

応用倫理—理論と実践の架橋—

Vol. 9
2016 年 3 月

北海道大学大学院文学研究科
応用倫理研究教育センター

目 次

コリングリッジの技術選択論——原子力発電を手がかりとして

寺本 剛（中央大学）…………… 1

自己抹消的な道徳理論の問題点は（あるとすれば）何か

秋葉剛史（千葉大学）…………… 12

書評 眞嶋俊造・奥田太郎・河野哲也編 『人文・社会科学のための研究倫理ガイドブック』

吉永明弘（江戸川大学）…………… 30

書評 文・伊勢田哲治／マンガ・なつたか 『マンガで学ぶ動物倫理——わたしたちは動物とどうつきあえばいいのか』

浅野幸治（豊田工業大学）…………… 33

コリングリッジの技術選択論 ——原子力発電を手がかりとして

寺本 剛 (中央大学) *

要旨

現代の科学技術社会では新たな技術が続々と登場し、しかも、一つ一つの技術の進歩が社会に大きなインパクトを与えるようになっている。こうした状況下では、リスクの小さい技術を選択する方法の探求が一つの重要な倫理的課題となる。デイヴィッド・コリングリッジによる技術についての意思決定理論はこの課題に対する一つの応答と見なすことができる。コリングリッジによれば、人間の知性、分析能力、情報収集力はきわめて乏しいものであり、どんな意思決定もそれが正しいものであるかどうかを前もって知ることはできない。こうした状況では漸進的な意思決定が最良の方法であり、技術は試行錯誤学習を通じて一步一步導入されるべきだとコリングリッジは主張する。また、コリングリッジはこうした試行錯誤学習を妨げる性質を持つ技術を「柔軟でない技術」と特徴づけ、そのような技術の選択は避けられるべきだとも主張する。本稿では、これらの主張を詳しく紹介するとともに、それが不確実性下における技術選択のためにわきまえておくべき基本的な洞察であること、またそれが技術選択の指針の探求の糸口となりうることを確認する。その一方で、その理論は万能ではなく、その有効性は限定的であることも合わせて指摘する。

David Collingridge's Theory of Technology Selection

New technologies appear constantly in rapid succession in the modern world of science and technology. Through development and diversification in technology, each technology poses large, diverse and unpredictable risks to society. From the ethical point of view, effective ways to discern risky technologies are required in such a situation, to avoid negative social effects in advance. David Collingridge's theory of decision-making provides a guiding principle to assess technologies. According to Collingridge, human understanding, analytical capacities and information acquisition are so limited that no decision can be known to be correct in advance. After a decision has been made, it is threatened by the discovery of unexpected facts that may cause those who made the original decision to regret their choice. Collingridge maintains that incrementalism is the most effective way under such circumstances; new technologies should be adopted step by step through "trial and error learning". However, technologies with their four properties of long lead-time, large unit size, capital intensity and dependence upon infrastructure, have a strong tendency to prevent trial and error learning. He calls these kinds of technologies "inflexible" and argues that such technologies should not be adopted, because they are highly likely to entail large economical and social costs. This paper aims to describe Collingridge's theory in detail and to evaluate it as giving a basic insight about how to recognize risky technologies under great uncertainty, even though the effectiveness of this theory is essentially limited.

1. 本稿の趣旨

科学技術は私たちの生活を便利で豊かなものにした。その多くは社会の基盤をなしており、それなしには私たちの「普通の生活」が成り立たなくなっている。しかし、生活を支え、よりよくするはずの技術が、思わぬかたちで人びとの生命や健康を脅かし、様々な経済的・社会的損失をもたらしてきたことも歴史の教えるところである。今後、技術開発のサイクルが速まることで新たな技術が続々と登場し、しかも、そうした一つ一つの技術が社会に大きなインパクトを与えるのだとしたら、新しい技術を適切に評価し、リスクの小さい技術を選択していくいわゆるテクノロジーアセスメントの仕組みづくりが一つの重要な課題となるに違いない。この課題について哲学者や倫理学者が貢献できることのの一つが、過去の技術の発展動向や技術選択のあり方を批判的に検討し、それを踏まえて適正な技術選択の指針を探求することだと筆者は考える。こうした大枠の問題意識のもと、以下では技術の柔軟性（flexibility）を技術選択の一つの基準として提起するデイヴィッド・コリングリッジの議論を紹介し、その有効性を見定めたい。もっとも、本稿の議論を通じて明らかになるように、技術選択の指針の探求は一筋縄では行かない難問であり、本稿がなしうるのはあくまでこの探求を今後深めていくための一つの糸口を提供することにとどまる。

2. 技術選択の困難

科学技術の負の影響を回避する理想的な方法は、それが生じないよう未然に対処することである。事前に対処できれば、損害や被害を最小限に抑えることができるからだ。そのためには新しい技術のリスクを見極め、できるだけリスクの小さい技術を選択することが必要となる。しかし、技術というもののあり方を考慮したとき、こうした未然の対処には本質的な限界がある。ドン・アイディはその難しさを次のように表現している。

技術が使用や方向、そして軌道の様々な可能性を「内に含んで」おり、それらの可能性が本質的に多様安定的であるため、その効果や使用や産物をはっきり予言することは、不可能ではないにしろ困難である。そして、いったんところを得ると、技術は多様な役割指定の文脈において、意図せざる、しばしば「報復」的な作用を生み出すが、これを予言することもまた不可能ではなくとも困難なのである。¹

この世界は多様な諸要因からなる複合体であり、新しい技術がそこに組み込まれ、定着する仕方には無数の可能性がある。ここでアイディが「多様安定的（multistable）」と言っているのは技術が本質的にもっているこうした性質のことである。技術が社会に定着するあり方が多様であるとしたら、ある技術が世に出てどのような物質連関の中に位置づけられ、どのような歴史的・社会

* 本論文は、科学研究費補助金（25370032）および（26770010）による研究成果の一部である。本論文を作成するにあたって、査読者をはじめ、多くの方から有益なコメントをいただいた。記して謝意を表する。

1 Ihde (1999), p.48, 引用文の翻訳に際しては邦訳を参考にしたが、地の文との整合性を重視して訳語等を変更した部分がある。

的文脈で利用され、どのような方向性と速度で普及・発展していくか、また、その過程でどのような影響をもたらすのかということは明確には予測できないことになる。期待とともに新技術を導入しても、その期待は裏切られ、都合の悪い結果がもたらされることも十分ありうることなのである。

こうした技術の予測不可能性は人間の意思決定能力の側面からも確認することができる。たとえば、コリングリッジは技術選択の問題を議論する前提として次のような現実を指摘する。

人間の知性、分析能力、情報収集力はきわめて乏しいものである。そのため、どれほど重要な意思決定でも、いや、おそらくどんな意思決定あれ、それが正しいものであるかどうかを知ることはできない。どんな選択をしたとしても、意思決定を行った人にその選択を後悔させるような予想外の事実が見出される可能性がいつも存在するのである。²

私たちの知的能力は自分たちが思っているほど優秀なものではない。どれほど努力して情報を集め、時間をかけてあらゆる可能性を想定したと思っても、有限な存在である私たちには全てを知ることはできない。世界は常に私たちの思惑を超えて変転するのであり、期待はいつ裏切られてもおかしくないと覚悟しておくべきなのである。

このように世界や技術の本質的な複雑さの点から見ても、またそれに関わる私たち人間の意思決定能力の点から見ても、技術が社会にもたらす影響について確実な予測をすることは不可能である。そうだとしたら、リスクの小さい技術を選択するという技術選択の理想は本質的に実現困難であり、技術がもたらす悪影響を未然に回避することも難しいということになる。

しかし、予想外の被害や損害が生じた場合には、後からその技術を放棄したり、修正したりすればよいのではないだろうか。だが、こうした事後的な対処の仕方にも一つの悩ましい問題がある。コリングリッジは新しい技術の導入に際して私たちが陥るディレンマについて次のように述べている。

テクノロジーの社会的影響はその初期段階には予測できない。しかし、望ましくない影響が見出された時には、その影響をもたらしたテクノロジーはたいてい経済構造や社会構造全体の大きな部分をなしており、それをコントロールすることは極めて困難な状態になっている。これがコントロールのディレンマである。変更が容易な時には、変更の必要性を予見することはできないが、変更の必要性が明白な時には、変更は高くつき、困難で、時間を要するものになってしまうのである。³

自動車が発明された当初は、交通事故で多くの人びとが命を落とすことになるなどということはいずれも想像できなかったに違いない。しかし、20世紀後半には自動車保有台数の増加にともなって、交通事故による被害が増加し、日本では「交通戦争」と呼ばれる状況が続いた。現在もその

² Collingridge (1992), p.4.

³ Collingridge (1980), p.11, 強調は原文による。

状況が根本的に改善されたとは言い難い。それでも私たちは自動車に依存する生活をやめることはなかった。その代わりに、法律の整備や歩車分離の推進、自動車の安全性能の向上をはかることで事故による悪影響を減らそうとしているが、これには長い時間がかかっており、成果もまだ十分とは言えない。同様のことは携帯電話についてもあてはまる。仮に携帯電話が身体に想定外の害をもたらすことが明らかになったとしても、私たちはその利用をなかなかやめることができないにちがいない。むしろ、害を減らすための技術開発を急いではじめるというやり方をとるのではないだろうか。このように、一度定着してしまった技術は、既存の経済的・社会的文脈と強固に結びついているため、それを完全に放棄することは困難であるし、仮に修正できるとしても、それには多大な時間と経済的・社会的コストがかかるのである。

以上のようなコリングリッジの指摘が正しいならば、私たちは未来が十分に予測できないにもかかわらず、技術選択に際して失敗は許されないという苦しい立場に立たされていることになる。もし選択に失敗すれば、負の影響をもたらす技術が社会に居座り続け、私たち自身を困らせることになるし、影響が長期にわたれば、未来世代にも迷惑をかけることになるからだ。このような苦境にあって私たちに何かできることはあるのだろうか。技術選択をめぐるコリングリッジの考察はこの問いに答えようとするものである。

3. 漸進的な意思決定

以上のような苦境に対処するための方策として、コリングリッジは、漸進的な (incremental) 意思決定を推奨する。これははじめから成功を前提として技術を導入するのではなく、選択が失敗する可能性をはじめから想定し、修正する余地を確保して事を進め、失敗した時にはすぐに引き返したり、その失敗から学んで進む方向を変えたりしながら、徐々に技術を展開していくようなスタイルの意思決定である。コリングリッジはそのプロセスを「試行錯誤学習 (trial and error learning)」という言葉で特徴づけ、それが適切に行われるための条件を以下のようなかたちでまとめている⁴。

- ① 失敗の代償は小さくしなければならない。効果的な試行錯誤は、試行が失敗したときの痛みが許容できる程度に小さい場合にのみ可能になる。代償が致命的に大きいと、そこから何かを学ぶ前に命を落としてしまう。また、失敗の代償が大きいにもかかわらず試行を断行することは、失敗が許されない状態で何かを試すことであり、無謀である。
- ② 一度に大きな変更をしない。急激に大規模な変更を行うと、それが失敗に終わった時の代償も大きくなる。また、一時に多くのことを変更してしまうと、失敗の要因を特定することが困難になるため、そこから学ぶことも困難になる。
- ③ 学習は速やかになされなければならない。ある技術についての試行が数年あるいは数十年かかると、その技術についての学習は妨げられ、遅れてしまう。何か教訓を得た時には、その技術によって解決しようとしていた元の問題が変質してしまっているかもしれない。

4 Collingridge (1992), pp.4-5.

- ④ 些末な選択ではなく、失敗の代償が大きくなるような重大な選択に意思決定や批判の労力が費やされなければならない。
- ⑤ 意思決定は単独で行うのではなく、多くの専門家による批判にさらされながら行われなければならない。
- ⑥ 意思決定に際しては、政治的に力のない人びとの声に耳を傾けなければならない。真の意味で有益な意見交換は政治的な力を分け合うことによって可能となる。
- ⑦ 特定の個人やグループのバイアスを防ぐために、意思決定には妥協が必要である。
- ⑧ 意思決定は、全ての情報を握るコーディネーターが主導するのではなく、意思決定に関わる当事者同士の相互調整によってなされる必要がある。

一時に大きな決定をしてしまうと、それが失敗した時の代償は大きくなる。将来について私たちの予測が当てにならない以上、このような意思決定の仕方は明らかにリスクが大きい。あるいは、ある問題を解決しようとして新しい技術を導入したにもかかわらず、別のより安全で安価な技術によって元の問題が解決できることがわかったり、そもそも元の問題が解決の必要のないものになってしまったりするということも考えられるだろう。この場合、予測の失敗によって無用な損失がもたらされたことになる。こうしたリスクや損失を避けるためには、安易な賭けに走るのではなく、小さな意思決定を着実に積み重ね、そのつどの結果を速やかに次の意思決定に活かしながら進んだ方がよい。こうしたプロセスを通じて、全体として健全な意思決定を実現しようというのがコリングリッジの狙いだと言えよう。たしかに、このようなやり方をすれば、一つ一つの意思決定がおよぼす影響のスパンが短く、意思決定の重みが各段階に小さく分散するため、想定外の事態が生じ、選択の失敗が明らかになったとしても、前の段階に引き返すことが容易であるし、代償も最小限に抑えることができる。これが「確実な予測はできないが、一度決めたら引き返すことが難しい」というディレンマに対する一つの有力な対処法であることは間違いないだろう。

また、注目すべきは、こうした漸進的な意思決定のプロセスの条件として、多様な人びとの見解を反映することが挙げられていることである（⑤,⑥,⑦,⑧）。これも意思決定の失敗を可能な限り回避するための工夫だと言える。ある技術について強い利害関係を持つ人びとだけが意思決定を主導する場合には、その技術の欠陥やリスクについて十分に公正な判断をすることができない可能性が高い。このような甘い吟味の結果導入された技術が結果的に失敗に終わり、しかも引き返すことが困難だとすれば、不当に大きなコストの支払いを強いられることになるだろう。また、社会のなかで技術が良好なかたちで受け入れられ続けるためには、できるだけ多くの人々がそれに賛同しなければならないし、またそれに付随するリスクを引き受ける人がある場合には、それをできる限り納得できるかたちで引き受けてもらうのでなければならない。そうでなければ、技術の導入後にそれに反対する運動が起きて計画が頓挫し、導入時に払われたコストが無駄になってしまうことも考えられる。また、その技術が定着し維持され続けるなかで、社会的弱者にそのリスクや被害が集中するといった状態が存続するかもしれない。倫理の観点から見て、これが技術選択の失敗であることは言うまでもない。社会のなかにある多様な声に耳を傾けることは、このような意味で技術選択の成功を強く条件づけていると見てよいだろう。

4. 技術の柔軟性

漸進的意思決定によって技術の不透明な未来に対処するという発想はコリングリッジだけのものではない。例えばアイディも、技術選択に際して「否定的な影響が現れはじめたら、影響を誇大に受け止め、できるだけ早く調査せよ」とか「多様な軌道に応じた代替案を増やせ」とか、「専門家以外の様々な利用者による使用試験を企画せよ」といった提案をしているが⁵、これらはコリングリッジの提案する試行錯誤学習の要素と内容的に重なるものである。

