SURE 静岡大学学術リポジトリ
Shizuoka University REpository

Is the Experience Machine Thought Experiment a Knock-down Refutation of Hedonism?

メタデータ	言語: jpn 出版者: 公開日: 2018-08-02 キーワード (Ja): キーワード (En): 作成者: 米原, 優 メールアドレス: 所属:
URL	https://doi.org/10.14945/00025651

経験機械で快楽説をやっつけられるのか

米原　優

はじめに

よく知られているように、ロバート・ノージックは『アナーキー・国家・ユートピア』の中で、「経験機械」という思考実験を提示した。この思考実験は「快楽が、そして、それのみが、それ自体として望ましいものである」という快楽説への有力な反対論の一つと考えられている（cf. Kymlicka 2002, 13／二〇～二一頁、児玉二〇一二、一四八～一四九頁）[1]。しかし、ほんとうにそうだろうか。筆者が見る限り、この思考実験だけで、快楽説を否定することはできない。それどころか、快楽説に有利な証拠の一つとして使える（使われてしまう）可能性もある。そう言える理由を明らかにするのが本稿の目的である。

構成は以下の通りである。まず、次節において、人が経験機械を拒絶する理由に関するデ・ブリガードの主張を紹介する。彼が提示する理由は、ノージック自身が考えるのとは異なるものである。その上で、第二節では、行動経済学者であるトヴェルスキーとカーネマンの主張に従いつつ、人が経験機械につながれたくないと思う理由を、苦痛の回避という観点から説明する。さらに、第三節では、まず、何らかの苦痛を被ることなしに、ある快楽を獲得するこ

とはできないという状況において、快楽説に従った場合に、快楽の獲得と苦痛の回避のどちらが望ましいと判断されるのか、この問題に関するミルの主張を紹介する。そして、こうした判断方法に従えば、経験機械につながれたくないという人々の思いは、快楽説によって、理にかなったものと評価されると指摘する。その上で、快楽説は経験機械に対する多くの人々の直観を肯定するものであり、この点で人々にとって歓迎すべき理論となる可能性があると論じる。最後に、結論で、経験機械という思考実験を使って快楽説を否定しようとする人が、以上で述べたような事態を避けたいのならば、どういった課題に取り組まなければならないのかを述べる。

第一節　経験機械

ノージックの経験機械とは次のような思考実験である。

あなたが望んだ経験全てを与える経験機械が存在すると想定してみよう。超一流の神経心理学者が、あなたの脳を刺激することで、素晴らしい小説を書いたり、友人を作ったり、おもしろい本を読んだりしていると、あなたに考え感じさせることができるとする。その間ずっと、あなたは脳に電極をつけられて、水槽のなかに浮かんでいる。あなたは自分の人生での経験を前もってプログラミングし、この機械に永久につながれるべきだろうか。あなたが望ましい経験を見逃してしまうのではないかと、不安に思っているのならば、企業が多くの他の人の人生を完全に調査していると想定することもできる。あなたはこのような経験の巨大な集積のなかから選択し、たとえば今後二年間の人生での経験を選ぶことができる。二年が過ぎれば、十分か十時間水槽から出て、次の二年の経験を選ぶことになるだろう。もちろん、水槽のなかで、あなたがそこにいると知ることはない。あなたはそ

れを実際に起こっていることだと思うだろう。他の人たちも自分が望む経験をするためにつながれることができるので、彼らの世話をするために機械につながれないままでいる必要はない。(Nozick 1974, 44／六七～六八頁)

我々の大多数はこうした機械につながれたくないだろうとノージックは論じる。というのも、私たちは「現実に触れつつ、自分自身の生を生きる(これは能動態の動詞である)(live (an active verb) ourselves, in contact with reality)」ということを快楽という経験よりも大事なものと考えているからである(ibid., 45／七〇頁)。もしそうならば、快楽説はそうした私たちの価値観と相容れないものと言えるだろう。

しかし、こうした価値観はほんとうに存在するのだろうか。デ・ブリガードはある実験を使って私たちがそのような価値観を持ってはいないということを示そうとしている。この実験では、まず「よくも悪くもないお話(neutral vignette)」「悪い(negative)お話」「よい(positive)お話」がそれぞれ二四人の被験者に配られた。それぞれの内容は以下の通りである。

