修士論文

高齢者におけるモデルフリー・モデルベース行動とそれに関わる脳ネットワークの特性

石橋 優希

2019 年 1 月 31 日

奈良先端科学技術大学院大学
情報科学研究科 情報科学専攻

本論文は奈良先端科学技術大学院大学情報科学研究科に
修士授与の要件として提出した修士論文である.

石橋 優希

審査委員 :
中村 哲(主指導教員)
松本 裕治(副指導教員)
川人 光男(副指導教員)
森本 淳(副指導教員)
田中 宏季(副指導教員)

高齢者におけるモデルフリー・モデルベース行動と
それに関わる脳ネットワークの特性

石橋 優希

内容梗概

人間は日常生活において常に意思決定を行いながら生活をしている．その意思決定の脳機能理論の基本には「model-free 強化学習」がある．エージェントは与えられた環境下において状態を観測し，選択肢の中から報酬を得るための行動を選択する．また，経験を通じて変化を脳内シミュレートするような内部モデルを学習する「model-based 強化学習」も，意思決定戦略の一つとして挙げられる．model-free 強化学習は，中脳ドーパミンと大脳基底核によって構成される神経回路によって実現されていると考えられており，model-based 強化学習は前頭前野を中心とする大脳皮質の回路によって実現されていると考えられている．このように，意思決定と脳活動は密接に関係している．また，個人の特性や状態によっても意思決定手法が異なることも明らかになっている．そこで本研究では，この高齢者の認知機能の低下における意思決定の変化に着目し，model-based / model-free の意思決定手法を用いて高齢者の行動手段を明らかにし，安静時の脳活動と関連付けようと考えた．高齢者の安静時の脳活動を用いて解析を行ったあと，行動データの指標と安静時の脳活動の結果から，2 つのデータの相関をみた．その結果，脳活動について，先行研究において model-free / model-based それぞれが関連すると言われている脳部位に関する相関は見られなかったが，model-free においては，報酬関連の活動と正の相関を示す楔部に相関が見られ，model-based においては，刺激履歴と相関がある後部頭頂皮質に相関が見られた．

Model-free / Model-based Behavior of Elderly People and Characteristics of the Brain Network

Yuki Ishibashi

Abstract

Human beings are constantly making their decisions in everyday life. “model-free” is the basis of the brain function and making decisions. The agent observes the state of their situation and selects their action from the options to obtain the reward. “model-based” is also one of the decision-making strategies. model-free is composed of neural circuit of midbrain dopamine and the basal ganglia. Model-based is composed of cerebral cortex circuit focused on the prefrontal cortex. Thus, decision-making and brain activity are closely related. It is also clear that the decision-making method differs depending on individual characteristics and situation. Therefore, in this study, I focused on the change of decision-making in the decline of cognitive function of elderly people. I decided to clarify behavioral measures of the elderly people using the model-free / model-based decision-making method and associate them with brain activity at rest. After analyzing using the brain activity at rest of the elderly people, correlation between behavior data index and the brain data at rest. As a result, there was no correlation that is shown by prior research between brain activity and model-free / model-based. In model-free, there was a correlation with Cuneal Cortex that shows the action of rewards. In model-based, there was a correlation with PPC(Posterior parietal cortex) that shows stimulation history.

目次

図目次

表目次

第 1 章　序論

1.1　背景

人間は日常生活において常に意思決定を行っている．意思決定は，個人の認知機能に基づくものである．認知機能は，理解・判断・倫理などの知的機能のことを言い，五感(視覚・聴覚・触覚・嗅覚・味覚)を通じて外部から入ってきた情報から，何かを記憶したり，問題解決のために深く考えたりなどを行う，人の知的機能を総称した概念である．我々はその認知機能に基づいて，日常的に意思決定を行いながら生活をしている．

日常生活における環境に対応するために必要な認知機能は，年齢を重ねるごとに低下する傾向にある．ここで言う認知機能の低下とは，理解力や判断力，記憶力や言語理解能力など，認知機能に関わる能力が低下している状態のことを指す．この認知機能の低下により，若年者と高齢者で意思決定の手法に異なる傾向がみられることが分かっている(Eppinger et al., 2013)．また，Decker(2016)は Eppinger と同様に，子供・青年・成人に分類し意思決定手法の違いを評価した．認知機能と意思決定の手法について述べている論文は多々あり，たとえば認知的負荷をかけた状態における意思決定の研究(Otto et al., 2013)や，個人の性格と意思決定の関係性に関する研究(Gillan et al., 2015)がある．

また一方で，認知機能と脳内活動を関連付け，脳の活動部位を調査する先行研究があり(Daw et al., 2011)，model-free / model-based 強化学習と神経基盤との相互作用を定量的に評価している．Dall(2012)は，model-free / model-based の両立について，行動と脳活動の両方の視点から評価を行った．また，Jan(2010)は，報酬関連の意思決定に関係する脳領域について調査している．

私はこれらのことから，model-free / model-based の意思決定手法を用いて高齢者の行動手段を明らかにし，安静時の脳活動と関連付けようと考えた．高齢者の安静時の脳活動を用いて解析を行ったあと，行動データの指標と安静時の脳活動の結果から，2 つのデータの相関をみる．

1.2　意思決定計算モデルの理論

1.2.1　強化学習理論

強化学習とは，試行錯誤を通じて価値を最大化するような行動を学習することである．教師付き学習のように正しい行動を教わり学習するのではなく，取った行動に対して得られる報酬を通じて自らの行動を評価することで学習を行う点が特徴的である．強化学習の説明のために，状態s_t，行動a_t，報酬r_tとして，説明する．意思決定者「エージェント」と制御対象「環境」は以下のやり取りを行う．

(1) エージェントは時刻 t において環境s_tに対して意思決定を行い，行動a_tを行う．

（2）行動a_tののち，環境はs_{t+1}へ遷移し，エージェントはその遷移に応じて報酬r_tを得る．

（3）時刻 t を t+1 に勧めて最初のステップに戻る．

強化学習において，エージェントはあらかじめ環境や正しい行動に関する知識を持たない．また，環境の状態遷移や報酬の与えられ方は確率的である．

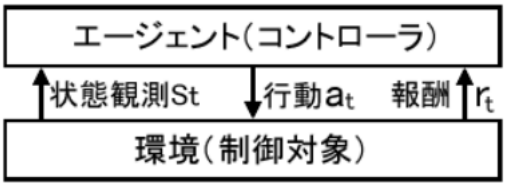

図 1.1 強化学習モデル

(Kitamura et al.,1999)

(Ⅰ)強化学習の基礎理論

強化学習理論では，即時報酬ではなく，将来における価値の最大化を行う．「s_tにおいてある行動a_tを取った時の価値」を参考に，その価値の一番高い行動を選択することで価値を最大化することができる．この，「s_tにおいてある行動a_tを取った時の価値」をQ値，もしくは状態行動価値と呼び，Q(s, a)と書く．現時点tにおけるQ値は次の時点t + 1のQ値によって表すことができる．

$$Q(s_t, a_t) = E_{s_{t+1}}(r_{t+1} + \gamma E_{a_{t+1}}(Q(s_{t+1}, a_{t+1})))$$

$E_{s_{t+1}}$はs_{t+1}における期待値，$E_{a_{t+1}}$がa_{t+1}に関する期待値である．γは割引率と呼ばれ，将来における価値をどれだけ割り引くかに関するパラメータである．強化学習におけるアルゴリズムは，このQ値の学習方法でいくつかに分類されている．(1)には期待値が含まれており，その期待値を得るためには次の時点の状態を正確に計算することが必要である．強化学習で用いる期待値は，試行錯誤により，状態tにおける環境sのもとで行動aをとったときに得られる報酬rを用いてQ値の更新を行っていく．ここでは，シンプルなQ学習について説明していく．

(Ⅱ) Q学習

Q学習では，次のようにQ値の更新を行っていく．

$$Q(s_t, a_t) \leftarrow (1 - \alpha)\mathrm{Q}(s_t, a_t) + \alpha(r_{t+1} + \gamma max_{a_{t+1}} Q(s_{t+1}, a_{t+1}))$$

