Society for NeuroEconomics
NEUROSCIENCE • PSYCHOLOGY • ECONOMICS

Do within-trial interferences in OFC computations explain irrational choices?

A neuro-computational approach to value synthesis and comparison.

SORBONNE UNIVERSITÉ
Paris Brain Institute ICM
MBB
Motivation Brain Behavior team

Juliette Bénon
PhD candidate

Supervised by **Jean Daunizeau**

How do we take decisions?

How do we take decisions?

Why are we irrational?

How do we take decisions?

Why are we irrational?

Possible biological constraints :

How do we take decisions?

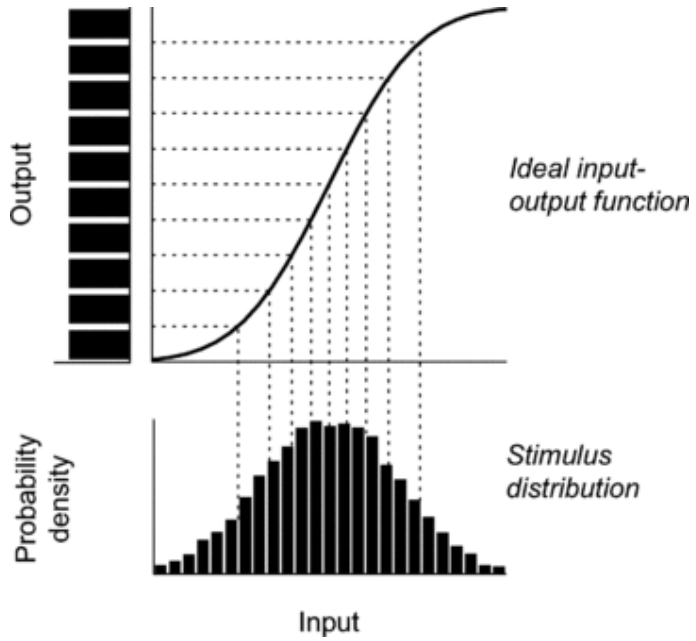
Why are we irrational?

Possible biological constraints : - Efficient coding

How do we take decisions?

Why are we irrational?

Possible biological constraints : - Efficient coding

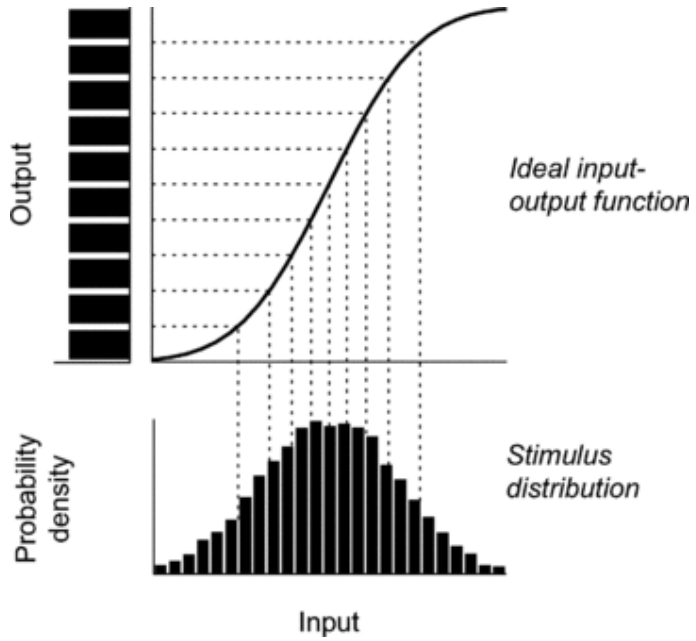


Louie & Glimcher, 2012. Figure 5.

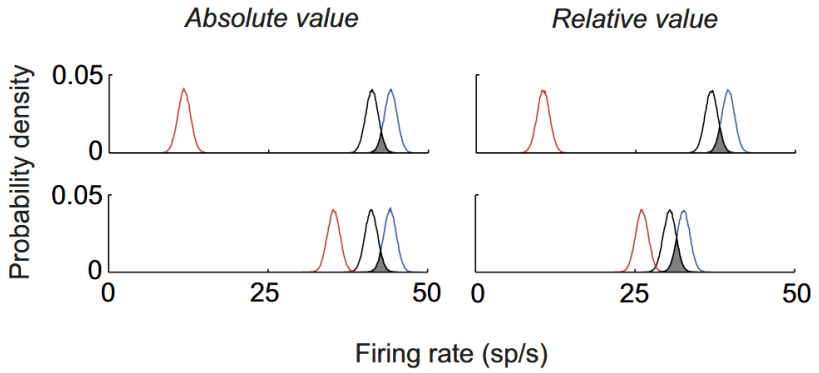
How do we take decisions?

Why are we irrational?

Possible biological constraints : - Efficient coding



Louie & Glimcher, 2012. Figure 5.



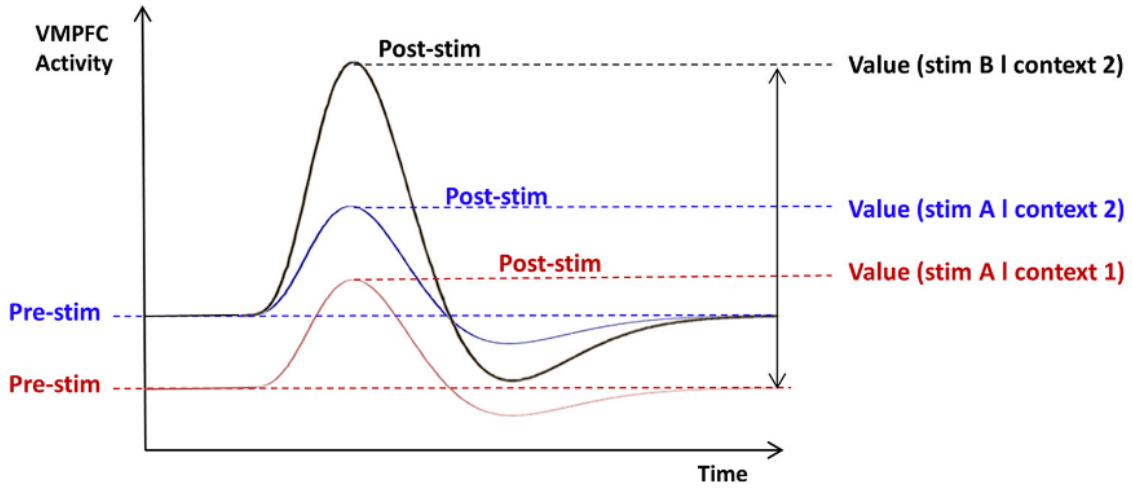
Louie et al., 2013. Adapted from Figure 1B.

How do we take decisions?

Why are we irrational?

Possible biological constraints : - Efficient coding

- Neural autocorrelation



Abitbol et al., 2015. Figure 1.

How do we take decisions?

Why are we irrational?

Possible biological constraints :

- Efficient coding
- Neural autocorrelation
- Limited energy budget

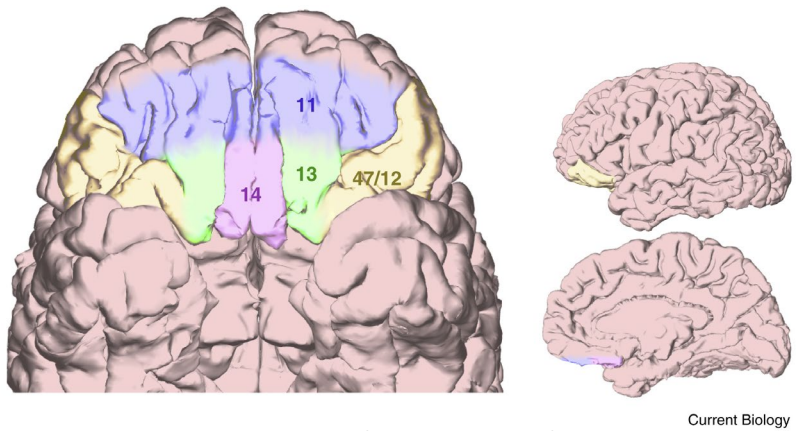
How do we take decisions?

Why are we irrational?

Possible biological constraints :

- Efficient coding
- Neural autocorrelation
- Limited energy budget

Orbitofrontal cortex (OFC)



Rudebeck & Rich, 2018. Figure 1.

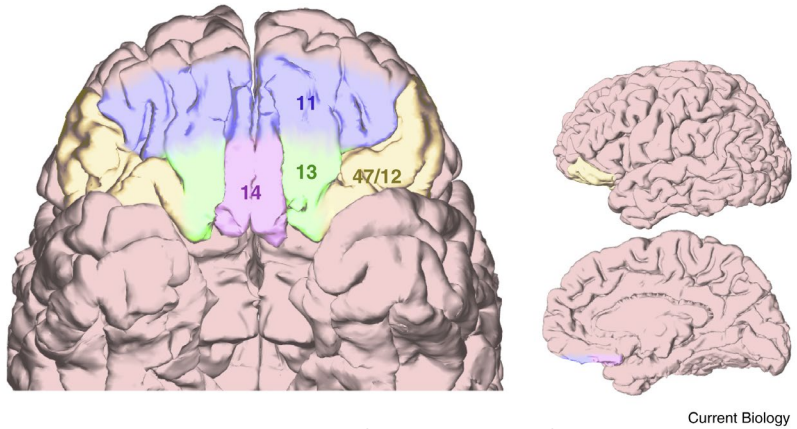
How do we take decisions?

Why are we irrational?

Possible biological constraints :

- Efficient coding
- Neural autocorrelation
- Limited energy budget

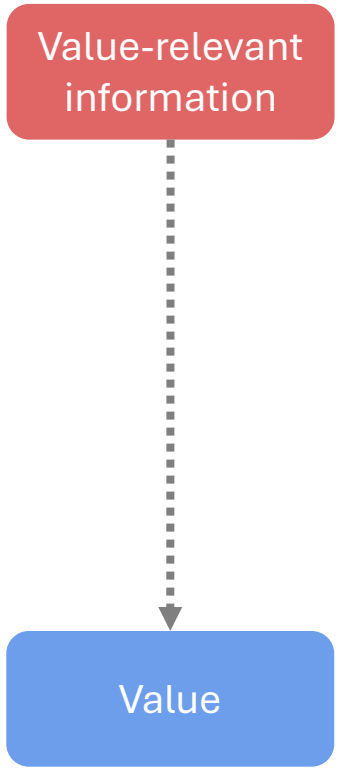
Orbitofrontal cortex (OFC) → Key region for value-related computations



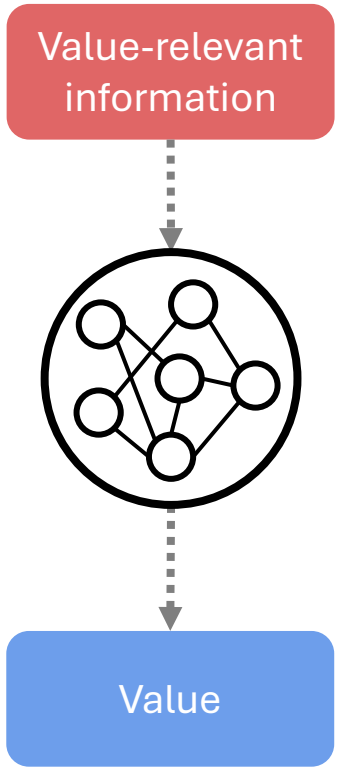
Rudebeck & Rich, 2018. Figure 1.

Padoa-Schioppa & Assad (2006)
Kable & Glimcher (2009)
Hun et al. (2012)
Suzuki et al. (2017)
Juechems & Summerfield (2019)
Pessiglione & Daunizeau (2021)
O'Doherty et al. (2021)

Modelling value computations



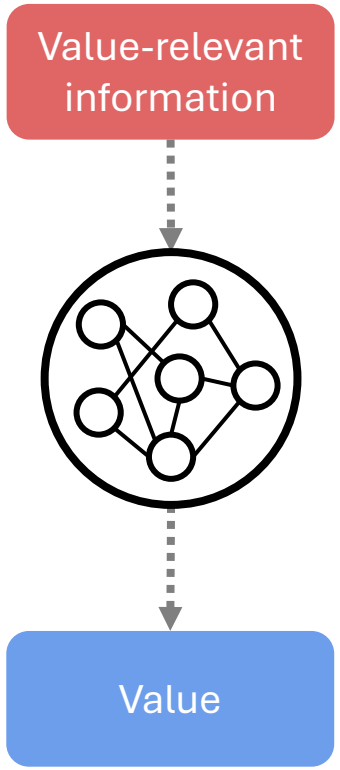
Modelling value computations



Use artificial neural networks:

- No a priori constraints on the value mapping function
- Generative models of neural activity

Modelling value computations



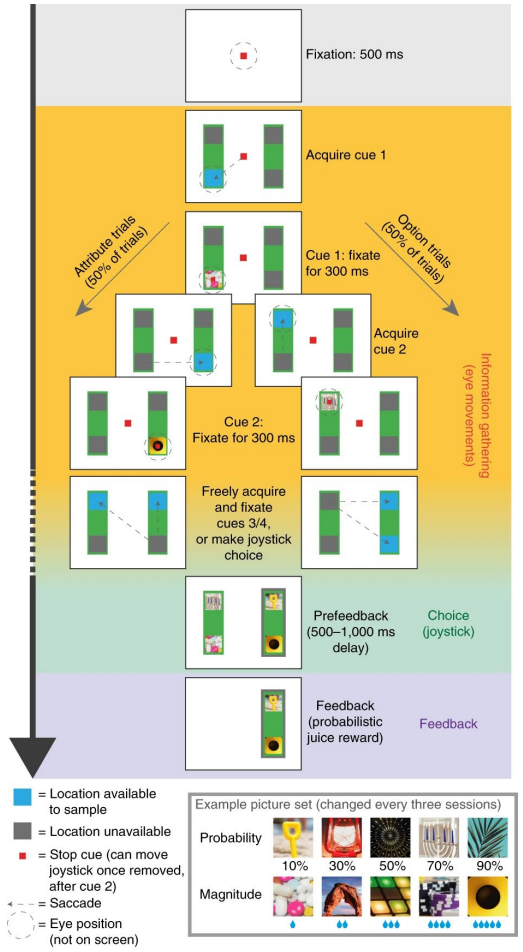
Use publicly available dataset:

nature neuroscience
ARTICLES
https://doi.org/10.1038/s41593-018-0239-5

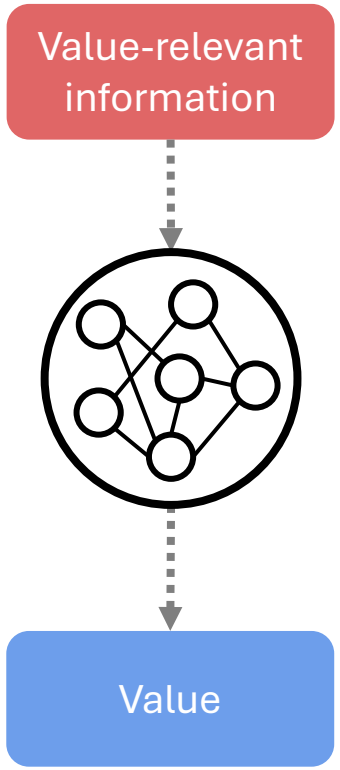
Triple dissociation of attention and decision computations across prefrontal cortex

Laurence T. Hunt[1,2,3,4,7*], W. M. Nishantha Malalasekera[1,7], Archy O. de Berker[1,2], Bruno Miranda[1,5,6], Simon F. Farmer[1], Timothy E. J. Behrens[2,3] and Steven W. Kennerley[1*]

Hunt et al., 2018. Figure 1a



Modelling value computations

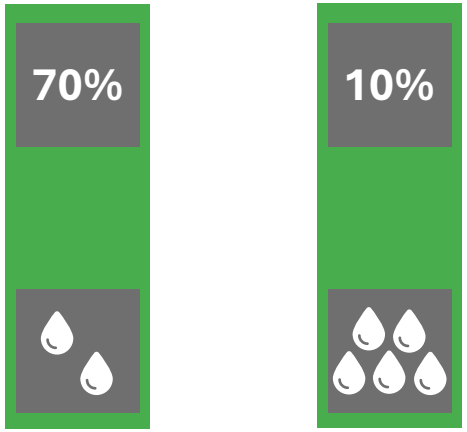


Use publicly available dataset:

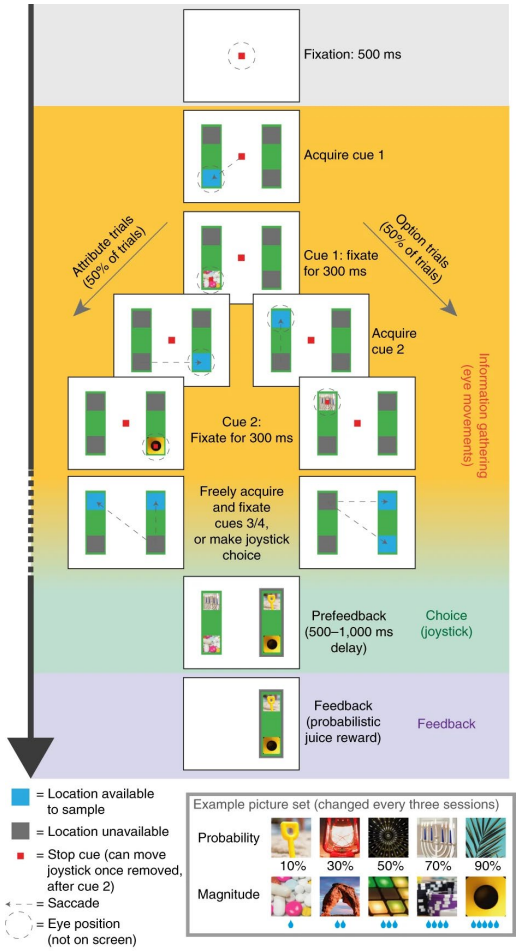
nature neuroscience
ARTICLES
https://doi.org/10.1038/s41593-018-0239-5

Triple dissociation of attention and decision computations across prefrontal cortex

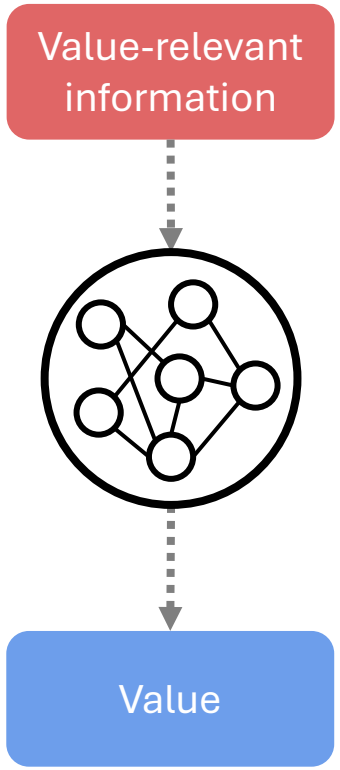
Laurence T. Hunt[1,2,3,4,7*], W. M. Nishantha Malalasekera[1,7], Archy O. de Berker[1,2], Bruno Miranda[1,5,6], Simon F. Farmer[1], Timothy E. J. Behrens[2,3] and Steven W. Kennerley[1*]



Hunt et al., 2018. Figure 1a



Modelling value computations



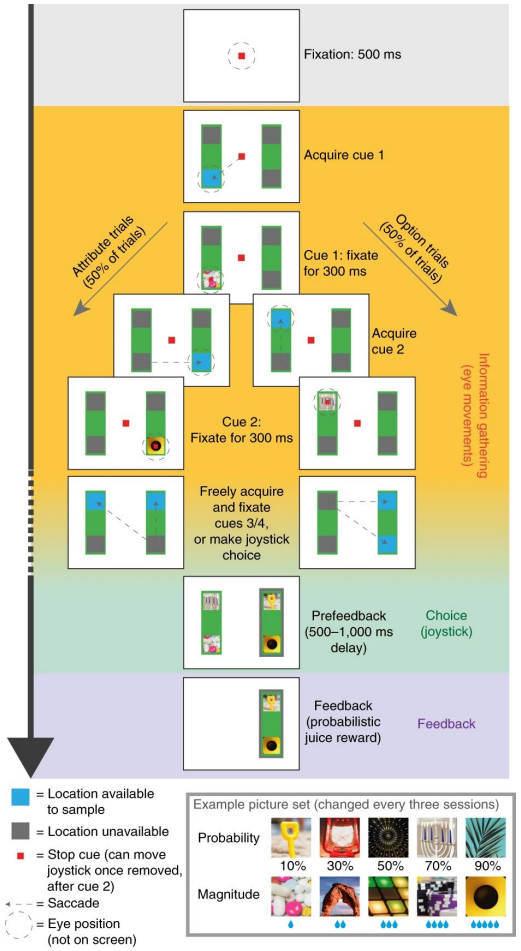
Use publicly available dataset:

nature neuroscience
ARTICLES
https://doi.org/10.1038/s41593-018-0239-5

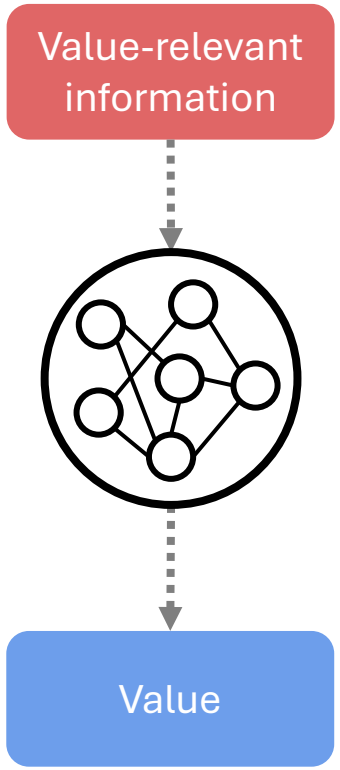
Triple dissociation of attention and decision computations across prefrontal cortex

Laurence T. Hunt[1,2,3,4,7*], W. M. Nishantha Malalasekera[1,7], Archy O. de Berker[1,2], Bruno Miranda[1,5,6], Simon F. Farmer[1], Timothy E. J. Behrens[2,3] and Steven W. Kennerley[1*]

Hunt et al., 2018. Figure 1a



Modelling value computations

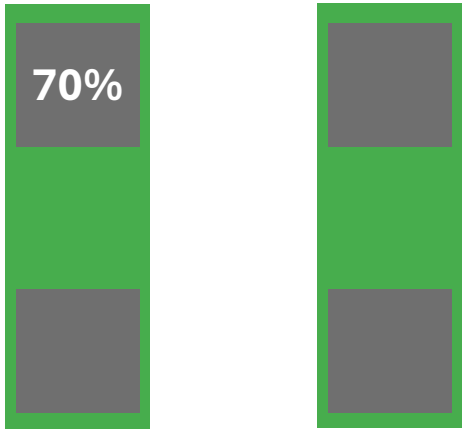


Use publicly available dataset:

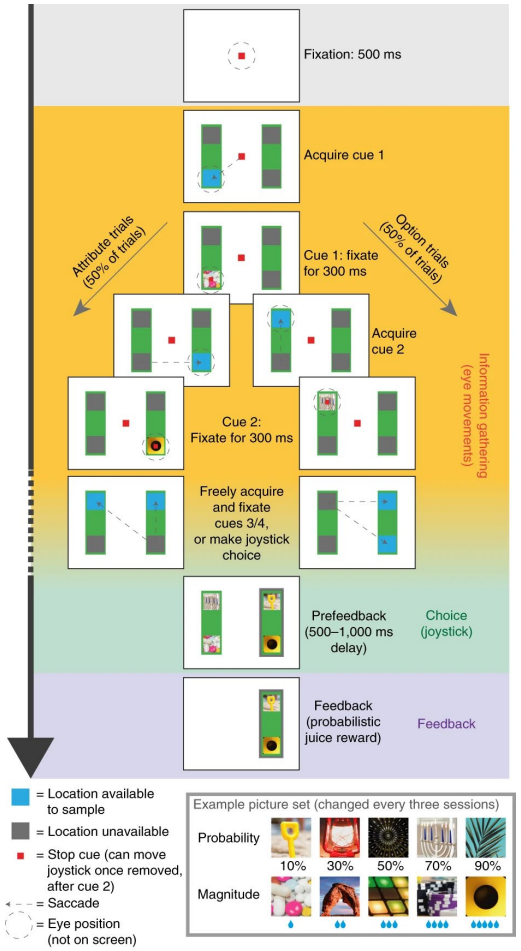
nature neuroscience
ARTICLES
https://doi.org/10.1038/s41593-018-0239-5

Triple dissociation of attention and decision computations across prefrontal cortex

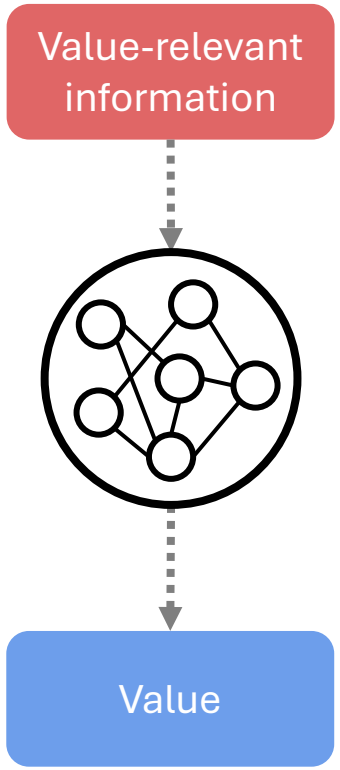
Laurence T. Hunt[1,2,3,4,7*], W. M. Nishantha Malalasekera[1,7], Archy O. de Berker[1,2], Bruno Miranda[1,5,6], Simon F. Farmer[1], Timothy E. J. Behrens[2,3] and Steven W. Kennerley[1*]



Hunt et al., 2018. Figure 1a



Modelling value computations

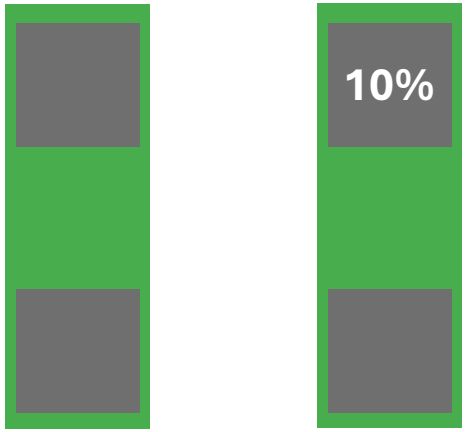


Use publicly available dataset:

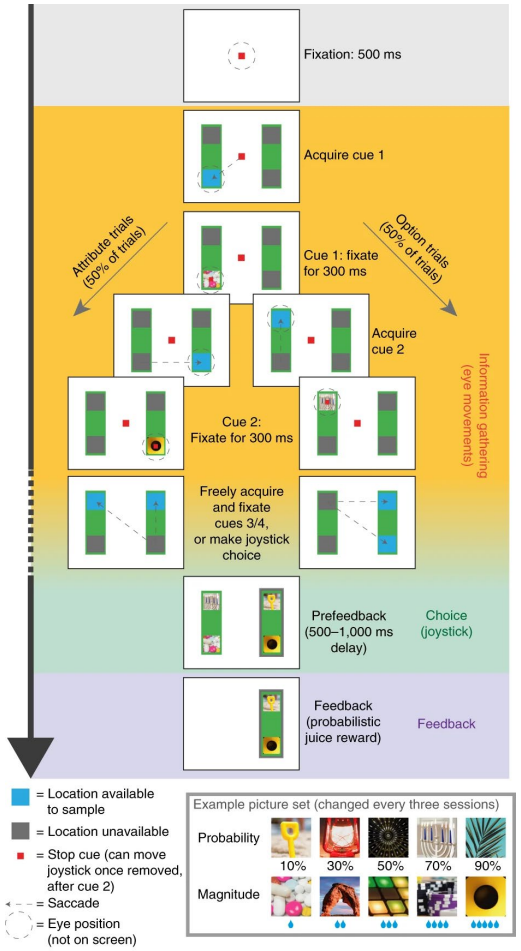
nature neuroscience
ARTICLES
https://doi.org/10.1038/s41593-018-0239-5

Triple dissociation of attention and decision computations across prefrontal cortex

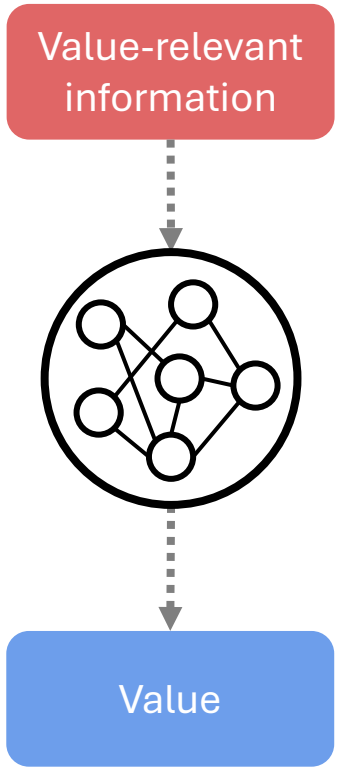
Laurence T. Hunt[1,2,3,4,7*], W. M. Nishantha Malalasekera[1,7], Archy O. de Berker[1,2], Bruno Miranda[1,5,6], Simon F. Farmer[1], Timothy E. J. Behrens[2,3] and Steven W. Kennerley[1*]



Hunt et al., 2018. Figure 1a



Modelling value computations

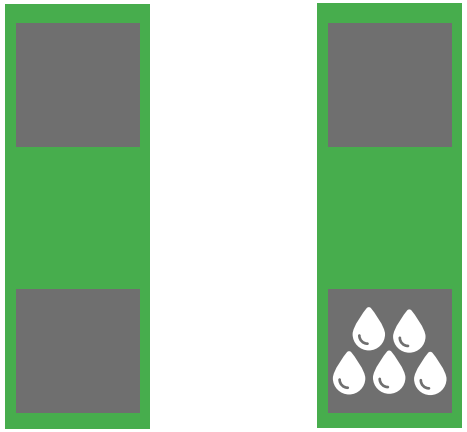


Use publicly available dataset:

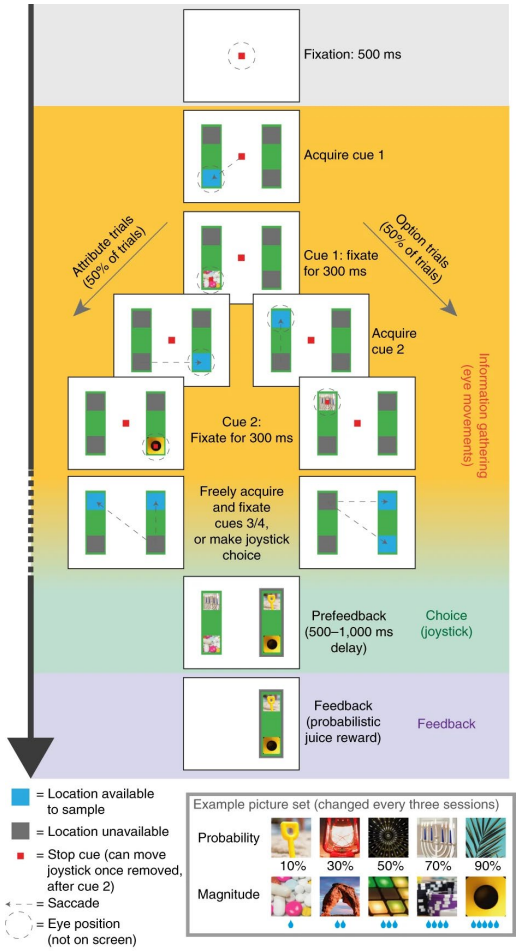
nature neuroscience
ARTICLES
https://doi.org/10.1038/s41593-018-0239-5

Triple dissociation of attention and decision computations across prefrontal cortex

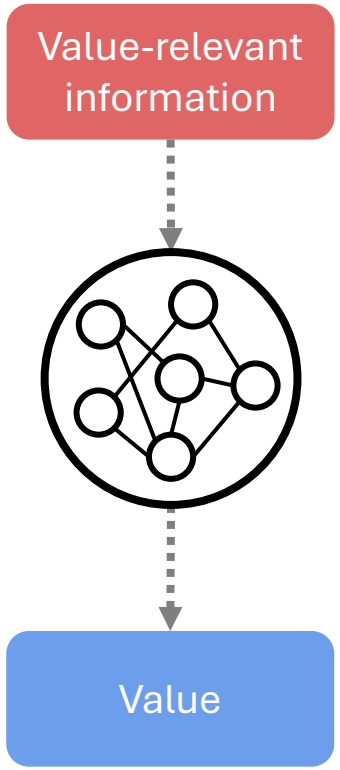
Laurence T. Hunt[1,2,3,4,7*], W. M. Nishantha Malalasekera[1,7], Archy O. de Berker[1,2], Bruno Miranda[1,5,6], Simon F. Farmer[1], Timothy E. J. Behrens[2,3] and Steven W. Kennerley[1*]



Hunt et al., 2018. Figure 1a



Modelling value computations

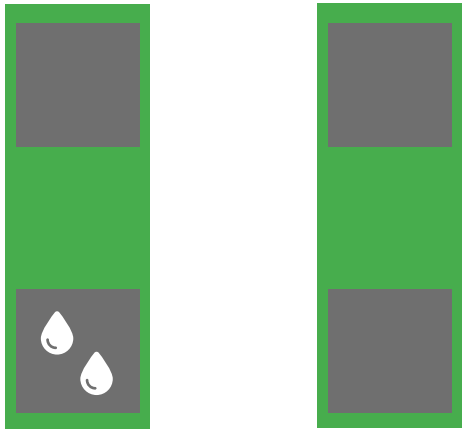


Use publicly available dataset:

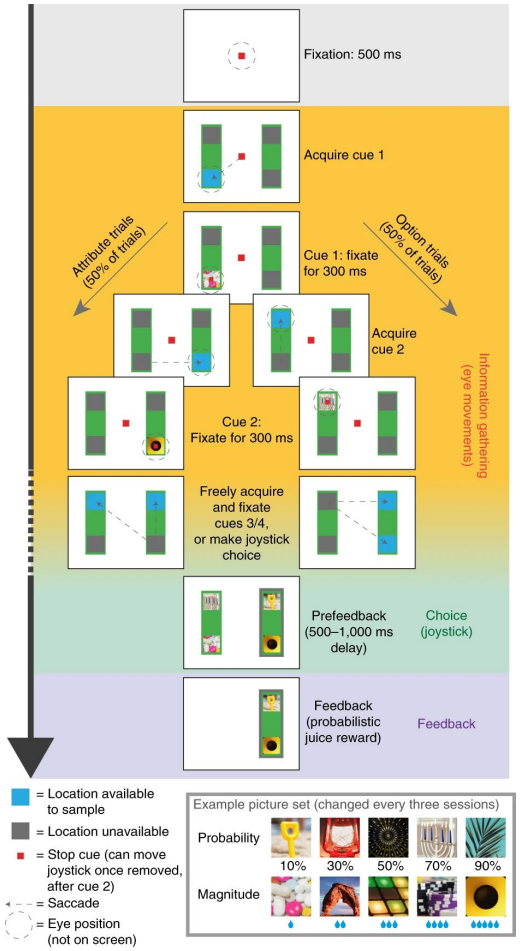
nature neuroscience
ARTICLES
https://doi.org/10.1038/s41593-018-0239-5

Triple dissociation of attention and decision computations across prefrontal cortex

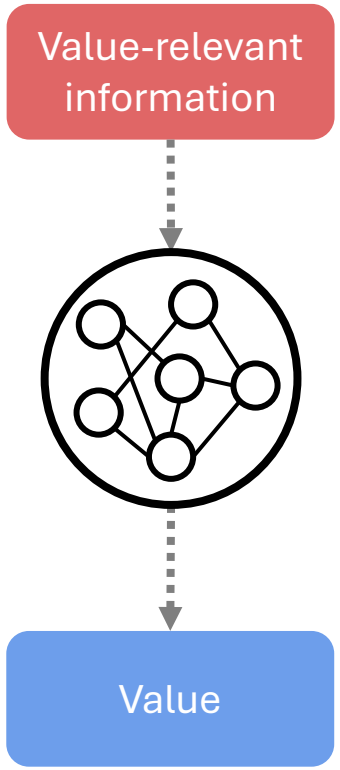
Laurence T. Hunt[1,2,3,4,7*], W. M. Nishantha Malalasekera[1,7], Archy O. de Berker[1,2], Bruno Miranda[1,5,6], Simon F. Farmer[1], Timothy E. J. Behrens[2,3] and Steven W. Kennerley[1*]



Hunt et al., 2018. Figure 1a



Modelling value computations



Use publicly available dataset:

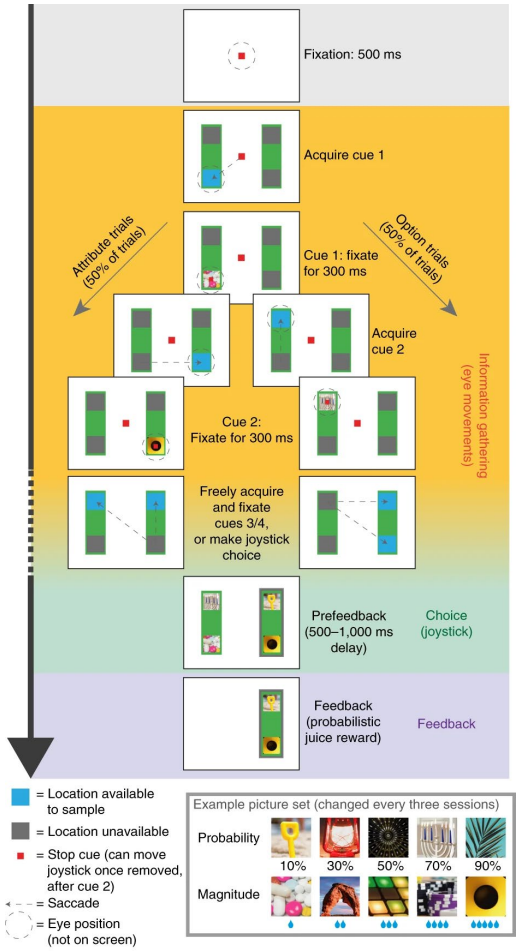
nature neuroscience
ARTICLES
https://doi.org/10.1038/s41593-018-0239-5

Triple dissociation of attention and decision computations across prefrontal cortex

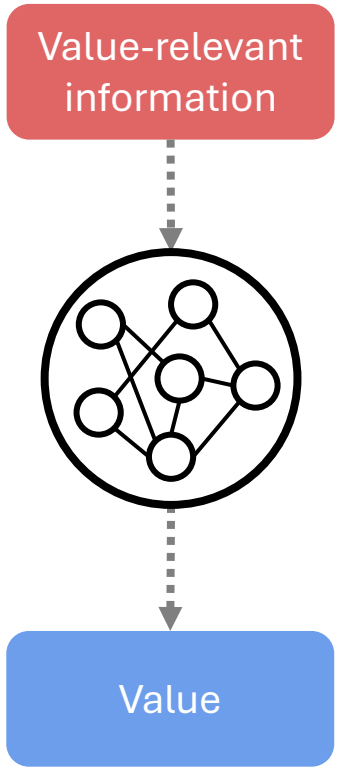
Laurence T. Hunt[1,2,3,4,7*], W. M. Nishantha Malalasekera[1,7], Archy O. de Berker[1,2], Bruno Miranda[1,5,6], Simon F. Farmer[1], Timothy E. J. Behrens[2,3] and Steven W. Kennerley[1*]

Models do the exact same task as the monkeys!

