

意思決定における認知バイアスと報酬学習の相互関係の検討

Investigation of Mutual Interaction Between Cognitive Bias and Reward Learning in Decision-making

高橋英之

Hideyuki Takahashi

玉川大学 脳科学研究所

Tamagawa university, brain science institute.

hideman@lab.tamagawa.ac.jp

Abstract

Decision-making is an important ability for living in various social situations. This ability is thought to be consisted of two processes, one is the reward learning process and another is the top-down process such as a cognitive bias. In this study, we investigated how these two processes for decision-making worked in healthy and autistic participants and try to argue the relationship between the cognitive bias and the reward learning in decision-making.

Keywords — Decision-making, Cognitive bias, Reward learning, autism

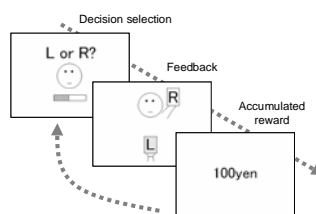
1. はじめに

様々な状況で適切に意思決定を行う能力は、生存のために必須である。人間や他の動物の意思決定は、行動の結果得られる報酬を利用して学習されることが多い。しかし人間を取巻く社会は複雑であり、完全に同一の状況が生じることは無い。従ってすべての状況に適切な意思決定を、報酬のみから学習することは不可能である。そこで人間は選好バイアスなどの認知バイアスをうまく利用しながら意思決定を行っていると考えられている。認知バイアスとは、部分的な情報を元に過去の経験や知識から近似的に意思決定を行うことである。認知バイアスにもとづく意思決定は、しばし不合理な行動を誘発するが、多くの場合、素早くある程度適切な意思決定を行うことが可能になる。その一方で、人間にも他の動物と同様に報酬学習を通して適切な意思決定を学習、実行していくシステムが備わっている。意思決定にかかわる近年の脳研究から、認知バイアスにもとづく意思決定には主に前頭前野などの大脳皮質が関わっており、報酬学習は大脳基底核や扁桃体などの皮質下の脳部位が重要な機能を担っていることが分かってき

た[1]。これらの知見は、脳内に二種類の性質の異なる意思決定の為の情報処理が存在し、並列的に働いていることを示している。しかし現在のところ、これら二つの情報処理がどのように互いに関わりあって働いているのかについては明確には分かっていない。特に、二つの情報処理が、どのように使い分けられているのかはまだわからないことが多い。そこで本研究では、報酬獲得ギャンブル課題を用いて認知バイアスと報酬学習がどのような相互関係をもって働いているのかを検討することを試みた。具体的には Matching pennies game という課題を用いた以下の検討を行った。

2. 実験設定

Experimental stimuli



Pay-off matrix

		opponent	
		L	R
subject	L	-50yen	50yen
	R	50yen	-50yen

図1 実験に用いた matching pennies game

ラップトップパソコンの画面で行う Matching pennies game(図1)を用いた実験を精神疾患や発達障害を有しない健康な成人 (healthy participant)と高機能の自閉症スペクトラム成人 (participant with high function ASD)を対象に行った。実験では、対戦相手を被験者の勝率が常に50%なコンピュータに固定し、被験者の相手に対する認識を人間 (Human opponent : HO 条件)、コンピュータ (Computer opponent : CO 条件)と交互に変化させた実験(Experiment A)と、二体

のコンピュータ(Robo A と Robo B)と交互にゲームを行い、被験者の勝率を実験セッションごとに操作した実験(Experiment B)の二つを行った。どちらの実験も 1 セッション 50 試行である。(Experiment A は HO 条件と CO 条件をこの順序でそれぞれ 3 セッションずつ、計 6 回繰り返した。Experiment B は Robo A と Robo B 相手を交互に 4 回ずつ、計 8 回セッションを行った。Experiment A の被験者の勝率はどのセッションも 50%、Experiment B における各セッションの勝率は前から順に、80%、20%、80%、20%、50%、50%、50%、50%であった。行動の解析は、過去の被験者と対戦相手の行動履歴にもとづく被験者の行動選択の条件付確率から以下の式で計算したエントロピー H を 0 から 1 の範囲で正規化したものを用いた (c は行動、 s はゲーム上の状況を示す)。

$$H = -\sum_c \sum_s P(s)P(c|s) \log_2 P(c|s)$$

この値が 0 に近ければ近いほど被験者の意思決定が定型的であることを意味し、1 に近いほど被験者が過去の行動履歴から独立した意思決定を行っていたことを示唆する。エントロピーはセッションごとに計算した。

