Aha! Moments

as meta-cognitive prediction errors

Hermish Mehta
University of California, Berkeley

Aha! Moments

as meta-cognitive prediction errors

in collaboration with
Rachit Dubey, Mark Ho, Tom Griffiths
Princeton University

outline

1. Motivation
2. Background
3. Theoretical Model
4. Experimental Results
5. Parting Thoughts

$\neq$

Not all inferences are equally rewarding.

Motivation

Why are some inferences more rewarding than others?

questions

1. Why do we have *Aha!* moments?
2. What makes these moments so rare?
3. Why are they so rewarding?

[Marr 1952]

computational

algorithmic

implementational

[Marr 1952]

computational

What is the goal of the system?
How can we formulate the objective?

algorithmic

implementational

[Marr 1952]

What is the goal of the system?

computational

How can we formulate the objective?

What is an appropriate algorithm to achieve this goal?

algorithmic

implementational

[Marr 1952]

computational

What is the goal of the system?
How can we formulate the objective?

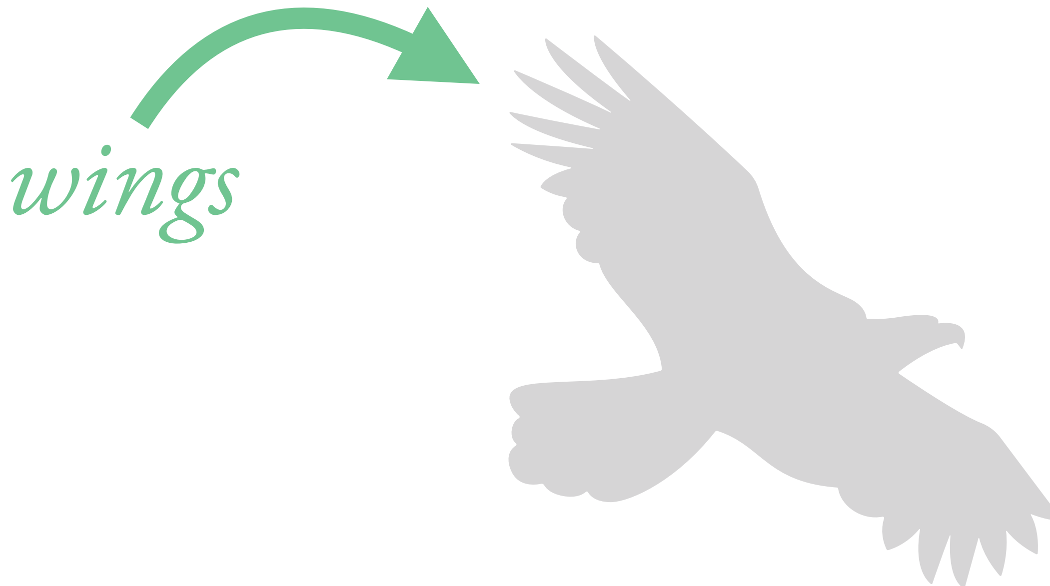
algorithmic

What is an appropriate algorithm to achieve this goal?

implementational

How is the algorithm implemented?

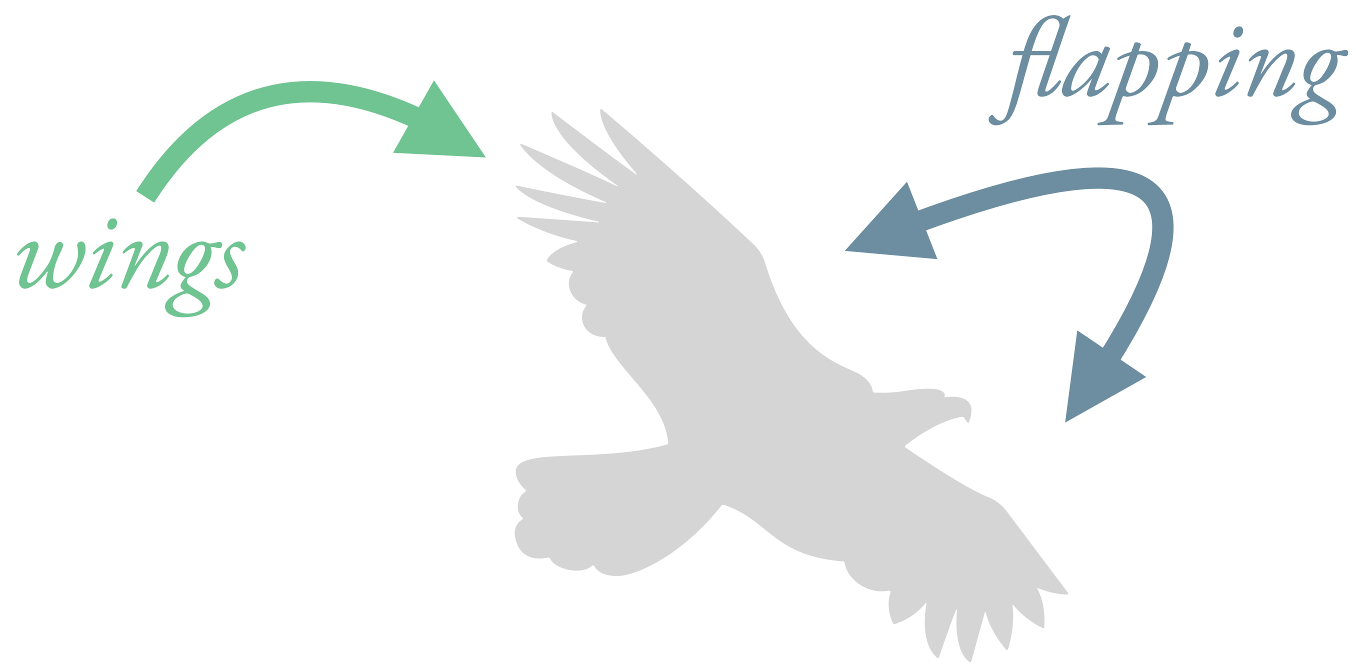
computational	Why? *(problem)*
algorithmic	What? *(rules)*
implementational	How? *(physical)*



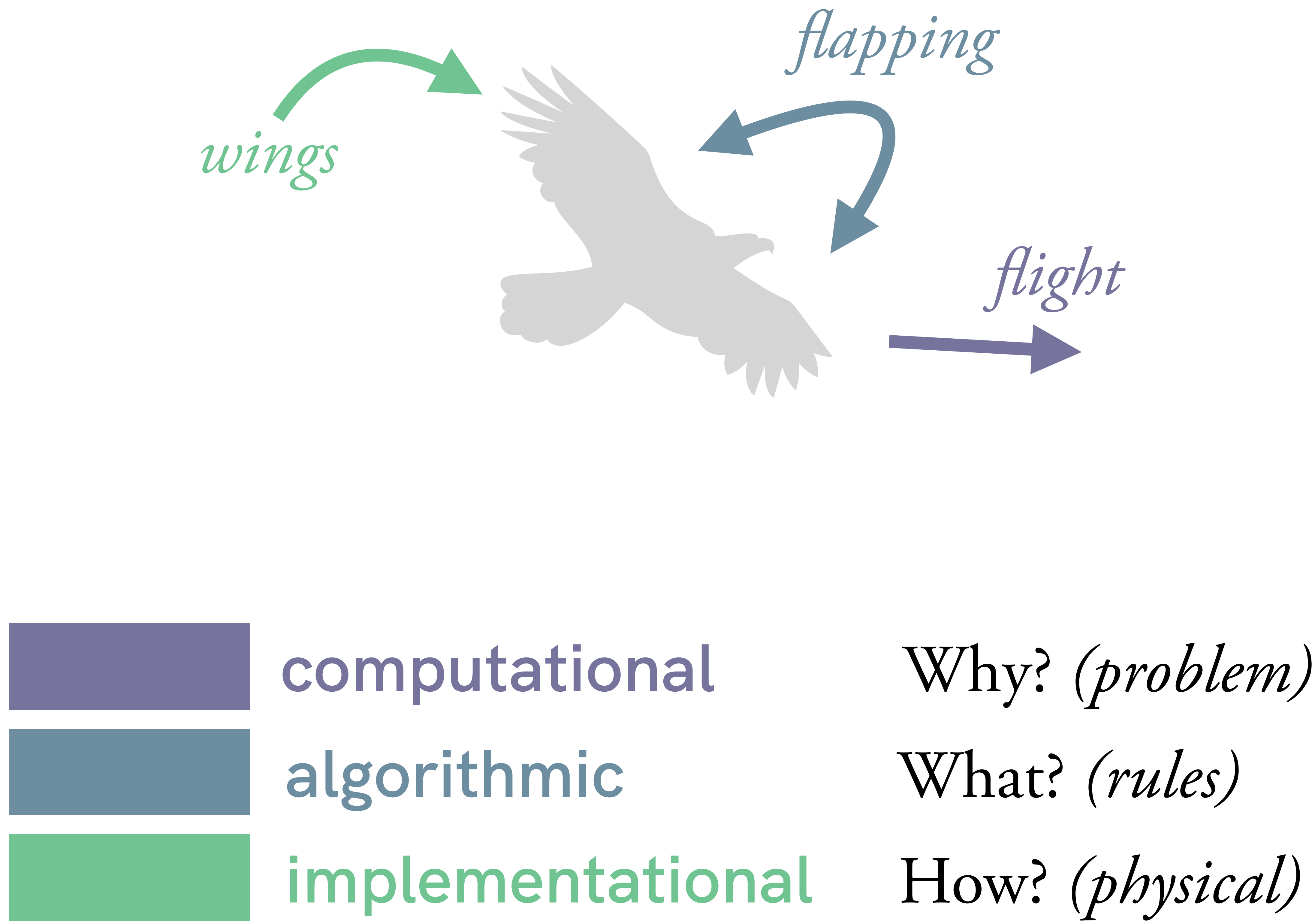
computational Why? *(problem)*

algorithmic What? *(rules)*

implementational How? *(physical)*



computational	Why? *(problem)*
algorithmic	What? *(rules)*
implementational	How? *(physical)*



wings
flapping
flight
computational
algorithmic
implementational
Why? (problem)
What? (rules)
How? (physical)

[Marr 1952]

"Trying to understand perception by studying only neurons is like trying to understand bird flight by studying only feathers: It just cannot be done."

key idea

Aha! moments arise under resource constraints and are meta-cognitive prediction errors which cause reward prediction errors.

meta-cognitive prediction errors

reward prediction errors

resource constraints

Background

learning |ˈlərniNG|

(noun)

A non-temporary change in the behavioral mechanisms engaged in a certain situation… that results from repeated experience with the situation… and providing the change can't be explained in terms of innate behavioral tendencies.

learning |ˈlərniNG|

The process of improving predictions of the future.

$$\left\{\begin{matrix}\textit{prediction}\\ \textit{error}\end{matrix}\right\} = \left\{\begin{matrix}\textit{actual}\\ \textit{outcome}\end{matrix}\right\} - \left\{\begin{matrix}\textit{predicted}\\ \textit{outcome}\end{matrix}\right\}$$

Critical to learning.

$$\left\{ \begin{matrix} \textit{prediction} \\ \textit{error} \end{matrix} \right\} = \left\{ \begin{matrix} \textit{actual} \\ \textit{outcome} \end{matrix} \right\} - \left\{ \begin{matrix} \textit{predicted} \\ \textit{outcome} \end{matrix} \right\}$$