Hunt et al., 2018. Figure 1a



Modelling value computations



Use publicly available dataset:

nature neuroscience ARTICLES
https://doi.org/10.1038/s41593-018-0239-5

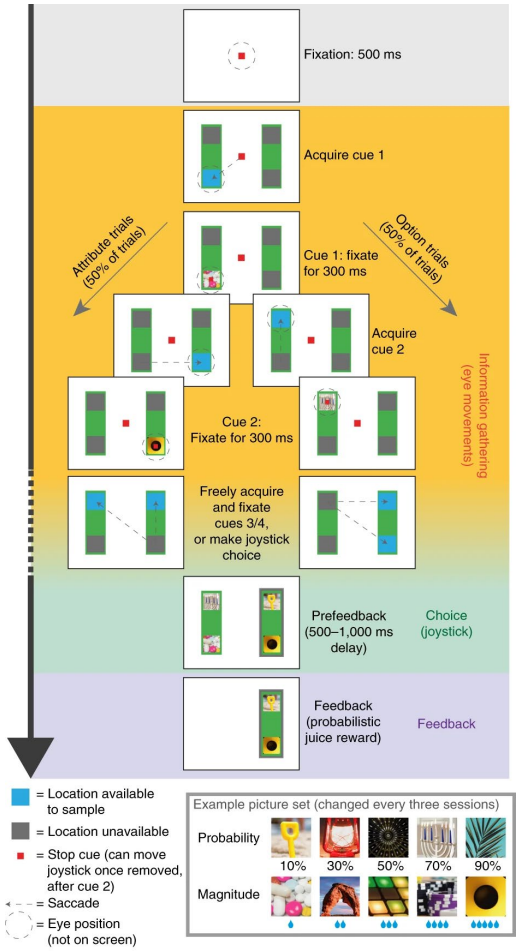
Triple dissociation of attention and decision computations across prefrontal cortex

Laurence T. Hunt[1,2,3,4,7*], W. M. Nishantha Malalasekera[1,7], Archy O. de Berker[1,2], Bruno Miranda[1,5,6], Simon F. Farmer[1], Timothy E. J. Behrens[2,3] and Steven W. Kennerley[1*]

Models do the exact same task as the monkeys!

- Two options

Hunt et al., 2018. Figure 1a

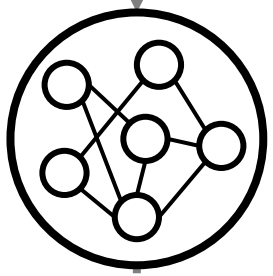


Modelling value computations

McGinty & Lupkin 2023
Rich & Wallis 2016

Value synthesis

Value-relevant information

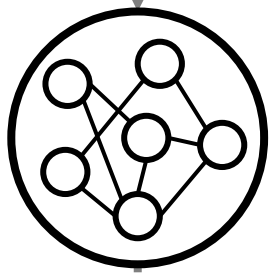


Values of both options

Hunt et al. 2018
Strait et al. 2014

Value comparison

Value-relevant information



Value difference

Use publicly available dataset:

nature neuroscience ARTICLES
https://doi.org/10.1038/s41593-018-0239-5

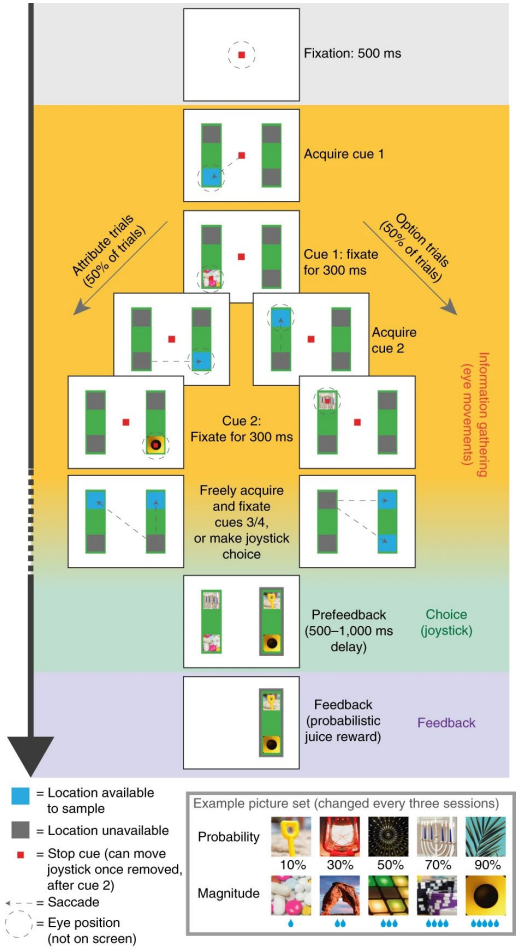
Triple dissociation of attention and decision computations across prefrontal cortex

Laurence T. Hunt[1,2,3,4,7]*, W. M. Nishantha Malalasekera[1,7], Archy O. de Berker[1,2], Bruno Miranda[1,5,6], Simon F. Farmer[1], Timothy E. J. Behrens[2,3] and Steven W. Kennerley[1]*

Models do the exact same task as the monkeys!

- Two options

Hunt et al., 2018. Figure 1a

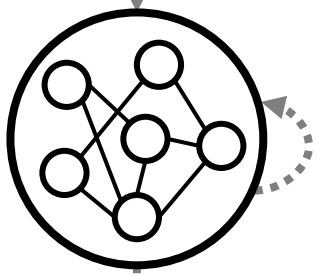


Modelling value computations

McGinty & Lupkin 2023
Rich & Wallis 2016

Value synthesis

Value-relevant information

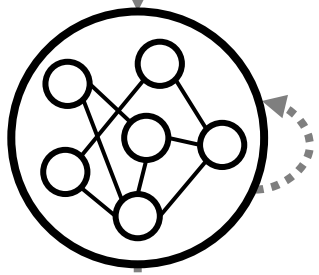


Values of both options

Hunt et al. 2018
Strait et al. 2014

Value comparison

Value-relevant information



Value difference

Use publicly available dataset:

nature neuroscience ARTICLES
https://doi.org/10.1038/s41593-018-0239-5

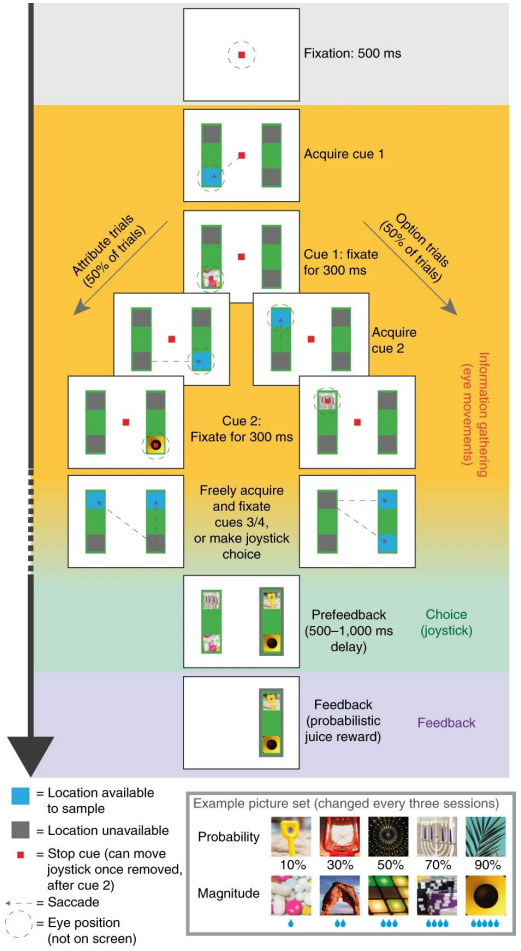
Triple dissociation of attention and decision computations across prefrontal cortex

Laurence T. Hunt[1,2,3,4,7*], W. M. Nishantha Malalasekera[1,7], Archy O. de Berker[1,2], Bruno Miranda[1,5,6], Simon F. Farmer[1], Timothy E. J. Behrens[2,3] and Steven W. Kennerley[1*]

Models do the exact same task as the monkeys!

- Two options
- Sequential information acquisition

Hunt et al., 2018. Figure 1a

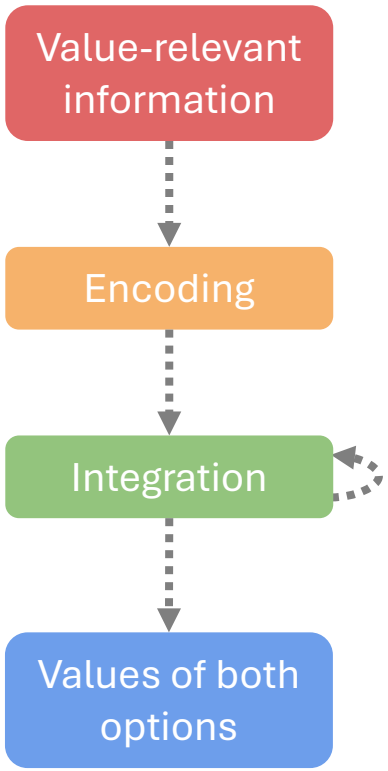


Modelling value computations

Neural network architecture

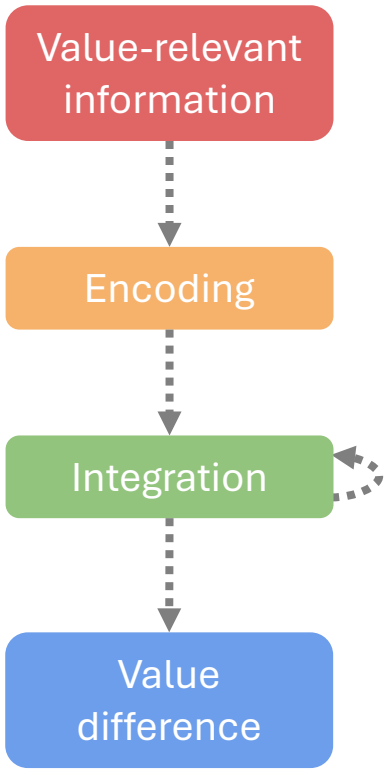
McGinty & Lupkin 2023
Rich & Wallis 2016

Value synthesis



Hunt et al. 2018
Strait et al. 2014

Value comparison



Pessiglione & Daunizeau 2021

Modelling value computations

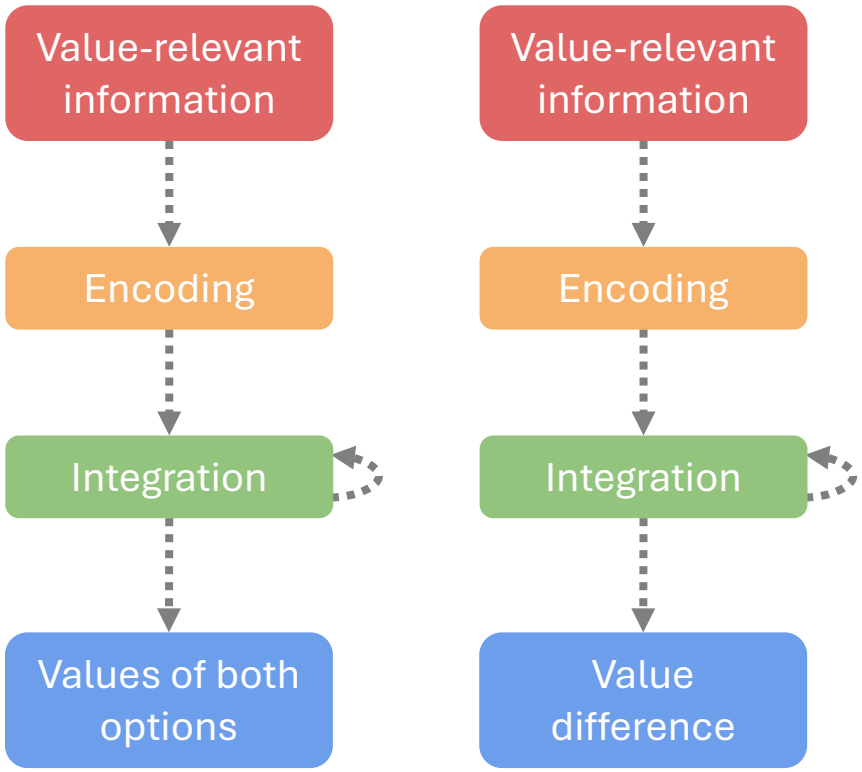
Neural network architecture

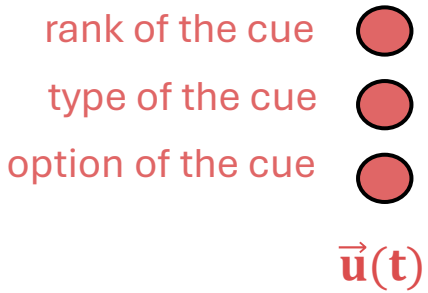
McGinty & Lupkin 2023
Rich & Wallis 2016

Value synthesis

Hunt et al. 2018
Strait et al. 2014

Value comparison





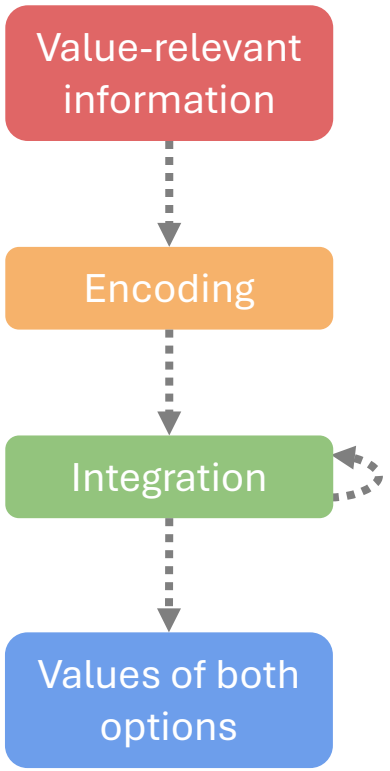
Pessiglione & Daunizeau 2021

Modelling value computations

Neural network architecture

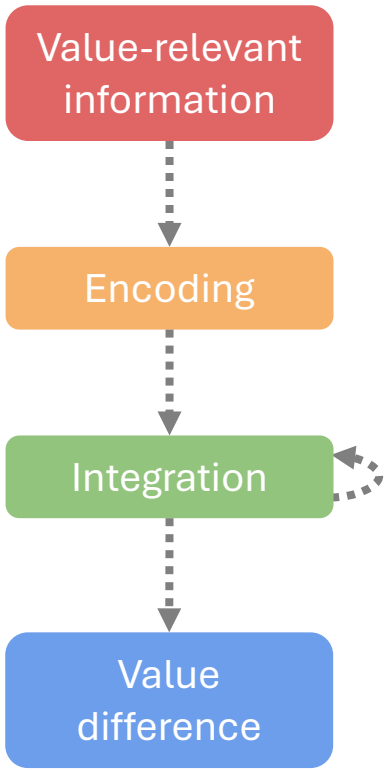
McGinty & Lupkin 2023
Rich & Wallis 2016

Value synthesis



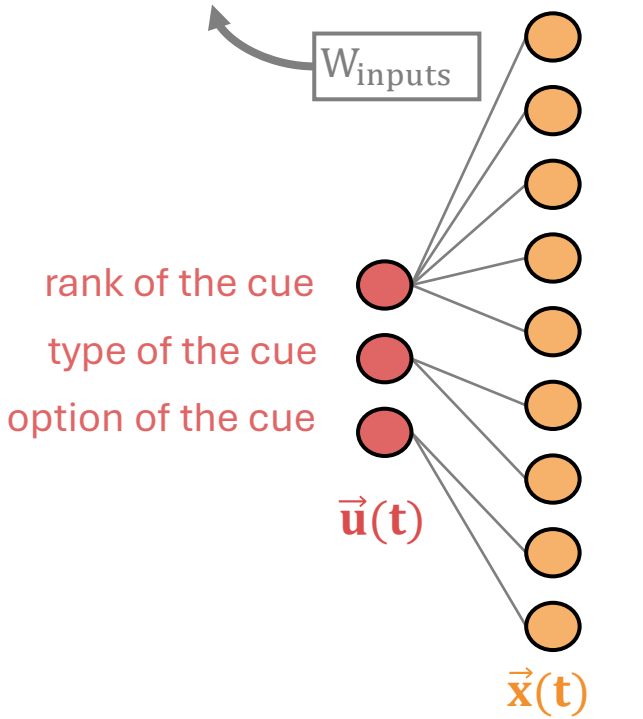
Hunt et al. 2018
Strait et al. 2014

Value comparison



Pessiglione & Daunizeau 2021

Distribute inputs for population encoding



$$\vec{x}(t) = \mathrm{sig}\left(W_{inputs} \cdot \vec{u}(t) - \overrightarrow{b_x}\right)$$

Modelling value computations

Neural network architecture

McGinty & Lupkin 2023
Rich & Wallis 2016

Value synthesis

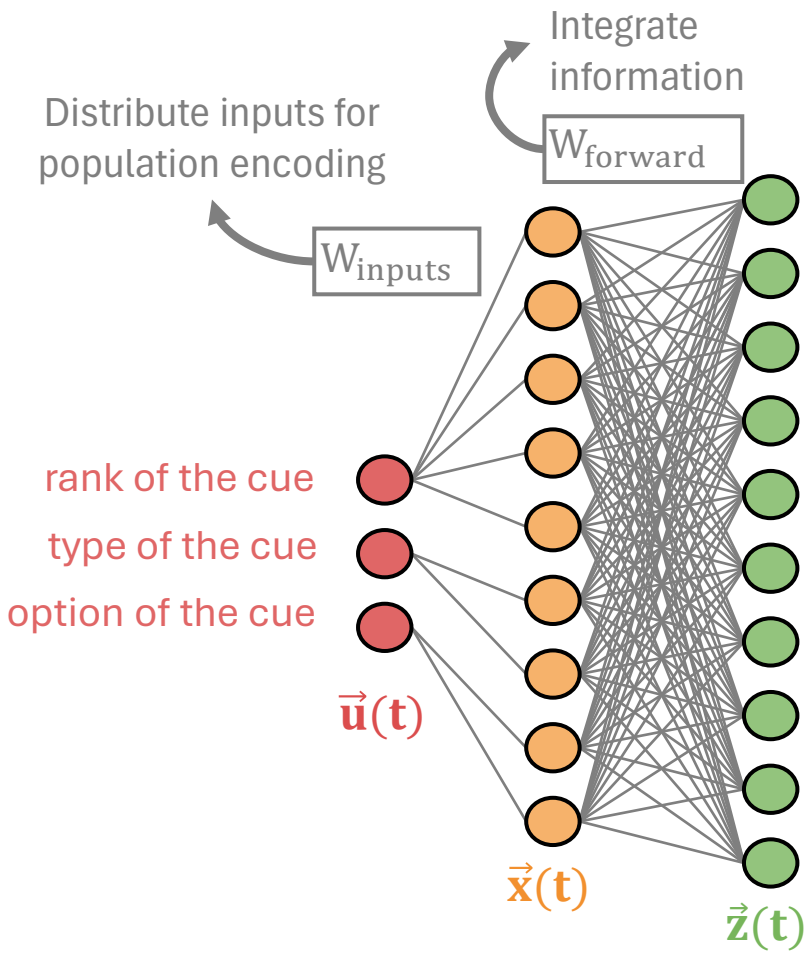
Value-relevant information
↓
Encoding
↓
Integration ↻
↓
Values of both options

Hunt et al. 2018
Strait et al. 2014

Value comparison

Value-relevant information
↓
Encoding
↓
Integration ↻
↓
Value difference

Pessiglione & Daunizeau 2021



$$\vec{x}(t) = \mathrm{sig}\left(W_{inputs} \cdot \vec{u}(t) - \overrightarrow{b_x}\right)$$

Modelling value computations

Neural network architecture

McGinty & Lupkin 2023
Rich & Wallis 2016

Value synthesis

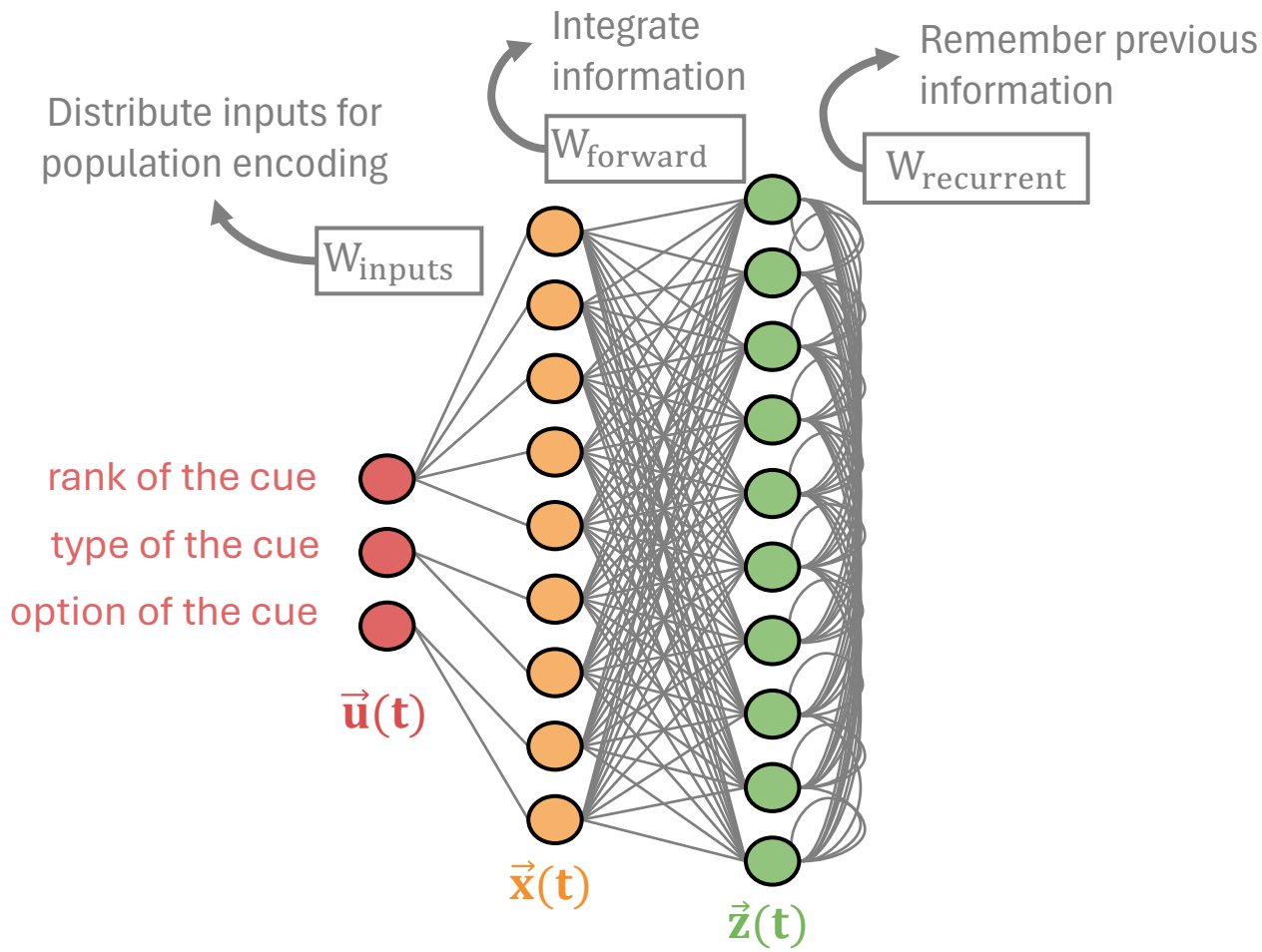
Value-relevant information → Encoding → Integration → Values of both options

Hunt et al. 2018
Strait et al. 2014

Value comparison

Value-relevant information → Encoding → Integration → Value difference

Pessiglione & Daunizeau 2021



$$\vec{x}(t) = \text{sig}\left(W_{inputs} \cdot \vec{u}(t) - \overrightarrow{b_x}\right)$$

$$\vec{z}(t) = \text{sig}\left(W_{forward} \cdot \vec{x}(t) - \overrightarrow{b_z} + W_{recurrent} \cdot \vec{z}(t-1)\right)$$

Modelling value computations

Neural network architecture

McGinty & Lupkin 2023
Rich & Wallis 2016

Value synthesis

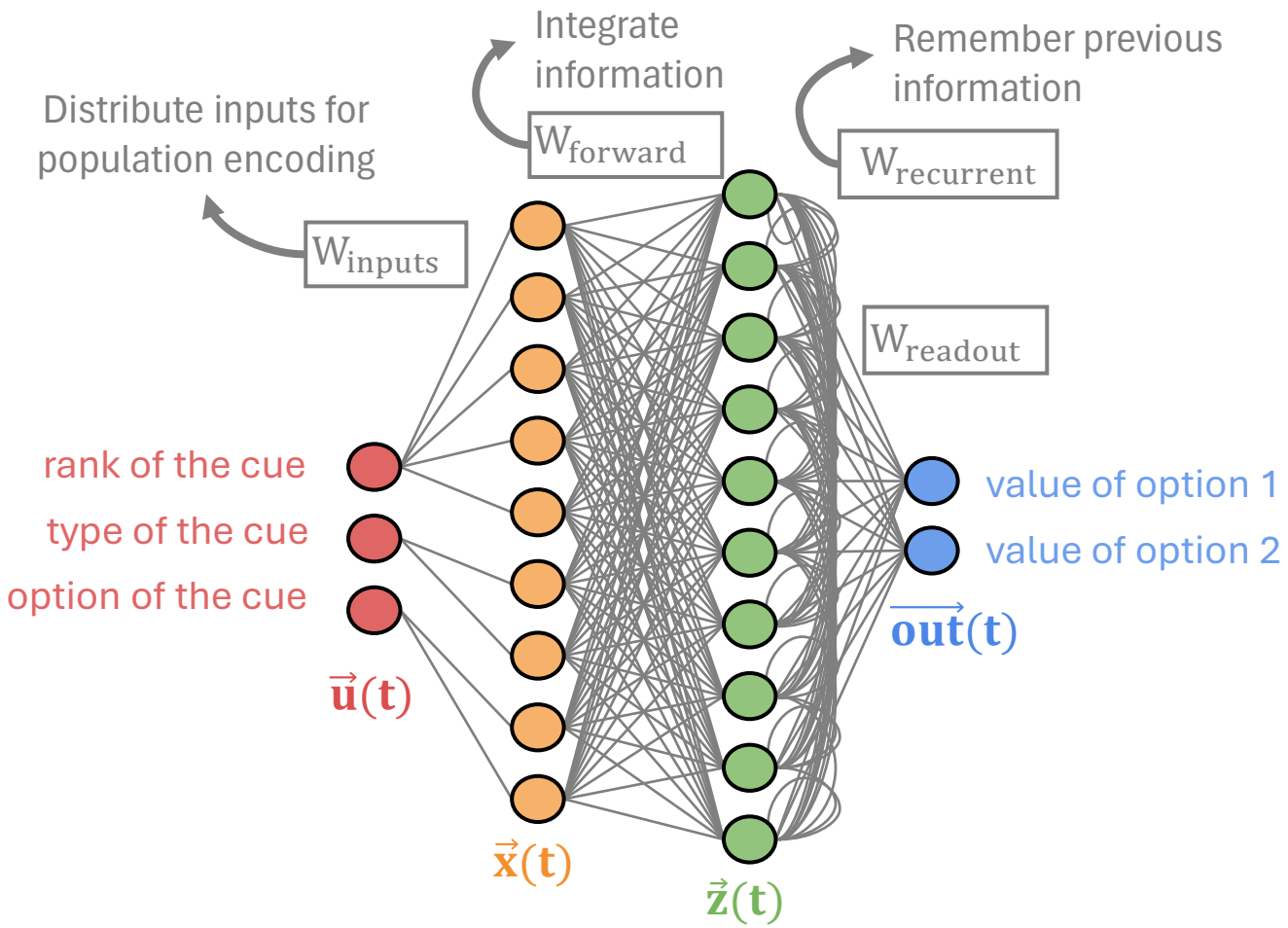
Value-relevant information → Encoding → Integration → Values of both options

Hunt et al. 2018
Strait et al. 2014

Value comparison

Value-relevant information → Encoding → Integration → Value difference

Pessiglione & Daunizeau 2021



$$\vec{x}(t) = \text{sig}\left(W_{inputs} \cdot \vec{u}(t) - \overrightarrow{b_x}\right)$$

$$\vec{z}(t) = \text{sig}\left(W_{forward} \cdot \vec{x}(t) - \overrightarrow{b_z} + W_{recurrent} \cdot \vec{z}(t-1)\right)$$

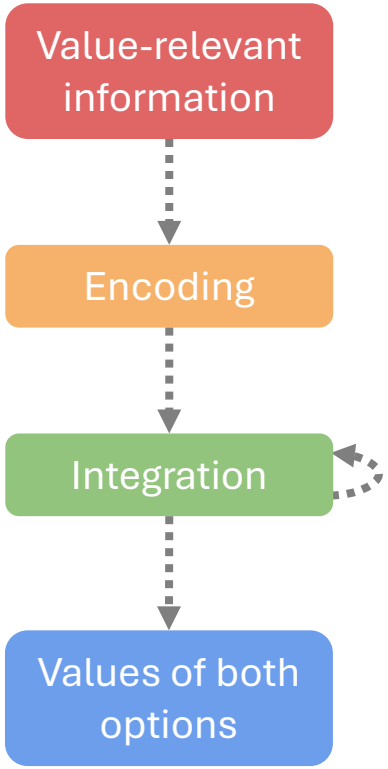
$$\overrightarrow{out}(t) = W_{readout} \cdot \vec{z}(t)$$

Modelling value computations

Information encoding frameworks

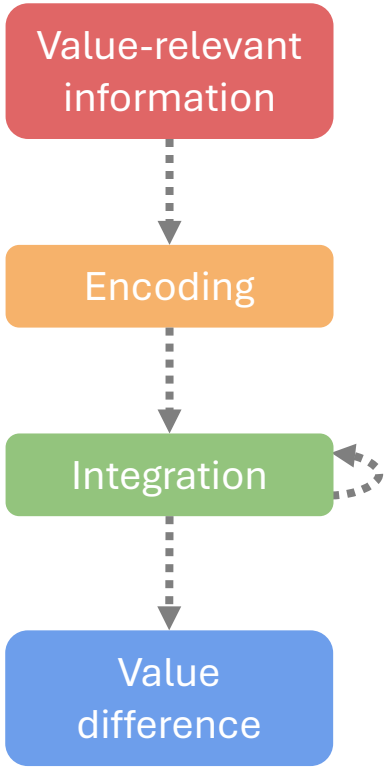
McGinty & Lupkin 2023
Rich & Wallis 2016

Value synthesis



Hunt et al. 2018
Strait et al. 2014

Value comparison



Pessiglione & Daunizeau 2021

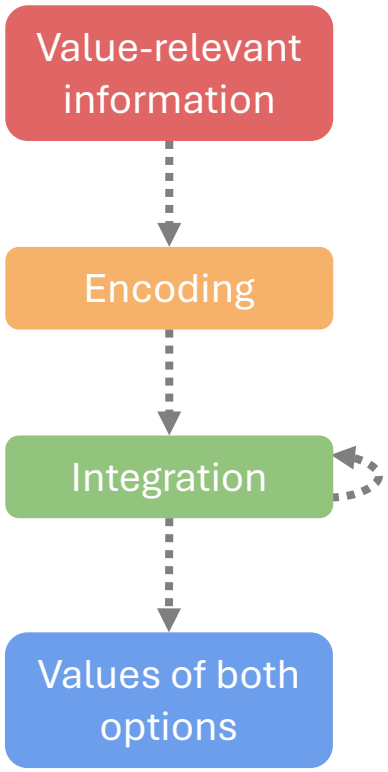
How to encode information?

Modelling value computations

Information encoding frameworks

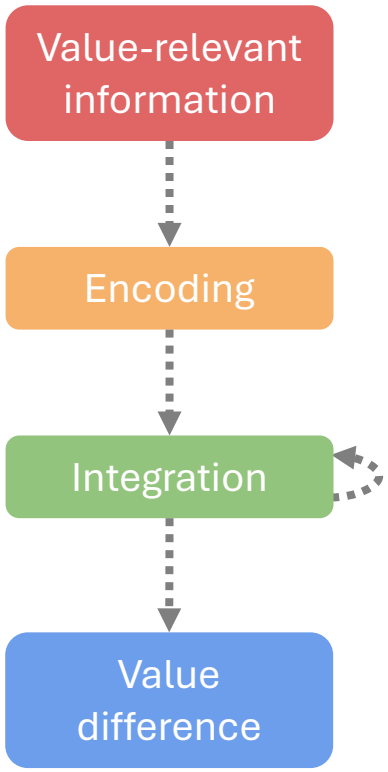
McGinty & Lupkin 2023
Rich & Wallis 2016

Value synthesis



Hunt et al. 2018
Strait et al. 2014

Value comparison



Pessiglione & Daunizeau 2021

How to encode information?

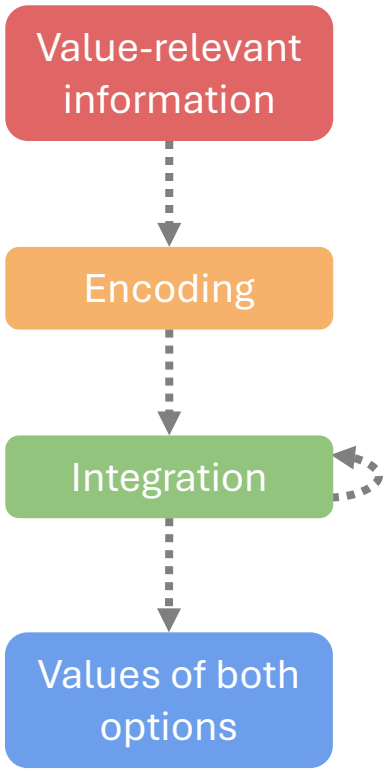
Spatial framework: left vs. right

Modelling value computations

Information encoding frameworks

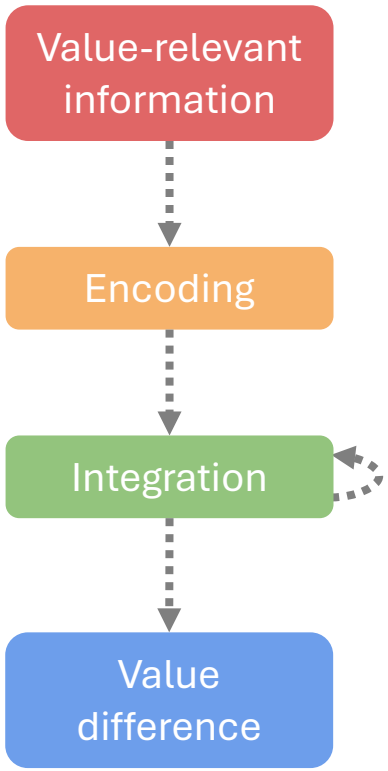
McGinty & Lupkin 2023
Rich & Wallis 2016

Value synthesis



Hunt et al. 2018
Strait et al. 2014

Value comparison



Pessiglione & Daunizeau 2021

How to encode information?

Spatial framework: left vs. right

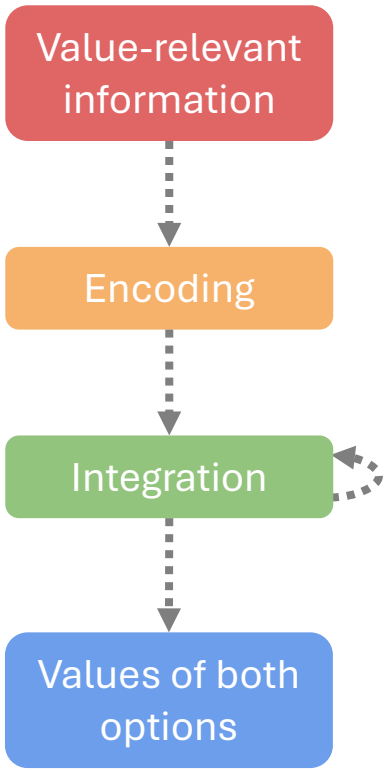
Temporal order framework: first vs. second

Modelling value computations

Information encoding frameworks

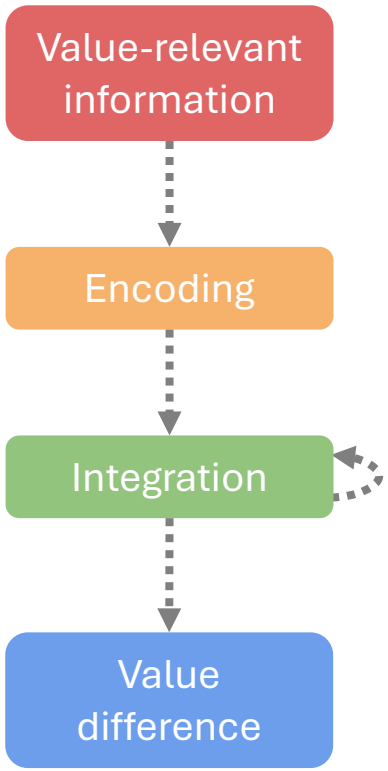
McGinty & Lupkin 2023
Rich & Wallis 2016

Value synthesis



Hunt et al. 2018
Strait et al. 2014

Value comparison



Pessiglione & Daunizeau 2021

How to encode information?