3. 実験結果

3.1. 対戦相手が人間かコンピュータかに応じた意思決定の変化 (Experiment A)

Experiment A における healthy participant 19 人のセッションごとのエントロピーと獲得報酬の平均のグラフを図 1-1 に示す。獲得報酬は常に勝率を 50% に固定している為、条件間に有意差は存在しない。しかし healthy participant での結果は、エントロピーの値が、統計的に有意に HO 条件の方が CO 条件よりも高くなった。図 1-2 は participant with high function ASD 19 人のセッションごとのエントロピーと獲得報酬の平均のグラフである。participant with high function ASD での結果は、獲得報酬、エントロピーの値には条件間で有意な差が無かった。被験者群×実験

条件でのエントロピーの値に関する分散分析を行ったところ、二要因に交互作用があった (ANOVA, $F[1, 36]=4.35, p<0.05$)。また被験者間、実験条件間での行動決定までの反応時間には統計的な有意差が存在しなかった。これらの結果は、healthy participant は対戦相手が人間かコンピュータかという認識に応じて自らの意思決定を変化させていたのに対して、participant with high function ASD ではそのような傾向が弱かったことを示唆する。また内省報告で、healthy participant も participant with high function ASD も HO 条件ではすべての被験者が、自らの対戦相手が人間だと最後まで信じていたと報告した。

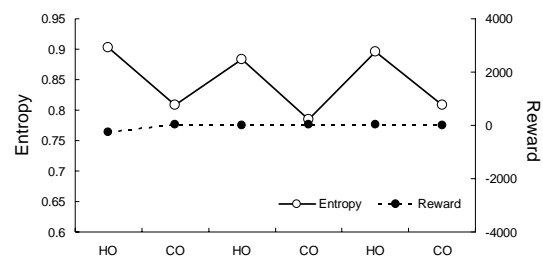


図 2-1. healthy participant におけるエントロピーと獲得報酬のセッションごとの平均 (Experiment A)

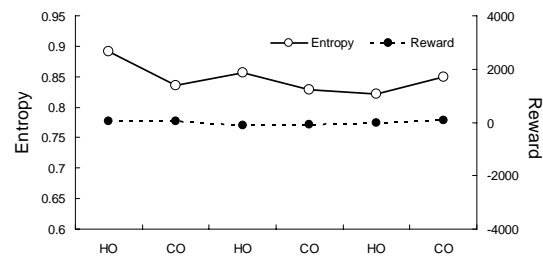


図 2-2. participant with high function ASD におけるエントロピーと獲得報酬のセッションごとの平均 (Experiment A)

3.2. 勝率に応じた意思決定の変化 (Experiment B)

Experiment B における healthy participant 11 人のセッションごとのエントロピーと獲得報酬の平均のグラフを図 3-1 に示す。獲得報酬は設定された勝率に連動してセッションごとに変化し、エ

ントロピーの値が獲得報酬に連動して変化していることがわかる。具体的には、勝率が高く獲得報酬が多いときにはエントロピーの値が小さく、勝率が低く獲得報酬が負の場合にはエントロピーの値が大きくなった。図 3-2 は participant with high function ASD 10 人のセッションごとのエントロピーと獲得報酬の平均のグラフである。participant with high function ASD での結果も、healthy participant と同様に獲得報酬は設定された勝率に連動してセッションごとに変化し、エントロピーの値が、獲得報酬に連動して変化していることがわかる。また被験者群×実験条件でのエントロピーの値に関する分散分析を行ったところ、二要因に交互作用はみられなかった。この結果は、healthy participant も participant with high function ASD も獲得報酬の履歴に応じて自らの意思決定の傾向を変化させていたことを示唆するものである。

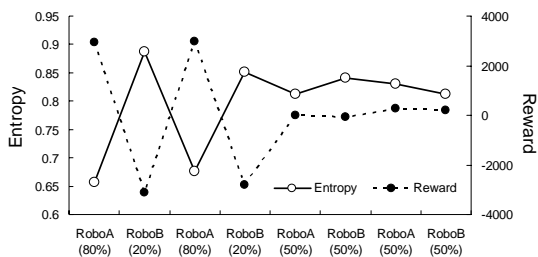


図 3-1. healthy participant におけるエントロピーと獲得報酬のセッションごとの平均 (Experiment B)