image *label*

x_i

y_i 0

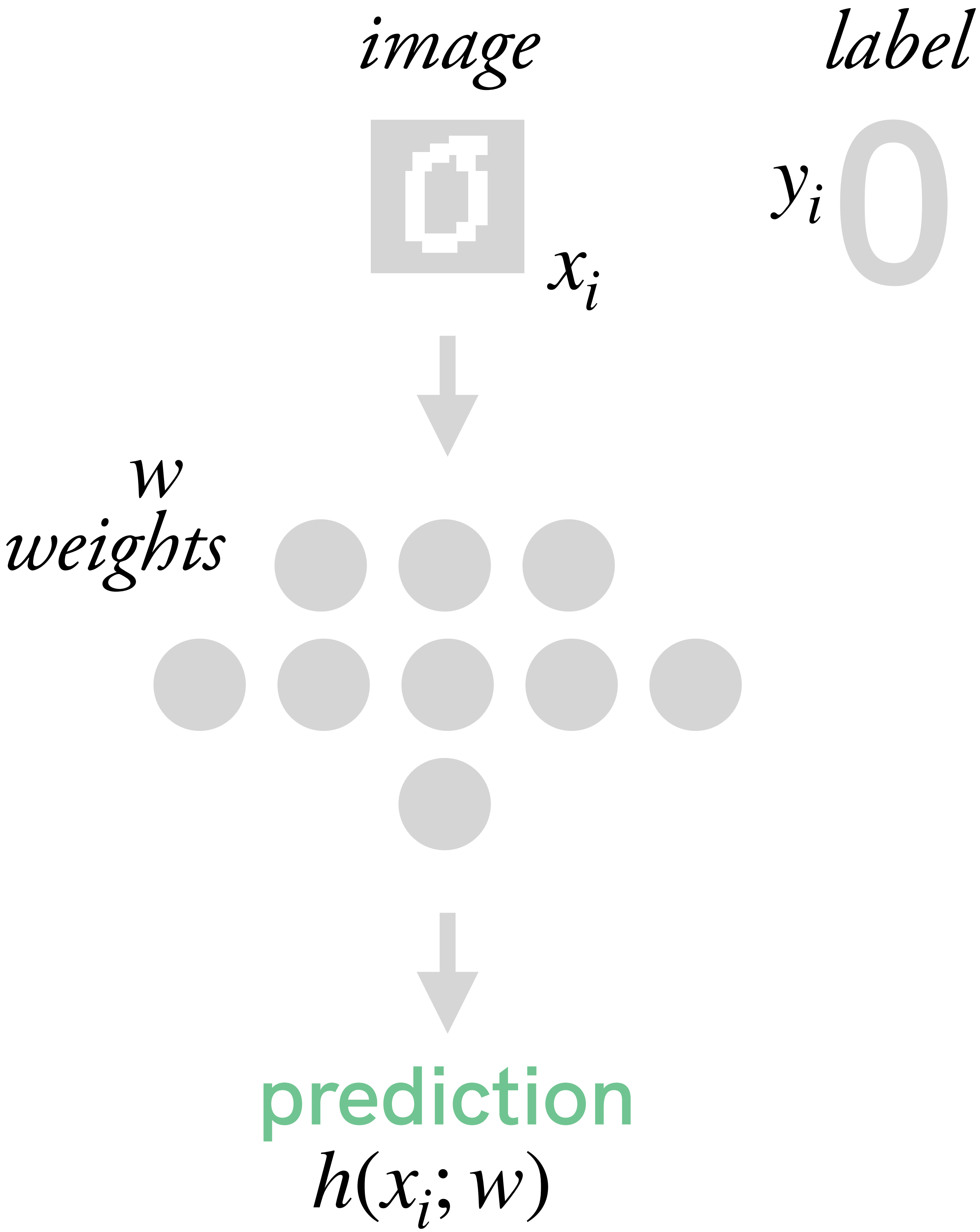


image
label
y_i
0
x_i
w
weights
prediction
$h(x_i; w)$

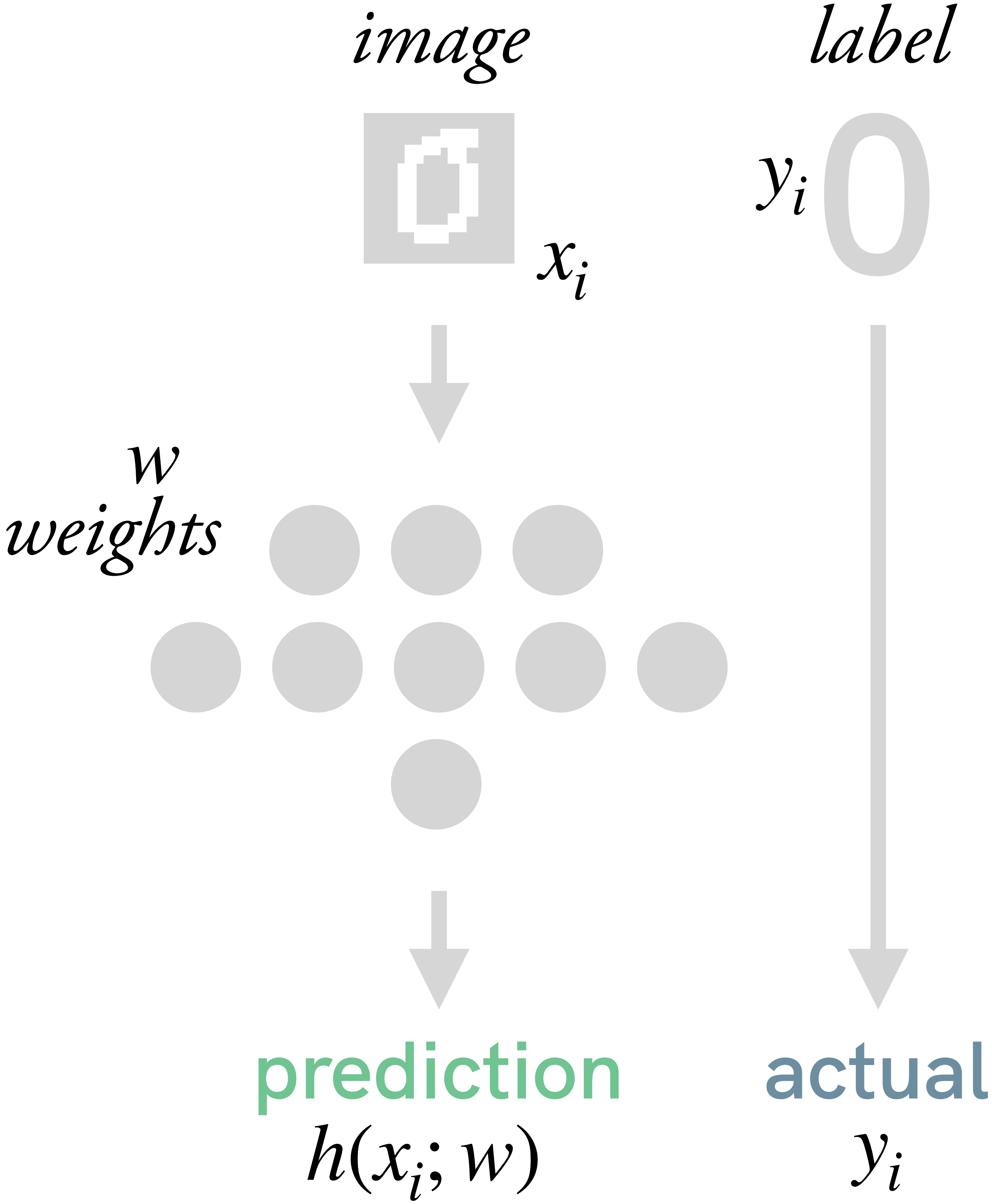


image
label
x_i
y_i
0
w
weights
prediction
$h(x_i; w)$
actual
y_i

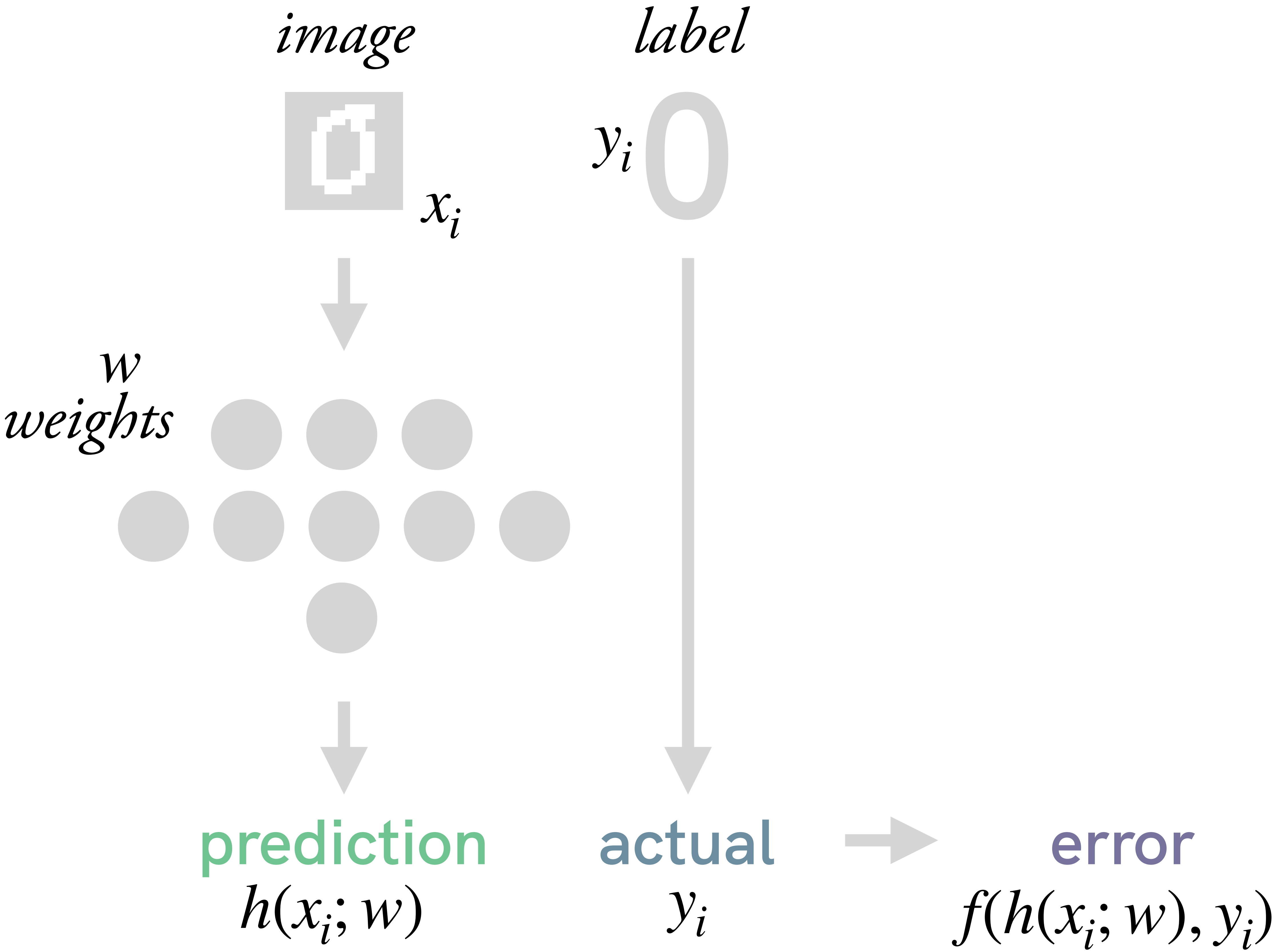
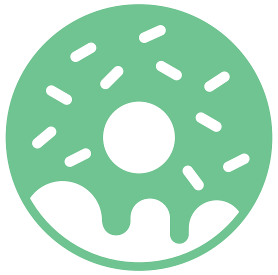


image
label
x_i
y_i
0
w
weights
prediction
$h(x_i; w)$
actual
y_i
error
$f(h(x_i; w), y_i)$

machine learning

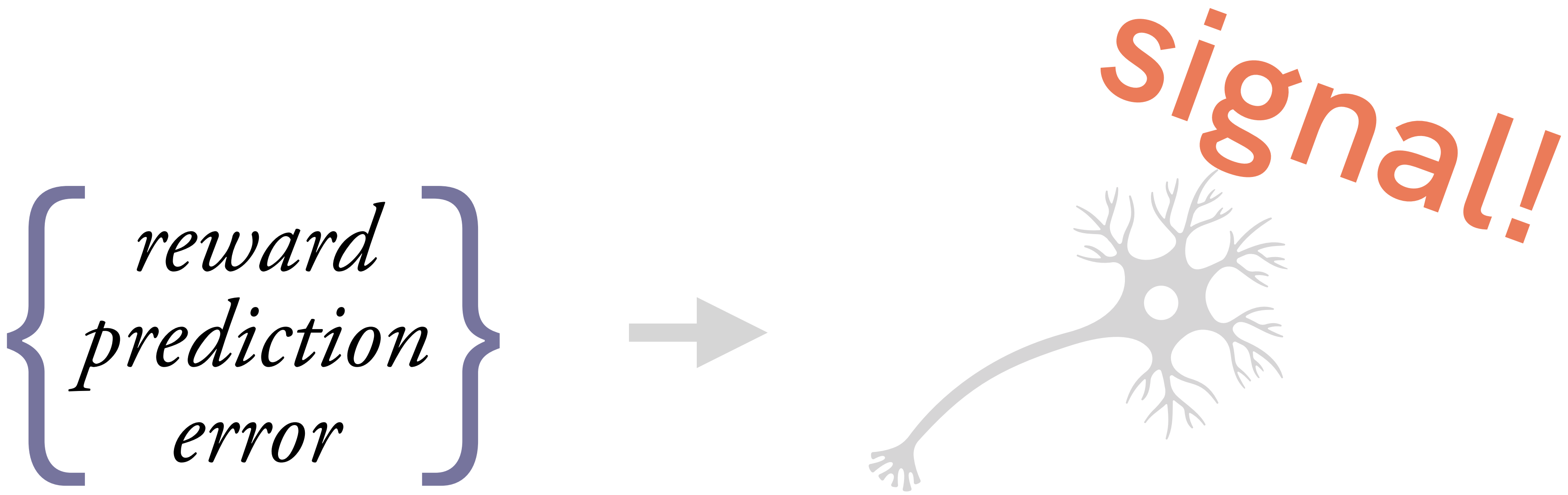
$$w^* = \arg\min_{w} \frac{1}{n} \sum_{i=1}^{n} \underline{f(h(x_i; w), y_i)}$$

Minimize average prediction error.



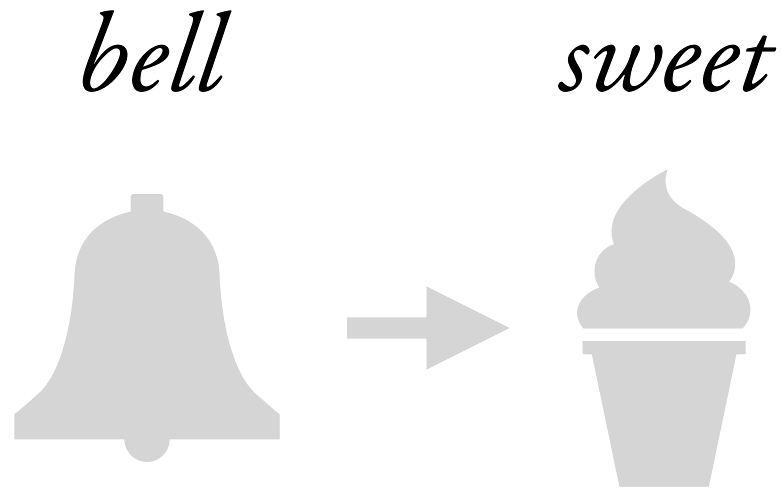
Rewards can also be predicted.

$$\left\{ \textit{reward prediction error} \right\} = \left\{ \textit{actual reward} \right\} - \left\{ \textit{predicted reward} \right\}$$



The phasic dopamine response signals prediction error.

bell *sweet*

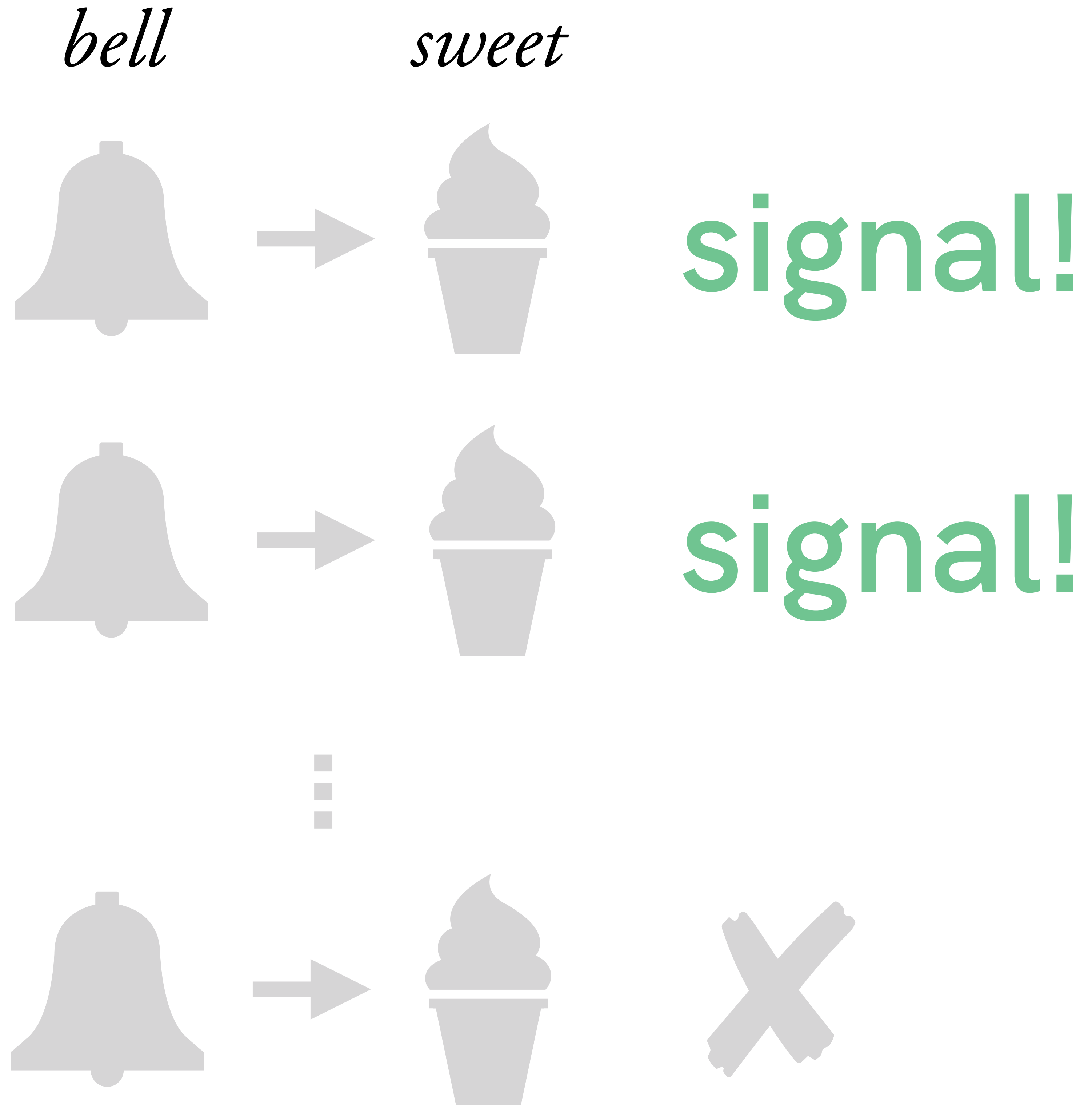


signal!

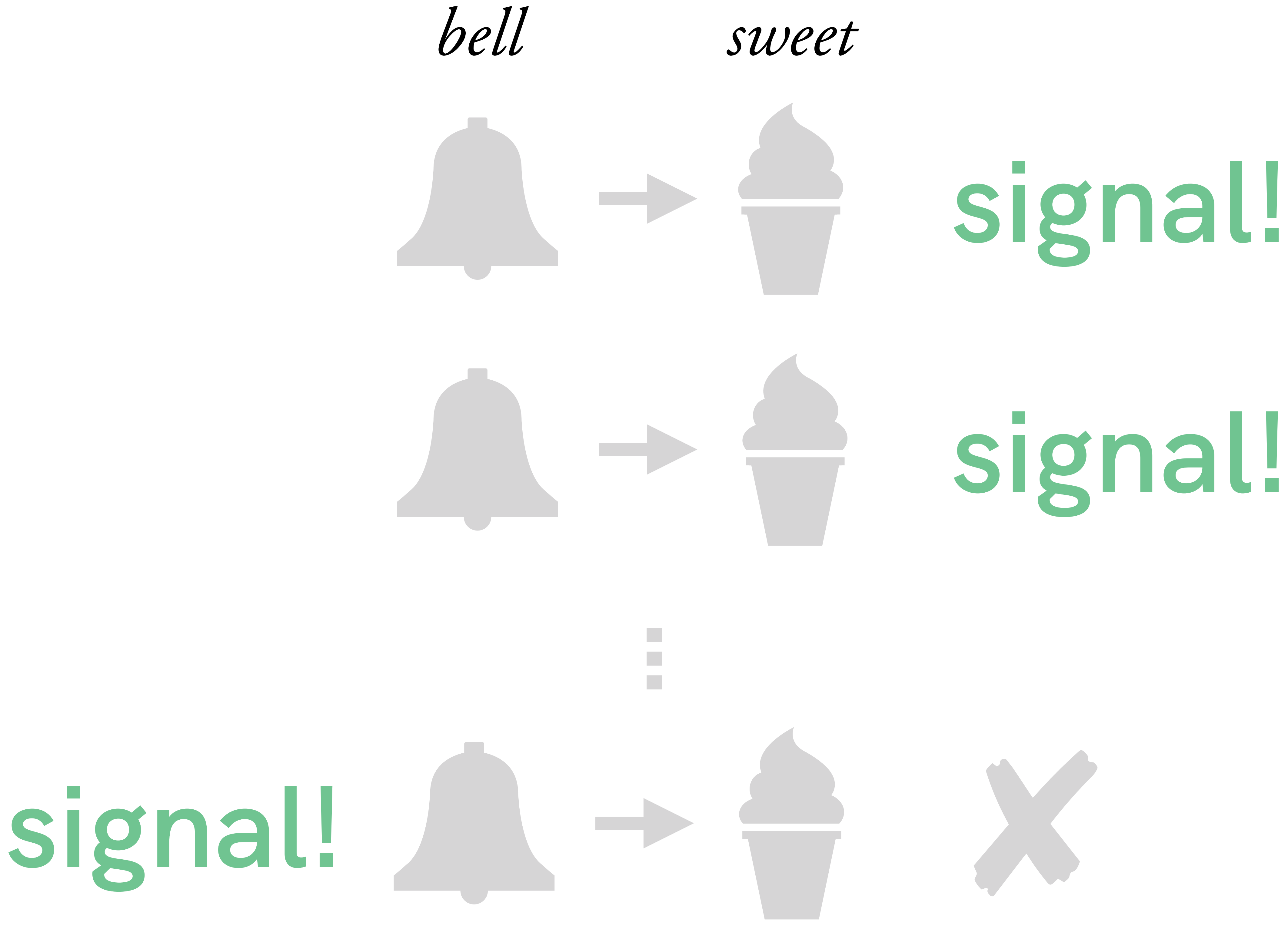
bell *sweet*

signal!

signal!



bell
sweet
signal!
signal!


bell
sweet
signal!
signal!
signal!

bell

bell

[Schultz 1952]

reward prediction error

positive

no error

negative

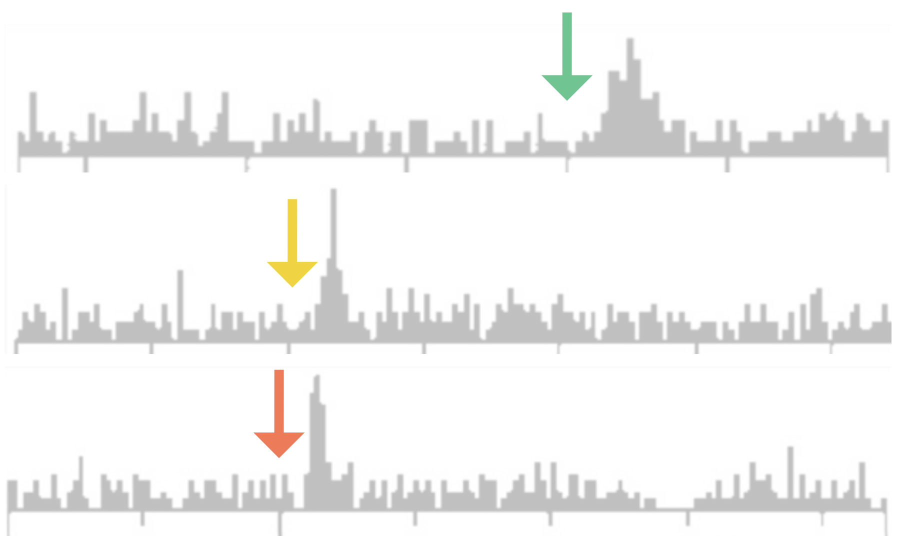
[Schultz 1952]

reward prediction error

positive

no error

negative



dopamine response

Bayer, Hannah M., and Paul W. Glimcher. "Midbrain dopamine neurons encode a quantitative reward prediction error signal." *Neuron* 47.1 (2005): 129-141.

