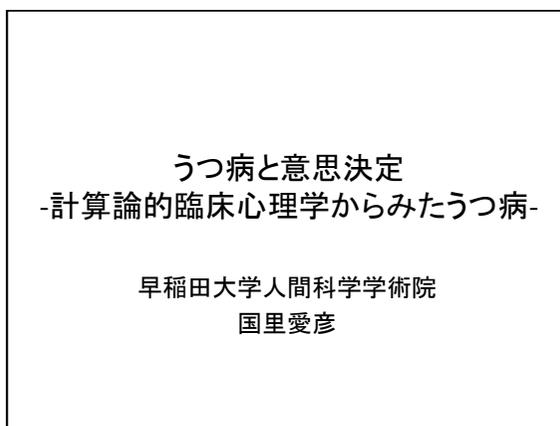


【講演2】

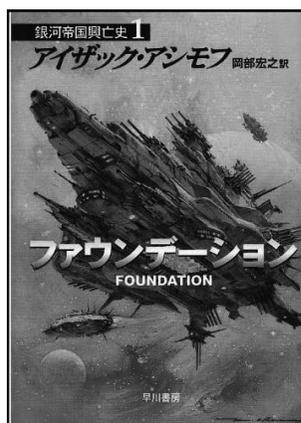
国里 愛彦（早稲田大学助手）

「うつ病と意思決定：計算論的臨床心理学からみたうつ病」

早稲田大学人間科学学術院の国里愛彦と申します。今日の内容としましては、「うつ病と意思決定：計算論的臨床心理学からみたうつ病」という内容でお話しさせていただきます。こんな大きなタイトルをつけてしまったことを後悔はじめているのですが、細かい研究内容についてもお話ししていきたいと思っています。



最初に、アイザック・アシモフの『ファウンデーション』という小説があります。SFで非常に有名なアシモフの小説です。この中でSFがお好きな方もおられるかもしれませんが、『われはロボット』などを書いている著名なSF作家になります。このアシモフの書いた『ファウンデーション』という小説は、主人公が数学者の小説です。数学者が人の集団行動をシミュレーションすることを可能にしたというのが物語の前提であります。その結果、銀河帝国が崩壊するという予想をしてしまい、逮捕されてしまうという話から物語がスタートします。あくまでアシモフの小説というのはSFの話ではあるのですが、実際にこういう人の意思決定というものが、現在の科学の中で実際に計算することができるのだろうか、そういう疑問が出てきます。こういった、人の意思決定



- 数学者が人の集団行動をシミュレートし、銀河帝国の崩壊を预言する。
- 人の意志決定は計算可能か？

を計算していく，数学的に解いていくということは，計算論的神経科学という分野で行われつつあります。

私はうつ病の臨床心理学的研究をしてきているのですが，計算論的神経科学での問題を臨床心理学的な問題に置き換えたとき，「うつ病患者さんの意思決定は計算可能か？」というリサーチクエスチョンに置き換えることができます。果たして，現在の段階で，うつ病患者さんの意思決定というのは数学的に計算可能でしょうか？今日は，このことについて，お話ししていきたいと思います。

うつ病患者さんの意志決定は計算可能？

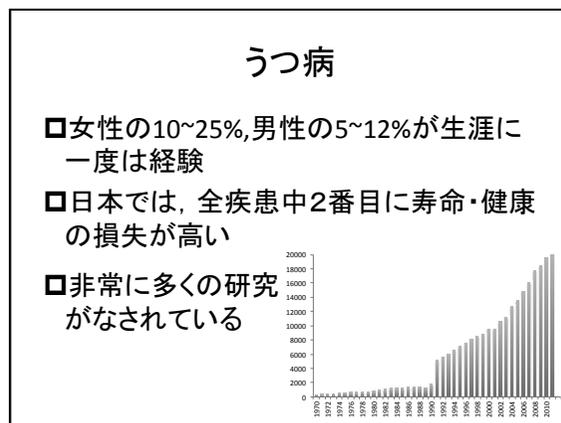
今日，お話しする内容になります。最初にうつ病に関する基本的な内容についてお話ししていきます。その後，うつ病に関する意思決定の研究について少しだけ触れて，今日のお話のメインテーマのうつ病の意思決定を考える上での枠組みとしまして，強化学習理論についてお話しします。こういった強化学習理論というものから導かれる銅谷仮説，さらにうつ病と銅谷仮説の関連についての話をしていきます。最後に副題にありました計算論的臨床心理学のことについて，まとめに代えてお話ししていきたいと思います。

本日の話題

- うつ病について
- うつ病の意志決定
- 強化学習理論とDoya仮説
- うつ病とDoya仮説
- 計算論的臨床心理学

それでは早速，うつ病についての説明になります。うつ病は，女性の10～25%，男性の5～

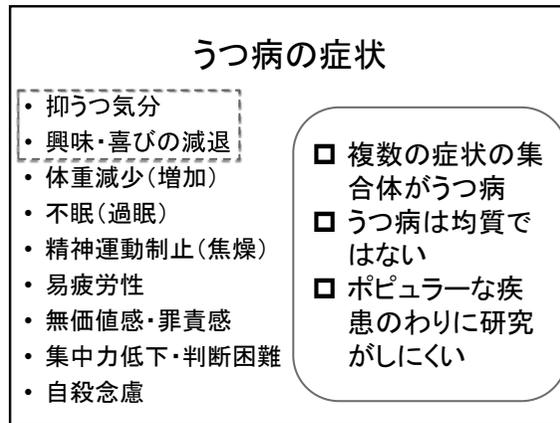
12%が生涯に1度は経験する、非常に多くの方がなり得る疾患になります。生涯罹患率において女性が2割、男性は1割程度という非常に高い確率でかかる可能性のある疾患になります。さらに罹患率が高いだけでなく、うつ病は、精神疾患やそれ以外の全疾患をまとめた中でも2番目に寿命や健康の損失が高いといわれています。2020年という近い将来、かなり高い確率でうつ病による寿命や健康の損失が高くなるであろうという試算が出されています。こういったこともありまして多くの研究が行われています。Web of Scienceなどで調べて、昔からたどっていきますと、どんどんうつ病に関する論文が増えていっています。現在は、年間2万本近く論文が出版されています。このように多くの研究がなされてきています。



