

## 二重過程理論—進化的に新しいシステムは古いシステムからの出力を修正しているのか?

### The Dual Process Theory: Does the Evolutionarily Current System Revise Outputs of the Evolutionarily Old System?

山 祐嗣

Hiroshi Yama

大阪市立大学

Osaka City University

yama.hiroshi1204@gmail.com

#### Abstract

The dual process theorists have assumed the evolutionarily old system and the evolutionarily current system. The former system shows evolutionary rationality, whereas the latter system embodies normative rationality. There are three views on the relationship between the two systems; if the old system revise outputs of the current system; (1) The current system revise the outputs of the old system, (2) outputs from the two systems are independently exist, and (3) the new system just rationalizes the outputs of the old system. I propose that the analysis of the historical record of human decision making can potentially be beneficial not only for the solution of this problem and the theoretical development of the dual process theory, but for the progress of human society.

**Keywords — Dual Process Theory, Rationality, Historical Analysis**

#### 1. はじめに

本稿では、進化的に古い認知システムと新しい認知システムという分類を提唱する二重過程理論に基づいて進化的合理性と規範的合理性を想定する。その上で、二重過程理論に内在する「進化的に新しいシステムが古いシステムからの出力に対して、修正・上書きを行っているのか」という問題を検討する。この可能性は、3通り考えられ、第一は、新しいシステムによって修正がなされているとするもの(古いシステムへの謀反)、第二は、古いシステムと新しいシステムからの出力が共存している可能性、そして第三は、新しいシステムは古いシステムからの出力を合理化しているに過ぎない(従僕)とするものである。ここでは、この問題を、心理学実験というよりは、歴史的な変化によって検討する。歴史的には、戦争だけではなく、殺人や残虐行為が減少している。また、この50年をみると、差別や偏見等も抑制されるようになった。このような社会的変化の分析が、この問題を検討する方法として有望である。

#### 2. 二重過程理論

二重過程理論は、すぐれた論理学、数学、数学、経済システムなどを作り上げた人間が、なぜ推論課題において非合理的なバイアスに容易に陥りやすいのかを説明するために提唱された(Evans & Over, 1996)。当初の分類は、それぞれの過程の特徴をまとめて、潜在的で自動的な過程と顕在的で制御的な過程という対比であった。ここでは、前者を進化的に古いシステム、後者を進化的に新しいシステムとして区別する。進化的に古いか新しいかは、たとえば、古いシステムとして脳幹や小脳などの大脳辺縁系、新しいシステムとして大脳新皮質という区分があるが、ここで重視するのは心理学的な構成概念で、重要な点は、情報処理に大きな容量が必要かどうかという点である。この容量は、ワーキングメモリの認知容量とほぼ同一視されており、自動的な情報処理では少なくても済むが、制御的な情報処理においては大きな容量が必要である。大脳新皮質は、大きな容量の必要条件だが、十分条件ではない。

推論課題などのバイアス現象から提唱されたこの理論は、同じような自動的と制御的という区分が発見された記憶領域や社会心理学の知見を取り込んで、人間の認識全体についての理論として発展している。Evans & Stanovich (2013)による表1に、両システムおよびそれぞれのシステムの処理の特徴を対比したものである。進化的に新しいシステムにおける処理は、ワーキングメモリの容量を費やすという点が、定義的特徴となる。

進化的に古いシステムは、領域固有的なモジュールの集まりと考えられ、新しいシステムは、それらのモジュールの中央実行系のようなものと想定されている。モジュールとは、精神というシステムにおける機能単位部品を表す概念として用いられ、特定の刺激(領域固有的)にのみ反応して出力までが自動的である。また、カプセル化されているので、少なくとも出力までは意

識的な修正はできない。進化的に古いシステムはモジュールの束とされ、現在、顔モジュール、素朴生物学モジュール、素朴物理学モジュール、社会的交換モジュール、心の理論モジュール、危機管理モジュールなどが想定されている。