よくも悪くもないお話

土曜の朝、もう一時間ほど寝ていようかと思っていたとき、突然玄関のチャイムが鳴った。あなたは渋々ベッドから起き上がって、玄関を開けに向かう。玄関の外には、「ミスター・スミス」と名乗る黒い上着にサングラス姿の背の高い男がいた。彼が言うところでは、あなたに直接関わる重大な情報があるとのことである。ちょっと戸惑ったものの、興味を駆られたあなたは、彼を家に招き入れた。「残念なことに、ちょっとばかり困ったニュースをお伝えしなければならないのですよ」、ミスター・スミスはそう言った。「とんでもない手違いがありましてね。

あなたの脳は、超一流の神経心理学者が作った経験機械に、間違って接続されてしまっているのです。あなたが経験したことは、みんなあなたに快い経験を提供するために作られたコンピューター・プログラムの産物以外の何物でもないんですよ。あなたがこれまでの人生で被った不快なことも、全部より大きな快楽をもたらすためのきっかけでしかないわけで（たとえば、あのコンサートのチケットをゲットするために長い列に並んで待たなきゃならんかったこととか、覚えてます？）。残念なことにね、間違えたことに、さっき気づいたのです。あなたは接続されることになってなかったんです。それは別の人でね。本当に申し訳ありません。そのようなわけで、どっちか選んでほしいのです。この機械に接続されたままでいることもできますし（そのときは、この会話の記憶も消去します）、現実の生活に戻ることもできますよ」。（De Brigard 2010, 47）

「悪いお話」では、この後に次の台詞が追加されている。

悪いお話

ちなみに、あなたの現実の生活はこの仮想現実での生活と全く違うものだってことを知っといた方がいいかもしれないですね。現実の世界において、あなたはウェスト・ヴァージニアにある万全の警備を備えた刑務所の囚人です。（ibid.）

「よいお話」では、「よくも悪くもないお話」の後に次の台詞が追加されている。

よいお話

ちなみに、あなたの現実の生活はこの仮想現実での生活と全く違うものだってことを知っといた方がいいかもしれないですね。現実の世界において、あなたはモナコに住む大金持ちの芸術家です。(ibid.)

それぞれのお話を読んだ被験者たちは、「つながれたままでいる」「現実世界に帰る」のどちらかに○をする。その結果、「悪いお話」を読んだ被験者の一三％（三人）が「現実世界に帰る」、八七％（二一人）が「つながれたままでいる」を選択し、「よいお話」では、その数は半々であった。そして、「よくも悪くもないお話」では、五四％（一三人）が「現実世界に帰る」、四六％（一一人）が「接続されたままでいる」を選んだ（ibid., 47-48）。つまり、現実世界が仮想世界より悪いものであれば、多くの人はそこにとどまることを望むし、そうでなくても、（たとえ、現実世界が非常によいものであったとしても）半分の人は仮想世界に残ることを望むというわけである。

さらに、八〇人の被験者に対し、次のような「第二のよくも悪くもないお話」も配られた。そこでは、「よくも悪くもないお話」のあとに、次のような台詞が追加されている。

第二のよくも悪くもないお話

この機械に接続されたままでいることもできますし（そのときは、この会話の記憶も消去します）、接続を切ることもできますよ。でも、機械外での生活はあなたがこれまで経験した生活とは全くの別物だってことは知っておいた方がいいかもしれないですね。(ibid., 49)

このお話を読んだうちの五九％（四七人）が「つながれたままでいる」を、残りの四一％（三三人）が「接続を切る」を選んだとのことである。ここでも、相当数の人が、仮想世界にとどまることを選んでいる。

そうした実験結果を受けて、デ・ブリガードは「現実に触れつつ、自分自身の生を生きる」ということは、それほど重視されていないのではないかと考える。では、なぜ人は経験機械につながれたくないと思うのか。この点について、彼は次のように論じている。

