

目的論的機能主義と日常的／科学的心理学

久保田 智也

1. はじめに

トマト売り場で立ち止まっている彼女は、一つ一つのトマトの状態を比較し、よく熟したものを選ぼうとしているのだろう。カレー粉をかごに入れて野菜売り場へ向かう彼は、今晚カレーを作るつもりだろうから、玉ねぎやにんじんを購入するにちがいない。ほとんどの人間は、こうして行動から心を読み取り、心から行動を予測する能力を有している。哲学において、この能力は「日常心理学」や「素朴心理学」などと呼ばれる。

他方で、行動や心の働きを支える情報処理メカニズムを、観察実験や数理モデルに基づいて詳細に理解することを目指すのが認知科学である。トマトの表面から反射した光が網膜に投影されてから色や形の知覚経験が生じるまでに、どのような計算処理が行われているのか（Marr 2010）。知覚したトマトが熟しているかどうかは、文的表象を用いた推論によって判断されるのか、それとも熟したトマトに関するプロトタイプとの類似度によって判定されるのか（Churchland 1995）。カレー粉売り場から野菜売り場に向かう人の脳内で、店舗全体の空間配置はどのように表象されているのか。こうした問いに答える認知科学の理論は、日常心理学との対比のために「科学的心理学」とも呼ばれる。

日常心理学と科学的心理学がいかなる関係にあるのかという問題は、特に 80 年代以降の心の哲学において盛んに論じられてきた（Bermudez 2005）。哲学者たちの立場は還元主義と反還元主義とに大別できる。還元主義によれば、「信念」や「欲求」などの概念による日常心理学的な説明は私たちの脳の働きに関する真なる記述であり、科学的心理学は信念や欲求が脳においてどう実現されているのかを明らかにするだろう。他方で反還元主義者は、日常心理学が脳の認知メカニズムについて正しい説明を与えているということを否定する。つまり、日常心理学は脳の働きについての誤った説明であり、放棄されなければならないと主張するか、あるいは、

日常心理学はそもそも脳のメカニズムに関する説明ではないと主張する⁽¹⁾。

そして現在の論争状況としては、還元主義の展望は暗いといえる。たとえばゴドフリースミスは、信念や欲求といった日常心理学の表象概念に還元的定義を与える試みがことごとく失敗に終わってきたことを理由に、還元主義について悲観的な見方を示している（Godfrey-Smith 2004, 147–8）。実際のところ、かつては信念や欲求などの還元について論じていた志向性の自然主義的理論（cf. Millikan 1984; Dretske 1988）は、近年ではもっぱら認知科学の表象概念に焦点をあてる傾向にある（Rupert 2018, 206）。日常心理学の還元主義にはこの20年ほど進展が見られない。

ところで、この論争に関わる哲学者たちは、信念（そしてときに欲求）を日常心理学的概念のモデルケースとして取り上げることが多い。信念に対する還元主義の成否によって、日常心理学全体に関する還元主義の成否を測ることができるかのような想定は、両陣営によって暗黙に共有されてきたと思われる。しかしながら、日常心理学は部分的に正しく、部分的に誤っている、ということはあるまいのだろうか（cf. Godfrey-Smith 2004; Bermudez 2006, ch. 5）。

本稿ではこの可能性を追求してみたい。具体的には、日常心理学の還元主義を意図概念に関して擁護することを目指す。すなわち、意図に関する日常的説明は、ある一定の場合には行動主体の脳状態を正しく記述しており、科学的心理学がその裏付けを与えるはずだということを主張する（他方で、信念や欲求などその他の命題的態度の還元可能性に関して本稿は何も主張しない）。この主張は、素朴な意図概念の分析とミリカンの目的論的機能主義を組み合わせることで導かれる。

議論は以下のように進行する。第2節では日常心理学における意図概念の分析を

1 「還元主義」という語は、心身問題の文脈では、心的タイプは物理的タイプに同定できるという心身同一説を意味することがある。しかし、ここでの「還元主義」は、心と物質の形而上学的関係というよりも、日常的心理学と科学的心理学という二つの理論のあいだの関係についての主張を表す。この意味での還元主義にはさまざまな特徴づけがありうるが（Bermudez 2005, 44）、還元主義の基本的な主張は、日常心理学による心や行動の説明は科学的心理学の言葉で言い換えることができる、日常心理学が定める法則は、科学的心理学が発見した法則によって説明することができる、といったものである。同様に、ここでの「反還元主義」はこれらの主張を何らかの理由で否定する立場を表し、心身問題における機能主義を意味するものではない。反還元主義の代表的立場には、消去主義と道具主義がある。消去主義は日常心理学は根本的に誤った理論である（それゆえ当然、科学的心理学によって言い換えられるものでもない）とし、道具主義は日常心理学は行動の予測にこそ役立つものの、脳状態を初めとした特定の実在を記述しているわけではないという立場である。

