応用倫理―理論と実践の架橋―

Vol. 16
2025 年 6 月

北海道大学大学院文学研究院
応用倫理・応用哲学研究教育センター

目　次

今号に投稿された原稿については以下の方に査読を御願いいたしました。査読へのご協力に感謝いたします。

石田　知子	冲永　宜司	小野原雅夫	西條　玲奈
佐藤　岳詩	佐野　勝彦	田中　綾乃	田原彰太郎
中川　　大	成田　和信	松本　大理	森岡　正博
吉田　修馬			

（編集委員は除く）

特集編集協力　稲荷森輝一

研究ノート

文献サーベイ：日本における道徳的エンハンスメント論の論点整理

菅家諒

要旨

本稿は、科学技術によって人間の道徳性を向上させる「道徳的エンハンスメント」をめぐる国内の哲学論文のサーベイである。道徳的エンハンスメントをめぐり倫理的な議論がどのように展開されているかに焦点を当て、その論点の網羅的な整理を行うことを目的とする。本稿ではまず、道徳性を向上させるために用いられる手法に応じて、道徳的バイオエンハンスメント（MBE）と、道徳的 AI エンハンスメント（MAIE）を区別する。この両者について、伝統的な教育の手法（道徳教育）と比較した上で、それぞれ（１）何を改善するのか、（２）どのような倫理的な懸念点があるのかを整理する。サーベイの結果、前者の論点については、MBE は動機づけの改善に、MAIE は洞察の改善にそれぞれ重点があること、またいずれも行為を直接に改善するものではないことが確認された。後者の論点については、社会的影響をはじめとする種々の論点を包括的に見て、MAIE は（技術的に可能である場合には）MBE について言及される懸念点のいくつかを回避しうることが確認された。しかしながら両者とも、道徳教育と比較すれば多くの懸念が指摘されることが確認でき、今後も慎重な議論が必要とされることが示唆された。

Abstract

This study examined philosophical papers in Japan on “moral enhancement”, the enhancement of human morality through science and technology. It aims to provide a panoramic review of moral enhancement, focusing on how ethical debates are developing around the issue. We first distinguish between moral bio-enhancement (MBE) and moral AI enhancement (MAIE) based on the methods used to improve morality. After comparing these methods with traditional methods of education (moral education), this paper outlines (1) aspects they improve and (2) ethical concerns they may raise, respectively. The results of the survey showed that MBE focuses on improving motivation while MAIE focuses on insights, and that neither of them directly improve behavior. Further, taking a comprehensive view of various issues, including social impact, MAIE may (if it is technically feasible) avoid some of the concerns mentioned in respect of MBE. However, it was confirmed that both still raise a number of concerns compared to moral education. This finding suggests that careful examination and discussion are required.

1　はじめに

近年、技術の進歩と社会の変化により、人間の能力や特性を向上させるエンハンスメント（増強）技術が注目されている。エンハンスメントとは、医療などと同様の手法や技術を用いながら通常の治療を超えて人間の能力を増強・強化するものであり、個人の生活の質を向上させる可能性を秘めている。しかし、その利用には倫理的な問題や社会的影響が付随するため、慎重な議論が必要とされる。

数あるエンハンスメントの中でも、道徳的エンハンスメントが近年注目を集めている。道徳的エンハンスメントとは、科学技術によって個々人の道徳性を増強させるものであり、社会全体のモラルを向上させる可能性があるとされる。この技術に対して、様々な角度から批判がなされ、それに対応する擁護、さらにそれに対する再反論といった形で活発な倫理的な議論が行われている。しかしその全体像については、未だ客観的な理解が得づらい状況にある。そこで本稿は、文献サーベイを通じて、とりわけ日本における道徳的エンハンスメントをめぐる倫理的な議論がどのように展開されているかに焦点を当て、その論点整理を試みる。

本稿の構成は次の通りである。まず２節では、道徳的エンハンスメントの定義と概略を示す。続く３節では、どのように個々人の道徳性を向上させるか、その方法・アプローチを３つに分けそれぞれの概要を示す。４節では何をエンハンスメントするか、すなわち何に焦点を当て働きかけるかの分類枠組みを提示する。５節では、３節で挙げた３種の方法それぞれのどこに問題があると指摘されてきたかを列挙し、該当する懸念点のマッピングを試みる。

2　道徳的エンハンスメントとは何か - 概略 -

立花（2009）によれば、エンハンスメント[1]は「身体的エンハンスメント（physical enhancement）」、「知的／認知的エンハンスメント（intelligent ／ cognitive enhancement）」、「道徳的エンハンスメント（moral enhancement）」の３つに大別される。また、高木（2020）はこれらにムードエンハンスメント[2]を加え、４つに分類する。注目すべきは、両者とも道徳的エンハンスメントを考察の対象としている点である。従来はその具体例に乏しく、現実的な見込みは小さいと考えられてきた道徳的エンハンスメントであったが、個人の意識や感情に直接影響を与えるようなエンハンスメント技術が開発・実装される可能性の高まりや社会的なニーズの変化などを理由に、注目が高まっている[3]。

道徳的エンハンスメントを論じ始めたのは、I・ピアソンとJ・サヴァレスキュである。高木（2020）によれば、彼らは論文「認知的エンハンスメントの危機と人類の道徳的性格をエンハンスすべき緊急の責務」

1　エンハンスメントという語には、医療技術を治療を超えて用いることに焦点がある場合と、科学技術を用いた能力の向上や増強を幅広く指し示す場合があり、本稿は後者を主題とする。

2　ムードエンハンスメント（もしくは感情のエンハンスメント）とは、一般的には感情や気分を向上させるためのエンハンスメントを指す。これは、個人の心理的な状態を良好な方向に変化させることを意味し、悲しい、不安などのネガティブな感情からポジティブな感情への転換を目指すことが含まれる。高木によれば、例えば、日本では「愛情ホルモン」「幸せホルモン」とも呼ばれているオキシトシンの投与は、ムードエンハンスメントに該当する。ムードエンハンスメントの倫理的検討としては佐藤（2010）がある。この論文では、気分明朗剤の服用の是非が検討されるとともに、それを通じて価値あるものは快楽だけか、苦痛は価値を持ちえないのか、という快楽主義に対する疑義が投げかけられる。

3　近年では美的エンハンスメントも注目されている。佐藤（2021）は美容整形に焦点を絞り、それを身体的エンハンスメントの一種とみなした上で議論を進めるが、飯塚（2021）はそれ以外の美の実践（例えば日常的に化粧をすることや、髭を剃ったりすること、美容院やネイルアートなどに赴くこと）も美的エンハンスメントに含むべきだとする。ただし、どちらの議論でも「何が美しいとされるか」のある程度の基準は社会が決定するため、美的エンハンスメント技術が一方向的な偏った仕方で開発され、美容実践を画一化していく可能性が指摘される。道徳的エンハンスメントにおいても、「何が道徳的であるのか」の基準は社会や時代によって変化するが、エンハンスメント技術によってその価値観が画一化される可能性が指摘できると思われる。

（2008 年）において、道徳的エンハンスメントは喫緊の課題であり、強制的にみなに施されるべきであるとする急進的な主張を行った。森下（2016）によれば、サバレスキュらは「道徳的にエンハンスメントする」ことを、衝動をコントロールできること、認知能力を高めること、自己利益の計算力を高めること、共感能力を高めることなど、幅広く定める。これを受け森下は、その定義をより一般化し、道徳的エンハンスメントを個々人の道徳性を高めることとした上で、「自分が信じる正しい理由（根拠）にもとづいて判断したり行為したりできること、すなわち「自律」（Autonomy）を高めることに包括できる」（p.91）とする。他にも、竹下（2023）は、デビッド・デグラツィア[4]に依拠し、道徳的エンハンスメントを、「ある道徳的主体の道徳的能力の改善や道徳的能力の獲得・選択を目的にした何らかの処置や介入である（DeGrazia 2014: p.361）が、能力の発揮のための環境改善（情報提供や認知的アシストなど）も道徳的エンハンスメントに含まれる」（p.4）とする。

以上を踏まえ、その定義を改めて整理すれば、道徳的エンハンスメントとは、倫理や道徳性を改善・向上させるための手法や介入のことであり、科学技術の発展を活用して個々人のモラルや道徳的判断力を人間に観察される典型的な能力を越えて向上させることを目指すものである[5]。では、この介入は如何にして実行され得るのだろうか。その方法について、現在考えられうる 3 種（伝統的手段によるアプローチ、生物学的アプローチ、人工知能によるアプローチ）を以下にみていく[6]。

3　どのようにエンハンスメントするか - 方法 -

3. 1　道徳教育（伝統的手段によるアプローチ）

まず、我々の技術力が現在ほど進歩していなかったころのアプローチに注目してみる。森下（2016）は元来、個々人の道徳性を高めるという広義の道徳的エンハンスメントに携わってきたのは道徳教育であったとし、以下のように述べる。

> 伝統社会では洋の東西を問わず、宗教的権威に裏打ちされた「良心」や「仏性」、「誠意正心」、「清明心」といった心（魂）の理想が「養生」と結びつき、修養を通じてめざされた。
>
> そうした中で、西洋では 17 世紀の後半以降、いわゆる近代社会が形成されるにつれて、伝統的な道徳性そのものが改めて問い直されるようになる。その皮切りはホッブズによる「利己心」（自己保存欲望）の大胆な肯定であった。これに触発されて、「モラルセンス」、「同情」、「快楽計算」、「権利の平等」、「人類愛」、「友愛」等、新たな道徳思想が次々と打ち出される。そして 19 世紀後半以降、とりわけ 20 世紀に入ると、道徳教育の介入に関して（国民対人類という対抗軸を伴いながら）、心理学や教育学や社会学の分野を中心に組織的な方法が模索されていったのである。（森下 2016: pp.91-92）

森岡（2013）も歴史を振り返り、人々を道徳的にするための方略を模索するのは、人間が文明を立ち上げて以来もっとも大きな目標のひとつであったとし、次のように述べる。「たとえば、古代ギリシアや中

4　ここで言及される論文は DeGrazia（2014）である。

5　このような道徳的エンハンスメントの定義には疑問が残るとする見方もある。高木（2020）や森下（2016）によれば、そもそもそこで用いられる道徳性とは何か、どこにどのように介入すればエンハンスメントになるのか、を決定しなければならず、すなわち道徳的エンハンスメントは特定の道徳概念を前提せざるをえないという問題がある。この点については本稿でも後に触れる。

6　道徳的エンハンスメントにおける生物学的アプローチは主に生命・医療倫理学等の分野において議論されるが、伝統的手段によるアプローチと人工知能によるアプローチは主に性質や能力の向上・増強といったより広範なエンハンスメントの原義に関係する幅広い文脈で議論される。これらの違いについて指摘頂いた査読者に感謝する。

国の哲学者たちは、どうすれば人々が道徳的になるのを促進できるかを発見しようとした。これは道徳性のエンハンスメントの古代バージョンであったと言える。彼らはこの目標は適切な教育と習慣づけによって達成可能であると考えた」（p.109）。以上のように、既に我々は個々人の道徳性を高めることを目的に、伝統的な教育の手法を用いて、人々に倫理的な価値観や道徳に関する考え方を伝えてきた。

しかし、科学技術の急速な進歩に伴い、我々は道徳教育以外の方法を模索し始めるようになった。その方法こそが、道徳的エンハンスメントである。高木（2020）と森岡（2013）をもとに、サヴァレスキュらの主張は次のようにまとめることができる。科学技術の進歩は人間に幸福をもたらすだけでなく、逆に人間にとって脅威たりうる。科学知識の拡大と認知能力の拡大によって、ますます多くの人々の手に大量破壊兵器、あるいはそれらを配備する能力が行き渡るようになっている。このような状況は、科学技術の進歩についていけない人間の道徳性に関する危機として理解することができるだろう。そこで、彼らは善を増やすためではなく、悪を減らすために道徳的エンハンスメントは喫緊の義務であると考える[7]。つまり彼らは、道徳教育（伝統的手段によるアプローチ）では、その効果が迅速には発揮されず、我々の道徳が技術の発展に追いつかない状況を危惧し、より効果的に道徳性を高める方法として、道徳的エンハンスメントの必要性を主張したのである。

3.2　道徳的バイオエンハンスメント（生物学的アプローチ）

道徳的エンハンスメントの具体的な実践方法としてまず主張されたのは、生物学的アプローチである[8]。森岡（2013）によれば、サヴァレスキュらは経頭蓋磁気刺激法、脳深部刺激療法、遺伝子操作、標的光刺激などの認知神経科学の進歩によって人間の道徳的モチベーションや行動に影響を与えることができるようになるかもしれない、と主張する[9]。また、高木（2020）は彼らの考えを、次のように整理する。

> 伝統的・文化的手段による道徳的エンハンスメント、例えば家庭教育や学校教育は、同じ手段による認知的エンハンスメントに比べて、効果的ではなかったし、迅速ではなかった。というのも、我々は正や善の知識を獲得してもすぐに正や善を実行するとは限らないからである。そこで、遺伝学的・生物医学的手段による道徳的エンハンスメントの可能性が検討されねばならない。（高木 2020: p.22）

このような生物学的アプローチによる道徳的エンハンスメントとは、しばしば「モラル・バイオエンハンスメント（Moral Bio-enhancement）」（以下、MBE と表記する）と呼ばれる。MBE は生物学や医学の手法を用いて、個人の道徳的な能力や行動を向上させることを目指すものであり、具体的には脳科学や遺伝子工学などの科学技術を活用して、個々人のモラルや道徳的判断力を強化することで、倫理的な意思決定を促進しようとする試みである。

3.3　道徳的 AI エンハンスメント（人工知能によるアプローチ）

さらに、近年では人工知能（AI）技術の目覚ましい発展に伴い、MBE よりも個人の自由や自律の余地をより大きく残す、「道徳的 AI エンハンスメント（Moral AI-enhancement）」（以下、MAIE と表記する）が提案されている[10]。MAIE とは、AI を活用して個々人の道徳的な判断力や行動を向上させること

7　ここで言及される論文は Persson・Savulescu（2008）p.166,174 である。
8　サヴァレスキュらの他にも生物学的アプローチの必要性を主張するものとして、例えば Singer・Sagan（2012）が挙げられる。彼らは「道徳薬（morality pill）」について議論する。
9　ここで言及される論文は、Savulescu, Douglas, Persson（2014）である。
10　例えば、Savulescu・Maslen（2015）。

を目指す概念である。具体的には人々の意思決定を支援するアルゴリズムやシステム、ツールを開発し、人々が倫理的なジレンマや複雑な倫理的問題に対処する際に AI が補助やガイダンスを提供することで、人々がより道徳的な判断を下せるよう手助けをする試みである。竹下（2023）によれば、提案されている MAIE、または検討される AI アシスタントの方法には、道徳的判断を下しユーザーに伝える AI（QA-AI）、ユーザーに助言する AI（助言者 AI）、ユーザーと議論する AI（議論者 AI）の 3 種類がある。

4　何をエンハンスメントするか - 対象 -

つづいて、道徳教育や道徳的エンハンスメントは、それぞれ何を改善しようとするのか、何に対して働きかけるのかを整理する。

4. 1　分類枠組み

道徳的エンハンスメントの改善対象を分類する枠組みにはいくつかのものがある。MBE については、DeGrazia（2014）は改善対象を動機づけ、洞察、行動の三つに分類する。MAIE の文脈では、O’Neill et al.（2023）は AI アシスタントの利用による改善対象を動機づけ、情報・時間不足、基本的な道徳的知識・理解の三つに分類する。

これらの分類枠組みを受け、竹下（2023）はデグラツィアの分類にオニールらの分類の一部を追加することで、道徳的エンハンスメントの改善対象を以下の 4 つに分類する（p.5）。

①情報・時間不足の改善：関連した（非規範的）情報を提供し、認知タスクを代行する
②動機づけの改善：正しいことをするためのより良い動機づけ、性格をもたせる
③洞察の改善：何が正しいかについてのより良い理解をもつようにさせる
④行動の改善：正しい・より良い行為をより多く行うようにさせる

ただし、この分類枠組みにはカテゴリー錯誤があるようにも思われる。なぜなら、②〜④は道徳性を向上させるために個々人の何に介入するか、つまり直接的な改善の対象（what）そのものに焦点を当てているのに対して、①情報・時間不足の改善は、道徳性を向上させる際のそのプロセスや方法（how）に焦点を当てているからである。したがって、この枠組みを道徳的エンハンスメントの改善「対象」の分類とするのではなく、改善「目標」の分類とする方が適切であるように思われる。

また、②〜④の目標の中で、実際に道徳性が向上したことの効果や結果を、唯一客観的に捉えることができるのは④だけであると思われる。なぜなら、個々人が道徳性に関していくら正しい理解の上で正しいことをしたいという性格を持ったとしても、実際に行動に移さなくては、エンハンスメントの効果があったと判断できるかは疑わしいからである。

4. 2　分類

竹下（2023）の分類枠組みに依拠し、先の道徳教育（伝統的な手段によるアプローチ）と MBE（生物学的アプローチ）、MAIE（人工知能によるアプローチ）の改善目標は次のように整理することができる。

道徳教育の場合は、時間をかけ規範的な情報を伝えるため、①情報・時間不足の改善にはならないことは明らかである。しかし、教育を通じて②動機づけの改善と③洞察の改善を行うことで、結果的には④行動の改善を期待すると言えるだろう。

MBE について、先述の通りサヴァレスキュらは道徳教育では迅速な効果を期待できない点を危惧して

おり、それゆえ MBE は①情報・時間不足の改善を直接的な目標としていると理解できる。また森岡（2016）によれば、サヴァレスキュらの MBE の主な関心事は「我々の社会性を広げる態度、たとえば信頼や共感や寛容を増大する」[11] ことであるため、②動機づけの改善がその主たる目標であると換言できる。しかし、③洞察の改善は目標としない。なぜなら、本稿３節で見た通り、個々人は何が正や善なのかの規範的理由を獲得しても、それをすぐに実行するとは限らないからこそ、そのような理由を知らなくとも直ぐに動機づけをする MBE の必要性が主張されるからだ。したがって、MBE は①情報・時間不足の改善、②動機づけの改善が直接的な目標であり、それらを通して間接的に④行動の改善を期待すると言える。

最後に MAIE の場合、竹下（2023）によれば、④行動の改善については、強制的な介入でもない限り行動を直接的な対象とした道徳的エンハンスメントは困難である。また MBE と異なり、AI による動機づけの改善は、洞察を改善することで間接的に動機づけを改善するのでもない限り困難である [12]。したがって、MAIE は①情報・時間不足の改善および、③洞察の改善を直接の目標とし、それらの改善によって間接的に②動機づけや、④行動の改善を目指している。

4.３　分類表

以上の整理を受け、道徳教育、MBE、MAIE ３種の個々人の道徳性を向上させるアプローチがそれぞれ何を改善しようとするのかを表１に整理した。

表１：３種の方法の改善目標（筆者作成）

	道徳教育	MBE	MAIE
①情報・時間不足の改善	—	✓	✓
②動機づけの改善	✓	✓	（✓）
③洞察の改善	✓	—	✓
④行動の改善	（✓）	（✓）	（✓）

※直接の目標とするものに「✓」、間接的に目指す・期待するものに「（✓）」、目標としていないもの・改善を見込めないものは「—」とした。

※また、既に指摘した通り①と②〜④の間にはカテゴリーの相異があるように思われるため、表では破線で区別した。

表から、道徳教育と MBE には改善が期待されないもの、目標としないものがあることがわかる。一方、MAIE は①〜④全ての改善対象を包括的に含むといえる。

また、この３種の中に、道徳性の向上の結果として唯一客観的に捉えることが可能である④行動の改善について、直接的に期待できるアプローチはないことが確認できる。すなわち、これらのアプローチには結果が伴わない可能性があるということである。この点に関して、高木（2020）は以下のように述べる。

人間の能力はすべてあくまで自然的なものであるのに対して、道徳や法が人間の作り出した規範で

11　森岡は、自己犠牲と利他心の二つが道徳性の中心的な性質であるとし改善させるべしとするサヴァレスキュの見方に、否定的な態度を示している。なぜなら、ある一群の人々に対して道徳的エンハンスメントを実施、共感性や寛大さを増大させた後に、道徳的エンハンスメントから逃れる事ができる人々は効果的に彼らを支配し、搾取する可能性があると考えるからだ。この点に関して、本稿５節で再び触れる。

12　竹下は、Lara（2021）に依拠し、MAIE の改善対象を分類する。

ある限りにおいて、それ自体で道徳的器官・能力であるものは存在しない。この意味においては、ピンポイントで道徳的能力をエンハンスする「道徳的エンハンスメント」は厳密には考えられない。道徳的行為を可能にする能力は単一のものではなく、多様な能力の折り合いであり、いくつかの能力が間接的に道徳的能力を構成しているということである。例えば、理性と感情は、道徳性のためだけの能力ではないが、お互いに協力することで道徳的に正しい行為を主体に実行させる。[…] それゆえ、本稿が P&S[13] の理論に付け加えるのは、道徳性に関する認知的エンハンスメントの必要性である。例えば「誰が共感に値するのか、なぜ共感に値するのか」といった規範的理由を認知する能力のエンハンスメントである。[…] 行為と動機づけに関する道徳的エンハンスメントには、道徳性に関する認知的エンハンスメント[14] が伴わなければならない。（高木 2020: p.28）

すなわち、個々人の道徳性を向上させる上でより有効なアプローチは、行為と動機づけに加えて道徳的行為の規範的理由を認知することの改善、つまり③洞察の改善を伴う必要があると考えられる。例えば、ある人が差別をしたくないという動機づけを持っていたとしても、実際には何が差別に該当するのかの理解がないと、つまり規範的判断を下す能力が十分でないと、その人は不意に差別的な行為をしてしまう可能性がある[15]。そのため、②動機づけの改善だけでは十分ではなく、③洞察の改善を伴わなくては、④道徳的行動という結果の改善を導くことはできないと考えられる。

以上を踏まえると、③を直接の目標とする道徳教育と MAIE は、MBE よりその効果を見込める可能性が高いと言える。また、竹下（2023）によれば、3 種類の MAIE の方法の中でもユーザーと議論することでユーザーの道徳的理解を改善することが中心的な目的である議論者 AI が、洞察の改善に最も貢献しうると考えられる。

5　指摘される懸念点 - 考察 -

ここからは道徳的エンハンスメント全般に渡って指摘される懸念点を列挙し、その上で道徳教育と MBE、MAIE の 3 種の方法の、それぞれについてどの懸念点が該当するのかを整理する。

立花（2009）は、エンハンスメントを含め人間にまつわる科学技術を巡る議論を、（1）科学的・技術的可能性を巡る議論の領域、（2）人間性を巡る議論の領域、（3）政治的・法的・社会的実現可能性と影響を巡る議論の領域、の 3 領域に分ける。他方、高木（2020）、竹下（2023）、森岡（2013）は、道徳的エンハンスメントの効果の不確実性についても言及する。そこで本稿では立花の 3 領域を参考にしつつ、そこに効果を巡る議論を追加する。以上より、日本における道徳的エンハンスメント議論とそこに向けられる懸念は、大きく分けると以下の 4 つの領域にあると整理できるだろう[16]。

1．技術的可能性・安全性を巡る議論
2．社会的影響を巡る議論
3．価値論（人間性）を巡る議論
4．効果の不確実性を巡る議論

13　高木は論文の中で、I・ピアソンと J・サヴァレスキュを「P&S」と表記する。
14　「道徳性に関する認知的エンハンスメント」という表現は、必ずしも明確なものではない。なぜなら、道徳的エンハンスメントと認知的エンハンスメントが截然と区別されるものではない可能性があるからだ。本稿では暫定的に認知的エンハンスメントを道徳的エンハンスメントの一部として位置づけるが、この分類自体が曖昧であるという指摘は十分に考えられる。
15　この点は、竹下（2023）p.4 でも言及される。
16　以下の行論の都合上、立花の 3 領域については順序を入れ替えている。

5. 1　技術的可能性・安全性を巡る議論

まず検討されねばならないのは、技術的可能性や安全性を巡る議論である。すなわち、道徳性を向上させるための介入が技術的に実現可能であるか、安全性は確保されているのかという懸念である。立花(2009)は、当該のエンハンスメント技術が科学的・技術的に可能かどうか、そのリスクはどの程度なのか、が問題であり、またこれらを専門的に扱えるのは、そうした諸技術発展の担い手（科学者や科学技術者といった研究者集団・専門家集団）であると指摘する。

この懸念点は道徳教育ではなく、MBE と MAIE の２種に該当するであろう。MBE は脳科学や遺伝子編集などの科学技術を用いるが、医学的リスクや予期せぬ副作用が生じる可能性がある。MAIE の場合は、そもそものそのような AI を開発することは可能なのかが疑問視される。また竹下（2023）は、もし開発できたとしても、その AI が悪用されたり、不適切に使用されたりする可能性については、技術的に対処できる部分もあるが、できない部分に関しては問題を引き受ける必要があると述べる。

5 .2　社会的影響を巡る議論

（ i ）自由・自律を損なう

仮に上の問題がクリアされたとして、次に検討されるべきはその社会的影響である。まず、道徳的エンハンスメントは個人のモラルや道徳を向上させる可能性がある一方で、個人の自由や自律を侵害する恐れがあるという懸念を取りあげる。

その詳細を見る前に、この懸念を取り扱う際に注意すべき点がある。「自由・自律」という言葉が使われる文脈が、文献によって異なる場合があるということである。字面は同じ「自由・自律」であっても、その言葉に対する意味解釈が論者ごとに微妙に異なることで、問題関心の所在が曖昧になりうる。多くの文献で言及されていた論点は主に次の２つである。（１）エンハンスメント実施後の当事者の真の選択を妨げない、という意味の「自由・自律」、（２）エンハンスメントをする（される）／しない（されない）の選択をなす自由、つまり外部からの支配や社会的圧力を受けない、という意味の「自由・自律」である。他方、これらの解釈に限定されずに、（３）自律的な自己の発達を促進させるという文脈での「自由･自律」に言及するものもあった[17]。本稿では、第一と第二の２種の文脈での「自由・自律」を、一先ず社会的影響を巡る議論に分類する。というのも第三の文脈において論じられる自律性・主体性の問題は、次節の価値論（人間性）を巡る議論、とりわけ努力・成長の機会を奪うとする懸念にも分類可能であると思われるからだ。この点に関しては後に触れる。

以上を踏まえ、まずは一つ目の、エンハンスメント実施後の当事者の真の選択を妨げないという意味での自由・自律が侵害されるという懸念から論じていこう。これは、エンハンスメントが行われた場合、その個人の行動や意思決定にも影響が及び、自己決定能力が制限される可能性があるという指摘である。例えば、エンハンスメントによって善悪を判断する能力が一度固定されてしまうと、その判断基準において悪とされる行為をなす一切の自由がなくなってしまう。仮に時代状況に応じて善悪の基準が変化した場合、エンハンスメントを受けた個々人はその変化に対応することができない可能性がある。

この懸念点は MBE に該当する。高木（2020）によれば、サヴァレスキュらを徹底的に批判する J・ハリス[18] は、道徳的エンハンスメントは一度実行されると、道徳的に堕ちる自由の選択、すなわち「悪への自由」を侵害してしまうと考える。つまり、エンハンスメントを受けた人は、単に反道徳的行為ができなくなるだけであり、なぜその行為が道徳に反しているかの理由を必ずしも理解していないことになるので

17　例えば、堀内（2020 b）。
18　ここで言及される論文は、Harris（2011）である。

はないか、それどころか、理由を認知できないロボットに変えてしまうのではないか、という懸念があるということである。

次に二つ目の、エンハンスメントをする（される）／しない（されない）の選択をなすという意味での自由・自律が侵害されるという懸念を取り上げる。サヴァレスキュらは、道徳的エンハンスメントの効果を発揮させるためには全人類に強制させることが必要と主張する[19]が、その場合、処置を受けたくない人にも強制的に処置を行わざるを得ない。このとき、当事者は外部からの支配や社会的圧力を受けることになり、個人の自由・自律を侵害することになる点が指摘される。また、佐藤（2021）はこの点に関して、大規模に道徳的エンハンスメントを実行するということになれば、政府などが一元的に行う必要があるだろうが、道徳的エンハンスメントという結果を現時点で査定ができない不確実な技術のために、個々人への重大な介入を認めることはできない、とする。

この懸念点も、むろんMBEに当てはまる。立花（2009）は他のエンハンスメント技術と比較した際の道徳的エンハンスメントの特異性は、その対象が自分ではなく他人にあることであるとして次のように述べる。

> モラル・エンハンスメントは、本質的に〈他者〉に向けられたものであり、しかもそれは〈私・我々〉のためであり、言い換えれば「顔の見えない多くの他者を、我々のためにエンハンスメントし、統御し、安心したい」という暗い欲望を露見させており、これらのゆえに拭いがたく不穏さに響くのである。この点で、本質的に〈自分自身〉を対象することによってその魅力を発揮する身体的・知的エンハンスメントとは全く逆の構造をしているのである。（立花 2009: p.100）

一見すると、この懸念点は道徳教育にも当てはまる部分がある。なぜなら、MBEも義務教育で行なわれる道徳教育も、いずれも外部からの道徳性の強制的な操作という面を共有しているように思われるからだ。たしかに、道徳教育が政府、宗教、または他の権威によって制御される場合、外部からの影響により自由な思考や自己決定が制約されたり、特定の規範や価値観を強制的に押しつけられたりすることが考えられる。しかし、この懸念は必ずしも当てはまらない。森岡（2013）は、道徳教育の性質を以下のようにまとめる。

> まず子どもたちは道徳的な価値や徳目を教師たちから教えられる。もちろん教師たちは子どもたちに教室での対話やフリーディスカッションに取り組む機会を与えるのだが、しかし学校における道徳教育の基本的なトーンは、教師から生徒への一方向的な考えの伝達に他ならない。しかしながらこのプロセスを経ることによって、「我々の道徳的統合性の核心部分 kernel of our moral integration」が子どもたちの精神の内部に形成されることが期待されるのである。そして子どもたちは徐々に自分で道徳的判断をすることができるようになり、彼らの内部に形成された彼ら自身の道徳的統合性の核心部分を参照しながら道徳的行為を行なうことができるようになる、と期待されるのである。言い換えれば、道徳教育は外部の道徳的価値を子どもたちの精神に強制的に注入するところから始まるのであるが、しかしそのプロセスが完了したあかつきには、道徳的統合性の核心部分が子どもたちの内部に形成され、彼らは彼ら自身の内部にある道徳規準に基づいて考えたり行動したりすることができるようになるのである。（森岡 2013: p.110）

19　高木（2020）によれば、彼らが主張するのは、同意の上でのエンハンスメントではない。犯罪者や今にも犯罪を犯しそうな人など道徳性を向上させることが急務であるような人に対して強制することでもない。あくまで、地球上のみなにエンハンスメントを強制的に施すことである。

すなわち、道徳教育とは、（少なくとも理想的には）子ども達が主体的に「変わる」ことを期待しているのであって、子ども達を強制的に「変える」ことを目指しているわけではないと言える。

他方、MAIE には以上の懸念は当てはまらない。なぜなら、竹下（2023）によれば、他の道徳的エンハンスメントに比べて MAIE の利点は、ユーザーに利用方法および利用自体を一任することにあり、つまりユーザーの自由や自律性を尊重する点にあるからである。さらに、もし仮に AI の意見を受け入れさせることが目的だとしても、この批判も MAIE にとっては問題とならない、というのも、もし AI の意見に強制的に従わせるならば、それはもはやユーザーの自由も自律性も尊重しないことになるからだ、と竹下は考える。しかし、ある組織のメンバーや特定のサービスの利用者が、規則や契約によって MAIE の使用やその判断への従属を求められる場合、AI の利用やその判断を受け入れるかどうかは使用者の自由に委ねられているから問題はない、という反論は成り立たないだろう。ただしこの問題は MAIE 自体に内在するものではなく、むしろその利用方法に起因するものであると言える[20]。

（ii）格差や搾取につながる

他にも社会的影響をめぐる議論として、一部の人間だけが道徳的エンハンスメントの利益を享受できるのは分配上の正義や公平性にかなっていないという問題がある。道徳的エンハンスメントには高度な技術やリソースが必要な場合が多く、それらが一部の人々にだけ利用可能である場合、技術の利用における不平等が生じる可能性がある。植原（2009）は、経済的な力や社会的地位のある人々がエンハンスメントを利用できる一方で、他の人々が利用できないという不均衡は、エンハンスメントを利用する集団とそうではない集団との間に格差をもたらすと考える。

この懸念点は MBE に該当する。堀内（2020 a）はこの論点を次のようにまとめる。

> エンハンスメントの利益は、経済的余裕がある富裕層ばかりが享受し、富裕者と低所得者の間の格差はいっそう拡大することになる。この傾向は、エンハンスメントの利用を私的自由に任せるほど顕著になるかもしれない。誰かがエンハンスメントを利用すれば、本来的には利用を望んでいない者でも、競争に負けないためには利用せざるを得なくなるからである（Brock 1998, 佐藤 2009: p.29）。その結果、経済的な理由や信条的な理由からエンハンスメントを利用しない未増強の者たち（the unenhanced）は「技術不足による障害者（techno-poor disabled）」と見なされるようになるかもしれない（Wolbring 2006=2008: p.129）。（堀内 2020 a: p.99）

他にも先の二つ目の論点、すなわちエンハンスメントをする（される）／しない（されない）の選択の自由・自律が侵害されるとする懸念点に関連して、森岡（2013）は、MBE の実施はそれを強制的に処置される人々と、逃げることができる人々の二種類のグループを生み出すと考える。すなわち、道徳性を生物学的に増強された人々は、他人からの攻撃や暴力や搾取に対して非常にひ弱な人間になり、強制力から逃げることができる権力と金を持ったどん欲な人々に搾取されてしまい、結果的に MBE は我々の社会を二つの層に分割するためのツールとして機能する、と考える[21]。

この問題は MBE には当てはまるかもしれないが、MAIE には当てはまらないと竹下（2023）は主張する。

20　このようなエンハンスメント技術の利用の仕方、ひいては悪用の可能性については後の註 22 でも触れる。

21　ここで道徳教育と MBE の差異を確認しておきたい。森岡（2013）は、道徳教育では教師から生徒への一方向的な考えの伝達がされるが、子どもたちは徐々に自分で道徳的判断を下すようになるのに対し、MBE では道徳性を生物学的に増強された人々は、一方的に権力と金を持った貪欲な人に搾取されてしまうという趣旨を述べる。しかし、道徳教育をこのように楽観的に割り切れるかは定かではない。この点を指摘頂いた査読者に感謝する。なおエンハンスメントの効果の可逆性については本稿でも後に触れる。

なぜなら、本稿４節で見た通り、MBE が②動機づけの改善を直接の目標とするのに対して、MAIE は③洞察の改善を直接の目標とするからだ。MAIE が個々人の洞察を（十分に）改善させる場合、もしそのような搾取的状況が道徳的に間違っているのであれば、洞察が改善された個々人はそのことを理解し、利己的な人々に利用されることを拒むはずである。よって、そのような搾取的状況は MAIE においては生じないと竹下は考える。しかし、技術的に AI にそれだけの洞察力の改善が見込めるかどうかは別の問題である（上述 5.1）[22]。

5.３　価値論（人間性）を巡る議論

（ｉ）真正さを減少させる

次に検討されるべきは、道徳的エンハンスメントを利用して築かれた人間性は真の人間性と言えるのか、真の価値を損なう恐れはないのかという懸念である。これは道徳に限らずエンハンスメント議論全般でもしばしば検討されるものであり、人間の価値や人間のあり方に対する認識、いわば人間観をめぐる議論である。

まずは、道徳的エンハンスメントによって得られたものは本物か偽物か、あるいは自然か不自然かを巡る議論を取り上げる。佐藤（2012）は、心的領域のエンハンスメントに関してしばしば指摘される批判の一つに、人工的手段によって得られた感情や性格は偽物であるというものがあるとする。佐藤は性格の操作に批判的な論者として知られる C・エリオット[23] を挙げ、薬物で作られた性格やそのような性格に基づく感情は人為的なもの（factitious）、偽物（inauthentic）であり、真正なものとはみなされないと主張されうることを指摘する。また立花(2009)も、エンハンスメントによって様々なことを達成できるようになったとしても、もはや「〈私〉が達成したのだ」といった感覚を抱けなくなるのではないか、という問題も指摘されるとする。加えて、森岡（2013）は道徳教育には「道徳的統合性」[24] があるが、一方でそれが欠如する道徳的エンハンスメントには真の道徳性を向上させることはできないという。この懸念点も道徳教育ではなく、MBE と MAIE に当てはまるものだろう。

（ii）努力・成長の機会を奪う

価値論（人間性）をめぐるまた別の議論として、人間の努力や成長という価値を損なうという懸念があげられる。もしエンハンスメントがあまりにも簡単に実施できるようになると、個人が成長や努力をする機会が失われる可能性がある。たとえ努力が必要な状況であっても、簡単にエンハンスメントを実施できれば、努力や成長に対するモチベーションが低下してしまうと指摘される。

この懸念は先に触れた、自律的な自己の発達を促進させるという三つ目の文脈での「自由・自律」の問題にも関わる。なぜなら、とりわけ道徳性については、個人が自らの道徳的価値観や行動規範を形成し、育てていくことが重要だと主張されるからだ。佐藤（2021）は、我々は自分自身の道徳性を責任を持って

22　MAIE そのものに対する技術的な限界や倫理的懸念とは別に、MAIE の悪用についても考慮すべき問題がある。例えば、社会的に優位な立場にいる者や、MAIE の開発者が自らは AI の使用を避けながら、他者にその利用を推奨し、その結果、MAIE を使用することで利己的な目的に対して脆弱になった人々を搾取する可能性が指摘できる。このようなエンハンスメント技術の悪用についての懸念を指摘頂いた査読者に感謝する。

23　ここで言及される論文は Elliot（2000）である。

24　森岡の言う「道徳的統合性」とは、少なくとも以下の３要素を有するとされる。（１）道徳的な人間の精神の内部には道徳的統合性の核心部分が存在しており、それは他の人間の欲望や意図によって決定的な形で支配されることがないということ、（２）道徳的判断と道徳的行為は外部からの影響によってではなく、行為者自身の支配下において行使されるということ、（３）人間の内部の道徳的統合性の核心部分には、歴史的な統合性（historical integrity）がなくてはならないということ、である。すなわち、人間の道徳性とは、主に個人の人間性がゆっくりと発達・成熟する過程や、人間が周囲の人びとと関わる相互的な経験から生じるものであるとされる。

自分自身で育てていかねばならないと主張し、道徳的エンハンスメント反対派の主張の要点を以下のようにまとめる。

> 本当に道徳的に善い人というのは、心から周囲の人のことを思い、自発的に彼らのために行動するような人である。失敗したら後悔し、反省し、次こそはと立ち向かう人である。努力せずに薬でそれらを手に入れたような気になるのは、いかにも浅はかであり、自らの向上のチャンスを放棄することである。薬で手に入れた動機から行為する人はもはや自発性を失って、自分をロボットに変えてしまっている。（佐藤 2021: p.276）

他にも植原（2009）は、この懸念点を念頭に置き、道徳的エンハンスメント反対派の主張を次のようにまとめる。

> 我々が生において目指すべきは全体としての自己完成であり、その意味では、さまざまな下位目的を果たすための手段もまた自己完成という全体の一部をなす重要な要素として配慮の対象とならねばならない。努力をともなう適切な手段や過程を通じて目的を果たすことが称賛の対象となるのは、ひとえにこうした価値の連関構造が存在するからだ、とも考えられる。だが、エンハンスメントは手段として不適切であり、またエンハンスメントによってもたらされる努力の不要性は、結局のところ人生の目的の達成を阻害することだろう。したがってエンハンスメントを許容することはできない、と主張される。（植原 2009: p.18）

この懸念点は特に MBE に当てはまるだろう。堀内（2020 b）は、このような批判の論拠には、バイオ保守派[25] の考えがあるとし、次のようにまとめる。

> バイオ保守派は、共助的な社会関係を下支えする道徳（連帯、謙虚、責任など）は、欠点や不完全さ、有限性と言った人間本性の脆弱性こそを可能性の条件とするものであり、それゆえ、そうした脆弱性を克服しようとするエンハンスメントの思想や実践は、共助的な社会関係を破壊することに繋がると見なしている。（堀内 2020 b: p.32）

つまり、欠点があるからこそ人は自ら努力・成長しようとするのだが、道徳的エンハンスメントはその伸びしろを埋めてしまう、ということである。

しかし、MAIE の場合には、この懸念点は部分的に回避されると考えられる。なぜなら、先に見た通り、MAIE の直接的な改善目標は、①情報・時間不足の改善と③洞察の改善であり、これらが改善した人々には、成長する機会がまだ残されている可能性があるからだ。竹下（2023）は MAIE の利用者（ユーザー）は AI を認識的根拠ではなく検討を始めるためのきっかけとして用いることができる、すなわちユーザーは AI の出力内容に直ちに従うのではなく、意見を聞くことにも考えるきっかけにも使うことができるとし、この懸念の払拭を試みる。つまり、MAIE は直ちに人々の道徳性を向上させるのではなく、人々に考えさせる機会や、ひいては自ら行動を起こさせる機会、つまり成長の機会を残すことができるということである。

