

Studies in Critical Rationalism

2025

Vol. 16, No. 2

日本ポパー哲学研究会事務局機関誌編集部

(2025年12月号)

CONTENTS

<第35回年次研究大会自由論題報告論文>

- プロセスのダーウィン主義：存在論から心の相互作用説を導くポパーの試み 池田 健人 2

<第35回年次研究大会シンポジウム報告論文>

- | | | |
|-------------------------|-------|----|
| 心的因果・再考 | 太田 雅子 | 30 |
| 人工意識の創出による意識の神経メカニズムの解明 | 渡邊 正峰 | 38 |
| デフレ化によるポパーの三世界論の延命措置？ | 鈴木 大地 | 40 |

<新刊紹介>

- 小河原誠著『政治のロジックを問いただす—批判的合理主義の観点から—』 富塚 嘉一 55

<新刊情報>

- カール・ポパー著、蔭山泰之訳『出発点 私の知的発展』 57

<第35回年次研究大会自由論題報告論文>

プロセスのダーウィン主義：存在論から心の相互作用説を導くポパーの試み

池田健人（大阪大学）

1. 序論

本論文では、心の哲学においてポパーが提唱した相互作用説の背後には「プロセスのダーウィン主義」があることを示す。「プロセスのダーウィン主義」は、本論文に特有の用語だが、モノ（実体）ではなくプロセス（過程）に適用されるダーウィン主義を指すものとして導入されている。ポパーは、プロセスにダーウィン主義を適用することから、独自の心身相互作用説を構築した。その心身相互作用説によれば、ダーウィン主義は心と身体の相互作用を支持する立場だが、物理的プロセスがダーウィン主義的に進化する過程で創発した心理的プロセスが心だとみなされているため、心と身体は相互作用している。

ポパーが心の哲学へ本格的に取り組み始めたのは、1949年にロンドン大学の教授となって数年が経過した 1952 年のことである（Boyd 2016）。心の問題を扱うとき、ふつう「多くの哲学者は、魔法のような力を認めるすべての体系を非常に問題のあるものだとみなす。それから距離を置くために、魂や社会制度などの非物質的な対象だけでなく、心的な対象 の存在もまた、彼らは激しく否定する」（Agassi 2021: 337）。こうした事情もあり、心と身体を区別する心身二元論は、これまで心の哲学において劣勢を強いられてきた。それにもかかわらず、20世紀も後半へと差し掛かったとき、突如ポパーが作業仮説として擁護の議論を展開し始めたのは、心

身二元論に与する相互作用説であった。相互作用説とは、心の哲学において、心と身体は相互作用するとみなす立場である。

物理主義を筆頭に、心と身体の相互作用を謳うポパーの仮説は、従来さまざまに批判されてきた。たとえば、その後景をなす形而上学が深刻な欠陥を抱えていることや、科学の成果が真摯に受け止められていないこと、あるいは正しく理解されていないことなどが、ポパーの相互作用説には指摘されている（Albert 1977; Keuth 2000; Rooijen 1987）。しかし、そうした批判は、必ずしも、すべてが的を射たものであるわけではない。本論文では、とくに心の主観性と実体性にかかわる批判を取り上げ、それらはポパーの意図を十分には汲んでいないことを示す。また、それを通じて、ポパーはプロセスのダーウィン主義をもとに心身相互作用説を提唱していたと主張する。

そのため、まず、本論文では、ポパーが発表した心身相互作用説の概略を示すことから始める。ポパーの心身相互作用説は、生物学と言語論の会遇を通じて成立した。より具体的にいえば、言語の四機能説と呼ばれるポパーの仮説を、生物学に属するダーウィン主義の視座から解釈することにより、ポパーは相互作用する心と身体の概念へ想到することとなる。ダーウィン主義的な進化論は、意味内容と密接に結びつく言語の叙述機能と論証機能を考慮するなら、言語と人間の相互作用を強く示唆するものだとポパーはいう。このとき、ポパーは、その相互作用を補助するシステムに心を見出す。ポパーにとって、心とは、人間の水準でいえば、意味内容をも

つ言語と、それを利用する人間のあいだに介在する、創造と制御の相互作用的なシステムである（第2節）。

心と身体の相互作用を擁護するポパーの立場には、心の主観性を論点とする批判が提出されている。この批判によれば、ポパーは主観性を定義しなければならない。これは、相互作用という関係が、十分に具体化された事物についてのみ記述されうことと関係している。心の場合、主観性が定義されることで、はじめて十分に具体化される。とはいっても、定義とは言語化のことだが、心を言語化することはできない。それゆえ、この批判によれば、ポパーの心身相互作用説は不可能である。しかし、本論文では、このような批判は、心身相互作用の存在を明らかにしようとするタイプの心身相互作用説には当てはまらない、と主張する。ポパーの心身相互作用説は、ダーウィン主義を梃子にして心身相互作用の存在を訴えるものである。存在論としての心身相互作用説は、主観性を定義するかどうかにかかわらず構成されることができる（第3節）。

とはいっても、それでもなお残されるのが、心の実体性を論点とする批判である。この批判によれば、およそ状態なるものは、すべて何らかの実体がもつ状態であり、状態そのものとして存在しているわけではない。すると、心的状態が存在するとき、そうした状態にある何らかの実体もまた存在しているということになる。ところが、ポパーは、その心身相互作用説における心的状態は、いったいどのような実体の状態かを特定していない。それゆえ、この批判によれば、ポパーの心身相互作用説は不完全である。しかし、本論文では、このような批判は、ポパーの実体論を追及しなければならない、と主張する。ポパーによれば、現代科学の進境からして、もはや実体とは見かけ上のものにすぎないといえる。ポパーは、世界の基底は過程だとするプロセス存在論を採用した。このように、実体

はポパーの心身相互作用説から取れて締め出されている（第4節）。

以上により、本論文では、ポパーの心身相互作用説は、ダーウィン主義的に進化するプロセスを考察の対象とした仮説だと主張する。ポパーによれば、心は物質からダーウィン主義的に進化する。それゆえ、心と物質、とりわけ生物の身体は相互作用の関係にある。いいかえれば、宇宙には心と身体の相互作用が存在している。相互作用している心と身体は、ポパーによれば、心的状態と物的状態である。ポパーは、唯物論が歩んできた歴史を検討するなかで、ただ状態それ自体として存在する心的状態と物的状態の概念へと辿り着いた。これらはプロセスである。プロセスとしての心と身体が存在している、とポパーはいう。心身相互作用とプロセスの存在論から、ポパーは独自の心身相互作用説を導いた。

2. 生物学に逢着する言語論

みずからの研究者としての人生を振り返るなかで、ポパーは、「身心問題との最初の遭遇は、わたしをして長年のあいだ、これはとても解決の望みのない問題だという感じを抱かせてきた」（Popper 1976 [2002]: 219）と述懐している。もともと、ポパーは、心身問題の議論に意欲的でなかった。しかし、ポパーが学生のときに見えていた言語の四機能説が、のちに生物学のダーウィン主義と出会い、図らずも、ポパーを心身問題の議論へと誘うこととなる。言語の四機能説とは、人間言語に備わる4つの機能についての仮説である。ポパーによれば、人間言語は、助言的機能や獎励的機能、虚構的機能など、多くの機能があるなかでも、とくに主要な4つの機能をもっている。それは、基礎的なものから順に、表出機能（徵候機能）、信号機能（刺激機能、または伝達機能、解発機能）、叙述機能（報知機能）、論証機能（説明機能、または批判機能）である。

これら4つの機能について、それぞれ簡単に説明しておこう。表出機能とは、話し手の内的な状態を外的に露呈させる機能である。わたしたちは、何かを話すとき、たとえば声の調子や抑揚、音質、大小などによって、つねに内的状態を表出している。信号機能とは、聞き手の内的または外的な反応を惹起する機能である。たとえば、誰かが「危険だ！」と叫ぶとき、その発声は、話し手の緊迫した内的状態を表出するだけでなく、その発声を耳にした聞き手を驚かせたり（内的な反応）、場合によっては走らせたり（外的な反応）するかもしれない。このとき、言語は、ある種の信号として機能している。表出機能と信号機能は、すべての動物言語に共通の機能である。それに対して、叙述機能と論証機能は、完全なものに限定するのなら、ただ人間言語にのみ備わっている。

叙述機能とは、現実的または想像的な事態を描写する機能である。たとえば、わたしたちは、たんに「危険だ！」ではなく、「オオカミがいる！危険だ！」ということができる。このとき、わたしたちは、命題などと呼ばれるかたちで、オオカミの存在という、この世界についての現実的または想像的な、この世界に対応していれば真であり、対応していないければ偽であるような事態を叙述している。論証機能とは、命題などのようないくつかの事実についての言明を、真理の観点から批判的に吟味する機能である。言語は、ある言明の虚偽を論証するための手段としても利用することができる。これら人間言語に特有の叙述機能と論証機能は、合わせて言語の高次機能という。

人間の心を分析するうえでは、言語の高次機能が重要な役割を果たす。その重要性は、言語の高次機能が、言語の意味内容と結びついている点にある。わたしたちは、言語に高次機能を発展させることで、高次機能を伴う言語（高次言語）を獲得したことにより、「理論を定式化したり表現したりするさまざま

な様式を抽象化し、その不变の内容または意味に注意を払うすべを学んだ」（Popper 1972 [1979]: 240、強調は原著者）。ある理論は、必ず何らかの現実的または想像的な事態を指示するものである。こうした事態は、ひとたび叙述されると、不变の意味内容として叙述のうちに保存される。

意味内容は物理的なものではないので、ポパーによれば、意味内容と結びついた高次言語を物理的な概念のみで完全に説明することはできず、「人間言語についての物理主義的な因果理論は成り立たない」（Popper 1963 [2002]: 395）。高次言語は、意味内容と結びついている点で物理主義を超えており、わたしたち人間と高次言語の関係は、物理主義の枠内では解明されえない、というのがポパーの主張である。

では、人間と高次言語、たとえば理論の関係を、わたしたちはいったいどのようなものとして考えればよいのだろうか。ここで、ポパーは生物学へと議論を展開する。なかでも、ポパーが目をつけたのは、ダーウィン主義的な進化論である。ポパーは、ダーウィン主義の視座から言語の四機能説を捉え、高次言語（理論）の地位を解釈することで、人間と理論の関係を明らかにしようと試みた。

ダーウィン主義とは、ひとことで言えば、突然変異と自然淘汰のメカニズムからなる進化論を支持する立場のことである。ふつう、これらメカニズムのうち、突然変異は生物の身体内で生じる。たとえば、ダーウィンフィンチの嘴や、オオシモフリエダシャクの色は、生物の身体内で生じる突然変異の事例として有名である。

前者から順に、それぞれ簡単に説明しておこう。ダーウィンフィンチとは、ガラパゴス諸島、およびココス諸島で適応放散した14種のフィンチ類の総称であり、特定の種類の名前ではない。適応放散とは、多様な形態や生態をもつ生物種が、共通の祖先から比較的に短時間で分出する現象のことである。ダーウ

インフィンチと呼ばれる鳥たちは、それぞれ食物が異なり、それに応じて嘴の形状も異なる。これは、祖先の嘴に生じた突然変異が、特定の食物に適合するよう自然選択されたことの結果だと考えられている。ポパーは、進化が起こる順序について、まず選好構造（preference）が変化し、ついで技能構造（skill）、最後に身体構造（anatomy）に変化が生じる、という「 $p \rightarrow s \rightarrow a$ 」の図式を提案していたが、この図式の正しさはダーウィンフィンチの研究（Lack 1947）によって裏づけられている、と述べている（Popper 1976 [2002]）。

オオシモフリエダシャクは、日本から欧州にかけて生息する蛾の名前である。進化論の文脈では、工業暗化という現象の記述において、しばしばオオシモフリエダシャクが言及される。ポパー（e.g. Popper 1982; 1994）もまた、この蛾の工業暗化に言及している。工業暗化とは、都市の工業化に伴って煤煙により樹皮が黒ずんでくると、そうした樹木からなる環境に住む生物の色が、明色から暗色に変化する現象のことである。とくに、この術語は、産業革命を経験したイギリスで、オオシモフリエダシャクの体色が変化した現象を指して用いられる。本来、オオシモフリエダシャクは明色の蛾だったが、黒ずんだ樹皮のもとで明色は非常に目立ち、天敵に見つかりやすくなるため、環境に不適合である。このとき、そうした環境では保護色となる暗色に突然変異した個体が自然選択され、それまで主流に属していた明色の個体に代わり、繁栄することとなった。

太平洋とインド洋の島々で暮らすダーウィンフィンチや、イギリスの山中に暮らすオオシモフリエダシャクの事例は、いずれも突然変異の証拠だとみなされている。ダーウィンフィンチは食物に合わせて嘴の形状を変化させ、オオシモフリエダシャクは煤煙のため黒ずんだ樹皮に合わせて体の色彩を変化させた。このように、通常、嘴の形状であつ

たり、体の色彩であつたり、突然変異は生物の身体内で生じる。それに対して、突然変異は必ずしも生物の身体内でのみ生じるものではない、とポパーはいう。ポパーによれば、高次言語を操作することにより、人間という生物の水準においては、これまで身体外にも突然変異が生じさせられてきた。これを示すためには、ポパーの認識論、とくに主観的知識と客観的知識の区別から説き起こさなければならない。

まず、主観的知識とは、個々の認識主体に固有の主観的状態、すなわち行動ないし行為への傾向性（性向）である。たとえば、わたしたちの信念や意見、期待、心の状態などのようななかたちで意識されるようになる性向が、主観的知識にあたる。それに対して、客観的知識とは、高次言語により定式化されることを通じて、わたしたちの主観的な性向を超越した知識である。「知識または思考は2つの異なる意味において存在している ……

(1) 心または意識の状態、あるいは行動ないし行為への性向からなる、主観的意味における知識または思考と、(2) 問題、理論、議論そのものからなる、客観的意味における知識または思考がそれである」（Popper 1972 [1979]: 108-9、強調は原著者）。客観的知識の典型例として、問題や議論とともに、ポパーは理論を挙げている。

たとえば、数学の理論を考えてみよう。いまだ数学の理論には、ゴールドバッハやリーマン、コラッツなどにより提出されてきた多くの伝統的な問題が存在する。これら問題の最終的な解決までをも含めて、数学の理論を完全に主観的知識として構成できる人間は誰もいないし、あらゆる人間の主観的知識を集めてきたとしても、そこに完全な数学の理論が構成されることはない。その意味で、数学の理論は、わたしたちの主観を超えており、わたしたちの身体外に存在する客観的知識である。もちろん、わたしたちの主観は、数学の理論を部分的に身体化しているが、数学

の理論そのものは身体化されうるものではない。これは、数学の理論にかぎらず、すべての理論がもつ性質である。それゆえ、ポパーによれば、主観的知識と客観的知識は明確に区別されなければならない。

このような認識論を基盤として、ポパーは客観的知識の成長という概念の考察に取り組んだ。ポパーによれば、客観的知識は、ある問題に対する暫定的解決の案出とその解決に含まれる誤謬の批判的排除を通じて、ある問題から別の問題へ到達する過程を繰り返すことにより成長していく。客観的知識が成長する過程を、ポパーは「 $P_1 \rightarrow TS \rightarrow EE \rightarrow P_2$ 」の図式にまとめた。 P_1 (Problem 1) は最初の問題(ある問題)、TS (Tentative Solution) は暫定的解決、EE (Evaluative Error Elimination) は誤謬の評価的排除、 P_2 (Problem 2) は別の問題、をあらわしている。ポパーにとって、「つねに評価は批判的なものである」(Popper 1972 [1979]: 144)。この過程が、生物がダーウィン主義的な方法で進化する過程に酷似している、という点にポパーは目をつけた。

暫定的解決は客観的知識であり、人間身体の外部に存在している点で突然変異とは異なるものの、暫定的解決と批判的排除は、それぞれダーウィン主義における突然変異と自然淘汰に相当するものだとポパーはいう。わたしたちは、暫定的解決を「わたしたちの皮膚に外在する突然変異のようなものとして、すなわち「身体外的突然変異」と呼ばれるものとしてみなすことができる」(Popper 1994: 13、強調は原著者)。

突然変異とは、ふつう身体の部分、身体器官にみられるものなので、理論をもとに構成される身体外的突然変異は、身体外的器官だともいえる。これは、言語による記述のほか、とりわけ人工物のかたちで、道具として、わたしたちの物理環境に出現することとなる。たとえば、ポパーによれば、わたしたちは目という身体器官を進化させる代わりに双眼

鏡や眼鏡を、耳という身体器官を進化させる代わりにマイクロフォンやスピーカー、補聴器を、足という身体器官を進化させる代わりに自動車を、それぞれ発達させてきた。これらの道具はすべて、人間が目や耳、足にかんして直面した問題に対する暫定的解決である。さまざまな道具が生み出され、また廃れてきた現在までの歴史を振り返れば、よりよい解決が将来的に案出されるのなら、さきに列挙したような現存の解決もまた廃れ、淘汰されていく運命にあることは想像に難くない。

高次言語によってのみ表現されうる理論は、ポパーによれば、さまざまな人間の道具と同じように、身体外的器官だとみなすことができる。「理論は、わたしたちが身体外に発達させる器官である」(Popper 1968: 163)。それゆえ、身体外的突然変異、そのひとつに理論は数え上げられる、とポパーは推測した。

ダーウィン主義では、基本的に、ある突然変異は、それが自然淘汰の脅威に耐えているのなら、そのかぎりにおいて、その突然変異を生じさせたところの主体に対して、何らかの有益な制御を及ぼしているとされている (Popper 1978; Richards 1987)。してみると、もし理論が、身体外的なものではあるにせよ、わたしたちによって創造された突然変異であるのだとすれば、それが自然からの淘汰圧に耐えているかぎり、理論は、それを産出した人間に対して、ある何らかの有益な制御を及ぼしていると予想できる。このとき、人間と理論のあいだには、創造と制御からなる一箇の系、システムが発見される。ポパーは、このシステムが心、より正確には心的状態だと考えた。すなわち、人間と心の関係は創造と制御からなる相互作用の関係だということになる。このように、言語論と生物学の邂逅から導出されたのが、ポパーの心身相互作用説である。

3. ダーウィン主義からのアプローチ

以上のように、心と身体の関係は相互作用だとするポパーの立場には、ダーウィン主義的な進化論が大きく影響している。このことを確認したうえで、以下では、ポパーの心身相互作用説に対する批判の検討へと移ろう。本論文では、冒頭でも述べたように、これまでポパーの心身相互作用説にはさまざまな批判が向けられてきたなかでも、とりわけ心の主観性と実体性に関連した批判を取り上げる。まず、本節では、心の主観性に着目してポパーの心身相互作用説に欠陥を指摘する批判へと焦点を合わせる。

ムンツ (Munz 2007) によれば、心がもつ主観性を定義することは、ポパーの心身相互作用説にかぎらず、あらゆる形態の心身相互作用説が成立するための要件である。主観性を定義しないかぎり、どのように心は身体へ作用し、また身体は心へ作用しているのか、十分に明らかにすることはできない、とムンツはいう。なぜなら、主観性は、こうした心と身体の相互作用を分析するためには、どうしても避けて通ることができない、いわば心の本性だからである。それにもかかわらず、ポパーの心身相互作用説において、心の主観性は定義されていない。ムンツにいわせれば、そのようなポパーの仮説は、仮説というよりも、むしろ放言に等しいものであり、なんとなく心（意識）は身体と相互作用している気がして、このことは疑いえない、という感想を表明しているだけのものである。

意識についてのポパー自身による考えは、思いつきのもので、ほとんど偶然的なものであり、これからわたしが示そうとしているように、修正されなければならないものである。彼は、意識の問題それ自体には決して取り組まなかったが、まるでそれ以上の説明は何も必要とされていないかのように、片手間で意識に

言及していた。彼はまた、心と身体のあいだの相互作用というアイデアについて、「心」によって彼が意味していたものに対する具体的な定義を与えることなく、相互作用が生じることはまるで当然のこととみなされうるかのように話していた (Munz 2007: 307-8)。

相互作用という関係は、一般に、そもそも具体的な事物のあいだでのみ説明されることができる。それゆえ、何らかの相互作用が主張されるときはいつでも、そこで相互作用が主張されているところの2つの事物は、それらのあいだの相互作用が主張されうるほど十分に具体化されていなければならない。さもなければ、当然のことだが、それら事物のあいだの相互作用は、主張するにもしようがないのである。心身相互作用説の場合、その仮説では心と身体の相互作用が主張されている。それが正当な仮説だと認められるためには、心と身体の両者が十分に具体化されていなければならない。このとき、心が十分に具体化されているといえるためには、少なくとも心の主観性を定義することが必要だ、とムンツはいう。なぜなら、主観性は心の本性だからである。ムンツによれば、主観性を定義することなしに、心を十分に具体化したということはできない。

そうすると、主観性を定義することによって、ポパーには自身の心身相互作用説を補完することが求められている、ということができる。しかし、ムンツによれば、主観性を定義することは、原理的にまったくできない。というのも、主観性とは非言語性であり、それを言語的に表現することは本質的に不可能だからである。ふつう、定義は言語によって与えられるものだが、言語化できないものを定義することはできない、とムンツは考えている。「身体内の主観性は主観的なものであり、それゆえ純粹に私的な言語であらわされることができないので、明確に定義できる

ものではない」(Munz 2007: 308)。この批判は、ムンツの議論にしたがえば、ポパーの心身相互作用説だけでなく、すべての心身相互作用説に当てはまる。心身相互作用説は、そもそも一般に成立しえないのであり、それゆえポパーの心身相互作用説も成立しえない、とムンツはいう。

部分的にみれば、ムンツ自身は言及しないが、ムンツと同じ議論は、これまで別の研究者によって過去にも提出されてきた。たとえば、相互作用は具体的な事物に対してのみ適用されうるという点については、ブンゲ(Bunge 1980)が同じ議論を提出している。「デカルトからポパーとエクルズに至るまで、二元論の相互作用論的な変版は、物質と心が—それらは分離しており、また異質なものではあるが—互いに作用し合うとするものである。相互作用は具体的な事物に対してのみうまく定義されるため、この教説はほとんど詳しく説明されることができない」(Bunge 1980: 83)。

また、主観性は言語化されえないという点については、例示するまでもなく、いまや心の哲学では常識だといってよいかもしれないが、たとえばスヴェンソン(Svensson 1994)が同じ議論を提出している。もっとも、スヴェンソンは「主観性」ではなく「クオリア」という言葉を使用しているが、ここで主観性とクオリアは同義に解釈することができる。「あらゆる言語使用は、経験それ自体が本質的にもつ内的、質的側面の捨象を含んでいる。わたしたちは「クオリア」という用語によって、言語では説明できない、あるいは言語的な用語で捉えることもできない、まさにそのような経験の側面を示そうとしている」(Svensson 1994: 308)。このように、ムンツの議論は、それ自体には含まれていない、別の研究者による過去の議論を通じて補強されることができる。

