PLOS | COMPUTATIONAL BIOLOGY

RESEARCH ARTICLE

Contextual influence on confidence judgments in human reinforcement learning

Maël Lebreton[1,2,3,4]*, Karin Bacily[1,2], Stefano Palminteri[5,6,7‡], Jan B. Engelmann[1,2,8‡]

Heesun Park
2020.08.20

Introduction

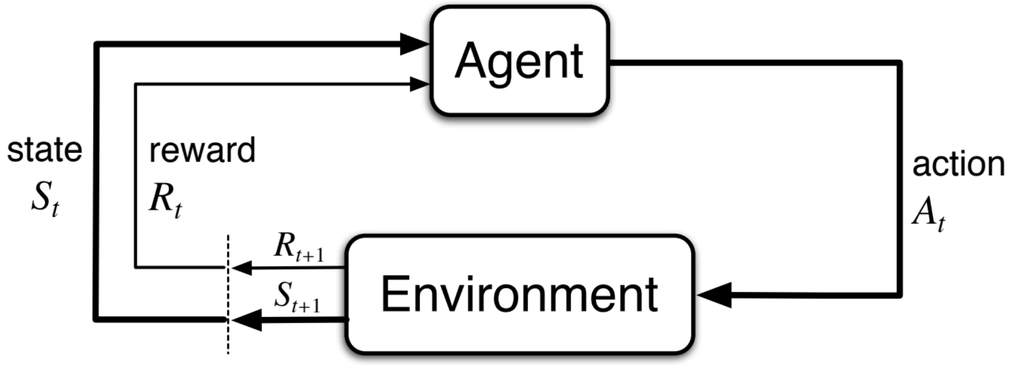
Metacognition and Confidence

- ***Metacognition***
  - Cognition about cognition
  - Often measured by introspective self-reports (e.g. confidence ratings)
- ***Confidence while Learning***
  - Confidence: a decision-maker's <u>*estimation of probability of being correct*</u> (Pouget et al 2016)
  - Comparisons and adjustments of *learning strategies*
  - Confidence bias

