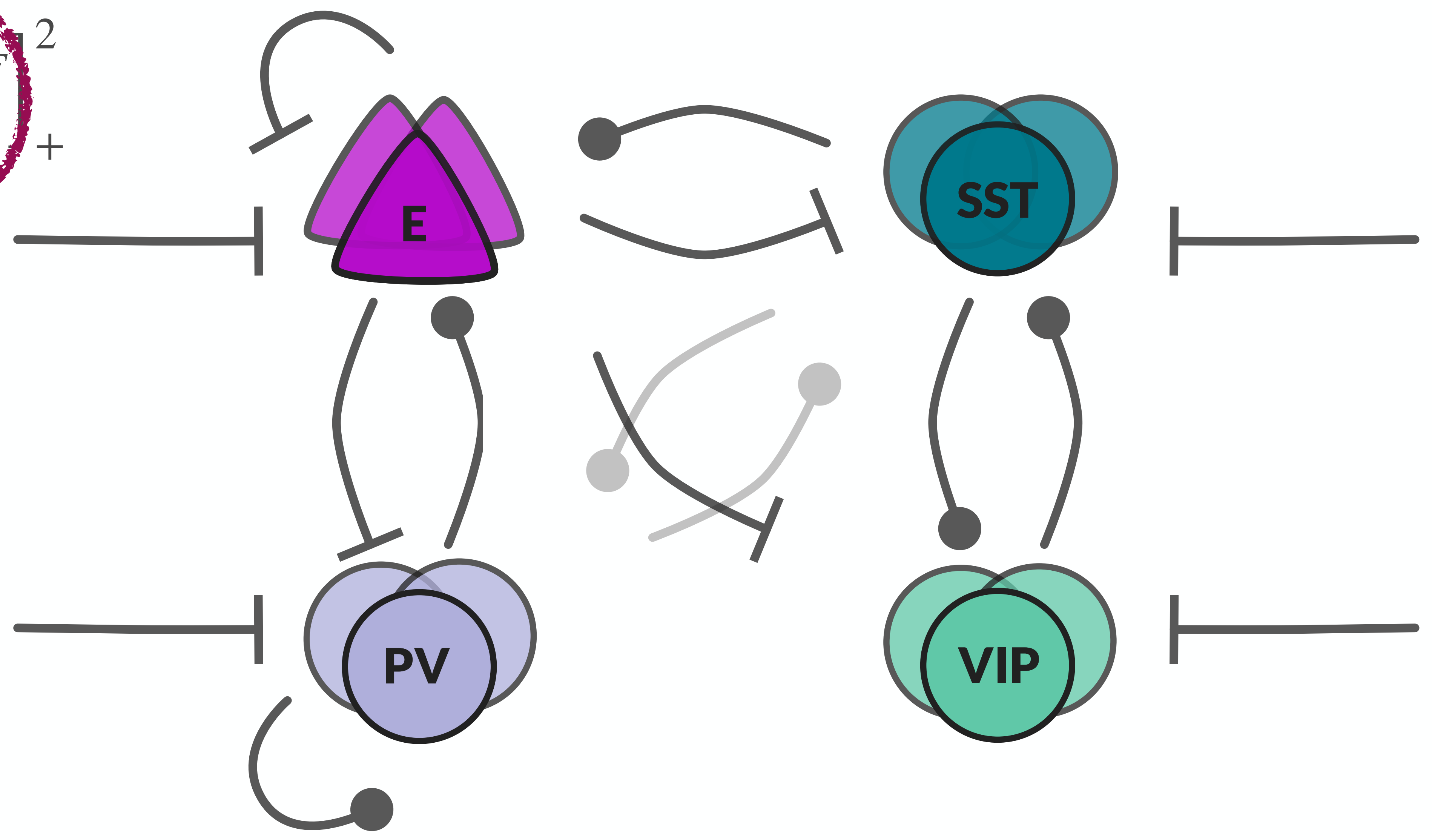
$$\tau_E \frac{\mathrm{d}r_E}{\mathrm{d}t} = -r_E + \left[w_{EE}r_E - w_{EP}r_P - w_{ES1}r_{S1} - w_{ES2}r_{S2} + I_E^B + I_E^{\text{stim}} + H^F\right]_+^2$$

$$\tau_P \frac{\mathrm{d}r_P}{\mathrm{d}t} = -r_P + \left[w_{PE}r_E - w_{PP}r_P - w_{PS2}r_{S2} + I_P^B + I_P^{\text{stim}}\right]_+^2$$

$$\tau_{S1} \frac{\mathrm{d}r_{S1}}{\mathrm{d}t} = -r_{S1} + \left[w_{S1E}r_E - w_{S1V}r_V + I_{S1}^B + I^{\text{pulvinar}}\right]_+^2$$

$$\tau_{S2} \frac{\mathrm{d}r_{S2}}{\mathrm{d}t} = -r_{S2} + \left[w_{S2E}r_E - w_{S2P}r_P + I_{S2}^B\right]_+^2$$

$$\tau_V \frac{\mathrm{d}r_V}{\mathrm{d}t} = -r_V + \left[w_{VE}r_E - w_{VS1}r_{S1} + I_V^B + I^{\text{kick}}\right]_+^2,$$



Baseline and stimulus input optimisation

$$\tau_E \frac{\mathrm{d}r_E}{\mathrm{d}t} = -r_E + \left[w_{EE}r_E - w_{EP}r_P - w_{ES1}r_{S1} - w_{ES2}r_{S2} + I_E^B + I_E^{\text{stim}} + H^F\right]_+^2$$

$$\tau_P \frac{\mathrm{d}r_P}{\mathrm{d}t} = -r_P + \left[w_{PE}r_E - w_{PP}r_P - w_{PS2}r_{S2} + I_P^B + I_P^{\text{stim}}\right]_+^2$$

$$\tau_{S1} \frac{\mathrm{d}r_{S1}}{\mathrm{d}t} = -r_{S1} + \left[w_{S1E}r_E - w_{S1V}r_V + I_{S1}^B - I^{\text{pulvinar}}\right]_+^2$$

$$\tau_{S2} \frac{\mathrm{d}r_{S2}}{\mathrm{d}t} = -r_{S2} + \left[w_{S2E}r_E - w_{S2P}r_P + I_{S2}^B\right]_+^2$$

$$\tau_V \frac{\mathrm{d}r_V}{\mathrm{d}t} = -r_V + \left[w_{VE}r_E - w_{VS1}r_{S1} + I_V^B + I^{\text{kick}}\right]_+^2,$$

Baseline and stimulus input optimisation

$$\tau_E \cancel{\frac{\mathrm{d}r_E}{\mathrm{d}t}} = -r_E + \left[\underline{w_{EE}} r_E - \underline{w_{EP}} r_P - \underline{w_{ES1}} r_{S1} - \underline{w_{ES2}} r_{S2} + I_E^B + I_E^{\text{stim}} + H^F\right]_+^2$$

$$\tau_P \cancel{\frac{\mathrm{d}r_P}{\mathrm{d}t}} = -r_P + \left[\underline{w_{PE}} r_E - \underline{w_{PP}} r_P - \underline{w_{PS2}} r_{S2} + I_P^B + I_P^{\text{stim}}\right]_+^2$$

$$\tau_{S1} \cancel{\frac{\mathrm{d}r_{S1}}{\mathrm{d}t}} = -r_{S1} + \left[\underline{w_{S1E}} r_E - \underline{w_{S1V}} r_V + I_{S1}^B - I^{\text{pulvinar}}\right]_+^2$$

$$\tau_{S2} \cancel{\frac{\mathrm{d}r_{S2}}{\mathrm{d}t}} = -r_{S2} + \left[\underline{w_{S2E}} r_E - \underline{w_{S2P}} r_P + I_{S2}^B\right]_+^2$$

$$\tau_V \cancel{\frac{\mathrm{d}r_V}{\mathrm{d}t}} = -r_V + \left[\underline{w_{VE}} r_E - \underline{w_{VS1}} r_{S1} + I_V^B + I^{\text{kick}}\right]_+^2,$$

Baseline and stimulus input optimisation

$$\cancel{\tau_E \frac{\mathrm{d}r_E}{\mathrm{d}t}} = -r_E + \left[\underline{w_{EE}} r_E - \underline{w_{EP}} r_P - \underline{w_{ES1}} r_{S1} - \underline{w_{ES2}} r_{S2} + I_E^B + \cancel{I_E^{\text{stim}}} + \cancel{H^F}\right]_+^2$$

$$\cancel{\tau_P \frac{\mathrm{d}r_P}{\mathrm{d}t}} = -r_P + \left[\underline{w_{PE}} r_E - \underline{w_{PP}} r_P - \underline{w_{PS2}} r_{S2} + I_P^B + \cancel{I_P^{\text{stim}}}\right]_+^2$$

$$\cancel{\tau_{S1} \frac{\mathrm{d}r_{S1}}{\mathrm{d}t}} = -r_{S1} + \left[\underline{w_{S1E}} r_E - \underline{w_{S1V}} r_V + I_{S1}^B - \cancel{I^{\text{pulvinar}}}\right]_+^2$$

$$\cancel{\tau_{S2} \frac{\mathrm{d}r_{S2}}{\mathrm{d}t}} = -r_{S2} + \left[\underline{w_{S2E}} r_E - \underline{w_{S2P}} r_P + I_{S2}^B\right]_+^2$$

$$\cancel{\tau_V \frac{\mathrm{d}r_V}{\mathrm{d}t}} = -r_V + \left[\underline{w_{VE}} r_E - \underline{w_{VS1}} r_{S1} + I_V^B + \cancel{I^{\text{kick}}}\right]_+^2,$$

Baseline and stimulus input optimisation

$$\cancel{\tau_E \frac{\mathrm{d}r_E}{\mathrm{d}t}} = -r_E + \left[\underline{w_{EE}}r_E - \underline{w_{EP}}r_P - \underline{w_{ES1}}r_{S1} - \underline{w_{ES2}}r_{S2} + I_E^B + I_E^{\text{stim}} \cancel{+ H^F}\right]_+^2$$

$$\cancel{\tau_P \frac{\mathrm{d}r_P}{\mathrm{d}t}} = -r_P + \left[\underline{w_{PE}}r_E - \underline{w_{PP}}r_P - \underline{w_{PS2}}r_{S2} + I_P^B + I_P^{\text{stim}}\right]_+^2$$

$$\cancel{\tau_{S1} \frac{\mathrm{d}r_{S1}}{\mathrm{d}t}} = -r_{S1} + \left[\underline{w_{S1E}}r_E - \underline{w_{S1V}}r_V + I_{S1}^B \cancel{- I^{\text{pulvinar}}}\right]_+^2$$

$$\cancel{\tau_{S2} \frac{\mathrm{d}r_{S2}}{\mathrm{d}t}} = -r_{S2} + \left[\underline{w_{S2E}}r_E - \underline{w_{S2P}}r_P + I_{S2}^B\right]_+^2$$

$$\cancel{\tau_V \frac{\mathrm{d}r_V}{\mathrm{d}t}} = -r_V + \left[\underline{w_{VE}}r_E - \underline{w_{VS1}}r_{S1} + I_V^B \cancel{+ I^{\text{kick}}}\right]_+^2,$$

Baseline and stimulus input optimisation

$$\tau_E \frac{\mathrm{d}r_E}{\mathrm{d}t} = -r_E + \left[w_{EE}r_E - w_{EP}r_P - w_{ES1}r_{S1} - w_{ES2}r_{S2} + I_E^B + I_E^{\text{stim}} + H^F\right]_+^2$$

$$\tau_P \frac{\mathrm{d}r_P}{\mathrm{d}t} = -r_P + \left[w_{PE}r_E - w_{PP}r_P - w_{PS2}r_{S2} + I_P^B + I_P^{\text{stim}}\right]_+^2$$

$$\tau_{S1} \frac{\mathrm{d}r_{S1}}{\mathrm{d}t} = -r_{S1} + \left[w_{S1E}r_E - w_{S1V}r_V + I_{S1}^B - I^{\text{pulvinar}}\right]_+^2$$

$$\tau_{S2} \frac{\mathrm{d}r_{S2}}{\mathrm{d}t} = -r_{S2} + \left[w_{S2E}r_E - w_{S2P}r_P + I_{S2}^B\right]_+^2$$

$$\tau_V \frac{\mathrm{d}r_V}{\mathrm{d}t} = -r_V + \left[w_{VE}r_E - w_{VS1}r_{S1} + I_V^B + I^{\text{kick}}\right]_+^2,$$

Baseline and stimulus input optimisation

$$\cancel{\tau_E \frac{\mathrm{d}r_E}{\mathrm{d}t}} = -r_E + \left[\underline{w_{EE}} r_E - \underline{w_{EP}} r_P - \underline{w_{ES1}} r_{S1} - \underline{w_{ES2}} r_{S2} + I_E^B + I_E^{\text{stim}} + H^F\right]_+^2$$

$$\cancel{\tau_P \frac{\mathrm{d}r_P}{\mathrm{d}t}} = -r_P + \left[\underline{w_{PE}} r_E - \underline{w_{PP}} r_P - \underline{w_{PS2}} r_{S2} + I_P^B + I_P^{\text{stim}}\right]_+^2$$

$$\cancel{\tau_{S1} \frac{\mathrm{d}r_{S1}}{\mathrm{d}t}} = -r_{S1} + \left[\underline{w_{S1E}} r_E - \underline{w_{S1V}} r_V + I_{S1}^B - I^{\text{pulvinar}}\right]_+^2$$

$$\cancel{\tau_{S2} \frac{\mathrm{d}r_{S2}}{\mathrm{d}t}} = -r_{S2} + \left[\underline{w_{S2E}} r_E - \underline{w_{S2P}} r_P + I_{S2}^B\right]_+^2$$

$$\cancel{\tau_V \frac{\mathrm{d}r_V}{\mathrm{d}t}} = -r_V + \left[\underline{w_{VE}} r_E - \underline{w_{VS1}} r_{S1} + I_V^B + I^{\text{kick}}\right]_+^2,$$

Baseline and stimulus input optimisation

$$\tau_E \cancel{\frac{\mathrm{d}r_E}{\mathrm{d}t}} = -r_E + \left[\underline{w_{EE}} r_E - \underline{w_{EP}} r_P - \underline{w_{ES1}} r_{S1} - \underline{w_{ES2}} r_{S2} + I_E^B + \cancel{I_E^{\text{stim}}} + \cancel{H^F}\right]_+^2$$

$$\tau_P \cancel{\frac{\mathrm{d}r_P}{\mathrm{d}t}} = -r_P + \left[\underline{w_{PE}} r_E - \underline{w_{PP}} r_P - \underline{w_{PS2}} r_{S2} + I_P^B + \cancel{I_P^{\text{stim}}}\right]_+^2$$

$$\tau_{S1} \cancel{\frac{\mathrm{d}r_{S1}}{\mathrm{d}t}} = -r_{S1} + \left[\underline{w_{S1E}} r_E - \underline{w_{S1V}} r_V + I_{S1}^B - \cancel{I^{\text{pulvinar}}}\right]_+^2$$

$$\tau_{S2} \cancel{\frac{\mathrm{d}r_{S2}}{\mathrm{d}t}} = -r_{S2} + \left[\underline{w_{S2E}} r_E - \underline{w_{S2P}} r_P + I_{S2}^B\right]_+^2$$

$$\tau_V \cancel{\frac{\mathrm{d}r_V}{\mathrm{d}t}} = -r_V + \left[\underline{w_{VE}} r_E - \underline{w_{VS1}} r_{S1} + I_V^B + \cancel{I^{\text{kick}}}\right]_+^2,$$

Baseline and stimulus input optimisation

$$\tau_E \cancel{\frac{\mathrm{d}r_E}{\mathrm{d}t}} = -r_E + \left[\underline{w_{EE}} r_E - \underline{w_{EP}} r_P - \underline{w_{ES1}} r_{S1} - \underline{w_{ES2}} r_{S2} + I_E^B + I_E^{\text{stim}} \cancel{+ H^F}\right]_+^2$$

$$\tau_P \cancel{\frac{\mathrm{d}r_P}{\mathrm{d}t}} = -r_P + \left[\underline{w_{PE}} r_E - \underline{w_{PP}} r_P - \underline{w_{PS2}} r_{S2} + I_P^B + I_P^{\text{stim}}\right]_+^2$$

$$\tau_{S1} \cancel{\frac{\mathrm{d}r_{S1}}{\mathrm{d}t}} = -r_{S1} + \left[\underline{w_{S1E}} r_E - \underline{w_{S1V}} r_V + I_{S1}^B \cancel{+ I^{\text{pulvinar}}}\right]_+^2$$

$$\tau_{S2} \cancel{\frac{\mathrm{d}r_{S2}}{\mathrm{d}t}} = -r_{S2} + \left[\underline{w_{S2E}} r_E - \underline{w_{S2P}} r_P + I_{S2}^B\right]_+^2$$

$$\tau_V \cancel{\frac{\mathrm{d}r_V}{\mathrm{d}t}} = -r_V + \left[\underline{w_{VE}} r_E - \underline{w_{VS1}} r_{S1} + I_V^B \cancel{+ I^{\text{kick}}}\right]_+^2,$$

Baseline and stimulus input optimisation

$$\cancel{\tau_E \frac{\mathrm{d}r_E}{\mathrm{d}t}} = -r_E + \left[\underline{w_{EE}}r_E - \underline{w_{EP}}r_P - \underline{w_{ES1}}r_{S1} - \underline{w_{ES2}}r_{S2} + I_E^B + I_E^{\text{stim}} \cancel{+ H^F}\right]_+^2$$

$$\cancel{\tau_P \frac{\mathrm{d}r_P}{\mathrm{d}t}} = -r_P + \left[\underline{w_{PE}}r_E - \underline{w_{PP}}r_P - \underline{w_{PS2}}r_{S2} + I_P^B + I_P^{\text{stim}}\right]_+^2$$

$$\cancel{\tau_{S1} \frac{\mathrm{d}r_{S1}}{\mathrm{d}t}} = -r_{S1} + \left[\underline{w_{S1E}}r_E - \underline{w_{S1V}}r_V + I_{S1}^B - \cancel{I^{\text{pulvinar}}}\right]_+^2$$

$$\cancel{\tau_{S2} \frac{\mathrm{d}r_{S2}}{\mathrm{d}t}} = -r_{S2} + \left[\underline{w_{S2E}}r_E - \underline{w_{S2P}}r_P + I_{S2}^B\right]_+^2$$

$$\cancel{\tau_V \frac{\mathrm{d}r_V}{\mathrm{d}t}} = -r_V + \left[\underline{w_{VE}}r_E - \underline{w_{VS1}}r_{S1} + I_V^B + \cancel{I^{\text{kick}}}\right]_+^2,$$

fractional mean-field equations

Baseline and stimulus input optimisation

$$\tau_E \cancel{\frac{\mathrm{d}r_E}{\mathrm{d}t}} = -r_E + \left[\underline{w_{EE}} r_E - \underline{w_{EP}} r_P - \underline{w_{ES1}} r_{S1} - \underline{w_{ES2}} r_{S2} + I_E^B + I_E^{\text{stim}} \cancel{+ H^F}\right]_+^2$$

$$\tau_P \cancel{\frac{\mathrm{d}r_P}{\mathrm{d}t}} = -r_P + \left[\underline{w_{PE}} r_E - \underline{w_{PP}} r_P - \underline{w_{PS2}} r_{S2} + I_P^B + I_P^{\text{stim}}\right]_+^2$$

$$\tau_{S1} \cancel{\frac{\mathrm{d}r_{S1}}{\mathrm{d}t}} = -r_{S1} + \left[\underline{w_{S1E}} r_E - \underline{w_{S1V}} r_V + I_{S1}^B - \cancel{I^{\text{pulvinar}}}\right]_+^2$$

$$\tau_{S2} \cancel{\frac{\mathrm{d}r_{S2}}{\mathrm{d}t}} = -r_{S2} + \left[\underline{w_{S2E}} r_E - \underline{w_{S2P}} r_P + I_{S2}^B\right]_+^2$$

$$\tau_V \cancel{\frac{\mathrm{d}r_V}{\mathrm{d}t}} = -r_V + \left[\underline{w_{VE}} r_E - \underline{w_{VS1}} r_{S1} + I_V^B + \cancel{I^{\text{kick}}}\right]_+^2,$$

fractional mean-field equations

$$w_{EE} r_E \rightarrow \gamma_E w_{EE} r_{Et} + (1 - \gamma_E) w_{EE} r_{Ed}$$

Baseline and stimulus input optimisation

$$\tau_E \cancel{\frac{\mathrm{d}r_E}{\mathrm{d}t}} = -r_E + \left[\underline{w_{EE}}r_E - \underline{w_{EP}}r_P - \underline{w_{ES1}}r_{S1} - \underline{w_{ES2}}r_{S2} + I_E^B + I_E^{\text{stim}} \cancel{+ H^F}\right]_+^2$$

$$\tau_P \cancel{\frac{\mathrm{d}r_P}{\mathrm{d}t}} = -r_P + \left[\underline{w_{PE}}r_E - \underline{w_{PP}}r_P - \underline{w_{PS2}}r_{S2} + I_P^B + I_P^{\text{stim}}\right]_+^2$$

$$\tau_{S1} \cancel{\frac{\mathrm{d}r_{S1}}{\mathrm{d}t}} = -r_{S1} + \left[\underline{w_{S1E}}r_E - \underline{w_{S1V}}r_V + I_{S1}^B - \cancel{I^{\text{pulvinar}}}\right]_+^2$$

