

## 監訳者解説

本書は、2021年に刊行されたラッセル・ポルドラックの『*Hard to Break: Why Our Brains Make Habits Stick*』の翻訳である。なぜ悪い習慣を断ち切ることが難しいのか、また、どのようにすれば断ち切ることができるのかというテーマを扱っており、タイトルも一見、自己啓発本やビジネス書のような趣である。しかし、本書を読んでも習慣を断ち切るための「魔法のような解決策」が与えられるわけではない。神経科学と心理学の歴史を繙きながら、学習や記憶、意思決定、自制心等についての最新の知見が紹介され、行動変容のための方策が議論される。

人間は合理的でない。これは読書人ならおなじみの、心理学や行動経済学の一般書によく出てくるテーマである。しかし本書は、本能と理性、あるいは、後で述べるような「システム1・システム2」のような単純な二分法によって説明するのではなく、「第二の天性」とも言われる習慣の観点から、人間の行動を司る複雑な仕組みを丹念に繙く点に特徴がある。

しかしそれだけなら、従来の一般向け科学書にもありそうな話ではある。本書が従来の科学書と決定的に異なるのは、心理学を中心とした従来の科学研究の再現性が低いことが明らかにになった「再現性の危機」を踏まえた、新しい一般向け科学書のスタイルを提示している点にある。

科学者は、自分の研究分野の話を面白く紹介したい。読者は、驚きの事実を知ったり、複雑な現象

を明快な原理で統一的に理解することで、知的快楽を得たい。しかし科学者は、つい話を盛ったり、話のわかりやすさを優先して事実を恣意的に選択（「チェリーピック」）したりしがちだ。科学者も、科学者である前に一人の人間である。自分の研究や研究分野の重要性をアピールすることは研究費やポジションの保証につながる。虚栄心もあるだろう。

一般書である以上、取り上げる研究の選択や説明の単純化は不可避であるが、本書は、エビデンスを吟味するためのリテラシーを提供しながら近年の知見を整理している。「ポスト再現性の危機」のナラティブの実験とも言えよう。神経科学の研究者であるとはいえ、学習や意思決定について専門外である私が監訳を引き受けたのは、本書のこの特質に心を動かされたからである。

## 著者について

本書のテーマを議論する前に、著者について触れておきたい。ラッセル・ポルドラックは、スタンフォード大学の心理学者で、ヒトの脳イメージングなどを用いて脳と心の関係を研究している。この分野は「認知神経科学」と呼ばれ、一般の人が抱く「脳科学」のイメージと一致する分野と言えるだろう。スタニスラス・ドゥアン、デイヴィッド・イーグルマン、リサ・フェルドマン・バレットなど、この分野の研究者による一般向け書籍が多数出版・邦訳されている。ポルドラックは、本書で紹介されている記憶や意思決定についての研究でよく知られているが、「オーブンサイエンス運動」のリーダーとしての顔ももっている。再現性が高く透明性のある研究を行うためのツールや知識を研究者に提供し、科学のあり方を改善するための活動に早くからコミットしてきた。

ポルドラックの名が専門分野を超えて広く知られることになったきっかけは、2007年11月に

『ニューヨーク・タイムズ』に掲載された神経科学者によるコラムをめぐる騒動である。「ミラー・ニューロン」の研究で知られるマルコ・イアコポーニはそのコラムで、当時の次期アメリカ大統領選の主要候補者の写真やビデオを被験者に見せたときのMRI画像をもとに「前帯状皮質の活性化はヒラリー・クリントンに対する複雑な感情の証拠であり、一方、扁桃体の活性化は共和党の候補者ミット・ロムニーに対する有権者の不安の証拠である」などと政治評論もどきのことをやったのである。

これの何が問題なのか。たとえば、扁桃体は、不安なときに活動を示すことがあるが、強い匂いでも活動は上がる。扁桃体の活動が見られたということだけで、有権者の心理状態を同定できるわけではない。