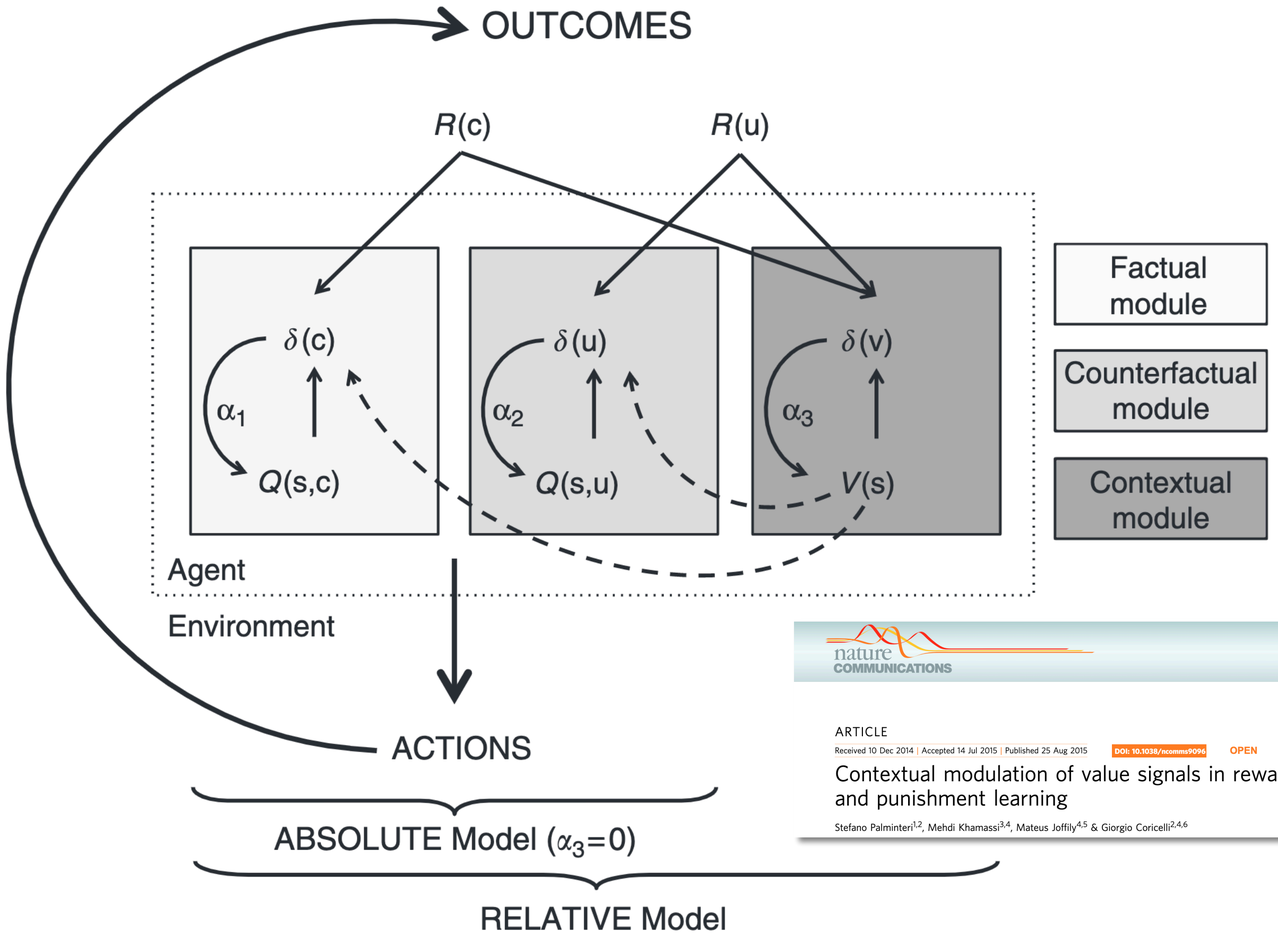
Norman, E., et al 2019

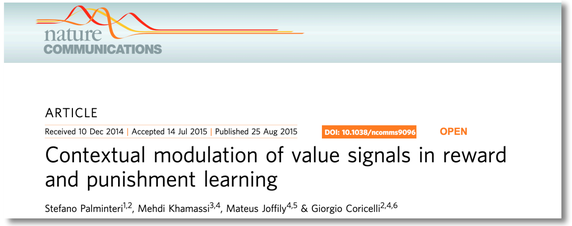
Introduction

Context-Dependent Reinforcement Learning



Sutton & Barto 1998




nature COMMUNICATIONS

ARTICLE

Received 10 Dec 2014 | Accepted 14 Jul 2015 | Published 25 Aug 2015 DOI: 10.1038/ncomms9096 OPEN

Contextual modulation of value signals in reward and punishment learning

Stefano Palminteri[1,2], Mehdi Khamassi[3,4], Mateus Joffily[4,5] & Giorgio Coricelli[2,4,6]

Palminteri, S., et al 2015

Hypotheses

1. ***Confidence bias***

(↓ Confidence in ***loss-avoidance*** than gain-seeking learning context)

2.

Context-Value

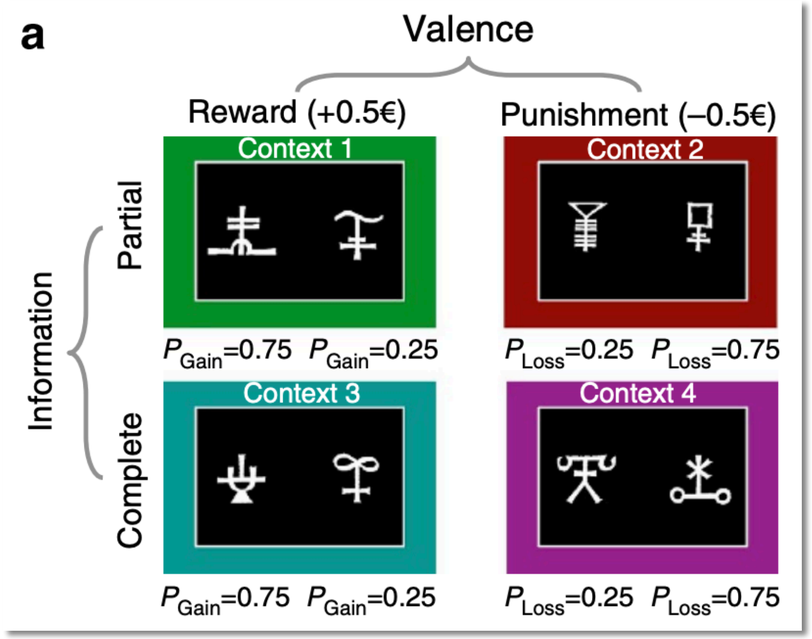
3.

