



大阪大学  
社会技術共創研究センター

## ELSI NOTE

### No. 06

# 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

2020 年 10 月 28 日

#### Author:

石田 柁（東京大学大学院総合文化研究科・博士後期課程）

※ 本 ELSI ノートに関わる文献調査は、日本学術振興会『課題設定による先導的人文学・社会科学研究推進事業』（領域開拓プログラム）「日本学術振興会 RRI の新展開のための理論的・実践的研究—教育・評価・政治性に注目して」（代表：標葉隆馬）性に注目して」（代表：標葉隆馬）ならびに JST 社会技術開発センター『人と情報のエコシステム』研究開発領域「情報技術・分子ロボティクスを対象とした議題共創のためのリアルタイム・テクノロジーアセスメントの構築」（代表：標葉隆馬）の一環として行ったものである。

## はじめに

神経科学・脳科学の研究の進展とともに、その成果の利用のありかたとして道徳・倫理的によいものはどのようなものかへの関心が高まっている。これは、もちろんひとつには科学研究一般についていえる論点である。他方で、神経科学・脳科学に特有の(もしくは他の研究領域に比べて神経科学・脳科学にとっての重要度が高い)道徳・倫理的問題がある。道徳的エンハンスメント(moral enhancement)にかかわる諸問題がその一例である。

本ノートでは、このような道徳的エンハンスメントの問題を考える上でとりわけ重要なテーマである「潜在的バイアス」について行われてきた倫理学・哲学分野における議論・視点に注目し、整理を行う。そのため、本ノートの第1部では、道徳的エンハンスメントの対象としてしばしば考えられる潜在的バイアス(implicit bias)に着目し、これにかかわる哲学的議論を整理する。第2部では、道徳的エンハンスメントそのものにかかわる哲学的議論を概観する。

## 第1部 潜在的バイアス

潜在的バイアスは、主に差別との関連で道徳的に問題のある現象だと考えられている。しかし、潜在的バイアスは、当の本人がそのバイアスを自覚ないしコントロールできるかどうか疑問が呈されている。

この点により、潜在的バイアス(を反映した言動)への道徳的によい対応が何かについては意見が分かれている。ひとつの立場は、主に心理学の知見にもとづいて、バイアスの自覚ないしコントロールが一定程度できる——それゆえ通常の差別的言動と同じように対応すべきだ——とする立場である。もうひとつの立場は、もっぱら倫理学の議論にもとづいて、仮に本当に自覚やコントロールができないとしても非難や帰責が適切でありうるとする立場である。

### 1.1. 潜在的バイアスの特有の論点とは何か

- **潜在的バイアス(implicit bias)**とは、「スティグマ化された社会集団のメンバーシップにもとづく人々を対象とする、無意識的・自動的になされる否定的評価の傾向」(Washington & Kelly 2016: 17)である。
  - 潜在的バイアスは、バイアス一般ないし偏見と同様に、差別との関連で道徳的に問題のある現象だと考えられている。差別にかかわる心理・認知的状態は、他の多くの心理・認知的状態とは異なり、それが「悪い」ということについて論争の余地がかなり小さいと考えられる(Harris 2011)。これによって、潜在的バイアスの軽減・除去が本当に「エンハンスメント」だといえるかどうかという問題を回避して、エンハンスメントの道徳的性質を論じることができる。
  - ✓ たとえば、「いかなるときも感情に左右されずに思考できるようにする」という介入を考える場合、このことがエンハンスメントにあたるかどうかは、対象となる人自身の考えに依存するかもしれないし、また感情が人間にとってどれほど大切かに関する一般的な道徳的立場にも依存しうる。そのため、このような介入の道徳的性質を論じるとき

には、これがそもそも「向上」なのかどうかという論点に加わることになる。

- ✓ 一方で、潜在的バイアスにかかわる介入は、エンハンスメントに関わるこのような論点を回避できると考えられている。ただし、後にみるように、たとえ差別にかかわるバイアスの道徳的な悪さが疑い得ないものであるとしても、そのバイアスが「潜在的」であることを理由として生じる道徳的論争がなされてきた経緯は見過ごしてはならない。
- 潜在的バイアスの典型的特徴としては、(1)それをもつ人が公言している立場と相反する内容でありうること、(2)それをもつ人の内省によってはバイアスを自覚できないこと、(3)それをもつ人の努力によってバイアスの発動を制御できないこと、(4)それをもつ人の言動や判断の広範囲にわたって影響を与えることが挙げられる(Washington & Kelly 2016)。
  - 潜在的バイアスの暫定的定義からみて最もストレートな事例は、潜在的連合テスト(Implicit Association Test; IAT)にかかわる研究から得られたものである。頻繁に言及されるものとして、黒人を白人よりも「危険である」「武器をもっている」等々と強く結びつけるバイアスが我々の多くにあるという研究がある(たとえば Payne 2001)<sup>1</sup>。
  - ほかの典型的事例として、公にはジェンダー平等を本気で支持している人が実際の言動において無意識のうちに女性を冷遇しているとき、この人は女性に対する潜在的バイアスをもっている可能性がある(Brownstein 2019)。これは、口で平等を唱えていても平等主義に真にコミットしていないというような場合に限られない。心の底から平等主義にコミットしている場合であっても言動がそれに対応していないということもありうる。そのような場合に何が問題なのかを考える上で、潜在的バイアスは有益な視点だ<sup>2</sup>。