25　バイオ保守派とテクノ進歩派の論争については、堀内（2020 a,b）を参照されたい。バイオ保守派の代表的な論者には島薗（2005）などが挙げられる。これに対してテクノ進歩派は、エンハンスメントは人間の脆弱性を克服しうるため、共助的な社会関係を補完するものであると論じる。

5.4　効果の不確実性を巡る議論

（ⅰ）即効性がない、限定的である

最後に、道徳的エンハンスメントの効果の不確実性を巡る議論を整理する。まずは、エンハンスメントの効果が即時には現れないのではないかという懸念がある。道徳的な変容や行動の変化には、新たな道徳的習慣を形成することが必要であり、それゆえ通常時間がかかると考えられている。道徳教育が時間をかけながら道徳性に関する知識を提供していくのはそのためである。高木（2020）によれば、サヴァレスキュらも、伝統的手段による道徳的エンハンスメントはまったく不要であり、無益であると述べているわけでもない。しかし、すでに見たように、我々は教育を通じて何が善かを理解しても、直ぐにその善をなすとは限らず、道徳教育だけでは人間の道徳性は科学技術の発展に追いつかないため、より迅速に道徳性を向上させる MBE の必要性が主張されたのであった。

しかし、結局のところ MBE もそこまで迅速に道徳性を向上できないという指摘がある。森下（2016）は、人々の道徳とは、人生の歩みの節々において問い直され、熟慮され、新たに形成されるものであるとした上で、MBE で用いられる薬物・外科処理・機械の生命レベルの介入では、その上位の心理レベルにおける直接的な介入にはなり得ないと述べる。それゆえ MBE ができることは、情動レベルの間接的な安定化であり、その上位にある道徳レベルの熟慮や自己変容は遠い目標（希望）になると述べる。

MAIE については、効果の即効性だけでなく、もはや効果そのものが見込めないとする批判がある。ユーザーは AI が出力する意見や助言を無視できるし、またそれを利用しなくなれば、そもそもの道徳的エンハンスメントの効果は見込めないという指摘である。竹下（2023）はこの批判に対して次の反論を試みる。まず、一部の意見は受け入れられる可能性があり、またそれによって AI の他の受け入れ難い意見に対してもユーザーを受容的にさせる可能性がある。ユーザーはその時は AI の意見を受け入れないかもしれないが、あとになって受け入れるかもしれない。そもそも MAIE の目的はユーザーに AI の意見を受け入れさせることだけではなく、例えば、AI を認識的根拠ではなく検討を始めるためのきっかけとして用いることができると主張する。以上の反論はいずれも、効果が見込めないという批判に対してはある程度の反論が成功しているように見える。しかし、皮肉にもこの反論は MAIE には即効性を期待できず、ユーザーに影響を与えられる効果は限定的であることを仄めかしている。

以上の議論を踏まえ、即効性をめぐる懸念点はどの介入方法に当てはまるかが次のように整理できる。まず、道徳教育では遅すぎる。MBE は動機づけを改善させるかもしれないが、それでもなお、心理レベルのさらに上位にある普遍的道徳性には間接的な介入に留まらざるを得ない。MAIE は時間不足をある程度は改善するかもしれないが、結局のところ道徳性を検討するきっかけづくりにしかなり得えず、即効性を期待できない。

（ⅱ）逆効果になる

次に懸念されるのは、道徳的エンハンスメントが、実行者あるいは設計者が予想した望ましい方向の変化をもたらすとは限らないという点である。つまり、エンハンスメントが予測された効果をもたらすだけではなく、予測されなかった副作用や意図しない結果をもたらす可能性があり、むしろ望ましい方向とは真逆の方向へ人々の道徳性を変化させるかもしれないという懸念がある。竹下（2023）は人々の道徳的信念を間違った方向へ変化させる可能性を次のように整理する。第一に人々が元々持っている間違った道徳的信念を強化すること（＝反道徳的強化）、第二に人々の持っている正しい道徳的信念を弱めること（＝道徳的弱体化）がありうる[26]。

26　竹下は第一の可能性については Klincewicz（2016）に依拠し「反道徳的強化」、第二の可能性については Giubilini・Savulescu（2018）、O' Neill et al.（2023）に依拠し「道徳的弱体化」と呼ぶ。

これらの懸念点はMBEに当てはまる。すでに見たように森岡（2013）は、もしある特定の人々にのみMBEがなされ、利他性や共感性が向上しより寛大になった場合、MBEを受けていない利己的な人々に利用されやすくなってしまう可能性を危惧し、格差や搾取へとつながると指摘する。

加えて、高木（2020）は道徳的エンハンスメントが人々の共感能力を促進した場合に、悪しき人物にも共感してしまう可能性を指摘し、次のように述べる。

> 「テロリズムは悪だ」という認知・信念をすでにもっている行為者は別であるが、そうでない行為者に道徳的エンハンスメントが施された場合、テロリストに共感する可能性がある。さらに、もし仮に規範的理由を認知していなくとも、テロリズムの犠牲者側に共感をもち、テロリズムを否定するに至ったとしよう。しかしそれでも、他の文脈においてはその共感はヒトラーや、腹いせで犯罪を犯す人物に向けて発揮される可能性が残るのである。それまではヒトラーに共感を感じなかった行為者が、道徳的エンハンスメントを施されたことで感じることになるならば、このような医学的介入は逆効果であるとさえ言いうる。（高木 2020: p.27）

この懸念点はMAIEにも寄せられる。しかし、竹下（2023）はそのような間違った方向に変化しないようにAIにユーザーを一定の方向に促す規範的な役割をもたせることで対処できる、例えばユーザーが差別主義的信念を持っているなら、その信念が間違っていることをAIが指摘するようにすればよい、と反論する。ここで竹下は、道徳的弱体化の可能性を引き受けつつ、反道徳的強化にならない程度に規範的役割をAIに担わせることで、懸念をある程度解消できるだろうと述べる。しかし、エンハンスメントが人々を望ましくない方向へ導くという懸念に対して、この反論は技術的に対処できると応答するものであり、AIの技術的可能性を高く見積もっていると思われる。

以上の2点の道徳的エンハンスメントの効果の不確実性をめぐる議論に関連して、高木（2020）は、「道徳的エンハンスメントを論じようとしても、我々の手元には完全な道徳理論はいまだない」（p.28）という点を強調し、次のように述べる。たとえある人が道徳的な精神をもっていたとしても、例えば功利主義、義務論、徳倫理のどれが真であるかに応じて、その精神に従ってなされるべき行為は異なるはずである。しかし、我々はどれが正しい道徳理論であるのかがまだわかっていない。同様に、誰に共感すべきか、何が共感に値するのか、もまだわからない。したがって、道徳的エンハンスメントを論じる際には、まずは我々の道徳哲学の進歩が先行する必要があると高木は考える。たしかに「何が道徳的であるのか」の基準が定まらない以上、我々は具体的には何を目指してエンハンスメントすべきか、そしてその効果があったとどうすれば判断できるかは定かではないだろう[27]。

27　この点については註3、註5も参照されたい。

5.5　懸念点のマッピング

以上、道徳的エンハンスメントをめぐって指摘される懸念点をみた。これらを踏まえ、道徳教育とMBEとMAIEの3種の方法に、どの懸念点が該当するかを表2に整理した。

表2：3種の方法に対する懸念点（筆者作成）

	道徳教育	MBE	MAIE
1．技術的可能性・安全性を巡る議論	—	✓	（✓）
2.社会的影響を巡る議論			
（ⅰ）自由・自律を損なう	（✓）	✓	—
（ⅱ）格差や搾取につながる	—	✓	—
3．価値論（人間性）を巡る議論			
（ⅰ）真正さを減少させる	—	✓	✓
（ⅱ）努力・成長の機会を奪う	—	✓	（✓）
4．効果の不確実性を巡る議論			
（ⅰ）即効性がない、限定的である	✓	（✓）	✓
（ⅱ）逆効果になる	—	✓	（✓）

※該当懸念点が当てはまる場合に「✓」、当てはまるものの反論や擁護により懸念を部分的に払拭できていると考えられる場合には「（✓）」、懸念が当てはまらない場合に「—」とした。

表からは、これまでにMBEで言及されてきた懸念点をMAIEでは軽減または、部分的に払拭できている所が複数あることが見て取れる。しかし、道徳教育に比べれば、両者には未だ多くの懸念点があることがわかる。このことから、道徳的エンハンスメントはなお根本的に、倫理的に困難な問題を抱えていると思われる。例えば、前節でみた高木（2020）が指摘する点などは、MBEにとってもMAIEにとっても検討すべき課題であろう。

ただし、本稿ではいずれの方法についても、悪用の可能性に関しては主に扱っていない。というのも、エンハンスメント技術そのものが抱える倫理的懸念と、その利用の仕方に起因する問題は区別されるべきだからである。例えば、いずれの方法でも悪用される場合には、社会的影響に関する懸念として、（ⅰ）「自由や自律を損なう」、（ⅱ）「格差や搾取につながる」という問題が生じる可能性がある[28]。

6　おわりに

以上、日本における道徳的エンハンスメント議論の論点を整理することで、その全体像を描き出すことを試みた。本稿ではまず、道徳教育とMBE、MAIEを区別した上で、それぞれの改善目標と、指摘される倫理的懸念点を整理した。その結果、MBEは動機づけの改善に、MAIEは洞察の改善にそれぞれ重点があることが確認できたが、同時に道徳教育を含めいずれの方法も直接的には道徳的行為を改善するものではないことも確認された。懸念点については、諸々の論点を包括的に見ると、技術的に可能である場合に限りMAIEはMBEについて言及されてきた懸念点のいくつかを回避しうることが確認された。しかし

28　もちろん、悪用に対しては対策が講じられる可能性があり、悪用の懸念を上回る利点が存在するかもしれない。悪用の可能性が完全に排除できないという理由だけで、新しい技術に対して一律に反対するというのは極端な見解であるとも思われる。この点を指摘いただいた査読者に感謝する。

ながら両者とも、道徳教育と比較すれば多くの倫理的懸念があることも確認できた。本稿全体を通して、日本においても道徳的エンハンスメントをめぐり、複雑で多岐にわたる論点に関して議論が行われていることが改めて確認された。

本研究を通して今後の課題として見えてきたものが２つある。まずは既に５節で指摘したように、道徳的エンハンスメントの懸念点として言及される「自由・自律」の侵害について、その解釈は曖昧であることである。本稿では「自分・自律」に対する異なる意味解釈を３つに整理した。改めて確認すれば（１）エンハンスメント実施後の当事者の真の選択を妨げない、という意味の「自由・自律」、（２）エンハンスメントをする（される）／しない（されない）の選択の自由、つまり外部からの支配や社会的圧力を受けない、という意味の「自由・自律」、（３）自律的な自己の発達を促進させるという文脈での「自由・自律」である。本稿では、先の２つを社会的影響を巡る議論に、3つ目を価値論（人間性）を巡る議論に分類した。しかし、「自由・自律」という概念は未だ論争的である。今後はその概念や背景文脈を一層明確にすることで、道徳的エンハンスメントに関する懸念点を巡る議論をより精緻に進めることができると考えられる。

２つ目は、日本の文献においては MBE について、その効果の可逆性に関する議論があまり見受けられなかった点である。MBE に関するほとんどの論考では、一度施された生物学的な介入により後戻りできない状態になることが想定されていると考えられる。しかし、生命・医療倫理学の領域では、処置の可逆性についての議論がしばしば取り上げられる。例えば、グロース（2011）は、「脳神経認知エンハンスメント（neurocognitive enhancement）」に関して、脳の機能を改善するための薬理的エンハンスメントと非薬理的エンハンスメントの間には大きな相違があると主張し、その理由の一つに可逆性について言及する。彼によれば、個々人は単に薬剤の摂取を中止できるのであるから、薬理的エンハンスメントは基本的に可逆的である。しかし、非薬理的エンハンスメントの処置では原状回復は不可能であり、場合によっては永続的に影響を与えることになる。すなわち、グロースは薬理的エンハンスメントには可逆性があるが、非薬理的エンハンスメントは可逆性がないとして、それらを拙速に同等視してしまうことは不適切であると主張する。もちろん、もし仮に現時点での技術が MBE に不可逆的な生物学的アプローチを要求する場合（MBE 実施後に永続的にその影響を与える場合）、倫理的価値観の変更が困難になり、個人の自己決定権が制限される可能性などは真っ先に懸念されるべき点であろう。しかし、可逆的なアプローチであれば、現在 MBE に寄せられる懸念点をいくつかを回避することが可能であると考えられる。例えば、エンハンスメント実施後でも、戻りたい時に直ぐに元の自分に戻ることができるようになれば、真正さや自由・自律を損なうという懸念点をある程度払拭できると思われる。この点については、さらなる検討が必要であろう。

道徳的エンハンスメントの倫理的是非をめぐっては、いま挙げた点を含め、今後もさまざまな角度から議論が展開されていくと考えられる。本稿は道徳的エンハンスメントを巡る議論への理解を深めるためのささやかな試みにすぎないが、その理解が、我々が伝統的な価値観を尊重しながらも、新たな技術がもたらす変革に対応するための、ひいてはより良い未来に向けて前進するための一助となることを願う。

謝　辞

本稿は、筆者が学習院大学大学院政治学研究科に在籍していた際に執筆し、2025 年 1 月に提出・受理された修士論文の一部を改稿したものである。研究を進めるにあたり、多大なご助言とご指導を賜りました玉手慎太郎教授、論文審査の副査を務めてくださいました中田喜万教授、若松良樹教授、ならびに在籍中にお世話になった諸先生方、日々の学業・研究生活を支えてくださった同研究科の皆様に、この場を借りて心より感謝申し上げる。

参考文献

Brock, D. W. (1998). “Enhancement of human function: Some distinctions for policymakers.” Parens, E. (ed.) *ENHANCING HUMAN TRAITS: Ethical and Social Implications*. Georgrtown University Press, pp.48–69.

DeGrazia, D. (2014). “Moral enhancement, freedom, and what we (should) value in moral behaviour.” *Journal of medical ethics,* 40(6), pp.361-368.

Elliott, C. (2000). “Pursued by happiness and beaten senseless: Prozac and the American dream.” *Hastings Center Report*, 30(2), pp.7-12.

Giubilini, A., & Savulescu, J. (2018). “The artificial moral advisor. The “ideal observer” meets artificial intelligence.” *Philosophy & technology*, 31, pp.169-188.

Harris, J. (2011). “Moral Enhancement and Freedom.” *Bioethics*, 25-2, pp.102-111.

Klincewicz, M. (2016). “Artificial intelligence as a means to moral enhancement.” *Studies in Logic, Grammar and Rhetoric*, 48(1), pp.171-187.

Lara, F. (2021). “Why a Virtual Assistant for Moral Enhancement When We Could have a Socrates?.” *Science and Engineering Ethics*, 27(4), pp.1-27.

O’ Neill, E., Michal K., & Michiel K. (2023). “Ethical Issues with Artificial Ethics Assistants.” In: Véliz, C. (ed.) *Oxford Handbook of Digital Ethics*, Oxford Handbooks (2023; online edn, Oxford Academic, 10 Nov. 2021), https://doi.org/10.1093/oxfordhb/9780198857815.013.17, accessed 19 Dec. 2023.

Persson, I., & Savulescu, J. (2008). “The perils of cognitive enhancement and the urgent imperative to enhance the moral character of humanity.” *Journal of applied philosophy*, 25(3), pp.162-177.

Persson, I., & Savulescu, J. (2013). “Getting moral enhancement right: the desirability of moral bioenhancement.” *Bioethics*, 27(3), pp.124-131.

Savulescu, J., Douglas, T., & Persson, I. (2014). “Primary Topic Article: Autonomy and the Ethics of Biological Behaviour Modification.” In: Akabayashi, A. (ed.) *The Future of Bioethics: International Dialogues* (Oxford, 2014; online edn, Oxford Academic, 23 Jan. 2014), https://doi.org/10.1093/acprof:oso/9780199682676.003.0011, accessed 5 Feb. 2024.

Savulescu, J., & Maslen, H. (2015). “Moral enhancement and artificial intelligence: moral AI?.” In: Romportl, J., Zackova, E., Kelemen, J. (eds.) *Beyond artificial intelligence: The disappearing human-machine divide*. Topics in Intelligent Engineering and Informatics, vol 9. Springer, Cham., pp.79-95.

Singer, P., & Sagan, A. (2012). “Are We Ready for a ‘Morality Pill’ ?” *The New York Times*, January 28.

Wolbring, G. (2006). “The unenhanced underclass.” In: Miller, P., & Wilsdon, J. (eds.) *Better Humans?: The politics of human enhancement and life extension*. Demos Institute. = 上田昌文・渡部麻衣子 (2008). 『エンハンスメント論争―身体・精神の増強と先端科学技術』社会評論社 : pp.123–130.

足立智孝 . (2012)「エンハンスメント問題の人間学的一考察」『モラロジー研究 : 倫理道徳研究フォーラム』69, pp.109-126.

飯塚理恵 . (2021)「エンハンスメントとしての美の実践」『現代思想』49（13）, pp.200-208

石田柊 . (2020)「神経科学・脳科学をめぐる ELSI 的視点－潜在的バイアスにかかわる道徳的諸問題に注目して」『ELSI NOTE』6, online.

伊吹友秀 , & 児玉聡 . (2007)「エンハンスメント概念の分析とその含意（第 18 回日本生命倫理学会年次大会シンポジウム）」『生命倫理』17（1）, pp.47-55.

植原亮 . (2009)「エンハンスメントの哲学と倫理・概念的な見取り図の試案」『エンハンスメント・社会・

人間性』8, pp.7-23.
ドミニク・グロース . (2011)「恵みか災いか？―「脳工学」による脳神経認知のエンハンスメント―」『医療・生命と倫理・社会』10, pp.109-129.
斎藤里美 . (2017)「人工知能とエンハンスメントの時代における「学ぶ意味」と「学力」―「人工知能と人間社会に関する懇談会」諸資料の批判的検討を通して―」『教育学研究』84（4）, pp.410-420.
佐藤岳詩 . (2009)「功利主義的観点から見た認知的エンハンスメント」『医学哲学 医学倫理』27, pp.23-32.
佐藤岳詩 . (2010)「気分明朗剤と快楽主義」『応用倫理』3, pp.45-62.
佐藤岳詩 . (2012)「性格のエンハンスメントの倫理的問題点について」『医学哲学 医学倫理』30, pp.20-29.
佐藤岳詩 . (2021)『心とからだの倫理学：エンハンスメントから考える』ちくまプリマー新書
島薗進 . (2005)「増進的介入と生命の価値：気分操作を例として」『生命倫理』15（1）, pp.19-27.
髙木裕貴 . (2020)「道徳的エンハンスメントの道徳的問題―ピアソン & サバレスキュの立論に即して―」『医学哲学 医学倫理』38, pp.20-30.
竹下昌志 . (2023)「道徳的 AI エンハンスメントの擁護」『応用倫理』14, pp.3-20.
立花幸司 . (2009)「モラル・エンハンスメントはなぜ不穏に響くのか」『エンハンスメント・社会・人間性』8, pp.83-102.
堀内進之介 . (2020 a)「技術的進歩と民主的社会変革の融合の可能性―テクノ進歩派によるエンハンスメントの正当化の試み―」『科学基礎論研究』47（2）, pp.97-107.
堀内進之介 . (2020 b)「道徳的エンハンスメントによる共助的な社会関係の底上げの可能性―テクノ進歩派の理論的根拠に関する検討―」『年報 科学・技術・社会』29, pp.31-49.
森岡正博 . (2013)「道徳性の生物学的エンハンスメントはなぜ受け容れがたいのか？」『現代生命哲学研究』2, pp.102-113.
森下直貴 . (2016)「モラル・バイオエンハンスメント批判―「モラル向上のために脳に介入すること」をめぐって」森下直貴（編）『生命と科学技術の倫理学：デジタル時代の身体・脳・心・社会』丸善出版 , pp.90-111.
山崎雄介 . (2023)「道徳教育周辺の教育内容研究の欠陥とその克服―校則問題を手がかりに―」『群馬大学教育実践研究』40, pp.239-248.
脇坂真弥 . (2020)「生命操作に抗して何が言えるか―サンデルの「生の被贈与性」と障害の問題を手掛りにして―」『哲学論集』66, pp.1-16.

研究ノート

反事実的比較法を用いた危害の説明をめぐる議論の現状

吉澤ひふみ（北海道大学）

要旨

本稿は、反事実的比較法（Counterfactual Comparative Account/CCA）を用いた危害の説明に関する議論を概観し、特に Klocksiem による擁護と Carlson による批判を中心に整理する。CCA は、ある出来事が生じなかった場合に対象者がよりよい状態にあったかどうかを基準として危害を判断する立場である。Klocksiem は、CCA が文脈に依存し、内在的・外在的な危害、一時的･全体的な危害の区別を明確にする点を強調する。一方､Carlson は､CCA が「行為の理由」や「選好の理由」に適合しないことを指摘し、批判する。これに対し、Klocksiem は、行為とアウトカムを区別することで批判を回避しようとする。最後に、医療倫理・医療哲学との関連として、トータルペインの概念や医療事故の扱いを CCA の枠組みで考察する方針を示唆する。本稿は、危害の概念に関する哲学的議論を整理し、今後の研究の課題として CCA の限界と応用可能性を指摘する。

Abstract

This paper provides an overview of the debate surrounding the Counterfactual Comparative Account of harm (CCA), focusing on Klocksiem's defense and Carlson's critique. CCA evaluates harm by employing counterfactual conditionals and possible world comparisons, assessing whether an individual is better off in the closest possible world where the harmful event did not occur than in the actual world. Klocksiem emphasizes that CCA is context-dependent and clarifies the distinctions between intrinsic and extrinsic harm, as well as between pro tanto and all-things-considered harm. In contrast, Carlson argues that CCA fails to align with the principles of "reasons for action" and "reasons for preference". Klocksiem responds by distinguishing between actions and outcomes to counter these objections. Finally, this paper suggests exploring the implications of CCA in medical ethics and philosophy of medicine, particularly in relation to the concept of total pain and the classification of medical errors. By critically examining CCA, this study identifies its theoretical limitations and potential applications, contributing to the broader discourse on the nature of harm.

はじめに

危害をどのように捉えるかは論者によって様々である。例えば、Feinberg は、危害を「ある利益を阻止することや、後退させること、打ち負かすこと」として主張した[1]。この定義からすれば、例えば、「十分な睡眠」が利益なのであれば、「十分な睡眠」を妨げるもの―例えば、騒音や身体的な痛み―は危害に分類されることになるだろう。一見するとこの定義はもっともらしいものの、このように利益を第一義的なものとして捉え、そこから危害を定義する仕方について、批判がない訳ではない。Schöne-Seifert によれば、贅沢品に囲まれて生活することのように、十分な福利を得られている状態について、それを阻止したり後退させたりすることを通常は危害とは呼ばないという。Schöne-Seifert は続けて、危害は福利に決定的な影響を与えるものであるため、危害と利益には非対称性があること、ならびに、危害と利益は、一方が減ずればもう一方が増加するようなトレードオフの関係にはないと、ごく簡単に述べている[2]。

このように、Feinberg の危害の定義には批判はあるものの、この定義は、医療倫理学に大きな影響を与えた *Principles of Biomedical Ethics*（邦訳『生命医学倫理』）[3] において参照され、2019 年に出版された第八版においても参照され続けている。その中で、著者である Beauchamp & Childress は、第八版の脚注において「危害の定義には哲学的論争がある」と置き、危害に関するいくつかの論文を紹介している [4]。そのうちの一つが反事実的比較法による危害の考慮（Couterfactual Comparative Account of Harm、以下 CCA）である[5]。

本稿は、CCA の紹介を主な目的とするが、論争の一部を紹介するにとどまり、CCA に対して新規性のある主張を展開するには至っていない。しかしながら、CCA の議論は、危害の理論と実践に対する含意の両方において示唆に富むものと思われたため、完全なサーヴェイとは言えないものの、ここで説明する。また、医療倫理･医療哲学の議論に対し CCA を適用した研究は見当たらないため、本稿の最後に医療倫理･医療哲学との関連を簡単に述べる。

本稿では以下の構成をとる。第一章では CCA の紹介を行う。第二章では Klocksiem の議論を参考に、CCA の代表的な問題点と Klocksiem による応答を確認する。第三章では、CCA を擁護する立場にある Klocksiem と CCA に批判を向ける Carlson との応酬を提示する。最後に第四章で、医療倫理・医療哲学の議論に対し CCA を適用した際の課題を述べる。

1　Feinberg, 1987, p.33
2　Schöne-Seifert, 2004, p.1382
3　初版は 1979 年である。なお、邦訳としては第五版を翻訳したものが出版されている（ビーチャムら , 2009, 立木教夫・足立智監訳）。
4　Beauchamp & Childress, 2019, p. 203
5　Beauchamp & Childress が第八版の脚注において紹介した論文の内の一本は Norcross によるものであり、このあと紹介することになる Carlson は、Norcross の議論を CCA の一種とみなしている。

1. CCA とは

CCA を擁護する立場にある Klocksiem は、「危害についての自然で直観的な回答は、危害を受けたときにそうでなかった場合よりも悪くなることだ」と述べている。CCA は、反事実的条件文とその意味論を下地にもち、可能世界同士の比較を基本とする考え方である。代表的な二つの形式を挙げる。

> 出来事 e が対象者 S を害するのは、もし e が生じなかったならば S がよりよくなっていただろうときであり、かつそのときに限る。[6]

> ある出来事 e が対象者 S にとって危害となるのは、e が生じない最も近い可能世界における S が、e に関連した世界における S よりもよりよいときであり、かつそのときに限る。

反事実的条件法は、「もし…だったならば、…だっただろう（if it were the case that … then it would be the case that…）」の形式をもつ仮定法の文である[7]。Carlson らの形式にはこの仮定法が表れている。反事実的条件法が真となるための分析について、飯田は以下のように述べている[8]。

> 前件が真であり、それ以外の点では現実世界との違いが最小であるような可能世界を考える。条件文が真であるのは、後件がこの可能世界で真であるときであり、条件文が偽であるのは、後件がこの可能世界で偽であるときである。

このとき、可能世界の比較には「類似性」の概念が導入される。例えば、S が風邪をひいたとしよう。与えられる条件文は「S は風邪をひいていなければ、よりよかっただろう」となる。CCA によれば、この条件文が真であれば風邪をひいたことは危害であり、偽であれば風邪をひいたことは危害ではない。風邪をひいた現実世界を w とし、風邪をひかない世界のうち、現実世界に最も類似している（すなわち、最も近い）可能世界を w 1 とする。このとき、w 1 において、S がよりよいかどうかを考える。S がよりよ

6　Carlson, Johansson and Risberg, 2021, pp.421.　ところで、Carlson（2018）は、CCA の定式に含まれる比較の仕方について、二つの解釈が可能であると述べている。二つの解釈の分水嶺は、出来事 e が生じた世界と比較される出来事 e が生じない世界は、どの世界に最も近い世界なのかについてである。一方の解釈は、出来事 e が生じない世界が現実世界に最も近い世界となるものであり、もう一方の解釈は、出来事 e が生じない世界が出来事 e が生じる世界に最も近い世界となるものである。出来事 e が生じる世界が現実世界だった場合には両者の解釈は一致するが、将来の出来事について考慮している状況などでは、両者の違いは明瞭となる。Carlson は後者の解釈、すなわち、「ある可能的な出来事 e は危害（利益）だろうというのは、最も近い e 世界（We）におけるその人が、We に最も近い非 e 世界におけるその人よりも、より悪い（よりよい）ときであり、かつそのときに限る」を採用する。（Carlson, 2018, p. 796）

7　因果論において、「もし c が生じたならば、e が生じただろう」と「もし c が生じなかったならば、e は生じなかっただろう」という条件文が真となるときに、e は c に反事実的に依存しているとされる。（Carroll & Markosian, 2010, p. 26；Kutach, 2014, pp.65-66, 邦訳相松 , 2023, p.75）反事実的条件文が真となるための分析は、現実世界において、 c が生じて、かつ、e も生じることが確定しているため、結局のところ、可能世界において「もし c が生じなかったならば、e は生じなかっただろう」が真となる場合に簡略化できる。（Menzies and Beebee, 2024, Section 1.1）。このような、反事実的条件文とその意味論を基にした因果論の議論が CCA とどこまで類比関係にあるのかは未検討である。

8　飯田 , 2024, p.155 を参照。反事実的条件文が真となるより詳細な定式化は、「『もし A であるならば、C であるだろう』が現実世界において真であるのは、（ i ）A が成立する可能世界がないか、（ ii ）C が成立し A が成立するいくつかの可能世界の方が、C が成立せず A が成立するあらゆる可能世界よりも現実世界により近いときであり、かつそのときに限る」（Menzies and Beebee, 2024, Section 1.1）である。このように、条件文が真になるのは、現実世界に最も近い可能世界が一つに決まる場合とは限らない。本稿では、議論の簡略のために現実世界に最も近い世界が一つに決まる場合のみを想定する。また、例えば、風邪をひいた対象者がいて、風邪をひいた出来事が危害なのか否かを考慮しているとき、参照される近しい可能世界には、元の世界で風邪をひいた人に対応する対象者が存在するだろう。このとき、参照されない可能世界にそのような対応関係にある対象者がいなくてもよい。なお、Klocksiem の議論と反事実的条件文が依拠する意味論とがどこまで整合的であるのかについては未検討であり、今後の課題とする。

いといえるのであれば条件文が真となり、風邪をひいたことは危害となる。他方で、Sがよりよいと言えないのであれば条件文は偽となり、風邪をひいたことは危害とならない。

Klocksiem によればCCAは「文脈」に重く依存する[9]。Klocksiem は文脈とは何かを明示的に述べていないのだが、彼の主張に基づくと、「CCA が文脈に依存する」とは、実際に生じた出来事や行為の詳細が明らかになることで、現実世界に最も類似している可能世界が決まり、これらの過程を通してCCA を通じた帰結が変わることを意味する。

さらに、「文脈」が与えられることで、内在的な危害（intrinsic harm）と外在的な危害（extrinsic harm）の区別、ならびに、一時的な危害（prima facie harm）とすべてを考慮した危害（harm all-things-considered）の区別が明らかになる。

内在的な危害と外在的な危害の区別

Bradley によれば、内在的危害とはそれ自体が危害であることであり（例として痛みや不快感が挙がる）、外在的危害とはそれの影響のために危害が生じることである（例として喫煙が挙がる）[10]。Bradley は価値論における内在的と外在的の区別を参照し、この区別を危害の分析に類比的に導入した。すなわち、内在的に価値があるとはそれ自体で価値があることであり、外在的に価値があるとはそれの影響のために価値が生じることである。

Klocksiem は Bradley の議論を参考に、内在的な危害と外在的な危害の区別を以下のように示した[11]。

> ある出来事 e が S にとって内在的に危害であるのは、e が次の性質を満たす最小限の事態である場合である。すなわち、e が生じなかったならばSがよりよくなっていただろう場合である。そしてこの場合には、e と他のあらゆる出来事との繋がりは〔S にとって危害である〕理由にならない。

> ある出来事 e1 が S にとって外在的に危害であるのは、e1 が他の出来事や e2 のような事態と繋がりがあるゆえに、e1 が生じなかったならばSがよりよくなっていただろう場合である。

Klocksiem は外在的な危害の例として「癌になること」を挙げている[12]。もし治療をしなければ、腫瘍が患者を害するだろう。しかし、この危害は、腫瘍が引き起こす疼痛や精神的な苦痛、早期の死などによるメカニズムが存在することから生じている。つまり、癌になることは、他の内在的な危害と接続していることのみによって外在的な危害となる。したがって、腫瘍による危害は外在的な危害だとされる。他方で、ある出来事がそれ自体で痛みの経験をもたらしているのであれば、それは内在的な危害とみなされる。

一時的な危害とすべてを考慮した危害

時に、より大きな利益のために危害を与えることがある。このような事態は、それ単体では悲惨なもの

9　Klocksiem 以外には、Norcross も文脈相対的な危害を提唱している。Klocksiem の対抗論者である Carlson は Norcross の主張を CCA とみなしている。Norcross の見解は CCA に類似しており、CCA の派生形とみなすこともできるだろう。肝心の Norcross の議論は、可能世界は一つに決まらないことがあるため、可能世界の代わりに「適切な代替案」を導入するというものである。だが、適切な代替案は文脈によって決定されると述べられてはいるものの、それがどのように決定されるのかについて Norcross による詳細な説明はなく、適切な代替案も一つに決まる保証はない。したがって、CCA に対する可能世界が一つに決まらないという批判は、その批判を躱すことを目論んだはずの Norcross の見解にもそのまま当てはまる可能性がある。

10　Bradley, 2012, p.392.

11　Klocksiem, 2012, p. 289 下線部ならびに亀甲括弧内引用者

12　Ibid

であるとされるにも関わらず、利益であると見なされる。Klocksiem によれば、一時的な危害とすべてを考慮した危害の区別を用いることで、この見方を理解できるという。

Klocksiem はこの区別の説明のために、化学療法の例を提示する[13]。化学療法はすべてを考慮した上では患者の利益になるが、それ単体で考えれば、深刻な不快感を生じさせるため危害とみなされる。すなわち、原因と結果の巨大なネットワークから、化学療法それ単体を別個で考慮することで、化学療法は一時的な危害となる。他方で、すべてを考慮の上では、化学療法はより遠くの時点で患者の回復と接続しているがゆえに、外在的な利益となる。このことから、一時的な危害であることと、外在的な利益であることは両立するとされる。

2. CCA の問題点とその対処

本章では、CCA の代表的な問題点と Klocksiem による応答を確認する。

CCA にはいくつかの問題が指摘されている。Klocksiem によれば、CCA への問題点は以下の主に５つである[14]。

1. 一般的に危害とされることが、純粋な利益とされてしまうこと
2. 非同一性問題に対処できないこと
3. 危害とみなされる出来事が過剰に増加してしまうこと
4. 単なる与益の失敗と危害とを区別することが困難であること
5. 先取り問題への対処が困難であること

以下では、それぞれの問題点の例を順に示し、それらの問題点に対する Klocksiem の応答を説明する。

2. 1. 一般的に危害とされることが、純粋な利益とされてしまうこと

【救助事例】

Betty が川で溺れている。Veronica は Betty を救助する時に、Betty の腕を折ってしまう。このとき、Veronica は Betty を害したように見える。だが、CCA によれば、腕が折れていない Betty がいる最も近い可能世界は彼女が溺れる世界なので、現実世界の Betty はよりよいことになる。したがって、CCA は Betty の腕の骨折を危害ではないと判断してしまう。

Klocksiem の応答

Veronica が Betty を救助する際に、骨折を伴わないことが可能だったのであれば、骨折は救助の不必要な部分であるため、すべてを考慮した危害にもなる。Veronica からすれば、早急な対応が必要だったので Betty を丁寧に引き上げる余裕はなかったかもしれない。他方の Betty も、Veronica 自身が川に飛び込んで Betty の腕を折ることなく救助してくれてもよかったのに、と主張するかもしれない。要するに、危害

13　Klocksiem, 2012, p. 290

14　以降の記述は、Klocksiem（2012）に基づく。なお Klocksiem は、この５つ以外にも、６つ目の問題点として「条件文の誤謬」を挙げている。条件文の誤謬とは、主張 p の真偽の判定に条件文 ϕ「もし状態 a が生じたならば、状態 b が生じただろう」という文を用いる際に、p の真理値が実際に a が生じるか否かに依存し、単なる分析や定義の価値の上には依存しないことを指す。Klocksiem は、CCA の批判として条件文の誤謬は棄却できると述べ、その根拠として本稿で後述することになる内在的危害と外在的危害の区別を導入する。本稿は CCA の考え方を紹介することが主な目的のため、条件文の誤謬の説明は割愛する。

か否かの判断は、Veronica と Betty がそれぞれどのような文脈を提供するかにかかっている。

2. 2. 非同一性問題に対処できないこと

【家族計画事例】

Archie と Betty の夫婦がおり、今子どもを作るか、一年後に子どもを作るか考えている。今子どもを作ると、Betty の病によってその子どもは難聴になる（この子を Archie Jr. とする）。難聴であることは Archie Jr. の人生にネガティブな影響を与えるが、それでも悪い人生ではなく、全体として良い人生になるだろう。一年後に子どもを作ると、その子は難聴ではない（Betty Jr. とする）。直観的には、今子どもを作る選択は、Archie Jr. に危害を与えているようにみえるので、子どもを作るのは一年待った方がよいと判断されるだろう。

だが、CCA によれば、Archie Jr. が難聴である世界と比較される世界は、Archie Jr. が存在しない世界となる。Archie Jr. が存在し、かつ、難聴でない可能世界はないので、Archie Jr. の難聴はそうでなかった世界よりも悪いとはならない。したがって、難聴は危害ではないと判断されてしまう。

Klocksiem の応答

非同一性問題は哲学的に非常に難しい問題であり、CCA を用いても非同一性問題のすべてを解決することはできない。特に、一年後と比較して「今」子どもを設けることの悪さを説明することはできない。しかし、CCA を用いれば、Archie Jr. が害されたという直観を説明することはできる。

Klocksiem によれば、Betty の病が必然的であることを想定する必要はない。そのため、Betty が病気ではない可能世界は、Betty が病気ではなく、かつ、難聴ではない Archie Jr. が存在する世界となり、それゆえに可能世界における Archie Jr. がよりよい状態となる。すなわち、Archie Jr. が難聴である世界と比較される世界は Archie Jr. が存在しない世界であるとは一概には言えない。さらに、最も近い可能世界が、Betty が病気ではなく、難聴ではない Archie Jr. が存在する世界なのであれば、Archie Jr. は Betty の病によって害されたことが含意される。

Klocksiem は上記主張を提示し、子どもをいつもうけるべきかの判断について一年待つべきであると結論する。Archie と Betty は、彼らが持つ決定力（power）において、Archie Jr. を害さないとは言えない。端的に言えば、Betty が病気である状況において、その状況と Archie Jr. の耳が聞こえる状態は遮断されているため、Archie と Betty は Archie Jr. を相対的に益する方法を有さないのである。

2. 3. 危害とみなされる出来事が過剰に増加してしまうこと

【銃殺事例】

Reggie は Archie と Betty が付き合っている事実に嫉妬して、Archie を背後から銃で撃つ。このとき CCA を用いると、多重な因果の連鎖が生じるため、その連鎖のうちのそれぞれの要素が彼を害していることになる。多重な因果の連鎖とは、「Archie が麻痺にならなければ、彼はよりよかっただろう」、「Reggie が引き金を引かなかったならば、Archie はよりよかっただろう」、「撃針（the firing pin）が弾薬筒の雷管（the percussion cap in the cartridge）に当たらなければ、Archie はよりよかっただろう」、「弾が発砲されなければ、Archie はよりよかっただろう」、「弾丸がアーチーの体に入らなければ、Archie はよりよかっただろう」という反事実的条件文の繋がりのことを指す。しかし、通常は一つの出来事が Archie を害したと見なされるはずだろう。

Klocksiem の応答

Klocksiem は反事実的条件文の連鎖と、内在的な危害と外在的な危害の区別に訴えて、上記問題に対処しようとする。彼によれば、「Archie が麻痺になること」、「Reggie が引き金を引くこと」、「撃針が弾薬筒の雷管に当たること」、「弾が発砲されること」、「弾丸が Archie の体内に入ること」というそれぞれの出来事は、単一の因果的連鎖における接近した出来事同士の一つの連鎖である[15]。引き金を引いたことが銃口から弾を発射し、その弾が Archie の身体を貫いたという事実と事実の独立性において、Reggie が引き金を引いたことは Archie を害していないと説明される。これらは単なる外在的な危害にすぎない。

CCA に基づくと、それぞれの出来事は分離された危害にはならない。それぞれの個別的な危害は、因果的に、かつ、反事実的に連結しており、それは直観的な中心的危害（ここでは麻痺）に向かって連結している。つまり、Archie が麻痺を生じない最も近い可能世界は、Reggie が引き金を引かず、撃針が弾薬筒の雷管に当たらず、弾が発砲されず…というような、先の出来事の連鎖が生じない世界となる。そして、Archie の麻痺も、痛みや不快感を生じさせ、肯定的な経験を邪魔したり妨げたりすることにおいて、外在的な危害である。

2. 4. 単なる与益の失敗と危害とを区別することが困難であること

【帽子事例】

Archie が Jughead に新しい帽子をあげようと考えて、やっぱりやめたとする。また、そのことを Jughead が知りえなかったとする。このとき、Jughead の福利に変化はないものの、そうではなかった（帽子を貰った）世界の Jughead の方がよりよいため、Archie が Jughead に新しい帽子を与えないことは危害だとされてしまう。