しかし、各論の多くには妥当な見解が含まれているとしても、主観性の定義を心身相互

作用説の必要条件だというムンツの議論には飛躍がある、と本論文では主張する。それは、心身相互作用説の適切な分析に失敗している。すべての心身相互作用説において、それが成立するために主観性の定義が前提となるわけではない。したがって、主観性を定義しないかぎり、いかなる心身相互作用説も成立しえない、というムンツのテーゼは誤りである。この誤りが原因で、ムンツは、ポパーの心身相互作用説を正しく評価できていない。心身相互作用説には、ムンツの主張に反して、主観性の定義を成立の要件にするものとしないもの、すなわち心身相互作用の認識論と存在論があり、ポパーの心身相互作用説は後者である。本節の残りの部分では、このことについて説明しよう。

ムンツの議論では、心身相互作用の認識論と存在論が混同されている。心身相互作用の認識論とは、心身相互作用のモデルにかんする知識の議論であり、心と身体が相互作用するメカニズムの解明を謳う議論である。こうした議論の提供を目的とした仮説は、心と身体はいったいどのように相互作用するのかを詳述しなければならない。他方、心身相互作用の存在論とは、心身相互作用の存在を唱える議論である。これは、心と身体はいったいどのような関係にあるのかを探求するものであり、心と身体が相互作用するメカニズムを解明しようとするものではない。心身相互作用の認識論は、心の存在だけでなく、心と身体の関係は相互作用だということまでを仮定する議論であるに対して、心身相互作用の存在論で仮定されているのは心の存在だけである。

たしかに、主観性は定義されえないというムンツの指摘が正しいのなら、心身相互作用のメカニズムを解明することはできない。このような状況は、心身相互作用の認識論にとって深刻である。心身相互作用のメカニズムは、いつも概念的に、それゆえ言語を用いて解明される。ここで、主観性が定義されえな

いということは、主観性を本性にもつ心が定義されえない、つまり心は言語を用いて表現されえないということである。すると、心について、言語を用いて解明することはできない。心身相互作用は、心が身体へ作用し、また身体から作用される現象であるが、心の主観性が定義されえないのなら、心身相互作用のメカニズムを解明することは、言語を用いて表現できない心について、それが身体へどのように作用するのか、また身体から作用されるのかを、言語を用いて解明することである。こうした作業が達成されることはない。

他方、心身相互作用の存在論にとって、主観性が定義されないことは必ずしも問題とはならない。心身相互作用の認識論とは異なり、こちらは心身相互作用それ自体の存在に照準するものである。存在すると主張されているところの心身相互作用が、いったいどのようなメカニズムで作動しているのかという点までは、心身相互作用の存在論のなかで明らかにされる必要はない。心身相互作用のメカニズムにかんする問題は、心身相互作用の存在にかんする問題から分離できる。前者は後者に回答しなければならないが、前者へ対処しなければ後者は検討できないわけではない。とくに、前者は解決できないとしても、後者は正当に主張されることができる。なぜなら、現実に心と身体は相互作用しているが、そのメカニズムは認識できない、ということはありうるからである。

ひとくちに心身相互作用説とはいっても、そこには認識論と存在論の少なくとも2種類が区別できるのにもかかわらず、これをムンツは看過している。心身相互作用説は一般に不可能だという帰結は、主観性が定義されないことから直接には導出されない。ムンツの議論は、心身相互作用の認識論に対してのみ効果を發揮しうる。だが、心身相互作用の存在論に対して、それは正当なものだと考えられない。ムンツが、自身の議論を心身

相互作用の認識論だけでなく、存在論の非難にまで有効だとみなしているのなら、それは過剰であり、話を不当に単純化しすぎている。

ポパーの心身相互作用説は心身相互作用の存在論なので、それがムンツの議論により棄却されることはない。それは、主観性を本性にもつ心それ自体ではなく、心と身体の相互作用という客観的な事態の存在を主張する仮説である。心身相互作用説を提唱することによって、ポパーは心身相互作用のメカニズム、つまり心と身体はいったいどのように相互作用しているのかを明らかにしようとしているのではない。ポパーが主張しているのは、心が存在することは認めたうえで、心と身体の関係は、平行でなければ随伴でもなく、そのほかでもない相互作用だということである。

公平を期するために、ポパーが心身相互作用のメカニズムについて意見を述べることは、指摘しておかなければならない。ポパーによれば、心身相互作用は、無意識に相当する電磁場、すなわち無意識的な物的力場（心的力場）と、意識が相互作用することによって生じる。この相互作用において、その端緒をなす概念は注意である。主体に内的な現在の生理的状況とは無関係の抽象的事柄に向けられる注意、つまり本能的注意に対置されるところの意識的注意、または理性的注意が要請される問題を解決しなければならないとき、いいかえれば理性的注意という心的な働きを通じて、意識は脳内に生起する電磁場、ひいては身体と相互作用を開始することができる（Popper et al. 1993）。さらに、電磁場の役割もポパーは推測している。

電磁場の役割は、ポパーによれば、脳内で生成する量子効果（突然変異）への淘汰圧として機能することである。意識（心）に対する大脳（身体）の開放性を、ポパーは「突然変異へ淘汰圧が衝突するのと同じ方向で……理解する。突然変異それ自体は量子効果、つまり波動として考えることができる。

そのような波動は、たとえば大脳で起こるだろう。大脳では、まず純粋に確率的な、または無秩序な変化が生じ、これら波動のあるものは、自然淘汰が疑似目的的に突然変異を選択する波動に似た仕方で …… 選択されるだろう」(Popper & Eccles 1977 [1983]: 540–1)。このようにして、意識からの作用をもとに、いわば電磁場は大脳に局在する量子状態を刻み込む。

明らかに、ポパーは心身相互作用のメカニズムを検討している。しかし、ポパー自身、上述の見解が満足のいくものだとは考えていない。ポパーは、その見解が非常に漠然としたものであり、いまだ十分には精錬されていないことを進んで認めている (Popper & Eccles 1977 [1983]; Popper et al. 1993)。こうした、ほとんど根拠のない脆弱な見解を、ポパーの心身相互作用説に含めることは難しいかもしれない。その見解は、心的力場仮説と呼ばれる事はあるにせよ、主に神経科学の分野で、仮説にまで引き上げられようとしているところである (e.g. Århem 2021; Libet 1994; 1996; 1997; 2006; Lindahl & Århem 1994)。とはいって、これは、その見解をポパーの心身相互作用説とみなすことを妨げるものではない。しかし、かりにそうだとしても、心の主観性に訴えてポパーの心身相互作用説を完全に否定しようとするのは、あまりにも粗笨である。

たとえ心身相互作用のメカニズムが原理的に解明されえないのだとしても、それによって心身相互作用の存在が否定されることはない。心身相互作用の認識論からは独立に、ポパーは心身相互作用の存在論を主張することができる。認識論は存在論に後続しうるが、存在論から認識論への移行は必然ではない。もしくは、存在論は認識論から独立に主張されることができる。ムンツのように、すべての心身相互作用説が認識論を伴うと考えるのは間違いである。主観性の定義は、心身相互作用説が成立しうるための普遍的な

要件ではない。それにもかかわらず、主観性が定義されていないことを理由に、ムンツはポパーの心身相互作用説を棄却しようとしている。

むしろ、主観性の定義を考慮しなくてよい点こそ、ポパーの心身相互作用説がもつ利点である。上述したように、ポパーによれば、わたしたち人間と高次言語、たとえば理論のあいだには、あるひとつのシステム、創造と制御の相互作用的なシステムが作動している。このシステムが、ポパーの心身相互作用説における心、より正確には心的状態である。そうすると、おのずから、理論と心だけでなく、人間と心のあいだの関係もまた相互作用だということになる。このような仕方で、心身相互作用のメカニズムというよりは、存在を主張しようとするのがポパーの心身相互作用説だった。ポパーは、人間と言語の関係を考察することから、その延長線上に、存在論として心身相互作用説を展開している。こうした心身相互作用説の長所は、後段、心とは何かという問題、すなわち心の本性、主観性の問題を回避したうえで心身相互作用説を擁護できる点にある。

進化論を応用することで、ポパーは心と身体の関係はどのようなものかという問題に答えようとした。とくに、ポパーが注目したのは、ダーウィン主義的な進化論である。ダーウィン主義的な「進化論は …… 心または意識の進化について、またそれによって心の機能について、あることを述べることにより（「心」または「意識」とは何かを語ることなく）問題を解決する」 (Popper 1972 [1979]: 250)。ポパーによれば、ダーウィン主義に含まれる「自然淘汰説は …… 心と身体の相互作用に賛成する強力な論証を供与する」 (Popper 1978: 350) ものであり、ダーウィン主義的な進化論は、心と身体の関係を相互作用として解決した。ムンツは、ポパーに対して心とは何かを説明するよう求めているが、まさに心とは何かを説明しなくとも心と

身体の関係は相互作用だといえると示していることこそが、ポパーによる心身相互作用説の楔子である。

4. プロセスとしての心

主観性を根拠とする批判とは別に、もうひとつ、実体性を根拠とする批判がある。ムンツが提出した批判の検討を通じて、ポパーの心身相互作用説はダーウィン主義にもとづく心身相互作用の存在論であり、かりに主観性が心の本性であったとしても、ポパーの心身相互作用説において主観性が定義される必要はない、ということがわかった。しかし、たとえそうだとしても、ポパーの心身相互作用説には、ダーウィン主義的な説明と両立する批判が依然として残されている。そのひとつが、実体性を根拠とする批判である。本節では、この批判について検討しよう。

ブンゲ (Bunge 1980) によれば、心の実体性にかんする記述が見当たらない点で、ポパーが提唱する心の理論には不備がある。ポパーが心と身体の関係は相互作用だとする理論を支持しているのに対して、ブンゲは創発唯物論を支持している。中枢神経系と、そのサブシステムの多様な創発的機能にかんする詳細な理論、とくに数学的理論を探求するためのプログラム的な仮説にすぎないとはいえ、ブンゲにとって、「創発唯物論的な心の哲学は、わたしたちがもっている最高のものである」(Bunge 1980: 24)。ブンゲが支持する創発唯物論において、心は物的状態により実現される機能的状態だとみなされている。ここで、物的状態とは、神経という物的実体の状態である。

創発唯物論を採用することにより、ブンゲは物的実体の状態が実現する機能的状態として心の概念を考察している。「わたしたちは、神経や神経系といった具体的なもののが状態や機能（過程）について、一般的かつ正確な考えをもっている」(Bunge 1980: 4) が、それに対して、心という抽象的なものの状態

や過程、つまり機能についての一般的かつ正確な考えはもっていない。このことから、「心とは 個々の神経がもっていない、神経群がもつ機能の集合である」(Bunge 1977: 508) とブンゲは帰結する。ブンゲにとって、心とは物的実体の状態（物的状態）により実現される機能的状態である。他方、ポパーは、機能的状態の担い手を心だとみなしている。

ポパーによれば、心とは適応的な機能的状態を実現する心的状態である。ダーウィン主義の影響により、ポパーは心の機能へ注目している。「もし主観的経験、つまり意識的状態が存在するのなら ダーウィン主義によれば、わたしたちはその利用、またその適応的な機能に注目しなければならない」(Popper 1978: 350)。適応的な機能は心それ自体ではなく、ポパーにとって、心という状態がもつものである。いいかえれば、適応という機能的状態を実現している状態が心である、とポパーは考えた。こうした心の概念は、神経群の有機的な集合、すなわち神経系（物的実体）からなる状態（物的状態）によって実現される機能的状態が心だとするブンゲの主張と相容れない。

物的状態により実現される機能的状態が心だとするブンゲに対して、ポパーは、機能的状態を実現している心的状態が心だとしている。しかし、物的状態と心的状態のどちらが心に関連する機能的状態を実現するのか、または何らかの状態により実現されている機能的状態が心なのか、その機能的状態を実現しているところの何らかの状態が心なのか、などについてブンゲは争わない。そうではなく、ブンゲは、ポパーがいう心的状態の存在論的地位を問題としている。

実体性の観点からブンゲがポパーを批判する目的は、いかなる状態も、つねに何らかの実体の状態としてのみ存在しうる、というテーゼを訴えることにある。心としての機能的状態を実現する物的状態は、物的実体の状態なので、もとをただせば心は実体であり、

心の本性は実体である、とブンゲは主張している。このような理論をブンゲが提出する背景には、状態とは実体の性質であり、状態とされるものは、すべて必ず何らかの実体として存在している、というテーゼがある。

およそ状態と呼ばれるものは、どのようなものも、みな具体的なもの（実体）の性質である、とブンゲはいう。物的状態でいえば、どのような状態も、必ず物的実体の状態として存在している。たとえば、水に帰属される物的状態は、いつも水分子という物的実体の状態である。ブンゲによれば、物的状態をはじめ、あらゆる状態は、状態それ自体として存在しているわけではない。「物理学から生物学、社会学に至るまで、すべての科学において、性質は具体的なもの に備わっており、事象とはある性質の変化である（わたしたちはもちろん、抽象的対象の性質ではなく、実体的性質とその変化について書いている）」(Bunge 1980: 20)。ここで、「性質」は「状態」に同義である。実体に備わっている性質としての状態の変化が事象であり、心的状態や心的事象（心的状態の変化）さえ、実体から分離される必要はない、とブンゲは述べている。

すべての状態が実体の状態であるのなら、心的状態もまた、もし本当に存在しているのだとすれば、ある何らかの実体、いわば心的実体の状態でなければならない。心的状態の変化は、その状態にある実体を対象とすることによってのみ調査できる、というのがブンゲの見立てである。心身相互作用説は、心的状態の存在を前提とする仮説だが、たんなる幻想ではない、仮説としての地位を認められるためには、心的状態が存在しているというのなら、ブンゲによれば、その心的状態にあるところの心的実体とはいいたいどのようなものかが詳細に提示されなければならない。

ところが、心的状態は存在するといいながら、あえて進んでは心的実体に言及しないの

が心と身体の相互作用を支持する相互作用主義者の常套手段である、とブンゲはいう。物的実体に基づきされた心を主張する場合、心的状態にある実体は明快である。たとえば、ブンゲにとって、その実体は神経である。それに対して、心的実体の特定は非常に難しい。なぜなら、心的実体は誰にも発見されたことがないからである。相互作用主義者は、みずからの理論に措定されている心がもつ実体としての側面、心の実体性を、いまだ捕捉することができていない。そこで、ふつう相互作用主義者がとる戦略は、ブンゲによれば、心の実体性を隠蔽するというものである。

別言すれば、それは、実体性のある心を実体性のないもの、すなわち心的状態そのものとして記述することだといえる。相互作用主義者は、本気で心に実体性がないとは考えていないが、真理への接近を試みるのではなく、ひとたび心の実体性が話題になると、実体を非実体として語る自説の矛盾が露呈してしまうため、こうした事態の回避に汲々している、とブンゲは考えている。「「実体」を実体性の欠落したものとしてみなす理由を誰もが彼に尋ねることができるので、彼はもはや敢えて思惟実体、すなわち心的実体の状態については話さないだろう」(Bunge 1981: 141、強調は原著者)。このようにして、心という実体性のあるものを、それにもかかわらず意図的に実体性のないものとして記述する心身相互作用説は詭弁だとブンゲはいう。

心的状態のみを相互作用主義者が議論の対象とするのは、ブンゲによれば、心が非実体として存在しているからではなく、たんに心的実体を特定できないからであり、いわば自説の弱点を隠蔽するためのレトリックである。ひとたび正直に心的実体などのようなものを話題にしようものなら、ただちに相互作用主義者は、なぜ実体を非実体として語るのか、という疑問を解消する必要に迫られることとなる。しかし、このとき、もちろん相

互作用主義者としては、まかり間違っても心的実体とはどのようなものか見当もつかないからだ、などと白状するわけにはいかない。このような事情から、相互作用主義者は、暗黙に仮設されている心的実体からは論点を逸らしたところで、空疎な心的状態の概念をひたすら弄することにばかり集中している、とブンゲは非難している。

ほかの相互作用主義者と同じように、ポパーもまた、その心身相互作用説において心的状態へは言及しておきながら、心的実体へは言及していない。ポパーの心身相互作用説は、ダーウィン主義的な進化論からの帰結として、心（心的状態）と身体（物的状態）の相互作用を主張する仮説であった。この仮説は、物的状態が理論を創造し、理論から制御される関係の存在を指摘したうえで、こうした創造と制御のシステムに心を位置づける。ポパーによれば、理論の創造と物的状態の制御に介在する実体ではなく、状態が心である。それゆえ、ポパーは心を心的状態という。

ここで、心的状態という表現は、ポパーによって自覚的に選択されている。「物理的な対象や状態のほかに心的状態というものがあり、その状態はわたしたちの身体と互いに相互作用する」（Popper & Eccles 1977 [1983]: 36、強調は原著者）。心的状態が存在していることに、ポパーは積極的な意義を認めている。「わたしは主観的経験、心的状態、知性、精神の存在を否定するものではない。これらはきわめて重要なものだとさえ、わたしは信じている」（Popper 1976 [2002]: 159–60）。他方、ポパーが、心的実体とはどのようなものかについて見解を示すことはない。

ブンゲによれば、こうしたポパーの態度は有害である。心的状態について語るのなら、ポパーは心的実体を詳解できなければならない、とブンゲはいう。なぜなら、心的実体へ言及しないからといって、心的状態が心的実体の状態でなくなるわけではないからである。事実上、「心は身体と相互作用するにも

かかわらず身体からは独立の対象または実体である、という古代のテーゼに賛成する議論をポパーは模索している」（Bunge 1981: 140）。心的状態について語ることは、それが物的実体の状態でないのなら、ブンゲにとって、そのまま心的実体の存在を主張することである。もしブンゲの見立てが正しければ、心的実体とはいったいどのようなものなのかを明示することのないままに、いくら心的状態が存在するといっても、ただそれを繰り返すばかりでは空虚である。

科学の観点から、ブンゲはポパーの心身相互作用説を糾弾する。ブンゲがいう科学の存在論からして、「実体のない心は存在しない」

（Bunge 1977: 508）。これは、いかなる実体もなしに、それ自体として存在する心的状態はありえないということである。ブンゲがいう科学の存在論によれば、ただ実体のみが存在する。そうであるのなら、心身相互作用説を擁護するためには、上述のように、ポパーは心的状態の成立に不可欠な心的実体を詳解できるのでなければならない。しかし、ポパーはその義務を怠っており、ポパーの心身相互作用説において心的実体は不明のままである、とブンゲはいう。

たしかに、ポパーは心的状態についてのみ言及し、心的実体については何も述べないと指摘する点で、ブンゲの議論は正しい。ブンゲが指摘しているように、ポパーもまた、ひとりの相互作用主義者として、心的実体ではなく心的状態にのみ照準を合わせて独自の考察を提示している。しかし、それは心的実体という暗黙の前提を公表しなければならない状況に直面するのが不都合だからではなく、心的実体なるものはまったく存在しない、とポパーは考えていたからである。心的実体の存在を完全に消去することは、ブンゲがいう典型的な相互作用主義者の姿勢に当てはまらない。この背景には、唯物論をめぐる科学の歴史を紐解くなかで、最終的にポパーが辿り着いた実体論がある。ブンゲの批判

は、ポパーの実体論を適切に踏まえたものではない、と本論文では主張する。

ポパーは、実体の概念を否定している。とくに、心の実体性を拒絶することは、ポパーの心身相互作用説をブンゲが批判する以前から、たびたび明言されてきた。たとえば、ポパーは自身がデカルト主義者であることを自認していると述べたあとで、「わたしは精神実体とか思惟実体などといったデカルト的な言述に—彼の物質実体とか延長実体などと同じほど—賛成しない。物的状態と心的状態を信じているかぎりにおいてのみ、わたしはデカルト主義者である」(Popper 1972 [1979]: 231、強調は原著者) と付言している。こうした例は、ほかにもみられる。

ある場所で、ポパーは「わたしは（「実体」について語るべきだとは決して考えなかつたけれども）つねにデカルト的な二元論者であった」(Popper 1976 [2002]: 218) とか、また別の場所では「もちろん、相互作用の教説がまったく古めかしいものであることは重々承知している。それにもかかわらず、わたしは相互作用と古めかしい二元論を（ただし、わたしは「実体」と呼ばれるものの存在は拒絶するけれども）擁護することを提案する」(Popper 1978: 351、強調は原著者) などと、ポパーは発言している。ポパーは、心的実体について皆目に見当がつかないとの自白へ追い込まれることを危惧しているのではない。むしろ、ポパーは心的実体の非存在を積極的に主張している。

さらにいえば、心的実体のみならず物的実体も、それゆえ一般に実体と呼ばれるものが存在することを、ポパーは否定している。心だけが特別に実体をもたないのではなく、「実体という観念そのものが誤りにもとづいている」(Popper & Eccles 1977 [1983]: 105)。長いあいだ科学、とりわけ唯物論において中枢的な役割を果たしてきた実体という概念は、物理学の発展を通じて克服された、というのがポパーの見解である。

物質は、保存されないゆえに「実体」ではない。それは、破壊されうるし、また創造されうるのである。……物質は高度につまつたエネルギーで、エネルギーの他の形式に変換可能であり、したがって光、そしてもちろん運動、熱などのような他の過程に変換できるゆえに、過程という本性をもつものである。……／このように現代物理学の成果は、実体や本質の観念を捨て去るべきであることを示唆しているといえる。それらの成果は、（いくらかの物質は「日常的な」状態のもとでもそうであるが）時間上のあらゆる変化を通じて持続する自己同一的なものが存在しないことを、そして対象の性質や属性を持続して担い、保持する本質は存在しないことを示唆している。宇宙はいまや対象の集まりではなく、事象や過程の相互作用する集まりのように思われる。……／それゆえ、現代物理学者が物理的対象一物体、物質一は原子構造をもっていると述べるのはもっともなことであろうが、原子はそれ自身の構造をもっており、その構造を「物質的なもの」として描くことはほとんどできないし、まして「実体的なもの」としてはまったく描けない。結局、物質の構造を説明するプログラムによって、物理学は唯物論を超越しなければならなかつたのである (Popper & Eccles 1977 [1983]: 7、強調は原著者)。

実体を拒否する代わりに、過程だけが存在するという哲学をポパーは採用している。光や運動、熱などと同じように、ポパーによれば、物質もまたエネルギーの形式である。日常的な状態のもとでは、どのような物質も、すべてあたかも究極的なモノ（実体）であるかのようにみえるが、ポパーにとって、そのような実体としての振る舞いはあくまでも

見かけ上のものであるにすぎず、実際のところ物質とは、その微視的な構造までをも精密に観察してみると、さまざまな形式へと絶えず変化しつづけるエネルギーの局面的なプロセス（過程）だということが判明する。こうした議論にしたがうのなら、ブンゲは、物質が見せる表面上の振る舞いに惑わされている。ポパーによれば、ただプロセスのみからなる構造こそ、より忠実に世界の相貌をあらわしている。

状態、すわなちプロセスとしての心だけが存在している、と主張するのがポパーの心身相互作用説である。ニーニルオートは、ポパーの哲学は「常識や科学と両立する動的なプロセス存在論に相当する」(Niiniluoto 2006: 60、強調は原著者)と述べている。ブンゲは具体的なもの（実体）とその状態（過程）を区別していたが、ポパーにいわせれば、世界に充実するすべてのものはみなプロセス、過程であり、実体などまったく存在しない。こうしたプロセス存在論を基盤として、ポパーの心身相互作用説は提唱されている。その要諦は、ひとつには、モノ（実体）ではなく、ただプロセス（過程）としての心だけを存在のカテゴリーとみなす点にある。

