

Title	欲求の反映と人称的バイアス
Sub Title	Reflection principles for desires and personal bias
Author	柏端, 達也(Kashiwabata, Tatsuya)
Publisher	三田哲學會
Publication year	2014
Jtitle	哲學 No.133 (2014. 3) ,p.1- 21
JaLC DOI	
Abstract	<p>The rational relation of an agent to its attitude at one time and to its attitude at another time is a general problem of rationality. Elizabeth Harman has perceptively suggested that 'Parfit's fourteen-year-old girl' case involves this type of problem. I sympathize with some of Harman's main conclusions: The girl's later preference for her loved one can be reasonable, and the prior overall judgment that she should prefer not to conceive a child can also be correct. However, I regard some of the principles that are abandoned or restricted by Harman as more defensible than she believes.</p> <p>In this paper, I show that one of the principles that Harman abandons, 'Reflection for Desires', which she regards as an analogue of the 'Reflection Principle' for subjective probability (cf. van Fraassen 1984), is not an exact analogue of it. The proper analogue of the consequence of the principle—that is, $P_a(t(Da, t+\alpha(p))=1 \supset Da, t(p))$—is not refuted by the fourteen-year-old girl case, since its antecedent is not satisfied in that case. In addition, I suggest that the principle that the better alternative should be preferred need not be restricted by this case, if we integrate personally biased, egocentric factors with others into our all-things-considered judgment. Which world is better as a whole for us can vary from time to time, especially when our new loved ones come into existence.</p>
Notes	投稿論文
Genre	Journal Article
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=AN00150430-00000133-0001

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.

欲求の反映と人称的バイアス

—— 柏 端 達 也* ——

Reflection Principles for Desires and Personal Bias

Tatsuya Kashiwabata

The rational relation of an agent to its attitude at one time and to its attitude at another time is a general problem of rationality. Elizabeth Harman has perceptively suggested that 'Parfit's fourteen-year-old girl' case involves this type of problem. I sympathize with some of Harman's main conclusions: The girl's later preference for her loved one can be reasonable, and the prior overall judgment that she should prefer not to conceive a child can also be correct. However, I regard some of the principles that are abandoned or restricted by Harman as more defensible than she believes.

In this paper, I show that one of the principles that Harman abandons, 'Reflection for Desires', which she regards as an analogue of the 'Reflection Principle' for subjective probability (cf. van Fraassen 1984), is not an exact analogue of it. The proper analogue of the consequence of the principle—that is, $P_{a,t}(D_{a,t+a}(p)) = 1 \supset D_{a,t}(p)$ —is not refuted by the fourteen-year-old girl case, since its antecedent is not satisfied in that case. In addition, I suggest that the principle that the better alternative should be preferred need not be restricted by this case, if we integrate personally biased, egocentric factors with others into our all-things-considered judgment. Which world is better as a whole for us can vary from time to time, especially when our new loved ones come into existence.

* 慶應義塾大学文学部教授 (哲学専攻)

1 存在そのものの価値

人はしばしば、特定の人物の存在そのものに固有の価値を見いだす。自分自身や、自分が愛する者などの存在のなかに、人生内部の出来事がもたらす諸価値へと還元できない価値を、それゆえ、人生において生じるさまざまな災厄や不幸によっては単純に相殺されない正の価値を、見いだす。たとえば障害をもつ子供の親が次のように感じることは、すくなくとも私には完全に理解可能である。すなわち、障害をもつということは、たしかに本人にとっても親にとっても大変なことだ。他の子供が普通に遊んでいる姿を見てうらやましく思うこともある。しかし、それはけっして、自分の子供がその別の健全な子供と入れ替わってほしいという意味ではない。自分はこの子を愛しており、この子がいなくなるということなど考えられない。あるいは、デレク・パーフィットは、次のような中年男性の憤慨を取りあげている。「数年前、十代の妊娠が前の年に減少したという事実を英国のある政治家が歓迎した。中年の男が怒って『タイムズ』に投稿した。彼は母親がたった十四歳の時に生まれたのだった。彼の母親がそれほど若かったので彼の若いころは母と子の両方にとって厳しかったということを彼は認めた。しかし彼の今の生は十分生きるに値するものだった。その政治家は、彼が生まれぬ方がよかったと示唆しているのだろうか。この示唆は彼には法外なものと思われた。」¹

上のいずれの例においても、自分自身や自分の愛する者が現にあるようなものとして存在しているということに対して、肯定的な評価が下されている。後者の例においてはおそらく総合的な比較考慮を経てそのように判断され、前者の例においてはおそらくある種のジレンマのなかで他方の角を制圧する形で存在の価値が肯定されている。そして他方で、いずれの例でも、自分自身や自分の愛する者が存在しないということを意味する可能性のある状況に対し、強い否定的評価が——法外な、恐ろしいものとして——

