

Title	自己欺瞞はいかにして可能か
Sub Title	How is it possible to be self-deceived ?
Author	浅野, 光紀(Asano, Koki)
Publisher	三田哲學會
Publication year	2003
Jtitle	哲學 No.109 (2003. 3) ,p.23- 46
JaLC DOI	
Abstract	In this essay, I shall deal with self-deception in its most literal sense, namely "to deceive oneself intentionally into believing what one knows to be false". Admittedly, this phenomenon has long been recognized to be psychologically perplexing and philosophically paradoxical. How could one person knowing the truth get himself to believe the opposite and, as a result, simultaneously hold the contradictory beliefs ? Moreover, the knowledge of one's own intention to deceive oneself should make the project self-defeating because no one can be deceived if he knows the deceiver's intention and motives whether or not they are one and the same person. What follows is an attempt to resolve these problems and clarify the nature and etiology of self-deception. Examining carefully a variety of ways in which desire can influence belief formation by biasing our cognitive processes, I propose a coherent description of self-deception which is real and common in our daily life. In due course it will be shown that our ability to know our own mental states is limited and we can be mistaken about and unaware of the reasons on which we act.
Notes	投稿論文
Genre	Journal Article
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=AN00150430-00000109-0023

慶應義塾大学学術情報リポジトリ(KOARA)に掲載されているコンテンツの著作権は、それぞれの著作者、学会または出版社/発行者に帰属し、その権利は著作権法によって保護されています。引用にあたっては、著作権法を遵守してご利用ください。

The copyrights of content available on the Keio Associated Repository of Academic resources (KOARA) belong to the respective authors, academic societies, or publishers/issuers, and these rights are protected by the Japanese Copyright Act. When quoting the content, please follow the Japanese copyright act.

— 投 稿 論 文 —

自己欺瞞はいかにして可能か

— 浅 野 光 紀* —

How is it possible to be self-deceived ?

Koki Asano

In this essay, I shall deal with self-deception in its most literal sense, namely “to deceive oneself intentionally into believing what one knows to be false”. Admittedly, this phenomenon has long been recognized to be psychologically perplexing and philosophically paradoxical. How could one person knowing the truth get himself to believe the opposite and, as a result, simultaneously hold the contradictory beliefs ? Moreover, the knowledge of one’s own intention to deceive oneself should make the project self-defeating because no one can be deceived if he knows the deceiver’s intention and motives whether or not they are one and the same person.

What follows is an attempt to resolve these problems and clarify the nature and etiology of self-deception. Examining carefully a variety of ways in which desire can influence belief-formation by biasing our cognitive processes, I propose a coherent description of self-deception which is real and common in our daily life. In due course it will be shown that our ability to know our own mental states is limited and we can be mistaken about and unaware of the reasons on which we act.

自己欺瞞はわれわれに親しい現象であり、いくつかの事例を列挙するの

* 慶應義塾大学大学院文学研究科博士課程（哲学）

はさして困難ではない。アルコール中毒であると医者に宣告された人が、そんなことはない、自分の意識はしっかりしてるし酒を手控えることも可能だと自分を納得させようとする。恋人が浮気している可能性が濃厚なのに、浮気しているならあんなに公然とやるはずがないなどと自分を宥めすかし、相手の身の潔白を信じてしまう。才能のないスポーツ選手が、自分に秀でたライバルの欠点をあれこれ探し回り、自分と同等か、さらにはそれ以下の存在とみなそうとする、等々…。

こうした事例が示しているのは、自分にとって望ましくない事態、特にわれわれの自尊心やプライド、こんな人間でありたいというセルフ・イメージを脅かすような事態に直面した時に、われわれが見せる反応の一つである。世界はいつもわれわれに親しげな表情を浮かべているとは限らず、自分の願いが適わなかったり、恐れていた事態が出来たりする。それがにわかに承服しがたいとき、われわれは正反対の事実を捏造しようと躍起になり、その結果、時にはそれを本気で信じてしまう。そうすれば、少なくとも心の安寧が得られるからである。

自己欺瞞は馴染み深い現象だが、それを整合的に説明するのは容易ではない。われわれの抱く信念や欲求や意図、そして行為が複雑に連携し、一見パラドキシカルな状況を形作っているように見えるからである。本論では、自己欺瞞が示している哲学的な問題を抽出し、その解決策を模索していきたい。その過程で、信念や欲求、意図と行為のような人間の心にまつわる諸概念、また自己知のあり方などについても考察していくことになるはずである。

1. 問題の提示：二つのパラドクス

冒頭に挙げたような事例について、その共通する特質について考えてみると、自己欺瞞という呼び名がいかに適切なものに思われてくる。通常の欺瞞、要するに他人を騙す場合に行われているのと同じ手続きが、そこ

でも踏まれているように見えるからである。真実を隠蔽し、詐術を弄して逆の偽なる命題を相手に信じ込ませる。自分に対して同じことをすれば、それは自己欺瞞になるというわけである。自己欺瞞においても、自分にとって望ましくない真実（「私はアル中だ」「恋人が浮気している」「私は才能がない」）が隠蔽され、逆に、自分にとって都合のいい、しかし偽なる命題（「私はアル中ではない」「恋人は潔白である」「私は優秀だ」）が信じられるに至っているからである。

これは正しいが、しかし、自己欺瞞を構成する最も基本的な事実の一つに思いを致すなら、どうしてこのようなことが可能なのか疑問に思えてくる。すなわち、自己欺瞞においては、騙す側と騙される側が同一人物であるという事実である。

通常の欺瞞、普通に他人を騙す場合について考えてみよう。そこでは、騙す側と騙される側は分裂している。両者は別人格である。騙す側の人には、ある命題を、それが偽であると知りながら相手に告げ、相手がそれを信用してしまったときに欺瞞は成立する。騙す側は真実を知りながら、自分では信じていないことを相手に植え付けるのである。欺瞞が成立したとき、真なる信念と偽なる信念はそれぞれ別人格に帰属されている。つまり真なる信念は、それをもともと知っている騙す側の人に、そして偽なる信念は、それを植え付けられた騙される側の人に。真なる信念と偽なる信念のこの分離が、欺瞞を成立させる条件である。騙される側の人があるとき真実に気づいていたら、彼は騙されることはなかったはずである。

