

Working memory guides action valuation in model-based decision-making

左肇煜

2023/09/09



中国科学技术大学
University of Science and Technology of China



ZY

Can you help me offer some paper that discusses how to use reinforcement learning to model human decision-making?



I can provide you with the titles and brief descriptions of a few papers that discuss the use of reinforcement learning to model human decision-making. Please note that the availability of these papers may vary, and you may need access to academic databases or libraries to obtain them.



- Title:** "A Survey of Models of Human Action Planning."
 - Authors: Tamar Flash and Daniel M. Wolpert
 - Description: This survey paper explores various models of human action planning and control, including those based on reinforcement learning principles. It discusses how reinforcement learning concepts can be applied to model human decision-making in motor control tasks.
- Title:** "Deep Reinforcement Learning for Human-AI Collaboration in Games."
 - Authors: Jacob Devlin, et al.
 - Description: This paper focuses on the use of deep reinforcement learning techniques to model human decision-making in the context of human-AI collaboration in video games. It explores how RL agents can learn to cooperate with human players effectively.
- Title:** "Modeling Human Decision-Making in Reinforcement Learning Using Neural

Google 学术搜索

A Survey of Models of Human Action Planning

文章

找到约 4,680,000 条结果 (用时0.37秒)

时间不限

2023以来

2022以来

2019以来

自定义范围...

按相关性排序

按日期排序

不限语言

中文网页

简体中文网页

类型不限

评论性文章

☐ 包括专利

☐ 包含引用

创建快讯

Action planning and predictive coding when speaking

J Wang, DH Mathalon, BJ Roach, J Reilly, SK Keedy... - Neuroimage

... Less is known about the neurophysiology of action planning processes in relation to subsequent sensory suppression, although in song birds and humans, action planning is associated with sensory suppression. This study investigated the relation between action planning and sensory suppression in humans using a task that required action planning and sensory suppression. The results show that action planning is associated with sensory suppression, and that this relation is modulated by the degree of action planning. This study provides evidence for the role of action planning in sensory suppression, and suggests that action planning may be a key component of the sensory suppression process.

☆ 保存 引用 被引用次数: 89 相关文章 所有 14 个版本

An empirical analysis of team coordination behaviors application to human-robot teaming

J Shah, C Breazeal - Human factors, 2010 - journals.sagepub.com

... In this study, we have investigated how human teammates incorporate team coordination behaviors into their decision-making processes. Ultimately, we envision that insights from our study will inform the design of human-robot teams.

☆ 保存 引用 被引用次数: 131 相关文章 所有 9 个版本

Action planning and the temporal binding of response

G Stoet, B Hommel - Journal of Experimental Psychology: Human ...

... action plan may be thought of as an action file that integrates all the information needed to execute an action. The action file is updated as the action is executed, and the updated action file is used to plan the next action. This study investigated the role of action planning in the temporal binding of response, and found that action planning is a key component of the temporal binding process.

☆ 保存 引用 被引用次数: 312 相关文章 所有 14 个版本

Business owners' action planning and its relationship three African countries.

M Freese, SI Krauss, N Keith, S Escher... - Journal of applied ...

... To our knowledge, action plans ... study elaborate and proactive action planning. Since cognitive ability is largely resistant to intervention and elaborate action planning is a key component of the action planning process, this study provides evidence for the role of action planning in the temporal binding of response, and suggests that action planning may be a key component of the sensory suppression process.

☆ 保存 引用 被引用次数: 378 相关文章 所有 21 个版本

Moving from agenda to action: evaluating local climate change action plans

Z Tang, SD Brody, C Quinn, L Chang... - ... environmental planning ...

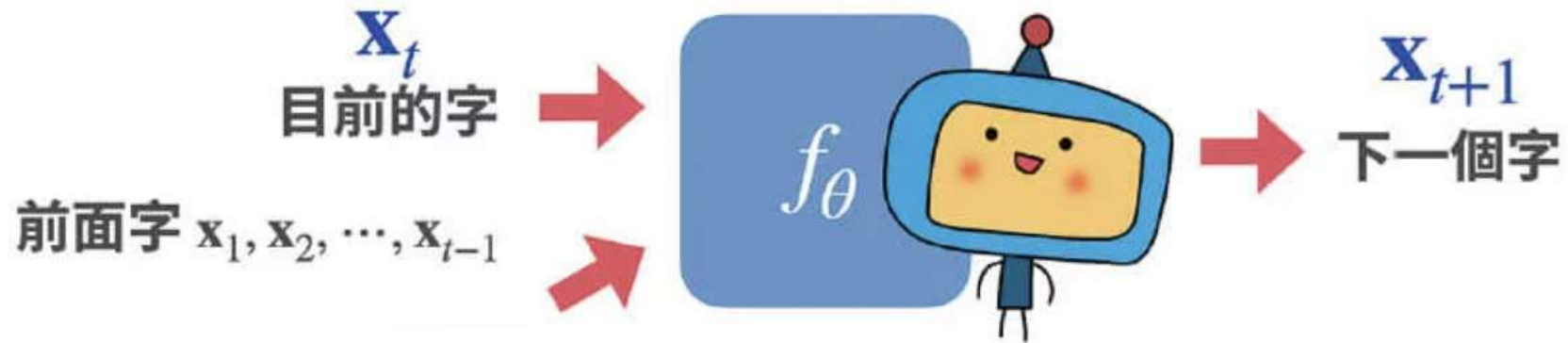
... One of the major findings of this study is that current local climate change action plans are predominantly on the built environment (eg energy, transportation, waste management, etc.). This study provides evidence for the role of action planning in the temporal binding of response, and suggests that action planning may be a key component of the sensory suppression process.

☆ 保存 引用 被引用次数: 378 相关文章 所有 17 个版本

Behavioral intentions and action plans promote physical activity study with orthopedic rehabilitation patients

... study with orthopedic rehabilitation patients

生成式AI—文字接龙



h_{t-1}

文字生成的 AI 模型設計原理，輸入前一個字之後便預測下一個字。

Source of image:

<https://blog.hamibook.com.tw/%E5%95%86%E7%AE%A1%E7%90%86%E8%B2%A1/%E6%8F%AD%E5%AF%86chatgpt%E3%80%81midjourney%E7%94%9F%E6%88%90%E5%BC%8F%E5%A6%82%E4%BD%95%E5%AD%B8%E7%BF%92%E5%86%8D%E9%80%B2%E5%8C%96%E7%BC%9F/?p=207118>



chrome 应用商店

WebChatGPT: ChatGPT 具备互联网访问功能

[webchatgpt.app](#) 精选

通过网络浏览增强您的ChatGPT提示与相关的网络搜索结果。

★★★★★ 1,090 效率



Can you help me find a paper that discusses how to use reinforcement learning to model human decision-making?

⚡ Quick Search

1. [ScienceDirect - How are reinforcement learning and deep...](#)
2. [Improving Human Decision-Making with Machine Learning](#)
3. [PLOS Computational Biology - The case of decision-making and its disorders](#)
4. [PubMed Central - Predicting human decision making in psychological tasks](#)
5. [Hugging Face - Illustrating Reinforcement Learning from Human Feedback](#)
6. [Journal of NeuroEngineering and Rehabilitation - Deep reinforcement learning for modeling human locomotion](#)

PLOS COMPUTATIONAL BIOLOGY

OPEN ACCESS PEER-REVIEWED
RESEARCH ARTICLE

Models that learn how humans learn: The case of decision-making and its disorders

Amer Dazibail, Kristi Griffiths, Fabio Ramos, Peter Dayan, Bernard W. Balleine

Version 2 Published: June 11, 2019 • <https://doi.org/10.1371/journal.pcbi.1006903>

Article	Authors	Metrics	Comments	Media Coverage
---------	---------	---------	----------	----------------

Reader Comments
Figures

Figures



As a library, NLM provides access to scientific literature. Inclusion in an NLM database does not imply endorsement of, or agreement with, the contents by NLM or the National Institutes of Health.
Learn more: [PMC Disclaimer](#) | [PMC Copyright Notice](#)

PLOS ONE

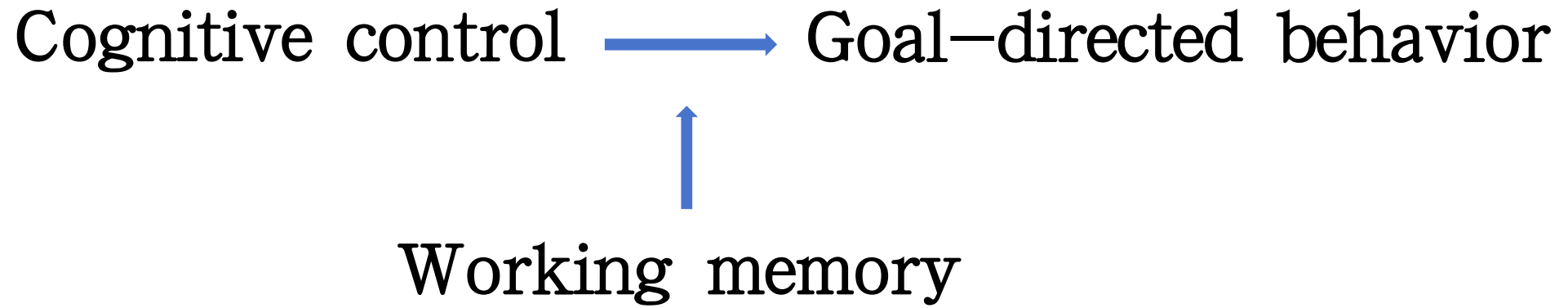
[PLOS ONE](#), 2022; 17(5): e0267907.
Published online 2022 May 31, doi: [10.1371/journal.pone.0267907](https://doi.org/10.1371/journal.pone.0267907)