このコラムは認知神経科学者の間で批判的となった。ポルドラックはすぐさまオープンレターを執筆し、16人の認知神経科学者が署名した。そのオープンレターは三日後に『ニューヨーク・タイムズ』に掲載された。ポルドラックは、イアコポーニらの脳画像の解釈は誤った「逆推論」だと批判した。なぜ「逆」かというと、多くの脳イメージング研究では、実験者が操作する刺激や課題（独立変数）の結果として脳がどのように反応するか（従属変数）を調べる実験デザインが想定されているのにも関わらず、イアコポーニらはそれは逆の向きに解釈していたからだ。しかし、（現在でも多くの人が誤解しているが）脳活動から心理状態を推定すること自体が誤っているわけではない。このことが私の研究とも関わってくる。

2005年に発表した私の論文で、人が目で見てもわからない微細な脳活動パターンから、機械学習を使って具体的な心理状態を解読する方法を提案した。現在「ブレイン・デコーディング」と呼ばれる手法で、2007年当時、認知神経科学の分野で注目され始めていた。ポルドラックも私の方法

を「妥当な逆推論」として、その後の意見論文等でも積極的に紹介してくれた。このコラムの騒動を含め、脳イメージングによる「読心術」とその諸問題については、ポルドラックの原著『新しい読心術 (The New Mind Readers: What Neuroimaging Can and Cannot Reveal about Our Thoughts)』（未邦訳）で詳しく議論されている。

ちなみに、私のアプローチでは、脳と心の関係を「コードするものとコードされるもの」としてとらえ、情報としての変換可能性に注目する。ここでは特定の因果の向きを想定するわけではない。したがって、「逆」とみなすこと自体が、特定の研究デザインに囚われたミスリーディングなラベルづけだと思った。今あらためて考えてみると、イアコポーニのコラムの問題は、推論の向きというよりは、脳を見れば心の背後にある要因やメカニズムが手にとるようになるかと考えがちな本質主義的な発想ではないだろうか。一般の人がもつそのような思考バイアスに便乗して、「脳科学者」がもっともらしいことを言っている光景は日本でもよく見られる。

その後ポルドラックは、認知神経科学研究の信頼性を高めるためのさまざまな取り組みを実践し、オープンサイエンスのリーダーと見なされるようになった。私の印象に残っているのは、研究方法の厳密性や透明性の基準を議論した彼のブログ記事に対して「Student」というハンドルネームでコメントした学生とのやりとりである。Studentは、厳密な方法論と透明性が、強い科学を構築するための重要な基盤であるという主張に同意するものの、ポルドラック自身の過去の研究に現在の基準を満たしていないものがあること、また、それにも関わらず新しく参入してきた若手研究者に高いハードルを課していることについて批判を展開した。同様の世代間のコンフリクトは、現代社会のさまざまな場面で見られる。ポルドラックはこのコメントを無視することなく、新たなエントリーを立て誠実

に議論している。参考文献でブログ記事へのリンクを示しているので興味のある方はご一読いただきたい。

このように、ポルドラックは、優れた研究業績をもつだけでなく、研究コミュニティを健全化するための行動力、オープンなマインド、誠実性をもった研究者である。私が学生の頃、研究を始めて驚いたことの一つは、研究者も平気でウソをつくことである。「1000円のを1000円で売ってどうするんだ」と真顔で指導する人もいた。著名な研究者を間近に見ながら研究する機会も学生のころからあったが、信頼できるシニア研究者に出会えることは少なかった。そのような中、私がアカデミアを去らずに研究を続けてこれたのは、ポルドラックや彼と同じ問題意識をもって研究コミュニティを改革する研究者が同世代にいたことが大きい。

### 本書のテーマとその背景

心のはたらきを議論する際、不確実な環境の中で適切に意思決定して行動しなければならない状況が想定されることが多い。しかし一方で、物理法則に従う世界は「それほど変化しない」のも事実である。世界の安定した側面には自動的に対処し、長期的な目標を常に意識的に考えたり、目下の欲求に左右されたりせずに、適切な行動ができるようにすることが習慣の機能と考えられる。

本書は、大脳基底核と大脳皮質のループ構造や強化学習など、習慣的行動を説明する切り口を提示しながら多くの研究知見を整理して紹介する。しかし、単純な図式でわかった気にさせるものではない。習慣的行動を通して脳と行動を統一的に理解することよりも、神経科学や心理学の知見が実際に行動変容を可能にするのかに重点を置く。