しかし自己欺瞞において起こっているのは、まさにこの不可能に思われる事態にはかならない。自分にとって望ましくない真理に気づき、それを信じてしまった一人の人間が、それを真っ向から否定する偽なる信念を自分に植え付けて心の安寧を得る。真理を知りつ、それを否定する偽なる命題をも信じること。そのとき人は、互いに打ち消し合う、矛盾する二つの信念を同時に抱くに至っているのではあるまいか。最初に挙げた例に即し

ていえば、自己欺瞞に陥っている人は、「私はアル中だ」と信じかつ「私はアル中ではない」と信じているように見えるのである。しかし一人の人が、同一の話題に関して、相矛盾する信念を抱くなどということが可能なのだろうか。真実を知らながらどうして自分を騙すことができ、またその結果、矛盾する二つの信念を抱くなどということが可能なのだろうか。自己欺瞞が提示するこの第一の問題を、ここでは「信念のパラドクス」と呼ぶことにしよう。

さて、矛盾する信念を抱くことが仮に可能だとしても、自己欺瞞にはもう一つの問題が存在する。こちらは「意図のパラドクス」とでも呼ぶべきものである。再び、通常の欺瞞、他人を騙す場合について考えてみよう。他人を騙すという行為には、様々な動機がつきものである。金銭目当てか、あるいはエイプリル・フールの冗談か、とにかく何らかの理由で相手に間違った考えを吹き込もうとするのである。そのとき騙す側の人は、その間違った考えをあたかも自分でも信じているかのように振る舞うだろう。自分は真摯に、相手のためを思って本当のことを告げているかのように装うだろう。騙そうという意図を相手に見抜かれたら、欺瞞はそこで失敗するからである。自分の本当の意図、相手を騙そうという意図を、先方に気づかれないようにする。これが欺瞞を成功に導く条件の一つである。こうしたことが可能なのも、騙す側と騙される側が別々の人間だからにほかならない。人が何を欲し、意図しているのか、当人以外には直接知る術がなく、他人は振る舞いから類推する以外にないからである。

さて、こうした考察を自己欺瞞の場合に当てはめてみると、それがまたしても不可能な過程に思われてくる。騙す側と騙される側が同一人物であった場合、前者の意図は後者に筒抜けのはずではないか。自分の意図は自分が一番よく知っているはずではないか。われわれは、自分自身の心の状態に関しては、特権的な知識をもっているはずではないか。自己欺瞞に至る過程とはこうであった。自分にとって望ましくない事態、それだけは

起こらないでほしいと恐れていた事態に直面した人が、それを否認して逆の事態を信じようとする。要するに自分を騙そうとする。そうすれば心の安寧が得られるからである。自分の欲するところを信じて安心したいというのが自己欺瞞を誘発する動機の一つであり、従って自己欺瞞もまた、この動機を実現するために意図的になされるものであると思われるのである。しかし自分を騙そうという意図を抱いてしまったら、まさにその時点で、自己欺瞞の企みは挫折するのではないだろうか。相手が他人ならともかく、自分の意図は自分から隠しようもないからである。そのとき人は、自分が単に自分を安心させたがっていることに気づいてしまう。現実の事態は、自分が信じようとしていることとは正反対であることが暴露されてしまう。自己欺瞞とは、意図的なものであるかぎり、その成立の条件をみずから奪っている過程としか言いようがない。自己欺瞞が提示するこの第二の問題を、「意図のパラドクス」と呼ぶことにしよう⁽¹⁾。

以下、本稿では、以上二つの問題の解決策を模索していくことにしよう。

2. 心の分割論について：若干のコメント

この問題を解決するためには、対照的な二つの道があるように思われる。一つめは、一人の人間の心の中に分裂を導入することである。騙す側と騙される側が同一人物であることに問題の根があるのだから、一人の人間の心を異なる部分に分割し、一方のシステムが他方のシステムに偽なる信念を植え付けると考えるのである。そうすれば、二つの矛盾する信念はそれぞれ別々のシステムに所属し、また、一方の意図が他方に気づかれる恐れもない。これは通常の欺瞞、他人を騙すケースにあくまで範を仰ぎ、似たような考え方を一人の人間の心の中にも徹底的に適用しようとする解決法である。

二つめの解決法はこれとは対照的である。こちらは、他人を騙す場合と

のアナロジーで自己欺瞞を考えることを緩和しようとする。このアナロジーをあまりに厳格に考え過ぎるところに問題の根があるのである。自己欺瞞においては、他人を騙す場合に成り立っている多くの事柄が成立していない。先走って言えば、この解決法においては、自己欺瞞が意図的な過程であることが否定され、また欺瞞者が最初に真実に気づいている必要もないとされるのである。

本論では、主に第二の解決法とその問題点について論じ、第一の解決法に関しては若干のコメントを述べるに止めておこう。後者が間違っていると思うからではない。それが提起する問題について論じることは、小論の範囲を大きく越え出してしまうからである。

自己欺瞞の問題に関して、心の分割仮説を明示的に唱えている論者は、私見の及ぶ限りでいえば、A. O. ローティアー、D. ペアーズ、D. デイヴィドソンの三人である。そしてこの順に、心の諸部分を実体的に考える傾向が強い。少なくともある時期までのローティアーは、統合された単一の自己という考え方に懐疑的であり、心を様々な問題解決のためのサブ・システムの集まりと考えていた⁽²⁾。ペアーズは、自己欺瞞に陥っている一人の人間、単一の人格を認めつつも、その心を二つの部分に分け、フロイトの意識と無意識の理論なども批判的に再検討しながら、かなり複雑な議論を展開している。心の諸部分も、自らの信念や欲求をもつ行為主体 (agent) であるとされる⁽³⁾。デイヴィドソンも、前節で述べた「信念のパラドクス」との関連で、心を分割する必要について示唆している。心を二つの部分に分け、矛盾する二つの信念をそれぞれに帰属させて衝突を防ぐ。デイヴィドソンにとって、これは合理的であるべき人間が、同一の主題に関していかにして矛盾した信念を抱けるのかという問題を解決するための概念的要請にすぎない。よって心の諸部分を実体的に考える必要はないこと、特にそこに、独立した人格性のごときものを付与する必要はない点が強調される⁽⁴⁾。