しかし、コリングリッジの議論のさらに興味深い点は、こうした漸進的な意思決定のあり方を手引きにして、選択を回避する方がのぞましいリスクな技術の特徴を明確にしていることにある。それをまとめると以下のようなになる⁶。

- ① リードタイム（製品開発や施設の建設のためにかかる期間）が長い。リードタイムが長いと、その技術の操業実績を評価する状態に到達するまでに長い時間がかかるため、失敗を同定し、素早くフィードバックして計画を修正することが困難になる。引き返すとなれば、また長い時間とコストがかかってしまう。
- ② 一単位あたりの規模が大きい。例えば、規模の経済を達成しようとして一つの機器や設備の規模を大きくすると、数少ない単位でしか試行できないため、失敗から学ぶ機会が少なくなり、計画を適切に修正する機会も失われがちになる。
- ③ 資本集約的（最初の設備投資が大きい）。運営コストを削減することは、技術の運営のなかで試行錯誤学習を通じて比較的容易になされうるが、資本コストはすでにかかってしまったコストであり、試行錯誤学習を通じて削減することは困難である。また、資本集約的な技術は失敗の代償が大きい。はじめに莫大な予算をつぎ込んでしまうと、計画が失敗したときに、それを回収できなくなってしまい、莫大な投資が無駄になってしまう。
- ④ インフラストラクチャーへの依存度が高い。ある技術が、その技術のために特別に開発される別の技術の支えによってはじめて利用可能になる場合には、試行錯誤学習が妨げられやすい。最も顕著な事例は、後でも指摘するが、原子力発電だろう。原子力発電のためには、ほとんどそのためだけにウランの採掘やウラン燃料の製造の技術が確立され、安定的に運用される必要がある。しかし、こうした支えとなる技術の確立や運用にも時間とコストがかかるし、また、計画が失敗した場合、その原因がどちらの技術にあるのかも特定がしにくくなるかもしれない。さらに、二つの技術が密接な相互依存関係にあるため、一つの技術の失敗が、事業全体の失敗となってしまい、失敗の代償もより大きなものになってしまう。

これらの性質を持つ技術は、種々の観点から見て「大きな」技術であり、その大きさゆえに融通が利かない、柔軟性に欠ける技術だと言えよう。こうした柔軟でない技術は、技術選択のリスクを回避するために求められる試行錯誤学習を妨げる傾向を強く持っているため、その導入はできるだけ回避するか、失敗の代償を十分に注意深く見積もり、覚悟した上で導入する必要があるこ

5 Ihde (1999), p.50-51.

6 Collingridge (1992), pp.12-14.

となる。

コリングリッジによれば、こうした柔軟でない技術の導入には、大きな組織が関与する場合が多い⁷。ある目的を達成するために、革新的な技術を開発し、導入することが巨大なプロジェクトとしてこれまでしばしば行われてきたが、その背後には国家をはじめとする大きな組織がそのプロジェクトの主体として存在していた。しかし、大きな組織の関与は機動的な意思決定を妨げる可能性が高い。大きな組織には技術力、財政力、政治力があるため、計画の失敗や運営の失敗が起きても、それをカバーし、計画を継続できてしまう。その結果、失敗の認知や失敗からの学習が遅れ、修正や後戻りが不可能になるところまで問題が深刻化してしまうのである。また、大きな組織は社会の中の多様な見解を反映してプロジェクトを進めるのではなく、その組織の利益や価値観に即して独断的に物事を決め、プロジェクトを進めてしまう傾向が強い。その場合にも、技術の社会的受容に問題が生じてしまい、計画が円滑に進まなくなったり、技術の強引な導入により無用な損害や被害がもたらされたりするのである。

5. 柔軟でない技術：原子力発電

柔軟でない技術としてコリングリッジは様々な技術を取り上げ、分析しているが⁸、なかでも彼が精力的に論じている原子力発電の事例についてごく簡単に紹介し、技術選択において技術の柔軟性が一つの重要な指標となりうることを確認しておきたい⁹。

コリングリッジのまとめによれば、アメリカにおいて原子力発電が導入されたのには次のような経緯があった。1945年から1960年代後半にかけて化石燃料の価格が持続的に低下し、それにもない電力価格も低下していた。これにより電力需要が喚起され、拡大傾向を見せる¹⁰。産業と経済が右肩上がり成長するなかで、電力需要はそのまま上昇しつづけるというコンセンサスが醸成される一方、発電のための化石燃料の不足が懸念されるようになり、電力の需給ギャップによる世界経済の崩壊がメディアで日常的に話題となったという¹¹。こうした時代状況のなかで新たな電力供給の手段として注目されたのが原子力発電という新技術であった。原子力発電の商業化を強力に押し進めようとする原子力委員会のプロパガンダが支えとなり、また、ジェネラルエレクトリック社やウェスティングハウス社が相次いでターンキー契約¹²という魅力的な契約形態でセールスを行ったこともあって、1963年から1970年代半ばにかけて原子力発電所の建設ラッシュが起ることになる¹³。

ところが、1960年代後半からの金利の上昇により燃料価格が上昇し、それにもなあって発電コストも増大していった。さらに、1973年の第一次オイルショックにより経済成長が鈍化し、電力

7 Collingridge (1992), pp.14-16.

8 Collingridge (1992) では、スペースシャトル、原子力発電、北海油田開発、ビデオテックス、開発途上国における巨大灌漑農地計画、イギリスにおける高層システム建築が柔軟性に欠ける技術の事例として取り上げられ、分析されている。

9 Collingridge (1992) に先だって、Collingridge (1983) ではより詳細に原子力発電が導入されたプロセスが検証されている。ここでは主によりコンパクトにまとめた前者の叙述にそくして原子力発電が柔軟性に欠ける技術であることを確認する。

10 Collingridge (1992), p.40.

11 *ibid.*, p.41.

12 建設及び試運転までの全業務をメーカーが一括して定額で請け負う契約で、工期延長や建設費の増大のコストもすべてメーカーが責任を負う。

13 Collingridge (1992), p.40, 49.

需要は1973年以降の10年間で毎年平均2%ずつ落ち込んでいく。それにともなって、電力の供給能力は過剰となり、既設ないし建設中の原子力発電所の必要性は低下した。さらに、こうした余剰プラントのコストを電力価格に転嫁せざるをえなかったため、電力価格がさらに上昇し、それが電力需要を冷え込ませるという悪循環が起こることになる。こうした影響から、1973年以降、契約のキャンセルや操業の遅れ、注文の減少が発生し、アメリカの原発産業は下火になっていったという¹⁴。

ここで重要なのは、問題の解決のために導入された原子力発電が先の4つの特徴を備えた、柔軟性に欠ける技術であったことだ。原子力発電は大きな技術であり、また、新しい技術でもあったため、もともとの建設期間は長かったが、それが予想以上に長くなり、操業するまでの準備期間も当然長くなっていった¹⁵。また、規模の経済を達成するために一単位当たりの設備規模は大きく設定されていた¹⁶。さらに、大規模の設備を建設するために膨大な初期投資が必要だったが、建設の長期化により、また、安全規制の強化により、建設費はさらにかさんだ¹⁷。そして、フロントエンドにおいてウランの採掘、ウラン燃料の製造、ウランの濃縮のための施設が、バックエンドにおいては使用済み燃料の処理、再処理、廃棄物の処分のための施設が、原子力発電のためだけに必要となった¹⁸。このように原子力発電は、リードタイムが長く、単位規模が大きく、資本集約的であり、インフラストラクチャーに大きく依存する技術であった。リードタイムの長さや単位規模の大きさは、その技術の実際のパフォーマンスやその技術がおかれる経済的・社会的環境についての情報把握を遅らせた。資本集約的な設備によって生産される電気のコストは資本コストの大きさと平均稼働率の影響を大きく受けるが、建設費の増加もあり、結果として発電コストの3分の2以上を資本コストが占め¹⁹、平均稼働率ははじめに予想した80%を大きく下回り、55%程度にとどまったため²⁰、コストも予想以上に高くなり、他の手段よりも割高な発電方式となってしまったと言う²¹。しかも、その時には当初懸念されていた需給ギャップによる経済への打撃という問題は消え去り、電力需要の鈍化による供給力過剰という正反対の状況が出来し、既設ないし建設中の原発に注がれた初期投資をいかに回収するかという新たな問題が生じてしまった。単位規模や初期投資が大きく、インフラストラクチャーへの依存度が高いこの技術は、一からすべてを作り上げるものであり、それだけに失敗の代償が大きくなったと言える。

また、コリングリッジは具体的に触れていないが、事故のリスクや廃炉および放射性廃棄物処分の問題を考慮するならば、この失敗のコストは経済的なものだけにとどまらず、作業員の被曝など重大な社会的コストももたらしたことになる。原子力発電の恩恵を受けない世代に廃炉や廃棄物処分についての意思決定、および具体的な実施の負担を負わせる結果を招いたことは、世代間倫理の観点から見ても大きな問題だと言えよう。柔軟性に欠ける新技術の導入という賭けは、

14 ibid, p.42.

15 ibid, p.43.

16 ibid, p.44.

17 ibid, p.42-43, 1960年代後半までに注文された施設は平均して当初の見積もりの2倍の資本コストがかかったと言われる。予想に反して、また、通常の技術とも異なり、建設の経験を経た後で資本コストはさらに増加し、1967年から1973年にかけておおよそ3倍にふくれあがったという。

18 ibid, p.44-45.

19 ibid, p.44.

20 ibid, p.47.

21 ibid, p.43.

このように社会的・倫理的な点でも大きな代償をもたらしたのであり、その意思決定の問題性を私たちは今一度噛み締める必要がある。

もっとも、以上のような失敗は決して予測の甘さだけによって起きたと見るべきではないだろう。もちろん、将来の経済環境についての予測だけでなく、施設の建設期間や建設費、具体的なパフォーマンスについて当初なされた予測の多くは外れてしまった。しかし、技術に失敗はつきものであり、予想が裏切られる可能性があるのは当然のことである。問題は予測を外したこと、あるいは失敗してしまったことそのものよりも、むしろ、その失敗が異常に大きなコストをもたらしたことである。そしてその原因は、「柔軟でない」技術を導入したことにある。そもそも、試行錯誤を行うまでもなく、計画の初期の段階から原子力発電の建設が様々な意味で「大きな」、「柔軟でない」プロジェクトであることは明かだった。それゆえ、その失敗がもたらすかもしれない代償の大きさについても初期段階から認知されて然るべきであった。この点についての認識が甘かったことが原子力発電の失敗に関しては本質的なものである²²。

6. コリングリッジの議論の有効性

漸進的意思決定を技術選択の理想的な在り方とし、技術の柔軟性を技術選択の基準として重視するコリングリッジの考え方は、有限な存在である人間が複雑な世界のなかで技術とともに生きていく際にわきまえておくべき基本的な洞察を示したものであり、今後の技術選択のあるべき姿を探求する上で有効な手がかりとなりうると筆者は考える。もっとも、コリングリッジの議論にも曖昧な点や不十分な点がないわけではない。最後にこのことについて指摘しておきたい。

まず、コリングリッジがどちらかというと経済的なコストの観点を中心にして議論を進める傾向があることは指摘しておかなければならない。この点については、科学技術が人びと（とりわけ社会的弱者）に与える悪影響こそ考えるべき問題であり、もっと倫理的な観点から技術の問題を分析すべきだという指摘がなされるかもしれない。もっとも、このような経済的観点への偏りがあるからといって、コリングリッジの議論の価値が著しく損なわれることにはならないと筆者は考える。経済的なコストは、とりわけ政府などの公的機関が深く関わる場合には、そのまま社会的なコストとなる。それが甚大な場合には、次の世代にまでその影響が及ぶだろう。これはこれで倫理的に由々しき問題である。また、コリングリッジ自身が社会的なコストの問題についてまったく考えていないわけでもない。たとえば、すでに触れたように、試行錯誤学習の議論において指摘されていることがら（特に⑤～⑧の論点）などは、社会的弱者や社会における多様な価値観を強く意識したものとなっていると言えるだろう。たしかにコリングリッジ自体はそれについてあまり具体的な議論を展開することはしていない。しかし、先の原子力発電の議論において

22 *ibid.*, pp.50-51, 技術の柔軟性や代償の大きさについて認識が甘くなってしまった要因の一つとしてコリングリッジは大きな組織の関与を挙げる。先にも指摘したように、原子力発電の商業化には、アメリカ政府や原子力委員会といった政治的・財政的に大きな力を持った組織が関与していた。これは、他の国でも同様で、原子力発電の導入には各国政府の意思決定が重要な役割を果たした。イギリス、フランス、ソヴィエト連邦では国営の電力会社が国営の開発組織から設備や施設を買い受けており、燃料供給、廃棄物処理、資金融資、セキュリティの分野にいたるまで政府と密接に関わっていた。日本や西ドイツでは私企業の電力会社がより大きな役割を果たしていたが、それでもやはり明確なエネルギー政策を持つ政府の方針のなかで大枠の意思決定がなされていった。そのなかであって、アメリカにおける原子力発電の商業化では政府の関与が相対的に小さかったため、学習がより速やかになされたコリングリッジは指摘している。経済状況の影響を受け原子力産業が下火になったことがそのことを示している。*ibid.*, pp.51-52, 69-70.

筆者が補足したようなかたちで、コリングリッジの議論を社会的・倫理的な側面からより具体的に補足することはそれほど難しいことではない。ある技術が未来においてどのような影響をおよぼすかは、その影響が経済的なものであれ、社会的・倫理的なものであれ、明確に予期できないのであり、そうである以上、技術の柔軟性に着目し、漸進的な体制で技術を導入していくことは倫理の観点から見ても有効な手だての一つである。

しかし、このことを踏まえた上で指摘しておかなければならないのは、技術の柔軟性に注意をはらい、新しい技術の開発や導入にあたって漸進的な姿勢で臨むという考え方は決して万能ではなく、それに過大な期待をかけてはならないということである。例えば、ある技術が少しずつ社会に順応し、それが広く普及した後で、潜んでいた問題が顕在化したり、その技術が広く普及したこと自体が原因で大きな被害が生じたりすることも十分に考えられることだ。柔軟性や漸進性という特徴に着目するだけでは、こうしたリスクを事前に把握し、回避することはできない。このことは市場に出される技術の多くをこうした指標によって事前に規制するのが困難であることも示唆している。資本主義経済のなかで流通する技術の多くは、経済的観点でも倫理的観点でもある程度リスクコントロールがなされた後で登場し、市場に出回った後も競争や批判にさらされ続けるという意味で柔軟で漸進的な性質を持っているため、コリングリッジの言う観点でふるいにかけてもそれをすり抜けてしまう可能性が高い。こうした技術がもたらす被害については今のところそのつど事後的に対処するほかにないことになるだろう。また、こうした被害を未然に防ぐためには、技術の普及のあり方とそれがもたらす負の影響について歴史的なデータを分析し、その結果を科学技術のリスクの一つのあり方として消費者に周知していくほかに、今のところ方法はないように思われる。

また、ナノテクノロジーや遺伝子工学のような技術は、いまだその物質的・社会的・倫理的影響がはかり知れず、新たな問題を引き起こしかねないこともあり、たとえそれが柔軟な性質を持ち、漸進的に進められうるのだとしても、導入には不安がつきまとうことになる。こうした技術はその進展の一步一步がいわば深く、重いものになってしまうのである。その場合には、柔軟性や漸進性を強調するだけでは十分なリスクマネジメントができないことになる。こうした「新しい」技術分野については、それが社会のあり方や私たちの価値観にも大きな影響を及ぼす可能性を考慮し、新しい技術に対する多様な価値観やその導入に伴う社会や人間のあり方の可能な変化について慎重な分析をし、そもそもそうした新たな技術を「徐々に試す」ことの是非から検討を行う必要があると思われる。

このように、柔軟で漸進的な性質を持つ技術であれば絶対に安全というわけではもちろんないし、その指標によってすべての技術を適切にふるいにかけることができるというわけでもない。しかし、このこともコリングリッジの議論の価値を著しく減じるものではないと筆者は考える。コリングリッジの議論は、技術にリスクがつきものであり、またそのリスクを事前に見極めるのが困難であることを認めた上で、認知が相対的に容易であり、避けることが可能な失敗を確実に捉えて、回避すべきことを主張したものと見ることができる。その意味で、技術の柔軟性という観点は、技術選択のための万能の指標ではないにせよ、そのための一つの重要なチェック項目であるには違いない。様々なビッグプロジェクトに巨額の予算がつぎ込まれるということが多く見られる現状においては、このチェック項目の存在を強調することはいまだに有効であり、また重

要であると筆者は考える²³。

また、技術の柔軟性の議論を糸口として、どのような技術がどのような意味でリスクが高いのかということについて、より詳細で、体系的なかたちで探求を広げていく可能性も開かれてくる。それによって、技術選択のために求められる指標をさらに精密にしていけることもできるかもしれない。すなわち、コリングリッジの議論は技術選択についてのより広範な研究へと進んでいくための出発点としても価値を持つと思われるのである。

文 献

- Ihde, Don (1999), *Technology and Prognostic Predicaments*, *AI & Society* 13. (中村雅之(訳)「技術と予測が陥る困難」『思想』926, 145-156 頁)
- Collingridge, David (1980), *The Social Control of Technology*, New York; St Martin, Palgrave Macmillan.
- Collingridge, David (1983), *Technology in the Policy Process: Controlling Nuclear Power*, Pinter; London
- Collingridge, David (1992), *The Management of Scale: Big Organizations, Big Decisions, Big Mistakes*, London; Routledge.

