

# Как человек принимает решения в условиях риска и неопределенности?

## РАЗБОР пупиллометрического исследования

Г.Л. Козунова  
к.п.н., ст. научный сотрудник МЭГ-центра

2024

Cognition, Attention, & Behavioral Neuroscience 2022 | 22:700–713  
<https://doi.org/10.1007/s10142-022-00962-z>

RESEARCH ARTICLE



### Pupil dilation and response slowing distinguish deliberate explorative choices in the probabilistic learning task

Galina L. Kozunova<sup>1</sup> · Kaznisa L. Soyfalina<sup>1</sup> · Andrey O. Prokofyev<sup>1</sup> · Vladimír A. Medvedev<sup>1</sup> · Anna M. Rytikova<sup>1</sup> · Tatjana A. Stegannova<sup>1</sup> · Boris V. Chernyshov<sup>1</sup>

Accepted: 7 March 2022 / Published online: 1 April 2022  
© The Author(s) 2022

#### Abstract

This study examined whether pupil size and response time would distinguish directed exploration from random exploration and exploitation. Eighty-nine participants performed the two-choice probabilistic learning task while their pupil size and response time were continuously recorded. Using LMM analysis, we estimated differences in the pupil size and response time between the advantageous and disadvantageous choices as a function of learning success, i.e., whether or not a participant has learned the probabilistic contingency between choices and their outcomes. We proposed that before a trial, value of each choice became known to a decision maker, both advantageous and disadvantageous choices represented a random exploration of the two options with an equally uncertain outcome, whereas the same choices after learning manifested exploitation and direct exploration strategies, respectively. We found that disadvantageous choices were associated with increases both in response time and pupil size, but only after the participants had learned the choice reward contingencies. For the pupil size, this effect was strongly amplified for those disadvantageous choices that immediately followed gains as compared to losses in the preceding choice. Pupil size modulations were evident during the behavioral choice rather than during the mental baseline. These findings suggest that occasional disadvantageous choices, which violate the acquired internal utility model, represent directed exploration. This exploratory strategy shifts choice priorities in favor of information seeking and its associated behavioral consequences are mainly driven by the contrast between the behavioral gain of the intended exploratory choice and its strong alternative, which has already proven to be more rewarding.

**Keywords** Pupil size · Response time · Direct and random exploration · Probability learning · Control

“A Nail in the hand is worth two in the bush” – will you follow this common wisdom, or will you ever abandon something good you already have and venture into the unknown in the vague hope of a bigger win? In a probabilistic environment, people usually tend to imagine hidden regularities in the outcomes of their actions, even when no such regularities actually exist (Eklund & Tunney, 2017; Lieder & Coricaria, 2007). Attempting to test these various and often a lucky break, people explore apparently disadvantageous options instead of just sticking to familiar profitable ones (Shanks et al., 2002). In fact, by doing so in probabilistic experimental tasks involving truly random and mutually

independent choice outcomes, they usually fail to maximize their profits (Guth & Harvi, 2005; Lieder & Coricaria, 2007; Vulkan, 2000).

In contrast to typical artificial experimental conditions, outcomes of one's actions in real life may be nonrandom and interdependent, e.g., the outcome of the next trial may be a consequence of the outcome of the previous trial. In this respect, exploring uncertain options instead of continuing to plow a worn-out rewarding alternative might bring new information about possible rewards, and thus might increase payoffs in the long run (Cognigni-Danza et al., 2017; Sijfalova et al., 2020). From this perspective, the occasional switches from the present value-driven advantageous response tendency to a disadvantageous choice in the probabilistic task may be considered as directed exploration—exploratory behavior that occurs when our desire for information overrides our need for reward. The crucial distinction between two qualitatively different types of

✉ Boris V. Chernyshov  
[b.chernyshov@mail.ru](mailto:b.chernyshov@mail.ru)

