

ヒトもサルも予想外の大当たりが、次の判断を狂わせる

行動経済学のプロスペクト理論と強化学習理論を統合した「動的プロスペクト理論」を構築し、不確実な状況下でのヒトとサルの判断・行動を統一的に説明することに成功しました。大当たりを経験するほど、ヒトもサルも「もう1回当たりそうだ」と判断しがちであることが、うまく説明できます。

結果が不確実なことについて、ヒトはどのように判断し、行動に移すのでしょうか。ギャンブルであれば、当たりの金額と当たりの確率をかけた期待値を計算し、その期待値が最も高いものに賭けることで、平均的に高い利得を得られます。このように合理的に自己利益を計算して判断することを前提に、伝統的な経済学はヒトの行動を説明してきました。

しかし、実際のヒトの行動は完全に合理的なわけではありません。1万円持っている状態で2万円を得るのも、100万円を持っている状態で2万円を得るのも、利得は同じ2万円ですが、1万円から増えた2万円の方が大きな価値があるように感じます。また、宝くじの1等の当選確率は極めて低いのに、当たるかもしれないと思ってつい買ってしまいがちです。このようなヒトの主観を普遍的に説明するのが、プロスペクト理論です。その一方で、ヒトの判断・行動を説明するもう一つの理論に強化学習理論があります。強化学習理論では、ヒトの主観は過去の報酬経験を踏まえて変わり、最も儲かる判断を選ぶように学習すると考えます。プロスペクト理論がヒトの主観は変わらないことを前提にするのは大きく異なります。

本研究チームは、これら二つの理論を統合した「動的プロスペクト理論」を構築し、不確実な状況におけるヒトの判断・行動を一つの統一理論で説明する事に成功しました。

具体的には、参加者70人にくじ引きを繰り返し経験してもらい、その行動変化を解析しました。その結果、予測したよりもずっと大きな報酬が得られた直後に、くじの当選確率を、プロスペクト理論が予想する以上に高く評価してしまうことが分かりました。つまり、“大当たり”を経験した直後には、「もう一回当たりそうだ」と感じてしまうわけです。この予想外の出来事からの学習は強化学習理論の根幹を成すアルゴリズムです。二つの理論を統合した動的プロスペクト理論は、毎回の判断・行動を、大当たりの直後も含め、よく説明できます。

同じ霊長類に属し、ヒトに類似した脳を持つサルでも、基本的に同じ結果が観察されました。今回の研究成果に基づいてサルの脳を調べる事で、私たち一人一人が抱く多様な金銭感覚や確率の感じ方、成功した時の喜びなどが生み出される脳の仕組みの理解につながる事が期待されます。

研究代表者

筑波大学医学医療系
山田 洋 准教授

研究の背景

ヒトがギャンブルを行う際には、さまざまな方法で当たりを予測します。最も単純な予測方法として、当たりの金額と当たりの確率を掛けた期待値を計算することが挙げられます。最も期待値の高い選択肢を繰り返し選ぶことで、平均的に高い利得を得る事ができるからです。伝統的な経済学では、ヒトは期待値に応じて合理的に判断することを前提として、行動を説明してきました。

しかし、実際のヒトの行動はそうではなく、お金の感じ方も確率の感じ方も、客観値からずれることが知られています。例えば、1万円持っていて2万円を得るのも、100万円を持っていて2万円を得るのも、利得は同じ2万円なのに、1万円から増えた2万円の方が大きな価値があるように感じます。また、宝くじの1等の当選確率は極めて低いのに、当たるかもしれないと思ってつい買ってしまいがちです。このような、経済行動に関するヒトの主観を普遍的に説明するのが経済学のプロスペクト理論^{注1)}です。