行う。私たちの意図帰属実践に着目することで、日常心理学における意図とは目標志向的行動を引き起こす一種の心的表象であることが明らかとなる。第3節では、本来機能と指令的表象というミリカンの概念、および科学的心理学の目的にまつわる彼女の見解を提示する。それに従えば、科学的心理学の目的は、行動や身体構造の目的論的性格を進化論的機能（本来機能）として捉えたうえで、それに対する機能分析を行うことである。第4節では、科学的心理学はある種の人間行動を説明するために意図を措定せざるをえないと論じることで、日常心理学における意図概念の還元可能性を示す。

2. 意図概念の分析

以下の分析の射程は限定的であるということをあらかじめ注意しておく。第一に、「意図とは何か」という哲学的問題への解答は与えられない。以下で行われるのはただ、たとえば「あの人はできるだけ熟したトマトを買うことを意図している」のような意図帰属において、私たちがいかなるものを「意図」と呼んでいるかの明確化のみである。意図に関する私たちの素朴な考えが哲学者をも満足させるのかという問題は、本稿の範囲を越える。

第二に、以下の分析は特定の仕方による意図帰属にのみ着目する。というのも、私たちが意図帰属の方法はただ一つではない。大きく分けて、心的状態から推論する、行動から読み取るという二つの方法があり、これらを組み合わせて利用することもできる。そして自己自身の意図については内観に頼ることもできる。「SはPを意図している」という判断が意味するものは、判断のために用いられた手段によって異なるかもしれない。素朴な意図概念は一貫した意味内容をもっていないかもしれない。しかし本稿は、意図帰属のあらゆる方法や、それらのあいだの（非）共通性について議論しない。以下の分析は、もっぱら行動観察にのみ基づく意図帰属を対象とする⁽²⁾。

2 なおこれに依じて、本稿全体の結論もまた限定的なものとなっている。たとえば、私たちはいくつかの意図概念を場面に依じて使い分けているということが判明したならば、それら全てについて還元主義が正しいのか、それとも一部は消去されるべきなのかといった問題が当然生じるであろうが、本稿はその点について何も主張しない。以下で論証されるのはただ、行動観察に基づく意図帰属において「意図」と呼ばれている内的状態の実在性は科学的心理学によって裏付けられるはずだということのみである。

2.1 意図は目標志向的行動から読み取られる

意図帰属の手がかりとなる運動の特徴にはどのようなものがあるかについて、心理学者たちはさまざまな見解を述べてきた。Baron-Cohen（1995）によれば、他者の視線や運動の自己推進性を検出するモジュールが意図帰属の基盤を提供しているという。また Csibra（2008）は、行動の可変性（同一の結果に対し、状況に応じてよりよい手段を採用できるか）によってある主体が意図的に動いているのかどうかを判断するという人間の生得的傾向について論じている。以下では、これらの心理学者たちの見解と、Trestman（2011, 211）のいう行動の非明示的目標志向性（implicit goal-directedness of behavior）の関係について検討したい。

非明示的目標志向性とは、状況の変化に対処しながら一定の状態に到達するという運動の様式をいう。トレストマンは、具体例として、タカが上空からウサギに襲いかかるといった動物の捕食行動をくりかえし引き合いに出す。捕食対象の位置、移動に応じて捕食者も自らの運動を変化させ、最終的には対象を仕留める。もちろん、この概念は人間の行動にも適用できる。疲弊した長距離トラック運転手が信号での停止・発進、先行車や後続車の動きに応じた車線の変更などを無意識的に行なっているとき、この行動は荷物を届けるという目標に向かう非明示的目標志向性を示している⁽³⁾。

