



НАУКА В ПУБЛИКАЦИЯХ

Дискуссионная площадка
ОТ ИДЕИ ДО ПРАКТИКИ



МОСКОВСКИЙ
ГОСУДАРСТВЕННЫЙ
ПСИХОЛОГО-
ПЕДАГОГИЧЕСКИЙ
УНИВЕРСИТЕТ



Московский
МЭГ-центр

Мозговые механизмы принятия решения у человека: Внутренний конфликт при переключении между стратегиями использования и исследования

Чернышев Б. В.

**Стратегии использования и исследования:
что это такое?**

Спикеры:

- [Борис Владимирович Чернышев](#), кандидат биологических наук, руководитель Центра нейрокогнитивных исследований МГППУ
- [Галина Леонидовна Козунова](#), кандидат психологические наук, старший научный сотрудник Центра нейрокогнитивных исследований МГППУ
- [Кристина Игоревна Пульцина](#), кандидат психологические наук, научный сотрудник Центра нейрокогнитивных исследований МГППУ

Публикации по теме мероприятия:

1. [Принятие решения в условиях неопределенности: стратегии исследования и использования](#) / К.Э. Сайфулина, Г.Л. Козунова, В.А. Медведев, А.М. Рытикова, Б.В. Чернышев // Современная зарубежная психология. 2020. Т. 9. №. 2. С. 93–106.
2. [Pupil dilation and response slowing distinguish deliberate explorative choices in the probabilistic learning task](#) / G.L. Kozunova, K.E. Sayfulina, A.O. Prokofyev, V.A. Medvedev, A.M. Rytikova, T.A. Stroganova, B.V. Chernyshev // Cognitive, Affective, & Behavioral Neuroscience. 2022. Vol. 22. № 5. P. 1108–1129.
3. [Losses resulting from deliberate exploration trigger beta oscillations in frontal cortex](#) / B.V. Chernyshev, K.I. Pultsina, V.D. Tretyakova, A.S. Miasnikova, A.O. Prokofyev, G.L. Kozunova, T.A. Stroganova // Frontiers in Neuroscience. 2023. Vol. 17. Article ID 1152926.

Центр нейрокогнитивных исследований МГППУ

Новости

Наши вакансии

Доступ к УНУ

О методе МЭГ

Сотрудники

Исследования

МЭГ-центр МГППУ

НИУ ВШЭ

Сколтех

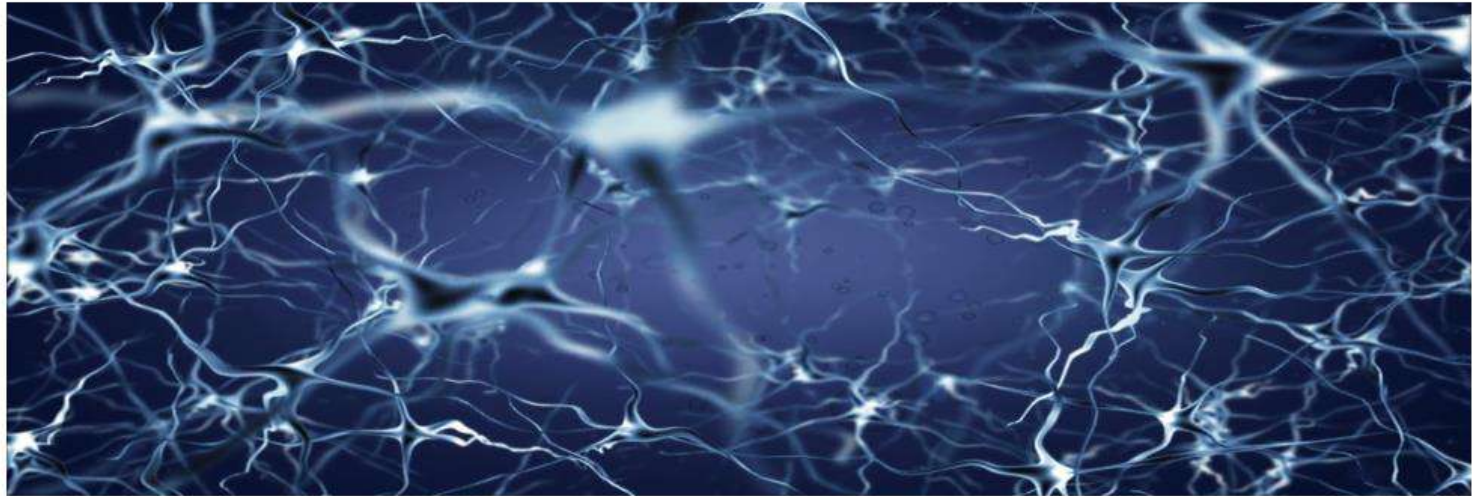
Публикации

Оборудование

Участие в
исследованиях

Полезные ссылки

Контакты



Добро пожаловать в МЭГ-центр!

Центр нейрокогнитивных исследований (МЭГ-центр) [Московского государственного психолого-педагогического университета \(МГППУ\)](#) — **первая и единственная** в России научно-исследовательская лаборатория для измерения активности нейронов головного мозга с использованием магнитоэнцефалографии (МЭГ).

МЭГ имеет высокое, миллисекундное временное разрешение. В сочетании с информацией, полученной при структурной магнитно-резонансной томографии (сМРТ), этот метод позволяет регистрировать изменения нейронной активности головного мозга с высоким пространственным разрешением (несколько

**Магнитоэнцефалография (МЭГ):
современный метод нейровизуализации работы
головного мозга с высоким временным и
пространственным разрешением**



