



意識をめぐる新たな生物学的自然主義の可能性

小草, 泰
新川, 拓哉

(Citation)

科学基礎論研究, 51(1-2):115-135

(Issue Date)

2024

(Resource Type)

journal article

(Version)

Version of Record

(Rights)

本著作物の著作権は著者に帰属します。
クリエイティブ・コモンズ [表示 - 非営利 - 改変禁止 4.0 国際]ライセンス

(URL)

<https://hdl.handle.net/20.500.14094/0100492319>



意識をめぐる新たな生物学的自然主義の可能性

小 草 泰* 新 川 拓 哉**

New Biological Naturalism on Consciousness

Yasushi OGUSA* and Takuya NIIKAWA**

Abstract

This paper examines new biological naturalist theories of consciousness, focusing on Ginsburg and Jablonka, and Feinberg and Mallat as their representatives. In doing so, we compare and analyze the two theories from three perspectives: the signature, the ontology, and the function of consciousness. We point out several significant theoretical differences between them, especially that Ginsburg and Jablonka propose a novel naturalistic teleological understanding of the ontological status of consciousness, whereas Feinberg and Mallat's theory conforms with so-called type-B physicalism.

1. イントロダクション

意識とは何か。意識と物理的現象はいかに関係しているのか。どんな生物（や無生物）に意識があるのか。意識は何の役に立っているのか。意識をめぐるこうした根本的な問いは、多くの研究者の関心を惹きつけながらも、長い間科学的なアプローチでは扱えない謎とみなされてきた。

どのようにして意識ほど比類なきものが神経組織を刺激することの結果として生じるのか。これは、アラジンがランプをこすったときに魔人が出てくるのと同じくらい説明のつかないことだ。(Huxley 1866, p.210)

そのため、意識は多くの科学者にとって一種の「タブー」であり、もっぱら哲学者が（アームチェアに腰かけて）形而上学的な観点から取り組む研究対象とみなされてきた。だが、ここ 30 年ほどで、意識研究をめぐる状況は大きく変化した。現代の意識研究は哲学・生物学・神経科学・心理学といったさまざまな分野の学際的な協働により推進されており、「意識の科学」というラベルでさえ、もはや奇異なものではなくなっている。

本論文は、こうした動向のうち、意識を生物学的現象として捉え、進化論をはじめとする自然科学の豊富な知見に基づいて意識を研究しようとする試みの概要と、またそれに関わる争点を明らかにすることを目指すサーベイ論文である。生物学的観点に重きを置くこのような方針の出発点は、今から 30 年ほど前に哲学者のジョン・サールが提起した「生物学的自然主義 (biological naturalism)」に遡ることができる。サールは、二元論と物的一元論の対立という心身問題をめ

* 高崎経済大学

E-mail: yasushi0810@hotmail.com

** 神戸大学

E-mail: niitaku11@gmail.com

本研究は、国立研究開発法人日本医療研究開発機構 (AMED)「脳とこころの研究推進プログラム (精神・神経疾患メカニズム解明プロジェクト)」(JP21wm0425021)」の支援を受けて実施されたものである。また本論文を完成させるうえで、山田圭一氏から多大なサポートを受けたこと、そして匿名査読者二名のコメントが有益であったことを記し、これらの方々に深く感謝したい。

本著作物の著作権は著者に帰属します。クリエイティブ・コモンズの「表示-非営利-改変禁止 4.0 国際 (CC BY-NC-ND 4.0)」に基づいて配布されたオープンアクセスの著作物であるため、適切なクレジットを表示し、非営利目的で、改変を加えていなければ、誰でもダウンロード、再利用、複写、転載、または配布することができます。

ぐる旧来の枠組みを乗り越える道として、意識を、脳内の神経活動に因果的に依存しつつも、それに還元することのできない創発的な生物学的現象とみなす立場を提示したのである (Searle 1992)。ただし、この先駆的な試みは、生物学や進化論に関する具体的な知見に基づいて展開されていなかったこともあり、科学的ないし学際的な意識研究の動向において大きな流れを形成するには至らなかった。

だが近年、生物学の知見をもつ哲学者や、哲学的関心のある生物学者らにより、進化生物学やその他の自然科学の成果を積極的に取り込んだ意識の理論が提案されてきた。彼らは、意識にとって本質的なのは、情報や機能といった非生物学的システムにも実装できるものではなく、神経生物学的なプロセスであると考え、サールの路線を受け継いでいる。しかし、彼らはそれにとどまらず、意識を生み出す高度な神経生物

学的プロセスは具体的にどのようなものであり、またそのようなプロセスはどのように進化してきたのかを探究したり、さらにどの生物に意識があるのかに関する具体的な仮説を提示したりといった試みを通じて、意識研究の一つの潮流を形成しつつある。このような動向に属する理論をまとめて「新たな生物学的自然主義」と呼ぶことにしよう。

本論文では、新たな生物学的自然主義のうち、特に著名な理論を二つ取り上げ、それらを比較分析しながら解説する。具体的には、神経生物学者のシモーナ・ギンズバーグと生物学の哲学者であるエヴァ・ヤブロンカの二人 (以下 G&J と表記) による理論 (Ginsburg and Jablonka 2019)、そして、精神医学と神経学の専門家であるトッド・ファインバーグと分子系統学および形態学を専門とするジョン・マラットの二人 (以下 F&M と表記) による理論 (Feinberg and Mallatt 2016b, 2018b) を取り上げる^{1,2}。彼ら彼女らはどちらも進化論的な観点から意識研究を推進する論者として大きな影響力をもつ。特に、彼ら彼女らの主著は日本語にも翻訳されており、本邦でも大きな学術的注目を集めていると言える。しかし、両理論の関係——類似点や相違点——はまだ十分に研究されていない (鈴木大地による両理論を比較する貴重な研究が最近出版されたが、その哲学的な差異については詳しく論じられていない (Suzuki, D. 2022))。また、それぞれの理論の批判的・応用的検討もまだ十分に進んでいない。以上のことを踏まえると、両理論を比較分析したうえで批判的検討を加えることには、大きな学術的価値があると考えられる。

議論の構成は以下である。まず 2 章において、両理論の共通点を確認したうえで、それぞれの理論を比較し相違を解明するための三つの視点／次元を設定する。続いて、3 章では G&J の理論の概要を、そして 4 章では F&M の理論の概要をまとめる。5 章では、G&J と、F&M それぞれの理論に関して暫定的な評価を与え、今後さらに議論すべき論点を指摘する。

¹ 以下で G&J の見解をまとめる際は、二人の主著である Ginsburg and Jablonka 2019 と、その後出版された論文 2020、さらに彼女らにジョナサン・バーチが加わった三人によって書かれた Birch et al. 2020 を適宜参照する。また F&M の見解をまとめるにあたっては、主著である Feinberg and Mallatt 2016b と 2018b、そして同時期に公表された論考である 2016a と 2018a を参照する。これら複数の著作において彼ら彼女らの論述は多少洗練されたり変化したりしている点もあるが、本論文ではそうした細部には立ち入らず、これらの著作を適宜参照しつつ、彼ら彼女らの見解のコアの部分をまとめることにする。また、邦訳のある著作 (Ginsburg and Jablonka 2019, Feinberg and Mallatt 2016b, 2018b) を引用する際は、大部分邦訳に依拠したが、ごく一部表記を改めたところもある。

なお、本論文で詳しく検討することはできなかったが、G&J は上記の文献の出版後、*Biosemiotecs* 誌の第 15 巻第 3 号において、これまでの研究を整理しさらに発展させたターゲット論文を発表しており (Ginsburg and Jablonka 2022)、同雑誌の同号とそれ以降の号には、その論文に対する多くのコメント論文も掲載されている。これらの論文において取り上げられているトピックは、「無制約連合学習 (UAL)」(本論文 3.2 を参照のこと) を意識の徴候とみなす G&J の見解の正否 (Sharov 2022, Olteanu 2022, Irwin 2023)、「意識の機能」や「行為主体性 (agency)」をめぐる G&J の議論の発展性 (Belardinelli & Pievani 2023, Noble 2022)、さらに G&J の理論とユクスキュルの「環世界」概念の関連など (Tønnessen 2022, Campbell 2022, Kull and Favareau 2022) 多岐にわたっている。これら一連の論文は、G&J の見解に対する注目度の高さを示すものであり、今後彼女らの見解を検討するにあたり参照されるべき文献となっている。(これらの文献情報については、査読者から有益なコメントを受けたことを記しておく。)

² これら二組による著作に加え、ピーター・ゴドフリー・スミスの意識に関する一連の著作も、新たな生物学的自然主義の代表的な成果とみなされている (Godfrey-Smith 2016, 2020)。本論文ではゴドフリー・スミスの見解を扱うことができなかったが、彼の見解もたいへん興味深いものであることを明記しておく。

2. 新たな生物学的自然主義者の間の共通点と違い

G&J と F&M はいずれも、意識を生命と連続的なものとして捉え、意識の問題に生物学的な観点からアプローチする。このような方針の土台となっているのは、「かつて自然科学の範囲を超えるものとされていた生命は、もはや神秘ではなくなり、生物学の枠内で扱いうるものとなった」という広く共有されている認識である。今日では、科学者の間で、生命を説明するために生氣論のような現代の自然科学になじまないものを持ち出す必要などないという合意が形成されているのである。そしてこのことを踏まえれば、意識を生命に連なるもの、あるいは生命現象の特殊な形態として捉えることで、意識も自然科学の範囲を超えるようなものを持ち出すことなく説明可能になるはずだ、と彼らは考えているのである。

このような大まかな方針に加えて、意識はどのように進化してきたのかという問いや、どの生物に意識があるのかという問いに対する具体的な仮説においても、彼らはかなりの程度一致している。彼らはともに、意識の進化はカンブリア紀にまで遡るという仮説を提唱している (Ginsburg and Jablonka 2019, ch.9, Feinberg and Mallatt 2018b, ch.7)。この時期に、初期の節足動物と脊椎動物（原始的な魚類）の間で捕食関係が生じ、効果的に食物を得ようとする節足動物と、首尾よく危険を察知し逃れようとする脊椎動物双方の感覚・運動能力が飛躍的に高まった。この「軍拡競争」の中で、初めて意識が生じたというのである。そして、彼らはこの仮説を擁護することを通じて、哺乳類と鳥類にだけ意識があるという従来の意識研究のパラダイムに挑戦し、脊椎動物一般と一部の節足動物や軟体動物にも意識があるとする、意識研究の新たなパラダイムを形成しつつある。

しかし、これらの一致点を超えて、彼らの見解の詳細に目を向けるとき、それぞれの見解の独自性と、それらの間の興味深い相違が浮かび上がってくる。一つには、様々な動物に意識があるかどうかを判定するための基準や目印——本論文ではそれらを「意識の徴候」と呼ぶことにしよう——は何であり、またそれらはどのような意味で意識の存在を示すものとなっているのかという点について、彼らの見解は重なり合いつつも重要な仕方で異なっている。また、意識が生物学的現象として自然界のうちに存在すると主張する点では一致していても、意識（ある動物）の存在論的身分に関

して、彼らの見解には興味深い相違がある。さらに、意識の「機能」は何なのか（あるいは、そもそもそのようなものがあるのかどうか）という点についても、検討に値する意見の相違がある。以下では、彼らの著作の中で論じられている豊富なトピックの中から、特にこれら三つ——意識の「徴候」「存在論」「機能」——に着目し、それぞれの見解の概要をまとめることにしよう。