一方、進化的に新しいシステムは、これらのモジュールを、大きな認知的容量に支えられた中央実行系が制御していると想定されている。この中央実行系は、汎用性があるとされる。さらに、Stanovich (2009)によれば、このシステムは、認知容量を備えた計算の実行を担当し、知能指数に表わされる能力を支えているアルゴリズム的精神(algorithmic mind)とそれらの実行を監視している内省的(reflective mind)に分類される。

表 1 二重過程についての分類と各々の諸特徴

タイプ1 過程(直観的)	タイプ2 過程(内省的)
定義的特徴	
ワーキングメモリは不要	ワーキングメモリは必要
典型的関連項目	
速い	遅い
高容量	容量に限界
並行的	系列的
無意識的	意識的
バイアスがある反応	規範的反応
文脈的	抽象的
自動的	制御的
連想的	ルール基盤的
経験則的意思決定	帰結主義的意思決定
認知的能力とは独立	認知的能力と相関
進化的に古いシステム	進化的に新しいシステム
古い進化	新しい進化
動物の認知に類似	人類独自
潜在的知識	顕在的知識
基本的感情	複雑な感情

### 3. 進化的合理性と規範的合理性

二重過程理論の、それぞれのシステムは種類が異なる合理性を示すとして、進化的に古いシステムでは進化的合理性を、進化的に新しいシステムでは規範的合理性を発揮していると考えられている。概して、進化的に古いシステムは、行動が適応的か否かで自然選択されている。したがって、このシステムの目標は、遺

伝子の乗り物である生体がどのように行動すべきかを処方することである。1000 万年単位の人間の脳の進化の中で、文明と呼ばれるものが発展した時期はわずかであり、その大半の環境は野性的であって、行われてきたことは狩猟採集である。そうすると、人間の脳が最も適応的な環境は文明社会ではなく、野生環境ということになる。また、現代のホモ・サピエンス以前は、認知容量も決して大きくはない。進化的合理性はそのような状況で示され、それがバイアスを生むとされてきた。それぞれのモジュールは、固定的という点で現代社会では不適応な場合があるが、それが進化した環境では適応的であるということで、進化的合理性を示す

一方、進化的に新しいシステムは、最適な行動のために世界を正確に記述することを目標としている。ここから必要なことは規範的合理性であり、これは、論理学や確率論などの規範に対して合理的、すなわち合致しているということを表わしている。論理学や確率論が考案されたのは、進化の歴史の中ではるかに新しいという点で、進化的に古いシステムはこれらに対応できていないと考えられている。しかし、これらを考案したのも人間で、少なくともこれらの規範を創り出すような能力を発揮できるのは、進化を遂げた人間の脳あるいは精神であるといえるのである。つまり、規範自体を考案したり、その規範を守ったりすることは、認知容量が大きくなったことによって可能になったといえる。

これまでは、少なくとも推論領域においては、命題論理学や述語論理学が規範であった。しかし、推論において用いられる命題等が、100%の確率で真であったり偽であったりすることは現実にはなく、また、条件文において前提が偽である場合に、条件文自体がすべて真となるという命題論理学の取り決めに対する疑義も生じ、確率論と不確実性を規範として取り入れたのが、新パラダイムである。その結果、推論を意思決定の枠組みで議論できるようになった。しかし、伝統的に意思決定の規範とされていた、期待効用最大化原理が果たして規範足りうるのかという問題がとり残されているままである。とくに、効用が本質的に主観性をもっているという点が、規範として問題なのである。

ただし、この合理性についての二分法も非常に曖昧である。二重過程理論も、かつての「氏か育ちか」や「遺伝か環境か」、「生物学的か文化的か」のような二分法をそのまま用いているわけではない。とくに文化