α (0< α <1)は学習率と呼ばれるパラメータで，Q値の更新具合を制御できる．

一般的なQ学習の流れは以下の通りである．

(1) 初期の$Q(s_t, a_t)$の値をランダムに決定する．

(2) 初期状態$\mathrm{t} = 0$から一定回数行動aを行い，$Q(s_t, a_t)$を更新していく．

(3) 更新を一定回数行ったら初期状態$\mathrm{t} = 0$に戻す．

(4) (2)と(3)を一定回数行ったら終了．

1.2.2 model-free model-based 強化学習

強化学習の理論的な背景を説明したところで，意思決定の脳機能理論の基本となる「model-free 強化学習」について説明する．エージェントは与えられた環境下において状態を観測し，選択肢の中から報酬を得るための行動を選択する．報酬予測誤差(実報酬 − 予測報酬)は将来の行動選択のための学習信号として機能する．この学習は誤差の減少を目指す。これが「model-free 強化学習」と呼ばれるのは，報酬予測と行動選択に内的モデルを一切必要としないからである．しかし，感覚からの連想だけで報酬予測と行動選択がいつも決定されるわけではない．我々は経験を通じて変化を脳内シミュレーションするような内部モデルを学習することができる．「model-based 強化学習」の研究も，この内部モデルが土台となっている．一般に model-based 強化学習では状態遷移と報酬関数両方の内部モデルを学習する．そして，この学習から，報酬予測と行動選択の意思決定を行う．

1.3 意思決定に関する先行研究

1.3.1 行動実験

Daw(2011)は，model-free 強化学習，model-based 強化学習の行動評価の違いと fMRI を用いて，神経基盤との関係性を定量的に評価した．Daw は，被験者は行動実験において model-free と model-based 両方を混合したハイブリット型の結果を出すのではないかと仮説を立てた．

まず，model-free 戦略と mode-based 戦略を分類するために，タスクを行った(図 1.2)．被験者は，2 回の選択が求められる．Stage1 では，図 1.2(b)に示された緑の画像ペアから 1 枚の画像を選択する．Stage1 で選択が終わると，ピンクの画像ペアもしくは青の画像ペアに遷移し，被験者は再度，遷移先の画像ペア 2 枚のうちから 1 枚を選択する．合計 2 回の選択が終了すると，被験者には報酬が与えられる．画像ペアの移り変わりは図に示された通り確率的に変化しており，Stage1 の選択に応じた確率で Stage2 の状態が決定する．報酬の有無は，ガウスのランダムウォークに従って時間的に変化した．画像の選択や遷移，提示は図 1.2(a)のように秒数を固定し，選択に時間制限を設けた．

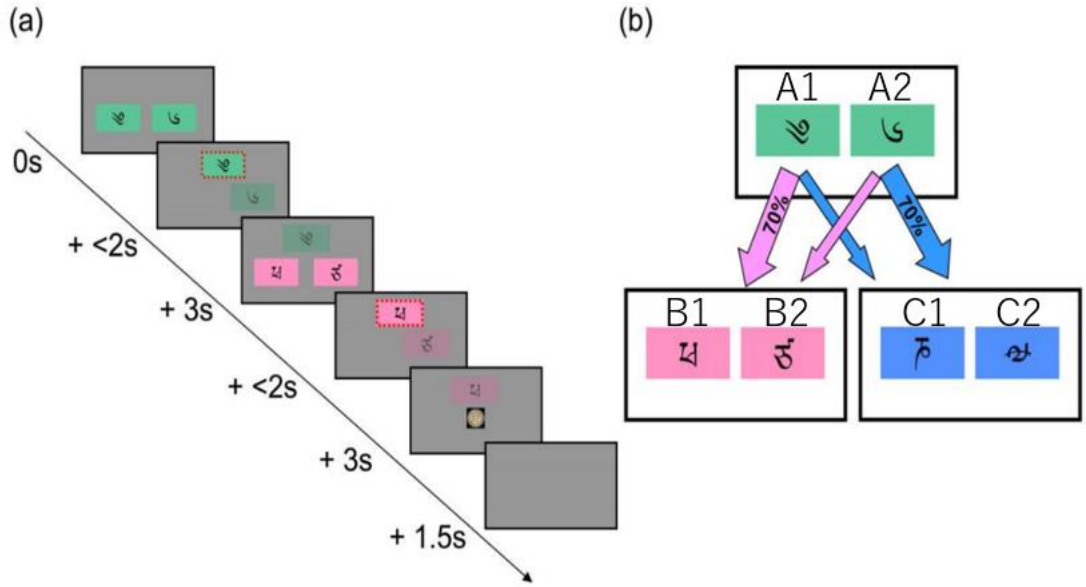

図 1.2 先行研究課題

(Daw et al., 2011)

(a)1 トライアルの流れを示す．A1，A2 の画像が表示され，再度 A1，A2 の画像が表示されるまでを 1 トライアルとする．(b)画像の遷移を示す．A1 を選んだ場合は 70%の確率で B1，B2 の画像ペアに，30%の確率で C1，C2 の画像ペアに遷移する．A2 を選んだ場合は 70%の確率で C1，C2 の画像ペアに，30%の確率で B1，B2 の画像ペアに遷移する．

1.3.2 Stay probability

先行研究である model-free 戦略と mode-based 戦略を分類する課題(Daw et al., 2011)において，報酬を得た場合と得ない場合，また Stage 1 から 2 への遷移が高確率(common)か低確率(rare)だったかで，その次の試行で同じ選択をする度合い(stay probability)が、model-based と model-free の戦略で異なることがわかっている．これは、model-free 戦略では，遷移確率によらず stage2 で報酬を得られた選択を繰り返すため，rewarded で高く unrewarded で低い stay probability となる(図 2.4 左)．一方 model-based 戦略では，遷移確率を考慮した選択を行うため，common で報酬が得られた場合と rare で報酬が得られなかった場合に高い stay probability となり，その逆が低くなるためである．この傾向を言い換えると，model-free 戦略では報酬の影響が，model-based 戦略では報酬と遷移確率の相互作用があると言える．

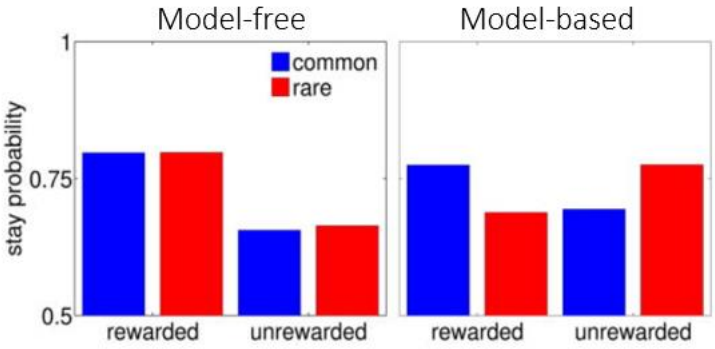

図 1.3 Stay probability

(Daw et al., 2011)

次の試行で同じ選択をする度合い(stay probability)を示す．model-based と model-free の戦略で異なる表示を示す．

1.3.3　結果

行動実験の結果から，被験者は単純に報酬が得られた選択肢を繰り返すだけでなく，報酬と遷移確率を考慮するモデルを含んだ選択も行っていたことが分かった．図 1.4 は，Daw の実験における被験者の stay probability を被験者間の平均でプロットしたものである．

また，fMRI の結果から，model-free 強化学習，mode-based 強化学習において，報酬予測は腹内側前頭前野と腹側線条体の 2 つの領域に相関があることが提示された．これまでは，ドーパミンと報酬予測が model-free 強化学習に関連していると言われていたが，実際は純粋な model-free 強化学習ではなく，mode-based 強化学習の戦略によって計算された評価も反映しているという事実が fMRI から得られた．

また，課題中に認知的負荷をかけた状態(Otto et al., 2013)や，衝動的な性格であれば model-free 戦略の比重が増すなど(Gillan et al., 2016)，個人の特性や状態との関連も明らかになっている．また，若年者と高齢者で意思決定の手法に異なる傾向がみられ，model-free/model-based の戦略の比重は年齢との関連も強いことも明らかになっている (Eppinger et al., 2013)．Eppinger の研究では，高齢者は全体としての報酬データではなく直前の報酬に頼り行動を選択しているため，model-based 強化学習の精度が良くないとある．つまり，報酬予測価値を model-based の選択に組み込む際の精度が年齢に関係しており，加齢が行動選択に影響を与えているからであると考えられる．

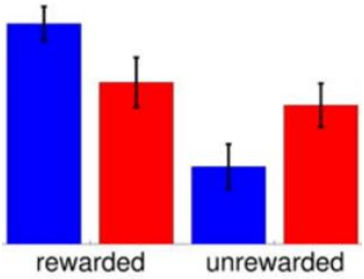

図 1.4 Stay probability

(Daw et al., 2011)

被験者全員の Stay probability の平均を示す.

1.3.4　状態行動価値

ここで，model-free / model-based それぞれについて，状態行動価値の求め方を記載する.