<sup>1</sup> 人間に限らず、AI の学習が人種差別とその歴史的経緯によって差別的なものに偏ることも指摘されている(e.g. 江間 2019)。

<sup>2</sup> たとえば Washington & Kelly (2016) の思考実験に登場する平等主義者はそのよう

- 潜在的バイアスは、バイアス一般ないし偏見と同様に、差別との関連で道徳的に問題のある現象だと考えられている。しかし、このような考えは、一見したところ理論的に一貫した正当化ができないように思われる。
  - 一方では、自覚的なバイアスや意図的な差別的行為ならまだしも、自分でコントロールできなかつたり気付かなかつたりするところで勝手に生じてしまうバイアスについて道徳的責任を負うべきであるのかについては、それほど自明ではないという指摘がある(Bargh 1999; Saul 2013)。
  - 他方では、自覚的であろうとなかろうと、また意図的であろうとなかろうと、被差別者が不利益を被っていることに変わりがない。その場合、一般的な差別への批判と同様に、潜在的バイアスに起因する差別的言動を咎められる、あるいは潜在的バイアス自体を矯正することが道徳的に求められることは(一定程度)正当性があるとも考えられる。
- 潜在的バイアスに起因する差別的言動を咎めたり潜在的バイアス自体を矯正したりすることを道徳的に正当化する道筋として、おおむね以下の二つの議論がなされてきた。
  - 【第一の議論】たしかに、コントロールできなかつたり自覚できなかつたりするものに起因する言動については、道徳的に悪かつたり非難対象であつたりすることがない。だが、人は自らの潜在的バイアスをコントロールしたり自覚したりすることができる。それゆえ、潜在的バイアスに起因するという理由で差別的言動が許されはしない(本ノート の 1.3 節)。
  - 【第二の議論】たしかに、人は自らの潜在的バイアスをコントロールしたり自覚したりすることができない。だが、コントロールや自覚ができないものが原因だとしても道徳的に悪かつたり非難対象になつたりすることがある。それゆえ、潜在的バイアスに起因するという理由で差別的言動が許されはしない(本ノート の 1.4 節)。

---

な人である。

## 1.2. 潜在的バイアスが「勝手に」生じたものであるのはいかなる意味においてか？

- 潜在的バイアスは顕示的バイアス (explicit bias) と対比され、何らかの意味で「勝手に」生じるものであることを基本的特徴とする。そして、1.1 節で見たように、この点が潜在的バイアスにかかわる特有の道徳的問題を引き起こす。
- しかし、バイアスが「勝手に」生じたものであるとは厳密にはどういうことか。おおむね三つの考え方がある。

① 第一に、バイアスをもつ本人がそのバイアスの発生を制御ないしコントロールできないことに注目する見方がある。この点は上の定義で触れられている。

- 【再掲】潜在的バイアス: 「スティグマ化された社会集団のメンバーシップにもとづく人々を対象とする、無意識的・自動的になされる否定的評価の傾向」(Washington & Kelly 2016: 17)
- 自分がある判断をする上で、その判断に自分のバイアスが反映されないようにすることができていれば、その人は自分のバイアスをコントロールできているといえる。そのうえで、どのようなコントロールが問題になるのか。たとえば、直接的で即座のコントロールができなければいけないのか、それとも間接的で長期的なコントロールができればよいのか (Holroyd 2012; Faucher 2016)。
  - 詳細は 1.3 節を参照されたい。
- バイアスをもつことをコントロールできるかどうかと、それが自らの言動に反映されることをコントロールできるかどうかは、別の問題である (Holroyd 2012)。