本書の構成は第1章でわかりやすく示されており、全体の論旨も明確であるため、章ごとの要約などは行わないが、従来の脳と心に関する議論を参照しながら本書の背景を整理したい。

「本能と理性の闘争」は本当か？

人の行動や心を説明するのに、本能と理性の葛藤という枠組みが古くから用いられてきた。プラトンは、生存本能や情動と理性の闘争として心を描いた。ダーウィンも「人間は理性的思考が手なずけている内なる太古の野獣を宿す」と述べている。近年では、ダニエル・カーネマンの『ファスト&スロー』で、「システム1（高速な無意識的・直感的処理）・システム2（遅い熟慮的処理）」の二分法が有名になった。

しかし、このように心を本能的な部分と理性的な部分に分ける考え方について、現在見直しが進んでいる。「システム1・システム2」の元となった二重過程理論（本書では「二重システム理論」と表記）の提唱者であるエヴァンスとスタノビッチは、脳の具体的な構造と同一視されることを恐れて、「システム1・システム2」の使用を中止し、「タイプ1・タイプ2」という古い用語に戻すことを提案した。また、「タイプ1は必ずしも進化的に古いとされる脳領域と対応するとは限らない」「意識的な思考が必ずしも行動を制御しているわけではない」「タイプ1のプロセスが常に認知バイアスの原因となるわけではない」など注意を促し、単純な二分法に警鐘を鳴らしている。

第2章で紹介されるポール・マクリーンによるトカゲの日常的な行動の記述は印象的である。きわめてルーチン化された爬虫類の一日の行動パターンは、現代人の日常をも彷彿とさせる。習慣的行動に関わる大脳基底核という脳部位の役割を明らかにしたマクリーンの功績をポルドラックは高く評価

するが、マクレーンが自身の研究から導いた「三位一体脳」説は現在の神経科学では否定されている。「三位一体脳」は、ヒトの脳が、(1)本能的・習慣的行動に関わる爬虫類の脳（大脳基底核や脳幹）、(2)情動に関わる古い哺乳類の脳（辺縁系）、(3)理性を司る新しい哺乳類の脳（大脳新皮質）という三つの基本的構造を段階的に追加しながら進化したという説である（本文27ページ参照）。この説は、近年の研究により根拠を失っている。たとえば、鳥類でそれまで基底核と考えられていた領域にも、大脳新皮質に相当する領域が多く含まれていることがわかり、2004年に多くの領域の名称が変更された。魚類や両生類、爬虫類にも大脳に相当する領域は存在する。いわゆる「辺縁系」が情動に特化した部位ではないことも明らかになっている。

リサ・フェルドマン・バレットは、著書『バレット博士の脳科学教室 7・½章』で「本能と理性の闘争」や「三位一体脳」といった考え方を厳しく批判し、身体に不可欠なリソースの管理（「身体予算管理」という統一的観点から合理性や脳の仕組みを説明しようとしている。同様に、変化する環境の中で恒常性（ホメオスタシス）を維持するために感覚入力を正確に予測するように脳が機能しているとする「予測符号化」や「自由エネルギー原理」も注目を集めている。このような新しい考え方を背景に、ニック・チェイターは『心はこうして創られる——「即興する脳」の心理学』で、意識的な心の背後に無意識的な心があると考え、「心の深さ」を否定し、「マインド・イズ・フラット」と喝破した。

#### 習慣的行動はいかに説明されるか

では、習慣は脳や心のはたらきの中でどのように位置づけられるだろうか。本書の前半の主役とな

る脳部位は大脳基底核である。正確に言うと、大脳基底核と大脳皮質の間のループ構造が鍵を握る。大脳皮質からの大脳基底核の線条体への入力には直接路と間接路に分かれて処理され、ドーパミンの影響を受けながら、大脳皮質に戻っていく。線条体の各部位は大脳皮質の異なる部位とつながっていて、学習にもなつて、前頭前野を含む「認知」皮質線条体ループから、運動野を含む「運動」皮質線条体ループへと活動がらせん状に移行していく。目標指向行動から習慣への移行は、このようなループの連続的な移行で説明される。