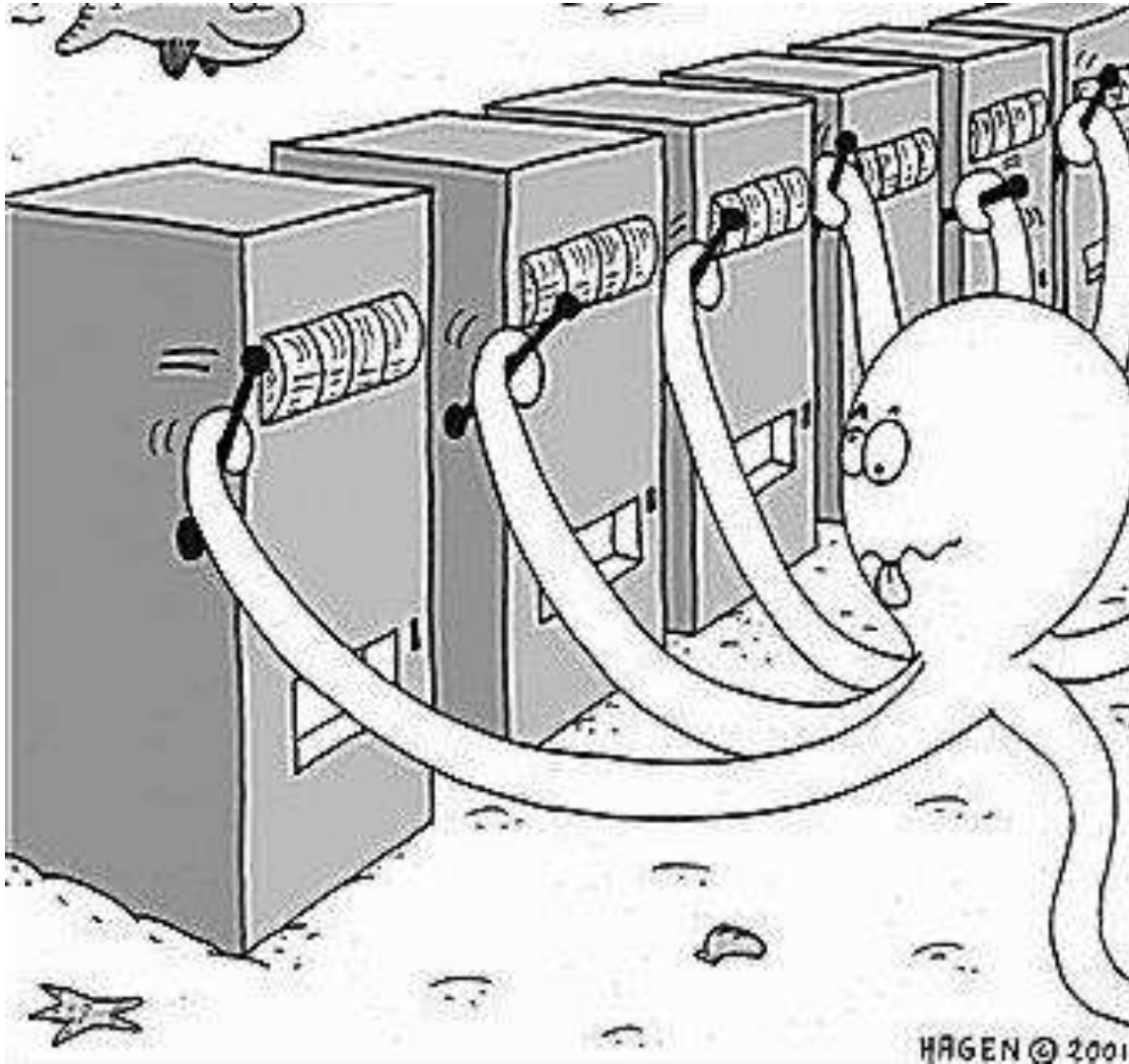
Вероятностная среда



Обучение с подкреплением и вероятностная среда

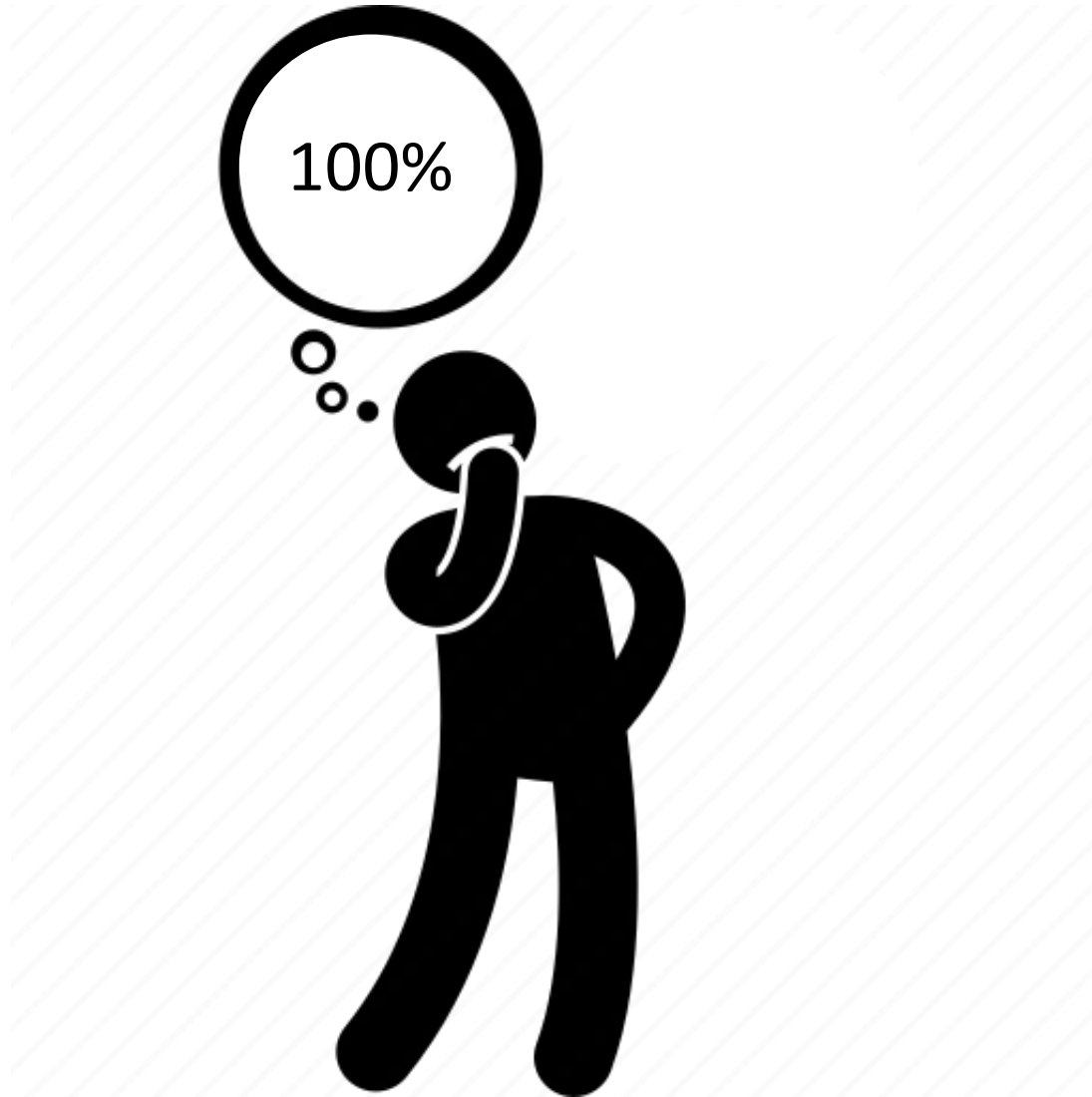


Обучение с подкреплением и вероятностная среда

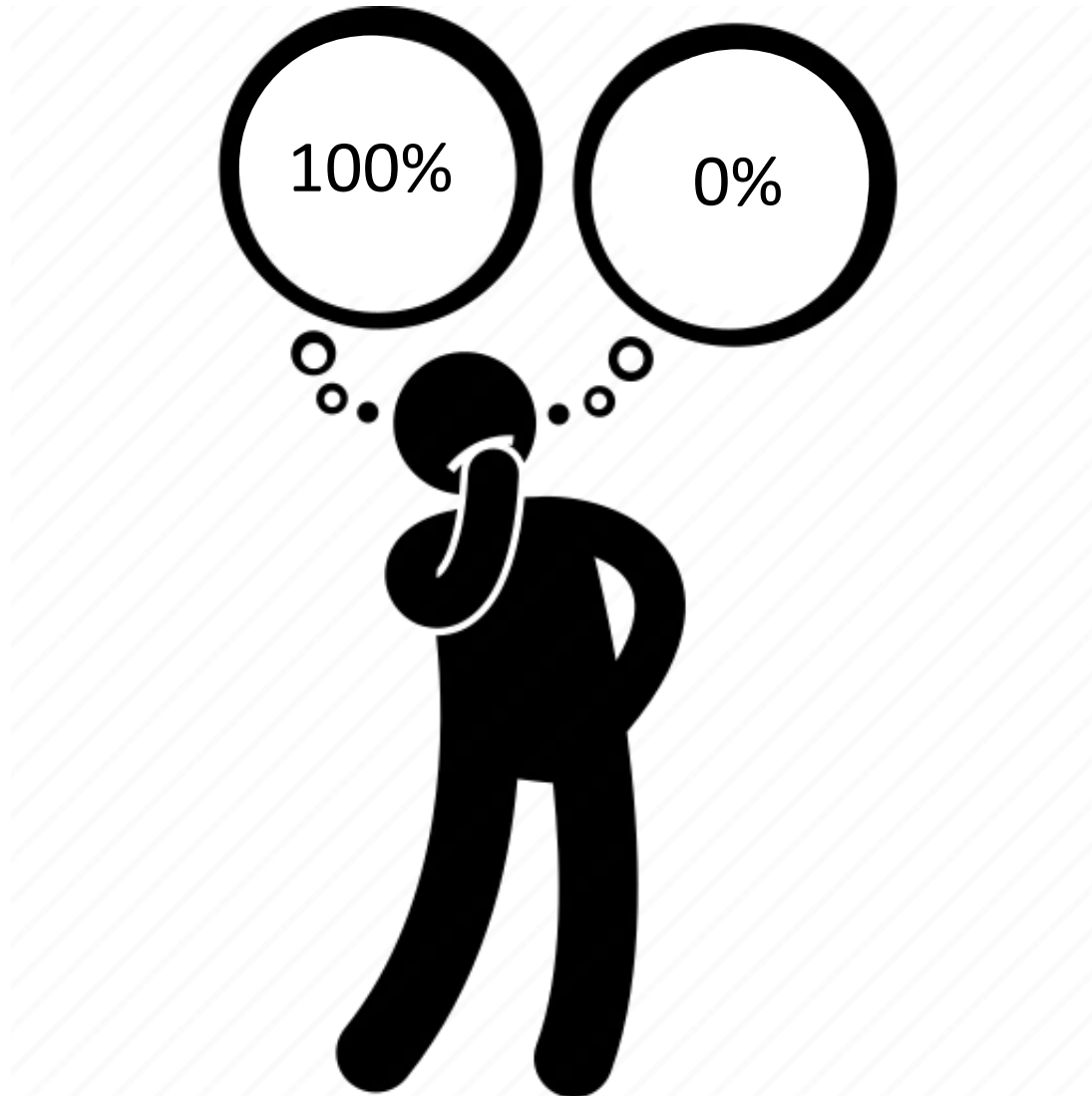


Экспериментальная задача
«**Многорукий бандит**»
(multiarmed bandit task)

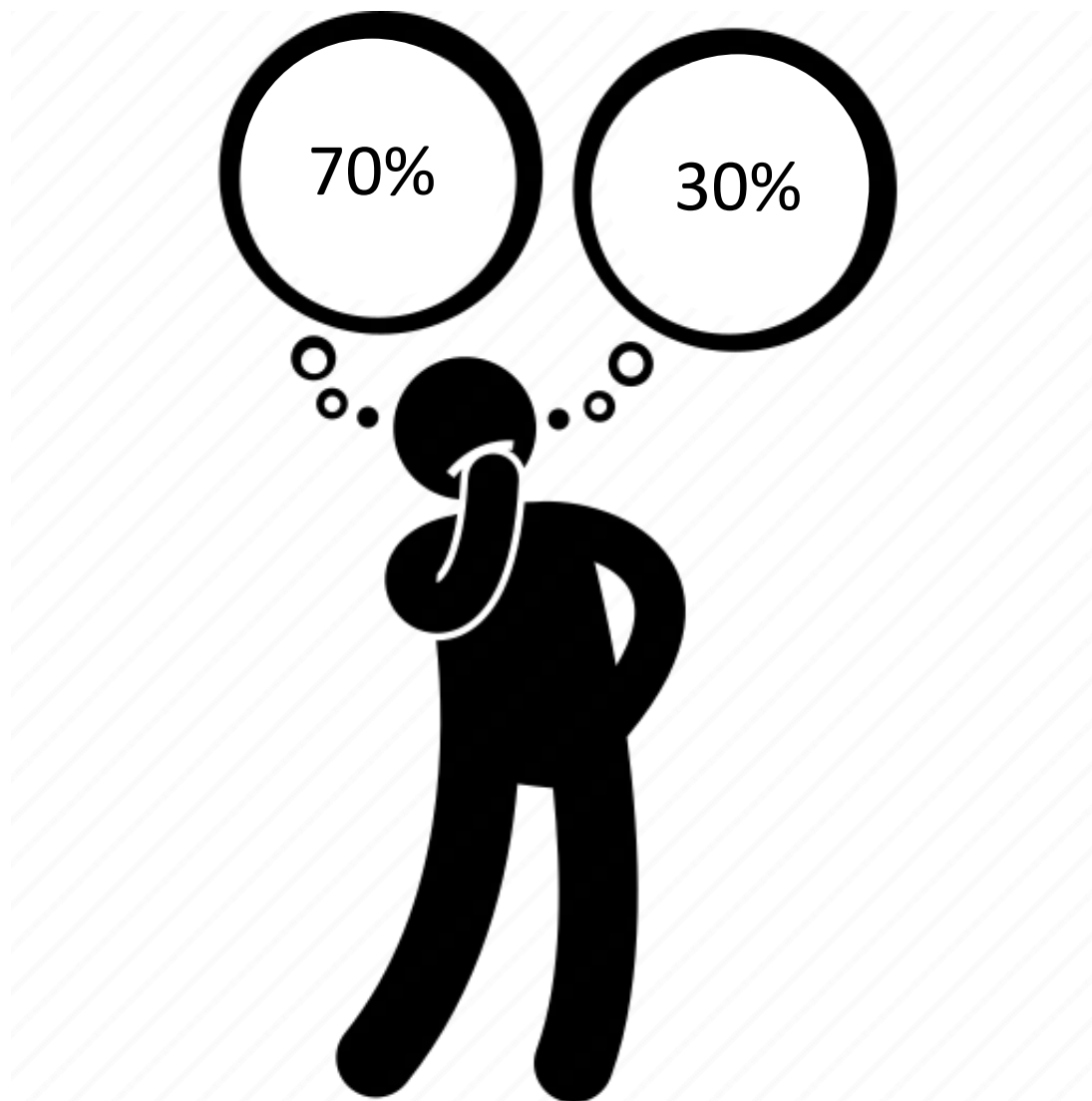
Обучение с подкреплением и вероятностная среда



Обучение с подкреплением и вероятностная среда



Обучение с подкреплением и вероятностная среда



Обучение с подкреплением и вероятностная среда



№ попытки	Кнопка 1: вероятность успеха 70%	Кнопка 2: вероятность успеха 30%
1	ПРОИГРЫШ	
2		УСПЕХ
3	УСПЕХ	
4		ПРОИГРЫШ
5	УСПЕХ	
6		ПРОИГРЫШ
7	ПРОИГРЫШ	
8		УСПЕХ
9	УСПЕХ	
10		ПРОИГРЫШ
11	УСПЕХ	
12		ПРОИГРЫШ
13	УСПЕХ	
14		ПРОИГРЫШ
15	ПРОИГРЫШ	
16		УСПЕХ
17	УСПЕХ	
18		ПРОИГРЫШ
19	УСПЕХ	
20		ПРОИГРЫШ

Обучение с подкреплением и вероятностная среда