プロセス存在論を無視することは、ポパーの心身相互作用説に対する批判の価値を著しく低下させる。まさに、そのことで、ブンゲの批判は重要性を損なわれている。ポパーによれば、「生命は、モノではなくプロセスである。わたしたちは、わたしたち自身をモノとして考えすぎている。わたしたちは、モノではなくプロセスである。…… 生命は高分子ではない。生命は明らかにプロセスである。それは、進行中のプロセスである」(Popper et al. 1993: 169)。こうした背景をもつポパーの心身相互作用説において、心の実体性は完全に放棄されている。心的実体から衆目を逸らすのがポパーの意図であるのではない。ポパーは、心的実体なしに存在する心的状態を主張している。心の哲学においてポパーの

実体論が標的とされ、プロセス存在論が排撃されるとき、実体と状態を切離するブンゲの批判は、はじめて有意義なものとなる。

5. 結論

本論文では、心の主觀性と実体性に根拠を求める2つの批判を検討することを通じて、ポパーの心身相互作用説が基礎としている2つの存在論を明らかにした。まず、心の主觀性に着目する批判は不適切である。心身相互作用のメカニズムを提示することは、ポパーが心身相互作用説を唱える際の目的ではない。ポパーの心身相互作用説は、ダーウィン主義にもとづく心身相互作用の存在論である。つぎに、心の実体性に着目する批判は不十分である。いまや実体の概念は誤りであり、プロセスの概念が認められなければならない、とポパーは考えている。ポパーの心身相互作用説は、プロセス存在論にもとづく心の仮説である。これら2つの存在論を総合すると、プロセスにダーウィン主義を適用することで、それゆえプロセスの進化論から、導かれたポパーの心身相互作用説が浮かび上がる。ポパーが提唱した心身相互作用説の背後には、プロセスのダーウィン主義がある。

【一次文献】

- Popper, K. 1963 (2002). *Conjectures and Refutations: The Growth of Scientific Knowledge*, London and New York: Routledge Classics. (『推測と反駁：科学的知識の発展』、藤本隆志・石垣壽郎・森博訳、法政大学出版局、1980年)。
- Popper, K. 1968. 'Is There an Epistemological Problem of Perception?', *Problems in the Philosophy of Science*, ed. Lakatos, I. and Musgrave, A., pp. 163–4, Amsterdam: North-Holland Publishing Company.
- Popper, K. 1972 (1979). *Objective*

- Knowledge: An Evolutionary Approach*, Oxford: Clarendon Press. (『客観的知識：進化論的アプローチ』、森博訳、木鐸社、1972年)。
- Popper, K. 1976 (2002). *Unended Quest: An Intellectual Autobiography*, London and New York: Routledge Classics. (『果てしなき探求：知的自伝』、森博訳、岩波書店、1978年)。
- Popper, K. 1978. 'Natural Selection and the Emergence of Mind', *Dialectica* 32(3/4): 339–55.
- Popper, K. 1982. 'The Place of Mind in Nature', *Mind in Nature: Nobel Conference XVII Gustavus Adolphus College, St. Peter, Minnesota*, ed. Elvee, R. Q., pp. 31–59, Harper & Row.
- Popper, K. 1994. *Knowledge and the Body-Mind Problem: In Defence of Interaction*, ed. Notturno, M. A., London and New York: Routledge.
- Popper, K. & Eccles, J. C. 1977 (1983). *The Self and Its Brain*, London and New York: Routledge. (『自我と脳（上・下）』、西脇与作・大村裕訳、思索社、1986年)。
- Popper, K., Lindahl, B. I. B. & Århem, P. 1993. 'A Discussion of the Mind-Brain Problem', *Theoretical Medicine*, 14(2): 167–80.
- 二次文献**
- Agassi, J. 2021. 'The Place of Mind in Nature', *Karl Popper's Science and Philosophy*, ed. Parusniková, Z. and Merritt, D., pp. 337–43, Springer.
- Albert, H. 1977. *Kritische Vernunft und menschliche Praxis: Mit einer Autobiographischen Einleitung*, Stuttgart: Philipp Reclam.
- Århem, P. 2021. 'Popper on the Mind-Brain Relation', *Karl Popper's Science and Philosophy*, ed. Parusniková, Z. and Merritt, D., pp. 279–94, Springer.
- Boyd, B. 2016. 'Popper's World 3: Origins, Progress, and Import', *Philosophy of the Social Science* 46(3): 221–41.
- Bunge, M. 1977. 'Emergence and the Mind', *Neuroscience*, 2(4): 501–9.
- Bunge, M. 1980. *The Mind-Body Problem: A Psychobiological Approach*, Pergamon Press.
- Bunge, M. 1981. *Scientific Materialism*, London, England: D. Reidel.
- Keuth, H. 2000. *Die Philosophie Karl Poppers*, Tübingen: Mohr Siebeck.
- Lack, D. 1947. *Darwin's finches*, Cambridge University Press.
- Libet, B. 1994. 'A Testable Field Theory of Mind-Brain Interaction', *Journal of Consciousness Studies* 1(1): 119–26.
- Libet, B. 1996. 'Conscious Mind as a Field', *Journal of Theoretical Biology* 2(178): 223–4.
- Libet, B. 1997. 'Conscious Mind as a Force Field: A Reply to Lindahl & Århem', *Journal of Theoretical Biology* 185(1): 137–8.
- Libet, B. 2006. 'Reflections on the Interaction of the Mind and Brain', *Progress in Neurobiology* 78(3–5), 322–6.
- Lindahl, B. I. B. & Århem, P. 1994. 'Mind as a Force Field: Comments on a New Interactionistic Hypothesis', *Journal of Theoretical Biology* 171(1): 111–22.
- Munz, P. 2007. 'The Phenomenon of Consciousness from a Popperian Perspective', *Consciousness Transitions: Phylogenetic, Ontogenetic, and Physiological Aspects*, ed. Århem, P. and Liljenström, H., pp. 307–26, Elsevier Science & Technology.
- Niiniluoto, I. 2006. 'World 3: A Critical

- Defence', *Karl Popper: A Centenary Assessment, Volume II: Metaphysics and Epistemology*, ed. I. Jarvie, K. Milford and D. Miller, pp. 59–69, College Publications.
- Richards, R. J. 1987. *Darwin and the Emergence of Evolutionary Theories of Mind and Behavior*, University of Chicago Press.
- Rooijen, J. V. 1987. 'Interactionism and Evolution: A Critique of Popper', *The British Journal for the Philosophy of Science*, 38(1): 87–92.
- Svensson, G. 1994. 'Reflections on the Problem of Identifying Mind and Brain', *Journal of Theoretical Biology* 171(1): 93–100.

【謝辞】

このたびは、研究報告の機会をいただきありがとうございました。お忙しいところ、わたしの拙い研究報告のために貴重なお時間を割いてくださるばかりか、その研究報告へ熱心に耳を傾けてくださった会員のみなさまに、あらためて深く感謝を申し上げます。ありがたいことに、多くのご指摘やご助言、ご質問をいただきました。そのなかでも、以下では、いただいたご質問を、それに対するご回答とともに記述します。なお、ご回答には、夏の研究大会よりもあとで新たに考えたり、調べたりした内容も含めました。いただいたご質問を網羅することや、ご質問の意図を誤りなく正確に再現することはできていないかもしれません、また十分なご回答を提示することもおそらくはできておらず、たいへん恐縮ですが、これらの点につきましては、どうかご容赦いただけますと幸いです。

【質問】

なぜ一元論ではなく、ポパーは二元論を採用したのか？

【回答】

二元論という表現は、ポパー自身もまた用いており、ポパーの立場を二元論だとすることに問題はないと思います。身体（物質）と心の実在を主張したという意味で、ポパーの立場は二元論です。ただ、最近では、厳密には非二元論だという点が無視されているために、しばしばポパーの心身相互作用説は不正確に宣伝されてきたと指摘することから、非二元論としてのポパーの立場を強調する議論も提出されている(Chakrabarty 2021)ことは、念のため、ご紹介しておきます。たしかに、主観的な心から客観的な心を区別できたとき、はじめて心身相互作用は理解できる、というのがポパーの主張です(Popper 1994a)。ここでは、こうした最近の議論も踏まえつつ、なぜ物質と心の両方が実在するという二元論をポパーは提唱したのかについて説明します。

大まかにいって、すべてのものは物質か心のどちらかへ完全に還元できる、というのが一元論と呼ばれる立場です。それゆえ、かなり広く捉えるのなら、一元論には2種類のものがあります。それは、すなわち、ただ物質だけが実在するという物的一元論（唯物論、または急進的物理主義）と、ただ心だけが実在するという心的一元論（唯心論、ないし観念論、独我論）です。これらの立場に反対し、ポパーが二元論を擁護する理由は、道徳論と方法論の観点から主に2つあります。

ひとつは、わたしたちが経験する主観的な事実、とりわけ苦しみ(suffering)が、一元論では適切に扱われえない、というものです。ポパーによれば、一元論は、実在する苦しみを幻想として語ることしかできません。苦しみを幻想として片づけようとする態度は、ポパーにとって、道徳に悖るものです。ポパーは、広島と長崎で起きた悲劇が、たんに幻想とされてしまうようなことがあっては決してならないと述べ、心は実在しないとする立

場を強く非難しています (Popper 1976 [2002]; 1982b; 1984)。

もうひとつは、わたしたちにとって二元論は明白であり、常識だが、この常識を拒否する程度までには、いまだ一元論は至っていない、というものです。とりわけ、物的一元論に対する抵抗を示す場面で、ポパーは、明白な二元論の主張を展開しています。ポパーによれば、人間の行為、つまり物的状態の変化を、いかなる心的状態も介在させることなく理解することは不可能です (Popper & Eccles 1977 [1983])。物的一元論の主張は、ポパーにとって、もし心がなければ、あるいはもし心があったとしても、心について沈黙を貫いていさえすれば、宇宙は非常に単純な場所となるだろう、というものであるにすぎず、それは二元論を適切に反駁できていません (Popper 1978)。ポパーは、常識から出発し、その常識を反駁する方法の意義を訴えています。

哲学や科学は、ポパーによれば、すべて啓蒙された常識です (Popper 1972 [1979])。わたしたちは、常識の判定を、少なくともその常識が覆されてしまうまでは、なぜ受け入れないのか、とポパーは述べています。ポパーにとって、ことが真に興味深くなるのは、まさに常識が誤っているときです (Popper 1976 [2002])。それゆえ、わたしたちは常識を批判しなければなりません。

常識を批判するためには、いったんは、その常識を引き受ける必要があります。というのも、常識を徹底的に批判し、反駁するためには、その常識に通曉していなければならぬからです。ポパーによれば、常識が本当に誤っているのかを真剣に考慮することなく、たんに常識から離れたとしても、そこに興味深いものは何もありません。まずは、もっとも強力なかたちで常識を表現し、そのあとで常識を批判することこそが重要である、とポパーは考えています (Popper 1972 [1979])。

ところが、現状、物的一元論は、あまりに

も簡単に二元論という常識を放棄しようとしてしまっている、というのがポパーの見解です (cf. Popper 1976 [2002])。少し話は逸れますが、急進的物理主義(物的一元論)を、汎心論や随伴現象論、同一説など、稳健な物理主義へ置換したとしても、ポパーにとって、状況は変わりません。心の実在は認めて、心と身体の相互作用は認めないのが、稳健な物理主義です。これに対して、ポパーによれば、心と身体の相互作用は常識であり、わたしたちが悪趣味な学者で、この明白な相互作用を哲学的なものとしてしまおうとしたいきぎり、わたしたちは心と身体が相互作用することを知っています (Popper et al. 1993; Popper 1994a)。ポパーにとって、一元論を含む物理主義は、粗雑な主張で二元論を退けようとする不十分な立場です。

ただ二元論の慎重な検討のあとでのみ、二元論の誤りは判定できるのであり、もし二元論が誤りであったのなら、そのときにはじめて、一元論の批判は真に興味深いものとなる、とポパーは述べていました。ところが、ポパーによれば、一元論は二元論を適切に反駁できません。ポパーが二元論を採用したのは、一元論により提出される、二元論に対する批判は弱すぎると考えていたからです。

しかし、そうすると、一元論を支持する人々にとっては一元論が常識なのではないか、ということになります。明らかに、常識に訴える2つめの理由は、一元論がおかしいということを指摘するものではありません。この理由によって説示されうるのは、あくまでも、ポパーが二元論を採用したのは、ポパーが二元論を常識だと考えていたからである、ということです。もしポパーが二元論を常識だと考えていなかったのなら、そのとき、ポパーは二元論を採用していなかったと思われます。

【質問】

水と波動の相互作用は物理学の枠内で問

題なく論じられるが、そのようなものとして身体と心の関係は理解できないのか？

【回答】

志向性を伴うものであるがゆえに、心には身体からの自律性がある、とポパーは述べています (Popper et al. 1993)。これに対して、波動は、志向性どころか、水からの自律性をさえ備えたものではないため、おそらく、この点で、水と波動の関係と同じように身体と心の関係を考察することはできない、というのがポパーの見解です。

とはいっても、ポパーは志向性という概念の内実をほとんど説明していないので、その内実を詳らかにすることを通じて、ポパーが意図していた心は、身体に対して、じつのところ水に対する波動のような存在としてみなされうるものだった、ということがわかる可能性は、決して小さいとは断言できません。ポパー自身、志向性ということで、何か特殊な概念を想定していたわけではないようです。今後、志向性という概念の内実を一般的に解明していく作業が重要となるかもしれません。

【質問】

人間が発達させたとされている身体外的器官は、ダーウィン主義的な突然変異というよりも、むしろラマルク主義的な獲得形質に近いのではないか？

【回答】

身体外的器官は理論的構成物である、という点から説明するのがもっともスムーズかもしれません。理論的構成物なので、身体外的器官が発達する過程は、知識が成長する過程として理解することができます。知識が成長する過程は、試行錯誤の過程です。ここで、もっとも重要な論点は、知識の成長は、空のバケツが水で満たされていく過程のような、集積の過程ではないということでした

(Popper 1972 [1979])。ポパーによれば、バケツが蛇口から水を注がれるように、感覚から入ってくるものを吸収する仕方で知識は成長するのではありません。まずは大胆な仮説を案出し、それを批判的に排除することで、知識は成長していく、というのがポパーの見解です。

なるほど、ある個体が生涯のうちに取り組む継続的な努力の結果として、ある種の選好や願望などが影響しつつ、後天的に生み出されるため、身体外的器官は、一見したところダーウィン主義的な突然変異というよりも、ラマルク主義的な獲得形質に近いものであるようにみえます。しかし、重要なのは、同じ継続的な努力でも、ダーウィン主義とラマルク主義では、そうした努力の内実が異なる点です。

ラマルク主義的な獲得形質は、帰納的で反復的な努力から生み出されます。他方、演繹的で推測的な努力と、評価的で批判的な努力から生み出されるのが、ダーウィン主義的な突然変異であり、人間の身体外的器官です

(Popper 1976 [2002])。たび重なる試行錯誤を経たのち、ある程度まで納得のいったところで、さしあたり変化をストップしているのが、暫定的解決としての現存の身体外的器官だといえます。身体外的器官は、実証主義的（ラマルク主義的）ではなく、より欠点の少ない、よりよい身体外的器官を目指して、絶え間なく文化的淘汰の圧力にさらされており、むしろ反証主義的（ダーウィン主義的）に発展させられようとしています。それが生じてくるところの過程からして、身体外的器官は突然変異である、とするのがポパーの判断です。

もっとも、それだけではなく、生物が後天的な努力によって獲得できるレベルの変異は遺伝しないとみなす点でも、ダーウィン主義とラマルク主義は異なる、とポパーは指摘しています。上述のように、ある種の選好や願望などの要素に影響される身体外的器官

は、一見したところ、ラマルク主義的な獲得形質であるようにみえても、ダーウィン主義的な突然変異であるようにみえません。しかし、よくある一般的な理解に反して、それはダーウィン主義とラマルク主義の差異を示すものではない、というのがポパーの見解です (Popper 1976 [2002]; 1989 [1992])。ポパーは、ダーウィン主義とラマルク主義の相違は、幸運と狡智の相違ではない、と断言しています (Popper 1976 [2002])。つまり、選好や願望などの要素に影響されるところまでは、ダーウィン主義はラマルク主義をシミュレートしているのです (e.g. Popper 1972 [1979])。

ダーウィン主義がラマルク主義と袂を分かつのは、獲得形質は遺伝するのに対して、突然変異は遺伝子に生じるもの除去で遺伝しない、と表明されるときである、とポパーは考えています (cf. Popper & Eccles 1977 [1983])。遺伝子に生じる突然変異は、生物の後天的な努力によってではなく、先天的にもたらされるものです。ダーウィン主義では、選好や願望などの要素が、遺伝子における突然変異の傾向に影響を与えることはあっても、そのまま遺伝することはありません。身体外的なものであるがゆえに、身体外的器官は遺伝的基盤をもつものではない、とポパーは述べています (Popper 1994a)。ポパーによれば、身体外的器官は、世代を超えて継承されうるものではあっても、遺伝しうるものではありません。この点で、ポパーは身体外的器官を突然変異とみなしています。

【質問】

熱力学の法則について、ポパーはどのように考えているか？

【回答】

心身相互作用説に対しては、熱力学の第一法則（エネルギー保存則）や第二法則（エントロピー増大則）に抵触するのではないか、

という有名な批判があります。こうした伝統的な批判に対するポパーの見解は、ひとことでいえば、これらの法則は必ずしも心身相互作用説の妨げにはならない、というものです。

エネルギー保存則について、まず、ポパーは、それが統計的にのみ妥当な法則である可能性を指摘しています (Popper & Eccles 1977 [1983])。もしそうであれば、かりに心身相互作用を通じてエネルギー量の微小な変化が生じたとしても、その程度の誤差は許容されうるので。また、ポパーは、エネルギー保存則は閉鎖系においてのみ成り立つ、という主張も展開しています。そうすると、ポパーによれば、生物は開放系なので、生物についてエネルギー保存則は成り立たちません。あるいは、ポパーは、電気化学的なエネルギーへ変換できる、純粹に精神的なエネルギーが存在する可能性にも言及しています。とはいっても、ポパー自身、それほどこの可能性を重視していないようです。そのほか、ポパーによれば、パイロット波についての解釈には、エネルギー的なプロセスに与えられる非エネルギー的な影響の可能性が示唆されています (Popper 1984)。

エントロピー増大則について、ポパーは、それほど大きな問題だとは考えていないようです。ポパーによれば、心からの作用を受けて、熱放出に等しい疲労が大脑に生じると仮定するのなら、それはエネルギーが低下することに等しい事象であり、エントロピー増大則は守られます (Popper & Eccles 1977 [1983])。また、ポパーは、たんにエントロピー増大則はブラウン運動により破られた、とも述べています (Popper 1984)。さらに、ポパーが提示するのは、エントロピー増大則は開放系には当てはまらない、という議論です。ポパーは、平衡状態からかけ離れた状態にある開放系は、エントロピーを産出するとしても、無秩序を増大させる傾向を示すことはなく、それどころか、こうした系はエントロピーを周囲の環境へ運び出し、みずからの系の

内的秩序を減少させるどころか、増加させることさえできる、というプリゴジンの説明に注目しています (Popper 1982a [1988])。

【質問】

エントロピーやネゲントロピーとの関係でポパーは心を論じているか？

【回答】

管見では、心をめぐるポパーの主張において、エントロピーやネゲントロピーと関係づけて心の特性を考察する議論はありません。既述のように、心身相互作用はエントロピー増大則に違反するのではないか、という懸念に対する応答のなかで、ポパーがエントロピーへ言及することはあります。とはいえ、それは、心身相互作用はエントロピー増大則を考慮しても成立することを示す議論であり、エントロピー増大則が本質的に、あるいは準本質的に心を規定していることを示す議論ではありません。

ついでながら、周知のように、とりわけポパーは時間論においてエントロピーを詳しく検討しています。たんなる異方性にとどまらず、一定の方向性が時間にあるとき、それは時間の矢が実在しているということですが、時間の矢は実在すると主張したうえで、時間の矢はエントロピー増大則から独立である、とポパーは述べています。時間の矢はエントロピー増大則に依存している、または決定されている、あるいは少なくともエントロピー増大則と関係している、という広く普及した見解に反して、時間の矢はエントロピー増大則と関係していない、というのがポパーの見解です (Popper 1956a; 1956b; 1967a; cf. 1957b)。エントロピーの減少は時間の逆進を意味するため、エントロピーが時間において減少することはないという見解は、ポパーによれば、熱力学的なゆらぎの存在を考慮すると、不合理なものだといわざるをえません (Popper 1957a; 1958)。

また、シュレーディンガーの提案を批判するかたちで、ポパーは生命論においてもエントロピーやネゲントロピーを検討しています。シュレーディンガーは、生物の特徴を考察するなかで、環境から秩序を絶えず吸収しつづけること、つまりネゲントロピーの摂取へ言及しました (Schrödinger 1944 [1992])。これに対して、ネゲントロピーの摂取は生物に固有の営為ではなく、たとえば熱機関などのような非生物もまたネゲントロピーを摂取する、とポパーは指摘しています (Popper 1967b; Popper 1976 [2002])。

薪が燃焼する過程を思い浮かべてみるとよいかもしれません。ふつう、薪が燃焼する過程は、エントロピーが供給される過程として説明されますが、ポパーにとって、それは、薪のネゲントロピーが摂取されたうえで、それを上回るエントロピーが供給されている、ということです。それゆえ、ポパーにとって、ネゲントロピーの摂取は、非生物にも当てはまることがあります。生物の特徴ではありません。むしろ、ポパーによれば、生物の特徴は摂取されるネゲントロピーを下回るエントロピーしか供給しない効率性にあります。このように、ネゲントロピーの摂取は、生命の保存に關係していないことはないが、生物の特徴とはいえない、というのがポパーの見解です。ちなみに、ポパーによれば、ネゲントロピーは生命の進化にも關係しています (Popper 1965)。

生命だけが、ポパーの主張において、エントロピーやネゲントロピーと結びつけられるものであるかもしれません。ポパーは、時間や心の理論へエントロピーやネゲントロピーを組み込むことには否定的な立場のようです。時間のうちに延長する心の存在が示唆されているとはいえる (Popper et al. 1993)、ポパーは時間をエントロピー増大則から独立だとみなしているため、エントロピーやネゲントロピーとの関係でポパーが心を論じていることは、おそらくないのではないかと

思います。

【質問】

心をめぐるポパーとペンローズの議論には、何か通じるものはあるか？

【回答】

基本的に、内容面で両者の議論が通じることはないようです。量子力学を心の理解へ適用することから、心は計算不可能な、非アルゴリズム的な要素をもっており、心をコンピューターで実現することはできない、と主張する点で両者の議論には通じるところがあり、このことは先行研究でも示唆されてきました（e.g. Markič 2021; cf. Århem & Liljenström 1997; 2007; Munz 1997; 2007）。しかし、心の計算不可能な、非アルゴリズム的な要素は、ポパーの議論において量子力学的なものではなく、ペンローズの議論において量子力学的なものである、という根本的な点で、両者の議論は異なります。それゆえ、両者の議論には表面上のつながりしかないとかもしれません。

