

自己欺瞞について

鈴木 美佐子

はじめに

「自己欺瞞」という問題は、信念、欲求、意図、推論や注意、動機、自己知識、自己の統一性、心理的防衛など多くの概念に関わる興味深い問題であるが、大きく分けて二つの問いに関わっていると言える。つまり、「それは何か」という問いと、「どのようにして可能なのか」という問いである。両者は密接に関係していて、分析的な側面と心理学的な側面をもつと思われるが、本論では主に分析的な面から自己欺瞞を考えていく。すなわち、自己欺瞞とは「自己」と「欺瞞」という言葉からできていることが示すように、文字どおり「自分に嘘をつく、自分をだます」ことなのか、そうでなければ、それは他の欺瞞とどう違い、どうやってその名に値するのかということ⁽¹⁾を考察する。自己欺瞞はまた、信念にまつわるパラドクスの外観を呈する。他人をだますケースをもとにして「自己欺瞞」を規定しようとするおぼつかってしまうパラドクスを回避しつつ、様々な哲学者がこの現象を説明しようとしている。主要なものとして二つの方法がある。第一に一方の信念を排除する立場（これはどちらの信念を残すかで二通りの考え方に分かれる）、第二に矛盾し合う信念の共存は可能だとする立場である。

本論では、これら二つの見解を中心に取り上げ、それぞれを批判的に検討していくことで、自己欺瞞を分析し

ていくと同時にそのパズルを解消する道を探っていきたい。

1 「他人をだます」をもとにした自己欺瞞の規定

まず、「自己欺瞞 self-deception」という言葉は欺瞞の特殊ケースを表現し、それは他人をだますように自分を自分をだますケースを述べているのだと考えることにしよう。「AがBをpについてだます」と言うとき、Bはだまされてpと信じることになるのだが、だます人Aはpが成立しているとは信じていないと言える。

へ1 AがpについてBをだましているなら、Aはnot pと信じている。

へ2 AがpについてBをだましているなら、Bはpと信じている。

自己欺瞞は自分で自分をだますケースなので、これを使うと、

「1」 AがpについてAをだましているなら、Aはnot pと信じている。

「2」 AがpについてAをだましているなら、Aはpと信じている。

ということになる。この時Aはpを信じていながら、not pとも信じていることになる。これは一見して不合理であり、一種のパラドクスになっているように見える。⁽²⁾

たとえば、「雨が降っていると信じているが、降っていないとも信じている」と言う人に対して、私たちはその人が論理的に誤っているか、「雨が降っている」、「信じている」という言葉の意味がわかっているかと思うだろう。言葉の意味を理解していて、論理的能力に欠けているのでもなければ、矛盾し合うことを同時に信じることはあり得ないと考える。しかし、自己欺瞞を「だます」の特殊ケースと考えると、自己欺瞞に陥っている人は、雨についておかしいことを言う人と同じ状態にあるように思える。にもかかわらず、彼の論理能力は正常である。この状態はどのように説明されるのだろうか。

2 「一つの信念による「自己欺瞞」の説明

矛盾し合う二つの信念が同時に存在することは、通常不可能な心の状態であると考えられる。自己欺瞞と呼ばれる現象を分析してみると、一つだけしか信念は現れないと考えるのがこの立場である。その場合、「だまされる」方の信念だけで十分だと考えるか、「だます」方がもつとされる信念による説明ができるとするかで、さらにその立場は二つに分かれることになる。

(1) 「だまされる」者の信念による説明

たとえば、AとBの夫婦がいて、夫Aは妻Bをずっと信じていたが、最近事情が変わってきた。Bは週2日くらい仕事から遅く帰ってくるようになり、夕食後や週末に家を空ける口実を見つけ出すようになった。さらにAは友人から「Bが同僚のC氏と一緒にいるところを見た」と聞かされる。しかしAはBの仕事が忙しいのだと考え、同僚と一緒にいたのもそのためだと思い、人目のあるところで会っているのはかえってやましいことがないからだ判断する。つまり「Bは貞淑だ」(p)と信じている。この時私たちは「Aは自分を欺いている」と言うだろう。こういったケースを(1)の立場は次のように説明する。