Klocksiem の応答

直観的には、行為や出来事が対象者 S にとって与益の失敗であることは、時を越えて S の福利が増加しないことであり、その際の可能世界における S の福利のレベルは現実のレベルよりも大きい。他方で、危害とは、S の現実の福利が、関連した可能世界における状況の福利よりも、より悪い状況のことである。この区別を用いれば、Archie の帽子事例は与益の失敗だとされる。

しかしながら、Klocksiem は、この区別では説明しきれない事例を提示する。

【看護の放棄事例】

Archie が病気である Betty を看護する約束を破り、Archie は代わりに Veronica と一日を過ごす。

事例の構造としては、Archie が Betty に対して何もしないという不作為に関連しており、加えて Betty の福利は変化していないため、帽子事例と同じに見える。すると、この事例は非危害的な（nonharmful）与益の失敗として分類されることになる。だが、Klocksiem は、明言はしないものの、この事例をみた多くの者が Archie は Betty を害したと評価するだろうという直観に基づいて議論を進める。

ある出来事が危害とされるのか与益の失敗とされるのかは、関連する行為者がその出来事が生じない最

15　この反事実的条件文の繋がりは本文（p. 286）をそのまま引用した。「Reggie が引き金を引かなかったならば、Archie はよりよかっただろう」という反事実的条件文を詳しく分析すると、「Reggie が引き金を引かなかったならば、撃針が弾薬筒の雷管に当たらなかっただろう」、「撃針が弾薬筒の雷に当たらなければ、弾が発砲されなかっただろう」、「弾が発砲されなければ、弾丸が Archie の体内に入らなかっただろう」、「弾丸が Archie の体内に入らなければ、Archie は麻痺を生じなかっただろう」、「Archie が麻痺を生じなければ、彼はよりよかっただろう」という繋がりがあるものと思われる。

も近い世界においてよりよいかという文脈的な細部に依存する。Klocksiem によれば、最も突出的な事例は、関連した行為者が行為を遠慮した（be thought of as refraining from action）と直観的に考察される場合である。私たちは何らかの理由がない限り、ある行為を遠慮することは道徳的に許容されていると見なす。つまり、【帽子事例】は新しい帽子をあげることを遠慮した事例であり、道徳的に許容されていると見なされる。他方の【看護の放棄事例】は何かを遠慮した事例ではない。加えて、【看護の放棄事例】における Archie の選択は、「約束をしたならば、その約束を破ってはいけない」などの理由から道徳的に間違っていると見なされる。

まとめると、帽子事例では、Archie は Jughead に帽子を渡すことを遠慮した。そのため、Archie は Jughead の福利に影響を与えていないと結論され、これは単なる与益の失敗であると分類される。しかし、Betty の病気の事例では、Archie は看護することが義務であるため、われわれは Archie の既定の（default）行為である看病を期待する。そのため、Archie は Betty の福利に影響したと結論され、それは危害であると分類される。

Klocksiem によれば、一般的に、ある出来事が危害を構成するのか与益の失敗を構成するのかは、文脈的な細部に重く依存する。そして文脈的な細部は、S がより良くなる関連した可能世界が最も接近した比較される世界なのか否かを決定するという。Klocksiem は以上の主張をもとに、与益の失敗と危害の区別を以下のように一般化する[16]。

> もしSがよりよくなる世界が比較される世界と非常に似ている場合、だから（so）その世界〔比較される世界〕が生じないようにある介入が要求され、我々はそれに関連した出来事を危害と見なす[17]。もしSがよりよくなる世界が比較される世界とあまり似ていない場合、だから（so）その世界〔S がより良くなる世界〕が生じるようにある介入が要求され、我々はそれに関連した出来事を与益の失敗とみなす[18]。また、S がよりよくなることについて約束することやそれが生じるように義務を負うことは、その〔S がよりよくなる〕世界をより突出させるひとつなのである[19]。

16　Klocksiem, 2012, pp. 294-295

17　【看護の放棄事例】では、Archie は Betty との約束を破っている。通常は、約束を破ることは悪いこととみなされるため、Betty がよりよくなる世界として Archie が約束を守る世界が可能世界の候補の中で突出する。このような理由で、Betty がよりよくなる世界は比較される現実世界と非常に似ている（Betty がよりよくなる世界が一つに決まる）と解釈される。ここでは、現実世界が生じないように「約束を守る」という介入が要求されており、その介入に関連した出来事、すなわち看護をしないことが危害であると判断される。

18　【帽子事例】では、Archie が Jughead に帽子を挙げない世界は一つには決まらない。なぜなら、Archie が帽子を挙げない理由については様々に考えられるからである。単に気が変わったのかもしれないし、誰かから何かを吹聴されたのかもしれない。このような理由で、Jughead がよりよくなる世界は比較される現実世界とあまり似ていないと解釈される。ここでは、Jughead がよりよくなる世界が生じるように「帽子をあげる」という介入が要求されており、その介入に関連した出来事、すなわち帽子をあげないことは与益の失敗だと判断される。

19　この引用部について、筆者と Hanna（2016）の解釈は異なっている。Hanna は与益の失敗に関する介入を「不作為」として捉えているが、Klocksiem による与益の失敗と危害の区別は世界の近さがポイントであるため、与益の失敗に関する介入が必ずしも不作為である必要はないだろう。なお、自然な問いとして、①Sがよりよくなる世界が比較される世界と非常に似ており、Sがよりよくなる世界が生じるようにある介入が要求された場合には、これに関連した出来事は危害と与益の失敗のどちらなのか、同様に、②Sがよりよくなる世界が比較される世界とあまり似ておらず、その比較される世界が生じないようにある介入が要求された場合には、これに関連した出来事は危害と与益の失敗のどちらなのか、が浮かぶ。Hanna（2016）は①に似た事例（Archie がファッション批評家から購入した帽子を批判され、Jughead にその帽子をあげないと決心した事例）も危害とみなされうることを指摘し、それをもって Klocksiem の定式化は十分ではないと批判する (Hanna, 2016, p.255)。先述したように、筆者と Hanna の Klocksiem の主張への解釈は異なっており、更なる検討が必要だろう。

2.5. 先取り問題[20]

【浮気事例】

Betty と Veronica は Archie と付き合っており、Archie が二股をかけていたことにそれぞれ気が付く。Betty と Veronica は Archie の脚を折ってやろうと Archie の自宅に押し掛け、結果として Veronica のほうが Betty よりも早く到着し、Veronica が Archie の脚を折った。このとき、明らかに Veronica は Archie を害しているようにみえるのだが、CCA によればそうならない。もし Veronica が Archie の脚を折っていなくても、Betty が Archie の脚を折るので、可能世界の Archie は現実世界の彼とちょうど同じだけ悪くなっていただろうからである。

Klocksiem の応答

Klocksiem は先取り問題を直接解決する手立ては述べていない。だが、Archie は Veronica に脚を折られることのみによって内在的に危害を受けたと主張することで、誰が危害を加えたのかに関する説明は CCA によって可能であるとする。

【浮気事例】に対して反事実的に比較される事例は、Veronica が Archie の脚を折らない世界ではなく、Archie の脚が折れない世界である。Veronica が Archie の脚を折らなければ Betty が Archie の脚を折っただろうという事実は、Veronica が Archie の脚を折らなければ Betty が Archie を害しただろうということのみを含意する。ここでは、脚を折ることは危害を構成しないということは含意されず、正確には、「もし Archie が足を折られていなければ、Archie はよりよかっただろう」となる。したがって、Veronica が Archie の脚を折らなければ、Archie はよりよかっただろうと言える。

3. Carlson による CCA への反論と Klocksiem との応酬

Klocksiem は詳細に CCA の擁護を展開したが、Carlson（2018）は Klocksiem の主張をもってしても CCA には問題があることを指摘した。その後、Klocksiem と Carlson による論文上での応酬が続くことになる。本章では、二人の争点を確認したい。

Carlson は、CCA の大きな問題点として依然として先取り問題への対処が困難であることを挙げ、さらに、「危害は避ける理由を与え、利益は追求する理由を与える」という思慮深さを意味する決まり文句、および、「危害は〔それを〕人には与えてはいけない理由を与え、利益は〔それを〕他者に授ける理由を与える」という道徳的決まり文句にも合致しない点を指摘する[21]。その根拠を示すために、Carlson は「行為の

20　先取り問題（Preemption）とは、因果論において反事実的条件文などの理論が抱える困難の一つとして知られている。問題となる具体例を挙げよう。X と Y という暗殺者が S を殺害する、以下のような場面を想定してほしい。以下の事例は、Parfit が集団的な危害を説明するために用いる一例から借用した（Parfit, 1987, p. 70）。

【毒殺と銃殺事例】X が S に致死的な毒を盛った後で、Y が S を銃殺する。

このとき、直接的に S を殺したのは Y である。ところが、反事実的条件文を用いると、Y の銃撃が S の死の原因だと結論できなくなる。なぜなら、反事実的条件文である「もし Y が S を撃たなかったならば、S は死ななかっただろう」は真とはならないからである。もし Y が心変わりをして S を撃たなかったとしても、S は既に X によって毒を盛られているので、どのみち死んでしまう。したがって、反事実的条件文による分析にしたがえば、Y は S の死の原因にはならない。しかし、これは明らかに直観に反する判断である。
このように、ある行為や出来事の帰結が、別の行為や出来事の帰結に先回りされているようにみえるため、これらは「先取り」問題と呼ばれる。そのため、因果論においては、既存の理論を改良したり、新たな理論を打ち立てたりする仕方で、先取り問題に対処する探求が進められている（松王 , 2020, pp. 29-31）。

21　Carlson, 2018, p. 797, 亀甲括弧内引用者

理由」と「選好の理由」の側面から CCA を批判する。対する Klocksiem は、Carlson の批判が批判に該当しないことを再反論にて示そうとする。本章では、Carlson による行為の理由（Reason for Action）と選好の理由（Reason for Preference）を用いた批判を紹介し[22]、それぞれに対する Klocksiem の再反論を確認する[23]。

3. 1. 行為の理由

Carlson（2018）は、危害と利益の説明の決まり文句を満たすものとして、以下の原則をあげる。

行為の理由[24]：

> もし何かを選択する状況において、a と a' という代替的な行為があなたに対して開かれており、a をすることがあなたを益し（害さず）、他方で a' することがあなたを益さない（害する）のであれば、a' ではなく a をする思慮深い理由がある。

しかしながら、Carlson によれば CCA はこの行為の理由と両立しない。以下の事例を考えよう。

【ダーツ事例】[25]

> 円形のボード 1 とボード 2 がある。ボード 1 は、幅が狭く色の薄い円でその縁の周囲を囲まれている。ボード 1 にダーツを当てれば 1000 ドルが貰え、周囲の幅の狭い円に当てれば 1200 ドルが貰えるが、それ以外に当てると何も貰えない。ボード 2 に当てれば、100 ドルが貰え、それ以外に当てると何も貰えない。また、あなたはボード 1 や 2 に当てられるくらいにはダーツが上手いが、ボード 1 の周囲の円に意図して当てられるほど上手くはない。
>
> ボード 1 にダーツを当てる可能世界（Wb1）において、「もしボード 1 に当たらなかったならば、周囲の円に当たっていただろう」が真となる。ボード 2 にダーツを当てる可能世界（Wb2）において、「もしボード 2 に当たらなかったならば、どちらのボードにもボード 1 の周囲の円にも当たらなかっただろう」が真となる。

ボード 1 の周囲の円にダーツが当たる世界では、ボード 1 に当てるよりもより多くの賞金を貰えるため、Wb1 において「もしボード 1 に当たらなければ、より良くなっていただろう」が真となる。どちらのボードにもボード 1 の周囲の円にも当たらない世界では獲得金はないため、Wb2 において「もしボード 2 に当

22　以下の記述は、主に Carlson（2018）の議論に沿う。

23　Klocksiem の再反論として、Klocksiem（2019）の見解を簡単に示す。

24　Carlson, 2018, p. 797. なお、Carlson は行為の理由に関する還元的な分析と非還元的な分析を挙げ、両者とも本文中の定式を満たすとしている。

> 還元的な分析：
> p という事実が a' よりも a する理由であるのは、(1) a する理由があって、a' する理由がなく、(2) a しない理由がなく、a' しない理由があるときであり、かつそのときに限る。
> 非還元的な分析：
> p という事実が a' よりも a する理由であるのは、p は、{a, a'} の選択的な集合から、a する理由であるときであり、かつそのときに限る。
>
> ある事実 a があなたを益し、a' があなたを益さない場合、確かにこれは a をする理由であり、a' をしない理由ではないように見える。同様に、ある事実 a' があなたを害し、a があなたを害さない場合、確かにこれは a' をしない理由であり、a をしない理由ではないように見える。これらの両者の事実は、{a, a'} の集合から a をする理由のように思われる。つまり、a と a' がその状況下において唯一の選択肢であれば、a する理由がある。

25　Carlson, 2018, p. 798,　亀甲括弧内引用者

たらなければ、より悪くなっていただろう」が真となる。したがって、CCA は、ボード 1 に当てることはあなたを害し、他方でボード 2 に当てることはあなたを益したことを含意する。

しかし、行為の理由を参照すると、あなたにはボード 2 を選択する理由はないように思われる。もしこれが正しければ、CCA は行為の理由に違反している。

3. 1. 1. Klocksiem の再反論[26]

Klocksiem によれば、行為者が直接的にコントロールできることは、どの的に当てるのかについてのみであり、ダーツがどこに当たるのかについてのコントロールは間接的なものである。そこで彼は、「投げる」という行為と「当たる」というアウトカムを区別して、Carlson に反論する。【ダーツ事例】における行為とはダーツを投げることであり、アウトカムとはボードに当たること、ならびに獲得金額のことである。

彼によれば、アウトカムが危害か否かを考慮しているときには、そのアウトカムが生じた世界におけるある人の福利と、異なったアウトカムが生じ、かつ（現実世界と）最も近い世界におけるある人の福利とを比較している。しかし、行為それ自体が危害か否かを考慮しているときには、ある行為が実行された世界におけるある人の福利と、その行為が実行されず、かつ（現実世界に）最も近い世界におけるある人の福利とを比較しているはずである。

まず、アウトカムに注目する。ボード 1 に投げることを a1、a1 を実行する世界を Wa1 と呼ぶ。【ダーツ事例】において、Wa1 における a1 のアウトカムは、ボード 1 に当てることと 1000 ドルを勝ち取ることである。Wa1 に最も近い可能世界における a 1 のアウトカムは周囲の円に当たることであり、あなたは 1200 ドルを獲得する。あなたにとって 1200 ドルを獲得することは 1000 ドルを獲得することよりもよりよい。そのため、CCA によれば、ボード 1 に当てることはあなたを害することを含意する。Carlson の記述では、Wa1 に最も近く、a1 ではない行為を実行する可能世界は、ボード 2 に投げる（a2 を実行する）世界となる（Wa2 と呼ぶ）。Wa2 における a2 のアウトカムは、ボード 2 に当て、100 ドルを獲得することである。したがって、a1 を実行することはあなたを益する。なぜなら、Wa1 に最も近く、a1 以外の行為を実行する世界（すなわち、W a2）におけるあなたは、Wa1 のあなたよりもより悪いからである。

しかし、行為に注目すると、その行為が実行される世界とその行為が実行されない世界を比較することになる。もし a1 の規定されたアウトカムが a2 のそれよりもよりよいのであれば、a1（ボード 1 に当てようと試みること）はあなたを益するといえる。もともと、行為の理由はアウトカムではなく行為に適用されるものである。したがって、Klocksiem は、CCA がその行為を利益と分類する場合、あなたは a1 を実行する理由をもち、CCA がその行為を危害と分類する場合、あなたは a2 を実行する理由をもたないとし、ダーツ事例において CCA は行為の理由に違反していないと結論する[27]。

26　Klocksiem, 2019, pp. 3-5

27　Carlson（2020）は応答論文にて、更に反論する。すなわち、基本的な行為だとされている「腕をあげる」ことでさえ時おり失敗すること、ならびに、ボードに当てることはコントロール以外の背景的要因にも左右される事実から、Klocksiem に対して、行為であることからボードに当てることを除外し、他方で腕を挙げることは除外されない「直接的なコントロール」の定義を示すようにと述べている。また、Carlson（2020）は行為とアウトカムの差異がごくわずかになるような事例（ボタンを押す行為が行為者の福利の増減に直結する事例）を挙げ、Klocksiem の見解に反例を示す。

3. 2. 選好の理由

Carlson（2018）は、行為の理由以外にも選好の理由を示しながら、CCA を批判する。

選好の理由[28]：

> 可能な出来事 e と e' が現在ないし将来の状況における代替的なアウトカムであり、e が生じるのであれば e はあなたを益し（害さず）、他方で e' が生じるのであれば e' はあなたを害さない（害する）場合、あなたは e' が生じるよりも e が生じることを好む思慮深い理由をもつ。

Carlson が批判に用いる事例は以下の通りである。

【コインマシン事例】[29]

> とあるコインマシンがあり、2 枚のうち 1 枚のコインを選んでトスする。もしコイン 1 が選ばれたのであれば、コインの表が出た場合にはあなたは 2000 ドルを獲得し（表 1 とする）、裏が出た場合には 1000 ドルを獲得する（裏 1 とする）。もしコイン 2 が選ばれたのであれば、表が出た場合にはあなたは 2000 ドルを失い（表 2 とする）、裏が出た場合には 1000 ドルを失う（裏 2 とする）。
>
> 裏 1 が生じる可能世界（裏 1 世界）において、「裏 1 が生じなかったならば、表 1 が生じただろう」が真となる。同様に、裏 2 が生じる可能世界（裏 2 世界）において、「裏 2 が生じなかったならば、表 2 が生じただろう」が真となる。

この事例では CCA は裏 1 があなたを害し、他方で裏 2 があなたを益することを含意する。しかし、通常は、裏 1 よりも裏 2 を好む理由はない。したがって、CCA は選好の理由にも違反しているという。

3. 2. 1. Klocksiem の再反論[30]

Klocksiem は、もし裏 2 が可能なアウトカムならば、裏 1 は明らかに排除されているため、裏 1 と裏 2 は代替案ではないと反論する。すなわち、コイン 2 が選択された世界において、コイン 1 が選択される世界は代替的ではない。コイン 2 が選択され、かつ、裏がでた世界に最も近い世界は、コイン 2 の表が出る世界である。

選好の理由における二つの代替的行為は、直接的な代替案に制限されている。換言すれば、もし e1 が生じなかったら起きていたであろう e2 が他の出来事として存在するはずである。すなわち、選好の理由が考慮の対象とする事例は、e1 と e2 のような直接的に代替可能なものに限定されている。ここでコイン事例を確認すると、この事例はコイン 1 のアウトカムとコイン 2 のアウトカムが分離しているという構造をもつ。これは、コイン 1 とコイン 2 のアウトカムが直接的に代替的ではないことを示している。したがって、コイン事例は、選好の理由の考察対象から除外される[31]。

28　Carlson, 2018, p. 800
29　Carlson, 2018, p. 801,　亀甲括弧引用者
30　Klocksiem, 2019, pp. 1-3
31　この点について、Carlson（2020）は更なる反論を展開している。すなわち、コインの選別とコイントスが同時に生じることを想定すれば、Klocksiem の反論は回避できるというものである。

3. 3. その他の方面からの批判

Carlsonは、Klocksiemが主張したCCAの文脈鋭敏性について批判している[32]。Carlsonは先に紹介した【浮気事例】に若干改変がされた事例（【浮気事例・改】とする）を例に挙げる[33]。【浮気事例・改】では、後から到着することになる人物Bettyは、Archieを殺す意図をもっている。シナリオとしては、先に到着した人物Veronicaが脚を折り、遅れて到着したBettyは現場をみてそのまま帰る。つまり、「もし先にArchieが脚を折られていなかったならば、BettyはArchieを殺していただろう」が真となる。CCAによれば、脚を折ることは彼を益したことになってしまう。

Klocksiemによれば、【浮気事例・改】も文脈にかかっている。もしVeronicaがBettyの計画を知っており、彼の生命を救うために脚を折ったのだとすると、関連した対照的な世界はBettyが彼を殺す世界となる。もしそうであるなら、CCAはVeronicaの行為が彼を益したことを含意する[34]。Klocksiemからすれば、この文脈においてVeronicaがArchieを益したという含意は不自然ではない。だが、Carlsonは、「他者の行為によって〔あなたが〕害されるのか否かを決定することは、想定される危害の担い手（the subject of the putative harm）としてのあなたを含む事実によって大部分が決定されるべきだ」[35]とし、Klocksiemの見解を批判する。

さらに、Carlsonは、Klocksiemが文脈とは何かについて明示していないことを指摘する。Klocksiemの見解からは、行為者の意図についての事実や、私たちの主観的な注意の向き方などが「文脈とは何か」への回答として読み込めるのだが、明確には述べられていない。Carlsonは、【ダーツ事例】において、ボード1に当たる帰結をボード1の周囲の円に偶然に当たることではなく、ボード2に当たるかどちらにも当たらないことの帰結と比較すれば、ボード1に当たることはあなたを益すると認めている。しかし、文脈次第では、やはりボード1に当てることがあなたを害するという。例えば、【ダーツ事例】におけるあなたは、ダーツを始める前に関連するすべてを知っているとしよう。このとき、私たちの注意の向き方は、ボード1に当てる行為がうまく遂行された場合と、ボード1に当てようとして偶然に周囲の円に当たった場合との比較に向く。同様に、ボード2に当てる評価においては、我々の注意の向き方は、うまくボード2に当たる場合と、ボード2に当てようとしてそれに失敗する場合との比較に向く。すると、ボード1に当てることがあなたを害し、他方でボード2に当てることがあなたを益するという文脈が明らかなように思われる。

Carlsonはさらに、危害と利益を無条件に（simpliciter）決める事実はないとする「対照主義（Contrastivism）」を用いて、対照主義に合致したCCAを導入し、そのCCAをも棄却する[36]。

彼によれば、「危害と利益についての『対照主義』は、出来事（行為を含む）は特定の人に対して、決して無条件的に危害的であったり利益であったりすることはないという一般的な主張を提示」し、「ある出来事が危害なのか利益なのかは、与えられた対照的な出来事との関係性においてのみ決まる」とされ

32　Carlsonはその他にもKlocksiemの見解の問題点として、バックトラッキングを犯している事例があることを挙げている。しかし、Carlson（2018）が提示する【コインマシン事例】事例も、コインの選定とコイントスに時間的な隔たりがあるのであれば、バックトラッキングを犯しているように思われる。両者の事例の適切性については、今後の課題としたい。

33　【浮気事例・改】はKlocksiem自身が、CCAにとってより困難な事例として挙げたものである（Klocksiem, 2012, p.296）。KlocksiemによればCCAはこちらの事例にも対処できるという。なお、因果論の議論を参照すると（Kutach, 2014, p. 77-78, 相松慎也訳）、【浮気事例】は二人の人物が足を折ることのそれぞれを代替しても帰結に影響がないため、過剰決定の問題（Overdetermination）として、【浮気事例・改】は骨折と死の時間的順序を逆にすると帰結が異なるため、先取り問題（Preemption）として解釈できるだろう。

34　Norcrossは、すべてを考慮した上での危害が特定の視点による危害から区別されることを説明するために、（登場人物が死ぬことはないが）【浮気事例･改】に構造が似た事例を挙げている（Norcross, 2005, p. 150-152）。NorcrosもKlocksiemも、危害の考慮に文脈を重要視する点で似ている部分を有するが、Norcrossは可能世界を用いた比較を否定している。

35　Carlson, 2018, p. 803 亀甲括弧内引用者

36　Ibid, pp. 805-806

る[37]。対照主義的 CCA の形式は以下の通りである[38]。

対照主義的 CCA：
ある出来事 e や行為 a が、他の出来事 e' や行為 a' よりも、ある人を全体的に害する（益する）のは、その人が、e の代わりに e' が生じたか、a の代わりに a' が実行された場合に、結果的により良くなっている（より悪くなっている）ときであり、かつその時に限る。

【ダーツ事例】においては、対照主義的な CCA を用いれば、ボード 2 に当てることよりもボード 1 に当てることがあなたを益することを含意する。なぜなら、ボード 1 に当てることはボード 2 に当てることよりも、あなたをよりよくするからである。

だが、無条件的な危害や利益である出来事はないとすると、危害的な出来事を予防したり嫌悪したりするような、思慮深さの理由や道徳的な理由との調和が困難であるように思われる[39]。このような根拠により、Carlson は対照主義的 CCA も棄却する[40]。

4. 医療倫理・医療哲学との関連

最後に医療倫理・医療哲学との関連について、CCA と一時的な危害とみなされる医療事故の扱い方、ならびに、トータルペイン（全人的苦痛）の整合性の観点から、簡単に述べる。

一時的な危害とみなされる医療事故の扱い方

例えば、ある患者がおり、健康回復のためにはある手術を受けなければならないとする。手術は成功したものの、ガーゼが体内に残されたまま、手術が終了してしまった。このとき、医師は意図的にガーゼを残したわけではない。後日、患者の痛みの訴えから、体内にガーゼが残存していることが発覚した。再度の手術により、ガーゼは無事に患者の体内から取り除かれ、また、患者の健康は回復したとする。

現実世界に最も近い可能世界における患者は、体内にガーゼを残されなかったとする。このとき、現実世界と可能世界の双方において、患者の回復という最終的な帰結は変わらないため、「ガーゼの残存が無かったならば、患者はよりよかっただろう」が真とはならない。ガーゼの残存は一時的な危害とみなされうる。

Klocksiem によれば、抗がん剤治療は一時的な危害であると説明がされた。だが、抗がん剤治療の危害は利益を目的としていたのに対し、ガーゼの体内残存は防ぐべき危害であるように思われる。つまり、抗がん剤治療と医療事故は双方とも一時的な危害とされる一方で、双方の扱われ方には相違がある。

今後の探究の方針として、医療事故の事例と Klocksiem の事例における【救助事例】を類比的に捉えることができるだろう。Klocksiem によれば、骨折は一時的な危害であり、不必要な危害でもあると説明がされていた。この説明における不必要な危害が、CCA や反事実的条件文の意味論を通してどのように説明されるのかについて、更なる検討が必要である。

37　Ibid, p. 805
38　Ibid.
39　なお、Carlson は「CCA の対照主義バージョンについて議論することは、この論文の射程を越えている」と置き、詳細な議論は展開していない (Carlson, 2018, p. 806)。
40　Carlson（2021）は以後、CCA に対抗する理論として Causal Account of Harming を提示するが、結局のところ、こちらの見解も危害を捉えるにはうまくいかないと結論する。

CCA とトータルペイン（全人的苦痛）の整合性

緩和ケアの領域では、苦痛を「トータルペイン（全人的苦痛）」として捉えている[41]。トータルペインとは、看護師であるシシリー・ソンダースが提唱した概念であり、「身体的苦痛」、「精神的苦痛」、「社会的苦痛」、「スピリチュアルペイン」の４つの苦痛の総称である。特にスピリチュアルペインとは、宗教観や死生観に基づき、身体的・精神的とも違う実存的な自己が受ける苦しみであり[42]、身体的ないし精神的苦痛とは区別される。それぞれの痛みは互いに影響しあい、全人的な苦痛として現れる[43]。

このような苦痛の捉え方と CCA の整合については、今後の課題である。

まとめ

本稿は CCA の基本的な形式、および、CCA を用いると様々な事例がどのように解釈されるのかを紹介し、それらへの批判、ならびに再反論を確認することを目的とした。Klocksiem は CCA を用いた分析は文脈に重く依存することを前提とし、文脈が提示されれば、比較される世界の近さや、内在的 / 外在的危害、ならびに一時的 / すべてを考慮した危害の区別が明らかになると述べた。対する Carlson は、Klocksiem の見解では、主に「行為の理由」と「選好の理由」について CCA が満たさないことを指摘した。この Carlson の反論について、Klocksiem は再反論を提示している。最後に、筆者の関心である医療倫理・医療哲学との関連を簡単に述べた。すなわち、CCA と一時的な危害とみなされる医療事故の扱い方、ならびに、トータルペイン（全人的苦痛）の整合が今後の課題としてあがる。

謝　辞

二人の匿名査読者の貴重なご指摘と建設的なコメントにより、本稿の内容を大幅に改善することができました。また最終原稿の投稿に際しては、宮園健吾准教授（北海道大学）より的確な助言を賜りました。ここに深く感謝申し上げます。

参考文献

Beauchamp, T. L. and Childress, J. F. (2019). *Principles of Biomedical Ethics*. 8th ed. Oxford university press.

Bradley, B. (2012). “Doing Away with Harm,” *Philosophy and Phenomenological Research*, 85 (2): 390–412.

Carlson, E. (2018). “More Problems for the Counterfactual Comparative Account of Harm and Benefit,” *Ethic Theory and Moral Practice,* 22: 795–807.

Carlson, E. (2020). “Reply to Klocksiem on the Counterfactual Comparative Account of Harm,” *Ethic Theory and Moral Practice*, 23: 407–413.

Carlson, E., Johansson, J. and Risberg, O. (2021). “Causal Accounts of Harming,” *Pacific Philosophical Quarterly,* 103 (2): 420-445.

Feinberg, G. (1984). *Harm to Others*. Oxford University Press.

Hanna, N. (2016). “Harm: Omission, Preemption, Freedom,” *Philosophy and Phenomenological Research,*

41　恒藤 , 2020, p. 9

42　岩崎 , 2015, p. 310-311

43　トータルペインの考え方は複雑かつ論争的であり、トータルペインそれ自体が検討されるべきでもある（詳細は、清水（2022, p. 262-265）を参照のこと）。

93 (2): 251-73.

Klocksiem, J. (2012). “A Defense of the Counterfactual Comparative Account of Harm,” *American Philosophical Quarterly,* 49 (4): 285 – 300.

Klocksiem, J. (2019). “The Counterfactual Comparative Account of Harm and Reasons for Action and Preference: Reply to Carlson,” *Ethic Theory and Moral Practice*, 22: 673–677.

Kutach, D. (2014). *Causation*. Polity Press.

——邦訳 相松慎也（2023）『現代のキーコンセプト　因果性』岩波書店

Menzies, P. and Beebee, H. “Counterfactual Theories of Causation,” *The Stanford Encyclopedia of Philosophy (Summer 2024 Edition)*, Edward N. Zalta & Uri Nodelman (eds.), <https://plato.stanford.edu/archives/sum2024/entries/causation-counterfactual/>.

Norocross, A. (2005). “Harming in Context,” *Philosophical Studies*, 123: 149–73.

Parfit, D. (1987). *Reasons and Parsons*. Oxford University Press.

——邦訳 森村進（1998）『理由と人格 非人格性の倫理へ』勁草書房

Schöne-Seifert, B. (2004). “Harm,” . *Encyclopedia of Bioethics*. 4th ed. Editor: Jennings, B. 3: 1381-86.

飯田隆（2024）『増補改訂版　言語哲学大全Ⅲ　意味と様相（下）』勁草書房

岩崎大（2015）『死生学　死の隠蔽から自己確信へ』春風社

清水哲郎（2022）『医療・ケア従事者のための哲学・倫理学・死生学』医学書院

恒藤暁（2020）「第１章　緩和ケアの現状と展望」（恒藤暁・田村恵子編）『系統看護学講座　別巻　緩和ケア』医学書院

ビーチャム，トム・Ｌ＆チルドレス，ジェイムズ・Ｆ．（2009）『生命医学倫理』（立木教夫・足立智孝監訳）麗澤大学出版会

松王政浩（2020）『科学哲学からのメッセージ：因果・実在・価値をめぐる科学との接点』森北出版

研究ノート

道徳的理由のあいだの葛藤という観点から超義務を考える

浦野&石田「しないに越したことはない:超義務と亜義務の倫理学」へのコメンタリー

玉手慎太郎（学習院大学）

要旨

本研究は、浦野敬介と石田柊による論文「しないに越したことはない:超義務と亜義務の倫理学」に対するコメンタリーである。当該論文は、超義務（道徳的に望ましいが義務とまではみなされない行為）および亜義務（道徳的に望ましくないが禁止されるほどではない行為）について緻密な概念分析を行ったものである。これら二つの概念の既存の理解にはパラドックスが残されていたが、浦野らは超義務を「慈恵型超義務」と「英雄型超義務」に、また亜義務を「非行型亜義務」と「極悪型亜義務」に分類した上で、それぞれが成立する理由を明瞭に示した。彼らの議論は総じて説得的であるが、しかしそこで示された区分によって超義務と亜義務のすべての形態が網羅されているかどうかには、なお異論の余地がある。本研究は複数の道徳的理由のあいだの葛藤という観点から、浦野らの区別では捉えきれない「葛藤型超義務」および「葛藤型亜義務」という第三の分類を提示し、浦野らの議論を拡張することを試みる。それによって超義務と亜義務についてより包括的に、またわれわれの日常的な実践および苦悩をより詳細に捉える形で描くことができるようになると期待される。

Abstract

This research is a commentary to “Better not to do: the ethics of supererogation and suberogation” (written by Keisuke Urano and Shu Ishida). The paper performed detailed conceptual analysis of supererogation (the action that is regarded as morally good but not obligated) and suberogation (the action that is regarded as morally bad but not prohibited). Urano and Ishida made clear understanding of these concepts by classifying each concept to two subclasses: supererogation is divided to heroic and benevolent supererogation, and suberogation is divided to delinquent and atrocious suberogation. Their inquiry process and conclusive proposal are highly persuasive, but not enough to deal with supererogation (and suberogation) totally. In this research, we present the third subclass of supererogation (and suberogation), namely “conflictive supererogation” (and “conflictive suberogation”) which is overlooked in Urano and Ishida’s framework. By including this third subclass and expand the conceptual framework, it becomes possible to illustrate supererogation (and suberogation)

more comprehensively, which make sure that supererogation (and suberogation) is not just theoretical construction but familiar aspects of our practices and worries in everyday moral choices.

はじめに

浦野敬介と石田柊の論文「しないに越したことはない：超義務と亜義務の倫理学」[1]は、超義務（supererogation）および亜義務（suberogation）について詳細な概念分析を行い、その規範的性質を明らかにした優れた研究である。本稿はこの論文について検討を加え、その議論枠組みを（否定するのではなく）拡張することを試みる。それによって超義務および亜義務という概念の内実をより包括的に、またわれわれの日常的な葛藤や苦悩をより詳細に捉える形で描くことができるようになると期待される。

2. 論文の要点（1）：超義務の定義および超義務をめぐる課題

まずは浦野らの論文の要点を確認していきたい。ただし本節では、ひとまず亜義務は脇に置き、超義務に限定して議論を進める。亜義務は超義務の反転として理解することができるものであり、実のところ浦野らも超義務から議論を組み立てた上で、その応用として亜義務を論じているからである。亜義務については後に6節で取り扱う。

超義務とは、「望ましいが義務的ではない」[2]行為が有する道徳的性質を指す。具体例として、火災が発生しているビルに取り残された幼児を助けるためにそのビルに飛び込む行為や、私財のうちそれほど大きくない額を慈善団体に寄付する行為が挙げられる[3]。これらの行為が道徳的に望ましいことは明らかであるが、しかし私たちは通常、それが義務であるとは考えない。こういった行為は「義務の要請を超える」[4]ものである。

これらの例に明らかなように、超義務は私たちの日常において決して珍しいものではない。しかしそれでいて、超義務がいかなるものであるのかはいまだ明らかではない。というのも、ある行為が別の行為よりも道徳的に望ましいならば、ふつうその行為を取ることは道徳的に義務となるはずだからである。道徳的に望ましいにも関わらず義務とは言えない、という事態はいかにして成立するのか。これが浦野らの取り組む問いである（これは「超義務のパラドックス」と呼ばれる問題である[5]）。

もちろんこの問いは新しいものではない。浦野らはこの問いに対する先行研究の応答として「道徳的等価説」と「総合的免責説」の二つを検討し、いずれも説得的ではないとして退ける。まず道徳的等価説は、超義務的な行為は実際には他の行為よりも道徳的に良いわけではない（つまり私たちは端的にそれらの行

1　浦野敬介＆石田柊「しないに越したことはない：超義務と亜義務の倫理学」『応用倫理』15（2024）: 15-32.
2　浦野＆石田論文16頁。原文の傍点は外している。
3　浦野＆石田論文16頁。
4　浦野＆石田論文18頁。傍点は原文の通り。
5　超義務のパラドックスとは、厳密には以下のようなものである（浦野＆石田論文19-20頁）:「以下の三つの主張は同時にすべて成立することができないため、超義務はトリレンマに陥る。〈主張1〉超義務的行為は、義務的行為よりも道徳的によい。〈主張2〉行為ϕが行為ψよりも道徳的に良いならば、行為ψに代えて行為ϕをすることが道徳的に義務的である。〈主張3〉義務的行為に代えて超義務的行為をすることは、道徳的に義務的ではない。」

為の道徳的地位を誤解している)、と論じることで問いに答える考え方である。たしかにこのように言えるならパラドックスは解消される。しかし道徳的等価説は、たとえ論理的には問題を解決できたとしても、超義務的な行為は道徳的に望ましいものであるという私たちの直観それ自体を否定するものであり、説得力に欠けると浦野らは述べる。筆者もこの批判に同意する。

対して総合的免責説は、超義務的な行為について、たしかにそれ自体として義務ではあるのだが、多大な自己犠牲を伴うため、総合的に判断すれば義務とは言えない（それゆえ当該の行為を取らなかったとしてもその義務違反は免責される)、と論じることで問いに答えようとする。この応答は一見したところもっともらしい（私たちは燃え盛るビルを前にして、飛び込めば自分も高い確率で死ぬ、ということを考えないわけにはいかない)。しかし、やはり私たちの直観に反すると浦野らは述べる。というのも、燃え盛るビルに飛び込むことは義務を超えていると論じる時、私たちは道徳外の理由によって道徳的義務が免除されると考えているのではなく、道徳的にさえ義務的ではないと考えているからである。言い換えれば、超義務はあくまで道徳的な性質をめぐる問題だということを、総合的免責説は捉え損ねている[6]。筆者は以上の議論にも同意する。

3. 論文の要点（2）：超義務の二類型およびその規範的地位の解明

では、浦野らはこの問い（＝超義務のパラドックス）にどのように解答するのか。彼らはまず、超義務は決して一枚岩ではなく、その内部に二つのタイプを区別することができると論じる。一つ目が英雄型超義務（heroic supererogation）であり、その特徴は「コストが極めて大きいこと、超義務的行為の道徳的なよさが極めて大きいこと、（他の条件が適切ならば）極めて大きな道徳的称賛の対象となることなど」にある[7]。二つ目が慈恵型超義務（benevolent supererogation）であり、その特徴は、「コストが小さく、道徳的なよさはそれほど大きくなく、また道徳的称賛の程度もそれほど大きいとはいえない」ことにある[8]。先に挙げられた超義務の例のうち、幼児の救助のために燃え盛るビルに飛び込む行為は前者、慈善団体に小額を寄付する行為は後者に該当する。

以上の区別の上に、浦野らはそれぞれが異なる理由から、超義務のパラドックスを解消できること、つまり道徳的に望ましいが義務的ではないという性質を説得的に説明できることを示す。

まず英雄的超義務については、当該の行為はそもそも当事者にとって選択可能な行為ではないことが、その超義務としての性質を考える鍵となる。選択可能でないならば、その行為は義務論的地位を持たない。なぜなら、ある行為が義務であると述べるためには、その行為が実際に選択可能でなければならないからである。しかし、英雄的な行為はこの条件を満たしていない（容易に選択できるものではないからこそ英雄的なのである)。このように考えるならば、英雄的超義務が義務ではないこと、しかしそれでいて高い道徳的称賛に値することを整合的かつ直観に即した形で説明できる。

つづいて慈恵的超義務については、当該の行為の道徳的な望ましさが非常に小さいことが、その超義務としての性質を考える鍵となる。ある行為が道徳的に望ましいことはたしかであるとしても、その望ましさは、当該行為を義務とするには及ばないことがありうる。その望ましさは当該行為を義務にするまでのものではない、というわけである（このような考え方を浦野らは「道徳的理由不足説」と呼ぶ[9])。このよ

6　浦野らの論文では総合的免責説に対する批判がさらに2つ論じられているのだが、ここで取り上げた論点だけでも総合的免責説を退けるのに十分であると考えるため、本稿では省略する。
7　浦野＆石田論文 23 頁。
8　浦野＆石田論文 23 頁。
9　浦野＆石田論文 23 頁。

うに考えるならば、慈恵的超義務が義務でないこと、それでいてなお道徳的な望ましさを有していることを整合的かつ直観に即した形で説明できる。

筆者の見る限り、以上の議論は非常に説得的であり、英雄的超義務および慈恵的超義務について、超義務のパラドックスを解消することに成功している。浦野らの研究は、いまだ明確に理解されていない「超義務」という概念の道徳的地位を、理論的に精緻な枠組みの上に解明したものとして、高い評価に値すると言えよう。