№ попытки	Кнопка 1: вероятность успеха 70%	Кнопка 2: вероятность успеха 30%
1	ПРОИГРЫШ	
2		УСПЕХ
3	УСПЕХ	
4		ПРОИГРЫШ
5	УСПЕХ	
6		ПРОИГРЫШ
7	ПРОИГРЫШ	
8		УСПЕХ
9	УСПЕХ	
10		ПРОИГРЫШ
11	УСПЕХ	
12		ПРОИГРЫШ
13	УСПЕХ	
14		ПРОИГРЫШ
15	ПРОИГРЫШ	
16		УСПЕХ
17	УСПЕХ	
18		ПРОИГРЫШ
19	УСПЕХ	
20		ПРОИГРЫШ

№ попытки	Кнопка 1: вероятность успеха 70%	Кнопка 2: вероятность успеха 30%
21	УСПЕХ	
22	УСПЕХ	
23	УСПЕХ	
24	ПРОИГРЫШ	
25	УСПЕХ	
26	УСПЕХ	
27	ПРОИГРЫШ	
28	УСПЕХ	
29	УСПЕХ	
30	УСПЕХ	
31	ПРОИГРЫШ	
32	УСПЕХ	
33	УСПЕХ	
34	ПРОИГРЫШ	
35	УСПЕХ	
36	УСПЕХ	
37	ПРОИГРЫШ	
38	ПРОИГРЫШ	
39	УСПЕХ	
40	ПРОИГРЫШ	

Обучение с подкреплением и вероятностная среда



№ попытки	Кнопка 1: вероятность успеха 70%	Кнопка 2: вероятность успеха 30%
1	ПРОИГРЫШ	
2		УСПЕХ
3	УСПЕХ	
4		ПРОИГРЫШ
5	УСПЕХ	

Человек начинает предпочитать кнопку 1, и из-за этого он перестает нажимать на кнопку 2, и в результате он перестает получать информацию о степени "выгодности" кнопки 2.