Aは「Bは貞淑でない」(not p)と信じたからpを信じるようになるのではない。Aは、何らかの動機で(この場合は「貞淑でない」(not p)と信じたくないということかもしれないし、貞淑であってほしいと思うからかもしれない)pに関係する情報を都合のいいように解釈する。その結果pだという信念をもつというのである。

自己欺瞞においては、not pと信じていることが原因で、pと信じていることになるのだという説明が一般的だが、自己欺瞞者はnot pと信じているわけではないというのがこの説明のポイントである。すなわち、Aがpについて自己欺瞞をおかしていると言えるのは、

〔2.1〕 Aはpと信じている。

〔2.2〕 Aはpの真理値に関連するデータを何らかの動機からバイアスをかけるやり方で扱う。

〔2.3〕 AがバイアスをかけることがAがpという信念を得ることの原因である。

〔2.4〕 Aがpと信じることは誤りである。

〔2.5〕 Aがそのときもっているデータはpよりnot pに大きな根拠を与える。

という五つの条件を満たしているときである。⁽³⁾

これには問題点がいくつかある。まず、当事者がnot pだと考えてもおらず信じてもないのなら、pを支持しない情報から注意をそらしたり、証拠を都合のいいように考えてpと信じたとのだしても、当人は何ら気づかず、信じたからそう信じたただけだということになりはしないだろうか。これでは自己欺瞞が、樂觀主義あるいは希望的観測と呼ばれる状態と変わらないものになってしまうだろう。

それでも、普通なら特定の情報から目をそらしたり都合よく解釈したりせずに、pではないと信じるはずのところを、ごまかしたり次々にエクスキューズしていくことでpと信じたということは、前もってpでないと信じていなくても、やはり自分をごまかして信念を得ているのだから自己欺瞞だと言い得るかもしれない。自己欺瞞のそういった特徴を捉え、それを保証しているのが〔2.5〕である。手持ちの情報を並べてみれば、not pと判断されるにもかかわらず、Aはpと信じるので自己欺瞞だというのである。

ここから二つの問題が生じる。一つは、「pよりnot pの方が証拠によって強く支持される」と判定しているのは誰かということである。Aが判断しているのなら、当人がnot pと判断せず、not pと信じないということは考えられない。そうすると、この説明の中にはもう一つの信念not pが現れてしまい、パラドクスが再登場する。それを避けるためこの立場では、証拠の強さを判定するのは「偏らない観察者 impartial-observer」⁽⁴⁾だと考える。

つまり判定者は公平な判断のできる他人だといふのである。

しかし、上記の条件を満たしている人は、おそらく他人がそれは違うよ、と言っても、なお本人は正しいといふのではない。証拠を都合よく解釈して、解釈後のデータをもとに彼はそう信じたのである。公平に客観的に見れば誤っていることを信じているだけで、はたしてその人は自己欺瞞者と言えるだろうか。その判定ミスに気づいていないのだとしたら、その人はただの頑固者、視野の狭い人、と言われるだけではないのか。⁽⁵⁾

さらに、「2.4」「pと信じていることが誤っている」のではない場合にも、自己欺瞞は生じ得る。普通なら獲得されないpという信念をそう信じたがために信じたけれども、実はpが正しく、手持ちの証拠だけからはとてもpと一般にも信じられない場合でも、自己欺瞞と呼べるケースがあるのではないだろうか。⁽⁶⁾

(2)「だます」側の信念による説明

次に、pと信じている必要はないのだ、という説明を簡単に見ていきたい。⁽⁷⁾たとえば、末期ガンの患者が心の中では「自分は治らない」(not p)と信じているのに、「私は治る」と真剣に言ったり、退院後の計画を話したりする。しかし、彼は「治る」と信じているわけではないというようなケースである。この説明では、自己欺瞞は、