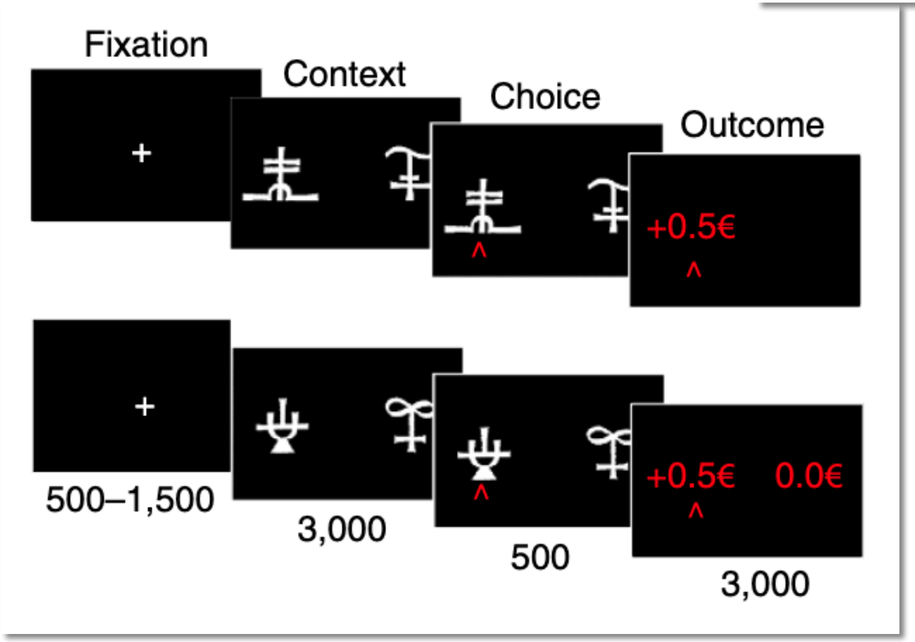
Performance in volatile environment

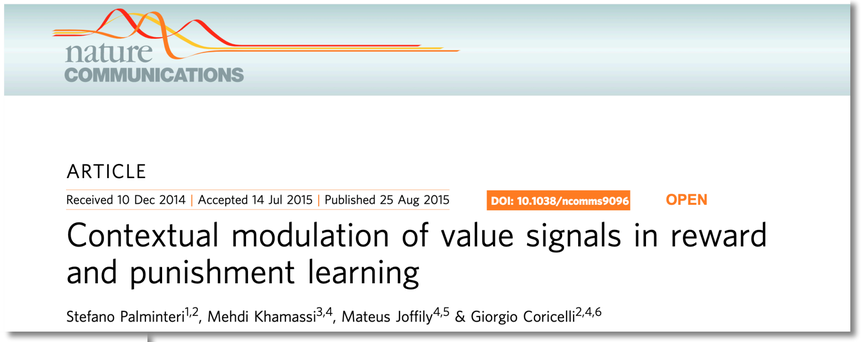
Methods

Methods

Probabilistic Instrumental-Learning task







Palminteri, S., et al 2015

Methods

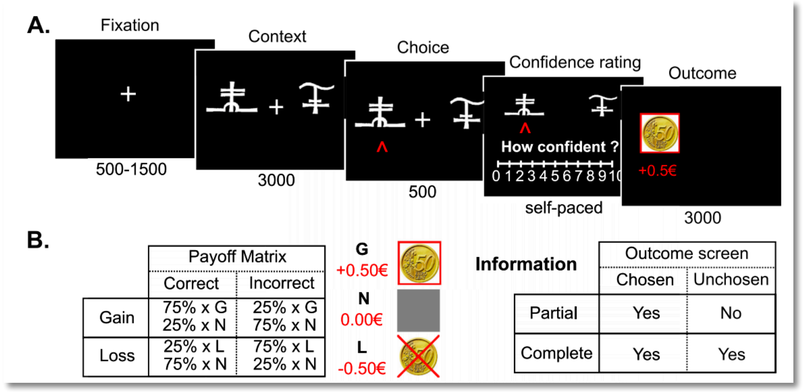
Experimental Design

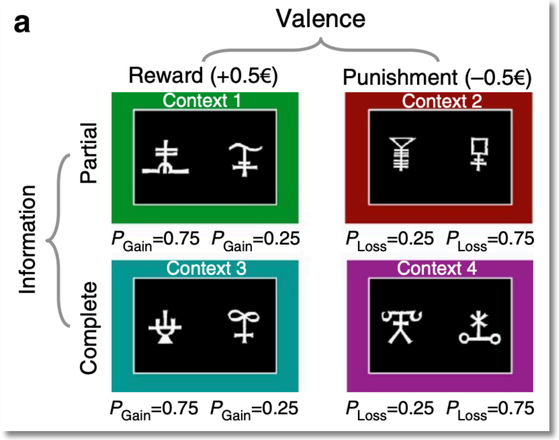
1. ***Confidence bias***
(↓ Confidence in ***loss-avoidance*** than gain-seeking learning context)