② 第二に、バイアスをもつ本人がそのバイアスの発生に自覚的でないことに注目する見方がある。これも上の定義で触れられている。

- どのような形式の自覚が問題になるのかが論点となる(Holroyd, Scaife & Stafford 2017)。
    - 自覚の形式の詳細は 1.3 節を参照されたい。
  - バイアスをもっていることに自覚的であるかどうかと、それが自らの言動に反映されることに自覚的であるかどうかは別の問題として議論される(Holroyd 2012)。
- ③ 最後に、バイアスの内容が、そのバイアスをもつ本人が真にコミットしているものとは異なるということに注目する見方がある。
- この考え方は深い自己(deep self)説や真の自己(real self)説などと呼ばれ、道徳的責任が問われるための条件についての議論でしばしば提案される(Faucher 2016; Glasgow 2016)。
  - 仮に、本人が公言しているものと本人が自動的ないし無意識にやっていることが真っ向から対立している状況を考えよう。そのうちどちらが「真の自己」だといえるだろうか。本人が公言しているもののほうが「真の自己」だと考えることには根拠がない(Fischer 2012; Faucher 2016)。むしろ、口ではジェンダー平等を主張している人が無自覚で男性を優遇しているとき、その無意識のふるまいについて「本性が出ている」と判断することはさほど不自然なことではないように考えられる。

### 1.3. 潜在的バイアスは実際にはコントロールや自覚の対象になるという立場

- 先に触れたように、潜在的バイアスは「スティグマ化された社会集団のメンバーシップにもとづく人々を対象とする、無意識的・自動的になされる否定的評価の傾向」(Washington & Kelly 2016, 17)として理解される。

- 潜在的バイアスにかかわる道徳的問題において重要なのは、問題となる否定的評価が「無意識的」ないし「自動的」になされるという点である。言い換えれば、問題となる否定的評価やそれに反映される可能性のあるバイアスを当人が自覚できたりコントロールできたりするか、これが問われる。
- 問題となる否定的評価にかかわる何もかもが「無意識的」ないし「自動的」であるかという点、必ずしもそうではないという考えがある。では、人はどういうことを自覚できたりコントロールできたりするのが問題となる。
- バイアスをもつこと(having implicit biases)をコントロールないし自覚できるかどうかと、それを自らの行動や判断において表すこと(manifesting implicit biases)をコントロールないし自覚できるかどうかは別の問題である(Holroyd 2012)。
- 「コントロール」について
  - バイアスをもつことと、それを自らの行動や判断において現すことの両方について、それぞれ、直接的で即座のコントロールと、間接的で長期的なコントロールがある。とくに、直接的で即座のコントロールの対象にならないが間接的で長期的なコントロールの対象になるようなものがある。たとえば、すくなくとも現時点においては、自分の語学力は、直接的で即座のコントロールの対象にはならないが、間接的で長期的なコントロールの対象にはなりうる例として挙げられる(Holroyd 2012)。
  - 潜在的バイアスをもつことを直接的で即座にコントロールすることは、できそうにない。潜在的バイアスが自らの行動や判断に反映されるかどうかの直接的で即座のコントロールは、できるかもしれないが、そのコントロールをやめた途端にバイアスが強く現れること(リバウンド効果)が知られており、コントロール可能性を強調すべきかどうかは自明でない(Holroyd 2012)。
  - 潜在的バイアスをもつことの間接的で長期的なコントロールは、1.4 節でみるようなしかたでできるかもしれない。潜在的バイアスが自らの行動や



判断に反映されるかどうかの間接的で長期的なコントロールとしては以下の三つのものが考えられる(Holroyd 2012)。

1. 潜在的バイアスを自らの行動や判断に反映させないような環境に身を置くことでコントロールができる。
  2. 実行意図(implementation intentions)すなわち「XであるならばYする」のような意図をもつことによって、潜在的バイアスの言動への反映が軽減されるということが示唆されている。
  3. 公言する信念や目標が潜在的バイアスのありかたに影響を与えるということが示唆されている。
- 「自覚」について：自覚とはどういうことかについて、以下の三種類を区別できるのだが、これまでは理論的研究でも経験的研究でもあまり区別されてこなかった(Holroyd 2015)。
    1. 内省的自覚(introspective awareness)：自分自身のことをふりかえって、自分は潜在的バイアスをもつと思うかどうか
    2. 推論的自覚(inferential awareness)：一般的傾向として自分(を含む人間一般)は潜在的バイアスもちがちだと思うかどうか
    3. 観察的自覚(observational awareness)：自分の行動の結果に、潜在的バイアスを疑わせるような不一致があると思うかどうか
    - 人は自分が潜在的バイアスをもっていることを自覚できそうにない。しかし、自分の潜在的バイアスが自分の行動や判断に反映されることであれば自覚できるということが示唆されている(Holroyd 2012)。