こうした提案はどれも興味深く、詳細な検討に値するが、私は本稿ではそれを行わないことにする。その理由は、まず第一に、こうした諸提案を検討するためには、次のような問題について考察する必要が生じる。すなわち、分割された心の諸部分をどの様な存在者として考えるのが適当なのか、それらは独立した意思と人格を備えているのかいないのか、そもそも人格性の基準とは何か。こうした問題について論じることは、小論の手に余る課題であり、他日を期したいと思う。また第二に、自己欺瞞を心の分割によって説明しようとする試みには、一般に次のような危険がつきまとう。心の諸部分を余りに実体的に考えすぎると、それは一つの実体が別の実体を騙すケース、つまり他者欺瞞そのものになってしまい、自己欺瞞でも何でもなくなるだろう⁽⁵⁾。自己欺瞞にいたる過程とは、自分にとって望ましくない事態に気づいた一人の人間が、何とか逆の事態を信じて心の安寧を得ようとする過程である。この個人の葛藤的状况という側面が、心の諸部分を余りに擬人化・実体化しすぎる説明からは失われてしまうのではないか。いずれにせよ、心の分割仮説は、自己欺瞞が提示する問題の解決策としては、最後の手段として検討されるべきものであり、それに訴えない解決法がもし他にあるのなら、そちらから先に探っていくべきだろう。しかしその前に、自己欺瞞の具体例を一つ示し、論点を整理しておこう。

3. 自己欺瞞の具体例と論点の整理

人がどのようにして自己欺瞞に陥るのか、今まであまり詳しく見てこなかった。自分を騙すに至る過程で、人はどのように振る舞い、また世界のどんな側面を認知しているのか。そこには本当に、心の分割を招来する余地はないのか。こうした点についてもっと具体的に見ていかななくてはならない。例えば次のような自己欺瞞の例について考えてみよう。

自分が最近太ってきているのではないかと、と考えている美男俳優がいるとしよう。体重は明らかに増えている。頬のあたりがたるんできているよ

うだし、ジーンズもきつくなった。さきほども、「ちょっと太った？」などと無神経に聞いてくる奴に会ったばかりだ。しかしこの俳優は、自分が太ってきているという考えに我慢ができない。俳優という職業に差し支えるし、周囲にちやほやされ、ずっと二枚目で通してきたプライドがそれを許さないのである。本当におれは太ってきているのか。いや即断は禁物だ。体重が増えているのは、ジム通いの成果が現れてきて、筋肉が増えているからに違いない。ジーンズがきつくなってきたのも、同じ理由で説明できる。頬のたるみにしても、鏡に写る顔の角度や照明をちょっと工夫すれば、以前と少しも変わらない。しかも、試しに病院で体脂肪率を一回だけ計ってみると、前よりも少し(0.0002%)落ちてきているではないか！こうしてこの俳優は次の信念を抱くにいたる―「おれは太っていない」。

この最後の段階、自分に偽なる信念を植えつけるに至ったときが、自己欺瞞が成立した段階である。自分を騙すのに、この俳優は成功したのである。そこにいたるまでは、自己欺瞞へといたる過程であり、自己欺瞞に陥り、偽なる信念を維持している状態とは一応区別される。この点を確認しておこう。またこの例で、私は曖昧性の問題、つまり人が太ったといえるのはどんな時かを問題にしたいわけではない。従って、この俳優が太ったことは端的に真理であるとしよう。例えば、利害関係のない他人が見れば、誰でも彼は太ったと判断すると考えるのである。

さて、この俳優が最後の信念にいたるまでには、様々な振る舞いや周囲の認知の仕方が介在しているが、その点については次節で詳しく見ていこう。ともあれ、先に提示した自己欺瞞の二つのパラドクスを、この例に即して整理しておこう。①まず信念のパラドクス。真実(「私は太った」)に気づいていながらどうして自分を騙し、偽なる信念(「私は太っていない」)を植え付けることができるのか。また欺瞞が成立した時点で、一人の人が互いに矛盾する二つの信念を抱いていることにならないか。②次に意図のパラドクス。普通に他人を騙す場合、騙そうという意図を相手に見

抜かれたら欺瞞は失敗する。しかし自分を騙す場合には、その意図は自分に筒抜けのはずである。われわれの俳優は、様々な振る舞いを通じて自分に偽なる信念を植えつける。こうした振る舞いは、自分に都合のよいことを信じて安心したいという意図のもとに行われているように見える。しかしこうした行為の際の意図は、本人の気づくところとなり、そこで自己欺瞞は挫折してしまうはずである。「私がこんなことをしているのは、自分を騙して安心したいからだ」と本人が知れば、自己欺瞞は不可能なものになる。にもかかわらず、どうして自己欺瞞と呼ばれる現象が存在するのだろうか。

この二つのパラドクスは、全く独立しているわけではなく、互いに関連している。二つを繋ぐものは、真なる信念である。この俳優がそもそも自分を騙そうという意図を抱くにいたったのは何ゆえだろうか。自分にとって望ましくない真実に気づいてしまったからであろう。この俳優は、多くの証拠が自分にとって望ましくない真実を示唆していることに気づいてしまった。体重計の数値、服のサイズ、見た目の変化、友人の証言（「最近太った？」）など、多くの事実が自分は太ってきていることを示唆している。一方で彼は、痩せていたいという欲求をもっている。実際この願いは切実であり、二枚目俳優という職業上からも、自分のプライドからいっても、容易にその挫折を肯定できるようなものではない。そこで彼は、自分を騙して心の安寧を得ようという作業に取りかかる。これがどのような作業であるかは、次節で詳しく見よう。とにかくこの時点で、自分を騙そうという意図が形成されているのである。すると事の発端には、真実への気づき、真なる信念があることになる。この信念を最初に抱かなければ、自分を騙そうという意図が導かれることもなかったはずである。

