

令和 3 年 5 月 7 日現在

機関番号：14101

研究種目：基盤研究(C) (一般)

研究期間：2018～2020

課題番号：18K00013

研究課題名(和文) 潜在的態度から見た道徳的責任の再検討

研究課題名(英文) Reconsidering moral responsibility from the perspective of implicit attitudes

研究代表者

薄井 尚樹 (USUI, Naoki)

三重大学・人文学部・准教授

研究者番号：50707338

交付決定額(研究期間全体)：(直接経費) 1,600,000円

研究成果の概要(和文)：本研究では潜在的態度から生じる行動の道徳的責任についての研究をおこなった。その成果として以下の点を明らかにした。(1) 潜在的態度のありようは、道徳的責任の帰属可能性を考えるうえで重要な役割を果たす、本当の自己という考えにうまく収まらず、このことは本当の自己という考えの捉え直しを迫る。(2) 潜在的態度のありようをマクロな観点から特徴づけることは、潜在的態度から生じる行動の道徳的責任を考えるうえで新たな可能性を提示する。(3) ある現代的な差別においては潜在的態度がはたらくとされるが、そうした現代的なものを含む差別において道徳的被行為者であるには特定の条件が必要である。

研究成果の学術的意義や社会的意義

潜在的態度と差別的な行動とのつながりが示唆されてきたこれまでの研究動向をふまえて、本研究では、潜在的態度から生じる行動の道徳的責任に焦点をあてた。その成果として、そうした道徳的責任の帰属可能性を明らかにするには、潜在的態度のありようと道徳的責任の両方について、既存の研究とは異なるアプローチが必要があることを明らかにした。また近年、人種差別からもたらされる社会問題があらためて注目されてきており、本研究はそうした問題の道徳的な意味を明らかにする一助となるものである。

研究成果の概要(英文)：In this study, I explored the moral responsibility for behavior caused by implicit attitudes. The results of this study illuminated the following points. (1) The nature of implicit attitudes do not match with the idea of the real self, which assumes a crucial role in clarifying the attributability of moral responsibility. Consequently, we have to review the idea of the real self. (2) Characterizing the nature of implicit attitudes in terms of a macro perspective opens up a new possibility of how we can consider the moral responsibility for behavior resulting from implicit attitudes. (3) It has been claimed that implicit attitudes operates in a contemporary form of discrimination. Moral patiency in discrimination including contemporary one requires particular conditions.

研究分野：哲学

キーワード：潜在的態度 道徳的責任

1. 研究開始当初の背景

潜在的態度と呼ばれる心の状態は、本人の直接的な自己報告ではなく、潜在連合テスト Implicit Association Test (Greenwald, McGhee, & Schwartz, 1998) や感情誤帰属手続き Affect Misattribution Procedure (Payne, Cheng, Govorun, & Stewart, 2005) といった手段によって間接的に測定される。研究開始当初の時点では、潜在的態度の存在からもたらされる哲学的含意をめぐる研究は、まだ初期段階にあったと述べることができる。たとえば、潜在的態度をテーマとするカンファレンスが開かれたのは 2012-2013 年のことであり、その成果が出版されたのは、本研究が開始される 2 年前の 2016 年のことだった (Brownstein & Saul, 2016)。さらに、国内の研究状況に目を向けると、潜在的態度について哲学の観点からアプローチした研究はほぼ皆無と言ってよい状況にあり、ほとんど未開拓の研究分野だった。

潜在的態度をめぐる哲学的研究のなかで、理論的にも実践的にも特に重要なのは、潜在的態度からもたらされる(主に差別的な)行動の道徳的責任にかんする問題である。理論的な観点からすると、自己報告によって測定される顕在的態度と異なり、潜在的態度は、コントロールが難しいことや、ときに本人が気づいていないことといった、道徳的責任の帰属可能性に問題をもたらしかねない特徴を有しているかもしれない点を挙げるができる(この論点については Holroyd, Scaife, & Stafford (2017) を参照)。また実践的な観点からすると、シューター・バイアス (Correll, Park, Judd, & Wittenbrink, 2002) の存在が示唆するように、潜在的態度は現実世界での人種差別にまつわる社会問題と密接なかわりがあるように思われる点を挙げるができる。

2. 研究の目的

以上の研究開始当初の背景をふまえて、本研究の大きな目的を、潜在的態度からもたらされる行動の道徳的責任の帰属可能性という論点に位置づけた。具体的には、以下の 3 つの問題を考察することを通じて研究の目的を達成することを目指した。