PMCID: PMC9154096
PMID: [35639730](#)

Predicting human decision making in psychological tasks with recurrent neural networks

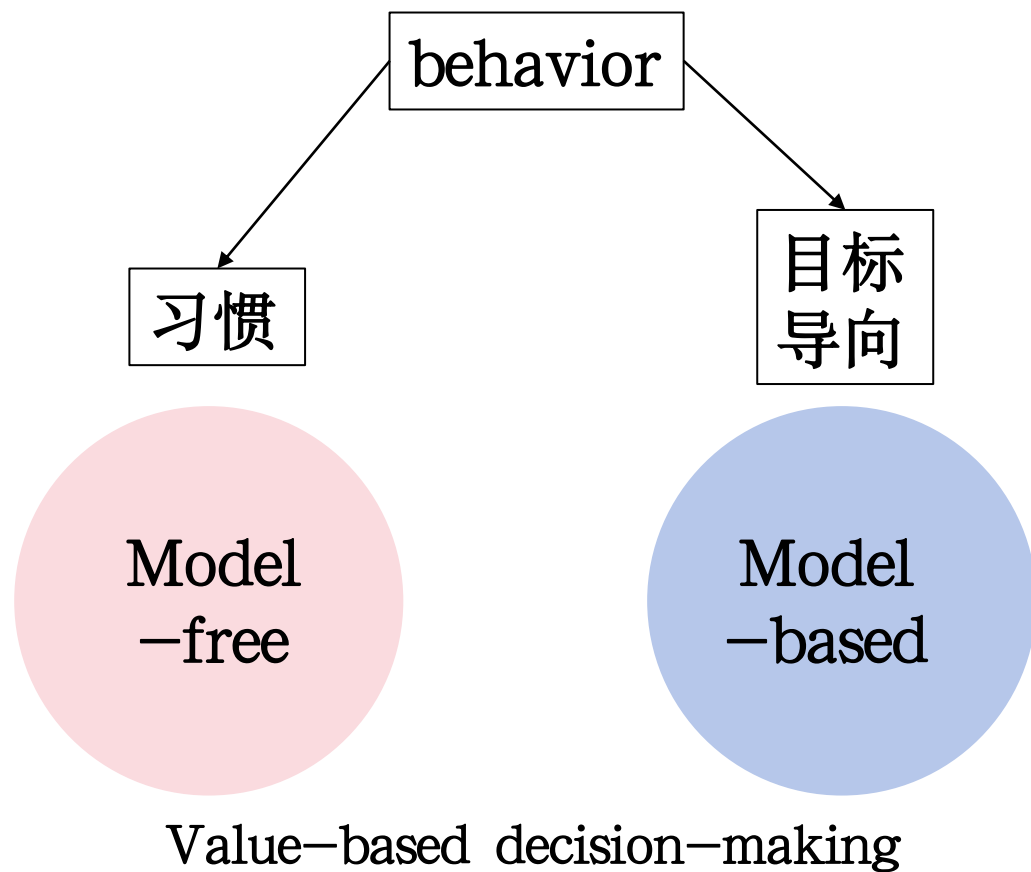
[Bahar Lin](#), Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing, ^{1, 2, 3, 4}
[Gabriel Bounieffouf](#), Conceptualization, Methodology, Project administration, Resources, Writing – review & editing, ⁴ and
[Guillermo Cecchi](#), Conceptualization, Methodology, Project administration, Resources, Supervision, Writing – review & editing ⁵

Background

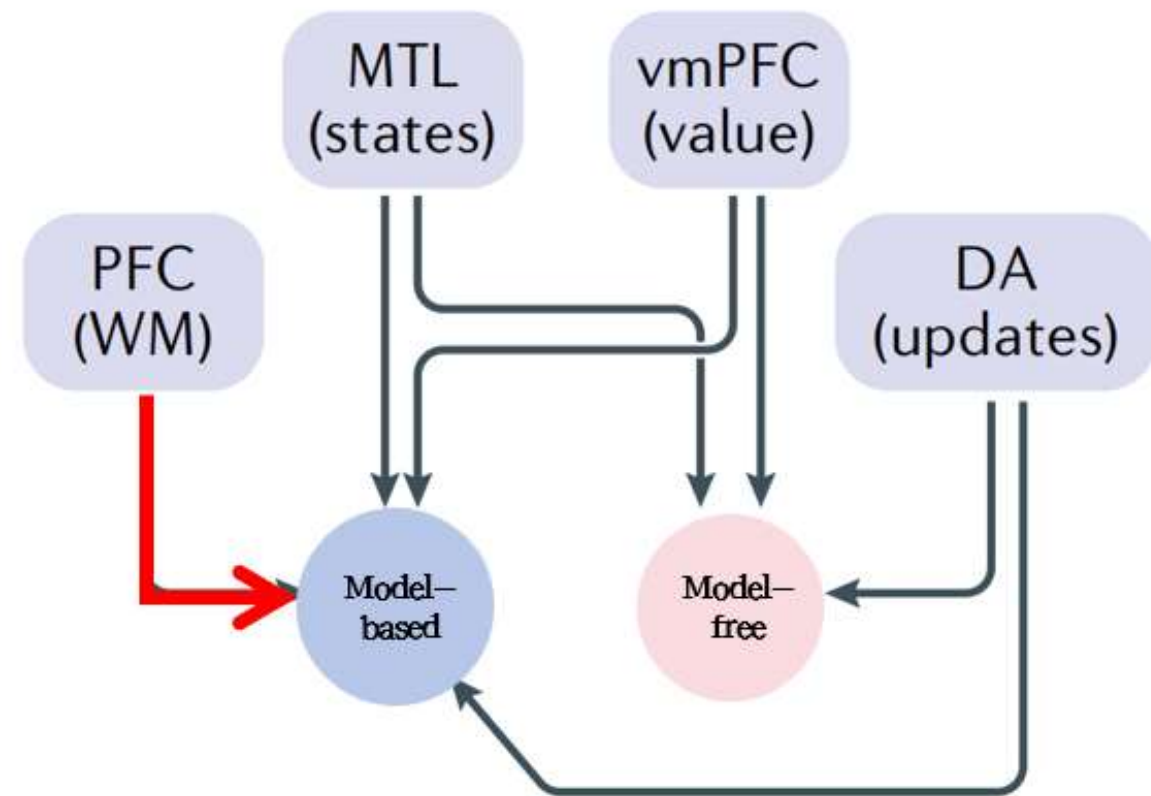


- Task-related information
- Capacity-limited and **delay-sensitive**

Dual-system



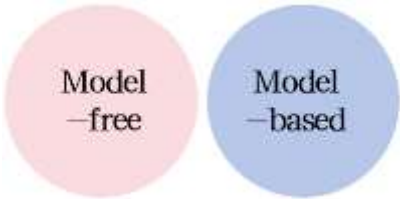
Daw et al., 2011, Neuron



Collins et al., 2020, Nat. Rev. Neurosci.