以上は正しい議論であると思う。しかしこのように考えると、パラドクスの解決策も見えてきたように思えるかもしれない。問題の根が摘出されたように思えるからである。つまり真なる信念、真実への気づきと

いったものを前提にしないで、自己欺瞞の成立を説明すればよいのである。そうすれば、自分を騙そうという意図が導かれることもなく、意図のパラドクスが発生する余地はない。また、矛盾する二つの信念の内の一方（真なる信念）が最初から無ければ、信念のパラドクスが発生する余地もない。このような考え方は上手くいくだろうか。次に、自己欺瞞は意図的な過程ではないと一貫して主張しているアルフレッド・メレの説を検討することにしよう。

4. メレの解決法とその問題点

メレは、次の点が自己欺瞞という現象の本質的な特徴、説明を要すべき重要な特徴だと考えている。すなわち、自己欺瞞に陥っている人は、多くの証拠が逆の事実を示唆しているにもかかわらず、自分にとって望ましい偽なる信念を抱くにいたっている。知的に欠陥があるわけでもない普通の人が、どうして多くの証拠によって棄却されるべき信念のほうをむしろ採用しているのか。この事実を整合的に説明することこそ、自己欺瞞がわれわれに突きつけている課題である。そしてそのためには、欺瞞者がもともと真実に気づいていたり、意図的に自分を騙すと考える必要は必ずしもない⁽⁶⁾。

この点は確かに重要であり、自己欺瞞が非合理的な (irrational) 現象であるといわれる根拠を言いあてている⁽⁷⁾。二つの互いに矛盾する命題のうち、多くの証拠によって支持されている命題と、証拠の支持が少ない命題の、どちらをわれわれは信じるにいたる傾向が強いだろうか。もちろん前者だろう。証拠が、どちらの信念にも同等に配分されているときには、われわれは判断を差し控え、当面の間、どちらの信念も抱かないように用心するだろう。ところが自己欺瞞においては、多くの証拠によって棄却されるべき命題の方が逆に信じられるに至っているのである。どうしてこうしたことが可能なのだろうか。

メレが強調するのは、われわれの欲求がいかに関世界の見方を歪め、客観的な判断を狂わせるか、である。自己欺瞞に陥っている人が信じるにいたるのは、自分にとって望ましい方の命題なのだから、欲求がわれわれの世界の認知の仕方にいかに影響を与えるのかを考察しなければならない。先の俳優の例でいえば、痩せていたいという欲求が、彼の信念の形成にどのような影響を与えるのかを考えなければならない。メレはこの点を、社会心理学における実験の成果なども参照しつつ、説得的に描いている。社会心理学が明らかにしてきたのは、欲求が介在していない場合でさえ、いかにして誤った信念が抱かれるに至るのかだった。偏ったデータから、ある社会事象について判断を下したり、データ自体が適性でも、先入観からそれらを誤って解釈したり⁽⁸⁾。自己欺瞞においても、似たような手続きが自分を対象に行われており、しかも欲求がそれを執拗なものにしているのではないだろうか。欲求の影響下、歪んだ証拠の収集や解釈に人は駆り立てられているのではないだろうか。実際メレに限らず、自己欺瞞に特徴的な振る舞いに関しては、多くの論者の間で意見が一致している⁽⁹⁾。自己欺瞞に典型的な振る舞いを、前節の俳優の例に即して見ながら、ここでは三点にまとめてみよう。

①まず自己欺瞞者は、証拠のインプットをコントロールしてしまう。つまり自分の欲求に適う、都合のいい証拠だけを集めるのである。俳優の例に即していえば、自分が痩せて見えるような角度と照明の下においてのみ鏡に写り、「私は太っていない」という信念を強化しようとする。また同じ例で、俳優が体脂肪率の検査を「一回だけ」受けていた点に注意しておこう。その結果は自分にとって望ましいものであった。体脂肪率はわずかながら落ちていたのである。しかし 0.0002% の下落などというのは誤差の範囲であり、もう一度検査すればどんな結果が出るか分かったものではない。しかし欺瞞者は、自分にとって望ましい証拠が得られれば、そこでもう調査を打ち切り、「私は太っていない」という結論に飛びついてしま

うのである。また欺瞞者は、他人を自分にとって都合のいい証拠の提供者に仕立て上げるという点も重要である。われわれの俳優は、相手が人のいい友人である場合に限って自分の体型に関する話題を持ち出し、「別に前と変わらないよ」という証言を引き出すよう仕向けるかもしれない。他人に相談することは、時にわれわれを誤りから救い出してくれる。しかし相談相手の選択にすでにバイアスがかかっているならば、他者は客観的な視点の提供者とはならないのである⁽¹⁰⁾。こうした様々な例が示しているのは、欺瞞者の「注意 (attention)」が、世界のどんな側面に向けられているのかという点である。自分が信じたいと望んでいることを支持してくれるような世界の側面には、必要以上に敏感になるくせに、逆の事態を支持する証拠はうっかり見落とし、また気づいても、ろくに考えもせず、特に考慮するにも値しないものと見なして意識の外に追いやってしまう⁽¹¹⁾。このようにして欺瞞者は、望ましくない真実からは目を逸らし、自分に都合のよい証拠だけを摂取し続けるのである。

②しかしこうした努力にも関わらず、自分にとって望ましくない真実を支持する証拠に気づいてしまった場合はどうするだろうか。証拠の身分はそのままにして、その意味するところを、解釈を変えればいいのである。自分にとって都合のいい、証拠の解釈。これが自己欺瞞者に典型的な振る舞いの第二のものである。俳優の例に即していえば、自分が太ってきたことを示す証拠、体重の増加や服のサイズの変化を、別の原因のせいにしていたのを想起しておけば足りる。実際人が、自分にとって耐え難い真実を示唆する証拠を都合よく解釈してしまう能力には、驚くべきものがある。彼らはそれを無視するのではなく、再解釈して説明し去ってしまうのである。例えば進化論を否定し、聖書が述べている通りの創造説を信じているキリスト教保守派のことを聞いてみればよい。彼らによれば、世界は約 5000 年前に神によって創造されたのである。そしてこの信念は、100 万年前の動物の化石が発見されても容易に揺らがない。彼らによれば、神