23 このことは公的な資金を取り扱う行政や自治体だけでなく、私企業の意志決定にも当てはまる。たしかに、私企業は収益を上げるために自由に活動し、時にリスクを冒して大きな事業に投資するといった賭けをする権利がある。その点を強調するならば、漸進的意思決定は安定した企業経営のための一つの指針たりうるとしても、企業経営を根本的に縛るものではありえないことになるだろう。しかし、私企業であっても、それが社会の中で営利活動を行っている以上、株主の利益に対してだけでなく、市民の安全や幸福に対しても責任を負わなければならないはずである。その意味では、長期的に見て社会全体に悪い影響を及ぼす可能性のある技術については、それを導入しようとしている私企業に漸進的な開発や実用化を求めることも許されるし、また必要だと思われる。たとえば、各国の電力会社は使用済み核燃料の安全な処分技術の開発を後回しにして、いわば見切り発車のようなかたちで原子力発電をはじめてしまった。その結果、核のゴミのリスクは本来その責任を負う義務のない世代にまで引き継がれざるを得なくなっている。このようにある技術の社会全体に対する長期的な影響が予想できる場合には、そこから利益を得る私企業にも技術導入とその帰結に対する責任が存するはずであり、公的機関は私企業の技術導入に関して漸進的な開発・導入を促すような規制を施していく必要がある。

自己抹消的な道徳理論の問題点は (あるとすれば) 何か

秋葉剛史 (千葉大学)

要旨

道徳理論は、我々がどのように行為すべきかについての指示を与える。そのため、日々の道徳的な意思決定において、自分の信じる道徳理論がどの行為を勧めているかを考えることは、我々にとってごく自然なことのように見える。しかしある種の道徳理論は、こうした直接的な意思決定法をそれ自身の基準から禁じる。なぜなら、その種の道徳理論に従おうとした場合、我々はかえってそうできなく（しにくく）なるからだ。このような一見逆説的な状況におかれた道徳理論は、しばしば「自己抹消的」と呼ばれる。そして少なからぬ論者によると、ある道徳理論が自己抹消的であるという事実は、場合によるとその理論を拒否するのに十分なだけの深刻な難点になる。しかし本稿ではこのような見方に対し、自己抹消性は少なくともそれ自体では、ある道徳理論を拒否するための説得的理由にはならないと主張する。そのため本稿では、従来この種の理論に向けられてきたいくつかの批判を検討する。より具体的に言うと、ここで考察されるのは、行為指導性の欠如、実践的指示の不確定性、心理的調和、望ましい道徳的熟慮、自己欺瞞、といった論点に関わる諸批判である。本稿の議論を通じて、自己抹消的な道徳理論に従来投げかけられてきたこれらの批判には十分応答可能であることが示される。

What are the Problems (if any) of Self-Effacing Moral Theories?

Moral theories generally tell us the right course of action. They deliver practical advices (or a series of commands) about how we should act, so it seems natural for those who sincerely accept a given moral theory to try to figure out, in each case where some practical decision is needed, what the theory recommends, and act accordingly. However, this is not what a moral theory always tells us to do. Certain moral theories in fact tell their followers not to consult them in daily decision-making. The reason for this is simply that if we consciously intend to act as the theories recommend, it would become difficult or even impossible for us to act in that manner. In recent literature, moral theories that satisfy this condition are called *self-effacing*, and have attracted some attention. Although quite a few authors seem to endorse the view that this character of self-effacement makes a moral theory highly problematic (or even unacceptable), in this paper I shall argue that this view is ungrounded. To do this, I will critically examine various objections to self-effacing moral theories found in the literature that concern the following points, respectively: lack of action-guidingness; threat of undermining psychological harmony and the desirable form of moral deliberation; an absurd requirement to have mutually contradicting beliefs; and an invitation to a kind of self-deception. It will be argued that none of these objections constitutes a serious threat to self-effacing moral theories as such.

道徳理論の主要な目的は、どのようなタイプの行為が道徳的に正しいかを明らかにすることだ。例えば功利主義なら、「幸福を最大化する」行為こそが正しいと主張し、カント的義務論なら、「定言命法と合致した」行為こそが正しいと主張するだろう。各々の道徳理論はこのような仕方、
 どういったタイプの行為をなすべきかを我々に指示するわけである。そして、一般に道徳理論がこうした実践的な指示を発するものだとなれば、ある道徳理論を受け入れるときに我々がどのような態度で実践に臨めばよいかは自ずと決まってくるように思われる。すなわち我々は、いくつかの行為の選択肢があるときにはいつも、自分の信じる理論が正しいとする選択肢はどれかを割り出しそれを実行に移すよう心がければよい、とそう思う。こうした意思決定の方法は、ある道徳理論を信じる者にとってごく自然なものだろうし、当の道徳理論によって勧められるものでもあるだろう。

だが実のところ、話はそう単純ではない。なぜならある種の道徳理論に関しては、その理論が正しいと言う仕方で行うと我々はかえってそう行為できなく（しにくく）なる、ということが成り立つからだ。このことは、以下で見るように様々な事情に起因しうるのだが、いずれにしても、このような状況におかれた道徳理論は我々に対して、その理論の主張を意識しながら行為することを（それ自身の尺度から）禁じることになる。言ってみればそうした理論は、行為に臨む我々に「私の主張を気にかけてはならない」と命じ、我々の思考から自らすすんで消え去ろうとするのだ。本稿で焦点を当てたいのは、しばしば「自己抹消的 self-effacing」と呼ばれる、一見奇妙なこの種の道徳理論である。

一部の論者が主張するところでは、自己抹消的な道徳理論は単に奇妙だけでなく、より実質的な困難を抱えている。彼らによると、ある道徳理論が自己抹消的であるという事実は、場合によってはその理論の拒否を正当化するのに十分なほどの深刻な難点になるのだ。じじつ関連文献の多くは、自己抹消性という性質が一般に道徳理論を受け入れがたいものにするという見方はほぼ前提とした上で、ある特定の道徳理論（e.g. 帰結主義や徳倫理の各バージョン）が実際に自己抹消的かどうかを検討することに主な労力を割いている¹。

しかしこのような見方は、どの程度の根拠をもつのだろうか。自己抹消性という性質が道徳理論にもたらすという問題とは具体的にどのようなものであり、それらは本当にこの種の理論にとって深刻な難点になるのだろうか——本稿で考えてみたいのは、従来の研究であり主観的に論じられてこなかったこれらの問題だ²。以下で見るように、実は自己抹消的な道徳理論は決して珍しいものではないから、これらの問題は、一般に我々が道徳理論とどう関わるべきかを考える上でも重要である。そして結論から言うと、（以上の流れから予想される通り）私は以下で、自己抹消性に関して提起されてきた諸問題は基本的に³、ある道徳理論を拒否するための説得的な理由にはならないと主張する（もちろんその理論が他の理由から受け入れがたいという可能性は残るが）。自己抹消性が困難をもたらすという印象の大部分は、道徳理論が我々の実践を導く際の多様な仕方を十分に考慮しないことから生じている、ということが本稿の議論を通じて明らかになるだろう。

1 例えば徳倫理については Keller 2007; Martinez 2011; Pettigrove 2011、帰結主義と義務論については Williams 1973: 128; Stocker 1976: 455f; Smith 2001 を参照。

2 ただしこれらの点にふれた議論がないわけではない。例えば Wiland 2007; Cox 2012 を参照。

3 ここで「基本的に」と留保した理由については、本稿の末尾（第5節）を参照のこと。

以下では、まず自己抹消性という特徴について改めて説明した後で（§1）、自己抹消的な道徳理論に対し最初に向けられる「行為指導性の欠如」という懸念をとり上げそれに応答する（§2）。続いて、その応答に対してさらに提示されうる三つの批判と（§3）、「自己欺瞞」という点に関連する一つの懸念（§4）に注目し、自己抹消的な理論をそれらから弁護する。

1. 道徳理論の自己抹消性

自己抹消性について改めて説明することから始めよう。後の議論の提示しやすさのため、ここでは特に D・パーフィットによる特徴づけ方 (Parfit 1984: ch.1) を参考にする。前述のように、各々の道徳理論はある特定のタイプの行為を正しいものとし、そのタイプの行為がなされることを我々に要求する。これは言い換えると、一般に道徳理論は、あるタイプの事態を、我々が自らの行為を通じて実現すべき目標として定めるということだ。以下では、任意の道徳理論を「T」によって表し、道徳理論 T が実現されるべき目標として我々に課す事態を（「Aim of T」という意味をこめ）「 A^T 」と表そう。例えば T が功利主義なら、それが我々に与える目標 A^T は「幸福が最大になること」である。また、T がある十種類の行為を禁じるような神の命令理論なら、その目標 A^T は「それら十種類の行為がなされないこと」である。（ここで「目標」は禁忌や義務に関する事態なども含む広い意味で使われていることに注意。この用法は以下でも同様である。）

さて道徳理論が一般にこのようなものだとすると、ある道徳理論 T を真なるものとして信じることのうちには、その理論の目標 A^T に対し一定の規範的なし動因的 conative な態度をとることが含まれることになる。すなわち T を信じる者は、 A^T の実現は望ましい（ A^T は実現されるべきだ）という評価的（規範的）判断を受け入れ、 A^T を実現しようという意志やそれを実現したいという欲求などをもつ者、つまり、 A^T の実現に対し少なくとも潜在的な動機をもつ者である⁴。しかしこの潜在的動機が、T を信じる者の日常的実践において顕在的な動機としてはたらくべきかと言え一概にそうは言えない。というのも導入部で述べたように、ある種の理論 T においては、その目標 A^T を実現しようという意図に導かれて行為することがかえって A^T の実現の妨げになる、ということがあるからだ。

本稿で関心を寄せるこの自己抹消的な状況を、もう少し丁寧に特徴づけておこう。そのために注意しておきたい点は三つある。第一に、自己抹消性を特徴づける上で重要になるのは、我々が一人称的な観点から行う意思決定の過程、つまり、自分はいかに行為すべきかを（通常いくつかの選択肢を前にして）考察し決定するという過程だ。以下ではこのような意思決定の過程を、短く「熟慮 deliberation」とも表すことにする。第二に、この意味での熟慮を進める仕方には様々なものがありうるが、自己抹消性と特に関係するのは、何らかの道徳理論 T への信念が顕在的・主導的な動機となって進められるような熟慮である。すなわちそれは、「当の理論 T が定める目標 A^T を、そのつど与えられた行為の選択肢と直接引き比べ、それらの選択肢の中でどれが A^T の実現に最も寄与するか（あるいはその実現であるか）を割り出す」という熟慮の方法だ。以下では、

4 本稿では Adams 1976: 467; Parfit 1984: 8 に倣い、「動機 motive」という語を、何らかの行為を生み出すものとしての心的な状態・態度を表す総称として用いる。なおここでは、何らかの道徳理論への信念（道徳的判断）と対応した動機づけの間には単に偶然的ではない結びつきがあると想定しているが、これは必ずしも、いわゆる「動機内在主義」をとることではない。

このような熟慮を「 A^T を直接的目標とした熟慮」と呼び、また、このタイプの熟慮に基づいて行為することを「 A^T を直接的目標として行為する」こととも表すでしょう。そして最後に、三つ目の注意点として、ある目標の実現にとってある熟慮の方法が適しているかどうかは、ほとんどの場合、我々について現に成り立つ様々な経験的事実——我々の心理的・身体的な特徴や、社会的・生物学的な環境など——に依存する。以下では、道徳理論が与える目標の実現に影響するこれらの経験的事実の総体を考え、それを「現実の経験的条件」と呼ぶことにする。

以上をふまえると、自己抹消性は次のように定義できる。すなわち、道徳理論 T が自己抹消的であるのは、 T が与える目標 A^T に関し、現実の経験的条件の下で、「我々が A^T を直接的目標として行為すると、 A^T の実現はより困難になる」という条件が成り立つときである⁵。ただしここで「困難になる」は、長期的・全体的な傾向を表す。すなわちこの条件が意味しているのは、もし我々が A^T を直接的目標とした熟慮に基づいて日々の実践を行うなら、 A^T はそうでない場合と比べ全体としてより実現されにくくなる、ということだ⁶。この定義が上の直感的説明と合致していることは、次のような考察からわかるだろう。一般に道徳理論 T は、 A^T を実現されるべき目標として我々に課す。それゆえ T によれば、もしあるタイプの行為や習慣・傾向性が A^T の実現を阻害するなら、我々はそれを行ったり身につけたりすべきでない。だが上の定義条件が成り立つ場合、「 A^T を直接的目標とした熟慮」はまさにそのような阻害要因になってしまう。よって T はそれ自身の基準に照らし、このタイプの熟慮を日常的に行わないよう我々に命じることになる。