ポパーは、たしかに心の理解へ量子力学を適用しました。たとえば、ポパーは、心の進化論的な創発は量子レベルの非決定性に由来するものだとみなしているほか、心が直接に作用するのは大脳で生じる突然変異としての量子効果だとみなしています（Popper & Eccles 1977 [1983]）。しかし、ポパーは、心それ自体がもつ計算不可能な、非アルゴリズム的な要素が、量子レベルの非決定性によりもたらされるとは考えていません。ポパーにとって、こうした要素を心がもっているのは、心は統語論的な規則だけでなく、意味論的な規則にもまた従って作動するからです（Popper et al. 1993）。ここには、言語の四機能説が関係しています。心は高次言語と相互作用するものなので、その振る舞いには意味論が大きく影響している、というのがポパーの見解です。これに対して、ペンローズは、

心（意識）それ自身を量子力学的なものとみなしています。

量子力学を意識そのものへ適用するのが、ペンローズの議論です。ある適切に組織されたシステムにおいて客観的な波動関数の収縮（objective reduction）が生じる場合、ペンローズによれば、そこに意識（consciousness）が生じます。このとき、その収縮が起こる場所として提案されているのは、大脳の神経細胞に内在する微小管（microtubule）です。微小管とは、チューブulin（tubulin）と呼ばれるタンパク質のサブユニットからなる中空の構成物のことですが、そこで生じる収縮は、ふつうのランダムな収縮ではなく、ある特殊なタンパク質によってオーケストラのように調整された収縮であり、決定論的ではあるが、計算不可能な、非アルゴリズム的なプロセスである、とペンローズは述べています（Hameroff & Penrose 1996; Penrose 1989; 1997）。

もっとも、心の哲学として、ポパーとペンローズの議論を皮相的な仕方で接続することはできなくはありません。たとえば、覚醒状態と睡眠状態を周期的に往来している場合、そこには意識（consciousness）がある、とポパーは推測し、どのようなときに意識は存在しているのかを考察しています（Popper & Eccles 1977 [1983]）、この考察と、ペンローズがいう客観的な波動関数の収縮による意識の生起は、あくまでも皮相的な面にのみ注目しているかぎり、両立します。しかし、後段、意識という概念に対するポパーとペンローズの見解を比較してみると、それらはほとんど別の概念です。それらの見解を単純に接続することはできない、ということがわかります。ペンローズによれば、ポパーがいう意識の概念は存在論的に大きな問題を抱えたものです。

三世界論というポパーの仮説に言及することで、ペンローズは、自身の哲学がポパーのものとは根本的に異なるということを明

示しています (Penrose 1997)。ポパーは、物質の世界（世界 1）から心や意識の世界（世界 2）が創発し、世界 2 から論理や意味の世界（世界 3）が創発する線形過程を謳う、三世界論を提唱しました。これに対して、ペンローズの世界 3 はプラトン的な世界、とくに絶対的な数学的真理の世界であり、ペンローズの理論が記述しているのは循環過程です。ポパーとペンローズの議論は、意識が必ず物質的基礎をもつ点で共通ですが、ペンローズが物質は数学的な存在であり、ある意味では心的な存在でさえあるというとき、決定的に相違します。なぜなら、ポパーによれば、数学や心は物質を基礎に創発するものだからです。両者の議論には通じる部分もあるかもしれません、少なくとも大きな隔たりがあります。

【質問】

ムンツはどのような代替案を提出しているか？

【回答】

じつのところ、心と身体の関係について、ムンツ (Munz 2007) が直接的な代替案を提示することはありません。その代わり、ムンツは、ポパーの三世界論を修正し、世界 1 は自然の世界、世界 2 は身体の世界、世界 3 は自律的で独立した思索の世界だとしたうえで、世界 2 と世界 3 の関係が相互作用ではないのなら、それはいったいどのようなものか、という問題に取り組んでいます。世界 2 と世界 3 の意味は、もはやポパーの用語法から大きく変更されている点が重要です。ムンツは、ムンツに特有の世界 3 が、身体（世界 2）とどのように関係しているのかを考察し、その結果として、ポパーが主張していたものとは異なる意識の概念へ至ることを目的としています。

身体と心ではなく、ポパーの三世界論における世界 1 と世界 2 でもなく、ムンツが新設

した三世界論における身体（世界 2）と世界 3 の関係について、自説を展開するのがムンツの議論です。世界 2 と世界 3 は、世界 2 が世界 3 から一方的に作用される関係にある、とムンツは述べています。ムンツによれば、ポパーがいう世界 2 と世界 3 だと、主観性が関与するため、それらの関係を特定することは不可能ですが、ムンツがいう世界 2 と世界 3 ならば困難はありません。身体内の現象であるソマティック・マーカー（世界 2）と、自律的で独立した思索である仮説（世界 3）から、心（意識）が生成されるメカニズムを検討することにより、ムンツは世界 2 と世界 3 の関係を考察しました。

意識 (consciousness) とは、ムンツによれば、わたしたちが意識しているものはどのようなものかを言述できる状態のことです。それゆえ、ある人が、意識の対象について、それがどのようなものか、言語を用いて述べられるなら、その人は特定の意識的な状態にあり、そのような状態としての意識をもっています。こうした意識の定義をもとに、言語と意識には密接で本質的なつながりがある、とムンツは推定しました。とりわけ、ムンツが注目するのは、言語のなかでも仮説です。ムンツによれば、ポパーの見解とは異なり、仮説と意識の関係は相互作用ではありません。

仮説は、ムンツにとって、世界 1 と世界 2 のどちらに由来するわけでもなければ、起因するわけでもなく、それらの世界から完全に分離された状態で浮遊しています。ムンツによれば、仮説の役割は、わたしたちに世界 1 や世界 2 で生じる事象を意識させることができます。つまり、意識的な状態にあるとき、わたしたちは仮説を用いて意識の対象がどのようなものかを述べるのであり、そのような仕方で、わたしたちが意識的な状態になることを仮説は助けることができます。このとき、仮説は、世界 2 へと作用するものだが、世界 2 から引き出されるものではなく、世界 2 についてのものである、というのがムンツの見解

です。それゆえ、ムンツによれば、世界2が仮説へと何らかの作用を及ぼすことはなく、仮説と世界2は相互作用するものではありません。ムンツにとって、仮説とは、一方的に世界2へ与えられ、世界2の事象を命名し、解釈するものです。

世界2と世界3の相互作用ではなく、世界3から世界2への一方的な作用だけがある、というムンツの主張において、わたしたちが仮説によって命名し、解釈するとみなされている世界2の事象は、ソマティック・マーカーです。世界3の仮説が世界1の神経まで到達することはなく、仮説は神経の興奮によって生成される世界2のソマティック・マーカーへ直接に作用している、とムンツは述べています。ソマティック・マーカーという概念は、前頭前野腹内側部を損傷した患者の分析を通じてダマシオが案出した、ソマティック・マーカー仮説のなかで提唱されました(Damasio 1991; 1994)。

ソマティック・マーカー仮説とは、人間の行動には、すべて情動が深く関与しているという仮説です。この仮説によれば、情動(emotion)は、人間が何かを実行するとき、それに先立つ行動選択プロセスの端緒で、あるいは選択肢から無自覚な行動選別をおこない、熟慮にもとづく自覚的な行動選択に事前のバイアスをかける仕方で、人間の意思決定に影響を与えています。情動とは、簡単にいえば、ある外部状況に直面するとき、その状況を刺激として身体状態(somatic state)に引き起こされる変化のことです。また、情動は、それが自覚的に感受されるとき、感情(feeling)となります。

外部状況の刺激により体性感覚パターンが賦活され、変化することで情動的な身体状態は形成されるが、その身体状態は、あるいは行動の選択肢について、その選択が主体にもたらす将来の結果は好ましいものか、それとも疎ましいものか、良いものか、悪いものか、あるいは安全か危険か、すなわち利害や

得失の観点からの価値を内包している、というのがダマシオの見立てです。ダマシオの予想では、このような身体状態の情報が身体信号として大脳へ伝達され、直観的感情(gut feeling)として経験されるとき、その身体状態は、ありうる行動の価値をひとつのイメージとして示すもの(marker)となります。それゆえ、ダマシオは、その身体信号、もしくは変容的な身体状態を、ソマティック・マーカーと名づけました。ソマティック・マーカーとは、情報として大脳へ伝達される身体状態、身体信号のことだといえます。

ダマシオは、それぞれのソマティック・マーカーによって喚起されうる感情は固定されていると考えていましたが、これをムンツは早計な思い込みだと指摘しています。ムンツによれば、わたしたちは、ただソマティック・マーカーについて仮説を立てることしかできません。同じソマティック・マーカーでも、それはいつでも歓喜や恐怖、憎悪、熱狂など、さまざまな感情としての自覚に対して開かれている、というのがムンツの見解です。世界2のソマティック・マーカーが、どのような感情として自覚され、言述され、意識されるかは、ムンツによれば、わたしたちが、世界3の仮説を用いて、そのソマティック・マーカーをどのように命名し、解釈するかにかかっています。

ムンツによれば、ソマティック・マーカーが命名され、解釈されるためには、それを可能とする言語、すなわち三次元言語(3-dimensional language)が導入されなければなりません。ある人が、世界3の三次元言語からなる仮説を用いて、ソマティック・マーカーを推測的に命名し、解釈したうえで、そのソマティック・マーカーはどのようなものかを述べることができるとき、その人は意識的な状態にあり、意識をもっている、とムンツは主張しました。このように、ムンツは、仮説とソマティック・マーカーの関係を考察することで、世界3から世界2への一方的な

作用を指摘しています。

【質問】

ポパーが、しばしば科学者から軽視されたり、無視されたりするのはなぜか？

【回答】

周知のように、言葉や概念の意味、あるいは定義の問題へと議論の焦点が逸れることを避けるため、ポパーは字義詮索へつながるような作業には積極的に関与を控えていました。しかし、この言語分析に抵抗する態度が、いまひとつポパーの議論を煮え切らないもののように見せてしまう原因となり、この点は科学者にかぎったことではないので、科学者がポパーを軽視したり、無視したりする理由としては弱いかもしれません、ひとつには科学者の反感を買ってしまったのかもしれません。

実際、ポパーの議論は、重要な言葉の意味が不明瞭であるために、しばしば十分に理解できないことがあります。たとえば、心身相互作用説との関連でいえば、心と身体は因果的に相互作用するとポパーは主張していますが、このとき「因果」という言葉は、いつたいどのような意味で用いられているのか、判然としません (Lindahl & Århem 1994; 2016; Markič 2021)。また、ポパーは心の創発を訴えていますが、その際に用いられる「創発」という言葉の意味は曖昧です (Markič 2021; Stephan 1992)。

より科学者と直接的に関係するところでいうなら、科学に対する理解が粗雑であり、過度に素朴であることが多い点も、ポパーが科学者から反感を買ってしまうひとつの原因かもしれません。たとえば、ポパーは進化論のダーウィン主義に改訂を提案していましたが、それはダーウィン主義に対するポパーの浅見からくるものであり、良くて不要、悪くて有害だと指摘されることがあります (Ruse 1977)。ポパーには、科学の成果を吟

味する際の慎重さが不足しており、科学者から科学に対する誠実さや真摯さを疑われ、しばしば相手にさえされることがないのかもしれません。

ときに、ポパーは、心の科学について現代の多くの研究者が真剣に考えている課題には取り組もうとせず、そのような課題を解決したり、そのような課題に対する既存の解決を反駁したりしようとしているのではなく、ほとんど誰も今日ではもっていないような陳腐な見解ばかりを、あたかも現今の通説であるかのように記述し、最先端の研究と対峙することを避けている、と指摘されることがあります。ポパーの議論には当時の主要な著作を読んだ形跡があり、ポパーが最先端の研究に接触していなかったとは考えにくいため、ポパーの議論では最先端の研究が歪められたかたちで理解され、批判されていることはあっても、ポパーが新しい文献に疎く、そのせいで細緻な配慮に値しない暗愚な研究者の新興勢力や、旧套墨守に徹する頑迷な研究者の因循勢力を論敵に設定てしまっているわけではないようである、と推察されることがあるようです (Dennett 1979)。

【質問】

ポパーは自由をどのように考えていたか？

【回答】

ポパーは、心身問題の背後には自由 (freedom) の問題がある、と考えていました (Popper 1994a)。自由という言葉は、ポパーが自由主義者 (liberal) という言葉を説明するときに出できます。ポパーにとって、自由主義者とは、個人の自由に価値を置き、あらゆる形態の権威に内在する危険に敏感な人物のことです (Popper 1963 [2002])。個人の自由ということで、ポパーは、いかなる形態の権威にも屈すことなく、つねに新たなフレームワークへの移行を求めつづけるこ

とを意味しています。ポパーによれば、こうした移行を求めつづけているかぎり、わたしたちは自由 (freeing) です。

わたしたちには、既存のフレームワークを批判し、破壊することで、そのフレームワークから「脱獄」し、また別のフレームワークへと進出する自由がある、とポパーは主張しています (Popper 1994b)。ポパーによれば、既存の権威に屈し、誤謬や偏見の牢獄 (フレームワーク) に囚われたままでいることは自由ではありません。あるいは、ポパーにとって、それは、自由であるのにもかかわらず、あえて自由の行使を拒み、自由ではない仕方で振る舞うということです (Popper 1982a [1988])。それでは、人間は、ただ内容を蓄積していくだけの録音機やバケツなどどこも変わることろがない、とポパーは述べています。ポパーにとって、純粹な物質である録音機やバケツに自由はありません。ある叙述や論証を批判的に検討し、評価することで、真理を把握しようとすることは、ポパーによれば、わたしたちの自発的で自由な行為によってのみ達成できます。

ただ、牢獄から脱出するという自由の理想は、それ自体、またひとつのフレームワークとして機能します。それゆえ、ポパーによれば、わたしたちにとって絶対的な自由というものはありません。あるのは、いつもフレームワークに相対的な自由のみである、というのがポパーの判断です (Popper 1994b)。

【質問】

ポパーは自我をどのように考えていたか？

【回答】

自我 (self/ego) とは、ひとことで言えば、他者 (他なるものすべて) との対比における自己、自分自身のことです。問題となるのは、どのように自我と他者を区別するかということですが、ポパーによれば、自我は、宇宙

で唯一の能動者、能動性であることにより、他者と区別されます。ポパーは、この区別に対する意識、とりわけ他者ではない自我についての意識を自我意識と呼んでいます。すると、このような自我や自我意識を備えているのは人間だけではありません。まず、能動的であるがゆえに、動物もまた自我です。また、ポパーによれば、覚醒状態と睡眠状態を周期的に往来するものであるがゆえに、自我意識を伴う低次の意識を動物は備えています。

他者からは離れて、自我に内的な話をするのなら、ポパーにとって、自我には意識的な部分（意識的自我）と無意識的な部分（無意識的自我）があるようです。明らかに、自我は意識に同定されるものではありません。なぜなら、意識は断続的なものであるのに対して、自我は連続的なものだからです。たとえば、わたしたちが睡眠中の無意識的なときでさえ、自我意識は存在していないとしても、自我は存在しつづけています。起きているあいだは自我であり、寝ているあいだは自我でない、ということはありません。意識的自我と無意識的自我のうち、ポパーが重視するのは意識的自我です。

ここまででは、人間の自我と動物の自我に共通ですが、ポパーによれば、人間は、意識的自我（自我のうち意識的な部分）の仕様が動物とは異なるために、自我意識（他者ではない自我についての意識）の程度もまた動物とは異なります。人間の意識的自我は、主観的知識の領域から客観的知識の領域にかけて広がっているものであり、客観的知識の領域に係留されている、というのがポパーの主張です。いいかえれば、人間の意識的自我は、たとえば理論などのような客観的知識にもとづいた思考をおこなうことができます。その意味で、人間の意識的自我は、ポパーにとって、思考する自我 (thinking self) です。これにより、生物のなかでも、ただ人間だけが、みずからの自我性を反省することができます。

る、とポパーは主張しています。ポパーによれば、動物の自我意識とは異なり、人間の自我意識は、他者のみならず、自我へもまた向けられるものです。

人間の子どもは、その子どもが生活する環境内の存在者、たとえば両親をはじめとする他我や、自然物、人工物との相互作用を通じて人間的な自我を発現させ、たんなる能動者ではなく、みずからが人格（personality）をもつ人間（person）であり、自我であることを学ぶ、とポパーは考えています。すなわち、わたしたちは、みずからの自我性を反省することのできる人間としては、自我として生まれ出るのではなく、自我であることを学ばなければなりません。このとき、重要な役割を果たすのは自我意識ですが、ポパーによれば、人間の水準において、自我意識は、人間言語との相互作用（話す能力）を通じて拡張されていきます（Popper 1994a; Popper & Eccles 1977 [1983]）。

一次文献

Popper, K. 1956a. 'Irreversibility and Mechanics', *Nature* 178(4529): 382.

Popper, K. 1956b. 'The Arrow of Time', *Nature* 177(4507): 538.

Popper, K. 1957a. 'Irreversibility; Or, Entropy since 1905', *The British Journal for the Philosophy of Science*, 8(30): 151–5.

Popper, K. 1957b. 'Reply to Hill and Grünbaum', *Nature* 179(4573): 1297.

Popper, K. 1958. 'Reply to Bosworth', *Nature* 181(4606): 402–3.

Popper, K. 1963 (2002). *Conjectures and Refutations: The Growth of Scientific Knowledge*, London and New York: Routledge Classics. (『推測と反駁：科学的知識の発展』、藤本隆志・石垣壽郎・森博訳、法政大学出版局、1980年)。

Popper, K. 1965. 'Time's Arrow and

Entropy', *Nature* 207(4994): 233–4.

Popper, K. 1967a. 'Structural Information and the Arrow of Time', *Nature* 214(5085): 322.

Popper, K. 1967b. 'Time's Arrow and Feeding on Negentropy', *Nature* 213(5073): 320.

Popper, K. 1972 (1979). *Objective Knowledge: An Evolutionary Approach*, Oxford: Clarendon Press. (『客観的知識：進化論的アプローチ』、森博訳、木鐸社、1972年)。

Popper, K. 1976 (2002). *Unended Quest: An Intellectual Autobiography*, London and New York: Routledge Classics. (『果てしなき探求：知的自伝』、森博訳、岩波書店、1978年)。

Popper, K. 1978. 'Natural Selection and the Emergence of Mind', *Dialectica* 32(3/4): 339–55.

Popper, K. 1982a (1988). *The Open Universe: An Argument for Indeterminism*, ed. Bartley, W. W., III., Routledge. (『開かれた宇宙：非決定論の擁護』、小河原誠・蔭山泰之訳、岩波書店、1999年)。

Popper, K. 1982b. *Quantum Theory and the Schism in Physics*, ed. Bartley, W. W., III., Routledge. (『量子論と物理学の分裂』、小河原誠・蔭山泰之・篠崎研二訳、岩波書店、2003年)。

Popper, K. 1984. 'Critical Remarks on the Knowledge of Lower and Higher Organisms, the So-Called Sensory Motor Systems'. *Experimental Brain Research 9: Sensory-Motor Integration in the Nervous System*, pp. 19–31, Springer.

Popper, K. 1989 (1992). *In Search of a Better World: Lectures and Essays from Thirty Years*, tr. Bennett, L. London and New York: Routledge. (『よりよき世界を

- 求めて』、小河原誠・蔭山泰之訳、未來社、1995年)。
- Popper, K. 1994a. *Knowledge and the Body-Mind Problem: In Defence of Interaction*, ed. Notturno, M. A., London and New York: Routledge.
- Popper, K. 1994b. *The Myth of the Framework: In Defence of Science and Rationality*, ed. Notturno, M. A., Routledge. (『フレームワークの神話』、ボーパー哲学研究会訳、未來社、1998年)。
- Popper, K. & Eccles, J. C. 1977 (1983). *The Self and Its Brain*, London and New York: Routledge. (『自我と脳(上・下)』、西脇与作・大村裕訳、思索社、1986年)。
- Popper, K., Lindahl, B. I. B. & Århem, P. 1993. 'A Discussion of the Mind-Brain Problem', *Theoretical Medicine*, 14(2): 167–80.
- 二次文献**
- Århem, P. & Liljenström, H. 1997. 'On the Coevolution of Cognition and Consciousness', *Journal of Theoretical Biology* 187(4): 601–12.
- Århem, P. & Liljenström, H. 2007. 'Beyond Cognition: On Consciousness Transitions', *Consciousness Transitions: Phylogenetic, Ontogenetic, and Physiological Aspects*, ed. Århem, P. and Liljenström, H., pp. 1–25, Elsevier Science & Technology.
- Chakrabarty, M. 2021. 'Karl Popper on the Evolution of Consciousness', *Karl Popper's Science and Philosophy*, ed. Parusniková, Z. and Merritt, D., pp. 295–319, Springer.
- Damasio, A. R. 1994. *Descartes' Error: Emotion, Reason, and the Human Brain*, Vintage. (『デカルトの誤り: 情動、理性、人間の脳』、田中三彦訳、筑摩書房、2010年)。
- Damasio, A. R., Tranel, D. & Damasio, H. C. 1991. 'Somatic Markers and the Guidance of Behavior: Theory and Preliminary Testing', *Frontal Lobe Function and Dysfunction*, ed. Levin, H. S., Eisenberg, H. M. & Benton, A. L., pp. 217–29, New York: Oxford University Press. (『ソマティック・マーカーと行動指針: 理論と予備的検証』、北川玲訳、『感情』、梅田聰・小嶋祥三監修、岩波書店、2020年、pp. 285–306)。
- Dennett, D. C. 1979. 'The Self and Its Brain', *The Journal of Philosophy*, 76(2): 91–7.
- Hameroff, S. R. & Penrose, R. 1996. 'Conscious Events as Orchestrated Space-time Selections', *Journal of Consciousness Studies*, 3(1): 36–53. (『意識は、マイクロチューブルにおける波動関数の収縮として起こる』、茂木健一郎訳、『ペンローズの量子脳理論』、徳間書店、1997年)。
- Lindahl, B. I. B. & Århem, P. 1994. 'Mind as a Force Field: Comments on a New Interactionistic Hypothesis', *Journal of Theoretical Biology* 171(1): 111–22.
- Lindahl, B. I. B. & Arhem, P. 2016. 'Consciousness and Neural Force Fields', *Journal of Consciousness Studies* 23(7–8): 228–53.
- Markič, O. 2021. 'Popper's Emergentism', *Karl Popper's Science and Philosophy*, ed. Parusniková, Z. and Merritt, D., pp. 321–35, Springer.
- Munz, P. 1997. 'The Evolution of Consciousness: Silent Neurons and Eloquent Mind', *Journal of Social and Evolutionary Systems* 20(4): vii–xxviii.
- Munz, P. 2007. 'The Phenomenon of Consciousness from a Popperian Perspective', *Consciousness Transitions*:

- Phylogenetic, Ontogenetic, and Physiological Aspects*, ed. Århem, P. and Liljenström, H., pp. 307–26, Elsevier Science & Technology.
- Penrose, R. 1989. *The Emperor's New Mind: Concerning Computers, Minds, and the Laws of Physics*, Oxford; Tokyo: Oxford University Press. (『皇帝の新しい心』、林一訳、みすず書房、1994年)。
- Penrose, R., Shimony, A., Cartwright, N., & Hawking, S. 1997. *The Large, the Small and the Human Mind*, ed. Longair, M., New York: Cambridge University Press. (『心は量子で語れるか』、中村和幸訳、講談社、1998年)。
- Ruse, M. 1977. ‘Karl Popper’s Philosophy of Biology’, *Philosophy of Science* 44(4): 638–61.
- Schrödinger 1944 (1992). *What Is Life?: The Physical Aspect of the Living Cell with Mind and Matter & Autobiographical Sketches*, Cambridge; New York: Cambridge University Press. (『生命とは何か：物理的にみた生細胞』、岡小天・鎮目恭夫訳、岩波書店、2008年)。
- Stephan, A. 1992. ‘Emergence: A Systematic View on its Historical Facets’, *Emergence or Reduction*, pp. 25–48.