[3.1] Aは not p と無意識に信じている。

[3.2] Aは p に対する賛成的態度 (pro-attitude) をもっている。

[3.3] Aは p と認める。

時であるとする。自己欺瞞者が pだと信じていないならAの状態はとくに説明が困難ではない。しかし、この説明は成功していないように思える。

第一に、「自分は治る」と心から認める人が「治る」とは信じていないとどうして言えるのか。私たちは人がある事柄を真剣に認めているのを見れば、そう信じていると判断する。そしてその人が認めたことに一致するように

ふるまえば、さらにその感を深くするだろう。自己欺瞞に陥っている人が、真剣に認め、信じていると言い、それに基づいて行動しているように見えても、本当は信じていないのだ、というのがこの説明だが、そうしたAの状態を「pを信じている」と言わないのだとしたら、さらに何が加われば「信じる」ことになるのか。その基準あるいは条件が示されなければならない。

さらに、Aは「治ると信じているか」と聞かれれば「信じている」と答え、回復後の話をしたり、見舞いの人に「心配ない」と言ったりする。それでもAが信じていないケースは確かにあり得る。ふりをしているだけ、ということもある。しかし、その場合は、他人をだましているだけということになる。どれほど熱心に「治る」と自分に信じ込ませようとしていたとしても、信じるふりをしたり、信じていると嘘をついているのなら、その欺瞞は成功してはいないと言える。

こうした反論に答えこの立場を救うために、「pと信じている」を条件に入れるけれども、二つの信念は意識的なものと無意識的なものなので両方が共存すると主張することもできるが、それは一方の信念を排除するものではなく別の説明となるので、次節で検討する。

3 心の分割

自己欺瞞に現れるとされる二つの信念のいずれかを否定するのではなく、両方をもっていることが可能だと言うための工夫をすることで、パラドクスの解決をはかろうとする方法もある⁽⁸⁾。ここではデイヴィドソンのものを取り上げて検討してみよう。

デイヴィドソンの自己欺瞞の規定はこれまでのものと少し異なるが、どうしてそれがパラドクスなのかをきちんと説明しているように思えるので、少し詳しく見ていこう。

Aがpについて自己欺瞞をおかすのは次のような過程を通じてである。

- [4.1] Aは手に入れている証拠が全体として not p を支持していると判断する。
- [4.2] Aは not p と信じる。
- [4.3] not p と信じることで、Aは不安や苦痛を感じる。
- [4.4] Aは p と信じる動機や理由をもっている。
- [4.5] Aは p と信じることを促すと思われること（前述のバイアスかけなど）を行う。
- [4.6] Aは p と信じる。

このようにしてAは自己欺瞞に陥るのだが、自己欺瞞が不合理なのは、二つの信念をもっているからだけではない。普通、pとnot pという対立し合う事柄のそれぞれを信じる理由をもっている人は、二つをつき比べ、それぞれを支持するそれぞれの証拠、さらに関連する証拠を考え合わせ、どちらか一方を採用する。それをデイヴィッドソンは「全体証拠の要請」と呼ぶ。⁽⁹⁾「全体証拠の要請」とは、「関係する入手可能なすべての証拠が支持する仮説を信じよ」というものである。自己欺瞞に陥る人は、ある命題に賛成する証拠も、反対する証拠ももっているが、証拠全体の支持する方の命題だけを受け入れることをしないという不合理をおかしているわけである。こうした不合理を彼は「意志の弱さ」と類比的に「根拠（保証）の弱さ（Weakness of Warrant）」と呼ぶ。⁽¹⁰⁾

証拠をもとに何かを信じるとき人はこの要請に従うはずである。従えば、ともかく片方の信念だけを採用し、片方は捨て去る。ところが自己欺瞞に陥っている人はそれぞれの命題を信じたままである。要請を認めればそれは不可能である。しかるに自己欺瞞者はそうしている。だからパラドクスだというわけである。

デイヴィッドソンが与えたパラドクス解決の要点は、人はそうした対立する信念を別々にしておくことができるし、自己欺瞞のような不合理に思える現象の場合はとくにそうしているのだということである。心の部分がある程度独