$$\tau_{S2} \cancel{\frac{\mathrm{d}r_{S2}}{\mathrm{d}t}} = -r_{S2} + \left[\underline{w_{S2E}}r_E - \underline{w_{S2P}}r_P + I_{S2}^B\right]_+^2$$

$$\tau_V \cancel{\frac{\mathrm{d}r_V}{\mathrm{d}t}} = -r_V + \left[\underline{w_{VE}}r_E - \underline{w_{VS1}}r_{S1} + I_V^B + \cancel{I^{\text{kick}}}\right]_+^2,$$

fractional mean-field equations

$$w_{EE}r_E \to \underline{\gamma_E} w_{EE}r_{Et} + (1-\gamma_E)w_{EE}r_{Ed}$$

fraction of
stimulated E cells

Baseline and stimulus input optimisation

$$\tau_E \cancel{\frac{\mathrm{d}r_E}{\mathrm{d}t}} = -r_E + \left[\underline{w_{EE}} r_E - \underline{w_{EP}} r_P - \underline{w_{ES1}} r_{S1} - \underline{w_{ES2}} r_{S2} + I_E^B + I_E^{\text{stim}} \cancel{+ H^F}\right]_+^2$$

$$\tau_P \cancel{\frac{\mathrm{d}r_P}{\mathrm{d}t}} = -r_P + \left[\underline{w_{PE}} r_E - \underline{w_{PP}} r_P - \underline{w_{PS2}} r_{S2} + I_P^B + I_P^{\text{stim}}\right]_+^2$$

$$\tau_{S1} \cancel{\frac{\mathrm{d}r_{S1}}{\mathrm{d}t}} = -r_{S1} + \left[\underline{w_{S1E}} r_E - \underline{w_{S1V}} r_V + I_{S1}^B - \cancel{I^{\text{pulvinar}}}\right]_+^2$$

$$\tau_{S2} \cancel{\frac{\mathrm{d}r_{S2}}{\mathrm{d}t}} = -r_{S2} + \left[\underline{w_{S2E}} r_E - \underline{w_{S2P}} r_P + I_{S2}^B\right]_+^2$$

$$\tau_V \cancel{\frac{\mathrm{d}r_V}{\mathrm{d}t}} = -r_V + \left[\underline{w_{VE}} r_E - \underline{w_{VS1}} r_{S1} + I_V^B + \cancel{I^{\text{kick}}}\right]_+^2,$$

fractional mean-field equations

rate of stimulated cells

$$w_{EE} r_E \rightarrow \underline{\gamma_E} w_{EE} r_{Et} + (1 - \gamma_E) w_{EE} r_{Ed}$$

fraction of stimulated E cells

Baseline and stimulus input optimisation

$$\tau_E \cancel{\frac{\mathrm{d}r_E}{\mathrm{d}t}} = -r_E + \left[w_{EE}r_E - w_{EP}r_P - w_{ES1}r_{S1} - w_{ES2}r_{S2} + I_E^B + I_E^{\text{stim}} + \cancel{H^F}\right]_+^2$$

$$\tau_P \cancel{\frac{\mathrm{d}r_P}{\mathrm{d}t}} = -r_P + \left[w_{PE}r_E - w_{PP}r_P - w_{PS2}r_{S2} + I_P^B + I_P^{\text{stim}}\right]_+^2$$

$$\tau_{S1} \cancel{\frac{\mathrm{d}r_{S1}}{\mathrm{d}t}} = -r_{S1} + \left[w_{S1E}r_E - w_{S1V}r_V + I_{S1}^B - \cancel{I^{\text{pulvinar}}}\right]_+^2$$

$$\tau_{S2} \cancel{\frac{\mathrm{d}r_{S2}}{\mathrm{d}t}} = -r_{S2} + \left[w_{S2E}r_E - w_{S2P}r_P + I_{S2}^B\right]_+^2$$

$$\tau_V \cancel{\frac{\mathrm{d}r_V}{\mathrm{d}t}} = -r_V + \left[w_{VE}r_E - w_{VS1}r_{S1} + I_V^B + \cancel{I^{\text{kick}}}\right]_+^2,$$

fractional mean-field equations

rate of
stimulated cells

$$w_{EE}r_E \to \gamma_E w_{EE} r_{Et} + (1-\gamma_E) w_{EE} r_{Ed}$$

fraction of
stimulated E cells

rate of
non-stimulated E cells

Baseline and stimulus input optimisation

$$\tau_E \cancel{\frac{\mathrm{d}r_E}{\mathrm{d}t}} = -r_E + \left[w_{EE}r_E - w_{EP}r_P - w_{ES1}r_{S1} - w_{ES2}r_{S2} + I_E^B + I_E^{\text{stim}} + \cancel{H^F}\right]_+^2$$

$$\tau_P \cancel{\frac{\mathrm{d}r_P}{\mathrm{d}t}} = -r_P + \left[w_{PE}r_E - w_{PP}r_P - w_{PS2}r_{S2} + I_P^B + I_P^{\text{stim}}\right]_+^2$$

$$\tau_{S1} \cancel{\frac{\mathrm{d}r_{S1}}{\mathrm{d}t}} = -r_{S1} + \left[w_{S1E}r_E - w_{S1V}r_V + I_{S1}^B - \cancel{I^{\text{pulvinar}}}\right]_+^2$$

$$\tau_{S2} \cancel{\frac{\mathrm{d}r_{S2}}{\mathrm{d}t}} = -r_{S2} + \left[w_{S2E}r_E - w_{S2P}r_P + I_{S2}^B\right]_+^2$$

$$\tau_V \cancel{\frac{\mathrm{d}r_V}{\mathrm{d}t}} = -r_V + \left[w_{VE}r_E - w_{VS1}r_{S1} + I_V^B + \cancel{I^{\text{kick}}}\right]_+^2,$$

fractional mean-field equations

rate of stimulated cells

$$w_{EE}r_E \rightarrow \gamma_E w_{EE} r_{Et} + (1-\gamma_E) w_{EE} r_{Ed}$$

fraction of stimulated E cells

rate of non-stimulated E cells

- Solve for inputs

Baseline and stimulus input optimisation

$$\tau_E \cancel{\frac{\mathrm{d}r_E}{\mathrm{d}t}} = -r_E + \left[w_{EE}r_E - w_{EP}r_P - w_{ES1}r_{S1} - w_{ES2}r_{S2} + I_E^B + I_E^{\text{stim}} \cancel{+ H^F}\right]_+^2$$

$$\tau_P \cancel{\frac{\mathrm{d}r_P}{\mathrm{d}t}} = -r_P + \left[w_{PE}r_E - w_{PP}r_P - w_{PS2}r_{S2} + I_P^B + I_P^{\text{stim}}\right]_+^2$$

$$\tau_{S1} \cancel{\frac{\mathrm{d}r_{S1}}{\mathrm{d}t}} = -r_{S1} + \left[w_{S1E}r_E - w_{S1V}r_V + I_{S1}^B \cancel{- I^{\text{pulvinar}}}\right]_+^2$$

$$\tau_{S2} \cancel{\frac{\mathrm{d}r_{S2}}{\mathrm{d}t}} = -r_{S2} + \left[w_{S2E}r_E - w_{S2P}r_P + I_{S2}^B\right]_+^2$$

$$\tau_V \cancel{\frac{\mathrm{d}r_V}{\mathrm{d}t}} = -r_V + \left[w_{VE}r_E - w_{VS1}r_{S1} + I_V^B \cancel{+ I^{\text{kick}}}\right]_+^2,$$

fractional mean-field equations

rate of
stimulated cells

$$w_{EE}r_E \rightarrow \gamma_E w_{EE}r_{Et} + (1-\gamma_E)w_{EE}r_{Ed}$$

fraction of
stimulated E cells

rate of
non-stimulated E cells

- Solve for inputs
- Optimise both inputs and a subset of weights

Baseline and stimulus input optimisation

$$\cancel{\tau_E \frac{\mathrm{d}r_E}{\mathrm{d}t}} = -r_E + \left[\underline{w_{EE}} r_E - \underline{w_{EP}} r_P - \underline{w_{ES1}} r_{S1} - \underline{w_{ES2}} r_{S2} + I_E^B + I_E^{\text{stim}} \cancel{+ H^F}\right]_+^2$$

$$\cancel{\tau_P \frac{\mathrm{d}r_P}{\mathrm{d}t}} = -r_P + \left[\underline{w_{PE}} r_E - \underline{w_{PP}} r_P - \underline{w_{PS2}} r_{S2} + I_P^B + I_P^{\text{stim}}\right]_+^2$$

$$\cancel{\tau_{S1} \frac{\mathrm{d}r_{S1}}{\mathrm{d}t}} = -r_{S1} + \left[\underline{w_{S1E}} r_E - \underline{w_{S1V}} r_V + I_{S1}^B - \cancel{I^{\text{pulvinar}}}\right]_+^2$$

$$\cancel{\tau_{S2} \frac{\mathrm{d}r_{S2}}{\mathrm{d}t}} = -r_{S2} + \left[\underline{w_{S2E}} r_E - \underline{w_{S2P}} r_P + I_{S2}^B\right]_+^2$$

$$\cancel{\tau_V \frac{\mathrm{d}r_V}{\mathrm{d}t}} = -r_V + \left[\underline{w_{VE}} r_E - \underline{w_{VS1}} r_{S1} + I_V^B + \cancel{I^{\text{kick}}}\right]_+^2,$$

fractional mean-field equations

rate of stimulated cells

$$w_{EE} r_E \rightarrow \underline{\gamma_E} w_{EE} r_{Et} + (1 - \gamma_E) w_{EE} \underline{r_{Ed}}$$

fraction of stimulated E cells

rate of non-stimulated E cells

- Solve for inputs
- Optimise both inputs and a subset of weights

$$\mathscr{C} = \sum (\text{residuals})^2 + \lambda \| W - W_0 \|_p$$

Baseline and stimulus input optimisation

$$\tau_E \cancel{\frac{\mathrm{d}r_E}{\mathrm{d}t}} = -r_E + \left[w_{EE}r_E - w_{EP}r_P - w_{ES1}r_{S1} - w_{ES2}r_{S2} + I_E^B + I_E^{\text{stim}} + \cancel{H^F}\right]_+^2$$

$$\tau_P \cancel{\frac{\mathrm{d}r_P}{\mathrm{d}t}} = -r_P + \left[w_{PE}r_E - w_{PP}r_P - w_{PS2}r_{S2} + I_P^B + I_P^{\text{stim}}\right]_+^2$$

$$\tau_{S1} \cancel{\frac{\mathrm{d}r_{S1}}{\mathrm{d}t}} = -r_{S1} + \left[w_{S1E}r_E - w_{S1V}r_V + I_{S1}^B - \cancel{I^{\text{pulvinar}}}\right]_+^2$$

$$\tau_{S2} \cancel{\frac{\mathrm{d}r_{S2}}{\mathrm{d}t}} = -r_{S2} + \left[w_{S2E}r_E - w_{S2P}r_P + I_{S2}^B\right]_+^2$$

$$\tau_V \cancel{\frac{\mathrm{d}r_V}{\mathrm{d}t}} = -r_V + \left[w_{VE}r_E - w_{VS1}r_{S1} + I_V^B + \cancel{I^{\text{kick}}}\right]_+^2,$$

fractional mean-field equations

rate of
stimulated cells

$$w_{EE}r_E \rightarrow \gamma_E w_{EE}r_{Et} + (1-\gamma_E)w_{EE}r_{Ed}$$

fraction of
stimulated E cells

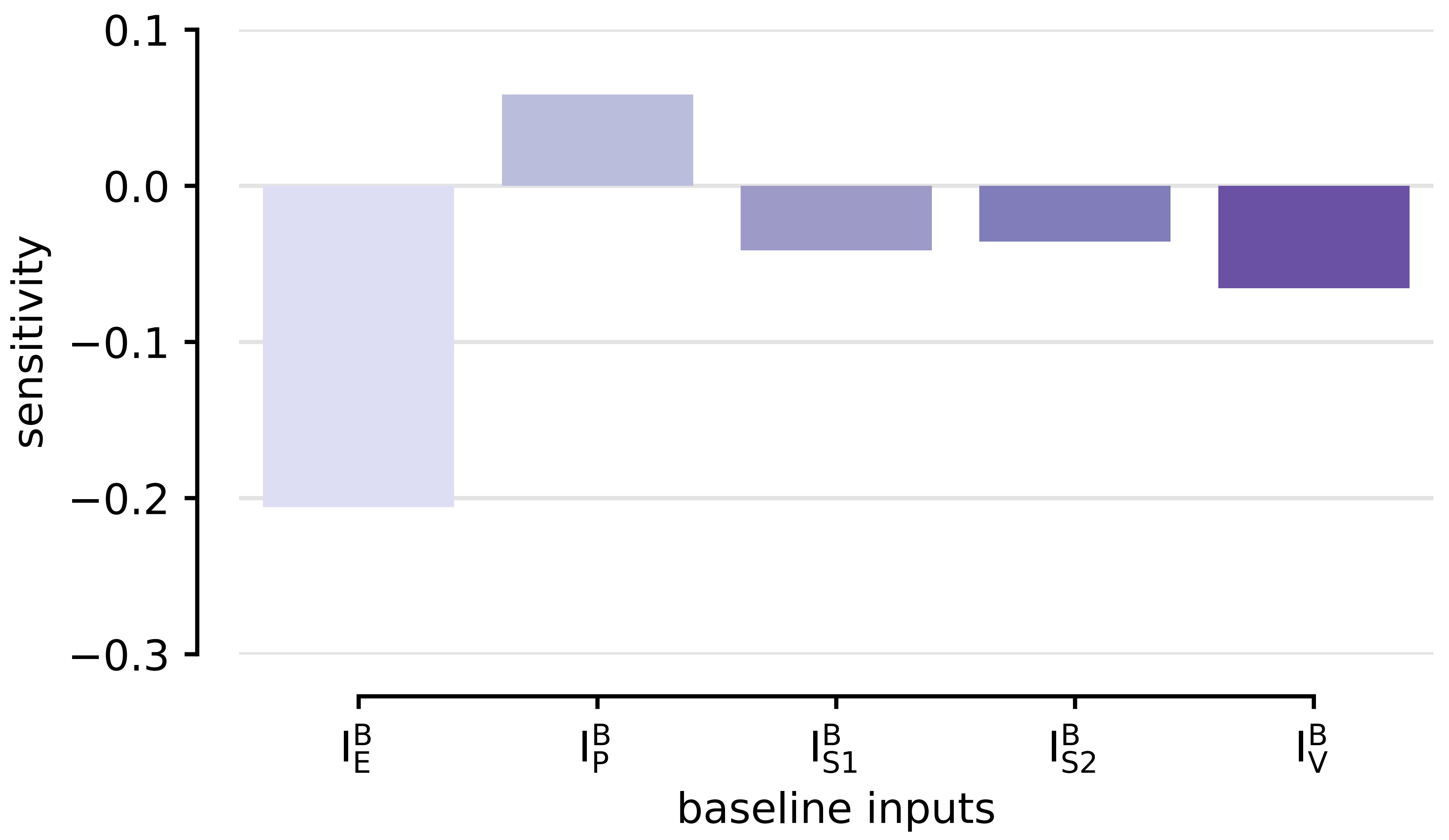
rate of
non-stimulated E cells

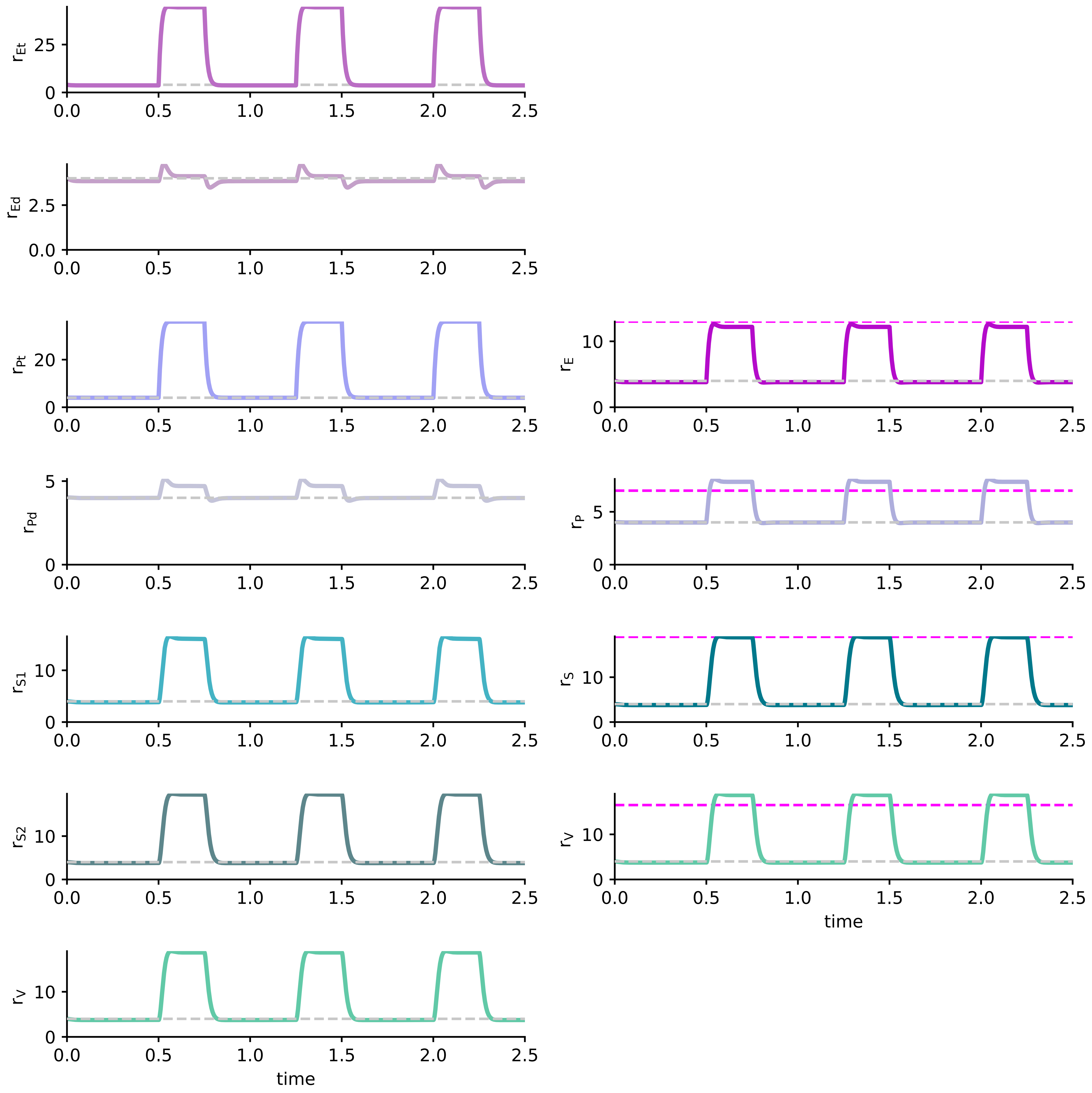
- Solve for inputs
- Optimise both inputs and a subset of weights