(Ⅰ) model-free

状態と行動は各試行 t において以下の式に従って更新される.

$$Q_{S_2MF}(a,t+1) = Q_{S_2MF}(a,t) + \alpha_2(\mathrm{r(t)} - Q_{S_2MF}(a,t))$$

ここで，α_iは現在の状態での学習率を表し，r(t)はその試行で受け取る報酬を表す．次のトライアルの Stage1 において model-free の値を更新するために，現在の Stage2 での状態行動価値と報酬が使用される.

$$Q_{S_1MF}(a,t+1) = Q_{S_1MF}(a,t) + \alpha_1(Q_{S_2MF}(a_{chosen},t) - Q_{S_1MF}(a,t) + \alpha_1\lambda((r(t) - Q_{S_2MF}(a,t))$$

(Ⅱ) model-based

Model-based の状態行動価値は，Stage2 での model-free 状態行動価値と遷移確率を考慮して計算される.

$$Q_{S_1MB}(a_1) = common * \max\left[Q^{MF}_{S_2-B}(\mathrm{a})\right] + rare * \max[Q^{MF}_{S_2-C}(\mathrm{a})]$$

$$Q_{S_1MB}(a_2) = rare * \max\left[Q^{MF}_{S_2-B}(\mathrm{a})\right] + common * \max[Q^{MF}_{S_2-C}(\mathrm{a})]$$

この式では，“common”は最高遷移確率として定義され，“rare”は最低遷移確率として定義されている.

1.4　本研究のアプローチ

前述した先行研究から，高齢者は認知機能の低下により，model-based の学習に関わる前頭前野を中心とする大脳皮質内の活動が低下し，model-free の傾向が見られるのではないかと仮説を立てた．

本研究の目的は，高齢者の認知機能の低下における意思決定の変化に着目し，model-based / model-free の意思決定手法を用いて高齢者の行動手段を明らかにし，安静時の脳活動と関連付けることである．高齢者の安静時の脳活動の解析を行ったあと，行動データの指標と安静時の脳活動の結果から，2 つのデータの相関をみる．

第 2 章　方法

2.1　被験者

13 名の健康な高齢者(男性 :8 名，女性 :5 名，最低年齢 61 歳，最高年齢 78 歳，平均年齢 66 歳±5.2，右利き)が被験者として研究に参加した．被験者には，実験前に実験の目的と内容について説明を行い，同意書へ署名いただいた．本実験は，株式会社国際電気通信基礎技術研究所(ATR)の倫理安全委員会の承認を得て行われた．すべての被験者は，本研究室において実施した別の実験において事前に安静時の脳活動データを取得済みである．

また，被験者の認知機能の指標として年齢，Education(教育年数)，IQ(JART 予測 IQ)と認知機能バッテリーCANTAB の 2 つの課題のスコア(詳細は「2.3.3 認知バッテリー(CANTAB)」参照)を用いた．被験者のデモグラフィックを表 2.1 に示す．

	平均	標準偏差
年齢	66.08	5.05
Education	14.85	1.79
IQ	111.46	7.07
IED error	25.69	32.03
SWM error	41.23	20.93

表 2.1 デモグラフィックデータ

2.2　課題

被験者が実施した行動実験の詳細は以下の通りである(図 2.1)．課題は前述の model-free 戦略と mode-based 戦略を分類する課題 (Daw et al., 2011) を元にした．Daw の課題では，報酬確率が時系列変化したが，本課題では報酬確率をセッションごとに固定して行った．被験者は Stage1 で 2 枚の画像(A1,A2)から 1 枚，画像の選択を行う．たとえば，Stage1 で A1 の画像を選択した場合，80%の確率(common)で Stage2 の画像ペア(B1,B2)に遷移し，20%(rare)の確率でもう片方の画像ペア(C1,C2)に遷移する．A2 の画像を選択した場合は，80%の確率で Stage2 の画像ペア(C1,C2)に遷移し，20%の確率でもう片方の画像ペア(B1,B2)に遷移する．Stage2 において画像ペアが表示されると，被験者は再度画像の選択を行う．すると，結果として被験者は確率的に 100 円もしくは 0 円を得る．Stage1 で画像の選択を行ってから 100 円または 0 円の報酬を得るまでを 1 トライアルとし，50 トライアルを 4 セッション実施した．4 セッションにかかる時間は約 40 分である。報酬確率は，セッションごとに固定し，セッションが変わる毎に変化させる．被験者はより多くの報酬を得るように Stage1 および 2 で画像の選択を行った．報酬を得るためには，報酬を得やすい確率

の画像ペアへ Stage2 の段階で遷移する必要がある．たとえば session1 と session4 では B1 と B2 が報酬を得やすい画像ペアで，session2 と session3 では C1 と C2 が報酬を得やすい画像ペアである．そのため，その画像ペアに遷移するために，被験者は Stage1 で画像の遷移確率を学び，各セッションにおいて報酬を得やすい画像ペアへ遷移しやすい画像を選ぶ必要がある．各画像ペアに移動後は，表 2.2 に示してある確率で被験者に報酬を与える．報酬は 0 円もしくは 100 円で，報酬確率は 100 円の出やすさを示している．画像の遷移確率（80%/20%）は，実験を通して固定した． 報酬の変動を学習できるかどうかを確認するために，実験はすべて同じ図を固定して使用した．

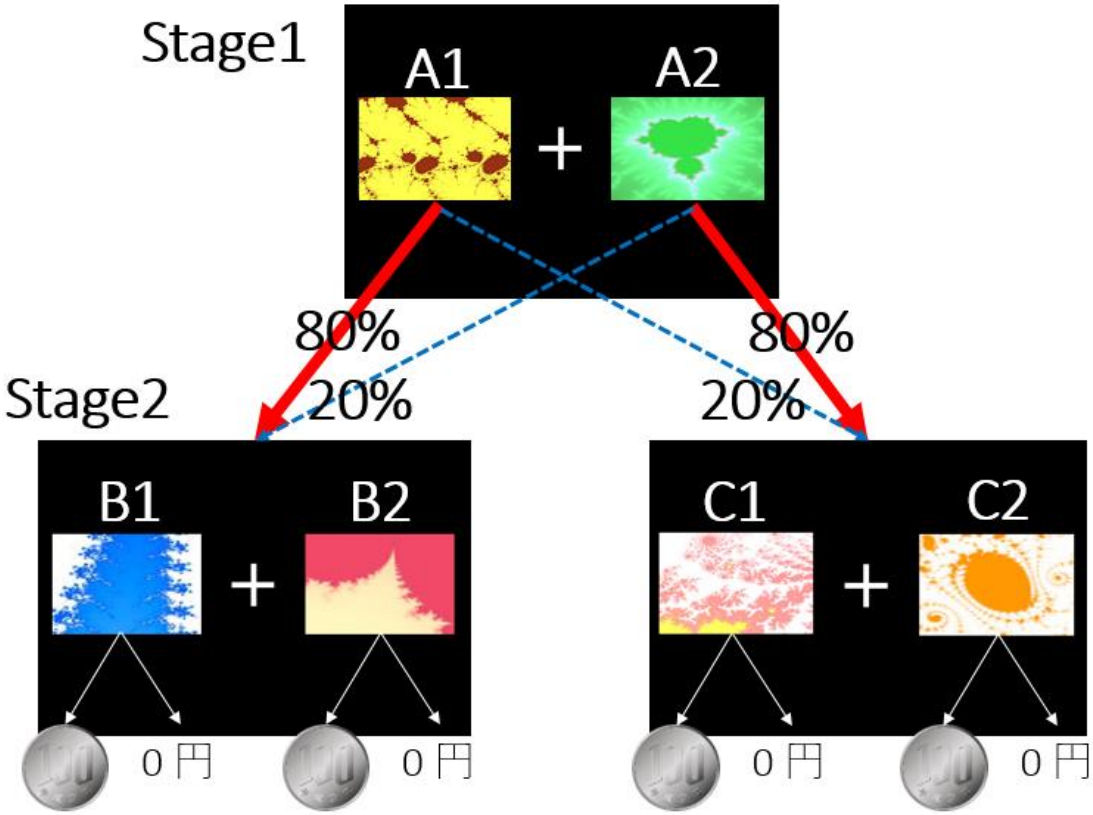

図 2.1 行動実験の流れ

実験における画像の遷移を示す．Stage1 において A1 を選んだ場合は 80%の確率で B1，B2 の画像ペアに，20%の確率で C1，C2 の画像ペアに遷移する．A2 を選んだ場合は 80%の確率で C1，C2 の画像ペアに，20%の確率で B1，B2 の画像ペアに遷移する．Stage1 で画像の選択を行ってから報酬を得るまでを 1 トライアルとし，50 トライアル 4 セッション行った．実験には図 2.1 で示している画像と同じものを使用した．

Session	1	2	3	4
B1	80%	40%	10%	70%
B2	80%	40%	10%	70%
C1	20%	60%	90%	30%
C2	20%	60%	90%	30%

表 2.2 各セッションの報酬確率

各セッションにおける B1，B2，C1，C2 画像の報酬確率を示す.