3. ギンズバーグ&ヤブロンカの意識理論の概要

まずは G&J の見解を確認していくことにしよう。意識の徴候、存在論、機能に関する彼女らの見解の要点だけをあらかじめまとめると、以下のようになる。

1. 意識の徴候——彼女らの用語では、意識への「移行の目印 (transition marker)」——は、「無制約連合学習 (unlimited associative learning)」である。
2. 意識は（生命、理性と並ぶ）ひとつの「ありさま (mode of being)」である。
3. 意識に機能はない。

これらの見解の内実を理解するには、まず彼女らの理論の土台となっている、〈魂のアリストテレス的な三段階〉という考え方や、また生命の自然化に関する彼女らの理解を確認しておく必要がある。3.1 ではその点を説明していこう。

3.1 魂のアリストテレス的な三段階と、生命の自然化

G&J の解釈によれば、アリストテレスは『魂について』において、心に相当するものとしての「魂 (プシュケー)」を、あらゆる動植物に見出される生命の原理とみなし、諸々の生物は魂を備えることで（外部から設定されたのではない）内的な目的（テロス）を指向する存在になるのだと主張した (Ginsburg and Jablonka 2019, pp.8–9, 邦訳 (上) pp.11–12)。アリストテレスの考えでは、魂は入れ子状になった三段階を形成する。第一段階は〈栄養生殖的な魂〉であり、これは植物を含むあらゆる生物に備わっている。栄養生殖的な魂によって、すべての生物は自己を維持し保存するという目的を指向する存在となる。第二段階は〈感覚的な魂〉であり、アリストテレスはそれをあらゆる動物に備わるものとみなした。感覚的な魂が備わっ

ていることにより、動物は知覚に基づいて欲求という主観的な感じ (feelings) を満たすという新たな目的をもつことになる。第三段階は、〈理性的な魂〉であり、人間だけがこれをもつ。理性的な魂が備わっていることで、人間は、真・善・美といった象徴的な価値を目的としてもつことになる。

G&J は、この枠組みを踏襲し、さらにそのうちに進化論的観点を組み込むことによって、感覚的魂に相当する意識を自然主義的に理解することが可能となると考えている (Ginsburg and Jablonka 2019, pp.12–13, 邦訳 (上) pp.15–16)。この考えを支えているのは、生命 (への移行) はすでに、生物学・化学・物理学の関連分野における取り組みを通じて、自然主義的に理解可能なものとなっているという認識である。生命に対するその取り組みと類比的な手法を意識にも適用することで、意識の自然主義的な理解が可能になるというわけである。

そのような生命の捉え方を理解する上で注意しなければならないのは、生命が自然主義的に理解可能なものと広くみなされているからといって、生命の本性について広範な合意が存在するとは限らないという点である。生命の本性については、現在も科学者の間で意見は一致していない (Birch et al. 2020, p.2)。にもかかわらず、生命が自然科学の枠内で扱いうるものとなったのは、生命の十分条件を形成する一連の特徴のリストに関して、科学者の間で概ね意見が一致しているからである。具体的にそのリストに含まれるのは、〈境界の維持〉〈代謝〉〈安定性〉〈情報の貯蔵〉〈内的環境の制御〉〈成長〉〈生殖〉〈死〉であり、これらはまとめて生命の十分条件となる (Ginsburg and Jablonka 2019, pp.17–19, 邦訳 (上) pp.22–24)。これらの要因の一つひとつが生命にとっての必要条件であるか否かについては意見が分かれるかもしれない。しかし、これらがまとめて見出されるならば、そこに生命があるといえるという点は広く認められているのである。

このように生命の十分条件を形成する一連の特徴について同意があるなら、そこからさらに、そうした特徴をすべて備えたシステムにしかできないようなことがあるかと問うことができる。(もしあれば) そのような何かは、それをすることができるものには生命があると判断できるという意味で、生命への「移行の目印」としての役割を果たしてくれるであろう。G&J がそのような目印になると主張するのは、有機化学

者ガンティらによって提案された「無制約遺伝」である (Ginsburg and Jablonka 2019, p.27, 邦訳 p.35, Birch et al. 2020, pp.2–3)。無制約遺伝とは、当初のシステムからオープンエンドな仕方では形を変えながら、オープンエンドな長さの系統を形成していく遺伝のことである。このような遺伝が生じるためには、上記の特徴をすべて備えたシステムが必要であり、それゆえ、無制約遺伝が可能なものには生命があると言えるのである。

こうして生命への移行の目印が設定されれば、それにもとづいて、生命に関する科学的な研究のための具体的なプログラムを立ち上げることができるようになる、と G&J は言う (Birch et al. 2020, p.3)。科学者たちは今や、無制約遺伝の起源はどの年代まで遡るのかという問いや、無制約遺伝が可能な最小のシステムのモデルはどのようなものかという問い等々を探索することを通じて、科学的な手法によって生命の問題に取り組むことができるのである。こうして、生命の本性に関する議論は依然として続いているとしても、生命はすでに自然科学的な枠内で扱いうるものとなったのだ、と G&J は主張する (Ginsburg and Jablonka 2019, pp.28–29, 邦訳 (上) p.37)。

ここで、以上の生命の自然化の手順に加えて、そのように自然化された生命それ自身に関する G&J の (アリストテレス的な) 捉え方を確認しておくことも重要である。アリストテレスは、あらゆる生物は栄養生殖的な魂を備えることで内的な目的へと向かう存在となっていた。これにならって、G&J は、複雑な化学的プロセスやメカニズムからなるシステムは、生命を備えることで、単なる化学的な現象のレベルには見出されない、自己維持や生殖という目的をもつシステムへと組織化されると主張する (Ginsburg and Jablonka 2019, p.28, 邦訳 (上) p.36)。この見方の中では、生命は、あるシステムの中の一部やプロセスのようなものではなく、システム全体をまさに当該の種類のある目的へと向かう——システムとするものとみなされている。G&J の表現によれば、生命はシステム全体を目的指向的なものとする、根本的に新たな「ありさま」なのである (Ginsburg and Jablonka 2020, p.154)。

3.2 意識への移行の目印としての「無制約連合学習」

G&J は、生命の場合と同様の手順を踏むことで、意

表 1 意識の十分条件を形成する諸特徴

(Birch et al. 2020, pp.6-7 の「表 1」の引用。ただし元の表に含まれている文献情報は除いた。)

	神経生理学的／認知的メカニズム	行動上の徴候	現象学的な マニフェステーション
バインディング／統一	同期的で継時的なバインディングメカニズムを通じた情報の統合	複雑なパターンの弁別	対象の異なる特徴が結びついて一つの知覚像になること
大域的なアクセス可能性	感覚・評価・記憶システムからの入力マルチモーダルな統合	マルチモーダルなパターンの識別・評価・記憶	経験が単一の統合されたまとまりであるということ
柔軟な価値づけシステム	異なる刺激を評価・再評価し、異なるニーズを重みづけする統合的システム	逆転学習、二次条件づけ、対立を伴う状況での意思決定	快苦の感覚、感情、気分
選択的注意と除外	注意ネットワーク	習慣形成と自動化された行動、気が散る条件での学習へのダメージ	経験の要素には注意の焦点にあるものと背景に退いているものがある
志向性	身体と世界の階層的マッピング	目標の表象に基づいた目的志向的行動	経験が世界や主体自身の身体を表象すること
通時的な統合	ワーキングメモリー、転変する短期記憶	痕跡条件づけの能力、遅延見本合わせ課題(delayed match-to-sample)、ビデオシーケンスから学習する能力	「見かけの現在 (specious present)」、意識の「流れ」
身体化と行為主体性	運動アウトプットに対するトップダウンの認知的コントロールメカニズム	運動-感覚-運動ループに導かれた探索	身体化された、目的志向的な行為主体性の感覚
自他の判別	自己、身体、動機づけられた行為の神経モデル間の相互作用、それによって、空間中を動く動物自身の自己中心的表象が生み出される	自己-モデルが（たとえば、卒中後に）損傷を受けると、身体部位が自身のものでないと感じられるようになる	自身の経験に対する所有感、世界に対する「観点」としての経験の構造

識も自然科学的な探求の対象とすることができると主張する (Ginsburg and Jablonka 2019, p.34-35 邦訳 (上) p.45)。つまり、(1) 意識の十分条件を形成するものとして研究者の間で広く認められている一連の特徴をつきとめ、(2) それらの特徴をすべて備えたシステムでなければできない何か（すなわち、意識への移行の目印）を特定するという手順によって、たとえ意識の本性に関する見解は依然として分かれていても、科学的な意識研究のための具体的なプログラムを立ち上げることができる、というのである。

(1) に関しては、G&J は、これまでの広範な意識研究のサーベイに基づいて、以下の八つの特徴が意識の十分条件を形成するものとして広く認められていると述べる (Birch et al. 2020, pp.6-7)。すなわち、(1) バインディング／統一、(2) 大域的なアクセス可能性、(3) 柔軟な価値づけシステム、(4) 選択的注意と除外、(5) 志向性、(6) 通時的な統合、(7) 身体化と行為主体性、(8) 自他の判別である。これらの特徴はいずれも、神経生理学的／認知的メカニズム、行動上の徴候、現

象学的なマニフェステーションという異なる三つの観点から記述することができるとされる（表 1）。

G&J は、意識の神経相関研究やグローバル・ワークスペース理論など多種多様な意識研究を精査したうえで、それらのうちに上記の八つの特徴への言及が共通して見出されると主張する (Ginsburg and Jablonka 2019, p.142 邦訳 (上) p.186)。これらの特徴の一つ一つが意識にとって必要であるかどうかについては意見が分かれるかもしれない。しかし、これらの特徴が合わさって意識の十分条件を形成するという点については、その他の点では意見を異にする意識研究者の間でも、すでに合意が形成されているというのである。

ただし、ここで、上記の特徴がまとまって十分条件となるという主張に関して気を付けなければならない点がある。それは、ここで主張されている十分性が「法則的な十分性」（つまり、この現実世界で成り立っている自然法則の下での十分性）であるとされており、また生体システムという特定の種類の存在者に関して成り立つものとして限定されているという点で

ある (Birch et al. 2020, p.5)。つまり、上記の特徴を備えたものは、生物か非生物かを問わず、あらゆる可能世界で意識をもつのだと主張されているわけではないのである。したがって、たとえば、上記の特徴をすべて備えているが意識のないロボットや AI を作ることができるとしても、G&J の主張が覆されるわけではない。また、われわれ人間と物理的・機能的に同一である（それゆえ上記の特徴（の物理的・機能的側面）もすべて備えている）が意識をもたない、いわゆる「哲学的ゾンビ」が形而上学的に可能である（現実世界と法則的に異なる可能世界において存在する）としても、彼女らの主張にとっては関連のないこととされる。

以上を踏まえて、次に、これら八つの特徴をすべて備えたシステムでなければできないこととは何か、という問いへと G&J は向かう。そして「無制約連合学習」がそれに当たると彼女らは主張する (Ginsburg and Jablonka 2019, p.225, 邦訳 (上) p.291)。つまり、無制約連合学習こそが、彼女らの理論における意識の徴候——彼女ら自身の言葉では、意識への「移行の目印」——とされるのである。