プロスペクト理論は、ヒトのギャンブルにおける予測を良く説明しますが、我々の多面的な価値観の一部を捉えるにすぎません。例えば、ヒトはギャンブルの勝ち負けに一喜一憂したり、勝ちが続いたり負けが混んだ状況で浮かれたり打ちひしがれたりします。先に述べたプロスペクト理論では、我々の主観は揺らぐ不変であるという前提にたって論理が構築されており、このような変化を説明しません。むしろ、このような主観の変化は、プロスペクト理論とは全く別の強化学習^{注2)}と呼ばれる価値の学習理論によって説明されてきました。予想よりも結果が良かったのか悪かったのかを意味する“報酬の予測誤差”に応じて、過去の経験から最も儲かるよう行動を学習する理論です。いずれも価値判断に関わる行動を説明する理論であるにも関わらず、その性質の違いからこれらの理論は別々に用いられて来ました。

今回の研究では、これら二つの理論を統合することで、これらの主観が生まれる原理を統一的に説明することに挑みました。更に、ヒトと、ヒトに近い実験動物のマカクザル（以下、サル）の判断行動を比較検証することで、サルがヒトに良く似た価値観を持つかを詳細に検証しました。

研究内容と成果

本研究では、ヒトとサルにできる限り良く似た当たり構造を持つギャンブルを行わせ、その行動を予測する理論を構築しました。サルを用いた実験では、実験動物として飼育するサルに、ヒトが行うのと同じようなギャンブル（くじ引き）を繰り返し経験させました（図1）。ここでは、ヒトで用いるお金の代わりに、ジュースなどの飲み物を報酬として用いました。このくじでは、報酬の量とその報酬がもらえる確率が、別々の色のパイの数で示されます。10カ月ほど訓練すると、サルはくじが意味する“確率”とジュースの“量”を理解し、できるだけ多くの報酬が得られるようにギャンブルをするようになりました。一方で、サルの実験とは別に行ったヒトを対象とした実験では、参加者は初めに説明文を読み、2種類の色のパイの数が当たり金額と確率を表す事を理解し、その後、繰り返しギャンブルを行いました。

このようにして取得したサルとヒトの行動データを解析しました。まず、確率と報酬の利得に基づいた判断行動が、どのような主観に基づいて行われているのかを、従来のプロスペクト理論に基づく数理モデルを用いることで推定しました。その結果、ヒトの主観もサルの主観も、報酬の利得とそれを得られる確率のどちらも直線で示す客観値からずれており、上に凸の数学的な特徴で表現されることが分かりました（図2）。

このようにサルの主観を従来の経済学の理論で説明した後に、従来のプロスペクト理論では説明されて来なかったギャンブルの当り外れに基づく主観の変化について検討を行いました。この際、プロスペクト理論と強化学習理論を統合した動的プロスペクト理論を構築しました。プロスペクト理論が仮定していた、主観が変化しないという前提を取り除き、予想外の度合いに応じて主観が変化する強化学習の性質を組み込みました。この新規理論を用いて行動データを解析した所、サルでもヒトでも、予想外に大きな

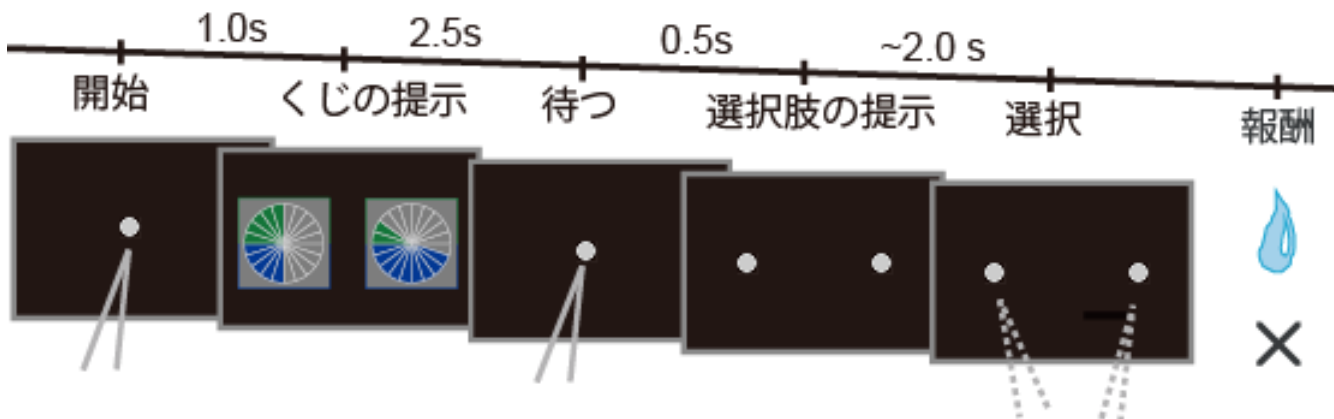
利得が得られた次のギャンブルで、確率に関する判断が全体に高まっている事が分かりました(図3)。これは、サルもヒトも同様に、予想外の大当たりを経験すると、何故か次のギャンブルも当たるのではないかと感じてしまう事を反映した主観と考えられます。今回の研究成果を基に、ヒトと良く似たギャンブルを行うサルの脳を調べる事で、私たち一人一人が抱く多様な金銭感覚や確率の感じ方、成功した時の喜びなどが生み出される脳のメカニズムを理解する研究へと発展する事が期待されます。