一つ例を挙げておくと（他の例は次節以降で）、古典的な行為功利主義はしばしばこの意味で自己抹消的な道徳理論の典型だとされる⁷。周知のように、この立場が我々に与える目標は「関係者の幸福が最大になること」だ。しかし功利主義者自身も認めるように、我々が実際にこの事態を直接的な目標として、つまり、熟慮のたびに幸福計算を行うという仕方で行う日々の実践に臨む場合、この事態が実現される見込みは逆に低くなると考えられる。というのもその場合、我々は計算に時間をとられて行為の機を逸したり、計算自体に苦痛を感じたり、時間的・情動的な制約や各種バイアスの下で計算することによってかえって大きく選択を誤ったり、他の行為者との協調行動をとりにくくなったり、等々の事情から、幸福の最大化を直接目指さない場合よりむしろ少なくとも幸福を実現できないということになりがちだからだ。こういったわけで行為功利主義は、それ自身の尺度に基づき、この理論への信念を主導的な動機として行為しないよう我々に命じることになる⁸。

5 ただし本稿の末尾の第5節でも触れるように、ある種の道徳理論においては、この定義条件（括弧内の部分）が現実の経験的条件に依存せずに成り立つと考えられる。

6 この意味での自己抹消性は「いつ誰にとって」が問題になる相対的な特徴であり（Parfit 1984: 5）、かつ程度差を許容する特徴である（Martinez 2011: 279）。本稿で「自己抹消的」として念頭におくのは、現在の我々（標準的な人間主体）にとって、問題の熟慮を行った場合 A^T の実現されやすさが無視できない程度に低下するようなケースである。（これは曖昧だが、以下の議論の害にはならないだろう。）

7 Mill 1998: 69f; Sidgwick 1966: 405; Adams 1976: 467; Hare 1981: chs. 2-3; Railton 1984: 153-4; Wiland 2007.

8 では反対に、自己抹消的でない道徳理論にはどんなものがあるか。一部の論者はその例として、ある種の規則功利主義（Wiland 2007: 276）や、道徳法則についての意識から行為せよと説く（カント的な）義務論を挙げている（Keller 2007: 223）。ただし本稿の末尾で触れる問題も参照のこと。

2. 自己抹消性と行為指導性

さてそれでは、以上の意味で自己抹消的であるような道徳理論にはどのような問題があるだろうか⁹。この点で真っ先に浮かぶのは、「行為指導性 action guidingness」に関する次のような不満かもしれない¹⁰。一般に道徳理論は、我々の行為実践を一定の方向へと導くことができなくてはならない。物理学や化学などの理論とは異なり、道徳理論はその本性からして、いかに行為すべきかを我々に教えることを使命とした実践理論だからである。そしてある道徳理論が真に行為指導的であるためには、それは、我々が何らかの（ときに困難な）選択に迫られる具体的な意思決定の場面で、我々の決定を主導するような実質的根拠や観点を提供できねばならない。しかし自己抹消的な道徳理論は、明らかにこの要件を満たさない。それが意思決定の場面で与えることができるのは、せいぜい、「私の言うことを考慮するな」という無責任で投げやりな指示だけだからである。ゆえにこの種の道徳理論は、そもそも道徳理論としての存在意義をもたない。

だがこの批判は、（少なくともそれ自体としては）深刻ではない。この批判への応答としてはまず、それが依拠する大前提、すなわち「道徳理論が受容可能であるためには行為指導的でなくてはならない」という前提を斥ける道がありうる（cf. Parfit 1984: §17）。これは私見では必ずしも無謀な道ではないが、この路線を多少なりとも満足な仕方で擁護するには、かなり立ち入った考察——とりわけ、道徳理論の真理（適合）性と実践性についてのメタ倫理的な考察——が必要であり、それに深入りする余裕は本稿にはないので、ここではより寛容な応答を探ることにしよう。以下で考察したいのは、道徳理論に行為指導性は必要だという前提は認めた上で、「自己抹消的な道徳理論であっても実際は十分に我々の行為を導きうる」と応じる道だ。この応答は次のように提示できる¹¹。

ある道徳理論 T が自己抹消的であるとしよう。このとき前節の定義より、我々が A^T を直接的目標とした熟慮を日々の実践の中で行うことは、現実の経験的条件の下ではむしろ A^T を実現されにくくする。つまりこのタイプの熟慮は、 A^T の実現に貢献しない。だがここで重要なのは、このタイプの熟慮が A^T の実現に貢献しないことが経験的条件によって決まるなら、同じ条件によって、他の可能なタイプの熟慮がどのくらい貢献するかも決まるはずであり、よって特に、それらのうちどのタイプが A^T の実現に最適であるかも決まるはずだという点である¹²。そしてそうだとすれば、理論 T はまさにそのタイプの熟慮を、 A^T を直接的目標とした熟慮に代わる「代替的」な熟慮方法として我々に勧めることになると考えられる。

もちろん、ここで言う最適性はあくまで相対評価によるものであり、あるタイプの代替的熟慮

9 念のため注意すると、自己抹消的な理論は通常の意味で「自己矛盾」しているわけではない。自己矛盾した理論からは何らかの命題とその否定の両方が帰結するが、自己抹消的な道徳理論 T から「 A^T は実現されるべきだ」という主張に加えてその否定が帰結する必要はないからである。また、自己抹消的理論が発するものとして本段落で挙げた「私の言うことを考慮するな」という指示は、一見すると遂行レベルでの矛盾をはらんだダブルバインドの指示（それに従うと従わなかったことになり、逆も然り）にもみえるがこれは単なる見かけにすぎない。なぜなら、 T は自分の言うことのすべてを考慮するなと指示しているわけではなく、あくまで日常的な意思決定の方法について指図しているだけだからである。

10 Williams 1973: 128; Jackson 1991: 466; Wiland 2008: 393; Martinez 2011: 280; Pettigrove 2011: 192.

11 Cf. Railton 1984: 152f; Parfit 1984: 29; Cox 2006: 508; Wiland 2007: 279f.

12 こうした「最適」な熟慮がただ一つに決まらない（極大的によい熟慮が複数ある）ケースも論理的には不可能ではないだろう。だが、前述した習得コスト等を考慮に入れるとそうしたケースはあまりなさそうなので、以下では最適熟慮が一つに決まる場合をもっぱら念頭におく。

が A^T の実現に最適であるとしても、それは必ずしも A^T の「完全な」実現を保証するわけではない。とりわけ、我々がある熟慮を日常的に行えるようになるにはそれに適した傾向性を身につける必要があり、その作業には相応のコストが伴うから (cf. Wiland 2008: 390)、ある熟慮が A^T の実現にもたらす貢献度を考える際にはそのコストを差し引くといったことも必要になる。しかしそれらを総合した上で、あるタイプの熟慮を我々が行う場合に A^T が最もよく実現されると言えるなら、 T は具体的場面での意思決定方法としてまさにそのタイプの熟慮を勧めていると言ってよい。つまり T は、たとえ熟慮において参照すべき目標を直接与えるのではないとしても、代わりに何を参照して熟慮すべきかを（経験的条件と共同で）決定することで、我々の日々の実践を間接的に導きうるのだ。

実際、道徳理論がこうした間接的な仕方で我々を導きうるという点は、早くから功利主義者たちによって強調されてきた¹³。例えば J・S・ミルによると、我々の日常的な道徳的思考は、「幸福の最大化」という本来の目標によってではなく、むしろ常識道徳を構成する諸々の単純な規則（嘘をつくな、等々）によって導かれるべきである。なぜならそれらの規則は、人類が長い実験の歴史を通じて、その順守が結果として我々の幸福の最大化を生み出すことを確かめてきた当のものだからである (Mill 1998: 69-71)。つまり功利主義は、我々がその実現を見越して日々の熟慮方法を選ぶべきところの目標（＝幸福の最大化）を特定することで、間接的に、日常的な我々の意思決定に導きを与えるのだ。

あるいは別の例として、徳倫理のあるバージョン、すなわち「そのつどの状況で有徳な行為者がそうするであろう仕方で行為すること」を目標として我々に課す立場を考えてみよう (Hursthouse 1999: 28)。そしていま仮の話として、我々の多くは（残念なことに）そのつどの状況で発揮されるべき徳をほとんど備えておらず、有徳者のように振る舞おうとすると混乱してしまいかえって有徳者とは程遠い行動に出てしまうが、非常に鋭敏な羞恥心をもっているため、何をしたら人に笑われずに済むかを考えるとおおむね本物の有徳者と同じ判断に達することができる、と想定しよう。そうするとこの場合、「人に笑われない行為を選ぶ」という熟慮方針が、上の目標の実現にとって最適となるかもしれない。

賢明な読者は気づいているように、ある実践理論がこうした間接的な仕方で導きを与えるというのは、実は道徳理論に限らずごくありふれた現象だ。我々はしばしば、ある目標を実現するにはそれを直接追求するよりも何か別のことを目指した方がよいということを知っており、その知識に基づいて日々の実践に臨んでいる。例えば、我々各人にとってよいこととは「最大限の快樂を得ること」だと考える快樂主義者も、この目標の実現のためには快樂とは別のこと (e.g. そのつどの活動の独自性を認識すること) に当面の意識を集中する方がよいことを知っている¹⁴。あるいは、外国語を流暢に話すこと、感じのよい人として他人に映ること、眠りにつくこと、テニスの試合に勝つこと、ゴルフクラブの芯にボールを当てること、等々の目標——その意識的な追求がよい結果を生まないことが経験的に知られているような目標——に関しても、その実現にとって最適なアプローチがそれぞれの場合でどのようなものかは多かれ少なかれ正確に知られているだ

13 前註 7 の諸文献を参照。なかでもヘアの二層功利主義 (Hare 1981) はよく知られているだろう。

14 快樂主義については Parfit 1984: 6; Railton 1984: 140-6; Martinez 2011: 279 を、また以下の他の例については、Railton 1984: 144, 154; Wiland 2007: 291 を参照のこと。

ろう（各種のノウハウ本はそういう知見であふれている）。そしてこれらの目標を実現しようとする場合、我々はまさにその種の知識を用いて、いかに行為すべきかの指針を得ることができている。だとすれば、ある道徳理論から（その自己抹消性のゆえに）直接的な導きを得られないとしても、そこに特段の問題はないのではないだろうか。

3. 間接的な導きにまつわる懸念

前節の議論によれば、道徳理論は自己抹消的であっても十分に行為指導的でありうる。なぜなら、たとえある理論 T が自己抹消的である——ゆえにその目標 A^T を直接追求することは T 自体により禁じられている——としても、 T を信じる主体は、現に成り立つ経験的条件の下で目標 A^T の実現に最も貢献する熟慮の方法を選び出すことで、日々の意思決定のための指針を得ることができるからだ。しかし、このような仕方で自己抹消的な道徳理論に行為指導性を認めることに対しては、いくつかの批判がありうる。本節ではそうした批判の中から、関連文献に見られる三つのものを取り上げ検討しよう。

指示内容が確定しない

一つ目のありうる批判は、自己抹消的な道徳理論は我々に確定した指示を与えてくれない、というものだ¹⁵。我々が具体的場面で自己抹消的な理論から引き出すことのできる指示は、必然的に代替的熟慮の選択に依存している。つまりその指示内容は、理論が与える目標 A^T をそのまま参照できる（自己抹消的でない理論の）場合とは違って、どのような代替的熟慮の方法を我々が先に選んでおいたかに応じて決まる。しかしここには、誤りの余地がある。これまでの議論ではあたかも、ある目標の実現にとって最適な熟慮が現実の経験的条件によって客観的に決まるなら、我々は自動的にそれを正しく選び出せるかのように語られていた。だがもちろん、実際にはそんな保証はない。つまり、 A^T の実現に最適だと思って我々が選んだ熟慮方法が実はそうでなかったという可能性は常に残るし、さらに言うところの可能性は、問題の最適性が経験的条件に依存する仕方がきわめて複雑であることを思えばかなり現実的な懸念でもあるだろう。しかしそうだとすると、我々は、たとえ T により最終目標は与えられていたとしても、個々の具体的な場面で何を目指すように T が指示しているのかについては、実は永久に確信をもてないことになる。つまり T を頼りにしても、実際の行為の場面では何ら確定的な指示を得られないことになる。これでは到底、理論 T が我々の行為に導きを与えているとは言えないだろう。

だがこの批判は説得的でない。指摘すべき点は二つある。第一に、たしかに我々は全知ではないから、最適な熟慮方法がどれかについて思い違いをする可能性はゼロにはならないけれど、そのことから、我々がこの点に関してもちうる意見のすべてが等しく無根拠だとするのは誤りだ。我々は（ミルが示唆していたように）様々な経験や科学的知識の蓄積をもとに、ある目標にとって最適だと信じるのがより理にかなっているような熟慮方法を割り出していけるし、その推測の精度を高めてもいける。つまりたとえ不可謬の確信ではないとしても、我々は理論 T が何を指

15 これを示唆するものとしては、Stocker 1976: 463, Wiland 2008: 390f を参照。

示しているかに関して十分に合理的な信念をもてるわけだ。さらに指摘の二点目として、こうした合理性以上のものを我々が望むべきでないと考えるべき理由がある。上の批判では、理論の指示内容を確実に知ることができないという点が自己抹消的な理論に特有の問題であるかのように論じられていたが、実は同様の問題は自己抹消的でない理論にも生じうる¹⁶。つまり、ある目標を熟慮の際に直接参照できる場合でも、与えられた行為の選択肢のうちのどれがその実現にとって最適かに関して我々が誤る可能性は十分にある。この可能性は、一般に道徳理論は十分に広い適用範囲をもった——それゆえ抽象的な——目標を与えようとするという事情からしてもある程度避けられないものだろう。（例えば、おそらく自己抹消的でないカント的義務論の場合でも、具体的な選択肢の中でどれが定言命法にかなっているかを確定することが容易でないケースはあるだろう。）最善の選択肢を選ぶという課題は、選択の対象が熟慮タイプではなく行為トークンであってもすでに十分難しいのであって、自己抹消的でない理論の場合も我々が望めるのはせいぜい合理的な選択だけだ。したがって自己抹消的な理論の場合にも、我々は十分な合理性をもってその指示内容を推測できればそれで満足すべきだろう。

何らかの望ましい事柄と衝突する

続いて二つ目の批判は、自己抹消的な理論による導きは何らかの望ましい事柄を損ねてしまう、というものだ。その理由としては、関連する二つのものが考えられる。

第一の理由は、我々の生における「心理的調和」という論点に訴えるものだ¹⁷。一般的に言って、我々自身の中で価値判断と動機づけが一致している状態は望ましいと言えるだろう。M・ストッカーが言うように、「我々は最低限、自分が主要な価値を認めているものによって動機づけられ、また自分の主要な動機が向かうものに価値を認めるべきだ。[...] このような調和は、よき生の一つのしるしなのである」(Stocker 1976: 454)¹⁸。だが——と反論は始まる——もし我々が自己抹消的な道徳理論によって導かれるべきだとするなら、こうした調和を手に入れることは難しくなる。なぜならその場合、我々はあるタイプの事態 (A^T) に価値を認めながらも、それを実現しようという動機によって導かれてはならず、むしろそれとは別のタイプの（代替的熟慮において目指される）事態の実現への動機が支配的になるよう自らを律さねばならないからだ。つまり自己抹消的な理論による間接的な導きは、我々の価値体系と動機体系の間に分裂を生じさせ、よき生の条件である心理的調和を実現困難にってしまう恐れがある。

これに関連する第二の理由は、「道徳的熟慮」というものに関する次のような（一見もっともな）見方から生じる¹⁹。道徳的熟慮は本来、そのつどの状況で事物が特定の道徳的価値をもつことを可能にし、その価値を基礎づける特性——幸福を最大にする、有徳者の振る舞いと合致する等の、各道徳理論がそれぞれに価値構成的だと主張する特性——に対する注視と応答という仕方で行われるべきだ。というのも、ある熟慮はまさにその種の価値構成的な特徴と関係することによって