トレストマンによれば、非明示的目標志向性の概念は、主体の意図や行動の進化論的機能とは独立に適用できる。それは特定の状態に収束していく可塑的なふるまい以上のものを意味しないからである。たとえば、私たちは、意図も進化の歴史ももたないことが明らかな幾何学図形の運動にさえ、非明示的目標志向性を見出す。また、タカの捕食行動は進化論的機能をもつ本能的行動かもしれないが、私たちはそれについて知ることなしに、タカの行動に非明示的目標志向性を認めることができる（少なくともダーウィン以前の人々はそうしていた）。また、人間の行動に対しても、それが意図的に行われたのかどうかについては意見を控えつつ、非明示的目標志向性を認めることができる（この点はこの論証において重要な役割を担う）。

ところでトレストマンが行動のこの性質を「非明示的」と呼ぶのは、意図によって引き起こされたという意味での明示的目標志向性との区別のためである。たしか

3 運転手の例はトレストマンによるものではなく、デイヴィッド・アームストロングによって導入され、意識の哲学においてしばしば言及されることとなった事例である（Armstrong 1980, 59）。

に二つの概念は異なるものだが、両者の外延は重複している。明示的目標志向性とは、非明示的目標志向性に意図という要素が加わった特例だからである。そこで、以下では前者を単に「目標志向的行動」と呼び、後者を「意図的行動」と呼ぶこととする。

さて、自己推進性、視線の固定、行動の可変性を伴う目標志向的行動の事例は枚挙に暇がない。たとえば、何かを手に取りするという目標志向的行動は、ほとんどつねにこれら3つの特徴を示すと思われる。机の上のコップを手に取りするという行動について考えれば、まず腕の動きに関して自己推進性が見られる。また、とりわけ遠くに置かれたコップを取ろうとするときには、視線はコップに対して固定されるだろう。また、椅子に座ったままではコップに手が届かないと気づけば、立ち上がって手を取ろうとするだろう。ここに行動の可変性が見られる。そして心理学者たちが論じるように、これらの特徴に気づけば、私たちは「この人はあのコップを取るつもりで腕を伸ばしたり、椅子から立ち上がったたりしているのだ」のように意図を読み取る。

まとめよう。いくつかの心理学理論によれば、私たちは運動の自己推進性、視線、可変性などを手がかりとして意図帰属を行う。しかるに、目標志向的行動はしばしばそれらの特徴を示す。それゆえ、私たちは目標志向的行動の主体に対して意図を帰属させることがある。なお、この帰結は、目標志向的行動に対して私たちはいつでも意図を読み取るとか、目標志向的行動に対する意図帰属はつねに正しいとかいったことを含意しないということには注意すべきである。

2.2 意図はいかなる内的状態だと考えられているか

では、目標志向的行動に対する意図帰属がなされるとき、意図はいかなる内的状態だと考えられているだろうか。ここでは二つの点を確認したい。

第一に、意図には充足条件があると想定されている。すなわち私たちは、特定の事態の成立・不成立に応じて、意図が実現した・しなかったと語る。たとえば、コップをとるという意図は、コップをとることができたときには実現したと言われ、手を滑らせてコップを落としてしまったときには実現しなかったと言われる。

第二に、意図の充足条件は行動の目標と一致する。というのも、充足条件を有する内的状態は、意図だけではない。私たちは、目標志向的行動の主体に対して、充足条件を有する複数の内的状態を帰属させることができる。たとえば主体は、椅子から立ち上がりたくないという欲求をもっているが、コップを手を取るために仕方

なく立ち上がったのだという解釈がありうる。このとき主体には「椅子に座ったままにいる」を充足条件とする内的状態（欲求）と、「コップを手取る」を充足条件とする内的状態（意図）の二つが帰属されている。これらは充足条件を有するという共通点をもつにもかかわらず、後者だけが主体の意図と見なされる。前者の状態は実現しなかった欲求だと考えられるにすぎない。私たちはいかなる基準で単なる欲求と意図を区別しているのだろうか。これに対する簡単な説明は、意図の充足条件は行動の目標と一致すると考えられているから、というものである。つまり、 p という意図をもつ人の行動は p という事態に向かう目標志向性をもつはずだと私たちは考えている。しかるに、具体例において、主体の行動はコップを手取るという目標への志向性を見せてはいるが、椅子に座ったままにいるという目標への志向性は見せていない（自発的に立ち上がったのだから）。それゆえ、充足条件を有するいくつかの内的状態のうち「コップを手取る」を充足条件とする内的状態だけが意図と見なされるのだ、と説明できる。