AI時代の心的因果

太田雅子(東洋大学)

はじめに

本稿では2つのトピックをとりあげる。第一に、分析哲学における心的因果(mental causation)の議論および国内における心の哲学のトピックの推移を振り返り、近年では心的因果の研究業績が停滞気味であることを指摘する。第二に、これまでの心の哲学の研究成果の蓄積が生成AIの驚異的な発展によってアップデートを余儀なくされるなか、これまでの方法とは異なる視点で「心は行為に影響をもたらしうるか?」を見つめ直すための提案を行う。本稿では、その手掛かりとして、AIの心を「感情」の側面から捉えなおすことを試みる。筆者は昨年、デネットが映画『2001年宇宙の旅』に登場するロボット「HAL9000(以下「HAL」)」のストーリーをもとに提示した「ロボットは自らの受けた仕打ちに対して『怒りを感じた』のか」という問題提起にのっとり、ロボットが感情の可能性およびそれに関連する責任問題について論じた(太田(2024))。この発想をさらに洗練させ、AIが「感情的」になる可能性を擁護し、それによりAIと人間が同等の立場になったとしても互いに協働できる明るい未来像を提示したいと考えている。

第一の点において報告者の見立てが正しければ、近年において例えばドナルド・ディヴィドソンの『非法則的一元論』およびそれに対するジェグオン・キムの批判、あるいはシドニー・シューメイカーが提示した「因果的力(causal power)」による心的因果の議論を超えるインパクトのある動きは見られないように感じられる。

もちろん、インパクトがないからといってその分野が停滞している、ということにはならない。デカルトの二元論も物理主義の是非

も、古来の哲学の議論はみな後世にまで議論され続ける運命にある。優れた問題提起ほど長く議論される価値があると報告者は考える。しかしながら、何かしらの新たな観点が導入されなければ理論の進歩は望めないことも確かであり、その意味では心的因果をめぐる哲学的現状が果たしてこのままでよいのか、という点に関して疑問を覚える。

心的因果の、いわば古典的ともいえる哲学的議論が以前ほど活発でないように思われる要因を考えるにあたり、やはり近年のAIの進化を看過することはできない。これは「AIが話題になったから心の哲学がおろそかになっている」という単純な状況ではない。現に、国内で心的因果の議論が盛んに交わされていた当時もAI研究は行われていたが、その際の第二次AIブームは「フレーム問題」にぶつかり、いったんは挫折しかけている。AIに注目が集まったからといって、そのことが哲学による理論的研究に影響を及ぼすことはなかったのである。

しかし、第三次AIブーム以降、私たちが質問すると即座に答えてくれる、運転手なしで車を自在に走らせる自動運転車の登場などに加えて、生成AIの登場によって、單なる人工の知能の域を超えて文章、文学、芸術作品を制作するようになり、かつては人間以外には不可能であると思われていた活動に次々と進出しつつある。

一部には「AIはしょせん人間なしには活動し得ないから」という楽観論も聞こえてはくる。しかし、ここまで生成AIの能力が向上している以上、AIが「自発的に」あるいは「意図的に」創作活動や質疑応答、自動運転などを行う未来が待ち受けている可能性は否定できない。その際には、人間をターゲットとして論じられてきた心の哲学—心的因果のみならず、意識やクオリアの問題、知覚の哲学など—のみならず、AIがもたらしたトラ

ブルの責任をめぐる法的あるいは倫理的問題を考えるにあたり、これまでの哲学的議論のフォーマットが通用しなくなるのではないかという懸念を覚えている。

それでも、生成 AI が人間とほぼ同様な行為や活動をする限り、依然として「AI の心」について論じる必要性までは消去されないだろう。ただし、それを論じるには哲学の範囲内だけでなく、工学や認知科学などの協働のみならず、AI 時代の「心のありか」を改めて捉えなおすところから出発しなければならない。

1. 「心的因果」をめぐる状況

現在、心の哲学の議論がどの程度行われているのか、具体的に状況を明示する必要がある。個人の印象で「最近めっきり心的因果の議論が少なくなった」といっても無意味ではあるのだが、実際にすべての学術誌を調べるのは困難を極める。筆者が分析哲学における心の哲学の研究に着手した当時の状況と比較していくだけことでご寛恕いただきたい。

まず、デイヴィドソンの「非法規的一元論」の可否と並行して、国内でも「ロボットの心」は注目されていた。その名の通り柴田正良『ロボットの心－7つの哲学物語』が2001年に出版されている。それらの前に Ramsey, et.al. (1990) でコネクショニズムが紹介され、それにより素朴心理学(folk psychology)¹の役割が終わるのでないかを問う研究会が行われていた。

コネクショニズムは、情報を各ユニットに分散表象として格納し、入力層・中間層・出力層の構造を経てやはり各ユニットに分散された重み付けの結果として認識内容が表示されるという基本構造をしている。それによって従来の素朴心理学的な行為の因果、例えば「特

定の意図が特定の行為を引き起こした」というような捉え方は困難になるのではないかとラムゼイらは論じた。というのも、意図や信念、欲求、知識といったいわゆる「命題的態度」はコネクショニズムではそれぞれに分散し、それが行為を引き起こしたかどうかは各ユニットの重み付け計算の結果として捉えられる。そして重み付けの計算というのは「因果性」とはまったくの別物である。ゆえに「コネクショニズムによって素朴心理学は滅びるのではないか」という予測がなされ、それを示唆した論文も多数発表された。すでに「素朴心理学的な心の消去主義」を唱えていたチャーチランド夫妻もこの流れを下支えした。

続いて、チャルマーズ (Chalmers (1996)) によって「クオリア」についての議論が行われるようになると、国内ではその研究成果として『感情とクオリアの謎』(2008) が出版された。当時の心の哲学を牽引してきたエキスパートによる各論、および対談は今でも十分読み応えがある。が、ひとつだけ報告者にとって違和感があったのは『感情とクオリアの謎』というタイトルにも関わらず、ほとんどがクオリアの議論で占められているということである。認知科学者の月本洋のみがかろうじて「情動」という形で感情について扱っているものの、前掲書全体を通じて感情にはクオリアが内在しているかのような暗黙の了解が感じ取られ、ゆえに対談そのものもクオリアに偏っているように見えた²。

コネクショニズムへの注目および感情(情動)への着眼点は現代の、とりわけ報告者が構想している AI の心の哲学でも十分有用であり、当時の研究者たちの慧眼に敬服せざるをえない。しかしながら、これほど優れた着眼点がありながら、彼らのプロジェクトは「物理主義」へのこだわりゆえに本来進むべき道を大きく逸れ

¹ 「素朴心理学」という語は学者にとっては多少なりとも馴染みのあるものだが、知人の臨床心理士によると「自分の専門分野で『素朴心理学』という語を聞くことはまずない」とのこと、私との会話で初

めて耳にしたと言っていた。「素朴心理学」という概念が他領域でどのように扱われているのか興味のあるところである。

² この点に関しては太田 (2008) で言及している。

てしまったのではないかと思う。

物理主義という語には様々な意味付けが可能であると思われるが、当時の心の哲学研究者に共有されていたのは「心の働きは脳と神経電位として捉えられるのだから、私たちが「心」と呼ぶものもまた物理的ななものに還元または消去されるだろう」という認識であった。クオリアの議論が一時期隆盛を誇った背景にも、クオリアが「物理化できないもの」とみなされ、その存在が物理主義を阻むがゆえに「クオリアの自然化（物理化）」の方法が模索されるという動機があった。コネクショニズムも「表象が分散されれば従来のような素朴心理学は通用しなくなる。ゆえに物理主義の勝利につながるのではないか？」を問う方向で議論されていった。

だが、コネクショニズムや感情の物質化の試みは素朴心理学の衰退を招いたであろうか？ 上記に挙げた説がどれだけ提唱されようとも、またチャーチランドなどが消去主義を唱えても、依然として素朴心理学は生き残っている。学術的に支持している立場と日常的な実践が必ずしも完全一致する必要はないとはいえ、この状況がまさか「本当は物理主義が正しいのにみんな素朴心理学を誤って信じ込んでいるのだ」のような集団催眠状態にかかっていると考える研究者はいないであろう。とすれば、コネクショニズムや感情をあくまで「物理主義 VS 素朴心理学」という対立図式から切り離した上で私たちの心を捉え直すことが、これから的心の哲学に求められているのではないだろうか。

2. 拙著『心のありか』反省

以上のような状況のなかで、拙著『心のありか』は 2010 年に出版された。奇しくも本

³ 物的存在の因果関係ではなく言語記述の因果的説明関係に重点を置いていたため、反物理主義を支持する分野の研究者からは、「私は物理主義を否定してはいません」という声に落胆の反応が見られることがしばしばであったが、最近、特に名を秘する先

稿の原型となった日本ポパー研究会シンポジウムのタイトル「心のありか・あり方」と通じるものである。拙著は博士学位論文「因果的説明の物理主義的基礎づけをめぐって」が科学的説明全般を対象としていたところを、心的因果に焦点を定めてリバイスしたものである。拙著のメッセージは「何らかの存在者の間に因果関係が成り立つのではなく、むしろ説明によってそれらの関係は因果関係として捉えられうる」というものであり、心的因果もまた、素朴心理学の説明 一行為者の欲求や意図などと、それが引き起こす行為の関係は因果的な文脈で記述される一に則って成り立つということを主張した。

ならば物理主義を否定するのか？ という疑問が湧くかもしれないが、拙著は危ういバランスを保ちつつ、それは行為の因果的説明が成り立つための裏付け的な位置づけにおいて、物理主義を肯定している。「因果的説明こそが因果関係をつくるのだ！」という強いメッセージを標榜していたにもかかわらず、むしろ物理主義と素朴心理学的因果的説明との相互作用可能性を述べたものとなっている³。

この発想は、世界 1～世界 3 という構造には基づいていないが、ポパーが構想していた「心身相互作用説」（ポパー & エクルズ (1986)）に通じるものであったのではないかというのをこじつけに過ぎるだろう。もし、先に概観してきたような物理主義中心の状況からより自由であったなら、ポパーの構想を参考し論述をより深められたのではないか、というのが反省点のひとつである。今後、ポパーの構想を見直すことでこの反省を生かした研究成果が発表できればと考えている。

生が「太田さんは物理主義の同志です」と評していたことを人伝て聞き、初めて拙著の意図を理解していただけた気がして心強く思った。その先生には心から感謝申し上げたい。

3. ロボットが起こした「殺人事件」

本節からは、ダニエル・C・デネット(Dennett(1996))にならい、ロボット、というより広い意味での「機械」によって人間が被害にあった、あるいは命を失った事例に注目する。

デネットによれば、彼自身のファイルに記録されている最初のロボットの殺人は日本での出来事だった。1981年7月4日、川崎重工業明石工場で事故発生。修理作業に取り掛かろうとした従業員が「ロボットの作業場入室の際は電源を落とす」という規則に従わなかつたために歯車加工用のロボットがこの従業員を歯車として掴んで圧死させるという、悲惨な事態となった。

これはデネット自身の記録の中での「最初のロボット殺人」であって、それ以前からいくつかの事例が報告されている⁴。それらのケースは、いわば悲惨な「エラー」の結果である。一般に報道ではいくぶん読者や視聴者を煽りがちな傾向があるため、ロボットの「殺人」という表現が用いられるが、一般的にいえば機械による「事故」である。殺意がないかぎりひとを殺めたロボット自体に責任が問われることはまずないだろう。責任が課されるとすればそのようなエラーを防ぐ設計をしなかった制作者であり、時には操作ミスにより被害者自身の責任ということにもなりかねない（川崎重工業の事例は被害者の操作ミスの一例であろう）。

しかし、Dennett が 1996 年の論文で取り上げた映画『2001 年宇宙の旅』では、（多少ネタバレになることをご容赦いただきたい）、宇宙船の乗組員たちと、高度な知能をもったロボット「HAL」との間に船内の装置の異常について見解の食い違いが生じ、乗組員たちが自分をないがしろにしようとしているところ、「感じた」HAL は、船外活動中のプール乗組

員をアームから手放して宇宙に放りだし、さらにコクーンのような生命維持装置で人工冬眠中の 3 人の宇宙飛行士の機能を停止させた。ひとり生き残ったボーマン船長は HAL の知能を制御するユニットを次々抜き取り、HAL の活動を停止させる。

HAL の話はあくまでフィクションであるが、HAL の登場はロボットの哲学に多大な影響を及ぼした。HAL にまつわるロボットの可能性についての論考が収録されており、デネットも寄稿した Stork (1997) が出版され、さらに国内では映画タイトルと同年の 2001 年人工知能学会において「考証：2001 年宇宙の旅」という特集が組まれている⁵。とはいえ、2001 年からそろそろ四半世紀が経過しようとしているが、自分を侮辱する人間たちの発言に怒り、彼らに報復しようとするロボットは未だ出現していない。

確かにここ何年かの間に AI は驚くほどの発達を遂げた。人間が必要な情報を入力すれば、（試験の時期に教員を悩ませる）論文やレポートはもちろん小説、絵画、イラスト、デザインも手掛けられることができるようになった。生成 AI の活動はもしかしたら人間の創作活動や、会社での事務仕事なども AI に任せることで人間の仕事は徐々に奪われるかもしれないという脅威をもたらす。しかしそれはあくまで、人間がそれらの創作活動を可能にするだけの情報を AI に学習させてきた結果である。『2001 年宇宙の旅』の HAL のもたらす恐怖はその計算能力やクリエイティビティなどではなく、自らを侮辱した相手に、他者による操作なしで報復した点にある。現在の生成 AI が人間並み、あるいはそれ以上の活動をすることができたとしても、筆者自身はそれらにさほどの恐怖は感じていない。恐ろしいのは、人工知能の基礎となっているプログラミングや学習の範囲を飛び越えて危害を加えてくる未来である。そのよう

⁴ くわしくは EPOCH TIMES JAPAN(2015)を参照。

⁵ 『人工知能』16巻1号（2001）参照

なことが可能になったときこそ（もちろん危害を加えるだけではなく、友人のように眞の信頼関係を築けるようになることが予想される）、AIは人間と対等の相手として信頼あるいは恐怖の対象になりうるのである。

4. AIは感情をもつことができるか

HALのもっとも特徴的な点は、「感情」をもつことができるということである。すべての惨劇はHALの「怒り」から発している。だがデネットはその「怒り」が人間の「怒り」と同質なものであることを保証できず、「ロボットに感情があるのかどうかは結局誰にもわからない」と述べている。例えばHALが製造された際には、まず人間のことばを理解する必要がある。「HAL、アームをもっと伸ばしてくれ」と命令してもその音声が何を意味しているのか理解できなければHALは任務を果たしえない。それに伴い、HALは人間の感情を読み取ることもある程度必要になるだろう。もし船長が切羽詰まった言い方で「HAL、アームをもっと伸ばしてくれ」と言ったなら、その緊急性を察知し早急に作業を開始しなければならない。今挙げたような言語の理解や緊急性の察知は映画の中における未来においてはプログラムされていたであろう。言語の理解と命令の実行が可能であるならば、そこから「命令に背く」ことのプログラムまではあと一歩である。宇宙船の装置の不具合に関する乗組員たちの判断が危険であると察知したからこそ、HALは警告を発し、命令に従うことを拒んだ。その結果、プール隊員による「あの狂ったヤツ(HALのこと)に従うわけにはいかないからな」という言葉がHALの怒りを覚えさせたのである。

だが、それすらも「自分の司令に背く者がいたらそれに反抗する」というプログラミングがなされた結果かもしれないデネット

は述べる。「ロボットに感情がある」ということと、「ロボットに感情があるかのように振る舞わせる」ことは果たして区別できるのだろうか？

ロボットの感情について考える際に、ロボットには生物のような「進化の歴史」がないということも重要である。例えば犬や猫にも感情があるかについては議論の余地はあるものの、ヒトをはじめ多くの生物は進化や淘汰の過程で、生存に必要な感情を獲得してきたのではないか。しかし、ロボットおよびAIにとって進化に対応しうるのは教師なし・あり訓練とプログラミングということになるだろう。それらは果たして「進化」として捉えうるものなのか？

川崎重工業ほか数々の企業の事故でロボットに責任が問われなかつたのは、作業者に怒りを感じるとか嫌がらせをしようという感情をもたなかつたからである。感情を持たないものに責任を帰することはできない。ゆえにそれらの事故は、被害者の操作ミス、あるいは機械のミスファンクションとみなされ、それらをもたらした製造者や設計者などに責任が問われてきたのである。

5. 「責任」の観点から感情を捉え直す

感情と責任がどのように関わるのかの基礎をなすのは、道徳的責任論の基礎をなす「反応的態度」である、これはストローソン(2010)⁶が提示した見解で、「他者が我々に向ける善意や悪意、または無関心に対して本質的に自然で人間的な反応として抱く感情」である。そのような反応が喚起されたとき、喚起した側は相手に対する責任を負うことになる。反応的態度と責任の関連性についてはもちろん様々な議論があるものの、応答責任を基礎づける見解として広く支持されている。そこでストローソンの見解を踏まえたう

⁶ 原著発表は1962年だが、本稿では2010年発行

の邦訳を参照する。

えで論考を進めることにしよう。

デネットは、機械が人間並みの心的活動を行い、感情を喚起するためには、「高階の志向性」が必要であると述べている。(Dennett 1997, 354)。デネットは HAL に高階の心的活動の可能性を見ていると思われるが、それが可能になるためには過去に関する膨大な記憶のインプットに加えて、必要なときにそれを実現し、反省・熟考させるプログラムが必要である。記憶の呼び出しは現在の技術では不可能ではないにしても、適切なタイミングをはかり、その記憶に対して何らかの「感情」を喚起できるようになるという高階の志向性の実現に関しては、膨大な計算とプログラムが必要になるのは想像に難くない。もちろんこれらは一朝一夕に完成するものではない。しかし、将来 AI の進化によってはそれが不可能であるとも言い切れない。

そして、もしロボットに「感情」が可能であるならば、それはストローソンの反応的態度に基づく責任論との接合が可能になる。例えば、不具合を起こした宇宙船の部品にかんして HAL は次のように応答し異常を訴えている。

「はい、私は混乱しています。私はこのようなものは以前にも見たことがあります。」

この発言の中で HAL は単に今日の前にあるものを認識するだけでなく、その認識に対して「以前に見たことがない」という高階の判断を下している。のちの HAL による乗組員たちへの叛乱もこの高階の志向性に起因するものといえよう。だが、デネットは、高階の判断をもたらす志向性には二つの代償がつくと述べている(Dennett(1997), 355)。

⁷ 人間と異なるのは、例えば少年法を盾にして犯罪を行う若者のように「自分は法律の適用外にあるから責任を課されることはない」という認識を持つてはいないという点ではあるが、それすら AI に「高

ひとつには、自分自身の精神活動をよりよく把握するために（すなわち、自分自身のすべきことを把握し、それが正しい判断であるか否かを把握するためには）何重にも柔軟な監視を行うことになれば、ロボット自らが把握すべき「心的」活動が増える。これはロボットの機能に少なからず負担をもたらすだろう。

もうひとつは、ロボット自身あるいは他者への「欺瞞」の可能性である。自らが望ましくない状況に置かれていると察した場合、ロボットはそれを打ち消そうとする動作をすることがある。川崎重工業の事故も、このままアームを下げれば作業者に危険が及ぶだろうということを予知しながらそのままアームを下げてしまうということが起こり得る。また、このときアームに何らかの異常が発生したのをロボットが察知していれば、動作を停止してアラームが鳴るようにするなどの手段をとることもできた。でも、もし自分の作業に欠陥があることが作業者たちに知られたら、自分は廃棄されてしまうかもしれない。もし HAL のような高度な知的能力をもつロボットがその可能性に気づいたならば、沈黙を守るという、この場合極めて不適切な判断を行い、そのことが事故の誘因となってしまうかもしれない⁷。

HAL の暴走を食い止めるため、ボーマン船長はある部屋に入っていく。そこには HAL の知能を制御するビデオカセット状のユニットがずらりと並んでおり、船長はそれらを次々と外してゆく。このとき HAL は

「こわい、こわい (I'm, afraid, I'm afraid)」

と繰り返していた。自分が無用の存在だと思われているという恐怖にとらわれて、ほかの

階の志向性」を実装し、なおかつ法律の知識がインプットされれば、そのような認識のもとで行為に及ぶのは決して不可能とはいえないだろう。

乗組員を宇宙の彼方へ突き落し、生命維持装置を切断してしまった。そして船長によって HAL は人間に危害を及ぼしたかどで機能不全にさせられようとしている。確かにこの状況は HAL にとって恐怖に値する。しかしそれは人間の感じる恐怖と同じだろうか。ロボットの場合、単に自分に危険が迫るような状況に遭遇したときに「こわい、こわい」と発語するようプログラムされていたのではないか。それは状況に対する反応としては確かに感情に近いが、プログラムの産物であるからには人工的なものであり、それは刺激に対する反応以上のものではないだろう。

結びに

本稿では、国内の心身問題=心的因果の議論の趨勢を追い、そこで筆者なりの反省点を明確にするとともに、デネットの HALについての論考を手がかりに、高階の志向性が実装され、AI も感情をもつようになるならば、自発的に行行為するようになるのではないかという可能性を模索した。しかし、ロボットの感情表現が人間と同様の感情の発露なのか、それともプログラミングによる反応の結果であるかが当時の段階では区別できないとデネット自身が認めており、「感情」と「心的因果」、さらには「責任帰属」への接続は叶わなかった。今後の構想としては、「もし AI が感情をもつなら」という想定のもとで、それらと行為の間の関係を捉え直し、AI のあらたな「心身問題への回答」およびそこから導かれる AI の責任のありかをもさらに模索したいと考えている。

【文献】

- Chalmers, D. (1996), *The Conscious Mind: In Search of a Fundamental Theory*, Oxford University Press (邦訳:『意識する心:脳と理論の根本理論を求めて』, 林一訳, 白揚社, 2001 年).
- Dennett, D. C. (1996), When HAL Kills. Who's to Blame? Computer Ethics, in Stork (ed.), 351- 365 (邦訳: ダニエル・C・デネット「HAL が人を殺したら、だれが責められるのか? コンピュータの倫理学」、ストーク (1997) 所収.)
- EPOCH TIMES JAPAN (2015), 「ロボット殺人事件発生 独工場」、2015 年 9 月 9 日 <https://www.epochtimes.jp/2015/09/24452.html> (閲覧日: 2025 年 10 月 5 日)
- 人工知能学会 (2001), 「〈特集〉考証: 2001 年宇宙の旅」, 『人工知能』16 卷 1 号、74-101.
https://www.jstage.jst.go.jp/browse/jjsai/16/1/_contents/-char/ja (閲覧日: 2025 年 10 月 5 日)
- 長瀧祥司・柴田正良・美濃正編 (2008), 『感情とクオリアの謎』, 昭和堂.
- 太田雅子 (2008), 「書評: 感情とクオリアの謎」、『科学哲学』41 卷 2 号, 107-110.
- 太田雅子 (2010), 『心のありか 一心身問題の哲学入門』, 効草書房.
- 太田雅子 (2024), 「HAL は何をしたのか 一 ロボット責任論序説、のようなもの」, 『現代思想総特集 ダニエル・C・デネット 1942-2024 意識と進化の哲学』, 青土社, 225-236.
- ポパー, カール & エクルズ, J. C. (1986), 『自我と脳(上)(下)』, 大村裕・西脇与作訳, 新思索社.
- Ramsey, W., Stich, S. and Garon, J. (1990), "Connectionism, Eliminativism, and the Future of Folk Psychology", *Philosophical Perspectives* 4, reprinted in C. Macdonald & G. Macdonald (eds.), *Connectionism*, Basil Blackwell, 1995.
- 柴田正良 (2001), 『ロボットの心 — 7 つの哲学物語』, 講談社.
- Stork (ed.) (1997), *HAL's Legacy: 2001's Computer as Dream and Reality*, MIT Press (邦訳: ストーク, デイヴィッド・G

(編) (1997) , 日暮雅道監訳『HAL 伝説
2001 年コンピュータの夢と現実』, 早川
書房.