無制約連合学習とは、次のような特徴をもつ連合学習である（以下は、Birch et al. 2020, p.8 をもとに、われわれがまとめたものである）。

- (1) 混成的な条件刺激——たとえば「黄色くて黒い、独特のにおいのするブンブン鳴る対象」のような、複数の知覚的特徴の自由な組み合わせ——を、無条件刺激（餌や電気ショックのような、特定の反射と結びつく刺激）と連合させることが可能である。
- (2) それらの条件刺激は当の動物にとって新奇な刺激——つまり、反射を引き起こすのでもなければ、あらかじめ無条件刺激と結びつけられているのでもないような刺激——であることが可能である。
- (3) 二階の条件づけ——無条件刺激と連合された条件刺激にさらに別の条件刺激を連合させること——も可能である。
- (4) 痕跡条件づけ (trace conditioning)、すなわち、条件刺激と無条件刺激との間に時間的な隔たりのある条件づけが可能である。
- (5) 柔軟で容易に組み替え可能な価値の連合——世界の変化に応じて、刺激のポジティブないしネガティブな価値がすばやく柔軟に組み替えられるこ

と——が可能である。

このような連合学習によって可能な連合のパターンは膨大なものとなり、その数は個体が一生の間に形成しうる範囲を大きく超え、実質的に無制限となる。この点で、無制約連合学習は、無制約遺伝と類比的な、オープンエンドな学習なのである。

G&J によれば、このような柔軟性の高い学習を行うには、意識の十分条件となる一連の特徴をすべて備えたシステムが必要であり、したがって、この学習は意識への移行の目印となるとされる。ただし、ここで注意しておかなければならないのは、無制約連合学習は、あくまで意識があることを示す——つまり「それが存在すれば意識がある」と言うことができる——「ポジティブな目印 (positive marker)」であって、意識がないことを示すためには使えない——「それが存在しなければ意識はない」とは言えない——と G&J が強調している点である (Ginsburg and Jablonka 2019, p.226, 邦訳 (上) p.292)。たとえば、新生児や無脳症の患者などは、無制約連合学習を行わないが、それでも無制約連合学習を支えるための進化を経た既存のシステムを備えている限りで、最低限の意識をもっている可能性があると言われ（ibid.）。さらには、無制約連合学習に依拠するアプローチは、この学習を行わない原初的な生物（や無生物）に意識がないことを含意するものではなく、ありとあらゆる生物に（さらには無生物にさえ）意識があるという説とも論理的には両立可能であるとされている (Birch et al. 2020, p.19)。

以上のように、ひとたび無制約連合学習が意識の目印として設定されると、（生命の場合と同様に）たとえ意識の本性に関しては意見が分かれていても、意識の問題に自然科学的な観点から取り組むための具体的なプログラムを立ち上げることができるようになる (Birch et al. 2020, p.10)。たとえば、「どの現存の動物種が無制約連合学習を行うことができるか」「無制約連合学習の起源はどこまで遡ることができるのか」「無制約連合学習のために十分な最小のシステムのモデルとはどのようなものなのか」といった問題に取り組むことにより、科学的な意識研究を推し進めることができるのだ³。

³ 実際 G&J 自身も、無制約連合学習という目印を手掛かりとして、「あらゆる脊椎動物、多くの節足動物、一部の頭足類」に意識があるという仮説を提起し (Ginsburg

3.3 ひとつの「ありさま」としての意識

先に見たように、G&J は、生命は自己を維持するという独自の内的な目的を伴うひとつの「ありさま」だと主張する。これと類比的に、意識もまた独自の内的な目的を伴うひとつのありさまであるとされる (Ginsburg and Jablonka 2019, pp.188–189, 邦訳 (上) pp.244–245)。この「ありさまとしての意識」というアイデアが、意識の存在論に関する彼女ら独特の見解を形作っている。

意識というありさまに伴う——意識が備わることによって動物がもつことになる——目的とは、快・不快の感じを伴う知覚的経験に基づいて、主観的に感じられる欲求を満たすというものである。意識をもつ動物は、あるものをしかじかのタイプのもの「として」、そしてまたしばしば快／不快といったプラスやマイナスの価値を有するもの「として」経験する。G&J は、意識ある動物に特徴的なそのような経験を「カテゴリー化する感覚状態 (categorizing sensory state: CSS)」と呼ぶ (Ginsburg and Jablonka 2019, p.375, 邦訳 (下) p.166)。たとえば、CSS をもつ雌鶏は卵を卵「として」、そしてまた抱くべきもの——抱くと快いもの——「として」見るのであり、それに伴って、卵を抱きたいという主観的な欲求を満たすことが、雌鶏にとっての目的となる。このような目的は、自己の維持という生命に伴う目的に（多くの場合寄与する一方で、しかし）必ずしも一致するとは限らない自律的な目的である (Ginsburg and Jablonka 2020, p.155)。一部の動物は、意識が備わることによって新たな目的の領域に生きることになったのである。

この目的論的な見解について特筆すべき点は、G&J が、意識というありさま（を備えた動物）の目的指向的な性格を還元不可能で実在的なものとみなしているという点である。彼女らによれば、意識を有する動物の行動（例：ヘビを見たカエルの跳躍）に十全な説明を与えるには、機械論的な説明（その行動を支えるメカニズムの説明）や、系統発生的な説明（当該のメカニズムの進化に関する説明）などに加えて、動物個体自身の目的や心的動機づけ（例：カエルの恐怖）に訴える目的論的な説明が不可欠である (Ginsburg and Jablonka 2020, p.153)。ここでは、意識を有する動物の目的指

向的な心的状態への言及は、他の何かへの言及によっては置き換えられない還元不可能なものとみなされている。彼女らによれば、目的指向的な状態は神経発火のような低次性質に対して「拘束をかけ (impose constraint)」、「因果的なトップダウンの効果」をもたらすものとして、説明上不可欠で還元不可能な役割を果たすのである (Ginsburg and Jablonka 2019, pp.162–167, 邦訳 (上) pp.209–215)。また G&J は、意識を備えた動物は単に目的に向かってふるまっている「かのよう」であるとか、あるいは、そのようにふるまうものとして記述することが実用的であると主張するような、反実在論的な見方をはっきりと退けている。彼女らは、目的指向的な生物の諸段階を進化論的観点から区別するデネットの枠組みを「アリストテレスの目的論的アプローチの拡張版」として評価しつつも、「生きものや感覚的生物に目的〔テロス〕があるかのように語るのはわかりやすさのためだ」とするデネットの考えに異議を唱えている。彼女らによれば、「選択に基づくシステムは本質的に (inherently) 目的論的」なのである (Ginsburg and Jablonka 2019, p.35, 邦訳 (上) p.46)。G&J の理論に見出されるこのような目的論的側面は、興味をそそるものであると同時に、議論を呼び起こすものでもあるだろう (5.2.1 を参照のこと)。

3.4 「意識の機能」という考え方はカテゴリー・ミステイクである

目的指向的なありさまとしての意識という以上のアイデアと切り離せないのは、「意識に機能があると考えるのはカテゴリー・ミステイクだ」という G&J 独特の主張である (Ginsburg and Jablonka 2019, pp.188–189, 邦訳 (上) pp.244–245)。彼女らによると、「機能 function」とは、生命や意識をもつシステム全体の目的指向的なふるまいのために、システム内の部分やプロセスが果たす役割である。この見方によれば、機能は、たとえば歯や心臓弁のようなシステムの部分や、またシステム内で生じるプロセスに帰属するものである。他方で、G&J によれば、生命や意識はシステム全体に帰される、当のシステムをまさに当の目的指向的なシステムとする「ありさま」である。それゆえ、生命や意識が機能をもつとするのはカテゴリー・ミステイクだと彼女らは主張する。生命や意識はそれ自身が何らかの機能を果たすのではなく、むしろシステム内の諸々の部分やプロセスがその実現に向けて

and Jablonka 2019, p.392, 邦訳 (下) p.187)、また無制約連合学習を行いうる最小のシステムのトイモデルを提示している (Ginsburg and Jablonka 2019, pp.360–368, 邦訳 (下) pp.148–158)。

機能を果たすような目的を設定するのである。それゆえ、「生命／意識の機能とは何か」と問うことは意味をなさないのだ。

4. ファインバーグ&マラットの意識理論の概要

次に、F&M の理論の概要を確認していくことにしよう。まず、意識の徴候、存在論、機能に関わる F&M の見解の要点だけをあらかじめまとめておくと、以下のようなになる。

1. 意識の徴候は、「外受容意識」に関しては、同型的で多感覚的な神経表象であり、「情感意識」に関しては、報酬と罰に基づく複雑なオペラント学習である。
2. 意識とその神経相関である「特殊な神経生物学の特性」の間には、ある種の認識的ギャップは存在するが、存在論的ギャップは存在しない。
3. 意識の機能は生存への寄与の観点から理解できる。

これらの主張をより具体的に理解するために、彼らの見解を見ていくことにしよう。

4.1 意識の四つの特性

F&M は、自分たちの説をサールの生物学的自然主義の発展版と捉え、「神経生物学の自然主義」と名づけている (Feinberg and Mallatt 2016b, p.14, 邦訳 p.18)。この理論は、(1)「外受容意識」(視覚や聴覚のような外的知覚に伴う意識)、(2)「内受容意識」(満腹感、吐き気、尿意など、体内——とりわけ内臓——からの刺激に関する意識)、(3)「情感意識」(ポジティブないしネガティブな感情価を伴う質的な感じ) という三つの種類の意識を、進化論的な観点から自然主義的に説明することを目指す (Feinberg and Mallatt 2018b, pp.11–21, 邦訳 pp.13–23)。F&M の著作では、主に外受容意識と情感意識に焦点が当てられており、内受容意識は両者の中間に位置するものとして中心的には取り上げられていないため、本論文でも外受容意識と情感意識に焦点を当てることにする。

F&M は、意識を謎めいたものにして以下の四つの特性を挙げ、満足のいく意識理論はそれらの特性を説明しなければならないとしている (以下は、Feinberg and Mallatt 2016b, pp.6–8, 邦訳 pp.7–10 をもとに、われわれがまとめたものである)。

参照性：意識は、脳活動によって生じているにも

かかわらず、外界の事物や身体の状態のような、脳の外部にある何かに「ついての」ものであり、脳内の現象として感じられることはない。

心的統一性：意識を生じさせているのは分割可能で不連続なニューロンの集まりであるにもかかわらず、意識は連続的で統一されている。

クオリア：意識は客観的な脳活動によって生じているにもかかわらず、主観的な質感 (クオリア) を伴っている。

心的因果：意識は心的現象でありながら、物質に因果的に働きかけ、人間やその他の動物の行動を引き起こす力をもつ。

彼らは、意識のこれらの特性を説明するには、生物学的特性を三つのレベルに区別し、各レベルに属する特性に言及することが不可欠だと考えている。そのため、まずはその生物学的特性の三区分について確認しておこう。

4.2 生物学的特性の三つのレベル

F&M は、意識 (がもつ上記の特性) を、意識を欠くものも含めた生物全般に見出される特徴の延長線上に位置するものとして説明しようとする。そうすることで、意識を一般的な生物学的現象と連続的なものとして理解しようというわけである。より具体的には、あらゆる生物に見いだされる、(1)「一般的な生物学的特性」、それに基づいて生じるより複雑な、(2)「ニューロン・反射・中核脳 [コア・ブレイン]」、さらにそれらに基づいて生じるより複雑な、(3)「特殊な神経生物学の特性」という三段階が区別され、意識は (3) のレベルにおいて成立するものとみなされる (Feinberg and Mallatt 2016b, ch.2, Feinberg and Mallatt 2018b, ch.6)。各レベルの生物学的特性の要点を確認しておこう。