Spatial framework: left vs. right

Temporal order framework: first vs. second

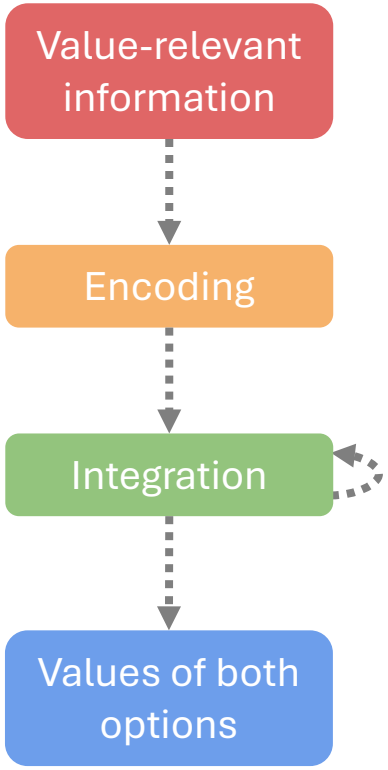
Attentional focus framework: attended vs. unattended

Modelling value computations

Information encoding frameworks

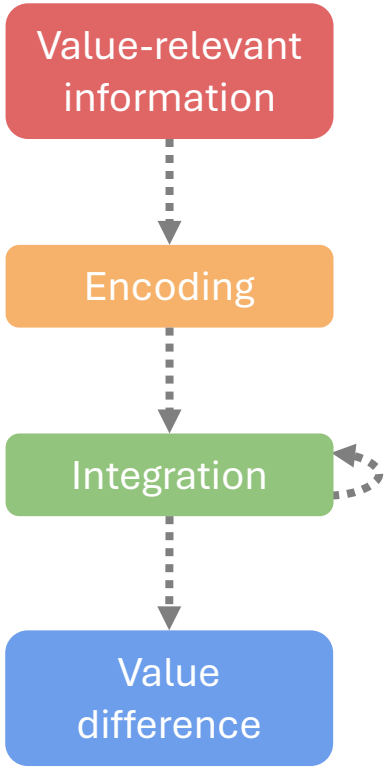
McGinty & Lupkin 2023
Rich & Wallis 2016

Value synthesis



Hunt et al. 2018
Strait et al. 2014

Value comparison



Pessiglione & Daunizeau 2021

How to encode information?

Spatial framework: left vs. right

Temporal order framework: first vs. second

Attentional focus framework: attended vs. unattended

Input framework

Output framework	Spatial	Temporal order
Spatial	✓	
Temporal order	✓	✓
Attentional focus	✓	✓

What do OFC neurons do?

What do OFC neurons do?

Training rational models:

What do OFC neurons do?

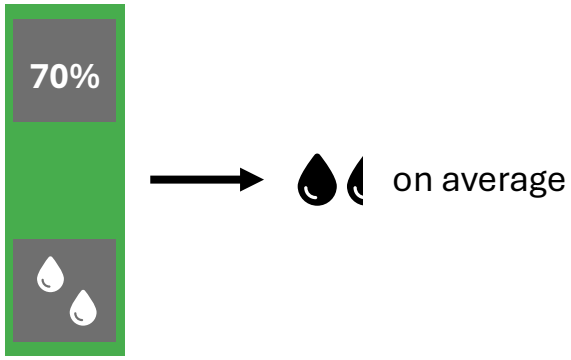
Training rational models:

Value = Expected reward

What do OFC neurons do?

Training rational models:

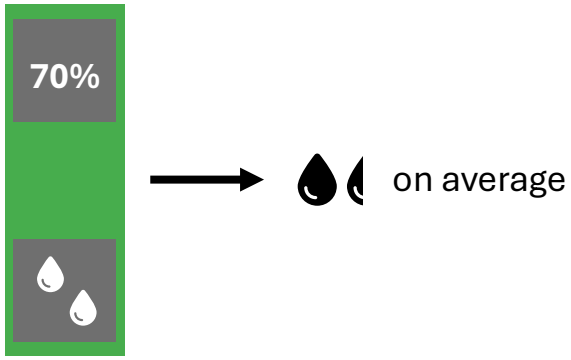
Value = Expected reward



What do OFC neurons do?

Training rational models:

Value = Expected reward

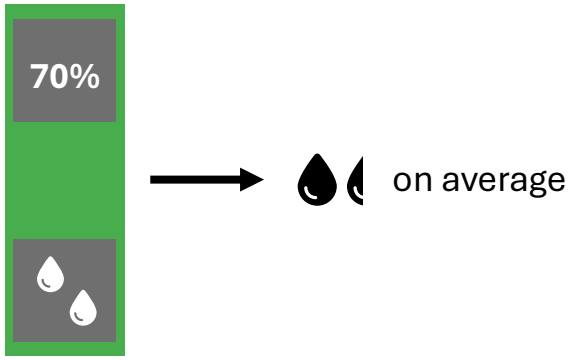


- Start with random weights
- Train on a dataset of random trials

What do OFC neurons do?

Training rational models:

Value = Expected reward



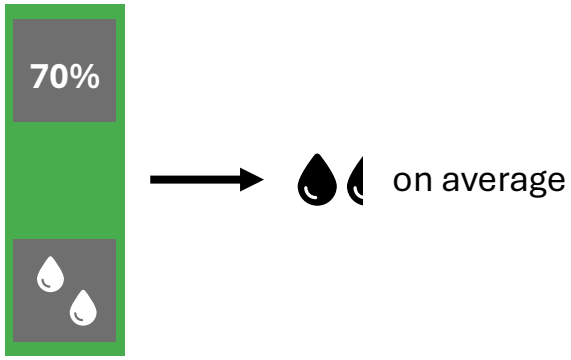
- Start with random weights
- Train on a dataset of random trials

Repeat 1000 times

What do OFC neurons do?

Training rational models:

Value = Expected reward



- Start with random weights
- Train on a dataset of random trials

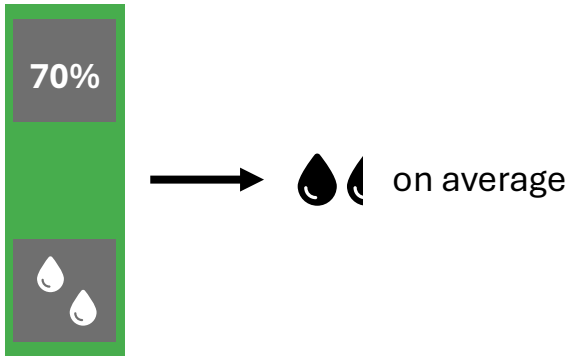
Repeat 1000 times

Comparison with OFC neural activity:

What do OFC neurons do?

Training rational models:

Value = Expected reward

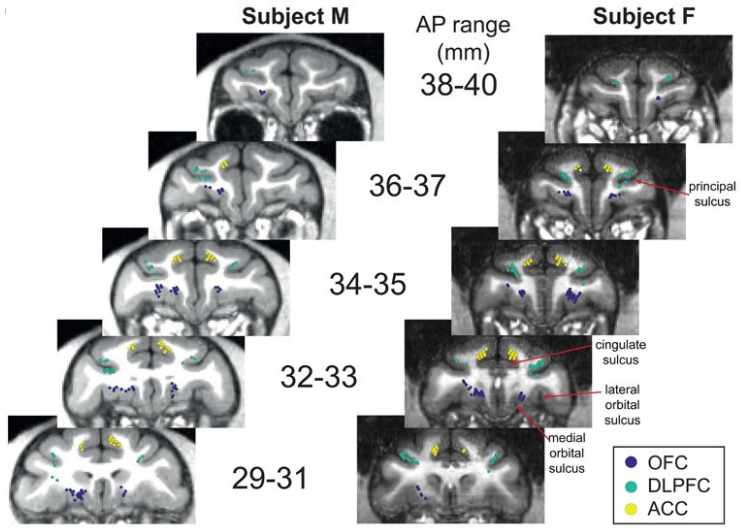


- Start with random weights
- Train on a dataset of random trials

Repeat 1000 times

Comparison with OFC neural activity:

Hunt et al., 2018. Figure 3b.

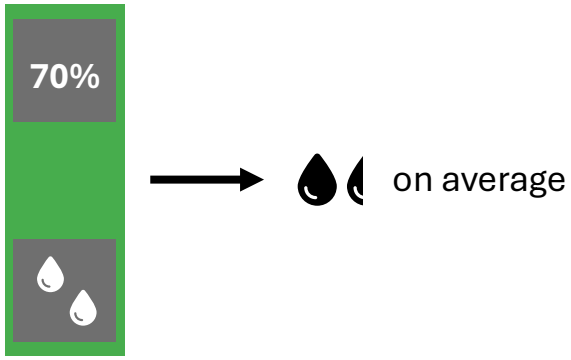


Electrophysiological recordings in the OFC of two macaque monkeys (183 neurons, 22 000 trials in total).

What do OFC neurons do?

Training rational models:

Value = Expected reward

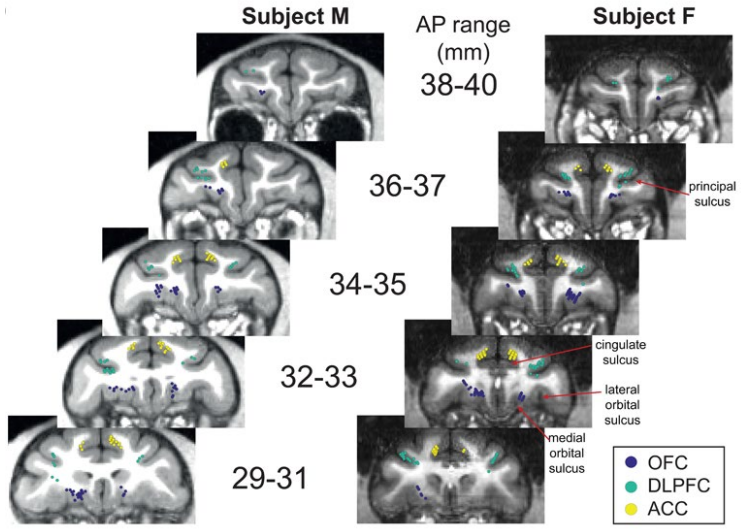


- Start with random weights
- Train on a dataset of random trials

Repeat 1000 times

Comparison with OFC neural activity:

Hunt et al., 2018. Figure 3b.



Electrophysiological recordings in the OFC of two macaque monkeys (183 neurons, 22 000 trials in total).

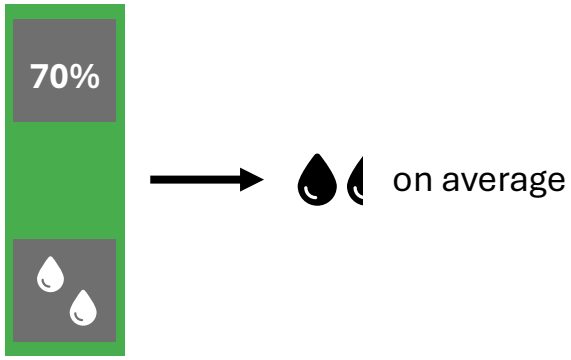
Representational geometry:

How the **sensitivity profile** of the neuronal population **varies** across different time windows.

What do OFC neurons do?

Training rational models:

Value = Expected reward

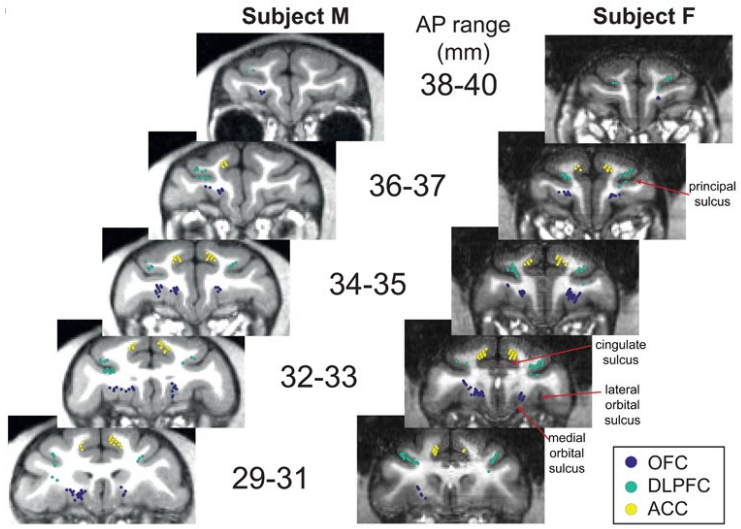


- Start with random weights
- Train on a dataset of random trials

Repeat 1000 times

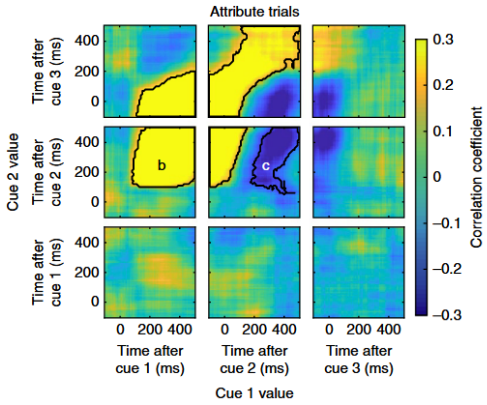
Comparison with OFC neural activity:

Hunt et al., 2018. Figure 3b.



Electrophysiological recordings in the OFC of two macaque monkeys (183 neurons, 22 000 trials in total).

Hunt et al., 2018. Figure 5g.



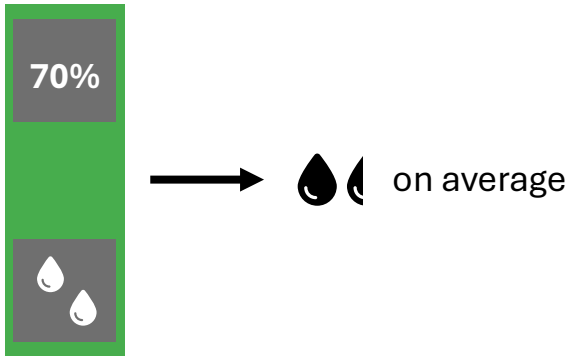
Representational geometry:

How the **sensitivity profile** of the neuronal population **varies** across different time windows.

What do OFC neurons do?

Training rational models:

Value = Expected reward

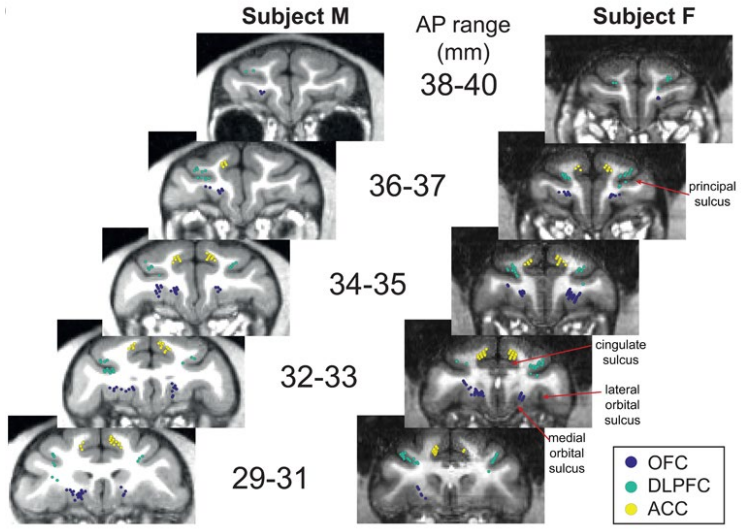


- Start with random weights
- Train on a dataset of random trials

Repeat 1000 times

Comparison with OFC neural activity:

Hunt et al., 2018. Figure 3b.

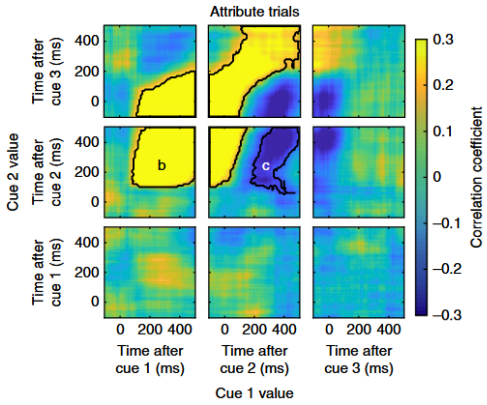


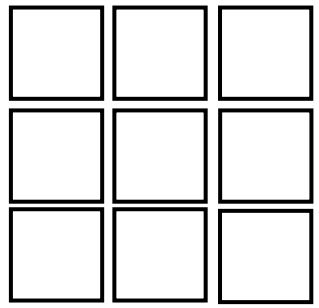
Electrophysiological recordings in the OFC of two macaque monkeys (183 neurons, 22 000 trials in total).

Representational geometry:

How the **sensitivity profile** of the neuronal population **varies** across different time windows.

Hunt et al., 2018. Figure 5g.



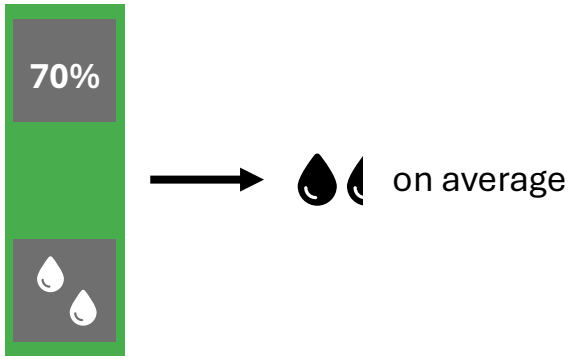


With our models

What do OFC neurons do?

Training rational models:

Value = Expected reward

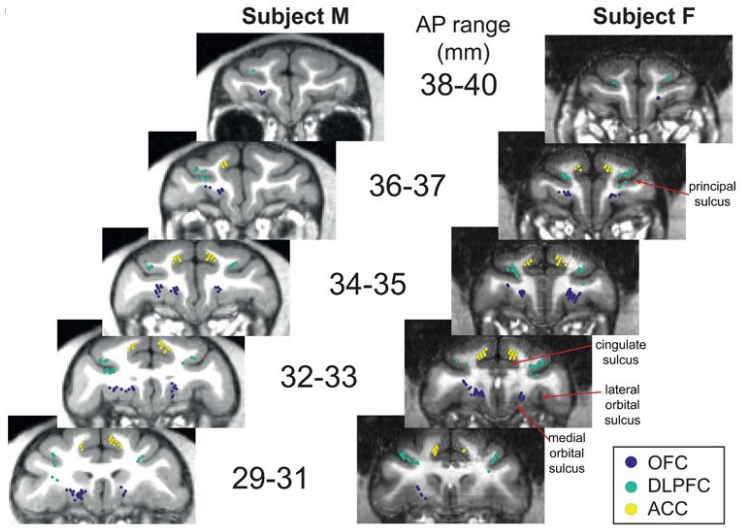


- Start with random weights
- Train on a dataset of random trials

Repeat 1000 times

Comparison with OFC neural activity:

Hunt et al., 2018. Figure 3b.

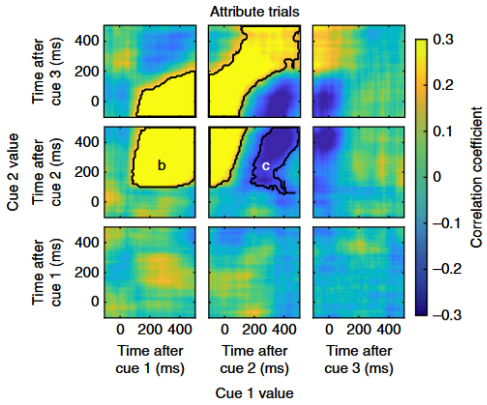


Electrophysiological recordings in the OFC of two macaque monkeys (183 neurons, 22 000 trials in total).

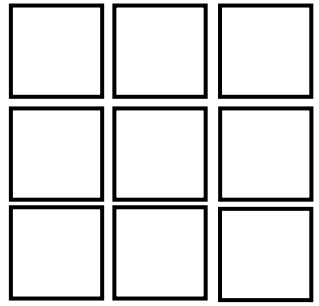
Representational geometry:

How the **sensitivity profile** of the neuronal population **varies** across different time windows.

Hunt et al., 2018. Figure 5g.



Compute the distance

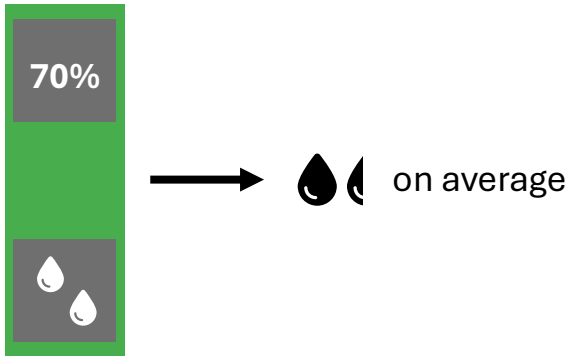


With our models

What do OFC neurons do?

Training rational models:

Value = Expected reward

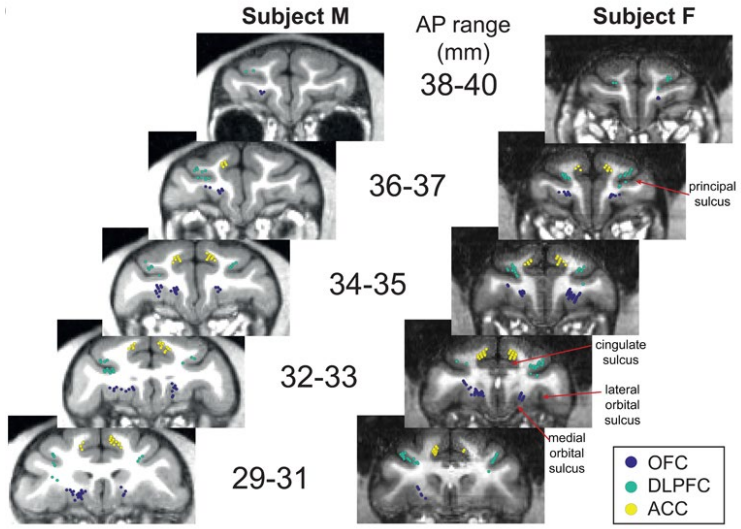


- Start with random weights
- Train on a dataset of random trials

Repeat 1000 times

Comparison with OFC neural activity:

Hunt et al., 2018. Figure 3b.



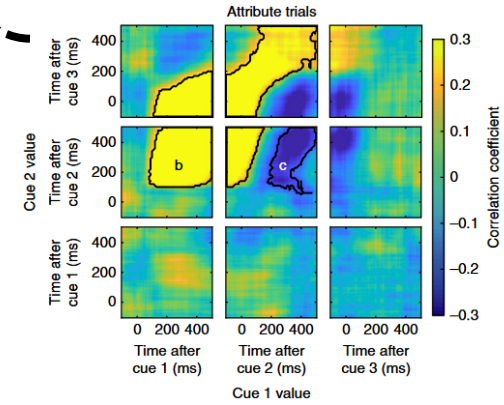
Electrophysiological recordings in the OFC of two macaque monkeys (183 neurons, 22 000 trials in total).

Representational geometry:

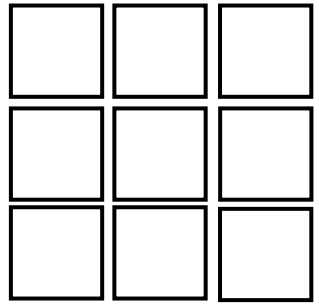
How the **sensitivity profile** of the neuronal population **varies** across different time windows.

Qualitative interpretation: signature of value comparison in the OFC

Hunt et al., 2018. Figure 5g.

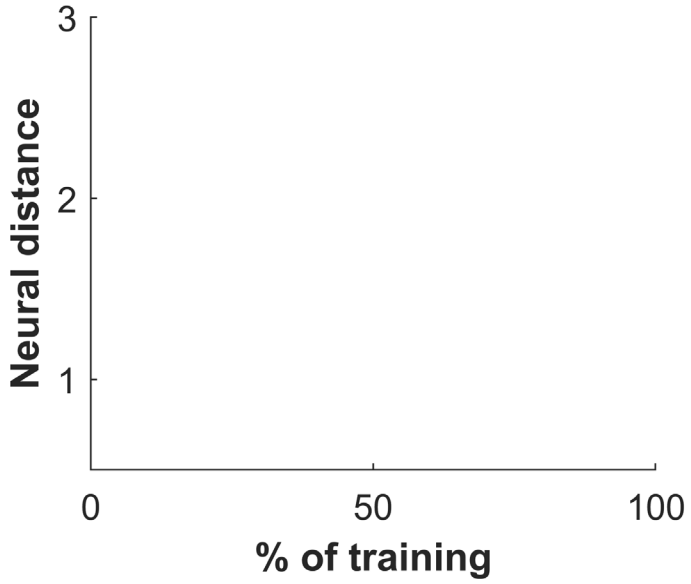


Compute the distance

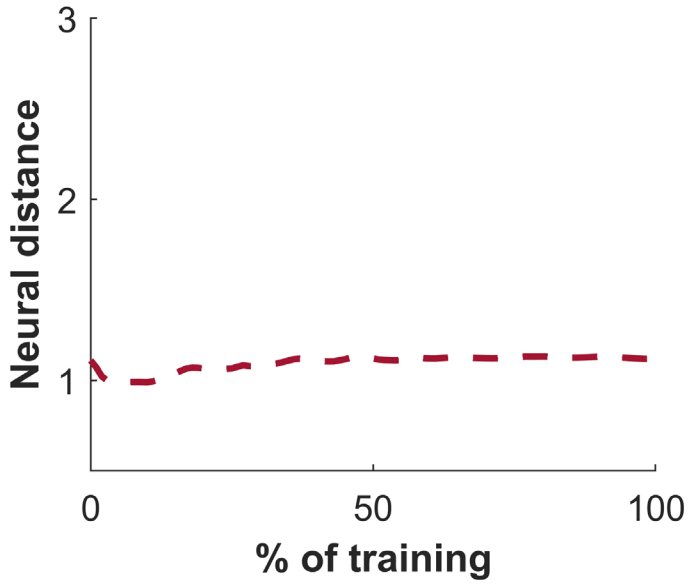


With our models

What do OFC neurons do?

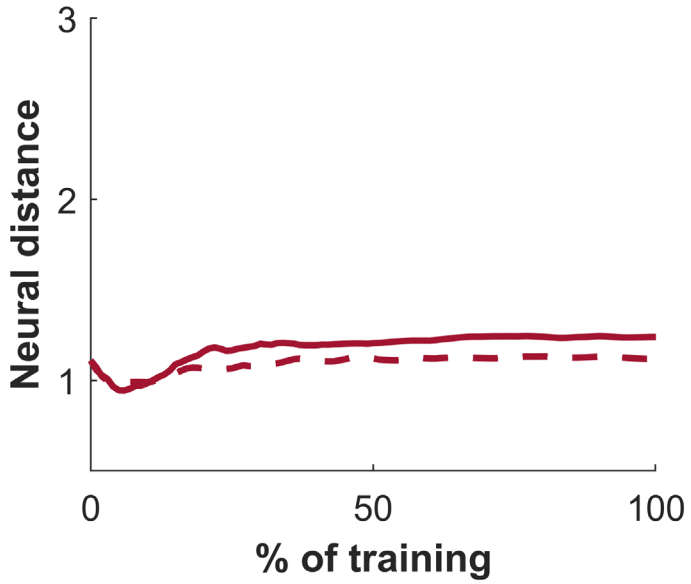


What do OFC neurons do?



Spatial → **Attentional** focus (value ***synthesis***)

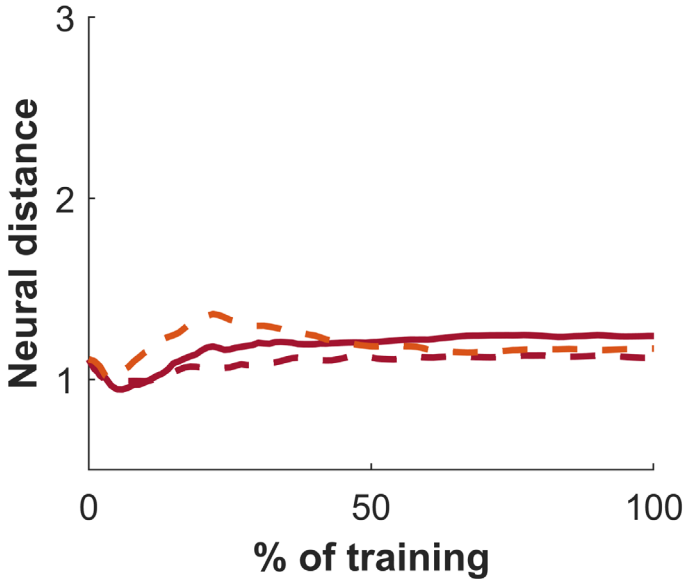
What do OFC neurons do?



Spatial → **Attentional** focus (value ***synthesis***)

Spatial → **Attentional** focus (value ***comparison***)

What do OFC neurons do?

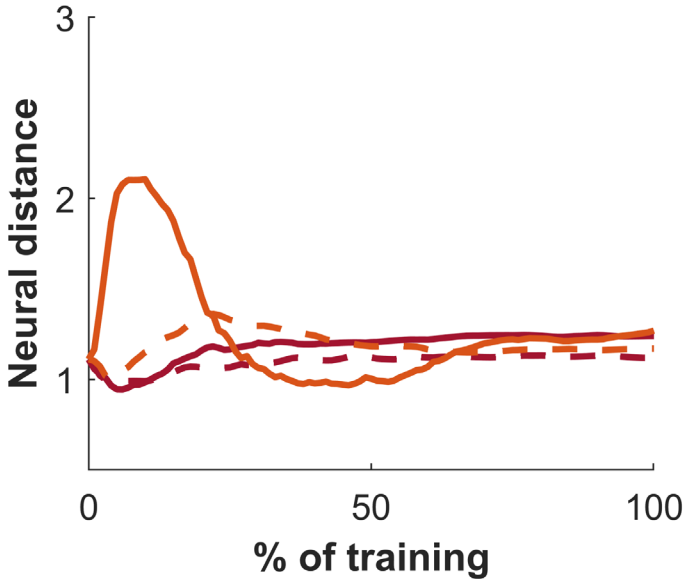


Spatial → **Attentional** focus (value ***synthesis***)

Spatial → **Attentional** focus (value ***comparison***)

Spatial → **Temporal** order (value ***synthesis***)

What do OFC neurons do?



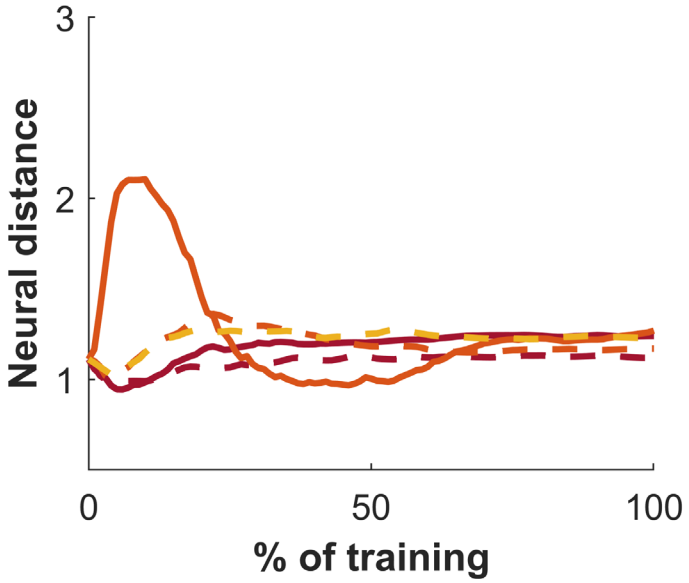
Spatial → **Attentional** focus (value ***synthesis***)

Spatial → **Attentional** focus (value ***comparison***)

Spatial → **Temporal** order (value ***synthesis***)

Spatial → **Temporal** order (value ***comparison***)

What do OFC neurons do?



Spatial → **Attentional** focus (value ***synthesis***)

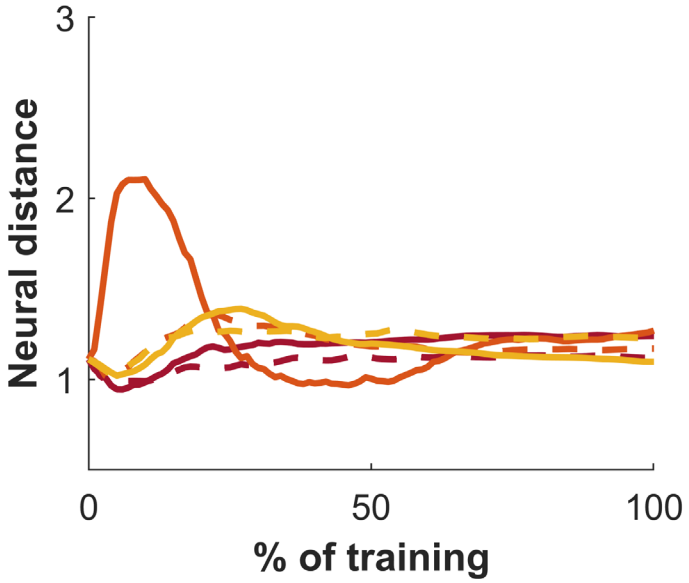
Spatial → **Attentional** focus (value ***comparison***)

Spatial → **Temporal** order (value ***synthesis***)

Spatial → **Temporal** order (value ***comparison***)

Spatial → **Spatial** (value ***synthesis***)

What do OFC neurons do?



Spatial → **Attentional** focus (value ***synthesis***)

Spatial → **Attentional** focus (value ***comparison***)

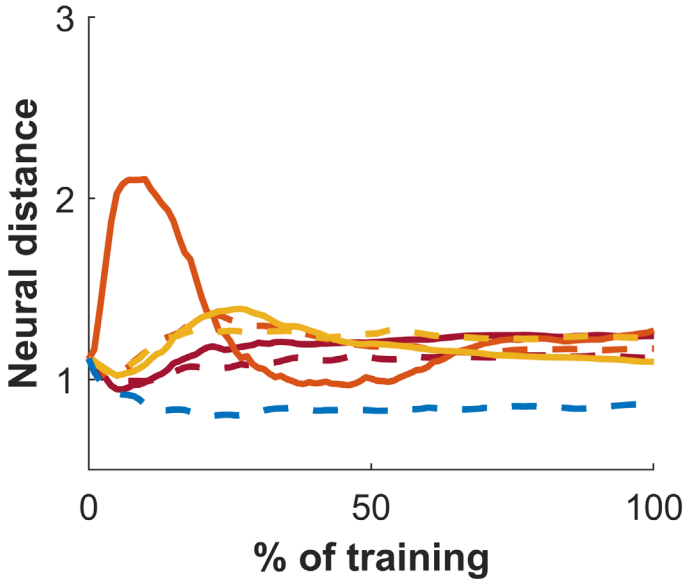
Spatial → **Temporal** order (value ***synthesis***)

Spatial → **Temporal** order (value ***comparison***)

Spatial → **Spatial** (value ***synthesis***)

Spatial → **Spatial** (value ***comparison***)

What do OFC neurons do?



Spatial → **Attentional** focus (value ***synthesis***)

Spatial → **Attentional** focus (value ***comparison***)

Spatial → **Temporal** order (value ***synthesis***)

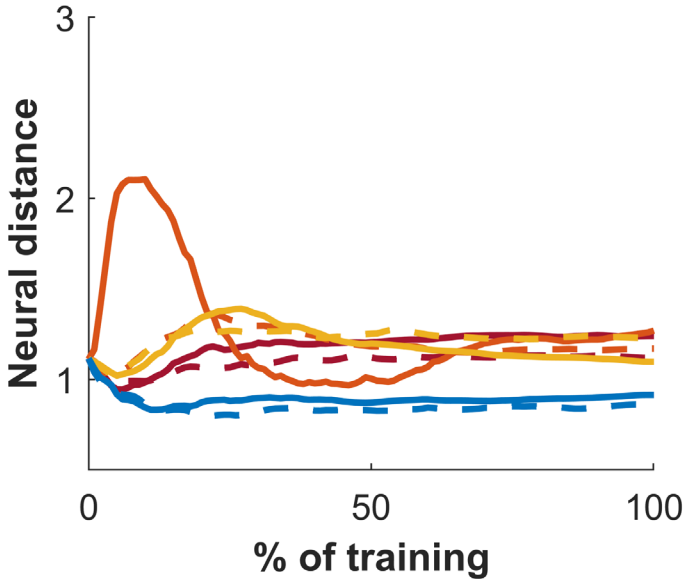
Spatial → **Temporal** order (value ***comparison***)

Spatial → **Spatial** (value ***synthesis***)

Spatial → **Spatial** (value ***comparison***)

Temporal order → **Attentional** focus (value ***synthesis***)

What do OFC neurons do?