2. ***Context-Value***

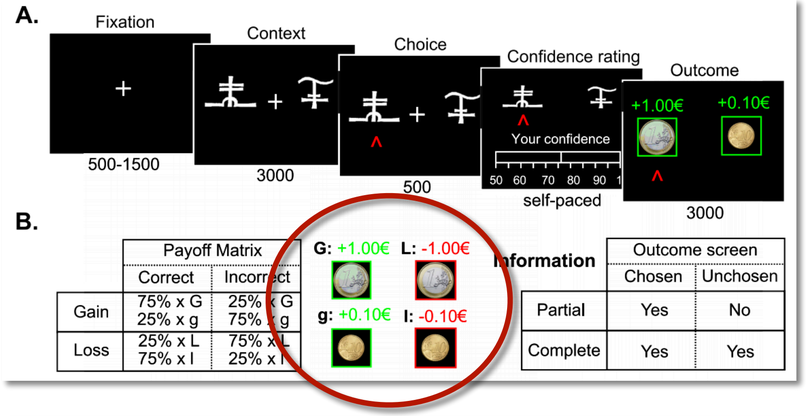
3. ***Performance in volatile environment***

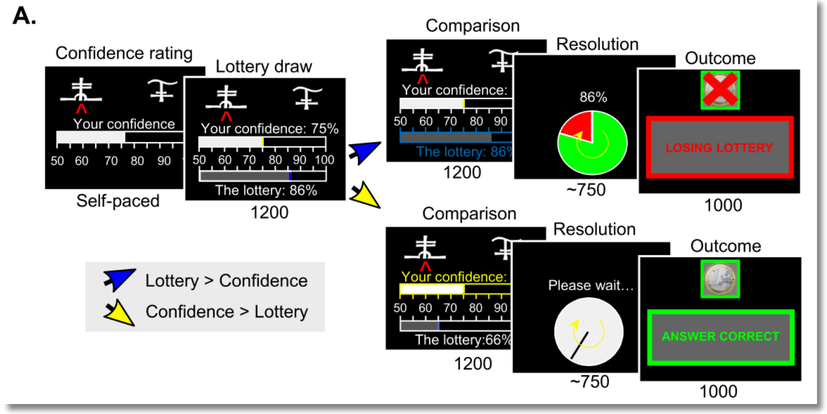
Experiment 1





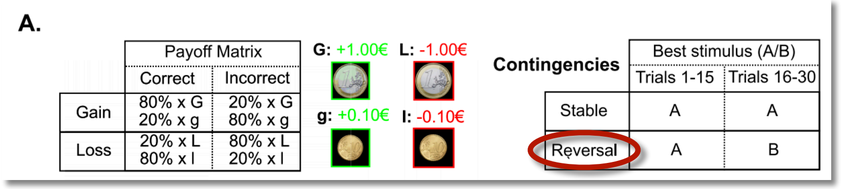
Experiment 2





Matching Probability (MP) Mechanism

Experiment 3



Methods

Modeling - RL Model

Absolute-value based learning	*Relative-value based learning*
ABSOLUTE Model	***RELATIVE Model***
$Q_{t+1}(s,c) = Q_t(s,c) + \alpha_c \delta_{c,t}$ *Learning rate* *Prediction error* $Q_{t+1}(s,u) = Q_t(s,u) + \alpha_u \delta_{u,t}$	$Q_{t+1}(s,c) = Q_t(s,c) + \alpha_c \delta_{c,t}$ $Q_{t+1}(s,u) = Q_t(s,u) + \alpha_u \delta_{u,t}$
$\delta_{c,t} = R_{c,t} - Q_t(s,c)$ *Partial & Complete* $\delta_{u,t} = R_{u,t} - Q_t(s,u)$ *Complete only*	$\delta_{c,t} = R_{c,t} - V_t(s) - Q_t(s,c)$ $\delta_{u,t} = R_{u,t} - V_t(s) - Q_t(s,u)$
	$V_{t+1}(s) = V_t(s) + \alpha_V \delta_{V,t}$ $\delta_{V,t} = (R_{c,t} + R_{U,t})/2 - V_t(s),$ *Complete* $\delta_{V,t} = (R_{c,t} + Q_t(s,u))/2 - V_t(s).$ *Partial*