$$\mathscr{C} = \sum (\text{residuals})^2 + \lambda ||W - W_0||_p$$

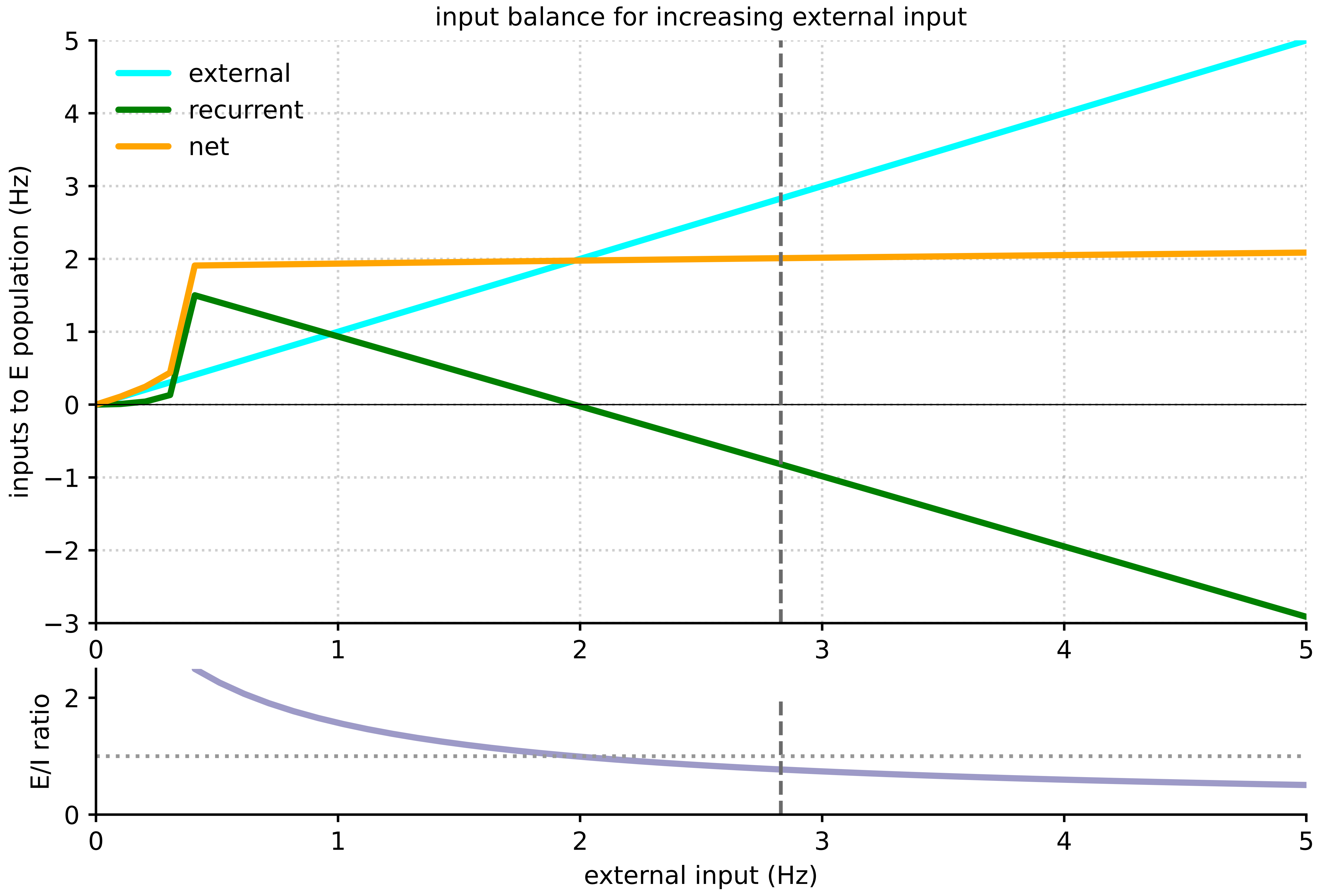
Campagnola
weights

Adjusting network to replicate naive responses



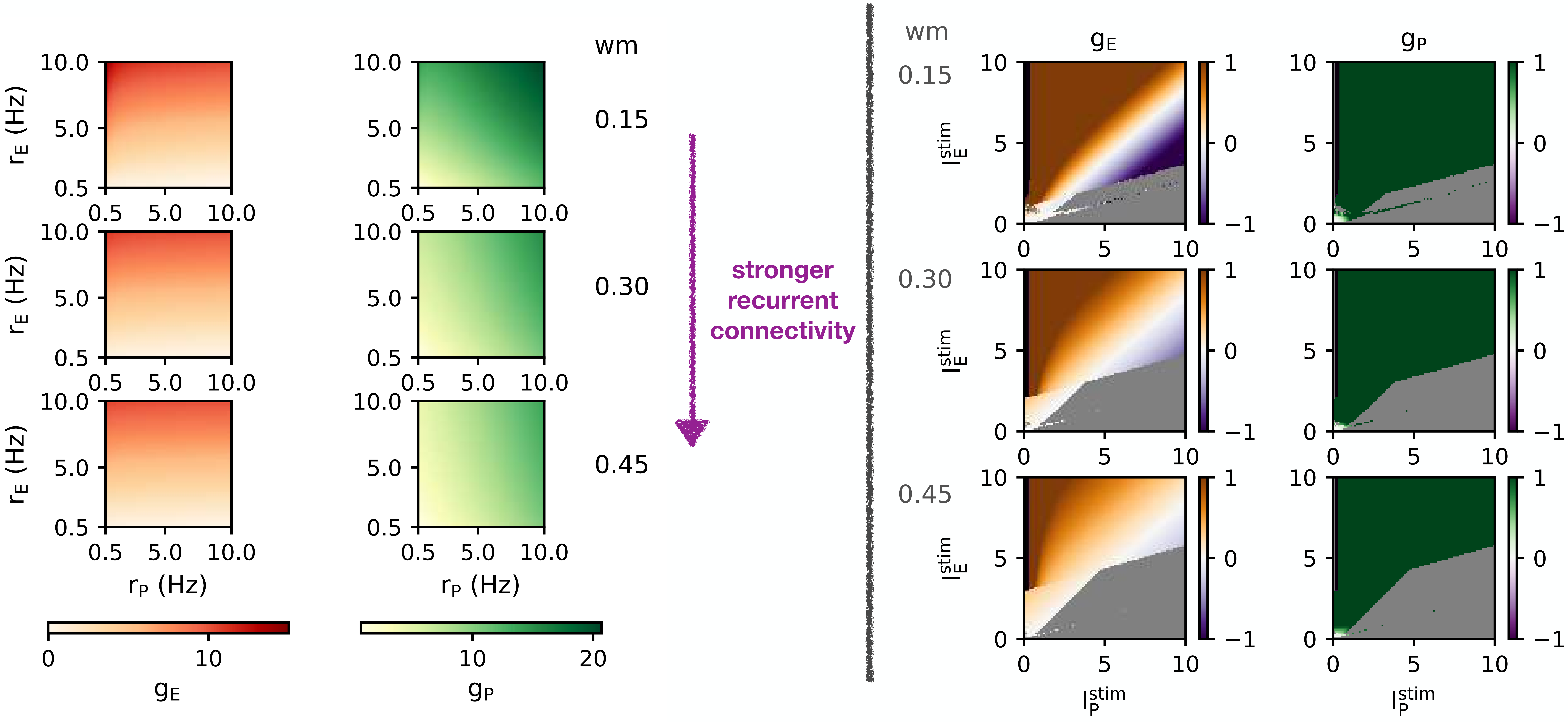


Inhibition stabilized regime



how shall we introduce the changes?

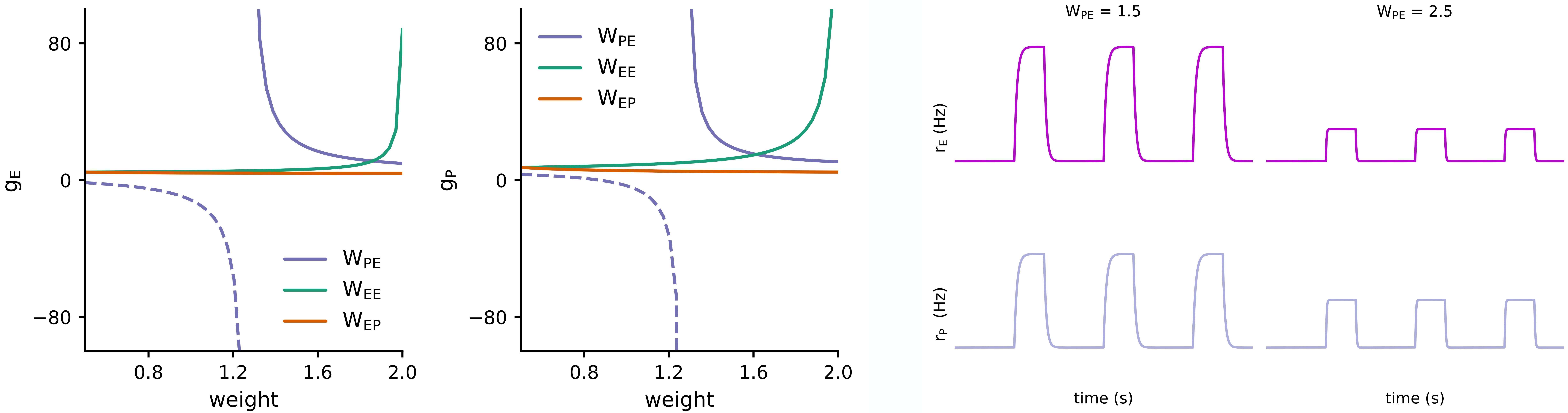
Network gain characterises sensitivity of responses to input changes



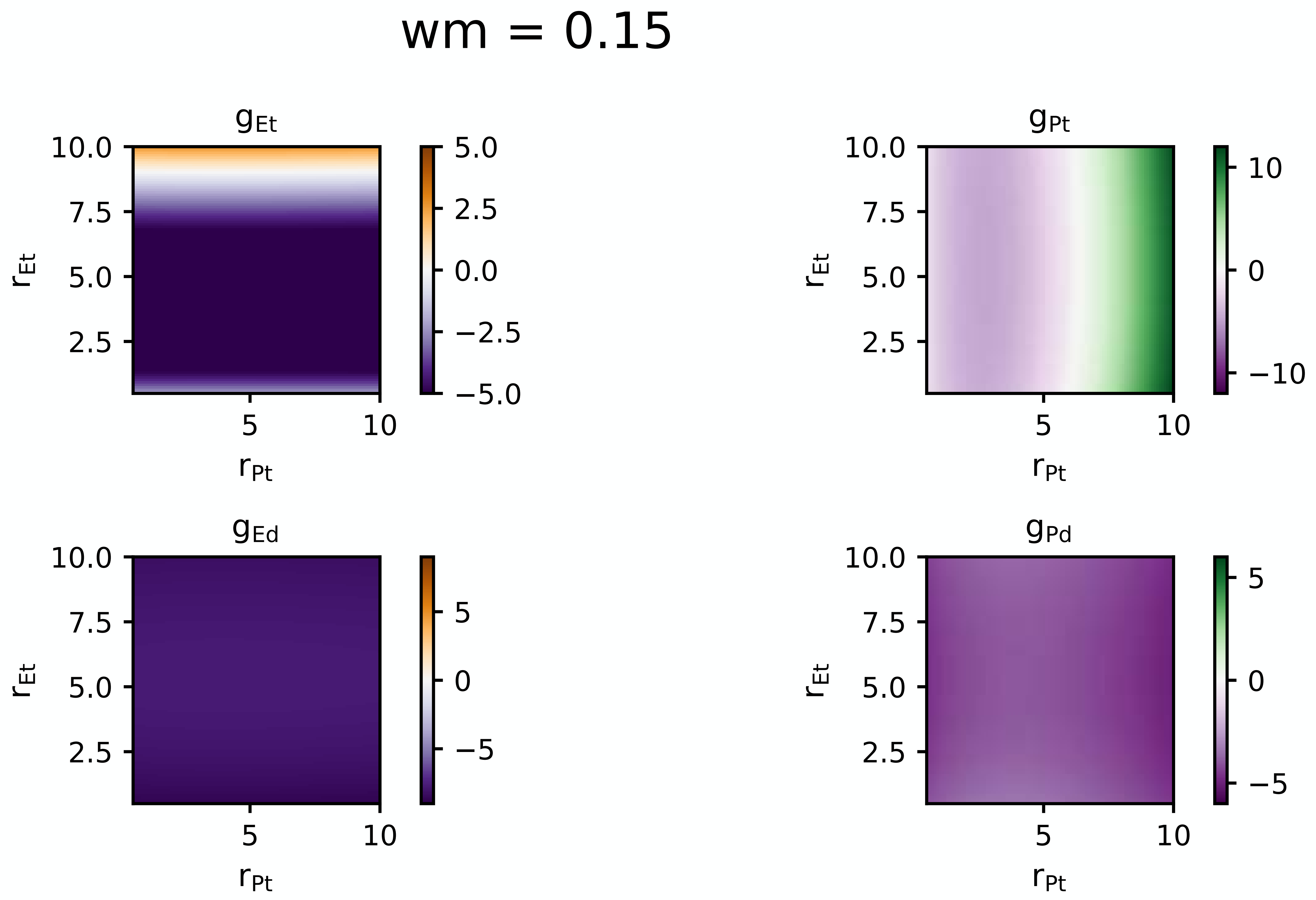
Potentiating E->PV weight reduces stimulus responses