Spatial → **Attentional** focus (value ***synthesis***)

Spatial → **Attentional** focus (value ***comparison***)

Spatial → **Temporal** order (value ***synthesis***)

Spatial → **Temporal** order (value ***comparison***)

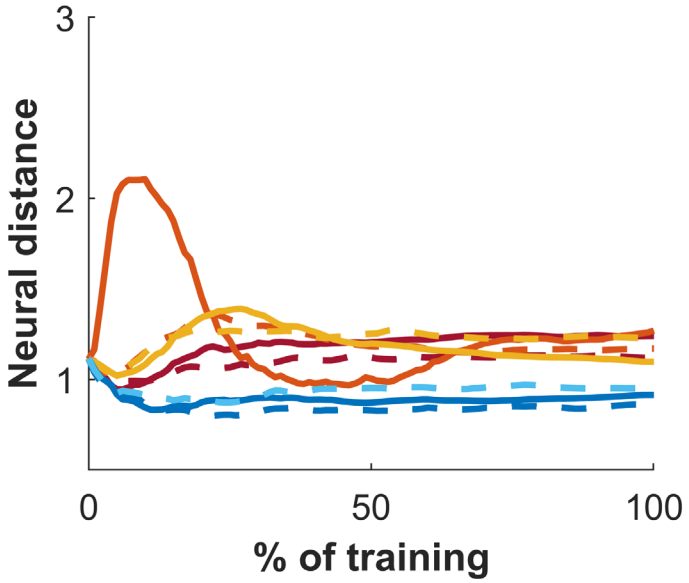
Spatial → **Spatial** (value ***synthesis***)

Spatial → **Spatial** (value ***comparison***)

Temporal order → **Attentional** focus (value ***synthesis***)

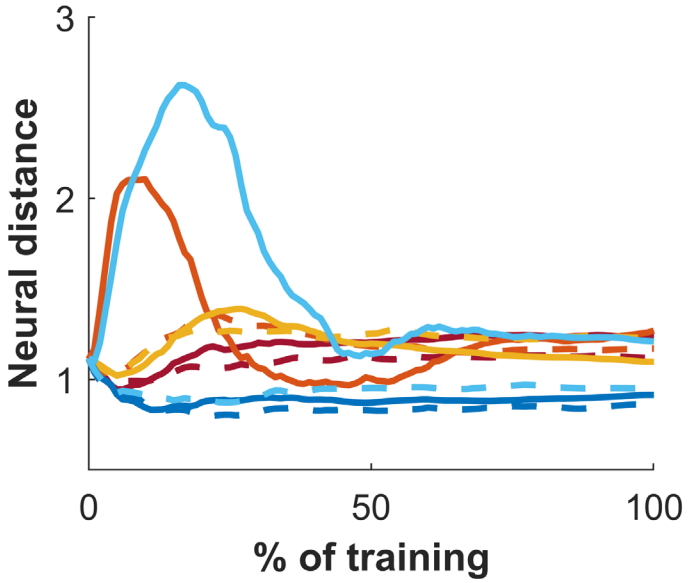
Temporal order → **Attentional** focus (value ***comparison***)

What do OFC neurons do?



- **Spatial** → **Attentional** focus (value ***synthesis***)
- **Spatial** → **Attentional** focus (value ***comparison***)
- **Spatial** → **Temporal** order (value ***synthesis***)
- **Spatial** → **Temporal** order (value ***comparison***)
- **Spatial** → **Spatial** (value ***synthesis***)
- **Spatial** → **Spatial** (value ***comparison***)
- **Temporal** order → **Attentional** focus (value ***synthesis***)
- **Temporal** order → **Attentional** focus (value ***comparison***)
- **Temporal** order → **Temporal** order (value ***synthesis***)

What do OFC neurons do?



Spatial → **Attentional** focus (value ***synthesis***)

Spatial → **Attentional** focus (value ***comparison***)

Spatial → **Temporal** order (value ***synthesis***)

Spatial → **Temporal** order (value ***comparison***)

Spatial → **Spatial** (value ***synthesis***)

Spatial → **Spatial** (value ***comparison***)

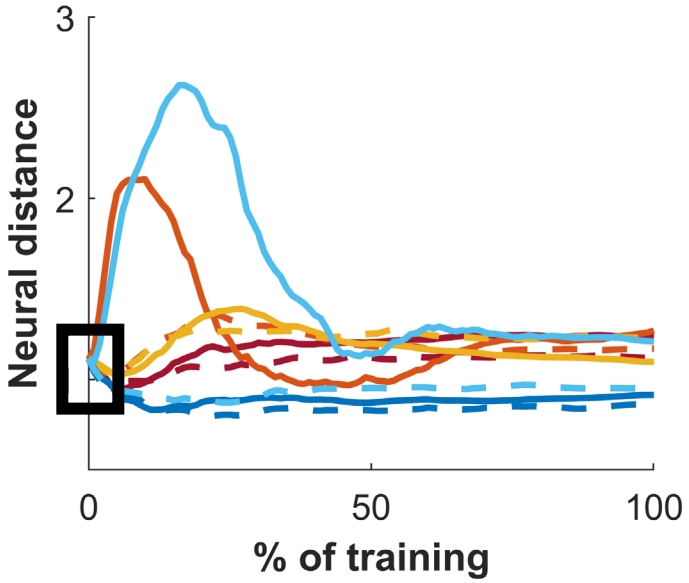
Temporal order → **Attentional** focus (value ***synthesis***)

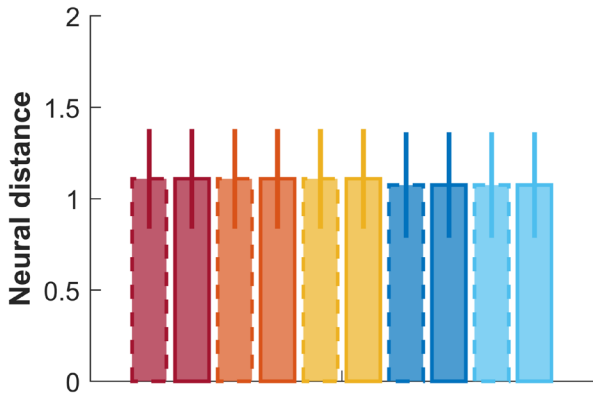
Temporal order → **Attentional** focus (value ***comparison***)

Temporal order → **Temporal** order (value ***synthesis***)

Temporal order → **Temporal** order (value ***comparison***)

What do OFC neurons do?

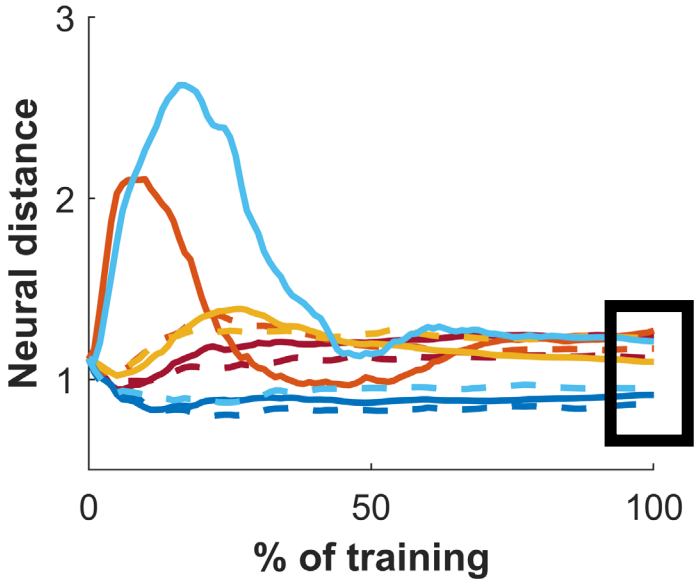


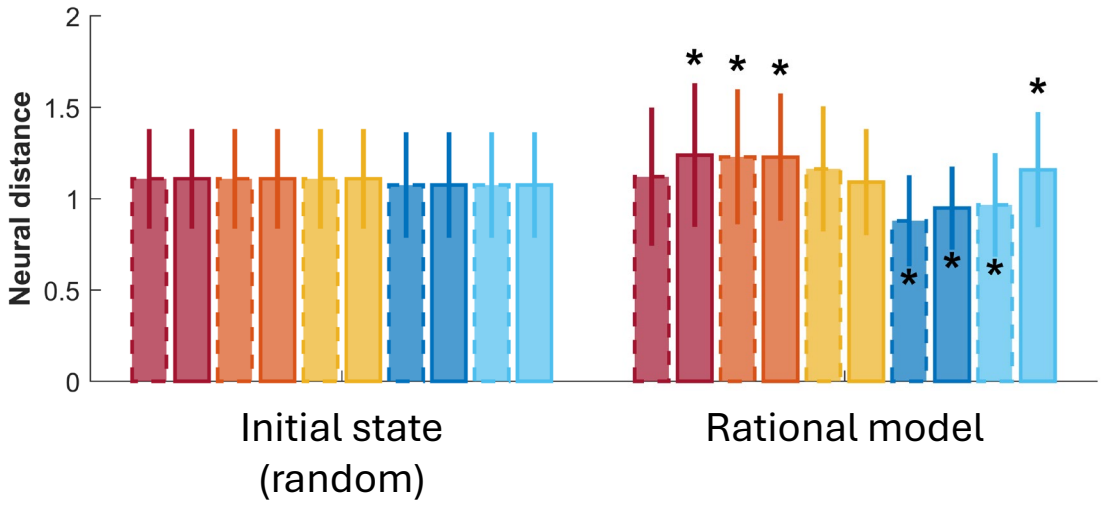


Initial state
(random)

- **Spatial** → **Attentional** focus (value ***synthesis***)
- **Spatial** → **Attentional** focus (value ***comparison***)
- **Spatial** → **Temporal** order (value ***synthesis***)
- **Spatial** → **Temporal** order (value ***comparison***)
- **Spatial** → **Spatial** (value ***synthesis***)
- **Spatial** → **Spatial** (value ***comparison***)
- **Temporal** order → **Attentional** focus (value ***synthesis***)
- **Temporal** order → **Attentional** focus (value ***comparison***)
- **Temporal** order → **Temporal** order (value ***synthesis***)
- **Temporal** order → **Temporal** order (value ***comparison***)

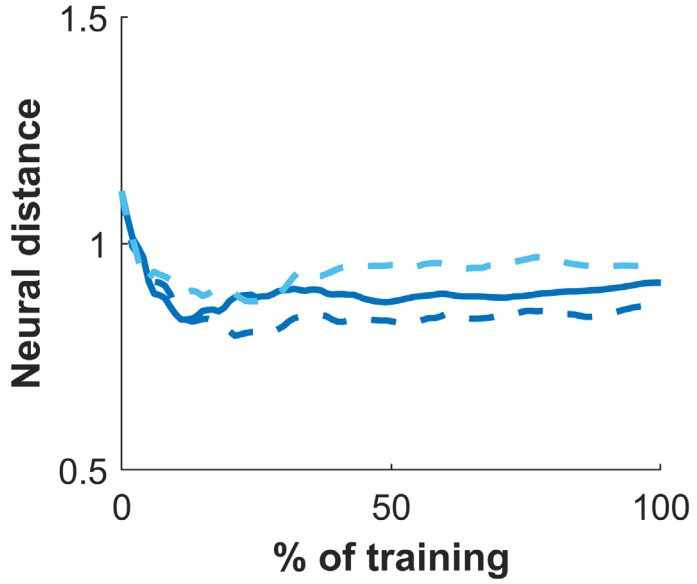
What do OFC neurons do?





- **Spatial** → **Attentional** focus (value ***synthesis***)
- **Spatial** → **Attentional** focus (value ***comparison***)
- **Spatial** → **Temporal** order (value ***synthesis***)
- **Spatial** → **Temporal** order (value ***comparison***)
- **Spatial** → **Spatial** (value ***synthesis***)
- **Spatial** → **Spatial** (value ***comparison***)
- **Temporal** order → **Attentional** focus (value ***synthesis***)
- **Temporal** order → **Attentional** focus (value ***comparison***)
- **Temporal** order → **Temporal** order (value ***synthesis***)
- **Temporal** order → **Temporal** order (value ***comparison***)

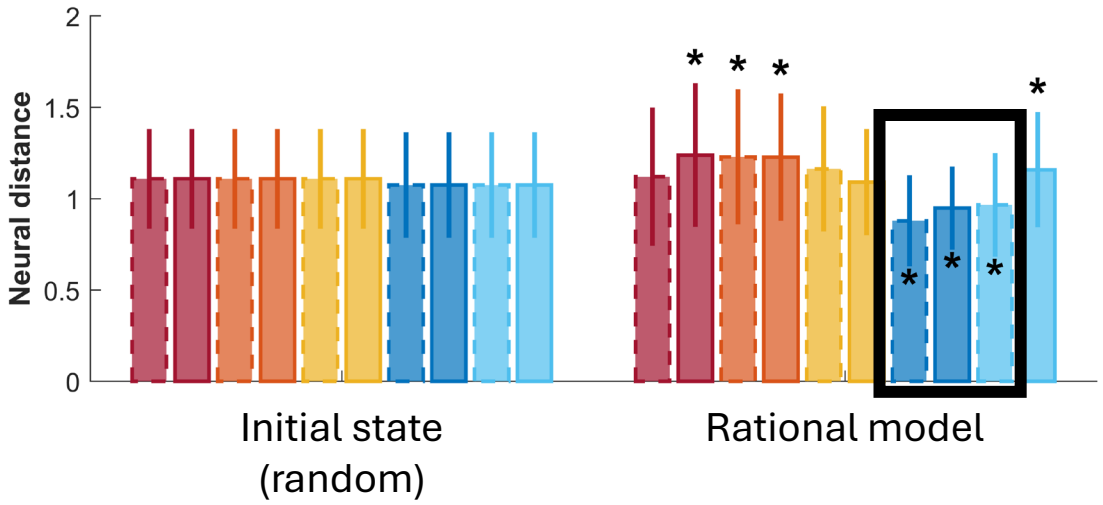
What do OFC neurons do?



Temporal order → **Attentional** focus (value ***synthesis***)

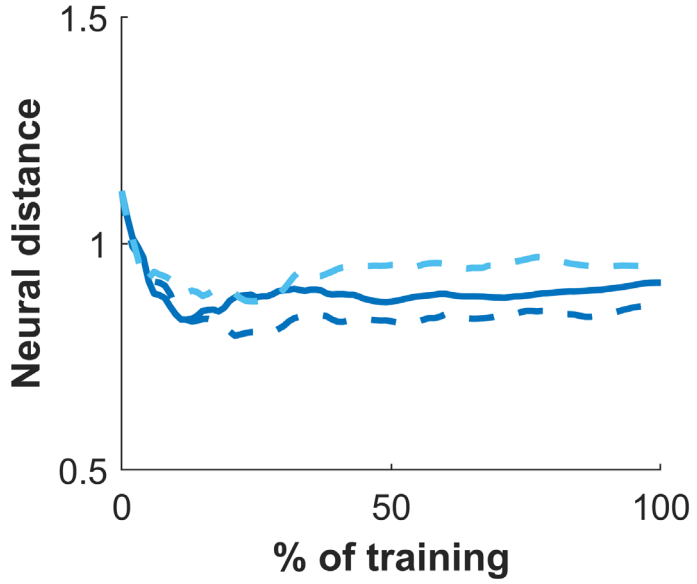
Temporal order → **Attentional** focus (value ***comparison***)

Temporal order → **Temporal** order (value ***synthesis***)



Three candidate models generate a realistic representational geometry.

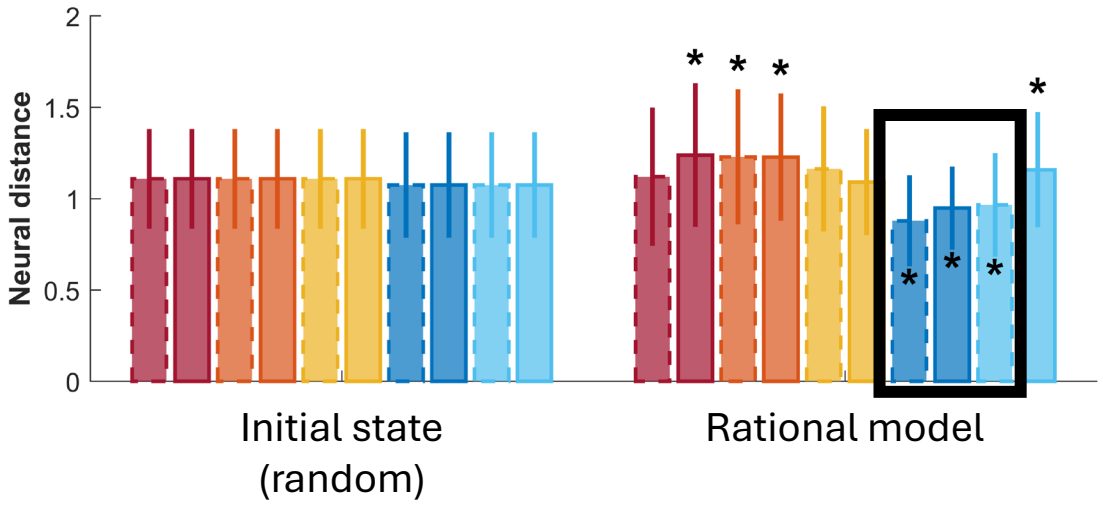
What do OFC neurons do?



Temporal order → **Attentional** focus (value ***synthesis***)

Temporal order → **Attentional** focus (value ***comparison***)

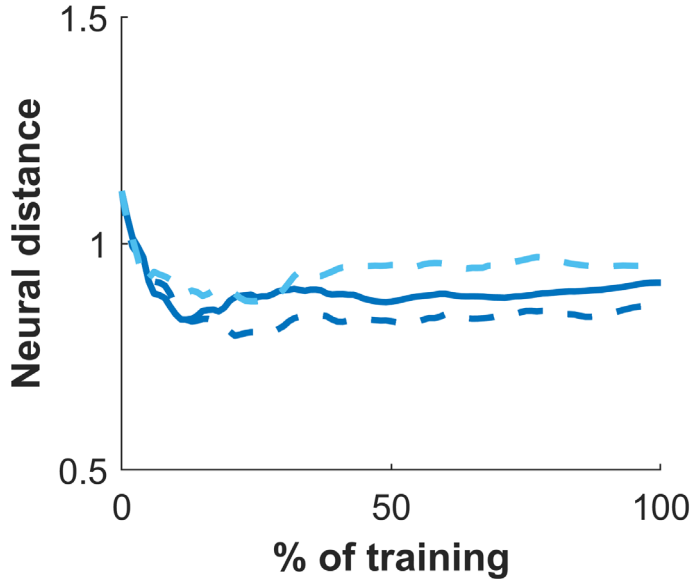
Temporal order → **Temporal** order (value ***synthesis***)



Three candidate models generate a realistic representational geometry.

Both **value synthesis** and **value comparison** scenarii generate key neural features of the OFC.

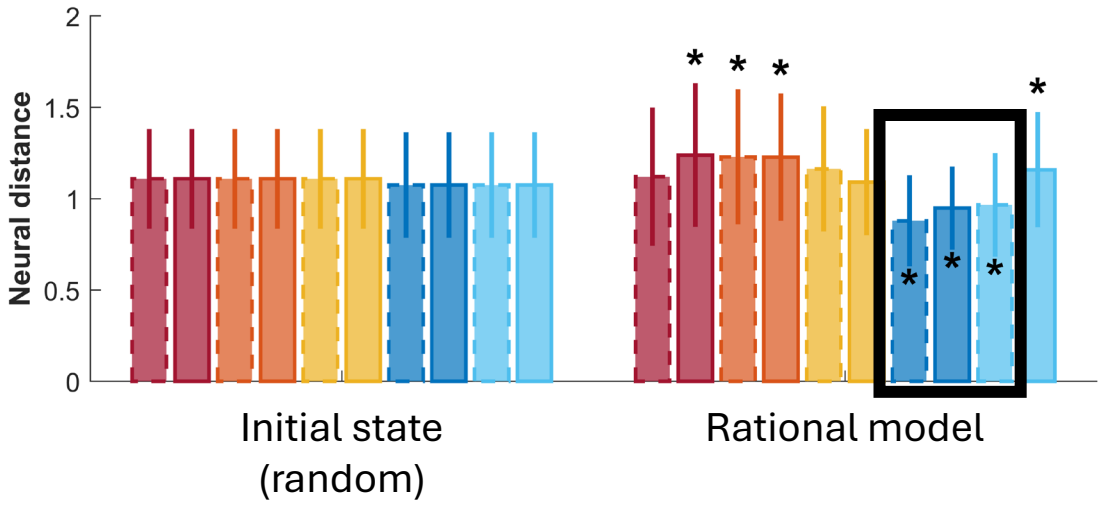
What do OFC neurons do?



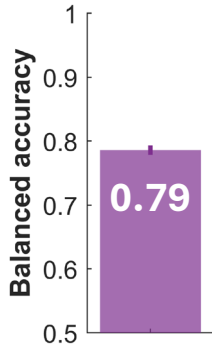
Temporal order → **Attentional** focus (value ***synthesis***)

Temporal order → **Attentional** focus (value ***comparison***)

Temporal order → **Temporal** order (value ***synthesis***)

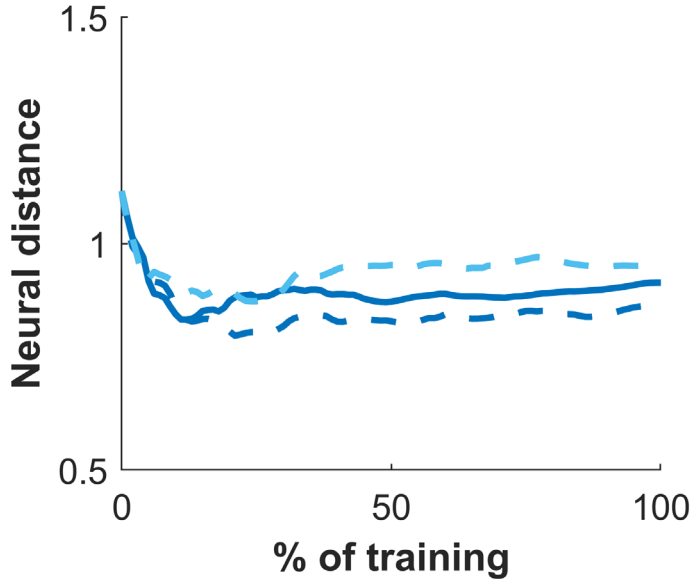


Three candidate models generate a realistic representational geometry.



Both **value synthesis** and **value comparison** scenarii generate key neural features of the OFC.

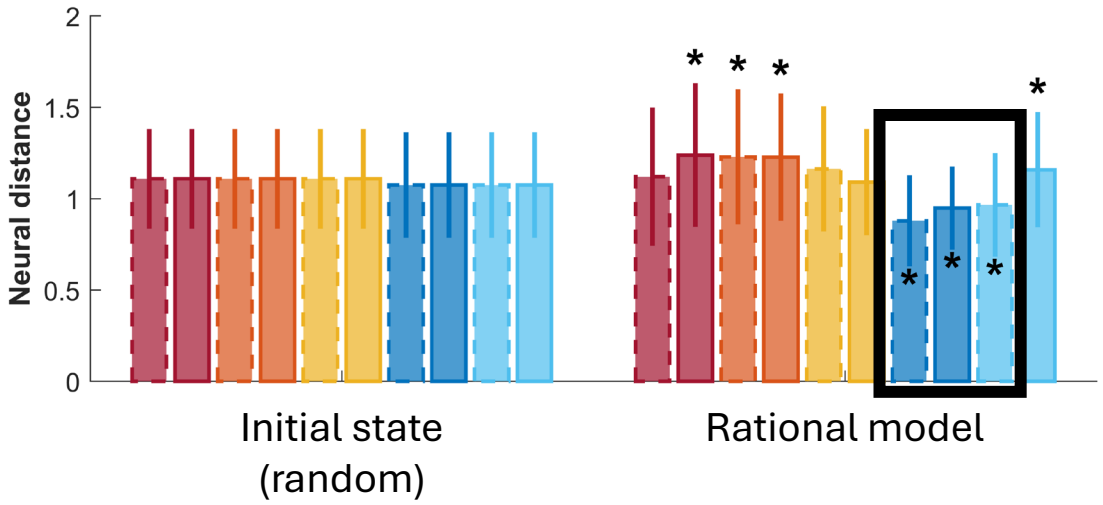
What do OFC neurons do?



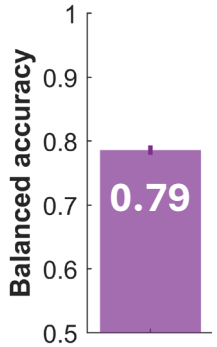
Temporal order → **Attentional** focus (value ***synthesis***)

Temporal order → **Attentional** focus (value ***comparison***)

Temporal order → **Temporal** order (value ***synthesis***)



Three candidate models generate a realistic representational geometry.



Both **value synthesis** and **value comparison** scenarii generate key neural features of the OFC.

What happens when we try to explain irrational behavior?

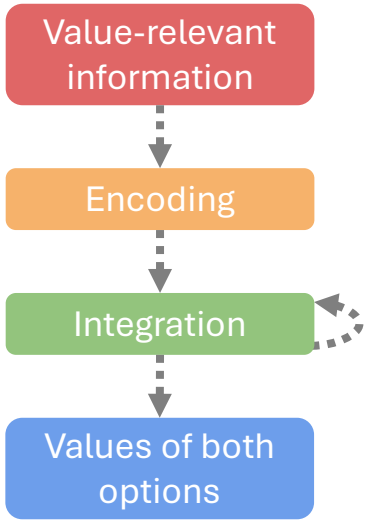
Modelling irrational decisions

Modelling irrational decisions

Distorting model architecture:

Modelling irrational decisions

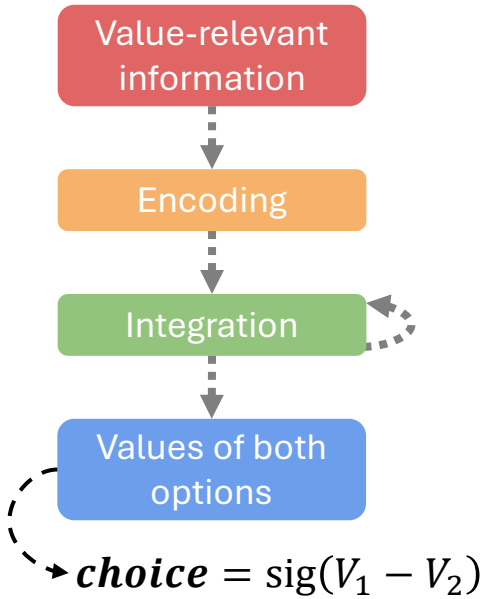
Distorting model architecture:



- Start from a rational RNN

Modelling irrational decisions

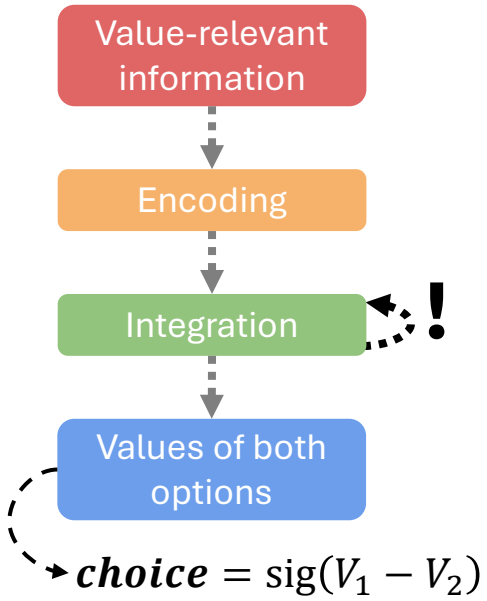
Distorting model architecture:



- Start from a rational RNN
- Generate choices from its output(s)

Modelling irrational decisions

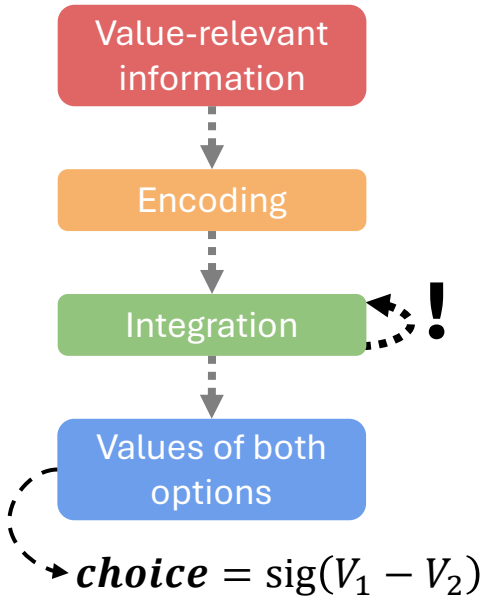
Distorting model architecture:



- Start from a rational RNN
- Generate choices from its output(s)
- Freeze all parameters except recurrent connections

Modelling irrational decisions

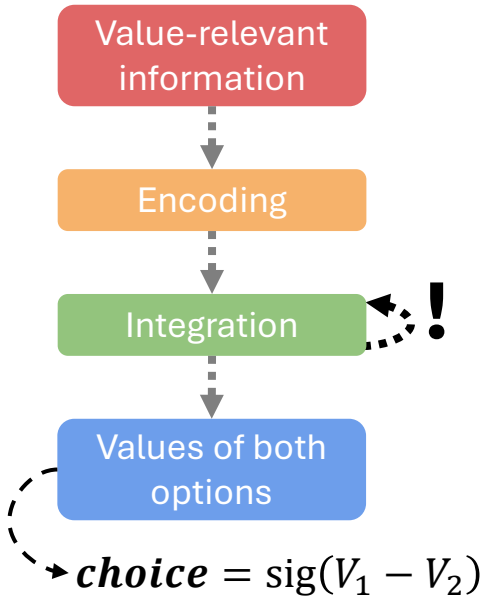
Distorting model architecture:

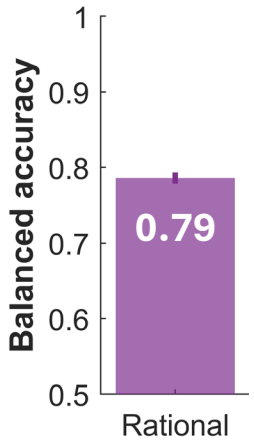


- Start from a rational RNN
- Generate choices from its output(s)
- Freeze all parameters except recurrent connections
- Fit the choices of a monkey on a subset of its trials

Modelling irrational decisions

Distorting model architecture:

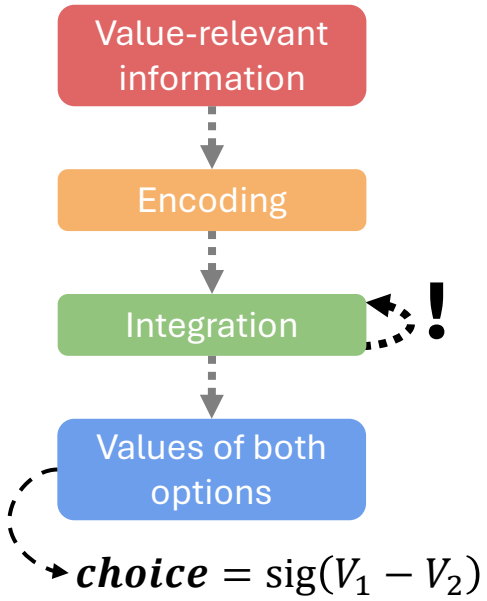




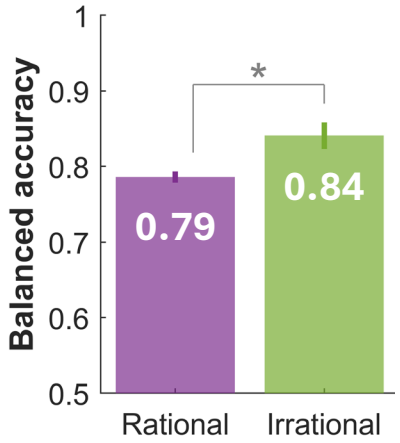
- Start from a rational RNN
- Generate choices from its output(s)
- Freeze all parameters except recurrent connections
- Fit the choices of a monkey on a subset of its trials

Modelling irrational decisions

Distorting model architecture:



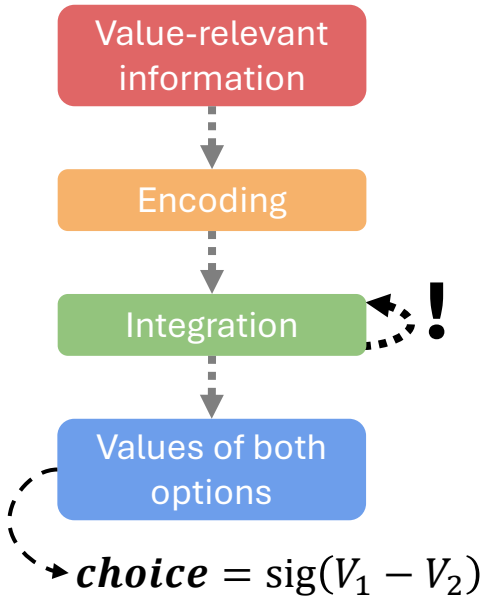
- Start from a rational RNN
- Generate choices from its output(s)
- Freeze all parameters except recurrent connections
- Fit the choices of a monkey on a subset of its trials



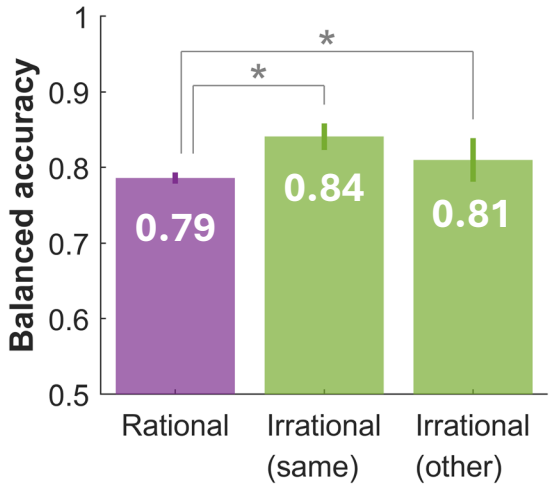
Additional mechanisms capture 26% of monkey irrational choices.

Modelling irrational decisions

Distorting model architecture:



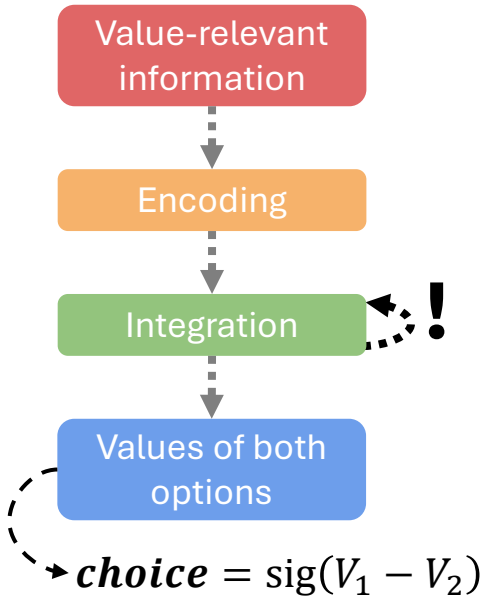
- Start from a rational RNN
- Generate choices from its output(s)
- Freeze all parameters except recurrent connections
- Fit the choices of a monkey on a subset of its trials



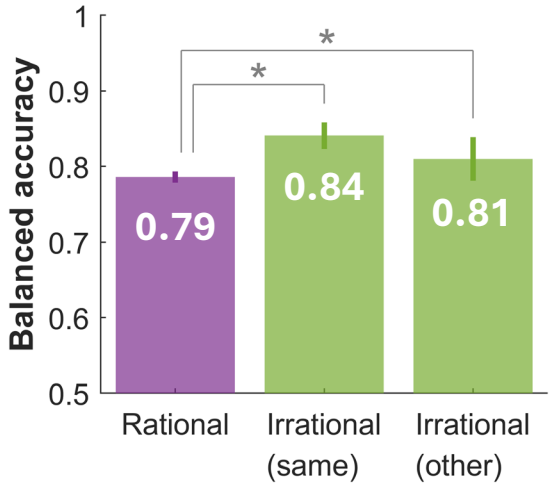
Additional mechanisms capture 26% of monkey irrational choices.

Modelling irrational decisions

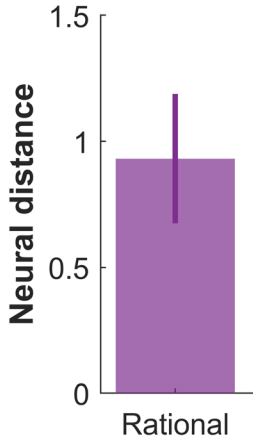
Distorting model architecture:

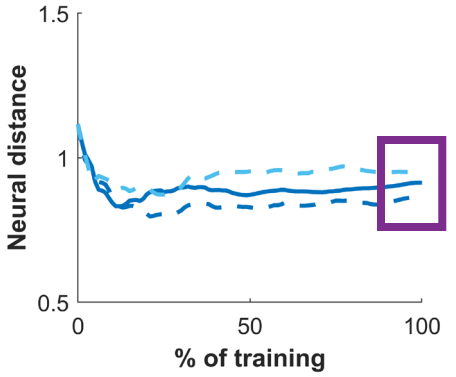


- Start from a rational RNN
- Generate choices from its output(s)
- Freeze all parameters except recurrent connections
- Fit the choices of a monkey on a subset of its trials



Additional mechanisms capture 26% of monkey irrational choices.