(1) 道徳的責任にかんする特定の理論のもとで、潜在的態度からもたらされる行動の道徳的責任を帰属することは可能か。

道徳的責任は哲学において長きにわたり論じられてきたトピックであり、その考えかたについては多岐にわたる (Talbert, 2016)。本研究の目的および研究期間をふまえると、そのすべてを網羅することは難しい。そのため、本研究では、道徳的責任の帰属可能性を本当の自己 real self あるいは深い自己 deep self と結びつけて考える、帰属主義 attributionism と呼ばれる立場 (e.g., Smith, 2005; Sripada, 2016) のもとで、潜在的態度からもたらされる行動の道徳的責任の帰属可能性を考察する(この課題に取り組んだ時点では、同様のアプローチとしてたとえば Brownstein (2016) があった)。

(2) 潜在的態度のありようは道徳的責任の帰属可能性にどのような影響をもたらすか。

潜在的態度が具体的にどのような心の状態なのかについては、ただの連合とみなす立場 (Blair, Ma, & Lenton, 2001) から、信念とみなす立場 (Mandelbaum, 2016) に至るまで、さまざまな見解が存在する。それぞれの立場に応じて、潜在的態度からもたらされる行動の道徳的責任の帰属可能性にかんする考えかたは異なってくるだろう。したがって、潜在的態度のありようにかんする見解と、潜在的態度からもたらされる行動の道徳的責任とのあいだの関係を明確にしなくてはならない。

(3) 潜在的態度がともなう差別的な行動の根底にあるメカニズムと、差別的な行動において道徳的主体であることは、道徳的責任の帰属にどのように寄与するか。

近年、古典的な差別の他にさまざまな現代的な差別のありかたが指摘され、そうした差別ではどのようなメカニズムがはたらいているのか、またときに潜在的態度がそこでどのような役割を果たしているのかが明らかになりつつある (Dovidio, Gaertner, & Pearson, 2017)。それをふまえて、潜在的態度がともなう差別のメカニズムと、そうしたメカニズムのもとでの差別的な行動をめぐる道徳的主体のありかた(道徳的行為者性/被行為者性)は、それぞれ、道徳的責任の帰属にどのように寄与するかを検討する。

3. 研究の方法

本研究は基本的に、先行研究を精査したうえで、潜在的態度からもたらされる行動の道徳的責任の帰属可能性を考察するという点で、文献の調査研究である。それに加えて、学会での発表を精力的におこなうことで、参加者との積極的な意見交換を通じて考察をブラッシュアップし、また参加者からもたらされる知見を採り入れることで研究を推進する。このように、文献の調査研究とそれをもとにした学会発表を繰り返すことで、最終的に論文として研究成果を公表することを目指すという方法を採用した。

4. 研究成果

(1) 道徳的責任にかんする特定の理論のもとで、潜在的態度からもたらされる行動の道徳的責任を帰属することは可能か。

この問いに取り組むために、そもそも潜在的態度は本当の自己とみなされうるのかという点を考察した。まず、本当の自己という考えがどのようなものかを明確にするために、Frankfurt による考察 (Frankfurt, 1988, 1999, 2006) を整理した。そこから「本当の自己」という考えにはふたつの側面があることが明らかになった。ひとつは共時的な側面であり、「本当の自己は... [中略]... 個別的に、あるいは断片的に構築された心的状態の寄せ集めではなく、対立がなく一貫性がともなうかたちで全体的に構成されたものとして理解される」(薄井, 2019, p. 205)。もうひとつは通時的な側面であり、本当の自己はその主題がさまざまな局面で一貫しており動かしがたいと理解される。このようにして、本当の自己を共時的な側面と通時的な側面から特徴づけたうえで潜在的態度について考えると、潜在的態度はどちらの側面も備えているとは言い難いことが明らかになった。しかし、このことから性急に、潜在的態度は本当の自己ではないと論じるのではなく、むしろ、本当の自己という考えのほうを再考する可能性を示唆し、潜在的態度からもたらされる行動の道徳的責任の帰属可能性には、まださらに議論の余地が残されていると結論した。以上の研究成果は、『倫理学年報』掲載論文 (薄井, 2019) として発表された。この論文は、潜在的態度について哲学からアプローチした論文としては、私の知るかぎり、国内でかなり初期に位置づけられると思われる。そのため今後の国内での研究の進展のための足がかりとなるだろう。