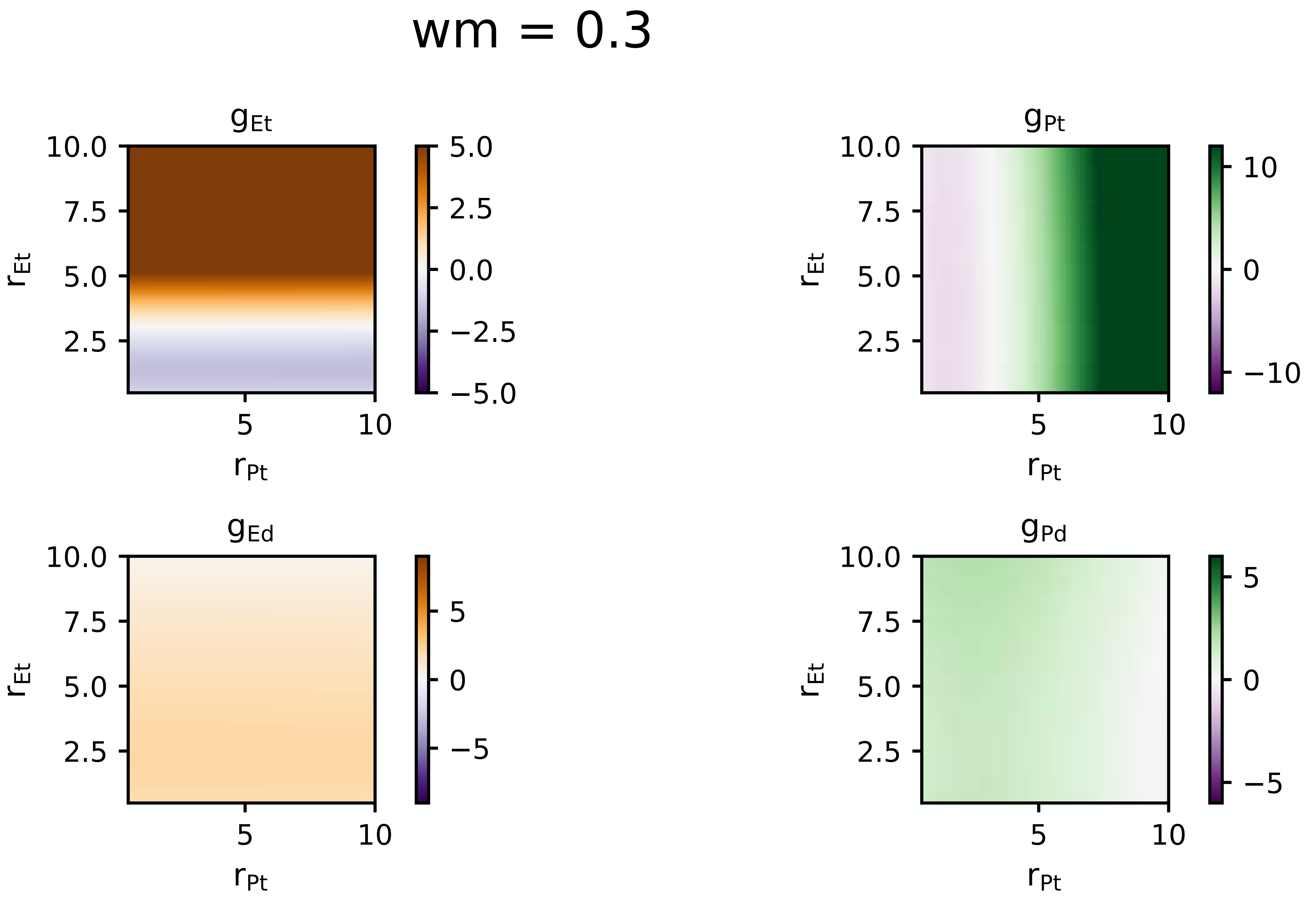
ISN

potentiating E->PV connection decreases response gain for incoming stimuli



Fractional network gain

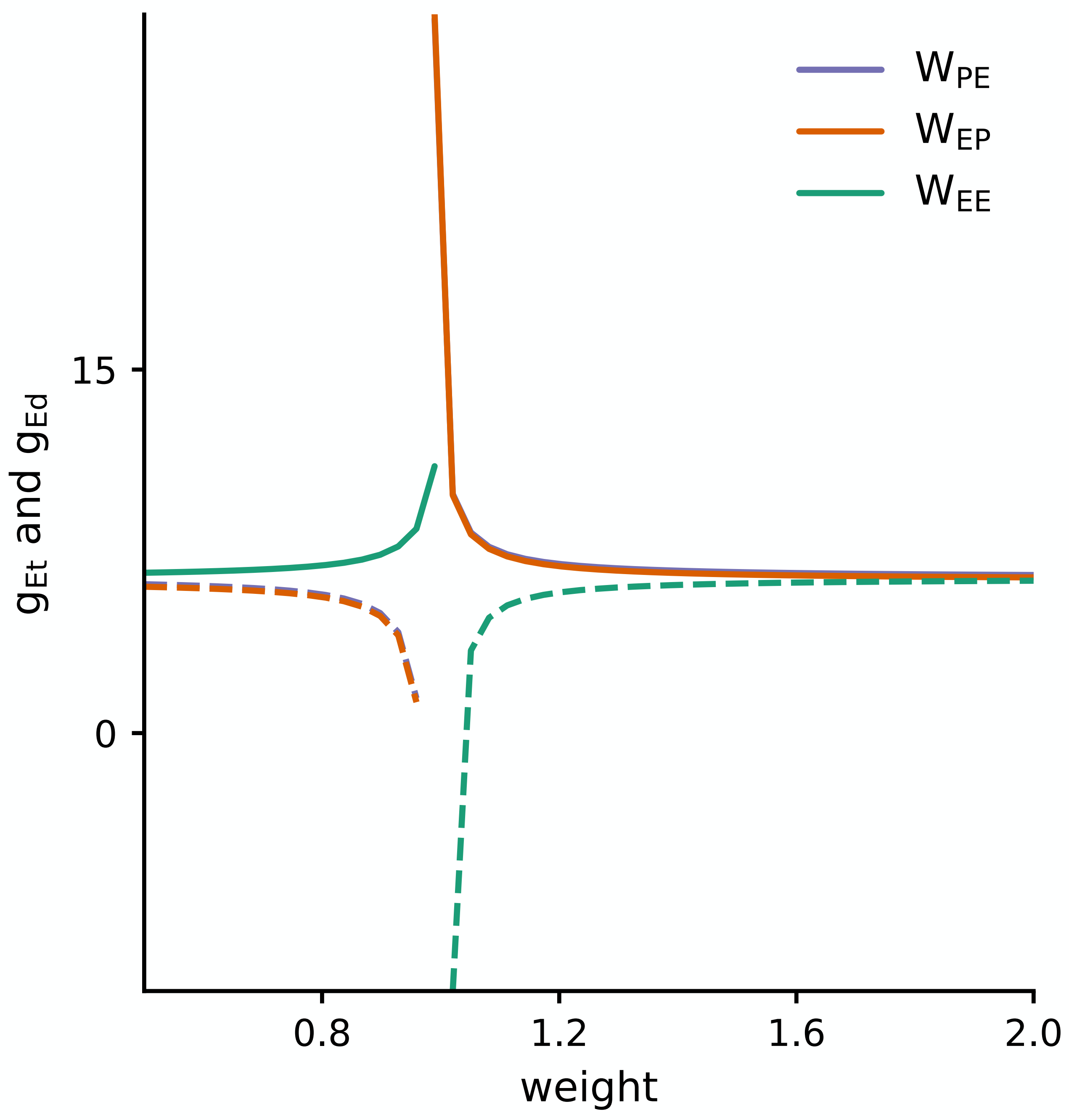


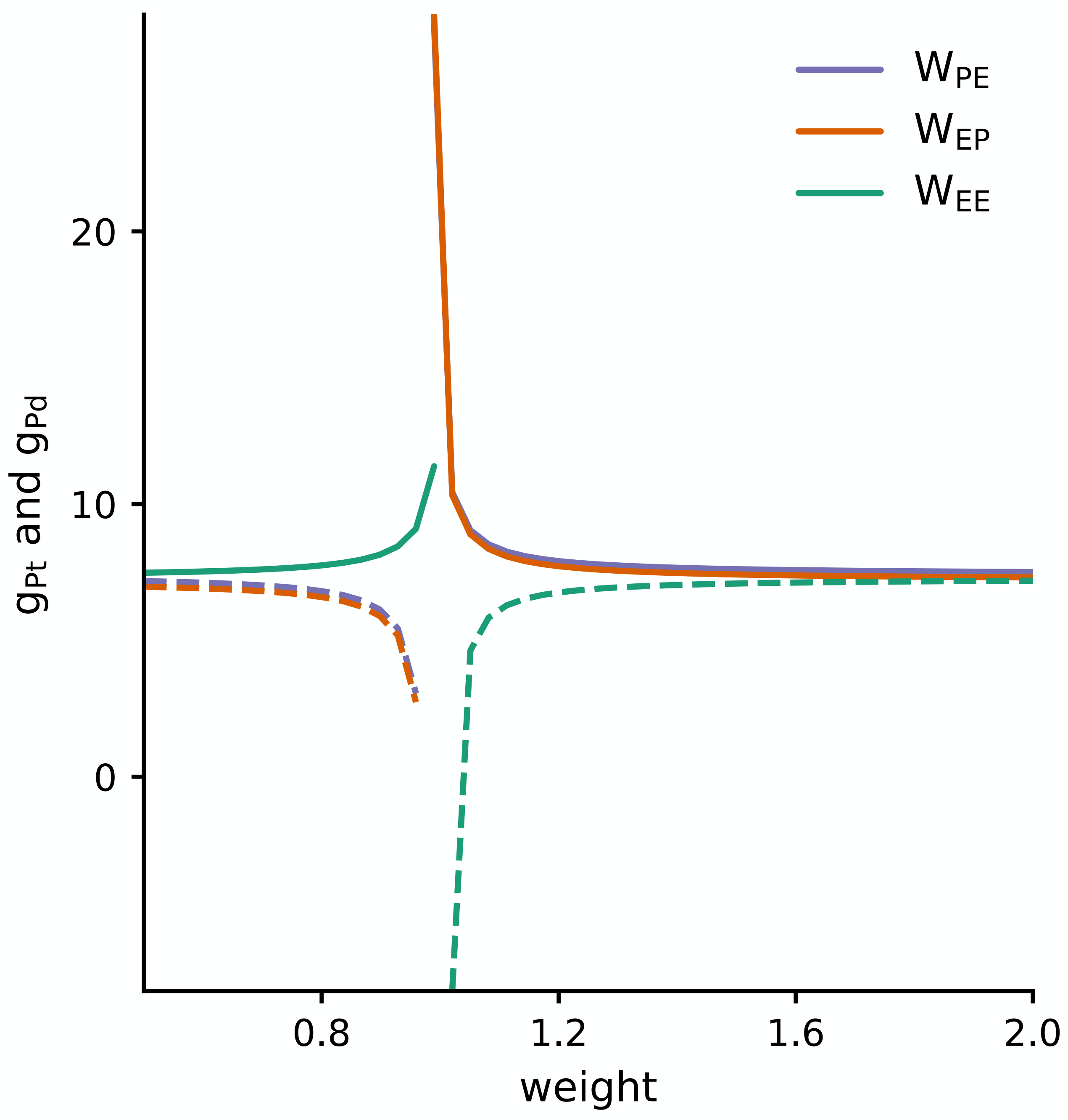


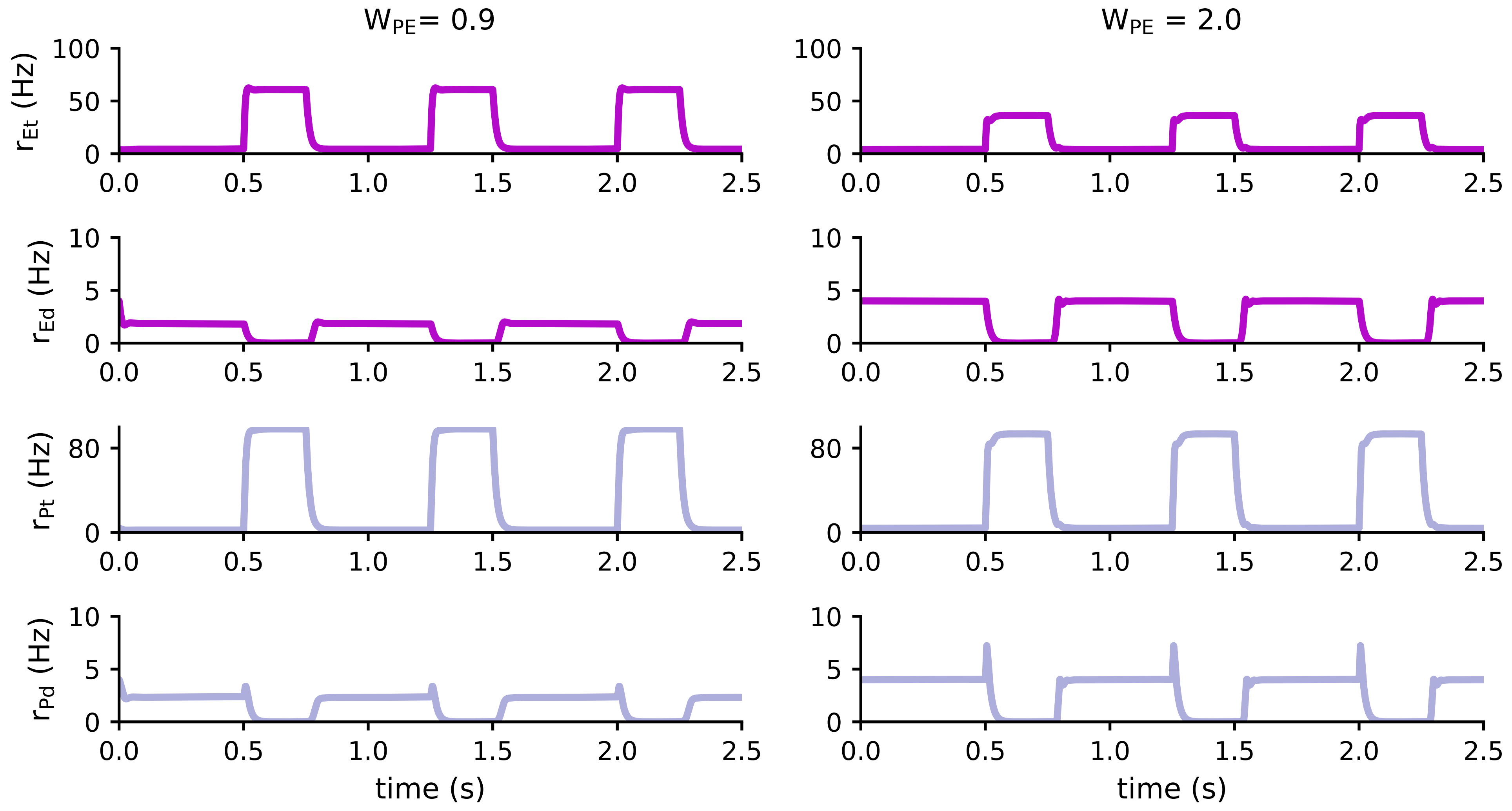
Counter-intuitive changes in response rates

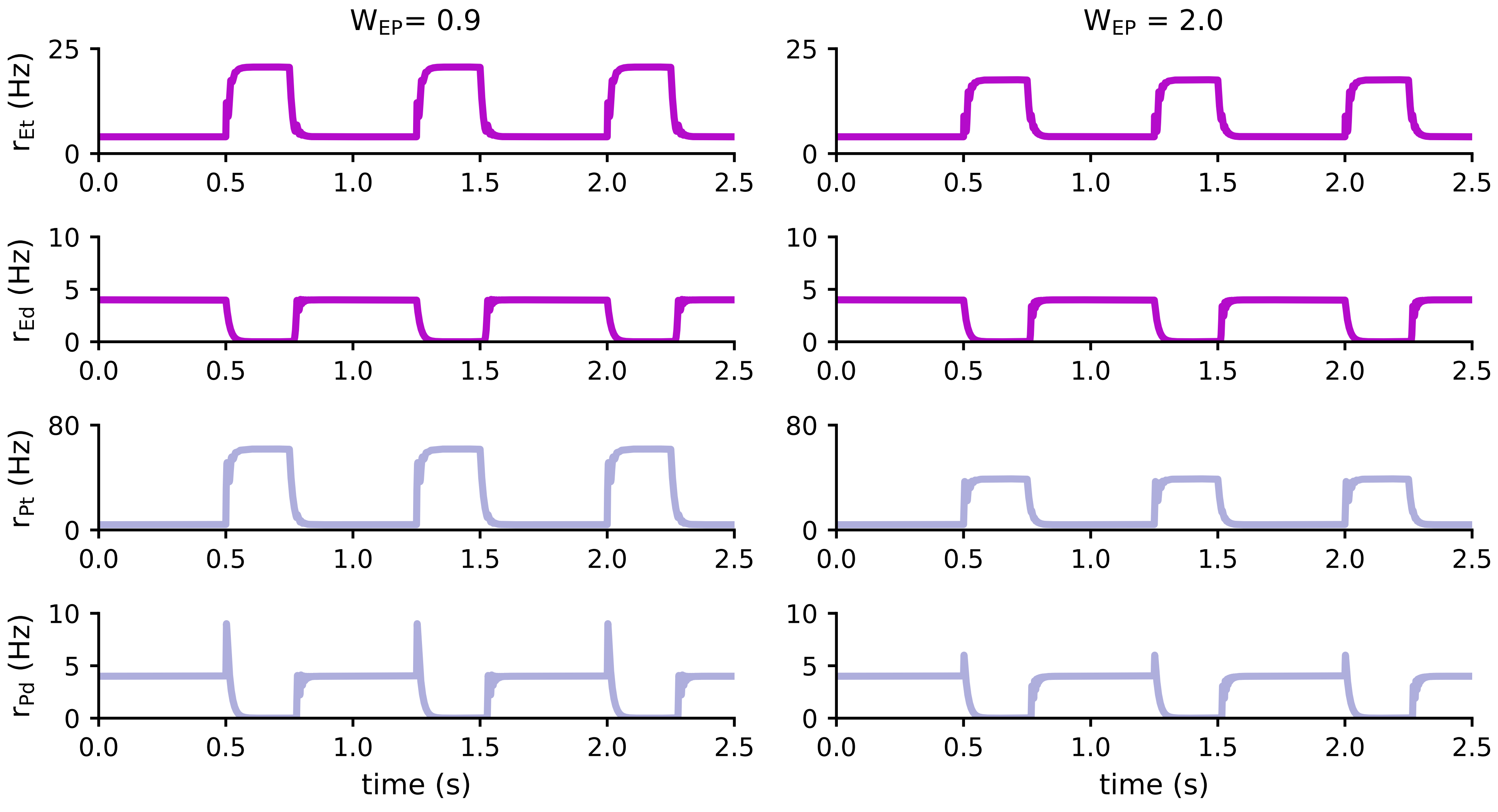
potentiating E<->PV connections decreases response gain for incoming stimuli

potentiating E->PV connections suppresses E responses



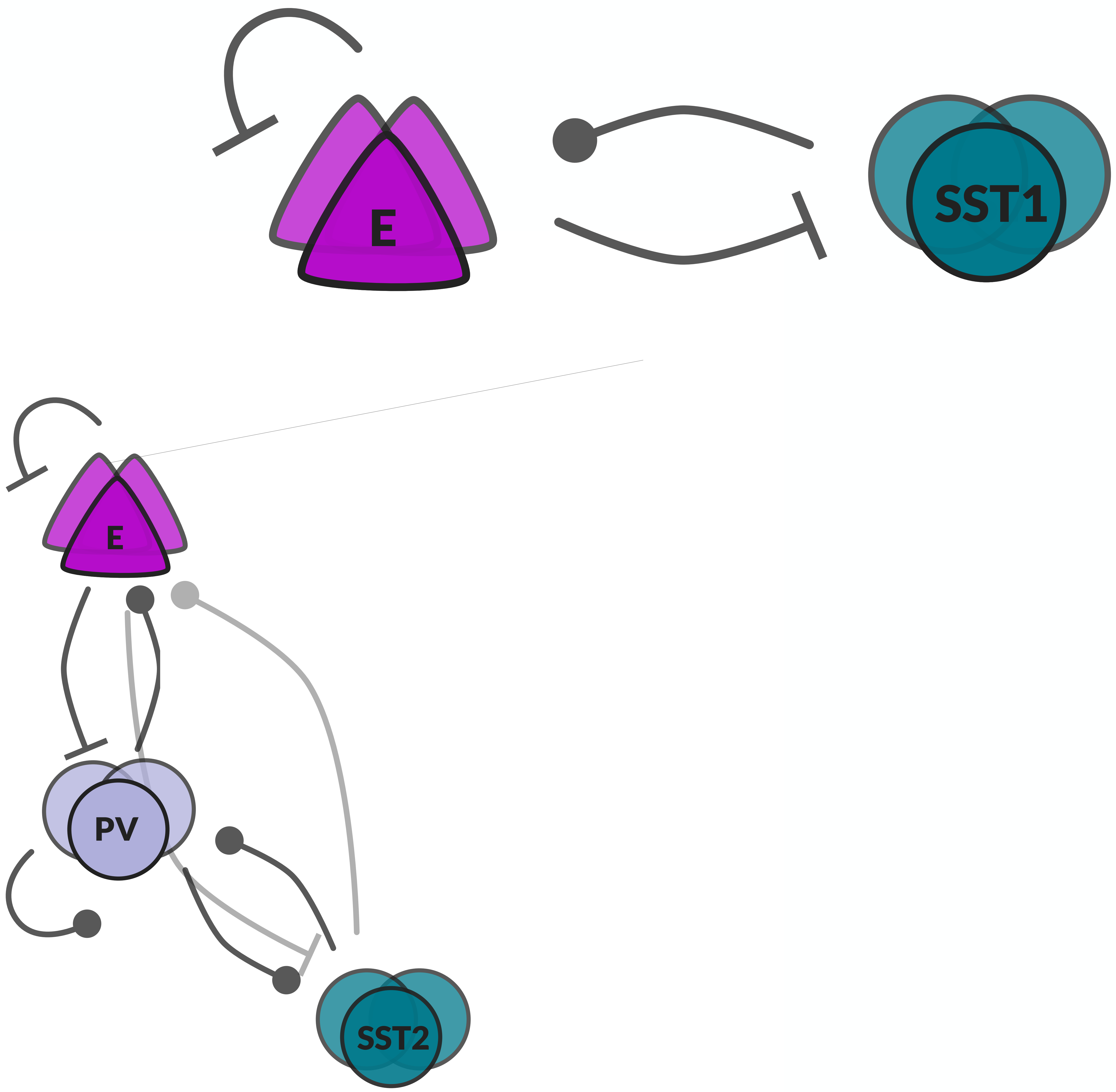






potentiating PV->E connections suppresses PV responses

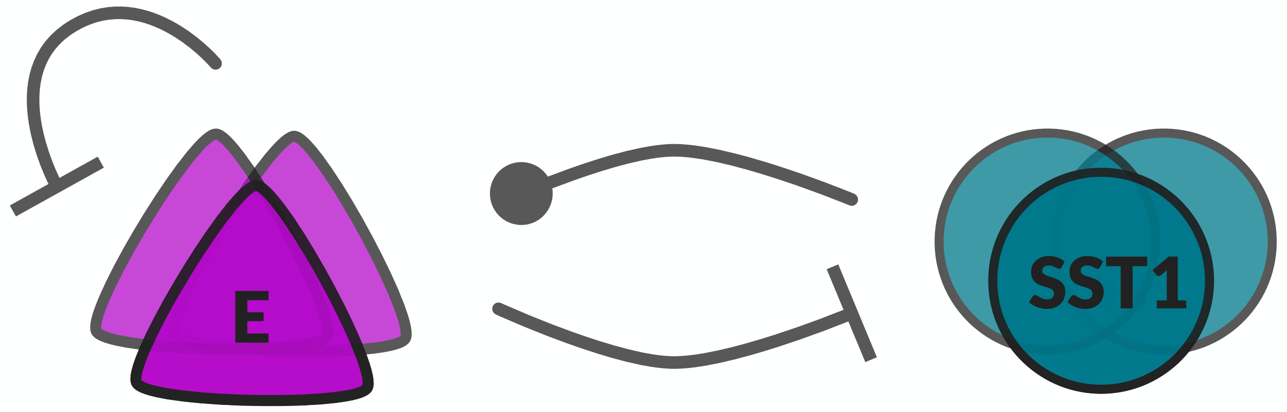
SST modulation

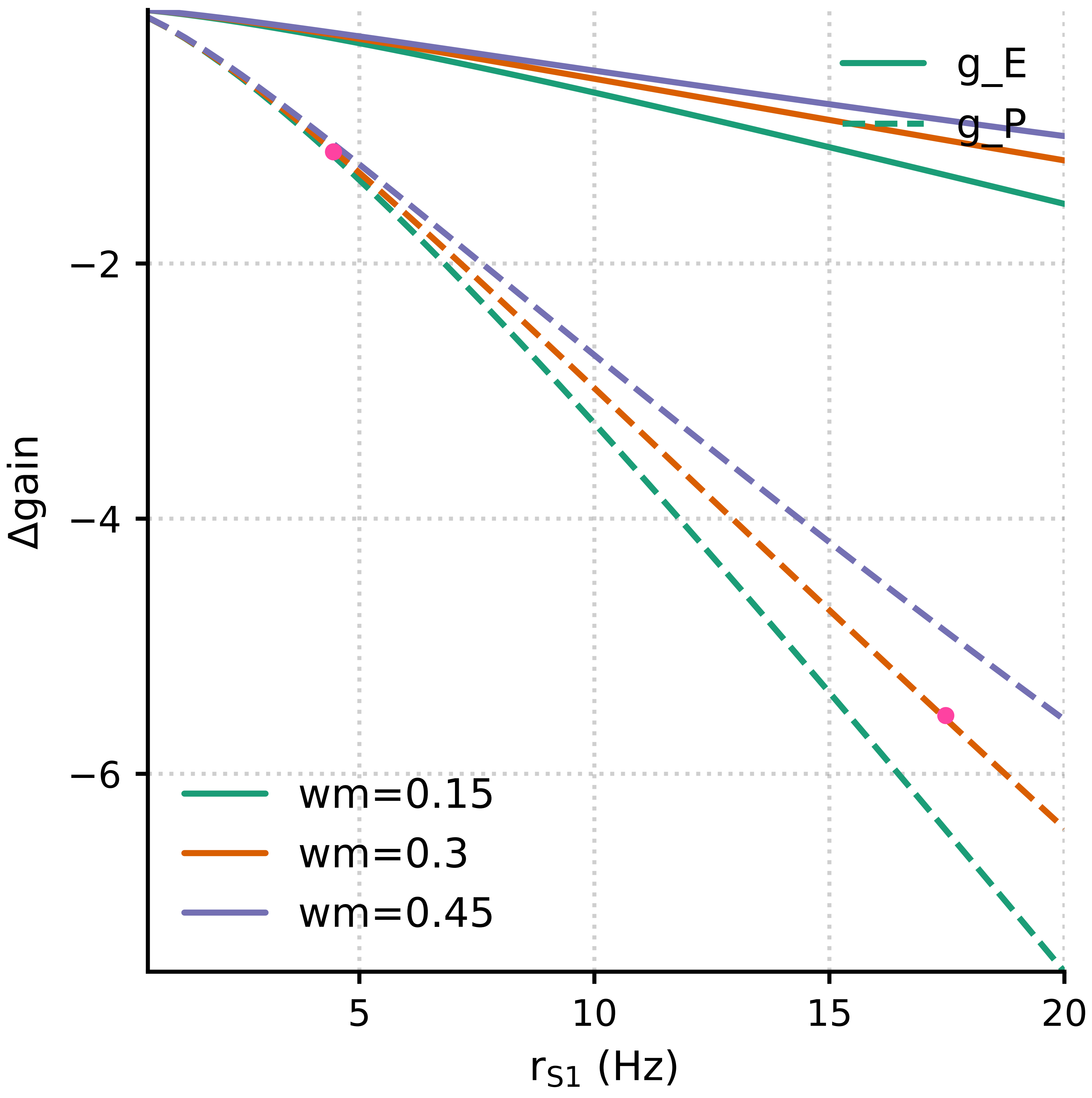


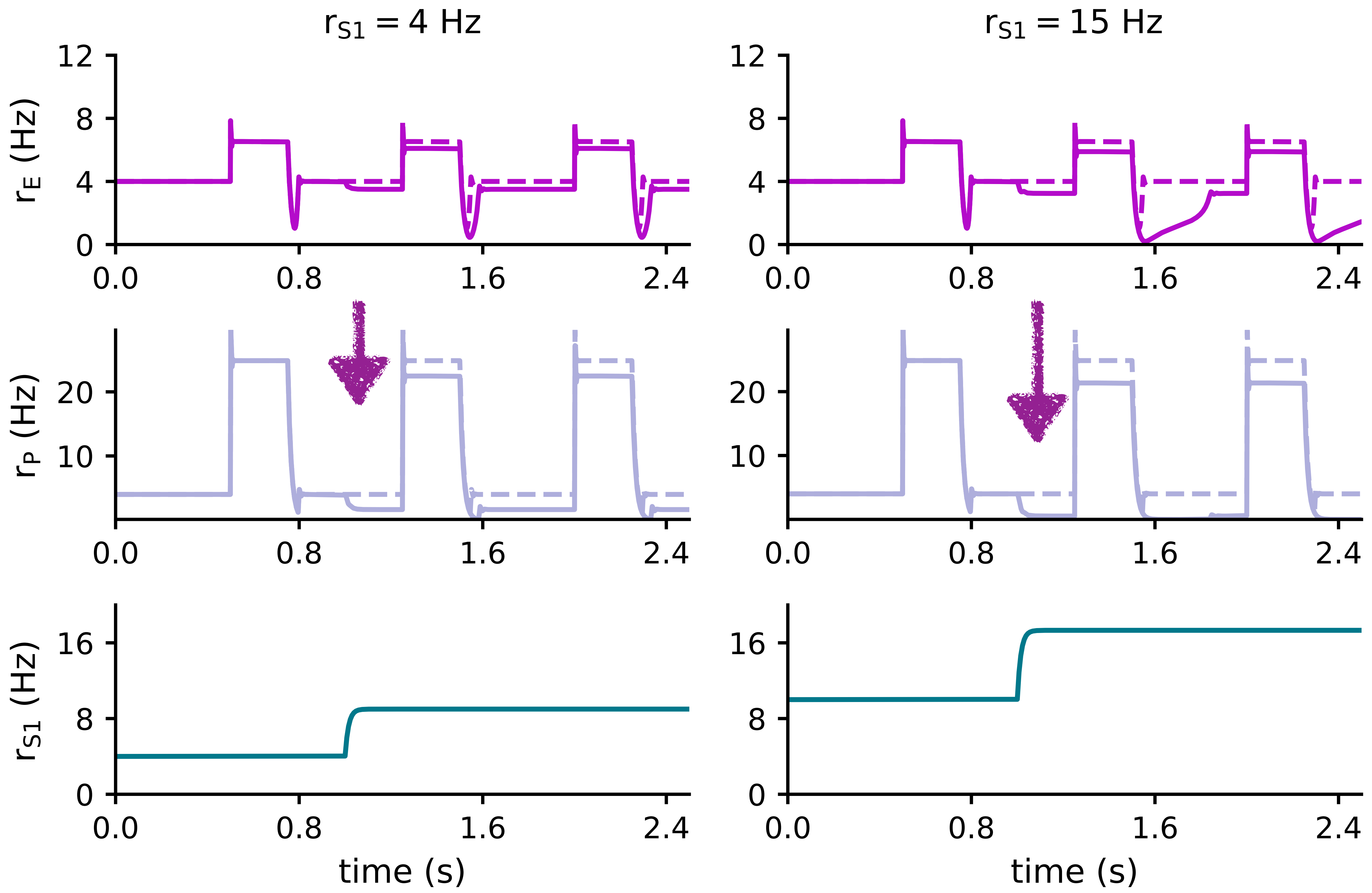
How does SST rate influence E and PV responses?