立していて、矛盾し合う信念の一方が意識に対して閉じていれば、矛盾を含みながら一方が原因となって他方を引き起こすことも可能であり、それはまるで別の二つの心の間で起こることのようなものであるとデイヴィドソンは言う。⁽¹¹⁾ 心全体から見れば矛盾する信念を含んでいるのだから、このような状態は不合理だが、両者の間に境界線があり、その境界を消さずに当事者が全体を見渡すことができなければ、二つの信念が総合され片方が打ち消されることもないと言うのである。

確かにこうした考え方は、自己欺瞞者が矛盾する命題を同時に信じていることができるのはどのようにしてかということの説明しているように見えるし、その不合理さが普通なら従っている「要請」に違反することから生じているというのをもつて得ているように思える。しかしこれにもいくつかの難点がある。

この説明によれば、だます部分はだまされる部分に p という信念を生み出すための様々な操作を行うことになる。それはだまされる部分が気づかないようになさなければならない。他人をだます場合と同じように、だまされる方がだます方の意図や戦略を知っていればそれは成功しない。狡知に長けただます部分はそうやってだまされる部分に信念 p をもたせる。これにはいろいろな疑問が投げかけられている。⁽¹²⁾ だます部分はその全体より不安や苦痛に耐性をもっているのか。だます部分がまずコントロールし、次にコントロールされた何も知らないだまされる部分の信念に、その後の推論や行為をコントロールさせるのはなぜなのか。また、最初に「全体証拠の要請」に従い信念 $not p$ をもった部分が欺瞞を始めるのだが、この要請はどうやって働かなくなり、無視されるのか。デイヴィドソンはこれらの難問に答えておらず、信念に関する問題を心理学的な問題に置き換えているだけのようにも見える。

だます部分が無意識、だまされる部分を意識と呼ぶことも可能であろう。そうするとフロイトの立場に対して提出されてきたのと同様の批判がデイヴィドソンにも与えられる。その一つは、全人格、行為者ということが意味をなさなくなるということである。心がともかく分割されると考えると、思考や欲求、行為さえ、全人格より小さい

部分に帰属させなければならぬことになってしまう。態度や行為が行為者のものではないなどといったことが考えられるだろうか。行為や意図がある一人の人間と結びつくのではなく、心の部分に結びつけられるのだとしたら、責任を問うとか、非難するとかといったことが難しくなる。私たちは自己欺瞞に陥っている人に多かれ少なかれ非難めいた気持ちをもっている。心の一部が誤りをおかしているだけで、その誤りについてその人自体を問題にできないのなら、糾弾する私たちの態度の方が理不尽なことになってしまうだろう。

また、私たちは人間を常識や規則に従った、一貫性のあるものとして捉えている。だからこそ他人の行為や信念、欲求というような心的なものについて語ることができ、他人に心的なものを帰属させることができるのである。不合理な現象や心の働きにだけ心の分割が限られるとしても、心の中で矛盾する出来事がそれぞれ勝手に生じるというものであれば、人間にもとより一貫性を期待することはできず、ふるまいや言葉から心の内容を予測するということもできなくなるのではないか。

さらに、そのような二重の構造は自己欺瞞からの回復をどのようにに説明するのだろうか。自己欺瞞の状態にあった私はあることを知って今まで信じていたことが間違っていたと気づく。それは not p であることを強く示す情報であるかもしれない。なぜ突然壁が取り払われるのか。その情報は私により一層の苦痛をもたらす。私が避けたかった信念を浮かび上がらせる。だます部分の心はなぜ一層手の込んだ戦略で私をだまさないのか。デイヴィドソンは、その説明を棚上げにしている。自己欺瞞が起こっている時には心に二重の構造が生まれているのだと言っているだけだと思える。ひとつの説明図式によって説明されることが多ければ、それは優れた説明だと言えるが、デイヴィドソンの見解は難問を増やしている。⁽¹³⁾