Methods

Modeling - Confidence Model

FULL Model	Context-Value $c_t = \beta_0 + \beta_{dQ} \times \Delta Q_t + \beta_V \times V_t + \beta_{c1} \times c_{t-1}$ Choice difficulty Confidence at the preceding trial
REDUCED Model 1	$c_t = \beta_0 + \beta_{\Delta Q} \times \Delta Q_t + \beta_{c1} \times c_{t-1}$
REDUCED Model 2	$c_t = \beta_0 + \beta_{\Delta Q} \times \Delta Q_t + \beta_V \times V_t$

$$\Delta Q_t = abs(Q_t(s,b) - Q_t(s,a))$$

Results

Results

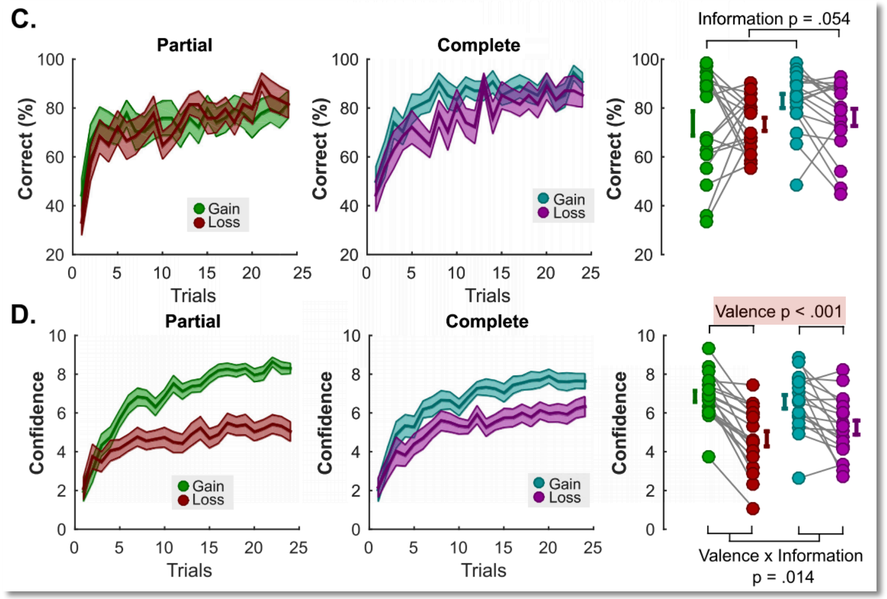
Confidence bias

1. ***Confidence bias***
(↓ Confidence in ***loss-avoidance*** than gain-seeking learning context)

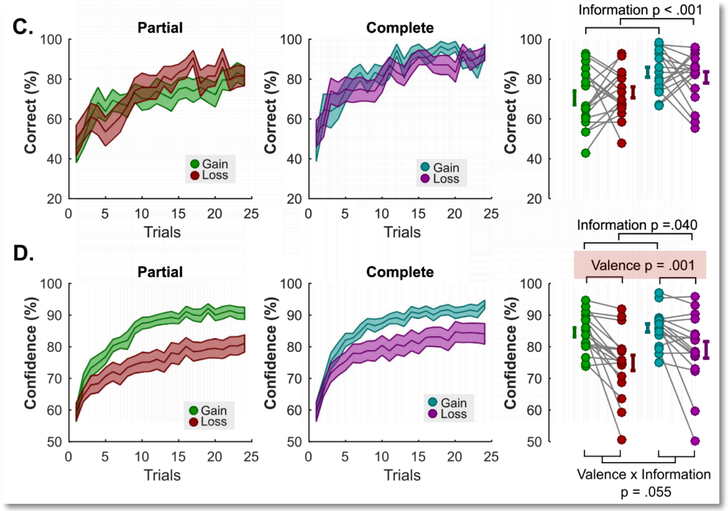
2. ***Context-Value***

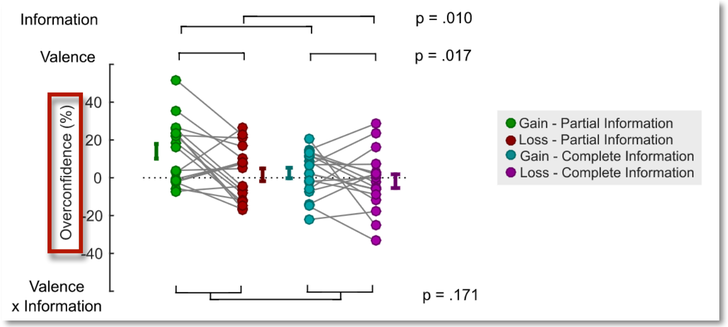
3. ***Performance in volatile environment***

Experiment 1



Experiment 2





Results

Modeling Results - RL

ABSOLUTE model

$$Q_{t+1}(s,c) = Q_t(s,c) + \alpha_c \delta_{c,t}$$
$$Q_{t+1}(s,u) = Q_t(s,u) + \alpha_u \delta_{u,t}$$

$$\delta_{c,t} = R_{c,t} - Q_t(s,c)$$
$$\delta_{u,t} = R_{u,t} - Q_t(s,u)$$

Best-fitting model!