## 1.4. 勝手に生じたものであっても道德的問題は消えないという立場

- 第一に、潜在的バイアス以外の事例に目を向ければ、勝手に生じたものに起因したとしても道德的問題が消えないというのはおかしい話ではない。たとえば、

機嫌が悪い状態でおこなった言動についても、道徳的に悪かったり非難対象であったりする。潜在的バイアスもそれと同じと考えられてきた(Madva 2017)。

- そもそも、道徳的な悪さや適非難性(非難に値するかどうか)が行為者のコントロール可能性や自覚に依存しないとする考え方は倫理学における主流な立場の一つである。典型的には帰結主義<sup>3</sup>が挙げられる。潜在的バイアスがかかっても同様であり、生じうる帰結が深刻であれば行為者側の事情にかかわらず道徳的に悪かったり非難対象であったりする(Glasgow 2016)。
  - ✓ 差別がかかわる問題については、行為者側の事情を問わない道徳理論の強みが近年注目されている。というのは、意図しない差別・間接差別・統計的差別・機械学習がかかわる差別的アウトカムについて考えるときにこの点が有用だからである(Lippert-Rasmussen 2013; Barocas & Selbst 2016; Zheng 2016)。
- 第二に、意図的・自覚的な言動とそうでない言動はどちらも道徳的に悪かったり非難対象であったりする。しかしながら、そうした言動に対する適切な対応にはいくつかのパターンがありうる。たとえば次のような区別が考えられる。
  - ある言動をした人はそれについて責任がある(being responsible)という点は両者に共通する。しかし、その言動についてその人の責任を問う(holding responsible)ことが適切であるのはその言動が意図的・自覚的である場合に限られるかもしれない(Washington & Kelly 2016; Brownstein 2019)。
  - ある言動をした人はそれについて説明責任(accountability)があるという点は両者に共通する。しかし、その言動をその人への帰属(attribution)

---

<sup>3</sup> 帰結主義とは、平たくいえば、ある行為の道徳的ステータス(よい行為かどうか、許される行為かどうか、非難に値する行為かどうかなど)はその行為がもたらす帰結(のみ)によって決まるとする立場である。より厳密な定式化や、上で示したものより複雑なバージョンの帰結主義の説明は、本ノートの趣旨に照らしておこなわない。

が適切であるのはその言動が意図的・自覚的である場合に限られるかもしれない(Zheng 2016; Watson 2004)。

- ✓ 責任のバリエーションとしての「説明責任」と「帰属」は、論者ごとにいろいろなしかたで使われる概念対である。ひとつのわかりやすい区別としては、帰属を「これはあなたがやったことだ」やそれに類する過去志向的責任概念と考え、説明責任を生じた帰結の解決や改善のコストを負うことにかかわる未来志向的責任概念と考えるという道筋がある(Zheng 2016; Holroyd, Scaife & Stafford 2017)。
- 第三に、潜在的バイアスにかかわる言動自体については道徳的責任が問われないとしても、別のしかたで道徳的責任が問われるかもしれない。
  - 潜在的バイアスについて知っておく道徳的義務(Washington & Kelly 2016; cf. Sher 2009)を考えることができる。たとえば、潜在的バイアスについての知見がまったくない 1980 年代の人々と、それらに容易にアクセスできる 2020 年の我々とは、この意味での道徳的義務の有無ないし程度が違いただろう。仮にそうだとすると、1980 年代の人々が潜在的バイアスに起因して無意識に男性を優遇するのと、2020 年を生きる我々が潜在的バイアスに起因して無意識に男性を優遇するのでは、他の条件が等しくても、道徳的責任の程度が異なることになる。
  - 潜在的バイアスを解消するような環境にいる道徳的義務(Rees 2016)を考えることができる。これは、自動的・無意識的な反応を徳(virtue)のひとつとみなして、潜在的バイアスという悪徳(vice)を解消するような環境に身を置くことを人々に求める立場である。ただし、これは個人の道徳的義務ではなく集団(共同体)の道徳的義務である。つまり、人々が潜在的バイアスをもたないようにしたり解消したりすることを促す環境づくりが、集団のメンバー全員に対して求められる。
  - 潜在的バイアスをもつことの道徳的問題とは別に、バイアスを言動に反映させることの道徳的問題を考えることができる(Holroyd 2012; 本ノートの 1.3 節も参照されたい)。



## 第2部 脳神経エンハンスメント

第2部では、道徳エンハンスメントを念頭におきつつ、脳神経エンハンスメントの道徳的な課題を論じた既存の議論を概観する。加えて、脳神経エンハンスメントへの道徳的な反論については、既になされた再反論を挙げる。