16		УСПЕХ
17	УСПЕХ	
18		ПРОИГРЫШ
19	УСПЕХ	
20		ПРОИГРЫШ

№ попытки	Кнопка 1: вероятность успеха 70%	Кнопка 2: вероятность успеха 30%
21	УСПЕХ	
22	УСПЕХ	
23	УСПЕХ	
24	ПРОИГРЫШ	
25	УСПЕХ	
26	УСПЕХ	
27	ПРОИГРЫШ	
28	УСПЕХ	
29	УСПЕХ	
30	УСПЕХ	
31	ПРОИГРЫШ	
32	УСПЕХ	
33	УСПЕХ	
34	ПРОИГРЫШ	
35	УСПЕХ	
36	УСПЕХ	
37	ПРОИГРЫШ	
38	ПРОИГРЫШ	
39	УСПЕХ	
40	ПРОИГРЫШ	

Обучение с подкреплением и вероятностная среда

Всем живым системам имманентно присуща тенденция к конструированию **внутренней модели** среды, с которой они взаимодействуют (Parr, Friston, 2017; Smith, Badcock & Friston, 2021).

Обучение с подкреплением и вероятностная среда

Всем живым системам имманентно присуща тенденция к конструированию **внутренней модели** среды, с которой они взаимодействуют (Parr, Friston, 2017; Smith, Badcock & Friston, 2021).

Каждое расхождение наблюдаемого результата с его субъективным прогнозом указывает на неопределенность ситуации. Неопределенность несет в себе потенциальную угрозу для выживания и сохранения целостности организма.

Обучение с подкреплением и вероятностная среда

Всем живым системам имманентно присуща тенденция к конструированию **внутренней модели** среды, с которой они взаимодействуют (Parr, Friston, 2017; Smith, Badcock & Friston, 2021).

Каждое расхождение наблюдаемого результата с его субъективным прогнозом указывает на неопределенность ситуации. Неопределенность несет в себе потенциальную угрозу для выживания и сохранения целостности организма. Адаптивным ответом на неопределенность является рекуррентная корректировка внутренней модели среды. Любой исход собственных действий, в том числе и негативная обратная связь на совершенное действие, несет в себе эпистемическую ценность для субъекта, снижая для него уровень внутренней неопределенности (Parr, Friston, 2017).

Обучение с подкреплением и вероятностная среда



В вероятностной среде внутренняя модель должна включать в себя вероятности разных исходов ("риски"): если при выборе "ожидаемо выгодной" опции Вы иногда проигрываете, то это не повод сразу же отказываться от нее в будущем.

Обучение с подкреплением и вероятностная среда



В вероятностной среде внутренняя модель должна включать в себя вероятности разных исходов ("риски"): если при выборе "ожидаемо выгодной" опции Вы иногда проигрываете, то это не повод сразу же отказываться от нее в будущем.

Однако модель требует постоянного уточнения, поскольку:

- Человек строит свою внутреннюю модель на основе небольшого ограниченного количества отдельных проб – поэтому модель может оказаться ошибочной (в силу случайных флуктуаций, или потому, что человек неправильно уловил закономерность как таковую).

Обучение с подкреплением и вероятностная среда



В вероятностной среде внутренняя модель должна включать в себя вероятности разных исходов ("риски"): если при выборе "ожидаемо выгодной" опции Вы иногда проигрываете, то это не повод сразу же отказываться от нее в будущем.

Однако модель требует постоянного уточнения, поскольку:

- Человек строит свою внутреннюю модель на основе небольшого ограниченного количества отдельных проб – поэтому модель может оказаться ошибочной (в силу случайных флуктуаций, или потому, что человек неправильно уловил закономерность как таковую).
- Условия (вероятности) могли измениться, и выгодный выбор мог стать невыгодным, а невыгодный – выгодным (волатильность).

Обучение с подкреплением и вероятностная среда



В вероятностной среде внутренняя модель должна включать в себя вероятности разных исходов ("риски"): если при выборе "ожидаемо выгодной" опции Вы иногда проигрываете, то это не повод сразу же отказываться от нее в будущем.

Однако модель требует постоянного уточнения, поскольку:

- Человек строит свою внутреннюю модель на основе небольшого ограниченного количества отдельных проб – поэтому модель может оказаться ошибочной (в силу случайных флуктуаций, или потому, что человек неправильно уловил закономерность как таковую).
- Условия (вероятности) могли измениться, и выгодный выбор мог стать невыгодным, а невыгодный – выгодным (волатильность).

Продолжая раз за разом выбирать одну и ту же опцию в соответствии со своей внутренней моделью, человек лишает себя возможности исследовать другие опции.

Обучение с подкреплением и вероятностная среда



В итоге человек оказывается перед дилеммой между **использованием** и **исследованием**: либо продолжать полностью полагаться на приобретенный ранее собственный опыт (внутреннюю модель), либо проверять малознакомые опции.

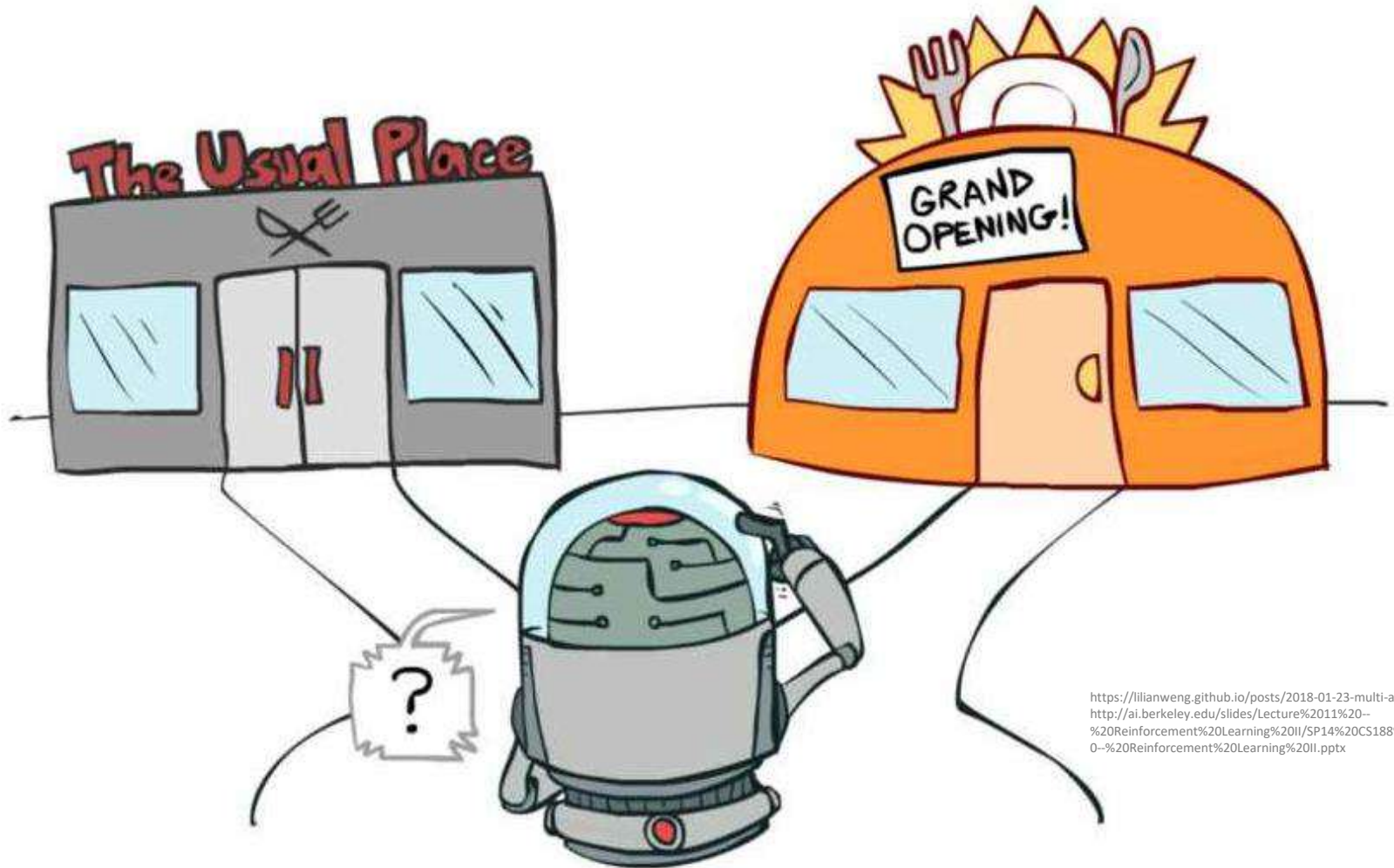
Обучение с подкреплением и вероятностная среда



В итоге человек оказывается перед дилеммой между **использованием** и **исследованием**: либо продолжать полностью полагаться на приобретенный ранее собственный опыт (внутреннюю модель), либо проверять малознакомые опции.