下されているように見える。

存在や存在のあり方それ自体がもつこの種の肯定的な価値は、私には実感できるものであり²、また実際、現代の多くの哲学的議論においても自然に前提されたり探究されたりしていると思われる³。しかしもちろん、こうした価値観がどの程度体系的に擁護可能であるかは別問題である。さらに、こうした種類の価値をかりに認めたとしても、人生における苦難が一般に（障害や貧困に関係なく）それらを圧倒してしまうだろうといった意見がありえないわけではない⁴。私の目的は、存在の価値を全般的に擁護することではない。ただ、あるパターンへの批判に対してはそれを擁護したいと思う。その批判とは、第一段落のような発言がどれも意地を張ったり自分を欺いたりした結果にすぎないという批判である。ずいぶんな物言いであるが、その批判はある一般的な理屈を拠り所としうる。それは例の政治家が示唆していたものである。すなわち、未来の苦難が十分に予想される人々については、そのような人々を存在させない選択⁵をする方が総合的に見てやはり賢明なのであり、そしてそのことは、ほんとうは誰にでも理解できる客観的な事実なのであって、一つの個別事例についてそれが過去形で語られている時点においても妥当する、という理屈である。もちろん、だからといって、存在の価値に対する冒頭のような直観が簡単に揺らぐことはないだろう。それどころか人々は、自らの直観をある一般的な教訓で補強しつつ、次のように反論したくなるにちがいない。すなわち、現在のこの価値ある状況をもたらしてくれた過去の選択が悪いものであるなどと、いったいどのようにして合理的に主張可能なのか。あるいは、価値ある未来をもたらすであろうこの選択がどうして悪いものでありうるのか。これらの背景にある「一般的な教訓」が何であるかについての議論は次節ですぐに開始する。ともあれ、ここには、合理性に関する基本的な原理の衝突があるように思われる。

2 十四歳の母

エリザベス・ハーマンは、パーフィットから引き継いだ次のような例によって、問題を定式化している。ある十四歳の少女が、子供がほしいと考えた。少女は、生まれてくる子供も彼女自身も確実に経済的な苦況に陥るであろうことを承知している。だが、彼女の予想では、生まれてきた子供に対して「この子が生まれてよかった」と心から思うようになることは確実だと考えられるし、また、その子自身もいずれ自分の存在を肯定するであろうことはまちがいないと思われる。だとすれば——と少女は続けるのだが——いま子供を作ることに決めるのを躊躇する十分な理由があるだろうか⁶。この例はちょうど、冒頭の怒れる中年男性の状況を彼の母親の視点からとらえたものに相当する（そして彼こそがこの少女の子供であったなら少女の予想は正確だったわけである）。

十四歳の母の話はいわゆる非同一性問題を生じさせる例として知られている⁷。しかしハーマンは、問題を、われわれが受けいれているように見える合理性の諸原理のあいだの齟齬の形で提起する。まず、例の少女は、実際に子供が生まれたならば自分がどう思うようになるであろうかを考慮しつつ、次の原理に訴えているようである。すなわち、

[1] あとで「ああしておいてよかった」と思える方の選択肢を選好せよ。

この原理は、とりわけ目先の快樂に惑わされないようにするための一般的な教訓としておなじみのものである。われわれは、こうした原理に従うからこそ、合理的にも、たとえばパーティに行くのを我慢して論文執筆に専念できるのである。

しかしながら、この例の状況に関して [1] の原理はいささか不都合な帰結をもつように見える。というのも、もし [1] が支配的であるならば、

われわれは少女に対し、子供を作るのはもうすこし待った方が良いという形の合理的な説得ができなくなると考えられるからである。だが、例のようなタイプの状況においては、考慮すべきすべての事柄を冷静かつ客観的に考慮すれば子供はむしろいま作るよりすこしあとで作った方が良いと判断されるケースが十分にありうるだろう。それはまず、母親にとってそうだとされるかもしれない。つまりもしもっとあとで子供を作ったならば、経済的、精神的に余裕のある環境で子育てができ、あなたはより幸せな家庭を築けるであろう、と。あるいはもっと普遍的に、子育てに悩み貧困に苦しむ一組の母子がいる代わりにそうでない一組の母子がいる社会の方がより良いという言い方で、その判断は表現されるかもしれない。いずれにせよ例の政治家は、そのような判断が十代の妊娠に関して一般に正しいと考えたのである。例の政治家が正しいかどうかはさておき、すくなくともわれわれのケースに関しては、そのような判断が正しいと仮定しよう。他方で、人は、次のような原理も合理性に関するあきらかに適切な原理と考えたいはずである。

[2] より良いと総合的に判断される方の選択肢を選好せよ。

それゆえ [2] によって、われわれのケースでは、少女は子供をいま作らないという選択肢の方を選好すべきである。ようするに、同じ状況に関して、ともにもっともに思われる原理 [1] と [2] が正反対の指令を出すように見えるのである。これはやはりある種のパラドックスであろう。

ハーマンはこのパラドックスを [1], [2] のいずれの原理をも放棄または制限することによって解消しようとする。正確には、彼女が問題にするのは、[1] に関してはより一般化された次の原理である。

[RD] もし、 p が真であることを自分が未来に合理的に選好しているであろう

欲求の反映と人称的バイアス

ということを合理的に予想することができ、かつ、そのとき自分が現在より劣化した認知的または評価の状況に置かれていないと合理的に予想されるならば、現在の時点ですでに、 p が真であることを選好すべきである。