Dual-system modeling

Algorithm



强化学习

Cognitive & neural
process

强化
学习

工作
记忆

Behavior

习惯

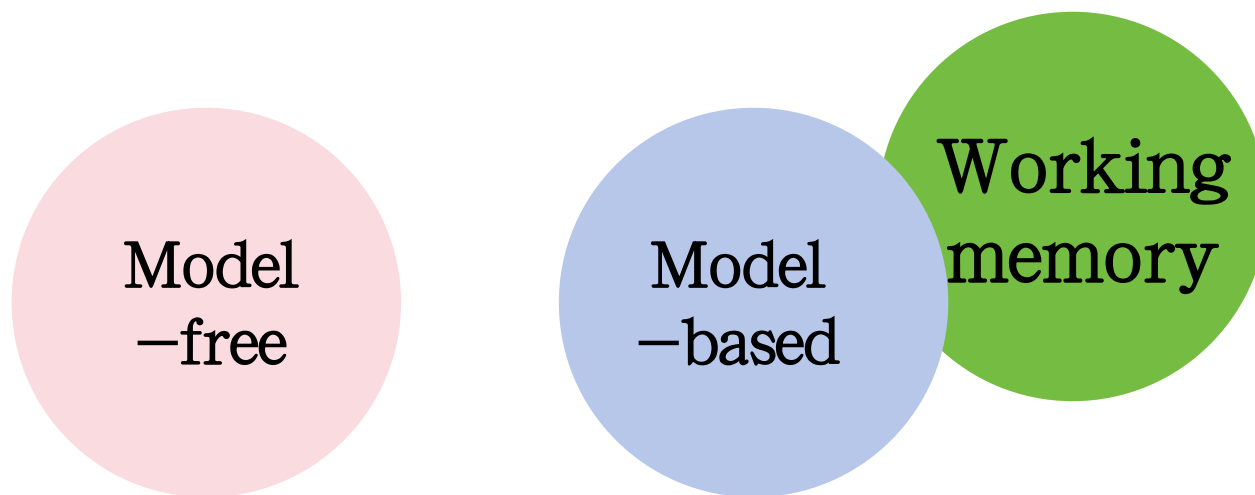
目标
导向

问题：

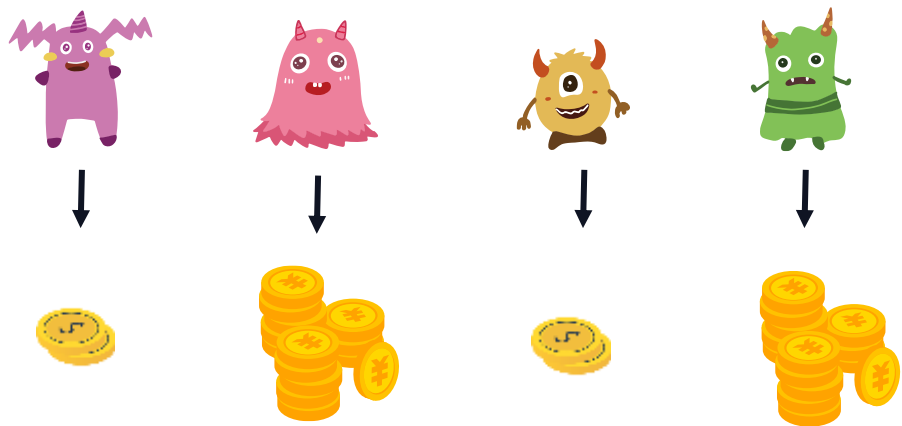
工作记忆中的奖励信息是否以及如何影响Model-based决策？

假设：

- 延迟敏感的工作记忆参与了目标估值，时间延迟将破坏估值过程
- 在Model-based中考虑工作记忆影响的模型可以更好拟合数据

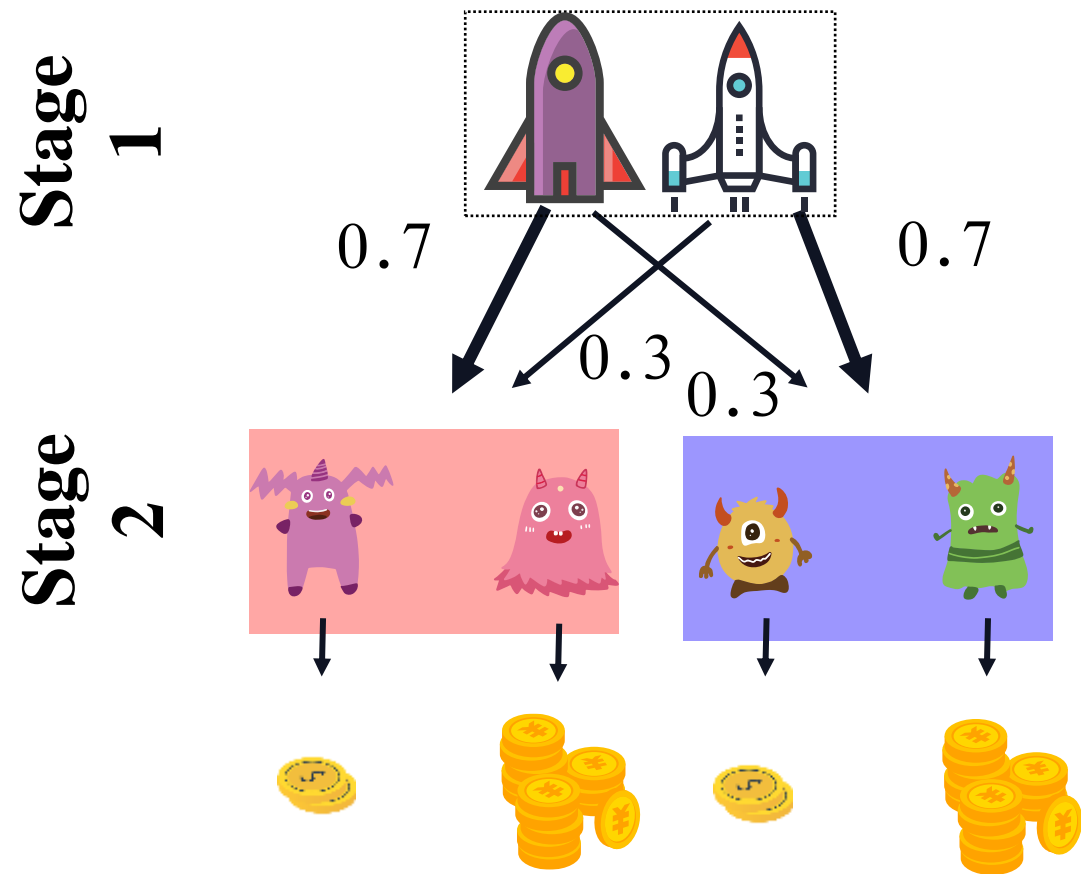


- 多臂老虎机任务



持续变化

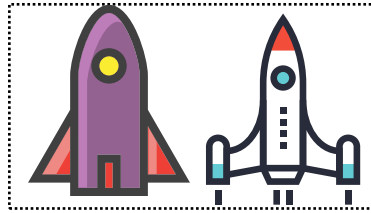