(1) に分類される特徴の一つに、「身体化」がある。すなわち、どんな生物にも身体があり、それによって、外的環境から切り離された〈内部〉が備わっている。このことは、それ自体で意識をもたらしわけではないが、私的で一人称的な意識が成立するための土台となる。また器官 (たとえば内臓) は組織によって構成され、組織は細胞から構成され、細胞は細胞小器官から構成され... という「階層構造」も (1) に含まれる。F&M によれば、このような階層構造の中で低次階層と高次階層が相互作用することにより、新奇特性

が創発する (Feinberg and Mallatt 2018b, p.71, 邦訳 pp.75-77)。たとえば、体内の原子や分子——それら自身は生きているわけではない——とそれらが構成するより高次の階層との相互作用により、「生きている (alive)」という性質が創発するという具合である。さらに、適応的機能——自然選択に基づく機能——も、生物全般に見出される重要な特徴である。F&M によれば、意識もまた適応的で生存に有益な生物学的プロセスの一種として説明されることになる (Feinberg and Mallatt 2018b, pp.71-73, 邦訳 p.77)。

レベル (2) において付け加わるのは、多細胞身体や、その中で広範囲にすばやく情報を伝達するニューロン、またニューロンからなる反射弓、そして (感覚受容器が頭部に集中することで生じた) 単純な中核脳である。このレベルでは、ニューロン間のすばやい情報伝達によって、互いに離れた身体部位を同調して動かすことが可能になる。そして高度に発達した動物においては、ニューロン間の結合がきわめて広範に及ぶことで、複雑な反射や、(リズムカルな反復的運動を司る) 中枢パターン生成器に基づく基本運動プログラムが備わるようになる。さらに、中核脳をもつに至った動物においては、中核脳によってホメオスタシスが調整されるとともに、(刺激に対する鋭敏さに関わるものとしての) 覚醒・注意・動機づけが制御されるようになる。

続くレベル (3) こそこが、F&M の意識理論にとって最も重要である。このレベルでは、10 万個以上の多様なニューロンによってもたらされる神経的な複雑さや、多様な刺激を検知する精緻な感覚器など、いくつかの重要な特徴が付け加わる。そのなかでも F&M が最も詳しく説明しているのが、高度に発達した神経階層である (Feinberg and Mallatt 2018b, pp.80-83, 邦訳 pp.86-89)。たとえば視覚処理は、点や短い線のような刺激に反応する低次段階のニューロンと、手や特定の顔など高度に統合された刺激に反応する高次段階のニューロン、そして両者の中間に位置する段階のニューロンの間での (そしてまた各段階内部での)、複雑な情報伝達によって行なわれる。こうした神経階層においては、各階層に属するニューロンは物理的には離れて位置しており、それゆえ低次段階は高次段階に物理的には含まれていない (細胞が器官に物理的に含まれるのとは違って)。しかし、それぞれの段階に属する互いに離れたニューロンがシナプス結合によってすばやく情報を伝達しあうことで、階層間の双方向的

で再帰的な相互作用が生じる。それによって、異なる感覚様相をまたいだ統合的な処理も可能となり、また次の時点の感覚的インプットを予測することもできるようになる。

これらの神経階層は、外部環境と動物自身の身体に関する「同型的な」神経表象を形成することで「感覚イメージ (sensory image)」を生み出すものである、と F&M は主張する (Feinberg and Mallatt 2016b, p.33, 邦訳 p.39)⁴。「同型的な」神経表象とは、たとえば視覚的な情報処理において、視野の各位置の空間的秩序が低次段階から高次段階まで保存されることによって、視野の各位置の地図が形成される (また聴覚的な情報処理において、周波数にしたがった音の配置が保存されることで、音の周波数の配置の地図が形成される) という具合に、刺激 (外界) のあり方を地図的に表わす神経的表象のことである。F&M は、神経階層のはたらきにより形成される多感覚的な同型的表象こそが、外受容意識をもたらすものだと考える。「同型的な神経表象は物理的、生物学的実体であるが、意識をもつ生物は同型的な神経表象の地図を、地図で描き出された感覚的な心的イメージとして経験することができる」というのである (Feinberg and Mallatt 2016b, p.33, 邦訳 pp.39-40)。

加えて、神経階層のはたらきは情感意識にとっても不可欠なものだと F&M は主張する (Feinberg and Mallatt 2018b, p.51, 邦訳 p.59)。彼らによれば、情

⁴ 「感覚イメージ」という語からは、知覚者の心の目に映る〈心的な対象〉のようなもの——哲学者が「センス・データ」と呼び、多くの場合忌避するもの——が連想されるかもしれない。しかし、F&M がこの語によって意味しているのはそのようなものではないと思われる。むしろ、「イメージ (image)」という語を用いるときに彼らが念頭においているのは、たとえば視覚に関して言えば、精巧な光学的装置としての眼によって形成される——すなわち、心的ではなく光学的な意味での——「視覚イメージ」(ないし「視覚像」) であると推測される。このことは、Feinberg and Mallatt 2016b の 5 章で、しばしばこの意味での「視覚イメージ」(や、その他の感覚様相に属する「イメージ」) への言及がなされていることからうかがえる。F&M の考えは、洗練された感覚器官を備えた生物のうちにこの意味での「イメージ」が形成され (それによって脳内で同型的表象が形成され) るということが、とりもなおさず、(「一人称的な視点」から言えば) 生物が環境のありようを意識的に経験することなのだ、というものとして理解するべきであろう。(F&M が用いる「イメージ」という語のニュアンスについては、査読者のコメントから示唆を得た。)

感意識は、同型的な神経表象の形成によってではなく、刺激に対する感情価が情感系のニューロンによって符号化〔エンコード〕されることでもたらされる。すなわち、異なる種類のニューロンが、電気的な発火パターンと放出する化学物質の種類とによって、特定の正や負のシグナルを運ぶことにより、動物は感情価を伴う様々な経験をもつことになるのである (Feinberg and Mallatt 2018b, pp.49–50, 邦訳 pp.55–56)。こうした情感系のニューロンのはたらきは、外受容意識に対応するのとは異なる神経回路において行われるが、しかし情感系のニューロンのはたらきもまた、階層をなすニューロンの双方向的で再帰的な情報伝達によって成立するものなのである。

4.3 意識の徴候に関する F&M の見解

F&M は、意識をもたらし生物学的特性に関する以上の見方に基づいて、外受容意識と情感意識の徴候について仮説を提案している。まず、外受容意識に関しては、同型的な神経表象が意識の徴候だと主張される (Feinberg and Mallatt 2018b, p.28, 邦訳 p.32)。彼らは、ある動物に同型的な神経表象を生み出す仕組みが備わっているか否かを調べることによって、その動物に外受容意識があるかどうかを判定できると考えるのである。

また情感意識については、「大域的〔グローバル〕オペラント条件づけ」と呼ばれる複雑なオペラント学習の能力が、意識の徴候だとされる (Feinberg and Mallatt 2018b, p.43, 邦訳 p.49)⁵。大域的オペラント条件づけは、経験に基づいて、プラスの価値をもつ刺激に近づき、マイナスの価値をもつ刺激を避けることを学ぶオペラント条件づけの一種であるが、その中でも特に、全身を使った新奇な行動に関わる条件づけである。このような条件づけは、単に体の一部分の動きを変えとか、あらかじめ固定的にプログラムされた行動を切り替えるなどといったふるまいに関わる学習に比べて、より高度で柔軟性のある学習である。しかし、これらの条件づけは、G&J が提起した無制約連合学習ほど高度ではないとされている。たとえば、無制約連合学習は、複数の知覚的特徴の混合である混成的

な条件刺激に関わる学習であるのに対して、大域的オペラント条件づけにはそのような条件は課されない。F&M は、無制約連合学習は「最低限の意識 (minimal consciousness)」の徴候とみなすには少し高度すぎるとし、代替案として大域的オペラント条件づけを提示しているのだ。

4.4 意識の存在論に関する F&M の見解

次に、意識の存在論に関する F&M の見解を確認しよう。意識の存在論的身分に関する彼らの見解は、必ずしも十分明確には展開されていないようにみえる。しかし、われわれが理解した限りでは、彼らの見解はつまるところ、意識をめぐる哲学的な議論において「タイプ B 物理主義」(あるいは「アポステリオリな物理主義」)と呼ばれている立場に近いものだと考えられる⁶。この点について、4.4 節および 4.5 節において順を追って説明していこう。

F&M によれば、意識を謎めいたものとする四つの特性 (参照性・心的統一性・クオリア・心的因果) は、生物学的特性の三段階 (「一般的な生物学的特性」「ニューロン・反射・中核脳」「特殊な神経生物学的特性」) の上に成り立つものとして捉えることで、既知の科学法則で捉えられないようなものを何ら措定することなく説明できるとされる (Feinberg and Mallatt 2016b, pp.222–226, 邦訳 pp.244–249)。以下、それぞれの特性に対する彼らの説明をまとめよう。

(1) 意識の参照性は、一般的な生物学的特性のレベルにおいて、身体化により〈生物の内部と外部〉が区別されること、またニューロン・反射・中核脳のレベルにおいて、反射によって外的刺激が環境に対する自動的反応と結びつけられることを基礎とする。そして、これらの土台の上に、特殊な神経生物学的特性のレベルで環境や身体に関する詳しい同型的な神経表象が築かれることで、参照性が成立する。

(2) 意識の統一性は、特殊な神経生物学的特性のレベルで、多様な感覚情報が複雑な神経階層における双方向的な相互作用を通じて統合された表象へとまとめ上げられることで成立する。意識の統一性をもたらしそのようなはたらきは、より基本的なレベルの統合的な特性 (生命一般に見出される恒常性〔ホメオスタシス〕など) に起源をもつ。

(3) 心的因果は、複雑な神経構造において多様な刺

⁵ 情感意識の徴候が (神経的なものではなく) 行動的なものとされるのは、情感意識の神経相関には (外受容意識の場合のような) はっきりとした同型性が見いだされないからだとされている (Feinberg and Mallatt 2018b, p.41, 邦訳 p.47)。

⁶ このような解釈は、Suzuki, T. 2022 においてもなされている。

激が処理された結果、文脈に応じた行動が組み立てられ、行動に関する指令が下されるというプロセスによってもたらされる。意識と結びつくこのような因果的効力は、動き回り、餌を摂取し、排泄物を排出し、反射的に反応するといった形で、意識をもたない生物が環境へと因果的に働きかける力の延長線上にある。

(4) クオリアは、統一性や参照性など他の特性と切り離して単独で分析することは困難だが、やはり、基本的なレベルの生物学的特性を基盤とし、そして複雑な神経階層における各段階内と各段階間での再帰的な相互作用に起因する。クオリアにとっては、多感覚的な同型の表象をはじめとする、感覚受容器から高次の感覚階層に至るまでのすべてが重要である。またクオリアは、さまざまな感覚刺激や正負の価値の区別に役立つものとして、適応進化の観点から理解可能である。

これらの説明において意識と生物学的特性の間に成り立つとされている関係を、F&M は「創発 (emergence)」と呼んでいる。彼らによれば、意識とは「複雑な神経階層を備えた生きている脳の創発的特質である」(Feinberg and Mallatt 2016b, p.195, 邦訳 p.213)。一見すると、この「創発」という語は、彼らの理論を物理主義——つまり、世界に存在するものはすべて物理的なものであるという立場——の一種とみなすことをためらわせるものであるように思われるかもしれない。というのも、意識は創発的であるという主張は、意識は物理的なタームによっては説明不可能な根本的に新たな存在者であるという考え方を含む意味でなされることもあるからだ。しかし、F&M による「意識は創発的である」という主張は、そうした考え方を含まない。彼らによれば、創発的性質とは、たとえば、個々の水分子は流体ではないにもかかわらず、水分子の集合体である水は流体であるというように、全体を構成する個々の要素には見出されないが、それらの要素から構成される全体には見出されるような性質のことである。そして、そのような性質は、構成要素とその間の相互作用によって説明可能であるとされる (Feinberg and Mallatt, 2016b, pp.8–10, 邦訳 pp.10–13)。したがって、意識 (の四つの特性) は創発的であるという主張も、意識を生物学的特性に基づいて物理主義的に説明することを脅かすものではない。実際、彼らは、これらの説明は意識と神経プロセスの間の「説明ギャップ (explanatory gaps)」を橋渡しするものだと主張している (Feinberg and Mallatt 2016b, p.225, 邦訳 p.248)。つまり、意識の一見した