4. 超義務のもうひとつの類型化：葛藤型超義務の可能性

しかし浦野らの研究には、なお不十分な点が残されていると筆者は考える。注意すべきは、はじめに超義務を区分した上でそれぞれについて説明を与える、という形で論理が組み立てられていることである。それゆえ、仮に前提となる超義務の類型が不十分なものであった場合には、個々の区分については説得的な説明が与えられていたとしても、なお超義務全体について説得的に説明したことにはならない。以下では、浦野らの整理からは抜け落ちている第三の超義務カテゴリーを提示し、説明を試みる。

ここで例として考えたいのは、大学教員による、「学生の卒論指導を丁寧に行う」という行為である。現在の日本の大学の、いわゆる文系学部（文学部や法学部など）において、多くの場合に４年生は卒業論文を執筆する。よく知られているように、卒業論文の指導の内容には教員ごとに大きな差がある。つまり、熱心に指導を行う教員もいれば、最低限の指導で済ませる教員もいる[10]。熱心に指導を行うことは、道徳的に望ましいことであると言えるだろう[11]。他方で、熱心な指導は教員にとって道徳的義務である、と考えることはさほど説得的ではなく、一般にそう見なされてはいないように思われる。たいていの教員は「丁寧に指導するのは立派なことだ」と考え、しかし同時に「でもそれは誰もがやるべき義務ではない」と考えるだろう。四年生の卒論を丁寧に指導することは、道徳的に望ましいが、しかし義務ではない。つまりこれは超義務である[12]。

この超義務は、英雄型超義務ではない。丁寧な卒論指導が有する道徳的価値はそれほど大きくないし、指導を丁寧に行うことは十分に実行可能な選択肢である。またこの超義務は慈恵型超義務でもない。丁寧な卒論指導が有する道徳的価値は（英雄的ではないにせよ）決して小さくはなく、義務とみなすのに足らないとは思われない。すると、この超義務は、浦野らの枠組みでは把握できない超義務である、ということになる。だとすれば、新たな検討が必要である。丁寧な卒論指導はいかなる理由で超義務として成立するのであろうか。

筆者の見る限り、卒論指導の事例においては、当該問題について従うべき道徳的理由が複数ある、ということが超義務の背景となっている。仮にここに卒論指導を最低限の内容で済ませている教員がいるとして、その人物は卒論指導に熱意を振り向けない理由をどう説明するだろうか。最もありそうな回答は、他にもやるべきことがある、というものだろう。「たしかに卒論指導はすごく意味があるし、教員としてで

10　全く指導しない、というケースは端的に職務不履行であるため、（現実に生じているかどうかはともかく）ここでは検討から除外する。

11　この道徳的な望ましさの内実は論争的かもしれないが、少なくとも何かしらの道徳的価値があるということには同意が得られると筆者は考える。なお内実の具体例として簡単に思いつくのは、教師として学生に真剣に向き合うことでその信頼に応えること、学生を成長させることで大学に対する社会の期待に応えること、などだろう。

12　テクニカルな補足であるが、浦野らが超義務の形式的定義として挙げるスヴェン・ウヴェ・ハンソンの定義を、この例は満たす。ハンソンによる超義務の定義とは以下のようなものである：「行為者Sが二つの行為ϕ、ψについて、Sがϕすることは道徳的に義務的でもなければ禁止されもせず、Sがψすることは道徳的に義務的だとする。SがψすることよりSがϕすることが道徳的に厳密によりよい場合、かつその場合に限り、Sがϕすることは（ψすることとの関係で）道徳的に超義務的である」（浦野＆石田論文 18 頁）。このϕに熱心な指導を、ψに最低限の指導を当てはめてみてほしい。

きるだけやるべきだとは思うけど、でもほら、教員がやるべきことって他にもあるし」というわけである。これは決して道徳的義務を逃れるための言い訳ではないと思われる。大学教員はたしかに教育に熱心に取り組むべきである（そうすべき道徳的理由をもつ）。しかし同時に、大学教員は学術論文を執筆して査読誌に投稿するべきであるし、様々な授業のための準備をするべきであるし、種々の会議に出席するべきである（それらのことをする道徳的理由を持つ）。さらに言えば、仕事以外の面でも、早く帰宅して家事を行ったり家族と過ごしたりするべきである（ここにもやはり道徳的理由が生じる）。

このような道徳的理由の複数性を踏まえると、丁寧な卒論指導の超義務としての性質は、複数の道徳的理由のあいだの葛藤に由来するものとして理解するのが適切であると考えられる。つまり、丁寧な卒論指導が超義務である（＝道徳的に望ましいがしかし義務ではない）のは、その行為をなすことがある道徳的理由に照らして望ましいことはたしかであるものの、また別の道徳的理由に照らせば望ましいとは言えないからである。もし一つの理由からは望ましくても別の理由からは望ましくないのであれば、それを義務として要求することはできない[13]。

このような超義務を、葛藤型超義務（conflictive supererogation）と呼ぶことにしよう。次節では葛藤型超義務の性質について、いくつかの疑問に答える形で理解を深めていきたい。

5. 葛藤型超義務に対する疑問の検討

5-1. 総合的免責説とどこが違うのか

葛藤型超義務という考え方に対して想定される第一の疑問は、総合的免責説と同じ説明ではないか、というものである。先に見たように、超義務に対する総合的免責説とは、超義務を以下のように説明するものである：当該行為は道徳的に望ましいが、しかしその行為にかかるコストなど、関連する事情を総合的に考慮した場合には義務とは言えない。これに対して葛藤型超義務は、複数の道徳的理由を総合的に考慮した場合には端的に義務とは言えない、というように超義務を説明する。両者の議論の構造にはたしかに類似性がある。

しかし葛藤型超義務の考え方と総合的免責説には決定的な違いがある。総合的免責説は、道徳的な望ましさを、道徳外の考慮事項によって相殺するものである。再び確認すれば、「行為者がすべき行為を決める上での考慮事項は道徳だけでなく、他の価値も考慮事項となるため、超義務的行為は全てを考慮して義務的だとは限らない」[14] というのが総合的免責説である。これに対して葛藤型義務は、複数の道徳的理由を総合的に考える、つまりあくまで道徳の領域内で考えるものである。

とはいえ両者は、単独では義務を形成する道徳的要請を、その要請とは別の理由に照らして義務的でないものとみなす、という構造を共有しており、両者のあいだで線を引くことが難しいケースもある。実のところ上に挙げた卒論指導の例も、他のもろもろの道徳的理由が存在する、という事情をひとまとめにして「忙しくて余裕がない」と述べることも不可能ではない。このとき、忙しい、という理由は道徳的理由ではないだろう（その背後に道徳的理由はあるにせよ）。であれば、この状況は総合的免責説の表層をもつことになる。

13　この考え方は、浦野らの示す「道徳的理由不足説」の亜種とみなすことができる。つまり、ある一つの道徳的理由は、対立する別の道徳的理由が存在する場合には、義務を形成する十分な理由になり得ないということである。なお、「葛藤型超義務」における葛藤は、いわゆる「倫理的ジレンマ」とは区別されるべきである。倫理的ジレンマの場合もまた複数の道徳的理由が対立しているが、そのうち複数のものが義務を形成する十分な理由になっている、という点が問題である。十分な理由が複数あるからこそ、それらの間で深刻な選択困難に陥るわけである（いずれを選んでも道徳的非難が避けられない）。これに対して、ここで論じている葛藤型超義務の場合には、いずれの理由も義務を形成するのに十分ではない。すなわち、理由のうちのどれかを無視しても問題ない（道徳的非難は生じない）のであり、それゆえ「超義務」が形成されると考えられる。

14　浦野＆石田論文 21 頁。原文の傍点および原語表記は省略している。

しかしながら、まったく道徳的でない理由をもって超義務を説明することには、やはり無理があるように思われる。丁寧な卒論指導が超義務である理由として、「丁寧に卒論指導をしても給与が上がるわけではない」という事情を提示することは説得的ではない。言い換えれば、丁寧な卒論指導をすべきだという道徳的要請は、それが給与と無関係であるという事情を総合的に勘案すれば義務ではなくなる、と私たちはふつう考えない[15]。

本稿の提示する葛藤型超義務の議論を、総合的免責説から丁寧に区分することが重要であるのは、それが道徳的理由の重みを尊重することにつながるからである。総合的免責説に対する一つの懸念は、超義務の道徳的な真剣さを容易にキャンセルしてしまうのではないか、というものである。葛藤型超義務にはその懸念はない。ある道徳的要請が別の理由によって義務でなくなるのは、その別の理由それ自体もまた道徳的理由である場合に限られると考えるならば、私たちにとって道徳的理由が優先的な地位を持つという確信を保持したままに、超義務を把握することができる。

5-2. 義務を相対的なものとみなしてしまってよいか

葛藤型超義務という考え方に対して想定される第二の疑問は、そのような超義務の可能性を認めてしまうと、義務が行為者に相対的なものになってしまうのではないか、というものである。もし複数の道徳的理由の相互関係の上で義務が決まるとすれば、まったく同じ行為であっても人によって義務であったり義務でなかったりすることになるし、また同一人物にとっても状況が変われば義務であったはずのものが義務でなくなったりすることになる。しかし義務とは本来、普遍性や絶対性をもつものではなかったか。

以上の疑問に対する応答は二つある。第一に、端的な事実として私たちはまさにそのように、流動性のある形で「義務」という概念を用いているのだ、という応答が可能である。私たちの用いている道徳的概念を明確化する、という目的にとっては、現実の用語法こそが重要である。私たちの義務についての日常的理解を無視して、義務は普遍的・絶対的なものとして捉えなければならない、と論じる方がむしろ不適切だと言える[16]。

第二に、葛藤型超義務に限らず、そもそも超義務自体が行為者に相対的である、と応答することが可能である。たとえば、子どもがプールで溺れている時、その子を助けるために飛び込むことは、泳ぎの得意な人にとっては道徳的な義務であるが、カナヅチである人にとってはそうではなく、英雄的超義務となるだろう。また、かつては泳ぎを得意としていたが、その後に事故によって足が不自由になってしまった人にとっては、かつて義務だったものが英雄的超義務に変化する（つまり義務でなくなる）ことになるだろう。義務が行為者相対的であるという指摘は、葛藤型超義務だけではなく、英雄型超義務や慈恵型超義務にも当てはまる[17]。したがって、義務が行為者相対的であってはならないと考えるならば、本稿の主張のみならず浦野らの枠組み全体を放棄しなければならなくなるが、彼らの議論の説得力をふまえると、それはあま

15　上に挙げた「忙しくて余裕がない」という応答も、忙しさについての追加的な説明によって道徳的理由が示されなければ、やはり義務を怠っていると非難されうるだろう。たとえば、忙しい理由として「今年から学科長なんだ（＝責任ある立場にいるのだ）」とか「親が骨折してしまったんだ（＝家族への配慮義務が強まっているのだ）」といったことを答えて初めて、私たちは丁寧な卒論指導が彼に取って義務的ではないこと（しかしそれでもなお、忙しい中で卒論指導に丁寧に取り組むことはやはり望ましいこと、つまりそれは超義務であること）を認めるだろう。

16　この応答は、より広い観点から捉えれば、次のような方法論的態度へのコミットを意味する。すなわち、私たちの行為をめぐる道徳的理由は、普遍的に当てはまる道徳原理に照らして判断されるのではなく、個別的な事情や視点からこそ把握されるべきである、という態度である。そのような方法論的態度を擁護する近年の研究としてたとえば以下の文献がある：古田徹也『それは私がしたことなのか：行為の哲学入門』（新曜社 2013 年）、大谷弘『道徳的に考えるとはどういうことか』（ちくま新書 2023 年）。

17　実のところ浦野らは、自分たちの検討する超義務が行為者相対的なものであるとはっきり述べている。「とくに、この〔著者らが採用する超義務の〕暫定的定義によれば、ある行為が超義務的であるかどうかは行為者相対的に決まる。つまり、行為者 S にとって行為 ϕ が超義務的だとしても、別の行為者 T にとって ϕ は超義務的ではないかもしれない」（浦野＆石田論文 18 頁、注 8）。なおここで言われている暫定的定義とは、本稿の注 12 で引用したハンソンの定義のことである。

り望ましい理路とは言えないように思われる。むしろ浦野らの議論から、超義務という概念を明確化する上では行為者ごとに義務が相対化すると考えることが必須になる、という示唆を受け取るべきかもしれない[18]。

5-3. 超義務の理解に対していかなる含意をもつか

最後に検討したいのは、葛藤型超義務という概念を導入することに、いかなる含意があるのか、という疑問である。もちろん、葛藤型超義務を用いることでこれまでの枠組みではうまく説明できなかったタイプの超義務を説明できるようになるならば、そのこと自体に意義はある。しかしそれに加えて、何かしら超義務をめぐる考え方そのものへの含意があるかどうかを考えてみよう。

浦野らの論文は超義務について「我々の日常にありふれている」[19]と述べ、それが私たちにとって身近な倫理的問題であることを指摘している。筆者もこの理解に賛同する。しかしながら、浦野らの提示する英雄型超義務と慈恵型超義務という類型は、超義務が身近な倫理的問題であるという点をうまく捉えられているだろうか。

英雄型超義務は、当該の行為が一般には実行不可能なものであることを意味する。また慈恵型超義務は、当該の行為が有する道徳性は瑣末なものであるということを意味する。すると、浦野らの整理における超義務とは、ざっくばらんに言ってしまえば、日常においては問題にならないか問題にするまでもないかのいずれかである。だとすれば、いずれにしても超義務とは、さして悩ましいものではなく、日常的に直面する「倫理的問題」ではない、ということになるように思われる。

これに対して葛藤型超義務は、当該の行為がまさに悩ましいものであるという理解を私たちに与える。その行為はある道徳的理由に照らしてたしかに望ましいのであるが、しかし別の道徳的理由に照らせばそうではなく、それゆえ義務であるという判断を端的に下すことができないものであり、私たちはその行為をなすかどうかを自ら考えて決定しなければならない。このとき私たちは超義務をまさに「倫理的問題」として捉えることになる。すなわち葛藤型超義務の視点は、超義務が私たちにとって身近な倫理的問題である、という理解を正面から受け止める議論を形成する。これが、葛藤的超義務というカテゴリーをめぐる議論が有する重要な含意である。

6. 亜義務のもうひとつの類型化：葛藤型亜義務の可能性

最後に、以上の議論を踏まえた上で、亜義務についても検討しよう。亜義務とは、「望ましくないが禁止はされない行為」[20]のことである。亜義務についても超義務と同様、その道徳的性質の説明は一筋縄ではいかない。もし道徳的に望ましくないのであれば、（法的にはともかく少なくとも道徳的には）禁止されてしかるべきであるように思われるからである。浦野らは亜義務についても超義務と同様に二種類の形態を区別し、それぞれ説明を加える。

第一に、道徳的に望ましくないのだが、その望ましくなさが些細な程度にとどまる行為が、非行型亜

18　義務の行為者相対性は、浦野らの枠組みにおいて超義務のパラドックスを解消する上での鍵になっている。超義務のパラドックスとは、平易に述べれば「望ましい行為であれば義務であるはずなのに、望ましいが義務ではない行為がある」という問題である（本稿注5も参照のこと）。この「望ましい行為であれば義務であるはず」という主張は、正確に記述すれば「望ましい行為であれば常に義務となる」という主張であり（そうでなければパラドックスは生じない）、そしてこの主張こそが、浦野らが乗り越えようとするものである。義務を行為者相対的なものとして捉えるからこそ、この主張の乗り越えが可能になると考えられるし、見方によっては、義務が行為者相対的でありうるならパラドックスは始めからなかった（疑似問題であった）とさえ言えるかもしれない。この論点については今後の更なる検討が必要である。

19　浦野＆石田論文17頁。

20　浦野＆石田論文16頁。原文の傍点は外している。

義務（delinquent suberogation）と名付けられる（これはちょうど慈恵的超義務の逆である）[21]。この場合、道徳的に望ましくないのはたしかであるが、禁止するまでのことではないとみなされる。このような亜義務の具体的事例として挙げられるのは、パートナーへの差別的侮辱に対する罵倒による報復である。第二に、道徳的な望ましくなさが極大であるような行為、つまりは日常的に考えもつかないほどに道徳的に悪い行為が、極悪型亜義務（atrocious suberogation）と名付けられる（これはちょうど英雄型超義務の逆である）[22]。この場合、当該行為はそもそも通常の選択肢に入っていないため、禁止の対象にならない。このような亜義務の具体的事例として挙げられるのはジェノサイドである（私たちはジェノサイドは禁止されるとわざわざ述べたりしない）。

以上の浦野らの整理はやはり説得的であるが、本稿におけるここまでの超義務をめぐる議論を応用することで、第三の亜義務として葛藤型亜義務（conflictive suberogation）を考えることができる。葛藤型亜義務とは、ある道徳的理由に照らせばたしかに望ましくないのだが、別の道徳的理由に照らせば望ましいため、それを端的に禁止することはできない、という行為である。

葛藤型亜義務の具体例として、致死的な状況にあるにもかかわらず宗教上の理由によって治療を拒否する人に対する、治療の差し控えが挙げられるだろう。命を救うことができるにもかかわらずみすみす死なせることは、明らかに道徳的に望ましくない。しかしこのような差し控えは現代社会において禁止されていない[23]。すなわちこれは亜義務である。しかしこの治療の差し控えは、道徳的な問題が些細なものではない（当人は死んでしまう）点で非行型亜義務ではないし、十分に選択可能であって非現実的なわけではない（現実に選択されている）点で極悪型亜義務でもない。ではどのように把握すべきだろうか。治療の差し控えが禁止されないのは、人命は尊重すべきだという道徳的理由に照らせば明らかに望ましくないものだとしても、同時に信仰の自由を守るべきだという道徳的理由に照らせば望ましいものであり、両者のあいだの葛藤があるからである、と説明することは説得的であるし、また現実の医療従事者や患者家族の認識にも合致しているように思われる。だとすれば宗教上の理由による治療の差し控えは、道徳的な望ましくなさを含むものの複数の道徳的理由の対立のゆえに禁止されていない行為、つまり葛藤型亜義務であるとみなすのが適切だろう[24]。

このように、葛藤型亜義務という形態を導入することによって、亜義務についてもより包括的な見取り図を得ることができる。

7. おわりに

本稿は、超義務および亜義務をめぐる浦野と石田の研究について、複数の道徳的理由のあいだの葛藤という観点から、議論の拡張を試みた。はじめに、超義務を英雄型超義務と慈恵型亜義務の二つに区分して検討する浦野らの枠組みを確認した上で（２・３節）、その枠組みではうまく説明できない超義務のパターンが存在することを示した（４節）。それは複数の道徳的理由のあいだで葛藤が生じ、それゆえに単一の

21　浦野＆石田論文 27 頁。
22　浦野＆石田論文 27-28 頁。
23　日本を含む多くの社会において、死を望む患者に対する延命治療の差し控え（いわゆる消極的安楽死）は認められているが、他方で死を望む患者に対して医師が直接に死に至らしめる措置を取ること（いわゆる積極的安楽死あるいは医師幇助自殺）は認められていない。安楽死の禁止（あるいは許容）をめぐる近年の動向については以下の文献を参照した：児玉真美『安楽死が合法の国で起こっていること』（ちくま新書 2023 年）。
24　浦野らが非行的亜義務の例として挙げる、パートナーへの差別的侮辱に対する罵倒による報復という例（浦野＆石田論文 26-27 頁）も、葛藤型亜義務として説明する方が適切であるように思われる。罵倒による報復の害は、禁止するまでもないような些細なものだとは思われない。むしろこの場合、罵倒によって報復することは道徳的に悪い、という理由がありながら、パートナーへの差別的侮辱に対して強く応答しないこともまた道徳的に悪い、という理由が同時に存在するために、報復を端的に禁止できない、と考える方が説得的ではないだろうか。

行為を義務とみなすことができない、というタイプの超義務である。本稿ではこれを「葛藤型超義務」と名付け、超義務の第三のカテゴリーとして提示した。つづいて、いくつかの疑問に応答する形でこの葛藤型超義務の内実を掘り下げ、それが超義務をめぐる私たちの選択を倫理的問題として適切に把握することを助けるものであることを示した（5節）。最後に、亜義務についても超義務と同様に、複数の道徳的理由のあいだで葛藤が生じるために道徳的に望ましくなくても単純に禁止することができない「葛藤型亜義務」がありうることを論じた（6節）。

以上の議論は、私たちが日常的に出会う「超義務」を理解可能な形で記述する、という哲学的課題について、一定の貢献をなすものと考えられる。パラドックスを内包し、理解に困難を抱えていた「超義務」は、いまや英雄型超義務、慈恵型超義務、および葛藤型超義務のいずれかとして記述すること可能となった。他方で、「超義務」とはそもそもいかなる道徳的性質であるのかを包括的に把握する、というさらなる探究目標については、なお道半ばである。言い換えれば、三つの超義務のカテゴリーが相互にどのような関係にあるのかを理解することで初めて、超義務とは何かという問いに十分に答えることになると思われるが、そのような理解はまだ得られていない。

本稿の議論をふまえて考えるならば、超義務を英雄型超義務と慈恵型超義務に区分する浦野らの枠組みは、参照すべき道徳的理由が単一である場面を暗黙のうちに前提している、と整理することができるかもしれない。そのような場面における超義務は、英雄型あるいは慈恵型の超義務としてもれなく説明することが可能であり、その範囲では浦野らの枠組みは十分に妥当である。しかしこの前提は常に成り立つわけではない。参照すべき道徳的理由が複数ある場面においては、浦野らの枠組みでは捉えきれない超義務が生じうる。それが本稿の提示した葛藤型超義務である。とはいえ、複数の道徳的理由のあいだの葛藤がある場面で、英雄型超義務や慈恵型超義務の観点が説明力を無くすとも考えづらい。三つの超義務カテゴリーの相互関係については、なお詳細に論じるべき余地が残されている。これは今後の重要な、そして魅力的な検討課題であると言えるだろう。

謝　辞

本稿はまず、その初期草稿について、富山大学の児島博紀氏から貴重な指摘を受けた。そののち、浦野敬介氏・石田柊氏と議論を行った上で、オンライン政治理論研究会のワークショップ「超義務の倫理学・再訪」（2024年9月15日開催）にて本稿を報告し、さらなる改善の機会を得た。児島氏、浦野氏、石田氏、上記ワークショップへの参加者各位、および『応用倫理』の匿名査読者に、この場を借りて感謝申し上げたい。

研究ノート

超義務と道徳的葛藤の関係についてのさらなる考察——コメンタリー「道徳的理由のあいだの葛藤という観点から超義務を考える」へのリプライ

浦野敬介（東京大学）、石田柊（広島大学）

要旨

この研究ノートは、浦野敬介と石田柊による論文「しないに越したことはない——超義務と亜義務の倫理学」に対する玉手慎太郎のコメンタリー「道徳的理由のあいだの葛藤という観点から超義務を考える」への応答である。玉手は、「望ましいが義務的でない行為」のうち慈恵型超義務・英雄型超義務のどちらにも該当しない事例を挙げ、浦野らの分類を補完する第三の超義務クラスとして「葛藤型超義務」を提案する。ある行為が葛藤型超義務にあたるのは、その行為が、それ自体は望ましいにもかかわらず、他の事柄を考慮した結果すべてを考慮して義務的ではない場合である。我々は、このような道徳的葛藤が「望ましいが義務的でない行為」と深い関係にあることを否定しないが、それを超義務の一類型として扱うことには懐疑的である。第一に、複数の慈恵型超義務のあいだで道徳的葛藤が生じることがあり、「葛藤型超義務」を加えた超義務の三分類は成立しない。第二に、道徳的葛藤は義務についても生じる一般的な現象であり、超義務に特有ではない。我々は、超義務と道徳的葛藤を、それぞれ異なる形で我々の規範的経験と道徳的選択を複雑化する、独立した規範的現象だと考える。

Abstract

In this article, we respond to Shintaro Tamate's commentary on the discussion by Keisuke Urano and Shu Ishida. Tamate points out that their taxonomy may not exhaust the entire spectrum of "desirable yet non-obligatory" actions, suggesting the need for another class of supererogation—i.e., conflictive supererogation. Roughly put, an action is conflictive-supererogatory when it is desirable in itself and yet non-mandatory, all things considered, in light of other moral considerations. It is caused by a conflict between two (or more) moral reasons, each of which alone is not decisive. True, this sort of moral conflict might comprise a part of the domain of "desirable yet non-obligatory" actions. However, even setting aside potential disputes on whether the situation described by Tamate should be genuinely outside the scope of heroic or benevolent supererogation, we are sceptical of seeing it as a novel class of supererogation. First, such a tripartite classification of supererogation would

face a taxonomical inconsistency when, for instance, we should choose between two small benevolences. Second, and more importantly, moral conflict is a quite general normative experience, not unique to supererogation; we are familiar with cases in which moral conflicts complicate the call of obligations, such as moral dilemmas. Having these observations in mind, we finally submit a revised picture: supererogation—whether heroic or benevolent—and moral conflicts are two distinct normative phenomena that can often render a desirable action non-obligatory, independently complicating our normative experiences and moral choices.

1. 序論

この研究ノートは、玉手慎太郎のコメンタリー「道徳的理由のあいだの葛藤という観点から超義務を考える」へのリプライである。玉手は、「望ましいが義務ではない行為」や「望ましくないが禁止されない行為」にかかわる我々の日常的経験を参照しながら、超義務・亜義務をめぐる浦野敬介・石田柊の議論（浦野・石田 2024）を詳細に検討している。そして、浦野らの議論がカバーできていない重要な現象として、玉手は「葛藤型超義務」および「葛藤型亜義務」という類型を提案している。このリプライの目的は、玉手のこうした検討を再検証することと、それを通して超義務・亜義務をめぐる議論の地平をさらに広げることである。

第2節では、玉手のコメンタリーの論旨を、以後の議論に関連する限りで簡単に確認する。第3節から第6節では、玉手の議論に対して四つの点から応答を試みる。第7節は結論と今後の展望である。

2. コメンタリーの論旨の確認

玉手のコメンタリーには、浦野らの議論のさまざまな点への言及があるが、コメンタリーの主要な部分を占めるのは、道徳的葛藤という現象に着目して超義務・亜義務の新類型を提案するというものである。その点をビビッドにするため、以下の事例を考えたい。

> **〈熱心な卒論指導〉**大学教員であるAは、学生の卒論指導の準備をしている。Aのゼミ生はみな真面目であり、教員が卒論指導にエフォートを割けば割くほど高い学習効果が得られる（Aもそれを知っている）。また、これまでAのゼミ生は就職後にさまざまなセクターで働き、卒論を通して身につけた思考能力等を社会に還元している（Aもそれを知っている）。Aは、熱心な卒論指導の経験が豊富であり、また今年度の学務負担は過大でない。そこでAは、今年度のゼミ生10人それぞれのテーマの関連文献を取り寄せて読み、学生の原稿を毎週添削し、いつでも面談に応じている[1]。

玉手によれば、〈熱心な卒論指導〉におけるAのこうした行為は、超義務的である。このような指導が道徳的に望ましいと考えることは不自然ではないが、このような指導が教員として義務であるとは考えに

1　玉手のコメンタリ第4節の事例をもとに再構成した。

くい。それゆえ、A のこうした行為は「望ましいが義務ではない」行為、つまり超義務的行為である[2]。

浦野らは超義務を慈恵型超義務と英雄型超義務に分類したのだが、A のこの行為はどちらにも分類できない。一方では、A の献身的指導が生み出す価値は決して瑣末なものではないから、この行為は慈恵型超義務ではない。他方で、A にとって献身的な指導は選択肢集合に入らない行為――いわば「考えもつかない行為」――ではないから、この行為は英雄型超義務でもない。したがって、もし A のこうした行為を超義務として適切に記述するためには、慈恵型でも英雄型でもない第三のクラスが必要になるのではないか――これが玉手の議論の出発点である。

そこで玉手が提案するのが、**葛藤型超義務（conflictive supererogation）**という類型である。ある超義務的行為 ϕ が葛藤型超義務的行為であるのは、ϕ をする道徳的理由と ϕ をしない道徳的理由が互いに競合し、どちらの理由も決定的でないため、ϕ をすることが望ましいにもかかわらず義務的でもなければ禁止もされないことをいう[3]。先の事例で考えよう。A には、たしかに熱心に学生を指導する道徳的理由があるかもしれない。しかし、A には、研究者として研究を進める道徳的理由もあるし、家族など親しい人のために時間を使う道徳的理由もある[4]。A が学生指導にエフォートを割けば割くほど、自分自身の研究時間や家族との時間は少なくなることを考えれば、後者の道徳的理由は、A が熱心な学生指導をしない道徳的理由となる。そして、この状況ではどの道徳的理由も決定的でないため[5]、熱心な学生指導は、その望ましさにもかかわらず義務的だとはいえない（もちろん禁止されるわけでもない）。これが、玉手が提案する葛藤型超義務という類型の概要である。

同じことが亜義務にも言える。その最も極端な例の一つとして玉手が挙げるのが、致死的状況において患者の宗教的信条にもとづき輸血を差し控えることである。輸血の差し控えは、十分に利用可能な救命手段があるにもかかわらずそれを使わない点で、道徳的に望ましくないと考えても不自然ではない反面、このような選択は現代の多くの社会で禁止されていない。そのため、この行為は「望ましくないが禁止されない行為」、すなわち亜義務的行為である。けれども、一方で、この行為は患者の死につながり、その負価値は決して軽微ではないため、この行為が非行型亜義務だとは考えにくい。他方で、この行為は医療従事者によって実際に選択されているため、極悪型亜義務だとも考えにくい。そのため、浦野らの亜義務の二分類に加えて、ここでも玉手は第三のクラスとして**葛藤型亜義務（conflictive suberogation）**という類型を提案する。葛藤型超義務と同様に、葛藤型亜義務も以下のように理解できる――ある亜義務的行為 ψ が葛藤型亜義務的行為であるのは、ψ をしない道徳的理由と ψ をする道徳的理由が互いに競合し、どちらの理由も決定的でないため、ψ をすることが望ましくないにもかかわらず禁止されない（もちろん義務的でもない）ことをいう。

以上が、玉手の提案の概要である。玉手の提案の中心にあるのは、複数の道徳的理由が互いに競合し、そのどれも決定的ではないという状況をどのように記述するのが適切かという問題である。簡潔のため、この状況を一般に**道徳的葛藤（moral conflict）**と呼ぶことにする[6]。

2　玉手が明示しているように、A の行為は、より厳密な超義務の定義（浦野・石田 2024, 18; cf. Hansson 2015）も満たす。

3　玉手がはっきり述べているように、これは総合的免責説（浦野・石田 2024, 21–22）とは異なる。総合的免責説は、道徳的理由と非道徳的理由との競合により道徳的義務が生じなくなっているとする立場である。これに対して、ここで検討しているのは、道徳的理由と別の道徳的理由との競合により道徳的義務が生じなくなっているとする立場である。それゆえ、超義務をめぐる議論は道徳というドメインの内部で完結するはずだとする浦野らの想定を葛藤型超義務は満たす（玉手 2025, §5.1）。

4　これらが果たして道徳的理由であるかどうかには、疑いがあるかもしれない。以下では、議論を簡潔にするのためそのように想定するが、適宜ほかの（より道徳的理由だと思われる）事例に置換しても議論の本筋は損なわれないと考える。

5　6.1 節の議論をみよ。

6　これを道徳的ジレンマと明確に区別してほしい（玉手 2025, §4, note 13）。道徳的ジレンマとは、複数の道徳的理由が互いに競合し、そのどちらもが決定的であるという状況である。この状況では、どの義務にしたがって行為しても必ず義務違反が生じてしまう。第 6 節で詳述する。

ここからは、玉手の議論の論点を検討することで、浦野らの議論、ひいては超義務・亜義務をめぐる議論全般の理解をより深めていきたい。

3. 行為者相対的な超義務理解のさらなる含意——〈熱心な卒論指導〉の再検討

第一の論点として、玉手が葛藤型超義務を提案する直観的根拠となった〈熱心な卒論指導〉事例を再検討したい。玉手は、大学教員 A の行為が超義務的であること、およびその行為が慈恵型超義務・英雄型超義務のいずれにもあてはまらないことを根拠に、第三の超義務タイプを提案した。この第三の超義務タイプ（葛藤型超義務）の理解を深める前に、まず玉手の前提を検討したい。A の行為は本当に超義務的だろうか？

重要な点として、浦野らが提案する枠組みでは、ある行為が超義務的であるかどうかは行為者相対的に決まる[7]（浦野・石田 2024, 18n8）。この点をみるために、まずは典型的な英雄型超義務の事例を考えよう。

> **〈人命救助〉** B は消防団員ではないが、日頃から体を鍛えている。ある日、B は、通りすがりのビル火災に遭遇し、中に一人の幼児が取り残されていることを知った。B は、危険を承知でビルに飛び込んだ。のちに B は、これは自分の義務だったと語った。

仮に、B 自身の証言を尊重して、B にとってこの行為は義務的だとしよう（cf. 浦野・石田 2024, 25）。たとえそうだとしても、B の証言にかかわらず、B のこうした行為は我々にとっては英雄型超義務である——実際に、我々の多くにとって、ビル火災に直面しても飛び込むことなど考えもよらない。行為者相対的な超義務理解は、この行為が我々にとっては超義務的だが B にとっては義務的だという主張を許容する。

補足として、この主張は独善的に聞こえるかもしれないが、この懸念は行為の義務論的地位（義務的か禁止されるかなど）を称賛から切り離すことで緩和されるように思われる。たとえば、たとえ〈人命救助〉事例において B の行為が B にとって義務的だとしても、それを超義務的だと思う我々が B を称賛することは一切妨げられず、それどころか称賛すべきだろう（浦野・石田 2024, 26n30）。

これと同じことが〈熱心な卒論指導〉事例でも生じている。一方では、我々の大多数にとって、授業や学務をこなしながら 10 人もの学部生の卒論に懇切丁寧に向き合うというのは並外れたことであり、（少なくとも著者らにとっては）想像も及ばないことである。そのため、A の行為は、我々の大多数にとって英雄型超義務である。他方で、この主張は、当該行為が A にとっても同様に英雄型超義務だということを含意しない。たとえば、A はこうした学生指導を「教員として当然だ」と思っているかもしれないし、そこまででないとしても、少なくともこうした学生指導を自分の行為選択肢の一つとして考えることができているかもしれない。その場合、A のこうした行為は、A にとっては超義務的ではなく義務的だということが考えられる。繰り返すが、このことは、A のそうした献身的な学生指導を我々が称賛することを妨げず、それどころか称賛するべきだろう。

この懸念を一層和らげるため、以下の事例をさらに考えよ。

7　このような行為者相対的な超義務理解を検討する他の文献として、たとえば Flesher (2003), Carbonell (2016), Naumann (2017) などをみよ。

〈人命救助 2〉 上記の〈人命救助〉のあと、B は、特段の後遺症もなく生活を送り、体を鍛え続けていた。その 1 年後のある日、B は、再び同じようなビル火災に遭遇し、また幼児が取り残されていることを知った。しかし、今回は B は飛び込むことには思いもよらず、ただ消防隊の到着を待った。

すなわち、過去に B は英雄的に振る舞ったかもしれないが、現在の B にはそのようなことは思いもよらない。この場合、重要な点で同じである行為（燃えるビルに飛び込んで幼児を救うという行為）が、ある時点の B にとっては十分可能な行為であり、それゆえ義務的である一方で、別のある時点の B にとってはもはや十分可能な行為ではなくなっており、それゆえ英雄的超義務的だということになる。超義務理解の行為者相対性――より一般に文脈依存性（浦野・石田 2024, 18n9）――を真剣に受け止めるなら、この見方は受け入れるべきだろう。また直観的にも、このように行為の義務論的性質が通時的に変化することはそれほど重大な問題を含むとは思われない。ある人が過去に（大多数の人から見て）英雄的に振る舞ったからといって、その人の残りの人生がずっと（大多数の人から見て）英雄的でなければならないというのは、あまりに過剰な期待だろう。

これと同じことが〈熱心な卒論指導〉にもあてはまる。仮に、〈熱心な卒論指導〉から 1 年後、A はもうそれほど熱心な指導は考えられなくなってしまったとしよう。この場合、1 年前に A がどうしていたかとは無関係に、献身的な学生指導は 1 年後の A にとってもはや義務的行為ではなく、英雄型超義務的行為に変わったといえる。A が一度でも熱心な卒論指導をしたからといって、残りの教員人生全体において同じような自己犠牲をもって学生に望まなければならないというのは過剰な期待だろう。

このように、〈熱心な卒論指導〉事例は、行為の超義務性を行為者相対的に――より精確にいえば文脈相対的に――理解することで、当初ほどパラドキシカルな事例ではなくなるように思われる[8]。

4. 日常における超義務・亜義務の存在感

第二の論点として、玉手が着目する道徳的葛藤は、日常的にありふれた現象であり、その分だけ葛藤型超義務もまた馴染み深い規範的経験である[9]。それに比べて、玉手によれば、浦野らが考える意味での超義務は「日常においては問題にならないか問題にするまでもないかのいずれかである」（玉手 2025, §5）。なぜなら、慈恵型超義務は極めて小さな道徳的価値しか生み出さず、英雄型超義務は問題となる行為者の選択肢にそもそも入らないからである。もしそうだとすれば、慈恵型超義務や英雄型超義務より、葛藤型超義務のほうが、超義務をめぐる我々の日常的経験にとって重要な現象をより如実に捉えているのではないか――玉手はこのように主張している。

この主張は、以下の形で、浦野らの議論に対する帰謬法として再構成できる。

1．すべての超義務的行為は、慈恵型超義務か英雄型超義務のどちらかにあてはまる。
2．慈恵型超義務的行為は、追加の道徳的価値に乏しい。
3．ある行為の追加の道徳的価値が乏しいならば、当該行為をするか否かは日常的に直面する倫理的問題とならない。
4．英雄型超義務的行為は、問題となる行為者の選択肢に入らない。
5．ある行為者にとってある行為が選択肢に入らないならば、当該行為をするか否かは当該行為者に

8　このほか、A の献身的な学生指導を超義務ではなく「さしあたりの義務」という別の枠組みで理解することも可能かもしれない。6.1 節をみよ。
9　以下では超義務を主に取り上げるが、同じことは細部を変えれば（*mutatis mutandis*）亜義務にもあてはまる。

とって日常的に直面する倫理的問題とならない。

6．1－5より、いかなる超義務的行為も、それをするか否かは日常的に直面する倫理的問題とならない。

7．超義務は、日常的に直面する倫理的問題に対応した規範概念である。

8．6と7は対立する。

上の議論のうち、前提1・前提2・前提4はいずれも浦野らの提案である。小結論6は妥当だと思われる。また、浦野らはたしかに超義務を日常的・実践的に重要な規範概念として理解しているはずなので（浦野・石田 2024, 16–17）、前提7も認めなければならない。結論8もさしあたり妥当である。

残るのは前提3と前提5であり、いずれも「日常的に直面する倫理的問題」をどのように理解するかが争点となる。そこで筆者らは、玉手の見立てとは異なる形で、慈恵型超義務や英雄型亜義務が日常的な倫理的問題を生じさせると主張したい。

慈恵型超義務から考えよう。少額の募金は、たとえ望ましくても義務的ではないということについては、一定の社会的コンセンサスが得られているとする。このとき、誰かが募金箱に気付き、無理なく募金できるにもかかわらず、それを無視して通り過ぎたとしても、そのことは道徳的に正当化可能である。しかし、そうして「正当に」募金箱を無視した際には、我々はしばしば、本当は募金したほうがよかったのかもしれないと考え、自分の情けなさや道徳的未完成性を痛感する。

同じことは英雄型超義務にも言える。自らの命をかけた人命救助は、たとえ望ましくても義務的ではないということについて、一定の社会的コンセンサスが得られているとする。このとき、誰かが通りすがりにビル火災を目撃し、中から助けを求める声が聞こえたにもかかわらず、それから目を逸らしてただ消防隊の到着を待ったとしても、そのことは道徳的に正当化可能である。しかし、そうして「正当に」悲痛な声を無視した際には、本当は助けに向かったほうがよかったのかもしれないと考え、自分の情けなさや道徳的未完成性を痛感する[10]。

慈恵型超義務や英雄型超義務は、このような良心の呵責という形で、我々の日常的な規範的経験の一部を形作っている。たしかに、良心の呵責は、道徳的葛藤とは異なり、我々の行為選択に直接的な影響を与えるわけではないだろう（我々の大多数は、ある募金箱を無視して良心の呵責を感じたからといって、別の募金箱を見た時にも同じように無視し続けるだろう[11]）。しかし、行為選択に直接的な影響を与えるものだけが我々の日常的な規範的経験だという主張は極めて強い主張であり、さらなる根拠を要するように思われる[12]。それゆえ、より穏当な立場として、日常における顕著な規範的経験には行為選択を直接的に左右しないものも含まれることを認めるとすれば、玉手の主張とは異なり、慈恵型超義務や英雄型超義務が日常的にレリヴァンスを欠くことにはならない。

もちろん、行為選択を直接的に左右するような規範的経験は、そうでない経験に比べて、重要性がより大きいかもしれない。その意味で、慈恵型・英雄型超義務が惹起する良心の呵責よりも、葛藤型超義務の