うつ病の症状は非常に多岐にわたるものになります。うつ病と聞きますと何となく典型的なイメージで憂うつそうに引きこもっている姿が頭に浮かんできますが、それ以外にもいろいろな症状をもっているわけです。よく出てくる症状としては、抑うつ気分と興味・喜びの減退、これがもっとも主な症状になります。この2つの症状のうち1つを満たしていないとうつ病という診断はされません。この2つがないうつ病はあり得ません。それ以外の症状ですと、例えば体重減少でありますとか、不眠に関するもの、なかなか夜寝つけないでありますとか、朝早く目が覚めてしまう、早朝覚醒がある、中途覚醒がある、そういったものがあります。さらに精神運動性の制止、動作が遅くなるとかいらいら感が強くなるといったものもあります。また、疲れやすくなるとか、自分の価値がないように感じてしまうというものもあります。さらに意思決定でありますとか、認知機能に関係するもの、先ほどの発表で注意の話がありましたが、例えば集中力が低下するとか、判断が難しくなるということもよくあります。「以前は献立を立てるのが簡単だったのに、うつ病になると献立がなかなか立てられなくなりました」ということを多くの患者さんがおっしゃいます。あとは希死念慮というのが非常に多くのうつ病患者さんにおいて認められます。多くの方が希死念慮をもっておられ、中には実際に自殺してしまう方も多いというのがこの疾患の特徴になります。

見ていただいたとおりですが、複数の症状が集まったものがうつ病であるということが出来ます。そういう意味ではうつ病というのはあまり均質ではない疾患ということが出来ます。いろいろなタイプの患者さんがおられて、それをひとくくりでうつ病と呼んでいるのが現状であります。そ

いう意味で、非常に多くの方が罹患するポピュラーな疾患の割に研究があまり進んでいないといううつ病研究の特徴であるということが出来ます。



ここから少し研究の話になっていきますが、うつ病が均質ではないということから、うつ病研究では何がコアなのか、といったことが問題になってきます。それに対して、中間表現型に注目した研究がここ10年行われるようになってきています。この中間表現型というのはどういうものかといいますと、まず私たちが実際にお見かけするうつ病患者さんの症状としてあらわれているものは表現型といいます。症状としてあらわれているのが表現型です。表現型には、遺伝子が影響するのですが、そういった遺伝子と表現型として実際に現れる症状との間にはギャップがあります。そこで、遺伝子と表現型との間に中間となる表現型というものがあるのではないかと考えられるようになりました。そのような中間表現型がどういうものかといいますと、ある程度うつ病の中でまとまりのある症状でありますとか、脳機能の異常といったものとして表現されます。

このうつ病の中間表現型としましては、Haslerがまとめています。だいたいこの4つが有力ではないかといわれています。その1つ目がアンヘドニアというものになります。これは最初に出てきました喜びや興味の喪失と同じもので、快感情を失ってしまうことを意味します。「これまで楽しいと思えていたことも楽しめなくなる」といった患者さんの報告がアンヘドニアになります。アンヘドニアは、有力な中間表現型の1つです。さらにストレスへの感受性もあります。これは、ストレスを負荷されたときに脆弱性が表出することになります。また、トリプトファン欠乏実験という手続きがあります。これは、セロトニンの前駆物質のトリプトファンというものがあるのですが、それを摂取しないということを人為的にする手続きです。私もこれをやっていたのですが、この手続きによってセロトニンを一時的に欠乏させることができます。こういったセロトニン欠乏手続きによって気分が悪化します。最後に、これもよく言われていますが、ストレスを感じたときの cortisol の反応も中間表現型の1つとして挙げられています。

今回はうつ病のアンヘドニアについてお話していきます。うつ病のアンヘドニアについてはいくつかの研究がなされています。こういったアンヘドニアの研究というのは、だいたい好きな飲み物ですとか、お金とか、好きな音楽、そういったものを見たり聞いたり楽しんでいるときの

研究上で何がコアなのか？

□中間表現型：表現型と遺伝子型の間となる表現型

□うつ病の中間表現型

- (1)アンヘドニア(快感情の喪失)
- (2)ストレス感受性(神経症傾向など)
- (3)セロトニン欠乏手続きによる気分悪化
- (4)過剰なコルチゾール反応

Hasler et al., *Neuropsychopharmacology*, 2004

脳活動を調べています。実験参加者の方にはMRI装置の中に入れていただいて、チューブでジュースを飲んだりとか、お金がもらえるようなゲームをしたり、好きな曲を聞いてもらったりしています。そのときの脳活動を調べる実験なのですが、うつ病患者さんと健康な方で比較していきます。その結果、好きな飲み物を飲んでいるときも、金銭的な報酬をもらっているときも、好きな音楽を聞いているときも、全部同じような脳領域でうつ病患者さんは活動が低いということがわかってきました。この領域というのは報酬処理にかかわる線条体という場所になります。腹側の線条体の活動が低くなることによって、うつ病患者さんはアンヘドニアになっているのではないかということがいわれています。なので、興味、喜びの喪失を感じてしまうというときに、やる気がないからというわけではなくて、脳機能の面でもなかなかそういうことが難しい状態なのだと考えることができます。

うつ病のアンヘドニアと脳機能

□好きな飲み物

(McCabe et al., *Psychopharmacology*, 2009)



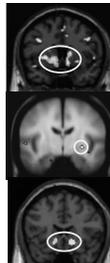
□金銭的報酬

(Pizzagalli et al., *Am J Psychiatry*, 2009)



□好きな音楽

(Osuch et al., *Neuroreport*, 2009)



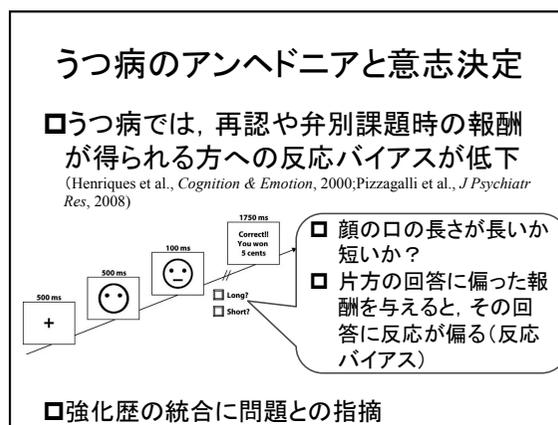
□報酬処理に関わる腹側線条体の活動がうつ病患者において低下

さらにこういったアンヘドニアというものがどのように行動面に影響してくるのか、どういう影響を与えるのかという、そういった意思決定にかかわる研究もなされてきています。HenriquesやPizzagalliなどが多くの研究を行ってきています。彼らの研究というのは、うつ病患者さんに対して再認でありますとか何か刺激を選択させる課題を行ってしています。そういった課題を行ったとき

に、片方の選択にだけ報酬が得られるようなやり方でフィードバックします。そうすると報酬がもらえるほうに我々は傾いていくのですが、うつ病患者さんでは、それがなくなるということが言われています。