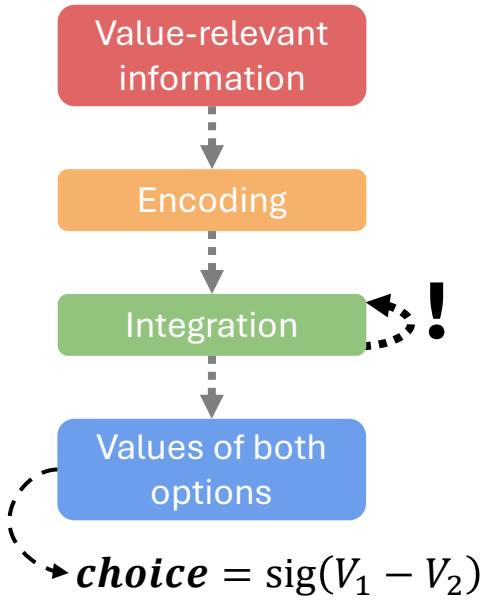
Temporal order → **Attentional** focus (value ***synthesis***)

Temporal order → **Attentional** focus (value ***comparison***)

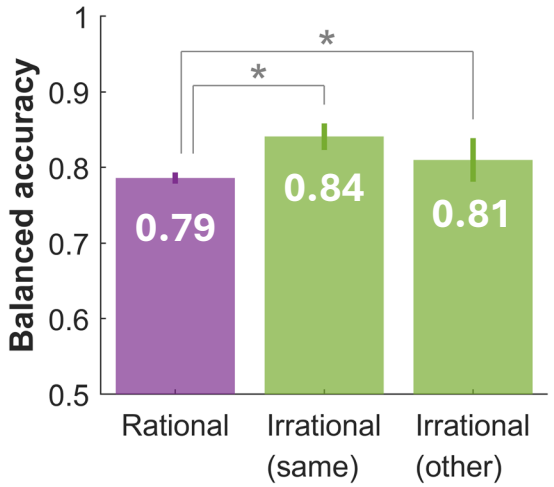
Temporal order → **Temporal** order (value ***synthesis***)

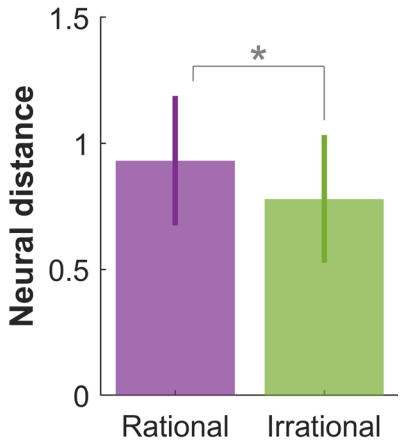
Modelling irrational decisions

Distorting model architecture:

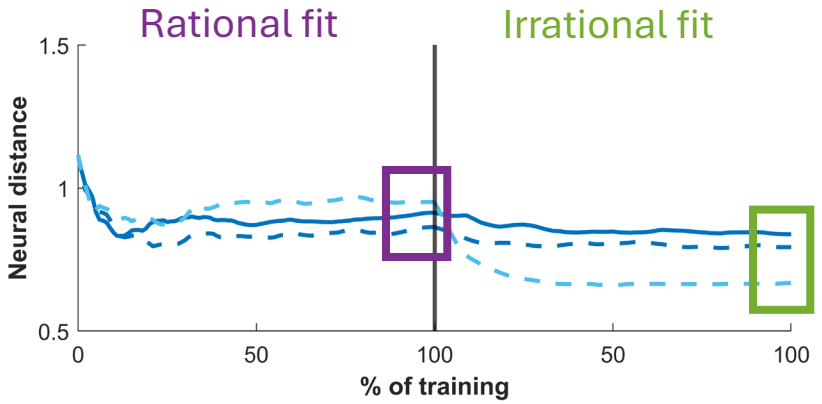


- Start from a rational RNN
- Generate choices from its output(s)
- Freeze all parameters except recurrent connections
- Fit the choices of a monkey on a subset of its trials





Additional mechanisms capture 26% of monkey irrational choices.



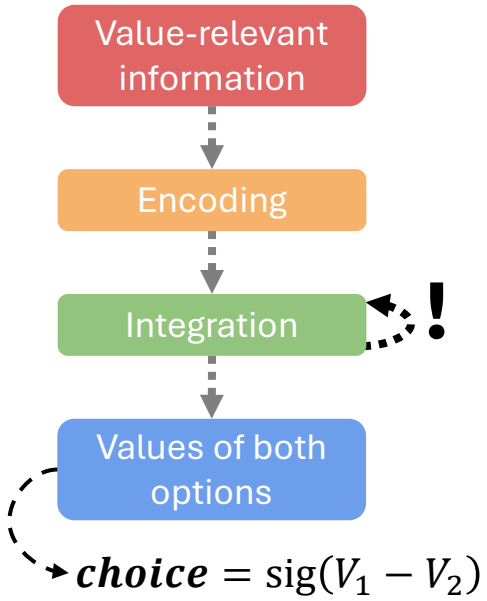
Temporal order → **Attentional** focus (value ***synthesis***)

Temporal order → **Attentional** focus (value ***comparison***)

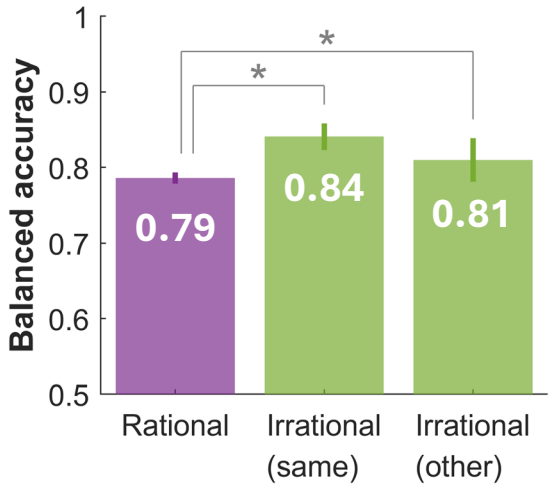
Temporal order → **Temporal** order (value ***synthesis***)

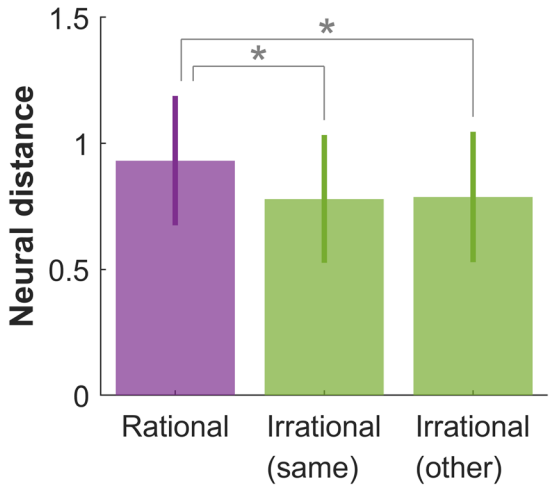
Modelling irrational decisions

Distorting model architecture:

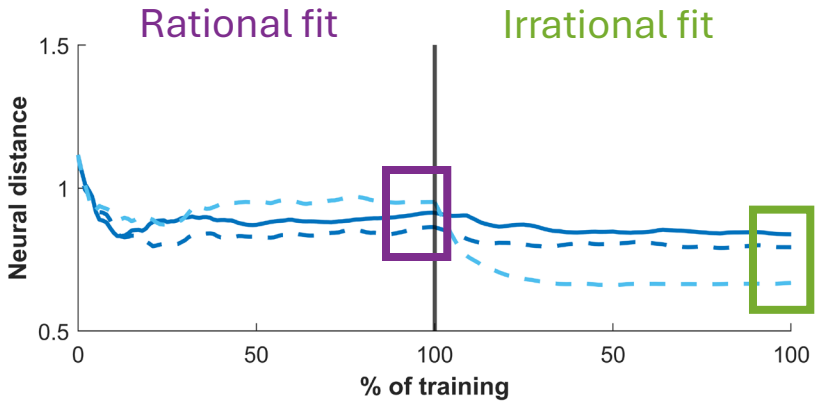


- Start from a rational RNN
- Generate choices from its output(s)
- Freeze all parameters except recurrent connections
- Fit the choices of a monkey on a subset of its trials





Additional mechanisms capture 26% of monkey irrational choices.



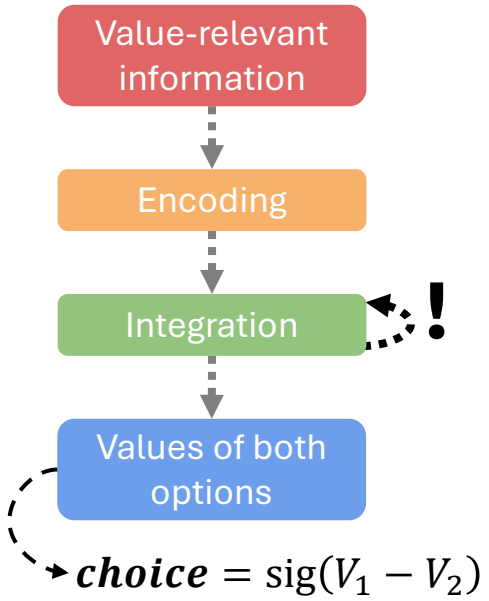
Temporal order → **Attentional** focus (value ***synthesis***)

Temporal order → **Attentional** focus (value ***comparison***)

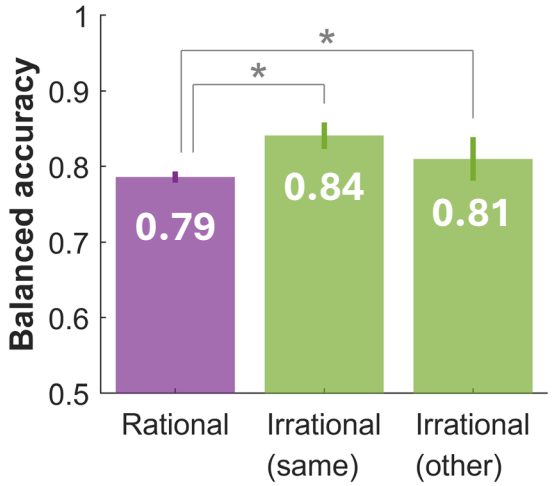
Temporal order → **Temporal** order (value ***synthesis***)

Modelling irrational decisions

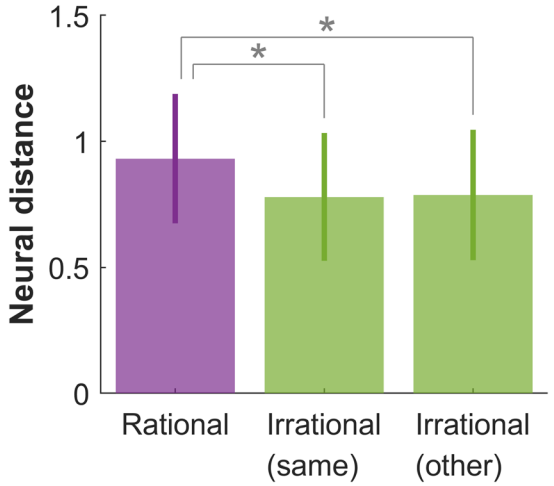
Distorting model architecture:



- Start from a rational RNN
- Generate choices from its output(s)
- Freeze all parameters except recurrent connections
- Fit the choices of a monkey on a subset of its trials



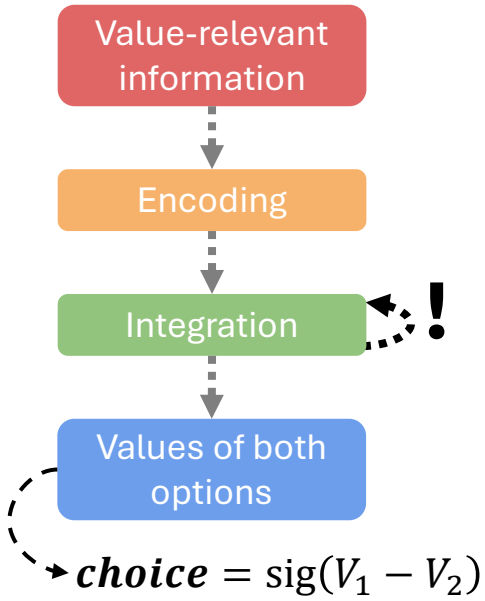
Additional mechanisms capture 26% of monkey irrational choices.



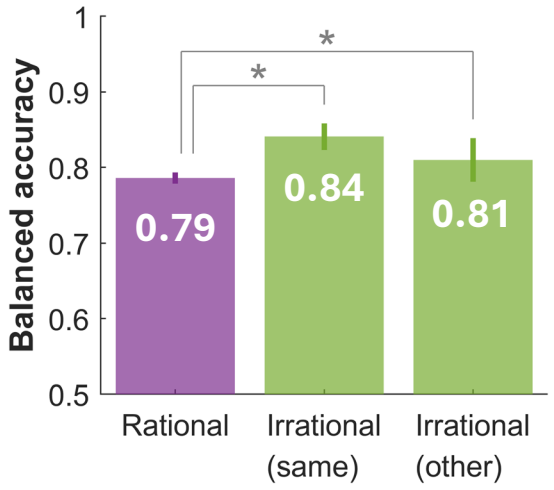
Models capturing part of the behavior generate neural geometries which are even closer to the OFC geometry.

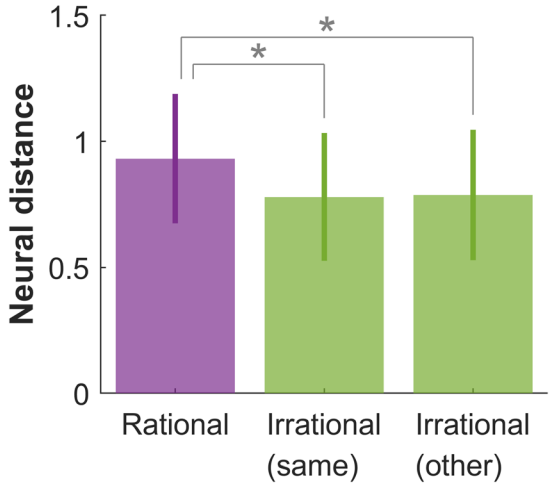
Modelling irrational decisions

Distorting model architecture:



- Start from a rational RNN
- Generate choices from its output(s)
- Freeze all parameters except recurrent connections
- Fit the choices of a monkey on a subset of its trials





Additional mechanisms capture 26% of monkey irrational choices.

Our models of irrational decisions are **behaviorally** and **neurally** realistic.

Models capturing part of the behavior generate neural geometries which are even closer to the OFC geometry.

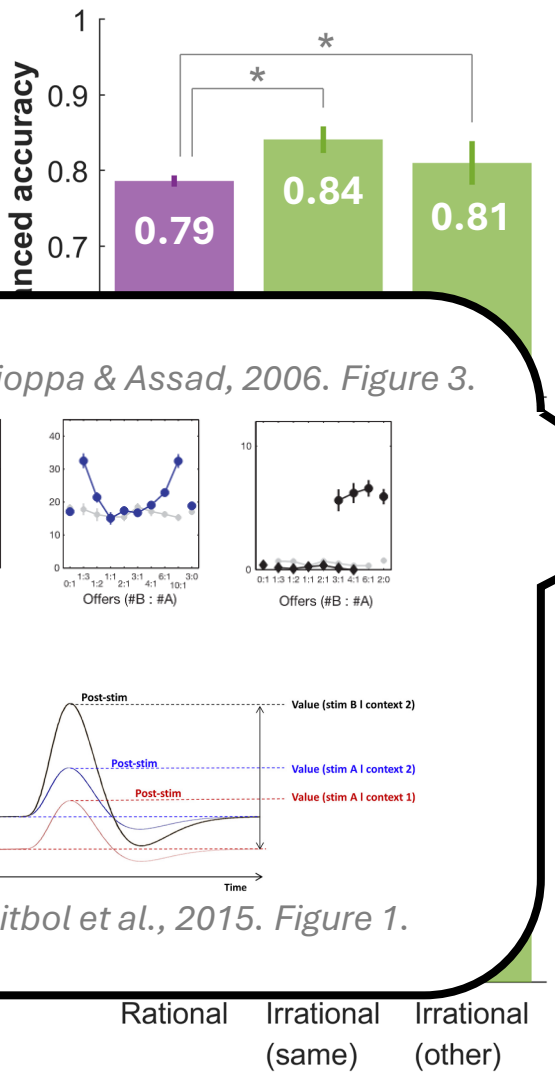
Modelling irrational decisions

Distorting model architecture:

Value-relevant information

Encoding

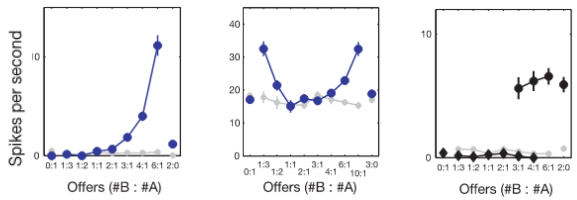
- St
- G
- Fr
recur
- Fit the choices of a monkey on a subset of its trials

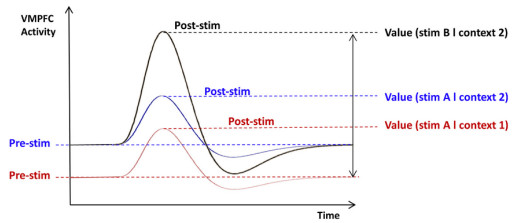


Rational and irrational models also generate **other OFC neural properties**:

- Offer value cells
  Chosen value cells
  Chosen offer cells
- Neural autocorrelation
- Confidence encoding

Padoa-Schioppa & Assad, 2006. Figure 3.





Abitbol et al., 2015. Figure 1.

Our models of irrational decisions are **behaviorally** and **neurally** realistic.

Additional mechanisms capture 26% of monkey irrational choices.

Models capturing part of the behavior generate neural geometries which are even closer to the OFC geometry.

What mechanisms underlie irrational decisions?

What mechanisms underlie irrational decisions?

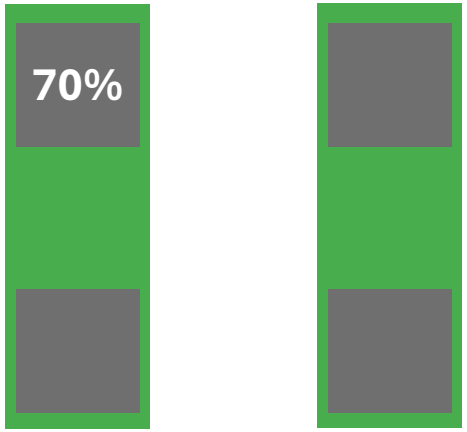
Interferences:

How much the **output** of a model **varies** with the **order** of information acquisition.

What mechanisms underlie irrational decisions?

Interferences:

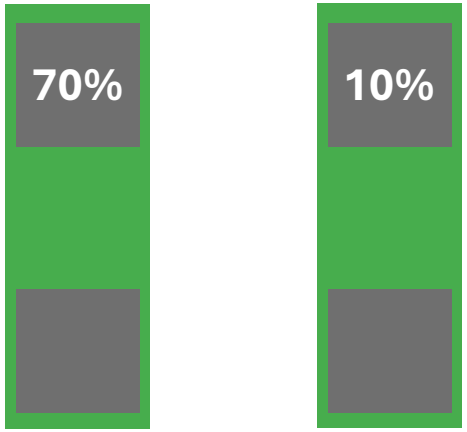
How much the **output** of a model **varies** with the **order** of information acquisition.



What mechanisms underlie irrational decisions?

Interferences:

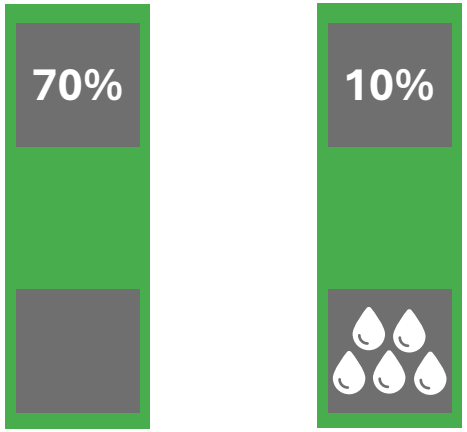
How much the **output** of a model **varies** with the **order** of information acquisition.



What mechanisms underlie irrational decisions?

Interferences:

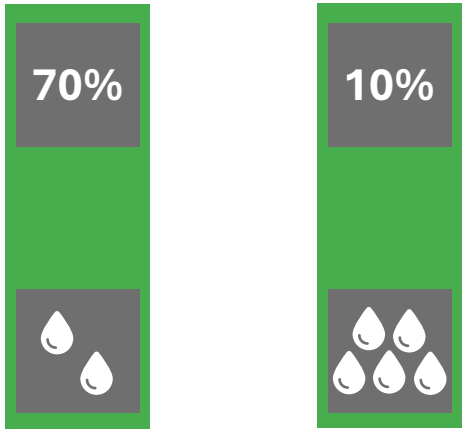
How much the **output** of a model **varies** with the **order** of information acquisition.



What mechanisms underlie irrational decisions?

Interferences:

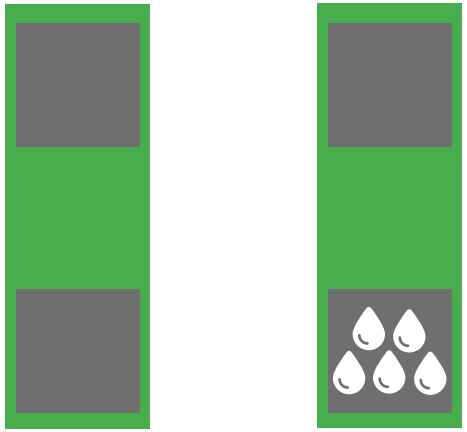
How much the **output** of a model **varies** with the **order** of information acquisition.



What mechanisms underlie irrational decisions?

Interferences:

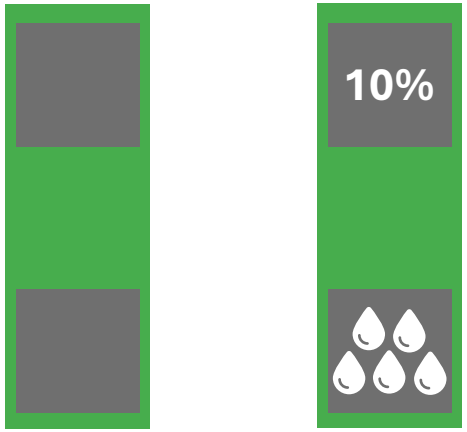
How much the **output** of a model **varies** with the **order** of information acquisition.



What mechanisms underlie irrational decisions?

Interferences:

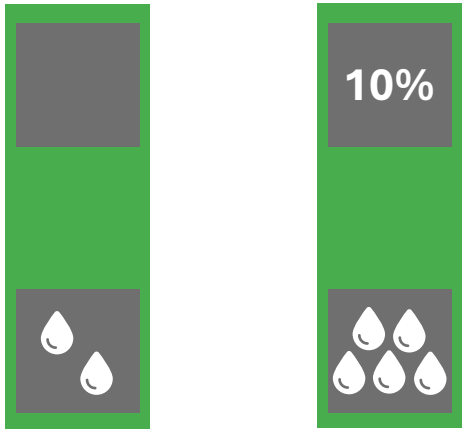
How much the **output** of a model **varies** with the **order** of information acquisition.



What mechanisms underlie irrational decisions?

Interferences:

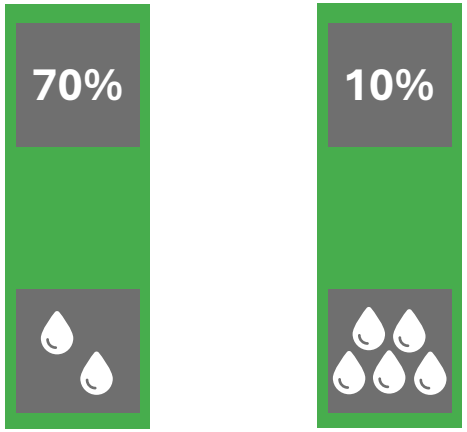
How much the **output** of a model **varies** with the **order** of information acquisition.



What mechanisms underlie irrational decisions?

Interferences:

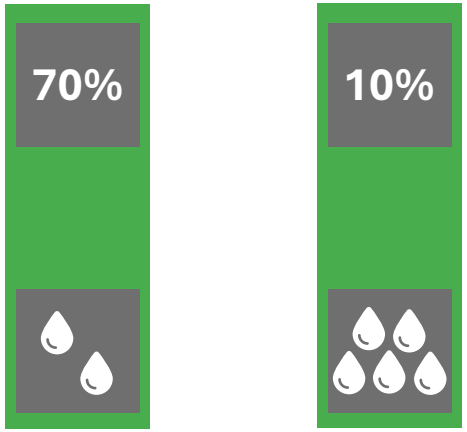
How much the **output** of a model **varies** with the **order** of information acquisition.

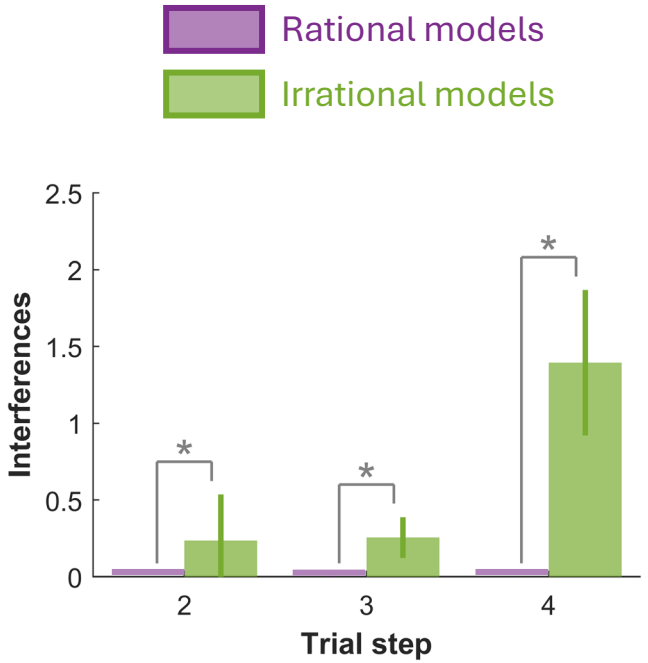


What mechanisms underlie irrational decisions?

Interferences:

How much the **output** of a model **varies** with the **order** of information acquisition.

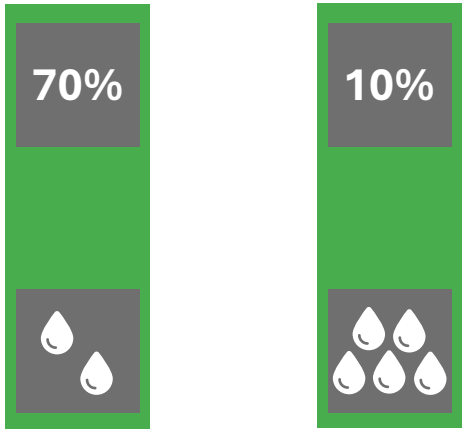


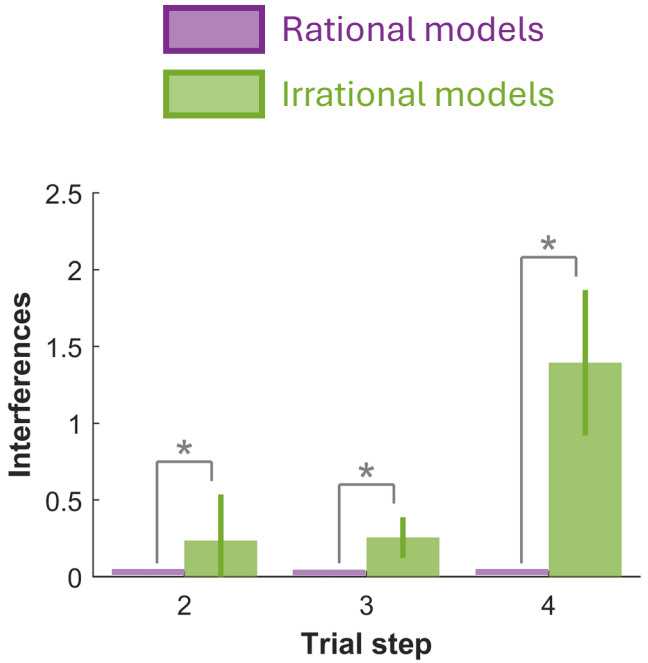


What mechanisms underlie irrational decisions?

Interferences:

How much the **output** of a model **varies** with the **order** of information acquisition.



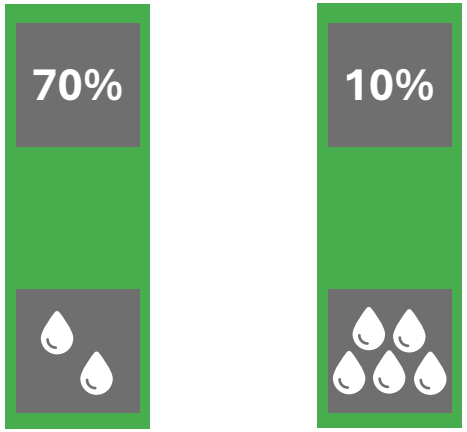


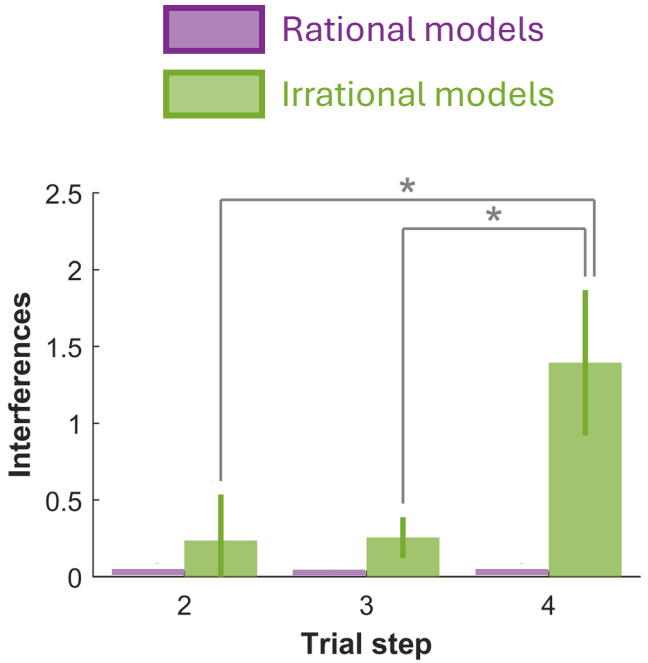
Irrational models develop **spill-over effects** between independant pieces of information.

What mechanisms underlie irrational decisions?

Interferences:

How much the **output** of a model **varies** with the **order** of information acquisition.





Irrational models develop **spill-over effects** between independant pieces of information.

These interferences **accumulate** over time.

What are the biological constraints on OFC value computations?

What are the biological constraints on OFC value computations?

Excitatory / inhibitory balance:

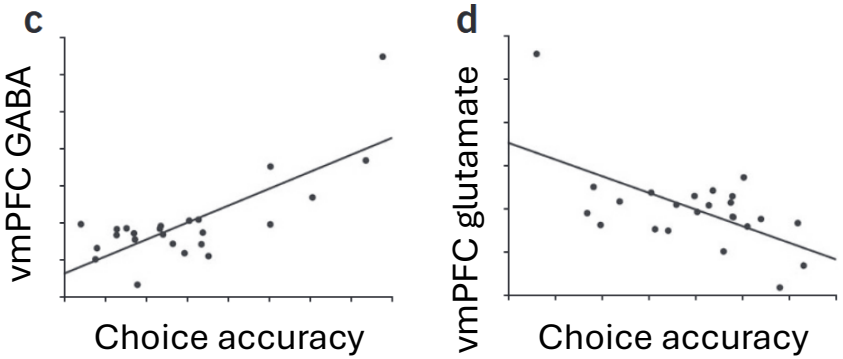
Ratio of positive versus negative connections between neurons.

What are the biological constraints on OFC value computations?

Excitatory / inhibitory balance:

Ratio of positive versus negative connections between neurons.

Jocham et al., 2012. Adapted from Figure 1.

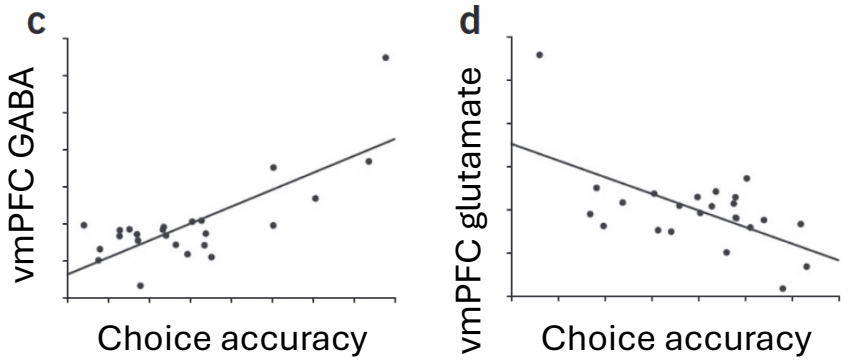


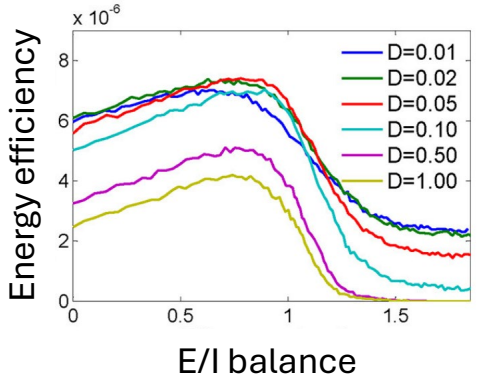
What are the biological constraints on OFC value computations?

Excitatory / inhibitory balance:

Ratio of positive versus negative connections between neurons.

Jocham et al., 2012. Adapted from Figure 1.





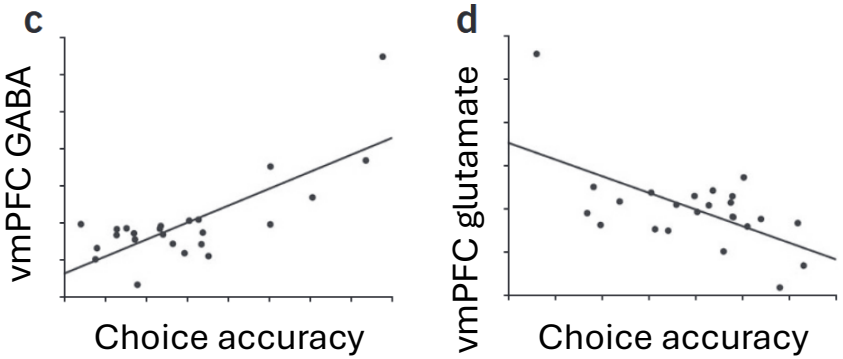
Yu et al., 2018. Adapted from Figure 5.

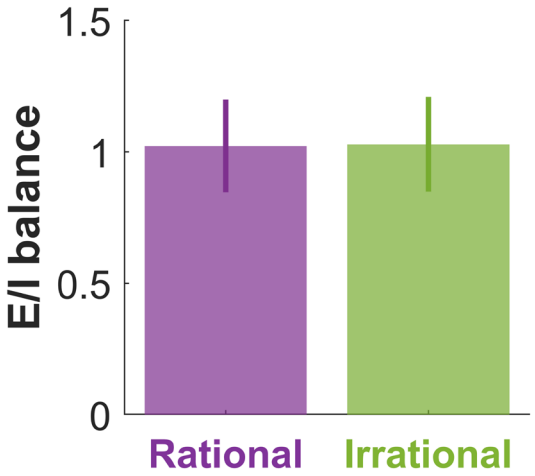
What are the biological constraints on OFC value computations?

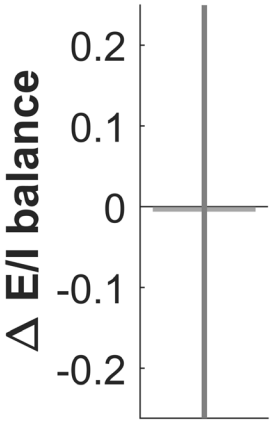
Excitatory / inhibitory balance:

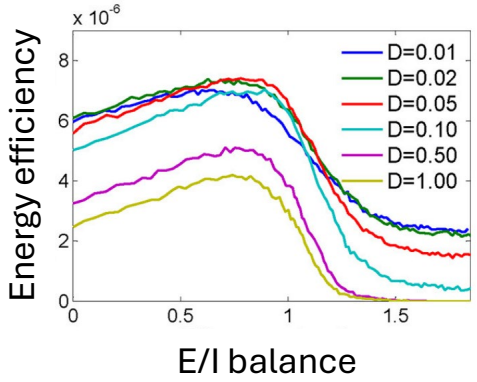
Ratio of positive versus negative connections between neurons.

Jocham et al., 2012. Adapted from Figure 1.









Yu et al., 2018.
Adapted from Figure 5.

What are the biological constraints on OFC value computations?

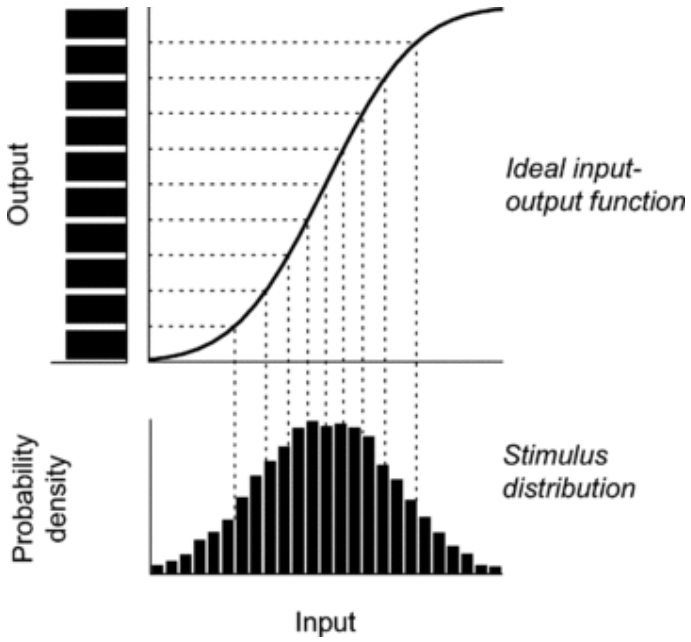
Information encoding

What are the biological constraints on OFC value computations?

Information encoding

Code efficiency:

How much individual units adapt their response to the statistics of their inputs.