結局のところ、私たちは意図ということで、目標志向的行動の目標と一致するような充足条件を有する内的状態を考えている（ことがある）といえる。第4節では、科学的心理学がある種の目標志向的行動を説明するためには、この種の内的状態を指定せざるをえないということを示す（それにより意図に関する素朴心理学の還元主義を擁護する）。第3節では議論の前提となるミリカンの目的論的機能主義を導入する。

3. ミリカンの目的論的機能主義

はじめに、ミリカン哲学のもっとも基本的な道具立てである目的論的機能の諸概念を導入する。次に、指令的表象の定義を述べる。最後に、科学的心理学の目的に関するミリカンの適応主義的見解を紹介し、それが考慮に足るものであるということの理由をいくつか提示する。なお、彼女による機能概念や表象概念の定義は非常に込み入っているため、それらを一言一句忠実になぞることはせず、本稿の議論に必要な点のみ解説する。

3.1 目的論的機能

心臓は血液を循環させるためにある。クジャクのオスのディスプレイはメスのクジャクとつがいになるために行われる。ドライバーはねじを回すための道具であ

る。このように、何かは何らかの目的をもつという事態を、ミリカンは（必ずしも生物のそれには限られない）進化に基づいて説明する。

ミリカンは、あらゆる種類の目的を直接機能（direct function）と派生機能（derived function）の二つに分類する（Millikan 1984, Ch.1-2; Millikan 2014）。直接機能とは、生物の形質、行動、あるいは人工物などが、その働きをもつがゆえに普及した（進化した）といえるような働きのことである。たとえば、心臓は血液を循環させるという働きにより、心臓をもつ生物個体の適応度をそれ以外の個体のそれよりも高めたことで、生物集団内に普及したといえる。したがって、心臓の直接機能は血液循環である。クジャクのオスは、羽根を大きく広げてメスにアピールすることで、それをしないクジャクよりも多くの子孫を残し、結果として正常なオス個体はみな同様の求愛行動をとるようになった。したがって、クジャクのオスがメスに対して羽根を広げてみせるのは、メスとつがいになるためである。ドライバーはねじを回すのに役立つからこそ大量生産されている。それゆえ、ドライバーの直接機能はねじを回すことである（ここでドライバーの設計者の意図に言及する必要はない）。これらが「直接」機能と言われるのは、遺伝的形質や本能的行動など、進化した当のものが有する機能だからである。

直接機能には、関係機能（relational function）と呼ばれる下位分類がある。これは、異なる環境に対応して異なる働きをする直接機能である。働きバチは、蜜を見つけて巣に帰ったあと、蜜の在処を仲間たちに知らせるためのダンスを踊る。ここで、ハチにはダンスを踊るための生得的メカニズムが備わっていると考えよう。このメカニズムは、ダンスを生み出すことで仲間たちに蜜の場所を知らせるという直接機能を有している。しかるに、どのようなダンスを踊るかは、どの方角の、巣からどれだけ離れた場所に蜜を見つけたかによって異なる。その点で、これは関係機能だといえる。

ところで、ハチが方角 X 、距離 Y の地点 (X, Y) に蜜を発見して帰巢したとき、ハチが踊るべきは（正確には、ダンスのための内部メカニズムが生み出すべきは）他でもなく (X, Y) に蜜があることを知らせるダンスであると、直観的には言いたくなる。つまりこの特定の状況において、ハチの内部メカニズムは、 (X, Y) に蜜があることを伝えるというより特定の機能を有すると言いたい。この直観を捉えるためにミリカンは適合機能（adapted function）という概念を用意した。関係機能を有するシステムが特定の環境に置かれたとき、それに応じた適合機能を獲得する。ハチの内部システムは上記の環境において、ダンスによって (X, Y) に蜜があ

ることを伝えるという適合機能を有する。

以上を前提として、派生機能の概念を導入することができる。派生機能とは、適合機能を有するシステムが機能遂行の過程で生み出した事物が当のシステムから継承する機能である。たとえば、上記の状況でハチが踊ったダンスは、(X, Y) に蜜があることを知らせるという派生機能を有する。派生機能の概念は、ハチのダンスが、(X, Y) に蜜があることを知らせるという目的のための手段であるという直観を反映している。