いまの説明だとわかりにくいと思うので、具体的な課題でご説明しますと、こんな漫画の顔が出てきます。この漫画の顔の口の長さが長いか短いかという判断をしてもらいます。長い場合は右ボタン、短い場合は左ボタンを押すみたいなことをします。実際のところ、長い口と短い口は同じ確率で出てきます。しかし、片方の、例えば長い口のほうだけ少し報酬をもらいやすくする。つまり、長い口を選んだときのほうがよりお金をもらいやすくしておくと、多くの人は気づかないうちにそういうものを選ぶようになっていきます。そういったことがうつ病患者さんでは生じにくくなるということがいわれています。

これができるのは、何かを選択したことで報酬がもらえたという履歴を利用できるためなのですが、そういった履歴をうまく統合できていないのではないかと考えられています。ただ、この課題には少し脆弱な面があります。私も1回実験で使ったことがあるのですが、実験協力者の方がおっしゃったのは、「なんかこっちのほうが報酬もらえるから押しました」という言い方をされていて、気づいちゃうと簡単にできちゃうという面がありますので、少し課題として脆弱かなという印象をもっています。より一般的な学習状況において、うつ病患者さんと健康な方との差を検出できるような実験課題はないか、そういう研究をする枠組みが何かないか探していました。



私は広島大学の大学院にいたのですが、その時は、広島大学精神科で強化学習を精神医学でちょっと使ってみようという機運が高まってきていた時期でした。そういうこともあって、強化学習というものをすんなり勉強することができました。それについて、いまからお話していきます。

こういったうつ病の研究は非常に多くなされてきています。年間2万本、もっと出ているかもしれません。そういった研究知見がどんどんたまっていくのはいいのですが、私たちはまだにうつ病というものがよくわかりません。研究をすればするほど、うつ病というのは見えなくなっていく面があります。いろいろな仮説が出てきています。古くは、学習性無力感がありますし、近いものですと例えば強化歴の統合みたいな話もあります。ただこういった複数の小さな仮説が

乱立していて、いまいち全体像がとらえられないなという印象があります。うつ病における意思決定の問題を考える上で、意思決定に関する基礎的な枠組みをきっちりおさえた上でうつ病の知見をちゃんと整理していった、新たに何か仮説を作る必要があるのではないかと考えています。

研究知見が増えて、うつ病は理解できた？



- 複数の小さな仮説が乱立し、全体像が捉えられない
- 意志決定の基礎的な枠組みから、うつ病の知見を整理し、新たに検討する必要がある

そこで、強化学習理論というものを基礎におくことにしました。強化学習と聞きますと、なんとなく私たち心理学者は、オペラント条件づけが頭に浮かぶわけですが、少し異なっています。ここで話をする強化学習というのは、Sutton & Bartoが提案した強化学習になります。これは、オペラント条件づけのような動物の学習実験から始まり、計算科学でありますとか機械学習のような、より工学的な分野で発展してきたものになります。スタートは心理学なのですが、工学部などで発展してきています。強化学習は、長期的に見て報酬を最大化するために、環境の状態に対して特定の行動を学習するエージェント、あまり使わない言葉ですが、エージェントが学習する上での計算論的な枠組みになります。ここで重要なのは、書いてあることはだいたいオペラント条件づけの話と変わらないのですが、計算論的な枠組みを使うということが重要になってきます。こういった計算論を使うことで、ヒトで分かったことを機械の学習に応用することもできるようになるということがあります。

強化学習

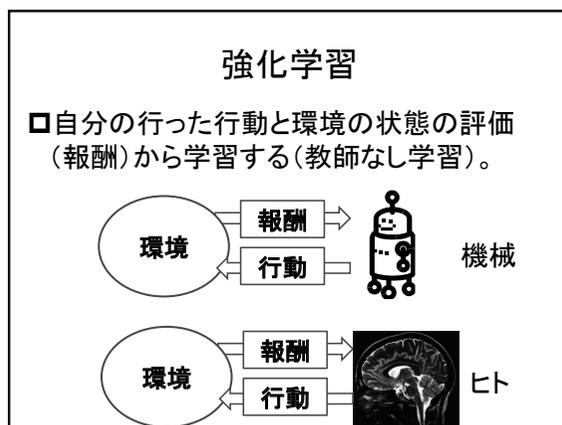
- 強化学習 (Sutton & Barto, "Reinforcement Learning", 2000)



オペラント条件づけ 機械学習: 計算科学

- 「長期的にみて報酬を最大化するために、環境の状態に対して、特定の行動を学習するエージェントについての計算論的な枠組み」

機械の学習では、身体をもった機械があつて何かしら環境に働きかけを行います。何かアクションをしたときに、報酬が出てくるみたいな、何か環境の変化が起こります。そういった変化をもとにして機械は学習します。これが強化学習の基本的な枠組みになります。ただし、これは我々ヒトがやっていることも一緒になります。我々も何か環境に対して行動を行って、そのフィードバックによって学習を進めていきます。そういった意味で、我々が脳内で行っていることは、機械が行っていることから類推することができると思います。つまり、我々の考えていることも機械が行っているのと同じようなやり方でなされており、ある程度シミュレーションできるのではないかと考えます。



もしかすると苦手という方もいらっしゃるかと思いますが、これから説明で数式が出てきます。強化学習の1つにQ learningモデルというものがございませう。ほかにもいろいろなモデルがあるのですが、比較的シンプルなモデルとして、このQ learningモデルが使われることが多いです。このQ learningモデルについて簡単にご説明していきます。数式の細かい部分はお説明しませんが、このQ learningモデルというのは環境の状態と行動を合わせた評価値を更新することで学習していくモデルになります。何のことだろうかという話ですが、ある環境の状態である行動をしたときにどれくらい報酬がもらえるかという、その期待値、そういったものをどんどん更新していくことで学習を進めていきます。この数式もそういったものになります。ここに、Q値というのがあるのですが、このQ値は、ある状態であることをしたときにどれくらいの見返りがあるかを表します。そのQ値をぐるぐる更新していきます。ここに、数式を言葉にしたものがありますが、{現在の行動Aの価値} というものが更新されていきます。その更新をするときに $a \times$ {予想した価値と実際に出た報酬との差} というものを計算していきます。ここが学習の上では重要になってきます。我々は学習するときにある程度こういったことをしたら、どれくらいの報酬がもらえるだろうという予期をもっています。そういった予期と実際に得られた報酬との差に基づいて、この選択肢はよかったな、これは悪かったなということを判断していきます。そういった実際に予想したものと実際にもらった報酬との差を、Temporal difference誤差といいます。そういった誤差に基づいてQ値を更新していきます。更新するときなのですが、この a が

重要になってきます。この掛け合わせた a によって、更新されるスピードが調節されます。例えば、TD誤差があったときに a が1ですとその誤差がそのままQ値の更新に反映されます。ただし a がゼロですと更新はゼロになりますので、Q値の更新はなされません。いわゆる学習が進まないという状況になります。この a というのは学習率と呼ばれ、学習のスピードを制御するものです。大きければ大きいほど学習が早く進みます。