今後の展開

今回の発見に基づいてサルの脳を詳細に調べる事で、私たち一人一人が抱く多様な金銭感覚や確率の感じ方、成功した時の喜びなどが生み出される脳の仕組みを理解する研究につながると考えられます。個人個人の価値観を生み出す脳の仕組みを理解する事で、将来的に、多様な価値観を育み、認め合い、より多くの人が幸せを感じられる社会の実現に近づくことができると期待されます。

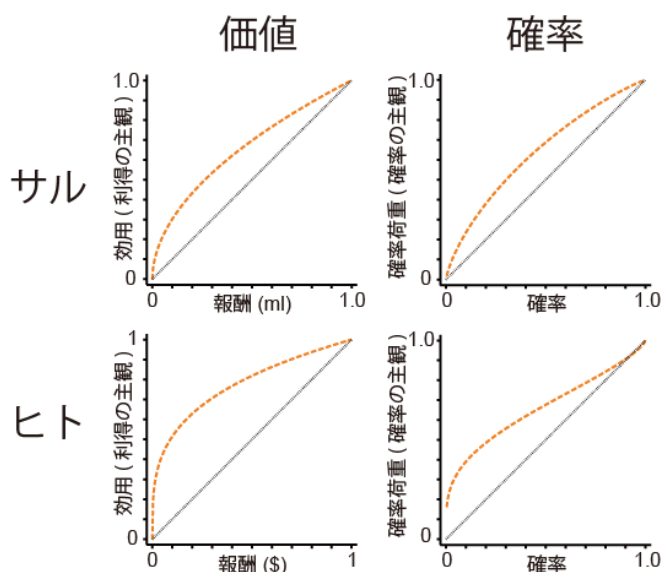
参考図

図1. ギャンブル課題



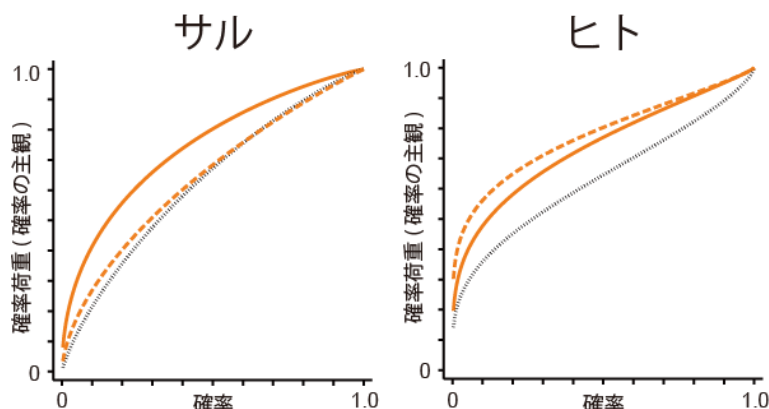
モニターを使ってサルに提示された行動課題の流れを示す。サルが中央の点を見ると、くじが二つ現れる。くじの上半分の緑色の区画はジュースの量を意味する。パイの区画1個分が0.1mlを意味し、0.1ml(1個)から1.0ml(10個)まで区画の個数が変化する。下半分の青色の区画は、提示されたジュースの量が当たる確率を意味する。その後、現れた二つの白い点のうちどちらかを選ぶと、そのくじの結果が与えられる。この図の左の例では、0.5mlのジュースが、50%の確率で当たることを意味する。残りの50%の確率でサルはジュースを得ることができない。ヒトの実験では、報酬にはお金(\$1~\$5)を用いた。くじの提示では、同様にパイの図形を用いた。