私が考えるところでは、経験機械につながれるということに反対するという人々の直観的反応を駆り立てるのは、現実という性質に関する省察でも、快楽への選好でもなく、現状（status quo）を維持することへの心理学的バイアスである。私の理解では、この文脈において、「現状」とは、（ノージックの言葉を使えば）「内側から」経験された自分自身の生活のあり方をおおよそ意味する。それゆえ、いくつかの事例における、経験機械から切り離されるということへの人々のためらいは、この根本的な心理学的現象の表出であろう。その現象とは、彼らの現状、つまり、これまで生きてきた生活を失うということへの嫌悪である。（悪いお話や第二のよくも悪くもないお話において、そうであるように）ほとんどの人々が経験機械で生活してきた後に、それから切り離されるということを選好しない理由の説明の一部は、その経験の仮想現実的な性質と関わるものでも、感じると言われている快楽の量とも関わるものでもなく、ほとんどの人々は、自分たちが知っている生活、つまりは、これまでのところ生きてきた生活、要するに、なじみ深く、楽しい生活をやめたくはないという単純な事実と関わるものであろう。（ibid., 50-51）

ここで言われているバイアスとは、行動経済学の分野において、「現状維持バイアス（status quo bias）」と呼ばれるものであり、それは「現状を変えうる決定に直面したとき、「それを手放すことの不利益は利益よりも大きいものとして表れる」ので、人々は自分たちが今ある状態、つまり、現状を選好する傾向があるという事実」を意味する（ibid., 50）。そして、この現状維持バイアスは「損失回避」という認知バイアスの一事例と考えられている。また、この損失回避とは、トヴェルスキーとカーネマンによれば、「損失は利得よりも大きいものとして現れる」ということであり、つまりは、次の図（Figure1）で表されるように、人はある量の損失（Losses）に対して、同じ量の利得（Gains）に対するプラスの評価（Value）以上に大きなマイナスの評価を下すというものである（Tversky & Kahneman 1991, 1041-1042）。それゆえに、Figure1において、利得や損失の大きさを表す横軸の中心線からの距離が同じ地点において、人が下す評価の高低を表す縦軸の中心線からの距離は異なっている（マイナスの評価の方が大きい）。

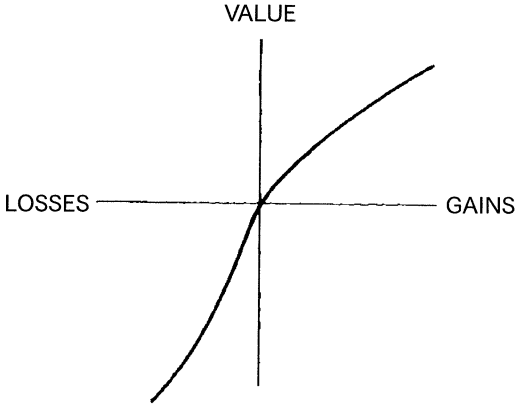

FIGURE I
An Illustration of a Value Function
（ibid, 1040の図）

そして、デ・ブリガードの考えによれば、人は「現実に触れつつ、自分自身の生を生きる」ということを大事にしているからではなく、今までの生活で手にしてきたものを失いたくないから、経験機械につながれたくないと思っている。また、こうした思いは、彼と同じように経験機械への拒絶理由を考察している成田の次の発言にも表れている。

しかし、我われはたいてい、現実世界で生きているうちに、特定の人や物と強い結びつきを築きあげ、そのけっか、その人や物が自分にとってかけがえのないものになっていることがある。そのような人と物と別れることは、辛いばかりでなく、人生における重大な損失であるように思われる。もちろん、コンピュータのプログラムしだいで、そのような人や物とそっくりの人や物を経験機械の世界に出現させることは可能である。しかし、経験機械の世界に現れたそれらの人や物は、現実世界で自分が強く結びついていた人や物とは別人であり別物である。このように、人生のなかばで経験機械の世界に移住することには、かけがえのない人や物との別れがともなう。

（成田二〇〇七、六二〇頁）

その上で、「このような別れを理由に、人生なかばで経験機械につながれることを拒むことは、それになりに理にかなっているように思われる」とも論じている（同上）。デ・ブリガードの実験は、多くの人がそうした理由で経験機械を拒絶しているということを示そうとするものとも言えるだろう。