もう1点、言葉だけの説明ですが、ここに γ というものがあります。これについてですが、これはどれくらい先の報酬を見通せるかどうかを制御するものになります。この γ のあとに、ごちゃごちゃと書いてありますが、これはその後の状態のことについて記述しております。 γ がゼロですと、その後の見通しがないということになりますし、 γ が1ですと先のことをちゃんと考えて判断するというようになります。ここで覚えていただきたいことは、この2点だけになります。 a というのはどれくらい学習が反映されるかということを規定しています。 γ というのはどれくらい先のことを見通せているかということを規定しています。

強化学習

□Qlearningモデル(Watkins, *Machine Learning*, 1989):
 状態と行動を合わせた評価値(Q値)を
 更新することで学習するモデル

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha \left(r_{t+1} + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t) \right)$$

現在の行動
Aの価値(Q_A)

←

現在の行動
Aの価値(Q_A)

+

$\alpha \times$

予想した価値と実際得た報酬
 との差(Temporal difference
 差: TD誤差)

α : 学習率(大きいほど学習が早い)
 γ : 減衰係数(大きいほど先を見通した行動が出来る)

強化学習では、複数の選択肢があるときに最も価値の高い行動を選択するようにモデル化しています。例えば、行動Aと行動Bがあったときに、Aのほうが価値が高ければそちらを選ぶようになります。これがそのモデルの式になりますが、重要なのはここだけになります。これは学習のパラメータの中での温度というものになります。この温度の高さ、低さによって、我々がどれくらい価値に基づいた意思決定をするかということが決まってきます。この温度というものが高くなると、私たちがそれまで学習した内容にかかわらず、よりランダムな反応をしてしまいます。温度が高いとランダムな反応をしてしまう。温度は τ で表せられます。 τ が高いとランダムな選択をしてしまう。逆に τ が低いとAとBの差により忠実に反応していくということがいわれています。いろいろと数式やら何やら出てきてわかりにくくなって来たと思いますので、簡単に説明をまとめていきます。

こういった話をまとめていくときに非常に有用な仮説として、銅谷仮説というものがあります。沖縄科学技術大学院大学におられる銅谷賢治先生が出した仮説になります。これは、これまでご説明してきました強化学習における学習のパラメータと神経修飾物質との関連性についてモデル

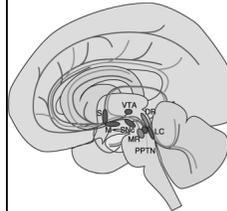
強化学習

- 最も価値の高い行動を選択することをモデル化
- 学習パラメタ(τ : 温度)によって、価値の差に敏感に反応する程度が制御される。
- 温度高いとランダムな選択をする

$$P(a_i | s) = \frac{\exp\left[\frac{Q(s, a_i)}{\tau}\right]}{\sum_{j=1}^m \exp\left[\frac{Q(s, a_j)}{\tau}\right]}$$

神経修飾物質と学習パラメタ

- 強化学習における学習パラメタと神経修飾物質に関する仮説(Doya, *Neural networks*, 2002; Doya, *Nature neuroscience*, 2008)



ドーパミン: TD誤差 (報酬予期と実際の報酬との誤差)
アセチルコリン: α (学習速度)
セロトニン: γ (減衰係数)
ノルアドレナリン: τ (温度)

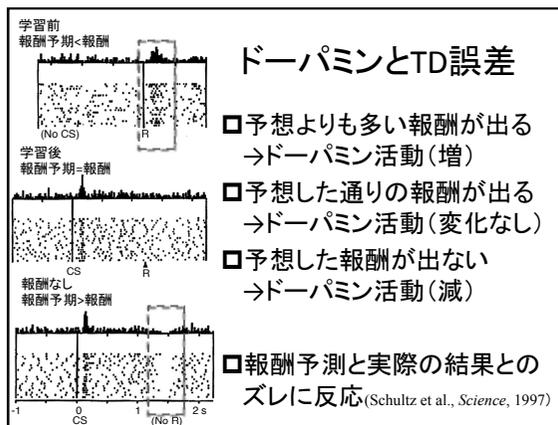
化したものになります。

このモデルは、最初に銅谷先生によって、2002年に提案されました。その後2008年に、この仮説について銅谷先生がレビューを行い、ある程度この仮説を支持するエビデンスが固まっています。銅谷先生の仮説では、ドーパミン、アセチルコリン、セロトニン、ノルアドレナリンという脳内の神経修飾物質というものが挙げられています。それらの神経修飾物質が、先ほど出てきた学習のパラメータに関係しているのではないかという仮説になります。1つ目はドーパミンです。これはTD誤差に関係してくるであろうと仮説が立てられています。つまり、報酬がこれくらいもらえるだろうという予期と、実際に報酬がもらえたときの、その誤差にドーパミンが関与しているのではないかという仮説です。2つ目はアセチルコリンです。これは学習の速度に関するであろうと仮説が立てられています。いわゆるTD誤差をどれくらい価値の更新に反映するかです。つまり、これくらい自分の予期と違ったよということをどれくらい次の行動に反映するか、これがアセチルコリンになります。3つ目がセロトニンです。これはどれくらい先の見通しをもっているかに関係するであろうと仮説が立てられています。最後は、ノルアドレナリンです。これはノルアドレナリンが選択における温度の制御に関係しているであろうという仮説です。つまり、ノルアドレナリンは、選択肢それぞれの価値に従って、意思決定することに関わるという仮説です。この銅谷仮説に基づいて考えていくと、うつ病はどういうふうに考えることができるのでしょうか？

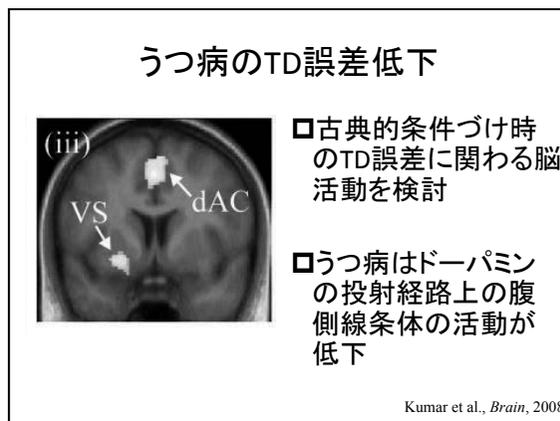
最初にドーパミンの話からしていきたいと思います。うつ病とドーパミン・TD誤差という研究はいくつかなされていますが、うつ病ではドーパミンが低下するというこれはこれまでの研究でいわれています。セロトニンのほうが治療上、ターゲットになっていたりしますが、ドーパミンも下がっていきます。そういったことからおそらくTD誤差も下がるのではないかということを予想することができます。