Inhibitory motif

SST modulation results in state dependent changes in network gain



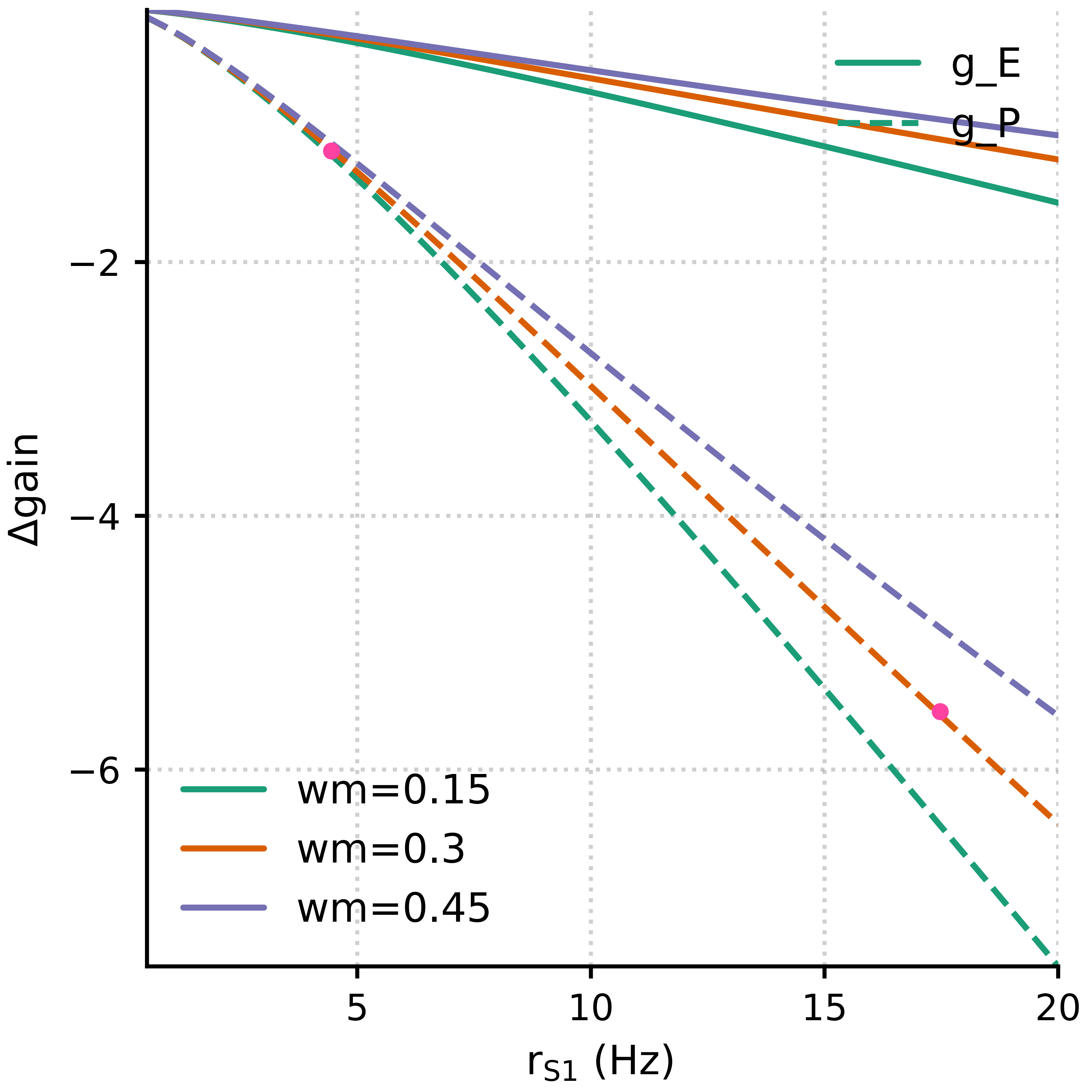




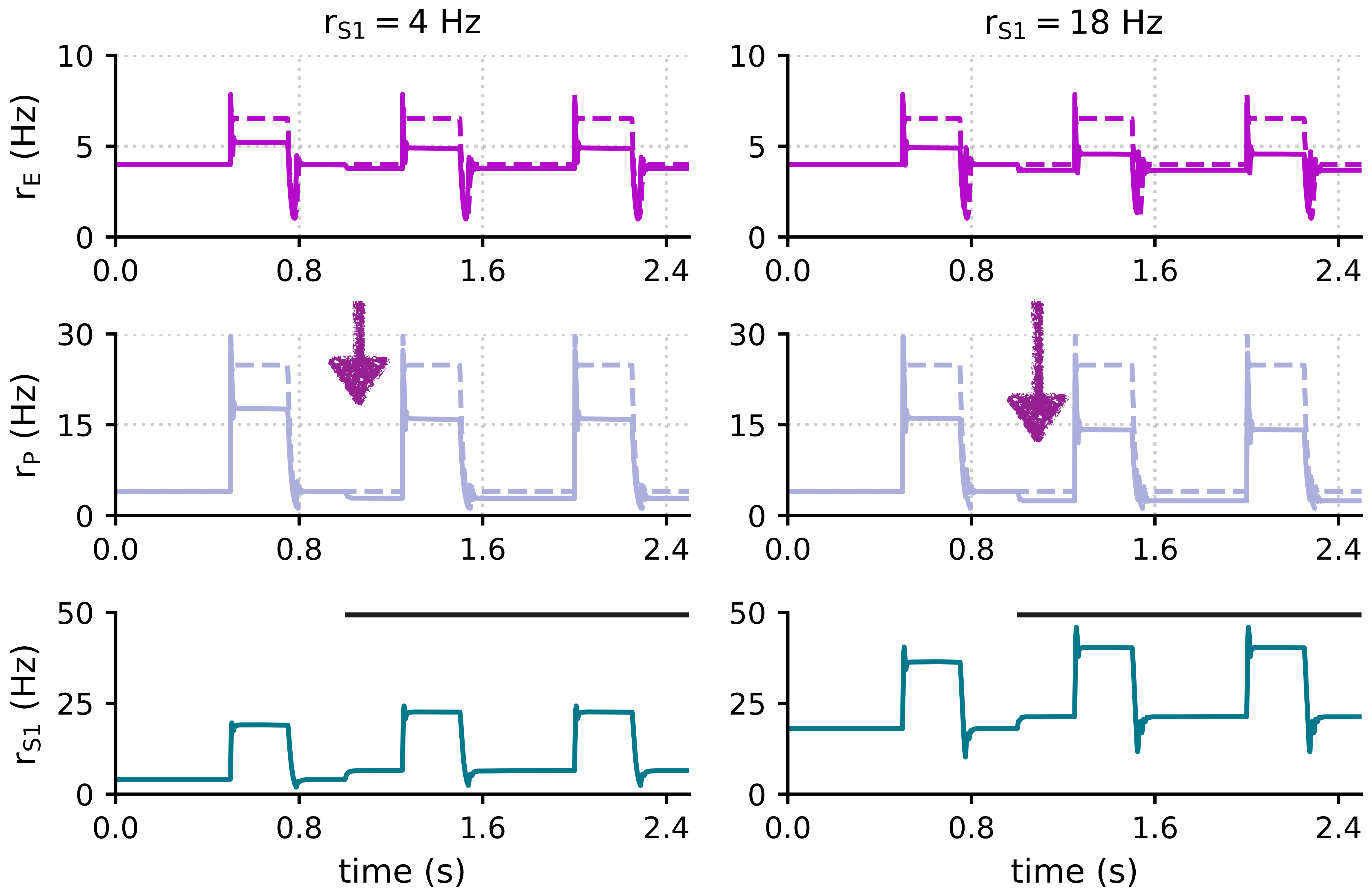
larger SST firing rates result in larger absolute changes in response amplitudes

Inhibitory motif

SST modulation result in state dependent changes in network gain

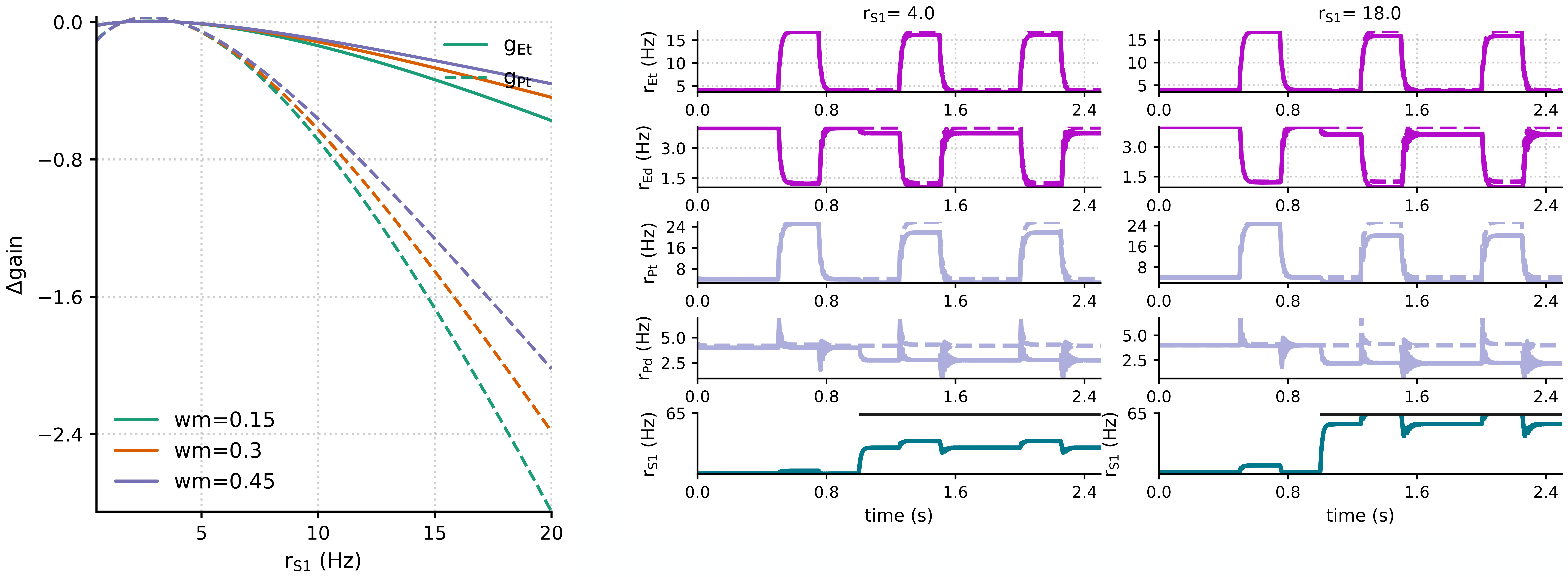


state dependent changes in
network gain



Inhibitory motif

Fractional gain change for modulation of SST1 sub-population



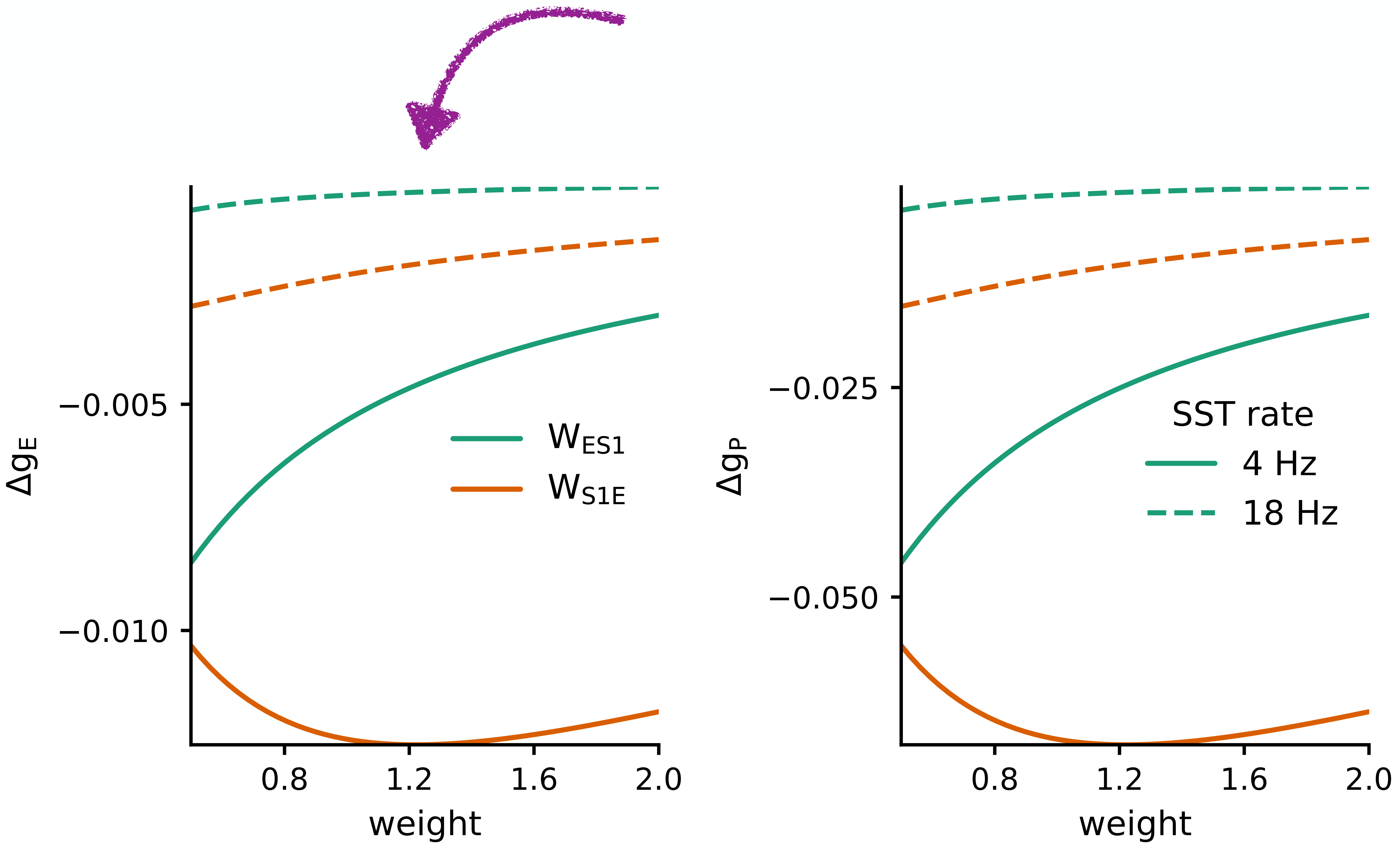
State-dependent influence of plasticity of E->SST on stimulus responses

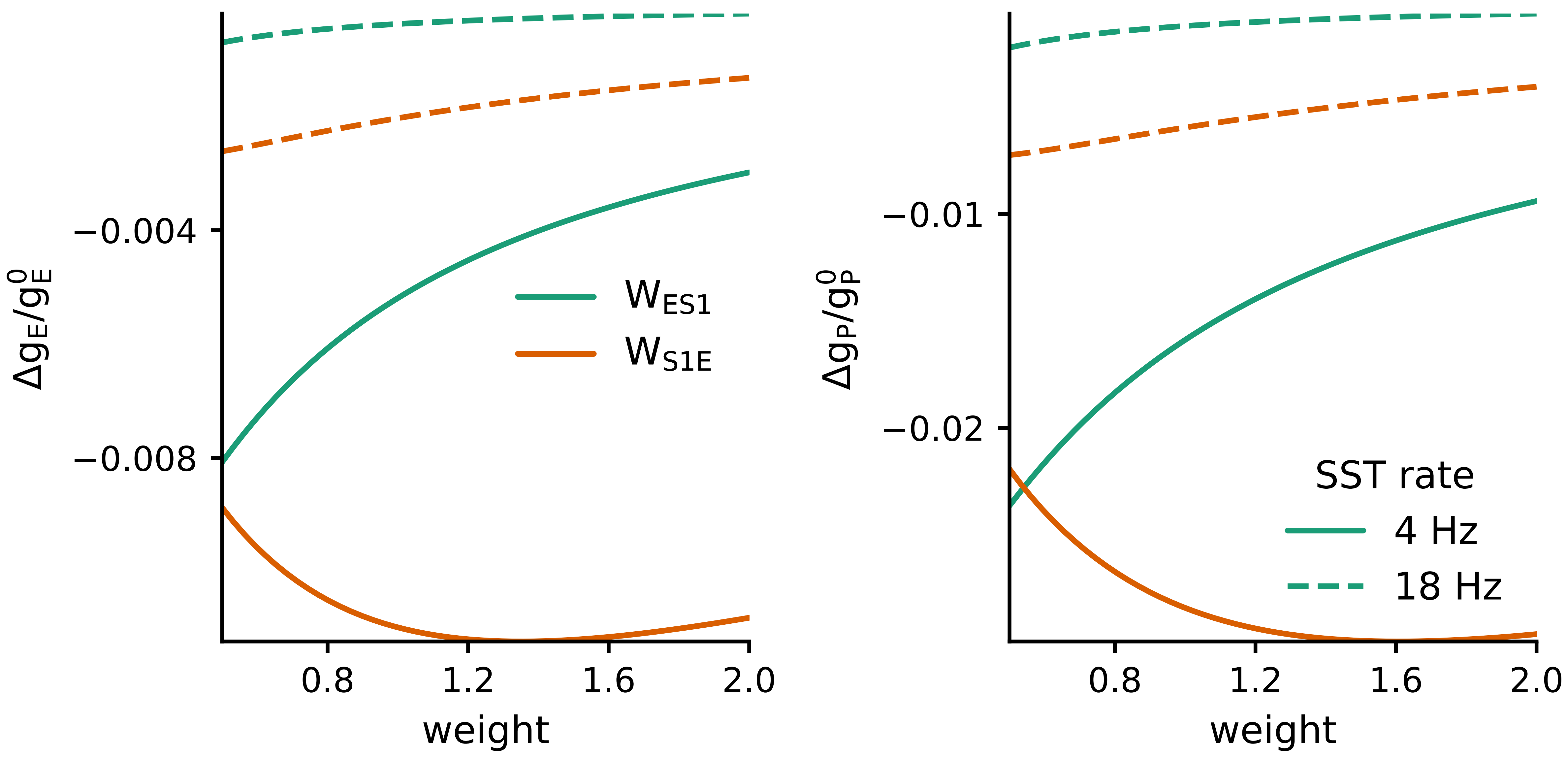
the effect of plasticity of E->SST1 synapse will
have state-dependent effects on the responses of E and PV cells

increasing SST rate reduces
the negative impact of SST1 on excitatory responses

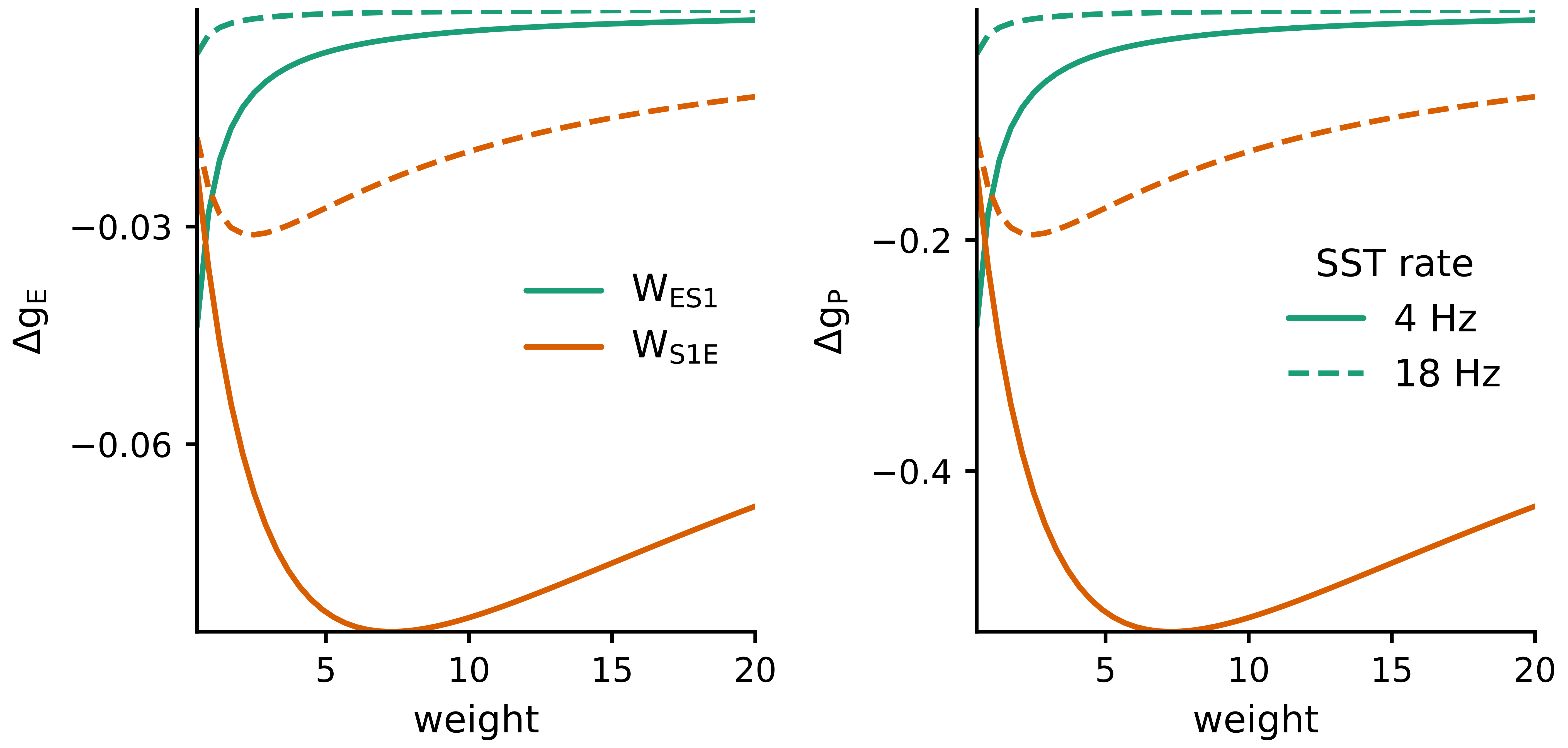
SST 1 population activity
can control the network gain
from what is expected
from feedback inhibition