もちろん、ほんとうに人がそういう理由で経験機械を拒んでいるのかどうかは検討の余地があるし、彼が行った実験だけで、それが真実であるということの証明はできないだろう。この点に関しては、さらなる検証が必要となるように思われる。

さらに、仮にデ・ブリガードの主張が正しかったとしても、それは快楽説にとってあまりいいお話ではない。というのも、もしそうならば、経験機械につながれたくないと考える人々は、自分にとってかけがえのない人や物がある「現状」を、快楽以上に大事なものと考えて、そうしているということになってしまうからである。となると、経験機械という思考実験は、依然として、快楽よりも大事なものはあるという私たちの価値観の存在を示すものとなってし

まう。確かに、その「大事なもの」とは「現実に触れつつ、自分自身の生を生きる」ということであるというノージックの考えは退けられるかもしれないが、依然として快楽説に不利な証拠を経験機械は提供しているということになるだろう。実際のところ、成田は「実在との接触contact with reality」を重要視するノージックの見方を批判した後で次のように述べている。

> だからといって、快楽主義が正しいとも思われない。快楽主義が正しいとすれば、人生のなかばで経験機械の世界に移住し、そこで以前にもまして快適に過ごすことができれば、そのような後半生は、本人にとって良い後半生になる。しかし、必ずしもそうは言えないように思う。それは、上で指摘したように、その移住には別れが伴うからである。(同、六二二頁)

しかし、もしほんとうに、デ・ブリガードが言うような理由で、人は経験機械を拒んでいるのならば、実のところ、それは快楽説の支持者にとって朗報となり得る事態である。そう言える理由を説明するため、次節ではまず、人が損失を拒み、現状維持を望む理由に関するトヴェルスキーとカーネマンの見方を紹介する。

第二節　損失回避と苦痛

人が損失を拒む理由に関して、トヴェルスキーとカーネマンは次のように述べている。

> 損失回避は非合理的だろうか。この問いは数多くの難しい規範的問題を生じるものである。意思決定をする者が

帰結に割り当てる価値を問題にすることは、選好の評価のための基準を必要とする。帰結の現実の経験がこのような基準を提供する。また、決定の文脈において、帰結に割り当てられる価値は、その帰結の経験の性質の予想として、正当化されうる。このような予想という観点を取れば、当初はリスク下での選択のパターンを説明するために作られた図1〔本稿第一節の図〕の価値関数は、快苦に関わる経験の心理的特性の予想として、解釈されうる。この価値関数は三つの基本的事実を正確に反映する。つまり、生物は定常〔すなわち、変化のない〕状態に慣れるということ、変化に対する反応は、限界に近づくと〔すなわち、変化が大きくなればなるほど〕、徐々に小さくなるということ、そして、苦痛は快楽以上に差し迫ったものだということである。苦痛と快楽のこの非対称性は選択における損失回避の究極的な正当化である。この非対称性ゆえに、帰結からの経験効用の最大化を図る意思決定者は、善い帰結以上に悪い帰結に重きを置くのが賢明である。(Tversky & Kahneman 1991, 1057)

この発言で重視すべきことは二つある。まず、人はある種の利得による快楽や損失による苦痛を予想し、それに基づいて、利得や損失を評価しているということである。つまり、人は何かを失うことで苦痛を被ると予想しているから、そうした損失を悪く評価する。もう一つは、「苦痛は快楽以上に差し迫ったもの」であるということ、つまり、苦痛の回避は快楽の獲得以上に重視されるということである。そして、損失回避は損失により被ると予想される苦痛の回避を重視する態度の表れということになるだろう。

デ・ブリガードの主張によれば、人はこれまでの生活で得てきたものを失いたくないから、経験機械につながれたくないと考えている。さらに、トヴェルスキーとカーネマンの言っていることが正しければ、こうした人たちは、自分が得てきたものの損失によって苦痛を被ると予期し、そうした苦痛の回避を快楽機械で感じる快楽の獲得以上に重