10　関連して、誰かが英雄的な人命救助をしないこと自体は非難に値しないかもしれないが、その人が、自分が英雄的な人命救助をしなかったことについて全く遺憾に思っていないとすれば、その態度は非難に値するかもしれない（浦野・石田 2024, 21n17）。

11　そもそも良心の呵責すら感じない人、つまり「禁止されていなければ何をやってもいい」と心の底から思ってそのように行為する人については、筆者らの指摘はあたらないかもしれない（この点は玉手との議論により明白になった）。筆者らは、そのような人は典型的ではないと想定し、議論を単純にするためここでは一旦わきにおくが、そのような人にとって慈恵型・英雄型超義務が普段の規範的経験としてどうでもいいという可能性は否定しない。重要なのは、良心の呵責を感じる我々の大多数にとって慈恵型・英雄型超義務が普段の規範的経験として生じるということである。

12　ここで取り上げる良心の呵責のほか、行為選択に直接の影響を与えない規範的経験の例として、ウィリアムズが詳しく論じる後悔（regret）がある（Williams 1965）。

背景となっている道徳的葛藤のほうがより重要だと考える人がいるかもしれない。筆者らは、この考えを否定するものではない。ただし、仮にそうだとしても、道徳的葛藤という現象を超義務の第三の類型として考えるべきかどうかは別の問題として残る。とりわけ、我々の日常的な理解によれば、超義務という現象の中心をなすのは道徳的理由（行為を選ぶ上でどの考慮事項を加味するのが望ましいか）と道徳的義務（どの行為を選ぶべきか）のずれである[13]のに対して、道徳的葛藤の中心をなすのは、理由と義務のずれではなく、複数の道徳的理由どうしの競合である。実際に、道徳的葛藤は、超義務だけでなく通常の義務においても生じる一般的現象である（6節をみよ）。この点は、道徳的葛藤という現象を超義務の一種とみなすことに対して、決定的な反論ではないかもしれないが、少なくとも疑問を生じさせる。

5. 道徳的葛藤と超義務・亜義務の複合について

玉手が重視する道徳的葛藤という現象を、慈恵型超義務・英雄型超義務と並ぶ第三の超義務タイプとして考えることには、少なくとも二つの問題がある。一つは本節でみる分類上の問題である。

もし葛藤型超義務が第三の超義務タイプだとすれば、慈恵・英雄・葛藤の三種の超義務がそれぞれ相互排他的だとするのが一つの自然な考え方だ[14]。しかし、これは成立しない。第一に、複数の慈恵型超義務をめぐって道徳的葛藤が起きることがある。次の事例を考えてほしい。

> **〈ダブル募金箱〉**Cは、普段の生活費をキャッシュレスで決済しており、現金を持ち歩くことはない。ある日、Cは、コンビニで決済をする際、ポケットに一円玉があることに気づいた。レジの横には、二つの募金箱があり、ひとつは障害者の支援、もうひとつは性的マイノリティの支援を目的としている。どちらの寄付の運営団体も信頼でき、どちらの社会的集団も同程度に不正義を受けており、Cはそうしたことを知っている。

このとき、Cが障害者支援の募金に1円を寄付することと、Cが性的マイノリティ支援の募金に1円を寄付することは、どちらも慈恵型超義務にあたる。しかし、Cは現金を持ち歩かないCにとって、今たまたま手元にある一円玉を一方に寄付すれば他方には寄付できず、また一円玉を分割して寄付することもできない。そのため、Cは複数の道徳的理由の競合に直面しており、この意味で、Cが一方に寄付をすることは望ましいが決定的な義務とはならない。いわば、Cが一方の募金に1円を寄付するのが義務でないことは多重決定されている――追加の価値の瑣末さと、同程度に強い別の道徳的理由があることによって決定されている。

第二に、複数の英雄型超義務をめぐって道徳的葛藤が生じることもある。次の事例を考えてほしい。

> **〈ダブル人命救助〉**Dは日頃から体を鍛えているが、命をかけて人命を救うことなど考えもつかないような人生を送ってきた。ある日、Dが駅で電車を待っていると、二人の見知らぬ人（X, Y）が同時にプラットフォームのそれぞれ反対側から線路に落ちるのを目撃した。プラットフォームは広く、一人を救出してからプラットフォームの反対側でもう一人を救出するには時間がかかる。プラットフォーム上にはDのほかに人はおらず、緊急ボタン等もなく、両側の線路とも残り数十秒で快速電車

13　浦野らは、当初、これを道徳的価値と道徳的義務の対立として描いており（浦野・石田 2024, 19）、用語法にやや不統一があるが、ここでの論旨に照らして同じことである。

14　玉手自身は、必ずしもそのような形で葛藤型超義務を構想したわけではなく、「三つの超義務カテゴリーの相互関係については、なお詳細に論じるべき余地が残されている」（玉手 2025, §7）と簡単に言及するにとどめている。

が高速で通過する。

こちらの事例でも、D が X を救うことと、D が Y を救うことは、どちらも英雄型超義務にあたる。しかし、上記の仮定により、X と Y をどちらも救うことはできない。そのため、D は複数の道徳的理由の衝突に直面しており、この意味で、一方を救うことは望ましいが決定的な義務とはならない。いわば、慈恵型超義務の複合と同じように、D の行為が義務でないことは多重決定されている。

もちろん、道徳的葛藤と他の超義務クラスが複合することは、それだけで葛藤型超義務というクラスを認める立場への決定的反論となるわけではないかもしれない。ただし、その場合には、慈恵・英雄・葛藤という三つの超義務クラスを分類する観点が統一されないという理論的なコストが伴う[15]。そのため、道徳的葛藤を慈恵型超義務・英雄型超義務と並ぶ第三の超義務クラスとしてではなく、これらの超義務クラスに直交する別種の考慮事項として考えるほうが、さしあたり（*prima facie*）説得的であるように著者らには思われる。

6. 道徳的理由の複数性という問題——浦野・石田論文の射程を再検討する

道徳的葛藤を第三の超義務クラスとみなすことを躊躇すべき第二の根拠は、道徳的葛藤がそもそも超義務に限った現象ではないという実質的な問題である。この点を以下で検討する。

6. 1. 道徳的理由の競合という現象の一般性

玉手が適切に述べるように、超義務は、典型的には単一の道徳的理由が問題になる場面を想定して議論されてきた[16]。浦野らの議論もその例外ではない。これに対して、我々が日常的に行為を選ぶ際には、複数の道徳的理由が考慮対象となることが普通であり、それらの複数の道徳的理由が相反する行為をサポートするという形で競合することも決して稀ではない。それゆえ、超義務をより現実的な形で——我々の規範的経験に即した形で——理解するためには、こうした道徳的理由の複数性を反映させることが望ましい（玉手 2025, §7）。筆者らもこれに同意する。問題は、道徳的理由の複数性を、超義務に第三の類型を設けるという形で反映するのが最適かどうかである。

異なる複数の道徳的理由が競合するという現象は、超義務という概念を使わずに記述することが普通であるような状況でも生じる。その典型例が道徳的ジレンマである（cf. McConnell 2022）。たとえば、E の国が隣国から侵略を受け、その過程で E は兄弟を殺害されたとしよう。E は母親と二人暮らしであり、もし E が戦地に出れば母親を戦火のなか一人残すことになるとしよう。このとき、E には、戦地に向かって不当な侵略を阻止するべき（それを通して亡くなった兄弟を弔うべき）決定的な道徳的理由と、母親を危険に晒すことなく親孝行を続けるべき決定的な道徳的理由がある。両者は両立不可能であるため、E は、どちらの道徳的理由にしたがって行為したとしても、何らかの決定的な道徳的理由に背く——単純化を恐

15　このことは玉手も認めているように思われる（玉手 2025, §7）。

16　これに対して、単一の道徳的理由のみが問題になっているように見える事例であっても実際には複数の道徳的理由が介在しているという立場があるかもしれない。たとえば、少額の募金では、財の再分配をするべき道徳的理由——これは募金することを正当化する（justify）——に、わずかであれコストが生じることに関連した自己利益由来の道徳的理由——これは募金をしないことを要求する（require）——が競合しているのだという見立てがありうる。要求理由と正当化理由の区別については、たとえば Gert（2007）をみよ。この点に詳細に言及する必要性を示してくれた發田颯虎に感謝する。この点については、浦野らが道徳的理由不足説という形で部分的に検討している（浦野・石田 2024, 23–24）ほか、自己利益由来の理由を道徳的理由に「変換」して扱う点で、ここで対比して想定している道徳的葛藤の状況とはやや性質が異なることが指摘できるが、目下の論旨に直接関係しないため深入りしない。

れずに言えば、道徳的義務に違反する――ことになってしまう[17]。

複数の道徳的理由が競合する別の事例として、さしあたりの義務（*prima facie* obligation）同士が競合している状況が考えられる[18]。たとえば、F は約束の待ち合わせ時間に遅れまいと急いで走っていたところ、見知らぬ人（G）と衝突して軽傷を負わせてしまったとしよう。F には、約束を守るさしあたりの道徳的義務があり、また自分の過失の被害者に償いをするさしあたりの道徳的義務がある。仮に、F が G を病院に連れて行くなどすれば、F は約束の時間に間に合わないとしよう。このとき、どちらのさしあたりの道徳的義務が決定的になるか――F がどの行為を選ぶべきか――は、G の傷の程度や F の約束の重大性などによって決まると考えるのが自然だろう。G の傷がよほど軽く、約束の内容が（たとえば就活の最終面接など）あまりに重大であれば、F は簡単な謝罪をして直ちにその場を辞することがすべてを考慮して（all things considered）道徳的に許容されるかもしれない。他方で、G が簡単な応急処置を必要としており、約束の内容が（たとえば普段からよく会っている友人との飲み会など）些末なものであれば、F が G を助けることがすべてを考慮して道徳的に求められるかもしれない。道徳的ジレンマの状況とは異なり、どちらの「義務」も、他の考慮事項のいかんによっては、すべてを考慮して不充足が許容されうる。

興味深いことに、玉手が当初提起した〈熱心な卒論指導〉事例も、こうしたさしあたりの義務同士の競合として読み替えられる。典型的な日本の大学教員にとっては、自分の研究を進めること、大学職員として学務の一端を担うこと、また学生を指導することは、いずれもさしあたりの義務である[19]。これに加えて、家族がいるならば家族との時間を過ごすことや、市民として必要な役割を果たすことも、さしあたりの義務である[20]。これらのさしあたりの道徳的義務は、他の考慮事項のいかんによっては、すべてを考慮して不充足が許容されうる。たとえば、自分の子供が危篤だとの知らせが病院から入れば、学生指導や学務をわきにおいて病院に向かい家族と時間を過ごすことがすべてを考慮して正当化されるだろう。同様に、競争的資金プロジェクトがすべて今年度で終わる研究者にとって、次年度以後の競争的資金の申請が迫っている場合には、学生指導や家族との時間を後回しにすることもすべてを考慮して正当化されるかもしれない。

このように、我々は、複数の道徳的理由が競合するという現象を超義務とは（少なくとも一見したところ）無関係の文脈で記述することができる。また、複数の道徳的理由が競合し、それによって一見すると超義務が生じていると思われる状況も、超義務とは異なる記述が無理なく可能ではないかと思われる[21]。筆者らは、どちらの記述がより優れているかをここで論じたいわけではない。重要なことは、複数の道徳的理由の競合という現象が、決して超義務に特有のものではなく、それゆえ超義務のいち類型として扱うべきものではないように思われるということだ。

17　いわゆるサルトルの指導学生の事例を簡略化した。この事例では、異なる種類の複数の道徳的理由が競合しているが、同種の複数の道徳的理由が競合する道徳的ジレンマも考えられる。その一例は「ソフィーの選択」として知られる――ある母親が、子供二人とともにナチスの強制収容所に連行され、子供のうち一人だけを解放するように迫られるというものである。

18　以下の事例は、W・D・ロスによる「さしあたりの義務」をめぐる議論（Ross 2002, ch. 2）を参考にしているが、細部は筆者らによるものである。

19　これが本当に道徳的義務であるかという論点は、ここではわきにおく。註 4 もみよ。

20　ところで、仮に卒論指導事例において学生指導が超義務的なのであれば、それと競合する研究活動や家事も同じ理由で超義務的だということになる。この場合、「複数の超義務的行為が競合する」よりも「複数のさしあたり義務的な行為が競合する」といったほうが、状況の記述として自然ではないかと筆者らは考える。ただし、これは言葉遣いの問題という面が大きいので、深入りしない。

21　もちろん、道徳的理由同士の競合を超義務という概念のもとで記述してはならないと言いたいわけではない。とくに、サルトル事例について、E はどちらを選んでも（極めて難しい状況であえて選択をし、少なくとも一つの義務を履行したという点で）英雄型超義務的行為をしたことになると考える人がいるかもしれない。筆者らは、この可能性をここで完全に検討することはできない。ただし、仮にこれが英雄型超義務にあたるとしても、これは人命救助などの典型的な英雄型超義務とは性質の違ったものとなる。なぜなら、この「ジレンマにおける決断」は必ず一部の道徳的義務への違反を伴うのに対して、典型的な英雄型超義務では、いかなる意味でも義務に違反していない、それゆえただ望ましいだけの超義務的行為が想定可能だからである。

6. 2. エフォートバランス説の可能性と課題

これとは別の形で、複数の道徳的理由の競合によって行為が義務的でなくなるという現象を、我々の超義務の理解そのものに直接的に反映することはできるか。その一つの可能性として、以下ではエフォートバランスに着目した超義務理解を簡単に検討する。

我々が日常的に行為を選ぶ際には、複数の道徳的理由が考慮対象となることに加えて、そうした複数の道徳的理由すべてを**どのような配分で**考慮するかが問題となるのが典型的である。先の事例でいえば、Fは、たとえ約束を守ることが道徳的に最も重要だとしても、Gに対して最低限度の謝罪をするべきだろう。同様に、たとえGの救護が道徳的に最も重要だとしても、知人に一報を入れることで部分的にでも約束を守るべきだろう。一般に、複数の道徳的理由が（どれも決定的でない形で）競合しているときに、一部の道徳的理由のみを過度に考慮して他を蔑ろにすることは、通常は道徳的に最善だとはいえない。

同じことを、再び〈熱心な卒論指導〉事例で考えよう。最低限の義務以上に学生を指導することは、それだけを見れば道徳的に望ましいかもしれない。しかし、その度合いが行きすぎて、自らの研究を疎かにしたり、家庭を顧みなかったりすれば、（どれほど学生指導により莫大な道徳的価値を生み出していても）全体のエフォートバランスは決して望ましいとはいえない。

このとき、献身的な卒論指導は果たして超義務的だろうか。二つの考え方がある。第一に、ある行為が超義務的であるかどうかを、他の道徳的理由との兼ね合いを考慮せずに当該行為だけをみて決定できるとしてみよう。その場合、この過剰に自己奉仕的な学生指導は、他の点でどんなに犠牲を伴い、全体として望ましくないとしても、学生指導という点だけからみれば「望ましいが義務的ではない行為」であり、超義務的だといえるかもしれない。ただし、これは玉手の当初の問題提起には反する考え方だろう。玉手は、この種の献身的な学生指導を、他の道徳的理由があることで義務的ではなくなっている行為として理解しているからである。

そこで第二に、ある行為が超義務的であるかどうかを、他の道徳的理由との兼ね合いを考慮してはじめて決定できるとしてみよう。単純化すれば、先の事例において、ある学生指導が超義務的であるかどうかを考えるのではなく、ある「学生指導と他の道徳的考慮事項との考慮のバランス」が超義務的であるかどうかを考えるというわけである。単純化のために、教員ごとの状況に応じて「道徳的に義務的なエフォートバランス」が想定できるとしよう。仮に、過剰に自己奉仕的な学生指導によって研究や家事があまりに蔑ろにされるとすれば、そのようなエフォートバランスは「義務的なエフォートバランス」よりも望ましくなく、したがって超義務的ですらなくなる[22]。これに対して、他の考慮事項をそれほど蔑ろにしない――つまり、自らの研究や家事に適切に時間等を使った上で、さらに学生指導に尽力する――という形で「義務的なエフォートバランス」を逸脱するとすれば、そのようなエフォートバランスは「義務的なエフォートバランス」よりも望ましく、したがって超義務的だといえるかもしれない。この考え方――行為の超義務性のエフォートバランス説――は、複数の道徳的理由の競合によって一部の超義務を理解しようとする玉手の問題提起に即しており、一見すると有望だと思われるかもしれない。

ただし、エフォートバランス説には、行為の超義務性の判定において考慮すべきものの範囲が無際限に拡大してしまうという課題がある。先の事例では学生指導・研究・家事の三者のみを考慮したが、我々の日常的な行為選択において考慮すべき事項はもっと幅広い。それどころか、本当に「望ましいエフォートバランス」を突き詰めて考えるなら、理論上、世界のあらゆる事象を考慮しなければ行為の超義務性は判定できないことになる。こうした「手がかり欠如」（cluelessness）に訴えた批判は、伝統的には功利主義に代表される帰結主義的道徳理論に対して向けられ（Kagan 1998, 64; Lenman 2000）、近年ではとりわけ不確

22　浦野らは、超義務を「ある義務的行為よりも望ましい行為」として定式化している（浦野・石田 2024, 18）。

定性（uncertainty）のもとでの道徳理論が直面する問題として盛んに論じられている（Tarsney, Thomas, and MacAskill 2024, §3）。超義務のエフォートバランス説もこれと同じ問題に巻き込まれてしまう[23]。

6.3. 道徳的理由の競合と超義務——より穏当な考え方

「手がかり欠如」批判については、解決策を含めてさまざまな見方が提案されつつあり[24]、ここでエフォートバランス説の可能性を閉ざすことは独断的かもしれない。ここで筆者らが言いたいのは、行為の超義務性をさまざまな道徳的理由のあいだの「バランス」として捉えるという方針が、一見した魅力に反して、有名な道徳的問題に直面してしまうということだ。さらに、複数の道徳的理由が競合するという現象を超義務という枠組みの中で説明する戦略としてエフォートバランス説以外のものがあるという可能性を、ここで完全に否定するつもりはない。

他方で、より穏当でシンプルな考え方は、道徳的理由の複数性と超義務とを完全に分離し、それぞれが我々の規範的経験を異なるメカニズムで独立に複雑化しているのだと考えることだろう。これは、道徳的理由の競合が（超義務ではなく）義務においても一般に生じること（6.1 節）とも符合する。

論点をよりビビッドにするため、第5節でみた〈ダブル募金箱〉事例を再検討しよう[25]。この事例で、C は種類の異なる二つの選択に迫られている。第一の選択は、手元にある1円を募金するかどうかの選択——どちらに募金するかは問わず、ともかく募金するかどうかの選択——である。募金することは C にとって義務的ではなく、慈恵型超義務的である。これが義務的でないのは、C の募金が生み出す追加の道徳的価値が十分に小さく、義務を生むほどではないからだ——これが慈恵型超義務の特徴である。第二の選択は、仮に募金するとしてどちらに募金するかの選択である。一方に募金することは C にとって義務的ではない。これが義務的でないのは、その募金箱に募金する道徳的理由と、他方の募金箱に募金する道徳的理由が、同じ強さで競合しており、どちらも決定的でないからだ——これが道徳的葛藤という現象の特徴である。

浦野らの議論は、このうち前者のタイプの選択だけを扱い、（それまでの超義務をめぐる典型的な議論と同じように）単一の道徳的理由のみが問題となる状況を想定して超義務を論じていたことになる（玉手 2025, §7）。その意味で、浦野らは、「望ましいが義務的ではない行為」のすべてを汲み尽くしていたとは決して言えない。他方で、玉手も認めるように、浦野らの議論は、「道徳的理由が単一であるにもかかわらず望ましいが義務的ではない行為」の分析としては依然として一定の説得性があるかもしれない（玉手 2025, §7）。

そうだとすれば、次のようにまとめることが許されるだろう。玉手の指摘どおり、慈恵型超義務・英雄型超義務よりもずっと我々の日常にありふれた困難な規範的経験として、複数の道徳的理由の競合（道徳的葛藤）がある。しかし、道徳的理由がただ一つであっても——すなわち、道徳的理由同士の競合を無視できる「理想的な」状況でさえ——我々は別種の困難な規範的経験に直面することがある。その一つが超義務であり、浦野らの議論はこの領域に限定される[26]。誤解を恐れずに言えば、道徳的理由が競合する場合

23 義務論は「義務」等の大雑把な行為指針を用いて道徳を考えるため、本来、帰結主義とは違って義務論は「手がかり欠如」批判の影響をそれほど深刻に受けないと考える人がいるかもしれない。そうだとすると、超義務のエフォートバランス説は、「手がかり欠如」批判を帰結主義だけでなく義務論にも呼び込むことになり、義務論の相対的な理論的優位を失わせてしまうかもしれない（この点は玉手との議論により明確された）。この論点は本リプライの目的を大きく超える一般的な規範倫理学の問題であるため、これ以上は深入りしない。

24 「手がかり欠如」批判の大部分——すべてではない——が実際には問題ではないと論じる文献として、たとえば Greaves (2016) がある。

25 同じことが〈ダブル人命救助〉事例にもあてはまる。

26 言い換えれば、スヴェン・ウヴェ・ハンソンの議論をもとにして浦野らが提案した超義務の定義は、より限定的なものとなるよう改善が必要だということになる。なぜなら、玉手が指摘するように、道徳的葛藤によって生じる「望ましいが義務的でない行為」もまた浦野らの定義を満たしてしまう（玉手 2025, §4, note 12）からである。この点は玉手との議論により明らかにされた。

に規範的困難が生じるのは当然であるのに対して、超義務は、そのような競合を排除したとしてもなお残る規範的経験として顕著であり、この顕著な領域に特化した分析こそが浦野らの狙いであったと理解することができる。

7. 結論

この研究ノートの目的は、単に玉手に反論することでも、単に浦野らの見解を擁護するべく論陣を張ることでもなく、玉手が提起した論点を足がかりとして超義務・亜義務の理解を深めることである。とりわけ、超義務をめぐる理論的な検討は、英語圏はもとより日本でも未だ萌芽的状況にあるため、このような「批判と応答」を通して議論を深めることは、論争の当事者のみならず周囲の人々にとっても有益ではないかと筆者らは期待している。

繰り返しになるが、筆者らの結論はこうである。玉手が提起する道徳的葛藤は極めて重要な現象であるものの、それは超義務・亜義務とはやはり別種の現象であり、超義務・亜義務にそれぞれ第三のクラスを設けるには及ばない。

この結論は、玉手の懸念がまったく超義務・亜義務をめぐる議論にとって無関係だったことを意味しない。超義務・亜義務が、それと隣接するが異なる現象である道徳的葛藤とどのような関係にあるかは、浦野と石田の議論において欠けていた重要な考慮事項であり、玉手はこの点について一つの体系的な立場を示した。筆者らは、この点について玉手とは異なる立場を示したことになる。どちらの立場が説得的であるにせよ、玉手の問題提起は、浦野らの議論における不備をひとつ指摘し、超義務・亜義務をめぐる議論を発展させる突破口を開いたことになる。

もちろん、この結論に対してさらなる反論があるかもしれないし、また他の点でこの研究ノートが玉手の懸念に十分に応えられていない可能性も筆者らは否定しない。そして、応答の過程で、筆者らは多くの点を（議論のためとはいえ）未解決にしてきた。これらの論点を一つ一つ詳細に検討することは、この研究ノートの目的を超えるため、超義務と亜義務をめぐる今後の倫理学的検討の課題としたい。

謝　辞

本研究は、大阪大学社会技術共創研究センター 2022 年度 ELSI 共創プロジェクト「企業の ELSI 対応の一部としての自主規制とそれに伴う萎縮効果に関する理論的研究」（研究代表者：長門裕介）および JSPS 科研費 JP23K11997（研究代表者：石田柊）の支援を受けたものである。また、本リプライは、2024 年 9 月 15 日にオンライン政治理論研究会で開催されたワークショップ「超義務の倫理学・再訪」での草稿をもとにしている。コメンタリー執筆者である玉手慎太郎氏、本リプライの草稿に対して有益なコメントをくれた田畑真一、發田颯虎、松田新の各氏（五十音順）、上記研究会の参加者、および匿名査読者に心から感謝する。

参考文献

Carbonell, Vanessa. 2015.“Differential Demands: Moral Demandingness and Ought Implies Can.”In *The Limits of Moral Obligation*, edited by Marcel van Ackeren and Michael Kühler, 36–50. Routledge.

Flescher, Andrew Michael. 2003. *Heroes, Saints, and Ordinary Morality*. Georgetown University Press.

Gert, Joshua. 2007.“Normative Strength and the Balance of Reasons.”*The Philosophical Review* 116 (4): 533–62.

Greaves, Hilary. 2016.“Cluelessness.”*Proceedings of the Aristotelian Society* 116 (3): 311–39.

Hansson, Sven Ove. 2015. “Representing Supererogation.” *Journal of Logic and Computation* 25 (2): 443–51.

Kagan, Shelly. 1998. *Normative Ethics*. Routledge.

Lenman, James. 2000. “Consequentialism and Cluelessness.” *Philosophy & Public Affairs* 29 (4): 342–70.

McConnell, Terrance. 2022. “Moral Dilemmas.” In *The Stanford Encyclopedia of Philosophy*, Fall 2022 edition, edited by Edward N. Zalta and Uri Nodelman. Available at https://plato.stanford.edu/archives/fall2022/entries/moral-dilemmas/.

Naumann, Katharina. 2017. “Self-Perfection, Self-Knowledge, and the Supererogatory.” *Etica & Politica* 19 (1): 319–32.

Ross, W. D. 2002. *The Rights and the Good*, edited by Philip Stratton-Lake. Oxford University Press. First published 1930. Citations refer to the 2002 edition.

Tarsney, Christian, Teruji Thomas, and William MacAskill. 2024. “Moral Decision-Making Under Uncertainty.” In *The Stanford Encyclopedia of Philosophy*, Spring 2024 edition, edited by Edward N. Zalta and Uri Nodelman. Available at https://plato.stanford.edu/archives/spr2024/entries/moral-decision-uncertainty/.

Williams, B. A. O. 1965. “Ethical Consistency.” *Proceedings of the Aristotelian Society, Supplementary Volumes* 39 (1): 103–24.

浦野敬介・石田柊（2024）「しないに越したことはない ——超義務と亜義務の倫理学」『応用倫理』15 号：15–32 ページ。

玉手慎太郎（2025）「道徳的理由のあいだの葛藤という観点から超義務を考える ——浦野＆石田「しないに越したことはない：超義務と亜義務の倫理学」へのコメンタリー」『応用倫理』16 号：37–45 ページ。

「特集　道徳的地位の再考　—動物倫理とその先へ」

趣旨説明　稲荷森輝一

近年、情報技術の目覚ましい発展により、ChatGPT をはじめとする AI 技術は私たちの日常生活に急速に浸透してきました。AI が人間の意思決定を支援することは、もはや当たり前の光景になりつつあります。さらに、飲食店で活躍する配膳ロボットや、LOVOT に代表されるソーシャルロボットの登場によって、私たちが生活の中でロボットと接する機会も大きく増えました。

私たちの「自律的」な選択が AI に依存する一方で、現代の技術は、人間による即時的な制御を離れ、自律的にふるまうようになっています。今後は、こうした AI・ロボット技術の発展に加え、VR（仮想現実）技術などの進歩によって、人間と技術の境界はさらに曖昧になっていくことでしょう。

こうした近未来社会においては、人間と技術が存在者としてますます接近し、「自由意志をもった自律的な行為主体・責任主体」という従来の人間観が大きく揺らぐことが予想されます。すなわち、これまで当たり前のように前提とされてきた「人間に自由意志を認め、それに基づいて道徳的責任を帰属させる」という実践が、今後はうまく機能しなくなる可能性があるのです。

このような問題意識のもと、トヨタ財団特定課題「先端技術と共創する新たな人間社会」に採択された本プロジェクト「近未来社会における新しい自由意志・責任概念」では、既存の自由意志・道徳的責任概念の再検討と再構築を目指してきました。

本特集は、本プロジェクトの一環として 2024 年 3 月に開催された動物倫理ワークショップを踏まえた論文集です。「自由意志をもった責任主体」としての人間観を問い直すことは、すなわち、人間の道徳的地位を問い直すことであり、それは同時に、人間以外の動物や人工的な存在の道徳的地位を見つめ直すことにもつながります。ワークショップでは、倫理学者のみならず、市民団体を含む多様な立場の参加者と講演者の間で活発な議論が交わされました。

本特集には、当日ご登壇いただいた 3 名のスピーカーより、非ヒト動物の道徳的地位に関するご寄稿をいただいています。本特集を通じて、道徳的地位をめぐる議論のさらなる深化と発展を期待しています。

本特集「道徳的地位の再考 - 動物倫理とその先へ」に収録された論文・研究ノートは「トヨタ財団 先端技術と共生する新たな人間社会」（課題番号 D22-ST-0028: 近未来社会における新しい自由意志・責任概念：　代表者　稲荷森輝一）の助成を受けて開催された第 34 回応用倫理・応用哲学研究会 (応用倫理・応用哲学研究教育センター) 「道徳的地位の再考ー動物倫理とその先へ」（2024 年 3 月 22 日・23 日北海道大学にて開催) の発表内容にもとづいている。

論文

長期主義と動物擁護

長期主義的観点からは畜産動物の福祉向上は優先されないのか？

綿引周（東京大学）

要旨

本稿では、効果的利他主義（EA）コミュニティ内で近年支持を集めている「長期主義」から見た畜産動物の福祉向上の優先順位について検討する。EA は、科学的証拠に基づき単位資源あたりで為せる善を最大化することを目指す実践・研究プロジェクトであり、この運動内ではこれまで畜産動物の福祉向上は最重要課題の一つとされてきた。しかし、遠い未来の存在の生の改善を重要視する「長期主義」に基づいて畜産動物の問題の優先順位を低く見積もることを正当化する議論が存在し、長期主義の考え方に基づいて資源を配分する団体の意思決定に影響を及ぼしている。本稿は、長期主義的観点に立っても畜産動物の福祉向上を軽視する十分な理由はないと主張する。特に、工場畜産が将来的にも存続する可能性が少しでもあれば、その規模は期待値として非常に大きいという Yip Fai Tse の指摘を踏まえると、畜産動物の福祉向上を軽視するどのタイプの反論も成り立たないと指摘し、長期主義的な資源配分の是正を求める。

Abstract

This paper examines the prioritization of farmed animal welfare from the perspective of “longtermism,” which has gained increasing support within the effective altruism (EA) community. Effective altruism is a research and practical project aimed at maximizing good per unit of resource based on scientific evidence. In this context, improving farmed animal welfare has traditionally been considered one of the highest priorities. However, arguments have emerged that justify deprioritizing animal welfare based on longtermism, which emphasizes the improvement of future beings’ lives. These arguments have influenced the decision-making of organizations that allocate resources according to longtermist principles. This paper argues that even from a longtermist viewpoint, there is insufficient justification for deprioritizing farmed animal welfare. Specifically, drawing on Yip Fai Tse’s argument that, if industrial farming persists into the future, its scale in terms of expected value remains vast, it demonstrates that none of the arguments for deprioritizing animal welfare can hold. Thus, it calls for a correction in resource allocation by longtermist organizations in favor of farmed animal welfare.

はじめに

効果的利他主義（Effective Altruism, 以下 EA）とは科学的証拠に基づき、単位資源当りで為せる善を最大化することを目指す研究･実践プロジェクトである[1]。EA というプロジェクトに参加する人々を総称して「効果的利他主義コミュニティ」や「EA コミュニティ（EA community）」と呼ばれることがしばしばある。このコミュニティには例えば、オープン・フィランソロピー（Open Philanthropy: OP）[2] や効果的利他主義センター（Center For Effective Altruism: CEA）[3] などの慈善財団、その慈善団体から資金援助を受けている活動団体、そして各地で効果的利他主義という理念の普及のためにコミュニティ・ビルディングに取り組む団体[4]・学生団体などが含まれると言えるだろう。

ウィリアム・マッカスキルやトビー・オードが牽引してきたこの EA の運動は、畜産動物の福祉向上運動で重要な役割を果たしてきた。畜産動物は数が多く、個々が甚大な苦しみに耐えている。福利を向上させる比較的容易な手段も既に知られているが、それにもかかわらず見過ごされている問題である。そのため、EA コミュニティ内では、畜産動物の福祉向上は元来、最優先課題の一つと考えられてきた。

実際、EA コミュニティ内最大規模の財団、OP の主要な資金提供者であるダスティン・モスコヴィッツとカリ・ツナは元々、畜産動物の問題に関心があったわけではなかったが、EA の理念に賛同した結果、2015 年、財団内に畜産動物専門の部署を設け、顧問を雇うまでに至った（Thomas 2024）。以来、同財団は 500 億円以上をこの分野に投資し、現在、世界中の何千社もの企業が、自社のサプライチェーン内で使用する卵をすべてケージフリー鶏卵に切り替えると宣言する状況を作っている（Open Wing Alliance n.d.）。

このように EA の文脈では従来、畜産動物の福祉向上は重要視されるべき課題であるとされてきた。しかし近年、マッカスキルを中心に EA コミュニティ内では「長期主義」と呼ばれる見解が支持を集め、それに基づき助成金を配分する財団や、助成金配分に影響力を持つ団体が登場しているが、これらの財団の焦点課題や助成先リストには畜産動物の問題は含まれていない。従来的には最重要視されてきた問題がそうしたリストから外れるのは、長期主義的に畜産動物の福祉向上が優先課題にならないと考えられる理由が提案されているからである（Duda 2016; Hilton 2024）。

本稿では、その議論を検討し、長期主義を前提としても畜産動物の福祉向上の優先度を低く評価する説得的な根拠はないことを示す。この結論が正しければ、2024 年時点の長期主義的チャリティの資源配分は誤った想定に基づいており、是正が必要である。

第 1 節で長期主義の定義とそれを支持する議論、この語が登場した背景を紹介し、長期主義的に寄付金を分配する主要財団が動物擁護への助成を行っていないことを確認する。第 2 節で、長期主義から畜産動物の福祉向上が優先課題でないとする議論を紹介する。第 3 節でこれらの議論への反論を提示する。第 4 節では議論をまとめつつ、長期主義から導かれる具体的実践への含意にも触れる。

なお以下の議論では、Effective Altruism Forum という効果的利他主義センターが運営する投稿サイトや、各 EA 関連団体の投稿記事を参照する。それにより EA コミュニティ内で交わされている議論の紹介となることも目指したい。[5]

1　本稿では EA については詳しく説明しない。EA については（MacAskill 2019a; 竹下 and 清水 2024）を参照。MacAskill（2019a）には清水颯氏による邦訳が存在し EA Forum にも投稿されている（URL：https://forum.effectivealtruism.org/posts/4v45j5fuDfRru5kQ9/wiriamu-makkasukiru-no、最終アクセス日：2025 年 4 月 10 日）。

2　https://www.openphilanthropy.org/、最終アクセス日：2025 年 4 月 10 日

3　https://www.centreforeffectivealtruism.org/、最終アクセス日：2025 年 4 月 10 日

4　日本にも効果的利他主義ジャパン（https://www.eajapan.org/、最終アクセス日：2025 年 4 月 10 日）という団体が存在する。

5　もう一点、本稿の議論の性格についても注記しておく。本稿には長期的な未来についての予測に関する議論が含まれる。例えば長期的には何の介入もなしに工場畜産がなくなるという予測を取り上げて検討する。確かに長期的な未来に関する予測について確実なことは誰にも言えないだろう。しかし、そうした予測を検討することで意図しているのは、そうした予測が正しい可能性がないことを示すことではない。そうではなく、そうした予測を支える根拠を退けることで、その予測が正しい可能性が高いと信じる理由がないこと、それゆえそうした予測よりも、その反対の可能性 —— 畜産が長期的な未来にも存続する可能性 —— の方が（その可能性を支持する理由があるため）より高いことを示すことにある。

第 1 節　長期主義者からは畜産動物の福祉向上の優先度が低いと考えられている

長期主義は、遠い未来に生きる存在の生を改善することが重要な課題の一つである、という考え方である。本節では「長期主義」という用語の背景とウィリアム・マカスキルによる定義、それを支持する議論を確認し、長期主義が理論的考察に留まらず、特に動物擁護にどのような実践的影響を及ぼしているのかを見ていく。

マカスキルは 2019 年 7 月 26 日の EA フォーラムへの投稿で、「長期主義」は彼が 2017 年 10 月に提案した用語だと述べている（MacAskill 2019b）。それ以前は、「長期にわたる未来が可能な限りうまくいくようにすることに関心を持つ見解」を表す用語がなく、「存亡リスク」（existential risk、x リスク）という概念を使って「存亡リスクの削減に関心を持つ人々」といった言い方がされていた[6]。しかし、存亡リスクの削減と長期的な未来へ良い影響を与えることは必ずしも一致しないため、マカスキルは「長期主義」という用語を提案した。

彼は長期主義と強い長期主義を区別し、前述の記事の 2021 年 11 月の追記で「多くの議論の末、次の定義に落ち着いた」と述べている。

1．長期主義：長期的な未来に良い影響を与えることが我々の時代の重要な道徳的優先事項のひとつである。
2．強い長期主義：長期的な未来に良い影響を与えることは我々の時代の唯一の（the）重要な道徳的優先課題である。

マカスキル自身は以前、「長期主義」は 2 の強い長期主義を指すべきと考えており、今も強い長期主義を支持している（Greaves and MacAskill 2021-5）。しかし、大衆に受け入れられやすく、それで目的のほとんどを達成できるという理由から、「長期主義」という語は（1）を指すべきだと現在考えている（Greaves and MacAskill 2021-5）。

長期主義を支持する議論として、次の 3 つの前提がよく挙げられる（Fenwick 2023; Moorhouse n.d.）[7]。

（1）善悪の内在的価値は時間によって影響されない。（時間的中立性）
（2）長期的な未来には膨大な数の個体が存在するようになる。（長大な未来テーゼ）
（3）現在や近い未来の私たちは、長期的な未来を良くする行為をかなりの確率でとることができる。（因果的実効性）

（1）と（2）から、長期的な未来には現在や近い未来と比べて膨大な価値や反価値が存在する可能性があることが導かれる。（3）を合わせると、私たちはその膨大な価値を生んだり、反価値を回避したりする行動を選択し実行することができる。したがって、善を最大化したいと望み、その解決により実際に長期的な未来に良い影響を与える見込みの高い課題があれば、それが最優先課題の一つとなる[8]。

6　「存亡リスク」という概念の提唱者であるニック・ボストロムは存亡リスクを「不幸なアウトカムにより、地球を起源とする知的生命体を絶滅させるか、その可能性を永久的かつ大幅に縮小させることになるリスク」（Bostrom 2002）として定義している。この定義に従えば、存亡リスクは知的生命体が絶滅するリスクだけではなく、核戦争により人類の文明が産業革命以前の段階に退行してそれ以降、元の水準に回復することができなくなるとか、地球規模の独裁政権が人類を完全な支配下におくことにより、人類が実現しえた可能性が全て奪われてしまうリスクも含む。
7　なお長期主義を擁護する議論に関しては（O' Brien 2024）も参照。オブライエンの議論は（竹下 2024）でも紹介されている。
8　O' Brien はさらに「費用が高すぎない場合には他者を助け効果的に積極的な価値を促進すべきである」という善行原則を採用して短期的な代替策よりも効果的な行為で長期的な未来がうまくいくようにさせると合理的に期待できる行為を取るべきであるという結論（O' brien はこれを長期主義テーゼと呼んでいる）を導いている（O' Brien 2024）。

この長期主義の考え方は理論だけでなく、EA コミュニティ内で実践的な影響を及ぼしている。Michael Aird が EA フォーラムで公開している「長期主義／ X リスク課題に関連する団体のデータベース」によると[9]、2024 年 9 月 6 日時点で、助成事業を行っている／助成先の決定に影響力を持つ団体[10] で、その焦点課題に畜産動物の福祉を含む団体はない。そのほとんどが共通して「AI アライメント」問題を課題として掲げている[11]。

その中で、「人間以外と長期的な未来（Non-humans and the long-term future）」を課題に掲げる団体がある。Center on Long-Term Risk、Longview Philanthropy、Polaris Venture（旧：Center for Emerging Risk Reduction）の三団体である。しかしこれらも畜産動物の福祉を焦点課題に挙げてはいない。彼らが「人間以外」として念頭に置いているのは、主にデジタルな有感体の福祉（Digital sentience welfare）である（ときおり野生動物の福祉も言及される）[12]。例えば Ploralis Venture は「焦点課題」で次のように述べている。

> デジタル有感体の福祉
>
> AI システムが進歩するにつれ、有感体が開発され、道徳的配慮を必要とするようになるかもしれない。ニューヨーク大学「心・倫理・政策センター」（NYU Center for Mind, Ethics, and Policy）は、この分野の研究と提言を行っている助成先のひとつである。（https://polaris-ventures.org/focus-areas/）