2.3　行動データ解析

2.3.1　model-based / model-free 指標

1.3.1 のように, 報酬(rewarded/unrewarded)と遷移確率(common/rate)でソートした stay probability の傾向で，model-based / model-free の戦略の度合いが評価できることから，今回も被験者の Stage 1 での選択(一つ前の施行と同じ選択であれば 1, 違う選択であれば 0)を被説明変数とし，一つ前のトライアルの報酬の効果(rewarded/unrewarded)と遷移確率(common/rate)およびその相互作用を説明変数にしたロジスティック回帰を行い, 報酬の効果の係数を model-free 指標に，相互作用を model-based 指標にした．ロジスティック回帰では，ある事象が起きるか(y=1)起きないか(y=0)を考えるとき，以下の式のように表すことができ，Y を応答変数，a_pを回帰係数，x_pを説明変数とする．Y の値が 1 に近ければ近いほど事象が起きやすく，0 に近いほど事象が起きにくくなる．

$$Y=\frac{1}{1+e^{-(a_1X_1a_2X_2\ldots+a_pX_p+b)}}$$

今回はこの式を以下のように用い,

$$Y=\frac{1}{1+e^{-(a_1X_1+a_2X_2+a_3X_1*X_2+a_4X_3+a_5X_4+b)}}$$

X_1を報酬の有無(reward / non reward)，X_2を遷移(rare / common)，X_1X_2を reward*transition，X_3を Stage1 での選択，X_4を Stage2 での選択として，説明変数に用いた．ただし，今回は先行研究の課題ではランダムウォークで変動していた報酬確率を固定にしている．最近の研究（da Silva and Hare, 2019）では，この課題の改定を考慮したモデルとして，説明変数に Stage1 での行動選択（A1/A2)と，Stage2 での状態（B/C）を入れるモデルが提唱されており，今回はこちらを採用した．したがって説明変数は，報酬と遷移確率およびその相互作用，Stage1 での行動選択と，Stage2 での状態となる．また今回は，R の lme4 パッケージの glmer 関数を用いた mixed effect model を用いた logistic

regression を行なった(R Development Core Team, 2010)．モデルは上記の説明変数を fixed effect とし，b を random effect としたモデルを採用した．

2.3.2　認知機能バッテリー(CANTAB)

脳活動解析には年齢，性別，教育年数，IQ，認知機能指標 2 つと，上記で求めた model-based / model-free 指標を説明変数として用いた．また，確率的環境における意思決定に関わると考えられる行動の柔軟性や記憶力を評価する課題から得られる．2 つの指標を用いた．この指標は，認知機能指標として，CANTAB(Cambridge Neuropsychological Test Automated Battery)で提供されている．

(Ⅰ)　Intra-Extra Dimensional Set Shift (IED)

CANTAB(Cambridge Neuropsychological Test Automated Battery)で収集するデータの一つで，注意の維持・転換・柔軟性に関する課題である(図 2.3, 2.4)．課題では，ピンク色の図形と白線で構成された図を使用し，正しい方を選択する．1 つの段階で 6 回正解すると次の段階へ進み，ルールが変更される．図 2.4 において，9 段階を達成するまでのエラー数“total errors”を，柔軟性の指標として用いた．

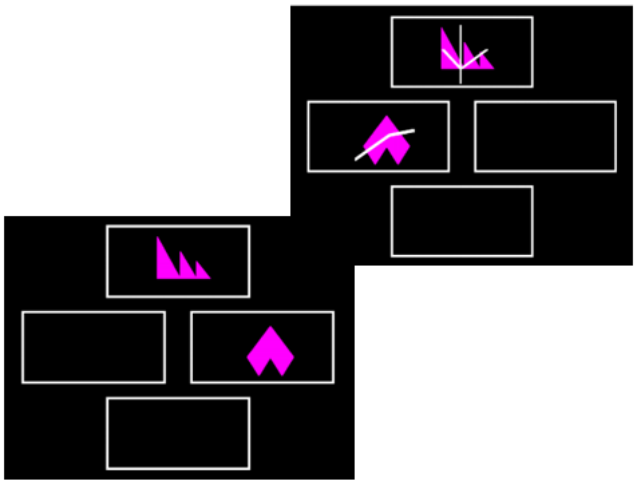

図 2.3 IED 課題

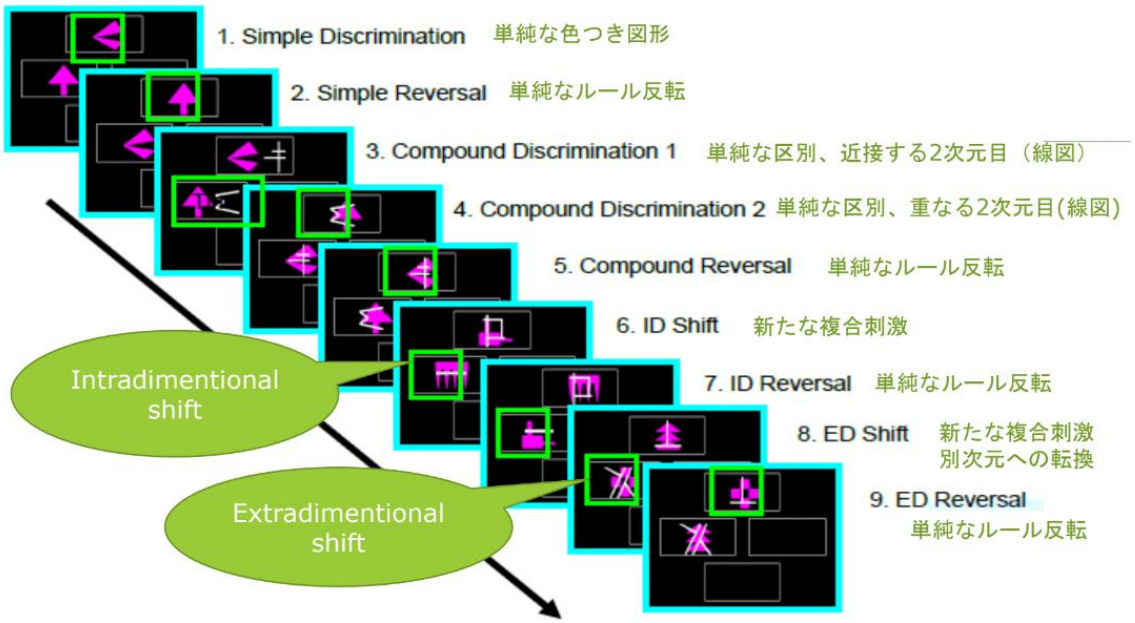

図 2.4 IED 課題の流れ

（II） Spatial Working Memory (SWM)

視覚情報の保持とワーキングメモリに関する課題である(図 2.5)．課題では，箱の中に隠されたトークンを見つける．一度見つかった箱からは二度とトークンは見つからない．トークンが以前に見つかった箱を再び見た数 “Between errors”をワーキングメモリ機能の指標として用いる．

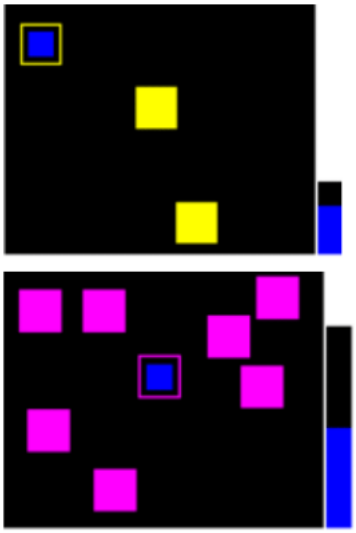

図 2.5 SWM 課題

各図は各タスクを示している．任意の四角(上画像 ： 黄色，下画像 ： ピンク)をタップし，箱を見つけていく．一度箱が見つかった四角からは二度と箱は見つからない．青色の箱が見つかったると右端のバーには見つけた個数分，箱がスタックされていく．黒い部分が無くなればすべての箱を見つけたことになり，そのタスクは終了する．

2.4　脳活動解析

2.4.1　撮像パラメータ

脳活動計測は，脳活動イメージングセンタにて行った．脳画像は，3T MRI 装置(MAGNETON Prisma, Siemens)を使用し，Gradient-echo planner imaging 法(TR : 2.5 秒, TE : 30 ミリ秒, Flip angle : 80 度, FOV 212×212mm: , ボクセルサイズ ：3.3×3.3×3.2mm, スライスギャップ ：0mm，スライス数 ：40)を用いて撮像した．また，脳構造画像を T1 強調画像法（ボクセルサイズ：1.0 x 1.0 x 1.0mm）によって撮像した．