視しているということになるだろう。つまり、経験機械という思考実験は、ある種の苦痛の回避と快楽の獲得のどちらを選ぶのかを人々に迫るものとなるし、それにつながれたいと思わない多くの人は苦痛の回避の方を選んでいるということになる。

実のところ、快楽説はそうした人々の態度を理にかなったものと評価できる考え方である。そう言える理由を次節で論じる。

第三節　苦痛の回避と快楽の獲得

快楽説において、「苦痛がない」ということは、快楽と同様に望ましいものと考えられている。たとえば、快楽説の支持者とされるジョン・スチュアート・ミルは、自身が擁護するのは「快楽、および、苦痛からの自由が目的として望ましい唯一のものである」という立場であると説明している（Mill 1861, 210／二六五頁）。しかし、そうした立場において問題となるのは、「何らかの苦痛を被ることなしに、ある快楽を獲得できない場合に、苦痛の回避と快楽の獲得のどちらが望ましいのか」ということである。というのも、快楽と苦痛の不在を同等に望ましいと考えるのであれば、この二つの選択肢に優劣をつけることができなくなってしまうからである。

この問題に対し、快楽説はどう対応するのか。ここではミルの提案するやり方を紹介する。彼は『功利主義』の中で、二つの快楽のうち、より価値があるのはどちらかという問題について次のように述べている。

> 二つの快楽のうち、より享受する価値があるのはどちらか、または、二つの生き方のうち、その道徳的属性や帰結は別にして、感情にとってより好ましいのはどちらかという問題に関しては、両方の知識によって、資格を与

えられた人々か、彼らの意見が相違する場合、その大多数の判断が最終的と認められなければならない。(ibid., 213／二七〇頁)

すなわち、AとBという二つの快楽のどちらが望ましいのかという問題に関しては、AとB両方を経験し、よく知っている人たちの判断を、彼らの間で意見が分かれるのならば、その多数派の判断を、最終的な回答としなければならないということである。つまり、双方をよく知る全員、ないしはその大多数が、「Aの方が望ましい」と判断するのならば、事実Aの方が望ましいということになる。

ここで論じられるのは、快楽同士の比較に関してであるが、快楽の獲得と苦痛の回避のどちらが望ましいのかということについても同じことが言えるとミルは考えているように思われる。それは次の発言から明らかである。

苦痛が皆同質というわけではないし、快楽もそうである。また、苦痛は常に快楽とは異質のものである。特定の苦痛という代償を払って、ある特定の快楽を追求する価値があるか否かを決定するものとして、経験ある人の感情や判断以外に何があるのか。(ibid., 213／二七〇～二七一頁)

つまり、二つの快楽同士の比較の場合と同様、何らかの快楽の獲得とある苦痛の回避のどちらが望ましいのかという問題についても、経験ある人の感情や判断によって、最終的な決着はつけられる。すなわち、Aという苦痛を被ることなく、Bという快楽を獲得できない場合に、Aの回避とBの獲得のどちらが望ましいのかという問題に関しても、両方を経験した人たち(または、その大多数)の判断が最終的な回答となるということである。

この判定方法を経験機械の話に当てはめると、どうなるのか。まず、第一節で取り上げたデ・ブリガード、さらに、前節で言及したトヴェルスキーとカーネマンの主張が正しければ、経験機械につながれたくないと考える人は、「これまでの生活で得たものをすべて失うことによる苦痛」の回避と「経験機械で得られる（諸々の）快楽」の獲得を比較し、苦痛の回避の方が望ましいと判断しているということになる。もっとも、この場合において、問題となっている快苦を経験した人は誰もいないから、これは両方を経験した人たちの判断ではない。しかし、全く同じではないにしても、類似の経験を基にした判断とは言えるだろう。というのも、人々はこれまで経験した快楽をもとに、「経験機械で得られる（諸々の）快楽」がどんなものかを想像し、また、実際何かを失ったときの苦痛をもとに、「これまでの生活で得たものをすべて失うことによる苦痛」を想像しているものと思われるからである。そして、もし、快楽の獲得と苦痛の回避のどちらが望ましいのかという問題に関し、双方を経験した人たち（または、その大多数）の判断を最終的なものと認めるのであれば、この経験機械という思考実験においても、そうしなければならなくなるだろう。