3.2 指令的表象

表象には記述的表象と指令的表象、および記述的かつ指令的表象の三種類がある。ミリカンは主張するが、ここでは指令的表象の特徴のみ押さえておく。端的に言えば、指令的表象とは、進化した送り手受け手構造⁽⁴⁾のもとで送り手によって生み出され、受け手から特定の反応を引き出すことを機能とする媒体⁽⁵⁾である（Millikan 1984, Ch. 6）。以下では、親が子に「部屋を掃除しなさい」と命じる場面を例にとつて、この定義を説明しよう。

送り手は一定の記号を生み出し、受け手はそれに応じて一定の行動をとる。「部屋を掃除しなさい」という発話は、親によって生み出され、子の一定の行動、つまり部屋の掃除を始める行動を引き起こす。ここで親は送り手、子は受け手である。

ところで、言葉によって他人に命令するという行為は、この親子が歴史上初めて行ったものではない。言語コミュニケーションは数十万ないし数百万年前に始まり、それをする個体たちの適応度を高めたためか、あるいはその利益を見抜いた他個体たちが模倣したためか、ともあれ何らかの経路で人類全体に普及した。この意味で言語を介する送り手受け手構造は（広い意味で）進化したのだといえる。

それぞれの言語表現もまた、送り手や受け手に一定の利益をもたらすことで普及

4 「進化した送り手受け手構造」という語そのものはゴドフリースミスがミリカン（やルイスやスカームズ）の立場を説明するさいに導入した用語である（cf. Godfrey-Smith 2013）。

5 ミリカン自身は「媒体」の代わりに「アイコン」という言葉を用いている。ミリカンはその語をパースから借用したと述べている（Millikan 1984, 95）。たしかに、ミリカンにおいて全ての表象体系はその対象領域と構造的に同型なので、彼女のいう表象にはパースのいう「アイコン」に通ずるところはたしかにある。しかし、本稿ではその同型性の内実について説明しない以上、「アイコン」という術語を用いることの正当性を疑う読者もいるかもしれない。これは以下の議論を理解することの妨げになる恐れがあるため、ここでは「アイコン」よりも含意の少ない中立的な表現として「媒体」を採用する。

してきたものと考えられる。「部屋を掃除しなさい」という命令文は、受け手が部屋の掃除をするという、送り手と受け手双方にとっての共通利益（親は部屋が整頓されていることを望んでおり、子は掃除することで説教を終わらせられる）をもたらすからこそ広く利用されている。仮に、多くの子どもたちがこの命令文を理解する能力をもたない、あるいは部屋の掃除をすることが子どもたちにとってきわめて大きな不利益であるなどの理由で「部屋を掃除しなさい」が部屋の掃除を引き起こせなかったとしたら、この表現は普及しなかっただろう。

しかるに、異なる命令文は異なる機能をもつはずだ。たとえば「部屋を掃除しなさい」と「廊下を掃除しなさい」では、受け手から引き出すべき反応が異なる。たとえば目的語を変更するといった、何らかの操作により機能を変化させられる記号は（本稿では）「媒体」と呼ばれる。日本語の命令文は全てこの意味での媒体である。

指令的表象の定義についての例解は以上である。最後に、のちの議論のため、指令的表象をやりとりする送り手受け手構造は単一個体の内部にも存在すると考えられることを確認しておこう。ミリカンによれば、欲求は指令的表象の一種である（Millikan 1995, 165ff.）。欲求は何らかの内的メカニズムによって生み出され、運動に一定の行動を生じさせることを機能とする。欲求は脳内で何らかの言語的媒体として実現されており、媒体の構造に応じて、異なる行動を生み出す機能をもつ。たとえば、「部屋を掃除したい」と「廊下を掃除したい」という欲求は、それぞれ物理的に異なる脳状態であり、前者は部屋の掃除、後者は廊下の掃除を引き起こすことを機能とする。そして以上のシステム全体が人間には生得的に備わっている。このようにして指令的表象とその送り手・受け手が個体の内部に存在すると想定することは十分に可能である。

3.3 科学的心理学の目的

ミリカンによれば、科学的心理学と動物行動学（ethology）は、行動がいかなる機能をもち、いかなる内的メカニズムに支えられているのかを説明するという目標を共有している。両者の違いはただ、人間とそれ以外の動物のどちらを研究対象とするかという点にある（Millikan 1995, 163–4）。