- 两阶段任务



持续变化

Model-free agent

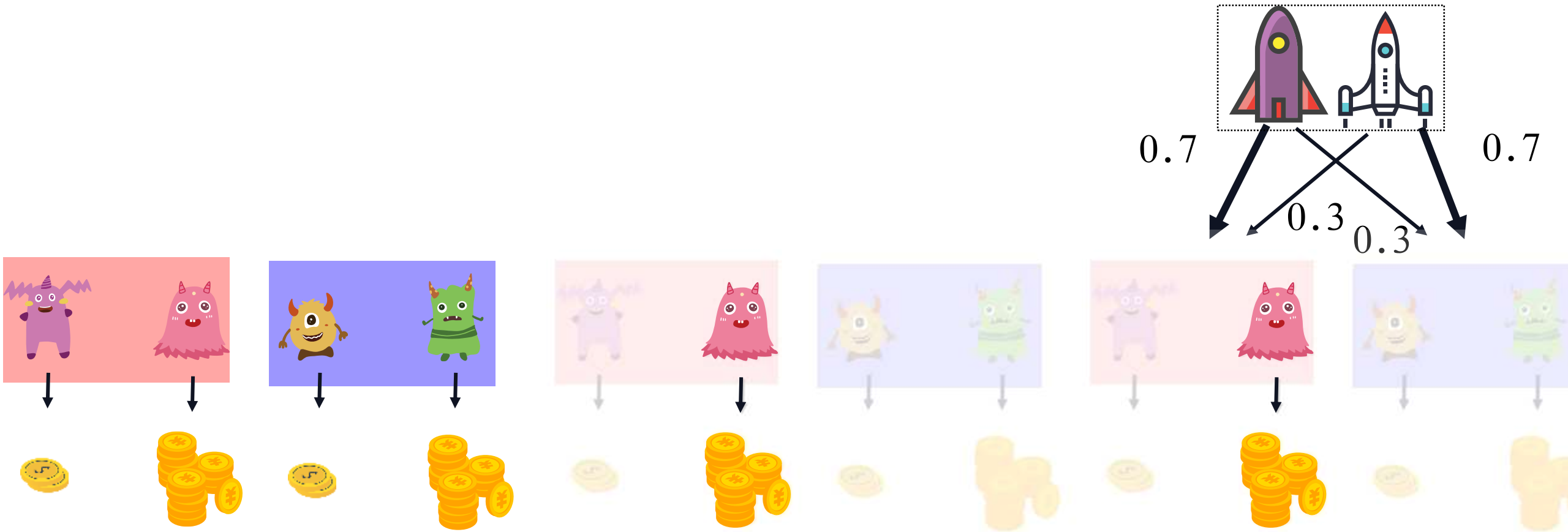
Stage
1



Stage
2

Model-free 不考虑目标价值

Model-based agent



强化学习



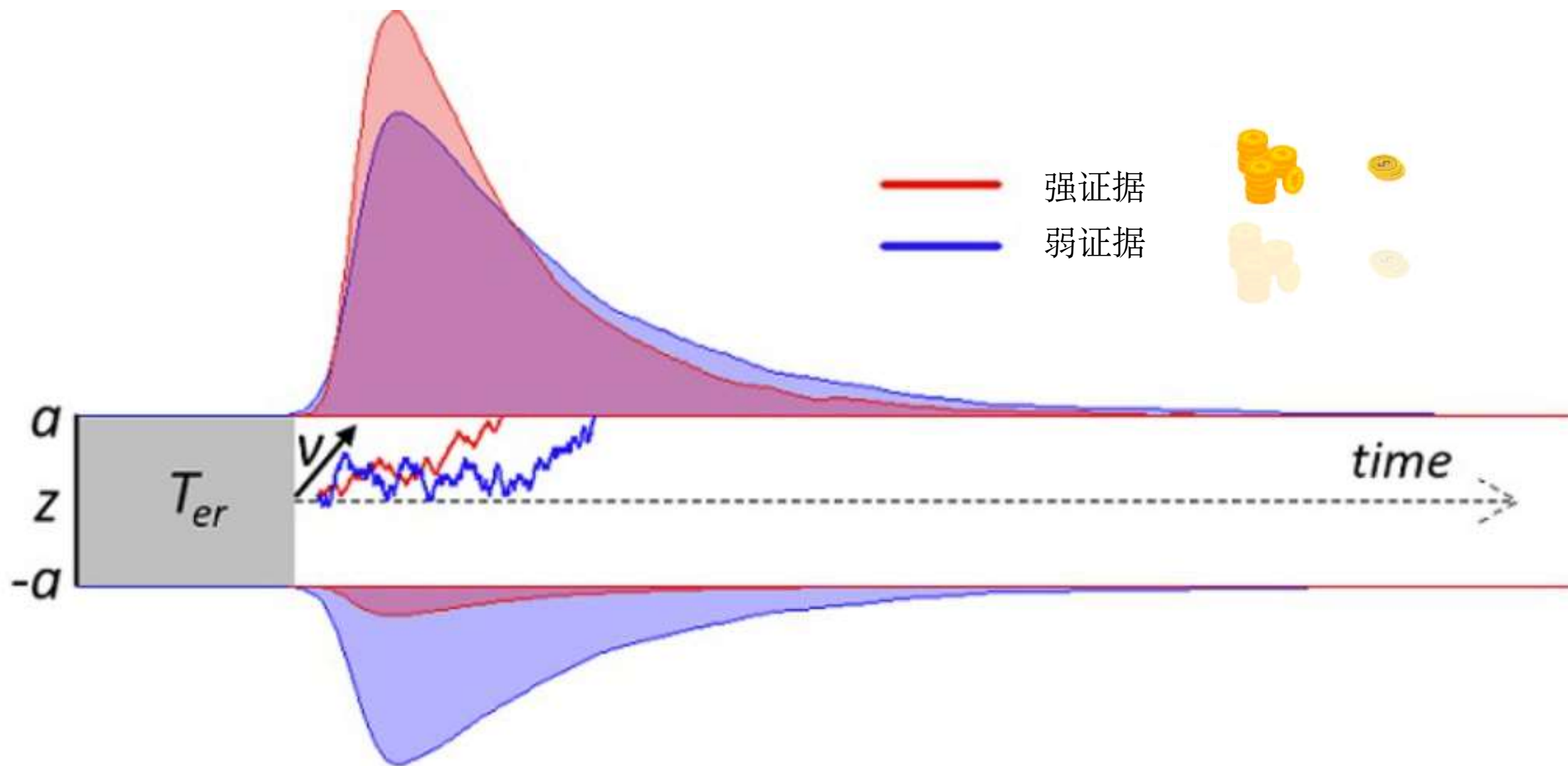
工作记忆



$T=0$

$T=1$

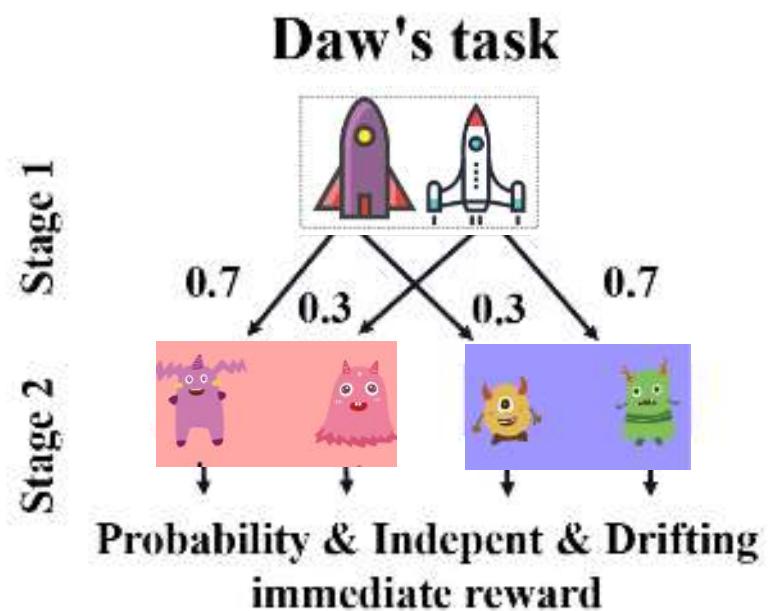
$T=10$



证据不足导致错误增加，反应时间增加

Analysis 1

在两阶段任务中分析延迟对价值估计的影响



Undelayed

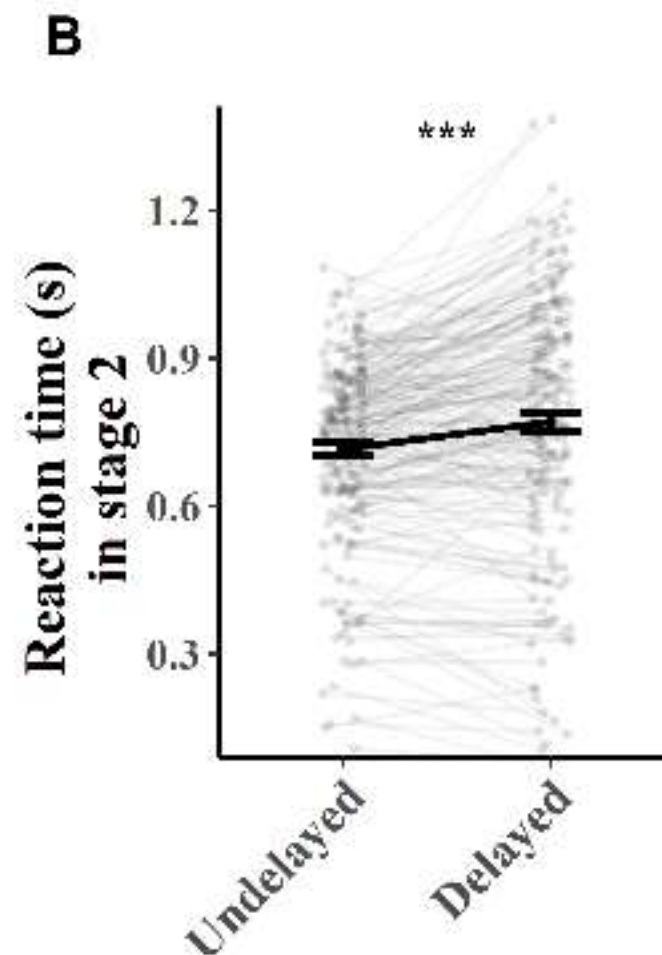
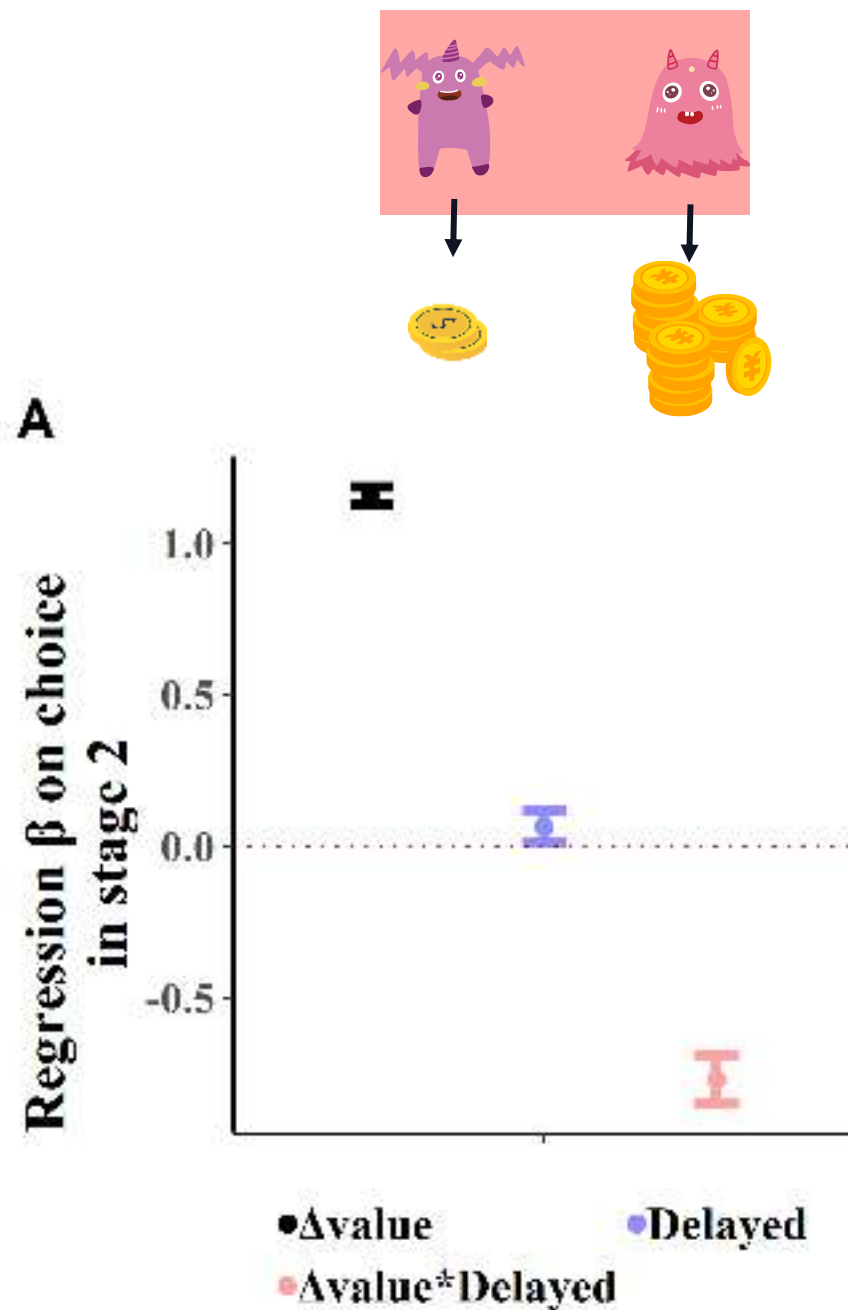