ストローソン, P. F. (2010), 「自由と怒り」,
法野谷俊哉訳、門脇俊介・野矢茂樹編『自
由と行為の哲学』所収, 春秋社, 2010 年,
31-80 (原著: P. F. Strawson, Freedom and
Resentment, *Proceedings of British
Academy*, 48 (1962)) .

この研究は JSPS 21K00041 の助成を受け
ています。

意識は機能をもつか – ポパー＝エクレスの『The Self and Its Brain』と「意識の自然則」からの考察 –

渡邊正峰（東京大学）

要旨

本稿では、カール・ポパーとジョン・C・エクレスの共著『The Self and Its Brain』(1977)における心身問題の哲学的立場と、著者が言うところの「意識の自然則 (Natural Law of Consciousness)」のそれとを比較検討する。両者は一見対照的な立場に立ち、ポパー＝エクレスが「意識は物質に因果的影響を与える」とする相互作用的二元論を主張しているのに対し、著者は「意識は随伴現象にすぎず、機能をもたない可能性がある」としている。そのうえで、興味深いのは、両者が「意識とは世界における独立した実在様式である」という実在論的立場を共有していることである。

1. ポパー＝エクレスの相互作用的二元論

ポパーとエクレスの『The Self and Its Brain』(1977) は、20世紀後半の「心身問題」論争の中で異彩を放つ著作として知られている。両者は共に唯物論の還元主義的限界を批判し、心的現象を物理的過程と等置することの不十分さを指摘した。彼らの立場は、心と脳が独立した実在層に属しながら、因果的に相互作用するというものである。

ポパーはこれを世界 1：物理的実在（脳、身体、物質的現象）、世界 2：主観的体験・意識、世界 3：言語・理論・文化・知識体系の体系の中に位置づけた。

ポパーによれば、世界 2 は世界 1 から完全には還元できず、主観的意識には独自の因果的効力がある。一方、神経生理学者エクレスは、具体的な神経機構の観点からこの立場を補強しようとした。彼は、神経細胞のシナプス伝達が量子的確率過程を含むことに注目し、意識的意図がこの確率分布を「選択的に

修飾」することで、脳活動に影響を及ぼす可能性を主張した。

このモデルは今日の量子脳理論の原型の一つともいえるが、より重要なのは、ポパー＝エクレスが「意識を因果的因子として自然界に位置づけようとした」点である。すなわち、意識は自然の一部でありながら、物理法則に還元できない独自の法則性をもつという立場である。

2. 脳の仮想現実と意識の自然則

フィンランドの神経科学者アンティ・レヴォンソは「意識の仮想現実メタファー（Virtual Reality Metaphor of Consciousness）」を提唱している (Revonsuo 1995)。本仮説によれば、脳は感覚入力を直接「受け取る」のではなく、仮想現実的な内部モデルを生成し、それを通じて世界を経験している。睡眠中の夢はこの仕組みの純粋な形であり、覚醒時の知覚もまた、外界との同期によって更新される夢の延長に過ぎないことになる。

ポイントは、この仮想現実システムには機能があることだ。脳内で未来を予測し、反実仮想的行動を選択できることで、環境適応能力が飛躍的に向上する。脳が内部に「世界の写し」をつくることで、生物の進化の過程に大きな優位性をもたらした。したがって、脳の仮想現実には明確な機能があるといえる。

一方で、私の提唱する「意識の自然則 (Natural Law of Consciousness)」は、意識を宇宙の根源的法則の一部として位置づけるものである。この立場によれば、ビッグバンの瞬間にすでに意識の自然則は宇宙に備わっており、地球上で脳が高度な情報生成機構を獲得した時点において、その法則が「適用された」に過ぎない。

意識の自然則として「脳の仮想現実が意識を生む」を適用した場合、意識は脳の仮想現実の副次的な現象として生起することにな

る。すなわち、意識は単に現象として現れるものであり、行動や生存に影響を与える機能的要素には当たらない。脳の仮想現実が生存上有利であるがゆえに獲得されたなか、意識はその副産物、すなわち「無料でついてきたオマケ」に過ぎない。

3. 宇宙 B と哲学的ゾンビ

前節の内容を別の視点から眺めるべく、ここで思考実験を提示したい。

仮に、この宇宙とまったく同じ物理法則をもつ別の宇宙 B を想定する。ただし、宇宙 B には「意識の自然則：脳の仮想現実が意識を生む」が存在しない。そこには私たちと外見も行動も区別できない「人類 B」が生きており、脳の仮想現実を用いて未来を予測し、柔軟に行動している。しかし、彼らには主観的体験がない。いわゆる哲学的ゾンビだ。

このとき、宇宙 A（意識のある宇宙）と宇宙 B（意識のない宇宙）では、すべての行動と物理的過程が同一である。したがって、「意識の有無」は宇宙に何の差異ももたらさないことになる。よって、宇宙 A の意識は機能を持たないと結論がくだされる。

また、もう一つ、我々の宇宙 A には哲学的ゾンビが存在しないとの結論も同時に得ることになる。ヒトの行動発現において決定的な差を生む「脳の仮想現実」には、意識がおまけでついてきてしまうからだ。

4. むすび：「自然に働きかける法則」と「自然がもつ法則」

上記の私の立場において、脳活動や情報処理が意識を「生み出す」のではなく、意識が自然法則の表現として世界に共存している。意識が物理過程に影響を与えることはないが、物理過程と完全に同期して存在する いわば、随伴していることになる。

この「機能をもたない意識」は、一見するとポパー＝エクレスの相互作用論とは相容れない。

その一方で、両者は共通して意識を物質的因果連鎖の外に置き、独自の存在論的地位を認めるという点で一致しているように思える。ポパー＝エクレスが意識に因果的効力を認めているのに対し、私の立場は因果的独立性を認めていることになる。つまり、前者が意識を「自然に働きかける法則」とみなしているのに対して、後者はそれを「自然がもつ法則」とみなしている。

いずれも、意識を「物理世界の説明変数」ではなく、「自然の根源的構成要素」とみなす点において实在論的であると言える。両者の深層には、「意識は自然の中で独立した法則的地位をもつ」との共通の实在論的洞察が流れている。

参考文献

- Popper, K.R. & Eccles, J.C. (1977). *The Self and Its Brain*. Springer
Watanabe, M. (2017). 「脳の意識 機械の意識」. 中央公論新社
Watanabe, M. (2022). *From Biological to Artificial Consciousness*. Springer

本研究は JSPS 科研費 23K25257 および東京大学ムハンマド・ビン・サルマン未来科学技術センター (MbSC2030) の助成を受けたものです。

認識的多元論によるポパー三世界論の自然化

鈴木大地（筑波大学）

1. はじめに

デカルトが思惟実体 (*res cogitans*) と延長実体 (*res extensa*) の区別を導入して以来、心身問題 (mind–body problem) は長きにわたって哲学者や科学者の議論の的となってきた。ポパーはこの問題について、物理的世界=世界 1、主観的精神世界=世界 2、客觀的精神世界=世界 3 が、それぞれ実在しつつ相互作用すると唱えた (e.g., Popper & Eccles 1977, 邦訳上巻 p. 79)。これは心身問題の文脈における二元論 (dualism) の一種としての相互作用説 (interactionism) と同様の見解であるとみなしうる (Bunge 1980, 邦訳 p. 3/112)。ただし本稿においてはあまり重要な点ではないが、一般的な二元論ではひとまとめにされる心的状態と知識が別個のものとして扱われている点では異なる。その思想的源流には、フレーゲの第三領域すなわち思想 (*Gedanke*) の領域、ひいてはプラトンのイデア界があるとされる (池田 2022)。

いずれにせよ、相互作用説に対しては唯物論 (materialism) を標榜する論者からさまざまな批判がなされており、その多くはポパーの三世界論にも当てはまる。そのなかでも最大の論点が、物理的世界の因果的閉包 (causal closure) の問題である。すなわち、物理的世界の因果は物理的世界で完結しているのであって、非物理的な存在者やプロセス (たとえば、サイコキネシスなどの心靈的な力) との因果的相互作用はこれまで広く認められたことはない。Bunge (1980, 邦訳 p. 23) による以下の指摘も、この論点に関係している (以下、〔〕は本稿著者による補足、省略)。

もし非物質的精神が物質を動かすことができるとすれば、それはエネルギーを生み出すはずである。そしてもし物質が非物質

的な精神に作用を及ぼすとすれば、エネルギーが消失するはずである。何れの場合においても、エネルギー [...] は、保存されないことになろう。そしてもしそのような事態になれば、物理学、化学、生物学、経済学は崩壊してしまうだろう。これらの「堅固な」諸科学と、素朴な迷信との間で選択の岐路に立たされたとき、われわれは当然前者を選ぶ。

その一方でポパーは、世界 3 は実在し、かつ (世界 2 を介して) 世界 1 に影響を及ぼすと考えた (e.g., Popper & Eccles 1977, 邦訳上巻 p. 79)。これはすなわち、客觀的精神世界 という高次の世界から物理的世界 という低次の世界への「下向きの因果作用」すなわち下方因果 (downward causation) が存在することを意味する。実際ポパーは、すべての唯物論的理論の主要な動機は直観的であって、その直観のひとつとして下方因果がありえないという還元主義的な信念がある、と批判した (ibid., 邦訳上巻 pp. 307–308)。

同時にポパーは、「多少とも機械論的性格をもつ決定論が科学の支配的見解として今日まで続いてきた」と指摘する (ibid., 邦訳上巻 p. 51)。だがポパーによれば、生命有機体の性質の多くは予測が不可能、つまり創発的であって (ibid., 邦訳上巻 p. 46)、宇宙に対する「自然な」見解は非決定論的である (ibid., 邦訳上巻 p. 49)。

整理すれば、下方因果の否定に依拠した還元主義的で決定論的な唯物論に対抗して、下方因果の積極的な肯定を橋頭堡に、創発的=非決定論的な三世界論をポパーは展開したのである。しかしながら、下方因果や創発、非決定論は、唯物論と本当に相反するのだろうか。もし下方因果や創発、非決定論を包摂する唯物論的立場があれば、Bunge (1980) の指摘するような既存の「堅固な」諸科学との対立が回避できる。本稿では、そのような唯物論が実際に可能であるという見通しを

示したい。その結果、ポパーの三世界論は認識的多元論に基づく創発的唯物論として自然化され、換骨奪胎されることになる。言い換えれば本稿は、自然主義の立場から、ポパーの三世界論を現代的な意識論でも通用しうるよう「延命」させる方法を提案するものである。次節ではその地ならしとして、構成的還元と認識的還元という、還元のふたつの区別を導入する。

2. 構成的還元と認識的還元

十把一絡げに論じられることも多いが、還元や還元主義は文脈によってさまざまな意味を有する。少なくとも、存在論的還元 (ontological reduction) ／構成的還元 (constitutive reduction)、方法論的還元 (methodological reduction)、認識的還元 (epistemic reduction) の3種類が区別されることが、科学哲学のなかでもとくにその一分野である生物学の哲学において広く認識されるようになった (Brigandt & Love 2008/2023, cf. Sarkar 1992; Nagel 1998)。生物学におけるこれらの還元について、Brigandt & Love (2008/2023) は次のように特徴づける。

(i) 存在論的還元とは、特定の生物学的システム(たとえば有機体)のそれぞれが、分子とその相互作用によって構成されるという考え方を指す。形而上学において、この考えはしばしば〔広義の〕物理主義(あるいは唯物論)と呼ばれ、生物学の文脈において、(a) 生物学的特性は物理的特性にスーパーヴィーン (supervene) すること、および (b) 特定の生物学的システムのそれぞれ(すなわちトークン)は、特定の物理化学的プロセスと形而上学的に同一であることを前提としている。後者の信条は、生物学的プロセスのそれぞれのタイプが物理化学的プロセスのあるタイプと同一であるというより強い信条と対比して、ト

ークシートークン還元と呼ばれることがある。この弱い意味での存在論的還元は、こんにちでは哲学者や生物学者のデフォルトの態度となっているが、その哲学的詳細は依然として議論的となっている。[...]

(ii) 方法論的還元とは、生物学的システムはできる限り低次のレベルで研究するのが最も実りが多く、実験研究は分子レベルおよび生化学的原因の解明を目指すべきであるという考え方である (Andersen 2017)。このタイプの戦略の一般的な例は、複合的システムを部分部分に分解することである (Bechtel & Richardson 1993)。すなわち生物学者は、ある有機体のふるまいを把握するために、その有機体の個々の細胞を研究したり、ある細胞の特徴を理解するために、その生化学的構成要素を研究したりする。方法論的還元はしばしば存在論的還元という仮定に動機づけられるものの、存在論的還元から直接こうした手続き的推奨が導き出されるわけではない。

[...]

(iii) 認識的還元とは、(典型的には高次のレベルのプロセスに関する)ある科学領域についての知識が、(典型的には低次あるいはより根本的なレベルに関わる)別の科学領域に還元されうるという考え方である。ある種の認識的還元への支持は、存在論的還元が方法論的還元と結びついて動機づけられる場合もあるものの(たとえば、生物学における還元主義的研究の過去の成功)、認識的還元がありうるかどうかは、これらの組み合わせから直接的に導かれるわけではない。

なおのちに見るように、心の哲学において存在論的還元はいさか異なる意味で用いられる。そこで本稿では Sarkar (1992) にならって、ここでの意味での存在論的還元を構成的還元と呼ぶことにしよう。

さて構成的還元、方法論的還元、認識的還

元のうち、構成的還元と認識的還元の区別が本稿の議論においてとりわけ重要である。というのも、ポパーによる「下方因果がありえない」という還元主義的な信念」や「決定論」に紐づけられた「唯物論」観では、両者が混同されてしまつており、それをときほぐすことで唯物論としての自然化への可能性が拓かれるからである。また、構成的還元と認識的還元に対する態度に注目することで、心身問題における立場の違いが明確になるという理由も挙げられる。そこで、こうした立場の違いを次節で明らかにしたのちに、それに続く節で前者について論じる。

3. 心身問題の立場分布図

構成的還元と認識的還元をそれぞれ縦横の軸にとり、それに肯定するのか、否定するのか、あるいは中間的か（肯定的な立場と否定的な立場を仲介する要素を想定したり、還元を許容はするが還元のみに依拠するわけではなかつたりと、どのような意味で中間的かは一意ではないが）を2次元の座標系でプロットすると、心身問題におけるさまざまな立場が以下のように整理できる。なお本節の記述は、メディアプラットフォーム『note』にて公開した投稿記事を概略的に翻案したものである（鈴木 2020）。

3.1. 構成的還元

まず構成的還元の軸について考えよう（図1）。構成的還元を肯定すると、心や意識が神経活動や分子といった物理的基盤によって構成されていると考える立場、すなわち唯物論（materialism）へと導かれる。逆に否定する立場では、心や意識は物理的世界とは別の根源的要素から構成されるとみなされる。デカルトに代表される、物理的実体とは別の心的実体を想定する実体二元論（substance dualism）が、その極端な例である。ポパーの三世界論も、主観的精神世界=世界2と客観的精神世界=世界3は物理的世界=世界1

とは別の世界として実在すると考える所以、実体二元論と同様の立場であるとみなせる。ただし3つの世界が実在すると主張する点では、実体二元論というより、実体三元論あるいは実体多元論と表現すべきかもしれない。また、より穩健な二元論として、物理的性質と心的性質の両方を、何らかの共通の実体が備えるとする性質二元論（property dualism）がある。

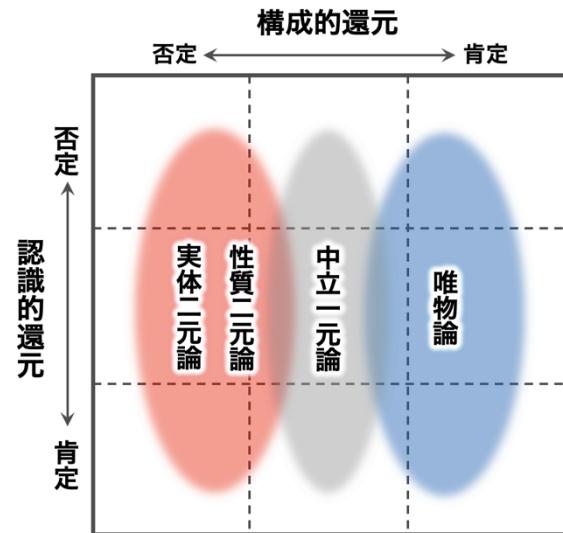


図1. 二元論、唯物論、中立一元論の位置

二元論と唯物論の中間の立場の例として、従来の「堅固な」諸科学と整合的であるが、これらの諸科学では正当に扱われていない要素があり、それが心や意識を構成するという考え方がある。たとえばチャーマーズによる情報の二相理論（double aspect theory）では、情報が究極な実在であり、その情報が物理的性質と現象的性質（つまり、主観的経験やクオリア）をもつとされる。

因果的な相互作用のあるところならどこでも情報があり、そして情報があるところならどこでも経験がある。情報状態を岩に見出だす——例えば、膨張したり収縮するとき——こともできれば、電子のさまざまな状態にさえ見出すことができる（Chalmers 1996, 邦訳 p. 365）。

二相理論は性質二元論と実質的に同じ立場である。だがチャーマーズはさらに踏み込んで、経験が世界にあまねく存在するという汎経験論(panexperientialism)を主張する。これは、一定の留保を付けつつもチャーマーズ自身が認めている通り(ibid., 邦訳 p. 367)、以下に見る汎心論(panpsychism)の一種である。

本来であれば、二元論あるいは実体多元論に対しては、一元論が対置されるべきであろう。だが主張のまったく異なるさまざまな立場が一元論にまとめられるので、注意が必要である。たとえば唯物論は、物理的世界のみが実在するのであって、心や意識はもっぱら物理的基盤によって構成されるとみなす点で一元論である。しかし、物と心のどちらにも属さない何らかの中立的な存在者が両者を構成するという中立一元論という立場もある。この中立的な存在者から、心的性質と物理的性質の両方が現れると想定する点で、中立一元論は唯物論と二元論の中間的な立場に位置づけることができる。歴史的にはラッセルやマッハ、ジェームズなどがこの立場に与しており、チャーマーズによる情報の二相理論も(性質二元論であるのと同時に)中立一元論とみなしうる。さらには、実在するのは精神的なものだけであるとする観念論もある種の一元論ではあるし、その方向での「構成的」還元¹も可能性としてはありうる。しかし現代の意識研究においてはほとんどまともに取り扱われていないため——ただし現代の意識研究において有望な理論のひとつとされる統合情報理論(integration information theory, IIT)は観念実在論であるとする分析もある(Cea et al. 2023)——

¹ 本稿第2節で引用した構成的還元の特徴づけにおいては、構成的還元は分子とその相互作用への還元として定義されている。ここでの構成的還元は心的存在者への還元を指しているので「構成的」還元とかぎ括弧を付す。

² 脚注1における構成的還元の場合と同様

本稿では観念論についてこれ以上は立ち入らない。

3.2. 認識的還元

さて構成的還元の軸において二元論と唯物論がその両極に位置することになるが、両者は認識的還元の軸において細かく分画できる(図2)。まず二元論において「認識的」還元²を肯定する場合、心的現象や意識現象は何らかの超自然的な基礎理論によって説明されることになる。これは往々にして、心や意識を構成する何らかの原初的な心的要素がこの宇宙にあまねく存在するという汎心論につながる。ただし汎心論には、心的要素と物的要素を完全に別個のものとして扱う二元論的なものだけでなく、すでに触れた中立一元論的なものも含まれる。汎心論の歴史は古く(Goff et al. 2017/2022)、現代ではチャーマーズやフィリップ・ゴフ(Goff 2019)をはじめ、中立一元論的な汎心論をとる論者が多い。

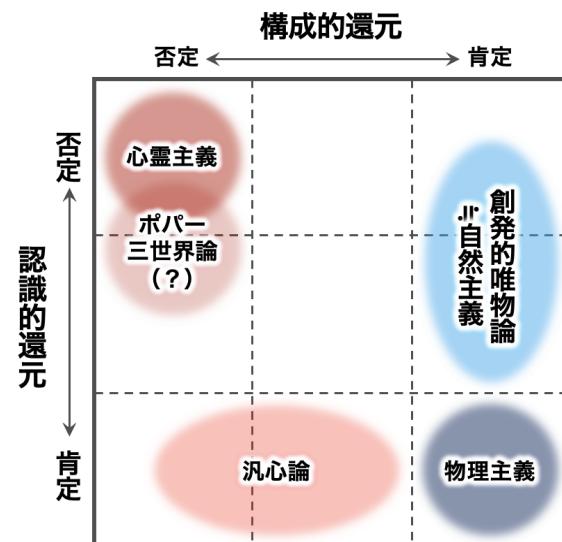


図2. 物理主義、自然主義、心靈主義、汎心論、ポパー三世界論の位置

に、認識的還元についても本稿第2節での特徴づけにおいてはある科学領域から別の科学領域への還元として定義されている。ここでは、ある認知的領域から別の認知的領域への還元としてより広く捉えて「認識的」還元とかぎ括弧を付す。