ハーマンはこれを「欲求の反映 (Reflection for Desires)」と呼び、信念 (主観的確率) に関するいわゆる「反映原理 (reflection principle)」の類比物と考えている⁸。また [2] は、可能世界のあいだの比較として、次のように定式化される。

[BP] もし、ある世界のあり方が別のもう一つの世界のあり方よりも良く、かつ自分がそのことを知りうる立場にあるならば、より良いあり方をしているその前者の世界を選好すべきである。

この原理は、ある選択肢がより良いものであるという判断から、その選択肢を選好すべきであるという結論を導き出すための、一般的な原理である⁹。

見て分かるように、[RD] にはいくつかの条件が、明白に重要でないケースを除去するために追加されている。たとえば、ある組織に捕まると、特殊な方法によって洗脳され、その結果その組織の一員になってよかったと思うようになるということが分かっているとす。その場合、後のその選好を理由に、その組織に捕まるのがよいという結論をわれわれは引き出さないであろう。現在の選択のための根拠となる未来の選好は、“正常”なものでなければならない。薬物の影響や、評価のための証拠に関する制限が予想されるものであってはならない。

一つの考えとして、少女の主張する未来の選好を、この条件によって排除することが提案されるかもしれない。つまり、生まれてきた子供への愛を端的に不合理なものと思なす考えである。未来に予想される選好が不合理なものであるならば、[RD] はそれを現在の判断に反映させることを要

求しないのだから、われわれは落ち着いて全体を考慮し、いまは子供を作らないことに決められるというわけである。これはもちろん、前節で示した存在に対する諸々の価値を否定する道に等しい。例の中年男性なら激怒するだろう。また、彼でなくても、この提案を飲むのは最後でよいと感じられる。われわれは誰かを愛するごとに合理性を失っていくのだろうか。愛はそこまでわれわれを盲目にするのだろうか。

ハーマンの方針は、反対に、後の時点の選好を理にかなったものと見なす余地を与えようとするものである。そのためにハーマンは [RD] を否定する。彼女によれば、信念はそれが合理的であるかぎり真理という不変的なものをめざすものなのに対し、合理的な欲求や選好はそうでないからである。それゆえ、信念について反映原理が成り立つとしても、欲求についてその種の原理が成り立つ必要はない。もしハーマンの言うとおり [RD] が妥当でないとすれば、未来の（合理的な）選好を現在の判断に反映させなければならないという合理的な制約は一般に存在しないということであり、したがって、とりわけこの場合、未来に予想される問題の選好は、子供を作ることを十四歳の時点で決断することの合理的な根拠にはならないことになる。

欲求や選好の対象の時間的な変動をもっともなものを見なす理由の一つとしてハーマンがあげているのは、「誰をわれわれが愛するかは時間とともに変わり、また人のどの本性をわれわれが愛するかも時間とともに変わる」という事実である¹⁰。そして彼女によれば、[BP] を制限する次のような原理が認められるのである。

[PLOO] 自分の愛する者が現にあるような人物になっているということや、自分の愛する者がまさに存在するに至ったということ自体を選好することは、合理的でありうる。たとえ物事がより良いあり方をしている別の可能性がそのときにあったとしてもである¹¹。

このように「愛するものに対する選好が合理的なものでありうるため、合理的な選好の対象は時間とともに変化する」のである¹²。[PLOO]は[RD]の基盤を掘り崩すものであると考えられていると同時に、明示的に[BP]の一般性とも衝突している。つまり[PLOO]は、存在に対する後の時点の選好を正当化するとともに、それとは別のより良い選択肢が存在していたことを（そして現実よりも良い可能的状況が存在していることを）後の時点において認めることを可能にするのである。

私は、ハーマンの議論がめざすところとその背景にある直観には共感する。すなわち、存在に対する後の時点の選好を不合理なものとして断じるべきではないと考える。存在の価値と愛との典型的なつながりに対する彼女の見解にも頷けるところがある。さらに、他方で、可能な世界の良し悪しに関する前の時点の一般的な判断には理があるとも考えたい。とはいえ、以上の考えを擁護するためにハーマンが制限したり退けたりした諸原理は、私の見るところ、彼女が考えるよりも擁護可能なものである。論文のこのあとの部分で私は、関連する合理性の原理を適切に解釈することによって、それらの原理またはその相当物が、問題となる事例に関して十分に維持しうるものであることを示そうと思う。そしてその過程の一つで、「総合的な判断」というものに関連して、まだ議論の余地はあるが検討に値する一つの考えを示唆することにしたい。

3 欲求の反映

正確にはどのようなものを「欲求の反映」の原理と考えるべきなのだろうか。つまり、いくつかの妥当な推論の背景をなすように見えるとされ、かつ重要な反例によって批判に晒されないかを検討されるべきなのは、正確にはどのような原理なのだろうか。それを知るにはまず、いわゆる「反映原理」からのいかなる類推が、「欲求の反映」のもっともらしさを示唆するものであるのかを明確にしなければならない。

自身の未来の態度のあり方によって、現在の自分が合理性に関して一定の制約を受けることは、たしかにあると考えられる。バス・C・ファン・フラーセンが指摘した「反映原理」はそのうちの一つであろう。その原理は次のようなものである¹³。

$$[3] \quad P_{a,t}(p \mid P_{a,t+a}(p) = r) = r \quad (\text{ただし } \alpha < 0 \text{ でない})$$