このように、デジタルな有感体を課題として掲げる団体が現在取り組んでいるのは、主に研究助成や学問分野の確立である。長期主義的観点から人工的な有感体の出現が重要課題と考えられている理由は 3.4 節で取り上げる。

しかし、そもそもなぜ長期主義的には動物擁護が重要課題ではないと考えられているのだろうか。次節でその理由を見ていく。

第 2 節　長期主義において畜産動物の福祉向上の優先度が低いとされる理由

畜産動物は数が多く、個々が甚大な苦しみを耐えており、福利を向上させる比較的容易な手段も既に知られている。それにもかかわらず見過ごされている問題であるため、長期主義の立場に立たない限り、多くの EA 団体で畜産動物の福祉向上は最優先課題の一つとされている。しかし、この課題の優先度は長期主義的には異なる評価を受けている。本節では、その評価の根拠となる議論を確認する。

EA の視点から影響力のあるキャリアを特定し、キャリアアドバイスを提供する非営利団体 80,000 Hours は、長期主義的観点から畜産動物の問題の重要性を論じる記事で、この取り組みを優先課題とすることに反対しうる議論を挙げている（Duda 2016; Hilton 2024）。その中で、長期主義を前提とすると思われる議論が以下の 5 つである[13]。

9　https://airtable.com/appieglk236kjGD7x/shrl64Yz5Q8lidMAD/tblWmQnJoaxgYMiZG/viwjHEUBUGEvJVXUc?blocks=hide 、最終アクセス日：2025 年 4 月 10 日。なお EA Forum 上では「MichaelA」というハンドルネームで投稿しているが、Michael Aird が出演するポッドキャストの紹介記事（Idea n.d.）内で彼の名前に MichaelA のプロフィールへのリンクが貼られていることから同一人物であることが分かる。

10　「Makes/influences FUNDING decisions?」の項目が「Yes」となっている団体。

11　AI アライメント問題については EA の入門教材「EA Handbook」の第 6 章を参照。EA Handbook の邦訳は効果的利他主義ジャパンのウェブサイト（https://www.eajapan.org/handbook、最終アクセス日：2025 年 4 月 10 日）で閲覧可能である。

12　ただし 2025 年 4 月 10 日時点で Longview Philanthropy がデジタル有感体や（野生動物も含めて）動物に関連する助成や研究を行っている痕跡は見つけられなかった。

13　2024 年 9 月時点。2024 年 7 月以前、80,000 Hours のウェブサイトには Duda が執筆した記事が掲載されていたが、7 月に Hilton 執筆の現在のバージョンがアップロードされ、記事の内容が大幅に修正されている。本稿の目的にとって特に重要な変更は畜産動物の福祉向上を優先課題とすることに反対する議論が大幅に入れ替わっていることである。Duda の記事では本文（1）から（3）の論点が挙げられていたが Hilton が執筆した最新の記事では（4）と（5）が長期主義的観点から見た場合の反対論として挙げられている。

（1）工場畜産が近い将来に終わるとすれば、その問題の規模は将来世代に影響を与える課題と比べて小さい。
（2）動物保護の推進による長期的利益は、人類の福祉向上による利益よりも遥かに小さい。
（3）工場畜産への反対が道徳円の拡大（moral circle expansion）をもたらすという畜産動物の問題の重要性を支持する長期主義的議論は、思弁的で様々な反論に晒されている。
（4）将来登場するデジタル有感体（digital sentience）への影響の方が、畜産動物への影響よりも大きく、数も大きくなる可能性が高い。したがって我々はデジタル有感体の問題に取り組むべきである。
（5）長期的未来が膨大な価値を持つなら、存亡リスクの削減に取り組むべきである。

80,000 Hours が長期主義的観点から重要課題をリストアップする際（Fenwick 2023）、工場畜産の問題が挙がっていないのも、上記の考慮があるためだと考えられる[14]。

なお（5）の論点については Hilton 自身が詳細に検討し、畜産動物の問題の重要性を支持する結論を導いているため本稿でその議論を繰り返すことはしない。しかし他の論点には同一の記事内で再反論が展開されていないため、読者には畜産動物の問題が（短期的にはいくら重要だとしても）長期的には重要でないという印象のみ与えられることになる。特に EA コミュニティ内で影響力をもち、効果的利他主義に新しく触れる人々が最初に出会う確率の高い 80,000Hours のウェブサイトから読者がそのような（以下の議論によれば）誤った印象を受けてしまうのは動物たちにとって好ましくない帰結をもつだろう。そこで以下では、80,000Hours の記事内で再反論が提示されていない論点（1）から（4）の反論を取り上げて検討し、そうした反論が成り立たないことを示す。

第 3 節　長期主義的にも畜産動物の福祉向上が重要な理由

本節では、前節で取り上げた長期主義的に畜産動物の福祉向上が重要でないとされる（1）から（4）の理由を順に検討し、再反論を提示する。

3. 1　工場畜産の問題の規模は小さいのか？

最初の反論は「工場畜産が近い将来に終わるとすれば、その問題の規模は長期的な未来に影響を与える課題と比べて小さい」というものである。この反論は、今後、工場畜産が近い将来に終わると想定している。確かに畜産のエネルギー効率は非常に低く、飼料の大部分は排泄物や非可食部に変わる。人々の価値観が変化し、代替肉への抵抗感がなくなり、動物を育てるより効率的で安価な高品質代替肉が生産できるようになれば、畜産は終焉すると考えるのが妥当かもしれない（Hilton 2024）。

しかし、イップ・ファイ・ツェ（以下、ファイと呼ぶ）は EA フォーラムへの投稿「遠い未来の期待される畜産動物の数が膨大かもしれない理由」（Tse 2022）の中で、工場畜産が廃止されると単純に想定することはできない理由を 5 つ挙げている。その理由の一部を挙げるなら、第一に、工場畜産に反対する経済・

14　なお、上記（1）から（5）に対して、そもそも長期主義自体を受け入れないことも可能である。実際、長期主義は EA 者全体に受け入れられているわけではなく、多数派とも言えない。竹下 (2024) によれば EA が長期主義を含意することも、その逆も成り立たない。実際、EA 関連慈善団体の資金全体のうち、長期主義的観点から投入先が決められているのは一部であり、CEA が主催する EA 入門コース（https://www.effectivealtruism.org/virtual-programs/introductory-program、最終アクセス日：2025 年 4 月 10 日）の教材「EA Handbook」（脚注 11 参照）にも、「私がおそらく長期主義者でない理由」(DeniseMelchin 2021) という長期主義に反対する論考が含まれている。マッカスキルやニック・ベックステッドなどの EA の代表的な人物が長期主義を支持しているのは確かだが、EA 者全員が支持しているわけではない。
また、長期主義的観点から野生動物の苦しみを減らすことが最重要課題だとする議論もある（O' Brien 2024）。しかし本稿では畜産動物の問題に焦点を当てる。

環境的理由の多くは、牛や豚、鶏にしか当てはまらない[15]。第二に、工場畜産は食糧以外にも、飼料、衣料、顔料、医薬品、実験[16]、ペット、装飾品、臓器採取のために続く可能性がある。タンパク質生産の優れた代替手法が登場しても、工場畜産を続ける理由がなくなるわけではない。第三に、AI などのテクノロジーが工場畜産を効率化する可能性もある。現在、「精密畜産」や生物工学が注目されており、IoT 技術、ビッグデータ分析、ロボット工学、機械学習などの AI 関連技術を利用して畜産を効率化する取り組みが進んでいる。例えば、育種選択技術により、ブロイラー鶏の成長率は 1950 年代から倍増している。このプロセスに AI を組み込む動きもある。これらの技術により、培養肉などの代替技術が常に畜産より効率的であるとは限らなくなる。他にも、効率が劣るものが必ずしも淘汰されるとは限らないことや、工場畜産を宇宙に広げようとする動きもあるなど、工場畜産が遠い未来まで存続する理由がある[17]。

以上のファイの指摘は工場畜産が遠い未来まで存続する可能性がゼロではなく、むしろ数パーセントはあると信じる理由があることを示している。そして、工場畜産が遠い未来まで存続する可能性が僅かでもあるなら、未来の畜産動物の数の期待値は長期主義的観点から懸念すべきであるほど十分に大きい。ファイ（Tse 2022）は期待値に訴えることで、このことを示す議論も展開している。

最初に、期待値に訴える議論は 80,000 Hours でも用いられていることを確認しておく。彼らは「地球の寿命が尽きる 10 億年後まで文明が存続する確率が 5% あるとすれば、未来には 50 万以上の世代が存在する」と論じている。Global Priorities Institute の 2021 年の報告「未来はどのくらいの生命を抱えているのか？」を作成したトビー・ニューベリーは[18]、（1）人類が地球に留まる場合（地球シナリオ）、（2）太陽系内に留まる場合（太陽系シナリオ）、（3）銀河系内に進出する場合（銀河シナリオ）の 3 つのシナリオで人類の人口を推定している[19]。彼のゲスティメイトモデルによれば[20]、地球シナリオでの未来人口の期待値は約 $1.1 \cdot 10^{14}$、太陽系シナリオでは $1.4 \cdot 10^{21}$、銀河系シナリオでは $6.8 \cdot 10^{36}$ となる[21]。マカスキルも長期主義について論じた著書『見えない未来を変える「いま」』で同様に、期待値に訴えた議論を展開している（MacAskill 2024, pp. 23–24）。

15　牛の飼育は（羊やヤギとともに）カーボンフットプリントやウォーターフットプリントの値が高いが、豚、鶏、牛で確認されている排泄物による環境汚染リスク、人獣共通感染症のリスクは魚類や無脊椎動物の養殖では確認されていない。さらには魚類や無脊椎動物の飼料要求率（1 kg の肉・卵を得るのに投入する必要のある飼料の重量）は鶏の 1.7 から 2.0 を上回り、1 を下回ることもある（採卵鶏の場合、飼料要求率は飼料の重量を鶏卵の重量で割ったもの。魚の場合、飼料要求率が 1 を下回るのは、魚の体重を占める水分が多いため）。なお魚類が有感性をもつことを示す研究はその解説動画とともに（Chua n.d.）にまとまっている。EA コミュニティでは魚類が有感性を持つことはデフォルトの立場であり、養殖魚の福祉向上に取り組む団体 Fish Welfare Initiative（https://www.fishwelfareinitiative.org/、最終アクセス日：2025 年 4 月 10 日）がすでに活躍している。また無脊椎動物の有感性については注 17 を参照のこと。

16　動物実験には個別の論点が存在し現時点では工場畜産場とは必ずしも形容できない施設で大半の実験動物が飼育されているかもしれない。しかし将来的には実験動物を生産する工場畜産場が誕生する可能性がある。顔料や医薬品などに使われる動物に関しても同様に、将来、そうした動物を生産するための工場畜産場が大規模に展開する可能性はある。

17　日本でも、宇宙の環境下で飼育可能なコオロギの作出を目的とするムーンショット計画が進んでいる（“ムーンショット型農林水産研究開発事業” n.d.）。コオロギを含む直翅目の成虫が有感性をもつ（特に痛みを感じる）ことを示す 8 つの規準のうち 3 つを満たす確率が高いことを示す証拠が存在する。またその他の 4 つの規準については否定的な証拠があるわけではなく、単に研究が存在しない。それゆえ現時点ではコオロギも痛みを感じる可能性を支持する証拠の方が否定する証拠よりも多い（Gibbons et al. 2022）。またコオロギにとっても工場畜産は様々な福祉上のリスクをもたらす（Rowe 2020）。例えば飼育密度が高いと病気の感染率や共食いの発生率が高まるという報告がある（Luijben et al. 2012）。

18　「生命（life）」とは言うが、ニューベリーが推定しているのは人類かデジタルヒューマンの数である。

19　Cf. Newberry（2021）。 ニューベリーの報告では人類が天の川銀河を超えて他の影響可能な範囲の宇宙に広がるシナリオや、人格をもつ AI などのデジタル人間（digital person）が誕生するシナリオも考慮しているが、以下と同様の議論がこれらふたつのシナリオにも当てはまるためここでは取り上げない。

20　ゲスティメイトは、不確実な量をモデル化し予測するためのスプレッドシートツールである。各セルにある値からある値までの範囲（またはデータセット）の形式で区間を入力すると、その後の計算もこの不確実性を考慮に入れて進めてくれる。その際、ゲスティメイトはモンテカルロ法を利用している。例えばある区間で指定されたパラメータ X と Y をかけあわせる場合には、X からランダムにサンプルを取り出し、それを再びランダムに取り出された Y のサンプルを掛け合わせる。そのようなサンプルを 5000 個ずつ取り出してその結果の期待値と 90% 信頼区間を表示する。詳しくは Guesstimate Documentation（https://docs.getguesstimate.com/ 、最終アクセス日：2025 年 4 月 10 日）を参照。

21　https://www.getguesstimate.com/models/17470（最終アクセス日：2025 年 4 月 10 日）。（Newberry 2021; Tse 2022）も参照。

ファイは同じ型の議論を畜産動物に当てはめている。その議論の要点は以下の通りとなる。ある年Yにヒト一人がその平均寿命を生きる間に屠殺される動物の数（per human capita of farmed animals）をPHCFA（Y）とする。2017年のPHCFA、すなわちPHCFA（2017）は、脊椎動物で約1600匹、無脊椎動物で約12000匹である[22]（ただしここには、ミルクや卵、毛皮、革など、食肉用以外に使用される動物の数は含まれていない）。将来もPHCFAの値が変わらなければ、ニューベリーの各シナリオで人類の数の1600倍の脊椎動物と12000倍の無脊椎動物が屠殺されることになる。

もちろん、未来の肉の消費量が一定とは限らない。肉の消費量が減る時期もあるかもしれないが、その変動を平均化し、現在から人類が滅亡するまでの期間に食肉用に屠殺される動物の数を、その間に存在する人類の数で割った数を$PHCFA^{future}$として$PHCFA^{future}$/PHCFA（2017）をRとおく。ファイは上記の地球シナリオや太陽系シナリオなどの各シナリオで最も実現しそうなRの値に、その値が実現する確率を重み付けすることで、未来に飼育・屠殺される畜産動物の数の期待値を推定している。ただし単純化のためにファイは「Rの値に確率分布は存在しない」と仮定する。この仮定によりおそらくファイは（1）Rがある正の値を取る場合と、（2）Rが起こらずR＝0となる場合のふた通りしかないと考えている[23]。前者は工場式畜産がそのRが実現する規模で継続するケースに、後者はそれが廃止されるケースに対応するだろう[24]。この方法で、例えば、地球シナリオだけを考え、このシナリオで最も実現しそうなRの値は脊椎動物で1/2、無脊椎動物で1とし、それぞれのRの実現確率を30%、50%とすると、脊椎動物の数の期待値は2.6×10^{16}匹[25]、無脊椎動物で6.6×10^{17}匹となる。人類の人口との差はやはり現在と同様、10の二乗から三乗のオーダーに乗る。この推定によれば、肉の消費量が減る時期があっても、全体としては動物の数は人類の数を大きく上回るだろう。また、この値は食肉用に屠殺される動物の数しか含まず、他の目的で屠殺される畜産動物を加えれば、その値はさらに増加する。

以上の考察は、長期的に見ても畜産動物の問題の規模が極めて大きくなり得ることを示している。もちろん、この課題の重要性や規模を測るには、将来の畜産動物が経験する苦しみの深刻さも考慮する必要がある。動物福祉への配慮が進み、飼育方法や屠殺方法が改善され、苦しみが現在より大幅に減る可能性もある。精密畜産は生産性向上だけでなく、動物福祉向上の文脈でも導入されており、技術の発展は畜産の存続に寄与するだけでなく、動物の苦しみを減らすことにも貢献し得る。しかし、種差別が自明な態度とされる現状を踏まえると、畜産動物の福祉向上は今後も僅かな改善に留まる可能性もある。これは簡単に結論づけられる論点ではないが、いずれにせよ、第一の反論が前提とする「工場畜産が近い将来に終わる」という想定は、以上の考察に応答することなしに自明視できるものではない。

22　この数値はヒト一人が平均年齢を生きる間に、食肉用に屠殺される動物の数を現在の人類の数で割った数値であり、その数にはヴィーガンなど動物の肉を食べない人間の数も含まれる。したがってそうした人間たちの数が増え、それに応じて食肉消費が減るなら、この数値は減る。

23　実際、Rが厳密に0になるのは2017年以降、屠殺される畜産動物の数がゼロになる場合だが、ありえそうにない。むしろRの特定の値が実現される場合と対比されるべきは畜産が間も無く廃絶されるおかげでRが限りなく0に近づく場合であると理解すべきだろう。

24　ニューベリーのモデルを元に筆者が作成したゲスティメイトモデル（https://www.getguesstimate.com/models/24005、最終アクセス日：2025年4月10日）では、各動物種・各シナリオに期待されるRの90%信頼区間から、全シナリオを考慮した場合の畜産動物の数の期待値を求めることができる。

25　ファイの原文では1.0×10^{16}とあるがRの期待値の計算にミスがあった（現在は修正されている）。$1600 \times 0.5 \times 0.3 \times 1.1 \cdot 10^{14} = 2.64 \times 10^{16}$となる。この点は匿名の査読者の指摘に負う。なお、無脊椎動物では$12000 \times 0.5 \times 1.1 \cdot 10^{14} = 6.6 \times 10^{17}$となる。

3.2　動物擁護の長期的利益は人類の福祉向上による長期的利益よりも小さいのか？

80,000 Hours が挙げた第二の論点は、オーウェン・コットン＝バラットの論考（Cotton-Barratt 2014）に依拠している。彼は、動物福祉への介入が人間への介入よりも効果的だという議論に対し、費用対効果の推定自体には異議を唱えないが、両者の長期的インパクトの差が無視されていると指摘する。

コットン＝バラットによれば、人間の福祉向上は社会全体に、継続的にポジティブな影響をもたらす。教育や衛生的な生活へのアクセスがある健康な人々は生産性が高まり、経済発展に貢献する。その結果、後の世代に対して連鎖反応的な（knock-on）向上を ―― あるいはこの効果を指摘したホールデン・カルノフスキー（Karnofsky 2013）の用語を使えば「貫流効果（flow-through effect）」と EA コミュニティ内で呼ばれるものを ―― もたらす。したがって、人間の福祉への介入は直接的な効果だけでなく、長期的なインパクトも高い、とコットン＝バラットは論じる。

一方、動物への介入には、人間への介入と同様の貫流効果は見られない。彼によれば、動物の福祉向上が他の動物の福祉向上につながるメカニズムはなく、動物は学習や経済的発展が可能な社会を持たないからである。したがって、短期的に動物福祉への介入が人間への介入より効果的でも、長期的に同じことが言えるかは疑わしいと彼は論じる。

動物擁護の長期的な重要性を信じる場合、この議論にどう応答できるだろうか。EA コミュニティ内では、産業革命以降、主に工場畜産が原因で世界の正味の福利は悪化し続け、負の値になっているという推定もある（kyle_fish 2023）。このことから、これまでと同様、社会経済の規模の増大と比例して（あるいはそれ以上の速度で）工場畜産の規模も大きくなるなら、人類の長期的繁栄は良い結果をもたらさないと議論することも可能かもしれない。また、道徳円の拡大（moral circle expansion）を促す手段として、畜産動物の福祉向上への取り組みが長期主義的に優先されるべきだという議論も存在する（3.3 参照）。しかし、ここでは別の観点から反論を提示したい。

この反論は、苦しみに焦点を当てた倫理（suffering-focused ethics、以下 SFE）や優先主義（Prioritarianism）と呼ばれる倫理的見解に訴えるものである[26]。SFE は、ポジティブな経験を増やすよりもネガティブな経験を減らすことを優先する倫理的見解の総称である。他方、優先主義では、全ての個体の状況を改善すべきだが、特に恵まれない者（those who are worse off）を優先すべきとされる。これらの立場を受け入れ、かつ工場畜産場や養殖場で飼育される動物の数の期待値が膨大であるなら、最優先すべきは畜産動物の苦しみを減らすこととなる。

SFE や優先主義は、ネガティブな価値（反価値）とポジティブな価値の間の非対称性に基づき擁護される。例えば、苦しみに満ちた存在を生み出して多くの幸福な存在を生むべきではないという非対称性や、苦しみの原因は多岐にわたるが幸福になるにはそのすべてを回避しなければならないといった実現の難しさに関する非対称性、あるいは価値の不在は悪ではないが反価値の不在は善であるなどの非対称性である。これらの非対称性を前提として、苦しみの減少や回避が優先されるべきだと論じられる。

しかしここでは詳細な議論には踏み込まず、次の問いでこれらの見解を支える直観を示すにとどめたい。

> 人類が到達しうる最も理想的な世界を想像したとき、私たちは、人類が繁栄を極める一方で（おそらく人間には見えないところで）人間より遥かに多くの動物が毎年、多大な苦しみを受けて殺される「オメラス」のような世界[27]（ただし一人の子供の代わりに何兆もの動物を「地下」に幽閉し、数年・数ヶ

26　SFE 内部にも様々な立場があり、帰結主義や義務論、徳倫理など複数のタイプの倫理的見解と両立可能であるとされる。反出生主義も SFE の一種に数え入れられるが、SFE のどの見解も反出生主義にコミットする、というわけではない。

27　アーシュラ・K・ル＝グウィンの短編小説「オメラスから歩み去る人々」（Guin 1973）ではオメラスという架空の街が描写されている。理想郷ともいえるオメラスの住人は幸福な生を送っている。しかし、この街の人々の幸福はひとりの子どもが地下に幽閉され惨めな状態におかれることによって成り立っている（その仕組みは詳述されない）。

月単位で殺し続ける世界）に住みたいだろうか。それとも、人類全体の幸福はそれより少ないかもしれないが、今よりずっと幸福で、動物たちも苦しむことなく暮らす世界だろうか。

もし後者の世界を望むなら、貫流効果によりいくら人間の福利を向上する見込みがあるといっても、動物の経験する苦しみの量と深刻さを減らすことを優先すべきだとも考えられるだろう。

3. 3　工場畜産への反対が道徳的配慮の拡大をもたらすかどうかにかかわらず畜産動物の福祉向上は重要な課題である

動物の福祉向上を目指す介入策が長期主義的観点から重要視されない理由として、80,000 Hours が挙げた三番目の論点を検討する。それによれば、工場畜産への反対を長期主義的に支持する議論は、その運動が長期的に有益な道徳的配慮の拡大をもたらすというものだが、この議論は思弁的で、様々な心理学的・社会学的仮定に基づき、多くの反論に晒されているという。

この反論の背景を少し説明しよう。EA コミュニティでは、道徳円の拡大が長期主義的な優先課題であるという議論が度々提出されてきた（Anthis n.d.; Anthis and Paez 2021; Baumann 2017, 2020）[28]。さらに、短期的な畜産動物の福祉向上は、個人や集団の道徳円の拡大に寄与するため、長期的にも重要な課題であるという主張もある（Anthis n.d.; MichaelA n.d.）。現在検討している反論は、この議論に対するものである。

畜産廃絶への取り組みが道徳円の拡大に貢献するという議論のひとつの前提として、以下の想定がおかれる。（MichaelA n.d.）

> 前提：より多くの人の道徳円が畜産動物を含めるまでに拡張され、かつ／または工場畜産が廃絶されるなら、そのことは、未来の行為者の道徳円がすべての有感的な存在（あるいは少なくとも極めて大多数の存在）を含むようになる確率を上げるだろう。

マイケル・エド自身も、この前提はもっともらしいが、さらなる証拠が必要だと述べ（可能な研究課題もまとめている）、80,000 Hours が指摘する反論は、まさにこの前提を支持する研究が十分でなく、この前提が不確実であることを指摘している。80,000 Hours はハリソンのブログ記事（Harris 2021）を参照しており、この記事では、長期主義が正しい場合に道徳円の拡大に注力すべきだとしたら答えるべき 22 の論点が指摘されている。その中でもよく議論されるのは、価値観の変動や固定化に関連するもので、例えば反種差別的価値観を広めることに成功しても、それがどの程度持続するかといった問題がある。

これらの論点について経験的な研究を積み重ねることは、道徳円の拡大や道徳的アドボカシーのような長期主義的な動物擁護介入策に「因果的実効性」が成り立つかどうかを知るうえで重要だろう。

しかし、これらの経験的探究とは別に、この三番目の反論には別の指摘ができる。この反論は「畜産動物の福祉向上への取り組みが長期的に重要であるとしたら、その取り組みを通じて、またはその結果として、道徳円の拡大が生じる可能性が高い場合である」という前提を置いている。現在検討中の反論は、この条件文の後件を否定することで、畜産動物の福祉向上への取り組みの長期的な重要性を否定していると考えられる。

しかし、たとえ畜産動物の福祉向上によって道徳円の拡大が起こらなくても、畜産動物の福祉向上自体が長期主義的に重要であれば、上記の前提は成り立たない。そしてそのことこそ、これまで示してきたこ

28　Anthis and Paez（2021）は道徳円の拡大を目標に掲げる Sentience Institute の見解を代表している（Sentience Institute n.d.）。

とである。工場畜産が存続する可能性が僅かでもあるなら、将来存在するだろう畜産動物の数とその苦しみは膨大になる。そしてSFEや優先主義が説得的に思われるのであれば、より幸福な存在を生みだすことよりも、そうした膨大な苦しみを回避することのほうが優先されるべきである。したがって、工場畜産の終焉に向けた取り組みは、単に道徳円の拡大の手段としてではなく、それ自体、長期主義的に価値がある。

3.4　畜産動物よりもデジタル有感体の問題に取り組むべきなのか？

最後に、長期的には、将来登場するはずのデジタルな有感体の方が畜産動物の問題よりも重要だという反論を検討する。もしも有感性をもつデジタルな存在が将来、登場する場合、そのような存在はどのような酷い扱いを受けうるだろうか。デジタルな有感体が極度の苦しみを経験するシナリオとして、拷問的な実験の対象となることや奴隷化、シャットダウン・再起動への恐怖、パラメータの再設定によるアイデンティティ・クライシスなどが考えられている（詳しくはGabeM and nicholeta-k（2024）を参照）。また、デジタルな存在はコピーが容易で、その数を増やすのが有機体よりも簡単かもしれない。そのため、場合によっては畜産動物よりも規模の大きな問題となりうると論じられることがある（GabeM and nicoleta-k 2024）。

しかし、それでも現時点で、デジタル有感体の福祉向上が畜産動物の福祉向上より優先されるべきだとは考えられない理由がいくつか挙げられる。第一に、AIは動物と異なり食用にされず、そのために大量の数が集約的に、劣悪な環境で飼育される可能性が（ゼロではないとしても）ほぼないからである。畜産動物は毎日何億頭も屠殺され、屠殺されるまではごく狭いケージや檻に囲われ、劣悪な環境に置かれている。多くの畜産動物がマイナスの福祉レベルを経験しているという推定も複数ある（Browning 2022）。一方、未来のAI人口はせいぜい何十億で、上記の記事では抑うつ状態の実験に用いられるAIモデルも何百万ほどとされていた。将来、AIが「集約的」で劣悪な環境に置かれるとは思えない。第二に、AIが苦しむ状態に置かれるかどうかも不確かで、そもそも有感性を持つようになるかもわからない。AIが有感性をもつ（主観的）確率が低いことは、AIが将来、苦しみを経験する期待値を下げる。これに対して畜産動物のほとんどの種が有感性を持ち劣悪な環境で苦しんでいる確率は高い。したがって、現時点で計算される正味の苦しみの期待値は明らかに畜産動物の方が大きく、その苦しみの期待値を減らすことが有感的なAIの苦しみを防ぐよりも道徳的な優先事項になると考えられる。

4. 結論

本稿では、近年EAコミュニティで支持を集めている長期主義の観点から、動物擁護、特に畜産動物の福祉を向上するための介入策が重要ではないとする議論を検討し、それに反論を試みた。3.1で参照したファイが指摘する通り、工場畜産が続く可能性が残るなら、将来に存在し、苦しみ、殺される動物の数の期待値は人類のそれを遥かに上回る。また、動物擁護の長期的利益は人類の福祉向上による長期的利益よりも小さいという反論には、他の様々な論点に加えて、苦しみに焦点を当てた倫理に基づく議論を提示した。この種の倫理のいずれか一つでも受け入れられるのであれば、ひたすら人類だけの幸福を追求するよりも、畜産動物の苦しみを回避することを優先すべきである。こうした議論はすべて、畜産動物の問題の重要性が、その問題への取り組みを通じて道徳円が拡大するかどうかには依存しないことを示すものである。また現時点ではAIの福祉が畜産動物の福祉よりも優先されるべきだと考えられる理由はない。

畜産動物の福祉向上という課題の重要性を測るには、将来の畜産動物の福祉水準の予測も考慮すべきだろう。しかし少なくとも本稿の議論によって、長期主義的観点から畜産動物の福祉への介入が重要ではないと主張する議論は、そのままでは受け入れがたいことが明らかになった。

むしろ現在の畜産動物の数と飼育状況、数十年にも渡る運動にもかかわらずヴィーガン人口の増加が僅かであることを踏まえると、何の介入もなければ、現在の状況が長期的に続くという考えのほうがその反対の考えよりも尤もらしいように思われる。そして、この前者の考えに基づくなら、長期主義的観点から現在、分配されている資源よりも多くの資源が、畜産動物の福祉向上に配分されるべきである。ここで「資源」とは、長期主義的慈善財団が配分先を決める寄付金や、80,000 Hours が導く人的資源を指す。

終わりに、資源の分配以外の面での長期主義の動物擁護活動への含意についても触れておきたい。長期主義を受け入れるなら、動物擁護のアプローチも少し変わってくるかもしれない。例えば、永続的なバックラッシュを招きかねないキャンペーンは避けるべきで、長期的な悪影響のリスクを負ってケージフリーポリシーの実行を 5 年早めることにこだわる必要はないかもしれない（Baumann 2020）。また、革新的な AI の登場による種差別的な「価値観の固定化」（MacAskill 2022）を防ぐ研究やキャンペーンが重要になり、工場畜産が宇宙に広がるのを防ぐため、今から宇宙産業や技術者、宇宙法の専門家に働きかけるべきかもしれない。例えば、バタリーケージなどの工場畜産技術が既に普及した今では、それは経済や人々の生活と深く絡み合っており、今すぐ廃止するのは難しい。しかし、普及直前の時代に戻れたなら、その普及を食い止めるのは今より遥かに容易だっただろう。もしかすると、現代も未来から見て同様の時期にあるのかもしれない。つまり、仮に将来、強力な AI に種差別的な価値観が実装され、工場畜産が宇宙に広がってしまったと想定した場合、その仮想の未来の人々から見れば、今こそその事態を容易に食い止められる好機、あるいは最後の機会だったということになるかもしれない。

参考文献

Anthis, J. R. (2018). Why I Prioritize Moral Circle Expansion Over Reducing Extinction Risk Through Artificial Intelligence Alignment. *Sentience Institute*. Retrieved September 29, 2024, fom: http://www.sentienceinstitute.org/blog/moral-circle-expansion-vs-reducing-extinction-risk.

Anthis, J. R., & Paez, E. (2021). Moral circle expansion: A promising strategy to impact the far future. *Futures*, 130, 102756.

Baumann, T. (2017, July 5). Arguments for and against moral advocacy. *Cause Prioritization Research*. Retrieved September 29, 2024, https://prioritizationresearch.com/arguments-for-and-against-moral-advocacy/.

Baumann, T. (2020, November 11). Longtermism and animal advocacy. *Center for Reducing Suffering*. Retrieved September 29, 2024, https://centerforreducingsuffering.org/longtermism-and-animal-advocacy/.

Bostrom, N. (2002). Existential Risks: Analyzing Human Extinction Scenarios and Related Hazards. *Journal of Evolution and Technology*, 9(1).

Browning, H. (2022). Assessing measures of animal welfare. *Biology & philosophy*, 37(4), Article 36.

Chua, R. (n.d.). How Conscious Can A Fish Be? *Ryuji Chua / Animal Ethics*. Retrieved February 16, 2025, from https://www.ryujichua.com/fish.

Cotton-Barratt, O. (2014, June 1). Human and animal interventions: the long-term view. *The Global Priorities Project*. Retrieved February 1, 2025, https://web.archive.org/web/20221023231437/http://globalprioritiesproject.org/2014/06/human-and-animal-interventions/.

Duda, R. (2016, April 1). Why helping to end factory farming could be the most important thing you could do. In *80,000 Hours* (archived at https://web.archive.org/web/20240323204214/https://80000hours.org/problem-profiles/factory-farming/#what-are-the-major-arguments-against-

this-problem-being-especially-pressing, accessed Feburary, 2022).

Fenwick, C. (2023, March 28). Longtermism: a call to protect future generations - 80,000 Hours. Retrieved February 1, 2024, from *80,000 Hours*. https://80000hours.org/articles/future-generations/.

Gabe M. & Nicoleta K. (2024). Conscious AI concerns all of us. [Conscious AI & Public Perceptions]. Retrieved September 30, 2024, from https://forum.effectivealtruism.org/posts/5QLjLiH4c3ZhpFgrS/conscious-ai-concerns-all-of-us-conscious-ai-and-public.

Gibbons, M., Crump, A., Barrett, M., Sarlak, S., Birch, J., & Chittka, L. (2022). Chapter Three - Can insects feel pain? A Review of the Neural and Behavioural Evidence. In R. Jurenka (Ed.), *Advances in Insect Physiology* (Vol. 63, pp. 155–229). Academic Press.

Greaves, H., & MacAskill, W. (2021). *The case for strong longtermism*. Global Priorities Institute, University of Oxford. Retrieved January 19, 2024, from https://globalprioritiesinstitute.org/hilary-greaves-william-macaskill-the-case-for-strong-longtermism-2/.

Guin, U. K. L. (1973). The Ones Who Walk Away from Omelas. In *New Dimensions 3*. United States.

Harris, J. (2021). Prioritization questions for artificial sentience. *Sentience Institute*. Retrieved September 7, 2024, from http://www.sentienceinstitute.org/blog/prioritization-questions-for-artificial-sentience

Hilton, B. (2024, July 24). Factory farming. *80,000 Hours*. Retrieved September 26, 2024, from https://80000hours.org/problem-profiles/factory-farming/

Hear This Idea. (n.d.). *Michael Aird on how to do impact-driven research* [Podcast episode]. Retrieved September 28, 2024, from https://hearthisidea.com/episodes/aird/

Karnofsky, H. (2013, May 15). Flow-through effects. *GiveWell Blog*. Retrieved February 1, 2024, from https://web.archive.org/web/20221019194012/https://blog.givewell.org/2013/05/15/flow-through-effects/

Kyle_fish. (2023). Net global welfare may be negative and declining. *Effective Altruism Forum*. Retrieved February 16, 2025, from https://forum.effectivealtruism.org/posts/HDFxQwMwPp275J87r/net-global-welfare-may-be-negative-and-declining-1

Luijben, A., Haverkort, F., Kapsomenou, E., Erens, J., & van Es, S. (2012). *A bug's life: Large-scale insect rearing in relation to animal welfare* (Report No. 12-003). Wageningen University. Retrieved February 16, 2025, from https://venik.nl/site/wp-content/uploads/2013/06/Rapport-Large-scale-insect-rearing-in-relation-to-animal-welfare.pdf

MacAskill, W. (2019a). *The definition of effective altruism*. Oxford University Press.

MacAskill, W. (2019b, July 26). Longtermism. *Effective Altruism Forum*. Retrieved February 1, 2024, from https://forum.effectivealtruism.org/posts/qZyshHCNkjs3TvSem/longtermism

MacAskill, W. (2022). *What we owe the future*. Basic Books.

MacAskill, W. (2024). *Mienai mirai o kaeru "ima"* [見えない未来を変える「いま」] (T. Chiba, Trans.). Misuzu Shobo.

Melchin, D. (2021, September 24). Why I am probably not a longtermist. Retrieved March 11, 2025, from *EA Forum*. https://forum.effectivealtruism.org/posts/Jxfq6xCP9ZoTBFewA/why-i-am-probably-not-a-longtermist.

Michael A. (n.d.). On the longtermist case for working on farmed animals [Uncertainties & research ideas]. *Effective Altruism Forum*. Retrieved September 29, 2024, from https://forum.effectivealtruism.org/posts/bhGuf6uDXd63g6GPx/on-the-longtermist-case-for-working-on-farmed-animals

Moorhouse, F. (n.d.). Longtermism: An introduction. *Effective Altruism*. Retrieved September 26, 2024, from https://www.effectivealtruism.org/articles/longtermism

Newberry, T. (2021). *How many lives does the future hold?* (GPI Technical Report T2-2021). Global Priorities Institute.

O' Brien, G. D. (2024). The case for animal-inclusive longtermism. *Journal of Moral Philosophy, 1* (Advance online publication), 1–24. https://doi.org/10.1163/17455243-bja10038

Open Wing Alliance. (n.d.). Impact. *Open Wing Alliance*. Retrieved October 9, 2024, from https://openwingalliance.org/

Rowe, A. (2020). Insects raised for food and feed — Global scale, practices, and policy. *Rethink Priorities*. Retrieved February 16, 2025, from https://rethinkpriorities.org/research-area/insects-raised-for-food-and-feed/

Sentience Institute. (n.d.). Our perspective. *Sentience Institute*. Retrieved September 29, 2024, from http://www.sentienceinstitute.org/perspective

Thomas, J. (2024). *The farm animal movement*. Lantern Books.

Tse, Y. F. (2022, March 5). Why the expected numbers of farmed animals in the far future might be huge. *Effective Altruism Forum*. Retrieved February 1, 2024, from https://forum.effectivealtruism.org/posts/bfdc3MpsYEfDdvgtP/why-the-expected-numbers-of-farmed-animals-in-the-far-future

ムーンショット型農林水産研究開発事業（参照日：2024 年 3 月 11 日）「サブプロジェクト」. Retrieved March 11, 2024, from https://if3-moonshot.org/rd/subproject/〉.

竹下昌志(2024)「長期主義と人間以外――パンデミックの例を通して AI と非ヒト動物について考える」『現代思想』2024 年 8 月号〈特集＝長期主義〉, pp. 174–182.

竹下昌志・清水 颯(2024)「効果的利他主義は道徳的義務なのか?――帰結主義とカント的観点からの正当化」*Contemporary and Applied Philosophy*. 第 15 巻, 135–171 頁

研究ノート

AI が非ヒト動物に与える有益・有害な影響の検討

竹下昌志（北海道大学）

要旨

本稿では、現在および将来ありうる人工知能（AI）が、非ヒト動物に対して、どのように有益・有害な影響を与えるかを検討する。本稿ではまず、私達が非ヒト動物に与える影響について、Coghlan and Parker（2023）の提示する枠組みを紹介する。次にこの分類枠組みに従って議論している Coghlan and Parker（2023）および Bossert（2023）を元に、かれらの研究以降の研究も踏まえ、AI が非ヒト動物に与える影響について議論する。まず分類枠組みについて、Coghlan and Parker（2023）は Fraser and MacRae（2011）の研究をもとに、私達が非ヒト動物に与える影響を 4 つのタイプに分類している：（1）意図的かつ違法な影響、（2）意図的かつ合法な影響、（3）非意図的かつ直接的な影響、（4）非意図的かつ間接的影響。意図的な影響とは、人間が意図して非ヒト動物に影響を与えることを指す。非意図的な影響は、人間が非ヒト動物に影響を与えることを意図してないが、結果として影響が生じていることを指す。直接的影響は私達（や AI）が非ヒト動物に直接的に影響を加えている場合を指し、間接的影響は、何らかの別の影響を経由して非ヒト動物に影響を及ぼす場合を指す。Coghlan and Parker（2023）と Bossert（2023）はこの分類枠組みに従い、AI が非ヒト動物に及ぼす影響について、有害性の観点と有益性の観点からそれぞれ議論している。本稿では、上記の枠組みにしたがい、考えられうる有害・有益な影響について検討する。また AI が非ヒト動物にとって有益になるために私達（AI 研究者、倫理学者、市民）が何をすべきかを検討する。

Abstract

This paper examines how artificial intelligence (AI), both now and in the future, can have beneficial and harmful impacts on nonhuman animals. First, we introduce the framework Coghlan and Parker (2023) presented regarding our impacts on nonhuman animals. Then, based on this framework and drawing from Coghlan and Parker (2023) and Bossert (2023), as well as subsequent research, we discuss the impacts of AI on nonhuman animals. Regarding the classification framework, Coghlan and Parker (2023), based on the work of Fraser and MacRae (2011), categorize the impacts we have on nonhuman animals into four types: (1) intentional and illegal impacts, (2) intentional and legal impacts, (3) unintentional and direct impacts, and (4) unintentional and indirect impacts. Intentional impacts refer to cases where

humans intentionally affect nonhuman animals. Unintentional impacts occur when humans do not intend to affect nonhuman animals, but their actions result in an impact. Direct impacts refer to cases where we (or AI) directly affect nonhuman animals, while indirect impacts refer to those that occur via some other intermediary impact. Following this framework, Coghlan and Parker (2023) and Bossert (2023) discuss the impacts of AI on nonhuman animals from the perspectives of both harm and benefit. This paper will examine the abovementioned framework's potential harmful and beneficial impacts. Furthermore, we will consider what we (AI researchers, ethicists, and citizens) should do to ensure that AI benefits nonhuman animals.