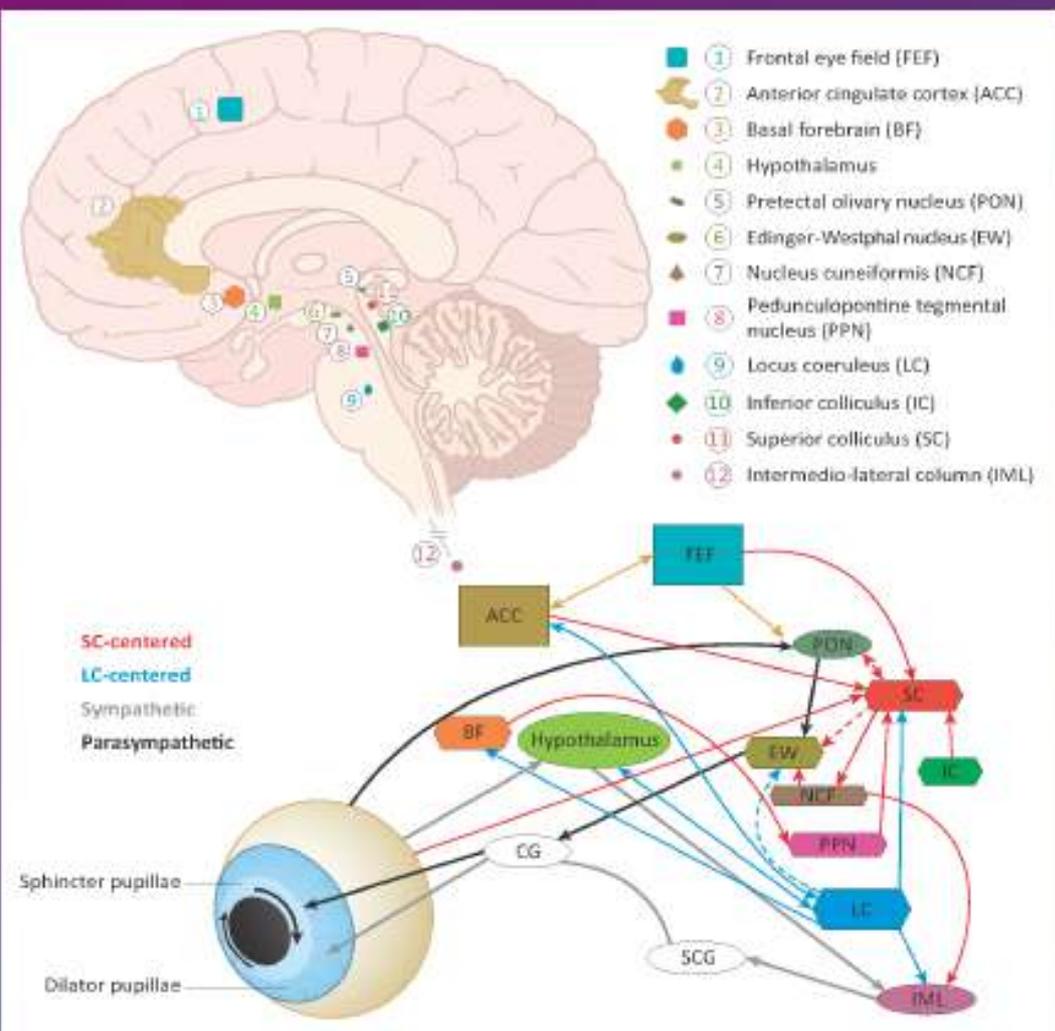
<sup>1</sup> Center for Neurocognitive Research (MNS-Center), Moscow State University of Psychology and Education, 20 Savitskaya str., Moscow 125645, Russia

# СОДЕРЖАНИЕ



- 1. Какие психологические процессы влияют на размер зрачка?**
- 2. Как возбуждение вегетативной нервной системы затрагивает высшие корковые механизмы принятия решений?**
- 3. Почему исследование новых стратегий решения знакомых задач сопровождается внутренним конфликтом?**
- 4. Что нового добавляет пупиллометрия в психологические исследования?**
- 5. Как мы моделировали принятие решений в условиях риска и неопределенности в оплачиваемом эксперименте на выборке из 63 взрослых добровольцев**