16 この点は、本稿での議論とは焦点の当て方は異なるが、Wiland 2008: 389f でも指摘されている。

17 Stocker 1976: 453-5, 461-3; Cox 2006: 508; Keller 2007: 222; Martinez 2011: 281; Pettigrove 2011: 193.

18 こうした調和の望ましさはその不在のケース、すなわち、自分が価値を認めないもの（薬物、児童ポルノ、等々）を欲してしまった、自分が価値を認めるもの（健康、社会的活動、等々）を欲せなかったり、といったケースと比べてみれば納得しやすいだろう。

19 Pettigrove 2011: 193. また Cox 2006: 510 にもこれと関連する議論がある。

道徳的熟慮になるのだし、また本来の道徳的熟慮は、そのつどの状況の中で行為の「理由」を構成するような特徴を見出し、その状況でなぜある振る舞いが正しいかの洞察をもたらすようなものであるべきだからである。しかし——ここからが反論だ——自己抹消的な理論によって導かれる主体の熟慮は、こうした本来の姿とはかけ離れたものになってしまう。なぜなら自己抹消性の定義より、そのような主体が日々の実践で行う代替的熟慮は、道徳的価値を構成すると彼が信じる特性——当の理論の目標 A^T に関わる特性——とは別の何かに目を向けるものとならざるをえないからだ。つまりその主体は日々の熟慮において、彼に行為の理由を与えるはずの事柄を洞察するどころか、絶えずそこから目を逸らしあさっての方向ばかりを気にかけることになってしまう。

しかし以上の批判も、自己抹消的な道徳理論にとって痛手にはならない。まず一つ目の理由に関して言うと、ある主体が価値づける事態と彼の動機づけが向かう事態が単に「同じでない」というだけで、そこに何か避けるべき不調和があるかのように語るのは明らかに大げさだ。もちろん、それらが同一でないケースの中には実際に望ましくない——ストッカーの言葉では「分裂症的」な——ケースもあるだろうが（註 18 を参照）、すべてのケースがそうである必要はない。とりわけ、自己抹消的な理論によって導かれる主体の場合のように、主体の動機づけが向かうタイプの事態が、彼が価値づけるタイプの事態の実現頻度や確率を高めるようなケースでは、その主体は前者のタイプの事態に対して少なくとも道具的な価値を認めることができる。つまりこのような主体においては、価値判断と動機づけの体系は無理なく調和するのであり、そこに避けるべき有害な衝突はない²⁰。

また二つ目の理由にも次のように応答できる。ある自己抹消的な理論 T によって間接的に導かれている主体を考えよう。そして、この主体が目標 A^T の実現のため身につけている代替的熟慮は、 A^T とは別のタイプのある事態「 E^T 」（「Ersatz aim of T 」の意）を直接的目標として行われるようなものだとしよう。（例えば T が行為功利主義なら、その代替的目標 E^T はミルが言うように「常識道徳の諸規則が順守されること」になるかもしれない。）そうするとこの主体は、たしかに日々の熟慮の中で A^T ではなく E^T の実現をもっぱら気にかけることになるが、このことは必ずしも、彼が本当に重要な（価値構成的で理由供給的な）事柄から目を背けているということを意味しない。というのも、この主体はまさに A^T の実現を重視するからこそ E^T の実現を気にかけるのであって、彼の E^T への関心はあくまで A^T への関心から派生したものである——ゆえに、もし A^T の実現に最善なのは別の熟慮方法だと信じていたなら彼は E^T への関心をもたなかった——ということは十分ありうるからだ。こうした場合に主体が決して A^T を軽視しているわけではないことは、次の類比からもわかるだろう。ある戦闘機のパイロットが遠方の標的を射撃する際、その最善の方法は、当の標的を直接見るのではなく、機内にあるモニターを見てそこに映る枠の中に（標的を表す）印が入った状態で発射ボタンを押すことだとしよう。この場合、機内のモニターに最大限の注意を集中しているパイロットは、たとえ標的に直接目を向けてはいないとしても、それを軽視しているわけでは当然ない（むしろ明らかに重視している）。

20 ただしここで、この主体が価値づける事態の実現にとって最善の手段が、彼の忌み嫌う何らかの事態の実現を目指すことだったらどうするのか、という懸念が生じるかもしれない。これに対しては二つの答えが考えられる。第一に、自分が忌み嫌う事態の実現を目指すことは相応の心理的負担になるから、そうした手段が最善となる見込みはあまり高くない。また第二に、たとえそれが現時点での経験的条件の下で最善になるとしても、自己抹消的理論はそれを勧めないとも解釈できる（以下の第 4 節後半を参照）。

もっともここで批判者はさらに、本来の道徳的熟慮において求められているのは、道徳的に重要な特性の単なる「重視」ではなく、その種の特徴に対するリアルタイムの「応答」だ、という点を再度強調するかもしれない²¹。だが、自己抹消的な理論の支持者がこの主張を認める必要はないだろう。第一に、この主張は実質的に論点先取である疑いが濃厚だ。そもそも自己抹消的な道徳理論とは、それが道徳的に価値ありとする事態 (A^T) を我々が日々の実践で直接意識すべきでないような理論のことだった。しかしここで批判者は単に、我々はそういう事態を直接意識すべきだと言い張っているだけのように見える。また第二に、上の批判者の主張は、道徳的熟慮や判断の本性に関するある特定のメタ倫理的立場を前提しているように思われる。すなわち前提されているのは、道徳的熟慮や判断をある種の知覚に似た心のはたらき、つまり状況内の特性の直接的認知の一種として捉える立場だ (cf. McDowell 1998: ch.7)。言うまでもなくこのような立場は（見込みがない訳ではないとしても）かなり実質的なものであり、上の主張もその分問題含みなものになる。

不可能なことを要求する

最後に、三つ目の可能な批判は、自己抹消的な道徳理論による間接的な導きが機能するためにはある不可能事が達成されねばならない、というものだ。その不可能事とは、同一の主体が同時に矛盾した信念をもつことである。この点は、M・スミスのものを簡略化した次のような例を通じて明らかにできる (cf. Smith 2001: §2)。道徳理論 T の例として通常の功利主義を考えよう。この T は、もっぱら行為者中立的（ないし不偏的）な観点から各人の幸福に着目し、世界の可能的諸状況の望ましさをその合計の大きさに基づき順序づける。それゆえ T によれば、「誰のものであれ幸福の総和が最大になること」が目標 A^T である。しかし（実際にありそうな話として）、T は自己抹消的であり、現実の経験的条件の下でこの目標が最もよく達成されるのは、我々が T そのものではなく行為者非中立的な順序づけ体系、つまり、身近な人の幸福をそうでない人の幸福より重く見積もるような（身内びいき的に偏った）体系に基づいて熟慮する場合である、と想定しよう。そうすると、T を信じる主体が身につけるべき代替的熟慮は、「自分に近い者の幸福が最大になること」を目標 (E^T) としたものになる。さてその上でいま、T を信じるあなたは、自分の行為 (e.g. 重い荷物をもって運んであげること) によって次の二つの状況のどちらか一方だけを現実化できるとしよう：①あなたの友人の幸福が5単位上昇する；②まったく見知らぬ人の幸福が6単位上昇する。このときあなたは、この二つの状況のうちどちらが望ましい（より大きな価値をもつ）と判断すべきだろうか。一方で、T を信じる（公平無私な）者としてのあなたは「状況① < 状況②」と判断すべきだが、他方で、代替的熟慮に従って行為する（身内びいきする）者としてのあなたは「状況① > 状況②」と判断すべきだ。しかし明らかに、あなたは同時にこれらの判断に同意することはできない。こうして、自己抹消的な理論による間接的な導きはある不可能事を要求することになってしまう²²。

だがこの批判も説得的でない。明らかにいまの議論では、あるタイプの熟慮をひとが実行する

21 Pettigrove 2011: 192; Martinez 2011: 280.

22 ここでの問題を、「T への信念」と「T が求める傾向性」の間の衝突から生じるものとすれば、以下の註 24 と 28 で触れる批判は、この同じ衝突を逆の角度から問題視するものと言える。

ときには、それに対応した命題を「本気で」信じていなくてはならないと想定されている。上の例で言えば、問題の主体が代替的熟慮に従って状況①の実現を状況②の実現より優先することから、彼は二つの状況の価値について「状況①>状況②」と判断しなくてはならない、とされていたわけである。しかし実際には、我々はある命題を本気で真と信じることなしにもそれを実践上の前提 *presupposition* として採用することができ、その前提に基づく意思決定を行うことができる。例えばある会社の苦情処理係は、「カスタマーが言うことはすべて正しい」という命題を（真とは信じないまま）実践的前提として採用することで、自分の業務を円滑に進められるかもしれない（Sainsbury 2011: 140）。あるいは、心についての厳格な消去主義者は、心的語彙が実際には系統的に誤っていると確信しつつも、実践的には（例えば「悲しんでいる人を慰めよ」と命じられた場面では）その種の語彙を用いてどう行為するかを決められるだろう。また同様のことは、プトレマイオス天文学の天球仮説、ニュートン力学、理想気体の法則などを（厳密には偽だと信じつつも）用いて実践的な意思決定を行うようなケースにも言えるだろう。これらのケースが示すように、一般に我々がある熟慮方法をとるために、それに対応した命題を信じている必要はない。それゆえ、ある道徳理論 T の判定とは矛盾した行いを勧めるような熟慮を行う主体（上の功利主義者のように）も、 T と矛盾した信念をもつ必要はない。

4. 自己欺瞞を命じるのではないかという懸念

前節では、「自己抹消的な道徳理論は間接的に我々の実践を導きうる」という主張に対するいくつかの批判を検討し、それらには十分応答可能であることを示した。しかしこの種の道徳理論への批判の余地はまだ残っている。本節で検討したいのは、自己抹消的な道徳理論はそれを信じる主体にある種の「自己欺瞞」を命じることになるのではないか、という懸念だ²³。この懸念の内容は、少し長くなるが以下のような考察により与えられる。

ある自己抹消的な道徳理論 T が、ある主体の実践を導いているとしよう。我々はこれまで、そのような主体は、 A^T の実現のための代替的熟慮を身につけているだけでなく、 T を信じ続けていると暗に想定してきた。つまりその主体は、代替的熟慮の傾向性と同時に「 T は真だ」という信念をもっている、と（前節の第二と第三の批判はまさにこの想定から生じていた）。だがもしかすると、これは T が誰かの行為を導く際の最良の仕方ではないかもしれない。もともとの選ばれ方からして、いま問題の代替的熟慮は、 T の目標 A^T とは異なる——そしてときにそれと食い違う——何らかの事態（ E^T ）を直接的な目標として進められるものだ。しかし第1節でも見たように、一般に T への信念をもつ主体は、少なくとも A^T の実現への潜在的な動機をもつ。それゆえ、問題の代替的熟慮を行うときに主体は、 T への信念をいったん脇におき、その動機的な効力をいわば「割り引く」といったことをしなくてはならない。だがこれは、もっぱら代替的熟慮にいそむべき主体にとっては余計な作業ではないだろうか。この主体はむしろ、 A^T に特に価値はなく、この事態が実現されようとされまいと重要な違いはないと思っているとき、つまり、 T を真ではなく偽だと信じているときの方が、代替的熟慮をうまく進められる——結果、本来の目標 A^T の実

23 Parfit 1984: 42; Cox 2012: §5.

現にもより近づける——のではないだろうか。

実際、この考えを支持するいくつかの根拠がある。その一つ目は、集中・注意力の問題だ。何らかの比較的単純なタイプの代替的熟慮——例えば第2節で示唆したような単純な規則に基づくもの——を行う場合でも、しばしば主体には、自分がおかれた具体的状況を注意深く観察し、その中で微妙な差異を見分けることが求められる。だがもし主体の中に、 A^T が実現されるべきだという（ T への）信念が残っていたとすれば、 A^T への意識はときに彼の注意力を奪い、代替的熟慮の観点から集中して状況を観察することを妨げるかもしれない。こうした集中の妨害は、とりわけ速断が必要な状況では有害になる恐れがある。

第二に、 T への信念を保持していることで、代替的熟慮の恣意的な中断への誘惑が生じる可能性もある。 T を信じる主体にとって、 A^T は定義からして、代替的熟慮における目標 E^T よりも本来は価値の高い目標である。それゆえ、 A^T と E^T が食い違うと思われた場合には、主体はつい代替的熟慮を中断して A^T を直接追求したくなってしまう、というのはかなりありそうなことだ。（例えば第2節で見た行為功利主義者は、「嘘をつくな」という規則を破った方が大きな幸福を生み出せることが確実だと思えるような状況でこうした誘惑を感じるだろう。）しかしこの誘惑は、主体の行動から一貫性を失わせたり、結果的に誤った判断へと導いたりといった仕方で、結局は彼を A^T の実現から遠ざけるかもしれない。

第三に、D・コックスらが指摘する興味深い可能性がある。それは、主体が T への信念をもっていることで、代替的熟慮とそれに基づく行為が無価値なものと感じられるようになってしまうという可能性だ（Cox 2012: 299f; Wiland 2008: 386）。これはとりわけ、 T による正しい行為の基準があまりにも高く厳格であるケースで起こりうる。例として、ある義務論は、非常に要求の大きな行い（e.g. 受けた恩は二倍にして返すこと）を我々が果たすべき義務として定めているとしよう。そして、ある時点でその理論を信じた一人の主体が、これよりも無理なく自分にも続けられる行い（e.g. 受けた恩をできる範囲で、せいぜいそれと同等に返すこと）を代替目標として設定し、普段はもっぱらその代替目標の実現に努めるようになったとする。さていま、この主体が当の理論の目標を改めて思い起こす機会をもったとしよう。このとき、もし彼がこの理論の高い要求を真正な義務として真に受けすぎるとすれば、彼はそれまで自分が行ってきた慎ましやかな善行（程々の恩返し）とその本来の義務の間にある大きな隔たりを強く意識せざるを得ず、それによって無力感を深めた結果、その慎ましやかな善行をばかばかしく思うようになる（そしてやめてしまう）かもしれない。この場合逆に、もしこの主体が当の義務論を偽だと信じ、自分の行いは非の打ちどころがなく善いと満足できていたら、彼の行いは結果的にその理論の目標により近づいていただろう。つまりこのケースは、ある道徳理論への信念がかえってその目標の実現から主体を遠ざけてしまう場合として理解されうる。（Cox 2012: 299f が論じるように、これと同様の議論は第2節で見た行為功利主義や徳倫理の理論に関しても可能だ。）

このようなわけで²⁴、我々の現実の経験的あり方をふまえる限り、ある自己抹消的な理論 T の目

24 以上に加えた四つ目の理由として、愛情や友情に関する論点も挙げられうる。多くの道徳理論（特に帰結主義や義務論など）は、本来の善悪の基準は不偏的なものだとして主張しながら、日々の熟慮においては愛情などの特別な人間関係に基づいて（偏った仕方）思考することを勧めるかもしれない。だがそうすると、当の不偏的な道徳理論への信念は、この特別な関係を損ねてしまう可能性がある（先の第3節で見た第三の反論と Smith 2001 を参照）。これに対する応答は、後註 28 と Railton 1984: 146f を参照。