4 おわりに

ここまで検討してきた説明方法以外に、注目すべき二つの見解がある。一つは、二つの信念が自己欺瞞に現れることは認めるが、時間的な経過ということを考慮に入れば、二つの信念が存在してもパラドクスにならないとする説明であり、もう一つは、自己欺瞞は他人をだますことの特例ケースではなく、別の形で言い換えられるべき現象だとする見解である。⁽¹⁴⁾これに関する詳しい検討は別の機会に譲りたいが、そうした見方が示唆することを少し述べて終わりたい。⁽¹⁵⁾

第一の見解は、自己欺瞞には始まる時点から達成される時点までの時間的経過があり、総称すれば自己欺瞞と呼ばれるが、複数の出来事が次々に起こっていると考える。過程全体としては二つの信念が両立していると言わなければならないが、最初にもたれた信念 *not p* は、自己欺瞞が達成されたときには消え去っている。で説明は困難ではないと言うのである。この見方は、二つの信念の間の揺れ動きが繰り返し起きるためにあたかも二つの信念をもっているように見えると述べ、自己欺瞞に見られる揺れや信念の不安定さといった特徴をよく捉えている。

第二の見解は、言葉の表面上の形式にとらわれてはならないことを示唆する。再帰的な表現であるのに非再帰的表現と同じ意味のまま使われていると思ひ込むことで、私たちは「自己欺瞞」を「だます」の一形態と考えるが、「自習する *teach oneself*」が「人の助けを借りずに学ぶ」と言い換えられるように、「自分をだます」も「もっている証拠がそれを信じることを保証しないような状況で信じる」とことと言ひ換えられればよいのだとする。

しかし、前者は信念の修正プロセスとの違いを説明することができないし、後者では愚かさや素朴さ、偏見、優柔不断などの状態から自己欺瞞を区別できない。またこの言い換えの中で「保証しないような状況」とはどのような状況なのかということが、再び問題になり、「証拠が *p* を信じることを保証しないと判断される状況」を判断す

るのは誰なのか、という問題が再燃する。

とはいえ、このような自己欺瞞の見方はこの問題の解決に対する重要な示唆を含んでいる。自己欺瞞は確かに信念にまつわる問題であるので、私たちはともすれば信念の内容やその矛盾にだけ目を向けがちである。その結果、心を分割して矛盾を許容できる形にするか、矛盾回避のために片方の信念を否定するか、のどちらかの道をとる論者が多い。それは「自己欺瞞 $\parallel$ 自分をだますこと $\parallel$ p」という信念と not p という信念の共存」と問題を考えるゆえである。それはミールの言う「語彙的アプローチ」であるが、⁽¹⁶⁾「だます」にも様々な意味があることを忘れ、「だます」を固定して自己欺瞞やそれと共通点をもった状態を考えようとしてしまう傾向がある。もしも「自己欺瞞」を「だます」のモデルから説明しようとするなら、その言葉の吟味が必要である。

また最後に簡単にふれた見解の難点が示しているのは、事例に基づいた検討の重要性である。自己欺瞞者は「こうあつてほしいことを信じる」という傾向にあるが、「そうであつてほしくないことを信じる」場合もある。また、自己欺瞞の周辺には区別の難しい、よく似た状態がたくさんある。その代表例が希望的観測であり、半信半疑の状態である。こうしたものの個別例を取り出し、自己欺瞞とは何であるかだけではなく、何でないのか、ということも十分に考察した上で、自己欺瞞を論じるべきである。

註

- (1) 自己欺瞞を心理学的側面から見ると、そこにどんな過程が含まれ、どうやって動機づけられているのか、というような問題になるが、もちろんその問題は分析的アプローチの中である程度論じられる。
- (2) このパラドクスは、「自己欺瞞の静的なパズル」(MELE [2001], p.6)、「表面的パラドクス」(JOHNSTON [1988], p.63)、「ド

クサのパラドクス」(BARNES [1999], p.14) などと呼ばれている。

(3) この説明は MELE [2001] Chap.3 および MELE [1985] を参考に筆者が単純化したものである。

(4) MELE [2001] Chap.5, §4.