# Механизмы пупиллярного ответа на разнообразные события:



\* Относительное снижение освещенности

\* Когнитивные усилия при решении предельно трудной задачи

\* Перегрузка рабочей памяти

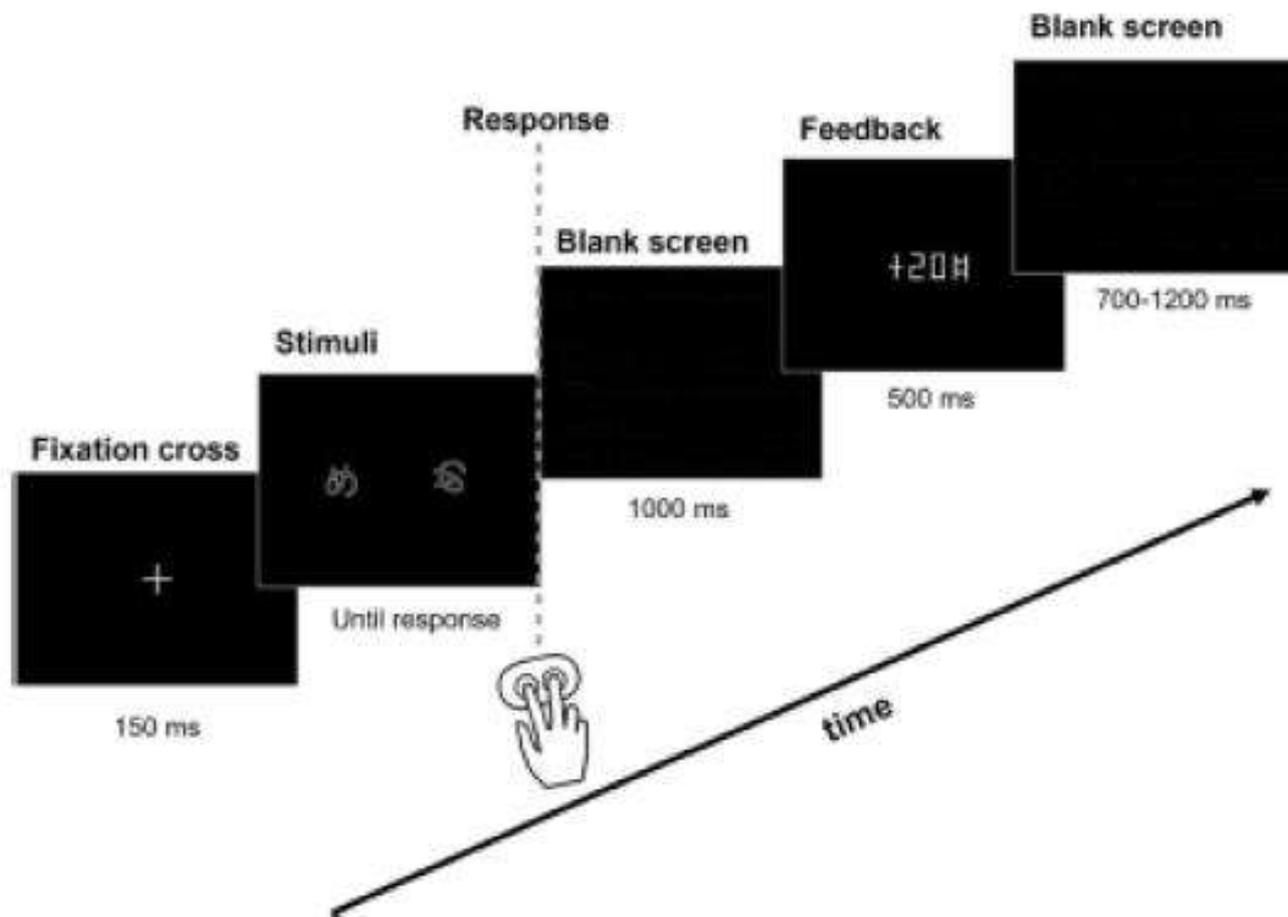
\* Высокая мотивация при увеличении размера ожидаемой награды или угрозе

\* Противоречивые стимулы или собственная ложь

\* Новая задача, незнакомая ситуация или внезапные изменения в привычном порядке

Strauch C. et al. Pupillometry as an integrated readout of distinct attentional networks //Trends in Neurosciences. – 2022.