Hammer, Martin. "The neural basis of associative reward learning in honeybees." *Trends in neurosciences* 20.6 (1997): 245-252.

Hollerman, Jeffrey R., and Wolfram Schultz. "Dopamine neurons report an error in the temporal prediction of reward during learning." *Nature neuroscience* 1.4 (1998): 304.

Schultz, Wolfram, Peter Dayan, and P. Read Montague. "A neural substrate of prediction and reward." *Science* 275.5306 (1997): 1593-1599.

Hart, Andrew S., et al. "Phasic dopamine release in the rat nucleus accumbens symmetrically encodes a reward prediction error term." *Journal of Neuroscience* 34.3 (2014): 698-704.

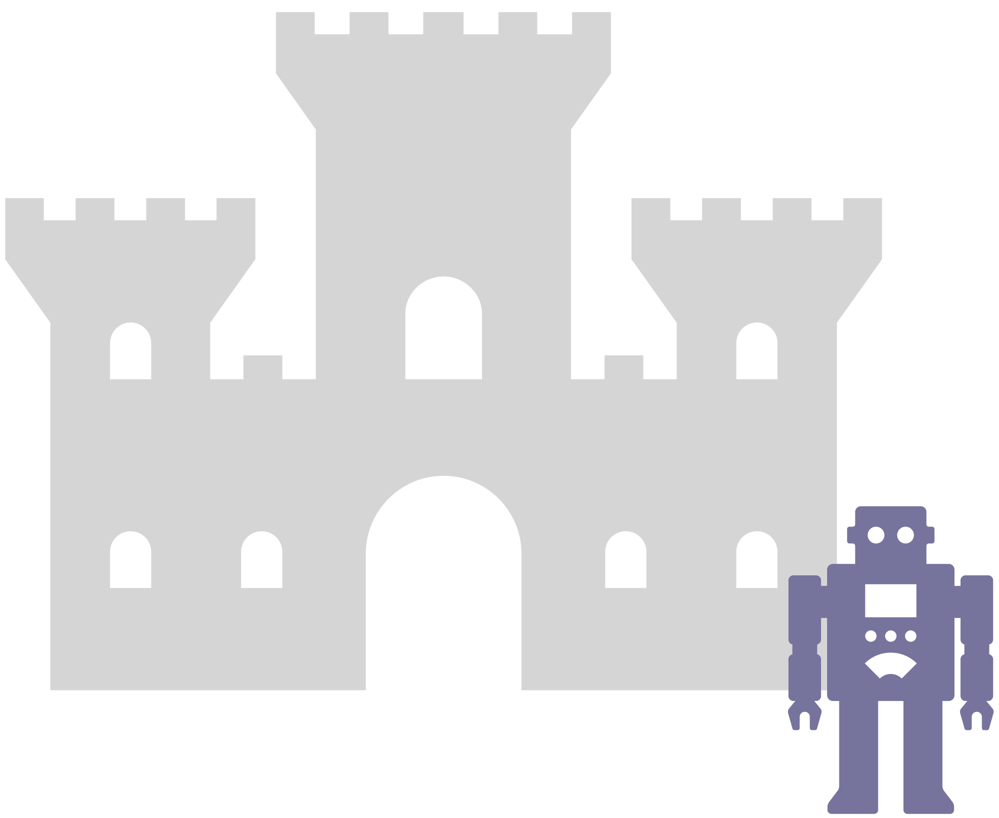
Colombo, Matteo. "Deep and beautiful. The reward prediction error hypothesis of dopamine." *Studies in history and philosophy of science part C: Studies in history and philosophy of biological and biomedical sciences* 45 (2014): 57-67.

[Horvitz 1988] [Russell 1997]

bounded optimality

Artificial agents need to tradeoff between efficiency and accuracy.

computational limitations & costs



agent

algorithm 1

algorithm 2

algorithm 3

algorithm 4

resource rationality

People need to make optimal use of their limited resources, which involves "meta-reasoning," thinking about thinking.

Griffiths, Thomas L., Falk Lieder, and Noah D. Goodman. "Rational use of cognitive resources: Levels of analysis between the computational and the algorithmic." *Topics in cognitive science* 7.2 (2015): 217-229.

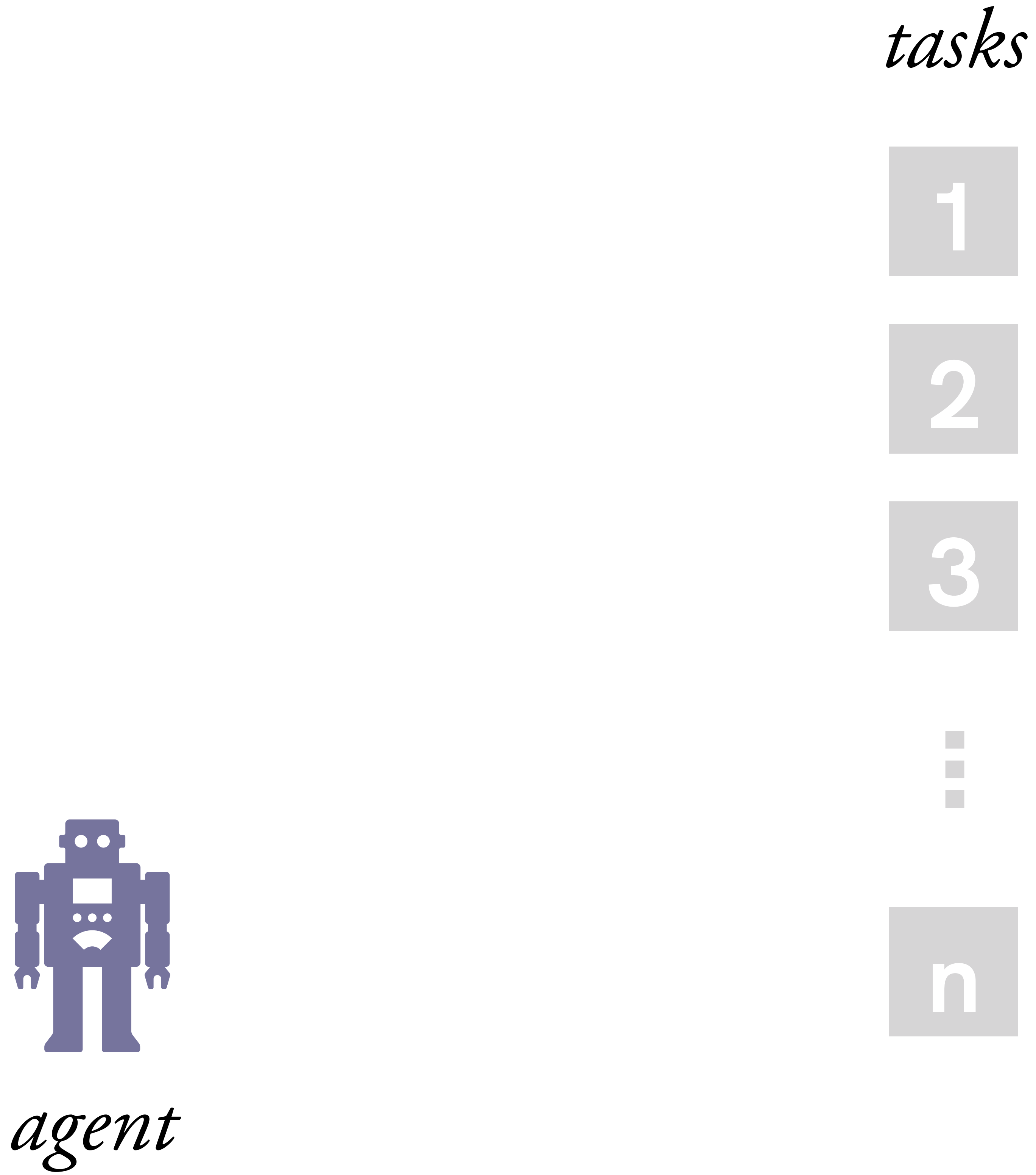
Lieder, Falk, Ming Hsu, and Thomas L. Griffiths. "The high availability of extreme events serves resource-rational decision-making." *Proceedings of the annual meeting of the cognitive science society.* Vol. 36. No. 36. 2014.

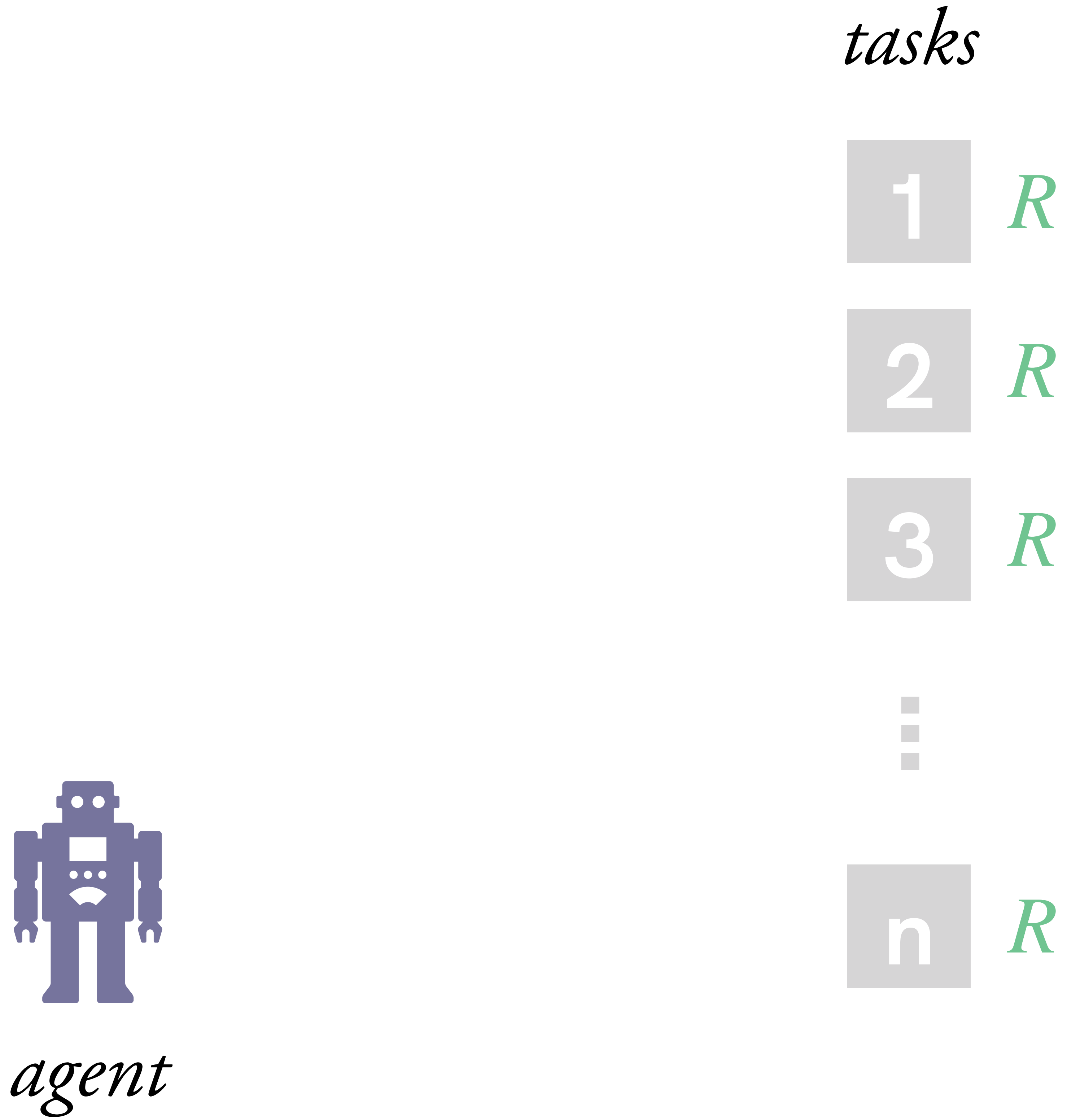
Lieder, Falk, et al. "The anchoring bias reflects rational use of cognitive resources." *Psychonomic bulletin & review* 25.1 (2018): 322-349.

Shenhav, Amitai, et al. "Toward a rational and mechanistic account of mental effort." *Annual review of neuroscience* 40 (2017): 99-124.

Lieder, Falk, and Thomas L. Griffiths. "Strategy selection as rational metareasoning." *Psychological Review* 124.6 (2017): 762.

Theoretical Model


tasks
1
2
3
n
agent


tasks
1
R
2
R
3
R
n
R
agent

goal

Maximize reward.

tasks

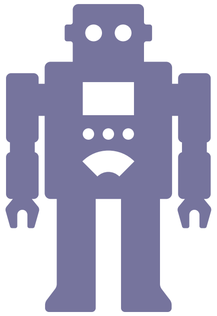
1 R

2 R

3 R

⋮

n R



agent

goal

Complete as many tasks as possible.