(5) 判定者を、実際に自己欺瞞に陥っている時の本人でも客観的な第三者でもなく、「普通の状態にある私」というような形で考えることができるかもしれない。たとえば、他人から指摘されて初めて、それまで自分が自己欺瞞に陥っていたと気づく場合がある。その時には not p という信念が獲得されているわけであり、普通なら not p と信じたはずだという判断をその人はもつわけである。普通であれば信念 not p をもつということをデイスポジションのようなものとして想定することで、この立場の説明を救う方法を考えていくことができるように思われるし、今後の課題としたいが、ミールはその点を全く考慮していない。

(6) たとえば、ある政治家のスキャンダルが表沙汰になり、様々な疑惑が報道される。偶然疑惑の一つについて自分が以前友人に違ったことを得々と解説したことがあったとする。いろいろな情報から判断してもその政治家がやったことのように思えるし、多くの人がそう考えている。しかしその人はそう考える証拠もないのに、彼がやったのではないと信じる。そしてたまたまその件だけはその政治家が関与していなかったというケースはどうだろうか。

(7) AUDI [1985], pp.172-178.

(8) ROTY [1972], PEARLS [1984], DAVIDSON [1985] などがそうである。

(9) DAVIDSON [1990] pp.449-464.

(10) DAVIDSON [1985] p.140. および DAVIDSON [1980]

(11) DAVIDSON [1990] p.459.

(12) JOHNSTON [1988] pp.63-66, BARNES [1997] Chap.2などを参照した。

(13) デイヴィッドソンとよく似た考え方にマクラフリンの「接近可能な信念」と「接近不可能な信念」の区別がある (MACLAUGHLIN [1988])。つまり、自己欺瞞のような現象の場合二種類の信念が含まれていて、だまされる方の信念には接近可能だが、だます方の信念には接近できないので、両方がもたれていても矛盾は生じないということである。この考え方もデイヴィッドソンと同様の困難をもつ。

(14) SORENSEN [1985] pp.64-69.

(15) CANFIELD & GUSTAVSON [1962] 及び MARTIN [1986]

(16) MELE [2001] p.5.

文獻

- AUDI, Robert [1985] : "Self-Deception and Rationality", in M.MARTIN (ed.), *Self-Deception and Self-Understanding*, University of Kansas Press.
- BARNES, Annette [1997] : *Seeing Through Self-deception*, Cambridge University Press.
- CANFIELD, John & GUSTAVSON, Don [1962] : "Self-Deception" *Analysis*, Vol.23.
- DAVIDSON, Donald [1980] : "How is Weakness of the Will Possible?" in *Essays on Actions & Events*, Oxford.
- [1985] : "Deception and Division" in E.LEPORE & B.MCLAUGHLIN (eds.), *Actions and Events*, Basil Blackwell, pp.138–148.
- [1990] : "Paradox of Irrationality" in P.K.MOSER (ed.), *Rationality in Action*, Cambridge University Press, pp.449–464.
- JONSTON, Mark [1988] : "Self-Deception and the Nature of Mind", in MACLAUGHLIN & RORTY(eds.), *Perspectives on Self-Deception*, University of California Press.
- MACLAUGHLIN, Brian P.[1988] : "Exploring the Possibility of Self-Deception in Belief", in MACLAUGHLIN & RORTY(eds.), *Perspectives on Self-Deception*, University of California Press.
- MARTIN, M.W. [1986] : *Self-deception and Morality*, University Press of Kansas.
- MELE, Alfred [1985] : *Irrationality*, Oxford University Press.
- [2001] : *Self-Deception Unmasked*, Princeton University Press.
- PEARS, David [1984] : *Motivated Irrationality*, Clarendon Press.
- RORTY, Amelie O. [1972] : "Belief and Self-deception", *Inquiry*, Vol.15, 1972.
- SORENSEN, Roy [1985] : "Self-Deception and Scattered Events", *Mind*, Vol.94, pp.64–69.