

Streszczenie

Zachowanie, które w danej sytuacji przynosi subiektywną korzyść, zostaje skojarzone z tą sytuacją i jeśli okoliczności powtórzą się, to wzrasta prawdopodobieństwo, że powtórzone zostanie także to zachowanie. Za wzrost prawdopodobieństwa powtórzenia działania odpowiada uczenie ze wzmocnieniem, które wskazuje się jako podstawowy mechanizm umożliwiający dostosowanie zachowania do zmian zachodzących w otoczeniu. Uczenie ze wzmocnieniem odgrywa kluczową rolę dla funkcji wykonawczych, dlatego też zaburzenia uczenia ze wzmocnieniem obserwowane są w szeregu zaburzeń neurologicznych i psychiatrycznych, np. w schizofrenii, uzależnieniu czy chorobie Parkinsona.

Celem moich badań było opracowanie testu behawioralnego pozwalającego na obserwację uczenia ze wzmocnieniem w warunkach możliwie zbliżonych do naturalnych oraz sprawdzenie czy obserwowane zachowanie w warunkach testu może zostać odwzorowane za pomocą modeli obliczeniowych. Wykorzystałam opracowaną metodykę do porównania wzmacniającego działania nagrody naturalnej (sacharyna) oraz substancji uzależniającej (etanol) oraz analizy zmian w uczeniu ze wzmocnieniem u genetycznie zmodyfikowanych myszy z selektywnymi mutacjami upośledzającymi działanie układu serotoninowego.

W badaniach wykorzystałam test probabilistycznego uczenia ze odwróceniem. Ten test został zaadoptowany do klatki IntelliCage, dzięki czemu przebiegał w naturalnym cyklu okołodobowym i bez deprywacji pokarmu. Co istotne, test przebiegał z minimalnym bezpośrednim udziałem eksperymentatora. Podczas testu przebywające w grupie kilkunastu osobników samice myszy miały dostęp roztworu sacharyny lub etanolu, zależnie od doświadczenia. Jednak dostęp do substancji był probabilistyczny, ponieważ drzwi prowadzące do butelek z nagrodą otwierały się z prawdopodobieństwem 30% lub 90%. W założeniu, uczenie ze wzmocnieniem powinno pociągać za sobą to że zwierzęta będą częściej wybierać opcję związaną z większym prawdopodobieństwem uzyskania nagrody. Istotnie, w przypadku grup testowanych z wykorzystaniem sacharyny, zaobserwowałam znamienne preferencję opcji związanej z pewniejszym uzyskaniem nagrody.

W celu odwzorowania zachowania myszy zaaplikowałam szereg modeli obliczeniowych z równaniami opartymi o założenia algorytmu *Q-learning*. Ze względu na efekty wpływu długości interwału na decyzję oraz potencjalne ograniczenia związane z pamięcią, włączyłam do badań rozszerzenia tego algorytmu. Spośród przeanalizowanych rozwiązań lepszym dopasowaniem do wyborów zwierząt charakteryzowały się modele zakładające aktualizację

oczekiwań na podstawie zarówno podjętego i niepodjętego działania oraz różnicujące wrażliwość na pozytywną i negatywną informację zwrotną.

Pomimo różnic w wzmacniających właściwościach alkoholu i sacharyny nie zaobserwowałam wyraźnych różnic w mechanizmie uczenia ze wzmocnieniem u testowanych zwierząt. Natomiast niezależnie od typu nagrody na prawdopodobieństwo powtórzenia wyboru istotny wpływ miał czas, jaki upłynął od poprzedniej próby uzyskania nagrody. Mutacje w postaci inaktywacji receptorów 5-HT₇ oraz pozbawienia receptorów NMDA na komórkach serotoninowych wykazały pomijalny efekt na uczenie ze wzmocnieniem, chociaż druga modyfikacja pociągnęła za sobą osłabienie tendencji do powtórzenia poprzedniego wyboru notowanej przy dłuższych odstępach czasowych.

Podsumowując, przedstawiłam nowy test uczenia ze wzmocnieniem oraz rozszerzyłam modele obliczeniowe przewidujące wybory dokonywane przez zwierzęta. Opracowana metodyka może posłużyć do badań w dziedzinie neuropsychofarmakologii i wnioskowania o zmianach powodowanych zaburzeniami neuropsychiatrycznymi na funkcje wykonawcze oraz skuteczności leków w przywracaniu normalnego funkcjonowania.

mgr Judyta Jabłońska | Praca doktorska | *Analiza modeli uczenia się ze wzmocnieniem związanym z działaniem nagród naturalnych i substancji uzależniających w zadaniu z probabilistycznym odwróceniem*

Abstract

If under given circumstances a specific behaviour is associated with subjectively positive consequences, the odds of repeating the same behaviour under similar circumstances will increase. The change in behaviour is driven by reinforcement learning and is essential for the ability to adapt to changes occurring in the environment. Reinforcement learning plays a key role in executive functions, and accordingly, impairments in reinforcement learning emerge in multiple neurological and psychiatric disorders, including schizophrenia, addiction and Parkinson's disease.

Here, I describe a task for the study of reinforcement learning under conditions resembling a near-natural environment and test if the observed behaviour can be explained with the use of computational models. Furthermore, I have used the novel task and models to assess differences in the reinforcing effects of natural (saccharin) and drug (alcohol) rewards, and also to study reinforcement learning in genetically modified mice with selective mutations in the serotonin system.

The task is based on probabilistic reversal learning using the IntelliCage, which permits to observe the behaviour of female mice remaining in a group, under a natural circadian cycle and without food or water deprivation. Most importantly, the task requires minimal direct interaction with the experimenter. Mice in the IntelliCage were offered access to bottles with saccharin or ethanol solution, in separate experimental groups. The availability of the reward was probabilistic as the doors guarding access to a substance were programmed to open with 30% or 90% probability each time the animal approaches them. A preference of the greater reward probability is the evidence of reinforcement learning, and it was robustly observed for animals with access to saccharin as a reward. The observed choices were fitted with computational models based on the *Q-learning* algorithm, which I have extended to account for the effects of time intervals between choices and potential limitations imposed by memory. The best fit to the observed behaviour was achieved by models that assumed distinct learning rates for the rewarded and non-rewarded outcomes and also included an update of the expected value of the option not taken.

I find that despite differences in apparent reinforcement strength between alcohol and saccharine solutions, there were no obvious changes of learning mechanisms. Conversely, irrespective of reward type, there was a robust effect of the time interval between choices on the probability of repeating the same choice again. Furthermore, complete inactivation of 5-HT₇ receptors or ablation of NMDA receptors on serotonergic neurons had only limited effects

on reinforcement learning, though the latter weakened the tendency of repeating the previous choice at longer time intervals.

In summary, I present a new reinforcement learning task and extended computational models that approximate animal behaviour in the task. The methodology is potentially useful for further studies in the field of neuropsychopharmacology, to model the effects of neuropsychiatric diseases on executive functions and the potential efficacy of therapies in restoring normal behaviour.