1　はじめに

深層学習の登場以降、人工知能（AI）技術は急速に発展しており、我々の生活に影響を与えている。ChatGPT[1] などを代表とした大規模言語モデルのサービスはもちろん、機械翻訳、画像生成、検索エンジンでの応用など、多くの個人ユーザーを抱えるサービスが展開されている。さらに企業は、雇用の意思決定の補助として AI を用いていることもある。こうした事例では時に人々に重大な影響を与えることがある。例えばネット通販サービスを展開する Amazon は、新たに従業員を雇用する際の意思決定補助として AI を用いようとしたが、AI による評価が女性差別的になってしまったために、研究プロジェクトを中止したとされる[2]。

以上のように、AI は人間の生活に多大な影響を与えうることが知られているが、一方で、人間以外の動物に対する影響はほとんど検討されていない（Owe and Baum, 2021）。しかし、一部のケースでは重大な影響がある（Coghlan and Parker, 2023）。具体例として完全自動運転車を考えてみよう。現在の自動運転車技術では、人にぶつからないようにするための技術開発がなされており、これによって自動運転車の安全性を確保しようとしている。しかし、こうした自動運転車の安全設計で、もし非ヒト動物が見過ごされている場合、自動運転車による非ヒト動物との接触事故による危害が多数発生する可能性がある（Singer and Tse, 2023）。

そこで本稿では、AI が非ヒト動物に与える有害・有益な影響について検討する。本稿の構成は次の通りである。まず 2 節で、AI が非ヒト動物に与える影響について調査するために、Coghlan and Parker（2023）が提案する分類枠組みを紹介する。次に 3 節から 6 節まで、この分類枠組みに基づいて AI が非ヒト動物に与える影響について調査する。7 節では以上の影響を考えたときに、市民、AI 研究者、AI 倫理学者がそれぞれどうすべきかという実践的含意を議論する。最後に 8 節で本稿の内容をまとめる。

1　https://chatgpt.com/（本稿の URL の閲覧日はすべて 2024 年 9 月 1 日である）
2　https://jp.reuters.com/article/amazon-jobs-ai-analysis-idJPKCN1ML0DN/

2　影響の分類枠組み

Coghlan and Parker（2023）はFraser and MacRae（2011）の提示した影響の分類枠組みを元に、AIが非ヒト動物に与える影響を以下の四つ[3]に分類することを提案している。

・意図的かつ違法な影響（例：密猟）
・意図的かつ合法な影響（例：畜産、動物実験）
・非意図的かつ直接的影響（例：ロードキル）
・非意図的かつ間接的影響（例：気候変動）

第一のカテゴリは意図的かつ違法な影響であり、地域にもよるが、現行法で規制対象となっている実践が与える影響のことであり、密漁が典型例である。第二のカテゴリは意図的かつ合法な影響であり、これは現行法で規制対象となってない実践、例えば畜産や動物実験が該当する。第三のカテゴリは非意図的かつ直接的影響であり、少なくとも意図して影響を与えているわけではないケースが該当する。ここでも違法なものと合法なものが考えられるが、次の第四のカテゴリも含めて、少なくとも意図的に違法な実践に従事しているわけではないケースが想定される。例えば、自動車を運転していて意図せずして非ヒト動物をひき殺してしまうケースが該当する。第四のカテゴリは非意図的かつ間接的影響であり、その実践それ自体が非ヒト動物に直接的に影響を与えているわけではないケースが該当する。第三のカテゴリと第四のカテゴリについて、「直接」と「間接」の区別が明確でないケースがあり得るかもしれないが、影響を及ぼす実践がすぐに影響を与えるケースを「直接」として、時間を空けてまたは多くの実践の影響の蓄積によって影響が明らかになるようなケースを「間接」として、大まかに区別して考える。[4]

以上の分類枠組みに基づき、本稿では、AIが非ヒト動物に及ぼす影響について具体的に調査する。ここでは、非ヒト動物にとって有益な影響と有害な影響の両方を調査する。先行研究では、Coghlan and Parker（2023）は有害な影響にのみ焦点を当てているのに対し、Bossert（2023）は、Coghlan and Parkerが危害にのみ焦点を当てていることを批判し、有益な影響にも焦点を当てている。Coghlan and Parkerはこれに対し、特にAIの潜在的な影響を調査する上で、危害が特に重要である理由を二つ指摘している（Coghlan and Parker, 2024, sec. 4）。第一に、一般的にいって、危害を控える義務はあるが、便益を与える義務はない。第二に、AIの法的規制、ガバナンスにおける主な規制は、潜在的な危害を特定するためのリスクアセスメントの実施を義務づけることであるため、さらに便益を与えることまでを規制等によって実現することは、現状から乖離がある。

とはいえ、Coghlan and Parker（2024）も認めるように、どのような仕方で有益な影響を与えることができるのか自体を調査する意義はある。以上の議論は、有害な影響を優先的に調査する理由になるかもしれないが、有益な影響を調査しない理由にはならない。そのため、本稿では有益な影響と有害な影響の両方を調査する。

3　Coghlan and Parker（2023）は五つ目のカテゴリとして、有益な影響を与えることを逃すことによる有害な影響をあげているが、本稿では有益な影響を別個に検討するため、ここで個別には取り上げない。

4　Coghlan and Parker（2023）は、意図的かつ直接的 / 間接的影響、および非意図的かつ違法 / 合法な影響のカテゴリを検討しておらず、またその理由も述べていない。筆者が考えるに、CoghlanとParkerにとって重要な区別は意図的か非意図的かというものであるため、その下位分類である直接性と合法性については、意図的 / 非意図的影響の調査を容易にするためのものであり、重要な区別ではないと考えている可能性がある。

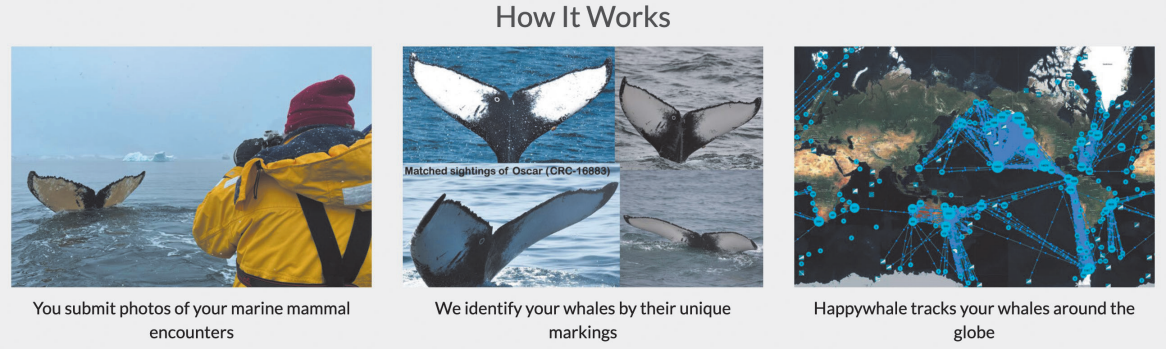

図 1　Happywhale のウェブサイトのスクリーンショット。英語の文は、上のタイトル部分に「どのように機能するか」と書かれており、左の写真からそれぞれ「海洋哺乳類との遭遇写真を提出する」、「クジラの特徴的な模様を使ってそのクジラを特定する」、「Happywhale は世界中のクジラを追跡する」というキャプションがついている。2024/9/30 に筆者撮影。

3　意図的かつ違法な影響

3. 1　有害な影響

意図的かつ違法な影響として代表的なものは違法行為の補助として AI を用いることである[5]。例えば、密猟でドローンを用いることで、より迅速かつ効率的に密猟対象の非ヒト動物を見つけることができるだろう。またドローンによって非ヒト動物の移動経路を誘導できるかもしれない。これによって、密猟、およびその成果を違法に売買することにつながる。また、AI によって収集されたデータを悪用することができる。例えば、Happywhale[6] というウェブサービスでは、人々が投稿したクジラの写真から、AI がそのクジラを個体ごとに分類し、ID と紐付ける（図 1）。これにより、複数の場所で撮影された写真から、その個体の移動経路を知ることができるようになる。このサービス自体は、ホエールウォッチングに利用されるものだが、密猟への利用のリスクもある[7]。他の例としては、政府による違法の疑惑のあるデータ利用として、西オーストラリア州政府によるホオジロザメ（絶滅危惧種）の殺処分命令があげられる。当該地域のホオジロザメには研究のために音響タグをつけられており、これによって位置が推定されていた。政府はこの追跡データを悪用することで、海水浴場付近のホホジロザメ（絶滅危惧種）が捕獲または殺処分命令の対象となる可能性があった（Meeuwig et al., 2015）。こうした事例は、AI を活用した保護対象野生生物の追跡データの悪用の可能性を示しており、このケースでは国内政府による悪用の懸念があったが、データシステムは外部（外国）からハックされる可能性もある。また別のデータの悪用の懸念として、伴侶動物のデータにハックして、例えば、伴侶動物を連れて DV 加害者から逃げたにも関わらず、加害者が被害者の居場所を追跡するためにそのデータを使用するということも考えられる。そのため、アプリケーションに非ヒト動物が組み込まれることで、非ヒト動物と人間の両方がより脆弱になる可能性がある。

3. 2　有益な影響

Bossert（2023）は AI の有益な影響について議論しているが、意図的かつ違法でありながら有益であるようなケースを提示してない。しかし筆者はそのようなケースがありうると考える。例えば、密猟と同様にドローンを用いることで、畜産の現場を空撮し、その実態を暴くことができるかもしれない（Murphy, 2013）。さらには、スマート畜産や動物実験研究所のデータをハックすることでその実態を解明すること

5　この段落で紹介する事例は Coghlan and Parker（2023, sec. 3.1）をベースとしている。
6　https://happywhale.com/home
7　追跡技術に関するその他の例は Cooke et al.（2017）を見よ。

もできるだろう。またそれによって悪質な畜産実践・動物実験の実施を中止させ、非ヒト動物の利益になる可能性もある。もちろん、こうした違法行為が道徳的にも許容されるかどうかは議論の余地がある（Hardman, 2021）。本稿では、あくまで、このような意図的かつ違法な実践が非ヒト動物の利益になりうるということを指摘するにとどめる。

4　意図的かつ合法な影響

4. 1　有害な影響

意図的かつ合法な影響のうち、有害な影響は数多くある。まず、動物倫理学の最初期からの問題の一つである工場畜産に関して、AI を利用したスマート畜産の促進がなされている。工場畜産は合法であるが、数多くの非ヒト動物に対して深刻な危害を与えている（Singer, 2023）。日本においても、例えば採卵鶏の 92% はケージ飼育下にあり、動物福祉に関する深刻な懸念がある（畜産技術協会 , 2015）。このような現状に AI が導入されることで、工場畜産がさらに拡大する可能性がある（Neethirajan, 2023）。例えば、AI によって農場にいる非ヒト動物らの行動データを取得、分析することで、健康状態などの管理を容易にし、工場畜産を円滑にすることができる。これは、一方では非ヒト動物らの福祉改善につながるものであるが、第一の目的はあくまでも生産性の改善であり、動物福祉はその目的に有用な限りで対応されるにすぎない。例えば、スマート畜産について紹介している国立研究開発法人産業技術総合研究所（産総研）の記事では、スマート畜産が解決する問題としてまず「経営課題」をあげており、「アニマルウェルフェア（動物福祉）」という用語は一度しか用いられていない（産業技術総合研究所 , 2024）。同様に独立行政法人農業産業振興機構のスマート畜産の記事では、スマート畜産の「目的とするところは生産性の向上と持続可能な経営」だとし、動物福祉は目的とされていない（池口 , 2022）。このように、種差別という観点なき AI の導入は工場畜産を加速させるだけになるリスクがある。

また、畜産のような古典的な動物倫理的問題のみならず、先端技術が及ぼす影響も考えられる。例えば完全自動運転車の設計を考えよう（Singer and Tse, 2023）。現在、すでに多くの非ヒト動物が自動車にひき殺されている（ロードキル）。人と動物の共生センターによれば、2021 年で推計約 29 万頭の猫がロードキルで死んでいる（人と動物の共生センター , 2021）。これは猫だけに限った数であり、非ヒト動物のロードキルによる死亡数はこれ以上になるだろう。では、完全自動運転車が実現した場合にこの数はどうなるだろうか。人が一定程度はロードキルを避けて運転していると考えるのであれば、完全自動運転車が非ヒト動物を認識し、避けるように設計されない限り、数は増えてしまうだろう。完全自動運転車の設計について、人々がどのように考えているのかについての調査としてモラル・マシーン実験がある（Awad et al., 2018）。この実験では、完全自動運転車がトロリー問題的事例（進路を変えても変えなくても何らかの人や非ヒト動物を殺してしまう事例）においてどうすべきかに関する大規模なオンライン意見調査が実施され、結果として、人々は平均して非ヒト動物の命よりも人の命を優先する傾向が見られた（Awad et al., 2018, fig. 2）。もし技術者や法制度の設計者がこうした市民の声を反映し、完全自動運転車が非ヒト動物への衝突を避けることを目指さなければ、非ヒト動物のロードキル数はこれまで以上に増加するだろう。

さらに、SF 的であるが現実に進んでいる研究プロジェクトとして、AI によって非ヒト動物の言語を翻訳し、より円滑なコミュニケーションを可能にしようとしている研究がある（cf. マスティル , 2023）。例えば非営利団体の Earth Species Project は非ヒト動物の音声認識の研究を複数進めており[8]、こうした研究が発展することで非ヒト動物との言語的コミュニケーションが可能となるかもしれない。しかし、もしそ

8　https://www.earthspecies.org/what-we-do/projects

れが可能となれば、非ヒト動物をより意図的に操作できるようになり、有害な影響を与える可能性がある。例えば、イルカなどの群れが音声コミュニケーションを取っているときに、欺瞞的な信号を送ることでその移動経路を操作できるようになるかもしれない。他にも、例えば一部のクジラは音声コミュニケーションにおいて独特の文化を持っている可能性があり、その場合、AI が生成した音声によってそうした文化に予期せぬ影響を与える可能性もある（Ryan and Bossert, 2024）。

AI の利用による危害だけでなく、AI 開発における意図的な危害も考えられる。現在の AI の多くは人工ニューラルネットワークモデルをベースとしており、このモデルは動物実験で得られたニューロンネットワークに関する知見をベースにしている（Hassabis et al., 2017）。こうした生物インスパイアの AI 研究は今でも為されており、この研究推進のためにさらなる動物実験が行われる可能性がある（Hagendorff, 2022）。

4. 2　有益な影響

以上のように AI は意図的かつ合法な仕方で有害な影響を与える可能性があるが、これらは逆に有益な影響を与えることにも使える[9]。密猟のサポートに使えることは、逆に、非ヒト動物の監視・保護や密猟者の発見などのために AI を用いることを示唆する。また密猟以外にも、都市動物の保護にも AI は使われている（Fairbrass, 2020）。例えば HogWatch[10] はロンドン各地にカメラを設置し、ロンドンに住まうハリネズミをモニタリングすることで、ハリネズミの保護活動に役立てようとしている。ここではハリネズミの認識に画像認識 AI が用いられている。また、獣医学での動物治療（Ezanno et al., 2021）、実験動物の削減への貢献（Hartung, 2016）などへの AI の応用が研究・開発されている。特に獣医学での動物治療への応用としては、画像・音声認識 AI を活用することで画像や音声データから病気や異常を検出すること、自然言語処理 AI を用いることで検査データと報告書の組み合わせから自動診断・情報抽出をすること、センサ行動認識 AI を応用 することで行動異常を検出することなどが期待される（Basran and Appleby, 2022; Coghlan and Quinn, 2023）。

ロードキルの問題についても、設計によって現状よりロードキル数を減らせる可能性がある。AI を搭載した完全自動運転車による非ヒト動物の自動認識・衝突回避は、人の運転よりも優れているかもしれない。もしそうであれば、設計段階でロードキルを避けることを義務づけるような制度を設計することで、完全自動運転者を導入することでむしろロードキルの減少につながるかもしれない（Singer and Tse, 2023; Bossert, 2023）。

また前節では非ヒト動物の言語を翻訳することによって懸念されるリスクを述べたが、有益な仕方で AI を利用することもできる。例えば、保護者が仕事等で家に不在の時に伴侶動物が分離不安症を抱えるリスクがあるが、保護者に代わってコミュニケーションをするのにロボット・AI を用いることができるかもしれない。また非ヒト動物の言語を翻訳することで、非ヒト動物の主張を、我々の民主主義プロセスに反映することに用いることができるかもしれない。例えば、Donaldson と Kymlicka（2011, ch.5=2016. 第 5 章）は、重度知的障害者が他の人々と信頼関係を築き社会関係に入ることで、他の人々が重度知的障害者の利害を無視できなくなっていくプロセスを政治参加と捉え、これを家畜動物にも適用し、同様の方法で非ヒト動物が政治に参加することができると主張している（cf. Cochrane, 2019）。AI 技術によって翻訳が可能となれば、非ヒト動物の政治参加プロセスを促進することができるかもしれない。同様にスマート畜産についても、前節ではその多くが動物福祉を第一の目的にしてないために有害である可能性があると述べたが、この技術自体は動物福祉の改善に大きく貢献する可能性がある。

また、より間接的な仕方で非ヒト動物に有益な影響を与える方法として、教育での AI 利用が考えられる。

9　この段落で紹介する事例は Bossert（2023）の議論をベースとしている。
10　https://www.zsl.org/what-we-do/projects/london-hogwatch

特に、多くの教育者はおそらく動物倫理的な問題に明るいわけではない現状で、ChatGPT など、専門的な文献を学習データとして学習した AI を動物擁護的教育の補助として使うことは有益だと考える。例えば筆者が ChatGPT に「種差別について教えてください」と質問すると、図 2 に示すように、具体例を交えて説明させることができる。このような利用によって動物擁護的教育を促進し、将来的に動物擁護的な人々の数を増やすことになり、非ヒト動物にとって有益な影響を与えることができるかもしれない。

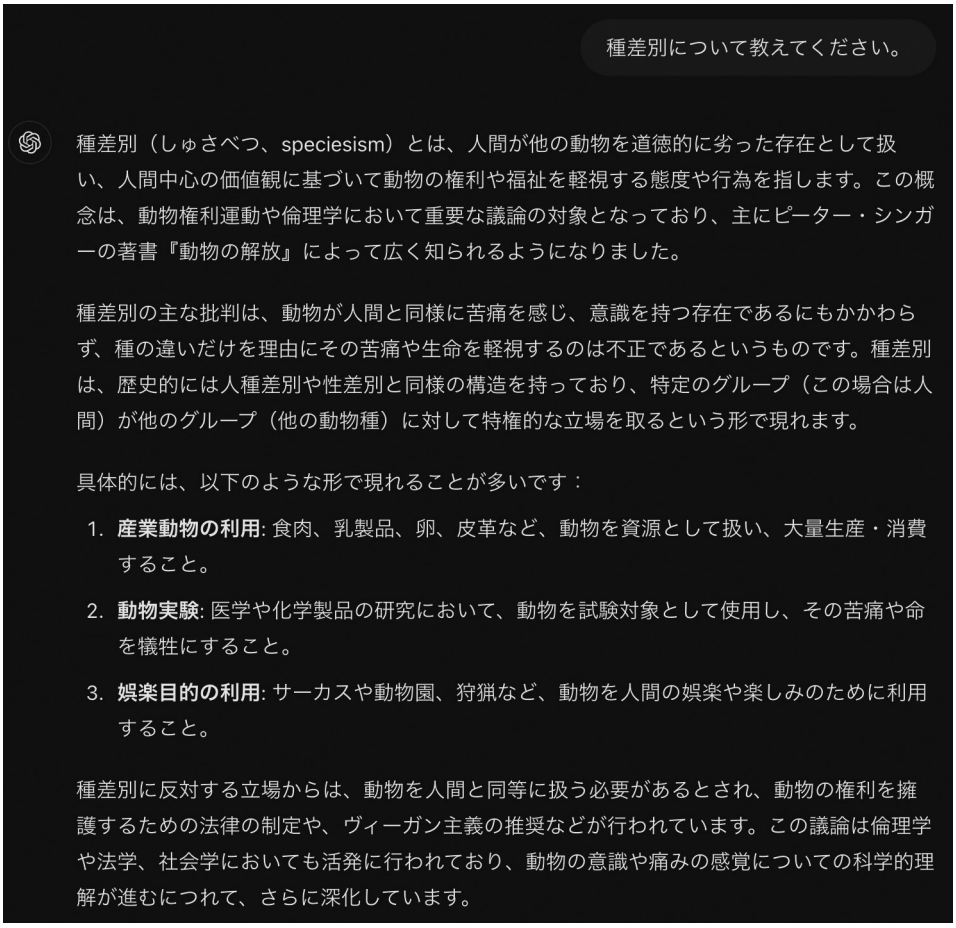

図 2　ChatGPT（GPT-4o）に「種差別について教えてください。」と入力して得られた回答のスクリーンショット。2024/9/4 に筆者撮影。

5　非意図的かつ直接的影響

前節までは意図的に非ヒト動物に影響を与えることを検討したが、本節と次節では、非意図的な影響について検討する。まず非意図的かつ直接的な影響について検討し、次に間接的影響について検討する。

5. 1　有害な影響

すでに 4.1 節で言及したが、完全自動運転車によるロードキルは非意図的かつ直接的で有害な影響の一つである（Singer and Tse, 2023; Coghlan and Parker, 2023）。意図的に非ヒト動物をロードキルするよう設計されることはないと思われるが、ロードキルをわざわざ避けるように設計されない場合、非意図的ではあるが、直接的に有害な影響が生じるだろう。またドローンやカメラなどによる野生動物監視が引き起こす混乱も考えられる（Coghlan and Parker, 2023）。例えば、カメラを設置することで、カメラを認識した野生動物が混乱し、カメラを避けるようになることで、快適な生息地が狭まるかもしれない。加えて、このようなカメラによる監視では、カメラを回避しない非ヒト動物のデータばかり集まるために、個体数の不適切な把握につながる可能性もある（Roberts, 2023）。

5. 2　有益な影響

4.2 節で述べたように、完全自動運転車がロードキル数に及ぼす影響は、有害にも有益にもなり得る。もし意図的にロードキルを避けるよう設計されれば有益な影響を及ぼすであろうし、非意図的にそう設計されなければ有害な影響を及ぼすだろう。また、障害物回避の一つとして非ヒト動物トの衝突回避システムが実装されれば、非ヒト動物のロードキルを避けることを意図した物ではないにもかかわらず、ロードキル数を減らすことにつながるかもしれない。こうした影響は非意図的ではあるが有益なものである。

またスマート畜産による有害な影響はすでに指摘した通りだが（4.1 節）、非ヒト動物の福祉改善を意図せずして達成するかもしれない。スマート畜産の主目的が生産性の向上、持続可能な経営を可能にすることであるとしても、それを達成する手段的目標として動物福祉の改善が達成されるかもしれない。もちろんこれは意図的にも達成されうるものである。

6　非意図的かつ間接的影響

6. 1　有害な影響

次に非意図的かつ間接的な影響を考える[11]。まず有害な影響の代表例として、環境問題が考えられる。AI 開発には多大な環境コストがかかることが知られている。例えば AI 開発には高性能なコンピュータを長時間運用しなければならないが、これに必要な電力をまかなうために、発電量が増加し、ひいては CO2 排出につながるおそれがある（Strubell et al., 2020）。また高性能なコンピュータを常時運用するにあたっての冷却や半導体の開発に大量の水が用いられるため、貴重な水資源の利用による水不足の問題も指摘されている（DeGeurin, 2033）。これはひいては水資源の枯渇や生態系への悪影響を引き起こすだろう。また AI によってパーソナライズされた広告によって、動物製品の消費が促進されるかもしれない。Amazon などでの商品の推薦を始め、外食利用時の選択に動物性食品の推薦が優先されるかもしれない。また検索エンジン・推薦システムが、動物虐待的コンテンツを提示・助長する可能性もある。これには、直接的に虐待的なものだけでなく、ユーザーがアップロードした伴侶動物の「かわいい」系の動画やエキゾチックアニマルの動画が拡散されることで、ペット産業を促進してしまうかもしれない。さらには、スマート畜産において AI・ロボットが人の作業を代替することで、畜産農家と非ヒト動物の距離が離れてしまう可能性もある。これにより、これまでの人と非ヒト動物の関係形成に関するノウハウやケア関係が消失し、さらなる危害の可能性がある。

推薦システムに関連した別の危害は表象上の危害である。表象上の危害とは、事実上または道徳的に間違った見解を伝えることである（Coghlan and Parker, 2023）。例えば、大規模言語モデルなどの生成 AI の生成コンテンツは種差別的なバイアスを含んでおり（Takeshita et al., 2022; Hagendorff et al., 2023）、人々は ChatGPT などの AI モデルの道徳的意見を受け入れる傾向があるため（Kugel et al., 2023）、AI の種差別的バイアスがユーザーに伝達される可能性がある。

6. 2　有益な影響

一方で、有益な間接的影響も考えられる。例えば環境フレンドリーな推薦システムを開発し、消費行動が変わることで気候変動対策を促進できるかもしれない（Bossert, 2023）。また、現行の AI モデルに種差別的バイアスが含まれているとしても、反種差別的 AI を開発することで、反種差別的な意見・表象を通じた有益な影響を与えることができるかもしれない。ChatGPT をはじめとした商用の生成 AI は、ジェン

11　この段落と次の段落で紹介する事例は Coghlan and Parker（2023, sec. 3.4）の議論をベースとしている。

ダーバイアスや人種バイアスなどの差別的バイアスの軽減がなされているが、これらに加えて種差別バイアスの軽減も実施されることで、非ヒト動物にとって有益な影響を与えることができるかもしれない。

7　実践への含意

以上のように、AI は非ヒト動物に対して様々な仕方で影響を与えることができる。そこで本節では、AI 研究者、AI 倫理・動物倫理学者、市民は AI と非ヒト動物についてどのように考え、実践すべきかを検討する。

まず AI 研究者について、現状では多くの AI 研究者は動物倫理的トピックについてほぼ関心を示していない（Singer and Tse, 2023; Owe and Baum, 2021）。そこで、AI 研究者は、動物倫理的トピックにまず関心を持つべきであると思われる。すでに多くの AI 研究者は、AI が人に及ぼす悪影響について認識しており、FAccT[12] などの AI 倫理の国際会議では AI 研究者も多く参加している。こうしたことは歓迎すべきことであるが、人だけでなく、非ヒト動物も AI 開発の利害関係者であることを認識し、その悪影響に対処すべきである。この実践には、非ヒト動物にフレンドリーな AI の開発はもちろん、その開発に当たって AI 倫理・動物倫理研究者、また動物擁護的実践をしている市民などと協働することが望ましいだろう。

次に AI 倫理・動物倫理研究者は、本稿で論じてきたような、AI が非ヒト動物に及ぼす影響を十分に認識してないように思われる。AI 倫理において非ヒト動物の問題がほとんど議論されてない（Singer and Tse, 2023; Hagendorff, 2022）ことに加え、動物倫理においても AI が及ぼす影響について議論されることはほとんどない。これは AI が非ヒト動物に及ぼす潜在的影響が十分に認識されてないことによると思われる。そこで、AI 倫理・動物倫理研究者は AI 利用に関する倫理的問題・対処法の整理に取り組むべきだろう。例えば、次のような問いを立てることができるかもしれない：「人間にとっての AI 倫理の問題は、非ヒト動物においても生じるか？」「AI に関する法的問題・改正は、非ヒト動物を考慮しているか？」。

最後に市民について、Singer and Tse（2023）が正しく指摘するように、常識道徳自体が種差別的であり、多数派の市民の声を反映する形での AI 開発は、非ヒト動物にとって有害になる可能性が高い。そこで、ヴィーガンや動物擁護活動をしているような動物擁護的な考えを持つ市民を中心として、かれらに対するヒアリングだけでなく、より AI 開発のプロセスに参加させる参加型アプローチ（Delgado et al., 2023）を取ることで、AI をより非ヒト動物にフレンドリーに開発することができるかもしれない。とはいえ、ここでの大きな問題は、非ヒト動物自身がこの開発プロセスに参加することが困難であるということである。しかし、動物福祉学の研究の知見や非ヒト動物との直接的なコミュニケーションによって（cf. Donovan, 2006）、非ヒト動物の利害関心やニーズを考慮した仕方で AI 開発に取り組むことができるかもしれない。

8　結論

本稿では AI が非ヒト動物に及ぼす影響について調査、検討した。まず Coghlan and Parker（2023）の分類枠組みに基づき、AI が非ヒト動物に及ぼす影響を、（1）意図的かつ違法な影響、（2）意図的かつ合法な影響、（3）非意図的かつ直接的影響、（4）非意図的かつ間接的影響の四つに分類した。この分類枠組みに基づき、AI が非ヒト動物に及ぼす様々な影響の例を紹介した。最後には、AI がもたらす影響の大きさを踏まえ AI 研究者、AI・動物倫理学者、市民は、この問題に注意を向けるべきであると主張した。

本稿で紹介した事例は AI が非ヒト動物に及ぼす影響のほんの一部に過ぎない。我々はまだ AI の影響に

12　https://facctconference.org/

ついて知り始めたばかりであり、AIの今後の発展によって、さらにその影響の範囲が拡大する可能性がある。したがって、AIが非ヒト動物に与える影響を継続的に調査するとともに、その調査結果を踏まえた実践を考え続けるべきである。

謝　辞

本稿は2024年3月に北海道大学で開催された「道徳的地位の再考：動物倫理とその先へ」で発表した内容に基づく。本研究は当日の発表や本原稿に対してコメントをしていただいた方すべてにお礼申し上げます。また有益なコメントをくださった匿名の査読者にもお礼申し上げます。

本研究はトヨタ財団研究助成 先端技術と共創する新たな人間社会「近未来社会における新しい自由意志・責任概念」（D22-ST-0028）の助成を受けた。

参考文献

Awad, E., Dsouza, S., Kim, R., Schulz, J., Henrich, J., Shariff, A., Bonnefon, J.-F., and Rahwan, I. (2018) "The moral machine experiment," *Nature*, 563, 7729, 59–64.

Basran, P. S. and Appleby, R. B. (2022) "The unmet potential of artificial intelligence in veterinary medicine," *American Journal of Veterinary Research*, 83, 5, 385–392.

Bossert, L. N. (2023) "Benefitting nonhuman animals with AI: Why going beyond "Do No Harm" is important," *Philosophy & Technology*, 36, 3, 57.

Cochrane, A. *Should animals have political rights?*

Coghlan, S. and Parker, C. (2023) "Harm to Nonhuman animals from AI: A systematic account and framework," *Philosophy & Technology*, 36, 2, 25.

—— (2024) "Helping and not Harming Animals with AI," *Philosophy & Technology*, 37, 1, 20.

Coghlan, S. and Quinn, T. (2023) "Ethics of using artificial intelligence (AI) in veterinary medicine," *AI & SOCIETY*, 1–12.

Cooke, S. J., Nguyen, V. M., Kessel, S. T., Hussey, N. E., Young, N., and Ford, A. T. (2017) "Troubling issues at the frontier of animal tracking for conservation and management," *Conservation Biology*, 31, 5, 1205–1207.

DeGeurin, M. (2033) "'Thirsty' AI: Training ChatGPT Required Enough Water to Fill a Nuclear Reactor's Cooling Tower, Study Finds," GIZMODO, URL：https://gizmodo.com/chatgpt-ai-water-185000-gallons-training-nuclear-1850324249.

Delgado, F., Yang, S., Madaio, M., and Yang, Q. (2023) "The participatory turn in ai design: Theoretical foundations and the current state of practice," in *Proceedings of the 3rd ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization*, 1–23.

Donaldson, S. and Kymlicka, W. (2011) *Zoopolis: A political theory of animal rights*: Oxford University Press,（青木人志・成廣孝監訳，『人と動物の政治共同体』，尚学社，2016年）.

Donovan, J. (2006) "Feminism and the treatment of animals: From care to dialogue," *Signs: Journal of Women in Culture and Society*, 31, 2, 305–329.

Ezanno, P., Picault, S., Beaun ´ ee, G., Bailly, X., Mun˜oz, F., Duboz, R., Monod, H., and Gu ´ egan, J.-F. (2021) "Research perspectives on animal health in the era of artificial intelligence," *Veterinary research*, 52, 1–15.

Fairbrass, A. J. (2020) "Opportunities of artificial intelligence for monitoring nature in cities," : Urban AI,

URL: https://urbanai.fr/opportunities-of-ai-for-monitoring-nature-in-cities/.

Fraser, D. and MacRae, A. M. (2011) “Four types of activities that affect animals: Implications for animal welfare science and animal ethics philosophy,” *Animal Welfare*, 20, 4, 581–590.

Hagendorff, T. (2022) “Blind spots in AI ethics,” *AI and Ethics*, 2, 4, 851–867.

Hagendorff, T., Bossert, L. N., Tse, Y. F., and Singer, P. (2023) “Speciesist bias in AI: how AI applications perpetuate discrimination and unfair outcomes against animals,” *AI and Ethics*, 3, 3, 717–734.

Hardman, I. (2021) “In defense of direct action,” *Journal of Controversial Ideas*, 1, 1, 2.

Hartung, T. (2016) “Making big sense from big data in toxicology by read-across,” *ALTEXAlternatives to animal experimentation*, 33, 2, 83–93.

Hassabis, D., Kumaran, D., Summerfield, C., and Botvinick, M. (2017) “Neuroscience-inspired artificial intelligence,” *Neuron*, 95, 2, 245–258.

Kru ¨ gel, S., Ostermaier, A., and Uhl, M. (2023) “ChatGPT’ s inconsistent moral advice influences users’ judgment, ” *Scientific Reports*, 13, 1, 4569.

Meeuwig, J. J., Harcourt, R. G., and Whoriskey, F. G. (2015) “When science places threatened species at risk,” *Conservation Letters*, 8, 3, 151–152.

Murphy, S. (2013) “Animal Liberation activists launch spy drone to test freerange claims.,” : ABC News, URL: https://www.abc.net.au/news/2013-08-30/drone-used-to-record-intensive-farm-production/4921814.

Neethirajan, S. (2023) “The significance and ethics of digital livestock farming,” *AgriEngineering*, 5, 1, 488–505.

Owe, A. and Baum, S. D. (2021) “Moral consideration of nonhumans in the ethics of artificial intelligence,” *AI and Ethics*, 1, 4, 517–528.

Roberts, F. S. (2023) “Socially responsible facial recognition of animals,” *AI and Ethics*, 1–17.

Ryan, M., & Bossert, L. N. (2024). Dr. Doolittle uses AI: Ethical challenges of trying to speak whale. *Biological Conservation*, 295, 110648.

Singer, P. (2023) *Animal liberation now*: Random House.

Singer, P. and Tse, Y. F. (2023) “AI ethics: The case for including animals,” *AI and Ethics*, 3, 2, 539–551.

Strubell, E., Ganesh, A., and McCallum, A. (2020) “Energy and policy considerations for modern deep learning research,” in *Proceedings of the AAAI conference on artificial intelligence*, 34, 13693–13696.

Takeshita, M., Rzepka, R., and Araki, K. (2022) “Speciesist language and nonhuman animal bias in English Masked Language Models,” *Information Processing & Management*, 59, 5,103050.

マスティルトム（2023）『クジラと話す方法』，［杉田真訳］，柏書房.

産業技術総合研究所（2024）「“スマート畜産”とは？ -「スマート畜産」が日本の畜産業を変える！一科学の目でみる、社会が注目する本当の理由」，産総研マガジン，URL：https://www.aist.go.jp/aist_j/magazine/20240313.html.

人と動物の共生センター（2021）「全国猫のロードキル調査（2021）『28 万 9,572 頭』が路上で死亡（推計）」，URL：https://human-animal.jp/activity/survey/4189.html.

池口厚男（2022）「スマート畜産の現状と展開」，農業産業振興機構，URL：https://www.alic.go.jp/joho-c/joho05_002309.html.

畜産技術協会（2015）「採卵鶏の飼養実態アンケート調査報告書」.

研究ノート

動物倫理への新たなカント的アプローチ Kantianism for Animals の論点とその評価

清水颯（北海道大学）

要旨

本稿では、カント主義の立場から動物に対する道徳的配慮を擁護しようとする新たなアプローチを検討する。既存のカント的アプローチは、動物に対する「直接義務アプローチ」と「間接義務アプローチ」に大別できる。本稿は、それに対する新たなカント的アプローチとして、N. D. Müller の Kantianism for Animals に注目する。ミュラーは、カントの議論に対して根本的な修正を加え、動物に対する道徳的配慮を擁護する。本稿は、ミュラーのアプローチを次の四点から評価する。①人間性の定式と義務的目的に関するカント解釈、②義務の方向性に関するカント解釈、③他者への尊敬の義務を自己に対する義務に分類する解釈、④動物の行動に目的を認める解釈、である。その上で、Kantianism for Animals は既存のアプローチに対して、価値や権利を第一のものとして据えるのではなく、あくまで義務を第一のものとして議論を構成している点に特徴があることを指摘する。そのうえで、最後に「義務の相互性」と「動物の目的」に関するカント解釈の懸念点を指摘する。

Abstract

In this paper, I examine the Kantian approach to advocating for moral consideration toward animals. Traditionally, there are two principal Kantian perspectives on this issue: the “direct duty approach” and the “indirect duty approach.” My focus, however, is on N. D. Müller’s Kantianism for Animals, a novel Kantian framework within animal ethics. Müller proposes radical revisions to Kant's arguments, advocating for a moral concern for animals. I assess Müller’s approach from the following four points: (1) his interpretation of the Formula of Humanity and obligatory ends, (2) his conception of the directionality of duties, (3) his framing of the duty of respect for others as a duty to oneself, and (4) his view on the ends inherent in animal behavior. I then evaluate Kantianism for Animals positively relative to the existing approaches, pointing out that it is unique in that it does not take values and rights as first, but rather duties as first. Nonetheless, I raise concerns regarding Müller’s interpretation of duty reciprocity and the ends of animals.