Delayed

Result 1

发现了延迟对动作估值的影响
(Dataset 1 and 2)

- 延迟使得选择变得随机而缓慢

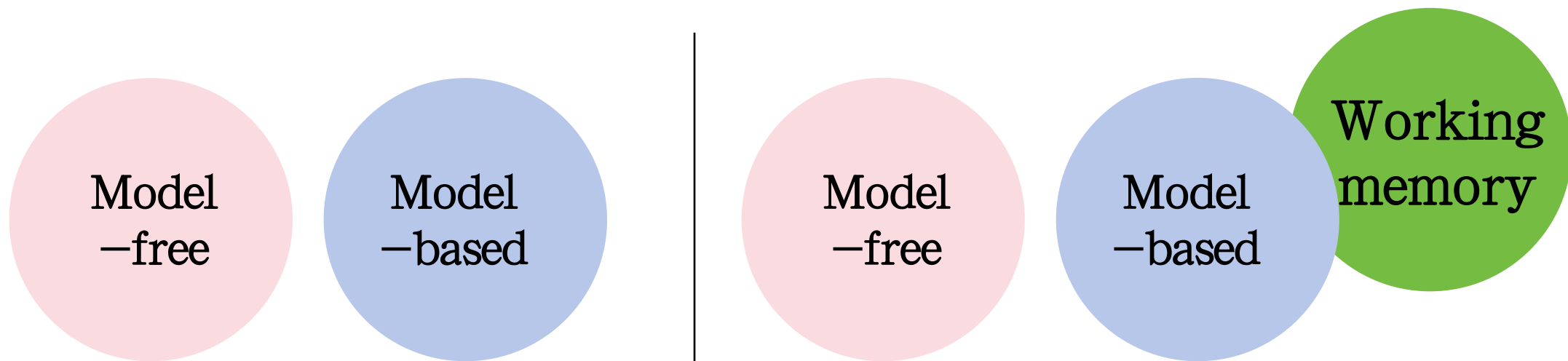


Interim summary

- 对目标动作的价值估计依赖于工作记忆中的奖励信息

Analysis 2

将工作记忆机制纳入决策建模，并与传统模型进行模型比较、行为仿真比较

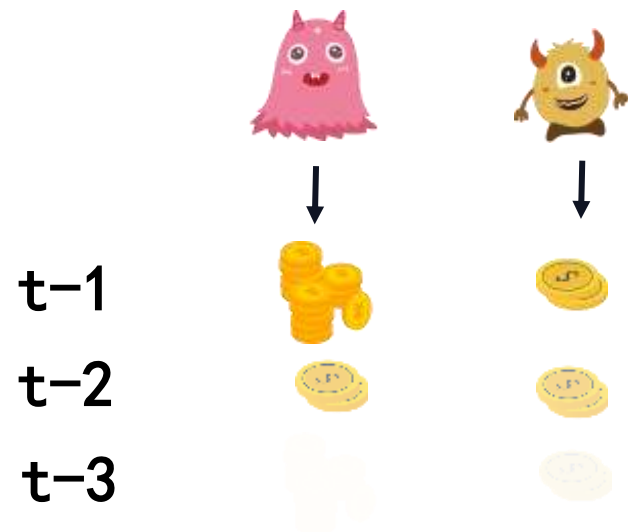


1. 工作记忆机制 2. 系统交互关系 3. 遗忘机制

工作记忆的价值处理机制

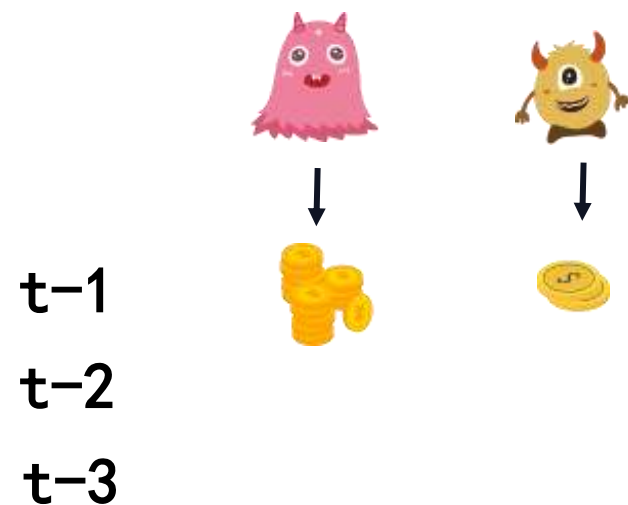
- **Sampler**

- 记住多个，但时间越久越记不清
- Bornstein et al. 2017, Nat. Commun.

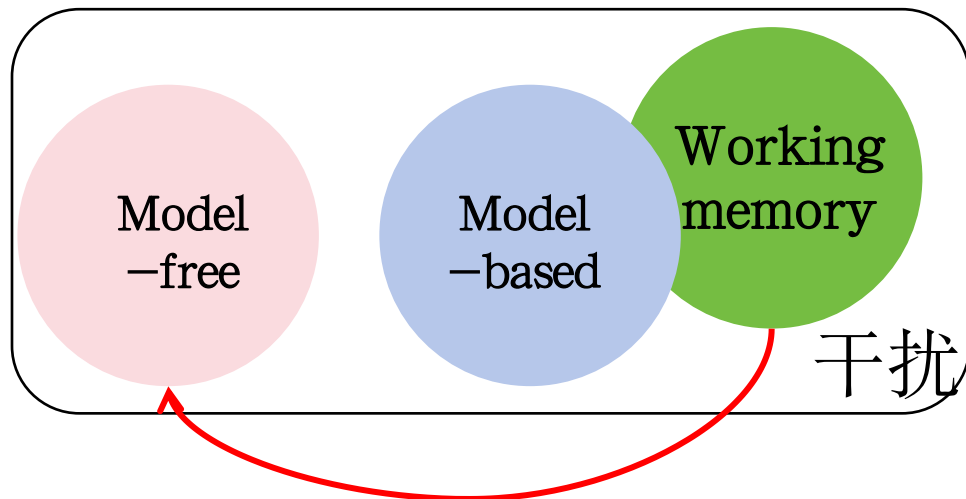
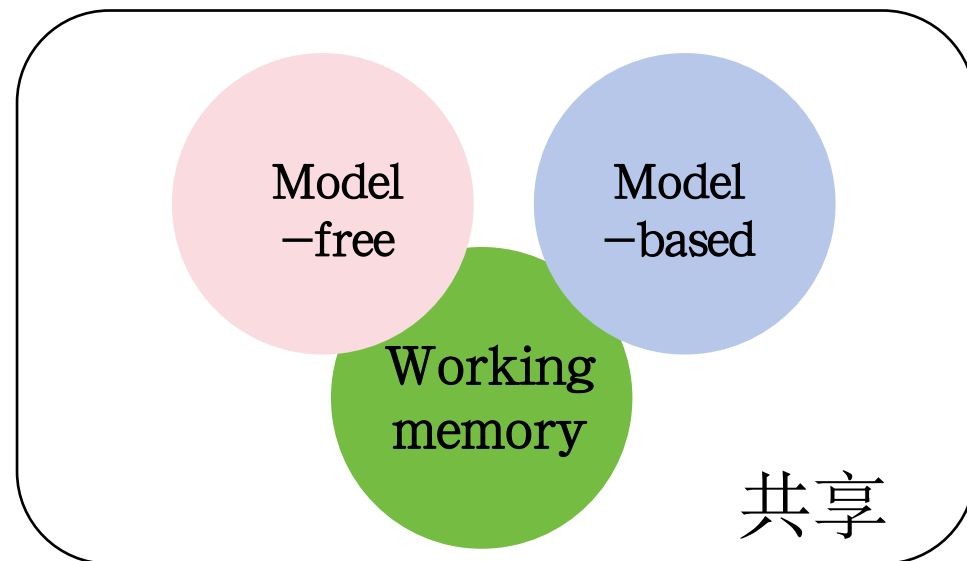
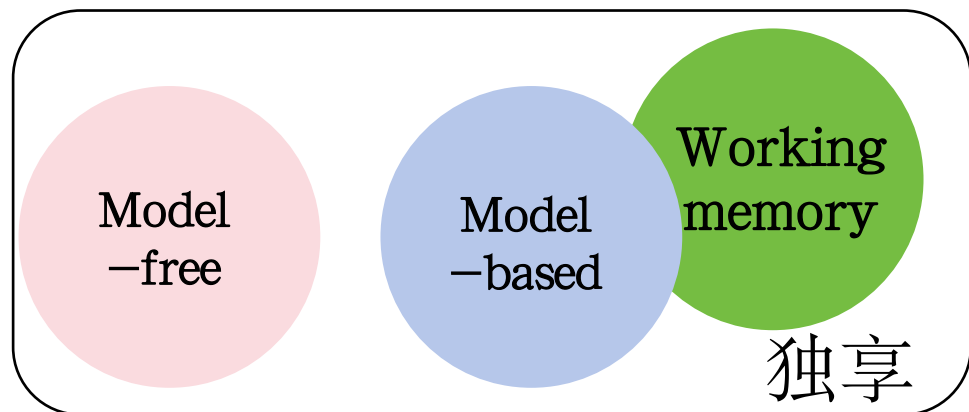


- **Refresher**

- 只记最后一次奖励
- Collins et al. 2017, J. Neurosci.