ここで「 $P_{a,t}(p) = r$ 」は、主体 a が t 時において事態 p に対して r の信念度合いをもつということを表している。[3] の式を、あえて日常言語に近い形で、 a の観点から表現するなら、「 p であることに対する自分の主観確率が現在または未来において r であるという条件のもとで実際に p であるというようなケースを考えた場合、そのケースに対する現在の自分の主観確率はやはり r である」ようになる。反映原理は規範的な原理であり、合理性に関する要請である。反映原理に反する心の状態に自然になることは実際にあるように思えるが、もし反映原理に従わないならばわれわれはダッチブックを作られてしまう¹⁴（この意味で上の日常表現の末尾は「 r であるべきだ」とするのがよいかもしれない）。

反映原理はいくつもの興味深い哲学的含意をもっており、その表現法にもさまざまなものがある。実際には [3] の帰結の一つにすぎないものが「反映原理」と呼ばれることも多い。また、いくつかの明白な「反例」を退けるため、[RD] におけるものと同様の条件を付加する必要があるだろう¹⁵。それらにここで立ち入ることはしない。目下の問題は、ハーマンが「欲求の反映」と呼ぶものに対応する反映原理の帰結が、正確にどのようなものであるかである。

次の原理は反映原理の帰結の一つである¹⁶。

$$[4] \quad P_{a,t}(P_{a,t+a}(p) = 1) = 1 \supset P_{a,t}(p) = 1$$

前件は特定の状況を表現することを意図したものである。その状況とはつまり、自分がある態度を未来にもつことを現在確実視しているような状況である ($\alpha=0$ のケースはいまは度外視する)。したがって [4] は、「未来に p を確信すると確実に予想されるならば、いますでに p を確信すべし」と読むことができる。私の考えでは、この [4]こそが、[3] の帰結のうち、欲求に関する反映原理のこの文脈における自然な類比物を生み出しうるものである。

欲求に関するもっともらしい反映原理がもしあるとすれば、それは、[4] において未来にもつことが確実視されている態度が、ちょうど欲求に置き換わったものである。すなわち、

$$[5] \quad P_{a,t}(D_{a,t+\alpha}(p))=1 \supset D_{a,t}(p)$$

[4] における「 $P_{a,t}(p)=1$ 」が p に対する t 時における a の確信を表していたのに対し¹⁷、[5] の「 $D_{a,t}(p)$ 」は、 p に対する t 時における a の欲求または選好を表している。(さらに、信念との類比を重視するなら、この「 $D_{a,t}(p)$ 」は「 $\neg p$ ではなく p を欲すること」と読むべきである。すなわちそれは「 $\neg p$ より p を選好する」ということに等しい¹⁸。)[5] は「未来に p を欲すると確実に予想されるならば、いますでに p を欲するべし」と読めるだろう。それは、未来に予想される自分自身の欲求が、現在の欲求のあり方に反映されていなければならないとする原理と見なすことができる。なお、簡潔さのためにここでは明示しないが、[5] の前件で触れられている未来の欲求も、適切な状況下での、合理的な欲求でなければならない。

もちろん [5] は [3] の帰結ではない。信念に関する反映原理の妥当性の話とこの [5] の妥当性の話は独立である¹⁹。しかし、ファン・フレーセンの議論とは別に、われわれの常識的な推論のなかに [5] のような原

理が合理性に関する教訓として示唆されているケースはたしかに存在すると思う。[1] が自然にあてはまるようなケースはおそらくすべてであろう。パーティに出ずに論文を書くという選択が望ましいものだったとあとで思うに決まっているのなら、いまその選択肢を好むべきなのだ。あるいは、未来に予期される欲望が現在の選択に理由を与えるように見える状況においても、[5] の原理を見いだせるかもしれない。たとえば、いまは満腹でしばらく柿など見たくもないが、明日にはまた食べたくなるだろうから、いま柿を購入し明日食べるための準備をしておく理由があるといった場合である²⁰。

重要なのは、われわれの十四歳の母の事例が、[5] のあてはまるケースではそもそもないという点である。その点で [5] は [RD] とは異なる。違いは [5] の前件のイコール1の部分、つまり「未来に p を欲すると確実に予想されるならば」の「確実に」の部分にある。まず [5] があてはまるケースについて言えば、パーティへの出席と論文執筆のいずれの選択をしたとしても、私は、あとで後者の選択肢を良いものと見なすであろう。すなわち、誘惑に負けてパーティに出席すれば「あのとき論文を書いていればよかった」とあとで思うだろうし、意志を強くして論文執筆に専念すれば「あのとき論文を書いておいてよかった」とあとで思うだろう。いずれの選択をするかは独立に、未来における後者の選択肢への選好が、決断の時点で（つまり未来の選好が決断のための根拠として用いられる時点で）予想されるのである。よってこの事例は [5] の前件を満たしている。柿を買うケースも同様である。だが、十四歳の母の事例はそれらとは異なる。決断しようとしている時点を考えられたい。かりにいま子供を作らないとしよう。われわれの例では、仮定により、その選択は総合的かつ客観的に見て正しい判断である。そして少女もまたその判断の根拠の重要な部分を認めうるだろう。つまり、彼女は、もしいま子供を作らなければ自分は経済的に困窮することもなく、よりあとで子供をもち、それな