標 A^T が最もよく実現されるのは、主体が適切な代替的熟慮の傾向性に加え、「T は偽だ」という信念——「T は真だ」という信念でなく——をもつ場合だという可能性は十分にある。しかも今みた論拠の一般性からして、このことは、稀であるどころかむしろ大抵の自己抹消的な理論について成り立ちそうにみえる。だがそうだとすると、実のところ多くの自己抹消的な理論は、それを信じる者にある受け入れがたい指示を発していることになるのではないだろうか。ここで思い出すべきは、一般に道徳理論 T は、我々の習慣や心的傾向性ができるだけその目標 A^T の実現に適したものになることを求めているという点だ。したがって、もし「T は偽だ」という信念をもつことが A^T の実現にとって最適であるなら、T はまさにその信念をもつよう我々に指示するはずである。だがこの指示は、すでに T を真と信じている者からすれば、自分の信じている理論を偽だと信じるようになれ、というものに他ならない²⁵。この指示の受け入れがたさは明白だろう。

ただしこの指示の受け入れがたさは、その「実行不可能性」に由来するものとして捉える必要はない。たしかに、我々は信念を自由自在にもったり捨てたりできるわけではないから、あることを信じるようになれという指示は、文字通り解釈すると心理学的に不可能な要求でしかないように見える。しかし我々は、ある程度の範囲内なら、何らかの因果的なはたらきかけ（e.g. 催眠術、洗脳、脳手術を受けるなど——cf. Parfit 1984: 41）によって、自分の信念状態を意図したものに变化させることはできる。したがって、ここで理論 T が発する指示は（実行不可能なものとして解釈しないとすれば）、「T は偽だという信念が自分の中に生じるよう自らにはたらきかけよ」というものとして解釈できる。

しかしたとえ実行不可能ではないとしても、この指示はやはり別の意味で受け入れがたい。というのも、それは端的に言って「自己欺瞞」の指示に他ならないからである²⁶。一般に、自分の目から見て偽であるような信念を他者に植えつけることは明らかに正しくない。だとすれば、そのようなことを自分に対して行うよう命じる道徳理論が正しいはずはない。もちろんここで、この手段によって得られる結果のよさ（ A^T のよりよい実現）が引き合いに出されるかもしれないが、そもそも我々は道徳的善を実現するための道具ではないのだから、この不正な手段は正当化されないだろう。しかも、いま植えつけが問題になっている信念は、道徳的善悪という我々の行為全般の価値に関わるきわめて根本的な種類のものだ。その根本的な部分に関して自分を欺けと命じることは、自分の生の全体をまやかしの上に築けと命じることに等しく、到底受け入れられない（cf. Cox 2012: 299）。したがって、そのようなことを命じる自己抹消的な道徳理論はやはり拒否されるべきである。

応答

さて以上の批判は、自己抹消的な道徳理論を拒否するための説得的理由になっているだろう

25 ただし T は、すでに T を信じている者以外に対しては、T をはじめてから信じさせよう命じることになるだろう。つまり T は、シジウィックが「秘教的 esoteric」（Sidgwick 1966: 490）な道徳理論と呼んだものになるだろう。この意味での秘教性が特に深刻な問題ではないという主旨の議論は、例えば Railton 1984: 155; Wiland 2007: 295f にある。

26 しばしば「自己欺瞞」という語は、ある信念をもったままそれと矛盾した信念を自分自身に植えつける過程（ないしそれにより生じた状態）を指すために用いられる（いわゆる自己欺瞞の「静的パラドクス」が問題になるときはこちらが念頭におかれている。Cf. 浅野 2012: 第 1 章）。だが本稿ではこの語を、ある時点でもっていた信念を放棄してそれと矛盾した信念を自分自身に植えつける過程（ないしそれに伴う状態）を指すために使う。自己欺瞞をめぐる議論において「パスカルの賭け」と呼ばれている過程は、この後者のタイプの過程を例示するものだ（cf. 浅野 2012: 第 3 章（特に 121-2 頁））。

か。私の見るところその答えはここでも否である。それを示すためにまず略記法を導入しておく、以下では、道徳理論 T は真だという信念を「 $B(T)$ 」と表し、また、 T は偽だという信念を「 $B(\neg T)$ 」と表すことにする。そうすると上の批判は、大部分の自己抹消的な道徳理論 T について次の二つが成り立つ、と主張するものとして整理できる：

主張 1. 現実の経験的条件の下では、 T の目標 A^T が最もよく実現されるのは、我々が信念 $B(T)$ をもたずに $B(\neg T)$ をもつ場合である²⁷。

主張 2. もし道徳理論 T について主張 1 が成り立つなら、 T はある時点で信念 $B(T)$ をもった主体に対して、その信念を捨てて信念 $B(\neg T)$ をもつよう命じる。

そして批判者はこれらに基づき、一般に自己欺瞞を命じるような道徳理論は受け入れがたいのだから、上の 1 と 2 を満たすような理論 T も受け入れがたい、と論じていた。

この批判に対してはもちろん、個々の道徳理論 T ごとに別個の対応を行うことも可能だろう。だがここでは、より一般的な応答を試みたい。そうした個々の対応の成否は、結局は以下でとり上げる論点に帰着することが見込まれ、個々の T に固有な事情に依存しない議論の方がその論点を際立たせやすいと思われるからだ。またここでは、上の批判へのそうした一般的な応答のうちでも、その最後のステップを拒否する——自己欺瞞を命じるような道徳理論が誤っているとは限らないと主張する——道には立ち入らないことにする。おそらくこの道をとることも不可能ではないが (cf. Parfit 1984: § 16)、自己欺瞞の道徳的地位に関してはすでにかなり議論の蓄積があり (cf. Jenni 2003)、この限られた紙幅で多少なりとも明確な結論を出すのは難しいと思われるからだ。

というわけで我々の問題は、上記の主張 1 と 2 が妥当かどうかという点になる。この点を考える上で重要なのは、この両者に共通して現れる主張 1 の読み方だ。いま注目すべき部分が際立つように要約すると、主張 1 が述べているのは、 A^T の実現にとって、「現実の経験的条件の下では $B(\neg T)$ をもつことが最適だ」ということである。しかしこの主張には、実のところ二つの可能な読み方がある。すなわち一方で、この主張は、我々の「経験的条件」を現時点でのそれとして厳密に捉えた上で、短期的・一時的な最適性を主張するものとして読まれうる。つまりこの場合、1 が述べるのは、「我々の経験的なあり方（心理的・身体的特性や社会的・生物学的環境など）がいま現在のままである限りは $B(\neg T)$ をもつことが最適だ」ということになる。この読みの下では、1 は我々の経験的あり方が現在のそれとは異なるような状況に関しては何も述べていないわけだ。しかし他方、主張 1 は我々の「経験的条件」というものをより長いスパンで緩やかに——現時点でのあり方を一つのフェーズとして含むが、それと異なったものへも移り変わりうるものとして——捉えた上で、長期的・総合的な最適性を主張するものとしても読まれうる。この読み方下では、主張 1 は、「我々の変化していく（しうる）経験的あり方に照らしてトータルでみれば $B(\neg T)$ をもつことが最適だ」ということを意味するものになる。

この区別を念頭において見直してみると、たしかに主張 1 は、いま言った一つの目の意味で読ん

27 より正確には、我々がこの信念状態に加え適切な代替の熟慮の傾向性を同時にもつ場合、であるが、この点はいまの論点に関わりがないので以下では省略する。

だ場合には真になるかもしれない²⁸。すなわち、上で挙げられたいくつかの論拠が示唆するように、我々のいま現在のあり方だけを見て評価する限りは、 A^T の実現にとって最適なのは $B(\neg T)$ をもつことだと言えるかもしれない。しかしここで重要なのは、上の批判が全体として成り立つには、主張1は前段落の二つ目の意味でも真にならなくてはならないという点だ。このことは、a) 上の批判が多義性の誤謬に陥るのでないとするれば、主張1はもう一方の主張2が真になるのと同じ意味で真でなくてはならず、b) 主張2が真になるのは、その前件である主張1を前段落の二つ目の意味において読んだときだけである、という二つの事実から帰結する。このうちa)についてはよいとして、b)が成り立つ理由についてはもう少し説明の必要があるだろう。すでに見たように、一般に道徳理論 T は、その目標 A^T の実現に最も貢献するような心的傾向性をもつことを我々に勧める。しかしもちろん、 T が我々にもつよう勧めるのはあくまでも長期的・総合的に見て A^T の実現に貢献するような傾向性であり、ある傾向性が単に特定の条件下でその実現に貢献するというだけでは、 T がその傾向性を我々に勧めるための条件として十分でない。つまり、主張2の前件(=主張1)が前段落の一つ目の意味で読まれたときには、その前件は必ずしも後件の成立を保証しない。その保証がある(=主張2が真になる)のは、主張1を前段落の二つ目の意味で読んだ場合に限るのである。

だがこの二つ目の意味で読んだ場合、主張1は真にならないと考えるべきもっともな理由がある。つまり長期的・総合的に見て、我々が $B(T)$ の代わりに $B(\neg T)$ をもつことが A^T の実現にとって最善であるということはまずありそうにない。その理由はひとこと言えば、そのような信念状態にある主体は、必要が生じた場合にも代替的熟慮の方法を適切に選び直すことができず、それは A^T の実現にとって致命的に不利だから、というものだ。

この点の説明のため、ある自己抹消的な理論 T を信じる主体が、現時点である代替的熟慮の方法 d を身につけているとしよう。こうした主体に関して、いま注目すべき点は二つある。第一に、彼はときに、 d 以外の熟慮方法を選び直すべき(それが合理的であるような)状況に直面しうる²⁹。例えばそうした状況としては、以前この主体が d を選んだときに成り立っていた経験的条件が何らかの仕方で変化し、もはやそこから d の最善性が帰結しないようなものになる(そして彼がそれに気づく)ケースがある。あるいは経験的条件そのものは不変でも、何らかの経験的発見により A^T の実現にとって d が最善だと信じることがもはや合理的でなくなるケースや、さらにはもっと単純に、主体が何らかの誤った経験的信念や誤った推論によって d を選んでしまっていたようなケースもありうる。おそらく、熟慮方法の見直しが必要になるこうしたケースはさほど稀でもないだろう。

しかし注目すべき第二点として、これらのケースにおいては、理論 T に関して主体がどういう信念をもっているかが、その後の A^T の実現されやすさを決定的に左右することになる。すなわち一方で、もし主体が $B(T)$ を保持しているなら、彼は d に代わる新たな熟慮方法を選ぶために必

28 ただし実はすでにこの点に関しても、大いに異論の余地がある。というのも第一に、行為功利主義の場合にそう考えられているように、ある道徳理論 T に於ける代替的熟慮が複数の規則を含むような場合、規則同士が衝突する状況が生じることが予想され、それらの葛藤を解決するためには本来の目標 A^T (幸福の最大化)を参照することが不可欠になるはずだからだ(cf. Mill 1996: 71; Hare 1981: ch.2)。また第二に、仮に A^T の実現にとって最善の代替的熟慮が愛や友情といった偏向的な傾向性に基づくとした場合も、 A^T が代表する不偏的な視点をもっている方が、愛情の過剰や不足の際でも安定した振る舞いが可能になるということは十分ありうる(Railton 1984: 146-8)。

29 もちろん第2節でも触れたように、熟慮方法の選び直しと身につけ直しには相応のコストがかかるだろう。ここで考えているのは、そうしたコストを差し引いてもなおそれらを実行すること(これは場合によると教育を通じた世代間作業になるだろう)が A^T の実現によりよく寄与するようなケースである。

要な参照点（ $= A^T$ ）をもつことができる。しかし他方、主体が $B(T)$ を捨て $B(\neg T)$ をもつようになっているとすれば、彼はもはや A^T を目標と捉えていないのだから、その実現に適した熟慮方法を彼が選り直せる見込みはごくわずかしかな（そもそも熟慮方法を選り直すべきであることに気づかない恐れもある）。よって主体が $B(\neg T)$ をもつ場合には、ある種の状況において A^T が実現される確率は著しく低下することになる。そしてこうした（おそらくさほど稀でない）状況への対応能力を奪うことになる以上、 $B(\neg T)$ をもつことが総合的・長期的にみて A^T の実現にとって最善であるということはまずありそうにない。むしろ A^T の実現のためには、我々はやはり $B(T)$ を保持しているべきだと考えられる³⁰。（ここでの事情はちょうど、ある生物種が現在の特定の生態環境に対してどんなに適応的だとしても、多少の環境上の変化に順応できないならば、種の存続という目的にとって総合的・長期的に有利だとは言えないのと類比的だ。）

もっともここで批判者からは、次のような応答があるかもしれない。たしかにいま論じられたように、信念 $B(T)$ を保持していることには A^T の実現に対し一定のプラスの効果があるかもしれない。だがここで我々は、この信念は同時に様々な仕方で——代替的熟慮への集中を妨げたり、代替目標を無価値なものだと感じさせたりといった仕方で—— A^T の実現に対しマイナスの影響も及ぼすということを忘れるべきでない。この影響は、やはり決して無視できないものなのではないか。つまり、 $B(T)$ をもつことによるマイナスの効果は場合によると、長期的に見てもそのプラス分を上回ることになり、その結果、 A^T の実現にとって最適なのはそれと反対の信念 $B(\neg T)$ をもつことになる（＝主張1が14頁で述べた二つ目の意味で真になる）、という可能性は十分あるのではないか。

繰り返しになるが、筆者の見立てでは、 $B(T)$ の代わりに $B(\neg T)$ をもつことの不利性は前述のようにならかなり致命的で、いま述べられた可能性も実際はあまりありそうにない。だが、これはたしかに経験的と言え経験的な問題なので、以下では議論の「保険」として、もし仮にその可能性を認めたとしても、上の批判を拒否する余地はなお十分に残る、ということを示しておきたい。より正確に言う以下で示したいのは、たとえ現実の経験的条件下では長期的に見ても $B(\neg T)$ をもつことが A^T の実現にとって最適であるとしても、理論 T が我々に $B(\neg T)$ をもつよう命じていると解釈する必要は必ずしもない（主張1が先の二つ目の意味で真になるとしても主張2を認める必要はない）、ということだ。

このことを示すため、まず次のように想定してほしい。いまあなたは、 A という事態を実現されるべき目標と捉えている一方で、 ϕ という行為をすることはどうしても避けたいと考えている。しかしあるときあなたは不本意にも、「現実の諸条件の下では、 ϕ をすることが A の実現のための最善の方法だ」という判断——例えば、「現在の収益率が続くなら、その海外支店をたたむことが会社の存続にとって最善だ」という判断——に至る³¹。このとき、もしあなたにとって、目標 A を

30 主張1の不成立を示そうとする以上の議論は要するに、 T が主体に対し $B(\neg T)$ をもつよう指示することはない、というものだ。しかしここで次のような疑問があるかもしれない——たとえいま問題の主体が $B(\neg T)$ をもつ（ $= A^T$ を道徳的に価値ある目標と捉えていない）としても、主体が A^T の実現を（道徳以外の観点から）望んでさえいれば、彼は d に代わる適切な熟慮方法を必要な場合を選ぶことができるはずだ。それゆえ、 T が主体に対し、 $B(\neg T)$ をもちながら A^T の実現を望むよう指示すること（したがって当然 $B(\neg T)$ をもつよう指示すること）は十分ありうるのではないかと、このような疑問である。しかしながら、この主体が A^T の実現を望むとすれば、彼の熟慮の過程はやはりその欲求によって影響を受けることになるだろう。そしてこの欲求は、本文でみたような理由（特にはじめの二つ）から、代替的熟慮の実行を妨げるだろう。よって実際のところ、もし理論 T が $B(\neg T)$ をもつよう指示するとすれば、 T は A^T の実現を望まないよう指示することになると考えられる。