# СХЕМА ЭКСПЕРИМЕНТА



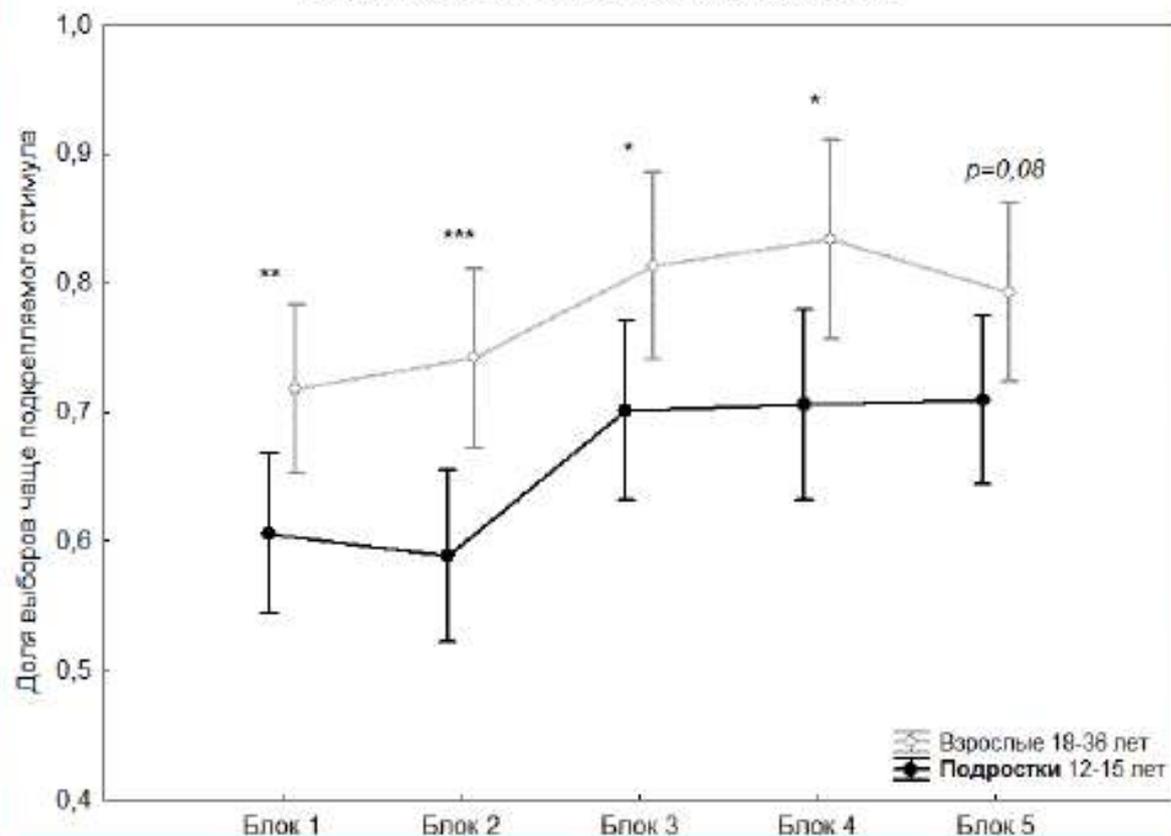
До эксперимента испытуемым сообщалось 2 факта:

1. обе картинки могут давать и положительный, и отрицательный результат, но одна из них будет приносить выигрыши чаще, чем другая