Выбор малознакомой опции может быть сопряжен с внутренним **конфликтом** (Cogliati Dezza et al., 2017; Daw et al., 2006): человек обрекает себя на повышенную вероятность проигрыша (ведь он поступает против своей внутренней модели, т.е. ему приходится подавлять "предоминирующую" реакцию выбора выгодной опции), – ради получения информации, которая потенциально может оказаться полезной когда-нибудь в будущем.

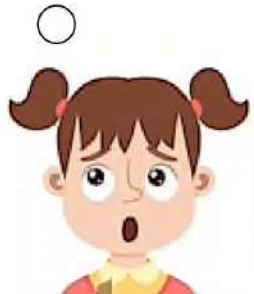
Дилемма между стратегиями использования и исследования



<https://lilianweng.github.io/posts/2018-01-23-multi-armed-bandit/>
<http://ai.berkeley.edu/slides/Lecture%2011%20-%20Reinforcement%20Learning%20I/SP14%20CS188%20Lecture%2011%20-%20Reinforcement%20Learning%20II.pptx>

Дилемма между стратегиями **использования** и **исследования** (exploration-exploitation dilemma).

Разновидности стратегии исследования



Chocolate



Hibiscus



Toblerone



Spinach

Разновидности стратегии исследования

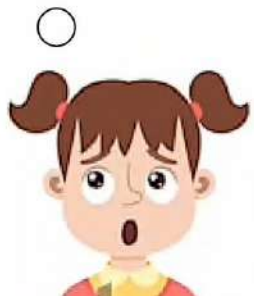


Exploitation



Разновидности стратегии исследования

Stick with what
is **known**
OR
try **something else**



Exploitation



Chocolate

Exploration



Hibiscus



Toblerone



Spinach

Разновидности стратегии исследования



Different strategies for exploration

- Go for novelty
- Go for what you expect to like
- Pick with 'closed-eyes'

Exploitation



Exploration



Разновидности стратегии исследования

"случайное" исследование – человек начинает выбирать доступные ему опции случайным образом, т.е. как бы отказывается от уже выученных им правил задачи, и пробует разные варианты.

- **без учета ценности** (чисто случайный выбор «с закрытыми глазами», это наименее затратный метод; к нему особенно часто прибегают маленькие дети)

- **с учетом ценности** (требует некоторых когнитивных затрат для учета предыдущих выигрышей и проигрышей)

"направленное" исследование – принятие решения, при котором человек сознательно жертвует сиюминутным выигрышем ради обретения новой информации – в надежде, что это поможет ему в будущем (наиболее затратный метод, так как требует учета степени неопределенности по каждой опции, и построения гипотез).

Разновидности стратегии исследования

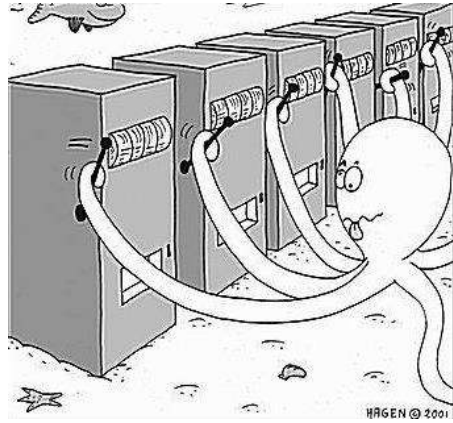
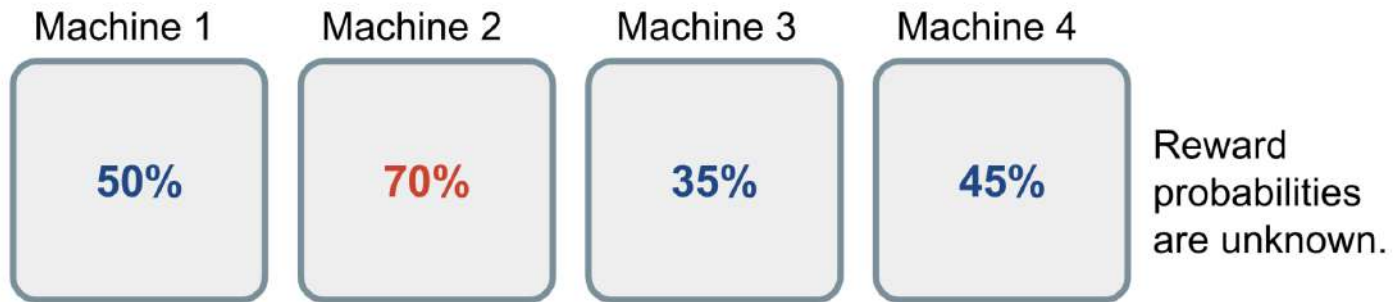
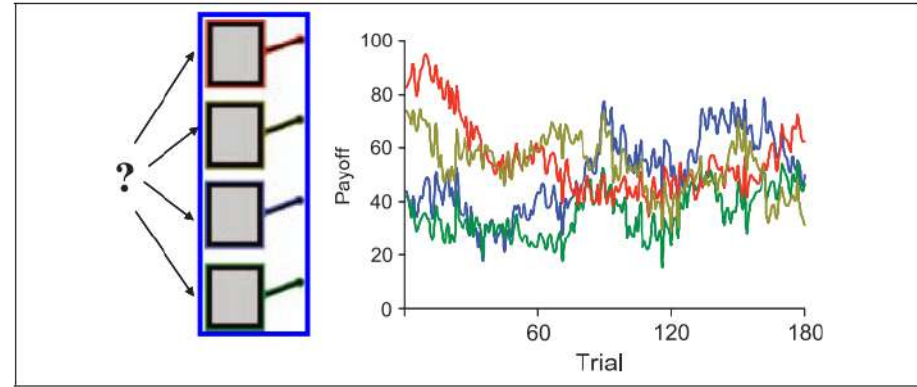


Figure 1. The four-armed bandit task. Participants made repeated choices between four slot machines. Unlike standard slots, the mean payoffs of the four machines changed gradually and independently from trial to trial (four colored lines). Participants were encouraged to earn as many points as possible during the experiment. After the experiment, each choice was classified as exploitative or exploratory using a computational model of reinforcement learning.

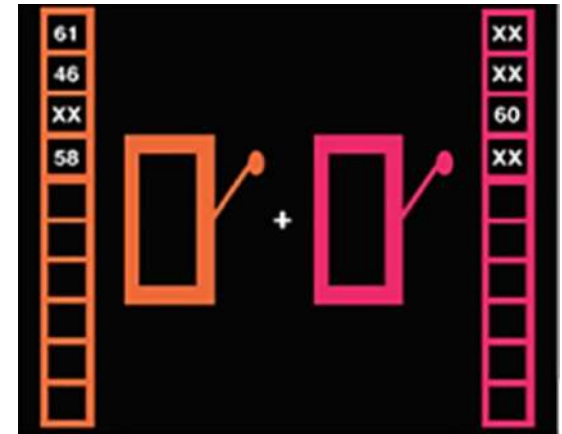


Which machine to pick next?