その前にドーパミンとTD誤差について少しだけ整理してみます。ドーパミンがTD誤差と関連するというのは、Schultzらの一連の研究からわかってきています。ちょっと煩雑になりますがご説明しますと、まず学習の初期段階、学習がスタートしたときは何かCSがあったとしても報酬が出てきたときにドーパミンが反応します。何か報酬が出てきたので、ドーパミンが活動します。

そのあと学習が進んでいきますと、CSが出たときにドーパミンがいちばん反応することがわかっています。何か報酬がもらえるらしいという期待が生じたときに、ドーパミンが反応しています。そして、実際に報酬がもらえたときは最初に比べてそんなに反応しなくなっています。最後ですが、ある程度学習して、CSの段階で、「あ、何か報酬がもらえるかもしれない」という期待を待っているときに、実際に報酬がなかったらドーパミンの活動はぐっと少なくなります。ここからわかることですが、もともとドーパミンというのは何か報酬に関係するのだろうかということは言われているのですが、実際に報酬そのものに対して反応しているというよりは、自分の予想した報酬ともらう報酬との差、そういうTD誤差に対して反応しているということがわかってきています。自分の予想と実際の結果とのずれ、それに対してドーパミンは反応しています。



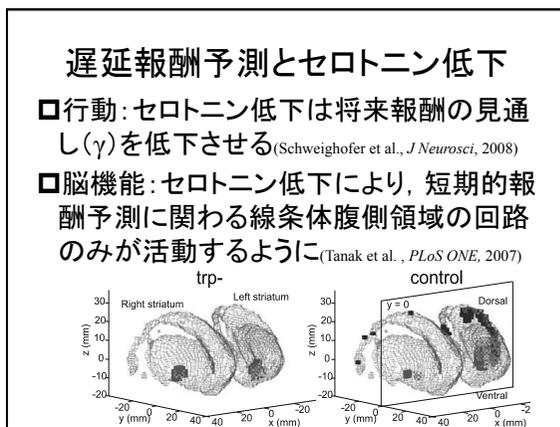
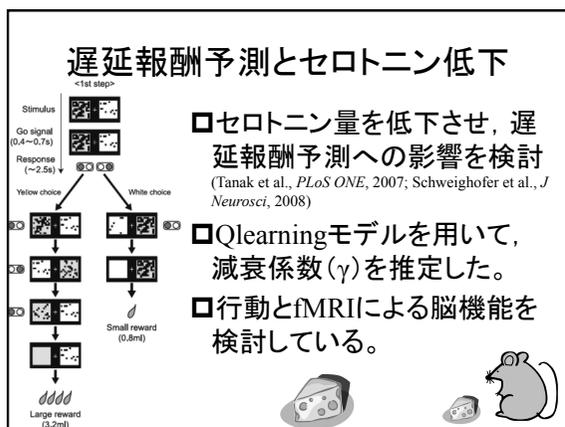
それでは、うつ病患者さんではどうでしょうか？ドーパミンは下がっているはずなので、TD誤差も低下するのではと思われます。これを、Kumarが研究しています。この研究では、古典的条件づけ時のTD誤差に関わる脳活動を調べています。ここでの古典的条件づけというのは非常に単純な実験になります。何か条件刺激となる、例えば図形で○とか△という図形が出てきたあとに、報酬として水が出てきます。喉をカラカラにした状態で水がちょろちょろと流れてきます。このときの脳活動を調べているのですが、うつ病患者さんというのはドーパミンの投射経路上の腹側



線条体の活動が低下しているということがわかりました。非常に間接的な結果ではありますが、この論文の中では一貫しない結果も出ています。ただ、部分的にうつ病とTD誤差の低下というのは関係しそうだということがわかりました。

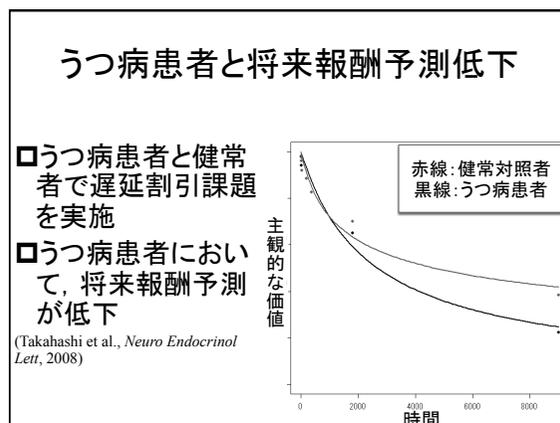
次にうつ病とセロトニン・ γ との関係になります。うつ病ではセロトニンの低下がずっと報告されてきています。セロトニンが低下して症状が出ていると思われるので、抗うつ薬でセロトニンを高めるということを治療ではやっています。そのため、セロトニンが関与すると考えられる減衰係数もたぶんうつ病では低下しているのではないかといったことが予想されます。そこでセロトニンが低下すると減衰係数もたぶん低下するだろうという仮説を立てて実験が行われています。

まず実際の患者さんで実験する前に、健康な方を対象にセロトニンを低下させる実験をしています。これは私が広島大学で研究を始める前に、行われていた実験になります。この実験では、セロトニンを人為的に低下させる手続きをしています。先ほどもご説明しましたが、セロトニンに合成される前の物質にトリプトファンがあります。それを人為的に減らすことでセロトニンを減らしています。その手続きを行った上で実験課題を行ってもらっています。セロトニンが、遅延した報酬の予測にどのように影響するのかということを調べています。実験課題はこのような課題になります。黄色の背景の上に黒いものが乗っています。白い背景の上にも黒いものが乗っています。そういった図形が2つ出てくるのですが、ボタンを押すごとに黒いものが除かれていきます。この黒いものが完全になくなったときに報酬がもらえるという課題になります。ぱっと見た感じ黄色のほうが黒いものがいっぱい乗っています。一方白いほうは黒いものが少ないです。白いほうは、1回、2回ボタンを押すだけで報酬がもらえます。一方、黄色のほうは、何回かボタンをポチポチと押さないと黒いものがなくなり、報酬が遅れてもらえます。ただし、もらえる報酬の量がちょっと違います。すぐもらえる白いほうを選択しますと報酬が少ないです。一方、黄色いほうは何回か押すので時間がかかりますが、何回か押したあとにたくさんの報酬がもらえます。ちょっとがんばってたくさんご褒美をもらうのか、すぐにももらえる少ないご褒美をもらうのか、そういったことを学習してもらいます。これは遅延報酬、遅れてもらえる報酬をどれくらい予測するかということを調べています。Q learningモデルなどを使って、行動データから減衰係数を推定したり、MRIによる脳機能の検討も行っています。