系统交互关系



遗忘机制

$$W_2(s_{2,t+1}, a_{2,t+1}) = W_2(s_{2,t}, a_{2,t}) + \varphi \left(W_0 - W_2(s_{2,t}, a_{2,t}) \right)$$

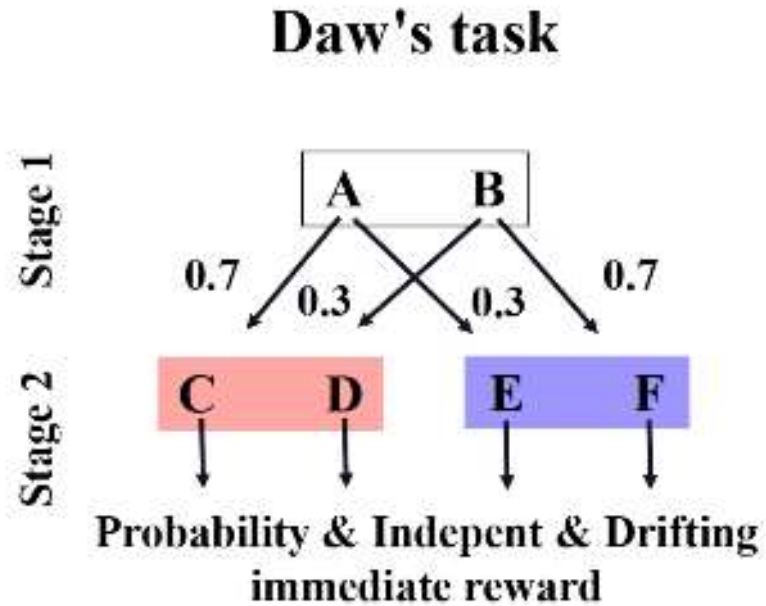
遗忘速率



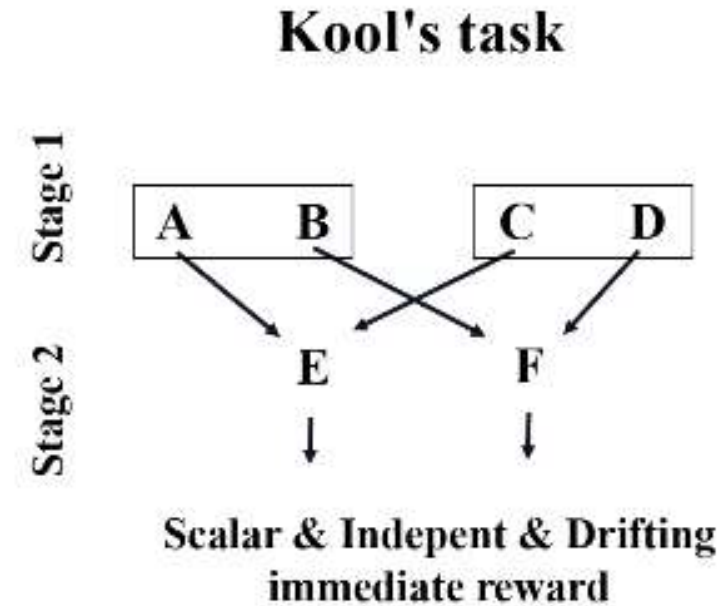
Alternatives models

- 共14 (12 + 1 + 1) 个备选模型
- 12 = 2(工作记忆机制) * 3(系统交互关系)* 2(是否包含遗忘机制)
- 1 = 经典的纯粹强化学习模型
- 1 = Win-stay Lose-swift

Dataset 1 and 2



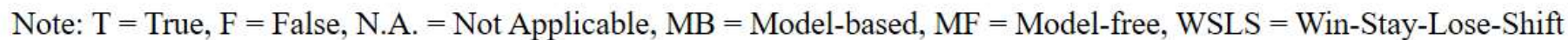
Dataset 3 and 4



- N=197, Kool et al. 2016, PLoS Comput. Biol.
- N=1108, Shahar et al. 2019, PLoS Comput. Biol.
- N=124, Bolenz et al. 2019, elife
- 一个replicated实验 (N=29)

Result 2.1:

考虑工作记忆的模型更好



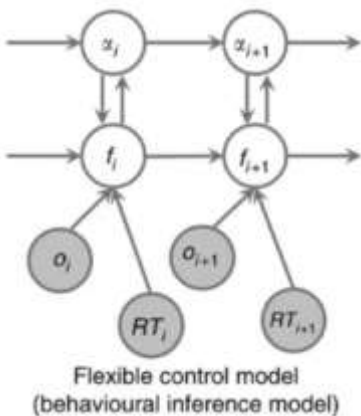
1 工作记忆机制

环境快速变化—refresher

$$W_2(s_{2,t}, a_{2,t}) = r_t$$

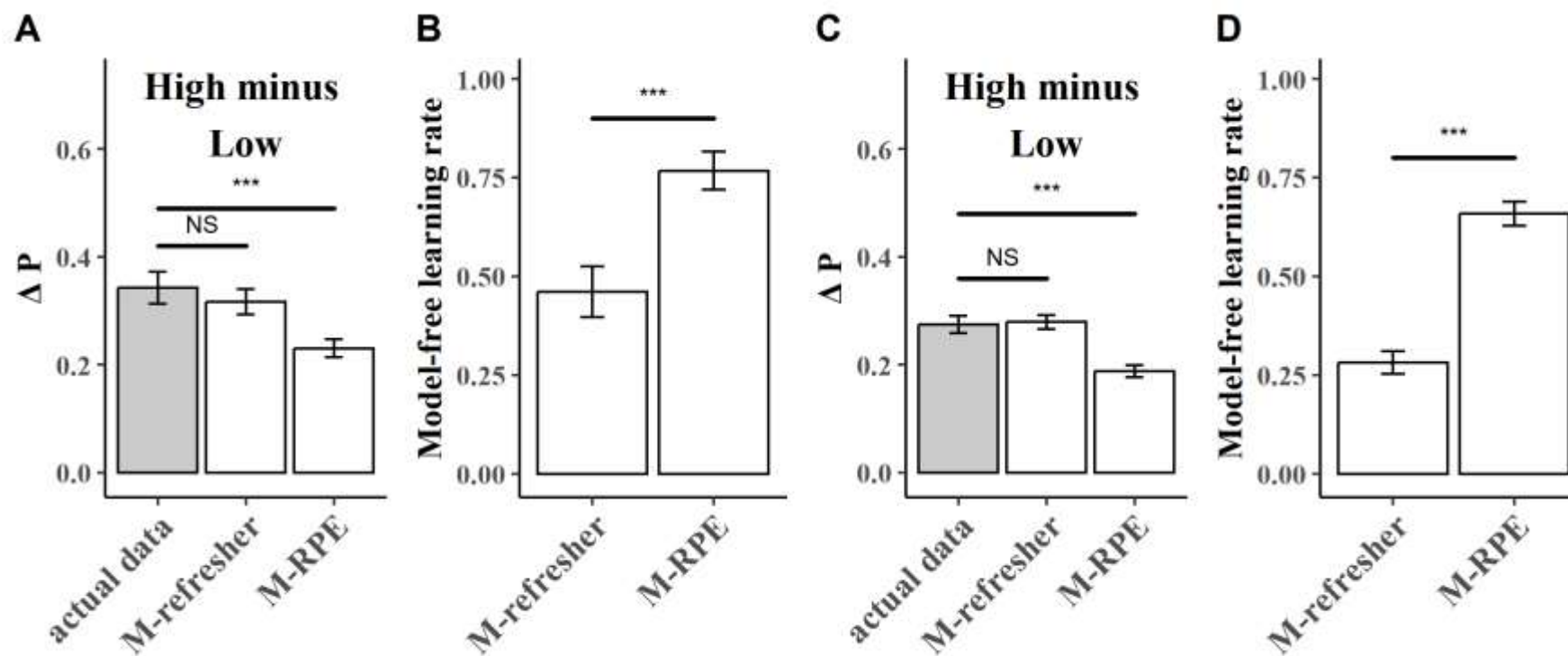
环境缓慢变化—sampler

$$\omega_s(q) = \begin{cases} \tau * (1 - \tau)^{q-1}, n_{s_{2,t}, a_{2,t}} > 1 \text{ and } q \in \{1, \dots, n_{s_{2,t}, a_{2,t}} - 1\} \\ 1 - \sum_{i \neq q} \omega_s(i), n_{s_{2,t}, a_{2,t}} > 1 \text{ and } q = n_{s_{2,t}, a_{2,t}} \\ 1, n_{s_{2,t}, a_{2,t}} = 1 \end{cases}$$