大脳皮質と大脳基底核が、「理性と本能」に対応するわけではない。一方が他方を支配しているわけでも、両者が「闘争」しているわけでもない。大脳皮質（前頭葉）から大脳基底核への「ハイパー直接路」が行動の抑制に関わることが知られているが、長期的な目標を達成するための自制心とはほぼ無関係らしい。自制心とは、衝動を抑える能力ではなく、衝動を抑える必要がある状況を回避する能力かもしれない。

習慣的行動は、自動的に実行される点で本能行動と似ているが、目標指向行動が学習によつて習慣化するし、大脳皮質の関与も大きい。原始的な能力とは程遠い、複雑な心の仕組みを反映したものである。「それほど変化しない世界」にうまく対応するための習慣システムは、変化や刺激に満ちた現代社会において依存症などの弊害を生み出している可能性がある。一方で、薬物乱用者が処方鎮痛剤を入手するために複雑な計画を立てるように、理性的とされる目標指向的行動が自分自身を苦しめることもある。第4章以降では、目標指向行動と習慣的行動は、現代のAIでも使われる強化学習の二つのタイプ（モデルベース型とモデルフリー型）と対比され、これらの強化学習をさらに階層的に統合する考え方が紹介される。

このように、習慣を通して脳と心のはたらきについて新たな視点を提示するのが本書の主題である。ただし本書には、もう一つ大切なテーマが含まれている。

### 再現性の危機

本書の裏テーマは、出版される科学研究の信頼性である。「ボックス」で、神経科学の実験手法とともに、研究の信頼性を評価するための統計学的概念について解説される。また、とくに後半の行動変容に関する多くの研究は、信頼性の度合いとともに紹介される。

2015年の大規模追試研究によると、心理学の代表的なジャーナルに掲載された研究のうち統計的有意性を再現できたのは36%にすぎない。人を対象とする研究だから結果がばらついて当然だという反応もあるが、同じような行動実験を用いる研究でも、経済学の行動実験の再現率は61%、実験哲学（哲学者による行動実験）では78%という報告もある。Pハッキング（実験手続きやデータを操作して統計検定のP値を低くする行為）やHARKing（Hypothesizing After the Results are Known 結果を見た後に仮説を作り、その仮説を事前仮説であるかのように装って統計検定し、論文に記載すること）など、研究分野で受け継がれる「疑わしい研究慣行（Questionable Research Practices, QRPs）」が影響していることは否定できないだろう。

しかし仮に適切に実験・解析を行ったとしても、統計的有意性だけでは再現性や真実性は保証されない。このことは、ボックス5・4で、ジョン・ヨアニディスの論文（「発表された研究結果のほとんどが誤りであるのはなぜか」）を紹介しながら解説される。

統計検定は、感染症の検査キットのようなものである。一つの仮説の統計的有意性・非有意性が、

一人の感染症検査結果の陽性・陰性に対応すると考えることができる。統計検定の有意水準（5%など）は、感染していない人を誤って陽性と診断する偽陽性率に相当する。新型コロナウイルスの感染拡大が始まったとき、積極的にPCR検査を行うべきかについて議論があったことを思い出してほしい。感染拡大初期には全体の感染率が低いので、検査で陽性であっても、そのうち実際に感染している人の割合（陽性適中率）は低いということが指摘された。

同様に、有意だった仮説が実際に正しい割合（陽性適中率）は、感染率に相当する基準率（全仮説のうち正しい仮説の割合）、あるいは、仮説の正しさの事前見積もりに依存する。神経科学や心理学の研究から推定される基準率や統計的検出力から陽性適中率を計算すると、その値は30%以下となる。疑わしい研究慣行によるバイアスも考慮すると、「発表された研究結果のほとんどが誤り」というのは、研究分野によってはあながち誇張ではないかもしれない。