結果になります。まず、セロトニンを低下させることで将来の報酬の見通し、これは減衰係数の γ ですが、 γ が低下するということがわかりました。実際にセロトニンというのが将来報酬予測、つまりどれくらい先を見通せるかということに関与していることがわかりました。さらに脳機能を調べた結果ですが、セロトニン低下により、短期的な報酬予測にかかわる線条体の腹側がわのみが活動しています。これは少し説明が必要になりますが、普通にセロトニンがちゃんとある状態ですと、線条体の腹側がわから背側がわにかけてグラデーションを描くような活動が得られます。一方、セロトニンを欠乏させてしまいますと、背側のグラデーションがなくなってしまって、腹側のみが活動するということがわかりました。線条体の腹側というのはより短期的な報酬にかかわるといわれてきていますが、セロトニン低下によって、こういった腹側がわのみが活動するようになって、非常に短期的な視点しかもてなくなっているのではないかとされています。こういった実験的な検討によって、セロトニンが低下すると長期報酬予測が低下するということがわかってきています。

これを、うつ病患者さんで調べるとどうなるのかについては、北海道大学の高橋先生が行った研究があります。この研究では、セロトニンが低下したうつ病患者さんにおいて長期報酬予測が低下しているか検討しています。実験課題は、遅延割引課題を用いております。今日もらえる1,000円と明日もらえる1,050円、どちらがいいですかということをお聞きするような課題です。明日貰えるほうがいいという方もおられますし、それくらいなら今日貰ったほうが絶対いいという方もおられます。こういった反応をプロットしていきます。このグラフの横軸は時間で、縦軸は主観的な価値になります。ヒトは時間がたてばたつほど主観的な価値が下がっていくということがわかっています。例えば、今日もらえる1,000円と1年後の1,000円だとはるかに1年後の1,000円のほうの価値が下がってしまいます。これを実際にうつ病患者さんと健康な方で調べますと、赤線が健康な方で黒線がうつ病患者さんです。ぱっと見てわかりますが、うつ病患者さんのほうが、将来の主観的価値を非常に割り引いているということがわかります。うつ病患者さんは、将来もらえる報酬の価値をより小さく見積もっているということがこの結果からわかります。このことから、セロトニンの低下というのが将来を見通す能力といますか、減衰係数に影響しているということがわかってきます。



最後に、ノルアドレナリンの話になります。ノルアドレナリンと温度、つまり行動のランダムネスが関連するということが仮説として挙げられています。ノルアドレナリンというものは高くなればなるほど1つの選択肢に反応が非常に集中して、温度が下がります。一方、ノルアドレナリンが低くなってくると、こんどは選択がランダムになって、温度が上がります。うつ病に関して、ノルアドレナリンの低下というものがこれまで報告されています。そこで、ノルアドレナリンが低下しているうつ病患者さんでは、ノルアドレナリンが関与する温度というものが高くなっているのではないかと考えられています。そういったことを検討してみようと思います。

動物実験ではノルアドレナリンと温度、つまり τ との関連性が示唆されてきています。ただし、うつ病とノルアドレナリンや τ との関連を調べた研究はこれまでなされてきていませんでした。これから、私が修士過程のときに行った研究についてご紹介します。まず、最初のリサーチクエストジョンとしては、抑うつは報酬に基づいた意思決定に問題を引き起こすか？まず、これを最初に検討しました。このように、抑うつにおけるアンヘドニアを確認できるかということと、もう1つは、抑うつが強化学習時の温度にどういうふうに影響するのか、その2つの疑問を調べました。大学生の抑うつの高い群と低い群で比較しています。比較するときに、Frankの確率選択課題という実験課題を用いています。抑うつの高い群と低い群の選定にはCES-Dのカットオフポイントを用いています。スクリーニングのときと実験に来られたときでどちらの時点でもカットオフポイントよりも高い、もしくはどちらの時点でもカットオフポイントよりも低いという方を対象にしました。この研究は非常に苦労したのですが、やっと論文になりました。4年かかってしまったので非常に生みの苦しみが大きかった論文になります。

うつ病とノルアドレナリン・ τ (温度)

□動物実験では、ノルアドレナリンと τ (温度)との関連性が示唆されている
(Doya, *Nature neuroscience*, 2008)

□うつ病とノルアドレナリンや τ (温度)との関連をしらべた研究はない

研究の疑問と方法

- 疑問1:抑うつは、報酬に基づいた意志決定に問題を引き起こす？
- 疑問2:抑うつは強化学習時の τ (温度)に影響する？
- 大学生の抑うつの高群・低群を対象にFrankの確率選択課題(Frank et al., *Science*, 2004)を実施

*Center for Epidemiologic Studies Depression scale(CES-D)のカットオフ値利用

Kunisato et al., *J Behav Ther & Exp Psychiat*, in press

先ほどもお話しましたが用いた課題は、Frankの確率選択課題というものになります。この課題は学習セッションとテストセッション、2つのセッションからなります。説明がすごくややこしいのですが、2つで一組の刺激対が出てきます。例えばこんな形の図形であるとか、こんな形の図形が2つ出てきます。その2つの図形のうちのどちらかを選択していただくと、報酬である「○ 10円」や、罰である「× -10円」といったフィードバックがなされます。実験参加者の方はどちらかの図形を選択して、こっちのほうが得だな、こっちは罰が少ないなということを学習しても

らいます。刺激は必ずAとB, CとD, EとFと, 2つで一組になって出てきます。これらのペアは、報酬のフィードバックの確率がそれぞれ違います。AとBの対ですと, Aのほうは80%の確率で報酬がもらえます。逆にいいますと20%の確率で罰が与えられます。一方Bのほうは20%の確率で報酬, 80%の確率で罰が出てきます。パッと見てわかるのですが, Aのほうが非常に得な選択になります。Bが非常に罰がよく出てくる選択です。CとDだと少し差が小さくなります。Cは70%の確率で報酬が出てきます。Dは30%の確率で報酬が出てきます。EFの対ですとEのほうは60%, Fは40%の確率で報酬が出てきます。EとFですとあまり差が感じられないので学習するのが大変です。このAB, CD, EFの対で, Aがより選べるよう, Cがより選べるよう, Eがより選べるように学習を進めていただきます。

確率選択課題:学習セッション

650ms reward (+10 yen) or punishment (-10 yen)
 feedback
 350ms +
 750ms choice
 800ms +

- 刺激対(AB,CD,EF)の選択に報酬か罰がFBされる。金額を最大にする選択をする
- 報酬のFB確率
 - ① A(80%)vs B(20%)
 - ② C(70%)vs D(30%)
 - ③ E(60%)vs F(40%)