はなんからの適応課題の解決のためにつくられ、そしてその文化がさらに新しい適応課題を人間に押し付けてくるというサイクルで形成され、ヒトの認知機構もそのように形成され、進化してきた。そして、ヒトの脳は社会的哺乳動物として進化してきたという特徴をもつ。そうしたことから、たとえば、周囲の顔色を伺いながら自分の欲求を抑制するような状況を、進化的に古いシステムの欲求を進化的に新しいシステムで監視するという構図は、現在では採用されていない。周囲の視線を気にして集団における調和を保つという合理性も、社会性哺乳類としての合理性である。他者への気遣いは心の理論モジュールの機能であり、もし自らの欲求が社会的交換における騙し行為であるならば、社会的交換モジュールがそれを抑制できるのである。

#### 4. 謀反、共存、あるいは従僕?

推論研究者が考えた初期の二重過程理論では、進化的に古いシステムからの出力を、新しいシステムが修正するということが想定されていた。確かに、もし新しいシステムがその認知容量を駆使して、汎用性がある機能を発揮すればそれも可能かもしれない。それによって、進化的に古いシステムからの認知バイアスを新しいシステムによって修正することが可能である(e.g., Stanovich & West, 1998)。しかし、その機能は、場合によっては、遺伝子に対しての謀反となる可能性もある。概していえることは、進化的に新しいシステムは、短期的な目標よりは長期的な目標を扱うのにすぐれており、短期的な目標を犠牲にすることは、遺伝子的な欲求を満たさないということも意味している。また、遺伝子とは独立して、あたかも自由意志があるかのように、自殺をそそのかしたりして(自殺が、期待効用最大化原理に照らして、規範的合理性があるかどうかはわからないが)、遺伝子の利益にならない行為を導くことがある。

一方、古いシステムは、強い感情を伴っていることが多い。たとえば、社会的交換モジュールの違反者に対する怒り、危機管理モジュールの恐怖が代表的なものである。さらに、心の理論モジュールからは、対象者からの悲しみや喜びに共感する強い感情が生まれる。このような場合には、新しいシステムによる修正や上書きは困難で、たとえば、古いシステムと新しいシステムからの出力は、それぞれ別個に作用するという見解に結びつく。経験的に、口の中の唾液は飲み込むことができるが、コップにためた自分の唾液を飲み込む

にはかなり抵抗を感ずる。進化的に新しいシステムは、コップの唾液も口の中の唾液も同じだと判断するのだが、進化的に古いシステムは、おそらく危機管理モジュールからの出力であろうが、その汚さへの恐怖をもたらすのである。また、たとえば「仏滅の日に結婚式をあげると不吉である」というような迷信についても、進化的に新しいシステムはそれに何の根拠もないという判断を下すのだが、古いシステムは不吉を信じて不安になってしまう。さらに、IAT (implicit association test) などによって、潜在的態度を測定している社会心理学者らは、偏見や差別等について意識的な修正が困難なことを主張している(たとえば、Banaji & Greenwald, 2013)。IATで測定されているものは、質問紙のような顕在的な態度を測定するテストで測定されたものとはあまり結びつきはなく、かつ、教育などによっても変容しにくいのである。つまり、進化的に新しいシステムが進化的に古いシステムからの出力を修正できないのである。

修正の対極にあるのが、進化的に新しいシステムは古いシステムからの出力の合理化を行っているに過ぎないという主張である。たとえば、Haidt (2007)は、道徳判断の領域においては、まず古いシステムからの直観的な好悪判断があって、新しいシステムはそれを合理化するに留まっていると主張している。

#### 5. 100万人の死より1人の死

進化的に古いシステムからの出力には、強い感情が伴うことが多いが、そのような感情は修正したり抑制したりされにくい。感情をかき立てられるものの1つに、心の理論モジュールあるいはマインドリーディングがある。これは心の中に感情を伴ったストーリーを作りやすく、人々を操作する力が強い。「1人の人間の死は悲劇だが、100万人の死は統計である」という有名な言葉がある。これは、ソビエト連邦の独裁者であるJ. Stalinの言葉とされているが、実際は、『西部戦線異状なし』で知られている作家のE. P. Remark、あるいはホロコーストにかかわったとされてイスラエルで処刑されたA. O. Eichmannの言葉ともいわれている。これは、1人の死では必ずなんらかのストーリーを伴い、心の理論モジュールからの出力である悲しみが喚起されるが、戦争などの100万人の死については、それぞれの人のストーリーを描くことは不可能で、統計として処理されるということとして解釈できる。