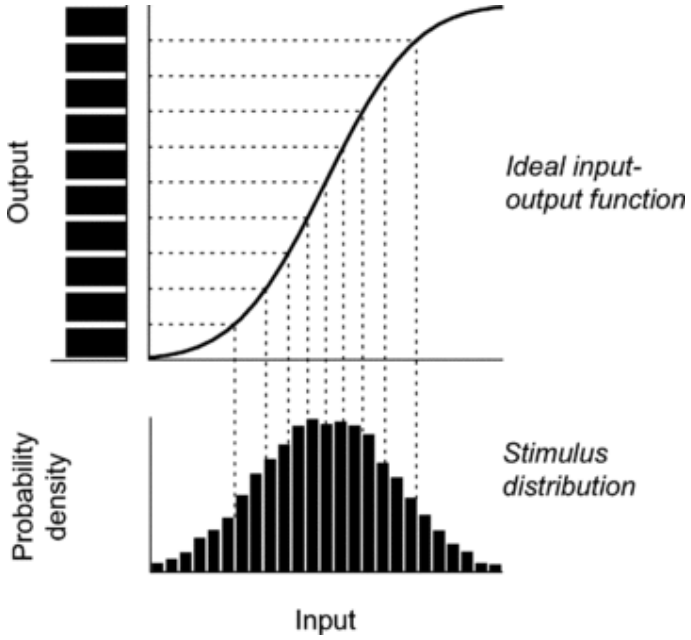
Louie & Glimcher, 2012. Figure 5.

What are the biological constraints on OFC value computations?

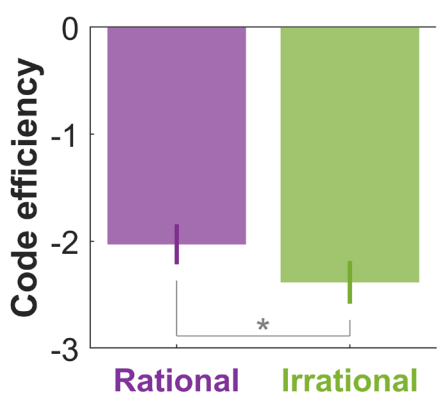
Information encoding

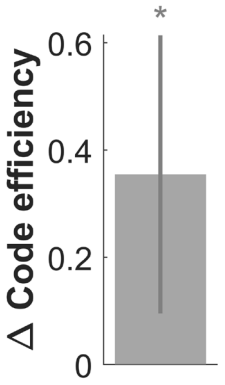
Code efficiency:

How much individual units adapt their response to the statistics of their inputs.



Louie & Glimcher, 2012. Figure 5.





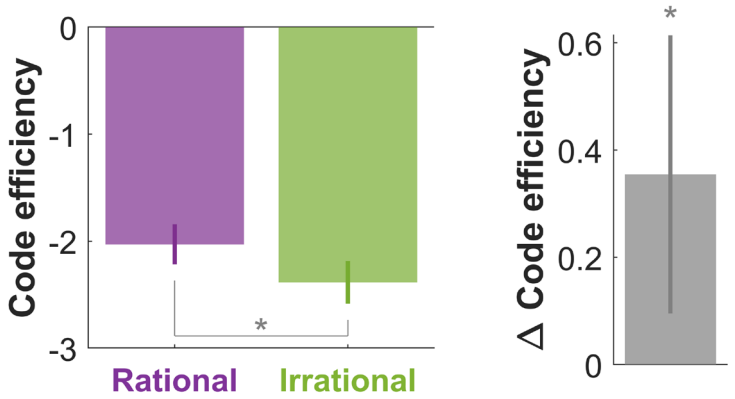
Units in irrational models are less adapted to their range of inputs.

What are the biological constraints on OFC value computations?

Information encoding

Code efficiency:

How much individual units adapt their response to the statistics of their inputs.



Units in irrational models are less adapted to their range of inputs.

Code sparsity:

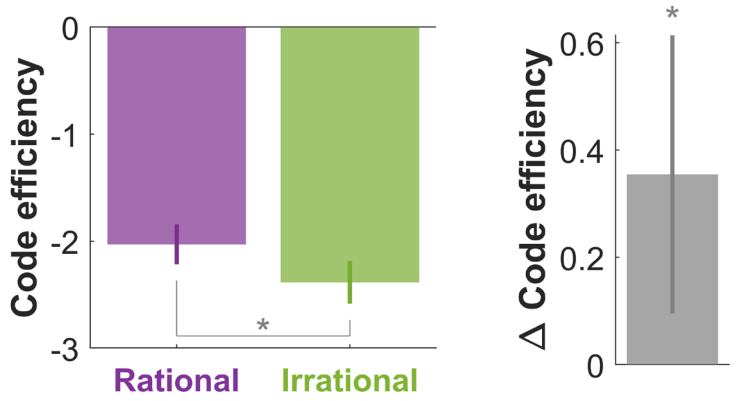
How many neurons are active at any one time.

What are the biological constraints on OFC value computations?

Information encoding

Code efficiency:

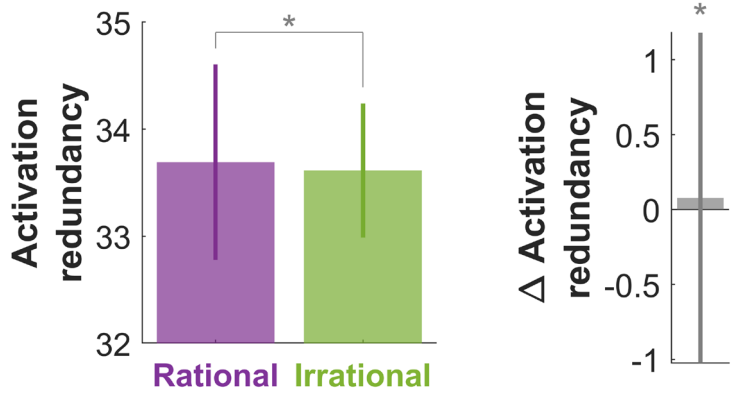
How much individual units adapt their response to the statistics of their inputs.



Units in irrational models are less adapted to their range of inputs.

Code sparsity:

How many neurons are active at any one time.



Irrational models use more sparse representations.

What are the biological constraints on OFC value computations?

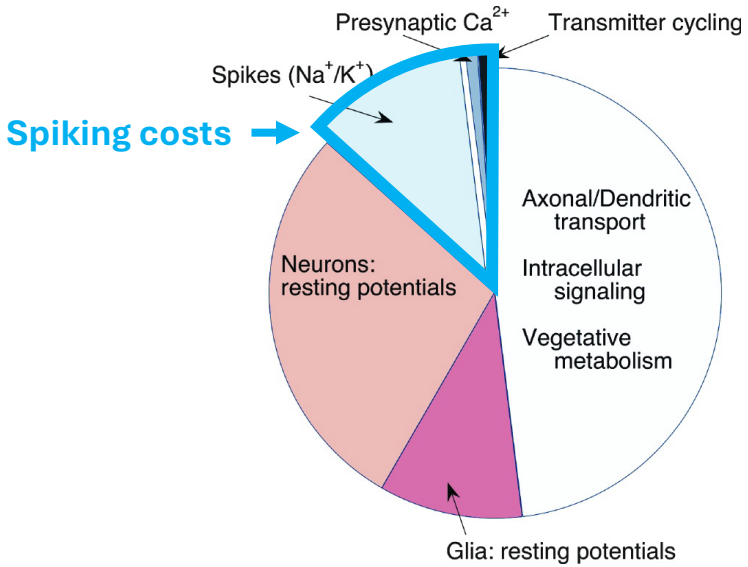
Energy budget

What are the biological constraints on OFC value computations?

Energy budget

Average firing rate (electrophysiological cost)**:**

How much neurons fire on average.



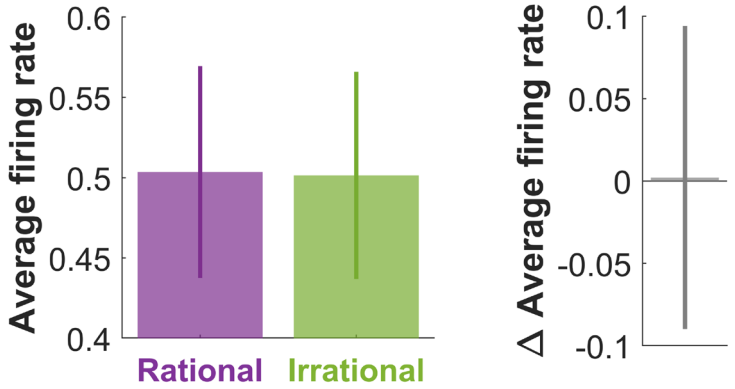
Lennie 2003.
Adapted from Figure 1B.

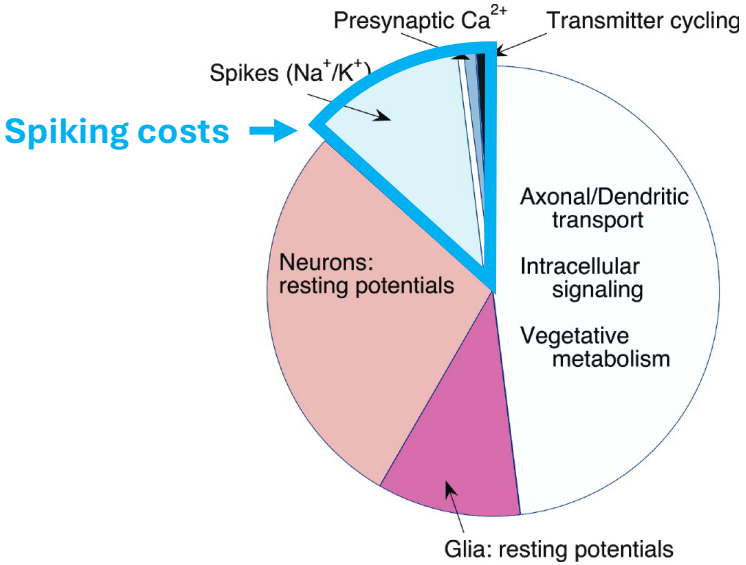
What are the biological constraints on OFC value computations?

Energy budget

Average firing rate (electrophysiological cost)**:**

How much neurons fire on average.





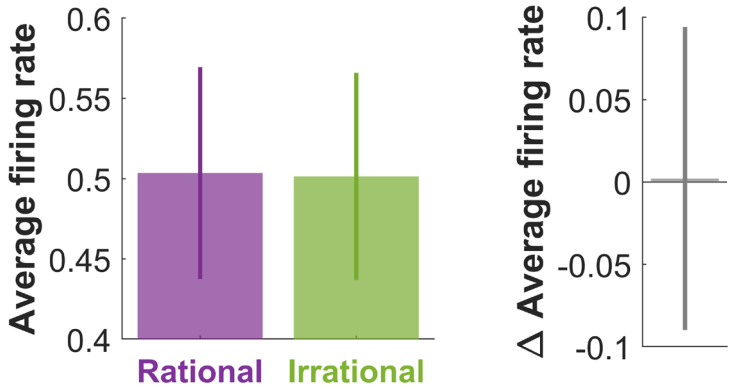
Lennie 2003.
Adapted from Figure 1B.

What are the biological constraints on OFC value computations?

Energy budget

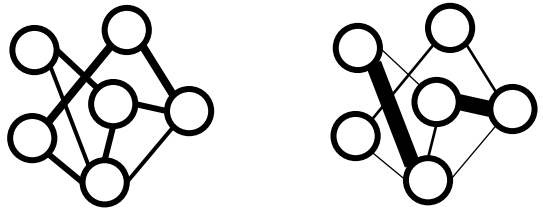
Average firing rate (electrophysiological cost)**:**

How much neurons fire on average.



Connections sparsity (structural cost)**:**

How inequal is the distribution of connections between neurons.



Low sparsity

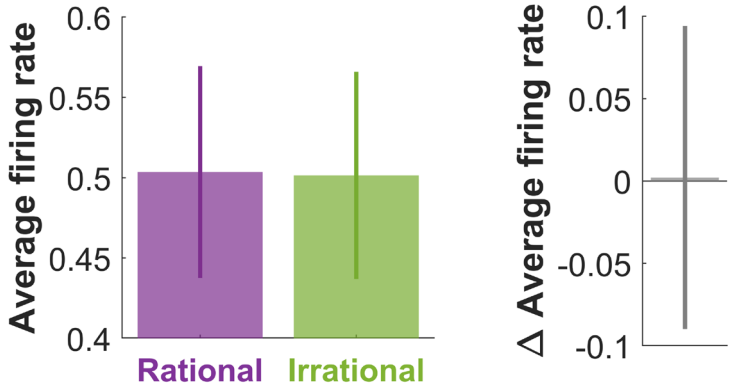
High sparsity

What are the biological constraints on OFC value computations?

Energy budget

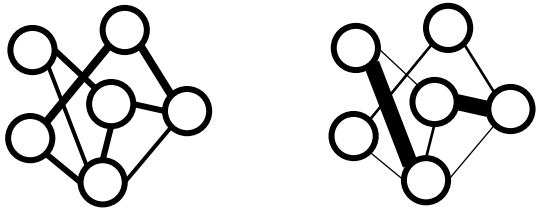
Average firing rate (electrophysiological cost)**:**

How much neurons fire on average.



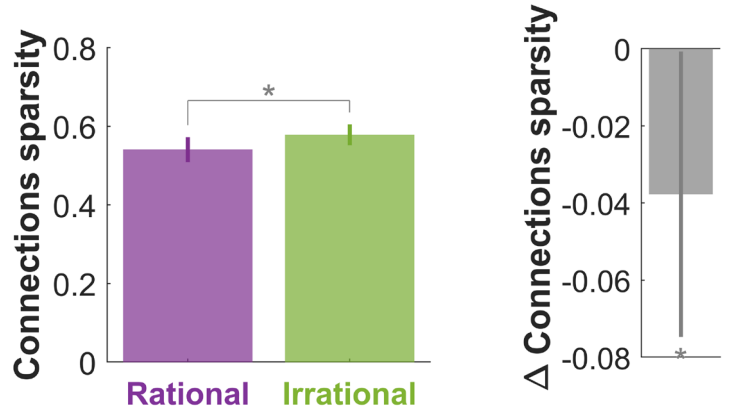
Connections sparsity (structural cost)**:**

How inequal is the distribution of connections between neurons.



Low sparsity

High sparsity



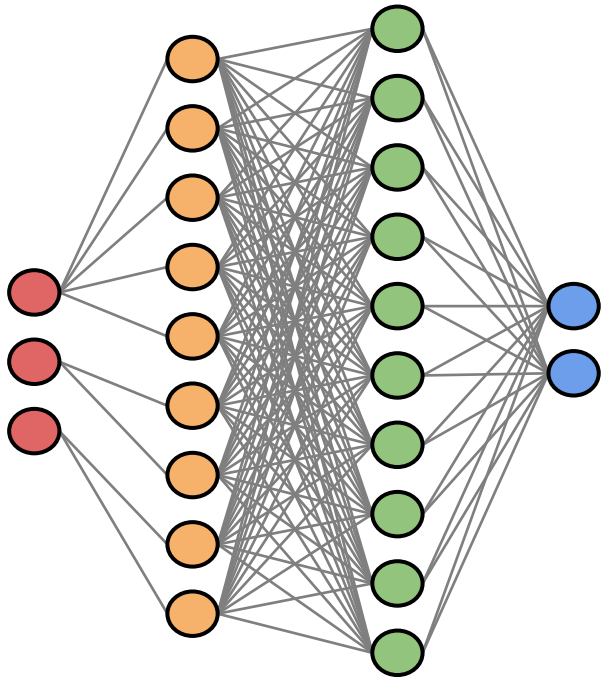
Irrational models rely on less connections.

What are the biological constraints on OFC value computations?

Robustness to lesions

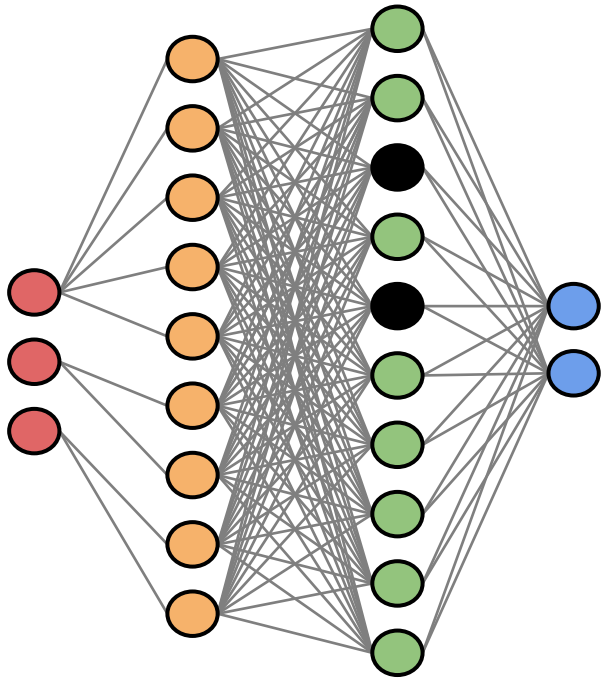
What are the biological constraints on OFC value computations?

Robustness to lesions



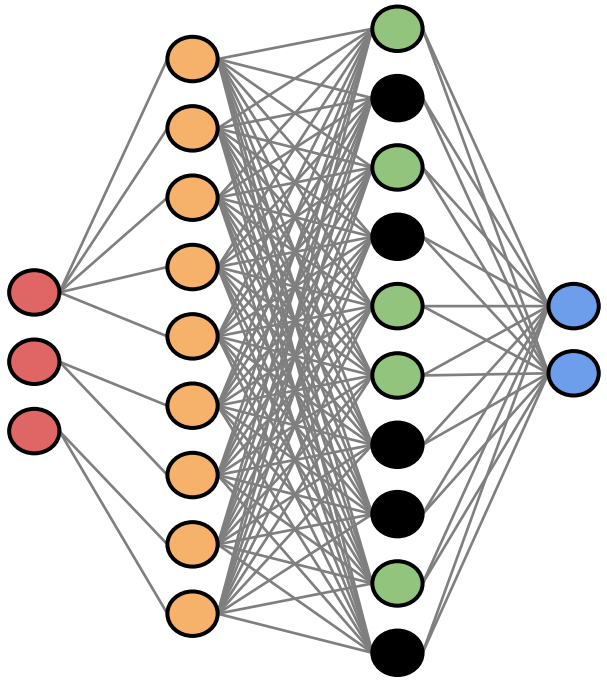
What are the biological constraints on OFC value computations?

Robustness to lesions



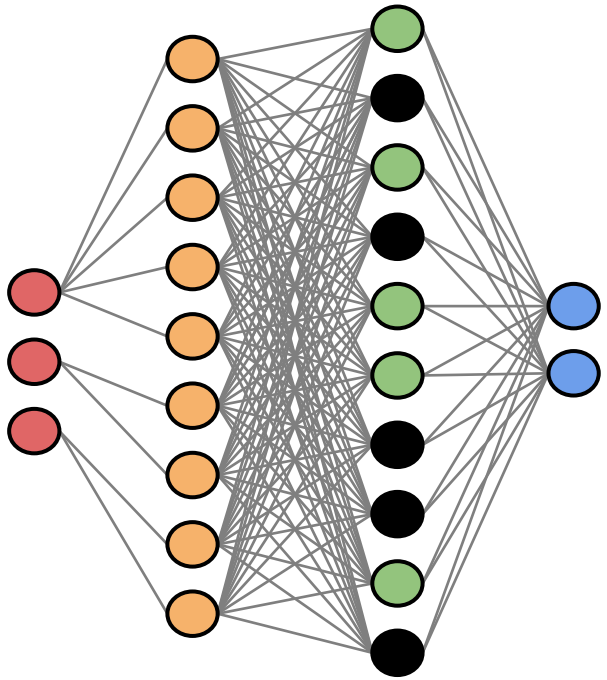
What are the biological constraints on OFC value computations?

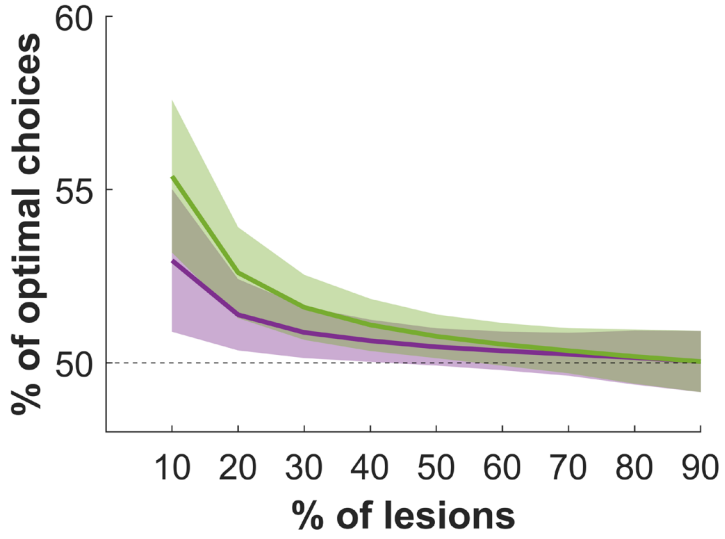
Robustness to lesions



What are the biological constraints on OFC value computations?

Robustness to lesions



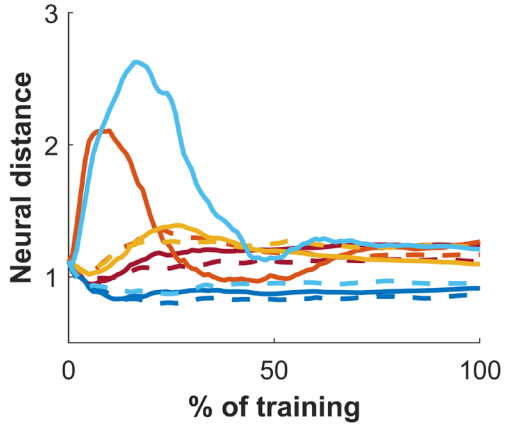


Irrational models are more robust to lesions.

Conclusion

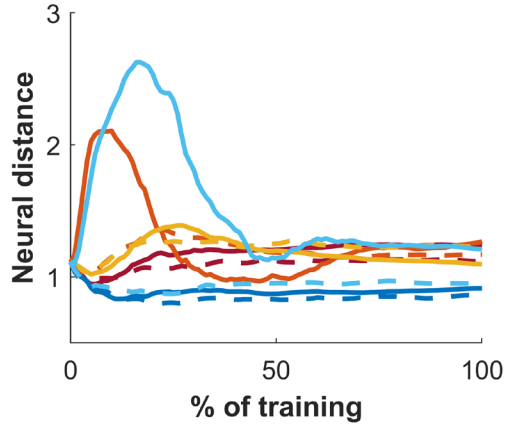
Conclusion

Three computational scenarii, including both **value synthesis** and **value comparison**, but using only a **non-spatial** encoding of offers, generate OFC-like neural activity.

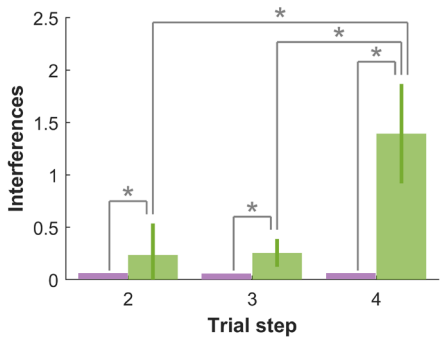


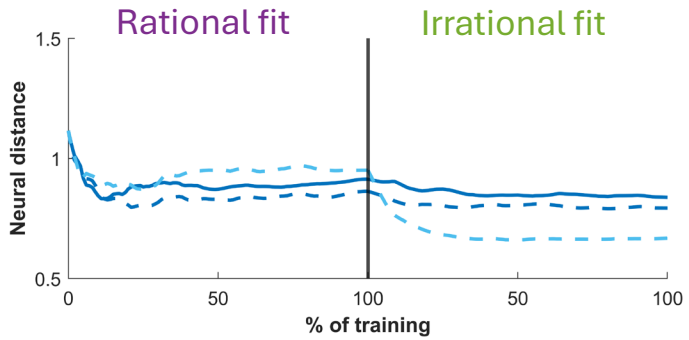
Conclusion

Three computational scenarii, including both **value synthesis** and **value comparison**, but using only a **non-spatial** encoding of offers, generate OFC-like neural activity.



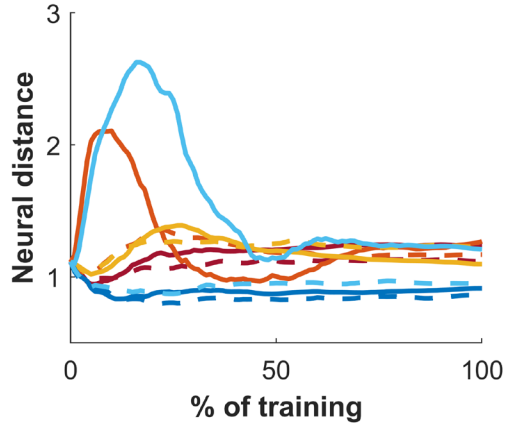
Interferences between and across options, accumulating through time, generate **realistic** irrational decisions and **realistic** neural activity patterns.



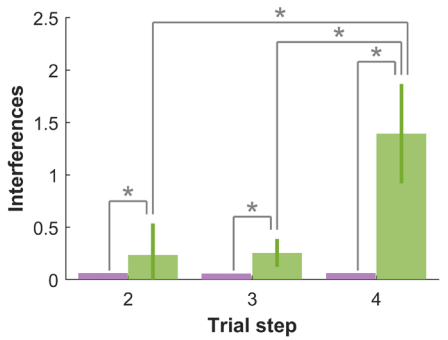


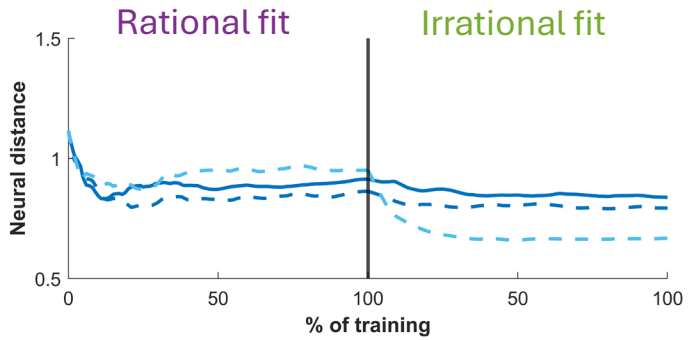
Conclusion

Three computational scenarii, including both **value synthesis** and **value comparison**, but using only a **non-spatial** encoding of offers, generate OFC-like neural activity.

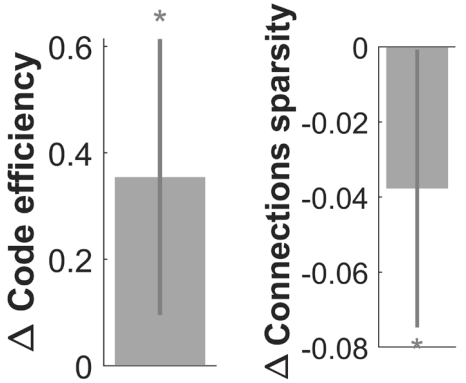


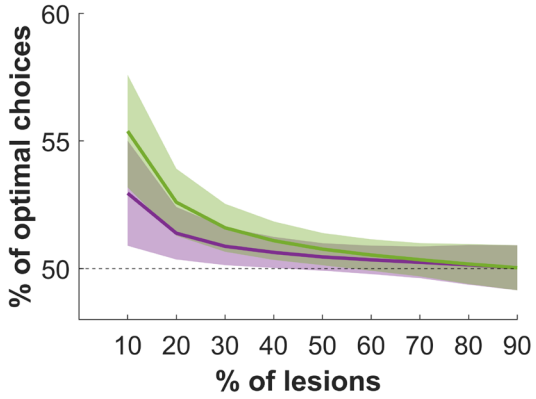
Interferences between and across options, accumulating through time, generate **realistic** irrational decisions and **realistic** neural activity patterns.





Biological constraints on the architecture of the OFC neural code might induce **interferences** causing **irrational** behavior.





Conclusion

Three computational scenarii, including both **value synthesis** and **value comparison**, but using only a **non-spatial** encoding of offers, generate OFC-like neural activity.

Interferences between and across options, accumulating through time, generate **realistic** irrational decisions and **realistic** neural activity patterns.

Biological constraints on the architecture of the OFC neural code might induce **interferences** causing **irrational** behavior.

What's next?

Conclusion

Three computational scenarii, including both **value synthesis** and **value comparison**, but using only a **non-spatial** encoding of offers, generate OFC-like neural activity.

Interferences between and across options, accumulating through time, generate **realistic** irrational decisions and **realistic** neural activity patterns.

Biological constraints on the architecture of the OFC neural code might induce **interferences** causing **irrational** behavior.

What's next?

Do specific biological constraints necessarily generate specific interferences?

Conclusion

Three computational scenarii, including both **value synthesis** and **value comparison**, but using only a **non-spatial** encoding of offers, generate OFC-like neural activity.

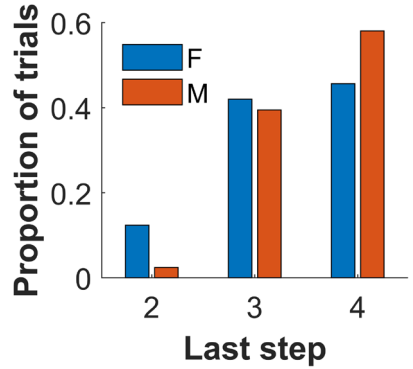
Interferences between and across options, accumulating through time, generate **realistic** irrational decisions and **realistic** neural activity patterns.

Biological constraints on the architecture of the OFC neural code might induce **interferences** causing **irrational** behavior.

What's next?

Do specific biological constraints necessarily generate specific interferences?

Why do monkeys stop early?



Conclusion

Three computational scenarii, including both **value synthesis** and **value comparison**, but using only a **non-spatial** encoding of offers, generate OFC-like neural activity.

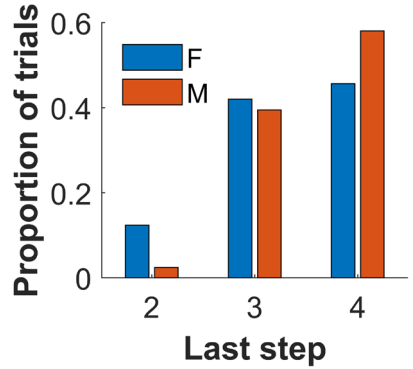
Interferences between and across options, accumulating through time, generate **realistic** irrational decisions and **realistic** neural activity patterns.

Biological constraints on the architecture of the OFC neural code might induce **interferences** causing **irrational** behavior.

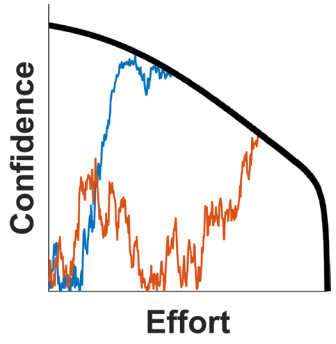
What's next?

Do specific biological constraints necessarily generate specific interferences?

Why do monkeys stop early?



Bénon, ..., Daunizeau, 2024.
The online metacognitive control of decisions.



Conclusion

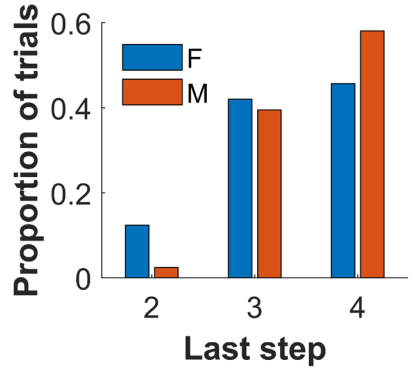
Three computational scenarii, including both **value synthesis** and **value comparison**, but using only a **non-spatial** encoding of offers, generate OFC-like neural activity.

Interferences be
accumulating thr
irrational decision
patterns.

Biological constraints on the architecture of the OFC neural code might induce **interferences** causing **irrational** behavior.

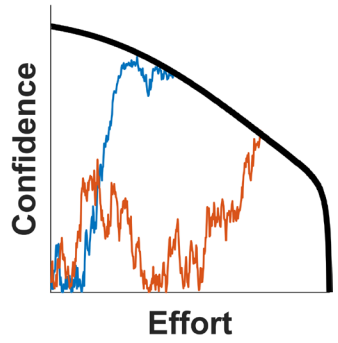
What's next?

Do specific biological constraints necessarily generate specific interferences?



Thanks for your attention!

Bénon, ..., Daunizeau, 2024.
The online metacognitive control of decisions.



Conclusion

Three computational scenarii, including both **value synthesis** and **value comparison**, but using only a **non-spatial** encoding of offers, generate OFC-like neural activity.

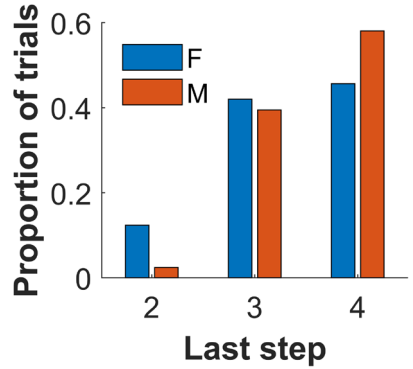
Interferences between and across options, accumulating through time, generate **realistic** irrational decisions and **realistic** neural activity patterns.

Biological constraints on the architecture of the OFC neural code might induce **interferences** causing **irrational** behavior.

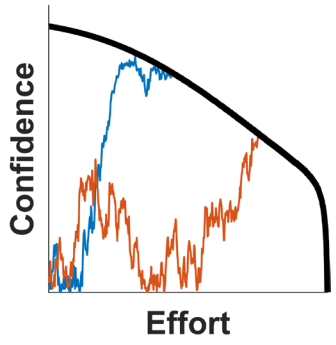
What's next?

Do specific biological constraints necessarily generate specific interferences?

Why do monkeys stop early?



Bénon, ..., Daunizeau, 2024.
The online metacognitive control of decisions.



Supplementary

OFC-like neural features

- Offer value / Chosen value / Chosen offer cells
- Confidence encoding
- Autocorrelation
- CCM features

Interferences

- Last attribute integration
- Attended vs. unattended value

Optimality vs. rationality

- Value functions
- Biological constraints

Biological constraints

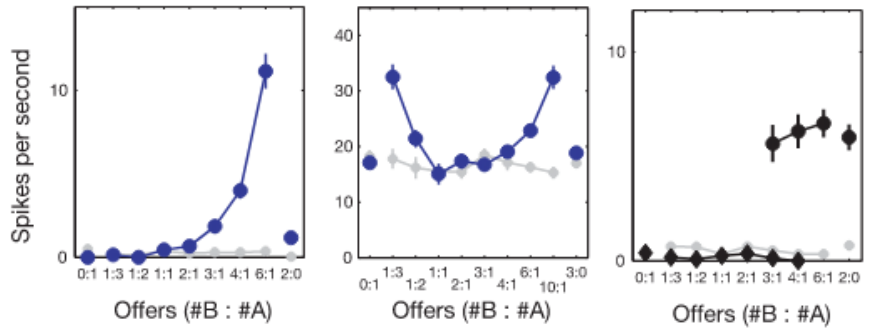
- Robustness (consistency)

Within-model and within-monkey variability

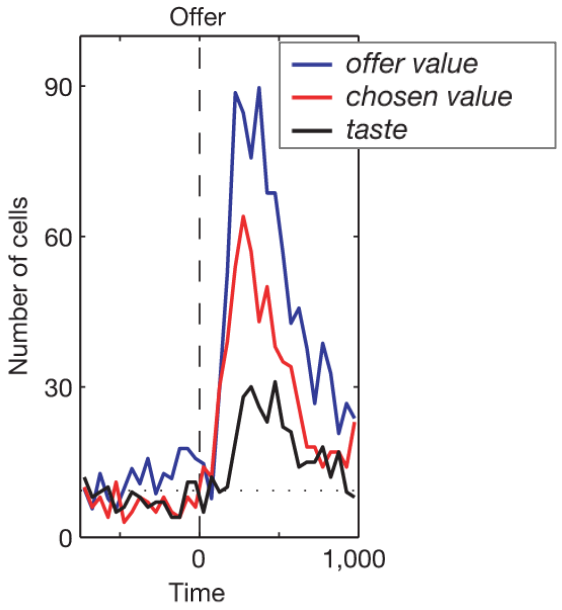
- E/I balance
- Code efficiency
- Code sparsity
- Electrophysiological cost
- Structural cost
- Robustness (optimality)

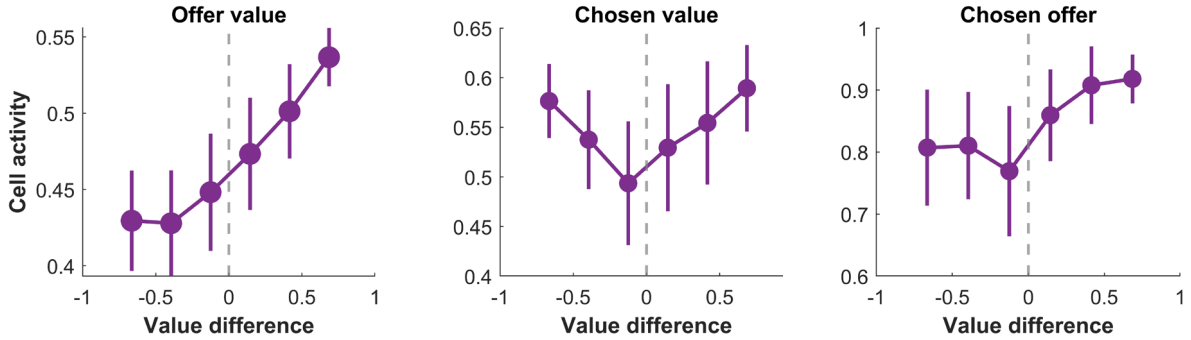
Supplementary summary

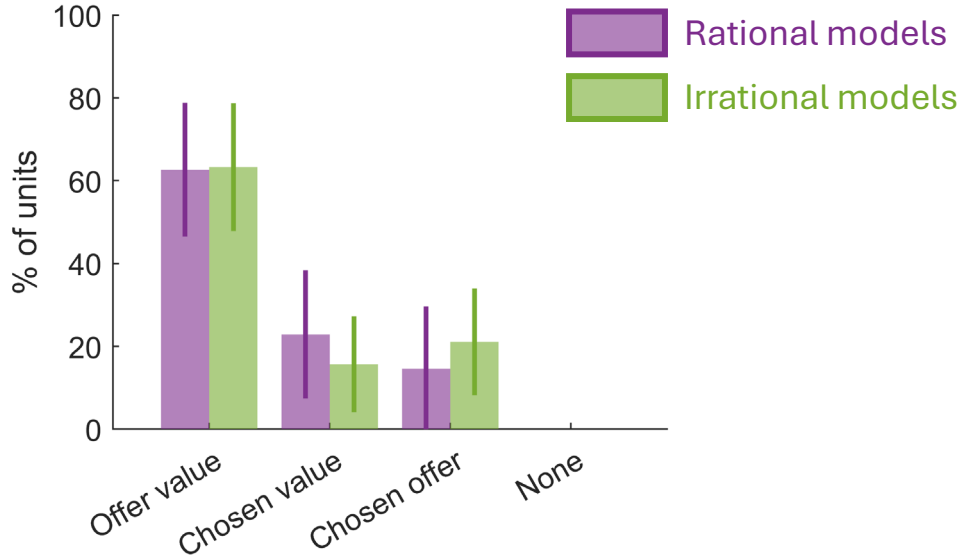
Padoa-Schioppa & Assad, 2006. Figure 3.