確率選択課題:テストセッション

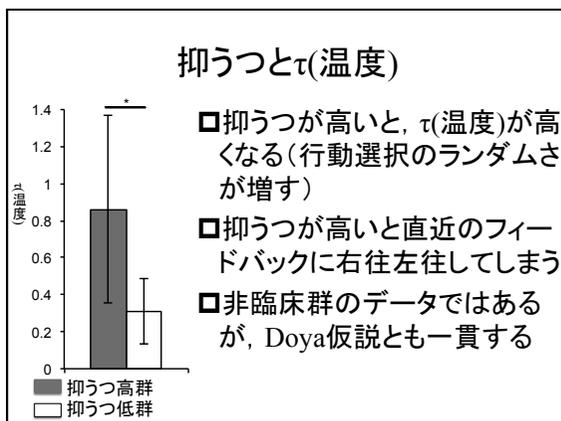
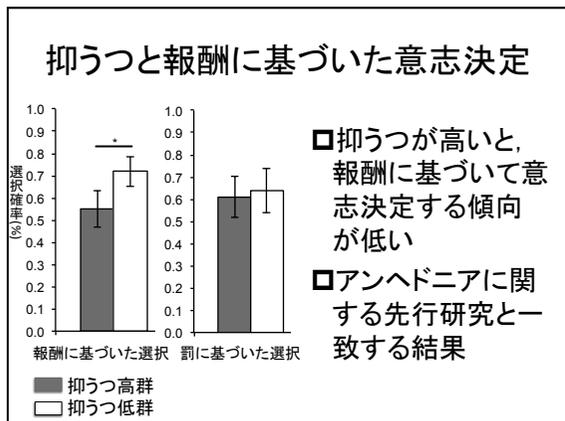
- 学習した刺激がランダムな組み合わせで呈示される(ex. AC,BD,AD...)※FBはなし
- A刺激対でAを選ぶ→高確率報酬刺激に基づいた意志決定の程度
- B刺激対でBを避ける→高確率罰刺激に基づいた意志決定の程度

750ms choice
 500ms +

ある程度学習が進んできたらテストセッションになります。テストセッションでは, 学習した刺激がランダムな組み合わせで出てきます。最初は必ずAとB, CとD, EとFという組み合わせで出てきますが, AとC, BとD, AとDといったように, ばらばらな組み合わせで出てきます。テストセッションは報酬や罰のフィードバックはありません。ここで重要になってくるのは, Aと何かが対になった刺激対と, Bと何かが対になった刺激対です。先ほどの刺激対を見ますと, AとBというのは非常に顕著な差があります。Aは非常に高い確率で報酬がもらえる刺激になります。一方Bは非常に高い確率で罰が与えられる刺激になります。ここで, Aとほかの刺激が対になったときに, Aをちゃんと選べるかどうかということにまず関心があります。Aとほかの刺激が対になったときに, Aをちゃんと選べるということは, 高い確率で報酬が出てくるAを選んでいることであり, 報酬に基づいた意思決定がちゃんとできていくことになります。一方, Bとほかの刺激が対になったときに, Bをちゃんと避けられたら, 高い確率で罰が出てくる刺激に基づいて意思決定をしているということになります。Aと対になった刺激対への反応と, Bと対になった刺激対への反応から, 報酬に基づいているのか, 罰に基づいているのかを確認することができます。これを抑うつの高い人と低い人で比較しました。これが結果になります。報酬に基づいた選択と罰に基づいた選択の程度を, それぞれ抑うつの高い群, 低い群で比較しました。こちらの罰に基づいた選択に関しては抑うつの高い人も低い人も変わりがありませんでした。一方, 報酬に基づいた選択では

抑うつの高い人において、そういった選択がより低いということがわかりました。このことから、抑うつが高いと報酬に基づいた意思決定する傾向が低くなるのではないかとこのようなことがわかります。この結果は、最初のほうでお話しましたPizaagalliなどの先行研究に一致するような結果といえます。つまり、抑うつが高い人において、アンヘドニア、つまり報酬への感受性低下が確認できました。

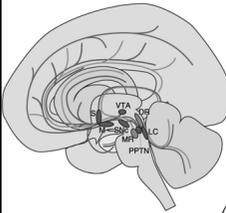
次に、こういった学習をしているときの学習のパラメータについて推定を行いました。これも抑うつが高い人と低い人で比較しています。結果として、温度つまり τ のみに有意な差が認められました。抑うつが高い人では、温度が高くなることがわかりました。これは抑うつが高くなると、温度、いわゆる行動の選択のランダムさがより増しているということを意味します。このことから、抑うつが高い人は、直近に何かフィードバックがあったときに、直近のことに右往左往してしまっ、これまで学習してきた履歴をちゃんと使えないのではないかとということが考えられます。あくまで非臨床群のデータではあるのですが、銅谷仮説とも一貫する内容かと思えます。



少しまとめになります。うつ病において銅谷仮説を検討する際に、ドーパミンやノルアドレナリンに関する研究は、まだまだ足りてないのですが、うつ病というのを強化学習理論から見るとどうなるのかというのをまとめました。これがまとめになります。最初ドーパミンが低くなるとTD誤差が低くなるというところから、うつ病患者さんは、予期した報酬と実際の報酬との誤差に鈍感になりやすくなっている。報酬の予期と実際の報酬との差といった随伴関係に鈍感になっているのではないかとということが考えられます。さらに、将来の報酬を低く見積もる傾向がある。先の報酬をあまり高く見積もらないということがあります。最後に、強化歴に基づいて、選択肢の価値にしたがった意思決定を行うというものができなくなっているのではないかとということが考えられます。このように強化学習理論から、これまでの研究を少しまとめていけるかなと考えています。しかし、なぜわざわざこのようなややこしい理論を使うのかということが気になる方もおられるかもしれません。それに関して、強化学習の枠組みを使う利点を述べたいと思います。

まず、強化学習理論を使う利点としましては、複雑な意思決定を強化学習の枠組みから比較的シンプルな数理モデルで整理することができます。いろいろな仮説があってそれをまとめるこ

強化学習理論からみたうつ病



↓ドーパミン: ↓TD誤差(報酬予期と実際の報酬との誤差)
 アセチルコリン: α(学習速度)
 ↓セロトニン: ↓γ(減衰係数)
 ↓ノルアドレナリン: ↑τ(温度)

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha \left(r_{t+1} + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t) \right)$$