りにその選択に納得してしまうであろうと予想するかもしれない。またすくなくとも、その場合近い未来に特定の子供は存在せず、それゆえ「この子が生まれていればよかった」という形で逆の可能性を渴望することが不可能であるということ、彼女は理解するはずである²¹。これがもし十四歳の母の例の状況だとすれば²²、それは、反映原理の欲求に関する適切な類比物を脅かすものではない。その前件がそもそも満たされないからである²³。

ちなみに、自分の未来の態度に確信がない場合には、もともとの信念に関する反映原理についても同様の、類比的なことが言える。たとえば冷蔵庫の中にケーキがあるかどうかについてまったく確信がないとしよう。すなわち冷蔵庫の中にケーキがあるということに対する私の主観確率はちょうど0.5である。これから私は冷蔵庫の中を確かめにいくつもりだ。それでどちらかははっきりする。いま、反映原理に従って、次のように考えてもよいだろう。冷蔵庫の中を見たあと私がそこにケーキがあると確信しているとすれば、それは冷蔵庫の中に実際にケーキがある場合であり、その場合に限られるはずだ、と。つまりもし冷蔵庫の中にケーキがあるならば、私は未来に確実に、そこにケーキがあると信じているであろう。だがもちろん、そのように予想されるからといって、冷蔵庫の中にケーキがあるといま信じるべきだということにはならない。[4]の前件は満たされていない。他方では冷蔵庫の中にケーキがないと私が未来に確信している可能性もあり、それは冷蔵庫の中にケーキがない場合だからである。以上のストーリーに反映原理に反する部分はない。反映原理はここではむしろ「ちゃんと見たとおりに信じる」ということを主張する原理として機能している。

私はこの論文で、欲求の反映を全般的に擁護する論証を提出したいわけではない。そうではなく、ハーマンの示唆する議論によってそのおそらく最も正確で有望なバージョン——それは [RD] ではない——に瑕がつく

ことはないということを示したいのである。そしてもし私が正しいならば、そのことが意味するのは、問題のケースにおいて十四歳の少女は、決断をするさいにそもそも反映原理の類比物には訴えられないということである。

4 総合的な判断

合理性の諸原理のあいだの齟齬は、後の時点においても生じているように思われる。例の少女がけっきょく十四歳で子供をもったとしよう。そのとき、一方で、愛する者の存在に対する選好がそれだけで理にかなったものでありうるという（[PLOO]の背景にあるような）原理を受けいれて、さらに欲求の現時点への反映（ $\alpha=0$ の場合の [5]）を考慮し、我が子の存在を積極的に肯定しつつ、他方で、この子は総合的に見れば存在させないのが正しい判断だったと主張することが、いかにして可能なのだろうか。というのも、かつての総合的判断は、もし正しいものであったとしたら、事情に変更がないかぎり現在も正しいはずだからである。ハーマンは、[PLOO]によって [BP] を制限し、これらの二つを両立させているように見える²⁴。ともあれわれわれは、よし悪しの総合的判断の無時間性に関して、一般に次のように主張することができるだろう。

[6] pであることがより良いと t_1 に総合的に判断されたならば、新しい証拠の出現といった重要な認知状況の変化がないかぎり、 t_2 でもそう判断すべきである。

一方で世界は刻々と遷移しており、主体が判断をするための条件も時間とともに変化する。したがって [6] の「重要な認知状況の変化がないかぎり」という但し書きは必要である。どのように選択すべきかを判断するのはわれわれなのだが、われわれは時間のなかに存在しており、またつねに

完全な情報を手にしうるわけでもない、いくつかの重要な事柄が選択のあとに生じたり、選択をしてはじめて判明したりする、一つの選択に対するすべてを考慮した最善の判断は、もちろん、時点に相対的に変化する余地を残していなければならない。

さて問題は、この [6] と [BP] によって、この子——冒頭の中年男性自身から見れば「私」——が生まれてきてここにいるというのはとても望ましいことだという事後の選好が否定されるかどうかである。見たところそれは否定されると思われるかもしれない。現在の状況は決断前にまさに予想されていた状況であり、その間、予想外のことはとくに何も生じていないからである。仮定により、子供がいま作られない世界の方がより良い世界であるという判断がすでになされているのだから、[6] によってそれは現在の正しい判断でもあり、そうすると [BP] によって、この子が（あるいは私が）いまここにいるということを選好すべきではない、という結論が導かれるように見える。