ここでいう機能とは、もちろん、ミリカンが定義した直接機能や派生機能のことである。したがって、動物行動学と科学的心理学の目標の一つは、行動（たとえばオスのクジャクのディスプレイ）がいかなる結果をもつ（メスとつがいとなる）こ

とによって進化したのかを明らかにすることだといえる。あるいは、行動そのものが進化したのではない場合は、行動を生み出した内的メカニズムがいかにして進化したのかを明らかにすることが目指される。

ところで、この種の機能帰属を行うためには、一定の行動パターンもしくは行動を生み出すメカニズムを同定し、それらの進化史について仮説を立てる必要がある。しかるに、人間やそれ以外の動物による行動の全てが機能をもつわけではない。たとえば「落ちる」、「転ぶ」、「滑る」などの言葉で記述される行動は、大抵において機能をもたない。他方で、「狩る」、「釣る」、「探す」、「挑戦する」、「脅迫する」などと記述される行動は、機能をもつと思われる（Millikan 1995, 144）。さらに、行動は自然言語の動詞と対応しない仕方で分類することもできる（Millikan 1995, 142-3）。たとえば「東京大学本郷キャンパスの正門からちょうど13メートル離れたところで立ち止まったまま2度まばたきする」という行動タイプは、過去に何度か例化されたことだろうが、およそ機能をもつとは考えられない。

要するに、一方で、自然なやり方と不自然なやり方を合わせれば、行動は際限ない仕方で分類することができ、大半の分類はいかなる機能ももたない。他方で、機能は進化史に依存する以上、行動観察から直ちに知りうるものではない。それでもなお人間にとって動物行動学と科学的心理学が遂行可能であるという事実は、機能を有する行動パターンを特定するための効率的な方法が存在するということを示唆する。

ミリカンは、行動が法則性をもつことや、繰り返し行われるということは、機能をもつことのよい手がかりではないと主張する（Millikan 1995, 143, 145）。他方で、科学者たちが実際にはいかなる手がかりを用いているのかについて、明確なことは語ろうとしない（Millikan 1995, 148）。しかしながら、直接機能と派生機能の導入を動機づけていた、あらゆる種類の目的は進化論的機能にほかならないというミリカンの想定がもし正しければ、さまざまな動物の行動から意図を読み取ってしまう私たちの生得的な傾向は、二つの科学を導くヒューリスティックの一つであるに違いない。実際、ミリカンが機能のない行動として挙げた「落ちる」、「転ぶ」、「滑る」について、私たちはこれらの行動から意図を読み取ることは稀である。他方で機能のある行動として挙げられた「狩る」、「釣る」、「探す」、「挑戦する」、「脅迫する」に関しては、意図帰属を控えることのほうが難しい。

ここから、第2節で見た目標志向的行動もまた科学的心理学の説明対象となると考えられる。なぜなら、そこで論じたように、私たちは目標志向的行動に対して意

図帰属する傾向にあり、かつ、たったいま論じた理由により、直観的な意図帰属を誘発するような行動は科学的心理学の研究対象となると思われるからである。次節では、科学的心理学が目標志向的行動に進化論的機能を帰属させうるいくつかの方法を確認し、ある種の目標志向的行動に機能を帰属させるためには指令的表象を指定せざるをえないということを示す。

4. 意図に関する還元主義

目標志向的行動とは、状況の変化によって異なる経路をとりながら同一の状態へ至るという様式をもつ行動であった。眼瞼反射や吸啜反射は、本能的かつ目標志向的な行動だと言えるかもしれない（阻害要因が生じたときにどのような反応を示すかに依存する）。そうである場合、これらの行動は自然進化に由来する直接機能を有する。眼瞼反射の機能は眼球を保護することであり、吸啜反射の機能は乳首に吸いつくことである。

目標志向的行動のなかには、何らかの学習によって進化したものもあるだろう（cf. Millikan 1984, 28-9）。たとえば、お菓子を食するという行動は、甘味を強化子とする条件づけられた目標志向的行動になっている場合がある。本人が間食を避けるという意志や、家族がお菓子を隠す、買ってこないようにするといった阻害要因に直面してもなお、あらゆる手段を用いてお菓子を手に入れようとするようであれば、これはたしかに目標志向的である。オペラント学習を経てお菓子を食べようとするという行動の頻度が増加したという意味で、この行動も進化したものだといえる。その直接機能は甘味を得ることである。あるいは、模倣学習によって集団内に普及した目標志向的行動というものも考えられる。前節で具体例に挙げられた狩猟行動はその一例と見なせるかもしれない。その直接機能は獲物を捕獲することである。