(2) 潜在的態度のありようは道徳的責任の帰属可能性にどのような影響をもたらすか。

(1) で結論したように、本当の自己という考えをそのまま潜在的態度に適用することは難しい。このことをふまえると、本当の自己という考えだけでなく、潜在的態度のありようについてもあらためて検討しなくてはならないだろう。本研究では、Payne, Vuletich, & Lundberg (2017) において提示される潜在バイアスのモデルを手がかりに、潜在的態度に社会的な次元をとりいれることで、潜在的態度のありよう、あるいはそこから生じる行動の道徳的責任について、集団/社会の観点から論じることができるのではないかと、いう考察をおこなった。さらに、このように潜在的態度のありようをマクロな観点から考察することで、社会的不正義をただすためには、個人の(ときに潜在的な)態度に焦点をあてるべきか、それとも社会構造を改革すべきか、という対立図式(この対立図式については Madva (2016) を参照)にたいして、そのふたつの立場を包括的に眺めうる視座を提供できることを明らかにした。以上の考察は、日本科学哲学学会第 51 回大会(2018 年) および応用哲学学会第 11 回年次研究大会(2019 年)において発表された。(1) と同様に、国内の哲学系の学会で潜在的態度に焦点をあてた発表としては、私の知るかぎり、かなり初期に位置づけられると思われる。また、これらの学会発表で得た知見をもとに、現在、英語論文の執筆を進めており、雑誌への投稿を予定している。

(3) 潜在的態度がともなう差別的な行動の根底にあるメカニズムと、差別的な行動において道徳的主体であることは、道徳的責任の帰属にどのように寄与するか。

この問いに取り組むにあたっては、回避的レイシズム *aversive racism* (Dovidio & Gaertner, 2004) と呼ばれる、現代的なレイシズムに焦点をあてた。回避的レイシズムは、「意識的で顕在的な平等主義的態度と、非意識的で潜在的な人種バイアスが結びついた組み合わせ」(Dovidio, Pearson, & Penner, 2019, p. 17) として特徴づけられる。このようなレイシズムは、古典的なレイシズムと異なり、つねに差別的な行動をもたらすわけではなく、ある特定の状況のもとで潜在的態度から差別的な行動が生じるとされる (Dovidio & Gaertner, 2004)。本研究では、ロボットへの回避的レイシズムは可能かという問いを通じて、人間でないものに目を向けることで、差別的な行動をめぐる道徳的主体であること(このばあい道徳的被行為者性)、および潜在的態度がともなう差別的な行動のメカニズムが、それぞれ、差別的な行動の道徳的責任の帰属可能性に寄与する境界を明らかにしようとした。この研究における考察は、Gray, Gray, & Wegner (2007) のおこなった心の知覚にかんする研究、および Gray & Wegner (2009) の提案する、モラル・タイプキャスティングという考えに依拠するところが大きい。それによると、道徳的行為者性/被行為者性(また、それにともなって道徳的責任の帰属可能性)は、独自の種類の道徳的知覚に依存するとされる。これらの研究に依拠することで、かりに潜在的態度がともなう回避的レイシズムの形式がロボットにあてはまるとしても、ロボットに道徳的被行為者性を認めることは難しいことから、差別が成立するとは述べがたいと結論した。以上の研究成果は、日本科学哲学学会第 52 回大会(2019 年)において発表され、またその発表を経て、『人文論叢: 三重大学人文学部文化学科研究紀要』掲載論文 (薄井, 2020) として公表されている。

今後の展望としては、(2) で論じたようなかたちで潜在的態度のありようをマクロな観点から特徴づけることからもたらされる哲学的問題について、さらに考察を進めていきたいと考えている。具体的には、3 つの問題に整理することができるだろう。第一に、潜在的態度をそのようにマクロな観点から特徴づけることによって、潜在的態度から生じる行動の道徳的責任は帰属可能なものとなるのか。第二に、帰属可能なものとなるとして、そのばあい道徳的責任が帰属

される主体はどのようなものだと考えられるのか。第三に、そのような特徴づけは、帰属主義においてカギとなる本当の自己という考えにどのような捉え直しを迫ることになるのか。以上の問題を検討することが、今後の研究において重要な課題になると考えている。