その上で残る脳神経エンハンスメントの道徳的問題は以下の通りである。(1)他の医学的介入と同様に、脳神経エンハンスメントの危険性の評価をエビデンスにもとづいておこなう必要がある。(2)脳神経エンハンスメントの恩恵を誰が受けるかにかかわる政策は、別途道徳的評価の対象となる。以上二点は脳神経エンハンスメントに特有でない一般的問題である。脳神経エンハンスメントに特有でありうるのは次の二点である。(3)結果だけでなく過程がまさに重要である活動については、活動の意義が脳神経エンハンスメントにより損なわれうる。(4)競争相手が脳神経エンハンスメントを受けることによって脳神経エンハンスメントを受けることが事実上強制される可能性を検証する必要がある。

### 2.1. 道徳エンハンスメントと脳神経エンハンスメント

- エンハンスメント(enhancement)または人間エンハンスメント(human enhancement)とは、「人間の形態や機能を、健康を回復したり保ったりするために必要な範囲を超えて向上させるために用いられる生医学的介入」(Juengst & Moseley 2019)である。
- 向上させる機能の種類によって、エンハンスメントはいくつかに分類できる。道徳エンハンスメント(moral enhancement)とは、エンハンスメントのうち道徳的に行動する能力を向上させるものである(Savulescu, ter Meulen & Kahane 2011)。
  - 道徳エンハンスメントと対比されるのは、他の機能を向上させるようなエンハンスメント、たとえば認知(cognitive)エンハンスメント、身体(physical)エンハンスメント、気分(mood)エンハンスメント、寿命(lifespan)エンハンスメントである(Savulescu, ter Meulen & Kahane 2011)。
  - 道徳エンハンスメントをよりはっきり特徴づけると、「単に教育したりトレーニングをしたり道徳的向上を動機づけたりするための戦略ではなく、道徳的向上を

- 工作 (engineering)・設計 (programming)・強制するための戦略」(Harris 2016)であるといえる。
- 認知エンハンスメントの典型例は、集中力を高める薬(現実の例としてはリタリンなど)の投与である。同様にして、道徳エンハンスメントの典型例としては、共感能力を上げる薬の投与や攻撃性を抑える薬の投与が考えられる。
    - ✓ ただし、第 1 部の最初に触れた問題がある。それは、こういう介入が本当に道徳的能力を向上させているかという問題である
    - ✓ たとえば、共感能力を高めるのが道徳エンハンスメントだというのは、共感することが道徳的によいことだというのを前提している。一方、共感ではなく理詰めで考えることこそ道徳的によいと考える立場からすれば、共感能力の向上は道徳的観点からみればむしろ悪くなっているかもしれない。
    - ✓ とはいえ、ある介入が本当に道徳的「エンハンスメント」だ(つまりその介入結果が道徳的によい)といえるかどうかは、介入をすること自体が道徳的によいかどうかとは別問題である。ここでは後者を扱う。
  - 第 1 部で差別との関連で潜在的バイアスを扱ったのは、まさに、他の事例に比べて「本当に道徳的観点からみて『エンハンスメント』になっているのか、むしろ悪くなっているのではないか？」という問いを回避しやすいとされるからである。脳神経エンハンスメント(neuroenhancement: Roskies 2016)というのは、介入する器官による分類名である。当然、他の器官に介入するようなエンハンスメントも考えられる (Juengst & Moseley 2019)。
    - 脳神経エンハンスメントによって向上しうる機能には、認知的なもの、道徳的なもの、さらには身体的なもの(典型的にはブレイン・コンピュータ・インターフェース)がある (Roskies 2016)。
    - エンハンスメントと治療 (treatment) の違いについては、専門家の伝統的区分に依拠するもの、「通常の機能」に依拠するもの、および介入の根拠となる疾患

を特定できるかどうかには依拠するものがある(Juengst & Moseley 2019)。どの説にもそれぞれ問題がある<sup>4</sup>。

## 2.2. 賛否(概観)

- 「脳や神経への介入による道德エンハンスメント」を考えたとき、既存の議論は、エンハンスメント一般にかかわるもの、(介入する器官を限定して)脳神経エンハンスメントに特有の問題を扱ったもの、および(向上させる機能を限定して)道德エンハンスメントに特有の問題を扱ったものに分けられる。ただし、常に明示的に区別して論じられてきたわけではない<sup>5</sup>。
- 脳神経エンハンスメントを許容可能だとする主な立場には以下のものがある(Roskies 2016)。
  - 非特別性: 脳神経エンハンスメントは、我々が既に慣れ親しんでいる教育や科学技術と質的に変わらない(Greely et al. 2008)。
  - 認知的自由(cognitive liberty): 自分自身の心理的状态や認知をどうするかについて我々には自由があり、エンハンスメントによってその機能を向上させることもそこに含まれる(Sententia 2013)。
    - ✓ 関連する立場として、変容の自由(morphological freedom)の一例として我々には脳神経エンハンスメントをおこなう権利があると考えることができる(Sandberg 2013)。