tasks

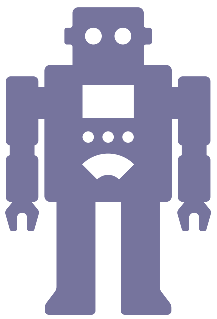
1 *R*

2 *R*

3 *R*

⋮

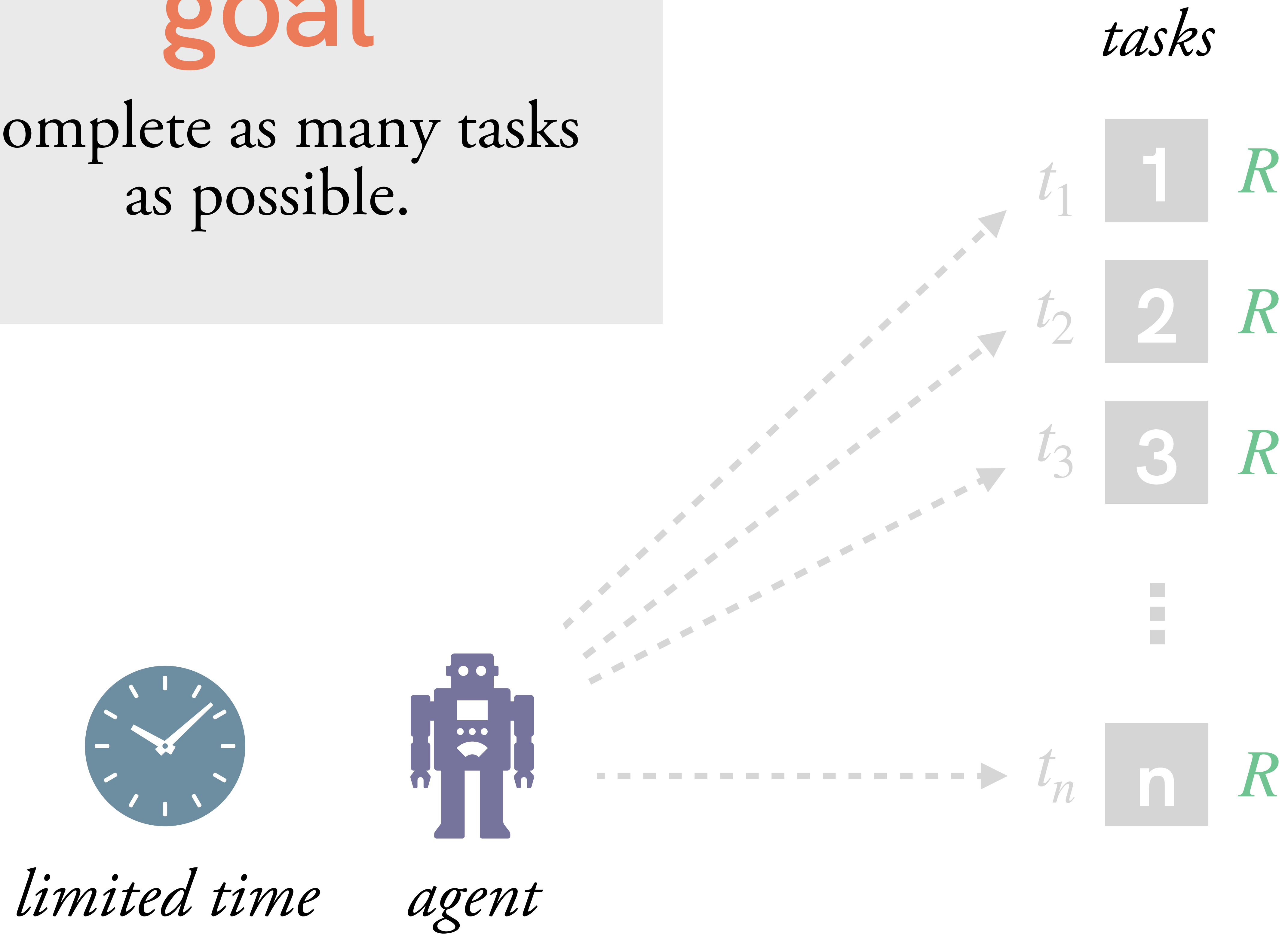
n *R*

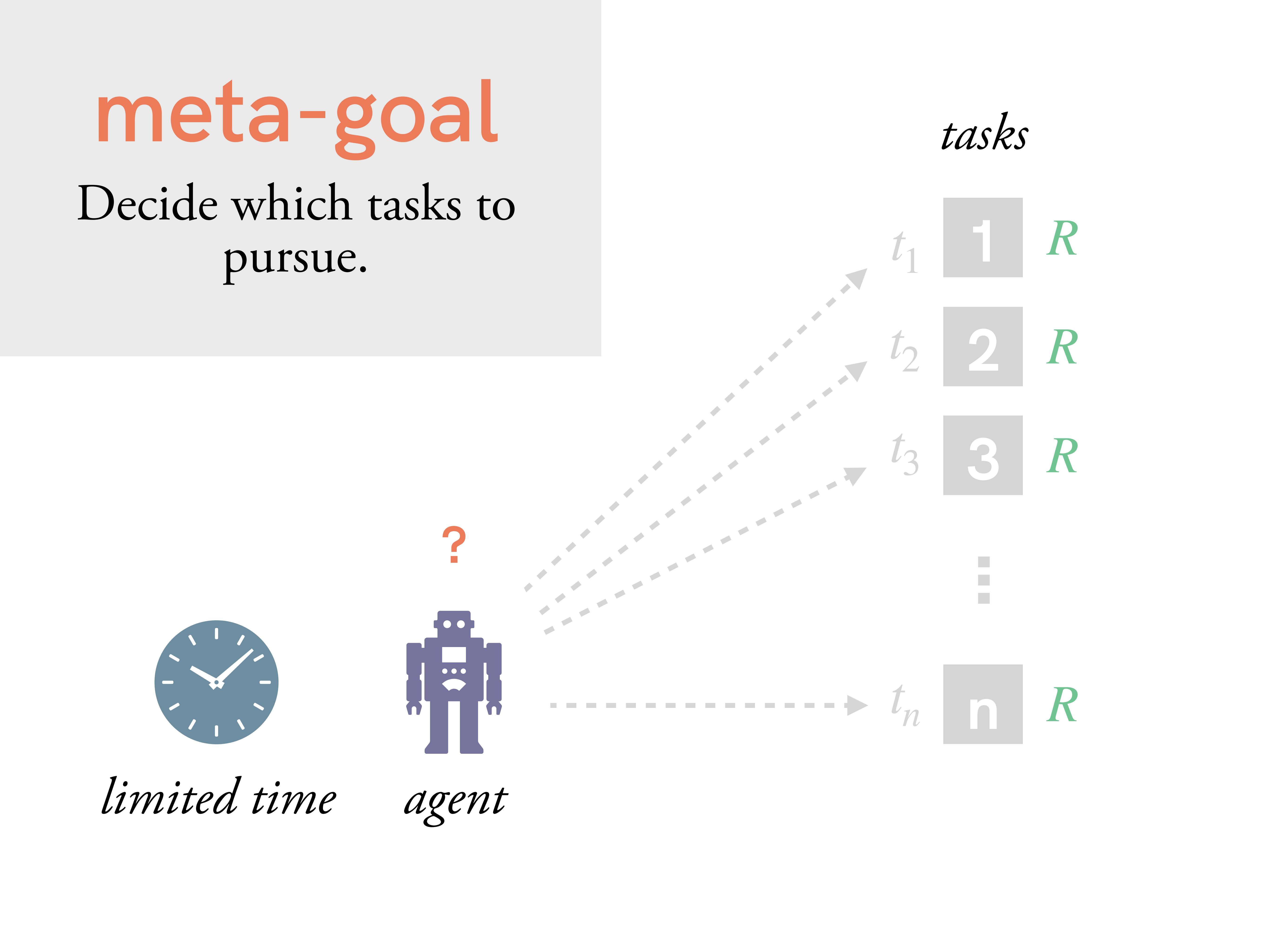


agent

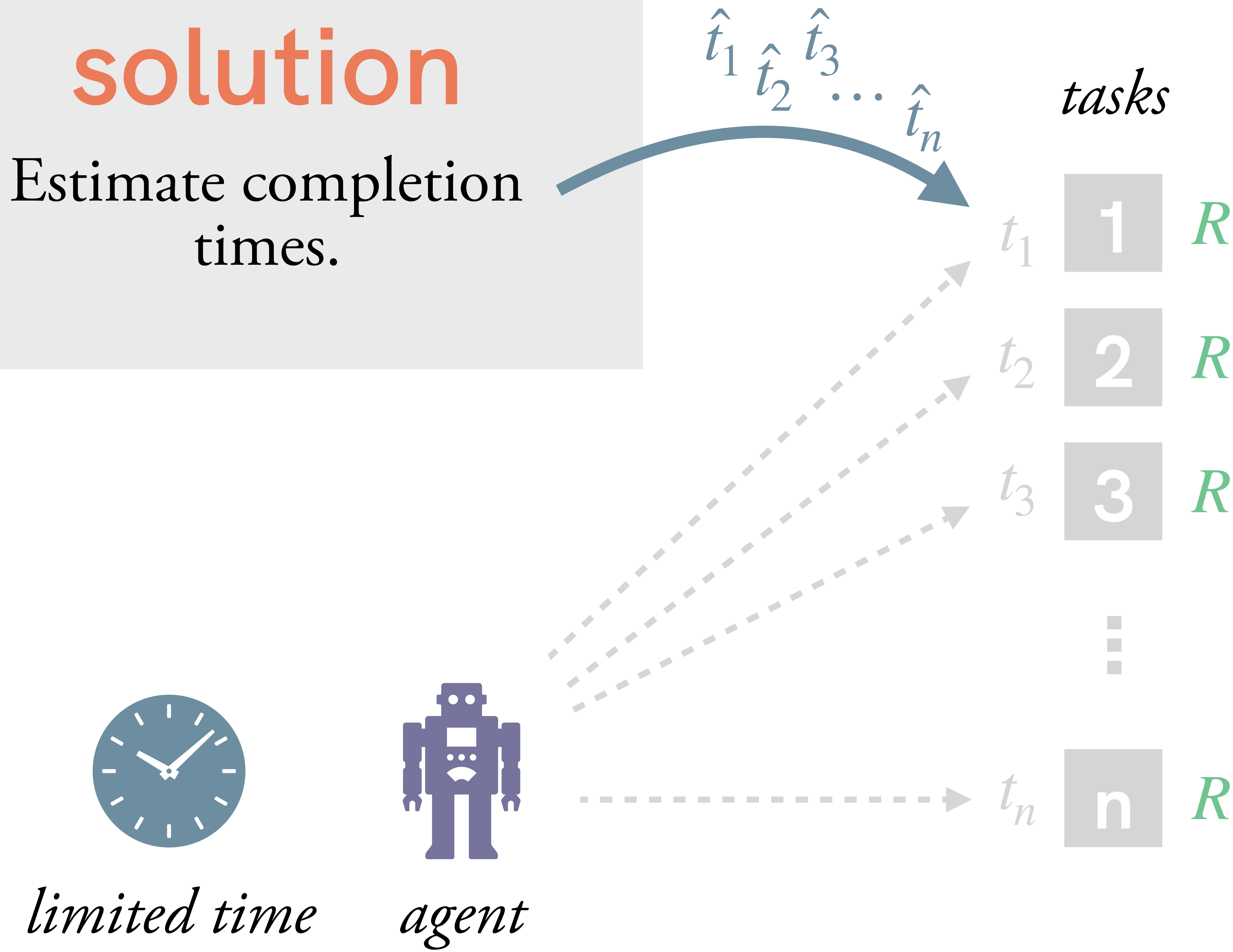
goal

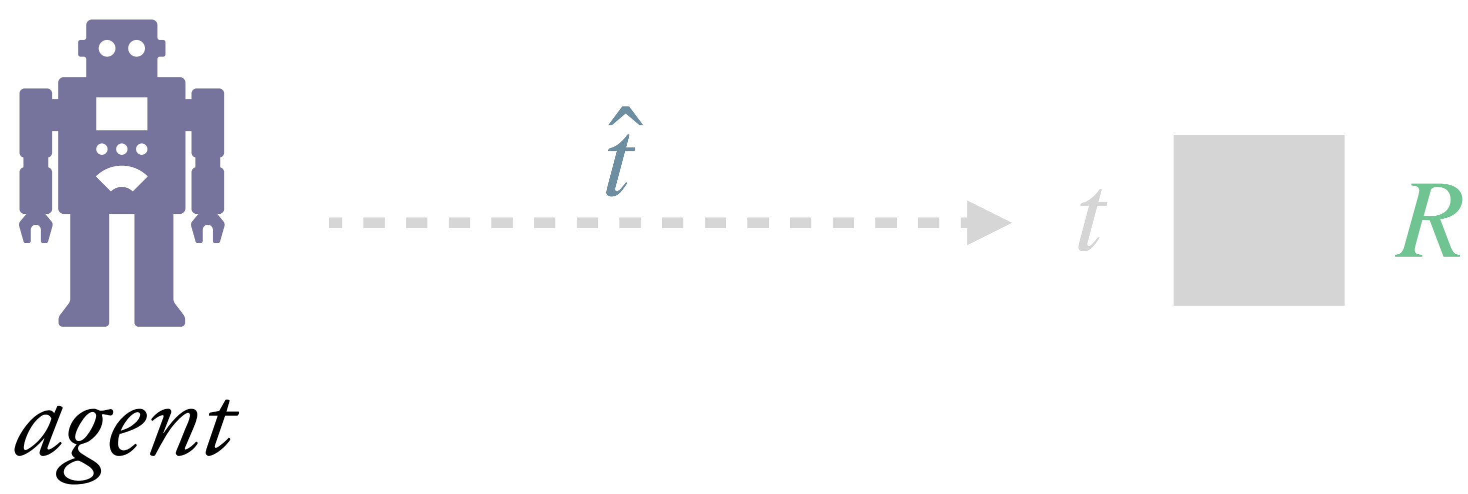
Complete as many tasks as possible.





meta-goal
Decide which tasks to pursue.
tasks
t_1
1
R
t_2
2
R
t_3
3
R
t_n
n
R
?
limited time
agent


solution
Estimate completion times.
$\hat{t}_1$ $\hat{t}_2$ $\hat{t}_3$ … $\hat{t}_n$
tasks
t_1 1 R
t_2 2 R
t_3 3 R
t_n n R
limited time
agent



$\hat{t}$
t
R
agent

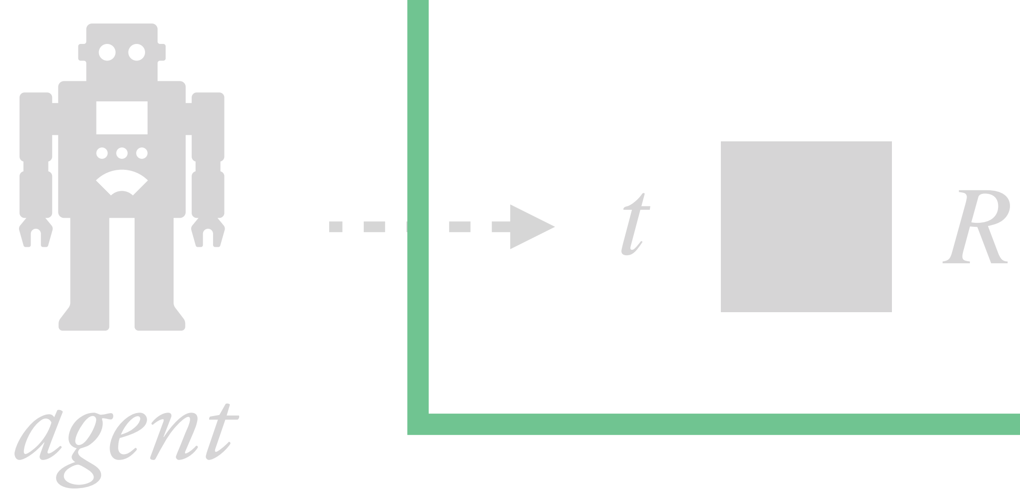
(meta-cognitive)
time prediction error

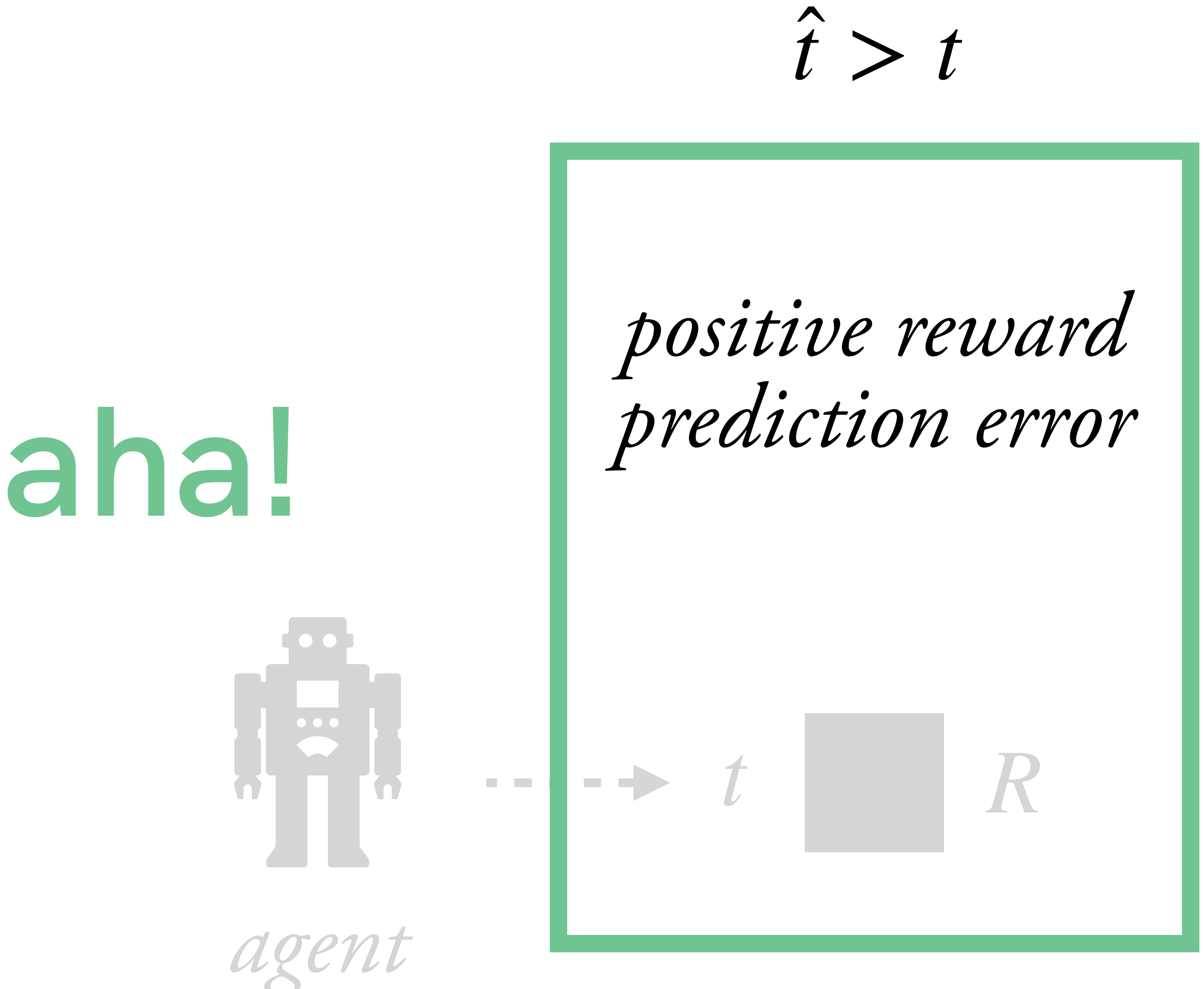
$\hat{t} > t$	positive
$\hat{t} = t$	no error
$\hat{t} < t$	negative