RELATIVE model

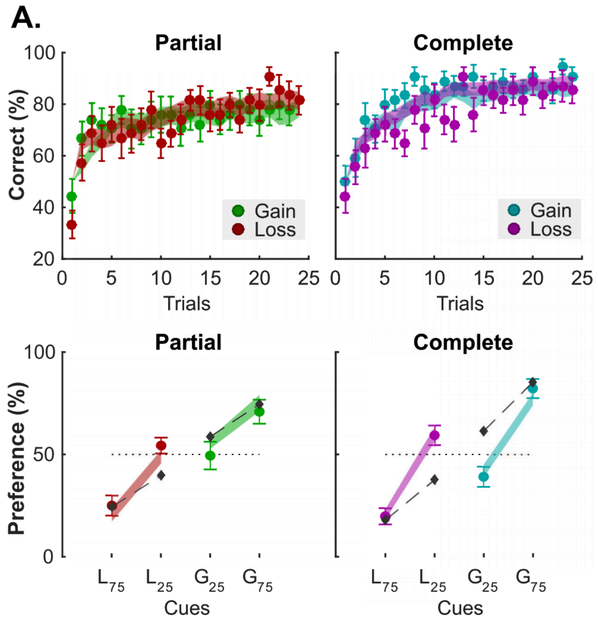
$$Q_{t+1}(s,c) = Q_t(s,c) + \alpha_c(R_{c,t} - (V_t(s) + Q_t(s,c)))$$
$$Q_{t+1}(s,u) = Q_t(s,u) + \alpha_u(R_{u,t} - (V_t(s) + Q_t(s,u)))$$

$$V_{t+1}(s) = V_t(s) + \alpha_V \delta_{V,t}$$
$$\delta_{V,t} = (R_{c,t} + R_{u,t})/2 - V_t(s)$$
$$\delta_{V,t} = (R_{c,t} + Q_t(s,u))/2 - V_t(s)$$

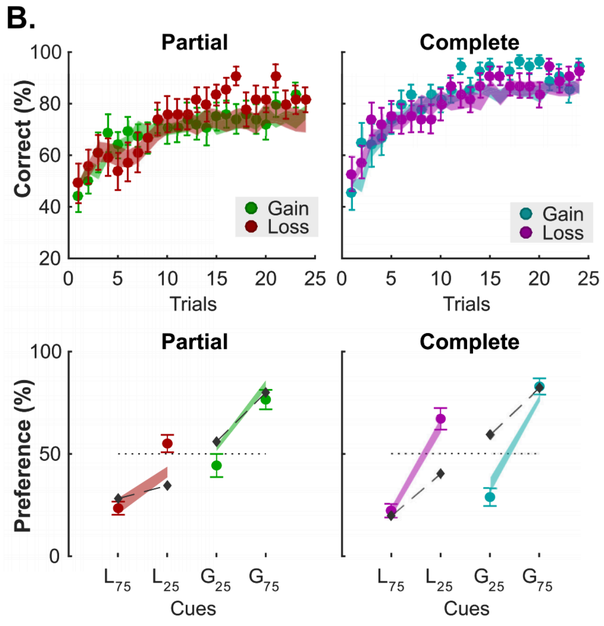
Table 1. Reinforcement-learning. Model comparison. AIC, Akaike Information Criterion (computed with nLL_{max}); BIC, Bayesian Information Criterion (computed with nLL_{max}); DF, degrees of freedom; nLL_{max}, negative log likelihood; $nLPP_{max}$, negative log of posterior probability; EF, expected frequency of the model given the data; XP, exceedance probability (computed using the Laplace approximation of the model evidence ME). The table summarizes for each model its fitting performances.

Exp. 1	**Model**	**DF**	**-2^*nLL_{max}**	**2^*AIC**	**BIC**	**-2^*nLPP_{max}**	**EF**	**XP**
	ABSOLUTE	3	385±20	392±20	404±20	391±20	0.28	0.02
	RELATIVE	4	345±24	353±24	369±24	354±24	0.72	0.98
Exp. 2	**Model**	**DF**	**-2^*nLL_{max}**	**2^*AIC**	**BIC**	**-2^*nLPP_{max}**	**EF**	**XP**
	ABSOLUTE	3	411±15	417±15	429±15	416±15	0.05	0.0
	RELATIVE	4	355±16	363±16	379±16	362±16	0.95	1.0