しかし、そうではない、と私は考える。またそのように考えるために [BP] を否定したり制限したりする必要もないと考える。決め手は、上で明示された [6] の条件節の一つ「重要な認知状況の変化がないかぎり」にある。私の考えでは、十四才の母のケースにおいて母親が決断する前の時点 t_1 と子供が誕生した後の時点 t_2 とでは、ある重要な認知状況の変化がある。たしかに t_1 時には、子供を作らない方が、一人の子供を作るよりも総合的に見て良かった。ところが後者が選択され t_2 になった時点では状況が異なる。 t_2 時において母親がもし総合的に判断するとすれば、すなわち、彼女が考慮しうるすべての事柄を考慮するとすれば、この子がいた方が良いのである。そしてその判断からは、まさに [BP] によって、そのような世界をもたらした選択が好ましいということが引き出せるのである。それぞれの時点において [BP] は正常に機能しうる。あるいは、子供である例の中年男性の視点からは、次のように述べることができるだ

ろう。すなわち「 t_1 時にはいなかった私が t_2 時にはいる。もし t_2 時のこの状況を離れて一般的に述べるならば、私のように苦勞をする人間は少ない方が良いでしょう。だが自分がいるというこの状況は、総合的に考えて、私にとっては良いものであり、それゆえに、私には好ましいものである」と²⁵。

一点、確認をすれば、「まさにあの境遇で生まれたこの私がここにいることは好ましい」から「あのような境遇で生まれる人物がいるということは好ましい」という一般的な主張は帰結しない。存在的一般化によって「まさにあの境遇で生まれたこの私がここにいる」から「あのような境遇で生まれる人物がいる」が帰結するとしてもである。したがって前段落の中年男性の言葉のなかに不整合なところはない。事情は「私が宝くじに当たったのは奇跡である」から「すくなくとも一人の何者かが宝くじに当たるのは奇跡である」が帰結しないのと同じである²⁶。

いったい t_2 の時点では認知状況の何が変化したのか。正直なところ私はそれを「認知状況」の変化と呼ぶことが適切であるかどうかについて自信がないのだが²⁷、とにかく、 t_2 の時点において、人称的バイアスをもたらすような新しい存在やそれとの新しい関係が生じていることはたしかである。「人称的バイアス」とは、たとえば次のようなものである。甲氏が痛い目に遭うか、乙氏が痛い目に遭うかのいずれかであるとする。完全に等質等量の痛みと損害がいずれかに与えられる場合であっても、あなたがもし甲氏であるならば、あなたは後者の選択肢を望ましいと考えるだろう。もし乙氏があなたの愛する人であったならば、あなたは逆の選好をもつかもしれないが、それもまた人称的バイアスの一つである。人称的なバイアスはさまざまな形で人間の推論を左右しうる。とりわけわれわれは、ほかならぬ自分自身や自分の愛する者の存在やその本質的ありようを左右する事柄に無関心ではいられないだろう。(こうした「バイアス」のもとでいかに判断すべきかについての標準的で統一的な理論を、われわれはま

だ手にしていないように見える²⁸。それどころか、人称的バイアスがそもそも合理的な判断を構成しうるかどうかにしてさえ意見が分かれるかもしれない。しかしすくなくとも、こうした選好の偏りはわれわれにとって完全に自然なものではあろう。)

人称的バイアスは、よし悪しについての三人称的、客観的な判断からは独立の選好を、われわれに与える²⁹。それゆえ、もし「総合的な判断」が人称的バイアスを含みうるのだとしたら、ここまで曖昧にひとまとめにして語ってきた総合的な判断と客観的な判断を区別しなければならない。私は、人称的バイアスを考慮した総合的判断は可能であると考え、われわれは通常それを考慮して判断を行なうものだとさえ思う。

この節の結論は以下である。例のより後の時点においても、合理性の原理のあいだの齟齬はじつは存在しない。上述の理由により、特定の存在である自分自身や自分の愛する者を誕生させない方の選択が正しかったと語りうる認知状況ではもはやないからである。後の時点では、自分自身や自分の愛する者の存在それ自体を総合的に考慮に入れて、現状を選好することが可能なのである。そうした選好が、他の重要な原理と衝突するがゆえに不合理だとされることはない。また他方で、一般にこのような境遇の人間はいいよりいることが好ましいという結論が、そこから導き出されることもない。

5 結び

判断や選択の主体は、判断や選択を行なうときのみ存在するわけではない。それは一定の期間にわたって同一人物として生きつづけるような存在者である。そのため、ある時点における自分の態度と別の時点における自分の態度とをいかに合理的に関係づけるかが、ときとして主体にとって重要な問題となりうる。いわゆる十四歳の母のケースも、ハーマンがあきらかにしたように、時間をまたぐ合理性に関するそうした興味深い問題を

含むように見える。その一つは、後の時点の選好を前の時点に「反映」させるための原理にまつわるものであろう。しかしそれに対する本稿の指摘は否定的な形のものであった。すなわち、その種の合理性の原理は、十四歳の母のようなケースについて言えば、そもそも重要な仕方では関わっていない、というものだった。関連して私が指摘したもう一つの論点は「総合的」な最善判断の時間性に関するものである。本稿では、その種の判断の時間的変化を正当化しうる条件の一つに、人称的バイアスを新たにもたらす状況の変化を数え入れることを提案した。十四歳の母のケースはまさしくそのような変化を典型的に見いだせる事例なのである。