almost no reduction in gain
for high SST rates

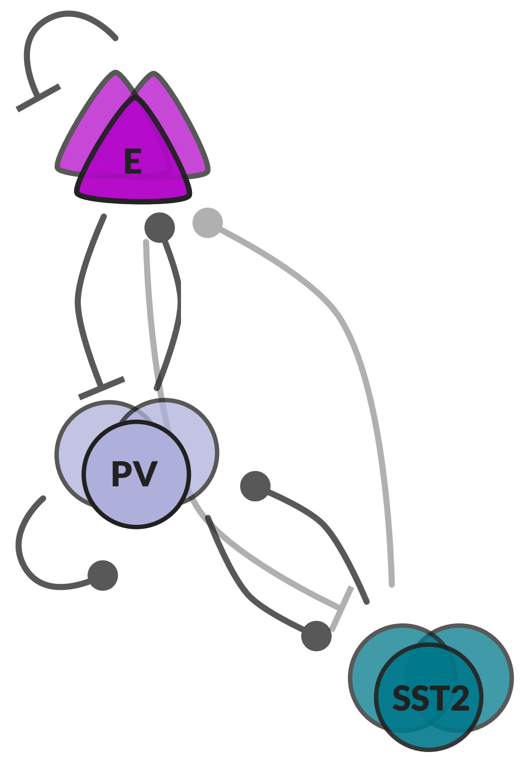


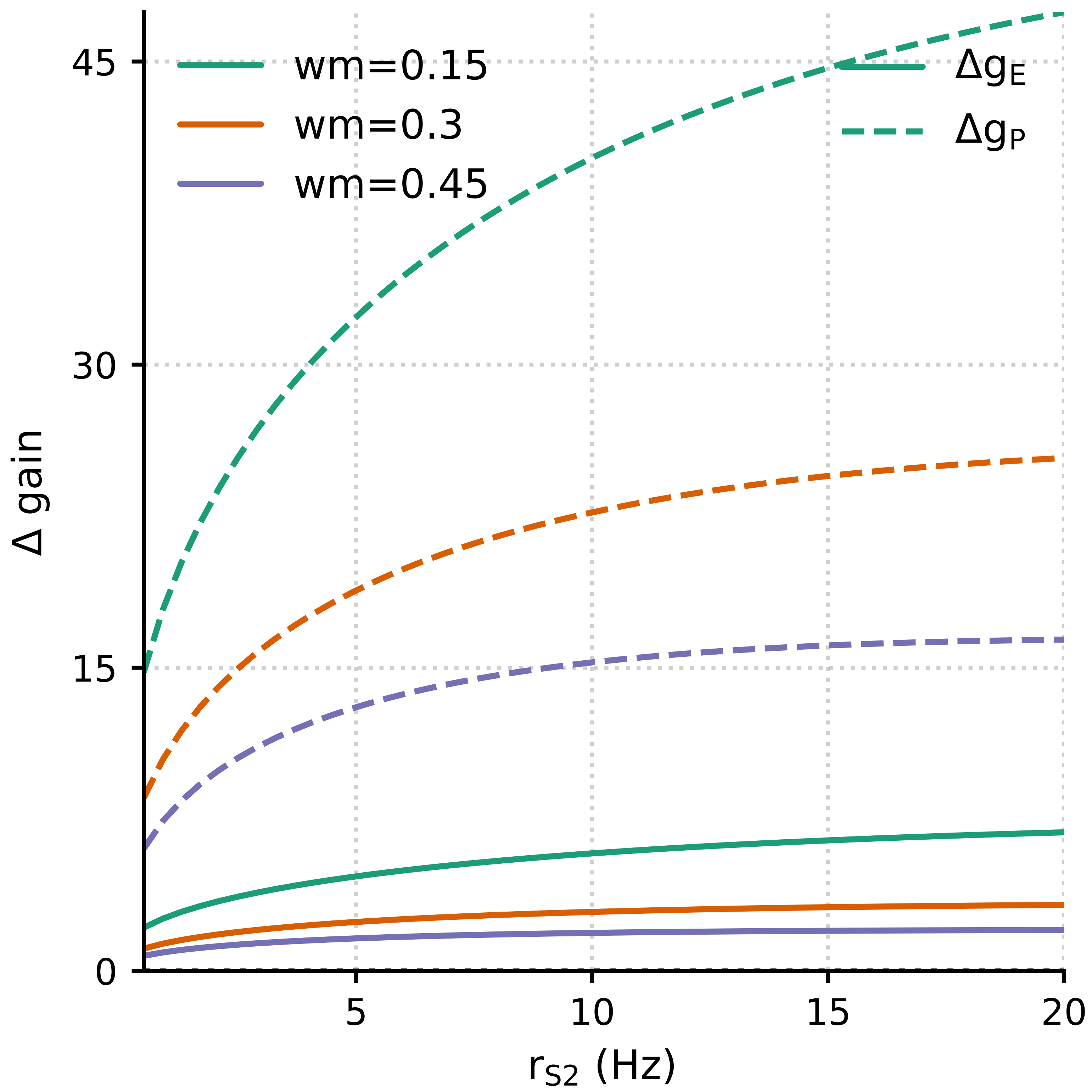


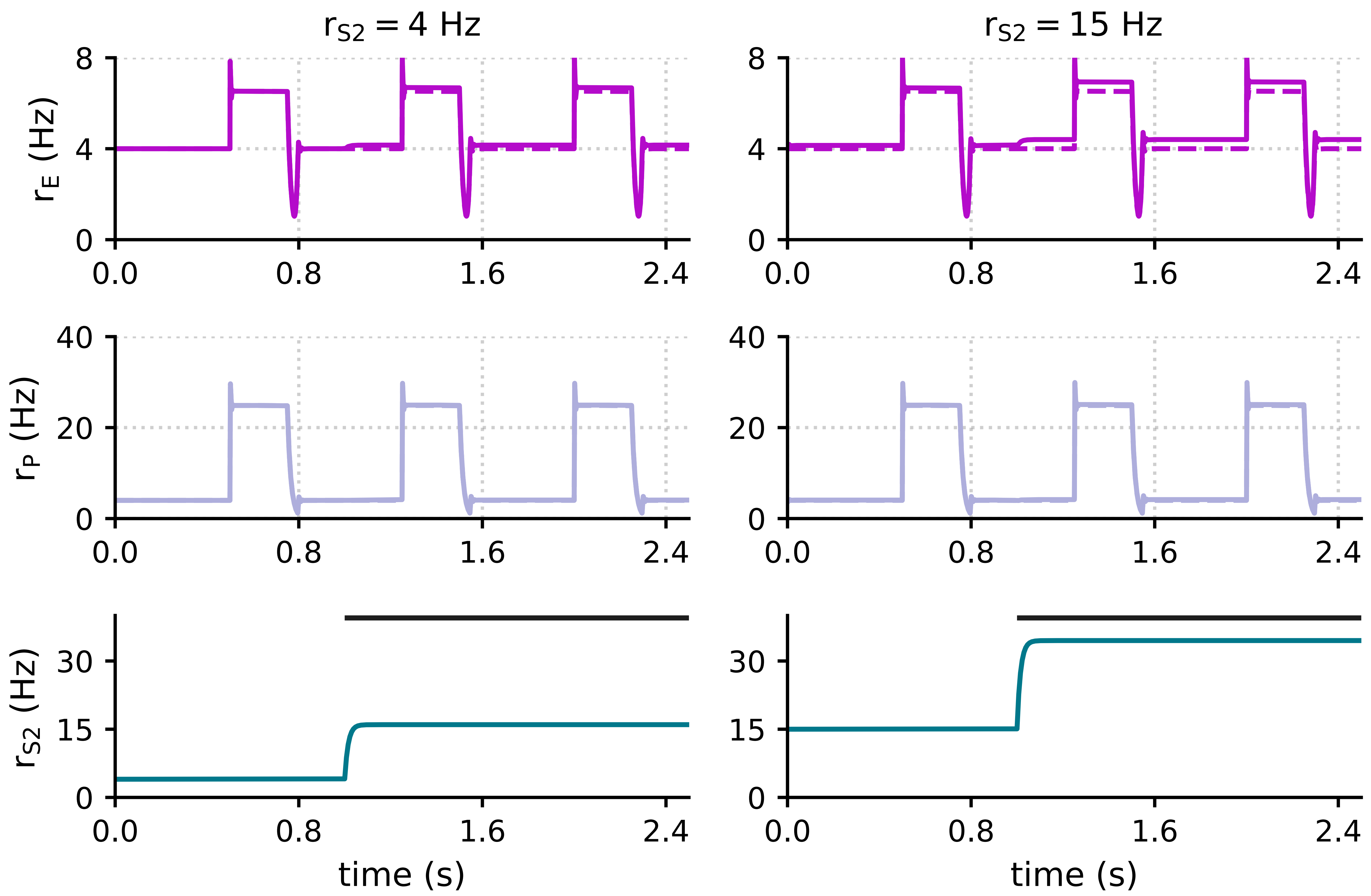
Fractional version



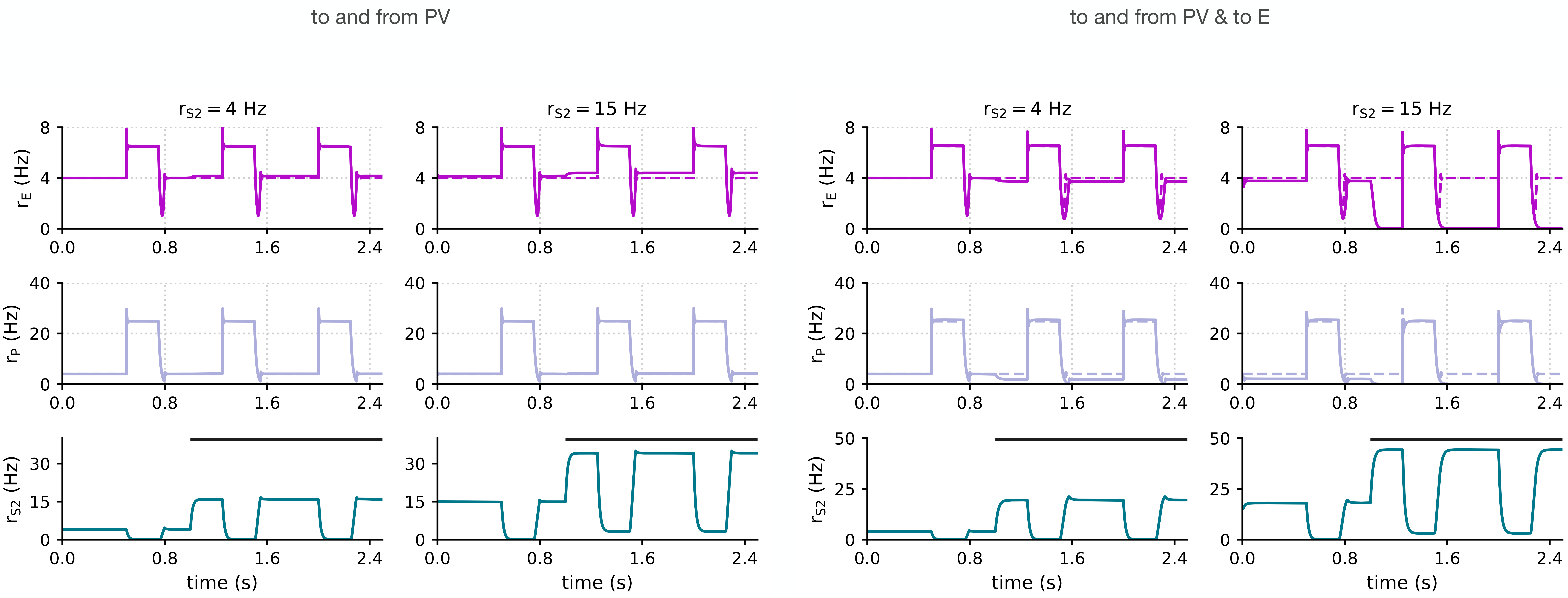
Disinhibitory motif





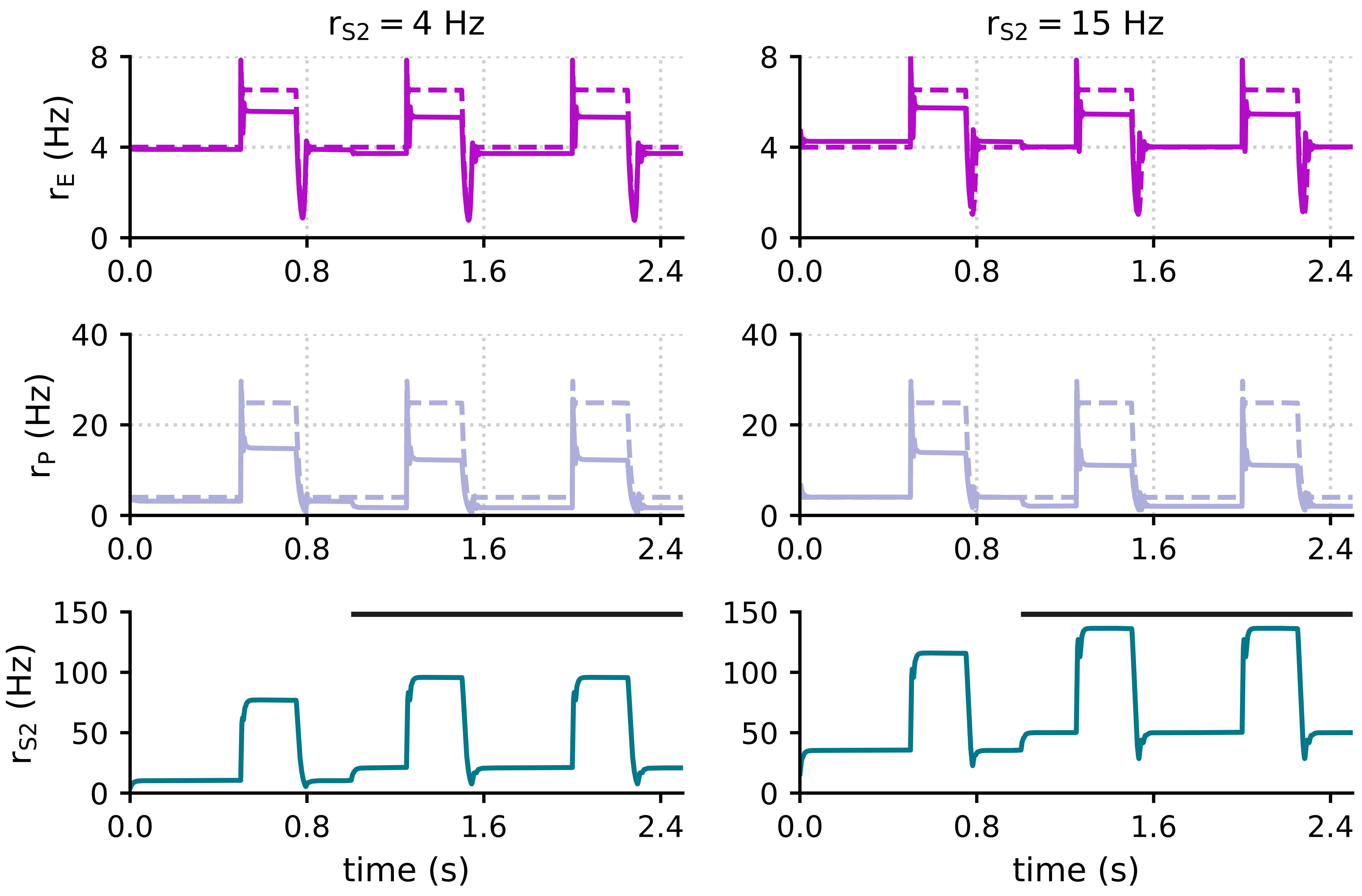


Disinhibitory motif



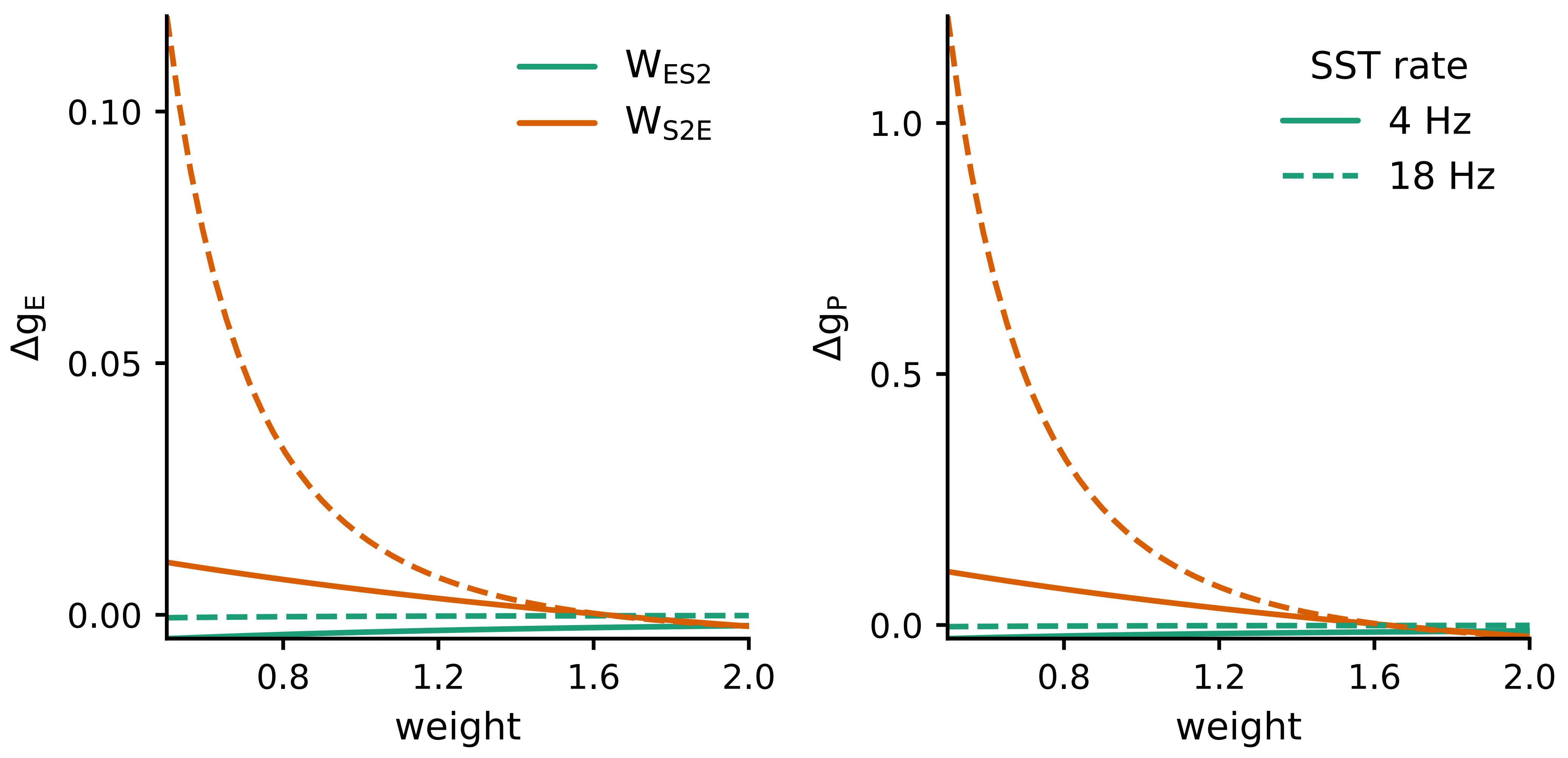
Disinhibitory motif

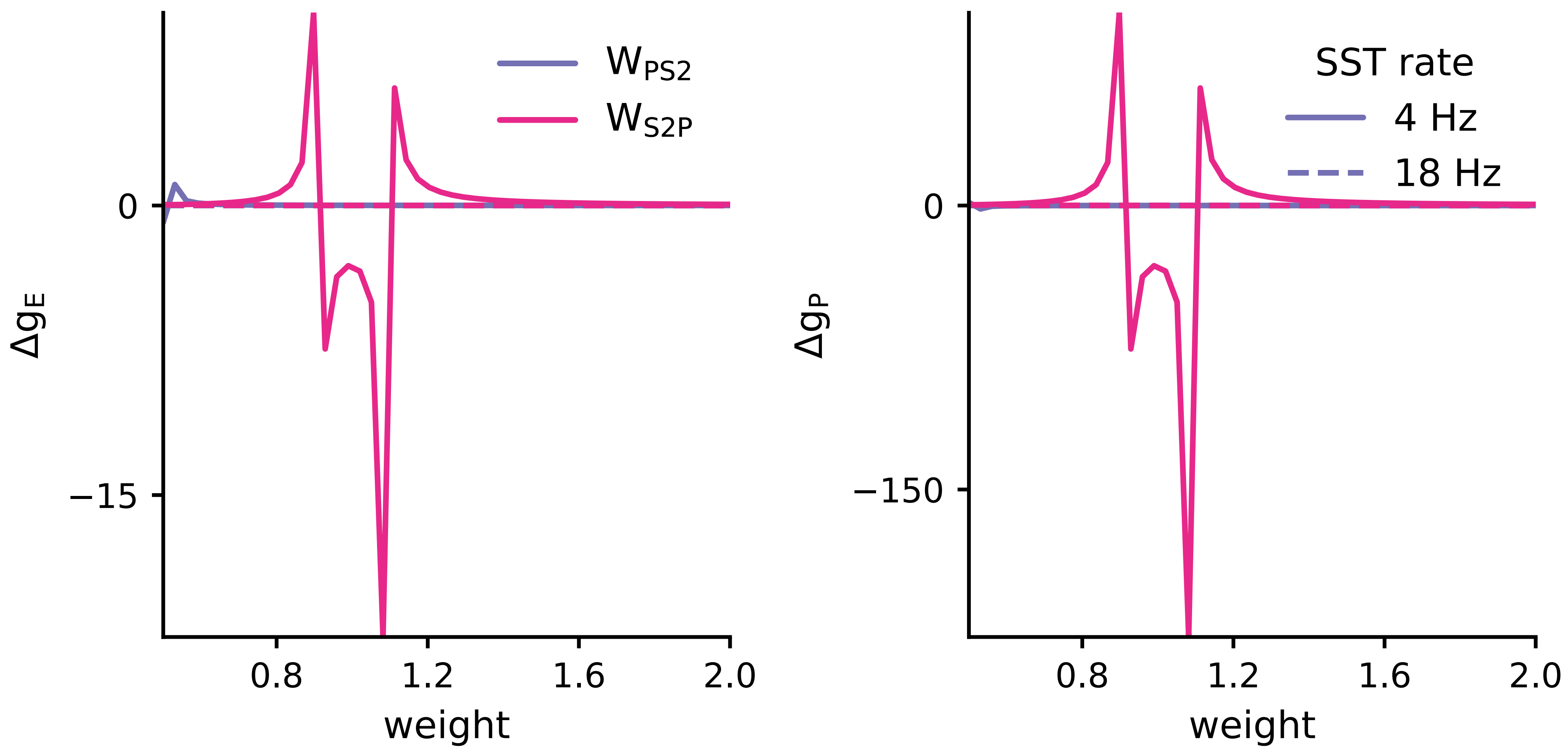
to everyone



Disinhibitory motif

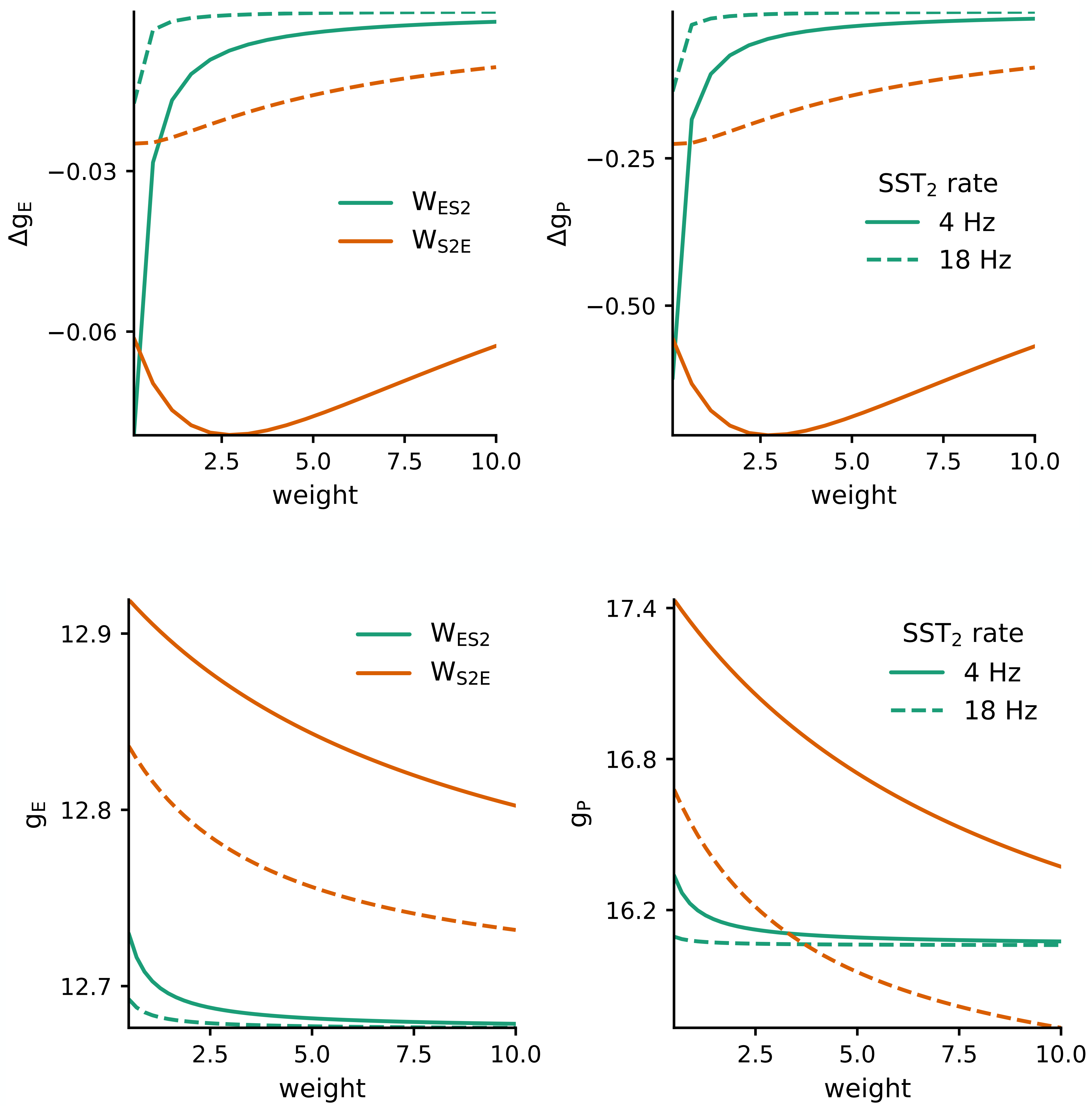
Gain change for modulation of SST2 sub-population

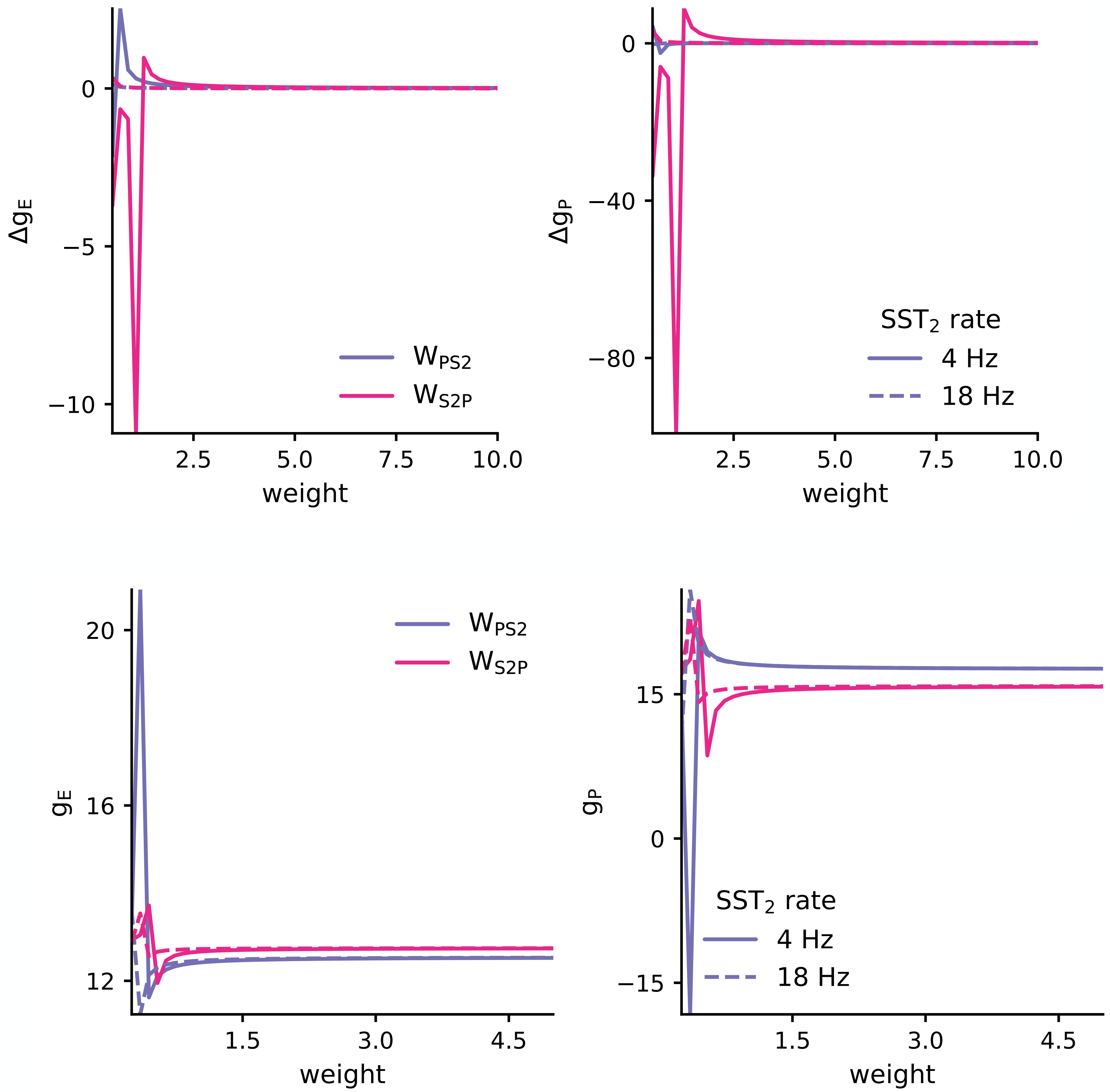




Disinhibitory motif

Fractional gain change for modulation of SST2 sub-population





Model

familiar no change

- SSTs suppress PV & E
- top-down feedback to E - (buffer)
