

目的論的機能主義——学習に基づく表象論の検討

勝亦佑磨 (Yuma Katsumata)

日本学術振興会特別研究員・東京大学

本発表の目的は、心の哲学における目的論的機能主義の立場から、私たちの心の持つ志向性がいかにして自然化されうるかという問題を考察することである。志向性とは、あるものがそれ自身とは別の何かを表象するという性質である。例えば、「目の前にヘビがいる」という信念は「目の前にヘビがいること」を表象しているとともに、「目の前にヘビがいる」という内容（表象内容）を持つという意味で、志向性を持つ。目的論的機能主義はこうした志向性を説明する有力な立場のひとつであり、心的状態を生物の心臓や肺のように「生存に有益な」機能を持つ状態として捉えることで、志向性の自然化を試みる立場である。目的論的機能主義の主流な見解は、ミリカンによる進化に基づく表象論（例えば Millikan, 1993）であり、それによれば、心的状態を生み出す形成機構（生産者）とそれに基づいて行動を引き起こす利用機構（消費者）は、それぞれの機能を持つために祖先の生存や繁殖に役立ち、それゆえに進化の過程で選択され、現在においても存続している。例えば、「目の前にヘビがいる」という信念の生産者と消費者の組は、目の前にヘビがいるときにその信念を生み出し、その信念に基づいて「逃げる」という行動を引き起こすといった機能を持つと考えられる。そして、こうした信念の生産者と消費者は、それぞれの機能を持つために祖先の生存に役立ち、選択され、現在においても存続していると考えられるのである。このようにして選択されてきた生物の器官や特質の持つ機能は目的論的機能と呼ばれ、目的論的機能主義の論者はこうした目的論的機能によって心的状態の持つ志向性を説明するのである。

このように目的論的機能主義の主流な立場は、ミリカンによる進化に基づく表象論であるが、本発表であえて注目するのは、ドレッツキによる学習に基づく表象論である（Dretske, 1988）。ドレッツキは、進化による説明と学習による説明の役割を区別することで、自身の学習に基づく表象論の必要性を主張する。ドレッツキによれば、前者は、生物の種や集合にまたがる表象を説明するものであるのに対し、後者は生物個体の表象を扱うものである。また、生得的行動や反射的行動に扱われるような表象は進化に基づく理論によって説明されるものであることに対し、信念や欲求といった意図的行動の「理由」でもありうるような表象は、学習に基づく理論によって説明されるという。ドレッツキはこのようにして、学習に基づく表象論の独自の役割を主張するが、彼の理論には、以下のような点において検討の余地がある。

第一に、ドレッツキ表象論は、誤表象を説明できているのかどうかという問題がある。私たちは、例えば、目の前にある枝をヘビだと間違えるときにもまた、「目の前にヘビがいる」という信念を持ちうる。このような誤表象のケースにおける「目の前にヘビがいる」という信念もまた、「目の前にヘビがいる」という内容を持つという意味で、志向性を持つ。このように、心が志向性を持つこと、すなわち表象することには、常に、誤って表象する可能性が伴っている。だがもしドレッツキがこうした誤表象を説明できな

いのだとすると、そもそも彼の表象論を、志向性を説明する理論として受け入れることすらできないであろう。

第二に、ミリカンによる進化に基づく表象論のほかに、学習に基づく表象論を支持する積極的な根拠があるのかどうかという問題である。ドレツキは、学習に基づく表象論と進化に基づくミリカンによる表象論による志向性の説明が異なるものであるとして、前者の必要性を主張するが、これらが本質的に異なるものとして単純に二分されるのかどうかという点には検討の余地がある。というのも、ある表象メカニズムが、ある種のレベルで選択され淘汰されるのか、それともある個体のレベルで選択され淘汰されるのかという違いに過ぎないとすれば、進化と学習に基づく表象論は、それほど差がないものであると考えられるからである。さらに、ミリカンによる進化に基づく表象論もまた、学習によって獲得されるような表象メカニズムに一定の説明を与えており、わざわざドレツキが提案するような学習に基づく表象論を持ち出さなくとも、それらを説明できる可能性があると考えられる。そうだとすると、ミリカンによる進化に基づく表象論のほかに、学習に基づく表象論を支持する積極的な根拠があるのかどうかという点が疑わしくなる。

本発表では、こうした第二の問題を中心に検討し、ドレツキによる学習に基づく表象論による説明は、ミリカンによる進化に基づく表象論による説明と本質的には差がないことを示す。そして、学習に基づくドレツキ表象論を支持する積極的な根拠はそれほどないということを示す。

《主な参考文献》

- Dennett, D. C. 1991, "Ways of Establishing Harmony," *Dretske and His Critics*, edited by Brian P. McLaughlin, Oxford: BasilBlackwell, pp.119-130.
- Dretske, Fred. 1988, *Explaining Behavior: Reasons in a World of Causes*, Cambridge, MA: MIT Press.
- Millikan, Ruth G. 1993, *White Queen Psychology and Other Essays for Alice*. Cambridge, Mass: MIT Press.
- Summerfield, Donna and Manfredi, Pat. 1998, "Indeterminacy in Recent Theories of Content," *Minds and Machines*, 8, pp.181-202.
- 勝亦佑磨 (2018) 「誤表象問題—学習に基づく目的論的意味論の検討—」、哲学・科学史論叢、第二十号。