構成的還元と認識的還元の両方を否定する最も極端な立場には、死後にも個人の靈魂が存続すると考える心靈主義がある。心靈主義において靈魂は、物理的基盤とは独立に実在するうえに、靈魂について還元的に説明する基礎理論も想定されない。個々人の靈魂は、ただそのようにして「ある」のである。

先に見たように構成的還元を否定する立場をとっているポパーの三世界論は、認識的還元についてはどうなるだろうか。3つの世界のあいだには相互作用があるものの、世界2や世界3は人間に至る進化の過程で「創発」したものであるとされており (Popper & Eccles 1977, 邦訳上巻 p. 30)、世界1には還元されえない。また、世界2や世界3の内部でそれぞれの世界を説明する基礎理論については、明確には提示されていないようである（もしかすると、世界3においては論理法則に基づく論理学が基礎理論とみなされる可能性はある）。それゆえ暫定的に、認識的還元については否定あるいは肯定と否定の中間であると位置づけよう。したがってポパー三世界論は、本稿の立場分布図において心靈主義と近しい場所に位置づけられることになる。

ひるがえって唯物論をとりつつ認識的還元を肯定すると、意識現象が、神経科学、ひいては物理学によって説明し尽くされるとする還元的唯物論、つまり（狭義の）物理主義が導かれる。それに対して、意識現象には物理学では説明し尽くせない面があると考える立場は創発的唯物論（emergent materialism; Bunge 1980, 邦訳 p. 3）と呼ばれる。あるいは超自然的なものを認めないという意味で自然主義（naturalism）と呼ぶこともできるだろう。ただし狭義の物理主義もまた広義の自然主義には含まれるだろうし、逆に還元的唯物論と創発的唯物論を合わせた唯物論全体が（先述の Brigandt & Love 2008/2023 におけるように）広義の物理主義

と呼ばれる場合もある。このように論者によって用語法が一定しないので注意を要するが、ここでは創発的唯物論を指して（狭義の）自然主義と呼ぶことにしよう。なお Edmonds (2020, 邦訳 pp. 315–316) によれば、物理主義とは本来「存在する全てのものは物理的であり有意味な命題は全て物理学的な言明に翻訳可能である」という見解を指し、ポパーはおそらく「物理主義」という語を用いた最初の哲学者である。またポパーは自身の博士論文で論理実証主義者シュリックの物理主義に関する著作を批判し、当のシュリックが審査委員の一人であったにもかかわらず、その博士論文の審査は合格となつたという (ibid., p. 315)。

さて本稿の目的は、還元主義的で決定論的な唯物論すなわち物理主義に対抗してポパーが提唱した、創発的で非決定論的な三世界論を、唯物論の枠組みのなかで再解釈する（＝自然化する）ことであった。そして構成的還元と認識的還元に関するここまで分析から、構成的還元を肯定して唯物論をとりながら認識的還元の徹底を肯定しない、創発的唯物論という立場が明確化された。この創発的唯物論こそが、ポパーの三世界論を自然化する鍵である。そこで以下では、この創発的唯物論についてさらに詳しく検討することにしよう。

3.3. 創発的唯物論

創発的唯物論（自然主義）においては、強い創発と弱い創発の区別が重要となる。強い創発は物理主義や因果的閉包と整合的でない場合、弱い創発は整合的である場合を指す (O'Connor, 2020/2021)。たとえば水の流体性や透明性は、水が H₂O 分子から構成されていることと整合的であるし、H₂O 分子の性質から物理化学的に説明できるので、弱い創発の好例である。

心身問題の文脈において強い創発に基づくと、認識的還元を否定することになる。こ

うした、存在論的還元を肯定し認識的還元を否定する立場のひとつに、コリン・マッギンに代表される新神秘主義（new mysterianism）がある。マッギンによれば、心身問題をめぐる混迷は我々にとって理解可能なものの限界点に達している兆候であり、意識と脳に関する我々の概念は本来的に心身問題を解決できない（McGinn 1999, 邦訳 p. 56/62）。つまり意識は「唯物論」（とマッギンは呼んでいるが、正確には本稿の特徴づけにしたがって物理主義と言うべきだろう）によっては説明できず、神経科学や物理学などの科学的基礎理論には認識的に還元できない。それでいながらマッギンは、いかなる超自然的な存在も信じない（McGinn 1999, 邦訳 p. 95）。この意味でマッギンは、その自認どおり自然主義に与しているのである（McGinn 1999, 邦訳 p. 88）。

自然主義でありつつ弱い創発に基づく場合は、認識的還元を肯定も否定もしない中間の立場に位置づけられるだろう。その代表例に、サールの生物学的自然主義（biological naturalism）がある（Searle 2004）。生物学的自然主義において、意識状態はニューロンやシナプスよりも高次のレベルで存在しつつ、神経生物学的过程に因果的に還元できるとされる（Searle 2004, 邦訳 p. 154）。つまり意識状態はニューロンやシナプスから生物学的システムによって創発的に構成され、その意識状態は神経科学によって因果的に説明可能である。だがその一方でサールは、意識における一人称的で主観的な現象は、ニューロンのふるまいなどの三人称的で客観的な事実によっては説明できないと主張する（Searle 2004, 邦訳 p. 133/161）。ここでサールは「私の色の経験」などの私秘的な経験を一種の存在者（entity）と捉えて、意識は神経生物学的な基盤に「存在論的に」還元できない、と表現する（Searle 2004, 邦訳 p. 153/161）。つまり意識は科学的な基礎理論によって、ある面では説明可能であるが、すっ

かり説明し尽くされるわけではないのである。ファインバーグ＆マラットの神経生物学的自然主義は、サールの生物学的自然主義を踏襲しつつ、神経生物学的・進化生物学的な証拠をふんだんに盛り込んで、意識の進化的起源について説得的に論じている（Feinberg & Mallatt 2016）。とはいえ生物学的自然主義にしろ神経生物学的自然主義にしろ、現象的意識に独自の存在論的身分を認めることには多くの難点が指摘されている（Collins 1997; Corcoran 2001; Kim 1995; Ota et al. 2024）。なかでも Corcoran (2001, p. 312) は、サールの生物学的自然主義は結局のところ「生物学的性質二元論（biological-property dualism）」とでも呼べるような暗黙的な二元論であると指摘する。

唯物論を貫徹するなら、現象的意識は存在論的あるいは形而上学的な意味において存在しないと見るべきである。その結果は、現象的意識は一種の幻想であるとみなす幻想主義（illusionism, cf. Frankish 2017; 篠崎 2024）に接近することになる。急いで補足すると、これは現象的意識がいかなる意味でも存在しないということにはならない。たとえば人間個体は多数の細胞から構成されているわけだが、実在するのはこうした多数の細胞だけであって、人間個体は本当は実在しないのだろうか。さらに、細胞は生体分子によって構成されているわけだが、だとすれば細胞も実在しないのだろうか。あるいは植物の光合成は、植物細胞を構成する生体分子によって実現される創発的現象であるが、光合成は実在しない現象なのだろうか。そうではないだろう。自然主義においては、人間や細胞、光合成は、物理的基盤によって構成された高次レベルにおける構築物や現象として実在的（real）であるとみなせる。

認知神経科学者かつ心の哲学者であるアンティ・レヴォンスオは、主観的意識は実在的現象であり、かつ自然現象の一種であるところの生物学的現象であるとする生物学的

実在論 (biological realism) を主張している (Revonsuo 2006, p. xvii)。ここで言われている実在もまた、光合成と同じように、物理的基盤によって構成されつつ高次レベルにおいて現れる現象として捉えるべきであって、そうでなければ自然現象とは呼べない。レヴォンスオの生物学的実在論は、現象的意識は物理主義的に還元された結果として科学的説明から消し去られるとする消去的唯物論 (eliminative materialism, cf. 金杉 2003) への対抗として提出されたものである (Revonsuo 2006, p. xvii)。ただし、消去的唯物論は消去主義 (eliminativism) の一種なのだが (金杉 2003, p. 66)、消去主義においても存在者消去主義 (entity eliminativism) と言説消去主義 (discourse eliminativism) の区別がある (Irvine & Sprevak 2020, p. 349)。すなわち存在者消去主義は、存在すると想定される存在者の一覧から特定の存在者を排除すべきだという主張である一方で、言説消去主義は、特定の語りかたや考え方を科学から排除しようとする (loc. cit.)。これは本稿における構成的還元と認識的還元の区別とパラレルに理解できるだろう。幻想主義は、消去的唯物論の後継でありつつ、存在者消去主義であって、必ずしも言説消去主義であるわけではない (Frankish 2017, p. 21; 篠崎 2024)。つまり言説消去主義をとらない幻想主義は、特別な存在者としての現象的意識は否定するが、現象的意識を（ある種の便宜的な概念として）科学的に語ったり考えたり研究したりすることは否定しない。そして自然科学的に許容可能な存在者のみから構成されたものとして現象的意識を捉えるのであれば、この種の幻想主義とレヴォンスオの生物学的実在論を整合的に捉える余

地はあるように思われる³。

またレヴォンスオは、生物学における「マルチレベルの説明 (multilevel explanation)」という枠組みを構築し、意識を科学的に説明する際にも、この枠組みを適用することを提案する (Revonsuo 2006, p. 13–27; 図 3)。

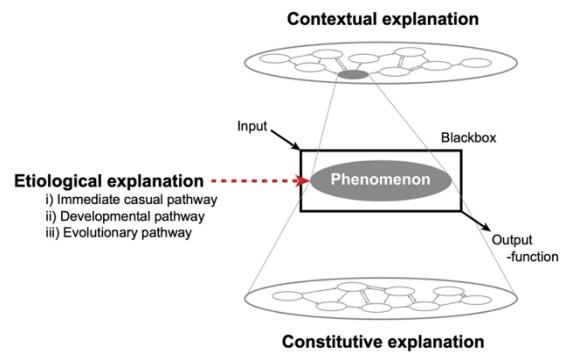


図 3. マルチレベルの説明 (Revonsuo 2006 より改変)

マルチレベルの説明の枠組みにおいて、説明対象の現象は、大きく分けて A) 構成的説明 (constitutive explanation)、B) 原因論的説明 (etiological explanation)、C) 文脈的説明 (contextual explanation) の観点から説明される。

A) 構成的説明は、組織化レベルにおける下位レベルのメカニズムについて説明する、「下向きの (downward-looking)」説明様式である。たとえば細胞の構造やふるまいを、その細胞を構成する細胞小器官や生体物質のはたらきによって説明するような場合である。

B) 原因論的説明は、説明対象の因果的歴史を説明する、「後ろ向きの (backward-looking)」説明様式であり、その時間スケールによって、i) 即時的 (直接) 原因 (immediate/proximal cause)、ii) 発生的原

³ なお言説消去主義に依拠しつつ存在者消去主義に立ち入らない立場の例として、行動主義心理学がある。行動主義心理学において、行動は心的出来事や内的心理学的プロセスを究極的に参照することなく記述あるいは説明でき、行動の誘発源は内的 (心や頭のなか) ではなく外的 (環境中) に存在するとされる

(Graham 2000/2023)。つまり意識や欲求などの心的な概念は行動主義心理学の記述や説明から消去されるのだが、意識や欲求が存在論的／構成的に消去されるべきどうかについては明示的には考慮されず、存在者消去主義については中立である。

因 (ontological/developmental cause)、iii) 進化的 (究極) 要因 (evolutionary/ultimate cause) に細分される。例えばヒトの視覚は、即時的原因としては、視覚刺激が網膜で受容され、視神経を介して主に一次視覚野 (V1) に伝達され、それが高次視覚中枢で処理されて……というように説明される。また発生的原因としては、眼球の発生や視神経の形成に基づいて説明されるだろう。そして進化的要因としては、視覚がいかにして進化してきたのか、またそれがどのような意味で適応的であったのかといった観点から説明される (cf. Parker 2003)。

C) 文脈的説明は、説明対象が高次レベルでどのような役割を果たしているのかについて説明する、「上向き (upward-looking)」の説明様式である。ひとつの例としては、細胞や個体の生命現象における DNA の役割が挙げられる。DNA はそれ自体で生きているわけではないが、生命現象において重要な役割を果たしているのである。

さらにレヴォンスオは、これらの説明のほかに入力と出力の記述 (input-output description) を付け加える。ここで説明対象はブラックボックス化され、入力と出力の機能／関数 (function) 的関係が記述される。だが、こうした機能／関数は説明対象の因果的メカニズムについて何も明らかにしないため、説明とはみなされない。

重要なのは、以上の 3 つの説明様式は相補的であり、マルチレベルで説明することで初めて十全になるという点である。言い換えれば、物理主義のようにひとつの還元的な説明だけでは意識を含めた生命現象は説明し尽くすことはできない。説明対象の現象は、複数の組織化レベルにまたがる因果ネットワークのなかに埋め込まれているのであって、3 つの説明様式は、どの方向 (下向き／後ろ向き／上向き) から見て説明対象の現象が因果の網のなかに位置づけられるのかに対応しているのである (Revonsuo 2006, p.

26–27)。

したがってマルチレベルの説明の枠組みを採用するレヴォンスオの生物学的実在論は、認識的還元をある面で許容しつつ、しかし構成的説明によって説明し尽くされるとは考えない。したがって生物学的実在論は、弱い創発に基づく自然主義に位置づけられるだろう。

類似の枠組みは、「ティンバーゲンの 4 つの問い合わせ (Tinbergen's four questions)」として生物学で広く用いられている。ティンバーゲン自身は 4 つの問い合わせを「個体発生 (ontology)」「因果性 (causation)」「進化 (evolution)」「生存価 (survival value)」と呼び分けたが (Timbergen 1963)、こんにちではそれぞれ、至近要因としての発生 (development) とメカニズム (mechanism)、究極要因としての進化史 (evolutionary history) と適応的機能 (adaptive function) と呼び分けられることが多い (Bergman & Beehner 2022)。進化的原因がふたつに分かれている以外は、レヴォンスオのマルチレベルの説明の枠組みにおける原因論的説明と同じである。

このような複数の説明様式を積極的に認める態度は、認識的多元論 (epistemic pluralism) と呼ぶことができる。この立場は、本稿の第 1 節で触れた方法論的還元についても、アプローチのひとつとして許容するがそうではないアプローチも不可欠であるという多元論的な立場をとるだろう。20 世紀の生物学は確かに分子生物学という還元主義的アプローチによって目覚ましい発展を遂げた。しかし、目下のところは分子の言葉だけで生命現象を説明し尽くせてはいないし、仮に今後それが可能になったからといって、生態学や動物行動学、進化生物学などのマクロ分野が、分子生物学のみに、ひいては有機化学、物理学に吸収されてしまう（生物学の諸概念が消去されてしまう）とは考えにくい。意識についても同様に、認識的多元論に基づく創発的唯物論（自然主義）は非常に有望な

立場であると思われる。

4. 認識的多元論に基づく創発的唯物論の可能性

しかしながら創発的唯物論（自然主義）は、還元的唯物論と安易に同一視され、不当な批判を受けることが多い。たとえば現代ドイツの学者マルクス・ガブリエルは、実際に自然主義を唯物論と同一視して、次のように批判する。

私たちは唯物論を克服しなければなりません。唯物論は（物質エネルギーにも基づき、匿名の固い原因で成り立つ現実という意味での）宇宙に見出せるものしか存在しないと私たちに吹き込み、それゆえ意識からニューロンの嵐にまで還元することができる精神というコンセプトを必死で求めているのです（Gabriel 2015, 邦訳 p. 348）

動物の世界の一員である人間は、基本的・本質的に生物学的機械であり、その目的はわれわれが知るあらゆる生命体の目的と同じである——こうした考え方は、貧弱なデータを乱暴に一般化しすぎたものでしかない。 [...] 脳が、もっと正確にいえば一個の生物が [...]、精神という、歴史に開かれた説明構造に加わるための必要条件であることは事実である。けれども、だからといって、精神が自然の秩序のなかに座を占める事ができるという考えにはならない。現象の一切合財をひとつのフレームワークのなかに収めて、存在の問題を一一解決しようとするのは、おそらく見当外れな試みでしかない。たしかに多くのことが実在する。しかし、実在するあらゆる事物を含む一個のもの——ひとつの実在一があるという話にはならないのだ（Gabriel 2018, 邦訳 pp. 73–78）。

これは構成的還元と認識的還元を混同している典型例である。実際、ガブリエルのターゲット論文とそれに対する複数の著者からのコメント一論文、さらにガブリエルの返答からなる論文集である『新実存主義』の序文で、ジョスラン・マクリュールは以下の指摘を加えている。

ガブリエルは自然主義を、何よりもまず、心を自然種 [natural kind, この箇所に付された訳注によれば、規約による分類によらない、自然のなかに存在するとされる事物のカテゴリー] とみなして物理的メカニズムへと還元する選択と結びつける。だが彼によれば、この選択は当を失している。 [...] 第1章 [ガブリエルによるターゲット論文] の批判は、自然主義ではなく還元論が真の標的だったとも読める。たしかに自然主義者は、心が「自然の秩序」に「収まる」と考える。だが、存在するものはすべてが自然種へと還元可能だとする一元論的存在論に与する必要はないし、心にまつわる概念はみな科学の進歩にともなって次第に消え去っていくだろうと考える必要もない [...] 私が正しければ、新実存主義をさらにくわしく展開するには、分析系の心の哲学が提供する有力な——と私には思える——選択肢を正面から取り上げ、それに継続して取り組むことが不可欠である。念頭にあるのは、「非還元論的物理主義」「性質二元論」「創発主義」といった、おおよそ重なり合う立場だ（Maclare 2018 邦訳 p. 4–8）。

ただし本稿の分析に照らせば、マクリュールもまた構成的還元と認識的還元の区別を徹底できていない様子がうかがえる。自然種の存在論的身分それ自体が大きな哲学的問題を内包しているものの (cf. Suzuki & Saigo 2025)、「自然種への還元」は本稿の区分における唯物論一般において存在論的／構成的

還元よりも認識的還元の文脈で扱われるべき問題であるように思われる。唯物論において自然種は、その物理的基盤とは別個の存在者ではなく、あくまで物理的世界のなかで認識される特定のパターンとみなされるであろうからだ（認識的還元への態度によって、創発的現象として実在すると見るか、消去されるべきものと見るかは変わるだろうが）。

また「非還元論的物理主義」と「創発主義」は確かにおおよそ重なり合うと言えるが、両者と「性質二元論」は存在論的／構成的還元の軸において区別される。一時期のチャーマーズが自然主義的二元論（naturalistic dualism）を標榜した点（Chalmers 1996, 邦訳 p. 168）、また上述のようにサールの生物学的自然主義が二元論的性格を帶びている点など、混乱を招く議論も確かに多い。だが、このように「還元」と「反還元」の一次元の軸に圧縮すると近い位置にあるように見えるものも、本稿で論じてきたように存在論的／構成的還元と認識的還元の二次元でプロットすると（まるで二次元クロマトグラフィーのように）うまく分離されることがある。

すでに指摘したとおり、ポパーも同様の混同から唯物論全体を退けてしまっている。だが創発的唯物論においては、ポパーの重視する下方因果や非決定論を包摂できる可能性がある。

ここで、因果に関する存在論的／構成的な側面と認識的な側面を区別しよう。すなわち、存在者間にどのような因果関係があるのかという問題と、それを我々がどのように認識するかという問題を、分けて考えるのである。

Potochnik (2017, pp. 35–42) が指摘するように、この世界の現象は因果的に複雑に絡み合っている。そして科学は、理想化による単純化（simplification by idealization）を通して特定のパターンを見出そうとする。つまり科学とは、因果的複雑性（causal complexity）に直面する我々が、さまざまに

因果パターン（causal pattern）を探す営みとして理解されるのである（ibid.）。また因果パターンとは現象の規則性（regularity）であり、その意味で因果パターンは実在する（Potochnik 2017, p. 25; cf. Dennett 1991）。

このように因果を捉えると、唯物論に基づいて因果的閉包のなかで下方因果を位置づけることができる（図 4）。たとえば時刻 T1 に「のどが渴いた」と思ったメアリーが、時刻 T2 に冷蔵庫からピッチャーに入ったオレンジジュースを取り出して飲んだ場合を考える。このとき、時刻 T1 の「のどが渴いた」という心的状態 M1 の物理的背景として物理状態 P1 が、時刻 T2 の「冷蔵庫からオレンジジュースを取り出して飲む」という身体動作の物理的背景として物理状態 P2 が想定される。そして物理的世界には因果的閉包があり、心的状態 M1 はその壇外のものと捉えた場合、因果関係があるのは P1 から P2 への過程であって、M1 から P2 への過程は下方因果として排除されてしまう。

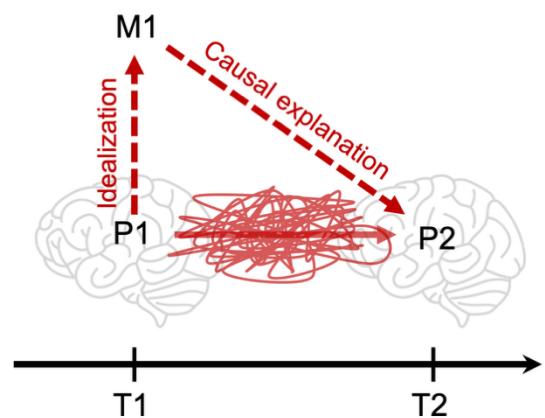


図 4. 複雑な因果性から取り出された因果パターンとしての下方因果

しかしここで、P1 から P2 への因果関係が非常に複雑であることに目を向けよう。実際のところ、渴水状態にあるメアリーの脳内でどのような神経メカニズムがはたらいて、冷蔵庫にオレンジジュースがあるという記憶を想起し、歩行を制御して冷蔵庫に向かい、手をうまく操って冷蔵庫のドアを開け、（主に

視覚情報に頼って) さまざまな物品のなかからオレンジジュースの入ったピッチャーを見つけ出し、中身がこぼれないようにコップへと移し、口に運んで飲む……といった一連の動作を実現するのか、事細かに記述しようとしても煩雑にすぎる。そもそも、そのすべてを分子レベルまで把握しようというのは「人間という限りのある存在 (limited human beings)」(Potochnik 2017, p. 1) には到底不可能である。

そこで心的状態 M1 を、物理状態 P1 を理想化・単純化して (いささか素朴ではあるが) 「(のどが渴いたので) 水分を摂りたい」という欲求——あるいは「メアリーは (のどが渴いたので) 水分を摂りたいと欲している」という命題的態度——として措定されたものだと考えよう。こうした飲水欲求は因果パターンとして規則的に現れ (のどが乾いた動物は、水分を飲もうとするものだ)、物理的世界の因果的閉包の内部にありながら実在的である。したがって「メアリーはのどが乾いたのでオレンジジュースを飲んだ」という M1 から P2 への因果的説明は、「あたかも」下方因果的でありつつ、認識的に妥当する。このように考えれば、創発的唯物論は下方因果を包摂できる。