は、どんな意図のもとでかは知らないが、偽の化石を最初から組み込んだ世界を創造したのである！⁽¹²⁾ このように、自分の欲求に沿った仮説を維持するために証拠の再解釈は有効であり、時にそれは限度を知らないように見えるのである。

③自己欺瞞に典型的な振る舞いの最後のものは、メレが‘Acting as if’と呼んでいるものである。つまり、とにかく自分にとって望ましい事態が成立しているかのように行動するのである。多くの場合行動は、その人がどんな人間であるかを示す指標になる。人のもつ信念や性格を、われわれはその人の行動から推察する。この事実を逆手にとって、自分にとって望ましい事態が成立しているかのように振る舞えば、自分でもついにはそうした事態の成立を信じるにいたるかもしれない。これは周囲の他者をも巻き込んで、自分にとって都合のいい証拠を自ら作り出す戦術である。例えば、自分は冷たくない、親切な人間であると信じたがっている人はどのように振る舞うだろうか。なるべく他人を助け、親切な人間であるかのように振る舞うだろう。そうすれば、他人の方でも彼をそのような人間として扱い、直接褒めたりもするかもしれない。そうした他者のフィードバックは、自分が親切な人間であることを示す証拠になる。また他者を巻き込まなくても、われわれは自分がどんな人間であるのかについて、自分の行動を観察することによって推論することもあるのではないだろうか。俳優の例でいえば、すでに体形に合わなくなっているにもかかわらず、性懲りもなく前と同じサイズの服を買い求め、自分は以前と同じく痩せていると思い込もうとするような場合がそれである。

以上が自己欺瞞に典型的な三つの振る舞いである。いずれも欲求の影響下、世界の認知の仕方が歪められることを示している。こうした欺瞞的な三つの行動を総称して、ここでは「証拠の操作」と呼ぶことにしよう。自分にとって望ましい証拠だけに注意を向けたり、証拠を都合よく解釈したり、さらには自分に都合のよい証拠を自ら作り出す。メレによれば、こう

した証拠の操作が、自己欺瞞における本質的な問題に解答を与えてくれるものである。自己欺瞞において人は、真実を示唆する多くの証拠が存在するにもかかわらず、どうして逆の偽なる事態を信じるに至れるのか。見てきたような証拠の操作は、自分にとって望ましい偽なる事態を支持する証拠を水増しし、同時に真実を示唆する証拠を自分から遮断するものである。こうした操作を続けていけば、われわれはもともと証拠の支持の少なかった偽なる命題でもそれを信じるに至るだろうし、その後も証拠の操作を止めなければ、その信念を保持し続けていくだろう。

以上のメレの説明は、自己欺瞞に陥っている人が見せる奇妙な振る舞いを実に適格に捉えており、また同時に、もともと証拠の支持の少なかった偽なる信念が獲得される理由の説明にもなっている。この点に関して、われわれはメレに同意することができる。

しかしメレの議論には、もう一つ重要な論点がある。それは自己欺瞞が意図的なものであることを否定することである。つまり上で見てきたような証拠の操作は、メレによれば、自分を騙して安心させようという意図なくして起こるのである。それは多くの場合、本人の全く気づかないうちに起るのであり、そのためには、本人が何らかの欲求をもっているだけで十分である。例えば、われわれが強い欲求の影響下にあるとき、われわれの注意は、われ知らず、物事の肯定的な側面にのみ向かい、それを強調し、否定的な側面を意識の外に遠ざける。またたとえ、それが意図的なものだとしても、その場合の意図とは、単に世界の肯定的な側面を自ら楽しんだり、自分がもともと抱いている信念をテストして確認するという以上のものではないというのである⁽¹³⁾。

こうした考えにわれわれは疑問をもつものではあるが、しかし、メレがこのように考えるのには訳がある。いうまでもなく、それは「意図のパラドクス」を回避するためである。証拠の操作が、単に自分を騙して心の安寧を得るためになされているとすれば、それは自滅的であろう。こうした

意図のもとに行為をすれば、その意図は本人の気づくところとなり、自分が信じようとしている事態が偽であることが意識されてしまう。自己欺瞞はそこで挫折してしまうのである。しかし本当にそうだろうか。ある意図を抱いていながら、それに本人が気づいていないということもあり得るのではないだろうか。自分の意図は自分に対していわば不透明な場合があり、われわれは自分自身の意図に関して間違ふこともあるのではないだろうか。このように考え始めたとき、われわれはメレとは違う自己欺瞞の解決法を考察し始めていることになる。しかしこれについては、後に述べるだろう。

またメレは、欺瞞者がもともと真実に気づいた上で証拠の操作をしているという考えも退けている。欺瞞者は、多くの証拠が自分にとって望ましくない真実を含意していることには気づいていない。むしろメレにとって、証拠の操作とは、欺瞞者が真なる信念を抱くに至るのを「防ぐ」ために行われるものなのである⁽¹⁴⁾。メレがこうした考えを採る理由も、前節の議論から明らかなはずである。自分にとって望ましくない真実に最初に気づいてしまえば、そうではありたくないという欲求と相まって、自分を騙そうという意図が導かれてしまうからである。また、もともと真実に気づいていなければ、自己欺瞞が成立した時点で、欺瞞者に矛盾する二つの信念を帰属する必要もなくなる。「信念のパラドクス」も解決されるのである。

問題は、メレが記述しているような現象が、自己欺瞞と呼ばれるに値するかどうかである。パラドクスの解決に急なあまり、彼は別の現象の整合的な説明でもって自己欺瞞の説明に換えてしまっているのではないだろうか。真なる信念も、自分も騙そうという意図もなく、単に欲求の影響下に自分の注意を世界の肯定的な側面に向け、自分にとって都合のよいことを信じていればよいのなら、それは楽天的な思い込みというべきではないだろうか。また、意図なしに、欲求だけで自己欺瞞に特有の振る舞いを説明