ところ不可解な特徴の根底にある生物学的基盤を特定できれば、その不可解さも解消される、というのが彼らの考えであろう。

以上より、意識の存在論に関する F&M の見解は、意識は物理主義にとって脅威になるような意味でその生物学的基盤を「超える何ものか」などではない、というものとして (ひとまず) まとめることができよう。彼らによれば、水 (の流動性) が水分子とその間の相互作用以上のものではないのと同様に、意識 (の四つの特性) もその基盤となる複雑な生物学的特性以上のものではないのである。

4.5 「自-還元不可能性」と「他-還元不可能性」⁷

しかし、F&M によれば、以上のように意識の四つの特性に関する「説明ギャップ」を橋渡しすることで話が終わるわけではない。彼らによれば、いわゆる「意識のハードプロブレム」には区別すべき複数のギャップが含まれており、それらをまとめて一挙に解決しようとするのは間違いなのである。具体的には、意識と物理的なものとの間には、説明ギャップに加えて、「経験ギャップ (experiential gap)」が存在する (Feinberg and Mallatt 2016a, p.123)。そして、それに対処して意識に関する十全な説明へ至るには、上記の四特性の説明に加えて、「自-還元不可能性」と「他-還元不可能性」という概念を導入しなければならない。

「自-還元不可能性」とは、「主観的意識は自身を生み出す客観的なニューロンを決して参照できないこと」を指す (Feinberg and Mallatt 2018b, p.113, 邦訳 p.120)。つまり、意識は (F&M によれば) ニューロンの活動のような物理的プロセス以上のものではないにもかかわらず、しかし、意識がその所有者に対してそうした物理的プロセスを開示することは決してないということである。それゆえ、意識主体としての一人称的な観点から——いわば意識の内側から——われわれが当の意識をもたらししている物理的プロセスに気

⁷ F&M 自身は「自-還元不可能性」と「他-還元不可能性」をそれぞれ、「自-存在論的還元不可能性 (auto-ontological irreducibility)」と「他-存在論的還元不可能性 (allo-ontological irreducibility)」と呼んでいる (Feinberg and Mallatt 2016b, p.220, 邦訳 p.242)。しかし、以下で述べるように、われわれの理解では、ここで問題になっているのは認識的なギャップであり、存在論的なギャップではない。したがって、ここでは「自-還元不可能性」と「他-還元不可能性」という表現を用いることにする (cf. Black and Allen 2017, p.1005; Suzuki, T. 2022, p.62)。

づくことはできないのである。他方、「他-還元不可能性」は、反対に、ニューロンの活動をどれだけ観察しても、その活動がもたらす意識そのものにアクセスすることはできないことを意味する。「客観的な第三者は、主体の経験を直接的には観測も測定もできない」のであり、「ひとたび神経状態が主観的意識の私秘的〔プライベート〕な状態を得ると、(...) 他者からは決して客観的に観測も経験もされなくなる」(Feinberg and Mallatt 2018b, pp.114-115, 邦訳 pp.121-122)。以上からわかるように、「自-還元不可能性」と「他-還元不可能性」という概念によって指摘されているのは、意識をもつ主体自身の視点と、脳の活動を観測する観察者の視点との間にある認識的なギャップであり、彼らが「経験ギャップ」と呼ぶものもつまりはある種の認識的なギャップとみなしてよいと考えられる。

F&M は、この認識的ギャップはリアルであり、決して埋められないものであると主張する。つまり、「意識の内側」から自分自身のニューロンの活動にアクセスできるようになることも、またニューロンの活動を観察することでそれがもたらす意識にアクセスできるようになることも決してないというのである。しかし、その一方で、そのような埋められない認識的ギャップが存在するという自体は自然科学の範囲内で説明可能であり、したがって、そのギャップの存在が自然科学に対して脅威となることはない、というのが彼らのスタンスである。

それでは、認識的ギャップに関する自然科学の枠内での説明とはどのようなものか。この点について彼らの記述はあまり豊富ではないが、まず、「自-還元不可能性の一部分は意識の進化過程から説明できる」とされている(Feinberg and Mallatt 2018b, p.114, 邦訳 p.121)。すなわち、動物の生存にとって重要なのは外界や自身の身体に注意を向けることであることを踏まえれば、意識が自分自身(の正体であるニューロンの活動)を参照することには適応上の利点は何もなく、それゆえ、意識は自-還元不可能性を伴うものとして進化したというわけである。また、他-還元不可能性については、参照性を初めとする意識の四つの特性はいずれも当の主体自身にしか経験できない特徴であり、他-還元不可能性はこれらの特性に由来するとされる。そして、それら四つの特性はいずれも特殊な神経生物学的特性に依拠して説明可能だったのだから、他-還元不可能性も自然科学にとって脅威となることはないというのである(Feinberg and Mallatt 2016a, p.123)。

以上より、意識の存在論+認識論に関する F&M の見解は、次のようにまとめられるであろう。

- (1) 意識は存在論的にはニューロンの活動をはじめとする一連の生物学的基盤以上のものではない。
- (2) しかし、意識とそれらの生物学的基盤の間には埋められない認識的ギャップが存在する。
- (3) ただし、そのギャップの存在は自然主義的に理解可能であり、それゆえ、自然科学にとっての脅威となることはない。

このような見解は、意識をめぐる哲学的議論において「タイプ B 物理主義」と呼ばれる存在論的立場と大枠において一致するように思われる。この立場は、意識は物理的なものである——意識と物理的なものの間に存在論的ギャップは存在しない——と主張する一方で、認識的ギャップの存在を認め、物理的なものに関するわれわれの知識から意識に関する知識を導き出すことは決してできないとする。そのうえで、当の認識的ギャップは、われわれが意識について考える際に用いる特殊な概念——「現象的概念 (phenomenal concepts)」——と、われわれが物理的なものについて考える際に用いる概念の相違に由来する(にすぎない)とされる(Loar 1990; Papineau 2002)。そして、現象的概念とその他の概念の相違を物理主義の枠組みの中で説明することで、この立場は物理主義を擁護しようとするのである。F&M は、タイプ B 物理主義者が「現象的概念」に訴えるところで、自/他-還元不可能性(とそれに関する進化論的・生物学的説明)に訴えているとみなすことができる。この類比が正しければ、F&M の見解は、ひとつには、タイプ B 物理主義に対して指摘されてきたのと同種の問題点に即して検討されるべきであろう(5.2.2を参照のこと)。

4.6 意識の適応上の価値と機能

最後に、意識の機能に関する F&M の見解を確認しておこう。まず注意しておきたいのは、彼らは「機能 (function)」という用語で何を意味するのかを明示したうえで、何が意識の機能なのかを明確に主張しているわけではない(し、またそもそも意識に関して「機能」という語をあまり用いない)という点である。よって、意識の機能に関する見解を彼らに帰属させることに関しては慎重でなければならないが、しかし、彼らが意識の適応上の価値について語っていることを踏まえれば、意識の機能は(それが厳密に何を意味す

るにせよ)、効果的に食べ物を獲得したり、捕食者やその他の危険から逃れたりといった活動を促進することで生存に寄与することにある、というのが彼らの考えだとみなしてよいと思われる。

彼らは、意識（の特性）がいかに生存にとって有利に働くかを強調する。たとえば、意識の参照性や統一性は、環境に関する統合された情報の獲得を可能にし、それによって環境内を生き延びる可能性を高める。またクオリアは、生存にとって重要なある刺激（たとえば捕食者の存在を示す刺激）を別の刺激から識別することを可能にし、それによって生存に寄与する、等々 (Feinberg and Mallatt 2016b, pp.217–220, 邦訳 pp.237–240)。意識はこうした適応的価値をもつものとして進化した現象として理解できる、と彼らは言うのである。これらの記述において、意識はもっぱら生存という目的にとって役立つものとして捉えられている。そして、このような考え方は、意識の機能は動物の生存の助けとなる一群の活動を可能にすることにある、という主張と自然に結びつくと考えられる。

このような見方下では、意識は、意識を欠く生物を含めた生物一般に共通の〈生存〉という目的に資するものとみなされている。この発想は、意識はその所有者であるシステム全体を生存とは別次元の目的をもつ存在にする（生命とは別個の）「ありさま」であり、そしてそのようなものとしての意識には機能はないと主張する、G&J の考え方とは対照的である。

以上で、G&J と、そして F&M の見解の概要を確認した。次節では、それぞれの説に関する暫定的な評価や議論となる点を見ていくことにしよう。

5. 評価と問題点の抽出

ここでは、意識の徴候 (5.1) と、そして意識の存在論および機能 (5.2) の二点に分けて、G&J と、F&M の見解それぞれに対する論点の洗い出しと暫定的な評価を行う。その際、両理論に対してこれまでになされてきた批判的なコメントなども適宜参照する。

5.1 意識の徴候について

5.1.1 ギンズバーグ&ヤブロンカ

意識の徴候に関する G&J の見解のポイントは、意識の十分条件を構成するものと広く認められている一群の特徴に目を向け、それらをすべて備えたもののみに可能な無制約連合学習を意識の「ポジティブな目

印」として設定する、というものであった。このアイデアは、意識の本性に関する論争を避けつつ、意識に関する共通のリサーチプログラムを打ち立てようとするものとして、よく練られた興味深いものである。その一方で、当然のことながら、批判的なコメントや、今後さらに検討されるべき点の指摘などもなされている。以下では、その一部を紹介することを通じて、彼らの見解をめぐる特に重要な論点を確認する。

第一に、G&J が想定している無制約連合学習と意識の結びつきは、現在のところ経験的な証拠によって十分にサポートされてはいないという指摘がなされている (Browning and Veit 2021; Birch 2020)。まず、無意識の（つまり意識下の刺激に対する）無制約連合学習というものが存在する可能性がある。もし無意識の無制約連合学習が存在するのであれば、無制約連合学習が意識のポジティブな目印となる——つまり「無制約連合学習があれば意識がある」——という主張は維持できないことになるであろう。現在そのような学習が存在するという経験的な証拠があるわけではない。だが、それは意識と学習の結びつきに関する実験的研究がまだ十分に進展していないからにすぎないかもしれない。今後の研究により、そのような学習の存在が後押しされる可能性は否定できない。また反対に、今後の経験的な探求により、無制約連合学習よりももっと単純な学習が意識の徴候として支持される可能性もある。もしそうなれば、意識の目印として無制約連合学習よりも適格なものがあることになる。たとえば、F&M は、無制約連合学習よりも単純な「大域的オペラント条件づけ」という学習を情感的意識の存在を示すものとみなしていた。どちらの連合学習が意識の指標としてより適格なのは、経験的な探求を通じてさらに追及されるべき点である。G&J の見解の成否は、このように意識と学習の関係をめぐる今後の経験的な探求にかかっている⁸。