2.4.2　安静時脳活動データの解析

SPM12(Statistical Parametric Mapping ： https://www.fil.ion.ucl.ac.uk/spm/) と CONN(functional connectivity toolbox ： https://web.conn-toolbox.org/)を用いて解析を行う．SPM は fMRI から得られる機能画像に対する解析手段である．fMRI では解析を行う前に解析画像に前処理を行う．今回の実験では，SPM は主に解析の前処理の段階で用いた．fMRI では, fMRI 内での被験者の頭部の微妙な動きや呼吸の仕方によってノイズが生じる．また, 脳構造の位置は個人によって差があるため, 解析を行う前に特定の座標値で表す必要がある．前処理では，このノイズの除去や座標値に対応する個人の位置特定を行う．

CONN は脳活動の結合の解析ための Matlab 上で動作するソフトウェアである．SPM で前処理を行った後, データを CONN で読み取り, 1st-level 解析ならびに 2nd-level 解析(詳

細は下部に記載)を行った

脳活動データの前処理について，脳活動計測で得られる EPI(Echo Planar Imaging)画像について, スキャン開始から 10 秒以内に撮像された画像は, 信号が不安定なため除去した．その後，頭部の運動補正を行ったあと，被験者ごとに T1 画像への位置合わせを行い，最後に平滑化を行った． 1st-level 解析では，各被験者のコントラストの評価を行う．2nd-level 解析では，1st-level 解析のあと，集団解析を行う．個人解析で評価した各被験者のコントラストを，集団解析で全被験者による総平均・標準偏差を算出し，課題に関連した脳活動が全体で一般的なものであるかを検証する．年齢，性別，教育年数，IQ，認知機能指標 2 つのデモグラフィックデータと，ロジスティック回帰により求めた model-based / model-free 指標を共変量として解析を行った．

第 3 章　結果

3.1　行動データ解析の結果

まず，被験者の行動パターンの傾向や学習ができているかどうかを調べるために，選択確率の正確さに関する移動平均(図 3.1)のプロットを行った．各図では，セッションごとに全被験者の平均と標準誤差をプロットしている．ここで言う選択確率とは，Stage1 で画像を選択するときに，報酬をもらいやすい画像ペアへ遷移しやすい画像を選択する確率である．縦軸が選択確率，横軸がトライアル数を表している．セッション 1 からセッション 4 までの報酬確率については表 2.2 に示してある．

次に，データに対して線形回帰を行い，グラフの傾きを調べた．傾きが右肩上がりで急であればあるほど学習スピードが早いことが分かる．図 3.2 に示す．また，y=ax+b として，各セッションにおける近似した直線のパラメータを表 3.1 に示す．この結果から，報酬確率の難易度が比較的簡単であった session1(80/20)と session3(10/90)は 4 セッションの中で傾きが右肩上がりで急になっている．次に，session4(70/30)では、session1 と session3 ほどではないが，右肩上がりの直線を示している．最後に，最も報酬確率の難易度が高かった session2(40/60)に関しては右肩下がりであり，最適行動を学習できていないと言える．

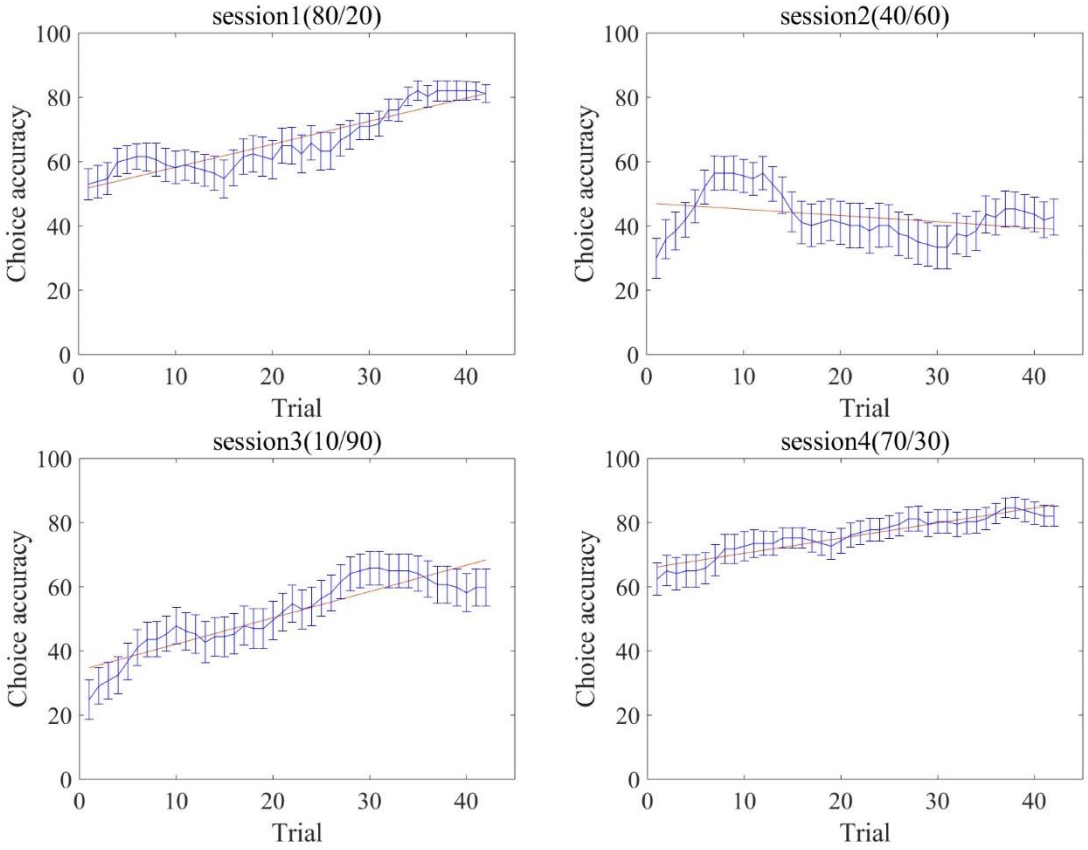

図 3.1 選択確率の正確さに関する移動平均

Session	a	b
1	0.7167	51.0736
2	-0.1947	47.1445
3	0.8194	33.9488
4	0.4718	65.7216

表 3.1 近似した直線のパラメータ

次に，stay probability の結果を示す.

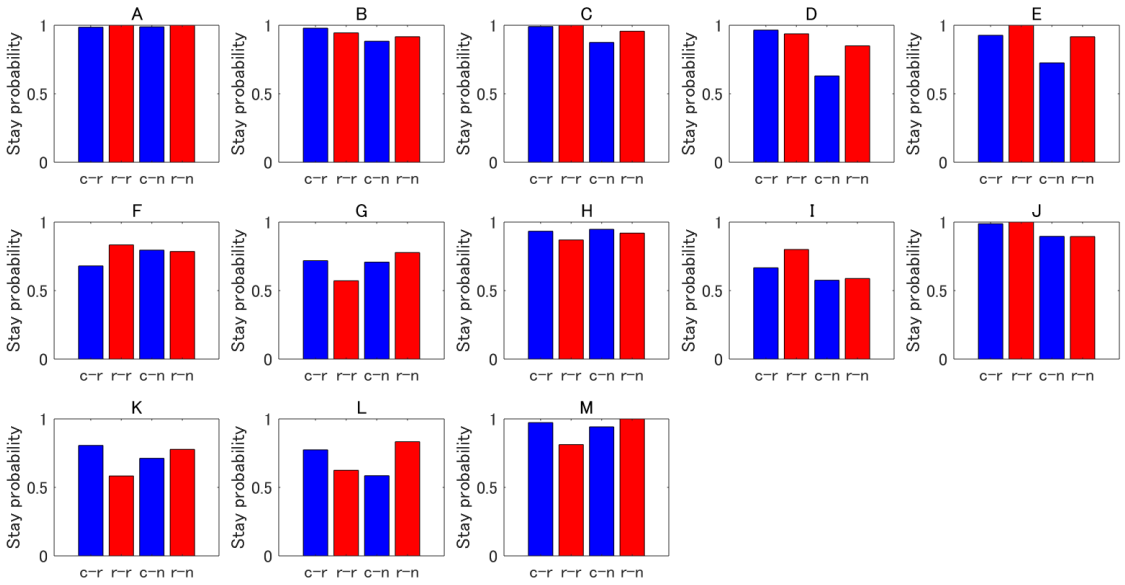

図 3.2 Stay probability

c-r : common-reward, r-r : rare-reward, c-n : common-no reward, r-n : rare-no reward

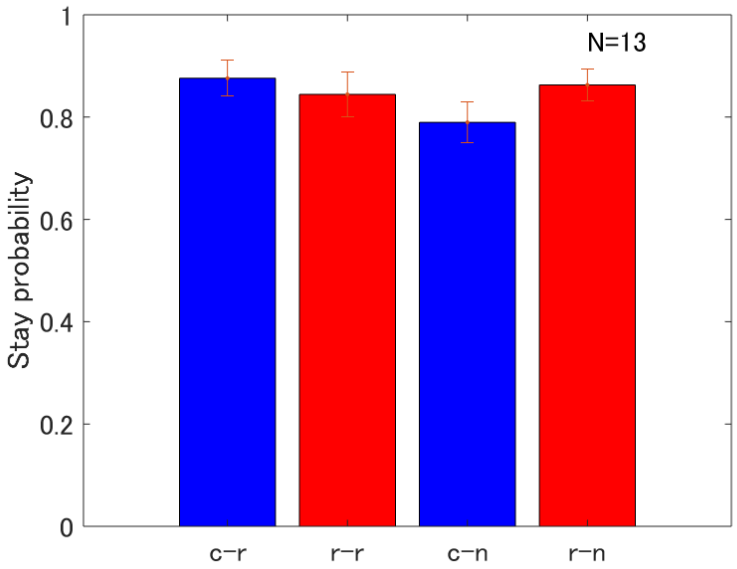

図 3.3 全セッションにおける stay probability 平均

c-r : common-reward, r-r : rare-reward, c-n : common-no reward, r-n : rare-no reward

ロジスティック回帰によって推定した model-free/model-based 指標の結果を表 3.2 に示す.