31 ここでは、「 ϕ = その海外支店をたたむこと」、「 A = 会社が存続すること」である。

放棄することはもちろん ϕ をすることもできない相談だとしたら、あなたはどうすればよいだろうか。その一つの答えは、この判断を真にしている現実の諸条件（e.g. 現在の収益率が続くこと）それ自体を変化させること、である。つまりこの場合あなたは、問題の諸条件に然るべき変化をもたらすこと（e.g. 収益率を改善すること）で、 ϕ を行うことがAの実現にとってもはや最善でないという状況をつくり出すよう試みることができる。要するに上の判断は、 ϕ を指示するものとしてではなく現状改革を促すものとしても解釈できるのだ（むろん一定の範囲内ではあるが）。

そしてこれと同様の解釈は、一見自己欺瞞を命じているように見える道徳理論の場合にも可能である。すなわち、たとえある理論Tを信じる主体が「現実の諸条件の下では、 $B(\neg T)$ を自分に植えつけることが A^T の実現のための最善の方法だ」という判断に至ったとしても³²、彼はこの判断を、必ずしも自己欺瞞の指示として理解する必要はない。むしろ彼はこの判断を、関連する経験的条件を変化させよという勧め、典型的には、自身の道徳的な思考能力を改善すべしという自己変革の勧めとして解釈することもできる³³。

以上より、本節で見た自己欺瞞に関する批判は説得的でない。なぜなら、それを構成する第一の主張は批判の成立に必要な意味では真でない公算が高く、また、たとえ第一の主張がその意味で真だとしても、第二の主張を拒否できる可能性は依然残るからである。

5. おわりに——非経験的な自己抹消性

本稿では、自己抹消的な道徳理論が直面するとされるいくつかの問題をとり上げ、それらが本当にこの種の理論にとって難点になるかどうかを検討してきた。ここまでの議論から得られたのは、それらは実際には深刻な問題を提起しないという結論である。

だが実のところ、道徳理論の自己抹消性に関しては以上で見てきたのとは若干異なる種類の問題がある。本稿の第1節で見たように、道徳理論の自己消去性は、「我々がその理論の与える目標を直接的目標として行為すると、その実現は逆に困難になる」という条件によって定義される。そして本稿ではもっぱら、この定義条件が、我々についての諸々の経験的真理に依存して成り立つようなケースを考察してきた。だが何人かの論者が指摘するように、この定義条件はときとして、そうした経験的真理なしでも成り立ちうる。つまり言ってみれば、「論理的・非経験的な自己抹消性」と呼びうるケースがあるのだ³⁴。例えば、ある種の状況において「親切な行為をすること」を目標として課すような道徳理論を考えることはできるだろう。しかしこの目標を実現しようという明示的意図は、当の目標の実現をいわば論理的に排除してしまうように思われる。というのも、ある他者Cの苦境に居合わせたとき、真に親切な行為者は「Cを何とか助けたい」といったCへの直接的配慮に導かれて手を差し伸べるのに対し、上の目標に直接導かれて行為する者は、同じ手助けをするにしても「親切な行為に該当することをしたい」といった考慮を経てそれをすることになる——そしてこれは親切な行為ではない——からである。もちろんこれと同様の議論は、

32 ここでは、「 ϕ =信念 $B(\neg T)$ を自分に植えつけること」、「 $A = A^T$ 」である。

33 もちろん、こうした現状改革がいつでも可能であるとは限らない（その可能性の有無は、Tの性質や主体のおかれた偶然的状況などに左右されるだろう）。しかしここでのポイントは、そうした改革の可能性がアプリアリには排除されない以上、上記の判断をそうした現状改革の勧めとして解釈できる可能性も一概には排除されないという点だ。

34 Cf. Stocker 1976: 462; Cox 2006: 511; Keller 2007: 225f; Martinez 2011: 279.

友情、愛情、勇敢さ、謙虚さ、等々についても成り立つ。だがそうだとすると、我々は上のようなタイプの目標を与える道徳理論とどう付き合えばよいのか。そうした理論を受け入れつつ、その目標の実現を排除しないことはどうやって可能なのか³⁵。これらの点についての考察は、また別の機会に譲ることにしたい³⁶。

参考文献

- Adams, R. 1976. Motive Utilitarianism, *Journal of Philosophy* 73: 467-81.
- 浅野 光紀 2012.『非合理性の哲学—アクラシアと自己欺瞞』、新曜社。
- Cox, D. 2006. Agent-Based Theories of Right Action, *Ethical Theory and Moral Practice* 9: 505-15.
- Cox, D. 2012. Judgment, Deliberation, and the Self-Effacement of Moral Theory, *Journal of Value Inquiry* 46: 289-302.
- Hare, R.M. 1981. *Moral Thinking: Its Levels, Method and Point*, Oxford UP.
- Hursthouse, R. 1999. *On Virtue Ethics*, Oxford UP.
- Jackson, F. 1991. Decision-Theoretic Consequentialism and the Nearest and Dearest Objection, *Ethics* 101: 461-82.
- Jenni, K. 2003. Vices of Inattention, *Journal of Applied Philosophy* 20: 279-95.
- Keller, S. 2007. Virtue Ethics is Self-Effacing, *Australasian Journal of Philosophy* 85: 221-31.
- McDowell, J. 1998. *Mind, Value and Reality*, Harvard UP.
- Martinez, J. 2011. Is Virtue Ethics Self-Effacing?, *Australasian Journal of Philosophy* 89: 277-88.
- Mill, J.S. [1871] 1998. *Utilitarianism*, fourth edition, edited by R. Crisp, Oxford UP.
- Parfit, D. 1984. *Reasons and Persons*, Oxford UP.
- Pettigrove, G. 2011. Is Virtue Ethics Self-Effacing?, *Journal of Ethics* 15: 191-207.
- Railton, P. 1984. Alienation, Consequentialism, and the Demands of Morality, *Philosophy and Public Affairs* 13: 134-71.
- Sainsbury, R.M. 2011. Fiction and Acceptance-relative Truth, Belief and Assertion, in F. Lihoreau (ed.) *Truth in Fiction*, Ontos: 137-52.
- Sidgwick, H. [1907] 1966. *The Methods of Ethics*, seventh edition, Dover.
- Smith, M. 2001. Immodest Consequentialism and Character, *Utilitas* 13: 173-94.
- Stocker, M. 1976. The Schizophrenia of Modern Ethical Theories, *Journal of Philosophy* 73: 453-66.
- Wiland, E. 2007. How Indirect Can Indirect Utilitarianism Be?, *Philosophy and Phenomenological Research* 74: 275-301.
- Wiland, E. 2008. On Indirectly Self-defeating Moral Theories, *Journal of Moral Philosophy* 5: 384-93.
- Williams, B. 1973. A Critique of Utilitarianism, in J. C. Smart & B. Williams, *Utilitarianism, For and Against*, Cambridge UP: 77-150.
- Williams, B. 1995. Acting as the Virtuous Person Acts, in R. Heinaman (ed.), *Aristotle and Moral Realism*, UCL Press: 13-23.

35 この点に関連する議論は、例えば Williams 1995; Hursthouse 1999: pt. II などにみられる。

36 本稿の執筆にあたり、匿名の査読者から多くの啓発的で有益なコメント・質問をいただいた。記して厚く感謝する。また本稿は、平成 27 年度文科省科研費の交付を受けた研究（課題番号 15H06088）の成果の一部である。

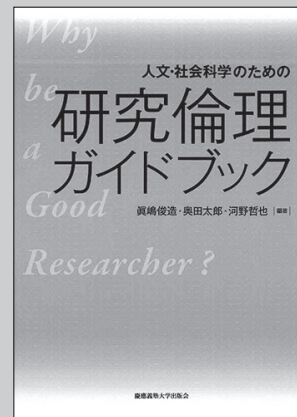
書 評

眞嶋俊造・奥田太郎・河野哲也編

『人文・社会科学のための研究倫理
ガイドブック』

(慶應義塾大学出版会、2015)

吉永明弘 (江戸川大学)



2014年のSTAP細胞に関する問題は、従来から問題視されていた「研究不正」に対する取り組みを加速させた。2015年から、日本学術振興会の科学研究費を申請する際に、申請者に研究倫理のテキストを読むことが義務づけられた¹。そこで「研究不正」とされているのは、捏造 (Fabrication)、改ざん (Falsification)、盗用 (Plagiarism) の問題であり、3つの頭文字をとってFFPと呼ばれている。その対極にあるのが「誠実な研究活動」であるが、その間には大きなグレーゾーンがあり、それは「好ましくない研究行為」(QRP: questionable research practice) と呼ばれている。「研究不正」の防止という観点からは、このあたりが研究倫理の主要なトピックとなるだろう²。

ここで紹介する『人文・社会科学のための研究倫理ガイドブック』の中でも、上記のような「研究不正」の枠組が解説されている。ただし、本書で扱われる「研究倫理」のトピックは多岐にわたっており、FFPやQRPに気をつけるためのマニュアルとして本書を手にとると、内容の豊富さ・多彩さに驚くことになる。以下では章ごとに、その内容を紹介する。

序章(眞嶋俊造)は、本書の導入部分という位置づけだが、独立した論考としても読める。第一に、研究倫理がテーマ化された背景、研究倫理の定義、研究不正とグレーゾーンの問題、研究倫理の射程(ミクロ、中範囲、マクロ)が述べられる。第二に、研究倫理は単なるコンプライアンスでもなく、また難しい学術研究でもなく、むしろ「処世術」と解されるべきものであると説かれる。第三に、専門職倫理の観点から研究倫理が考察される。専門職の倫理綱領は①コンプライアンスモデル(禁止の倫理)、②社会的責任モデル(大志の倫理)、③自己実現モデル(喜楽の倫理)に分類できるとした上で、現在の日本の研究倫理の議論には喜楽の倫理が欠けていることが示唆される。第四に、研究者の徳目が列挙される。第五に、倫理テスト(専門職テスト、普遍化可能性テスト、危害防止テスト、公開可能性テスト)が紹介され、その具体的な運用の仕方が述べられる。

1 日本学術振興会「科学の健全な発展のために」編集委員会編(2015)『科学の健全な発展のために——誠実な科学者の心得』丸善出版。

2 前掲書、46頁～52頁。本書では、序章第四節・第五節(4頁～6頁)、および第二章第三節(95頁～103頁)に、この枠組が示され、解説されている。ただし「誠実な研究活動」に代えて「責任ある研究活動」(responsible conduct of research: RCR)という言葉が用いられ、またQRPは「好ましくない研究活動」と訳されている。

第1章（神崎宣次）では、資料の収集に関する倫理問題が論じられる。文献資料の引用等については他の執筆者に譲る形で、ここでは歴史的価値を持つ資料の扱い方や、人間を対象とした場合の研究方法などに紙数が割かれている。具体的には研究者による資料の持ち去りや、現地調査の場での研究者の行きすぎた行動（被災地での過剰なアンケートなど）が問題視されている。研究者は「単なる調査や資料収集の対象ではない存在として接しなければならない」のであり、「得られる情報やデータという学術的価値以外の価値がそこには存在していることを、われわれは研究者として決して忘れてはならないのである」（47頁）。さらには「研究成果の還元」という言葉が免罪符のように使われている状況が批判され、研究者と対象者の共同的な社会実践が求められている。この論点はコラムで論じられる「アクションリサーチ」へとつながっていく。

第2章（金光秀和）では、高度科学技術社会における研究不正の影響、不正行為を誘う社会環境、CUDOSという理想とPLACEという現実の乖離、不正行為（FFP）／責任ある研究活動（RCR）／好ましくない研究活動（QRP）の区別といった、基本的な議論の枠組が要領よくまとめられている。その上で、「FFPは確かに起こった場合の影響は大きい、それよりも研究者が日常的に直面することが多いのは好ましくない研究活動（QRP）」（100頁）であるという認識のもとで、QRPの具体例（虚偽の陳述、不正確、偏向）が紹介される。最後に、研究行為のマネジメントの問題が、計画段階と実行段階に分けて論じられる。計画段階で考慮すべき点として、被験者保護、動物実験、利益相反の問題が、実行段階で留意すべき点として、データの管理や、共同研究における責任分担が挙げられている。

第3章（土屋敦）では、研究不正に関して、FFPに加えて、オーサーシップ問題と二重投稿、研究費の流用が論じられる。また研究不正の背景に若手研究者の就職問題などがあることが指摘される。そして人文社会系に特徴的な研究不正は、引用ミスや参考文献の抜け落ち、「既存研究で公表されている概念やアイデアを自覚的か無自覚かにかかわらず自分が考えたものと思い込んでしまう」（124頁）ことにあるとされる。それらを防ぎ、人文社会科学が健全に発展するためには、研究不正事象を広く公開するしくみが必要とされる。

第4章（新田孝彦）では、研究環境に関わる倫理問題が、管理職（トップマネジメント）と指導教員（ミドルマネジメント）としての研究者の立場から論じられる。ここでは組織倫理・企業倫理の枠組が援用され、研究不正を防止するには、組織風土に注目し、組織のマネジメントを問題にしなければならないとされる。そこで管理職には、ミッションと倫理規範の明確化、適切な研究目標の設定、適切なインセンティブの付与、適切な研究者の雇用、研究者の適切な管理（目標管理と秩序管理）、研究者への適切な教育訓練が求められる。また指導教員には、トップマネジメントの下で研究倫理を遵守しながら研究し、同時に学生に対して良好な研究環境を提供し、学生の成長に責任を負うという、「プレーイングマネージャー」の役割が期待されるという。

第5章（河野哲也）では、人間社会一般に対する「研究者の社会的責任」について、事例をもとに論じられる。第一に、学会ジェンダー問題として、『人工知能』の表紙に描かれたイラストの問題性が分析される。第二に、イラク戦争や3.11などの時事問題に対する研究者の情報発信のあり方が論じられる。そこでは、研究者は、政治判断や価値判断から中立的な事実の発信にとどまらず、規範的な主張を行うべきであり、批判よりも提案を中心にした発信を行うべきだと主張される。第三に、脳ブームにおける科学の過剰な通俗化と不適切な一般化を例にして、研究者間の

相互批判・相互チェックの必要性が語られる。以上をふまえて、「研究者の社会的責任とは、研究を研究者のためのものにしないこと」とされ、その点から「「専門家意識」などは、硬直した自己意識や高慢、公共性を欠いた仲間びいきに陥る可能性を常にはらんでいる」ので気をつけなければならないとされる（180 頁）。