补充分析：

纯粹强化学习建模无法模拟两阶段任务中的学习决策



1 工作记忆机制

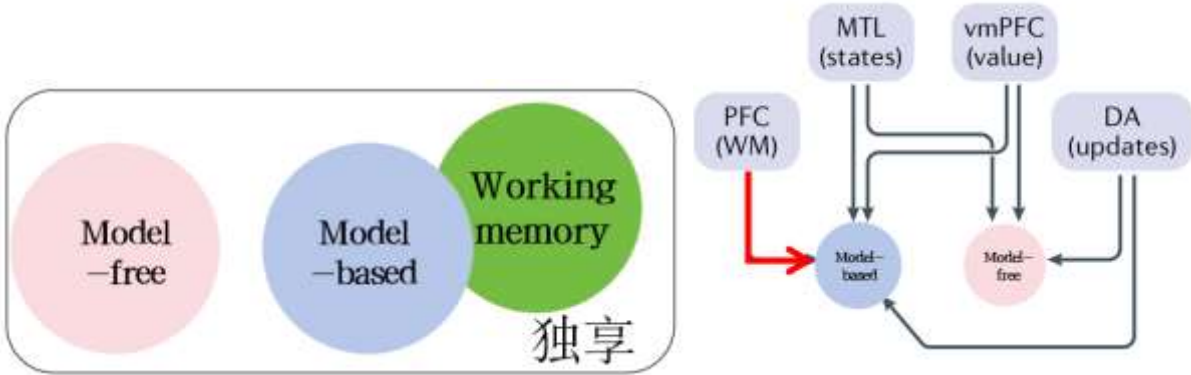
环境快速变化—refresher

$$W_2(s_{2,t}, a_{2,t}) = r_t$$

环境缓慢变化—sampler

$$\omega_s(q) = \begin{cases} \tau * (1 - \tau)^{q-1}, n_{s_{2,t}, a_{2,t}} > 1 \text{ and } q \in \{1, \dots, n_{s_{2,t}, a_{2,t}} - 1\} \\ 1 - \sum_{i \neq q} \omega_s(i), n_{s_{2,t}, a_{2,t}} > 1 \text{ and } q = n_{s_{2,t}, a_{2,t}} \\ 1, n_{s_{2,t}, a_{2,t}} = 1 \end{cases}$$

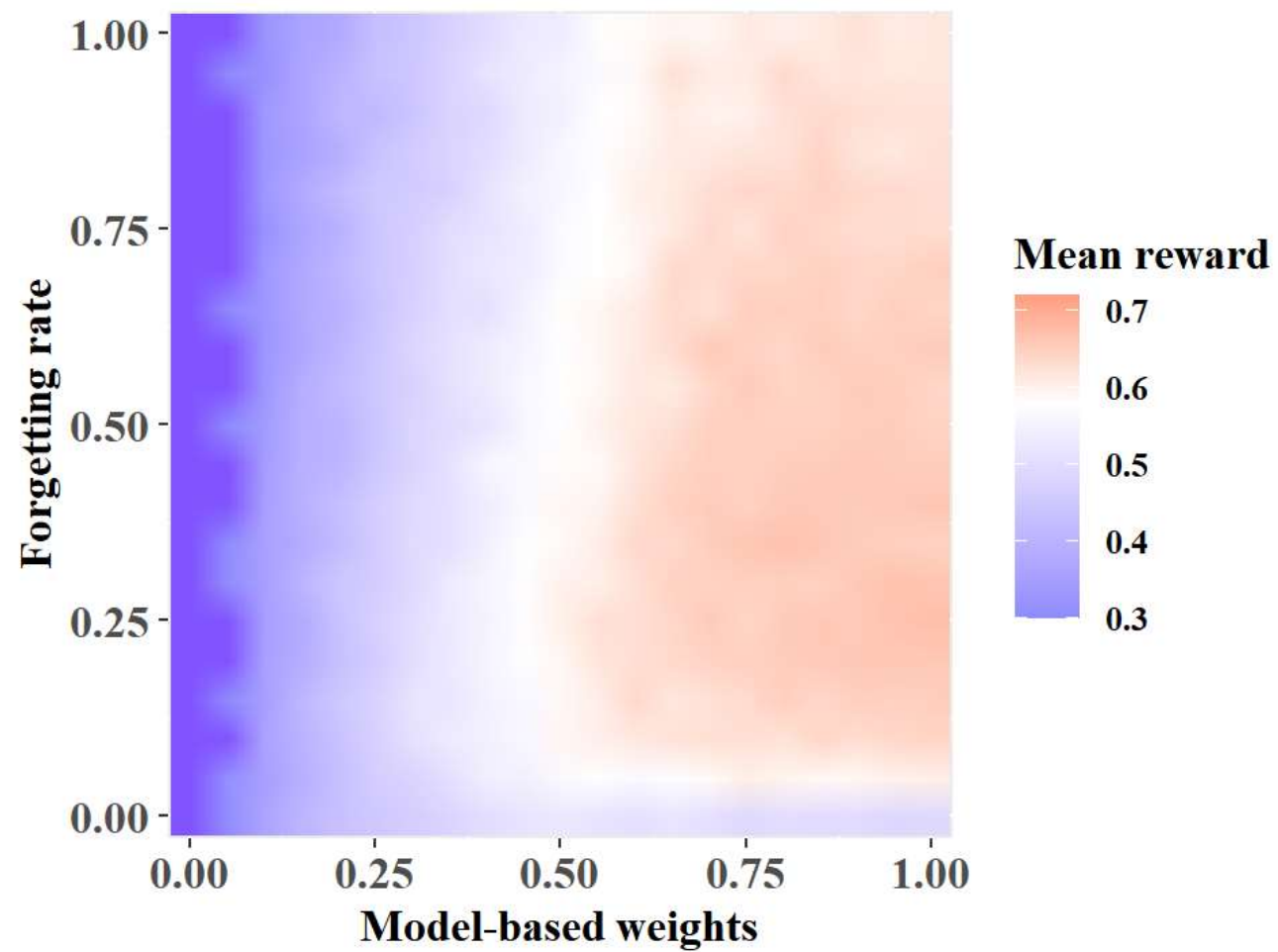
2 系统交互关系



3 遗忘机制

$$W_2(s_{2,t+1}, a_{2,t+1}) = W_2(s_{2,t}, a_{2,t}) + \varphi(W_0 - W_2(s_{2,t}, a_{2,t}))$$