⁸ ただし、このような経験的な探求の実行に関しては厄介な問題がある。すなわち、無制約連合学習と意識の結びつきを経験的に検証する——たとえば、意識がなくても無制約連合学習は可能かどうかをテストする——には、まず無制約連合学習とは独立に、ある動物に意識があるかないかを決定しておかなくてはならないが、その決定を行なうための理論中立的な方法が存在しないように思われるという問題である (Browning and Veit 2021)。意識理論から独立の意識の測定法は存在しないという問題は、意識理論を検証しようとする試みの本質的な難点だとみなされている (Irvine 2013; Michel 2019)。

第二に、関連する点として、G&Jの理論において「学習」が中心的な位置を占めていることに批判を向ける論者もいる。そうした論者は、意識の十分条件を構成するとされる一群の特徴の中には、学習と本質的に結びついているわけではないものが含まれていると主張する。ゴドフリースミスは、意識の十分条件を構成するとされた諸特徴のうち、たとえば、知覚的特徴のバイディングや選択的注意などは、学習とは無関係に、動物が周囲の出来事に対してリアルタイムで反応する際にも重要な役割を果たしているとして主張する (Godfrey-Smith 2021)。また、アーヴァインによれば、志向性、行為主体性、身体性、自他の区別といった特徴は、学習よりも、空間内で行為する自己についての自己中心的表象を生み出す能力とより密接に結びついている (Irvine 2021)。これらの論者によれば、意識の十分条件とされた諸特徴をもつば無制約連合学習と結びつけるのは適切ではなく、むしろそれらの特徴そのものや、またそれらと結びつく学習以外の能力 (自己中心的表象など) に目を向けるべきなのである。以上を踏まえると、意識の諸特徴を学習と結びつけるという着想それ自体も、今後詳しく検討されるべきであろう⁹。

第三に、意識の十分条件を構成する諸特徴に関して、ほんとうに G&J が主張するほど広範な同意があるのかを疑う論者もいる。たとえば、マシャリーとカラザースは、そのような合意は存在しないと主張する (Masciari and Carruthers 2021)。彼らによれば、(G&J の主張に反し) この点で意識と生命をめぐる状況には顕著な相違がある。たしかに生命の場合には、どの範囲のものに生命があるか、また、生命あるものにどのような特徴が共有されているかについて、研究者の間で広範な同意がある。だからこそ、それらの特徴を手がかりとして生命の目印 (無制約遺伝) を設定することもできる。しかし、どの生物に意識があるか、また、意識あるものにどのような特徴が共有されているかについては、現状では広範な不一致がある。そして、このような状況で特定の諸特徴を意識の十分条件とみなし、その諸特徴を手掛かりに意識の目印を設定することは、結局意識をもつものの範囲に関して勝手な取り決めを行なうことにすぎない、というのである。

マシャリー&カラザースによれば、たしかに、われわれ人間には間違いなく意識があり、そして、われわれ人間には G&J が挙げたような諸特徴が備わっている。その意味で、問題の諸特徴がわれわれの意識に「付随する (accompany)」ということには多くの人が同意するだろう。しかし、だからといってそれらの特徴のすべてが意識の説明において重要な役割を果たすという点で広範な同意があることにはならない。たとえば、諸特徴の二つ目に挙げられた大域的なアクセス可能性が人間の意識に見いだされることは、たしかに広く認められている。だが、マシャリー&カラザースが指摘するように、大域的アクセス可能性が意識に構成的に寄与しているかどうかについては論争がある。こうした点を踏まえると、G&J の見解の土台とされた同意がどれくらい広範囲のものであるのかとともに、またその同意がどの程度実質的なものであるのかについても、検討の余地があると言えるだろう。

5.1.2 ファインバーグ&マラット

意識の徴候に関する F&M の見解は、外受容意識の徴候は「同型的な神経表象」であり、情感意識の徴候は「大域的オペラント条件づけ」であるというものだった。意識の徴候に関するこの見解は、哺乳類や鳥類だけでなく、より原初的な生物にも意識があるという、彼らの興味深い仮説を支えるうえで決定的な役割を果たしている。しかし、そもそものいかなる根拠に基づいて、またはいかなる方法論的戦略にしたがって、同型的な神経表象や大域的オペラント条件づけが意識の徴候とされているのかは、十分に明らかとは言えない。それらを意識の徴候とするという考え方自体がもっともらしくないというのではない。同型的な神経表象や大域的オペラント条件づけが意識の徴候であるというのは、有望な仮説のうちのひとつであろう。しかし、単に「有望な仮説のうちのひとつ」ではなく、他の仮説に比してより説得力のある選択すべき理論になるには、それらを意識の徴候とみなすべき明確な根拠や方法論的正当性を述べなければならないであろう。だが、F&M は、そのような根拠や正当性を十分明確な形で提示していないように思われる。

たとえば、彼らは次のように述べている。

第一に、感覚的インプットを、世界と身体に関する詳細で多感覚的な地図の表象 [= 同型的な神経表象] へと組織化するような脳をもつあらゆる動

⁹ 以上の批判に対する G&J (とバーチ) の応答については、Birch et al. (2021) を見よ。

物は、心的イメージを意識的に経験することができる、とわれわれは仮定した (we assumed)。実際、そのような神経地図は (...) しばしば原意識 (primary consciousness) [高次の反省的意識と対比される基本的な意識] の目印であるとみなされている (...) 第二に、大域的オペラント学習を行うことのできるあらゆる動物はポジティブおよびネガティブな結果を感じることができるのであり、それゆえ、(人間がもつ感じられる感情に似た) 情動的なタイプの意識をもつ、とわれわれは仮定した。その際のわれわれの論理は、動物が報酬や罰となる刺激に対して最初に反応し、その後複雑な行動を通じて当の刺激を記憶していることを示すならば、その動物はそれらの刺激を最初の段階で感じていたのに違いないというものだった。(Feinberg and Mallatt 2018a. [] 内は引用者補足)

しかし、ここで述べられていることだけでは、外受容意識や情感意識の徴候に関する彼らの仮定 (assumption) を意識研究の出発点として受け入れるべき十分な理由が説得力をもって示されているとは言えない。これらの仮定に関する F&M の記述は少なく、この点は彼らの見解の未発展な部分の一つであると言えよう。

ちなみに、外受容意識の徴候が同型の神経表象であるとされていることに関しては、F&M が意識に関するある種の表象説を採用していることが手掛かりになるのかもしれない (Feinberg and Mallatt 2016b, p.265, n.1, 邦訳 p.263, 注 1)。当該の種類の表象説は、「主体が意識をもつとは、主体が環境 (や自身の身体) のありように関する表象をもつことである」と主張する。この種の表象説を前提すれば、同型の表象が意識の徴候となるという考え方も、たしかに自然なものとなるであろう。しかし、もし意識の徴候に関する F&M の見解がそのような仕方では表象説を前提しているのであれば、彼らはその説のより詳しい内実と、それを支持する理由についてより明確に述べなければならないはずである。

5.1.3 人工生命体への応用可能性

意識の徴候に関わる考察の締めくくりとして、G&J と F&M のどちらにも関係する発展的な論点について述べておこう。彼らは、意識にアプローチするにあ

たって、意識をもちうると現時点で確信できる唯一の存在者である生物に焦点を絞り、生物の意識の徴候となる学習能力や表象能力について進化論的観点から論じた。その試みの中では、ロボットや AI のような人工物の意識の可能性は意図的に脇に置かれている。たとえば G&J は、無制約連合学習の能力が意識の徴候であるというのはあくまで生物に関して言えることであるとし、AI やロボットに関しても同じことが成り立つと主張することは明示的に差し控えている (Ginsburg and Jablonka 2019, p.467, 邦訳 (下) p.284)。

このように、まずは「私たちが知っている意識」である生物の意識から話を始め、人工物の意識はひとまず脇に置くという戦略は、たしかに意識研究を手堅く進めるやり方として理解可能である。しかし近年になり、生命科学技術が飛躍的に進歩した結果、人工物と生物の境界が曖昧になりつつあることは注目に値する。たとえば、試験管内で人間の多能性幹細胞から三次元脳組織を生成する技術の進歩によって、(人間の脳にもみられるような) 同期的な神経活動を示す脳オルガノイドを生成することが可能になった (Sakaguchi et al. 2019)。こうした脳オルガノイドは、自然な生殖／複製プロセスを通じて生み出されたわけではないが、遺伝子を備えていて代謝などの生命現象もみられるという意味では、AI やロボットとも異なる。そして最近では、試験管内で生成された脳組織をシンプルなコンピューターゲームと接続し、電極によってそのゲームへの出力とそこからの入力を実装したところ、脳組織がゲームのやり方を学習したという報告がなされた (Kagan et al. 2022)。同様に、脳オルガノイドに光刺激に反応する眼に似た組織を発現させることができたという報告もある (Gabriel et al. 2021)。

こうした人工的生命体が意識をもつかどうかという問いは、現在大きな注目を集めつつある。そしてこの文脈において、無制約連合学習や同型の神経表象などが意識の徴候として機能しうるかどうかは、興味をそそる問題である。たとえば、仮想的な環境に結びつけられた脳オルガノイドが意識をもつか否かは、それが無制約連合学習や大域的オペラント条件づけの能力をもつかどうかを調べることによって検証できるかもしれない。あるいは、眼に似た組織をもち光刺激に反応する脳オルガノイドが意識をもつかは、それが同型の神経表象を形成できるかどうかを調べることによって検証できるかもしれない。

もちろん、こうした人工的生命体が本当に「生物」と言えるのかはさらに検討する必要がある。また、G&J や F&M は、こうした応用を意図していたわけではないだろう。それでも、最先端の科学技術によって生成された新たな生命体の意識の有無を評価するうえで、この応用可能性はきわめて重要である。

5.2 意識の存在論と機能について

5.2.1 ギンズバーク&ヤブロンカ

自然主義と目的論

意識をひとつの「ありさま」とみなすという方針は、G&J の見解を存在論的な点でも独特で興味深いものにしてている。ありさまとしての意識は、一方で、(デカルト的自我などとは違い) 進化論的な観点から理解できる生物学的なもの、したがって、自然的なものともみなされている。他方で、それは真に目的論的なものであるともされている。すなわち、意識というありさまを獲得した動物は、快不快といった価値的側面を含む経験に基づいて自らの欲求を満たすという、意識あるものに特有の目的を指向する存在となるというのである。意識が目的論的かつ生物学的な現象として自然界に実在するというこのような立場は、意識をめぐる現在の哲学的な論争においてユニークな位置を占めていると言えるであろう。一般に、意識や心の自然化をめぐる現在の哲学的議論においては、生物学的現象を含む自然的なものは最終的に目的論をもち出すことなく説明可能であることが前提されている¹⁰。自然的なものは非目的論的に説明できるという前提のもとで、その意味での自然的なものによって意識や心を説明できるかどうかが論じられているのである。これに対して、G&J の理論は、自然的なものに関するその前提自体の見直しを迫るものとして解釈できる。

この「アリストテレス的自然主義」とでも呼ぶべき立場は興味深いものである。しかし、一般に現代の自然科学においては、目的論的なものが還元不可能な形で実在するというような考え方は認められていない。したがって、G&J の立場がいかにして「自然主義」としての内実を確保できるのかは問題になるだろう。とりわけ、進化の総合説に基づく現在の標準的な進化生物学では、そもそもく目的に向かっ