参考文献

- Blair, I., Ma, J., Lenton, A. (2001). Imagining stereotypes away: The moderation of implicit stereotypes through mental imagery. *Journal of Personality and Social Psychology*, 81(5), 828-841.
- Brownstein, M. (2016). Attributionism and moral responsibility for implicit bias. *Review of Philosophy and Psychology*, 7(4), 765-786.
- , & Saul, J. (Eds.) (2016). *Implicit Bias and Philosophy. Volume 1-2*. Oxford University Press.
- Correll, J., Park, B., Judd, C., & Wittenbrink, B. (2002). The police officer's dilemma: Using ethnicity to disambiguate potentially threatening individuals. *Journal of Personality and Social Psychology*, 83(6), 1314-1329.
- Dovidio, J., & Gaertner, S. (2004). Aversive racism. *Advances in Experimental Social Psychology*, 36, 1-52.
- , Gaertner, S., & Pearson, A. (2017). Aversive racism and contemporary bias. In C. Sibley & F. Barlow (Eds.), *The Cambridge Handbook of the Psychology of Prejudice* (pp. 267-294). Cambridge University Press.
- , Pearson, A., & Penner, L. (2019). Aversive racism, implicit bias, and microaggressions. In G. Torino, D. Rivera, C. Capodiupo, K. Nadal, & D. Sue (Eds.), *Microaggression Theory: Influence and Implications* (pp. 16-31). Wiley.
- Frankfurt, H. (1988). *The Importance of What We Care About*. Cambridge University Press.
- (1999). *Necessity, Volition, and Love*. Cambridge University Press.
- (2006). *Talking Ourselves Seriously and Getting it Right*. Standard University Press.
- Gray, H., Gray, K., & Wegner, D. (2007). Dimensions of mind perception. *Science*, 315(5812), 619.
- Gray, K., & Wegner, D. (2009). Moral typecasting: Divergent perceptions of moral agents and moral patients. *Journal of Personality and Social Psychology*, 96(3), 505-520.
- Greenwald, A., McGhee, D., & Schwartz, J. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, 74(6), 1464-1480.
- Holroyd, J., Scaife, R., Stafford, T. (2017). Responsibility for implicit bias. *Philosophy Compass*, 12(3), e12410.
- Mandelbaum, E. (2016). Attitude, inference, association: On the propositional structure of implicit bias. *Noûs*, 50(3), 629-658.
- Madva, A. (2016). A plea for anti-anti-individualism: How oversimple psychology misleads social policy. *Ergo, an Open Access Journal of Philosophy*, 3(27), 701-728.
- Payne, B., Cheng, C., Govorun, O., & Stewart, B. (2005). An inkblot for attitudes: Affect misattribution as implicit measurement. *Journal of Personality and Social Psychology*, 89(3), 277-293.
- , Vuletich, H., & Lundberg, K. (2017). The bias of crowds: How implicit bias bridges personal and systemic prejudice. *Psychological Inquiry*, 28(4), 233-248.
- Smith, A. (2005). Responsibility for attitudes: Activity and passivity in mental life. *Ethics*, 115(2), 236-271.
- Sripada, C. (2016). Self-expression: A deep self theory of moral responsibility. *Philosophical Studies*, 173(5), 1203-1232.
- Talbert, M. (2016). *Moral Responsibility: An Introduction*. Polity.
- 薄井尚樹 (2019). 「潜在的態度は「本当の私」なのか」『倫理学年報』68, 201-214.
- (2020). 「ロボット / AI は差別の対象となりうるのか」『人文論叢：三重大学人文学部文化学科紀要』37, 15-26.

5. 主な発表論文等

〔雑誌論文〕 計2件（うち査読付論文 1件/うち国際共著 0件/うちオープンアクセス 2件）

1. 著者名 薄井尚樹	4. 巻 37
2. 論文標題 ロボット/AIは差別の対象となりうるのか	5. 発行年 2020年
3. 雑誌名 人文論叢：三重大学人文学部文化学科研究紀要	6. 最初と最後の頁 15-26
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 無
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

1. 著者名 薄井尚樹	4. 巻 68
2. 論文標題 潜在的態度は「本当の私」なのか	5. 発行年 2019年
3. 雑誌名 倫理学年報	6. 最初と最後の頁 201-214
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

〔学会発表〕 計3件（うち招待講演 0件/うち国際学会 0件）

1. 発表者名 薄井尚樹
2. 発表標題 ロボット/AIにたいする態度はなにを意味するのか
3. 学会等名 日本科学哲学会第52回大会
4. 発表年 2019年

1. 発表者名 薄井尚樹
2. 発表標題 潜在的態度は独自の道徳的責任をもたらすのか
3. 学会等名 応用哲学会第11回年次研究大会
4. 発表年 2019年

1. 発表者名 薄井尚樹
2. 発表標題 群衆バイアスモデルは帰属可能性としての道徳的責任を配慮しうるか
3. 学会等名 日本科学哲学学会第51回大会
4. 発表年 2018年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------