

【報酬を得るためのコストがその報酬の価値を高める】

世界初！「働いた後のビールはうまい」脳内メカニズムを発見！
—報酬を得るための努力がその報酬の価値を上げる脳メカニズム—

玉川大学脳科学研究所（東京都町田市 所長：小松英彦）の坂上雅道（さかがみまさみち）教授、田中慎吾（たなかしんご）特別研究員（新潟大学助教）、John P. O’Doherty（じょん ぴー おどはてい）客員教授（カリフォルニア工科大学教授）は、努力をして（コストを払って）得た報酬の方が、何もしないで得た報酬よりも主観的価値が大きくなる脳メカニズムを世界で初めて明らかにしました。

本研究成果は、科学雑誌“Nature communications”に掲載されました。

掲載論文名：The cost of obtaining rewards enhances the reward prediction error signal of midbrain dopamine neurons

和訳：報酬を得るためのコストは、中脳ドーパミンニューロンの報酬予測誤差情報を高める

本研究は、ニホンザル2頭を使って、報酬を得るためのコストがその報酬の価値を高めることを行動実験で示すと同時にその時のドーパミンニューロン※の活動を測定し、価値が高まる脳メカニズムを明らかにしました。

※ドーパミンニューロン… 中脳にあるドーパミンを神経伝達物質として放出する神経細胞。報酬に基づく強化学習に重要な役割を果たしていると考えられている。

<この研究のポイント>

ニホンザルを使った実験により、コストを払って得た報酬の方が何もしないで得た報酬よりも主観的価値が大きくなることを示した。中脳ドーパミンニューロンが、このような主観的価値を作り出すために重要な働きをしていることを解明。視覚刺激と報酬確率の関係を学習するのに、コストがあった方が学習は早くなる（努力したほうが学習は進む）ことを世界で初めて発見。

【実験の成果】

これまでなんとなく、「働いた後のビールはうまい」と言われていた現象がなぜ起こるのかという脳メカニズムを明らかにした。

ドーパミンニューロンが、報酬予測学習に関与する際に、まず報酬とコストの情報を統合することで、ハイコスト試行その物の価値（報酬予測）を下げる。その状態で報酬が与えられると、コストがあったために報酬予測が小さかったハイコスト試行では、ローコスト試行に比べて相対的に報酬予測誤差情報（報酬予測と実際に与えられた報酬の差分）が大きくなり、結果的に報酬価値が大きくなると考えられる。

このことは、我々の価値が相対的なことの基本的な脳メカニズムであり、価値や意思決定に関わる多くの現象を説明できる原理であると考えられる。また、コストがある方が学習は進むという現象の発見は、社会や家庭、学校での教育のあり方を再考する契機となりうるものと思われる。

脳大成理論 補足_7

【研究の背景】

物の価値とは、その物によって予測される報酬の量のことである（実験的には、刺激の呈示によって予測される報酬の量や確率）。大脳皮質の下側に位置する大脳基底核のニューロンが、この報酬予測に関わっていると考えられている。

中脳ドーパミンニューロンは、実際に与えられた報酬が、予測していた報酬とどのくらい違っているかをシグナルしていることが知られている。このドーパミンニューロンの報酬予測誤差情報が大脳基底核に伝えられ、大脳基底核ニューロンは報酬予測（価値）を現実的なものに修正する。

また、価値は相対的である。同じ報酬でも、文脈によって価値は大きく異なる。そのような例に、「働いた後のビールはうまい」効果がある。多くの人が似たような経験があると思うが、ヒト以外の動物でも、そのような効果は実験的に示されている。

我々は、ドーパミンニューロンの報酬予測誤差が、価値の相対化に重要な役割を果たしていることを以前報告したが、このメカニズムが「働いた後のビールはうまい」効果を作り出しているのではないかと考え、ハイコスト VS ローコスト（HLC）課題遂行中のサルのドーパミンニューロンの活動を解析した。

さらに、コストにより報酬の価値が高まれば、その報酬を学習に使った時、学習は促進されるはずである。そのことも調べてみた。

【実験方法】

ニホンザル2頭にハイコスト VS ローコスト（HLC）課題を訓練した（図1）。

この課題は、2種類の試行からなり、ハイコスト試行では、最初にハイコスト試行であることが刺激（ハイコスト刺激）により示された後、報酬であるジュースを得るために、目の前の画面に呈示される点を1秒以上見つめ続けなければならない（固視）。

ローコスト試行では、ローコストを示す刺激（ローコスト刺激）の呈示後、ジュースが与えられるまで、画面のどこを見ても構わない。

どちらの試行でも、与えられるジュースの量は同じであった。これまでの行動実験から、動物は好きなものには早く反応することがわかっているが、この課題でもハイコストあるいはローコストを示す刺激や、その後の報酬が来ることを示す刺激に反応させて、その反応時間を調べた（実験1）。

また、この課題遂行中のサルの中脳ドーパミンニューロンから電気信号を記録・解析した（実験2）。

さらに、ハイコストとローコストの設定は同じままにして、2つの刺激の中から報酬刺激を選ぶことを学習する課題（1つを選べば報酬、もう1つを選べば無報酬、図3）を導入し、サルが報酬刺激を学習することとコストの有り無しの関係を調べた（実験3）。