図2. サルとヒトの主観的な価値（利得）と主観的な確率の感じ方



プロスペクト理論の数値モデルを用いて求めた、サルとヒトの主観（A:利得、B:確率）。実験に用いたサル（一頭分）とヒト(70人の集合)を対象に、従来のプロスペクト理論を用いてそれぞれで求めた。

図3. サルとヒトの主観確率の変動



プロスペクト理論と強化学習理論を統合した数値モデル（動的プロスペクト理論）を用いて求めた、サルとヒトの主観確率。予想外に大きな利得を得た直後（橙実線）に予想通りの直後（黒破線）と比べて、主観確率が全体に高まっている事がわかる。高い確率で得られると思っていた大きな利得を逃した直後（橙破線）は、サルとヒトで主観確率の変化は異なっていた。サルの個体別にこの影響は異なったこともあり、自身にとってマイナスの出来事は、種間と個体間でさまざまな違いがある可能性が考えられる。

用語解説

注1) プロスペクト理論

ノーベル経済学賞を受賞したこの理論の一部は、確率加重関数と価値関数から成り立っている。プロスペクト理論を構成する論理的特徴のうち、確率加重関数は宝くじの例のように確率判断が主観的に歪むことを示す。価値関数は金銭の価値が主観的に歪み、正確には金額に比例しないことを示す。

注2) 強化学習理論

1990年代に Richard Sutton と Andrew Barto らによって体系化された理論。予測誤差学習と呼ばれる、予測からの誤差の大きさに応じて、予測を更新するアルゴリズムが理論の中心をなす。

研究資金

本研究は、以下の研究資金の支援を受けて実施されました。山田：日本学術振興会（JSPS）科学研究費(Grant Numbers JP:15H05374, and 21H02797), 武田科学振興財団、成茂神経科学研究助成基金、科学技術振興機構（JST）ムーンショット型研究開発事業・目標9「2050年までに、こころの安らぎや活力を増大することで、精神的に豊かで躍動的な社会を実現」「脳指標の個人間比較に基づく福祉と主体性の最大化」（PM 松元健二 玉川大学(JPMJMS2294)）。Agnieszka Tymula: Australian Research Council Discovery project (190100489)。

掲載論文

【題名】 Dynamic prospect theory: two core decision theories coexist in the gambling behavior of monkeys and humans
(動的プロスペクト理論：二つの核となる意思決定理論がヒトとサルのカンブル行動に共存する)

【著者名】 Agnieszka Tymula, School of Economics, University of Sydney, Professor
Xueting Wang, School of Economics, Finance and Marketing, College of Business and Law, RMIT University, Lecturer

Yuri Imaizumi, Medical Sciences, University of Tsukuba, student

Takashi Kawai, Division of Biomedical Science, Institute of Medicine, University of Tsukuba, Post doctor

Jun Kunimatsu, Division of Biomedical Science, Institute of Medicine, University of Tsukuba, Assistant Professor

Masayuki Matsumoto, Division of Biomedical Science, Institute of Medicine, University of Tsukuba, Professor

Hiroshi Yamada, Division of Biomedical Science, Institute of Medicine, University of Tsukuba, Associate professor

【掲載誌】 Science Advances

【掲載日】 2023年5月19日

【DOI】 10.1126/sciadv.ade7972

問い合わせ先

【研究に関すること】

山田 洋 (やまだ ひろし)

筑波大学医学医療系 准教授

URL: <https://trios.tsukuba.ac.jp/researcher/0000003502>

【取材・報道に関すること】

筑波大学広報局

TEL: 029-853-2040

E-mail: kohositu@un.tsukuba.ac.jp