上記のような観点から、多くの目標志向的行動に機能を帰属させられるだろう。会社に出勤する、論文を書くといった行動さえ、条件付け学習と模倣学習の観点から直接機能を帰属させる余地はあるかもしれない。

しかしながら、その行動自体はいかなる意味でも進化の歴史をもたず、それゆえ直接機能の帰属が不可能であるような目標志向的行動のカテゴリーが存在する。本稿ではそれを「気まぐれな行動」と呼ぶ。

たとえば、立ち上がって今いる場所からまっすぐ5メートル歩く、障害物は動か

すか乗り越えることとし、何としても5メートル先の地点まで進む、という目標志向的行動を気まぐれに起こすことは可能である。もちろんこれは本能的行動でも、条件付け学習によって獲得された行動でもない。また、誰かがそれを実行しているのを見て模倣したものでもない。今からこの行動をとることと、過去の同様な行動（そのようなものがあつたとして）がもたらした帰結とのあいだには、いかなる因果的な関係もない。それゆえ、ミリカンによる直接機能の定義を、この気まぐれな行動は満たさない。

気まぐれな行動とは、歴史上初めて行われるか、あるいはそうでないにしても、過去の試行とはいかなる因果的連関ももたない（それゆえ進化の歴史をもたない）目標志向的行動である。上記よりも自然な具体例としては、子供の行動に見られるような、即興で思いついた遊び（1つ飛ばしと2つ飛ばしを交互にしながら階段を降りる）、理由のないこだわり（家に帰るまで左右の手を同じ形のまま保つ）、などが挙げられる。

気まぐれ行動は進化の歴史をもたない以上、直接機能をもつことがない。それゆえ、科学的心理学が機能の概念によって気まぐれ行動を説明するためには、それに派生機能を帰属させるほかない。そのためには、気まぐれ行動を引き起こすことを直接機能とする何らかのメカニズムが存在すると想定する必要がある。このように想定することでのみ、気まぐれ行動を当該メカニズムの産物として位置づけ、派生機能を帰属させうるからである。

気まぐれ行動のためのメカニズムが人間に備わっているという想定は経験的仮説に過ぎないが、十分に見込みのある仮説だと思われる。というのも、気まぐれな行動を引き起こすメカニズムは、次のようにして適応に寄与する可能性があるからである。気まぐれ行動は、その定義上、遺伝的にプログラムされているわけでも、学習によって獲得されたものでもない。それゆえ一方で、少なくとも種が生息する標準的な条件下では、生得的行動や学習された行動と比べて、気まぐれ行動が個体に利益をもたらす保証は小さいといえる。しかし他方で、気まぐれ行動は、生得的行動や学習された行動を繰り返すだけでは到達できない、新たな利益を生物にもたらしうる。これまでに誰も食べたことのない植物を食べてみる、誰も行ったことのない場所に行ってみるという行動は現在以上の繁栄の可能性に開かれている。身体的な資源に余剰があり、危険に陥る可能性の小さい状況で気まぐれ行動をとることは、ほとんどの場合は利益も損害も生まず、まれに利益を生む（それゆえ選択されうる）と考えられる。

ともあれ、気まぐれ行動のためのメカニズムが存在するという想定のもとで議論を進める。このメカニズムは、運動系に神経伝達を行うことで気まぐれ行動を引き起こす。ここで運動系に送られた信号は、特定の気まぐれ行動を引き起こすことを機能とする指令的表象だといえる。というのも、仮定より、気まぐれ行動のメカニズムは、さまざまな信号を運動系へ送り、運動系がそれに応じた気まぐれ行動を引き起こし、その行動が十分な頻度で適応度を高めたことにより選択されてきた（このとき、個々の気まぐれ行動のタイプが選択されたわけではないということに注意せよ）。したがって、メカニズムは（3.2 節で導入した意味における）送り手、運動系は受け手、神経の活性化パターンは媒体であり、これらは気まぐれ行動を引き起こすという共通の機能を果たすことで共進化してきた。この意味で、メカニズムが運動系に送っているのは、一定の気まぐれな行動を引き起こすことを機能とする指令的表象だといえる。