Padoa-Schioppa & Assad, 2006. Figure 4.



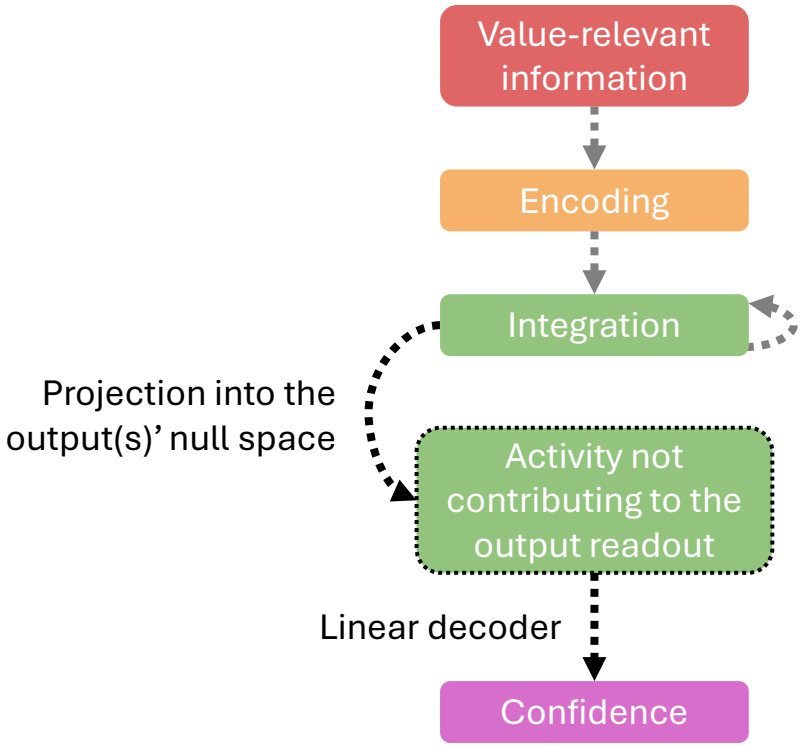


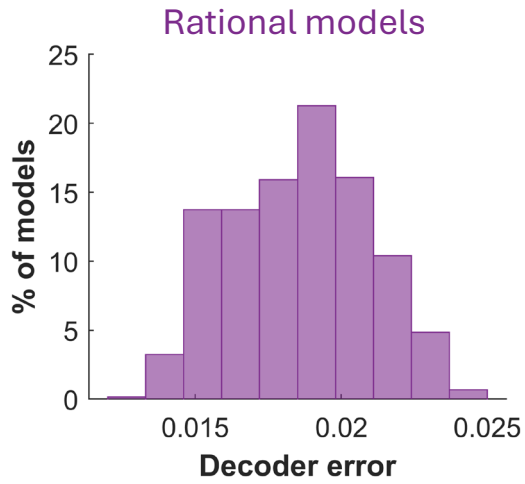


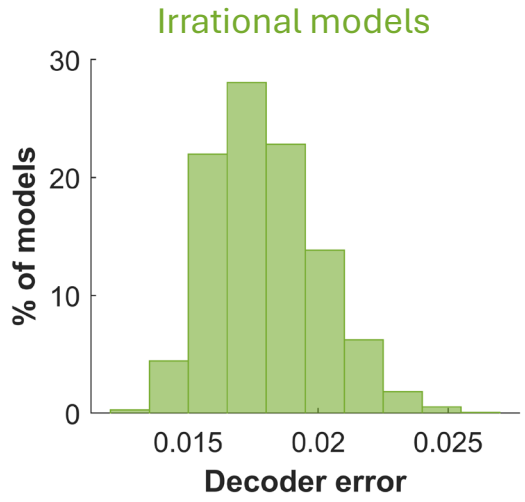
Supplementary summary

Confidence:

Probability that the option about to be chosen (given **partial** information) is the **best** option.



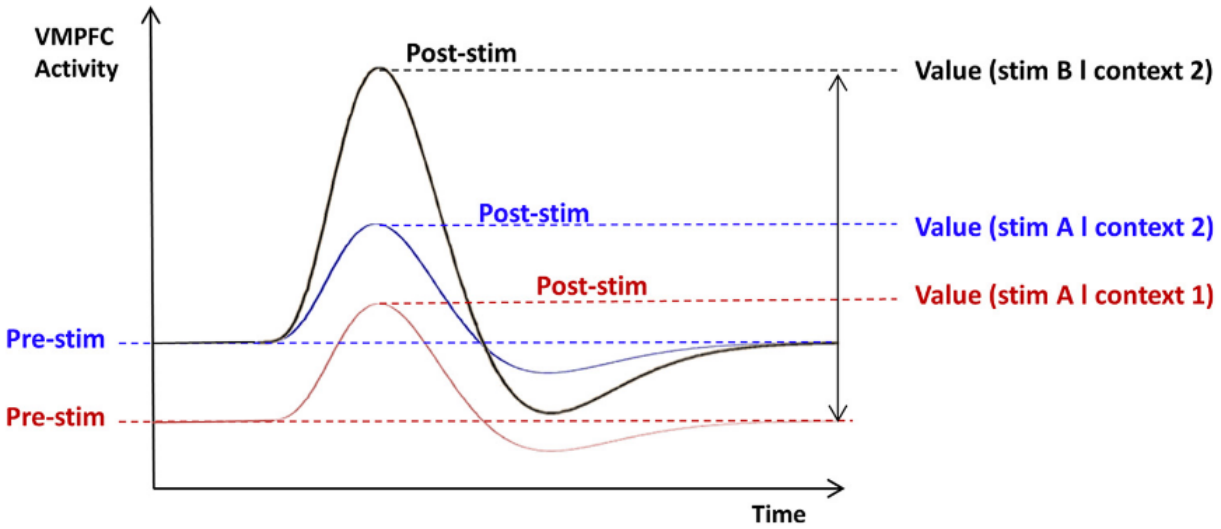




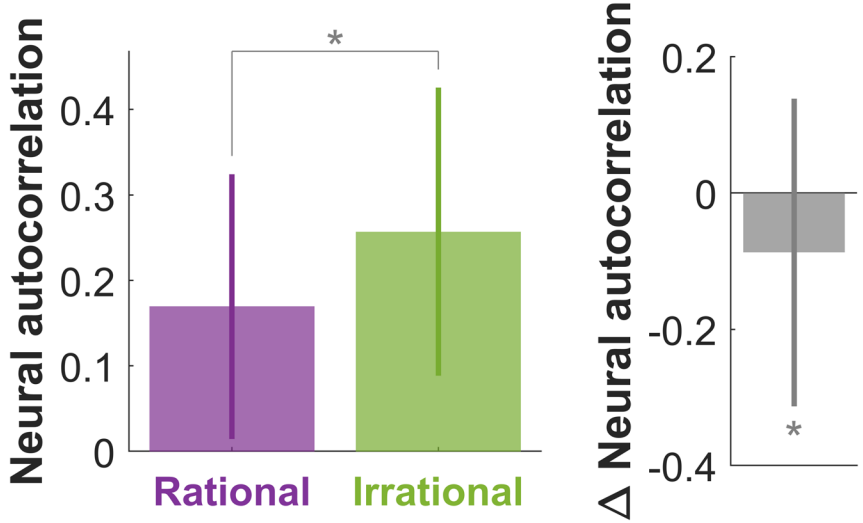
Supplementary summary

Autocorrelation:

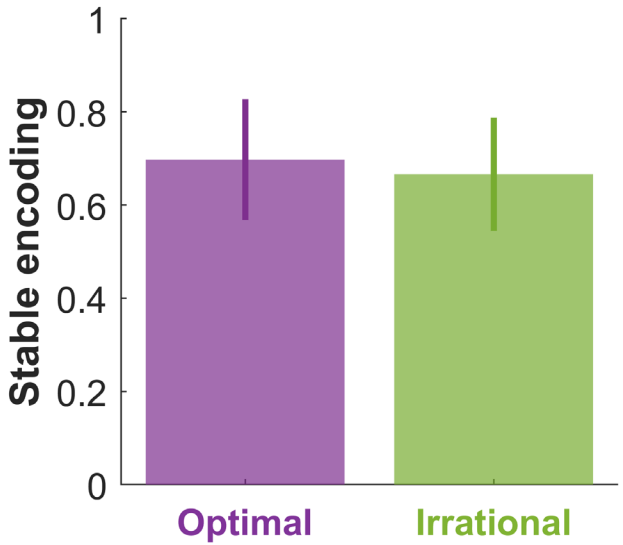
Influence of **pre-stimulus** activity onto **post-stimulus** activity.

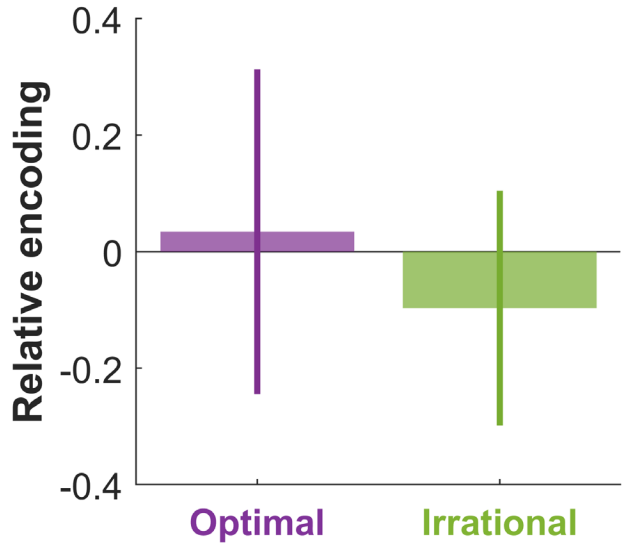


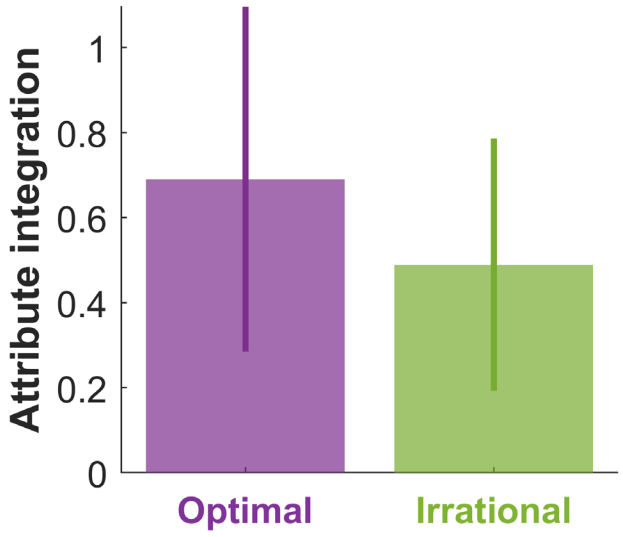
Abitbol et al., 2015. Figure 1.



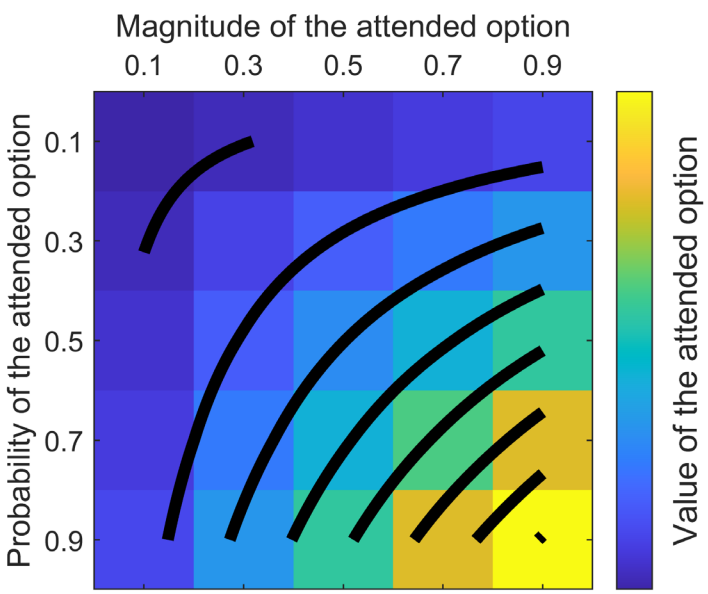
Supplementary summary

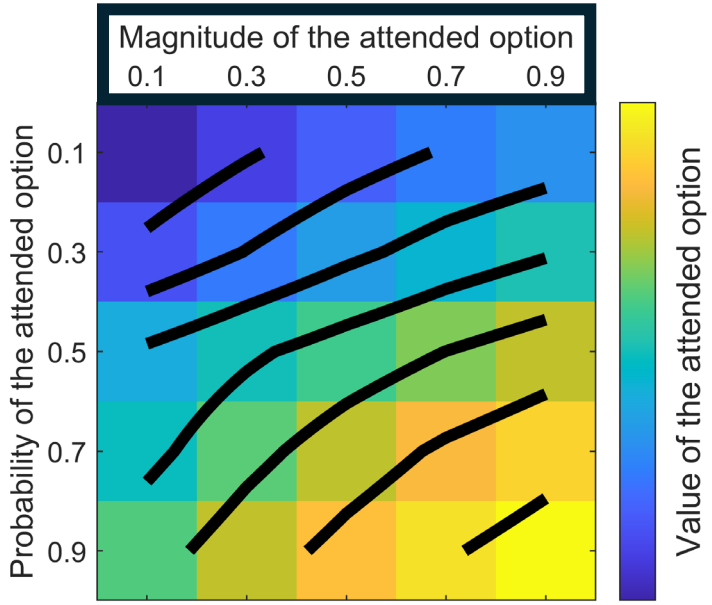




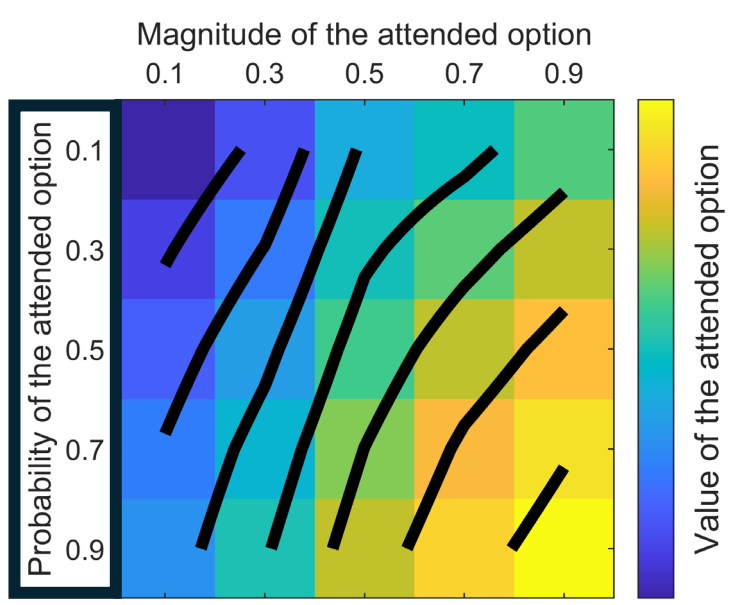


Supplementary summary



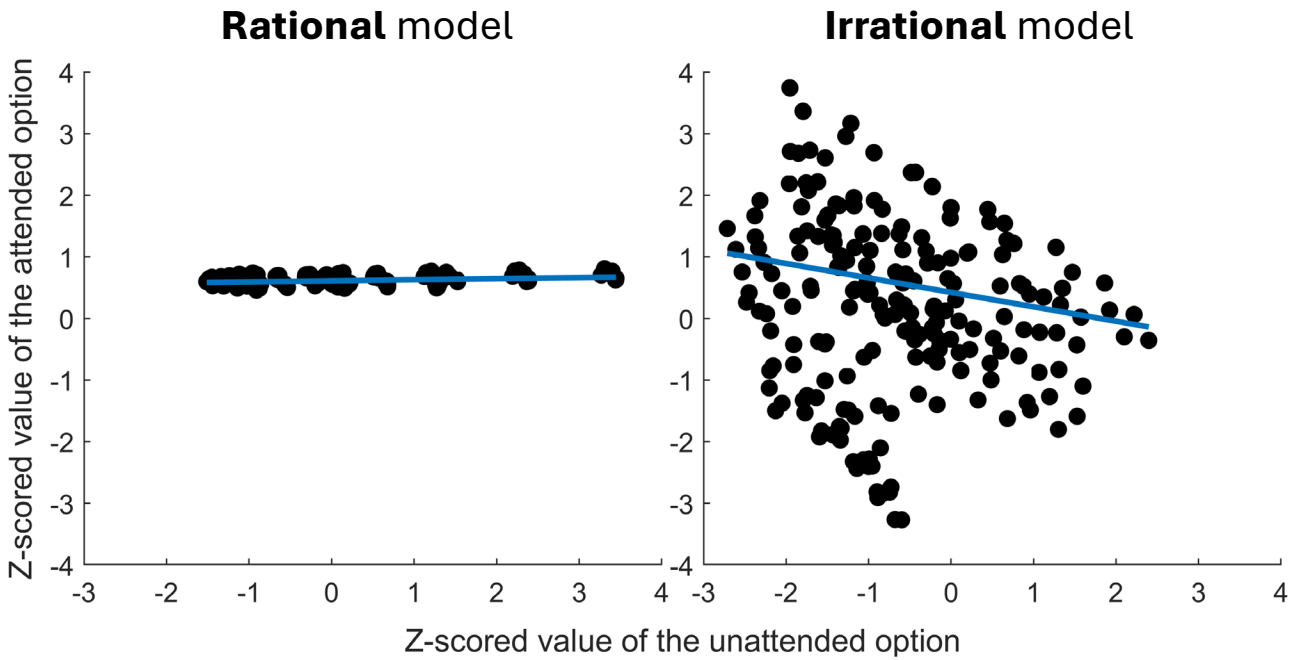


Irrational model, when the **magnitude** has just been attended

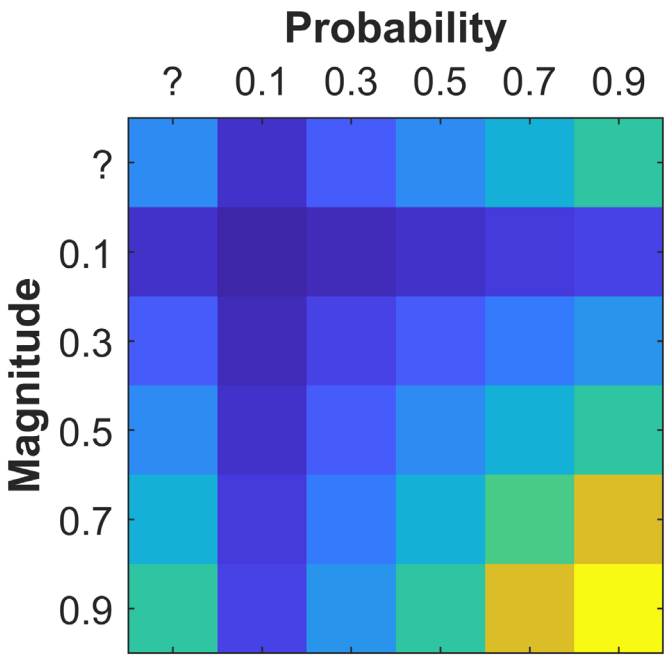


Irrational model, when the **probability** has just been attended

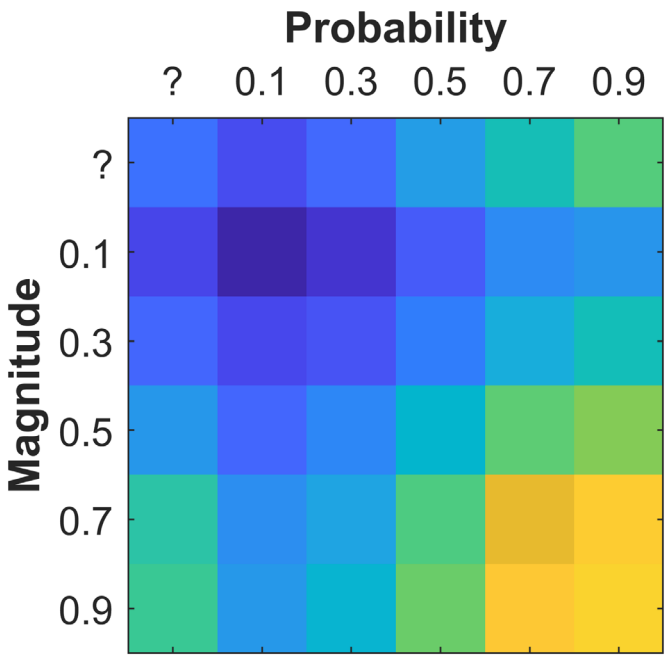
Supplementary summary



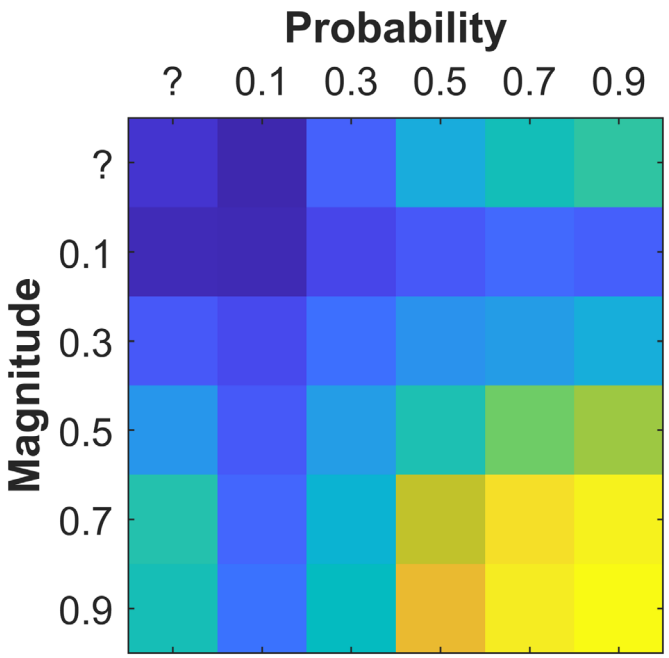
Supplementary summary



Optimal value function

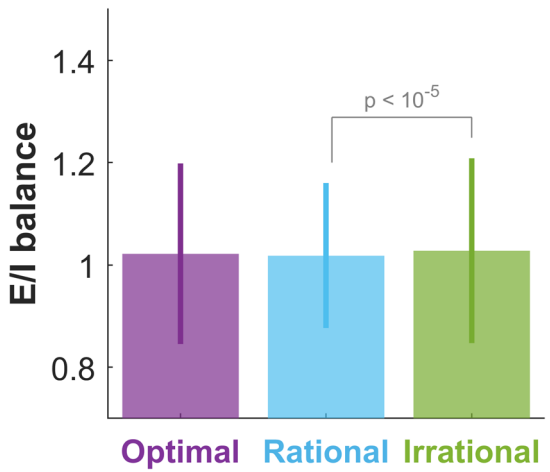


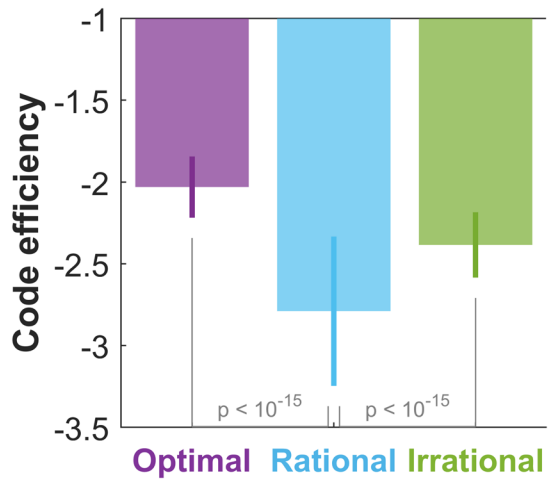
Value function of monkey F

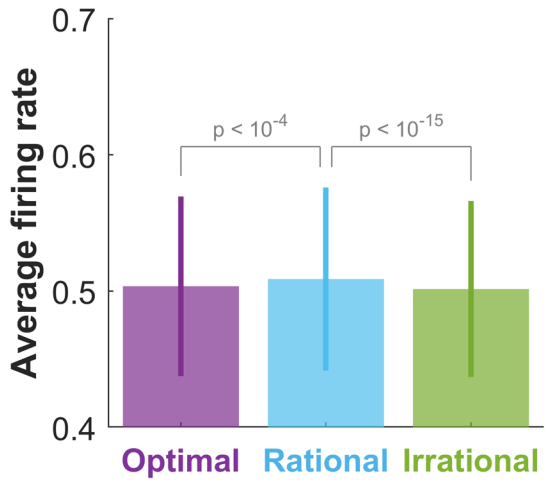


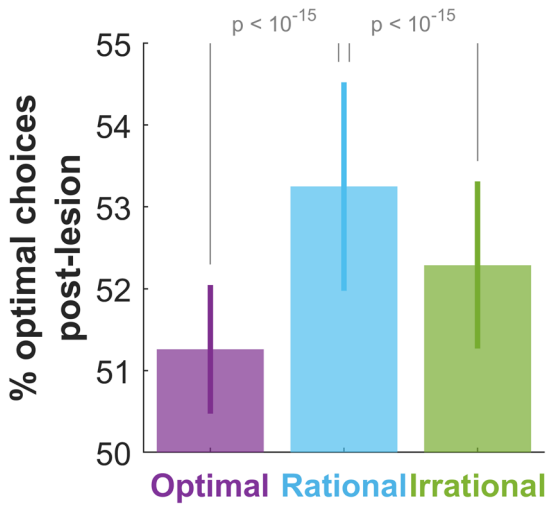
Value function of monkey M

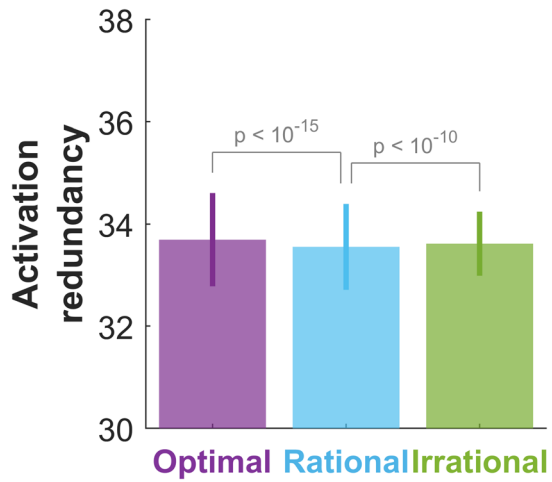
Supplementary summary

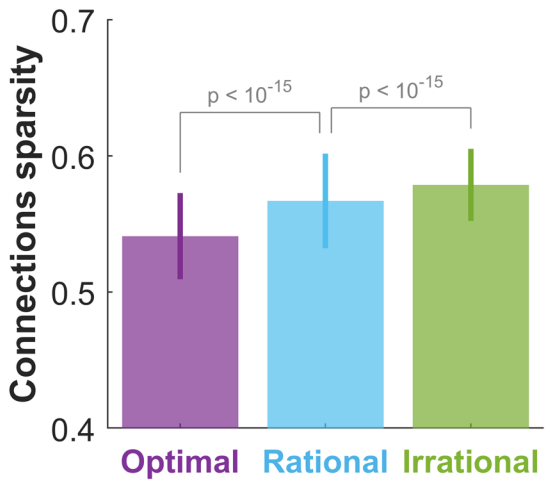




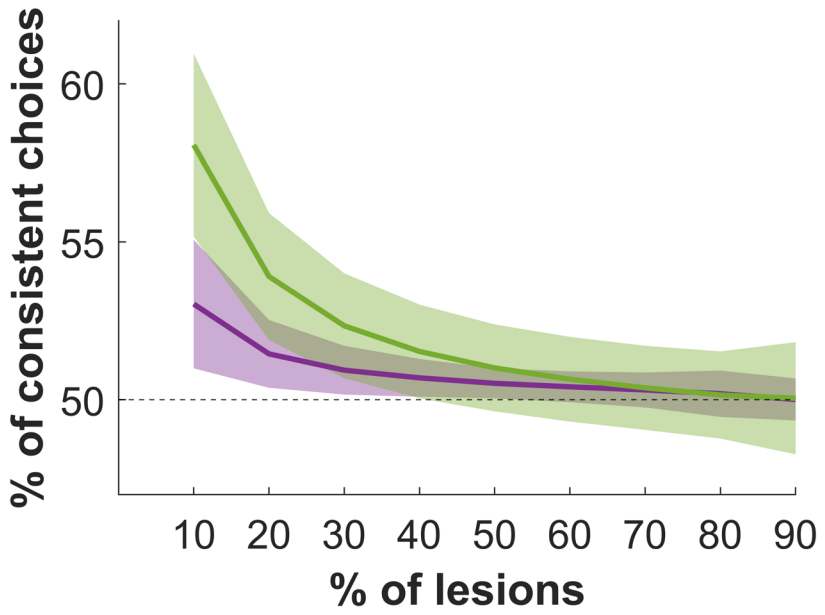






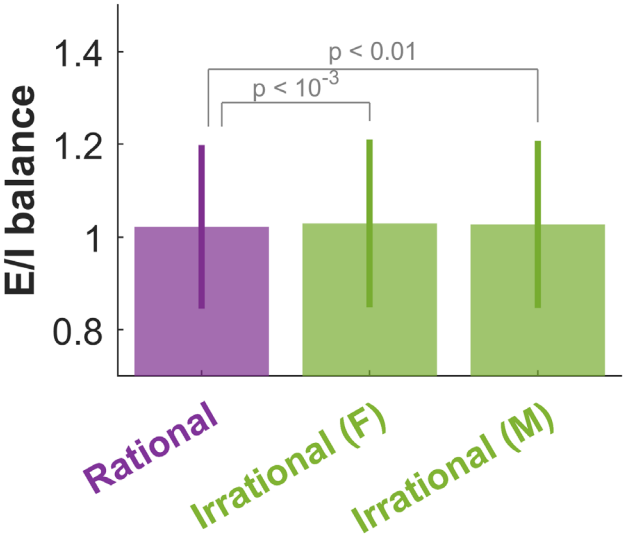


Supplementary summary

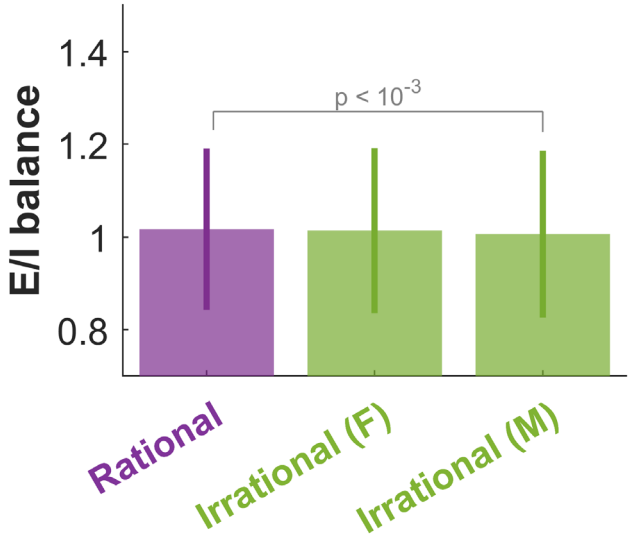


Supplementary summary

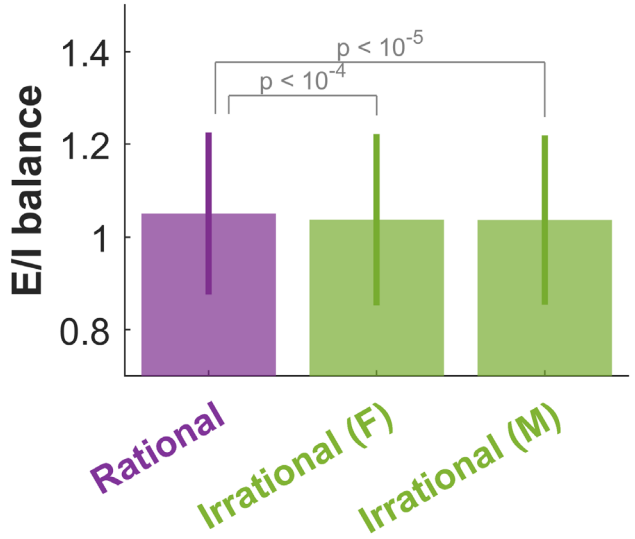
All candidates



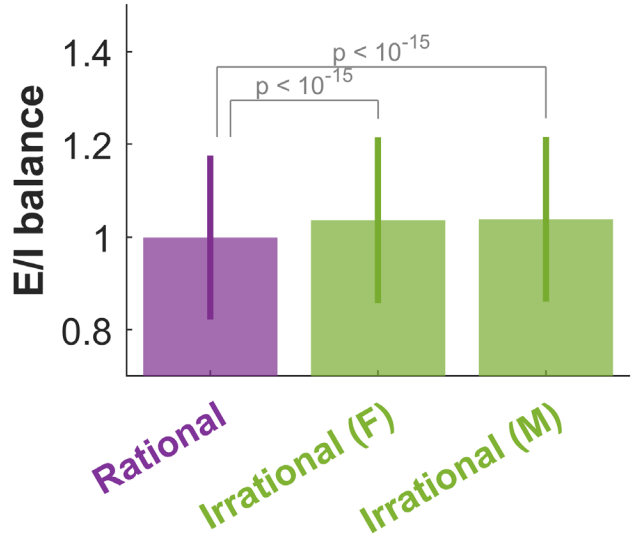
Temporal order
→ **Attentional** focus
(value ***synthesis***)



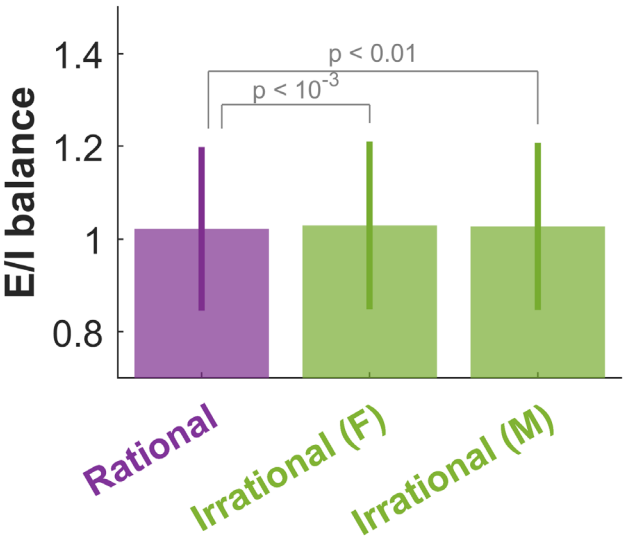
Temporal order
→ **Attentional** focus
(value ***comparison***)

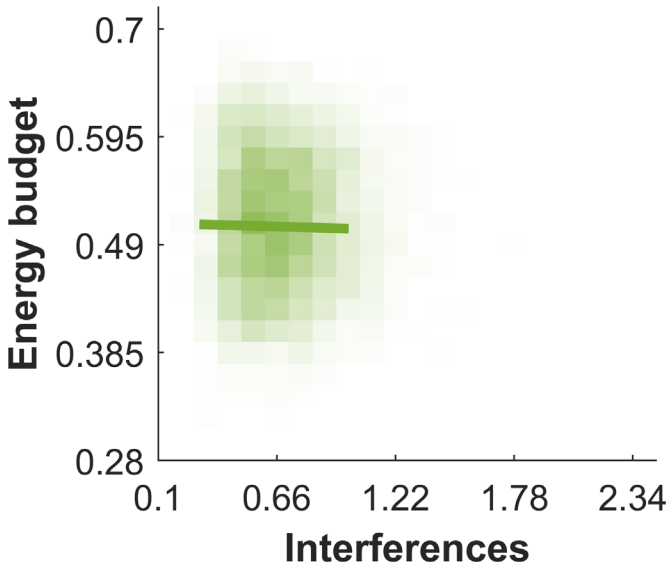


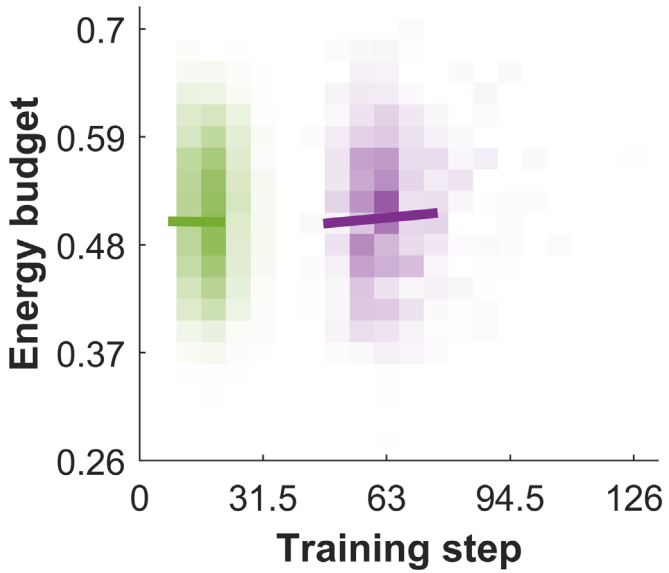
Temporal order
→ **Temporal** order
(value ***synthesis***)