---

<sup>4</sup> ただし、これはエンハンスメント一般にかかわる問題であること、および本ノートの主題(潜在的バイアスおよびそれへの介入の道德的性質)に直接かかわるものではないから、深入りしない。

<sup>5</sup> 以下では、おおむね脳神経エンハンスメントにかかわる議論をまとめるけれども、必ずしも上の区別に完全に則るわけではない。したがって、エンハンスメント一般の問題や道德エンハンスメントに特有の問題を含む点に留意されたい。

- 功利主義的理由: 認知的エンハンスメントは、コストを考慮して、よい帰結をもたらす (e.g. Sandberg & Savulescu 2011)。
- 脳神経エンハンスメントを許容不可能だとする主な立場とそれへの応答は以下のとおりである。
  - 危害: 脳神経エンハンスメントには特有の危険性がある。たとえば依存性が挙げられる (Heinz et al. 2012)。
    - ✓ 可能な応答: たしかに、脳神経エンハンスメントが道徳的に許容されるためには、エビデンスにもとづいてそのリスクが明らかにされなければならない。しかし、それは他の医学的介入の場合と同様であって、脳神経エンハンスメントに特有のことで、安全性の基準が脳神経エンハンスメントの場合にだけ高いわけでもない (Greely et al. 2008)。
  - 不自然さ: 脳神経エンハンスメントは不自然な能力向上方法だ。
    - ✓ 可能な応答: 我々が既に受け入れている医学的介入をはじめかなりのものが「自然」でないのに、脳神経エンハンスメントだけ自然性に訴えるのはおかしい (Greely et al. 2008)。
    - ✓ したがって、単に人工物であることをもって脳神経エンハンスメントを道徳的に許容不可能だと断じることは、かなり難しい。脳神経エンハンスメントが「不自然だ」と言いたい反対者が本当に考えているのは、次の「人間性の毀損」のようなことだろう (President's Council on Bioethics 2003)。
  - 人間性の毀損: 脳神経エンハンスメントは、人間性にとって重要なものを損なう。たとえば、努力を通じた目標達成をショートカットすることを可能にしてしまう (President's Council on Bioethics 2003)。
    - ✓ 可能な応答: 目標達成のためのショートカットはありふれており脳神経エンハンスメントを特別視する理由にならない (Farah 2002)。



- ✓ 目標達成のためのショートカットが問題になるのは、結果ではなく過程がまさに重要である場合に限られる。そういう場合には、脳神経エンハンスメントのショートカットとしての側面が倫理的に問題だという余地がある(Schermer 2008)。
  - 努力(effort)自体にはどういう価値があるか、また達成(achievement)が価値をもつためには努力が必要かについては、哲学的議論の蓄積がある(cf. Bradford 2015)。本ノートでは深入りしない。
- 不平等(1): 脳神経エンハンスメント技術が使われた社会では、認知的能力が高い人が低い人を見下すようになる(Lippert-Rasmussen 2012)。
  - ✓ 可能な応答: まず、これは事実にかんする根拠のない推測だ。また、仮にこれが正しいとしても、他人を平等なしかたで扱う傾向性を高めるような道德エンハンスメントをすればよい(Lippert-Rasmussen 2012)。
- 不平等(2): 他の新しい技術と同様に、脳神経エンハンスメントの恩恵を実際に受けられるのは既に恵まれた人々だ。つまり、脳神経エンハンスメントは既存の不平等を拡大させるおそれがある(Greely et al. 2008; Academy of Medical Sciences 2012)。
  - ✓ 可能な応答: 脳神経エンハンスメントの恩恵を誰が受けるかは、医療政策にかかわる道德的問題であって、脳神経エンハンスメント技術自体の問題ではない。とくに、ニーズや費用対効果によって分配方法を考えるなら、脳神経エンハンスメントはたとえば認知的能力が著しく低い人々にまず向けられるので、この技術を使うと既存の不平等が悪化するというのは自明ではない(Lippert-Rasmussen 2012)。
- 強制(1): 他の医療技術と同じように、脳神経エンハンスメントが公衆衛生の一貫として強制される(Roskies 2016)。