□報酬予期と報酬との誤差に鈍感になり、将来の報酬を低く見積もり、選択肢の価値に従った意志決定が難しい

とがなかなか難しいのが現状としてありますので、それがシンプルな数理でまとめられて、上手く整理できるのかなと考えています。

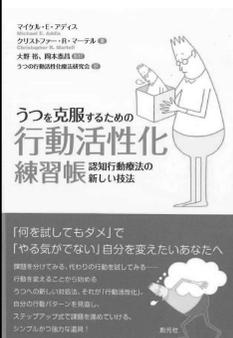
さらに学習時や意思決定時の内潜在的な学習パラメータの推定・評価をすることができます。なかなか学習や意思決定をしているときのパラメータというのは直接観察することができないのですが、この理論を使うことで推定することができます。さらにモデルベースな研究になりますので、モデルから新たな仮説や研究方法の創出も期待することができます。

さらに臨床的な応用を考えたときに、これまで、私はうつ病に対する行動活性化療法も研究してきました。行動活性化療法に非常に高い効果があるということがわかってきているのですが、理論がまだ仮説の域を出ない部分があり、この強化学習理論が行動活性化療法の理論的な基盤になるのではと考えています。強化学習理論をうまく使うことで行動活性化療法もしくは認知行動療法の理論的な基盤をさらに補強することができるのではないかと考えられます。さらに意思決定がしにくい状況において、それをうまく支援できるような治療的な工夫、そういったものをもたらすことができるのではないかといいことも期待しています。

強化学習理論の利点

- 複雑な意志決定を、強化学習の枠組みから数理的に整理できる
- 学習・意志決定過程の内潜在的な学習パラメータを推定・評価できる
- モデルベースな研究によって、モデルから新たな研究知見や研究方法の創出が期待できる

知見の臨床応用



- うつ病に対して、行動理論に基づく行動活性化療法の有効性が確認(岡島・園里他, 心理学評論, 2011)
- 強化学習は行動活性化療法などの認知行動療法の理論的な基盤を補強し、さらなる治療工夫をもたらす可能性がある

ただ数理モデルに関しては現在も進化しています。2012年に、『European Journal of

Neuroscience』誌上で、強化学習の特集が行われました。その特集では、Beyond simple reinforcement learningとして、単純な強化学習のモデルを超えて、より複雑なモデルも考えていく方向性が示されており、モデルの進化が検討されています。我々が臨床にいるときによく感じるものとしては、うつ病患者さんがもっている環境に対する信念、それが非常に学習に影響を与えているということがあります。そういった環境に対する信念を組み込んだ研究もできるのではないかと考えています。

数理モデルの変化

EJN EUROPEAN JOURNAL OF NEUROSCIENCE

European Journal of Neuroscience, Vol. 35, pp. 997-999, 2012 doi:10.1111/j.1460-9568.2012.02074.x

EDITORIAL
Beyond simple reinforcement learning: the computational neurobiology of reward-learning and valuation

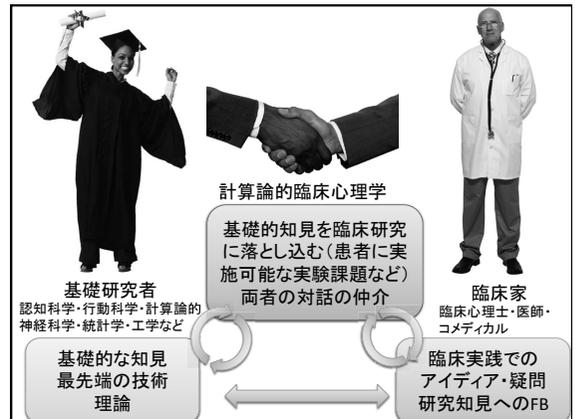
John P. O'Doherty

□2012年European Journal of Neuroscience誌で強化学習の特集が組まれた
□シンプルなモデルを超えて、環境に対する信念なども含めた複雑な臨床状況にも対応できるモデルを！

最後に、計算論的臨床心理学という新たな研究分野をここで提案します。こういった研究分野は、まだ存在しません。私自身がこういうものができるといいなと考えている内容になります。簡単にいうと、計算論の枠組みから臨床心理学的な問題に取り組む学問分野になります。例えば、今日お話しした強化学習理論でありますとか、グラフ理論、ゲーム理論などを臨床上の問題に適用するような研究分野です。こういった試みというのは精神医学で行われはじめているのですが、計算論的な基礎研究と臨床実践との乖離が非常に大きいと感じています。何とか基礎と臨床をうまくつないでいくことが、計算論的臨床心理学の使命と考えています。具体的には、認知科学、行動科学や、統計学などの基礎研究者の方がおられます。そういった基礎的な知見や最新の技術、理論というものと、臨床家の方の実践におけるアイデアや疑問とをうまくつなぐことができると良

計算論的臨床心理学 (Computational Clinical Psychology)

- 計算論の枠組みから臨床心理学的問題に取り組む学問(グラフ理論, 強化学習理論, ゲーム理論などを臨床へ)
- 精神医学への計算論の応用が始まったが、基礎と臨床実践との乖離は大きい
- 基礎と臨床をつなぐことが計算論的臨床心理学の使命



いかと思います。基礎的な知見をどうにか臨床の現場に落とし込めるような研究を行ったり、もしくは臨床家の方のアイデアを基礎研究の方にフィードバックしてつないでいくことが考えられます。そういった仲介の仕事が、計算論的臨床心理学においてできると良いのではないかと思います。

今回の発表の最初に、うつ病患者さんの意思決定は計算可能かとお質問したのですが、現段階ではイエスとは言い難いのですが、これがイエスとなるように研究を進めていくことと、少しでもいま苦しんでおられるうつ病患者さんに役に立つ研究を行っていきたいと思います。

以上になります。

(大久保) どうもありがとうございました。簡単に1点、2点確認のご質問があればお受けいたしますが、いかがでしょうか。

(質問者) セロトニンというものと、ノルアドレナリンの結果についてですが、セロトニンというものは将来の報酬を、抑うつが高いと将来の報酬に基づいて、直近の報酬を選ぶと理解しています。ただノルアドレナリンの結果を見ると、選択にばらつきがあるという話だったと思うのですが、結果に若干矛盾があるのではないかと感じたのですが、そのあたりについて。

(国里) ご質問ありがとうございます。少し内容が非常にざっくりしたものになってしまって、少しこちらの説明がわかりにくかったかと思います。抑うつが高いと将来の報酬に基づいて決定ができなくなります。この時のセロトニンとノルアドレナリンの違いが不明確だったかと思います。セロトニンは先の見通しにかかわるというのでいいと思うのですが、ノルアドレナリンに関していうと、そういった先のことというより、これまでの学習歴をちゃんと反映できているかというものになります。ノルアドレナリンが低いと選択がランダムになる、これまで学習してきた内容をうまく活用できていないことが、この結果からいえることになります。

(大久保) よろしいでしょうか。それでは国里先生、ありがとうございました。