被験者	model-free	model-based
A	-0.053138	0.00652
B	0.294819	0.003962
C	0.325005	0.005581
D	-0.391671	0.017513
E	0.114028	0.003843
F	0.021229	-0.0048
G	-0.30887	0.002707
H	-0.200211	0.003885
I	0.011054	-0.000574
J	0.302971	0.005434
K	-0.989716	0.013932
L	-0.425961	0.006878
M	-0.670314	0.011719

表 3.2 model-free/model-based 指標

図 3.4 に示された平均の stay probability は，先行研究(Daw et al.,2011)の結果で示されていたハイブリット型の stay probability と同じような形のプロット(図 1.4)となっており，全体の平均で見ると，ハイブリット型の行動戦略を行っていた．また，高齢者の行動戦略は model-free に偏ると仮説を立てていたが，ロジスティック回帰によって推定した model-free/model-based 指標からも分かるように，model-free に偏るということはなかった.

また，行動戦略と個人特性の関係を調べるために，model-free 指標と model-based1 指標の年齢と IQ に対する相関係数と p 値を求めた．表 3.3 に示す.

	相関係数	p 値
年齢・model-free 指標	-0.0641	0.8352
年齢・model-based 指標	0.6207	0.0236
IQ・model-free 指標	0.2374	0.4349
IQ・model-based 指標	0.1531	0.6176

表 3.3 相関係数・p 値

このことから，年齢と model-based index が有意に正の相関があることが分かる．

最後に，ロジスティック回帰により得られた model-free と model-based の mixed effects パラメータについて，本研究と先行研究における値の比較を行った．表 3.4 に示す．

	Model-free(P-value)	Model-based(P-value)
本研究	-0.05(0)	0.006(7.74e-13)
Otto et al., 2013	0.80(<0.0001)	0.25(<0.0001)
Gillan et al., 2016	1.06(<0.0001)	0.24(<0.0001)
Eppinger et al., 2013	0.26(0.05)	1.62(<0.001)

表 3.3 mixed effects パラメータ

先行研究と比較すると model-free / model-based いずれも極めて値が小さいが，報酬の効果はゼロとは有意に異なっており，また，報酬と遷移確率との相互作用もまた有意であり，多くの被験者が両者の戦略を用いていたことが分かる．

3.2　脳活動データ解析の結果

安静状態の脳活動から，安静時機能結合を計算し，model-free/model-based 指標と相関のある結合を求めた．model-free 指標と有意な相関があった場合，model-based 指標と有意な相関があった場合を表 3.5-表 3.6 に記載する．特に相関の高い結合を図 3.5-図 3.6 に示す．今回の解析では，ボクセル毎に一般線形モデルをあてはめ，推定されたパラメータから算出される統計量を図示しており，この統計量に閾値を設けた上で脳画像上にプロットしている．model-free 指標と有意に正に相関する結合として，“下側頭回後部右側 - 楔部左側”が見つかった． model-based 指標と有意に正に相関する結合として，“小脳 1 野左側 - 前頭頭頂 PPC”，“前頭頭頂 PPC - 小脳 1 野左側”が見つかった．

また，説明変数として，性別，年齢，教育年数，IQ，柔軟性指標，ワーキングメモリ指標を用いており，それぞれと有意に相関の見られた結合を表 5.1-表 5.6(第 5 章)に示す．特に相関の高い結合を図 5.1-図 5.6(第 5 章)に示す．

Analysis Unit	T-value	p-FDR
下側頭回後部右側 - 楔部左側	5.71	0.0143
楔部左側 - 下側頭回後部右側	5.71	0.0160
小脳 8 野左側 - (146)	-3.91	0.0482
小脳 8 野左側 - 前頭皮質右側	-3.98	0.0482
小脳 8 野左側 - 縁上回後部右側	-4.01	0.0482

表 3.5 model-free 指標

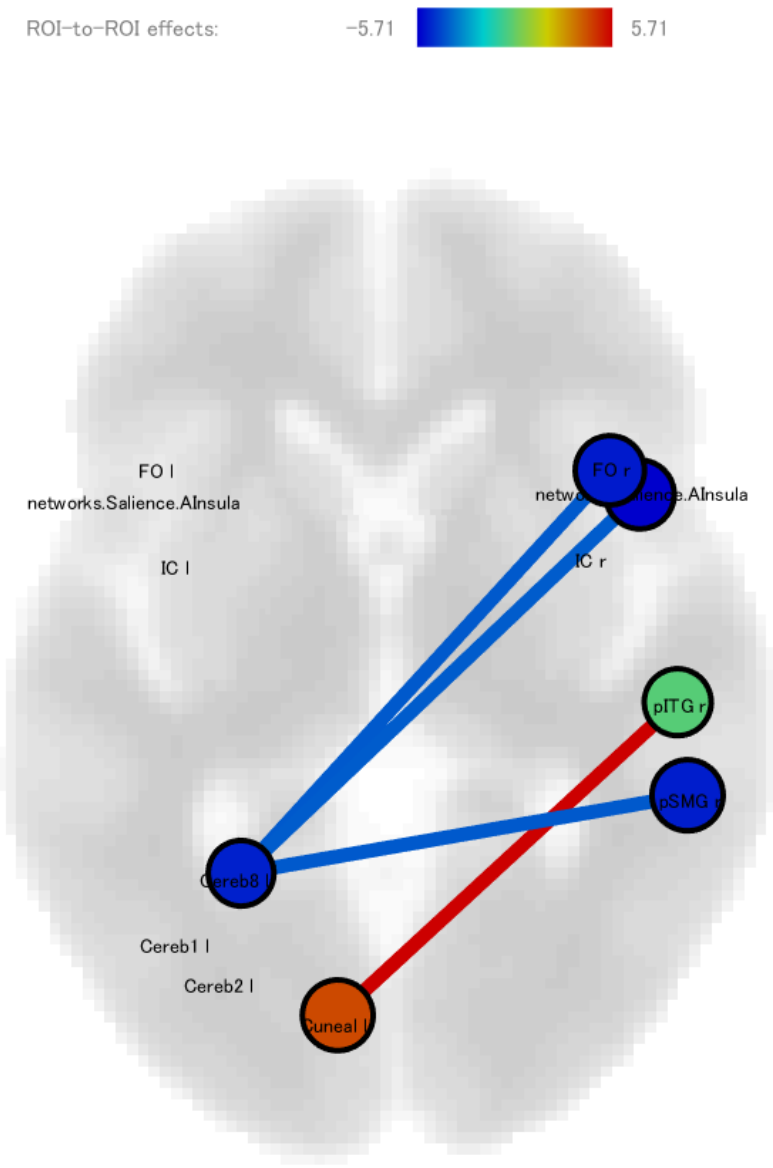

図 3.4 model-free

Analysis Unit	T-value	p-FDR
小脳 1 野左側 -前頭頭頂 PPC	9.92	0.0001
前頭頭頂 PPC - 小脳 1 野左側	9.92	0.0001
小脳 1 野左側 - 角界左側	7.95	0.0003
角界左側 - 小脳 1 野左側	7.95	0.0004
中側頭回前部右側 - 中側頭回後部左側	7.90	0.0006