agent
t
R

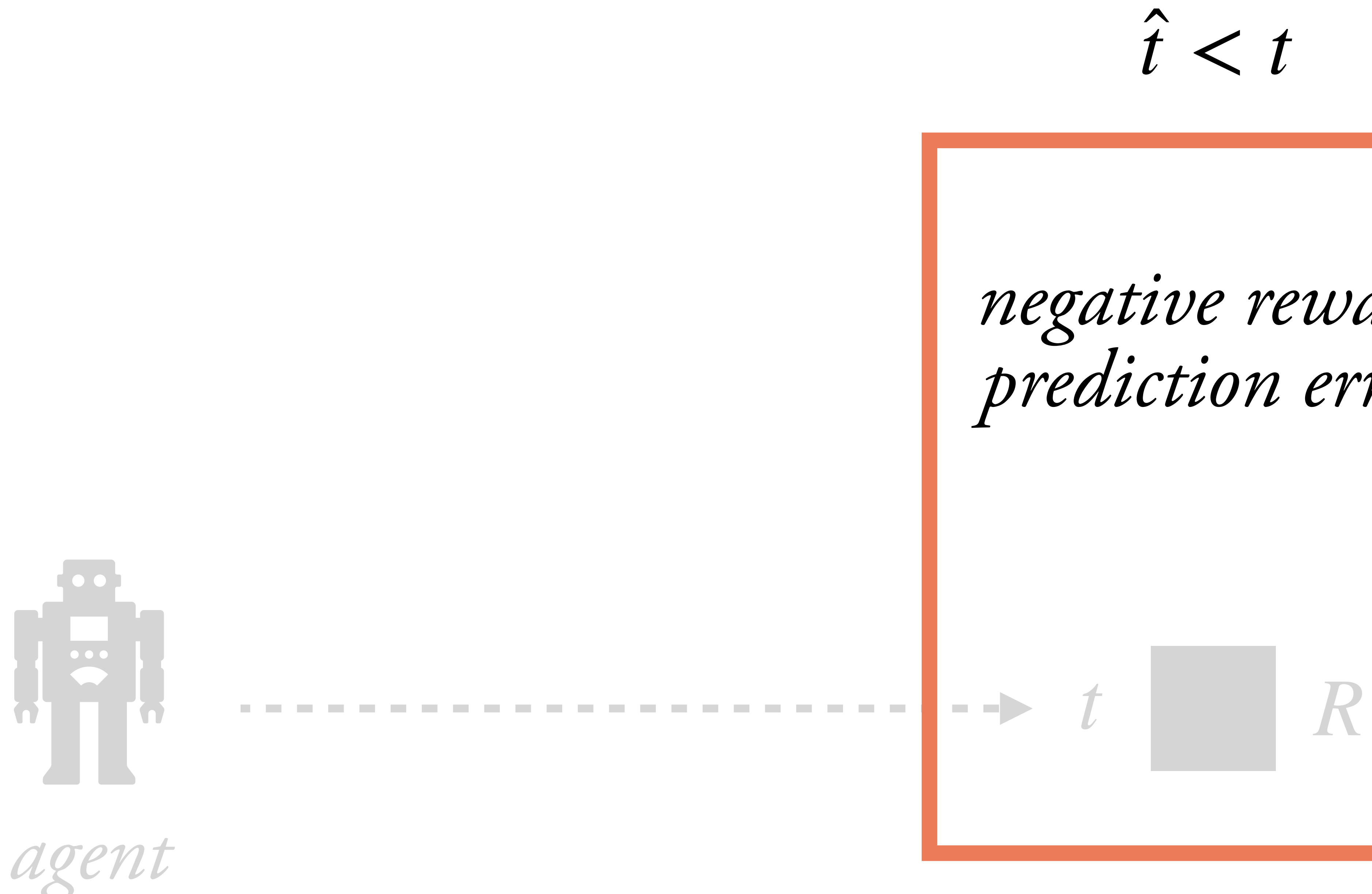
$\hat{t} > t$

positive reward prediction error

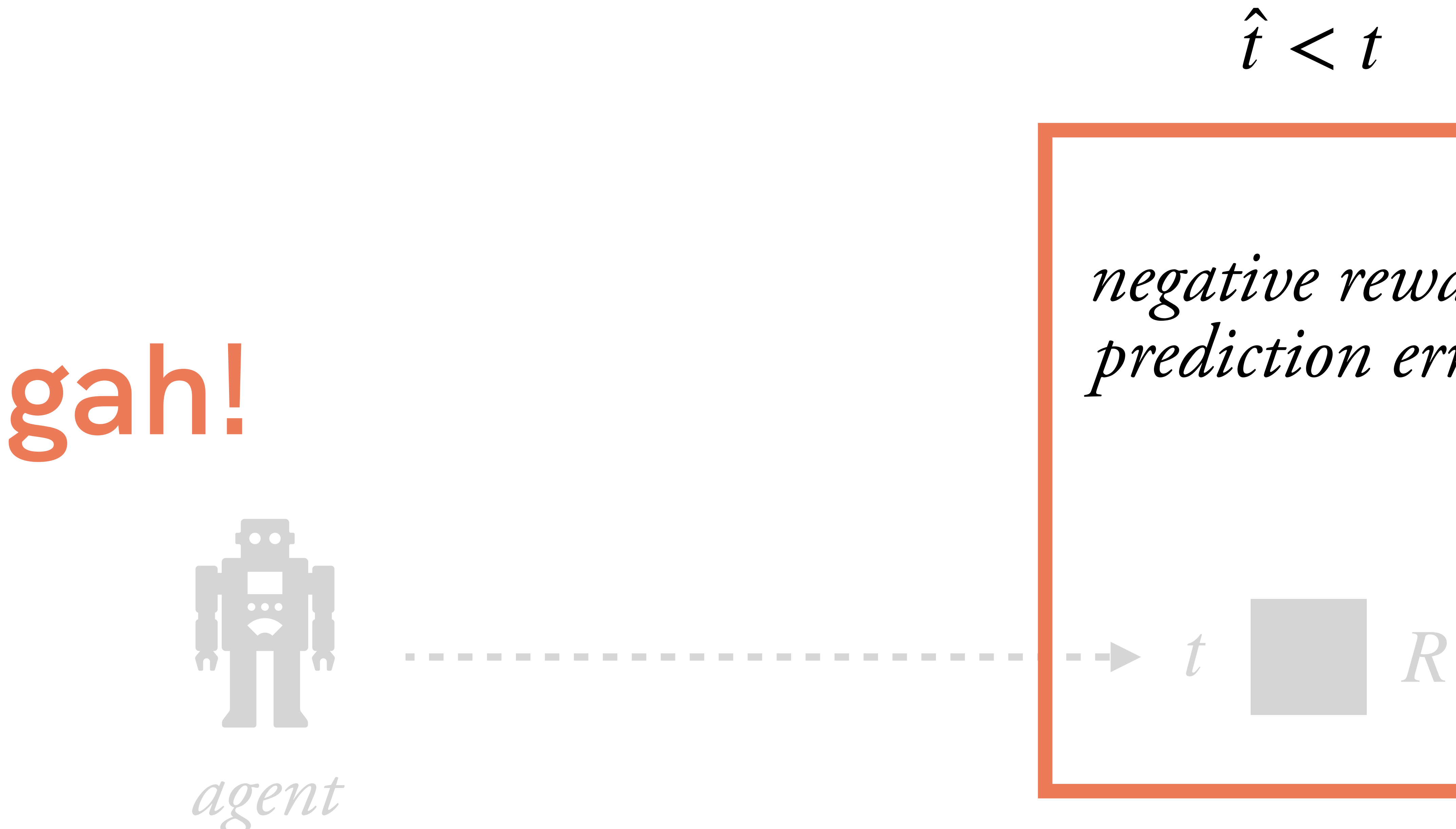



$\hat{t} > t$
positive reward prediction error
aha!
t
R
agent

agent
t



$\hat{t} < t$
negative rewa
prediction err
t
R
agent



$\hat{t} < t$
negative rewa
prediction err
gah!
agent
t
R

constraints

time predictions

constraints

time predictions

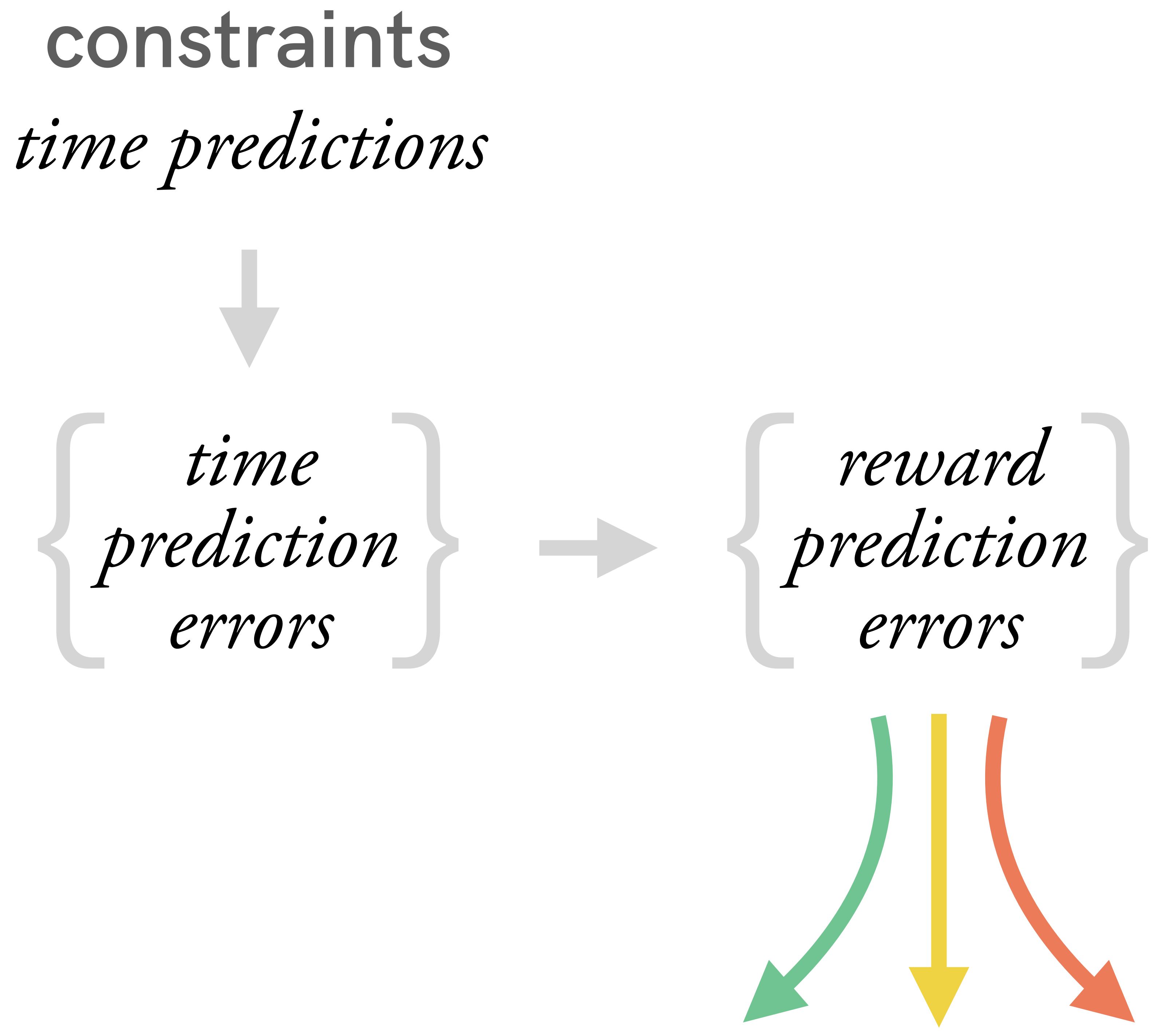
time prediction errors

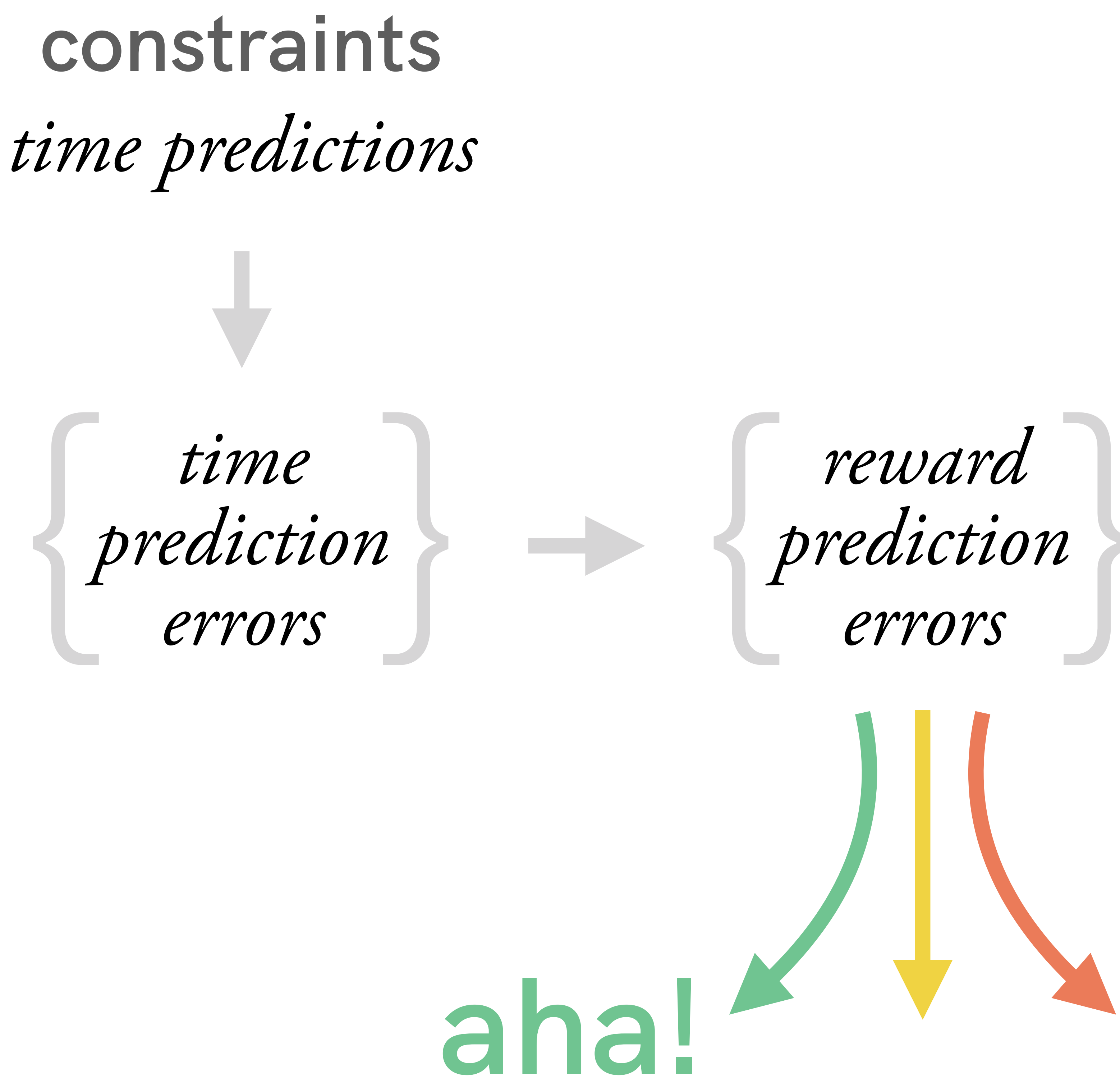
constraints

time predictions

↓

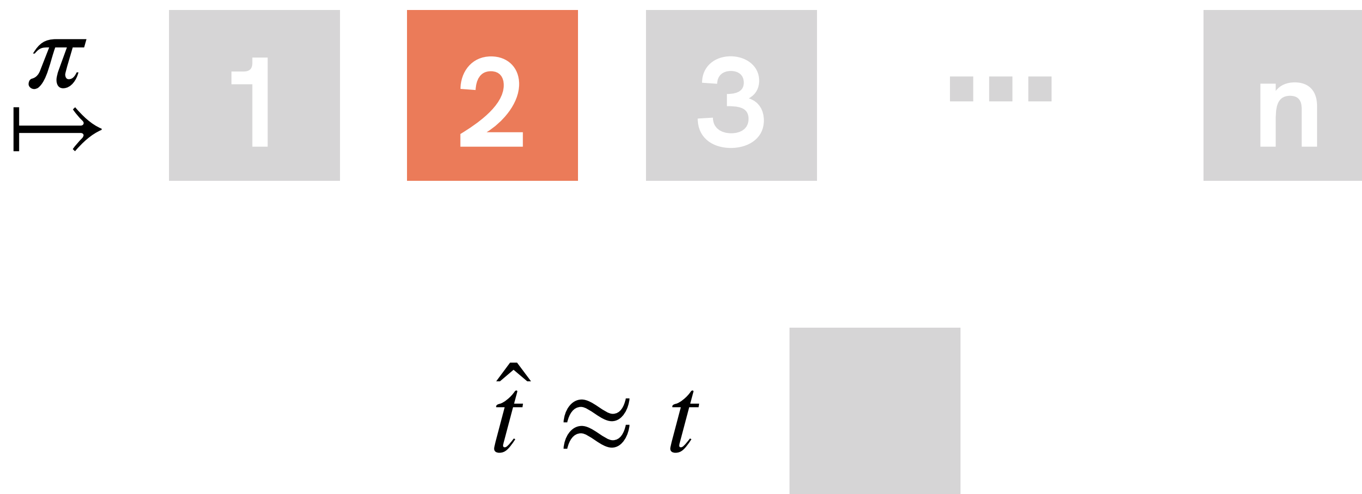
{ *time prediction errors* } → { *reward prediction errors* }