ストーリーとして人々に提示されると、より大きな

影響があるということは、ヒロインの少女が白血病で亡くなるという筋書きの「世界の中心で、愛を叫ぶ」によって、2005年から、日本骨髄バンクへのドナー登録者数が数年間飛躍的に増加していることがわかる。骨髄移植は、白血病や悪性リンパ腫などの血液の病気に有効な方法だが、そのためには、ドナーがいなければならない。そのため、政府広報などで、日本骨髄バンクへのドナー登録の呼びかけがあるが、あまり効果があるとはいえなかった。政府広報などのメディアによる説得が、進化的に新しいシステムに働きかけたと思われるのに対し、映画やドラマは古いシステムに働きかけて、心の理論モジュールを喚起させたのだと推定できる。

同じような現象が、人気歌手の **Kylie Minogue** の名前をとってカイルー効果 (**Kylie effect**) と呼ばれている。彼女は、2005年に乳がんが診断されて手術を受けて回復した際に、がん撲滅の啓もう活動などを行ったのだが、彼女の影響で多くの若い女性が乳がん検査を受けた。多くの女性は、頭では、つまりおそらく進化的に新しいシステムでは乳がんが怖いと理解しているが、実際に予防的な行動に移るには、心の理論モジュールによる **Kylie Minogue** という患者への共感の力が大きかったといえるのではないだろうか。

この区別は、トロッコジレンマにおける効用論者と義務論者の対比とも似ている。すなわち、5人を救うために1人を犠牲にしてもよいかどうかというジレンマにおいて、連結器でトロッコの進路を変更して1人を犠牲にするのはかまわないが、1人を歩道橋から突き落としてトロッコを止めて5人を救うという方法は選好されにくい。これを説明するのに、前者では功利主義的な効用論的判断なされやすいが、後者では「人を直接手にかけて殺してはいけない」という義務論的な判断が行われやすいと想定されている。この義務論的判断は、心の理論と関係していると推定され、直接手を下すことに嫌悪を感じると推定される。

## 6. 社会的意思決定—歴史的变化の分析

二重過程理論における新しいシステムが古いシステムからの出力を修正するのか否かという問題は、歴史的経緯を分析して議論すると興味深い。現代は、歴史的に稀有な豊かな時代である。とくに経済活動については、社会的交換によって生産の専門化と効率化が促進され、さらに、直接的な互惠性から間接的な互惠性へと発展してきた。間接的な互惠性は向社会的行動を

促進し、これがさまざまな制度(たとえば宗教や道徳、教育)によるものなのか、向社会的行動を互いにとるような集団において包括的適応度が高まったことによるものなのか議論されている。

歴史的経緯を見ると、進化的に古いシステムからの出力が修正される例は、論理学や科学の発展である。現代の科学的思考が、素朴物理学モジュールや素朴生物学モジュールとどの程度連続性があるのかどうか議論されているが、科学においては知識の蓄積によって、これら素朴モジュールからの出力への明らかな上書きがうかがわれる (**Carruthers, 2002**)。この蓄積された知識、たとえば論理学の知識は、習得されれば規範に一致していない出力を修正してくれる。