て行動する有機体」というような存在者は(説明項としても被説明項としても) 重要なものとはみなされておらず、もっぱら遺伝子(という「有機体以下の(suborganismal)」存在者)と個体群(という「有機体以上の(supra-organismal)」存在者)が説明上基礎的なものとされている(Nicholson 2014)。そのうえで、適応進化という現象は、個体群の遺伝子頻度の変化として(有機体の目的指向的なふるまいなどに訴えることなく)説明されるのである。すると、G&J のアリストテレス的自然主義は、現在の進化生物学と合致しない——それゆえ、「自然主義」と名乗ることが適切でない——時代錯誤で風変わりな説と考えられるかもしれない。それゆえ、彼女らの見解は、まずは進化生物学をはじめとする現在の自然科学の実践との整合性という観点から検討されるべきであろう。

しかしその一方で、近年、生物学や生物学の哲学において、進化の総合説に基づく従来の進化生物学は修正される必要があるとか、さらには他の見方にとって代わられるべきであるとの意見が提起されており、盛んな議論になっているということは興味深い事実である(Walsh 2015; Moosavi 2018)。特に、進化発生学の分野では、有機体の発生が適応進化において果たす役割を正當に扱うには、遺伝子と個体群のみに注目する見方を改め、有機体という存在者を進化論において重要なものとして復権させるべきだとの意見もある。そしてそれに伴い、生物学における目的論的な説明の正当性や役割を認めるべきだとする議論も展開されている(Walsh 2015)。こうした論争は今まさに進行中であり、その行方は明らかでないが、今後新たに形成される進化生物学のパラダイムのうちに、G&J の見解が足場を見出す可能性もあるだろう¹¹。

意識の価値や意義をめぐる議論との関連

意識というありさまを獲得した動物は新たな目的の空間のうちに生きることになるという G&J のアイデアは、意識をもつことの価値や意義をめぐる哲学的な議論に対してもインパクトをもちうる。というのも、彼女らのアイデアは、このテーマに関わる特定の立場と結びつき、相互に支持しあう可能性があるからだ。たとえば、Andrew Lee (2019) によれば、意識はその所有者が享受するよい(または悪い)ものの一

¹⁰ 生物学的な「目的」や「機能」に訴えて心を自然化しようとする理論においても、それらの目的や機能は最終的には目的論的なものへの言及なしに還元的に説明可能とみなされている。cf. Millikan 1989.

¹¹ ヤブロンカは自身でも、進化に関する従来の「遺伝子中心の(gene-centered)」見方に修正を迫る見解を提示している(Jablonka and Lamb 2006)。

つであるのではない。そうではなく、意識はその所有者を福利の主体とし、物事が主体にとって「重要である (matter)」——つまり、主体にとってよかったり悪かったりする——ということ自体を成り立たせるものである (Kriegel forthcoming, sec. 5)。この説によれば、意識はそもそも何かが主体にとってよい (悪い) ものとなるためのいわば根源的でニュートラルな土台としての役割を果たすのである。

このような考え方は、G&J の「ありさま」としての意識というアイデアと親和的である。彼女らの考えでは、意識というありさまを獲得した動物は、快を求めて不快を避け自身の主観的な欲求を満たすという自律的な目的を指向する存在となり、そして (その欲求を満たしたり満たさなかったりする) あれやこれやの事物を正や負の価値をもつものとして経験するようになる。ここでは、意識は、何かがそれにとってよい／悪いと言えるような主体をもたらし、そして主体の主観的欲求に応じてさまざまな事物に価値が付与されるためのまさに土台とみなされている。このように、ありさまとしての意識というアイデアは、意識の価値に関わる特定の立場と相互支持の関係に立つと言えるだろう。

またより一般的に、G&J の見解は、「福利」や「賢慮的価値 (prudential value)」に関する議論——すなわち、「道徳的な意味で善い」こととは区別される、「われわれ個人や生物個体にとって内在的によい」こととは何かをめぐる議論——に対して、興味深い手がかりを与えてくれるかもしれない。彼女らによると、意識ある動物にとって、快を獲得し不快を避けることで主観的欲求を満たすことがよいことである。この点で、彼女らの主張は福利に関する快樂主義と一致する¹²。しかし、G&J の理論においては、単に「意識ある動物にとっての福利は快である」という点で説明がストップするわけではない。というのも、彼女らの理論によれば、快を獲得し不快を避けることで主観的欲求を満たすことは意識ある動物にとっての「目的」であり、そのような目的に向かってふるまうことがそれらの動物の「ありさま」となっているからである。すなわち、快を得て不快を避けることは、意識ある動物がいわば

自らの「生き方を全うする」ことなのであり、その意味で快はよく不快は悪いのである。快のよさに関するこのような説明は、目的の達成という (アリストテレス由来の) 観点から福利に関する快樂主義の正当化を試みるものとして、さらなる検討に値するであろう。

以上の点を踏まえると、意識と道徳的地位 (moral status) の関係に対しても、G&J の見解は興味深い示唆を与えてくれる。一般に、意識をもつものは特別な仕方では道徳的な配慮の対象となるという考え方は、多くの人が直観的にもっともらしいと感じるものであり、また理論家によってもしばしば支持されている (Niikawa 2018; Shepherd 2018; Kriegel forthcoming)。この考え方によれば、たとえば、刺激に対して全く無反応な二人の患者の一方には意識があり、他方にはないとした場合 (また、ある動物には意識があり、それと似た別のある動物にはないとした場合)、前者は後者に比べてより高い程度の道徳的配慮 (または後者には向けられない独特な種類の道徳的配慮) の対象となる。G&J の見解は、このような考え方の一つの根拠を示唆していると言える。というのも、彼女らによれば、意識をもつものは快を求め不快を避けることで主観的欲求を満たすという (単なる生存や生殖を超えた) 自身にとっての目的を指向し、その目的に照らして「重要である」——よかったり悪かったりする——様々な事物に対処しつつ生きる存在だからである。つまり、意識を有するものは (意識を欠くものには見られない仕方) 自身の「関心を追求する」存在なのである。このような見方に基づけば、意識あるものが特別の道徳的地位をもつという考えも理にかなったものとなるであろう。意識あるものに特徴的な目的や関心の追求を妨害したり促進したりすることが道徳的に重要だと考えるのはもっともらしいからである (もちろん、このことは「意識と直接関係がない目的や関心の追究は道徳的に重要ではない」という主張を含意するわけではないが)。このように、G&J の見解は、目的や関心という観点から、意識と道徳的身分を結び付ける考え方に一定の根拠を与える可能性がある。

G&J の理論のねらいは、直接には意識を自然主義的に理解することにあり、意識をめぐる価値的・倫理的テーマを説明することにあるわけではない。しかし、以上のように、アリストテレス流の目的論を下敷きとする彼女らの理論は、意識の価値論に対しても貢献しうることを見込みがある。当該テーマに対する自然主義的な意識理論からの貢献は貴重であり、彼女らの理論

¹² G&J の見解によれば、意識に加えて、理性という「ありさま」を獲得した人間の福利に関しては快樂主義的に理解することはできないかもしれない。しかし、ここではひとまず意識というありさまのみをもつものとしての福利を問題とすることにしよう。

のこの側面を解明し発展させることは、今後の重要な研究課題の一つであるだろう。

5.2.2 ファインバーグ&マラット

G&Jの理論とは対照的に、F&Mの理論は、意識の存在論的な身分に関してタイプB物理主義に類する立場を取る点でも、また意識の機能を生存の観点から捉える点でも、従来のスタンダードな自然主義に沿っていると言える。彼らの見解によれば、意識を新たな目的の次元をもたらし特異な現象として理解する必要などなく、意識を欠いた生物にも広く見いだされる、生存に寄与するさまざまなもの——たとえば歯や心臓のような——と同列のものとして理解することができるのである。(そしてさらに、生存への寄与という一見すると目的論的に思える現象も、最終的に非目的論的に説明可能であるとF&Mは考えている。Feinberg and Mallatt 2016b, p.24, 邦訳 p.30.) しかし、もちろんそのような彼らの見解にも検討されるべき点はある。ここでは、彼らの存在論の見解に向けられる懸念について簡単に述べたのち、意識の機能を生存と結びつける考え方に関わる争点を指摘することにしよう。

認識的ギャップを認め、存在論的ギャップを否定するという方針の見込み

タイプB物理主義は、意識と脳活動の間の認識的ギャップを認めつつ、存在論的ギャップを否定する。この方針に対しては、次のような根本的な懸念が提起されている。すなわち、認識的ギャップを真正のものとして認めそれについて十分な説明を与えることと、存在論的ギャップを否定することは、やはり相反する方向へ向かっており、両立させることは難しいのではないかというものである(Chalmers 2006)。一方で、現象的概念が本当に意識に特有の認識的ギャップを説明するのに十分なものであるならば、そのような概念を導入することで結局物理主義の枠内に収まらないような何かを(暗に)もち出すことになるのではないかという疑いが生じる。他方で、現象的概念が物理主義の枠組みの内側で問題なく説明できるようなものであるならば、それが本当に意識に特有の認識的ギャップを説明しうるかどうかが明らかでない、というわけである。

こうした懸念は、F&Mの存在論の見解にも向けられる。一方で、自-還元不可能性や他-還元不可能性が真に認識的ギャップを説明するものだとするならば、彼らはそれらを導入することで結局物理主義の枠組み

を超える何かをもち出しているのではないだろうか。たしかに彼らはそうではないことを示そうとし、たとえば、自-還元不可能性は適応進化の観点から説明可能だと述べている。だが、そのような進化的説明が可能であることは、自-還元不可能性が強い意味で創発的な——物理的なタームで還元的に記述できない——性質であるという可能性を排除しない。ある性質が適応進化の結果として選ばれることは、その性質の存在論的地位に関する特定の見解を含意するわけではないのである。他方で、自/他-還元不可能性が明確に物理主義の枠内で説明可能なものだとするならば——たとえば、自-還元不可能性とはつまるところ(意識をもたない)AIやロボットにも見出されるようなものとしての「志向性」に依拠して説明可能なものとするならば——、それが果たして意識に特有の認識的ギャップを説明しうるかどうかが明らかでないだろう。先にも触れたように、自-還元不可能性や他-還元不可能性に関するF&Mの考察は、現在のところはまだ十分に展開されているとは言えない状態にある。しかし、認識的ギャップを認めつつ存在論的ギャップを否定するという彼らの方針が説得力をもつには、これらの還元不可能性について、より詳しい説明や議論が必要である。

意識の機能は生存の観点から理解できるのか

意識の機能を生存の観点から理解することは、進化的観点から意識理論を構築する試みの中ではきわめて自然なことに思われる。しかし、太田敏史・鈴木大地・田中泉史は、(少なくとも視覚的意識に関して)このような路線を現在の意識研究の実践にそぐわないものとして批判している(Ota et al. 2022)。彼らによると、現在の意識研究では、広義の報告可能性——言語報告に限らず、その他の行動による報告も含む意味での報告可能性——を基準として意識を研究することが一般的である。そして、そのような実践と合致するのは、意識の機能を生存ではなく「行為主体性 agency」の観点から理解するという方針である(Ota et al. 2022, p.464)。ここで「行為主体性」と呼ばれているのは、(あらかじめ固定された目的ではなく)自分自身にとっての可変的な目的のために自発的にそして合理的にふるまう能力のことである。そのような能力は、たとえば、視覚刺激に対して人から促されなければ行動できず、また視覚刺激に対する自身の行動の理由や意味を理解できない盲視の状態にある人(の

視覚)には欠けているものだとされている。視覚的意識はそのような行為主体性のための「バックグラウンド」であり、主体が行為の目標を特定し、遂行すべき行為を選択することを可能にするのだ、と太田らは主張している。