表 3.6 model-based 指標

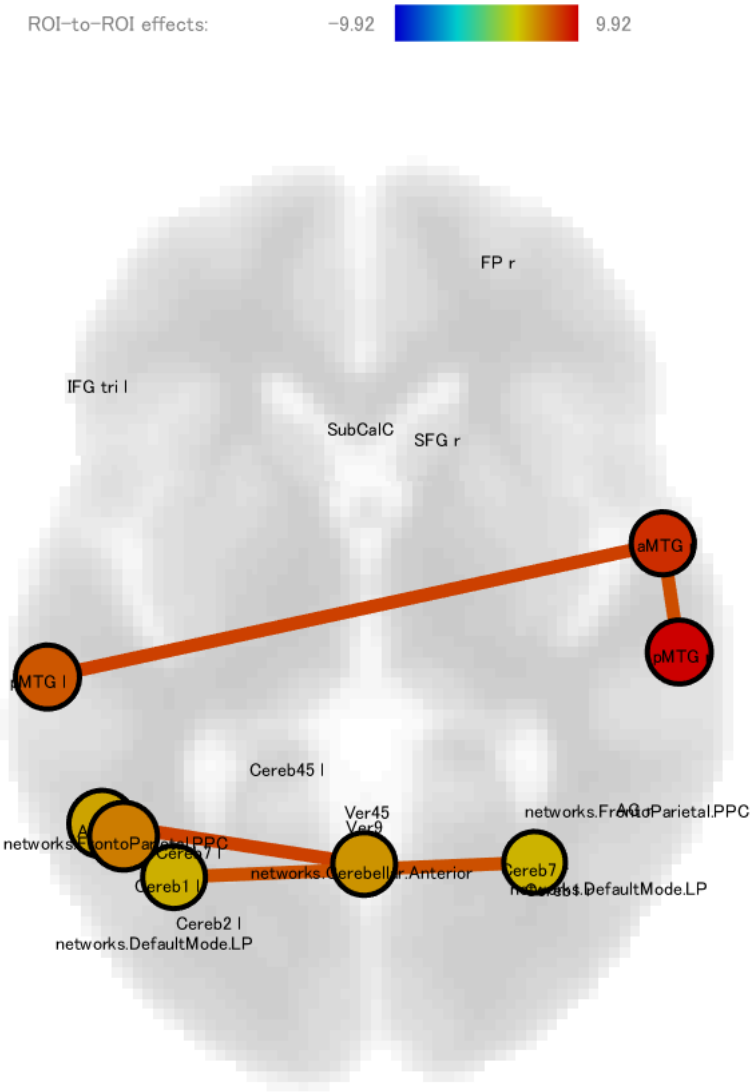

図 3.5 model-based

第 4 章　考察と結論

4.1　考察

4.1.1　行動結果

選択確率の正確さに関する移動平均の結果から，最適行動を最も学べていたのは session1(80/20)と session3(10/90)であった．また，最適行動を学習できていなかったのは session2(40/60)のみであった．4 セッション中 session2 が報酬確率の幅が狭く，モデルを構築するのが難しかったためであると考えられる．Stay probability の全体の平均と照らし合わせてみると，model-free / model-based のハイブリットのプロットとなっており，仮説のように全体として model-free に偏っているということはなかった．今回はサンプルサイズが小さく高齢者全体の傾向を捉えられていない恐れがあるため，被験者の年齢と行動戦略の関係を調べたところ，年齢と model-based 指標に有意な正の相関が見られ，仮説と逆の結果を得た．どちらの係数も小さい被験者もいたため，model-free / model-based の枠組みで説明できない被験者も存在する可能性がある．

Stay probability とロジスティック回帰によって推定した model-free/model-based 指標からは，被験者は model-free / model-based のどちらの戦略も用いていることが分かった．また，表 3.2 で示した mixed effects のパラメータ比較から，先行研究(Eppinger et al., 2013)のように, model-based の精度が良くないことに関しては今回の実験でも明らかとなったが, 仮説のように， model-free 寄りの行動戦略を取っていたとも言えない．

また，個人差も大きく，課題に対する理解が追い付いていない被験者がいた．そこで，IQ と行動戦略の関係を調べたところ，IQ と model-free / model-based 指標の間には有意な相関は見られなかった．

4.1.2　脳活動データ結果

Model-free 強化学習は中脳ドーパミンと大脳基底核によって構成される神経回路によって実現されていると考えられており, model-based 強化学習は前頭前野を中心とする大脳皮質の回路によって実現されていると考えられている．また，先行研究(Daw et al.,2011)では，腹内側前頭前野と腹側線条体の 2 つの領域が報酬予測と相関があると言われている．今回の脳活動データの解析では，model-free / model-based ともに先行研究と一致するような結果を得られなかった．

報酬関連の活動と脳活動調査した研究(Cox et al.,2008)では，高齢者は楔部と報酬関連の活動に正の関係があることが示されている．楔部は，視覚処理が行われる場所である．今回の解析では，model-free 指標に楔部における正の相関が見られた．報酬予測から行動選択を行う model-free では，model-based と比べてより報酬に対する依存があると考えると，楔部において正の相関が見られたのではないかと考えた．

また，Akrami(2018)では，過去の環境からの入力と後部頭頂皮質(PPC : Posterior parietal cortex)における表現と行動の間に強い相関が見られることが明らかになっている．PCCは，デフォルトネットワークの一つで，行動や注意力において重要な役割を果たす．今回の解析では，model-based指標に後部頭頂皮質における正の相関が見られた．過去の行動や刺激から現在の行動に至る際に後部頭頂皮質が関連していることから，model-basedにおいて後部頭頂皮質に正の相関が見られたのではないかと考えた．高齢者では，PCCのネットワークが低下する傾向がある．PPCが損傷すると，記憶の欠如など，様々な感覚障害を生じる可能性があるが，今回被験者にmodel-basedの被験者もみられたことから，他の被験者と比較してPPCが低下していない被験者がmodel-basedの行動を示したのではないかと考察する．

4.2 結論

本研究では， model-based / model-free の意思決定手法を用いて高齢者の行動手段を明らかにし，安静時の脳活動と関連付けた．高齢者は認知機能の低下により model-based の精度が悪くなり，model-free 寄りの行動を示すと仮説を立てたが，行動実験の結果から，被験者全体が model-free 寄りの行動戦略を取るわけではなく，先行研究ほど精度は良くないが，model-based の戦略を示す被験者もいた．脳活動について，model-free / model-based それぞれが関連すると言われている脳部位に関する相関は見られなかった．しかし，model-free においては，報酬関連の活動と正の相関を示す楔部に相関が見られ，model-based においては，刺激履歴と相関がある後部頭頂皮質に相関が見られた．

4.3 今後の課題

本研究では高齢者に絞って実験・解析を行ったため，若年者でも同様の実験・解析を行い比較をすることで今後の考察や見通しがしやすくなると考える．また，被験者ごとに課題の達成度に大きな差がみられたため，被験者の数をもっと増やすことで，全体としての傾向をよりつかむことができるようになると考える．

謝辞

本研究を行うにあたり，研究に関して指導していただいたATR脳情報通信総合研究所 数理知能研究室 田中沙織室長ならびに吉田和子先生に深く感謝致します．お二人には，研究に関してのアドバイスを始め，研究に対する姿勢についても教えていただきました．また，神経科学に対してバックグラウンドが全くなかったにも関わらず，快く計算神経科学研究室の学生として受け入れてくださった川人光男教授にも深く感謝致します．加えて，数理知能研究室の酒井雄希先生，井上佳奈さん，舘香織さん，泉原延幸さんにも，セミナー等で研究に関するご助言をいただくなど，多大なるご支援，ご協力をいただき，心より感謝申し上げます．基幹研究室である知能コミュニケーション研究室の中村哲教授ならびに田中宏季先生にも，研究室合宿ならびにセミナーにおいてご助言いただき，大変感謝しております．知能コミュニケーション研究室の秘書の松田真奈美さん，ATR 数理知能研究室の前任秘書内田抄代子さん，現任秘書美並優子さんには学生生活全般をサポートしていただきました，ありがとうございました。そして，日々の学生生活において時には相談に乗り，時には励ましてくれた友人たち，先輩方，本当にありがとうございました。最後に，大学院進学の道を快諾した上で二年間の大学院生活を支えてくれた家族に，心から感謝の意を表します．

参考文献

[1] Eppinger et al. “Of goals and habits : age-related and individual differences in goal-directed decision-making” Frontiers in neuroscience 7(2013)

[2] Daw et al. “Model-Based Influences on Humans’ Choices and Striatal Prediction Errors” Cell Press(2013):1204-1215

[3] Decker et al. “From Creatures of Habit to Goal-Directed Learners: Tracking the Developmental Emergence of Model-Based Reinforcement Learning” Free PMC Article(2016)

[4] Otto et al. “Working-memory capacity protects model-based learning from stress” PNAS (2013)

[5] Gillan et al. “Model-based learning protects against forming habits” Free PMC Article (2015)

[6] Dall et al. “The ubiquity of model-based reinforcement learning” Free PMC Article (2012)