謝辞：この論文は、MIPS（三田哲学会）哲学・倫理学部門2012年年次例会、および応用哲学会第5回（2013年）年次研究大会において行なった発表に基づいている。それぞれの会場においてコメントをくださった方々にお礼を申しあげたい。また本研究は、文部科学省私立大学戦略的研究基盤形成支援事業（平成24～26年度）による「社会的行動判断・意思決定のダイナミズムに関する人文科学、行動科学、神経科学の分野横断的統合研究」プロジェクト（慶應義塾大学「思考と行動判断」研究拠点）からの援助を受けたものである。

注

- ¹ Parfit (1984), p. 364 [邦訳書 p. 497]. 訳文は、鍵括弧と句読点に関する二箇所の表記と、また一箇所の時制表現を本稿で用いる形に改変した以外、邦訳書に従っている。
- ² もちろん、あらゆる人が類似のケースでこの種の価値づけを行なうだろうとは思っていないし、行なわなければならないと主張するつもりもいまの私にはない。たしかに、愛する対象の存在そのものを肯定することは、愛の概念と密接な（ひょっとすると本質的な）結びつきがあると論じることではできるかもしれない。しかし、すくなくとも、自分自身の存在自体に何らかの肯定的な価値を見いださなければならないかどうかについては（いかに「自己愛」がおなじみのものであろうと）議論の余地があるだろう。そして、より明白なことには、自分の存在自体に特段の価値を見いださない人がいることは事実である。

- ³ たとえば, Nagel (1970a), p. 2, p. 7 [邦訳書 pp. 2-3, p. 11]; あるいは, 森岡 (2007). それらのテキストで主張されたり主張の前提となったりしている存在の肯定的価値が, すべて正確に同じ種類のものであるとは私は考えていない.
- ⁴ たとえば Benatar (1997).
- ⁵ ここで「存在させない選択」とは, 身籠らないという選択のことであり, 中絶するという選択のことではない. 中絶の選択が賢明かどうかは, 私の考えでは, 存在を消去することが賢明かどうかと同様, まったく別の要素が関わる問題である.
- ⁶ Harman (2009), pp. 181ff. ハーマンはこれと同種または類似のいくつかの例を取りあげて論じている. 母になることを決意するこの十四歳の少女の例は, もともとは Parfit (1984), pp. 358ff [邦訳書 pp. 489ff] においてやや別の観点から論じられていたものである.
- ⁷ Parfit (1984), p. 359 [邦訳書 p. 490]. 前段落の少女はさらに「別の選択をしたとして, その場合にあとで生まれてくる子供は別人なのだから, これはいづれにせよ特定の子供にとってより悪い選択とならない」と続けることができるだろう. 非同一性問題については Kavka (1982) 等を見られたい.
- ⁸ Harman (2009), pp. 182-3.
- ⁹ Harman (2009), p. 185. ハーマンは「より良い」のいくつかの意味を分析し [BP] をさらに詳細化している (*ibid.*, p. 186). しかし, とりわけ (非同一性問題と関わりうるような) 十四歳の母のケースを扱うとすれば, ある選択肢が, 非人称的な観点から, そしてある種総合的で客観的な観点から母親にとって, より良くそして好ましいものだと [BP] のような原理から言える, という点を押さえておけばよい.
- ¹⁰ Harman (2009), p. 188.
- ¹¹ Harman (2009), p. 186.
- ¹² Harman (2009), p. 188.
- ¹³ van Fraassen (1984), p. 244.
- ¹⁴ ファン・フラーセンの考える巧妙なダッチブックは van Fraassen (1984), p. 258 にある.
- ¹⁵ 予想される薬物の影響, 錯乱, 記憶の自然な減衰等は, 意図された合理性の原理を主張するために除外されなければならないだろう. それに関しては van Fraassen (1995) などを参照されたい.
- ¹⁶ [4] の前件が成り立つとする. すると, [3] の $r=1$ のケースに条件付き確率の定義を適用したもつから, $P_{a,t}(p \wedge P_{a,t+a}(p)=1)=1$ が得られる. いかなる事象の確率も 1 以下なので, そこからは (乗法定理を考慮して) $P_{a,t}(p)=1$,