そして、私たちが気まぐれな目標志向的行動から読み取る意図は、この指令的表象に還元できる。そのことを示すには、意図と指令的表象の同一性を仮定したうえで、指令的表象についての記述から意図に関する直観的な記述が導かれることを示せばよい⁶⁾。よって、まずは気まぐれな行動の主体が p という意図をもつということは、 p を引き起こすことを機能とする指令的表象をもつということにほかならないと仮定しよう。ここから導かれるべき直観的主張は二つある（第2節を参照）。第一に、気まぐれな目標志向的行動の主体は必ず意図を有するという。これは、気まぐれな行動は指令的表象によって引き起こされているという科学的心理学の（前段落で見た）説明と上記の仮説を合わせることで帰結する。第二の直観的主張は、意図の充足条件は行動の目標であるということである。これは上記の仮定から直ちに導かれる。こうして指令的表象と意図の同定は科学的心理学の主張から日常心理学の主張を導くことを可能にするので、その同定は正当なものである。すなわち、少なくとも気まぐれ行動の主体に関しては、彼が有する指令的表象は日常的な意味での意図にほかならないといえる。こうして意図概念に関する還元可能性が示された。

5. おわりに

6 目的論的機能主義による還元は多くの場合（暗黙的に）こうした手続きをとる。cf. Schulte 2020.

本稿の目的は、素朴心理学の還元主義を、意図概念に関して擁護することであった。上記の議論がもし正しければ、科学的心理学は気まぐれな目標志向的行動に対する機能帰属のためにミリカンのいうところの指令的表象を指定せざるをえず、またこの指令的表象こそが、私たちが気まぐれな行動に対して直観的に帰属する意図なるものの正体だといえる。

上記の議論に不十分な点があるとすれば、その一つは、ミリカンの適応主義的心理学観の正当化であろう。この点は、認知心理学者の用いる機能概念はミリカン流の定義によってもっともよく捉えられるということを示すことにより補いうるだろう。これは今後の課題とする。

参考文献

- 引用文中の [] は引用者による補い、[...] は中略を表す。
- Armstrong, D. M. 1980. *The Nature of Mind and other Essays*. Cornell University Press.
- Baron-Cohen, S. 1995. *Mindblindness: An Essay on Autism and Theory of Mind*. The MIT Press.
- Bermudez, J. L. 2005. *Philosophy of Psychology: A Contemporary Introduction*. Routledge.
- Churchland, P. M. 1995. *The Engine of Reason, the Seat of the Soul: A Philosophical Journey into the Brain*. The MIT Press.
- Csibra, G. 2008. “Goal Attribution to Inanimate Agents by 6.5-month-old Infants.” *Cognition*, 107: 705–717.
- Dretske, F. 1988. *Explaining Behavior: Reasons in a World of Causes*. The MIT Press.
- Godfrey-Smith, P. 2004. “On Folk Psychology and Mental Representation.” In: *Representation in Mind: New Approaches to Mental Representation*, edited by Clapin, H., Staines, P. and Slezak, P. Elsevier.
- . 2013. “Signals, Icons, and Beliefs.” In: *Millikan and her Critics*, edited by Ryder, D., Kingsbury J., and Williford K. Wiley-Blackwell.
- Marr, D. 2010. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. The MIT Press.
- Millikan, R. G. 1984. *Language, Thought, and Other Biological Categories: New Foundations for Realism*. The MIT Press.
- . 1995. *White Queen Psychology and Other Essays for Alice*. The MIT Press.
- . 2014. “The Tangle of Natural Purposes That Is Us.” In: *Contemporary Philosophical*

-
- Naturalism and Its Implications*, edited by Bashour, B. and Muller, H. D. Routledge.
- Rupert, R. D. 2018. "Representation and Mental Representation." *Philosophical Explorations*, 21(2): 204–225.
- Schulte, P. 2020. "Why Mental Content is not Like Water: Reconsidering the Reductive Claims of Teleosemantics." *Synthese*, 197: 2271–2290.
- Trestman, M. A. 2012. "Implicit and Explicit Goal-Directedness." *Erkenntnis*, 77: 207–236.