- top-down feedback is inhibitory on E to cancel out feedforward excitation

familiar change

- SSTs suppress PV & E
- top-down feedback to E - previous buffer
- error in top-down feedback suppresses some E and the E that correspond to the changed stim show pronounced responses

kick on VIP and SST1

novel no change

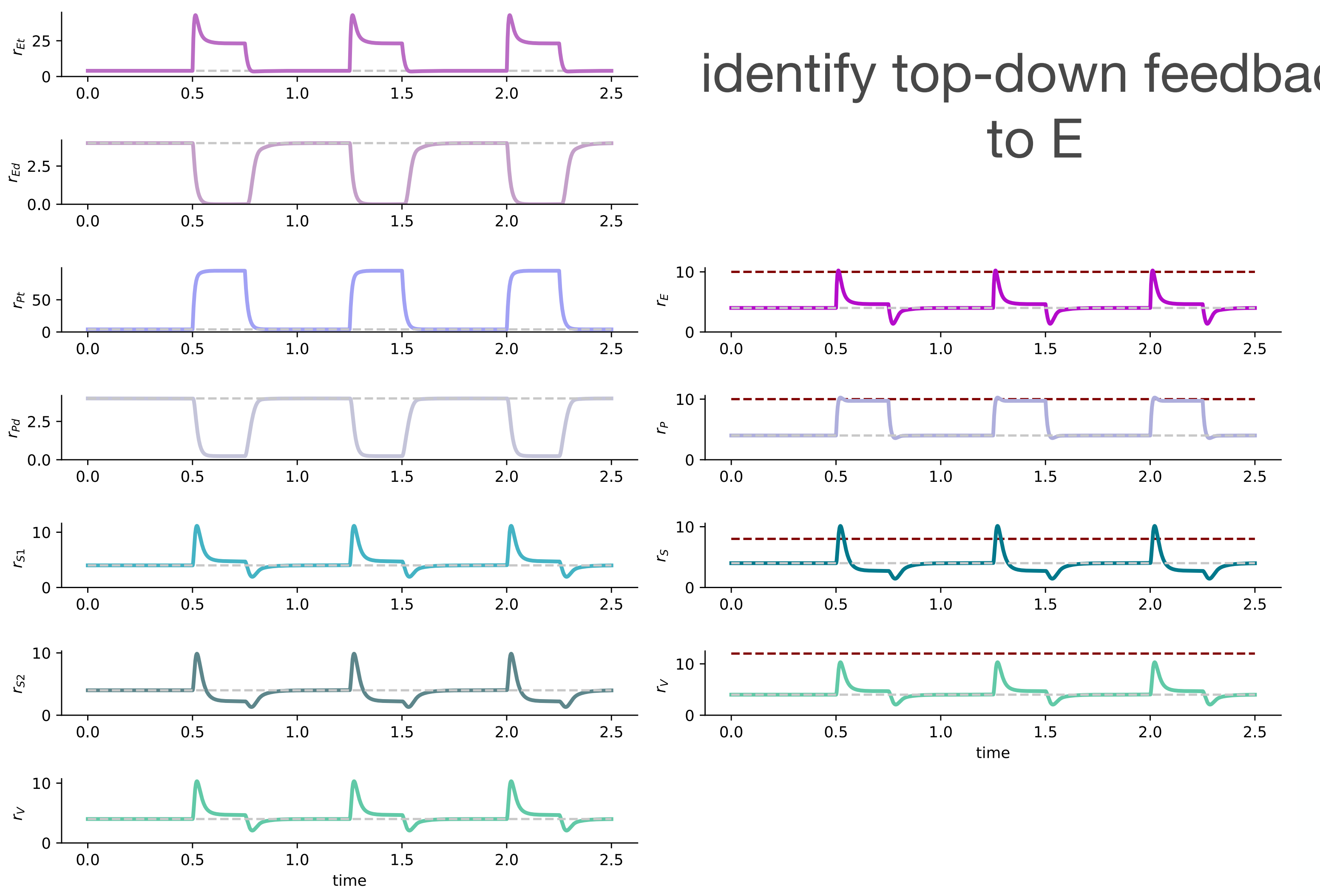
- PV and E are active - not yet suppressed by SSTs
- top-down feedback to E - previous buffer - but with less amplitude
- top-down feedback is inhibitory on E to cancel out part of feedforward excitation
- the small peak on VIPs may be due recurrent excitation from E and not due to top down

novel change

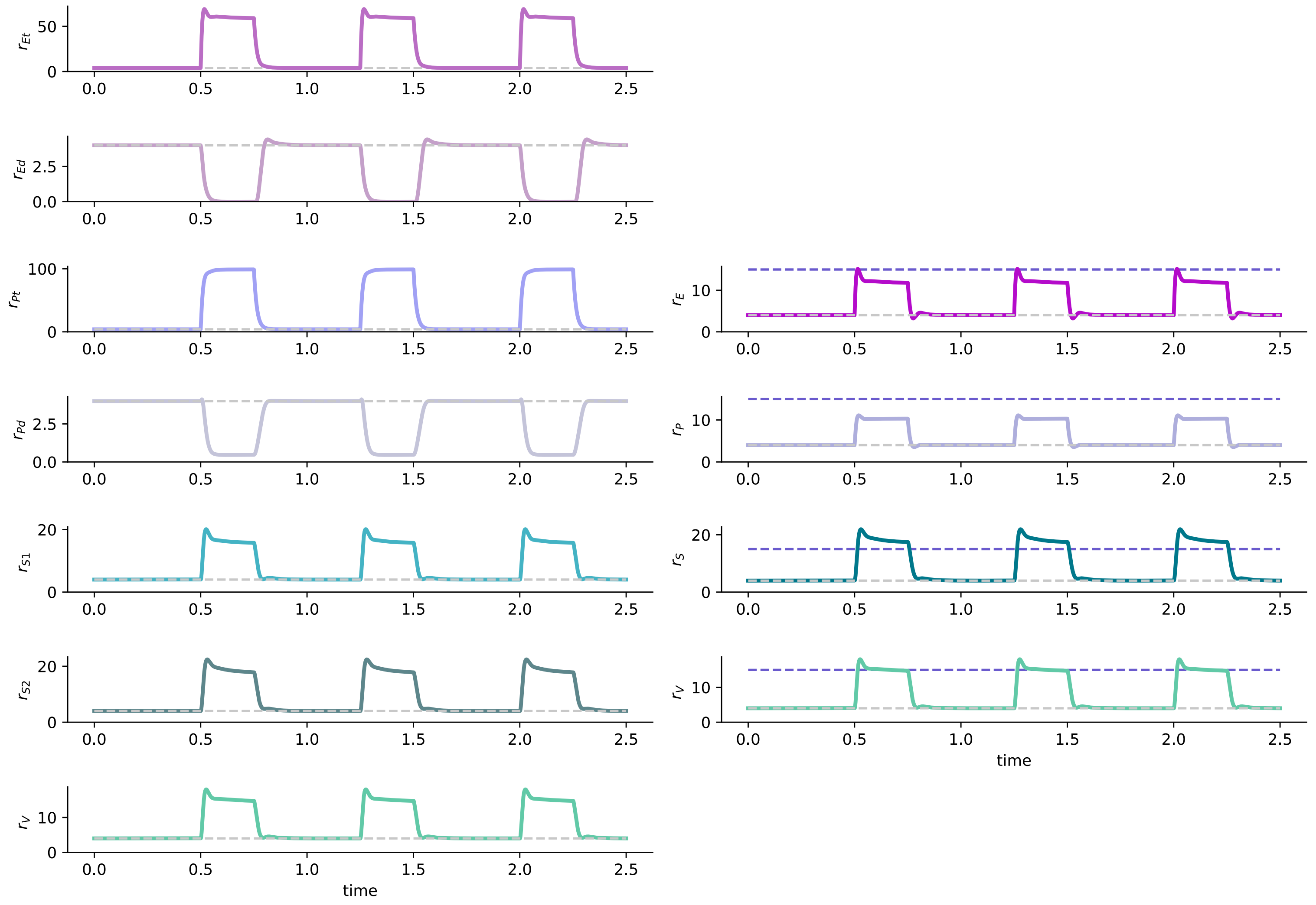
- PV and E are active - not yet suppressed by SSTs
- top-down feedback to E - previous buffer - now does not match input
- kick on VIP and SST1