すなわち [4] の後件が帰結する。

- 17 また、「 $P_{a,t}(p)=0$ 」は p でないことに対する t 時における a の確信を表し、「 $P_{a,t}(p)=0.5$ 」は p かどうかに関して t 時に a が信念をもたないことを表すと考えられるだろう。その他の中間的な値の定義の仕方には選択肢があるが、「 $P_{a,t}(p)>0.5$ 」を p を信じている状態とするのはおなじみのやり方である。
- 18 この論文では欲求の程度というものを考慮しない。そうしたものが信念の度合いとの類比でどこまで整合的に定義できるかが明白でないからである。ただしここでは、 $D_{a,t}(p) \wedge D_{a,t}(\neg p)$ という状態を排除するような意味で「欲求」の語を使う必要はある。同様の配慮は Harman (2009) でも前提とされている (*ibid.*, pp.183-4)。
- 19 驚いたことに、以上の諸点は複数の聡明な質問者によって誤解された。きつと発表時における私の提示の仕方が拙かったのだろう。ここまでの私の関心事はあくまで反映原理の「類比物」の名に値するものが何かである。
- 20 Nagel (1970b), p. 40. ネーゲルによれば「理由の影響の及ぶ範囲は時間を超えて拡大しうる。というのも、理由のある、または未来に理由が生じるであろう事柄に向けて、現在準備を進める理由があるからである」(*ibid.*, p. 45)。あるいは、明日にもまた柿を食べたいという気がいまはしているが、その欲望が明日になって消滅していることはあきらかなので、いま柿を購入しない選択をする理由がある、といった場合も考えられるだろう (cf. *ibid.*, p. 39)。ちなみに、ネーゲルのテキストで柿の例えが直接的に使われているのは、以上とは異なる状況に対してなので、注意されたい。
- 21 ここには非対称性があり、逆の決断をして子供を作った未来においては、生まれてきた特定の子供がいない世界について自分は否定的な評価を下すであろう、と現時点で予想することができると思われる。この非対称性は、たしかに問題のケースの興味深い特徴の一つであると言えるが、ここで示そうとしている論点を左右するものではない。
- 22 逆に少女が、子供をいま作らなければ自分はきつと後悔するにちがいないと確信しているならば、それは、特定の子供の非存在を根拠とするものではありえない。おそらくそれは、一般的に予想される寂しさや、あるいは妊娠可能性に関する特殊な事情に基づくものであろう。この場合彼女は、いま子供を作るべきでないとする事前の判断に反論しているのである。よってこれはいわゆる「十四歳の母」のケースではない。もしも少女のその総合的な判断が事態を正しくとらえているならば、彼女は、[1] のパターンの理屈に依ることなく大人たちを説得しうるだろう。
- 23 十四歳の母のケースにおいて逆の選択肢がもつこの（「それを選択した場合も

悪くはなかったと考えるであろう」という)特徴はハーマン自身も指摘している (Harman (2009), p. 182, pp. 189-91). しかし私は, その特徴を, [1] のような原理がなにか根本的な問題を抱えているということを示すものとしてではなく, 本文で論じたように, 問題のケースが適切に理解された [1] の適用例でないということを示すものとしてとらえたい.

²⁴ パーフィットも両方の主張を行なうことになるだろう (Parfit (1984), pp. 360-1 [邦訳書 pp. 492-3]).

²⁵ もちろんこの中年男性は, t_1 の時点においてはいかなる評価も下すことができない. 存在しなければ評価は下せないからである. 彼にできるのは, t_2 の時点において, t_1 の時点における母親の決断を評価することである.

²⁶ この種の奇跡の分析については柏端 (2005) を参照.

²⁷ 自分が乙氏ではなく甲氏であるといった事実は, 世界の構造に関わる事実 ([BP] の定式化が第一に想定するであろうような事実) とは通常考えられないだろう. そのような要素を「状況」の語でカバーしようとするのには無理があるかもしれない.

²⁸ 人称的なバイアスに対する注目すべき理論的アプローチとして, Hare (2009) がある. キャスパー・ヘアはそこで, 人称的バイアスに対して (時間的バイアスに対するのと類比的な議論に基づいて) 形而上学的な根拠づけを行なうことを試みている.

²⁹ もっとも人称的バイアスがつねに完全に自己中心的な判断をもたらすわけではない. たとえばまったく知らない誰かが事故死するよりは, 自分や自分の愛する者の爪が剥がれる方が, 総合的に見て (つまり人称的バイアスを考慮に入れても) かなりましであると私には思える. こうした判断はヘアの言う「穏健な自己中心的快樂主義者」において可能となる (cf. Hare (2009), pp. 35-7).

文 献

- Benatar, D. (1997), "Why It Is Better Never to Come into Existence," *American Philosophical Quarterly* 34(3), 345-55.
- Hare, C. (2009), *On Myself, and Other, Less Important Subjects*, Princeton University Press.
- Harman, E. (2009), "'I'll Be Glad I Did It' Reasoning and the Significance of Future Desires," *Philosophical Perspectives* 23, 177-99.

- 柏端達也 (2005), 「23.2576 インチの魚と微調整された宇宙と永遠の謎について」,
横山輝雄編, 平成 13~15 年度文部科学省科学研究費補助金研究成果報告書
『宇宙論における人間原理に関する自然哲学的研究』, 39-47 (http://pssj.info/program/program_data/37/37ws/Kashiwabata-Fish.pdf).
- Kavka, G. (1982), "The Paradox of Future Individuals," *Philosophy and Public Affairs* 11, 93-112.
- 森岡正博 (2007), 「生命学とは何か」, 『現代文明学研究』 8, 447-86.
- Nagel, T. (1970a), "Death," in *Mortal Questions*, Cambridge University Press,
1979 (永井均訳, 『コウモリであるとはどのようなことか』, 勁草書房,
一九八九年), 1-10.
- (1970b), *The Possibility of Altruism*, Oxford University Press.
- Parfit, D. (1984), *Reasons and Persons*, Oxford University Press (森村進訳, 『理由
と人格』, 勁草書房, 一九九八年).
- van Fraassen, B. C. (1984), "Belief and the Will," *Journal of Philosophy* 81, 235-56.
- (1995), "Belief and the Problem of Ulysses and the Sirens," *Philosophical Studies* 77, 7-37.