2. от Ваших выборов напрямую зависит сумма Вашего вознаграждения

# Как испытуемые учились методом проб и ошибок зарабатывать себе вознаграждение

Динамика обучения у подростков и взрослых

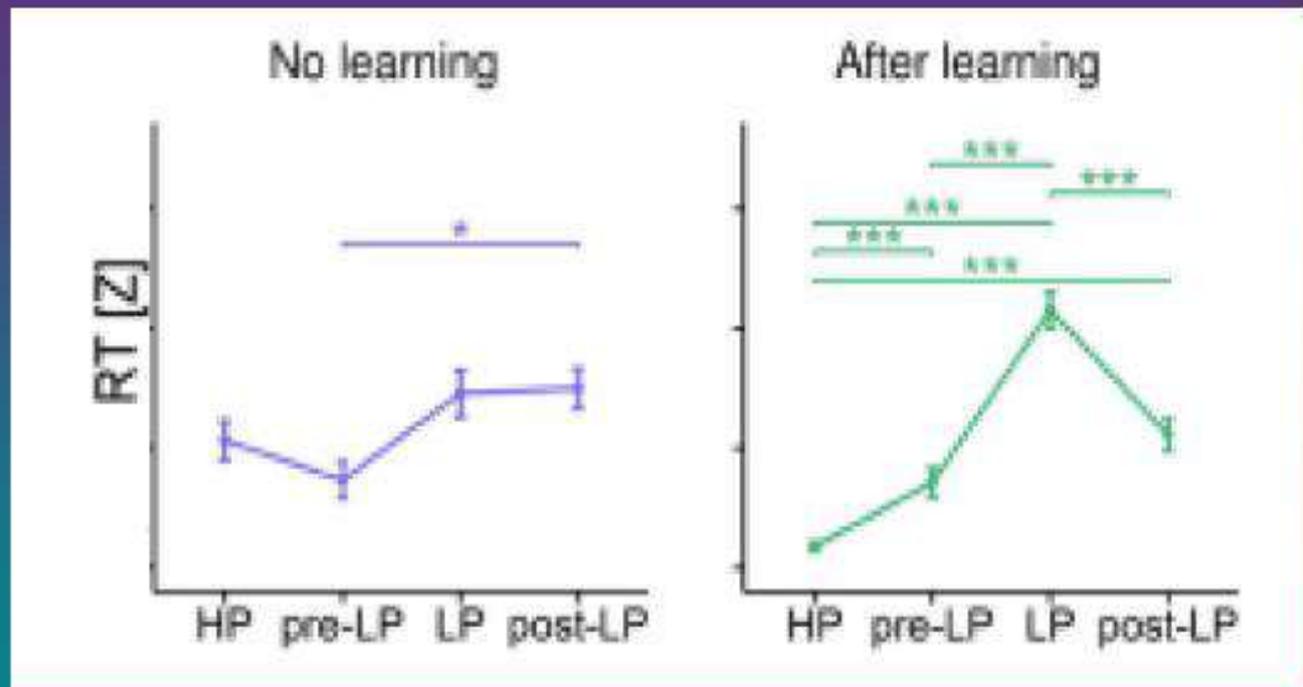


Эксперимент состоял из 5 аналогичных блоков, которые отличались друг от друга только картинками

К 4 блоку практически все испытуемые (61 из 63) научились выбирать ту картинку, которая чаще вознаграждается

Однако до конца эксперимента оставалась небольшая доля «необъяснимых» невыгодных выборов

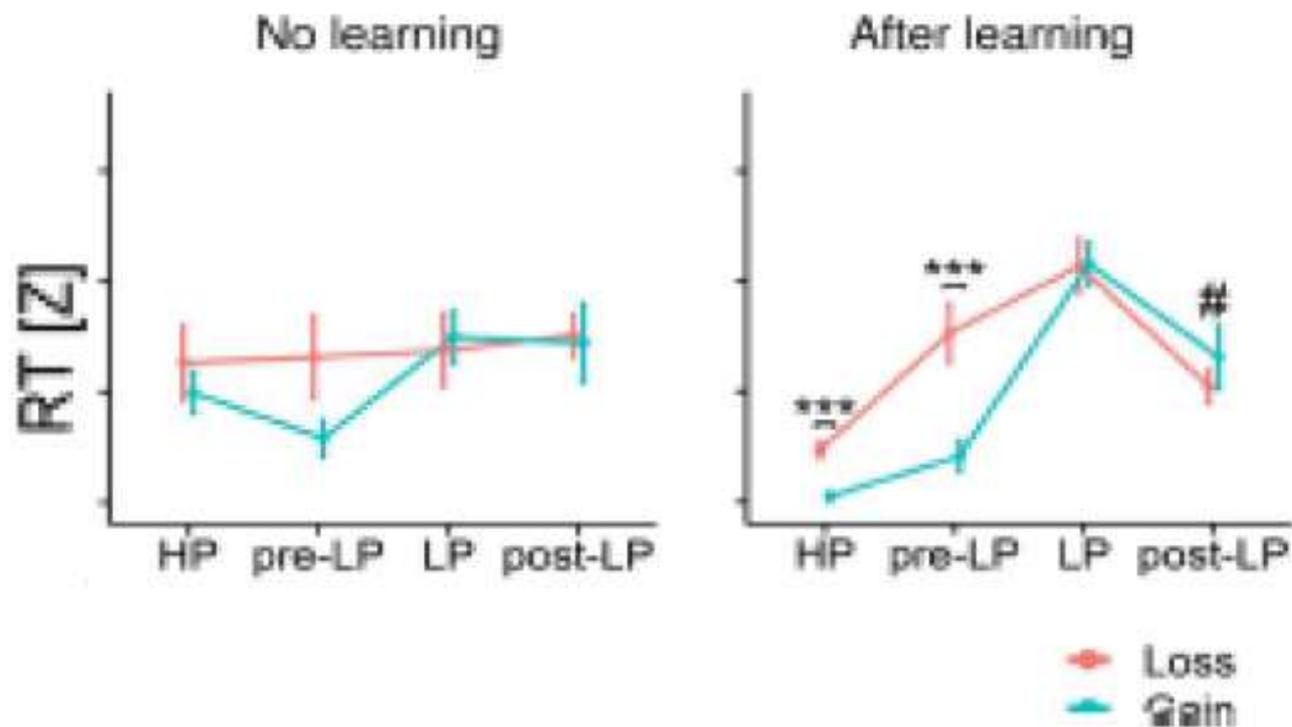
# Удлинение времени реакции как маркер неуверенности в положительном исходе