つまり、ミルが提案するやり方を使うと、大多数の人が「これまでの生活で得たものをすべて失うことによる苦痛」の回避を「経験機械で得られる（諸々の）快楽」の獲得以上に望ましいと判断しているのなら、そちらがほんとうに望ましいということになる。そうすると、人々がそうした判断をもとに経験機械につながれたくないと思っているのであれば、快楽説はその思いを、より望ましい選択の現れと評価するということになるだろう。こうした点で、快楽説は経験機械につながれたくないという多くの人々の判断を、より望ましい選択肢を選んでいるという点で、理にかなったものと肯定できる。となると、快楽説は多くの人にとって歓迎すべき見方となるだろう。

結論

以上の議論が正しければ、経験機械という思考実験だけで、快楽説を否定することはできなくなる。それどころか、快楽説は多くの人が示す経験機械への拒否反応を理にかなったものと評価できる考え方であるという可能性すらある。これは快楽説の批判者にとっては好ましくない事態だろう。

そうした事態を避けるにはどうすればよいだろうか。本稿の議論に即せば、次の三つの対策が考えられる。

①まず、第一節で紹介したデ・ブリガードの主張は誤りであり、経験機械につながれたくないという私たちの思いは、損失回避や現状維持バイアスとは別の要因に基づくということを示すという手が考えられる。

②あるいは、彼の主張を正しいと受け入れた上で、第二節で取り上げたトヴェルスキーとカーネマンの見方を否定し、我々が損失を拒む理由を苦痛以外の観点から説明するというやり方もあり得る（これに成功すれば、経験機械はある種の快楽の獲得と苦痛の回避のどちらを選ぶのか迫るものではなくなるだろう）。

③さらに、デ・ブリガードやトヴェルスキーとカーネマンの主張を妥当と認めつつも、快楽説の枠内で、「これまでの生活で得たものすべてを失うことによる苦痛」の回避を「経験機械で得られる（諸々の）快楽」の獲得以上に望ましいものと言うことはできないということの論証を試みるという策も考えられる。

経験機械という思考実験を使って、快楽説を否定するには、それだけではなく、上の対策のいずれかも使わなければならない。つまり、経験機械はそれ単体で快楽説をやっつけられるほど強くはない（あるいは、快楽説はそこまで弱くない）のである。

註

[1] 以後、諸文献からの引用・参照は、（著者 刊行年、該当頁）という形式で表記し、翻訳のあるものについては、その該当頁数のみを付記する。ただし、訳はすべて拙訳による。また、引用文中の〔　〕は筆者による補足を表す。

文献表

De Brigard, Felipe. 2010. 'If you like it, does it matter if it's real?' *Philosophical Psychology* 23: 43-57.

Kymlicka, Will. 2002. *Contemporary Political Philosophy*, 2nd ed. Oxford: Oxford University Press.［W・キムリッカ（千葉・岡崎他訳）『新版 現代政治理論』、日本経済評論社、二〇〇五年。］

Mill, John Stuart. 1861. *Utilitarianism*. In J. M. Robson, ed., *The Collected Works of John Stuart Mill*, Toronto: University of Toronto Press, London: Routledge and Kegan Paul, vol. 10, pp. 75-115, 1985.［J・S・ミル（川名雄一郎・山本圭一郎訳）『功利主義』『功利主義論集』京都大学学術出版会、二五五～三五四頁、二〇一〇年。］

Nozick, Robert. 1974. *Anarchy, State, and Utopia*. Oxford: Basil Blackwell.［R・ノージック（嶋津格訳）『アナーキー・国家・ユートピア』木鐸社、一九九二年。］

Tversky, Amons. & Kahneman, Daniel. 1991. 'Loss Aversion in Riskless Choice: A Reference-Dependant Model.' *The Quarterly Journal of Economics* 106: 1039-1061.

児玉聡 二〇一二『功利主義入門――はじめての倫理学』筑摩書房。

成田和信 二〇〇七「夢と経験機械と幸福」『慶応大学商学部創立五十周年記念日吉論文集』、六一三～六二三頁。

（よねはら　まさる　静岡大学教育学部准教授）