第 6 章（奥田太郎）では、文部科学省が 2014 年 8 月 26 日に公表した研究不正ガイドラインの内容が、専門調査員の総評やパブリックコメントをふまえて分析される。パブリックコメントの中には、研究者の負担増を心配する声や、不正防止強化・厳罰化および調査・措置の公正さを求める声もあるが、不正が誘発される環境についての言及もあった。つまり「過剰な成果主義ゆえに雇用が不安定化し、研究費確保の競争が激化すること」（202 頁）が不正の土壌となっているのではないかという疑義があった。この点について、2006 年の旧ガイドラインにおいては、このような研究不正が起こる背景についての記述があったが、2014 年の新ガイドラインではそれがなくなっていることが指摘される。また研究不正が新聞で話題にされるようになったのは 2000 年代以降であり、その背景には、研究不正を誘発する環境が生じただけでなく、「内部告発を受容する社会的な土壌が整ってきた」ことがあったと分析されている（206 頁）。

以上、各章の内容を概観した。本書ではこれら以外に、たくさんのコラムが収録されている。これらのコラムは単なる付録ではなく、味読すべき内容が詰まっている。「研究方法別 column」では、文献研究、インタビュー調査、フィールドワーク、アクションリサーチ、社会調査、実験のそれぞれに倫理問題があることを知ることができる。「分野別 column」では、ジェンダー、非民主主義体制の地域の研究、経済学を用いた政策提言、ジャーナリズム研究による二次被害、法学における研究倫理が扱われる。それぞれについて、当事者である執筆者の切なる声が聞こえてきて、読んでいる私も身が切られる思いがする。

このように本書には多彩な内容が盛り込まれているが、その中で特筆すべき点を四点挙げてみたい。第一に、「人文・社会科学のための」研究倫理として、FFP だけでなく、フィールドや社会調査において直面しがちな倫理問題が幅広く取り上げられていること。第二に、研究不正を単に研究者の心の問題と捉えるのではなく、研究不正を生み出す土壌となっている社会環境に論及していること。第三に、各章の議論が、執筆者によるこれまでの応用倫理研究の蓄積によって下支えされていること（例えば第 2 章には科学技術倫理、第 4 章には組織倫理・企業倫理、第 5 章には脳神経倫理、第 6 章には内部告発の倫理の議論が織り込まれている）。第四に、特にコラムにおいて、執筆者の肉声が聞こえてくること。

最後に私見を述べる。応用倫理学の研究者は、対象となる倫理問題から距離を置いて分析・検討するだけでなく、倫理問題を主体的に引き受け、社会に対して何らかの応答を行うべきと考えるが、研究者倫理はまさに研究者自身の問題であり、応用倫理学にとっての試金石となる領域と言えるだろう。本書の事例を読んでドキドキし、本を置いてわが身を振りかえって考える、それを繰り返すことが、本書の正しい読み方かもしれない³。

3 本稿執筆にあたって、2015 年 10 月 10 日に上智大学で開催されたシンポジウム「研究公正と研究倫理を問い直す——理念・制度・教育」における、伊勢田哲也氏、笹川光氏、奥田太郎氏の講演を参考にした。

書 評

文・伊勢田哲治／マンガ・なつたか

『マンガで学ぶ動物倫理——わたしたちは動物とどうつきあえばよいのか』

(化学同人、2015 年)

浅野幸治 (豊田工業大学)



本書は、ひじょうに良い、驚くほどよい、感動的な入門書である。主たる登場人物が高校生であることから考えて、本書が想定する読者層は主に高校生であろう。主人公は今上淋太郎という高校2年生であり、淋太郎の相方を勤めるのが級友であり幼なじみでもある根岸清音である。2人は、生き物探偵を始める。そこに舞い込んでくる依頼人からの依頼によって、物語が展開していく。そして各章のマンガ頁の後には、文章による講義がついている。巻末には、文献案内があり、参考となる映画・小説・マンガの紹介もある。

本書の章立ては、以下の通りである。

第1章 ペットのしつけ

第2章 殺処分と去勢

第3章 化粧品動物実験

第4章 肉食と集約的畜産業

第5章 動物園

第6章 外来生物

第7章 医療のための動物実験

第8章 野生動物による被害

第9章 イルカ・クジラ漁問題

第10章 人間と動物の権利

一見して明らかなように、第1章と第2章がひとまとまりであって、本書は、伴侶動物という身近な問題から動物倫理の話に入っていく。次いで第3章と第4章は、歴史的に見て現代の動物解放運動の出発点となったような問題群である。第5章と第6章、第8章と第9章は、動物の権利という考え方の応用問題と言ってよいだろう。この章立てに関して1つ気になるのは、医療のための動物実験の位置である。第3章と第4章で動物実験や集約的畜産業について考えた淋太郎が「動物をせまい所に閉じこめることについて」疑問を感じて、その関心を動物園の問題（第5章）へ向けるのは分かる。その意味で、第3章と第4章と第5章は一続きである。それに対して第6

章では、野生動物の問題に転じるように見える——実際に、第1章から第5章までは、基本的に人間の管理下にある動物の問題であった。ところがなぜか、医療のための動物実験に関わる章が、第6章外来生物の問題と第8章野生動物による被害の問題の間に差し挟まれている。たしかに、化粧品のための動物実験と医療のための動物実験を別建てで論じることは、十分に理解できる。けれども、医療のための動物実験の章は動物園の章のすぐ後に置けばよいのであって、動物園の章と医療のための動物実験の章の間に外来生物の章を割り込ませる理由はないだろう。最後の第10章では、淋太郎が、「人間と動物の関係にまつわる謎」に対して、「人間だけを特別扱いする根拠はない」「人権とか人間の尊厳なんていうのは根拠のない約束事、作り事なのです」という回答を披露する。ただし、この回答は皆の賛同を得られない。ということは、こういう答えはダメですよ、でも皆さんも自分で考えてみてくださいね、という終わり方と言ってよいだろう。入門書だから、答えを読者に押し付けないわけである。

上記のように、伊勢田は、動物倫理という書名の下で期待される話題をほぼ網羅的に取りあげている。表立って論じられていないのは、動物曲芸ぐらいである。そして各章の講義の部分で、伊勢田は、その章の問題を分かりやすく、非常に適切に解説している。そのように立派な本について、私に一体何を言うことができるだろうか。本書は入門書なので、動物倫理にまつわる問題や考え方を紹介し解説している。その点では、誰が書いても同じような内容になる。そうした内容については特に言うことはない。私が注目すべきなのは、伊勢田に特徴的な部分であろう。それは第10章の講義の部分である。そこで伊勢田は、人権を認めた場合に動物の扱いについて私たちが取りうる5つの選択肢と、もう1つ別の選択肢との合計6つの選択肢を提示する。それらは以下の通りである。

- 1、人間だけが人権をもつと、説明なしに言い張る。種差別の立場。
- 2、人間のもつある特徴が人権の根拠になり、他の動物にはその特徴がないと論じる。
- 3、言語能力など他の動物がもたない特徴を人権の根拠として挙げたうえで、人権をもたない人間がいることも認める。
- 4、人権のもとになる特徴を共有している他の動物にも人権と同等の権利を認める。動物権利論の立場。
- 5、人権よりも1つ格下の権利として動物の権利を考える。
- 6、権利という考え方を止めて、功利主義で行く。

以上6つの選択肢を概観したうえで、伊勢田は結局、次のように述べる。

いろいろな立場の長所短所を見比べたうえで淋太郎の答えを考え直してみると、実はそれほどの外れではないかもしれません。

もしこれが伊勢田の本音であれば、「不評であるかもしれないけれども、人権に根拠はない」というのが伊勢田の立場であるのかもしれない。だからといって、伊勢田の立場が危険思想だということになるかどうかは分からない。伊勢田も、すでに成立している約束事としての人権は認める

だろうからである。言い換えると、「人権が根拠のない約束事にすぎない」という言い方が否定しているのは、自然権としての人権である。

もう少し追及してみよう。伊勢田自身の立場はどこにあるのか。それが垣間見られるのは、「おわりに」と題する部分である。そこで伊勢田は、動物慰霊や動物供養を手がかりにして「日本人にも心から腑に落ちる動物倫理を構築する」ことを示唆している。実際に伊勢田は、「動物実験の倫理——権利・福祉・供養」と題する論考を別のところで書いている。けれども本書の範囲内では、淋太郎の回答——すなわち人権に根拠はない——に対する2つの相反するような評価が到達点である。一方で、人権は棄てがたいので、淋太郎の回答は正しくない。他方で、前節で述べた6つの選択肢はどれも受け入れられないので、淋太郎の回答は間違いでもない。そういう行き詰まり状態である。だから、「そのいずれでもない動物供養の倫理ができたらいいな、でもまだできない」というのが偽らざる心情だろうか。

最後に、批判的なことも書いてみよう。伊勢田は、上記の第4の選択肢を退ける。たしかに、動物の権利という考え方を受け入れるならば、肉食や動物実験だけではなくて伴侶動物も動物園も否定される。だから、動物の権利という考え方からは、動物解放論が出てくる。この場合の動物への基本的態度は、「危害を加えるな」というものである。こうして畜産も動物実験も、動物園も伴侶動物もなくなったら、動物との関係で残るのは野生動物の問題だけである。野生動物と人間の関係の基本は、棲み分けであり、野生動物による被害に対して田畑は柵などによって守ることができる。（しばしば、シカなどが増え（すぎ）ていると言われる。「増える」というのは、時間軸上の比較である。では、どの時期と比べて増えているのだろうか。たしかに、絶滅の危機にあったような時期と比べるならば、シカなどは個体数を回復しているだろう。けれども、例えば江戸時代と比べた場合、シカなどは本当に増えているのだろうか。疑わしい。他方、江戸時代と比べて人間の数は驚くほど増えている。なにも私は、人間が増えすぎているから人間を駆除せよなどと言いたいわけではまったくない。ただ人間がこの150年ほどの間に非常に増えて、野生動物の生存圏を侵食してきたということも、伊勢田には指摘しておいてほしい。）このような動物権利論は、「一本筋の通った態度」であろう。では、どうして伊勢田は動物権利論を退けるのか。

動物の権利という考え方への道行きは、2つある。1つは、人権から動物の権利へという道行きである。この道行きを伊勢田は正面切って取りあげ検討している。しかしながら、もう1つの道行きは、動物福祉から動物の権利へという道行きである。例えば、3Rの1つである洗練原則は、「できるだけ動物にとって苦痛の少ない方法」にすることを要求する。苦痛が少ないほうがより良いのであれば、苦痛がないのが一番良いのではないか。10の苦痛よりも5の苦痛、2の苦痛よりも1の苦痛のほうが良いのであれば、同じ論理で1の苦痛よりも0の苦痛のほうが良いということにならないか。たしかに伊勢田は、動物福祉と動物の権利とをきちんと説明している。動物福祉の考えの中心は、苦痛の削減である。ブランベル報告書にある「5つの自由」を見たときに明らかのように、そこに欠如しているのは「死からの自由」である。これが動物福祉と動物の権利との重要な違いである。しかし、このような動物福祉の立場は整合的に維持可能だろうか。動物福祉という考えには、動物の権利へと向かう論理が内在していないだろうか。例えばブランベル報告書の5つの自由の中には、「正常な行動を表現する自由」がある。動物が殺されたとき、動物は「正常な行動を表現する自由」を奪われないだろうか。奪われるだろう。ということは、「正常な

行動を表現する自由」を認めたならば、死からの自由も認めることになるだろう。要するに、動物福祉という考えが動物の福祉を気にかける以上、死（殺害）が動物の福祉に反しないと考えることは極めて困難である。実際に、福祉に反する危害としては、さまざまな傷害よりも死が最大である。その証拠に、傷害よりも殺人のほうが罪は重い。だからいやしくも動物の福祉を気にかけるならば、動物の命を気にかけることになる。伊勢田は動物権利論を退けるけれども、動物福祉という考えは認めるようである。そうすると、この動物福祉から動物の権利へという道行きを、どのようにして伊勢田は避けることができるのだろうか。

『応用倫理——理論と実践の架橋』第10号 論文公募のお知らせ

『応用倫理——理論と実践の架橋』編集委員会では、応用倫理学に関する研究論文、研究ノート、書評を下記の要項・投稿規定において公募いたします。なお、投稿は随時受け付けておりますが、第10号への掲載は2016年5月31日までの投稿を目安とします。皆様の御投稿をお待ちしております。

1. テーマは応用倫理学に関わるものとする。
2. 論文は独創性を有する学術研究成果をまとめたものとし、研究ノートは萌芽的研究の中間報告等とする。
3. 応募論文および研究ノートは未発表のもので、本『応用倫理』以外に同時投稿していないものに限る。二重投稿の場合、審査対象としない。

※ただし、外国語で既に発表された論文を著者本人が日本語に訳したものを投稿する場合は二重投稿とは見なさない。また著者本人が外国語で発表した論文に基づいて執筆されたために、もとの外国語論文と内容的に重複が多い場合も二重投稿とは見なさない。その際には投稿する際に、当該外国語論文の翻訳であること、ないしは当該外国語論文に基づくものであることを別紙により記載して申し出ることとし、掲載が決定した際には、その点を論文の中で明記することとする。

4. 使用言語は日本語とする。英語論文については *Journal of Applied Ethics and Philosophy* にて受け付ける。
5. 論文および研究ノートの分量は1万～2万字を目安とする。書評は2000～4000字程度とする。
6. 論文または研究ノート投稿者は『応用倫理』編集事務局に、①論文または研究ノートの原稿、②論文または研究ノートの和文要旨（500字程度）および英文要旨（250語程度）、③著者略歴（100字程度）の電子媒体テキスト（MSワードのファイルを記録したCD-Rを添付）およびハードコピー3部を送付する。また電子媒体のものは本センター事務局宛にメールでも送付すること。
7. 書評投稿者は、『応用倫理』編集事務局に書評原稿を電子テキスト（MSワードによる添付ファイル）にて送付する。
8. 投稿された論文及び研究ノートは、編集委員会が定める査読者2名により審査され、編集委員会において選考される。
9. 編集委員会は査読者の審査の結果を踏まえ、投稿者に対して修正・書き直しを求めることができる。修正・書き直し後に再投稿されたものについては、必要に応じて再査読を行う。
10. 掲載可となった論文及び書評は、ウェブページ及び冊子体により公開する。
11. 掲載の可否については編集委員会が最終決定を行う。

※ 本誌の査読はダブル・ブラインドで行っているため、論文本体には著者氏名は書かず、「拙論」等の表現も使わないこと。

○ 過去の本誌の内容は、北海道大学応用倫理研究教育センターのウェブサイト上及び北海道大学学術成果レポジトリ「HUSCAP」でご覧いただくことができます。 <http://ethics.let.hokudai.ac.jp>

<http://eprints.lib.hokudai.ac.jp/dspace/bulletin.jsp>

論文送付先／問い合わせ先

〒060-0810 札幌市北区北10条西7丁目 北海道大学大学院文学研究科 応用倫理研究教育センター

（電子媒体テキスト送付アドレス） E-mail: caep@let.hokudai.ac.jp

応用倫理——理論と実践の架橋 vol. 9

2016 年 3 月 31 日発行

編集委員長

蔵田伸雄

編集委員

柏葉武秀、近藤智彦、田口茂

眞嶋俊造、宮嶋俊一、山田友幸

©2016 応用倫理研究教育センター

ISSN 1883-0110

〒 060-0810

札幌市北区北 10 条西 7 丁目

北海道大学大学院文学研究科

応用倫理研究教育センター

Tel : 011-706-4088

E-mail : caep@let.hokudai.ac.jp

URL : <http://ethics.let.hokudai.ac.jp/>