constraints
time predictions
time prediction errors
reward prediction errors

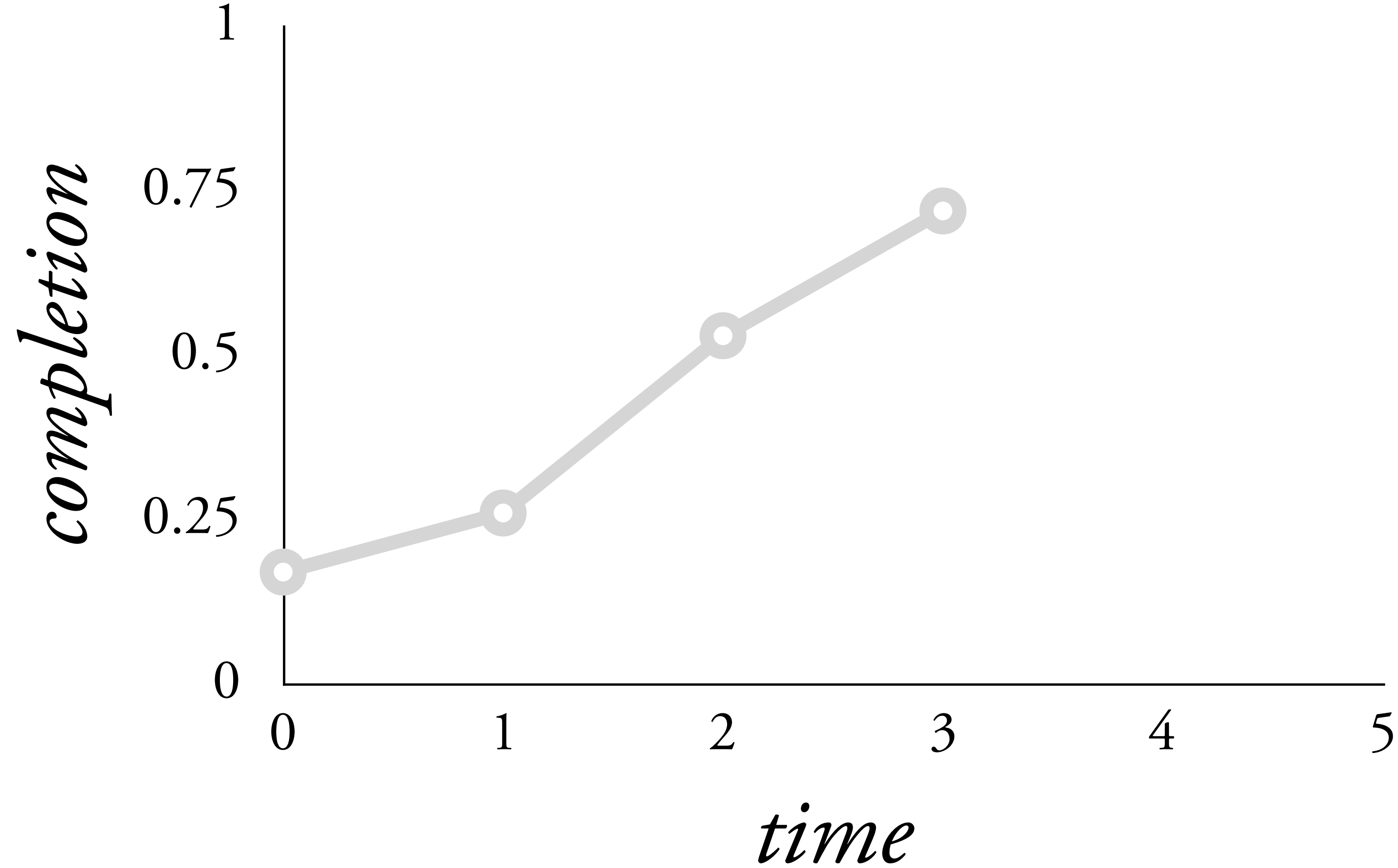

constraints
time predictions
time prediction errors
reward prediction errors
aha!

key idea

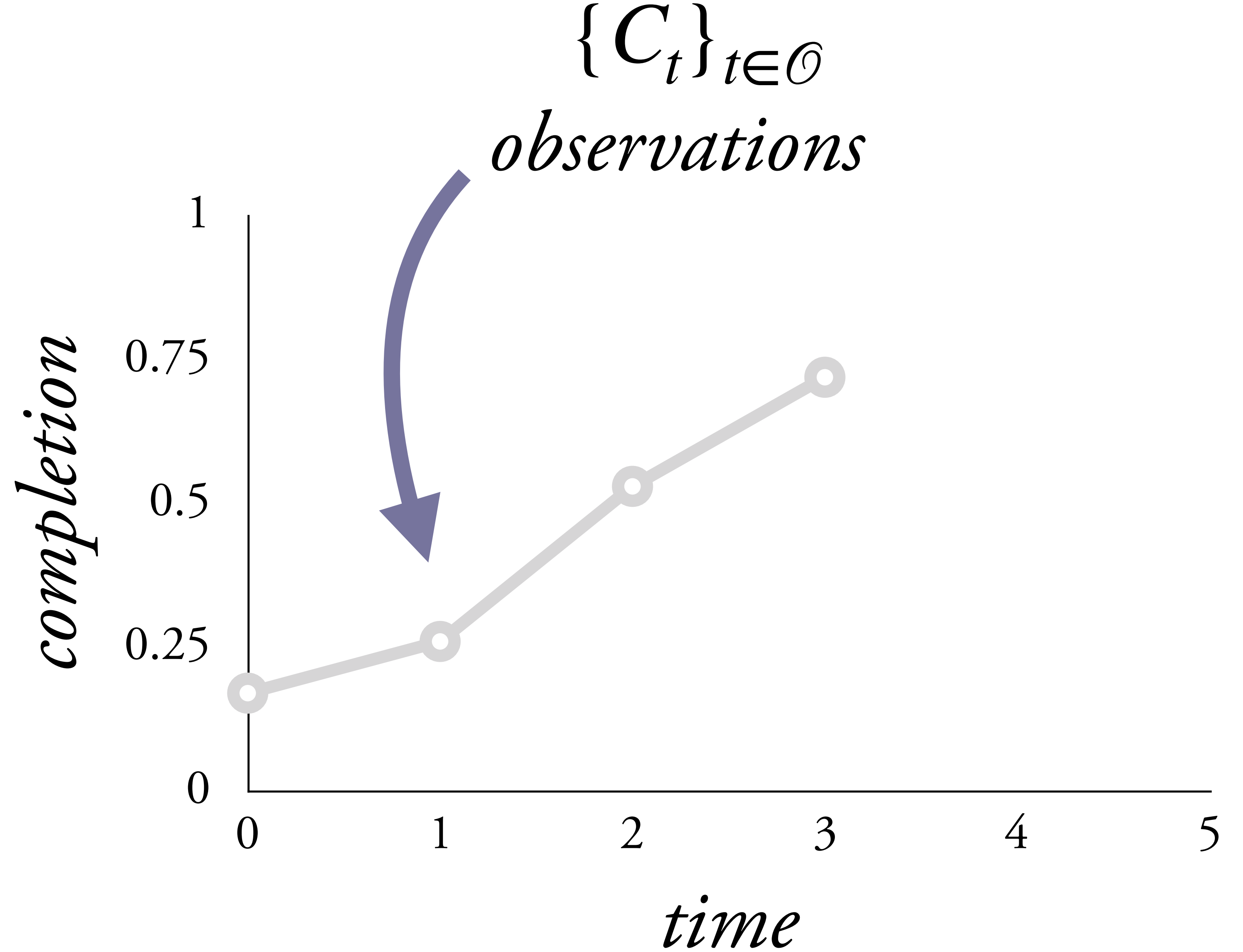
Aha! moments arise under resource constraints and are meta-cognitive prediction errors which cause reward prediction errors.



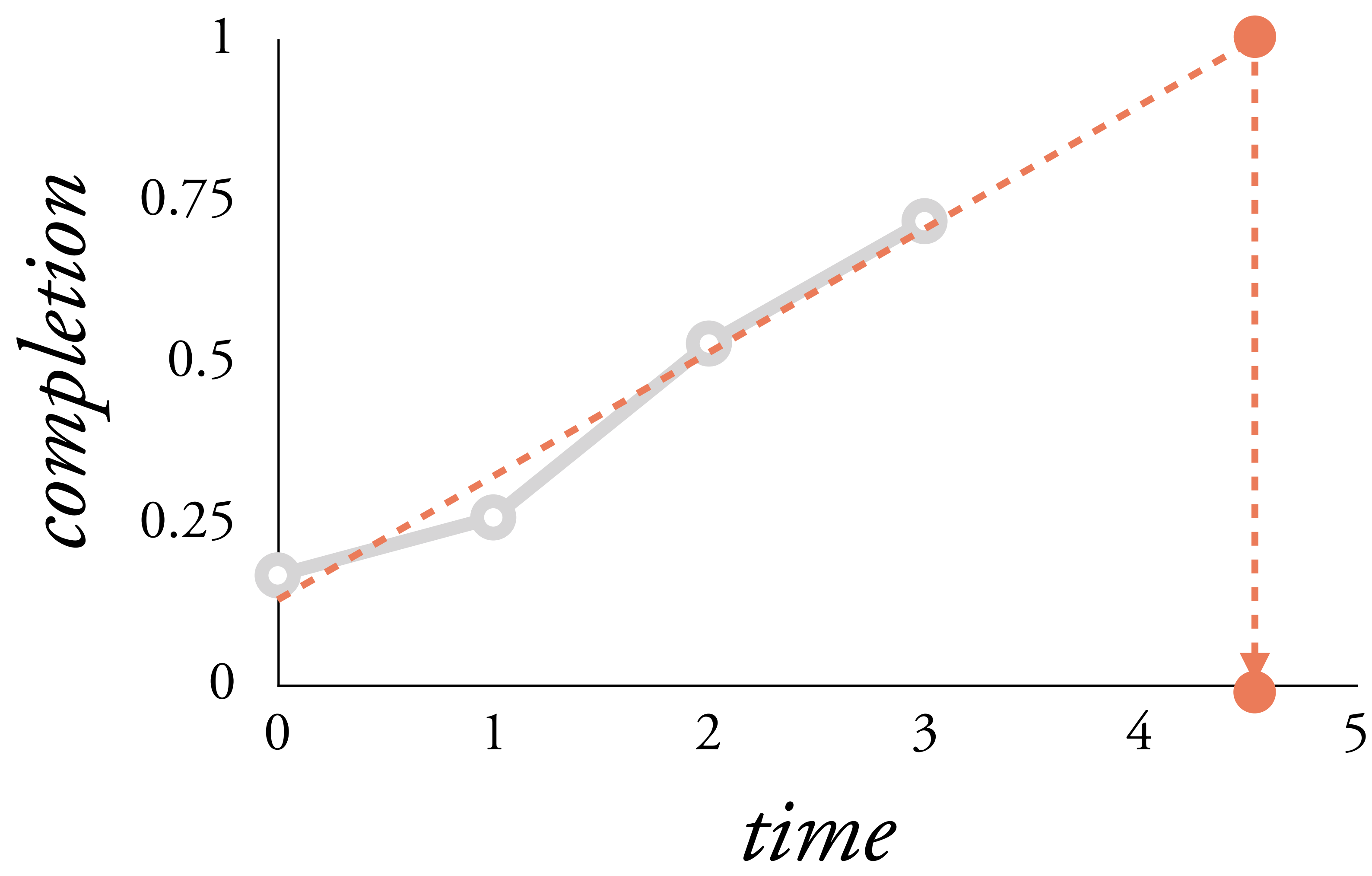
Our framework leaves out two essential pieces: task selection and time estimation.



completion
1
0.75
0.5
0.25
0
0
1
2
3
4
5
time



$\{C_t\}_{t\in\mathcal{O}}$
observations
completion
time
0
0.25
0.5
0.75
1
0
1
2
3
4
5


1
0.75
0.5
0.25
0
completion
0
1
2
3
4
5
time

step	*completion*
d_1	$c_1 = 0.05$
d_2	$c_2 = 0.15$
d_3	$c_3 = 0.24$
	$\vdots$
d_k	$c_k = 0.67$

step	*completion*
d_1	$c_1 = 0.05$
d_2	$c_2 = 0.15$
d_3	$c_3 = 0.24$
	$\vdots$
d_k	$c_k = 0.67$

Understanding increases by some random variable

$$X \sim f_X(\,\cdot\,;\theta)\,.$$

step	*completion*
d_1	$c_1 = 0.05$
d_2	$c_2 = 0.15$
d_3	$c_3 = 0.24$
	$\vdots$
d_k	$c_k = 0.67$

Understanding increases by some random variable

$$X \sim f_X(\,\cdot\,;\theta)\,.$$

goal

Estimate time to completion.

bayesian estimate

probability given parameters and history…

$$\underline{\Pr(C_m \geq 1)} = \int_{\mathbb{R}^d} \underline{\Pr(C_m \geq 1 \,|\, \theta, \mathbf{C} = \mathbf{c})} \cdot \underline{f_{\boldsymbol{\Theta}|\mathbf{C}}(\theta \,|\, \mathbf{c})\, \mathrm{d}\theta}$$

probability completed at step m

…weighted by posteriors

predictions

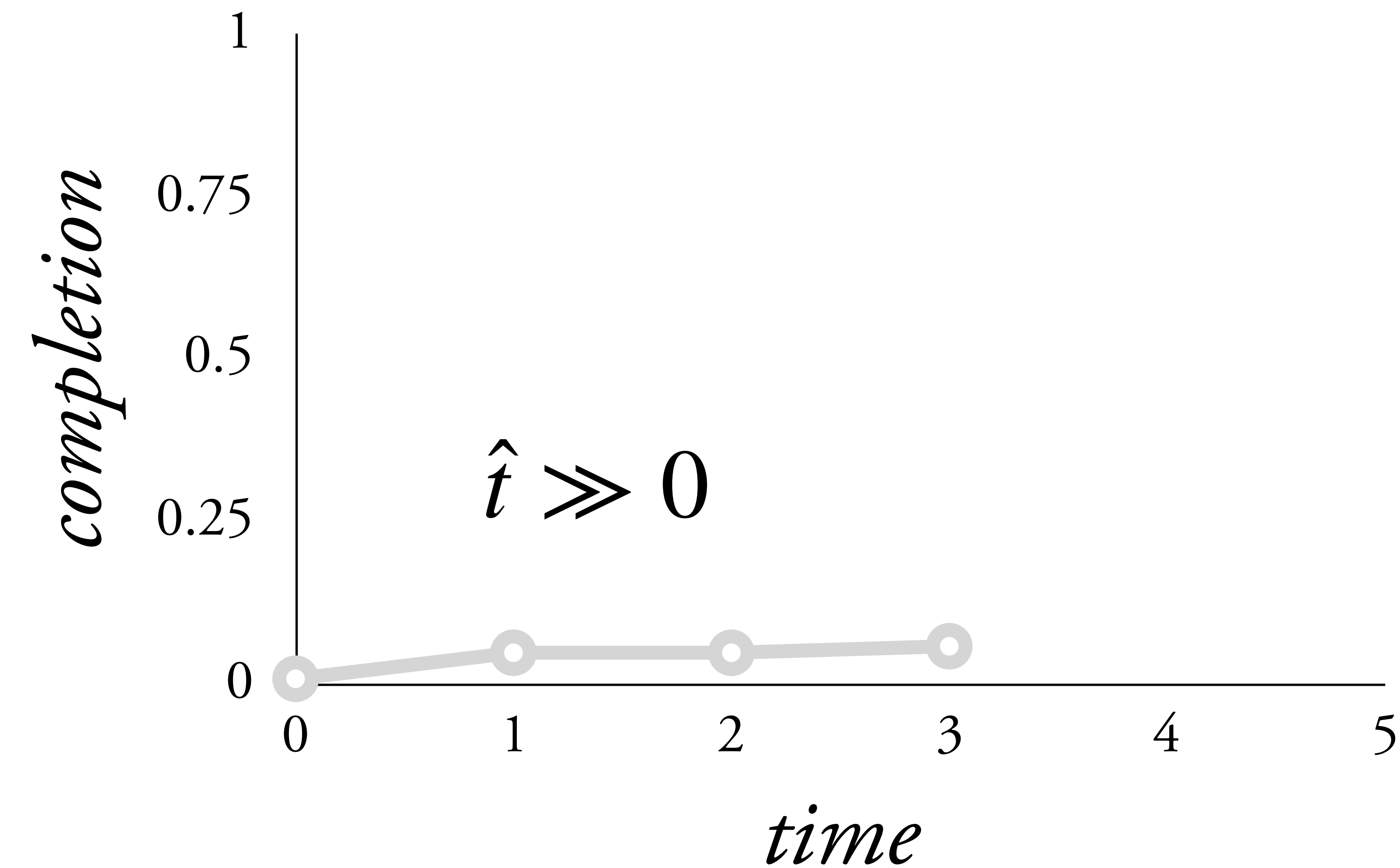
1. *Aha!* moments emerge when an impasse is suddenly broken.
2. *Aha!* moments do not emerge when learning occurs gradually.