- ✓ 可能な応答: ワクチンの強制接種が正当化されるのと同じ理由で、ある特殊な場合には脳神経エンハンスメントを受けることの強制も正当化されるかもしれない(Roskies 2016)。
  - ✓ 一部の職業(外科医、航空機のパイロットなど)について、集中力を高めるために脳神経エンハンスメントを受けることが義務付けられるということはあるだろうか。そのような強制は、脳神経エンハンスメントの効果にかんするエビデンスが乏しい現状では自己決定権を侵害する程度があまりに大きく、正当化されにくい。そもそも、コーヒーの摂取でさえ義務化している雇用者がいないことから、現実的ではない(Maslen, Faulmuller & Savulescu 2014)。
  - ✓ このように、そのようなストレートな強制は、現実には考えにくい。現実的問題になりそうなのは、次のような事実上の強制だ(Roskies 2016)。
- 強制(2): 競争相手がみな脳神経エンハンスメントを受けているとき、競争力を保つために、望まないエンハンスメントを受けることが事実上強いられる(Greely et al. 2008; Academy of Medical Sciences 2012; Maslen, Faulmuller & Savulescu 2014)。
- ✓ 可能な応答: 本当にそうなるかどうかは社会科学の研究によって検証されるべきだ(Maslen, Faulmuller & Savulescu 2014)。

## 参考文献

Academy of Medical Sciences. 2012. *Human Enhancement and the Future of Work*. London: Academy of Medical Sciences.

Bargh, John A. 1999. 'The cognitive monster: The case against controllability of automatic stereotype effects'. In *Dual Process Theories in Social Psychology*, edited by Shelly Chaiken and Yaacov Trope, 361–82. New York, NY: Guilford Press.

- Barocas, Solon and Andrew D. Selbst. 2016. 'Big Data's Disparate Impact'. *California Law Review* 104 (3): 671–732.
- Bomann-Larsen, Lene. 2012. 'A Liberal View on Liberal Enhancement'. In *The Posthuman Condition*, edited by Kasper Lippert-Rasmussen, Mads Rosendahl Thomsen, and Jacob Wamberg, 103–114. Aarhus: Aarhus University Press.
- Bradford, Gwen. 2015. *Achievement*. Oxford: Oxford University Press.
- Brownstein, Michael. 2019. 'Implicit Bias'. *The Stanford Encyclopedia of Philosophy* (Fall 2019 Edition), edited by Edward N. Zalta.  
<https://plato.stanford.edu/archives/fall2019/entries/implicit-bias/>
- Brownstein, Michael and Jennifer Saul, eds. 2016. *Implicit Bias and Philosophy, volume 2: Moral Responsibility, Structural Injustice, and Ethics*. Oxford: Oxford University Press.
- 江間有沙. 2019. 『AI 社会の歩き方——人工知能とどう付き合うか』化学同人.
- Farah, Martha J. 2002. 'Emerging Ethical Issues in Neuroscience'. *Nature Neuroscience* 5 (11): 1123–29.
- Faucher, Luc. 2016. 'Revisionism and Moral Responsibility for Implicit Attitudes'. In Brownstein and Saul, *Implicit Bias and Philosophy Volume 2*, 37–61.
- Glasgow, Joshua. 2016. 'Alienation and Responsibility'. In Brownstein and Saul, *Implicit Bias and Philosophy Volume 2*, 115–44.
- Greely, Henry, Barbara Sahakian, John Harris, Ronald C. Kessler, Michael Gazzaniga, Philip Campbell, and Martha J. Farah. 2008. 'Towards Responsible Use of Cognitive-Enhancing Drugs by the Healthy'. *Nature* 456 (7223): 702–705.
- Harris, John. 2016. *How to Be Good: The Possibility of Moral Enhancement*. Oxford: Oxford University Press.