Experiment 1



Experiment 2

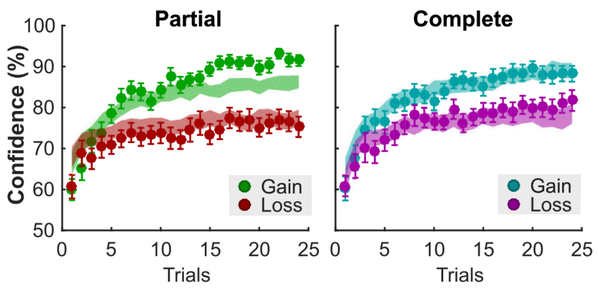


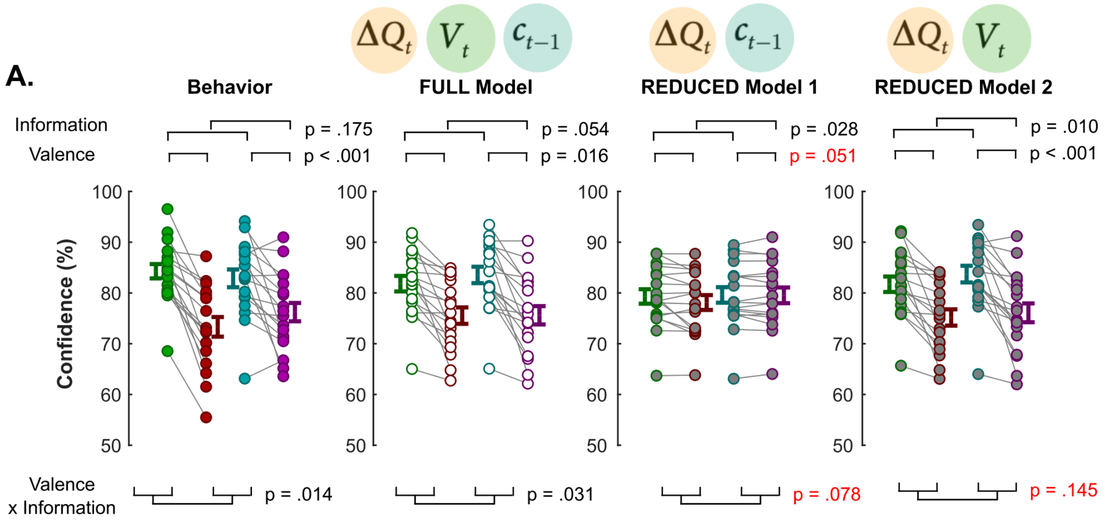
Results

Modeling Results - Confidence

Experiment 1

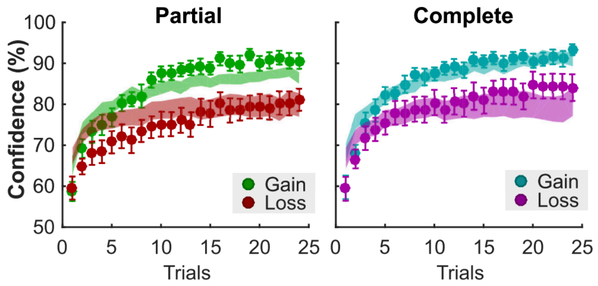
ΔQ_t V_t c_{t-1}

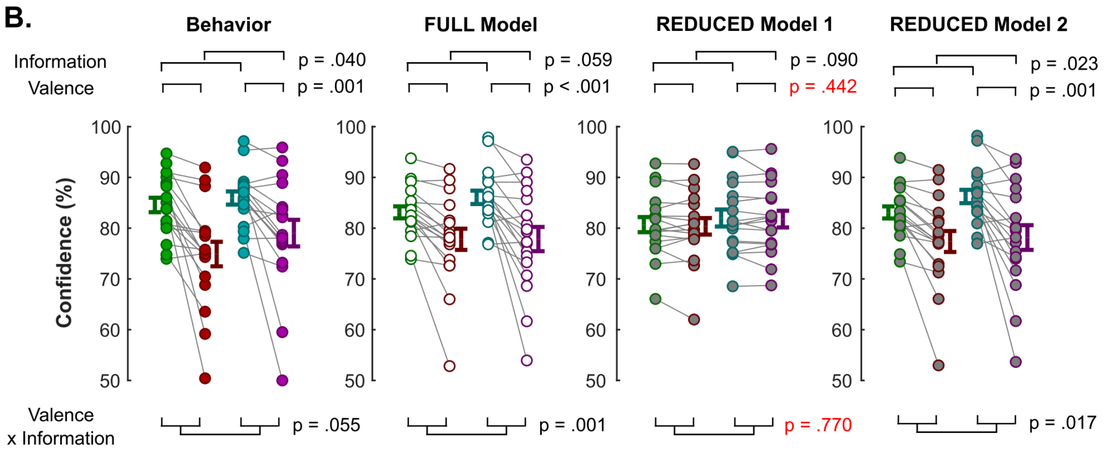




Experiment 2

ΔQ_t V_t c_{t-1}

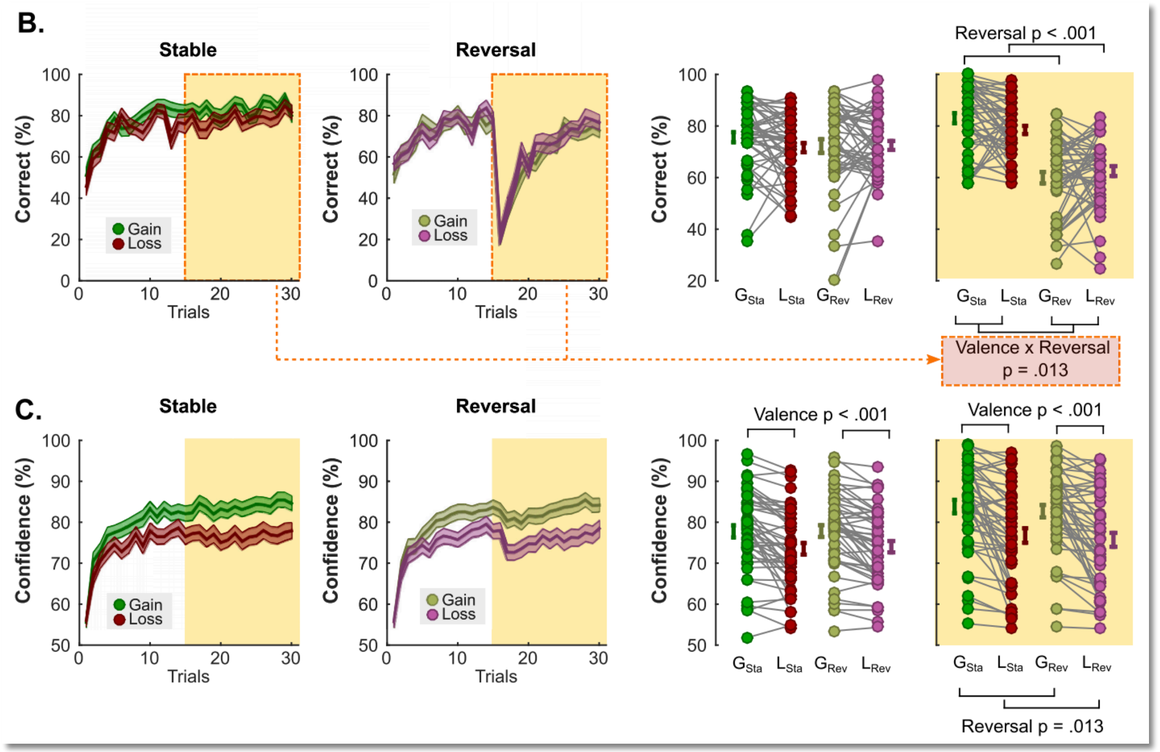




Results

Learning in the Reversal Condition

Experiment 3



1. ***Confidence bias***
(↓ Confidence in ***loss-avoidance*** than gain-seeking learning context)

2. ***Context-Value***

3. ***Performance in volatile environment***

Discussion

Discussion

Implications

1. *The effect of outcome valence on confidence*

   - *Two possible interpretations: <u>simple bias</u> vs <u>improvement</u>*

2. *The confidence can be modeled as a linear combination of 3 variables.*

   - *Previous confidence rating, choice difficulty, and **<u>the context value</u>***

Discussion

Limitations

1. *Little control on choice reaction times*
   - *The relationship between RT and confidence*
2. *The modeling approach informed and constrained by the previous papers*

Cold Spring Harbor Laboratory

bioRχiv

THE PREPRINT SERVER FOR BIOLOGY

Robust Pavlovian-to-Instrumental and Pavlovian-to-Metacognitive Transfers in human reinforcement learning

Chih-Chung Ting, Stefano Palminteri, Jan B. Engelmann, Maël Lebreton

doi: https://doi.org/10.1101/593368 In Library View PDF

This article is a preprint and has not been certified by peer review [what does this mean?].

Thanks!

References

- Norman, E., Pfuhl, G., Sæle, R. G., Svartdal, F., Låg, T., & Dahl, T. I. (2019). Metacognition in psychology. *Review of General Psychology*, *23*(4), 403-424.
- Lebreton, M., Bacily, K., Palminteri, S., & Engelmann, J. B. (2019). Contextual influence on confidence judgments in human reinforcement learning. *PLoS computational biology*, *15*(4), e1006973.
- Palminteri, S., Khamassi, M., Joffily, M., & Coricelli, G. (2015). Contextual modulation of value signals in reward and punishment learning. *Nature communications*, *6*(1), 1-14.