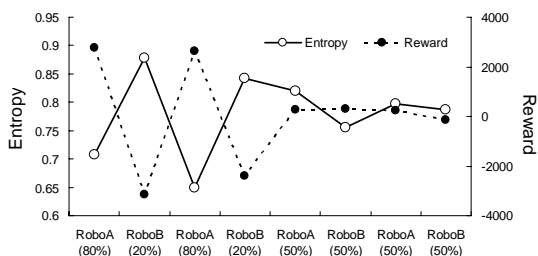


図 3-2. participant with high function ASD におけるエントロピーと獲得報酬のセッションごとの平均 (Experiment B)

4. 議論とまとめ

Experiment A では、実際には被験者の勝率は常に 50%であり、対戦相手が人間かコンピュータかという情報は被験者の意思決定の結果には影響を与えない。しかし、healthy participant の意思決定におけるエントロピーは、対戦相手の情報に応じて変化した。具体的には、相手が人間だと信じている条件では複雑な意思決定を、相手がコンピュータだと信じている条件では定型的な意思決定を行う傾向が見られた。これは対戦相手に心を想定することで、相手がこちらの行動を読んでくると意識がバイアスとして働き、被験者の意思決定を相手が人間だと思っている条件で複雑にしたものと思われる。またこの傾向は、報酬獲得量に条件間に有意差が無いにも関わらず、実験の後半まで持続した。これは相手が人間かコンピュータかという情報が強力な認知バイアスとして、healthy participant の意思決定に影響を与えたものと思われる。それに対して、participant with high function ASD ではこのような認知バイアスの働きは弱かった。

Experiment B では、被験者の対戦相手は常にコンピュータであると教示し、二種類のコンピュータとゲームを行っているように意識させる条件を用意し、各条件を交互に行った。しかし実際にはブロックごとに実験者側で勝率を設定し、それに伴って被験者の各ブロックでの獲得報酬が変化した。その結果、どちらの被験者群においても獲得報酬の変化に応じてエントロピーが変化することが示された。さらに今回の実験では最初の 4 ブロックは条件間に勝率の差を設けたが、後半 4 ブロックはそのような勝率の差を設けなかった。もし被験者が、前半にロボットの勝率の差を知識として学習して、それを後半に認知バイアスとして働かすのであれば、後半にも条件間にエントロピーの差が生じるはずである。しかし後半には条件間にエントロピーの差が生じず、短期的に獲得した知識は認知バイアスとして機能しないことが分かった。また participant with high function ASD も勝率に応じてエントロピーが変化していたこと

から、獲得報酬に応じて意思決定を調整する能力
そのものには、participant with high function
ASD には特異的な問題が無いのではないかと示唆
された。

以上をまとめると、人間や機械といった長期的
に獲得した知識は認知バイアスとして意思決定に
強力に働きかける。一方で短期的に獲得した知識
(相手が強い、弱いなど) は認知バイアスとして
働かず、その際には意思決定は報酬学習ベースと
なることを示した。これは私たちの脳に備わって
いる意思決定システムが、短期的に獲得した知識
に対しては可塑性を高め、長期的に獲得した知識
に対しては可塑性を下げること大局的なバラン
スを維持していることを示唆するものである。

また成人高機能自閉症を対象とした同様の検討
では、意思決定に認知バイアスが与える影響が弱
い一方で、報酬学習は健常者と同様に行うことが
示唆された。これらの結果は、意思決定に関わる
二つの情報処理のバランスの偏りが、自閉症の社
会的な問題の背後にある可能性を示唆するもので
ある

今後は、認知バイアスと報酬学習を同時に行わ
せることで、これらのシステムがどのように相互
作用しあっているのかについて、より具体的に検
討していければと考えている。また現在は同様の
課題遂行中の脳活動を fMRI を用いて計測してお
り、これらのシステムの働き背後にある神経基
盤について明らかにできたらと考えている。

参考文献

- [1] Daw, N.D., Niv, Y., and Dayan, P. (2005)
“Uncertainty-based competition between
prefrontal and dorsolateral striatal systems for
behavioral control,” Nature Neuroscience
8:1704-1711.