- Heinz, Andreas, Roland Kipke, Hannah Heimann, and Urban Wiesing. 2012. 'Cognitive Neuroenhancement: False Assumptions in the Ethical Debate'. *Journal of Medical Ethics* 38 (6): 372–75.
- Holroyd, Jules. 2012. 'Responsibility for Implicit Bias'. *Journal of Social Philosophy* 43 (3): 274–306.
- Holroyd, Jules. 2015. 'Implicit Bias, Awareness and Imperfect Cognitions'. *Consciousness and Cognition* 33: 511–23.
- Holroyd, Jules, Robin Scaife, and Tom Stafford. 2017. 'Responsibility for Implicit Bias'. *Philosophy Compass* 12 (3): e12410.
- Juengst, Eric and Daniel Moseley. 2019. 'Human Enhancement'. *The Stanford Encyclopedia of Philosophy* (Summer 2019 Edition), edited by Edward N. Zalta. <https://plato.stanford.edu/archives/sum2019/entries/enhancement/>
- Kelly, Daniel and Erica Roedder. 2008. 'Racial Cognition and the Ethics of Implicit Bias'. *Philosophy Compass* 3 (3): 522–40.
- Lippert-Rasmussen, Kasper. 2012. 'Treating Symptoms Rather than Causes? On "Enhancement" and Social Oppression'. In *The Posthuman Condition*, edited by Kasper Lippert-Rasmussen, Mads Rosendahl Thomsen, and Jacob Wamberg, 88–102. Aarhus: Aarhus University Press.
- Lippert-Rasmussen, Kasper. 2013. *Born Free and Equal? A Philosophical Inquiry into the Nature of Discrimination*. Oxford: Oxford University Press.
- Madva, Alex. 2018. 'Implicit Bias, Moods, and Moral Responsibility'. *Pacific Philosophical Quarterly* 99 (51): 53–78.
- Maslen, Hannah, Nadira Faulmüller, and Julian Savulescu. 2014. 'Pharmacological Cognitive Enhancement: How Neuroscientific Research Could Advance Ethical Debate'. *Frontiers in Systems Neuroscience* 8: 1–12.

- Payne, B. Keith. 2001. 'Prejudice and Perception: The Role of Automatic and Controlled Processes in Misperceiving a Weapon'. *Journal of Personality and Social Psychology* 81 (2): 181–92.
- President's Council on Bioethics. 2003. *Beyond Therapy: Biotechnology and the Pursuit of Happiness*. Washington, DC: President's Council on Bioethics.
- Rees, Clea F. 2016. 'A Virtue Ethics Response to Implicit Bias'. In Brownstein and Saul, *Implicit Bias and Philosophy Volume 2*, 191–214.
- Roskies, Adina. 2016. 'Neuroethics', *The Stanford Encyclopedia of Philosophy* (Spring 2016 Edition), edited by Edward N. Zalta.  
<https://plato.stanford.edu/archives/spr2016/entries/neuroethics/>
- Sandberg, Anders. 2013. 'Morphological Freedom – Why We Not Just Want It, but Need It'. In *The Transhumanist Reader: Classical and Contemporary Essays on the Science, Technology, and Philosophy of the Human Future*, edited by Max More and Natasha Vita-More, 56–64. Blackwell.
- Sandberg, Anders and Julian Savulescu. 2011. 'The Social and Economic Impacts of Cognitive Enhancement'. In Savulescu, ter Meulen, and Kahane, *Enhancing Human Capacities*, 92–112.
- Saul, Jennifer. 2013. 'Implicit bias, Stereotype Threat and Women in Philosophy'. In *Women in Philosophy: What Needs to Change?*, edited by Katrina Hutchison and Fiona Jenkins, 39–60. New York, NY: Oxford University Press.
- Savulescu, Julian. 2012. 'Enhancing Equality'. In *The Posthuman Condition*, edited by Kasper Lippert-Rasmussen, Mads Rosendahl Thomsen, and Jacob Wamberg, 184–203. Aarhus: Aarhus University Press.
- Savulescu, Julian, Ruud ter Meulen, and Guy Kahane, eds. 2011. *Enhancing Human Capacities*. Blackwell.

Schermer, Maartje. 2008. 'Enhancements, Easy Shortcuts, and the Richness of Human Activities'. *Bioethics* 22 (7): 355–63.

Sententia, Wyre. 2013. 'Freedom by Design: Transhumanist Values and Cognitive Liberty'. In *The Transhumanist Reader: Classical and Contemporary Essays on the Science, Technology, and Philosophy of the Human Future*, edited by Max More and Natasha Vita-More, 355–60. Blackwell.

Sher, George. 2009. *Who Knew? Responsibility without Awareness*. Oxford: Oxford University Press.

Washington, Natalia and Daniel Kelly. 2016. 'Who's Responsible for This? Moral Responsibility, Externalism, and Knowledge about Implicit Bias'. In Brownstein and Saul, *Implicit Bias and Philosophy Volume 2*, 11–36.

Zheng, Robin. 2016. 'Attributability, Accountability, and Implicit Bias'. In Brownstein and Saul, *Implicit Bias and Philosophy Volume 2*, 62–89.

ELSI NOTE No. 06

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

2020 年 10 月 28 日

大阪大学 社会技術共創研究センター

〒565-0871 大阪府吹田市山田丘 2-8

大阪大学吹田キャンパステクノアライアンス C 棟 6 階

TEL 06-6105-6084

<https://elsi.osaka-u.ac.jp>