Экспериментальная задача «**беспокойный многорукий бандит**» (restless multiarmed bandit).

Задачи, в которых вероятность выигрыша изменяется, провоцируют участника эксперимента только на **случайное исследование**.

Разновидности стратегии исследования



horizon 1	1	2	3	4	1											
horizon 6	1	2	3	4	1	2	3	4	5	6						
horizon 11	1	2	3	4	1	2	3	4	5	6	7	8	9	10	11	
	forced-choice trials				free-choice trials											

Экспериментальная задача «Горизонт» дает возможность изучать **направленное исследование**.

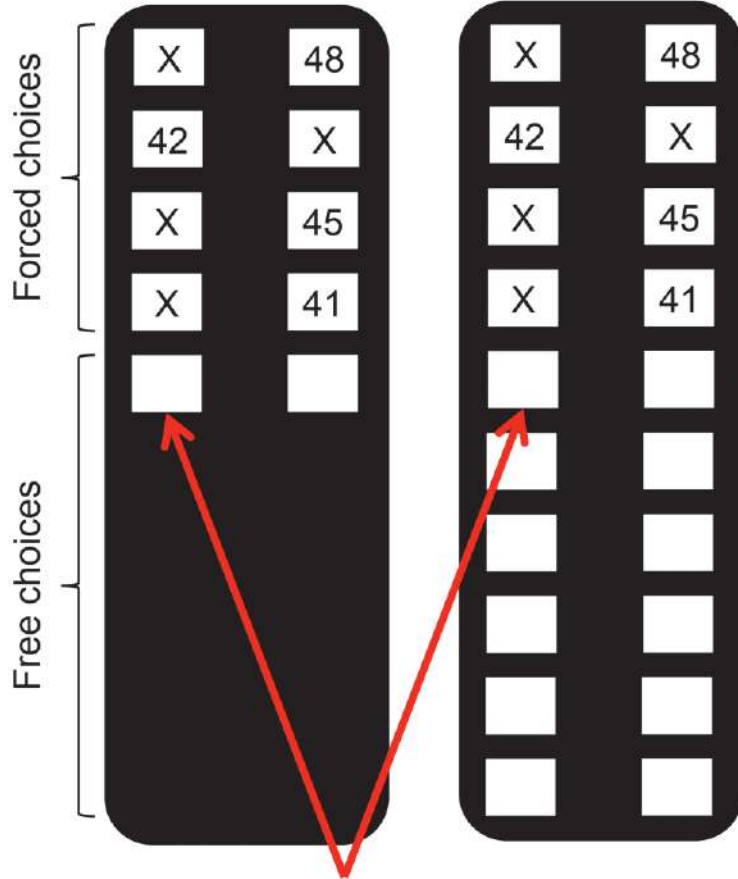
При коротком горизонте человек будет следовать стратегии использования, а при длинном – стратегии исследования.

Разновидности стратегии исследования

A

Horizon 1

Horizon 6

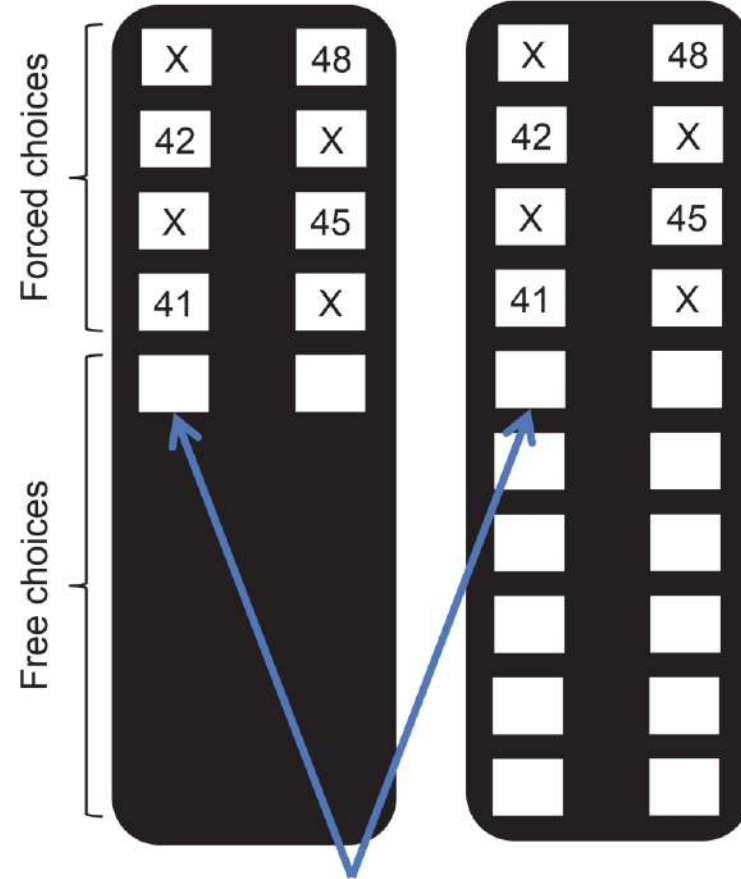


High information choice
Directed exploration

B

Horizon 1

Horizon 6



Low mean choice
Random exploration

Экспериментальная задача «Горизонт»

Figure 1. The horizon task. Participants make a series of decisions between two one-armed bandits that pay out probabilistic rewards with unknown means. At the start of each game, ‘forced-choice’ trials give participants partial information about the mean of each option. We use the forced-choice trials to set up one of two information conditions: (A) an unequal (or [1 3]) condition in which participants see 1 play from one option and 3 plays from the other and (B) an equal (or [2 2]) condition in which participants see 2 plays from both options. A model-free measure of directed exploration is then defined as the change in information seeking with horizon in the unequal condition (A). Likewise a model-free measure of random exploration is defined as the change choosing the low mean option in the equal condition (B).

Разновидности стратегии исследования

Bandit task

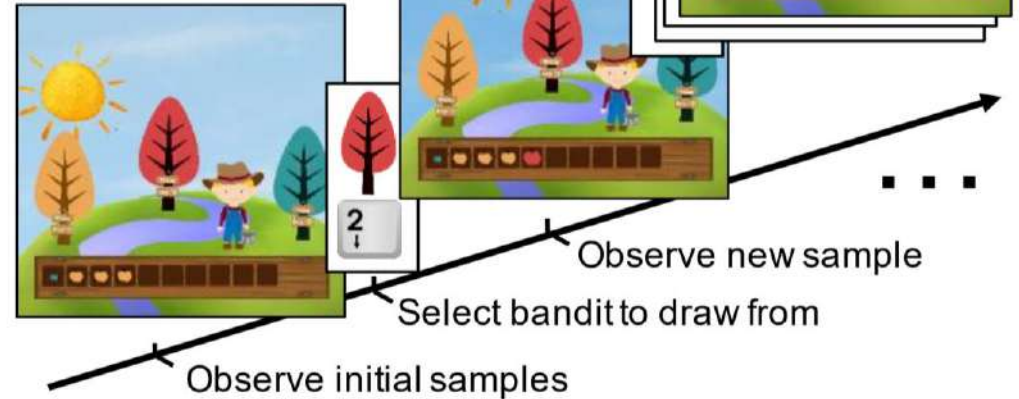
Bandits → Trees

Draw from a bandit → Apple

Reward → Apple size

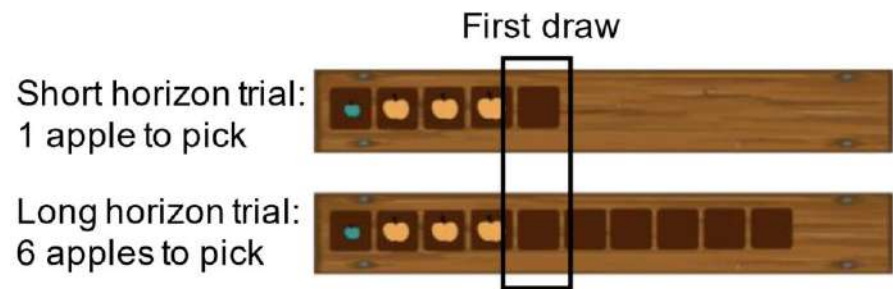


Game START: 3 of the 4 bandits are used



Game END: Reward is the sum of sample sizes

Экспериментальная задача «Ферма Мэгги» (Maggie's Farm task) – модификация задача «Горизонт»



Как изучать стратегию исследования?