После появления предпочтения выгодной картинки редкие отклонения от стратегии использования правила сопровождались явным удлинением времени реакции — т. е., это был трудный выбор, требующий усиления внимания и самоконтроля

Задержка ответов наблюдалась уже за 1 ход до рискованного выбора и после него — это отражение латентных процессов подготовки к рискованному решению и анализа его последствий

Относительно редкие, но неизбежные штрафы при совершении  
выгодного выбора  
приводили к замедлению ответа в последующей пробе **ТОЛЬКО ПОСЛЕ**  
**ОБУЧЕНИЯ**

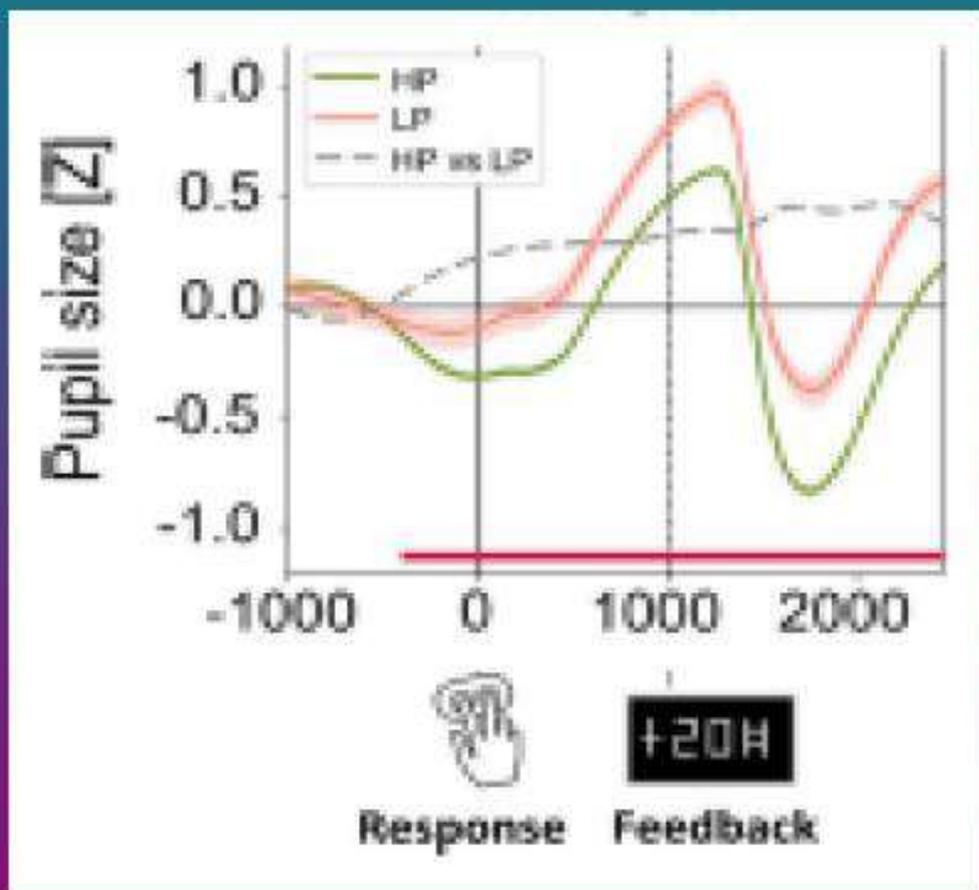


Чем чаще человек  
выбирает выгодную  
картинку, тем больше  
у него накапливается  
опыта наказаний за  
оптимальное в  
данных условиях  
поведение.

Почему не  
наблюдается  
признаков  
привыкания к этим  
неизбежным  
издержкам выгодной  
стратегии?