ここで付け加えておかなければならないのは、この批判はF&Mの見解を全面的に退けるためのものではなく、むしろ最終的には彼らの見解(の一部)を擁護するためのものだという点である。というのも、太田らによれば、行為主体性は人間のような高い概念能力を有するごく一部の動物だけでなく、より原初的な動物にも広く見出されるものだからである。行為主体性はそのように多くの動物に見出されるものであるゆえに、意識の機能を行為主体性と結びつける方針は、結局のところ、比較的原初的な種類も含む多くの動物に意識があるというF&Mの見解と親和的だといえる(太田 et al. pp.465–466)。また、行為主体性が生存だけには縛られない可変的な目的に関わるものであるということは、行為主体性が生存のために発揮されることもあることを排除するわけではない。実際、行為主体性(そしてそれゆえ意識)を伴う動物はしばしば、食べ物を獲得したり、危険を回避したりといった生存に関わる目的のために(も)行為主体性を発揮するであろう。この点で、意識の機能を生存と結びつけるF&Mの見解は部分的に——とりわけ、動物個体にとっての可変的な目的と、生存という目的が大部分一致すると思われる原初的な動物に関しては——正しいとも言える、と太田らは指摘する(太田 et al. p.466)。このように、意識の機能を生存ではなく行為主体性と結びつけたとしても、F&Mの見解の重要な部分を保持することはできるかもしれない。

F&Mの理論をどのように擁護/改良できるかという問題とは独立に、意識の機能をいかなる観点から理解するのかという問いは、それ自体重要で興味深いテーマである。特に、太田らの議論の中で、行為主体性が「動物個体自身にとっての可変的な目的」というものに関連付けられていることは注目に値する。このことは、意識の機能を問題にする際に、(意識をもたない生物のレベルには見出されない)「新たな目的の次元」を導入すべきかどうか重要なポイントの一つとなることを示唆している。この問いは、「意識に機能はない」とするG&Jにおいても焦点となっていたものである。意識を有する動物は、「単なる生存」を超えた、自律的な目的——快やその他の自身にとって

の価値——を追求する存在であるというような考え方には、たしかに何かしら魅力や説得力があるようにも思われる。このような考え方が、結局は古風な哲学的幻想にすぎず、進化論的な観点に基づく自然主義的な意識理論の中に居場所をもちえないものなのか、それとも何らかのやり方で進化論的な観点とも両立させることができるものなのか。この問いに取り組むことを通じて、新たな生物学的自然主義をさらに発展させることができるだろう。

参考文献

- Belardinelli, Sofia. & Pievani, Telmo. 2023. “What if Consciousness has no Function?” *Biosemiotics*. 16: 259–267.
- Birch, Jonathan. 2020. “In Search of the Origins of Consciousness.” *Acta Biotheoretica* 68 (2): 287–94.
- Birch, Jonathan, Simona Ginsburg, and Eva Jablonka. 2020. “Unlimited Associative Learning and the Origins of Consciousness: A Primer and Some Predictions.” *Biology & Philosophy* 35 (December): 56.
- . 2021. “The Learning-Consciousness Connection.” *Biology & Philosophy* 36 (5): 49.
- Black, Dylan, and Colin Allen. 2017. “Out of the Sludge: How Vertebrates Came to Have Subjective Experience.” *Bioscience* 67 (11): 1004–7.
- Browning, Heather, and Walter Veit. 2021. “Evolutionary Biology Meets Consciousness: Essay Review of Simona Ginsburg and Eva Jablonka’s *The Evolution of the Sensitive Soul*.” *Biology & Philosophy* 36 (1): 5.
- Campbell, Cary. 2022. “Embracing the Learning Turn: The ecological context of learning.” *Biosemiotics*. 15 (3): 469–481.
- Chalmers, David. 2006. “Phenomenal Concepts and the Explanatory Gap.” In *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism*, edited by Torin Alter and Sven Walter. Oxford University Press.
- Feinberg, Todd E., and Jon M. Mallatt. 2016a. “The Nature of Primary Consciousness. A New Synthesis.” *Consciousness and Cognition* 43 (July): 113–27.
- . 2016b. *The Ancient Origins of Consciousness: How the Brain Created Experience*. The MIT Press. [鈴木大地訳『意識の進化的起源』, 2017年,

- 勁草書房]
- . 2018a. “Unlocking the ‘Mystery’ of Consciousness.” *Scientific American*. <https://blogs.scientificamerican.com/observations/unlocking-the-mystery-of-consciousness/>.
- . 2018b. *Consciousness Demystified*. The MIT Press. [鈴木大地訳『意識の神秘を暴く』, 2020年, 勁草書房]
- Gabriel, Elke, Walid Albanna, Giovanni Pasquini, Anand Ramani, Natasa Josipovic, Aruljothi Mariappan, Friedrich Schinzel, et al. 2021. “Human Brain Organoids Assemble Functionally Integrated Bilateral Optic Vesicles.” *Cell Stem Cell* 28 (10): 1740–1757.e8.
- Ginsburg, Simona, and Eva Jablonka. 2019. *The Evolution of the Sensitive Soul: Learning and the Origins of Consciousness*. The MIT Press. [鈴木大地訳『動物意識の誕生 上・下』, 2021年, 勁草書房]
- . 2020. “Consciousness as a Mode of Being.” *Journal of Consciousness Studies*. Volume 27, Numbers 9–10, 2020, pp. 148–162.
- . 2022. “Learning and the Evolution of Conscious Agents.” *Biosemiotics*. 15 (3): 401–437.
- Godfrey-Smith, Peter. 2016. *Other Minds: The Octopus, the Sea, and the Deep Origins of Consciousness*. Farrar, Straus and Giroux. [夏目大訳『タコの心身問題』, 2018年, みすず書房]
- . 2020. *Metazoa: Animal Life and the Birth of the Mind*. Farrar, Straus and Giroux. [塩崎香織訳『メタゾアの心身問題』, 2023年, みすず書房]
- . 2021. “Learning and the Biology of Consciousness: A Commentary on Birch, Ginsburg, and Jablonka.” *Biology and Philosophy* 36 (5): 1–4.
- Huxley, T. H. 1866. *Lessons in Elementary Psychology*. London: Macmillan.
- Irvine, Elizabeth. 2013. “Measures of Consciousness.” *Philosophy Compass* 8 (3): 285–97.
- . 2021. “Assessing Unlimited Associative Learning as a Transition Marker.” *Biology & Philosophy* 36 (2): 21.
- Irwin, Louis N. 2023. “How Minimal Can Consciousness Be?” *Biosemiotics*. 16 (1): 21–26.
- Jablonka, Eva, and Marion J. Lamb. 2006. *Evolution in Four Dimensions: Genetic, Epigenetic, Behavioral, and Symbolic Variation in the History of Life*. The MIT press.
- Kagan, Brett J., Andy C. Kitchen, Nhi T. Tran, Forough Habibollahi, Moein Khajehnejad, Bradyn J. Parker, Anjali Bhat, Ben Rollo, Adeel Razi, and Karl J. Friston. 2022. “In Vitro Neurons Learn and Exhibit Sentience When Embodied in a Simulated Game-World.” *Neuron*, September. <https://doi.org/10.1016/j.neuron.2022.09.001>.
- Kriegel, Uriah. forthcoming. “The Value of Consciousness to the One Who Has It.” In *The Importance of Being Conscious*, edited by Geoffrey Lee and Adam Pautz. Oxford: Oxford University Press.
- Kull, Kalevi. and Favareau, Donald. 2022. “There is Umwelt Before Consciousness, and Learning Transverses Both.” *Biosemiotics*. 15 (3): 491–495.
- Lee, Andrew. 2019. “Is consciousness intrinsically valuable?” *Philosophical Studies* 176: 655–671.
- Loar, Brian. 1990. “Phenomenal States.” *Philosophical Perspectives. A Supplement to Nous* 4: 81–108.
- Masciari, Christopher F., and Peter Carruthers. 2021. “Perceptual Awareness or Phenomenal Consciousness? A Dilemma.” *Biology & Philosophy* 36 (2): 18.
- Michel, Matthias. 2019. “Consciousness Science Underdetermined: A Short History of Endless Debates.” *Ergo: An Open Access Journal of Philosophy* 6. <https://doi.org/10.3998/ergo.12405314.0006.028>.
- Millikan, Ruth. 1989. “Biosemantics,” *Journal of Philosophy* 86 (6): 281–297. [前田高弘訳「バイオセマンティクス」, 信原幸弘編『シリーズ心の哲学 III 翻訳篇』所収, 2004年, 勁草書房]
- Moosavi, Parisa. 2018. “Neo-Aristotelian Naturalism and the Evolutionary Objection: Rethinking the Relevance of Empirical Science.” In *Philippa Foot on Goodness and Virtue*, 277–307. Cham: Springer International Publishing.
- Nicholson, Daniel J. 2014. “The Return of the Organism as a Fundamental Explanatory Concept in Biology.” *Philosophy Compass* 9 (5): 347–59.
- Niikawa, Takuya. 2018. “Moral Status and Consciousness.” *Analele Universității Din București – Seria Filosofie* 67 (1): 235–57.
- Noble, Denis. 2022. “The Evolution of Consciousness and Agency.” *Biosemiotics*. 15 (3): 439–446.
- Olteanu, Alin. 2022. “Learning as Becoming Conscious: A note on Jablonka and Ginsburg’s Notion of Learning.” *Biosemiotics*. 15 (3): 457–467.
- Ota, Koji, Daichi G. Suzuki, and Senji Tanaka. 2022. “Phylogenetic Distribution and Trajectory

- ries of Visual Consciousness: Examining Feinberg and Mallatt's Neurobiological Naturalism." *Journal for General Philosophy of Science / Zeitschrift Für Allgemeine Wissenschaftstheorie* 53 (4): 459–76.
- Papineau, David. 2002. *Thinking about Consciousness*. Oxford University Press.
- Sakaguchi, Hideya, Yuki Ozaki, Tomoka Ashida, Takayoshi Matsubara, Naotaka Oishi, Shunsuke Kihara, and Jun Takahashi. 2019. "Self-Organized Synchronous Calcium Transients in a Cultured Human Neural Network Derived from Cerebral Organoids." *Stem Cell Reports* 13 (3): 458–73.
- Searle, J. 1992. "The Rediscovery of Mind (Representation and Mind)." *Cambridge, MA: Bradford Book*. [宮原勇訳『ディスカバー・マインド』, 2008年, 筑摩書房]
- Sharov, Alexei A. "Consciousness and Learning from the Biosemiotic Perspective." *Biosemiotics*. 15 (3): 483–490.
- Shepherd, Joshua. 2018. *Consciousness and Moral Status*. Routledge Focus on Philosophy. London, England: Routledge.
- Suzuki, Daichi G. 2022. "A General Model of and Lineage-Specific Ground Plans for Animal Consciousness." *Annals of the Japan Association for Philosophy of Science* 31: 5–29.
- Suzuki, Takayuki. 2022. "The Evolutionary Origins of Consciousness: A Key to the Solution of the Hard Problem of Consciousness?" *Annals of the Japan Association for Philosophy of Science* Vol.31: 55–73.
- Tønnessen, Morten. 2022. "The Evolutionary Origin(s) of the Umwelt." *Biosemiotics*. 15 (3): 451–455.
- Walsh, D. M. 2015. *Organisms, Agency, and Evolution*. Cambridge University Press.
- (2023年3月20日投稿、2023年7月31日再投稿、2023年7月31日掲載決定)