事前見積もりで可能性が低い、すなわち、意外性のある仮説の場合、統計的に有意であっても真である確率は低くなる。したがって、驚くべき仮説を証明するには、統計的有意性の基準をより厳しくしたり、独立なデータで再現性を確認するなど、エビデンスのより慎重な吟味が必要である。カール・セーガンが言ったように「途方もない主張には、途方もない証拠が求められる」のである。

統計検定は効果の大きさ（効果量）を示すものではない点にも注意が必要である。サンプルサイズ（データ数）が大きければ推定精度が上がるので、現実には意味がないようなわずかな効果量であってもゼロとの違いが検出され有意になってしまう。また、脳イメージング研究では、被験者のグループ平均について検定を行うことが多いが、グループ平均の脳マップがどの個人の脳マップとも似ていないこともよくある。どこにも存在しない平均脳に基づく「脳と心の理解」に意味があるだろうか。

健康問題や地球温暖化問題に対処するため、人の習慣を適切に変化させる方法へのニーズが高まっている。現時点で有効な介入法はきわめて限られているが、ポルドラックは、脳の理解に基づいて有効な方法が開発されることに期待を寄せる。しかし、脳に関する知見の少なからぬ部分が、偽陽性や、実質的に意味のない小さな差、どこにも存在しない平均脳などに基づいているとすれば、期待は裏切られることになる。実世界での活用というテストをパスできるように基礎研究の知見を頑健にしておくことができるかが試されている。

本書の翻訳は児島修氏が行い、神谷が専門用語や内容の確認を行った。児島氏の翻訳は完成度が高く、大きな変更は加えていない。原書に、網膜と活動電位についての一部不正確な記述があったので、ポルドラック氏に確認の上、修正した。教科書や論文を確認しながら作業を行ったが、とくに、学習や意思決定についての専門知識の不足による誤解があるかもしれないので、お気づきの点があれば指摘していただきたい。最後に、私に本書の監訳の機会を与えてくださったみず書房の武石良平氏に感謝したい。研究の信頼性に関するSNSでの私の発信を見てコンタクトしてくださったのだと思う。本書を通して、信頼性を高めるための改革とともに進展する脳研究のリアルな姿が読者に伝われば幸いである。

## 参考文献

- Marco Iacoboni, et al., "This Is Your Brain on Politics," *New York Times*, Nov. 11, 2007, <https://www.nytimes.com/2007/11/11/opinion/11freedman.html>.
- Adam Aron, et al., "Politics and the Brain," *New York Times*, Nov. 14, 2007, <https://www.nytimes.com/2007/11/14/opinion/web/14brain.html>.
- Poldrack, R. A. "The New Mind Readers: What Neuroimaging Can and Cannot Reveal about Our Thoughts". (Princeton University Press, 2018).
- Poldrack, R. Having my cake and eating it too? [russpoldrack.org](https://russpoldrack.org) (2016). <https://russpoldrack.blogspot.com/2016/07/having-my-cake-and-eating-it-too.html>.
- 篠塚一貴・清水透「比較神経科学からみた進化にまつわる誤解と解説」(『心理学ワールド』75号17-20ページ、2016)。
- リサ・フェルドマン・バレット『バレット博士の脳科学教室 7・½章』高橋洋訳、紀伊國屋書店、2021。
- Evans, J. St. B. T. & Stanovich, K. E. Dual-Process Theories of Higher Cognition: Advancing the Debate. *Perspect. Psychol. Sci.* 8, 223-241 (2013).
- ヤロブ・ホーヴィ『予測する心』佐藤亮司監訳、勁草書房、2021。
- ニック・チェイター『心はこうして創られる——「即興する脳」の心理学』高橋達二・長谷川珈訳、講談社、2022。
- クリス・チェインバース『心理学の7つの大罪——真の科学であるために私たちがすべきこと』大塚紳一

郎訳、みすず書房、2019。

Richie, S. *Science Fictions: How Fraud, Bias, Negligence, and Hype Undermine the Search for Truth* (Metropolitan Books, 2020).

Ioannidis, J. P. A. *Why Most Published Research Findings Are False*. *PLoS Medicine* 2, 6 (2005).

ステイブ・ピンカー『人はどこまで合理的か(上・下)』橋明美訳、草思社、2022。