Supplementary summary

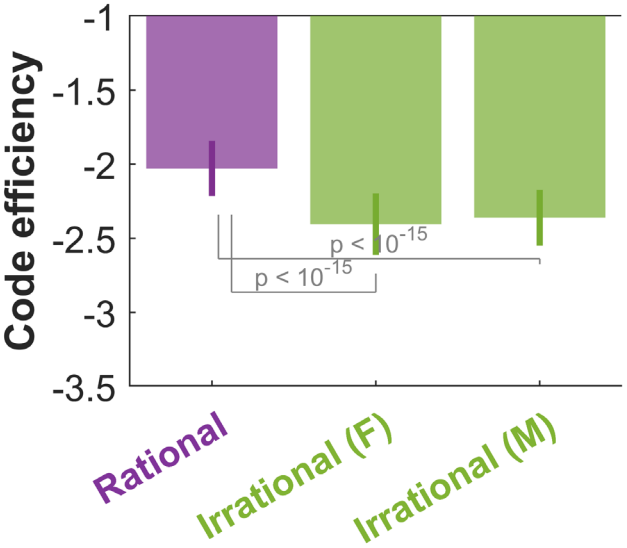




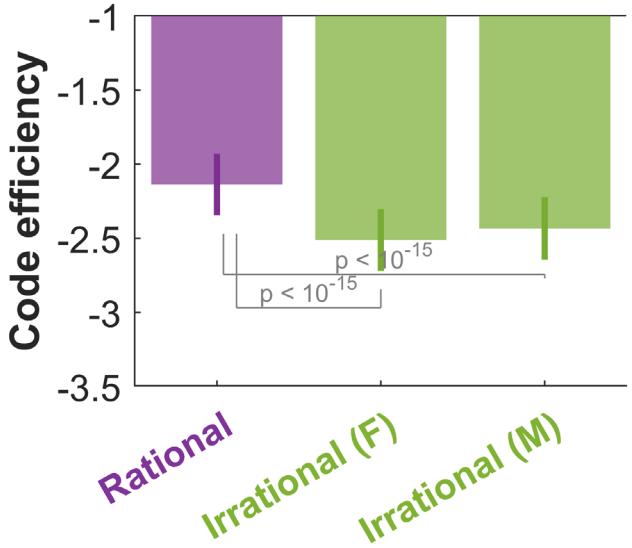


Supplementary summary

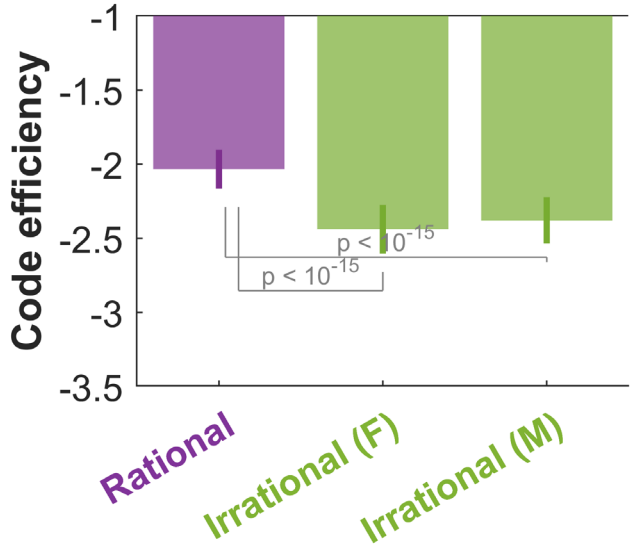
All candidates



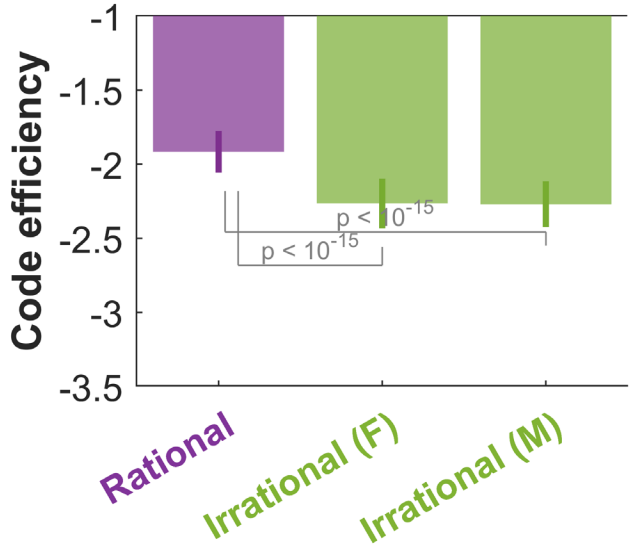
Temporal order
→ **Attentional** focus
(value ***synthesis***)



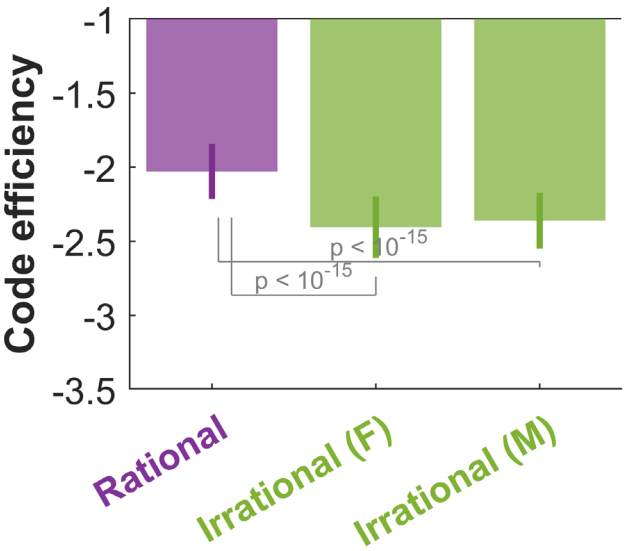
Temporal order
→ **Attentional** focus
(value ***comparison***)

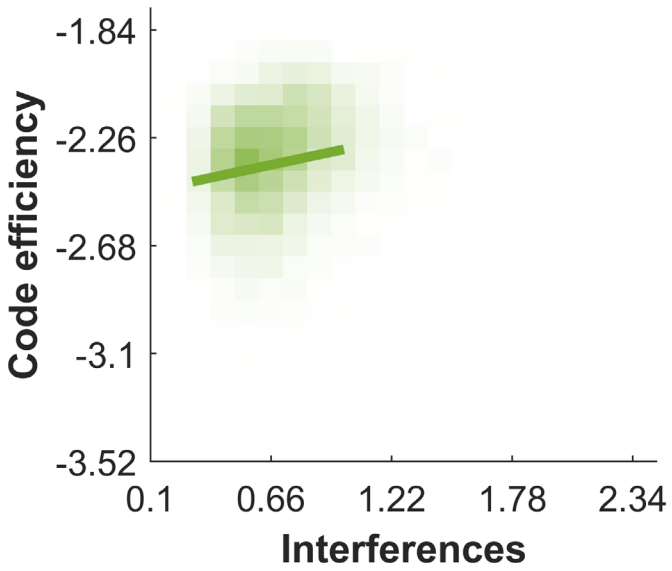


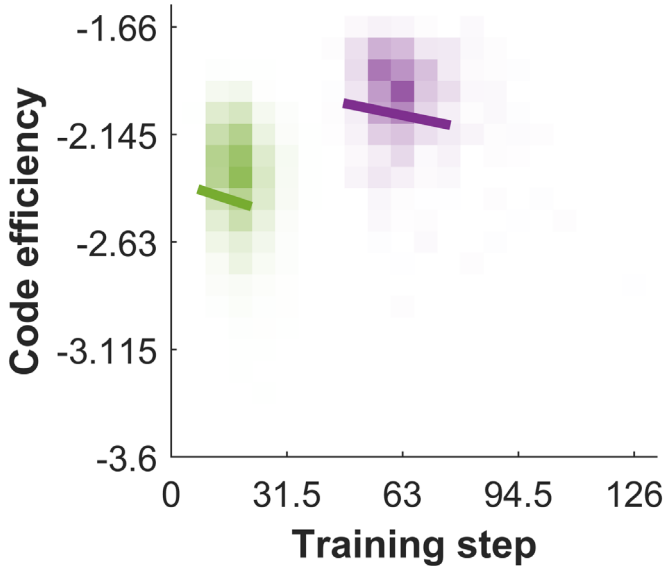
Temporal order
→ **Temporal** order
(value ***synthesis***)



Supplementary summary

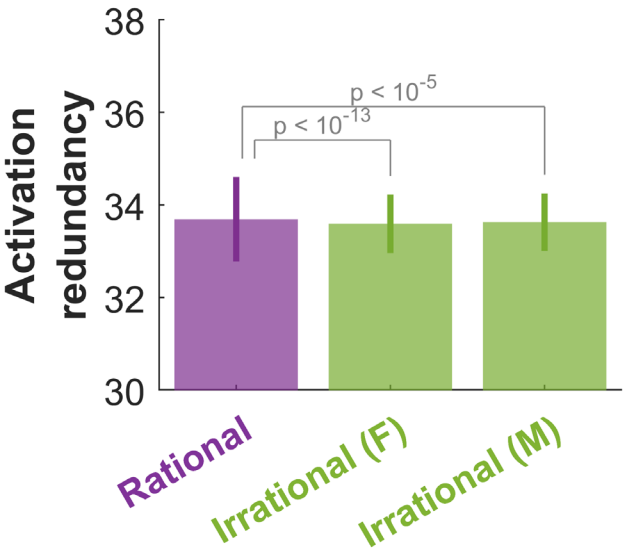




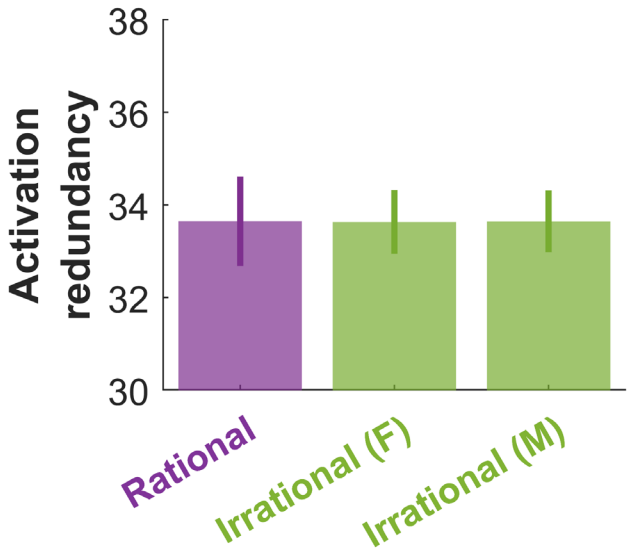


Supplementary summary

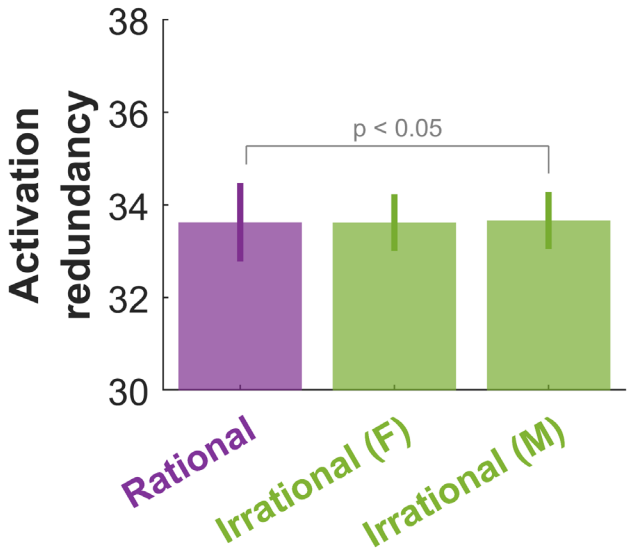
All candidates



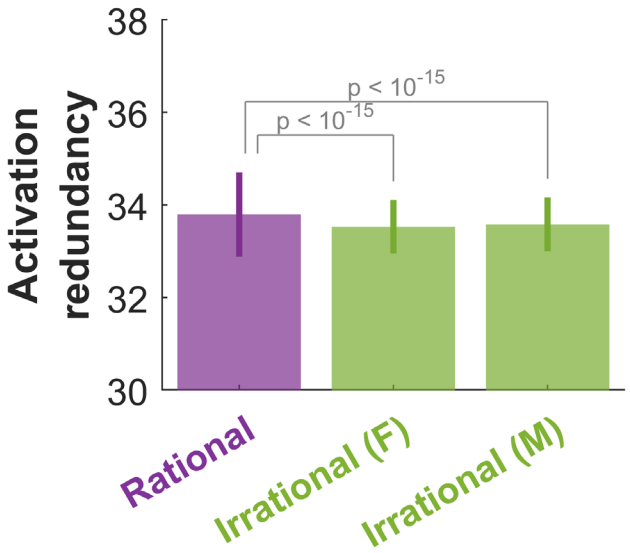
Temporal order
→ **Attentional** focus
(value ***synthesis***)



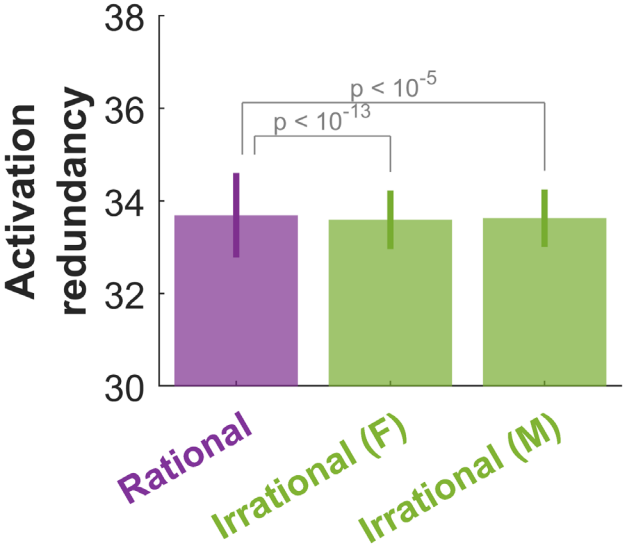
Temporal order
→ **Attentional** focus
(value ***comparison***)

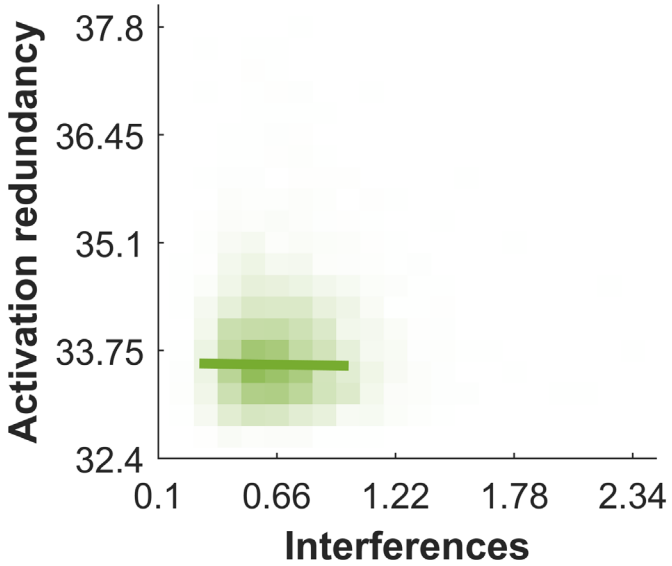


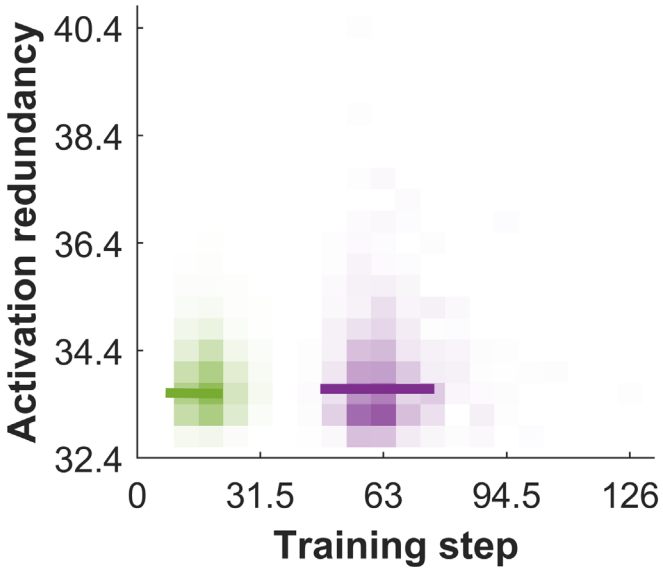
Temporal order
→ **Temporal** order
(value ***synthesis***)



Supplementary summary

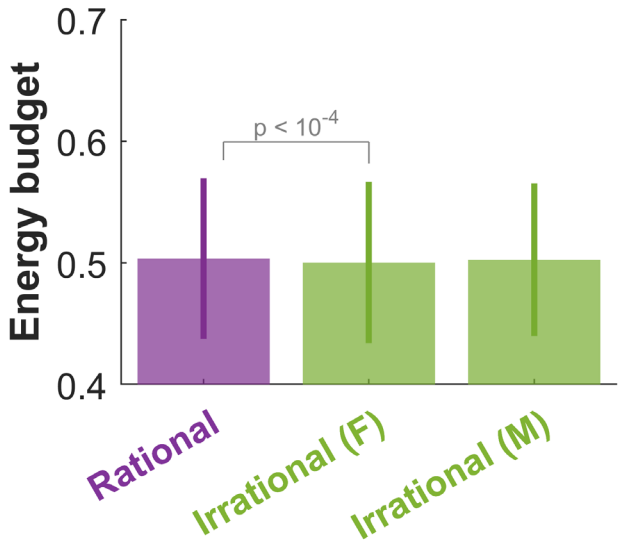




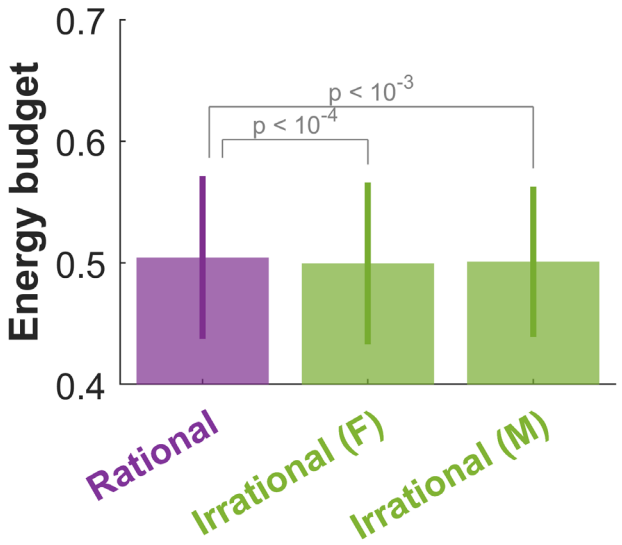


Supplementary summary

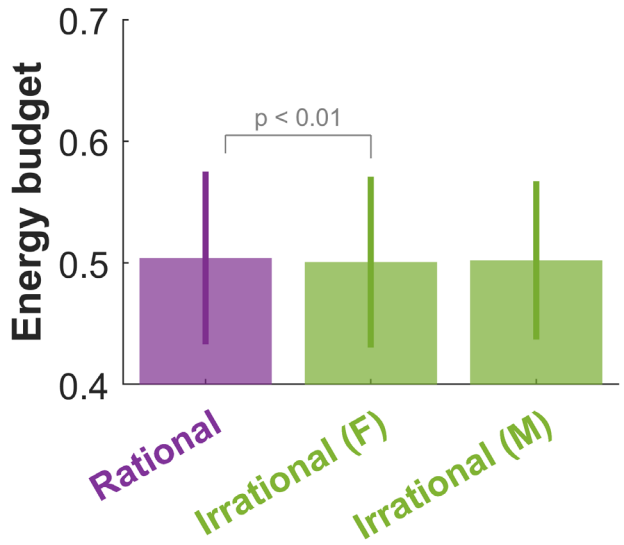
All candidates



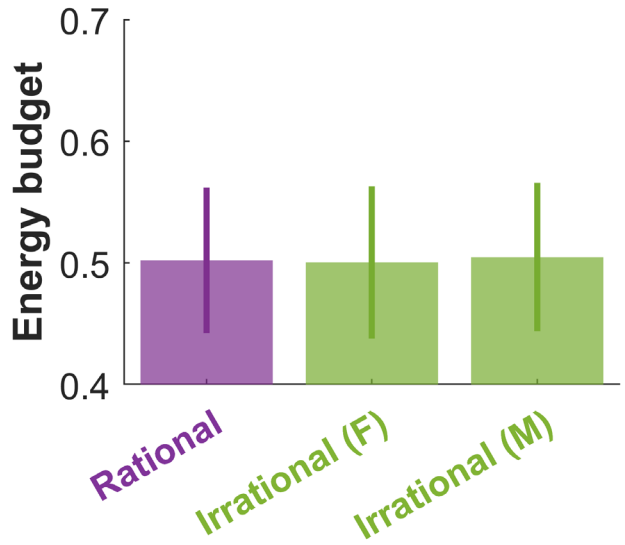
Temporal order
→ **Attentional** focus
(value ***synthesis***)



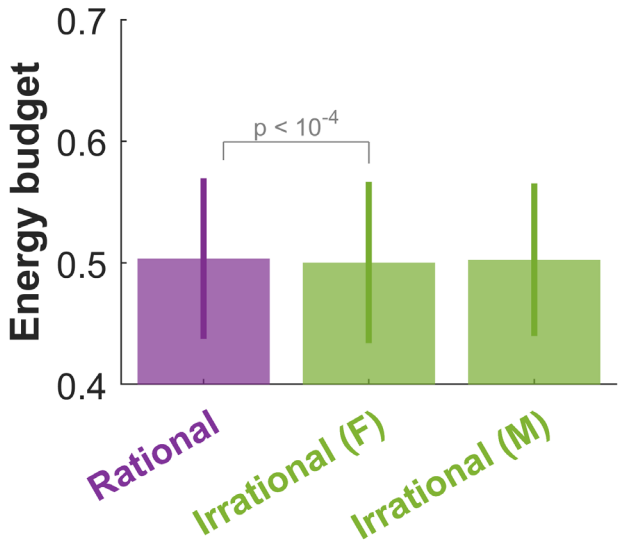
Temporal order
→ **Attentional** focus
(value ***comparison***)

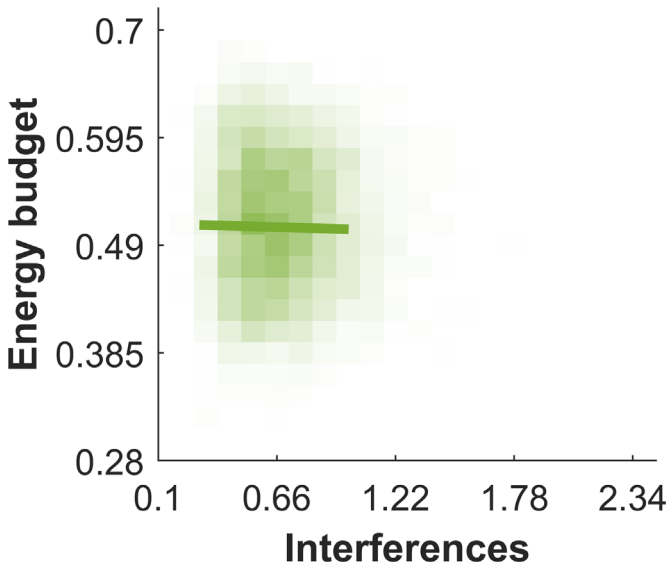


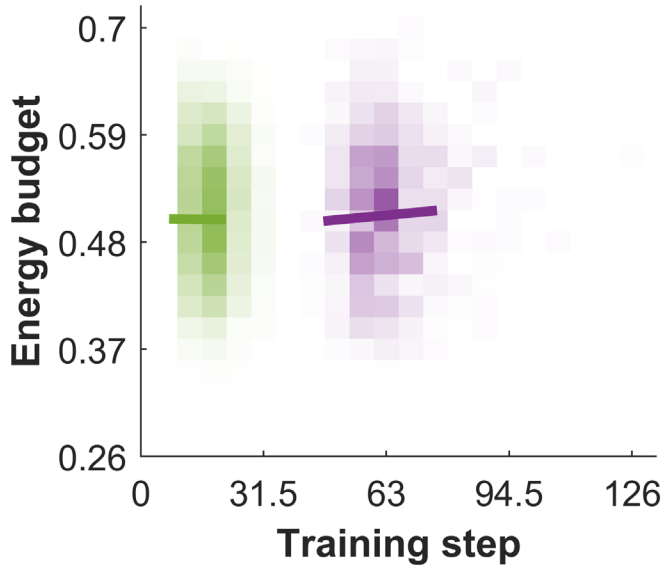
Temporal order
→ **Temporal** order
(value ***synthesis***)



Supplementary summary

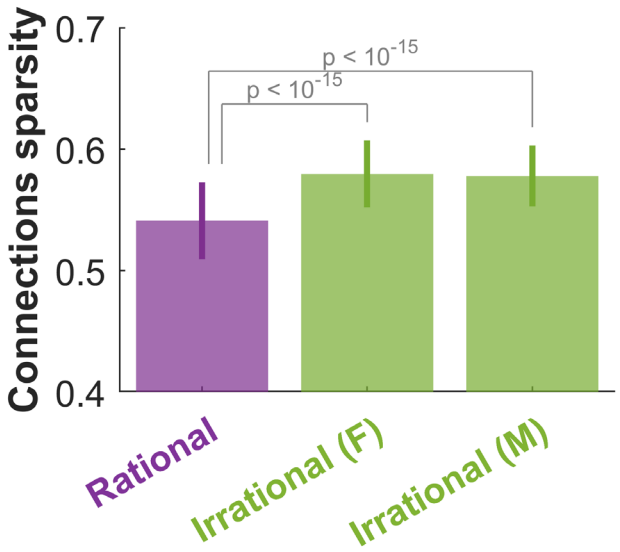




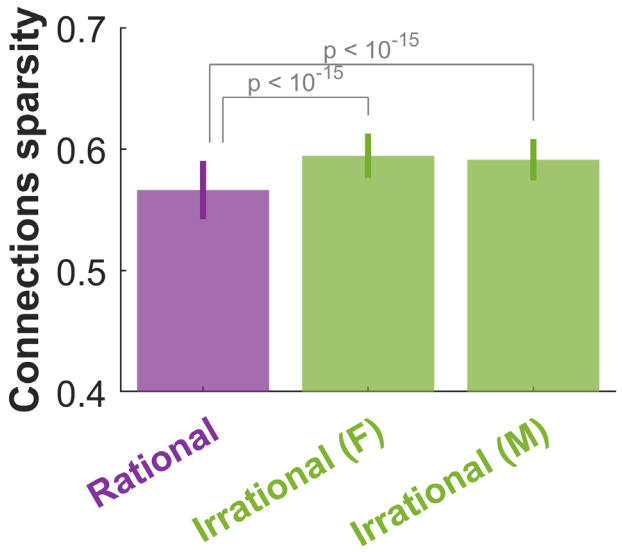


Supplementary summary

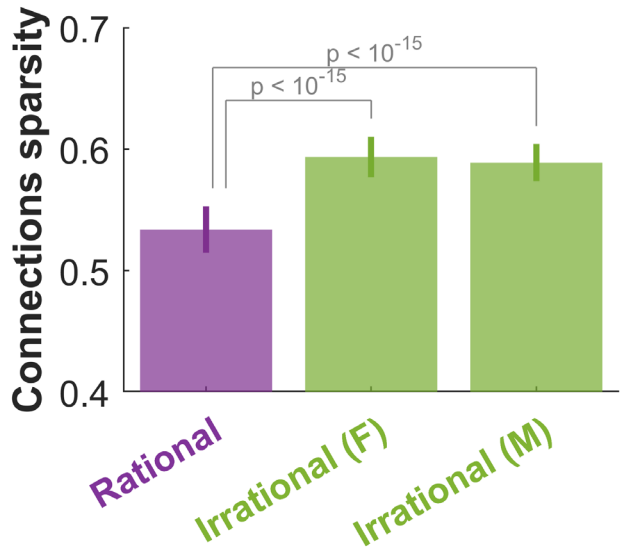
All candidates



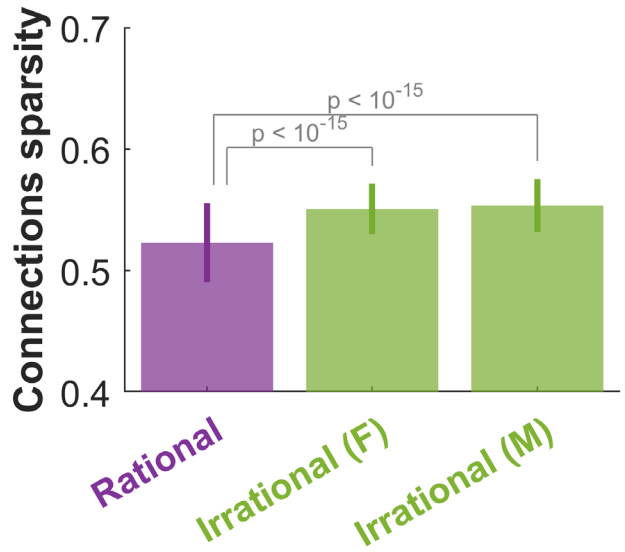
Temporal order
→ **Attentional** focus
(value ***synthesis***)



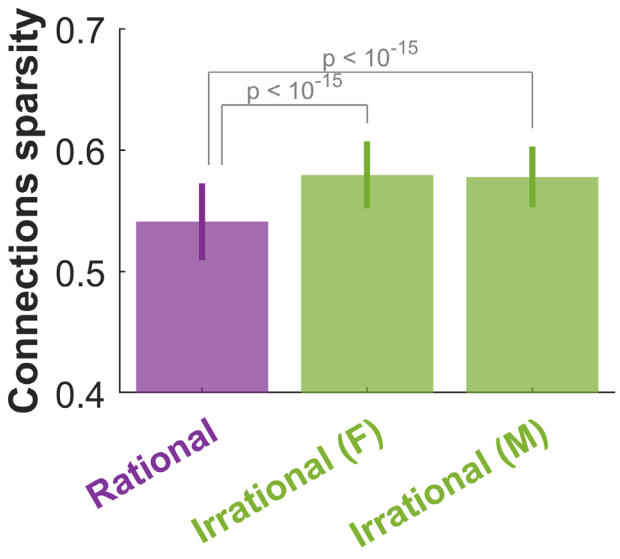
Temporal order
→ **Attentional** focus
(value ***comparison***)

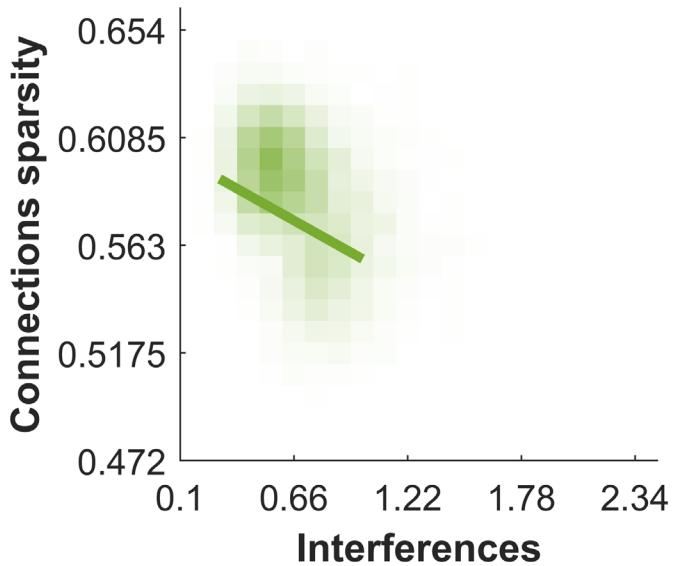


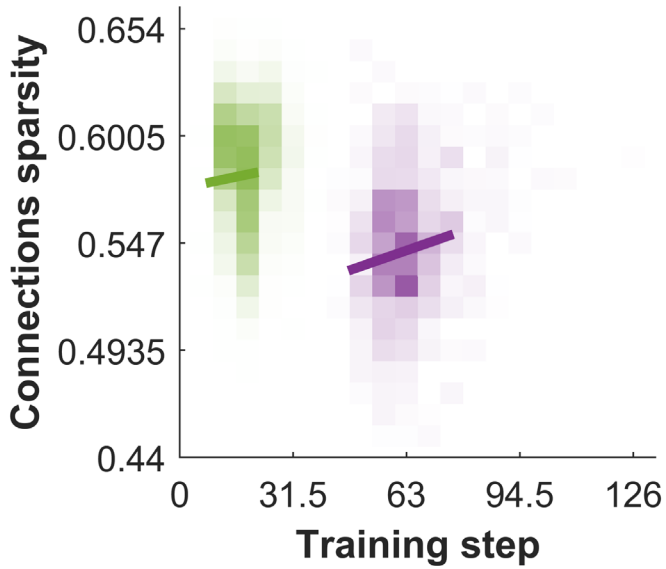
Temporal order
→ **Temporal** order
(value ***synthesis***)



Supplementary summary

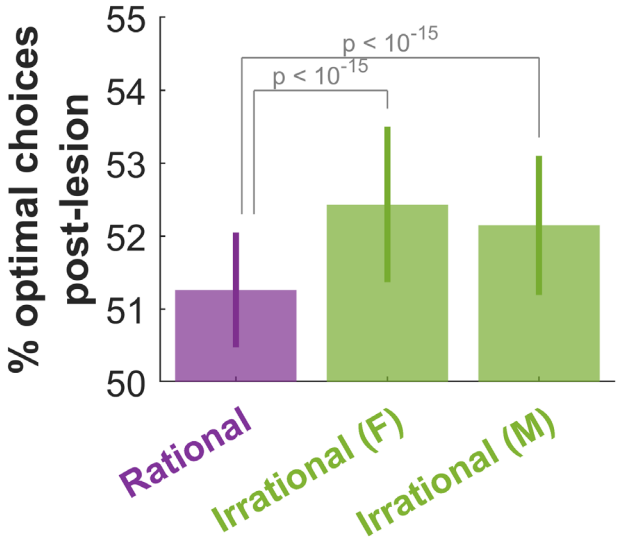




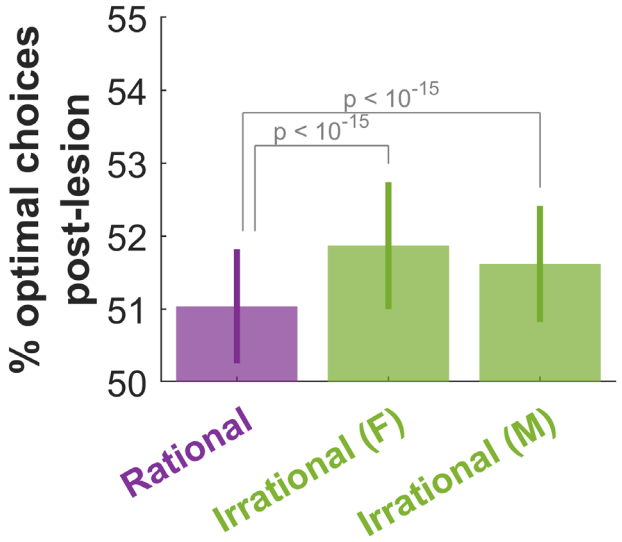


Supplementary summary

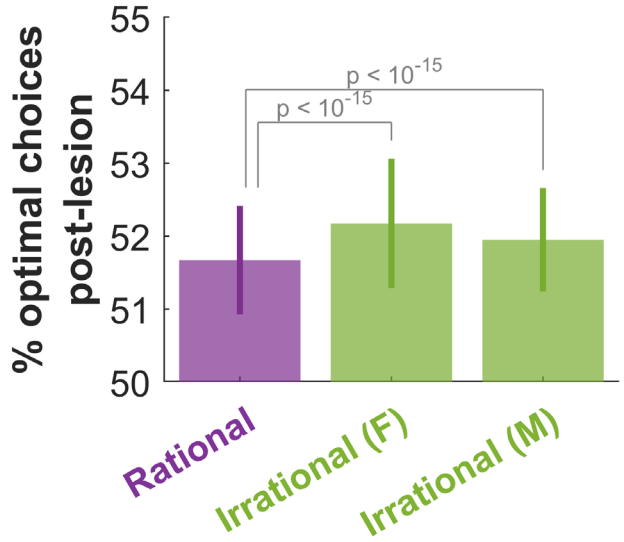
All candidates



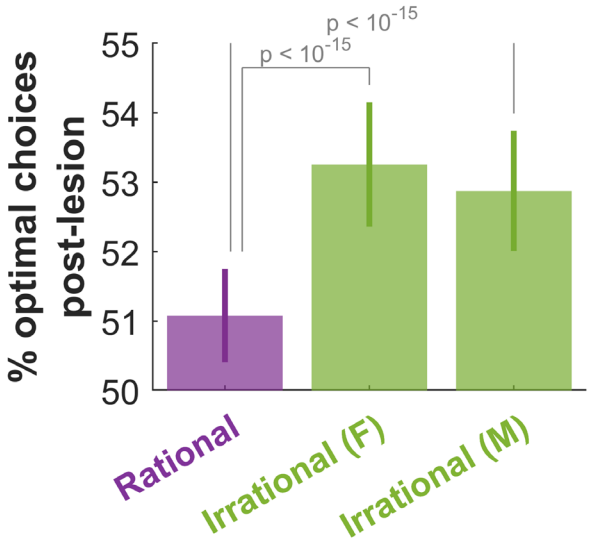
Temporal order
→ **Attentional** focus
(value ***synthesis***)



Temporal order
→ **Attentional** focus
(value ***comparison***)



Temporal order
→ **Temporal** order
(value ***synthesis***)



Supplementary summary