Model

familiar no change

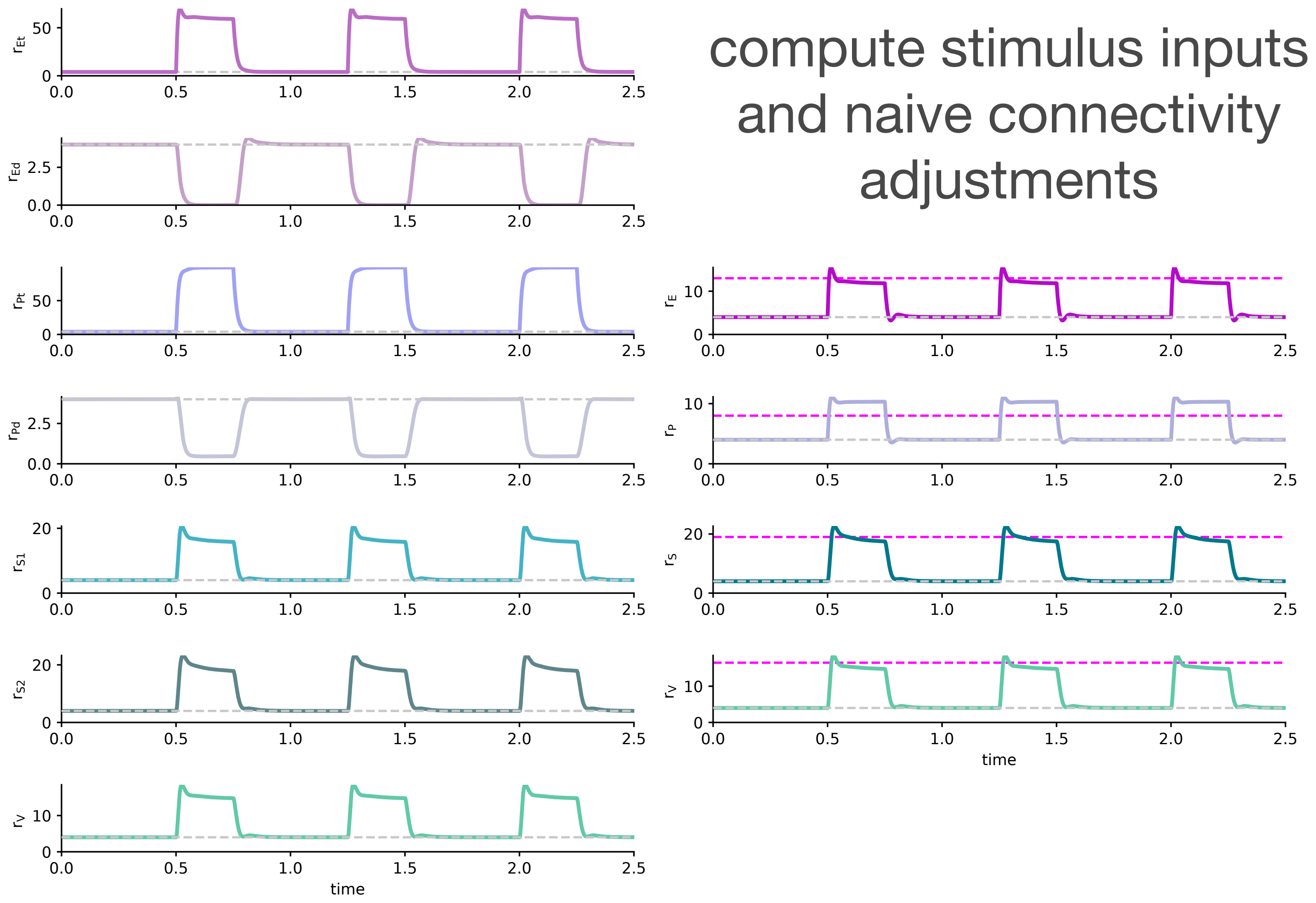


identify top-down feedback to E

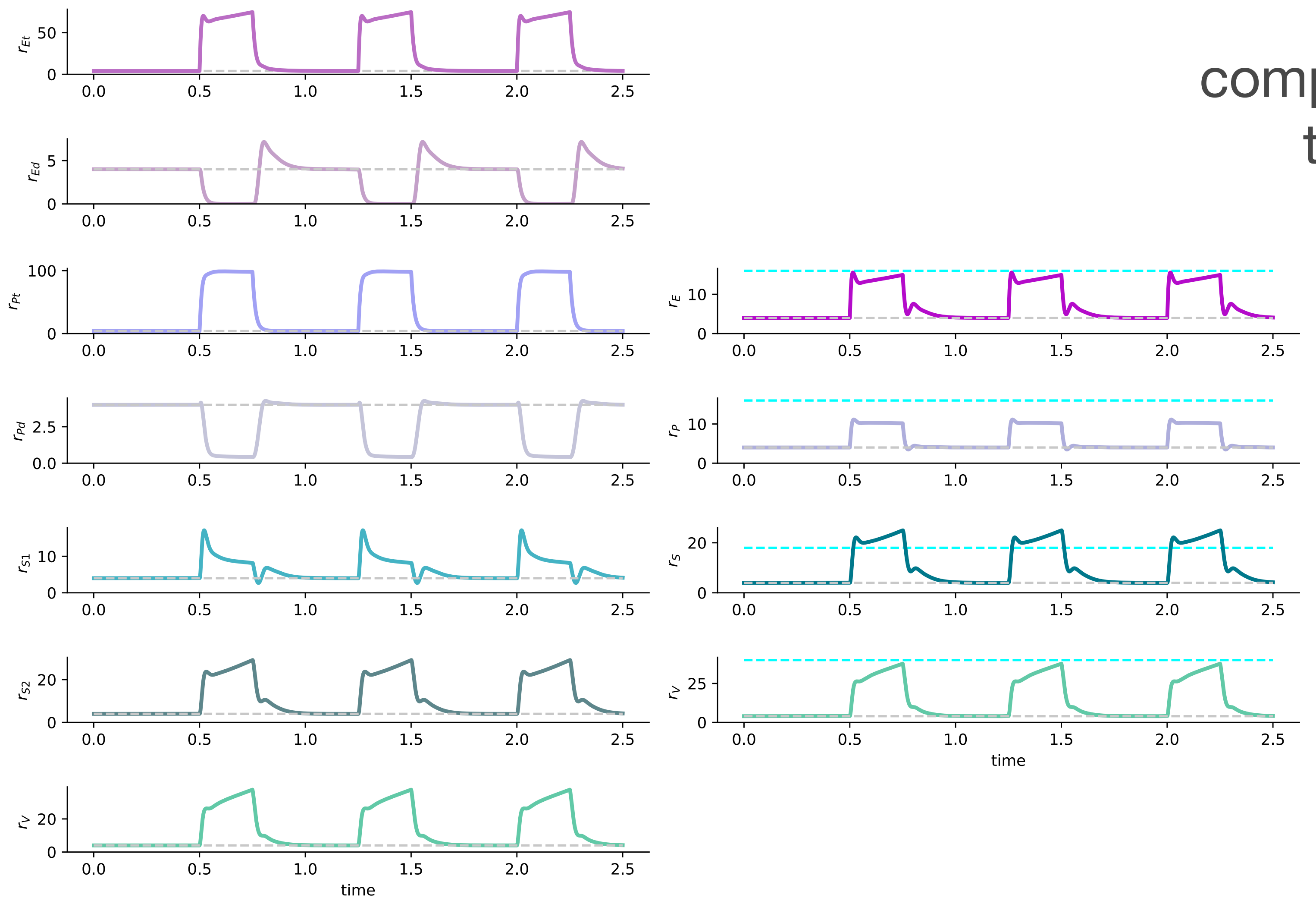


familiar change

novel no change



compute stimulus inputs and naive connectivity adjustments



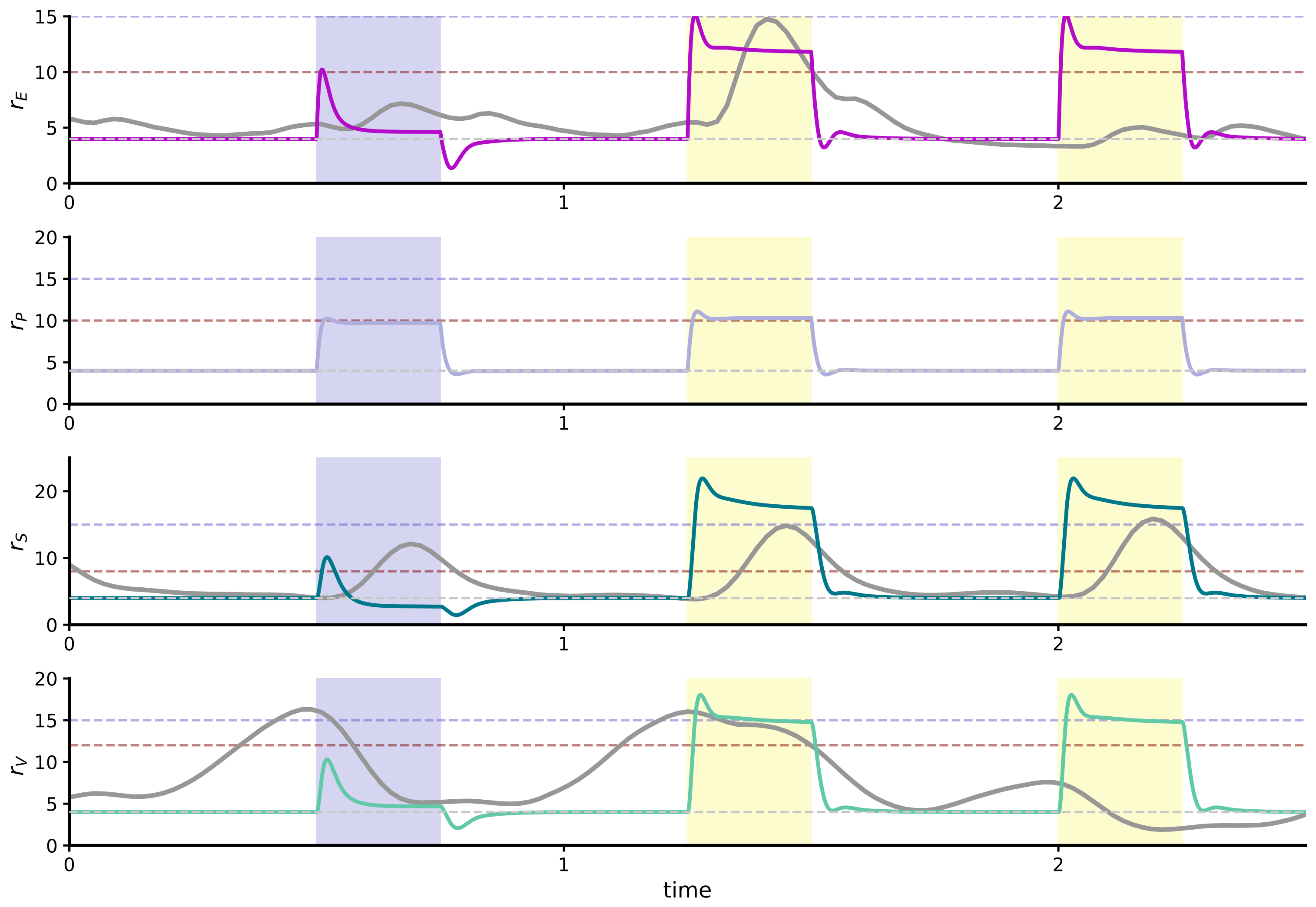
compute thalamic input to VIP and SST

novel change

Model

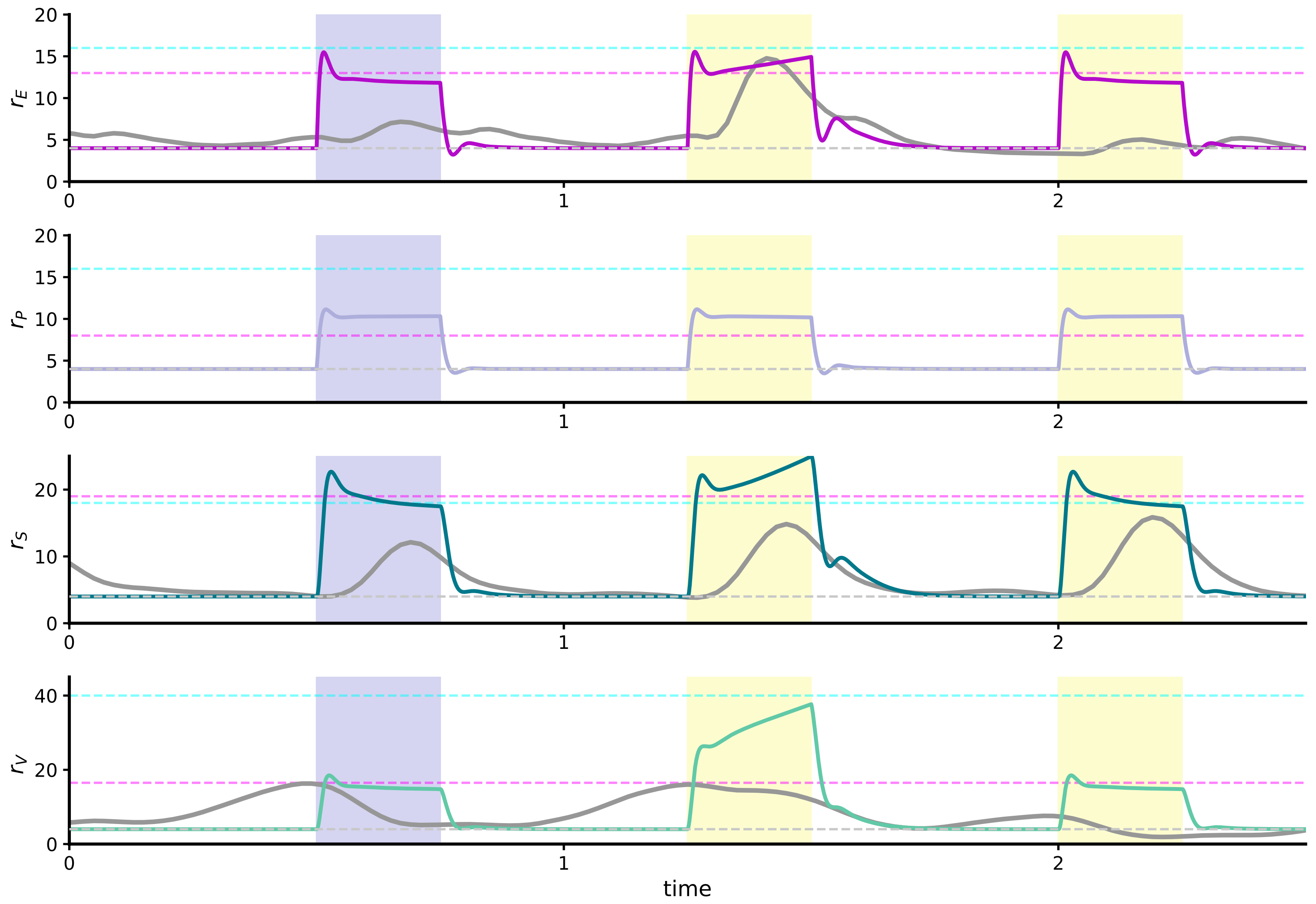
familiar
no change

familiar
change

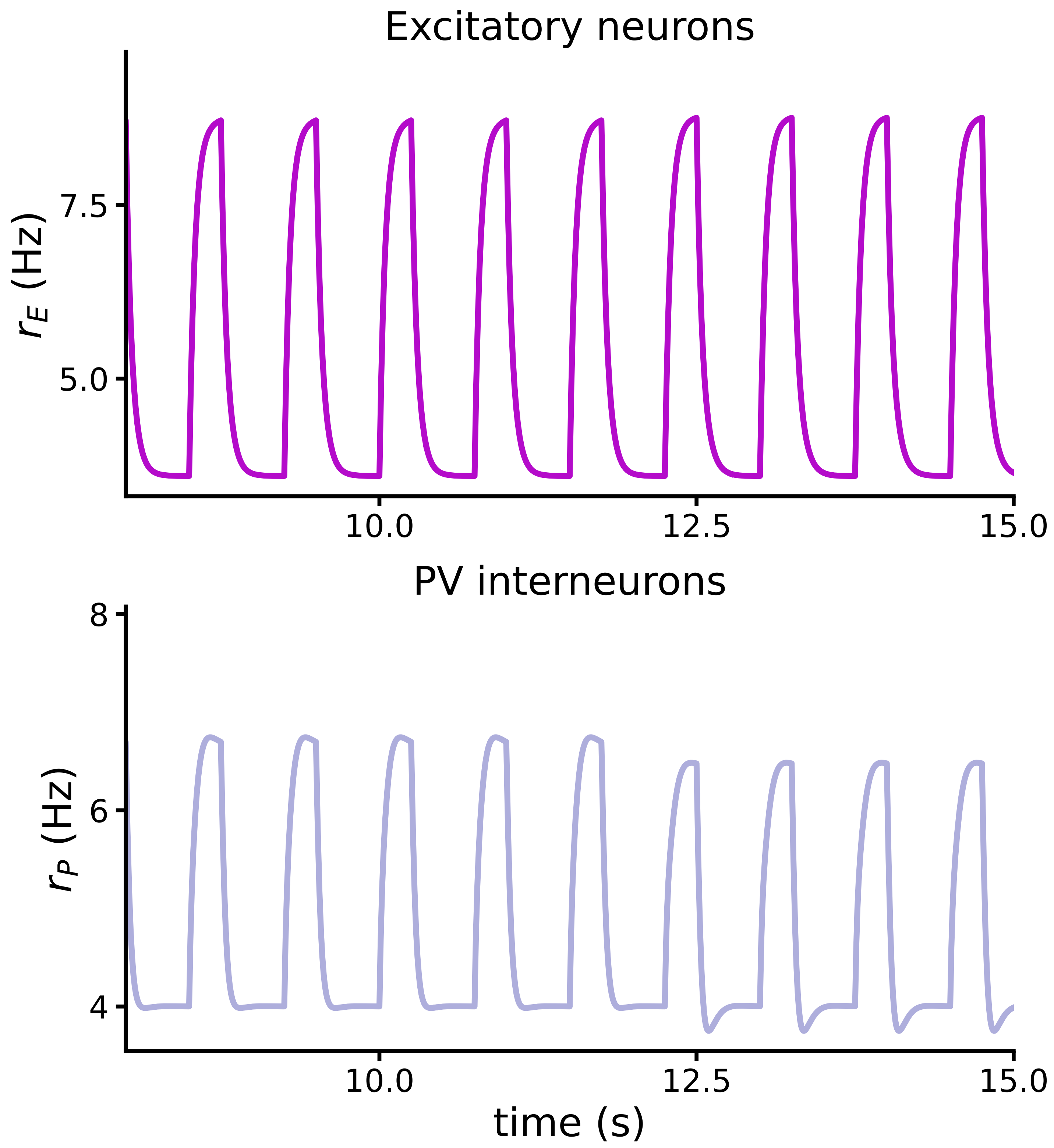


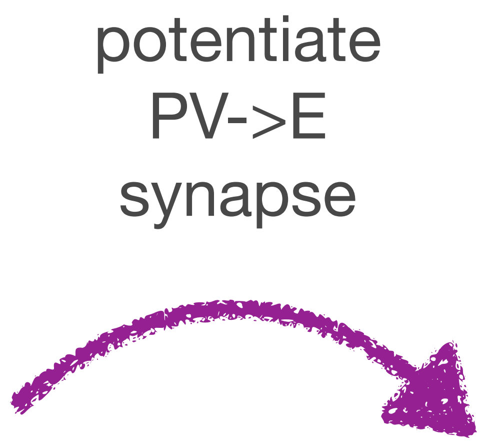
novel
no change

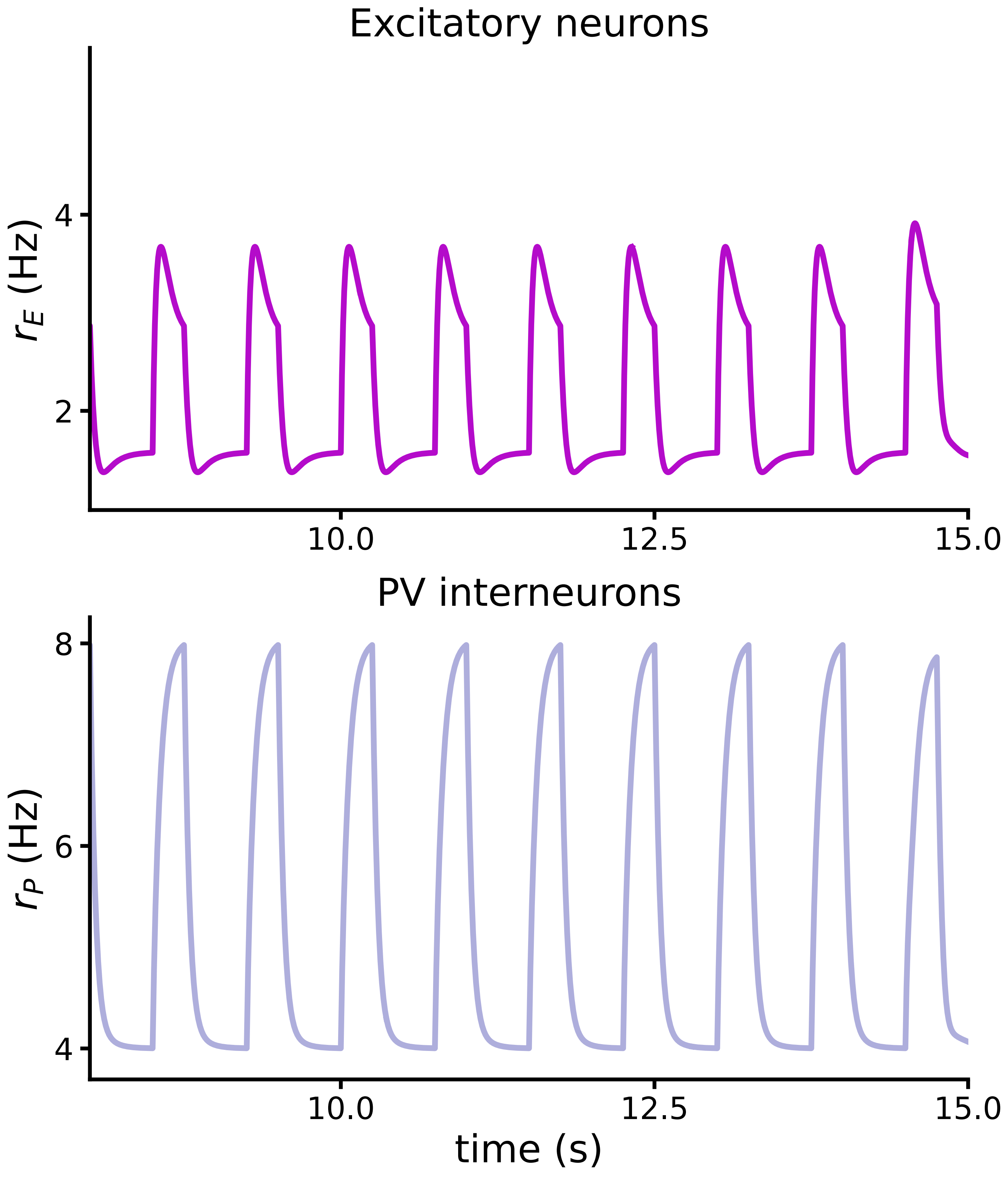
novel
change



Microcircuit model
start small







matches responses of
naive network

Summary

- Adaptation to stimuli and increased responses to novel stimuli require learing and plasticity
- Learning could be happening both recurrently in V1 or/and in higher order cortical areas
- Ablation/perturbation experiments required to disambiguate different predictive coding hypotheses

Up until now

What have we done

- Adjusted the initial connectivity and stimulus amplitude to match the naive responses
- Matched the baselines
- Identified connections that require update to match familiar non-change responses
- To obtain the constraints for the input to the VIP population upon image change, we matched the familiar change
- Novel change as validation that our choices were reasonable