[7] Jan et al. “Model-based approaches to neuroimaging: combining reinforcement learning theory with fMRI data” WIREs COGNITIVE SCIENCE(2010)

[8] 木村元・宮崎和光・小林重信 “強化学習システムの設計指針”(1999),[online] http://sysplan.nams.kyushu-u.ac.jp/gen/papers/sice99.pdf、参照(2018-12-15)

[9] 片平健太郎 “学習の理論から強化学習，計算論モデリングへ”(2017),[online] https://psych.or.jp/wp-content/uploads/2017/09/78-17-20.pdf、参照(2018-12-27)

[10] 脳画像計測入門．“SPM 操作法”脳画像計測入門．(2009),[online] http://www.geocities.jp/fmri_lab_1968/spm_howto.html、参照(2019-1-5)

[11] Platinum Data Blog. “強化学習入門〜これから強化学習を学びたい人のための基礎知識〜“Platinum Data Blog.(2015),[online]http://blog.brainpad.co.jp/entry/2017/02/24/121500、(参照2019-1-20)

[12] 坂上雅道・山本愛実 “意思決定の脳メカニズム-顕在的判断と潜在的判断-”(2009),[on

line]https://www.jstage.jst.go.jp/article/jpssj/42/2/42_2_2_29/_pdf、(参照2019-1-21)

[13] K-Lab. “脳画像解析ドキュメント”K-Lab.(2012),[online]http://www.nemotos.net/?page_id=127、(参照2019-1-23)

[14] Cox et al. “Striatal outcome processing in healthy aging” CABN (2008):304-317

[15] Akrami et al. “Posterior parietal cortex represents sensory history and mediates its effects on behaviour” Nature554 (2018):368-372

第 5 章　付録

説明変数として用いた性別，年齢，教育年数，IQ，柔軟性指標，ワーキングメモリ指標それぞれに有意に相関の見られた結合を表 5.1-表 5.6 に示す．特に相関の高い結合を図 5.1-図 5.6 に示す(可視化するために，性別，Total error，Between errors に関しては閾値を 0.001 に，年齢，IQ，Education に関しては閾値を 0.0000005 とした)．

Anotysis Unit	T-value	p-FDR
ICC r - 舌状回左側	53.48	0.0001
舌状回左側 - ICC r	53.48	0.0001
下前頭回三角部右側 - (160)	41.28	0.0003
(160) - 下前頭回三角部右側	41.28	0.0003
(159) - 下前頭回弁蓋部左側	38.83	0.0004

表 4.1 性別

Analysis Unit	T-value	p-FDR
Precuneous - (136)	37.08	0.0000
(136) - Precuneous	37.08	0.0000
舌状回右側 - (140)	26.87	0.0000
(140) - 舌状回右側	26.87	0.0000
PreCG l - (137)	26.64	0.0000

表 5.2 年齢

Analysis Unit	T-value	p-FDR
Precuneous - (136)	34.04	0.0000
(136) - Precuneous	34.04	0.0000
ICC l - SCC r	30.20	0.0000
SCC r - ICC l	30.20	0.0000
SCC r - ICC r	25.82	0.0000

表 5.3 IQ

Analysis Unit	T-value	p-FDR
Precuneous - (136)	24.05	0.0000

(136) - Precuneous	24.05	0.0000
楔部左側 - 楔部右側	22.34	0.0000
楔部右側 - 楔部左側	22.34	0.0000
舌状回右側 - (140)	20.65	0.0000

表 5.4 Education

Analysis Unit	T-value	p-FDR
横側頭回右側 - 小脳 45 野左側	7.78	0.0008
小脳 45 野左側 - 横側頭回右側	7.78	0.0008
傍帯状回右側 - 島皮質右側	7.32	0.0012
島皮質右側 - 傍帯状回右側	7.32	0.0015
傍帯状回右側 - PP r	6.99	0.0012

表 5.5 Total error

Analysis Unit	T-value	p-FDR
(134) - 高次物体処理領野左側	9.47	0.0001
高次物体処理領野左側 - (134)	9.47	0.0001
偏桃体左側 - 偏桃体右側	9.29	0.0001
偏桃体右側 - 偏桃体左側	9.29	0.0001
偏桃体左側 - 側頭極右側	9.22	0.0001

表 5.6 Between errors

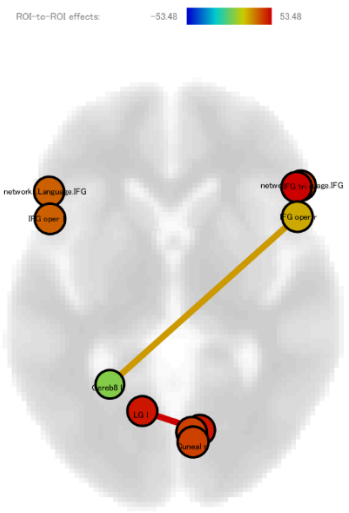

図 5.1 性別

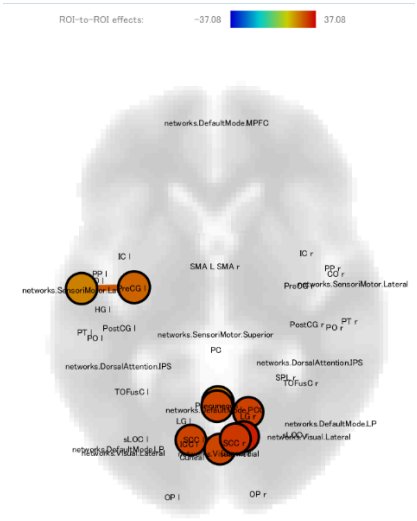
ROI-to-ROI effects:
-37.08
37.08
networks.DefaultMode.MPFC
IC l
IC r
SMA L SMA r
networks.SensoriMotor.Lateral
HG l
networks.SensoriMotor.Superior
PC
networks.DorsalAttention.IPS
TOFusC l
networks.DefaultMode.LP
networks.Visual.Lateral
OP l
OP r

図 5.2 年齢

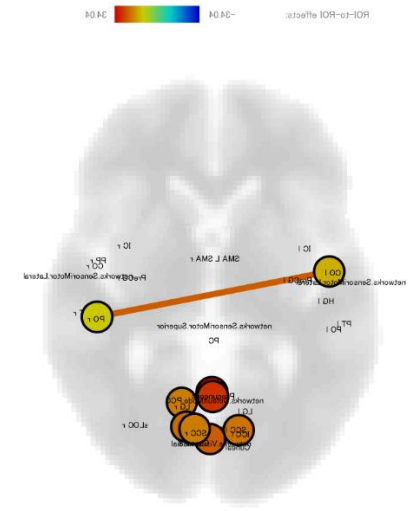
networks.SensoriMotor.Superior
PC

図 5.3 IQ

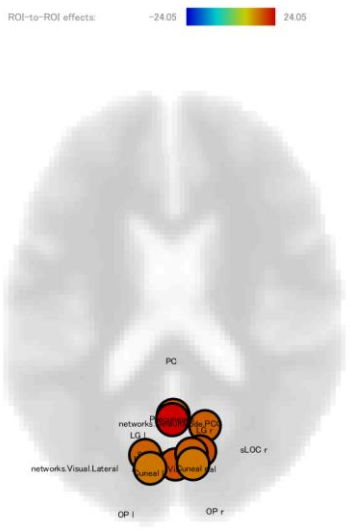

図 5.4 Education

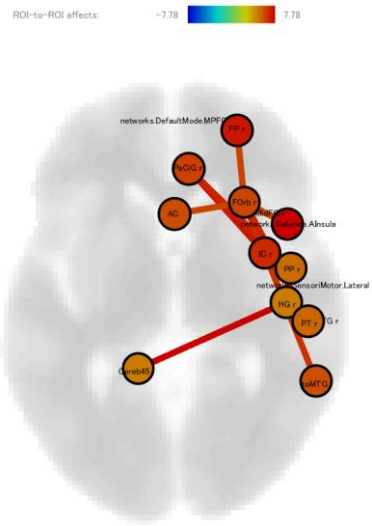

図 5.5 Total error

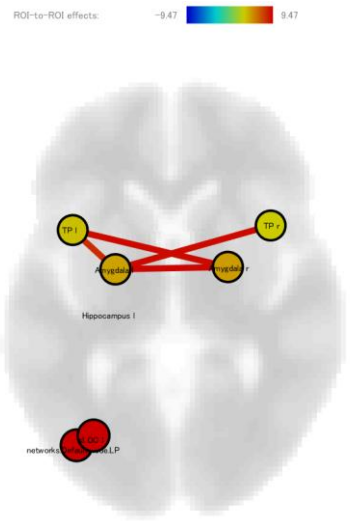

図 5.6 Between errors