するのには明らかに無理がある。つまり、欺瞞者が示す執拗なまでの証拠の操作である。俳優の例で言えば、自分が痩せて見えるような角度から鏡に写り、体重が増えてきたのを筋肉のせいにしたたりして、自分を納得させようとする。こうした欺瞞的な行動に、彼を駆り立てているものはなんだろうか。メレによれば、それは痩せていたいという欲求のなさしめるところであった。しかし、欲求だけでは、欺瞞者のこうした奇怪な行動を説明するのにあまりに不十分である。痩せていたいという欲求は、この俳優が、自己欺瞞に陥る遥か以前、実際に痩せていたときから、もっていたものである。しかしそのころ彼は、こうした奇怪な行動には走らなかったのである。明らかに、自己欺瞞を特徴づける執拗なまでの証拠の操作には、欲求以外の要因があるはずである。そしてそれは、真なる信念以外ではあり得ない。つまり証拠の多くが、自分が太ってしまったことを示しているのに気づいたからこそ、自分を騙して安心させようという執拗な行動に彼を駆り立てているのである。多くの論者が、自己欺瞞には真実への気づきが最初にあると考えているのも、それがなければ、執拗なまでの証拠の操作の動機を十分に説明できないからなのである⁽¹⁵⁾。自分にとって望ましくない真実に気づいた人が、何とか逆の事態を信じて心の安寧を得ようと努力する。自己欺瞞におけるこの葛藤的側面が、メレの議論からは抜け落ちてしまっているのである。もちろんメレが記述するような事態も存在する。つまり欲求の影響下、意図せずして世界の肯定的な側面にばかり注意が向かい、そのため自分にとって都合のよいことを信じ続けるということはあるだろう。しかしこれは、楽天的な思い込みというべき事態であり、自己欺瞞とは別の現象であるように私には思われる。

しかしこうした批判は、メレによって肩透かしを食わされてしまいそうである。というのも、メレ自身が、自己欺瞞と単なる願望的思考 (wishful thinking) をあまり区別しないと述べているからである。例えば B. サバドスによれば、二つは次のように区別される。願望的思考とは、ほんの

わずかな肯定的な証拠だけで結論にジャンプし、自分にとって望ましい偽なる事態を楽天的に信じている状態であり、一方自己欺瞞とは、多くの証拠が逆の事態を示唆しているのに気づいていながら、それにあえて逆らって自分に望ましい事態を信じている状態である⁽¹⁶⁾。しかしメレによれば、これは自己欺瞞という現象を余りにも狭く捉えすぎることであるという。また欲求の影響下に行われる証拠の操作という点から見れば、両者の境界は曖昧で、そう截然と区別できるものではない。むしろ日常的で、それゆえ中心的に論じられるべきなのは、願望的思考のようなケースだということである⁽¹⁷⁾。

こう見て来ると、メレとわれわれの対立点は、どんな現象をもって自己欺瞞と呼ぶかというトリヴィアルなものであることになる。しかし私には、メレは自己欺瞞のパラドクスを回避しようとするあまり、論じる対象をひそかに変えてしまったように思われてならない。自己欺瞞をその名に相応しい本来の意味で捉えるなら、パラドクスは今だに解決されていないのである。すなわち、自分にとって望ましくない真実に気づいてしまった一人のひとが、意図的に自分を騙すにいたる過程としては⁽¹⁸⁾。

5. 信念のパラドクスの解決

しかし以上の考察の中には、パラドクス解決のための手掛かりとなる考え方がすでに含まれている。最も重要なのは、自己欺瞞に陥っている人に見られる特徴的な振る舞い、つまり証拠の操作という考え方である。これが意図的に成されるものであっても、「意図のパラドクス」には陥らないことを私は次節で示すつもりである。従ってこの節では、証拠の操作が意図的に成されることを仮定した上で、この考えがどのように「信念のパラドクス」を解決するかを見よう。信念のパラドクスには二つの側面があった。第一に、真実に気づいていながらどうして自分を騙すことができるのか。また第二に、欺瞞に成功した時点で、真なる信念とそれを否定する偽

なる信念、この二つの矛盾する信念を、同時に一人の人間に帰属しなければならぬのではないだろうか。

この第一の側面に対しては、証拠の操作という考え方が直接的な答えを与えてくれよう。その内実は、すでに前節で詳しく見ている。証拠の操作とは、真実を示唆する証拠から注意を逸らし、あるいはそれを自分に都合のいいように解釈し、さらには偽なる事態を支持する証拠を自ら作り出してしまおうというものだった。このようにして、偽なる事態を支持する証拠を水増しし、逆に真なる事態を支持する証拠を意識の外に遮断する。こうした操作を繰り返していれば、本人の意識の上では、偽なる事態を支持する証拠のほうがその数と頻度において優勢に見えてくるはずである。そうすれば、たとえわずかな間だけだとしても、自分にとって都合のよい、偽なる信念を抱くに至るだろう。

さて問題は、信念のパラドクスの第二の側面である。もともと真実に気づいていたところに、自分を騙して偽なる信念を植えたのだから、その時人は、真なる信念と偽なる信念、矛盾する二つの信念を抱くに至っているのではないか？ 必ずしもそのように考える必要はない。この問題に関しても、証拠の操作という考えが役立ってくれる。つまり証拠を操作して偽なる信念を自分に植え付けようとしている過程を、同時に、真なる信念を自分から放逐する過程として捉えるのである。この点に関して重要なのは、ソーレンセンが述べているように、自己欺瞞とは瞬時にして成立するものではなく、漸進的な過程であるということだ⁽¹⁹⁾。真実に気づいて欺瞞を開始した時点と、欺瞞が成立して偽なる信念を抱くに至る時点との間には、時間の幅がある。そこで行われているのが証拠の操作であり、真実を示唆する証拠を意識の外に遠ざけるために、様々な方策が採られているのはこれまで見てきた通りだ。そのため自己欺瞞が成立し、偽なる信念を自分に植え付けるに至るまでには、真なる信念の方は棄却されていると考えることができるのである。同じ時点で、一人の人に、相矛盾する信念