## Анализ пупиллометрии.

Как выбрать временной интервал для измерения зрачкового ответа?



В отличие от произвольных движений рук, зрачковые реакции управляются не корой головного мозга, а подкорковыми структурами

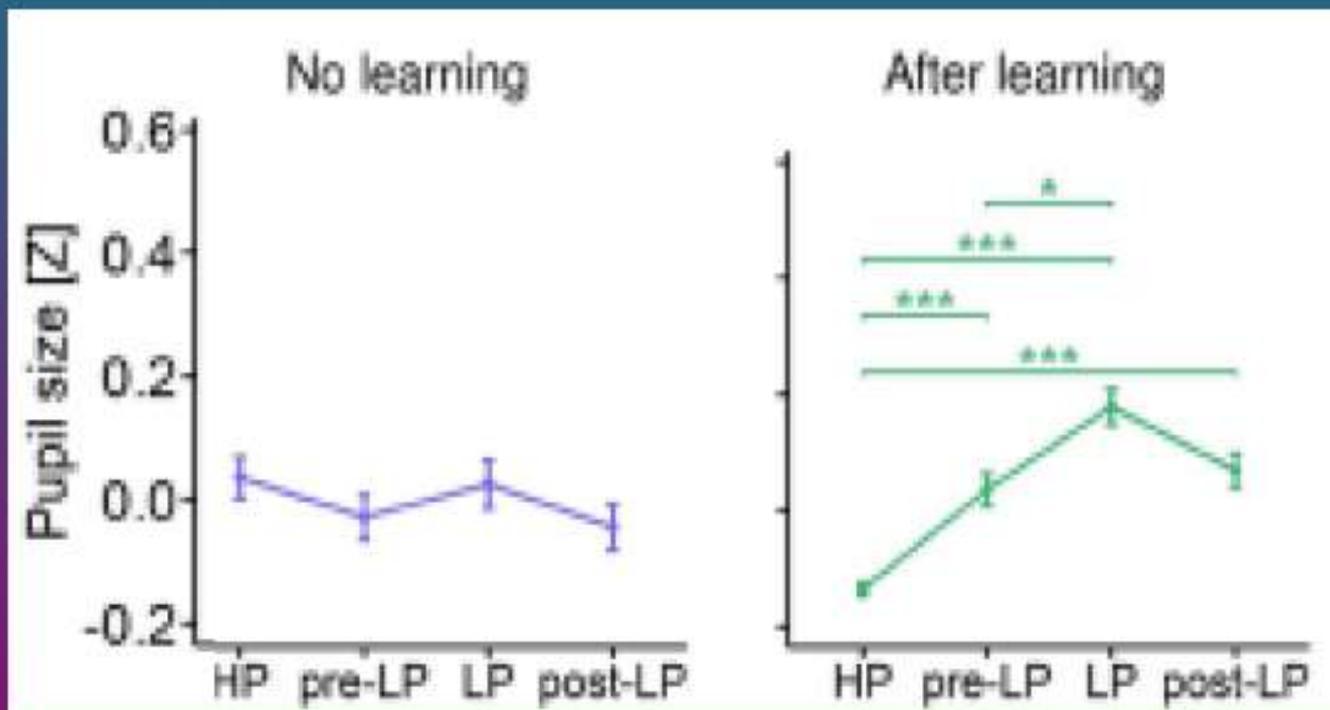
Поэтому зрачковые реакции не осознаются и не поддаются контролю

Как и все вегетативные реакции, зрачковые ответы управляются гладкими мышцами, и из-за этого расширение зрачка

- развивается с большей задержкой, чем моротный ответ

- длятся до нескольких секунд, и могут перекрывать по длительности несколько разных моротных ответов