补充分析：
适当的遗忘有利于任务表现



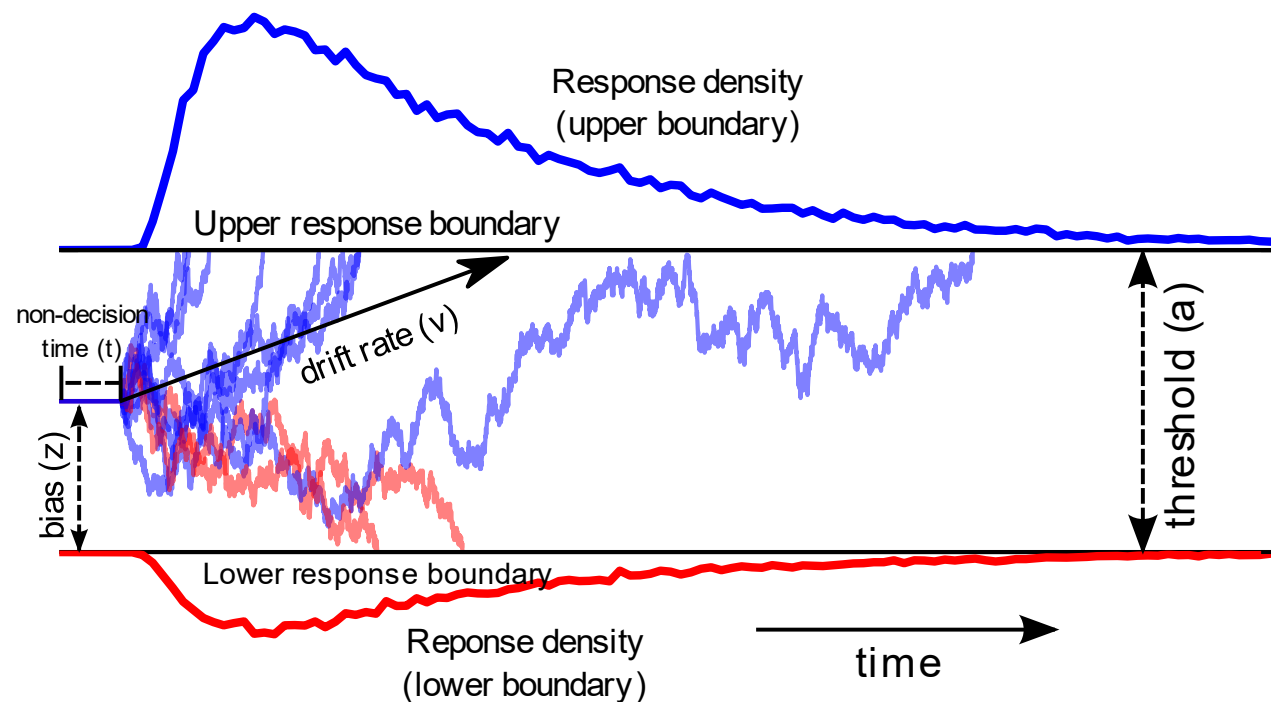
结合漂移扩散模型 (DDM) 拟合反应时间

Model
-free

Model
-based

Working
memory

$$Q_{net}(s_1, a_1) = \omega Q_{MB}(s_1, a_1) + (1 - \omega) Q_{MF}(s_1, a_1)$$

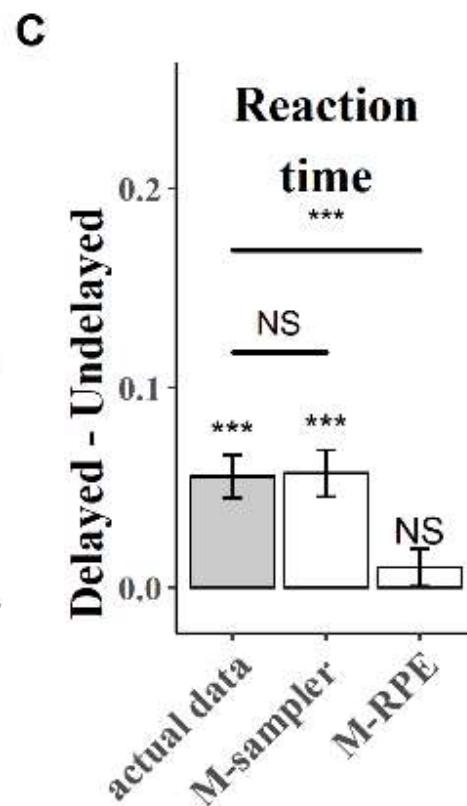
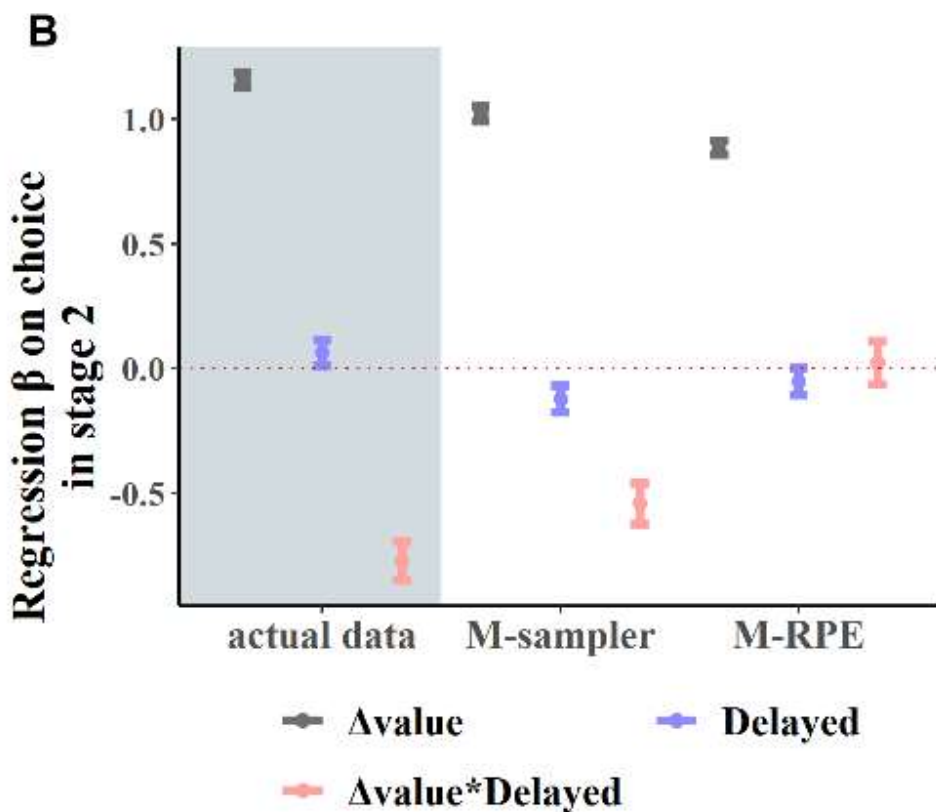


$$v_t = \tau(Q_{net}(s_1, a_1) - Q_{net}(s_1, a_2))$$

Result

Result 2.2:

模型定量地模仿了延迟对人类决策的影响



Result

Result 2.3:

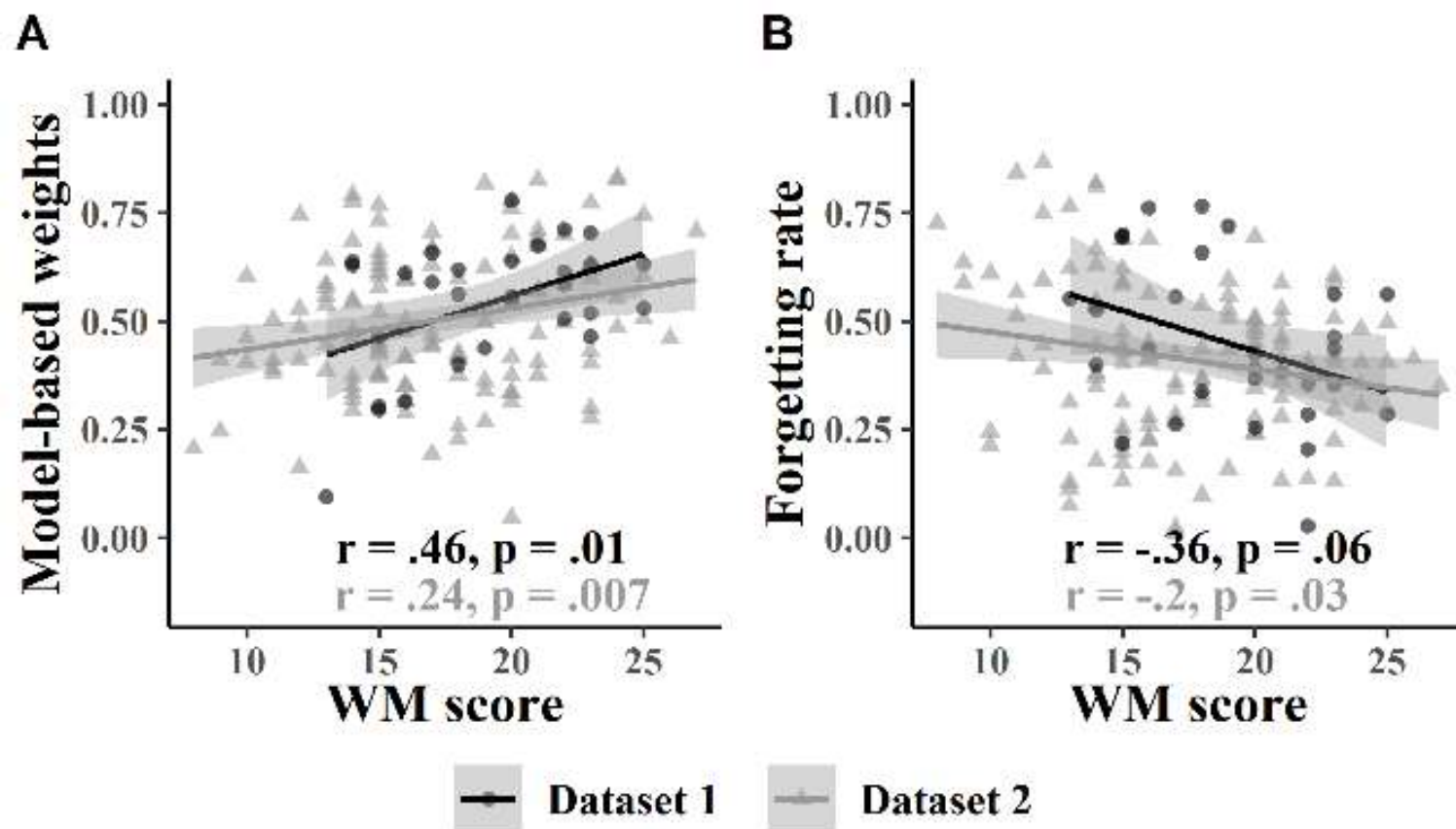
参数与工作记忆得分的关系

- 工作记忆得分越高，Model-based越强
- 工作记忆得分越高，遗忘速率越低

Model
-free

Model
-based

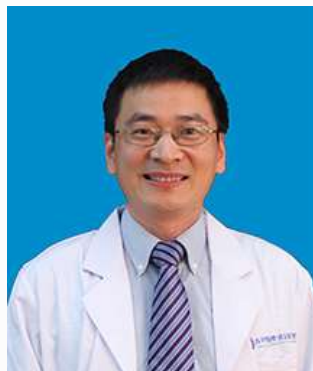
Working
memory



Summary

- 工作记忆中的奖励信息指导了model-based决策中的动作估值
- 提出了可以捕捉学习动态和个体认知特性的决策新模型

Thank you!



Hai Li



Li-zhuang Yang