を帰属する必要はないのである。

しかし注意しておくべきことがある。自己欺瞞は、それに成功したとしても、極めて不安定な状態だということである。つまり植えつけられた偽なる信念を、そのまま維持するのは容易ではない。それは当人の欲求に沿った信念、自分にとって都合のよい、望ましい事態への信奉である。だとすれば、それを得るや否や、当人はもう安心してしまって証拠の操作を一時的にせよ止めてしまうかもしれない。証拠の操作を止めてしまえば、当然予想されるのは、真なる事態を支持する証拠が再び当人の視界に目立ち始めるという状況である。日常的な場面では、真実を支持する証拠は逆の事態を支持する証拠よりもその数と頻度において優勢であり、それを視界から一掃することは難しいはずである。それは他人の証言や、見たくもない世界の側面として当人の注意を引きつける。真実に再び気づき始めた当人は、再び自己欺瞞のプロジェクト、証拠の操作に乗り出さざるをえない。おそらくこうした堂々めぐりの状況が、自己欺瞞者は矛盾する二つの信念を抱いているという考えを助長するのである。しかしこうした時間幅のあるプロセスを説明するのに必要になるのは、真なる信念、偽なる信念、そしてどちらの信念を抱くべきか決めかねている状態（どちらの信念も抱いていない状態）の三つであり、矛盾する二つの信念を同時に当人に帰属する必要はないのである。信念のパラドクスは解決されたのである。

6. 意図のパラドクスの解決

さて、残る問題は意図のパラドクスである。普通に他人を騙す場合、騙そうという意図を相手に見抜かれれば欺瞞は失敗する。しかし自己欺瞞の場合には、自分の意図は自分に筒抜けのはずではないか。だとすれば、自己欺瞞とは不可能な過程にならないだろうか。一方で、自己欺瞞は意図的なものではないとするメレの見解が陥る困難を、われわれはすでに4節で見ている。自己欺瞞を意図的な過程として整合的に説明する道はあるだ

ろうか。

この問題も、具体的には、証拠の操作に関わることである。自分に偽なる信念を植えつける過程で人は様々な証拠の操作を行うが、そうした振る舞いを、本人自身がどのように理解しているのかという問題である。われわれはそれを意図的な過程として捉えたい。つまり証拠の操作は、自分にとって望ましくない真実に気づいた人が、自分の欲するところを信じて安心しようと意図して行う様々な行為として捉えたい。しかし次の点が問題である。もしも当人が、「私がこんなことをしているのは、自分を騙して安心させたいからだ」などと気づいてしまったら、その場で自己欺瞞は挫折するだろう。従って、意図のパラドクスを解決するためには、自分の行為の意図に本人が気づかないことがありえるという点を論証しなければならない。行為の際、その本来の意図を当人が間違えたり、それに気づかないなどということがありえるのだろうか。

これはまさに自己知の問題、自分自身の心の状態に関して、当人自身に不透明な部分が存在するかどうかという問題にほかならない。われわれは自分自身の欲求や意図や信念に関しては、特権的な接近の仕方をもって見るように見える。他人における場合には、相手の振る舞いや発言からそれらを類推する必要があるのに対して、自分の場合には通常この手続きは必要とされない。よって人が何を意図して行為したかに関しては、多くの場合、本人自身の証言に一番の重みをおいて判断が下されるべきである。以下の私の議論は、こうした考えに疑いを挟むものではない。通常の場合、われわれは自分がどんな意図の下で行為しているのかをよく知っているのである。しかしこの事実、人が自分自身の意図に関して下す判断が常に正しいということを意味してはいない。自分自身の意図に関して、人は誤りえない (infallible) ということを意味してはいない。

人はどんな時に自分の意図に関して間違えるのだろうか。例えば、一つの行為で二つ以上の目的が達せられる場合について考えてみよう。些細なこ

とで、子供をひどく叱る親がいたとする。どうしてそんなに叱るんだと問われれば、「躾のためだ、立派な人間を育てるつもり（＝意図）でだ」と答えるだろう。そしてそれはまんざら嘘ではない。しかしこの親は全く関係のない理由、例えば偏頭痛か何かで苛々しており、その鬱憤を晴らすために子供を叱りつけたというのが実情だとしてみよう。そうすると、この親は、少なくとも自分の行為の主要な意図に関しては間違っていることになる。自分の行為を引き起こすに際して最も力のあった意図に言及することなく、二次的な意図のみでそれを説明しているからである。さらにこの親が、自分の説明に自分で納得してしまい、それを信じてしまえば、本来の意図は意識の外に追いやられ、二度と顧みられることはない。ここで重要なのは、二次的な意図の方が、自分や他人に受け入れられやすい、社会的規範に沿ったものだということだ。われわれは何らかの意図を抱く。しかしそれが自分の良心や社会的規範に背くものだった場合、自分の行為にとって本質的ではない、二次的な意図を持ち出してそれを説明してしまうのである。さらに甚だしい場合には、もともと抱いていなかった意図を捏造して説明してしまうかもしれない。人の行為の説明として、社会的に広く受け入れられ、ステレオタイプ化されたようなものがあれば、自分の行為の説明に際しても、人に納得されやすい意図を捏造し、さらに自分で信じ込むのも容易になるはずである。

こうした論点は、意図の問題に限らず、より広い射程をもつものである。社会心理学の分野では、人が自分自身の様々な心の状態について誤りうるということが実験的に明らかにされている。自分の感情や行為の原因がどこにあるのか、あるいは、ある対象や人物を自分が好きか嫌いかといった判断に関して、われわれは間違いうるのである。こうした際にわれわれが判断の拠り所とするのは、注意深い内省によってではなく、広い意味での「理論」、人の感じ方や振る舞い方に関する社会通念のごときものであることが多いからである。男（女）はどのように感じ、振る舞うべきか、親は

どのように感じ、振る舞うべきか、などといった社会通念である。また社会通念とは別に、われわれは誰でも、自分がどんな人間であるかに関する「理論」をもっている。自分の性質や才能に関する思い込みや、こんな人間でありたいというセルフ・イメージは、それが正しい場合でさえ、特定の判断に関しては人を誤りに導くことがあるのである⁽²⁰⁾。