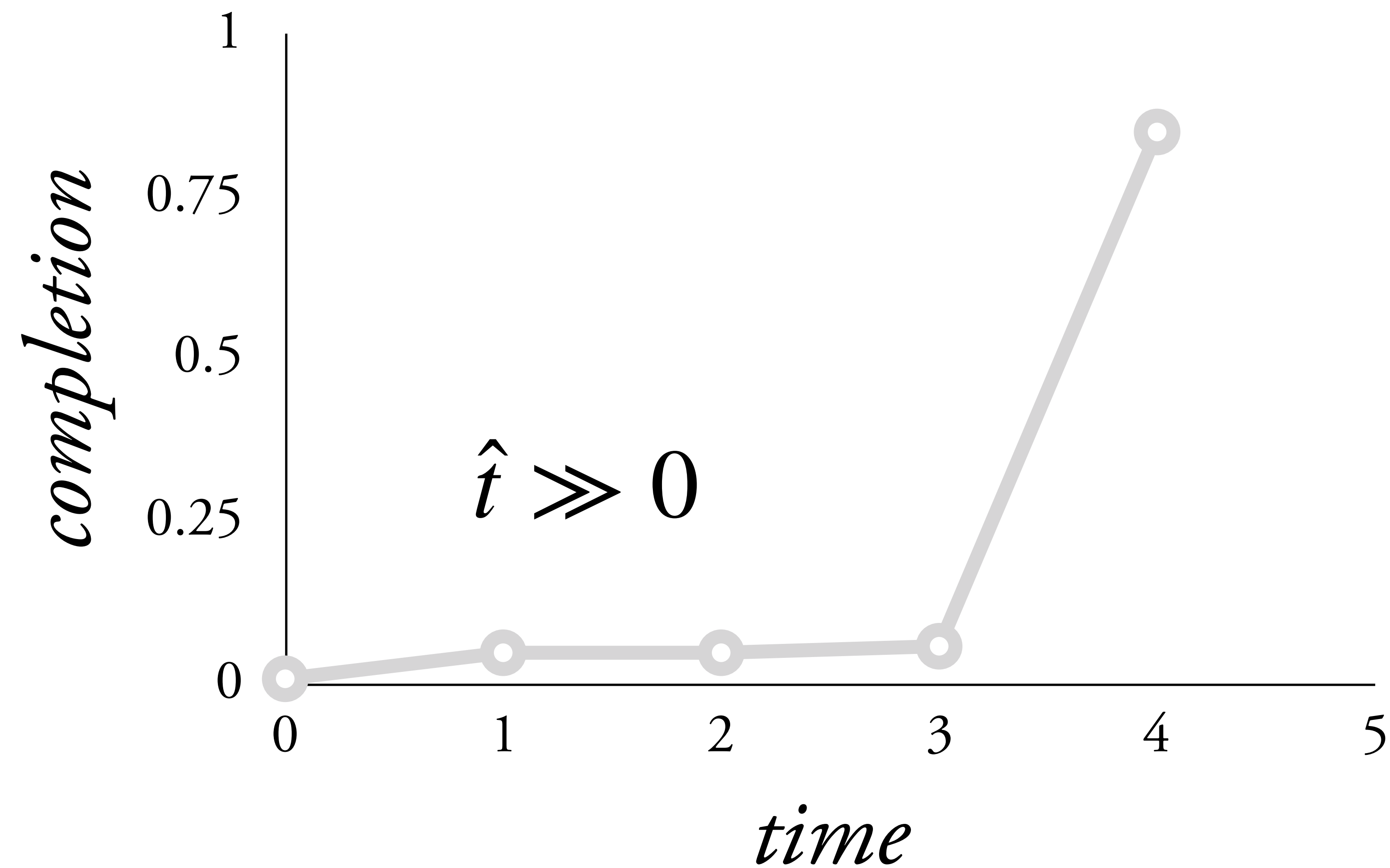
step	*completion*
d_1	$c_1 = 0.01$
d_2	$c_2 = 0.05$
d_3	$c_3 = 0.05$
	$\vdots$
d_k	$c_k = 0.06$

step	*completion*
d_1	$c_1 = 0.01$
d_2	$c_2 = 0.05$
d_3	$c_3 = 0.05$
	$\vdots$
d_k	$c_k = 0.06$

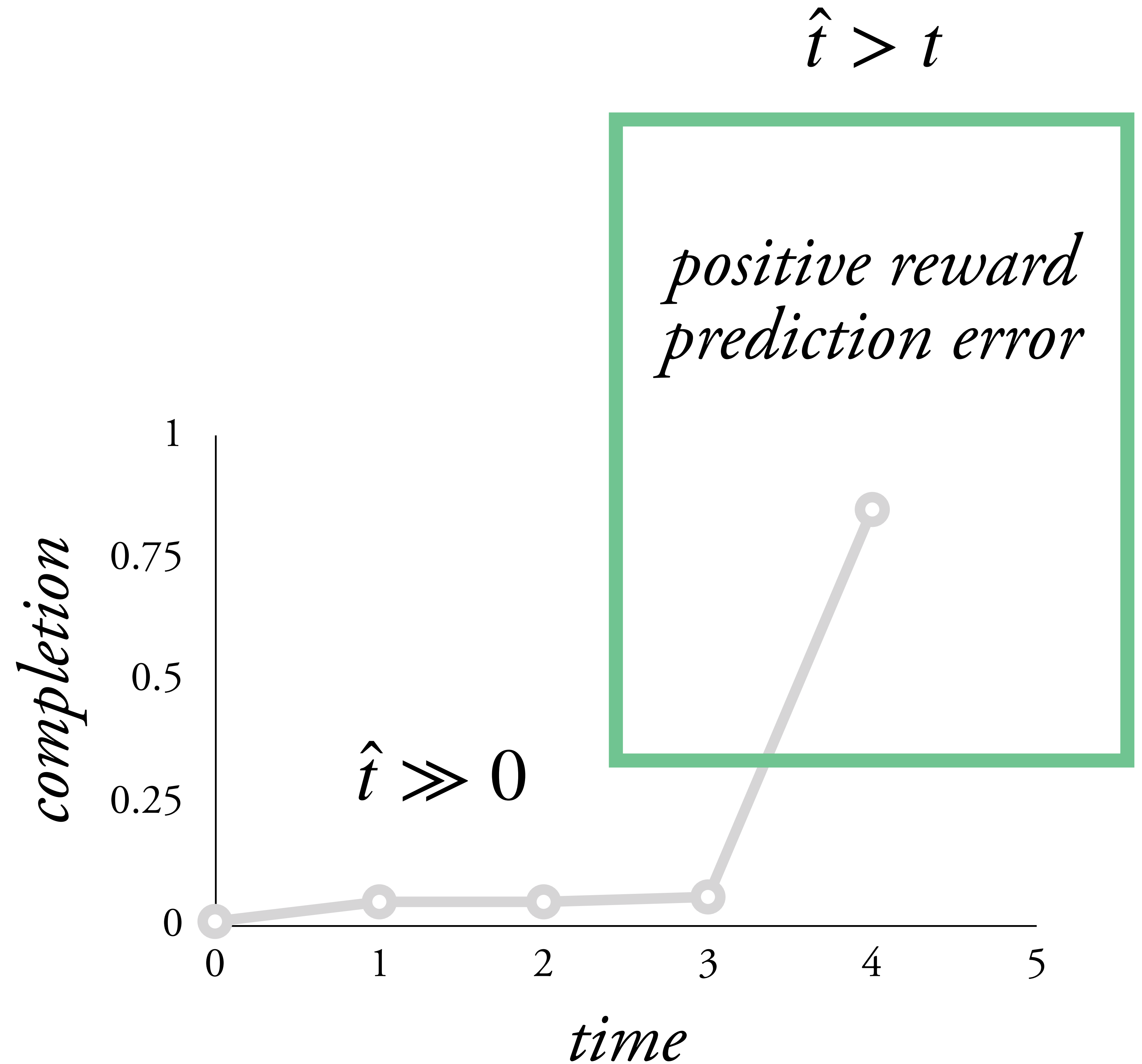
estimate

$$\hat{t} \gg 0$$

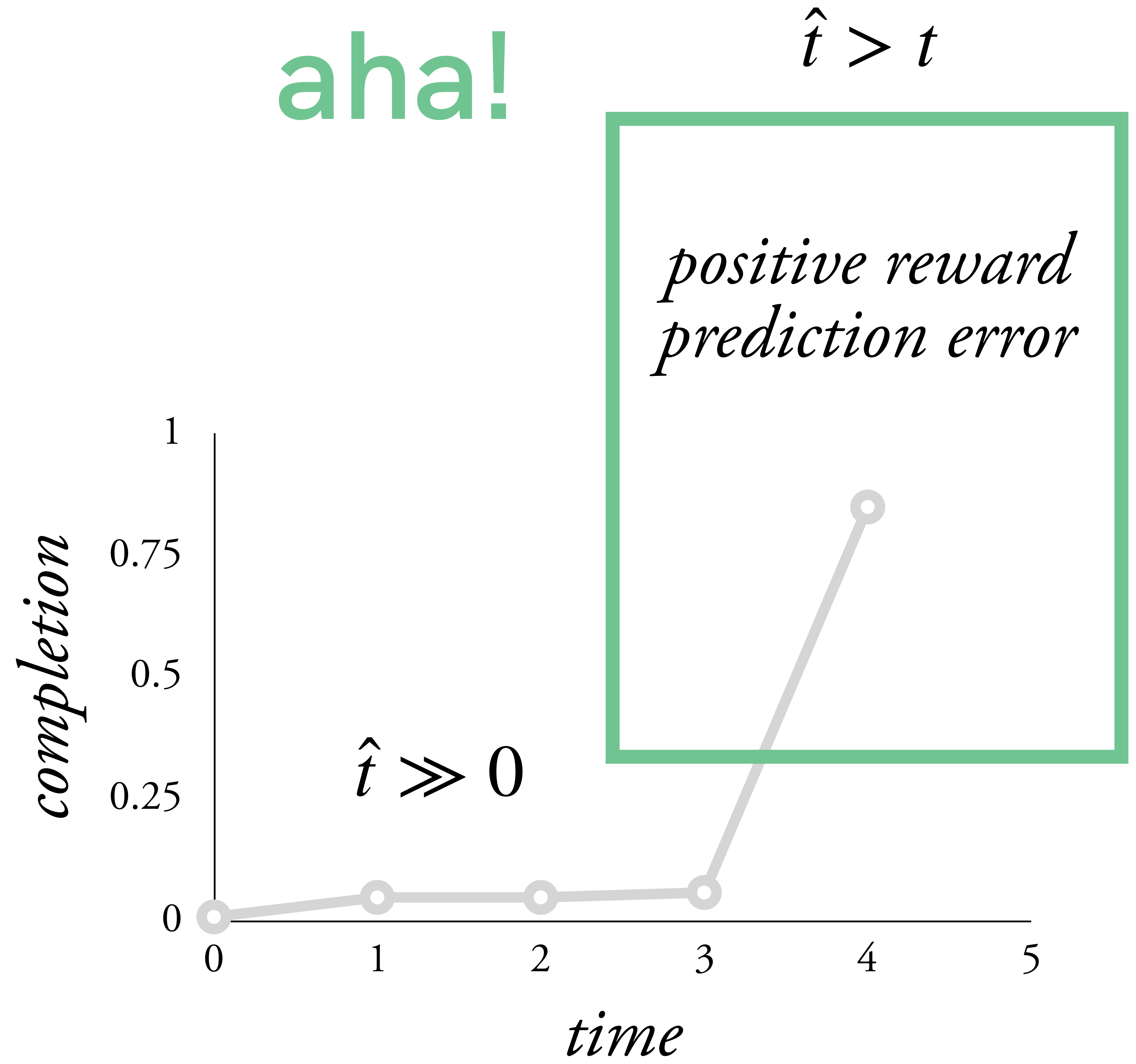

completion
1
0.75
0.5
0.25
0
$\hat{t} \gg 0$
0
1
2
3
4
5
time



1
0.75
0.5
0.25
0
completion
$\hat{t} \gg 0$
0
1
2
3
4
5
time



$\hat{t} > t$
positive reward prediction error
$\hat{t} \gg 0$
completion
1
0.75
0.5
0.25
0
0
1
2
3
4
5
time

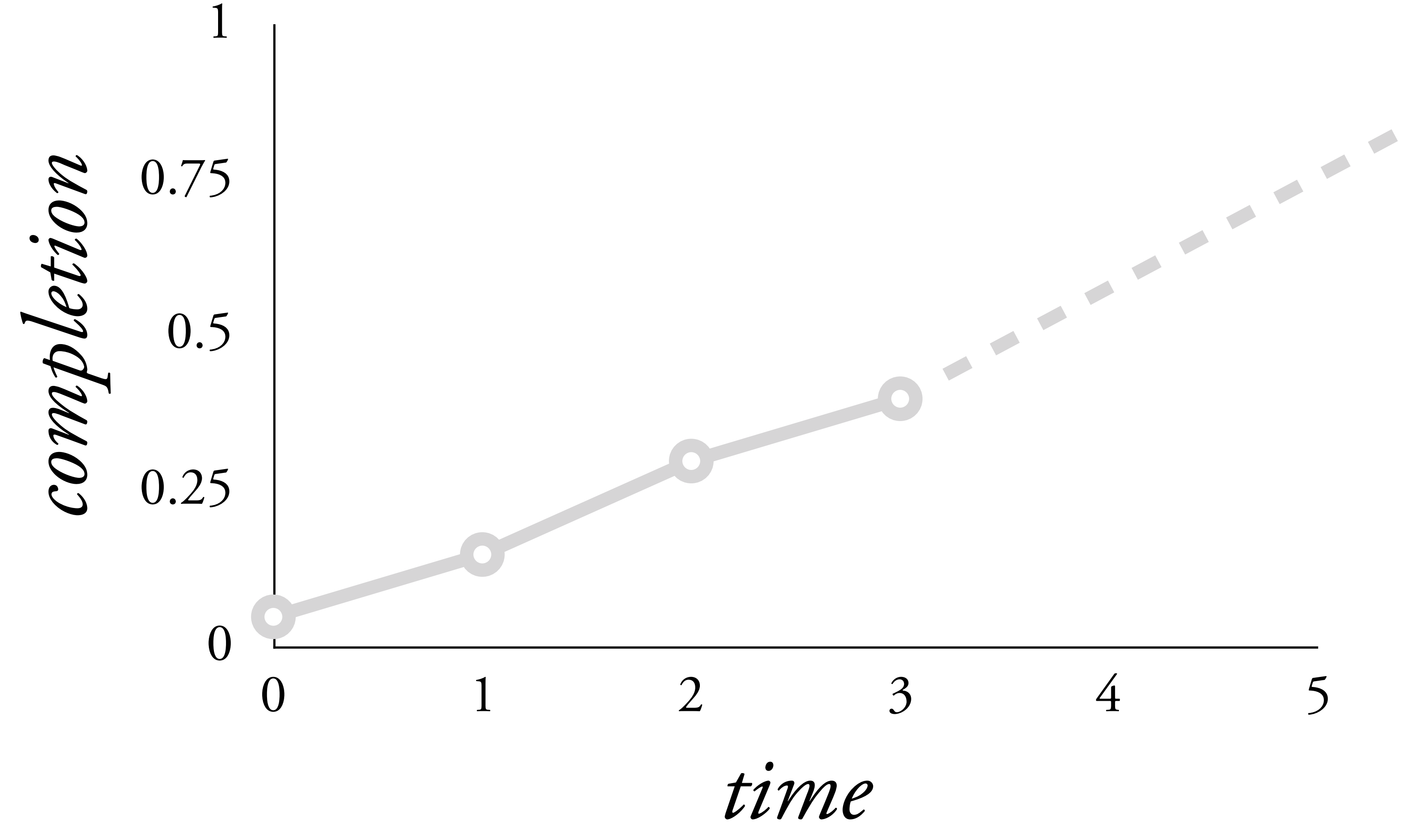

aha!
$\hat{t} > t$
positive reward
prediction error
$\hat{t} \gg 0$
completion
1
0.75
0.5
0.25
0
0
1
2
3
4
5
time