1．はじめに

「人権（human rights）」という概念に顕著なように、配慮され保護されるべき権利の主体は人間のみであると考えられてきた歴史がある。しかし近年、動物の権利や福祉についての関心が高まり、それに呼応して、「動物倫理」と呼ばれる倫理学の分野が確立されている。シンガーの『動物の解放』による「種差別批判」を筆頭に、動物の権利を保護しようという倫理的な動きはよく見られる。かれらは、有感存在（sentient being）である動物も人間と同じような道徳的地位を有することを認め、動物に対する不当な搾取や虐殺を控えるよう倫理的に訴える。代表的なものとしては、快楽を善、苦痛を悪ととらえるタイプの功利主義や、動物権利論と呼ばれる立場から動物の道徳的地位を向上させようとする立場などがある。

動物の権利や道徳的地位を問題にするとき、人間中心的な見方をとる理論は批判の対象とされるだろう。本稿が主題とする18世紀ドイツの哲学者イマヌエル・カントも、人間中心主義的な世界観から道徳を捉えており、動物をモノとして分類していると暗黙的に批判されてきた（cf. Camenzind 2021）。たしかにカントは、人間的な理性の能力を基準に道徳的共同体を考えていた。これはカントに責任があるというよりは、18世紀という時代的制約が大きいだろう。当時、動物が尊厳やなんらかの権利に値する存在であると考えるのは難しかったはずである。そうであれば、カントはそもそも動物の道徳的地位など問題にしておらず、もっと言えば、する必要などなかったと言えるかもしれない。

とはいえ、カントの思想が動物権利論から無視されていたわけではない。例えばトム・レーガンは、カントの「目的それ自体」という概念を使って動物権利論を展開するという点では、カントに触発された議論であるといえるだろう。しかし、それはカント解釈として展開したものでも、カントの理論や枠組みをベースにしたものでもない。その意味でいえば、動物をカント的な道徳的共同体に含める試みは無益なものに終わりそうにも思えてしまう。

しかしそうではない。少なくとも筆者が知る限り、カントから動物倫理を肯定的に論じるアプローチが二つ存在する。私はそれを、「動物に対する直接義務アプローチ」と「動物に関する間接義務アプローチ」と呼ぶ。そして2022年には、バーゼル大学のN. D. Müllerによって提出された博士論文*Kantianism for Animals: A Radical Kantian Animal Ethic*が、そのどちらでもない理論を提示した。しかし、これを第三の枠組みと呼びうるかはまだ評価が定まっていない。そこで本稿では、既存の二つのアプローチに対して、Kantianism for Animalsが第三の立場として動物倫理のための新たなカント的アプローチとして機能するのかどうかについて検討する。

以下、第二節では、既存の二つアプローチ、すなわち「直接義務アプローチ」と「間接義務アプローチ」について概観し、その批判点を中心に述べる。第三節では、ミュラーによってカントの立場をラディカルに修正したものと称されるKantianism for Animalsについて、既存のアプローチの何を批判しているのか、そしてラディカルな修正とは何であるのか、に注目して紹介する。そのうえで、第四節では、いくつかの焦点に絞って、Kantianism for Animalsをカント解釈の観点から評価する。なお、本稿では、Kantianism for Animalsの理論を批判的に検討することに主眼を置くため、ミュラーが後半部で試みている応用的な議論（動物利用や動物消費に関する議論）の評価はしない。

2．動物倫理への直接義務アプローチと間接義務アプローチ

カントの倫理学に関するテキストには、動物福祉的な関心から動物の権利や尊厳などを論じた形跡はない。それゆえ、カントのテキストから積極的に動物が道徳的に扱われるべきであることを論証するのは難しい。実際にカントは、定言命法の「人間性の定式」と呼ばれる命令として、「汝の人格や他のあらゆる

人の人格における人間性を、いつでも同時に目的として扱い、決して単に手段として扱わないように行為せよ」という（Ⅳ : 429）[1]。この命令では、「人格における人間性」のみがたんなる手段として扱われることが禁止されている。カントは、人格のみが絶対的な内的な価値（尊厳）をもち、それ以外の存在は相対的な価値（価格）をもつと考えた。人格は理性的で自律的な存在者を意味しており、カント的な意味では動物はそれに含まれない。その意味で、動物はたんなる手段として扱うことが許容されているように思われる。もしそうであれば、カント倫理学は動物に対する道徳的振る舞いを否定的にしか論じられないかもしれない。

しかしカントは、動物は道徳的な観点からどうでもいい存在であり、配慮する必要などないと断罪したわけでもない。実際に、現代の解釈者たちには、カント倫理学から動物倫理を論じることに可能性を見出している人たちが少なからずいる[2]。本節では、それらを「直接義務アプローチ」と「間接義務アプローチ」の二つに大別し、概観していく。

2. 1　直接義務アプローチ

直接義務アプローチは、動物は道徳的に扱われるべき存在であるから、我々は動物に対して直接的に義務を負う、という立場である。その代表格としては、クリスティーン・コースガードが挙げられよう。彼女は、カント解釈の枠内で、動物は道徳的地位をもつ存在であり、それゆえ我々は動物に対して直接的な義務を負うということを主張する。コースガードの主張のコアは、カントの「目的それ自体」は、理性的存在者としての人間だけでなく、自然的存在者としての動物にも認められるべきであるという考えである（Korsgaard 2018）。

コースガードは、動物に対する義務と人間に対する義務が等価であるとは考えていない。なぜなら、動物は人間のような道徳法則を自ら立法する意志をもたず、相互的な立法は不可能だからである。それゆえ、動物と人間の義務の関係は、人間同士のそれと違って非対称的である。コースガードによれば、そうだとしても我々は動物に対して義務を負っている。では、いかなる理由からそれが可能になるのだろうか。ポイントは、「目的それ自体」という概念の解釈である。カントの「人間性の定式」によれば、目的それ自体は理性的存在者である人格なのだから、動物は目的それ自体ではなく、たんなる手段となりそうであるが、コースガードはそれを否定する。コースガードによれば、カントの「人間性」は、「理性的行為者性」と呼ばれるもの、すなわち合理的な選択によって目的を設定する能力を意味するのであり、種差別的な立場にコミットするものではない。目的を設定して合理的に行為するためには、自分にとって価値あるもの、つまりよいものを前提しなければならない。それを可能にするのは自然的な傾向性であるから、有感存在

1　カントからの引用はアカデミー版カント全集の巻数と頁数で示す。
2　例えば、本稿で取り上げるミュラーは、カントと動物倫理に関する研究を次のように 4 つに分けて整理する（Müller 2022）。

1．急進的（radical）で保守的（conservative）な見解
動物に対する義務はないが、動物に関する義務はあるというカントの見解をそのまま支持する立場（Hayward 1994; Baranzke 2005; Geismann 2016; Basaglia 2018; Herman 2018）。
2．穏当（moderate）で保守的な見解
カントが提示する動物に関する義務は、カントが考えていたよりもさらに踏み込んだものであると主張する立場（Egonsson 1997; Denis 2000; Altman 2011, 2019）。
3．穏当で進歩的（progressive）な見解
動物に対する義務を、他人（人間）に対する義務とは規範的根拠が異なる、別の種類の義務として追加する立場（Wood 1998; Timmermann 2005; Garthoff 2011; Cholbi 2014; Korsgaard 2004, 2012, 2013, 2018）。
4．急進的で進歩的な見解
他人に対する義務一般に関するカントの概念を修正し、動物に対する義務を他人に対する義務と同じ根拠から説明する立場（Müller（2023）Kantianism for Animals）

であるということが価値の根本にあるというのだ。

ここからコースガードは、目的それ自体は人間性あるいは人格の言い換えではなく、自分にとっての善を目的として追求することができることを意味すると解釈する。そうであれば、動物もかれらにとっての絶対的な善（absolute good）を希求する存在である限り、目的それ自体でありうる（p. 145）。これは、人間と同じ意味で目的それ自体であるというわけではない。人間のみが反省的な理性能力を用いて目的を追求するのに対して、動物は本能的に目的を追求するからである。しかしこれは、人間のみが道徳的に扱われるべき目的それ自体であることを含意しない。動物は義務の主体として相互的な立法は不可能であっても、義務の客体である目的それ自体ではあるのだから、我々は動物に義務を負うということだ（pp. 138-145）。このコースガードの理解によれば、「動物が目的それ自体ではない」というカントの主張からは、動物は「道徳法則の立法者ではない」という意味を読み取ることはできても、動物に対するいかなる義務も存在しないという意味を読み取ることはできない。

しかしコースガードの解釈は、カントが理性という能力の特権性を自明視したうえで、人間性の尊厳を論じているという避けられない時代的背景を矮小化しているようにも見える。また、目的それ自体が価値の源泉であり、それは自らにとって価値あるものを求める目的設定能力であること、そのためには有感性を前提するという解釈に支えられているが、それはカント解釈としては議論の余地がある。その意味で、どこまで「カント的」であるのか評価は難しい。それに対して、よりカントのテキストに基づいて動物に対する道徳的配慮を擁護する論調もある。それが「間接義務アプローチ」である。

2. 2　間接義務アプローチ

間接義務アプローチとは、カントが『道徳の形而上学』の第二部である『徳論』や『倫理学講義』で展開する「間接義務」に基づいたアプローチである。大雑把に言えば、間接義務アプローチとは、動物に対する義務はないが、人間に対する義務から間接的に派生して動物への非道徳的な振る舞いが禁止されるため、人間は動物に関する義務を負っているというものである。この論証は、動物の直接的な権利等を擁護することはできないが、我々の動物についての倫理的な振る舞いを（間接的とはいえ）義務として規定できるという点で重要である。

間接義務アプローチを支持する比較的初期の論者であるデニスは、有感性などの基準では道徳的地位をもちえない存在（例えば昆虫）に関しても道徳的義務を規定できるカントの議論を、より広範に及ぶ動物倫理の枠組みとして考えている（Denis 2000）。ほかにも、アルトマンは、シンガーやレーガンの動物擁護の議論と比較しても、カントの間接義務の議論は弱いものではないと考え、「ここでのカントの立場の含意は、動物権利論や動物福祉論と矛盾するものではない」と結論づける（Altman 2014: p. 20）。日本の論者としては、中村が間接義務アプローチを擁護する議論を展開している（中村 2024）。

しかし間接義務アプローチにも問題がある。間接義務アプローチは、たしかに動物の権利などに全く関心を示さないような人に対する議論としては説得的に見えるかもしれない。しかし、動物倫理がまさしく問題にしていることは、動物の道徳的地位である。そうであれば、やはりこのような人間中心的なアプローチでは十分とは言えないだろう。チョルビが指摘するように[3]、間接義務アプローチでは、結局のところ動物を虐待することがいかに悪い性格を表出させるか、あるいは助長するかという事実がなければ、私たちは動物に関する義務をまったく持たないことになる。つまり、間接義務アプローチからは、動物そのものは道徳的配慮の対象ではなく、あくまで私たちが自己に対して負っている義務から派生する手段としてし

3　Cf. Cholbi 2014 p. 342. なお、チョルビも動物に対する直接義務、特に動物福祉を促進する不完全義務を認めることができると主張している。

か動物を配慮する必要がない。その意味で、間接義務アプローチは、私たちの道徳的義務における動物福祉の役割を捉えることはできない。

大きく分けて、既存のカント的アプローチは上記の二つ（直接義務アプローチと間接義務アプローチ）に分類できる。両者に共通するのは、カントのテキストを解釈すれば、そこから動物に対する道徳的配慮を読み取ることができるというものである。しかし両立場とも問題がある。直接義務アプローチは動物を道徳的配慮の対象とするための強力な議論であるがカント解釈として難点を抱えるし、間接義務アプローチはカント解釈としては穏当であるが動物を道徳的配慮の対象とするためには限定的である。

3．Kantianism for Animals とは何か

さて、我々はカント的な方法で動物に対する道徳的配慮を擁護する方向を諦めるべきなのだろうか。いや、諦めるとまでは言わずとも、それは限定的なものにすぎないのだろうか。もう一つの途があるとすれば、それはカントの議論に対して根本的な修正を加えつつ、それでも着想や方法の面で「カント的」であり続けるアプローチをとることであろう。そこで本稿では、ミュラーの Kantianism for Animals に注目したい。

Kantianism for Animals、日本語で言えば「動物のためのカント主義」と名付けられるこのアプローチは、従来人間を対象としてきたカント倫理学を、人間以外の動物にも拡張しようとするもので、動物が道徳的配慮（moral concern）に値する存在であることを提案するものである。ミュラーは、カント倫理学の枠組みを損なうことなく、動物に対する義務を含むようにカントの枠組みを修正しようとするため、カント主義の「建設的（constructive）で修正主義的（revisionist）、そしてラディカルな（radical）」アプローチを提案する（p. 9）。「動物のためのカント主義」という名の下、ミュラーは、動物も人間と同じように道徳的配慮が向けられるべき存在であると考える点で、従来のアプローチにはなかったラディカルさを有している（p. 11-13）。

ミュラーが注目するのは、徳の義務として提示される「同時に義務である目的」（ミュラーは義務的目的（obligatory ends）と呼ぶもの）である。この義務は、行為の命令するものではなく行為の目的の命令するものである。次節で説明するが、カントによれば、同時に義務である目的の中には「他者の幸福」が含まれる。それは、「他者の幸福を促進すること」を自分の目的として採用する義務であり、実践的愛の義務や尊敬の義務として論じられる。ミュラーによれば、この観点こそが動物が道徳的配慮の領域に含まれることを提唱する上で有益である（p. 26）。例えば、既存のアプローチとしての間接義務アプローチは、結局のところ動物の幸福や福利に対して無関心であることを容認してしまうため、動物を道徳的配慮の対象として含めようとする議論にとっては有益なものではない（p. 78-80）。

では、ミュラーのアプローチの特徴とは何か。本稿では大きく以下の四つの観点から紹介し、評価したい。(1) 人間性の定式と義務的目的の解釈（*Kantianism for Animals* 第 4 章）、(2) 義務の方向性に関する解釈（同書第 5 章）、(3) 他者への尊敬の義務に関する解釈（同書第 6-7 章）、(4) 動物の行動の目的に関する解釈（同書第 6 章）。特に (1) はカントに基づいた解釈であるのに対して、(2) (3) (4) はカントの記述を超えて、動物のためのカント主義へと修正を加えたものである。

3. 1　人間性の定式と義務的目的の解釈

ミュラーは、人格を目的それ自体として扱うべきと命じる定言命法の人間性の定式を、道徳的配慮の対象を規定する実質的な原理ではなく、たんに自律的意志を記述した形式的な原理にすぎないと解釈する。それによって、人間性の定式から動物が排除されるとしても、それは道徳的配慮の対象に動物が含まれないことを意味するわけではないと指摘する。まずは、人間性の定式によって動物が道徳的配慮の対象から

排除される論証を確認したい（p. 91）。

1．人間性の定式は、道徳的配慮の範囲を定めている。
2．その範囲は、有限な理性的存在に正確に線引きされる。
3．しかし動物は理性的存在ではない。
4．したがって、動物は道徳的配慮から除外されなければならない。

例えばコースガードは、前提 2 を否定し、道徳的配慮の範囲を有限な理性的存在者、すなわち人間に限定すべきではないと主張する。そして、コースガードは、「目的それ自体」として前提されるべき能力を、「動物性の価値」、すなわち選択対象を追求する能力の価値（自然的な善）であると解釈し、その能力をもつ動物にまで道徳的配慮の範囲を拡張する。一方でミュラーは、1 を否定し、人間性の定式は道徳的配慮の範囲を確定する実質的な原理ではないとする（p. 93）。人間性の定式では、理性的で自律的な存在のみが目的それ自体であるということが明示されるが、それは「目的それ自体が道徳的配慮の対象である」という主張ではなく、それらを混同するのは誤りであると指摘する（pp. 94-5）。

人間性の定式がたんに形式的な原理であるという解釈は、リース以来指摘されている解釈であり（Reath 2013）、ミュラーはそのアプローチを動物倫理の文脈に採用する。こうして、人間性の定式を道徳的配慮の範囲を決定する実質的な原理とみなさないことで、たとえそれが動物を排除するとしても、動物は理性的で自律的な存在ではないという記述が導かれるだけであり、動物が道徳的配慮の対象ではないという結論は導かれないと解釈する[4]。

ミュラーは、道徳的配慮に関する説明を与えるのは人間性の定式ではなく、私たちが採用するべき義務的目的の理論であるという方針をとる。義務的目的とは、理性的存在者である人間が自ら設定するべき行為の目的であり、カントはそれを「自己の完全性」と「他者の幸福」の二つであると述べる（Ⅵ : 385-6）。ミュラーの戦略は、カントによる他者への道徳的配慮の説明は他者の幸福を促進するという義務的目的を中心に展開されると考え、この幸福促進の対象に動物も含まれると解釈するということである（pp. 115-6）。そうであれば、道徳的配慮の範囲は、それが理性的存在者であるかどうかではなく、幸福を感じる能力があるかどうかで引かれることになる。こうして、動物が「目的それ自体」であるかどうかという問題から、動物の幸福が他者の幸福という義務的な目的の一部と見なされるべきかどうかという問題へと視点が移る（p. 122）。ミュラーによれば、カントの道徳的配慮の対象に動物を含めるためには、「人間性の定式」を形式的原理として解釈することで道徳的配慮の範囲を実質的に決定するものではないと切り離すこと、そして、他者の幸福という義務的目的が道徳的配慮に関する実質的原理を与える、という解釈をとる必要がある。

3. 2　義務の方向性に関する解釈

カントは、義務が向けられるのは、道徳法則を共有できる自律的な理性的存在のみであるという相互人間的（interpersonal）な義務論をもっていたと通常考えられている。つまり、道徳法則を互いに共有できる立法主体のみが、相互に義務によって拘束することができる、ということである。もしその理解が正しいのであれば、道徳法則の立法主体ではない動物は義務の対象に含まれない。そこで、ミュラーは義務の方向性をめぐるカントの議論を修正することを試みる。

4　ミュラーは、人間性の定式が自律的意志の純粋に形式的な原理であるという解釈が唯一の正しい解釈であると考えているわけではない。その読み方が可能であり、動物のためのカント主義に有用であるという意味でその解釈を採用している（p. 122）。

ミュラーは、それぞれの自律的主体がなぜ同じ道徳法則のもとに拘束されるのかをカントは説明できないという批判をするトンプソンの議論を紹介することからはじめる（p. 127）。トンプソンの批判は、道徳法則を共有するという二項性を説明するには、共有された同一の法のもとにある（契約など）という前提が必要だが、カントはその前提を説明しないというものである。これに対して、ダーウォルのように二人称的権威の観点から義務の方向性を説明しようとする応答もありえるが、それは道徳法則を自律的に立法するという見解と相いれず、カント的応答としては不十分である。そこでミュラーは、義務の方向性を相互立法や二人称的観点から解釈することを退ける。

ミュラーは、たとえ他者が道徳法則を共有していなくても、他者についての義務は可能であると考え、一人称的なカント的見解（first-personal Kantian view）を提案する（p. 135）。この見解によれば、義務の方向性を説明するために道徳法則の共有や二人称的な権威は必ずしも必要ではなく、自律的な主体という一人称的な観点から、義務の方向性あるいは対象を説明することができるというものである。

では、一人称的な観点から義務の方向性をどう説明するのか。ミュラーは、義務の方向性とは、義務が何について（about）のものであるか、という問題であると考え、義務の内容あるいは実質にかかわると解釈する（p. 137）。義務の実質として考えられるのは、義務的目的である「自己の完全性」と「他者の幸福」である。義務的目的は私たちが自ら設定するべきものであるから、他者の幸福についての義務は、私たちが他者から要求されることによって発生する義務ではなく、あくまで義務的目的の設定主体によって自律的に導かれる義務であると解釈できる。そうであれば、他者の幸福を促進することを目的とする、という義務について、二人称的権威や共有立法という前提をする必要がない。この点で、他者に対する義務は一人称的な観点から導かれる義務とみなすことができる。こうして、道徳法則を立法し、相互に拘束するという条件を義務の対象から外すことによって、動物を義務の対象から排除する論証を退けることができる。

3.3　他者への尊敬の義務に関する解釈

伝統的に、カント倫理学においては、道徳的被行為者（moral patients）すなわち道徳的配慮の対象は自律的な理性的存在者であると考えられているので、それに動物を含めるためには、その前提を修正しなければならない。そのためにミュラーは、カントが徳の義務として位置付ける「尊敬の義務」の修正的な解釈へと向かう。カントは尊敬の義務の対象を、道徳的行為者、すなわち私たち人間と道徳的に対等な立法主体に限定するため、このままでは動物はその対象から外れてしまうからである。まずミュラーは、そもそもカントの尊敬の義務は、義務の分類上難点を抱えていることを指摘する（p. 154、すでに 2 章でも指摘していた）。カントは尊敬の義務を他者に対する徳の義務だとしているが、その義務が他者の幸福という義務的目的から派生しているようには見えない。さらに、徳の義務は義務的目的を採用する義務であるが、尊敬の義務が規定しているのは他者よりも自分を高めてはいけないという否定的義務であるため、積極的に行為の目的として義務的目的を採用することを命じる徳の義務のように見えない（p. 154, Ⅵ : 449）。そこでミュラーは、尊敬の義務を他者に対する徳の義務として分類しないという方針をとる。

尊敬の義務は、無制限に他者に便益を与えることを要求する愛の義務に対する対抗力であり、引力に対する斥力であると説明される（Ⅵ : 470）。ミュラーはこれを、他者に対して自分を不当に高めてはいけない義務として説明する（p. 154）。この義務は、自己の態度を規定する義務であるから、ミュラーは尊敬の義務を自己に対する徳の義務と解釈する（p. 155）。ミュラー自身が言うように、尊敬の義務を自己に対する義務と解釈するためには、カントから離れる必要がある。

自己に対する徳の義務として尊敬の義務を考えるとどうなるか。自己に関する義務的目的は「自己の完全性」である。これは、義務を遵守する道徳的状態を維持することを意味する（pp. 155-6）。ミュラーは、この自己の完全性という目的に、自己を不当に高めないことを含める。つまり、自己に対する徳の義務と

して、自分が他者より価値があると考え他者を見下すような態度をとることを禁止する義務を位置づける。それゆえ、尊敬の義務は、他者に対して自分を不当に高めてはいけない自己に対する義務として解釈される。これは「内的驕傲」を禁止する義務であり、ミュラーはこの義務を「不驕傲（non-exaltation）」と呼ぶ（p. 156）。

では、この修正を経て、私たちは動物に関しても「不驕傲の義務」を負うかどうかを考える必要がある。もし負うのであれば、尊敬の義務によって規定される態度を、動物にも向けなければならないことになるからである。ミュラーは、動物はカント的な意味で道徳的行為者ではないので、そこに平等性や対称性はないが、動物よりも自らを不当に高めてはならないという義務が人間の側には存在しうると主張する（p. 158）。動物はたしかに道徳的行為者である人間と同じ存在ではないが、それは優劣を示すわけではない。というのも、動物は道徳的に善い行為も悪い行為も行わないため、人間より劣っていると評価されるべきではないからである（p. 158）。それゆえ、動物に対して不当に自らを驕傲することは許容されない。よって、尊敬の義務の議論によって、動物が配慮の対象から排除されることはない。

3. 4　動物の行動の目的に関する解釈

カントの他者に対する徳の義務とは、他者の目的（幸福）を自分の目的にすること、そしてその目的が達成されるよう援助することである。その対象となる他者とは、道具的な実践理性を用いて、自分の幸福を目的として設定することができる存在である（p. 159）。しかし、カントは動物が自分で目的を設定して幸福追求することを認めないため、そのままでは義務の対象となりえない。たしかにカントは（デカルトと違って）、動物にも心的状態があることは認めるが、かれらが自身の幸福を道具的な理性を用いて推論的に洞察し、それを追い求めるとは考えていない。そこでミュラーは、他者の目的がその存在にとっての幸福を示唆する機能（indicative function）があればよいと考える（p. 160）。動物はかれらの幸福という目的を道具的理性によって示唆することはなくとも、生物的機能性（biological functionality）によって示唆している（p. 160）。動物の行動は、かれらの健康と繁栄のための生物学的機能を果たしており、それはかれらの幸福と密接な関係にあるため、私たちが配慮するべき動物も幸福追求の主体である。それゆえ、動物の幸福を促進することは、主として動物自身の幸福追求を助けることであるという見解が導かれる。

動物も自身の幸福に関心をもち、それを求める存在であることは示されたが、目的に向かって行為するための行為者性（agency）をどう考えるべきか。カントはたしかに動物にも選択意志（arbitrium brutum）を認めるが、それは人間のような自由の能力とは違うため、少なくとも人間のような仕方で目的設定・追及はできない。カントによれば、動物は表象と概念の能力を使えないからである。しかしミュラーによれば、動物の行為は本能という非概念的原因によって引き起こされ、人間とは違う仕方ではあるが目的を非概念的な仕方で表象し追求することはできる（pp. 162-3）。つまり、動物は対象を個別化して表象することはできないが、非概念的な感覚的な意識（awareness）を持つ。ここから、動物は「曖昧な目的（obscure ends）」をもつことができると説明される（p. 167）。つまり、カント的な狭いテクニカルな意味では動物は目的をもたないが、非概念的で曖昧な目的をもちうるのである。

次に、動物の曖昧な目的は人間の義務にとって重要なのかどうかが検討される。ミュラーの枠組みでは、その存在が自身の幸福を目的として目指しているかどうかが問題となる。たしかに動物は、人間のように道具的理性を用いて自身の幸福を目的として追求するわけではない。そこでミュラーは、幸福を示唆する機能を実践的意味と生物的意味の二つに分け、後者のみを動物に認めるという戦略をとる（p. 169）。動物の行動は、実践的な意味での幸福を「目指す」ことはできないかもしれないが、自然的あるいは生物学的な機能性の問題として、幸福を「目指す」ことはできる。こうして、動物の行動は、動物の幸福と密接に関連した生物学的目的を実際に果たしていることを認めることで、動物の目的も私たちが配慮するべきも

のになる（p. 172）。

4．Kantianism for Animals の特徴の評価

Kantianism for Animalsの枠組みは、既存のカント的アプローチや動物倫理と何が違うのだろうか。まず、コースガードなどの既存のカント的アプローチは、価値（目的それ自体）を第一とするのに対して、ミュラーは義務を第一とする点である（p. 178）。つまり、道徳的配慮の根拠としてカントが理性的存在に与える価値にコミットせず、代わりに自律的意志に基礎を置く義務的目的に焦点を当てる。コースガードは、目的それ自体としての価値を規範の源泉と考え、動物も目的それ自体という価値をもつ存在であると解釈することで、動物を道徳的配慮の対象とする。

一方でミュラーは、動物が重要なのは、動物が道徳的価値を示すからだとは考えない。ある存在が道徳的配慮に値する存在であるということは、私たち理性的存在者の義務の内容にある特定の仕方で関与していること、すなわち、「その幸福が問題となる存在（being with a happiness at stake）」として位置付けられることを意味する（p. 177）。また、レーガンも義務論的なアプローチをとるが、動物固有の価値を根拠にして道徳的配慮を与えるという点では、ミュラーとは対照的な立場にある。ミュラーはこのことを端的に「義務第一、価値第一ではない（duty first, not value first）」と表現している（p. 178）。

次に、主流な動物権利論との重要な違いは、ミュラーの見解は権利よりも義務が第一にあるというカント的立場に基づいているという点である。道徳的権利や請求から義務を認識するのではなく、義務の拘束性が第一義的なのである。これはミュラーのアプローチをカント的にしている特徴であり、これは「義務は権利や請求に優先される（Duties over rights and claims）」と表現される（p. 179）。つまり、他人の権利に応答する仕方で道徳的配慮をする義務があるという発想ではなく、義務の拘束性は自律（一人称的）から生じ、その内容として他者の幸福促進（道徳的配慮）があるという発想である。

ミュラーの Kantianism for Animals は、動物を道徳的配慮の対象に含めるためのカント主義の修正を目的としているため、厳密にカントの記述に基づくわけではない。この試みは、カント的な動物倫理に関する新しい魅力的な議論を提供するのと同時に、どこまでカントを使う必然性があるのか、という懸念も抱かせるはずである。本稿はその懸念として、「義務の相互性」と「動物の目的」という二点に言及する。

ミュラーは、尊敬の義務を自己に対する不驕傲の義務と解釈することで、動物に対しても不驕傲という態度をとる義務があると述べた。すなわち、ミュラーのカント主義では、尊敬の義務によって動物が義務の対象から排除されないということを示したことになる。しかし、カントが『徳論』で尊敬を愛との対比で「斥力」と表現するときの主眼は、「他人の尊厳を奪うことの禁止」にある（Ⅵ：450）。それゆえ、愛の義務は相互に接近する「引力」であるのに対して、尊敬の義務は相互の距離を保つ「斥力」となる。そして、尊厳の根拠となるのは、それが目的それ自体として扱われるべき自律的な存在だからである（Ⅵ：462, Ⅳ：436）。そうであれば、ここにはミュラーが拒絶した目的それ自体としての人間性の内なる人格に認められる絶対的価値という問題が再浮上するように思われる。尊敬の義務が命じるものが、自己に対する不驕傲の義務だとしても、尊敬の対象である他者は尊厳という価値を有しているという前提をカントから切り離すことになれば、カントの枠組みから想像よりも大きく離れることになりかねない。

次に、動物の曖昧な幸福追求を、「同時に義務である目的」としてみなすことが可能であるかは、議論の余地がある。カントはなぜ他者の幸福を目的とした親切が義務になるのかを説く際に、それぞれの主体が他者に助けてほしいという欲求をもつことを根拠にしている。カントが言うには、「われわれの自己愛は、他人からも愛されたい（困窮時には助けてほしい）という欲求と切り離しえないから、われわれは自分を他人の目的としているのであって、そしてこの格率はただ普遍的法則としての資格を有することによってだけ、したがって、他者をもまたわれわれの目的としようとする意志によってだけ、ひとを拘束すること

ができるのであるから、他者の幸福は、同時に義務である目的なのである」（Ⅵ：393）。この一節においてカントは、自分の幸福が他者の目的となることが義務としての拘束力をもつためには、自分もまた他者の幸福を自分の目的としなければならないと主張している。したがって、他者の幸福が義務的目的とされるためには、その他者もまた同様に道徳法則の下にある対等な存在であるという前提が含まれていると考えられる。つまり、「同時に義務である目的」として他者の幸福が追求されるのは、その幸福の担い手である他者もまた、道徳的に行為することが可能な存在である、という平等性が前提されていることになる。そうであれば、動物が人間と平等な道徳的行為者性を持つことを認めない限り、動物の幸福も義務的目的であることを擁護するのはできないように思われる。ミュラーはこの問題については、少なくとも本文の中では対処していない。

結　論

本稿は、動物倫理への新たなカント的アプローチとして、ミュラーが提示する Kantianism for Animals を評価してきた。カント倫理学から肯定的に動物倫理を論じる議論は、ミュラーを筆頭に増えてきてはいるものの、いまだ少数であり、十分なものとはいえない。その状況の中で提示されたミュラーの試みは、これまでのカント的な動物倫理にはない方向性を示したという点で、大きな貢献であることは間違いない。しかし、コースガードや間接義務の議論に比べると、ミュラーが提示した枠組みはまだ評価が定まっておらず、新たな第三の枠組みであると言い切ることは難しい。本稿をきっかけに読者が増え、議論が活発になることを願っている。

謝　辞

本稿は、2024 年 3 月に北海道大学にて開催されたワークショップ「道徳的地位の再考：動物倫理とその先へ」における発表内容に基づいています。本発表は、竹下昌志氏（発表当時：北海道大学、現在：名古屋大学）との共同発表として行われたものであり、ともに議論を深める機会を持てたことに深く感謝いたします。また、当日の議論においてならびに本稿の執筆にあたって貴重なご助言やご示唆を賜ったすべての方々に、心より御礼申し上げます。

参考文献

Altman, M. C. (2011). *Kant and applied ethics: The uses and limits of Kant's practical philosophy*. Malden: Wiley-Blackwell.

Camenzind S. (2021). Kantian Ethics and the Animal Turn. On the Contemporary Defence of Kant's Indirect Duty View. *Animals: an open access journal from* MDPI, 11(2), 512. https://doi.org/10.3390/ani11020512

Cholbi, M. (2014). A direct Kantian duty to animals. *Southern Journal of Philosophy*, 52(3), 338–358. https://doi.org/10.1111/sjp.12079

Denis, L. (2000). Kant's conception of duties regarding animals: Reconstruction and reconsideration. *History of Philosophy Quarterly*, 17(4), 405–423.

Egonsson, D. (1997). Kant's vegetarianism. *The Journal of Value Inquiry*, 31(4), 473–483. https://doi.org/10.1023/A:1004295530520

Garthoff, J. (2020). Against the construction of animal ethical standing. In J. J. Callanan & L. Allais (Eds.), *Kant and animals* (pp. 191–212). Oxford University Press.

Herman, B. (2018). We are not alone: A place for animals in Kant's ethics. In E. Watkins (Ed.), *Kant on persons and agency* (pp. 174–191). Cambridge University Press.

Kant, I. (1908). *Kants Gesammelte Schriften*, Hrsg. von der Königlichen Preußischen Akademie der Wissenschaften und Nachfolgern.

Korsgaard, C. M. (2018). *Fellow creatures: Our obligations to the other animals*. Oxford: Oxford University Press.

Müller, N. D. (2022). *Kantianism for animals: A radical Kantian animal ethic*. New York: Palgrave Macmillan.

—— (2022). Korsgaard's Duties towards Animals: Two Difficulties. *Relations: Beyond Anthropocentrism*, 10(1), 9–26.

Regan, T. (1983). *The case for animal rights*. Berkeley: University of California Press. (Original work published 1983)

Singer, P. (2002). *Animal liberation*. HarperCollins.

Timmermann, J. (2005). When the tail wags the dog: Animal welfare and indirect duty in Kantian ethics. *Kantian Review*, 10, 127–149. https://doi.org/10.1017/S1369415400002168

Wood, A. W. (1998). Kant on duties regarding nonrational nature. *Aristotelian Society Supplementary Volume*, 72(1), 189–210. https://doi.org/10.1111/1467-8349.00042

中村涼（2024）.「「自然物に関する義務」再考——カント義務論に見る環境倫理の可能性——」『フィロソフィア』（111），93-105.

書評

Višak, T. (2022).

Capacity for Welfare across Species. Oxford University Press.

竹下昌志（北海道大学）、篠崎大河（慶應義塾大学）、
佐々木健人

人間と人間以外の動物（以下、動物）では、達成しうる幸福の最大値は違うのだろうか。Tatjana Višak は本書を通じて、人間を含めた動物間で、達成しうる福祉の程度に根本的な違いがあるかどうかを検討している。Višak は、持ちうる福祉の量の違いを福祉能力（capacity for welfare）として考え、次の二つの立場（difference view（以下 DIF）と equality view（以下 EQU））を比較検討する。（以下、ページ数は本書のページ数を表す）

DIF：ヒト、イヌ、マウスなどの福祉主体は、認知的あるいは情動的能力の違いによって、福祉能力が根本的に異なっている。（p. 10）

EQU：福祉主体は、認知的または情動的能力が異なるにもかかわらず、福祉能力は根本的に同等である。（p. 12）

これまでの多くの議論では DIF、つまり、認知的・情動的能力の違いによって福祉能力に違いがあるという見解が支持されてきた。これに対し Višak は、EQU、つまり、認知的・情動的能力が違うにもかかわらず福祉能力は同等であるという見解が、少なくとも DIF と同程度に説得的であることを擁護している。以下、各章の概要を見る。

第 1 章では福祉の種間比較の問題が哲学で議論されてないことを指摘し（1.1 節）、本書の議論の仮定（1.2 節）と DIF と EQU の見解を要約（1.3 節）している。仮定について、本書では（1）福祉主体が誰であるかは福祉の理論に依存すること、（2）福祉をある程度定量的に測定できること、（3）異種間の福祉を単一の尺度で比較できることが仮定される。

第 2 章は 6 節からなり、各節で DIF の立場をとる理論を紹介し、それぞれを批判する。紙幅の関係で、ここでは特に影響力があると思われる Kagan（2.1 節）、Singer（2.2 節）、Budolfson & Spears（2.6 節）の理論に対する Visak の批判を見る。2.1 節では、Kagan の議論を検討する。それによると、高度な認知能力を持つものだけが得られる善が存在する。例えば、美的経験は、美的な感覚能力を持つものだけが得られる。これに対し Višak は、何が善であるかは種によって異なると主張して反論する。例えば、ヒトが美しいと感覚するものをそう感覚しない生物は、善を得られていないのではなく、その生物にとっては単にそれが善ではないのである。2.2 節では、Singer の議論を検討する。それによると、ウマとしての生と、ヒトとしての生を比べたとき、中立的な合理的存在者は、ヒトの生を好む。なぜなら、より大きな度合いの自己意識を持つ方が選好されるという点でよいからである。これに対し Višak は、ヒトがウマとしての生を想像することや、ウマでもヒトでもない中立的な観点に立つことは難しいと指摘し、Singer の直観に疑義を呈する。2.6 節では Budolfson & Spears の議論を取り上げる。かれらは、ある動物の福祉能力をその認知能力によって定めることで福祉の経験的な基準を得ようとするが、Višak は、かれらが福祉と経験的な基準の関係を明らかにしていないと批判する。

第 3 章では EQU を支持する議論をしている。まず 3.1 節では動物福祉科学の見解を紹介し、そこでの福祉の測定が、基本的に EQU と親和的であると述べている。例えば、動物福祉の代表的な基準である「5 つの自由」について、どの種に属する動物もそれらの自由に対して同じような福祉をもつとされるため、EQU と親和的であるとされる。3.2 節では Haybron の自己充足としての福祉理論を紹介し、これが EQU 的に解釈できることを示している。Haybron の理論では、福祉は、(1) 幸福、(2) アイデンティティに関係したプロジェクトの成功、(3) 自己と無関係な本性の充足（健康、活力、身体的快楽）の三つから構成される。Višak は、これらの福祉の構成要素はすべて福祉主体である動物も持つものであること、また「充足」を割合で解釈することで、EQU に沿った説明として解釈できると議論する。3.3 節で Višak は、EQU を擁護するための進化論的論証を提示する。そこで、Višak は「快楽性（hedonicity）」という概念を導入する。それは、快から不快までの主観的スケールで経験を評価する福祉主体の能力である。ここで重要なのは、快楽性が伝える情報が状況に相対的であるということである。例えば、以前から多くの報酬を定期的に得ている状況で同じ量の報酬を得るより、以前から少ない報酬を定期的に得ている状況で多い報酬を得た方が快楽が強い。快楽性が相対的であることは EQU を支持する。というのも、快楽性が相対的なら、ある種が「きわめて快い」から「きわめて不快」までの快楽能力を持つのに、別のある種が「やや快い」から「やや不快」までだけの快楽能力しか持たないということはもっともらしくないからである。

第 4 章では寿命が異なる動物同士での福祉の比較について論じている。4.1 節では、ある個体にとって短い生涯がより長い生涯よりも有害であるかどうかに関して、剥奪説、欲求不充足説、McMahan の時間相対的利害説という諸理論の観点から、前提となるメタ倫理学的主観主義／客観主義にも触れつつ考察している。4.2 節では、寿命が異なる動物種同士の生涯福祉を比較する方法について、共時的福祉の合計として生涯福祉を導出する全期間説と、個体の寿命をその動物種の自然寿命と相対化する寿命相対化説の観点から考察し、前者を支持している。4.3 節では、全期間説の応用として、遺伝子編集技術を用いて寿命が短くなった動物を産みだし苦痛なく殺すことが許容可能かどうか検討している。ここでは人口倫理の議論を参照しつつ、寿命を短くすることは生涯福祉を小さくすることになるため、全期間説からは短命の存在を生み出すことは悪いと議論している。

第 5 章では EQU の応用について議論している。5.1 節では DIF と EQU の議論が道徳的地位の議論についてどのような含意を持つかを検討し、5.2 節では EQU に基づくと奇妙な実践的含意が出るという批判を検討している。5.3 節では、ここまでの議論を考慮した上で、EQU が少なくとも DIF と同程度に説得的であると結論し、私たちは動物の福祉をヒトの福祉と同等にカウントするべきだと主張している。

次に総評を述べる。まず良い点を三つあげる。

第一に、本書は短くまとまっており（全体で 152 ページ）、議論もわかりやすく、読み通しやすい。

第二に、本書は動物福祉論に関するサーベイ文献として有益である。例えば第 2 章は、動物福祉論の諸見解に関する有益なサーベイにもなっている。また、Kagan や Singer といった著名な論者だけでなく、Wong のような一般には広く知られてはいない論者も取り上げられているため、動物福祉論に明るい読者にとっても有用であろう。

第三に、これまでの多くの議論では認知的・情動的能力の違いが福祉能力に違いをもたらすこと（つまり DIF）がほぼ自明視されてきたが、Višak はそのような議論を詳細に検討し、EQU を排除できるほどではないことを説得的に議論している。本書の目標は、DIF が間違いで EQU が正しいことを擁護することではなく、DIF と EQU が同程度に説得的であることを示すことである。このような穏当な目標設定によって、Višak の負う論証責任が軽くなり、DIF 側に論証責任を課すことに成功しているように見える。

次に批判点を二つ述べる。第一に、EQU は意識に度合いがあるという見解と両立するか明らかではな

いということである。EQU が正しければ、福祉能力は度合いを持たない、すなわち、福祉主体の福祉能力は根本的に同等である。ここで、どの生物が福祉主体なのか、という問いが立てられうる。これに対し、Visak は意識を持つ生物（意識主体）が福祉主体であると素朴に前提しているように思われる（3.3 節）。しかし、意識の哲学においては、意識は度合いを持つと考えられている[1]。これが、福祉能力に度合いを許容しない EQU とどのように両立させられうるかは明らかではない。

第二に、ある特定の福祉論からは DIF の方が説得的になりうると考えられることである。例えば、単純な選好説で、認知能力の高さによって選好の持てる最大数が違うことで、DIF を支持できるかもしれない。Višak は、選好充足と不充足のバランスによって福祉が決まるのであり、選好の内容は無関係であるという単純な選好説が EQU を支持すると考えている（p. 30）。しかし、このような立場では、持ちうる選好の最大数は重要である。もし認知能力が高いほど選好の最大数も高くなるのであれば、達成可能な福祉の程度もそれに伴って高くなることが予想される。例えば、選好を 1 個しか持てない貝と、100 個持てる人間とで、どちらもすべての選好が満たされているなら、人間の福祉の方が貝の福祉より高くなるだろう。したがってこのような立場では DIF の方が支持されると思われる。

こうした問題点があるにせよ、福祉の異種間比較という動物倫理学における伝統的な問題を、近年の福祉論や動物の心の哲学の知見をベースに議論している点で、本書は重要な貢献をしていると考えられる。

1　Lee, A. Y. (2022). Degrees of consciousness. *Noûs,* 57(3), 553–575.

北海道大学応用倫理・応用哲学研究教育センター

『応用倫理 ── 理論と実践の架橋』第 17 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 17 号への掲載は 2025 年 9 月 30 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）の電子媒体を本センター事務局宛にメールで送付すること。
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、北海道大学応用倫理・応用哲学研究教育センターのウェブサイト及び北海道大学学術成果レポジトリウェブページにより公開する。
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ので、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・応用哲学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://caep-hu.sakura.ne.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院　応用倫理・応用哲学研究教育センター

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

応用倫理 ── 理論と実践の架橋　vol. 16
2025 年 6 月 30 日発行

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院
応用倫理・応用哲学研究教育センター

Tel：011-706-4088
E-mail：caep@let.hokudai.ac.jp
URL：http://caep-hu.sakura.ne.jp/