THE JOURNAL OF COMPARATIVE NEUROLOGY 493:99–110 (2005)

Adaptive Gain and the Role of the Locus Coeruleus–Norepinephrine System in Optimal Performance

GARY ASTON-JONES^{1*} AND JONATHAN D. COHEN^{2,3}

¹Laboratory of Neuromodulation and Behavior, Department of Psychiatry, University of Pennsylvania, Philadelphia, Pennsylvania 19104

²Department of Psychology, Center for the Study of Brain, Mind and Behavior, Princeton University, Princeton, New Jersey 08540

³Department of Psychiatry Western Psychiatric Institute and Clinic University of Pittsburgh, Pittsburgh, Pennsylvania 15260

ABSTRACT

Historically, the locus coeruleus–norepinephrine (LC-NE) system has been implicated in arousal, but recent findings suggest that this system plays a more complex and specific role in the control of behavior than investigators previously thought. We review neurophysiological, anatomical, and modeling studies in monkey that support a new theory of LC-NE function. LC neurons exhibit two modes of activity, phasic and tonic. Phasic LC activation is driven by the outcome of task-related decision processes and is proposed to facilitate ensuing behaviors and to help optimize task performance. When utility in the task wanes, LC neurons exhibit a tonic activity mode, associated with disengagement from the current task and a search for alternative behaviors. Monkey LC receives prominent, direct inputs from the anterior cingulate (ACC) and orbitofrontal cortices (OFC), both of which are thought to monitor task-related utility. We propose that these prefrontal areas produce the above patterns of LC activity to optimize the utility of performance on both short and long time scales. *J. Comp. Neurol.* 493:99–110, 2005. © 2005 Wiley-Liss, Inc.

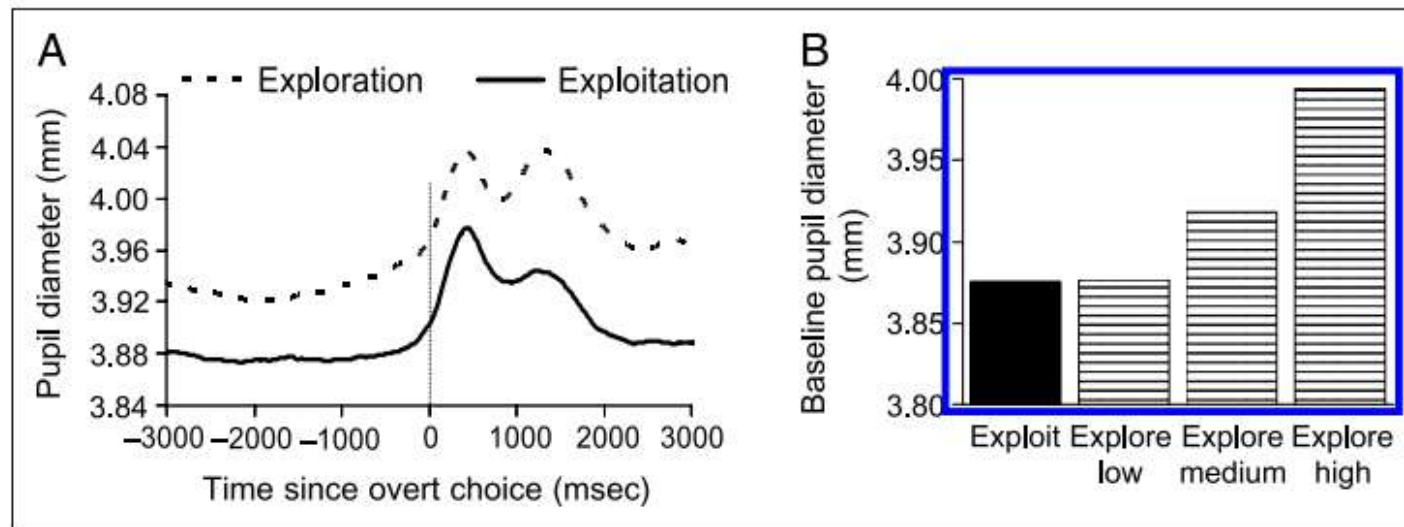
В этой работе была предложена теория, согласно которой причиной перехода к стратегии исследования заключалась в тонической норадренергической активации в головном мозге.

Как изучать стратегию исследования?

Pupil Diameter Predicts Changes in the Exploration–Exploitation Trade-off: Evidence for the Adaptive Gain Theory

Marieke Jepma and Sander Nieuwenhuis

Figure 2. Pupil diameter on exploration and exploitation trials. (A) Time course of grand-average pupil diameter aligned to the keypress indicating the selection of a slot machine for exploratory and exploitative choices. (B) Average baseline pupil diameter for exploitative choices (black bar), and exploratory choices with a low, medium, and high degree of exploration (striped bars).



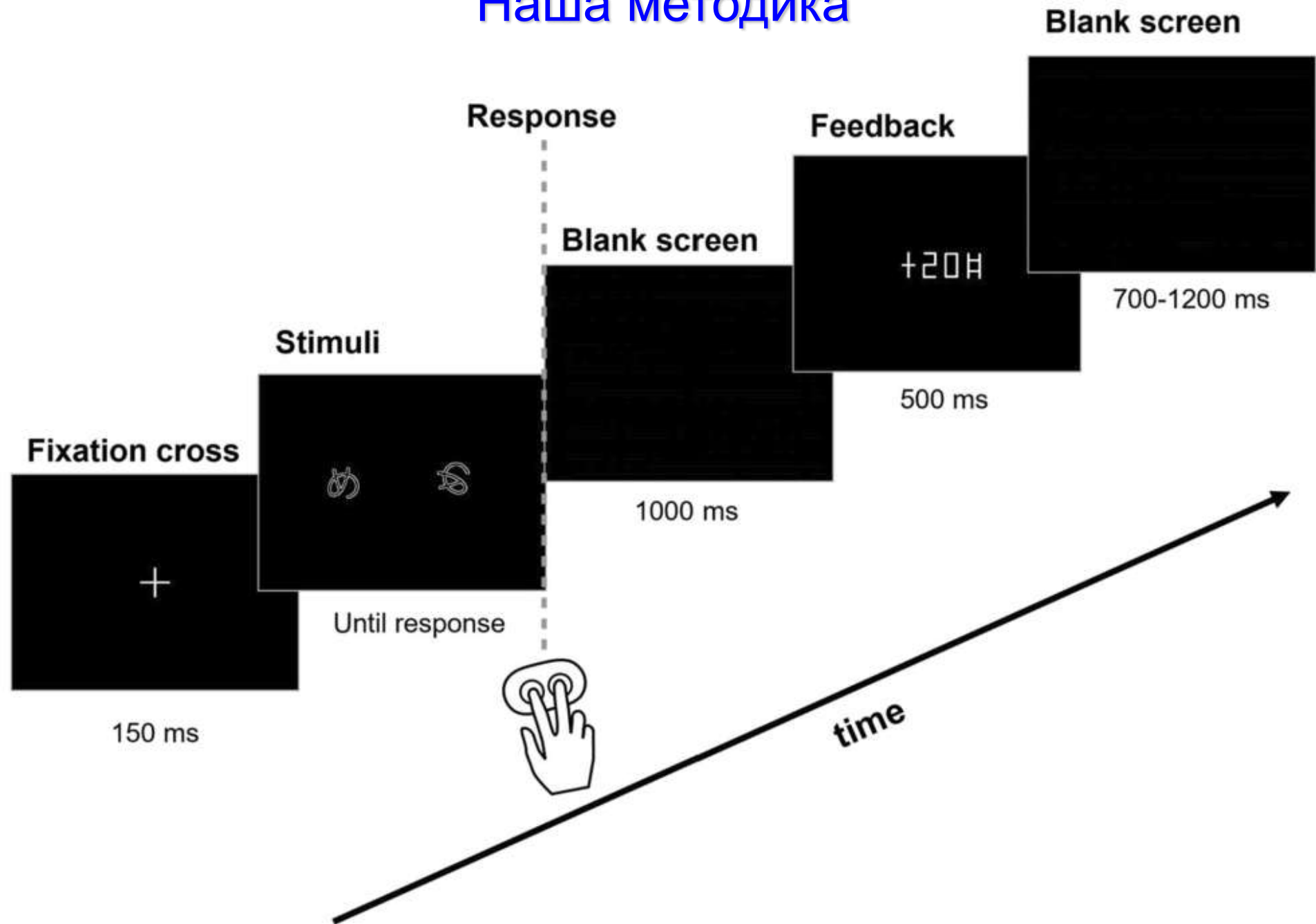
Показано тоническое расширение зрачка при стратегии исследования.
В этой работе использовали экспериментальную задачу «**беспокойный
многорукий бандит**»

Как изучать стратегию исследования?