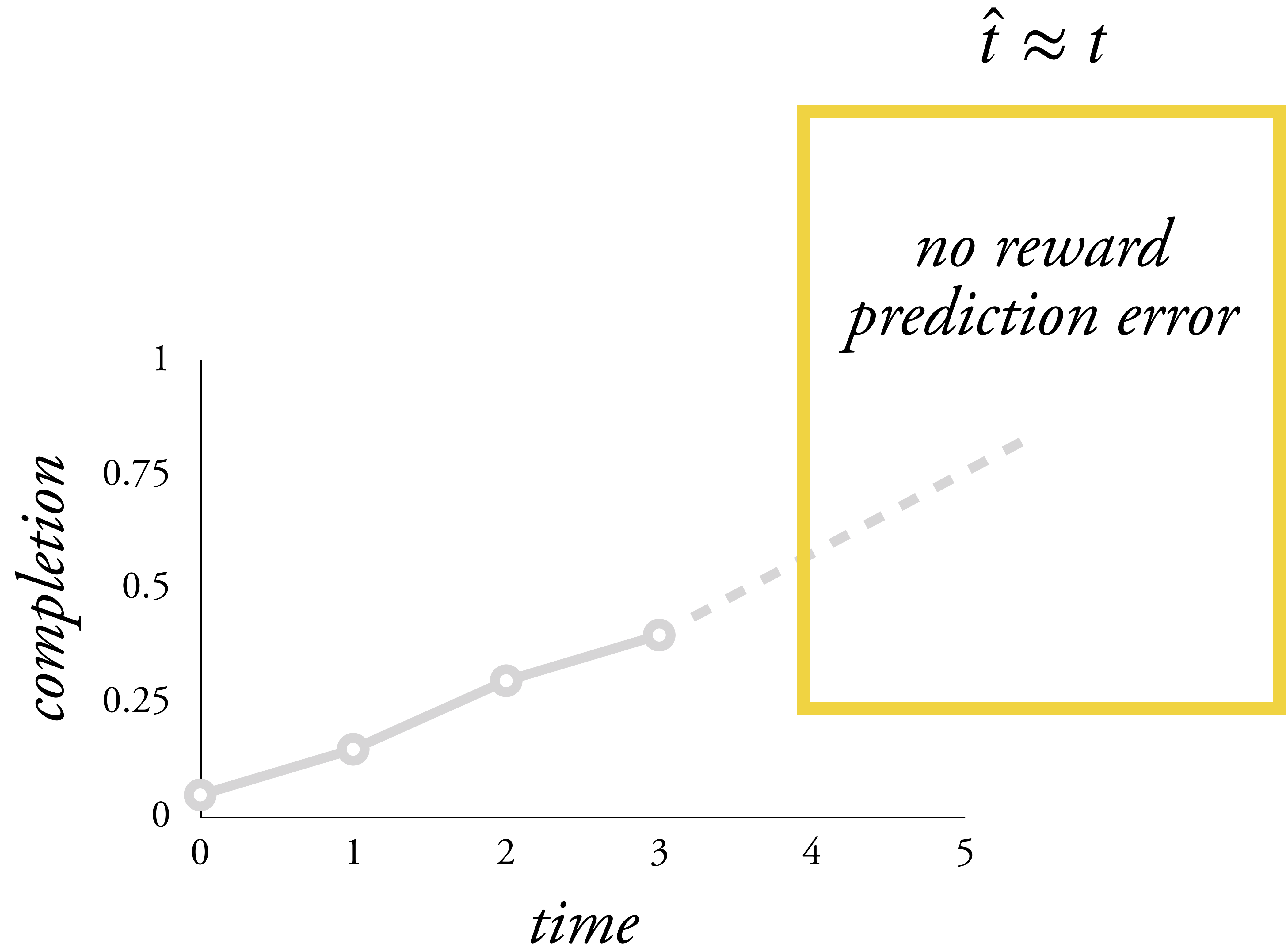
step	*completion*
d_1	$c_1 = 0.05$
d_2	$c_2 = 0.15$
d_3	$c_3 = 0.30$
	$\vdots$
d_k	$c_k = 0.4$

step	*completion*
d_1	$c_1 = 0.05$
d_2	$c_2 = 0.15$
d_3	$c_3 = 0.30$
	$\vdots$
d_k	$c_k = 0.4$

estimate

$$\hat{t} \approx t$$


completion
1
0.75
0.5
0.25
0
0
1
2
3
4
5
time



$\hat{t} \approx t$
no reward prediction error
completion
1
0.75
0.5
0.25
0
0
1
2
3
4
5
time

[Danek, Wiley & Öllinger 2016]

other results

3. *Aha!* moments will be influenced by our prior expectations and are a product of our subjective experience.

[Friedlander & Fine 2018]

other results

4. *Aha!* moments will be highest after we have spent considerable time on a problem and have been stuck on it for a while.

Experimental Results

question

Can solely manipulating people's perceived difficulty of a task influences their experience of an *Aha!* moment?

Participants were given 10 seconds for each of 5 anagrams.

SROT

anagram 1 of 5

Each anagram was immediately followed with an *Aha!* self-report.

Did you experience an *Aha!* moment (on a scale of 1-7) after the **last** anagram was shown to you?

anagram 1 of 5

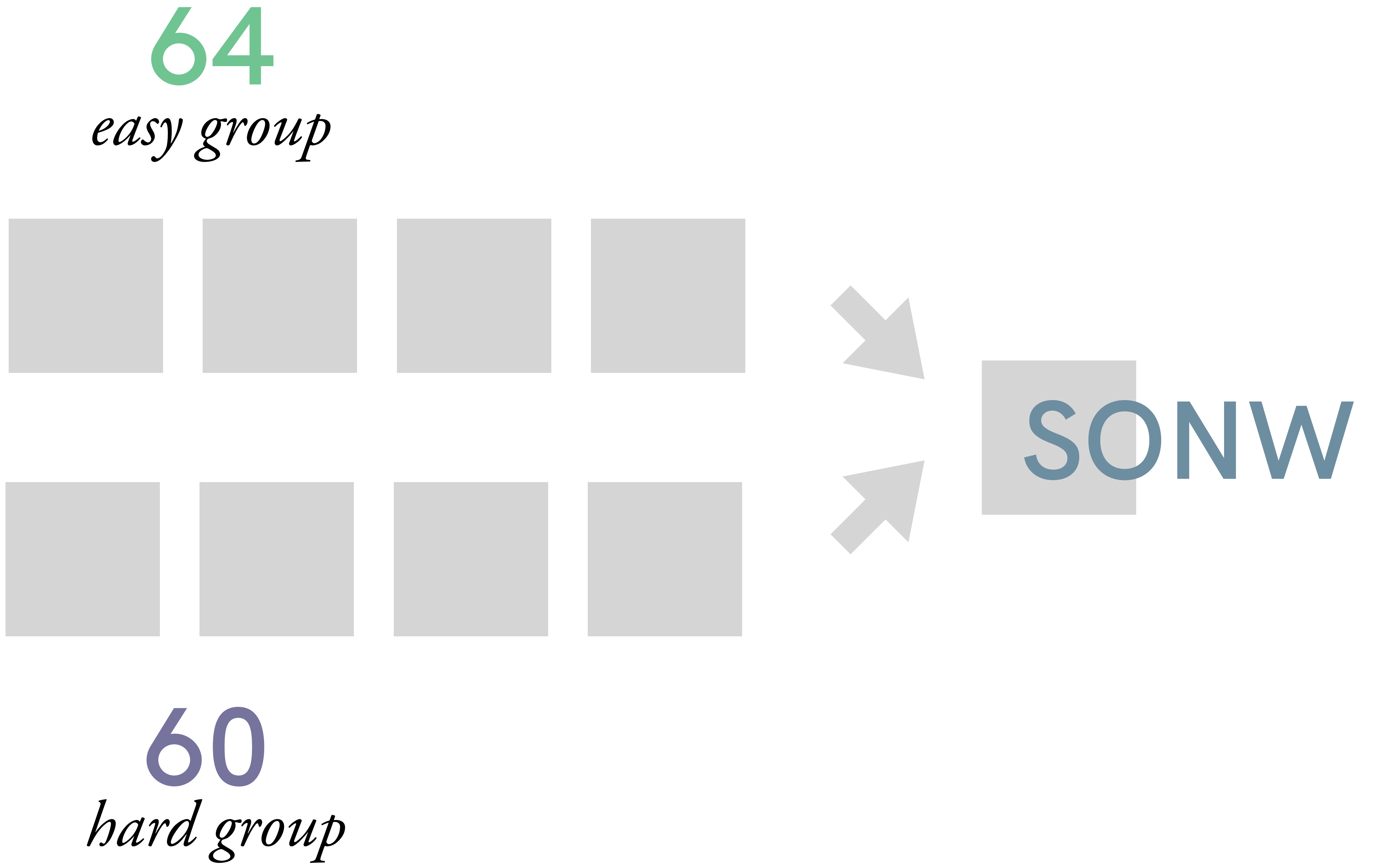

124
participants
64
easy group
60
hard group

64

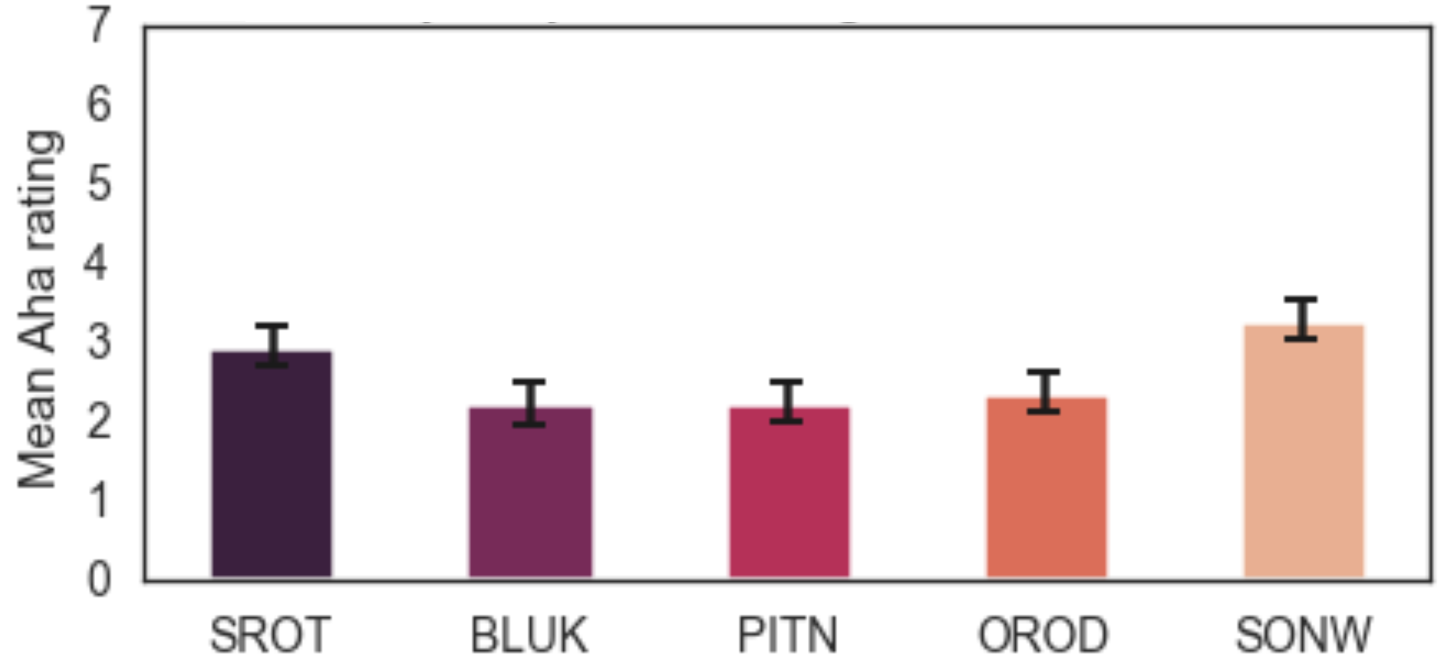
easy group

60

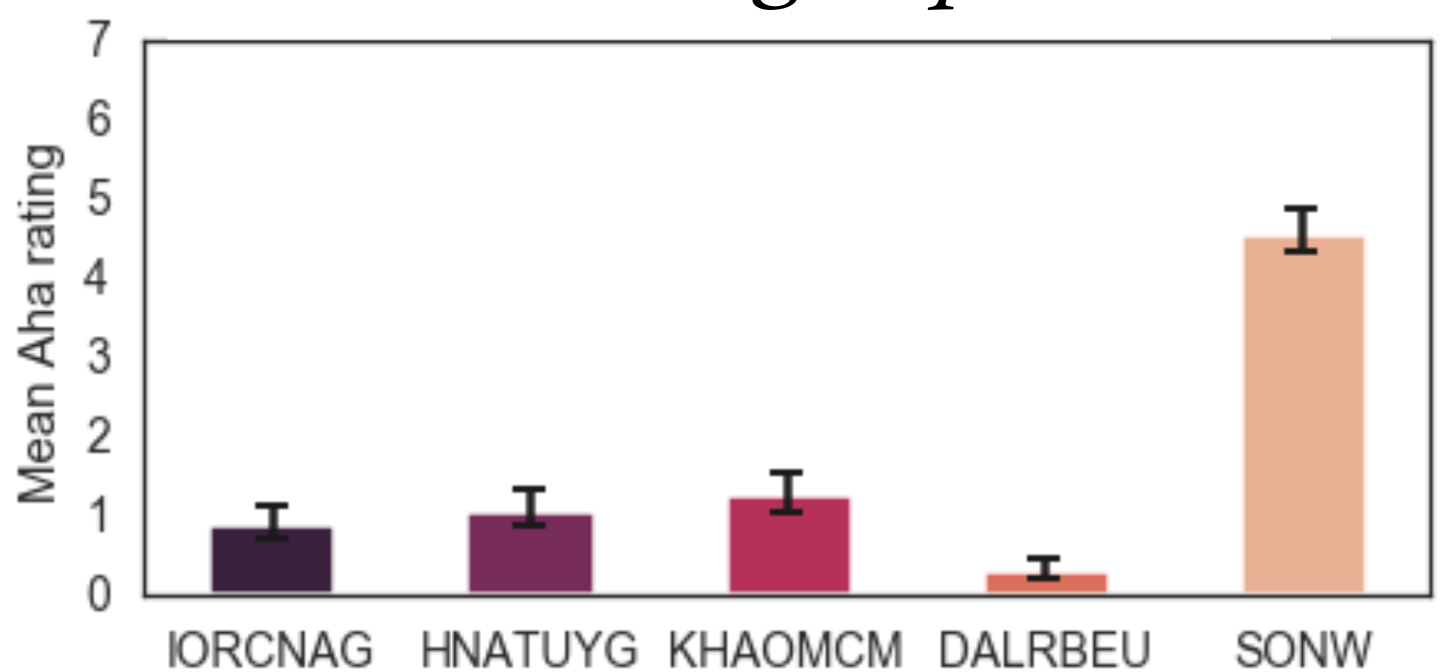
hard group


64
easy group
60
hard group
SONW

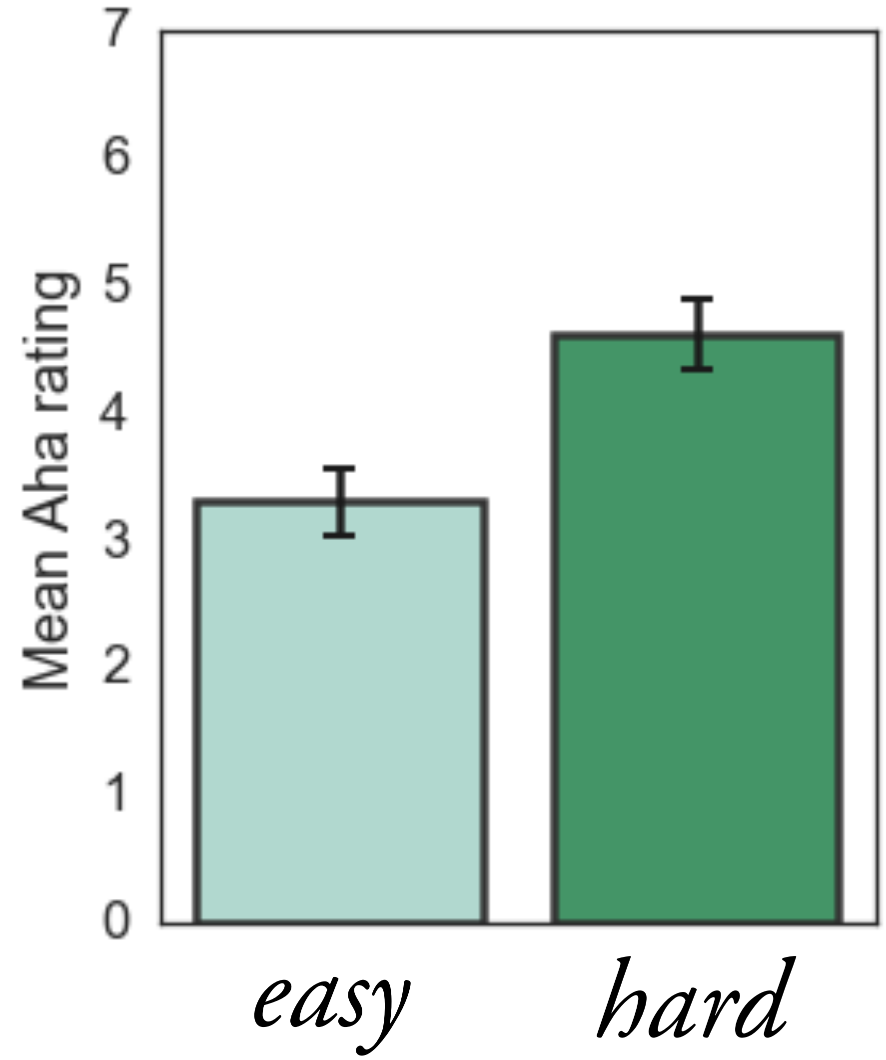
easy group



hard group



($t = 3.3$, $p < 0.01$)



Parting Thoughts

questions

1. Why do we have *Aha!* moments?
2. What makes these moments so rare?
3. Why are they so rewarding?

2. What makes these moments so rare?
   Positive time prediction errors are rare (planning fallacy).

further work

We are currently working on more robust behavioral experiments to test our theory.

Aha! Moments

as meta-cognitive prediction errors

Aha! Moments

as meta-cognitive prediction errors

in collaboration with
Rachit Dubey, Mark Ho, Tom Griffiths
Princeton University
University of California, Berkeley