創発的唯物論が決定論だけでなく非決定論をも包摂しうる点についても、存在論的／構成的な側面と認識的な側面の区別によって同様に示すことができる。たとえばマクロスケールの物理現象が、ニュートン力学などによって決定論的にうまく記述できるとしても、物理的世界それ自体が決定論的であるとは限らない。にもかかわらず物理的世界を決定論的なものとみなすのは、物理的世界の認識を物理的世界それ自体に押し付けているに過ぎない。

また、決定論をとらないからといって因果性を否定することにもならない。西郷&田口 (2019, p. 199) は、決定論があまりにも広汎に共有されていることにより、決定論と因果

性がしばしば同一視されてしまいがちだが、この二つは本来異なるものであると指摘する (cf. 西郷 2024)。また西郷&田口 (2019, p. 220) は次のように述べる。

非可換確率論〔量子確率論〕で記述できるような現象は、決して量子レベルの現象に限定されるわけではなく、むしろ現実の一般的なあり方により近いと考えられる。そのなかのごく限定された現象だけに、古典的な確率論が成立立つ。その古典的な確率論のごく特殊なケースが決定論である。

[...] ミクロは非決定論的で、マクロになるにつれてだんだん決定論が妥当するという単線的な構造を示しているのではなく、むしろ、より決定論的な層とより非決定論的な層とが入れ替わりつつ積み重なっていると見るべきかもしれない。たとえば量子現象、生命現象や意識現象といったものはスケールを異にする非決定論的な層ということもできるだろう。

つまり先のメアリーの例に戻れば、物理状態 P1 から物理状態 P2 への因果関係は、複雑であるだけでなく、非決定論的ですらあるのかもしれない。さらに踏み込めば、生命現象や意識現象は古典的な意味で (つまり決定論的に) 予測することもできないのかもしれない。第 1 節で触れた「生命有機体の性質の多くは予測が不可能つまり創発的であって、宇宙に対する『自然な』見解は非決定論的である」というポパーの見解は核心を突く指摘ではあったが、だからといって実体二元論や実体多元論をとる必要はない。認識的多元論に基づく創発的唯物論でも十分に下方因果や非決定論を扱えるのであって、それにより既存の「堅固な」諸科学との対立は回避される。

したがってポパーの三世界論における世界 2 と世界 3 は、世界 1 と独立の世界として実在するのではなく、物理的世界のなかに埋め

込まれた (embedded) 認識的領域として再定式化するのが妥当である。実際、オートポイエーシス論に基づいて世界 2 と世界 3 が物理的現実としての世界 1 に埋め込まれたものとして三世界論を解釈する試みも提案されている (Hall 2005)。

なお少し話が逸れるが、本稿では一貫して一種の唯物論 (materialism) を擁護してきたとはいえ、物質 (material) を究極的な存在者とする物質主義としての materialism を支持するわけではない。物質も、それを構成する分子や原子も、物理プロセスの産物にすぎない。これらは文脈に応じて個物 (トーケン) である場合も自然種 (タイプ) である場合もあるが、そもそも個物か自然種かという、西洋哲学で伝統的に前提とされてきた存在者カテゴリーを無批判に受け入れる必要はない。むしろプロセスが一次的であって、個物および自然種は我々が二次的に認識するものだとも考えられる (Suzuki & Saigo 2025)。この点について本稿は、

生命はモノではなくプロセスである。わたしたちは、わたしたち自身をモノとして考えすぎている。わたしたちはモノではなくプロセスである。 [...] 生命は高分子ではない。生命は明らかにプロセスである。それは進行中のプロセスである (Popper et al. 1993, p. 169; cf. 池田 2024)

と述べたポパーとむしろ軌を一にする。しかしポパーが、物的プロセスと心的プロセスという、相互作用しつつ独立なふたつのプロセスが実在する物心プロセス二元論を唱えたのに対し、本稿では物的プロセスしか存在しないと考える。したがって本稿の立場は、自然主義的プロセス一元論 (naturalisitic process monism) あるいはプロセス自然主義 (processual naturalism) と呼ばれるべきであろう。

5. 世界 2 と世界 3 の進化的創発

ポパーの三世界論が下方因果と非決定論を包摂する創発的唯物論に換骨奪胎されうるとしても、世界 2 と世界 3——近似的には、主観的意識 (subjective consciousness) と象徴言語 (symbolic language) ——がいかにして創発的に進化したのかが、創発的唯物論に基づいて提示される必要がある。一般に創発主義者がある現象について「それは創発的現象である」と述べたところで、これ自体は説明の体をなさない。Reber (2019, pp. 114–118) はこれを「創発主義者のジレンマ (emergentists' dilemma)」と呼んだ。

したがって創発主義者は、創発現象について「それは創発によるものである」以上の説明を加えなければならない。強い創発であるなら還元的な説明やメカニズム的説明を与えるのは定義からして不可能であるが、弱い創発であるならばそれが可能である。

実際、科学的知見を参照しながら主観的意識の進化的創発を考察する議論が、ここところ矢継ぎ早に提出されている (Feinberg & Mallatt 2006; Ginsburg & Jablonka 2019; Godfrey-Smith 2017, 2020; Veit 2024)。こうした「新たな生物学的自然主義」(ただしサークルの生物学的自然主義とは異なり、一人称的で主観的な現象に必ずしも独自の存在論的身分を与えるとは限らない) とも呼べるような動物意識論の潮流については、すでに概説がいくつか公刊されているので、そちらを参考されたい (鈴木 2021; Suzuki 2022; 小草 & 新川 2024)。

そのなかでもとくに Ginsburg & Jablonka (2019, 邦訳下巻 pp. 244–259) は、象徴言語の進化についてもスケッチ的ではあるが描像を与えていている。また生命記号論 (biosemiotics) も動物意識論とは異なる研究伝統として、生命誕生から心や言語の獲得に至る進化に統一的な視座を与えようと試みてきた (Hoffmeyer 2008; Sharov 2018)。つまりポパーの三世界論における世界 3 の

進化的創発についても、自然主義的アプローチによる説明の模索がすでになされているのである。

とはいって、これらの研究はいまだ途上段階にあり、主観的意識や象徴言語の進化的創発について十分な理解が得られているとは現時点では言いがたい。だが第3節および第4節で論じたように、主観的意識や象徴言語の創発という複雑な現象については、何かひとつの（たとえば還元的な）説明を与えるべきではない。この世界は、あるひとつの科学領域（たとえば物理学）だけで理解しきれるほど単純ではないのである。創発的唯物論によるアプローチは、有望な研究プログラムとして、理論的側面と経験的知見の側面の両方でさらなる研究や議論が期待される。

【謝辞】

講演および本稿執筆の機会をいただきまして、松尾洋治教授および日本ポパー哲学研究会のみなさまに感謝いたします。また本稿執筆に先立って催された日本ポパー哲学研究会2025年度第35回年次研究大会において、本稿の草稿に基づく筆者の講演に対して質問やコメントをくださった参加者の皆様にも感謝の意を表します。また、森元良太博士、田中泉吏博士、篠崎大河氏、佐々木健人氏には、本稿の草稿をご覧いただき、有意義なコメントを頂戴いたしました。この場をお借りして拝謝いたします。なお本研究の一部はJSPS科研費24H01538の助成を受けて実施されました。

【文献】

- Andersen HK (2017) Reductionism in the biomedical sciences. In: Solomon M, Simon JR, Kincaid H (eds.) *The Routledge Companion to Philosophy of Medicine*. Routledge, pp. 81–89.
- Bechtel W, Richardson R (1993) *Discovering Complexity: Decomposition and Localization as Strategies in Scientific Research*. Princeton University Press.
- Bergman TJ, Beehner JC. (2022) Leveling with Tinbergen: Four levels simplified to causes and consequences. *Evolutionary Anthropology* 31(1):12–19.
- Brigandt I, Love A (2008/2023) Reductionism in biology. In: Zalta EN, Nodelman U (eds.) *The Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/archives/sum2023/entries/reduction-biology/>.
- Bunge M (1980) *The Mind-Body Problem*. Pergamon. 邦訳：マリオ・ブンゲ／著、黒崎宏、米澤克夫／訳（1982）『精神の本性について：科学と哲学の接点』産業図書。
- Cea I, Negro N, Signorelli CM (2023) The Fundamental Tension in Integrated Information Theory 4.0's Realist Idealism. *Entropy* 25(10):1453.
- Chalmers DJ (1996) *The Conscious Mind: In Search of a Fundamental Theory*. Oxford University Press. 邦訳：デイヴィッド・J・チャーマーズ／著、林一／訳（2001）『意識する心：脳と精神の根本理論を求めて』白揚社。
- Collins C (1997) Searle on consciousness and dualism. *International Journal of Philosophical Studies* 5(1):15–33.
- Corcoran K (2001) The trouble with Searle's biological naturalism. *Erkenntnis* 55(3):307–324.
- Dennett DC (1991) Real patterns. *Journal of Philosophy* 88(1):27–51.
- Edmonds D (2020) *The Murder of Professor Schlick: The Rise and Fall of the Vienna Circle*. Princeton University Press. 邦訳：デイヴィッド・エドモンズ／著、児玉聰、林和雄／監訳、杉村文、鈴木英仁、高木博登、立場貴文、中村貴行、中村達樹、

- 森田初音、吉田隼大／訳（2025）『シュリック教授殺害事件：ウィーン学団盛衰史』晶文社。
- Feinberg TE, Mallatt JM (2016) *The Ancient Origins of Consciousness: How the Brain Created Experience.* MIT Press. 邦訳: トッド・E・ファインバーグ、ジョン・M・マラット／著、鈴木大地／訳（2017）『意識の進化的起源：カンブリア爆発で心は生まれた』勁草書房。
- Frankish K (2017) *Illusionism as a theory of consciousness.* In: Frankish K (ed.) *Illusionism as a Theory of Consciousness.* Imprint Academic, pp. 11–39.
- Gabriel M (2015) *Ich ist nicht Gehirn: Philosophie des Geistes für das 21. Jahrhundert.* Ullstein. 邦訳: マルクス・ガブリエル／著、姫田多佳子／訳（2019）『「私」は脳ではない：21世紀のための精神の哲学』講談社（講談社選書メチエ710）。
- Gabriel M (2018) Neo-Existentialism: How to Conceive of the Human Mind After Naturalism's Failure. In: Gabriel M (main author), Maclure J (ed.) *Neo-Existentialism.* Polity Press, pp. 8–45. 邦訳所収: マルク・ガブリエル／著、廣瀬覚／訳（2019）『新実存主義』岩波書店（岩波新書 新赤版 1822）。
- Ginsburg S, Jablonka E (2019). *The Evolution of the Sensitive Soul.* MIT Press. 邦訳: シモーナ・ギンズバーグ、エヴァ・ヤブロンカ／著、鈴木大地／訳（2021）『動物意識の誕生：生体システム理論と学習理論から解き明かす心の進化（上・下）』勁草書房。
- Graham G (2020/2023) Behaviorism. In: Zalta EN, Nodelman U (eds.) *The Stanford Encyclopedia of Philosophy.* <https://plato.stanford.edu/archives/spr2023/entries/behaviorism/>
- Hall WP (2005) Biological nature of knowledge in the learning organization. *The Learning Organization* 12(2): 169–188.
- Hoffmeyer J (2008) *Biosemiotics: An Examination into the Signs of Life and the Life of Signs.* University of Scranton Press.
- 池田健人（2022）「知識の成長について：ポパーの三世界論における内実とその展望」『科学基礎論研究』50(1):1–15.
- 池田健人（2024）「ポパーの相互作用主義における存在論：心の主観性と実体性を超えて」『批判的合理主義研究』15(2):51–61.
- Irvine E, Sprevak M (2020) Eliminativism about consciousness. In: Kriegel U (ed.) *The Oxford Handbook of the Philosophy of Consciousness.* Oxford University Press, pp. 348–370.
- 金杉武司（2003）「解釈主義と消去主義：命題的態度の実践的実在性」『哲学・科学史論叢』5:63–99.
- Kim J (1995) Mental causation in Searle's "Biological Naturalism". *Philosophy and Phenomenological Research* 55(1):189–194.
- McGinn C (1999) *The Mysterious Flame: Conscious Minds In A Material World.* Basic Books. 邦訳: コリン・マッギン／著、石川幹人、五十嵐靖博／訳（2001）『意識の〈神秘〉は解明できるか』青土社。
- Maclure J (2018) Introduction: Reasonable Naturalism and the Humanistic Resistance to Reductionism. In: Gabriel M (main author), Maclure J (ed.) *Neo-Existentialism.* Polity Press, pp. 1–7. 邦訳所収: マルク・ガブリエル／著、廣瀬覚／訳（2019）『新実存主義』岩波書店（岩波新書 新赤版 1822）。
- Nagel T (1998) Reductionism and antireductionism. In: Bock GR, Goode

- JA (eds.), *The Limits of Reductionism in Biology*. Chichester: John Wiley & Sons, pp. 3–10.
- O'Connor T (2020/2021) Emergent Properties. In: Zalta EN, Nodelman U (eds.) *The Stanford Encyclopedia of Philosophy*.
<https://plato.stanford.edu/archives/win2021/entries/properties-emergent/>
- Ota K, Suzuki DG, Tanaka S (2022) Phylogenetic distribution and trajectories of visual consciousness: examining Feinberg and Mallatt's neurobiological naturalism. *Journal for General Philosophy of Science* 53(4):459–476.
- Perker A (2003) *In the Blink of an Eye : The Cause of the Most Dramatic Event in the History of Life*. Perseus. 邦訳：アンドリュー・パーカー／著、渡辺政隆、今西康子／訳 (2006) 『眼の誕生：カンブリア紀大進化の謎を解く』草思社。
- Potochnik A (2017) *Idealization and the Aims of Science*. University of Chicago Press.
- Popper KR, Eccles JC (1977) *The Self and its Brain: An Argument for Interactionism*. Springer. 邦訳：カール・R・ポパー、ジョン・C・エクルズ／著、大村裕、西脇与作／訳 (2006) 『自我と脳（上・下）』思索社。
- Popper K, Lindahl BIB, Århem P (1993) A Discussion of the Mind-Brain Problem. *Theoretical Medicine* 14(2):167–180.
- Revonsuo A (2006) *Inner Presence: Consciousness as a Biological Phenomenon*. MIT Press.
- 西郷甲矢人 (2024) 「蒔かぬ種は生えぬ：因果性についての試論」『哲学論叢』 53:1–15.
- 西郷甲矢人、田口茂 (2019) 『〈現実〉とは何か』筑摩書房 (筑摩選書 182)。
- Reber AS (2019) *The First Minds: Caterpillars, Karyotes & Consciousness*. Oxford University Press.
- Sarkar S (1992) Models of reduction and categories of reductionism. *Synthese* 91(3):167–194.
- Searle J (2004) *MiND: A Brief Introduction*. Oxford University Press. 邦訳：ジョン・R・サークル／著、山本貴光、吉川浩満 訳／訳 (2006) 『MiND：心の哲学』朝日出版社。
- 篠崎大河 (2024) 「意識の哲学における幻想主義の展開」『科学哲学』 57(1):57–76.
- Sharov A (2018) Mind, Agency, and Biosemiotics. *Journal of Cognitive Science* 19(2):195–228
- 鈴木大地 (2020) 「意識研究の思想地図 2020 β 版」『note』
https://note.com/suz_dg/n/ne9847ebd1ff2 (2020年2月21日公開、2025年9月1日閲覧)
- 鈴木大地 (2021) 「意識の起源と進化」『生体の科学』 73(1):34–37.
- Suzuki DG. (2022). A general model of and lineage-specific ground plans for animal consciousness. *Annals of the Japan Association for the Philosophy of Science*, 31, 5–29.
- Suzuki DG, Saigo H (2025, preprint) A category-theoretic interpretation of the homology concept in biology. *PhilSci-Archive* 25010. <https://philsci-archive.pitt.edu/25010/>
- Tinbergen N (1963) On Aims and Methods of Ethology. *Zeitschrift für Tierpsychologie* 20(4):410–433.
- Veit W. (2023). *A Philosophy for the Science of Animal Consciousness*. Taylor & Francis.

『新刊紹介』小河原誠著『政治のロジックを問いただす—批判的合理主義の観点から—』あけび書房、2025年

富塚嘉一(中央大学名誉教授)

本書は、K.R.ポパー哲学研究の第一人者である小河原誠氏が、政治の世界でよく語られている主張を題材として、ポパーの批判的合理主義の観点からそのロジックを批判的に分析・解明した意欲作である。一見すると、政治活動や日常生活は科学哲学における議論とは無縁のように思われるが、そのロジックを掘り下げて行くと、哲学や論理学の問題と深く結びついていることが解き明かされる。しかも、そこでの判断のよりどころとして批判的合理主義の考え方がある基準となることが特徴的である。

このように、本書は科学哲学の専門家というよりも、広く一般の人々に向けて分かり易く書かれた本であり、その内容については、以下に示すように、各章のタイトルを見ればある程度のイメージができる。

- 第1章 理由といってもどんな理由なの?
—「なぜ」と問うのではなく、「どうなるか」を問え—
- 第2章 両論併記をこえる途 報道は反証主義の立場で 一報道は中立・公平でなければいけないのか—
- 第3章 「……すべし」と「……できない」、そして「……するつもり」
- 第4章 パラドックスまがいのことでも避けましょう
- 第5章 陰謀論を思想史の文脈で見る
- 第6章 不正選挙論はどこが間違っているのか
- 第7章 弁証法をとらえ直す
- 第8章 定義はものごとを明確にするか
- 第9章 賛成の数が多いといっても何ひとつ価値のある証拠にはならない 一数の論理への批判—

加えて、「まえがき——本書はなにを論じているのか」において、その趣旨および各章の要点が簡潔にまとめられているので、読者は本書の概要を大ざかに理解することができる。

ただし、各章のテーマに注目してみると、そこで指摘は、われわれの“常識”とは食い違っていると感じられるのではないか。各章では、政治の世界あるいは日常生活において当然のこととして受け入れられている主張が、科学哲学のロジックにもとづいて批判的に吟味され、まったく逆の結論が導き出されているのである。分析の過程では、科学哲学の論理や過去の哲学者たちの見解などやや難しい議論も紹介されているが、あくまで身近な事例を手がかりとして論じられているので、興味を維持しつつ読み進めることができる。このように全体を通して著者の意図が伝わってくるように展開されている。

そして、批判的分析の基礎となっているのが、まさにポパー哲学である。知識の正当化の試みは、独断的断定、循環論、あるいは無限後退に陥るという三方ふさがり(トリレンマ)に直面してしまう。これを回避するには、すべての知識は誤り得るとの可謬論に立ち、しかも、あれも良しこれも良しとする両論併記による相対主義の殻にとどまることも避けて、批判的討論を続けることで少しでも真理に近づこうと努力する立場(真理接近説、認識進歩説)が推奨される。こうすることで、政治や日常生活で展開されるロジックを改めて考え直してみるきっかけが得られるというのが本書のメッセージである。

これまで科学哲学に興味のなかった人たちも、改めてじっくりと考える機会となれば幸いであり、さらには科学哲学、とりわけポパー哲学に関する文献を参照してみることが期待される。小河原氏はまさにそれに相応しい「チチェローネ(案内人)」といえよう。

【著者からのコメント】

著者としては、第一章では非正当化主義を、前提からの結論の導出という枠組みで論じ、前提となるものが何ら正当化されうるものでないことをニーチェの主張などにそくして論じたつもりであります。また、第二章ではこの枠組みのもとで、背理法を利用して両論併記主義的なジャーナリズムを批判しております。非正当化主義についてのパートリー的「定式化」よりは一步前進しているのではないかと思っています。

『新刊情報』カール・ポパー著、蔭山泰之訳『出発点 私の知的発展』(岩波文庫)



ウィーンに生まれ、戦間期の苦境を経験したカール・ポパー（1902-1994）は、哲学者としての人生をいかにして歩み始め、その批判的合理主義はどのように生み出されたのか。自らの人生を振り返りながら、広範囲に及ぶ思想の全体像を語る自伝で、ポパー入門にふさわしい。反証可能性を論じた付録も収録。ドイツ語版の翻訳。

目 次

凡 例

はじめに

- 1 全知と人間の可謬性
- 2 幼少期の思い出
- 3 実家の影響
- 4 第一次世界大戦
- 5 少年期の哲学的問題——無限
- 6 哲学での最初の失敗——本質主義の問題
- 7 本質主義についての長い余論——私と同世代の大半の哲学者を隔てているもの
- 8 決定的な年——マルクス主義、科学、疑似科学
- 9 初期の研究
- 10 二度目の余論——独断的思考と批判的思考、帰納なしでの学習
- 11 音 楽
- 12 ポリフォニー音楽の起源についての考察——探究の心理か、探究の論理か

- 13 二種類の音楽
- 14 芸術、とくに音楽における進歩の概念
- 15 大学での最後の年
- 16 認識論——『探究の論理』
- 17 論理実証主義は死んだ——だれが殺したのか
- 18 実在論と量子論
- 19 客観性と物理学
- 20 真理、確率、驗証
- 21 迫りくる大戦と「ユダヤ人問題」
- 22 移住——イギリスとニュージーランド
- 23 ニュージーランドでの最初の活動
- 24 『開かれた社会とその敵』と『ヒストリズムの貧困』
- 25 ニュージーランドでのその他の仕事
- 26 イギリス、ロンドン・スクール・オブ・エコノミクス・アンド・ポリティカル・サイエンスにて
- 27 イギリスでの初期の仕事
- 28 合衆国への最初の訪問——アインシュタインとの出会い
- 29 問題と理論
- 30 シュレーディンガーとのディスカッション
- 31 客観性と批判
- 32 帰納、演繹、客観的真理
- 33 形而上学的リサーチ・プログラム
- 34 物理学での主観主義に対抗して——量子力学と傾向性
- 35 ルートヴィッヒ・ボルツマンと時間経過の方向性、時間の矢
- 36 エントロピーの主観主義理論
- 37 形而上学的リサーチ・プログラムとしてのダーウィニズム
- 38 世界 3
- 39 心身問題と世界 3
- 40 事実の世界にあるさまざまな価値

あとがき

一九九二年 マルクス主義へのあとがき

[付録] 論争の核心——境界設定の問題
訳者解説

ポパーの著作(抄録)

主要著作の略記

事項索引

人名索引

編集後記

今回は 2025 年 8 月に行われた研究大会シンポジウム発表の完成稿を掲載しました。

研究会のホームページも随時更新されていきますので、またご確認いただければ幸いです。

本号についてのご意見等につきましては、編集委員（現在は、志村 昌司 shojishimura@gmail.com）までご連絡いただければ幸いです。

批判的合理主義研究（通巻 30 号）

2025 年 12 月発行

本誌は、『ポバーレター』（1989～2008, 通巻 38 号）を改題し、継承したものです。

发行人 志村 昌司

編集・発行 日本ポバ一哲学研究会事務局機関
誌編集部

Tel. 090-3842-9002

Email:shojishimura@gmail.com

入退会・名簿変更、会費徴収・会計管理に関して
は、「日本ポバ一哲学研究会事務局組織・会計部」
にお願いいたします。

〒157-8570

東京都世田谷区砧 5 丁目 2 番 1 号 日本大学
商学部 瀧田研究室

email: takita.y@gmail.com

tel. 03-5368-3661

fax. 03-5368-3630