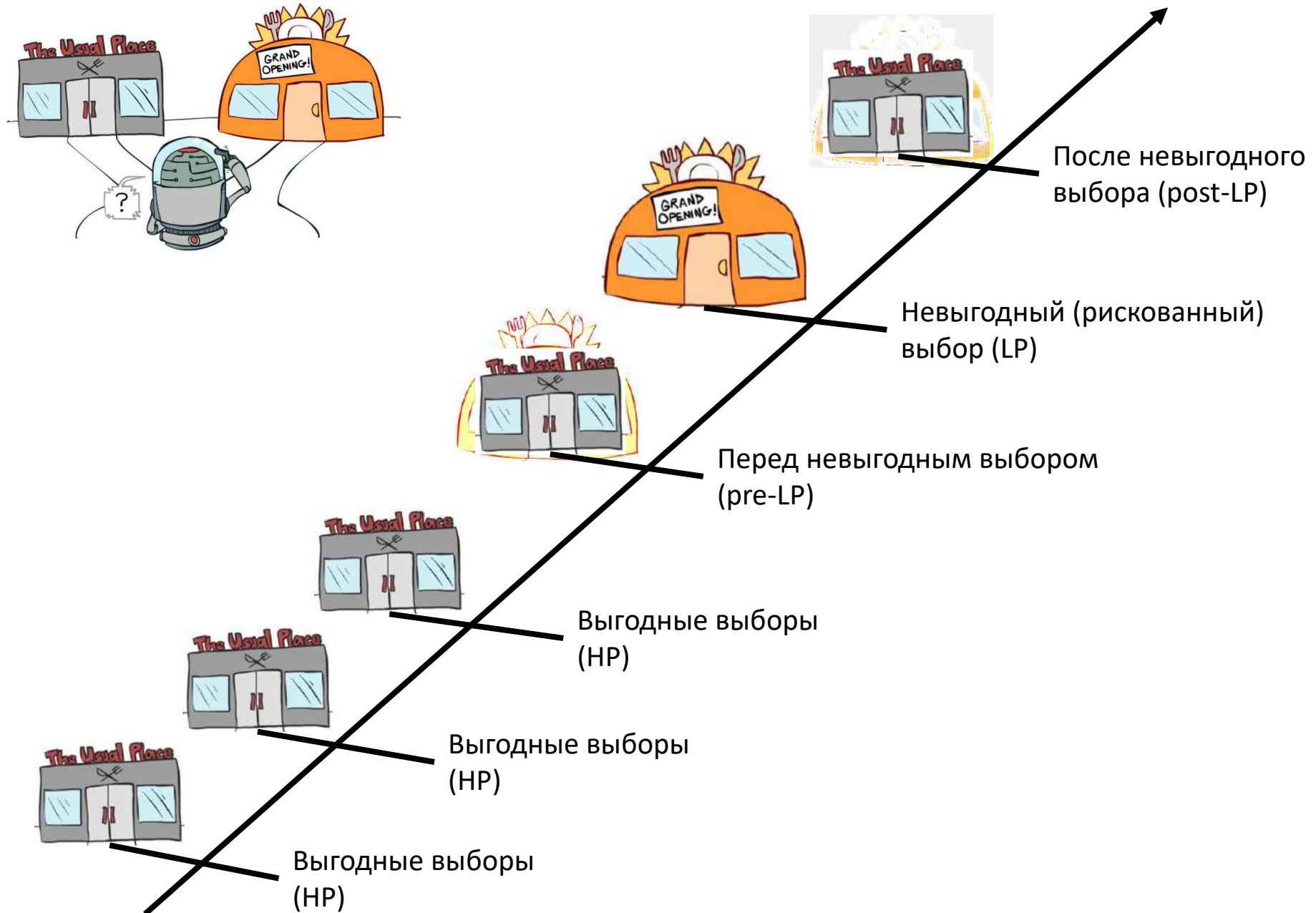
В подавляющем большинстве научных работ за последние полтора десятилетия фокус внимания ученых был направлен лишь на одну из форм стратегии исследования – на так называемое "случайное" исследование. В этих работах экспериментаторы поступали так, что привычная человеку выгодная опция неожиданно становилась невыгодной. В результате человек начинал хаотически выбирать разные доступные ему опции, т.е. как бы отказывался от уже выученных им правил задачи, и начинал пробовать все варианты наугад.

Хотя такая ситуация "случайного" исследования действительно нередко встречается в жизни, интереснее и полезнее проанализировать стратегию "направленного" исследования. Так обозначают принятие решения, при котором человек сознательно жертвует сиюминутным выигрышем ради обретения новой информации – в надежде, что это поможет ему в будущем.

Наша методика



Наша методика



Спасибо за внимание!

<http://megmoscow.ru/>

https://vk.com/Moscow_MEG_center/

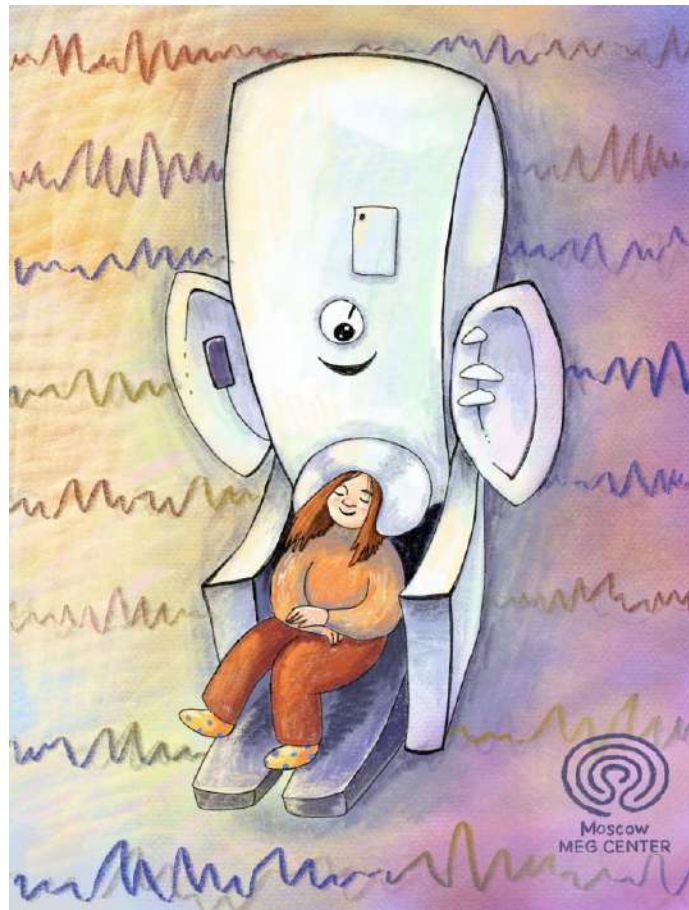


Рисунок К.Сайфулиной