脳大成理論 補足_7

【実験結果】

(実験1)

試行の最初に提示されるハイコストあるいはローコスト刺激に対する反応時間とそれぞれの試行でのエラー率から、サルは明らかにハイコスト試行を嫌っていることがわかった。しかし、コストの後の報酬が来ることを指示する刺激に対する反応時間は、ハイコスト試行の方がローコスト試行より有意に早く、報酬量は同じであるにもかかわらず、サルはハイコスト試行の報酬の方を好んでいることがわかった。

(実験2)

ドーパミンニューロンのコスト刺激に対する反応は、ローコスト刺激の方が大きく、ハイコスト刺激の方が小さかった。ドーパミンニューロンは、報酬には正の反応（活動量を上げる）を、罰には負の反応（活動量を下げる）を示すことが知られているが、上の結果は、ドーパミンニューロンが試行でのコストの情報と報酬の情報を統合した反応をコスト刺激に対して行っていることを示している。

さらに、報酬が来ることを指示する刺激に対しては、ハイコスト試行での反応の方がローコスト試行での反応より大きかった（図2）。

つまり、ドーパミンニューロンの報酬予測誤差応答は、ハイコスト試行での報酬の方が価値は大きいという反応を示していた。

(実験3)

報酬刺激と無報酬刺激の2種類から、1つを選ぶ選択課題（試行錯誤で報酬刺激を探す、図3）で、コストがある方が有意にこの学習を促進することがわかった（図4）。

図1
ハイコストvs.ローコスト課題

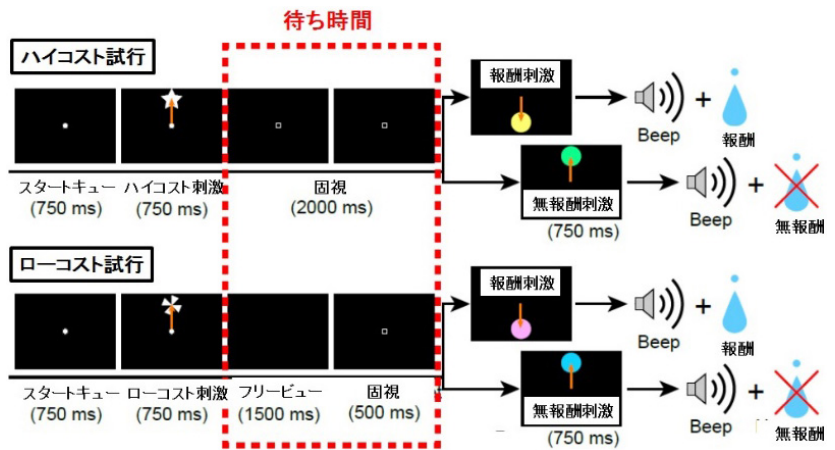


図2
ドーパミンニューロンの報酬刺激に対する応答
右図の点は、活動電位の発生時点。
次図はその加算ヒストグラム。

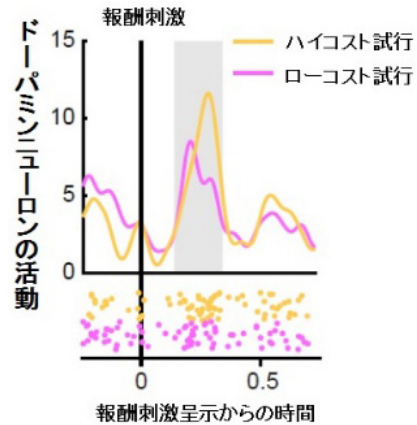


図3 報酬刺激学習課題
2つの刺激のうち報酬刺激
(○の刺激) を選択すれば報酬
が与えられる (実際には○
はない)。

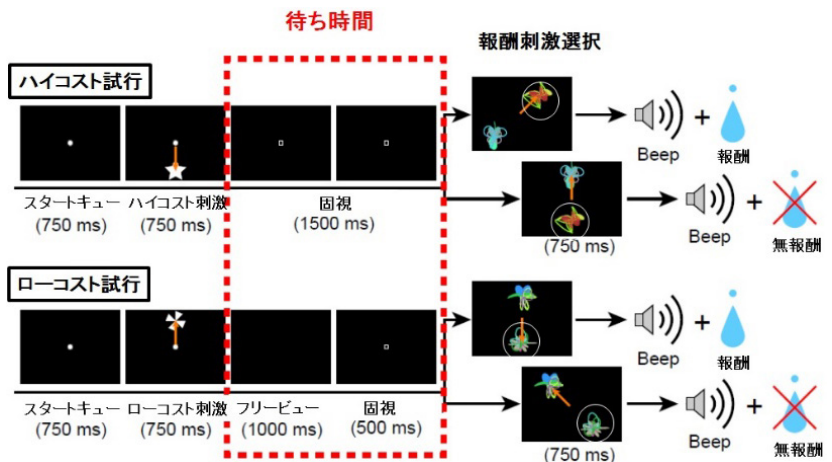


図4 報酬刺激の学習