## Редкие рискованные выборы на фоне уверенного применения выгодной стратегии вызывали расширение зрачка

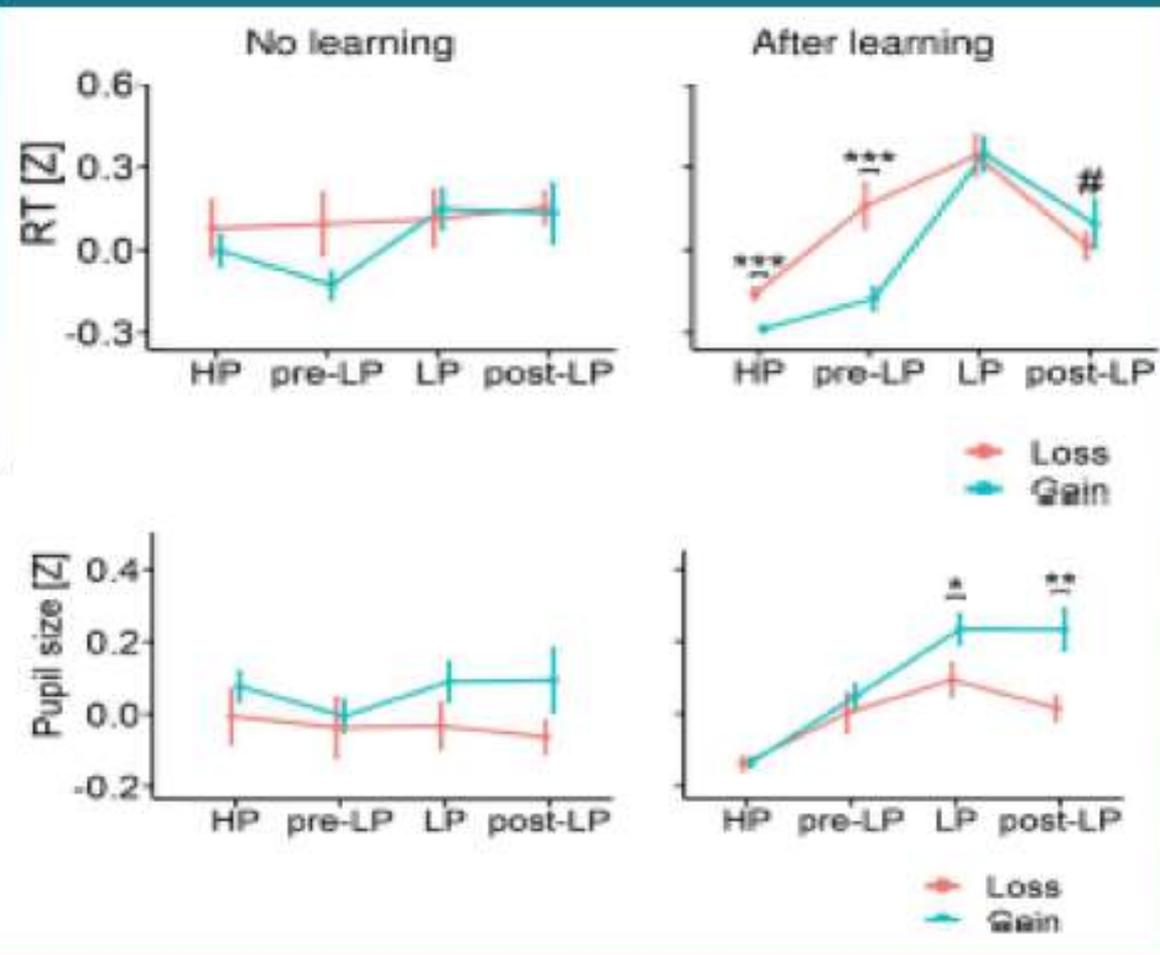


Усиление пупиллярного ответа начиналось в пробе, предшествующей переходу к риску

и продолжалось при возвращении к заведомо безопасной стратегии.

Точно такой же паттерн, как в отношении замедления времени реакции.

# Сравним, как влиял знак обратной связи на зрачковые ответы и на время реакции



Когда человек следовал выгодной стратегии, редкие неизбежные штрафы не вызывали у него вегетативной стрессовой реакции, но усиливали внимание и самоконтроль.

Когда человек отклонялся от выгодной стратегии, положительная обратная связь усиливала внутренний конфликт между использованием знания о вероятностной структуре задачи и рискованным поиском альтернативных способов получения награды.

# Выводы из проведенной работы и перспективы для дальнейших исследований



1. С помощью анализа размеров зрачка можно изучать не только состояние хаотичного поиска решения задачи, но и преднамеренный точечный поиск недостающей информации
2. Внутренний конфликт при отказе от заведомо выгодной стратегии в пользу проверки альтернативных гипотез вызывает фазическое расширение зрачка
3. Рассогласование результата уверенных автоматизированных действий с прогнозом усиливает внимания и когнитивный контроль в последующих попытках

**СПАСИБО ЗА ВНИМАНИЕ!**