さて、もし以上のような議論が正しく、人が自分の意図に関して間違えることがあるのなら、それは自己欺瞞における「意図のパラドクス」の解決に道を開くものとなる。欺瞞者が、証拠を操作して自分に偽なる信念を植えつけようとしている時、それは自分を騙して心の安寧を得ようとしてそうしているのである。美男俳優の例を思い出そう。彼は、自分は太っていないと信じるために、自分が痩せて見える角度や照明の下においてのみ鏡に写るなど、様々な証拠の操作を行っていた。彼は自分の欲するところを信じて安心してそうしているのである。こうした意図に、本人自身が気づいてしまえば自己欺瞞はそこで失敗する。しかし上で見たように、われわれは自分が抱いている意図に常に気づいているとは限らない。抱くに至った意図が、自分の良心や社会的規範に背くようなものである場合、あるいは自分のセルフ・イメージを傷つけるようなものである場合、われわれは自分の行為を、自分にとって受け入れやすい別の意図によって説明し、本来の意図を意識の外に追いやることがあるからである。例えば件の俳優は、自分が行っている行為についてこう説明できる、「自分が痩せているという真実（これは実際には偽であるわけだが）を確かめるために、証拠の確認（操作ではなく）をしているのだ」と。このようにして、欺瞞成立の妨げとなるような意図に気づくことなくわれわれは証拠の操作に勤しみ、ついには、自分に偽なる信念を植えつけることに成功するのである。われわれは自分自身の心の状態について、特に意図に関して無謬 (infallible) ではないという事実が、自己欺瞞を整合的なものとして記述することを可能にするのである⁽²¹⁾。

注

- (1) パラドクスのこの二つの定式化に関しては A. R. Mele, Two Paradoxes of Self-Deception in J. P. Dupuy, ed., *Self-Deception and Paradoxes of Rationality*, Sanford, California: CSLI Publications, 1998 を参照.
- (2) A. O. Rorty, Belief and Self-Deception, *Inquiry* 15, 1972. ただし Rorty, The Deceptive Self: Liars, Layers, and Lairs, in B. P. McLaughlin and Rorty, eds., *Perspectives on Self-Deception*, Berkeley: University of California Press, 1988 では, 統合された人格の必要性が説かれるなど, ローティ어의考えは変化を見せている.
- (3) D. Pears, *Motivated Irrationality*, Oxford: Oxford University Press, 1984, ch. 5.
- (4) D. Davidson, Deception and Division, in J. Elster ed., *The Multiple Self*, Cambridge: Cambridge University Press, 1988, pp. 91-92.
- (5) この論点に関しては, S. L. White, Self-Deception and Responsibility for the Self, in McLaughlin and Rorty, 1988, pp. 453-454; R. Sorensen, Self-Deception and Scattered Events, *Mind* 94, 1985, pp. 65-66. などを参照.
- (6) A. R. Mele, *Irrationality: An Essay on Akrasia, Self-Deception, and Self-Control*, New York: Oxford University Press, 1987, ch. 9, 10.
- (7) Mele, *ibid.* p. 136; Davidson, *op. cit.* pp. 81-82.
- (8) Cf. R. Nisbett and L. Ross, *Human Inference: Strategies and Shortcomings of Social Judgment*, Englewood Cliffs, New Jersey: Prentice-Hall, 1980, Pt. 1, 2; T. ギロビッチ『人間この信じやすきもの—迷信・誤信はどうして生まれるか』守一雄・守秀子訳, 新曜社, 1993年 (T. Gilovich, *How We Know What Isn't So*, New York: Macmillan, 1991) 第1部.
- (9) 例えば Mele, *Irrationality*, pp. 125-126, p. 143ff; K. Bach, An Analysis of Self-Deception, *Philosophy and Phenomenological Research* 41, 1981, pp. 357-362; Pears, *op. cit.* pp. 59-63. などを参照. 私のここでの三区分は, メレ自身のものというよりも, 多くの論者に共通する論点を抽出したものである.
- (10) ギロビッチ, 前掲書, 132-133頁.
- (11) Mele, *op. cit.* p. 126.
- (12) Cf. S. J. Gould, Adam's Navel, in his *The Flamingo's Smile*, Harmondsworth: Penguin Books, 1985.

- (13) Mele, op. cit. pp. 148-149.
- (14) ibid. pp. 128-129.
- (15) 例えば, Davidson, op. cit. pp. 88-90; F. F. Schmitt, Epistemic Dimensions of Self-Deception, in McLaughlin and Rorty, 1988, p. 190, p. 195; A. W. Wood, Self-Deception and Bad Faith, in McLaughlin and Rorty, 1988, pp. 222-223などを参照.
- (16) B. Szabados, Whishful Thinking and Self-Deception, *Analysis* 33, 1973, pp. 204-205.
- (17) Mele, op.cit.pp.134-136.
- (18) メレに近い立場から自己欺瞞を論じたものとしては, ほかに A. Lazar, Division and Deception: Davidson on Being Self-Deceived, in Dupuy, 1998; B. P. McLaughlin, Exploring the Possibility of Self-Deception in Belief, in McLaughlin and Rorty, 1988などを参照.
- (19) Sorensen, op. cit. p. 69.
- (20) こうした論点については, Nisbett and Ross, op. cit. ch. 9を参照. 自分の行動の原因に関する判断の誤りについては R. E. Nisbett and T. D. Wilson, Telling More Than We Can Know: Verbal Reports on Mental Processes, *Psychological Review* 84, 1977., 自分の選好や感情に関する間違いについては T. D. Wilson, Strangers to Ourselves: The Origins and Accuracy of Beliefs about One's Own Mental States, in J. H. Harvey and G. Weary, eds. *Attribution: Basic Issues and Applications*, Orland, Fla.: Academic Press, 1985; A. R. Hochschild, Emotion Work, Feeling Rules, and Social Structure, *American Journal of Sociology* 85, 1979.なども参照せよ.
- (21) 本稿は, 2002年10月26日に行われた三田哲学会研究例会(MIPS)での発表原稿に, 加筆・修正してここに掲載するものである. 発表の際には多くの質問が寄せられた. その内の幾つかについては, 発表の時には省いた部分(第4節の後半部分)がお答えしているはずであるが, 今回論じることができなかった問題に関しては, 今後の研究の参考にさせていただき, これからの仕事, 論考の中でお答えしていくと申し上げるほかはない. 質問者の方々に深く感謝する.