一般に、進化的に古いシステムからの出力で、強い感情を伴うものは修正されにくいとされる。しかし、そのような出力のうち、否定的な側面のもの、たとえば戦争、暴力、差別等は実際には減少している。**Pinker (2011)**の分析によれば、歴史的に、拷問や残虐な刑罰、殺人などは確実に減少している。たとえば、18世紀から19世紀にかけての、ヨーロッパでは、戦乱が縮小化し、拷問や魔女狩り、異教徒への弾圧などが減少した。また、16~17世紀には公開で行われて人々の娯楽という側面を持っていた死刑が、その後、恐怖と嫌悪の対象になっている。その時点までは、異教徒や魔女、犯罪人は危機管理モジュールを機動させる刺激であった。彼によれば、この背景には、王権の強化(警察権力の強化)と人道主義革命があると推定されている。王権は、封建的領主間の勢力争いを沈静化させるリヴァイアサンとして機能した。**I. Kant**に代表される人道主義は、内省に基礎を置く弁証法的方法から理性を導こうとする考えに基づいており、二重過程理論の用語を用いれば、進化的に新しいシステムの内省的機能の結果である。もう一つ重要な点は、18世紀に、活版印刷による書籍の生産が急上昇し、小説が人々の間に普及したことである。小説を読むことは、自分自身とは異なる他者が、どのように考え、どのように感ずるかを想像することを喚起する。すなわち、進化的に古いシステムの心の理論モジュールを刺激し、他者への共感を高める。この共感が高まれば、拷問や惨殺される側の恐怖や苦しみを自分のものとして想像することが可能になり、それらに対する拒否感が生じてきたといえる。いいかえれば、小説は、そこに登場する人物の死を、統計から悲劇へと変えてくれるのである。このようにして、危機管理モジュールからの、罪人や異端に対す

る恐怖や嫌悪という出力が、協同という適応課題を解決するのに有利だった心の理論モジュールからの共感という出力と競合したり、進化的に新しいシステムでの人道主義思想から上書きされたりして、減少したと推定できる。

1960年以降の変化は、第二次世界大戦のような大きな戦争の終結と、殺人を含む犯罪の世界的な減少に特徴づけられる。さらに今世紀に入ってから、ジェノサイド(たとえば、1970年代後半のPol Potによるカンボジアでの虐殺や、1994年のルワンダでの虐殺、1995年のボスニア内戦でのスレブレニツァの虐殺)が消滅している。そして、それと並行するように進んだのが、知能指数の増大と人権意識の高まりである。知能指数は認知容量と結びついていて、変動しにくいと考えられていたが、実は、知能テストの素点平均自体がかなり上昇している。この上昇を最初に指摘したFlynn(2012)の名をとって、フリン効果と呼ばれている。この原因ははっきりと明らかになっているわけではないが、親による教育の意識の変化、子どもの労働からの解放、公教育や高等教育の普及は、この50年で大きく向上していて、これらの要因が知能指数を押し上げているのではないかと推定されている。とくに、米国において、知能指数が低いとされたアフリカ系米国人の生活や教育環境が改善されるにつれて、彼らの知能指数が劇的に上昇している。2節で述べたように、知能指数は、二重過程理論では、進化的に新しいシステムの認知容量と直結している(Stanovich, 2009)。したがって、このような時流の変化は、認知的に新しいシステムが、古いシステムからの出力をより抑制できるようになった結果とも解釈できるのである。

また、この50年間の変化として、人権意識の高まりも無視することはできない。1964年に米国において、人種差別を禁ずる公民権法(Civil Rights Act)が制定された時でも、まさかその50年以内にアフリカ系から大統領が選ばれるとは誰も予想していなかった。また、オーストラリアにおいても、ヨーロッパ系の最優先主義とそれに基づくアボリジニを含む非ヨーロッパ系への排除政策である白豪主義の放棄があった。1973年の移民法、1975年の人種差別禁止法の制定によって、原則的に移住手続きや、移民の国内での生活・教育・雇用における人種差別が禁止されるようになった。また、この50年の間に、フェミニスト運動の高まり、動物への虐待の禁止、同性婚への許容などの、弱者とされた側の権利を尊重するという潮流が並行して続いていて、

オバマ大統領の誕生も、その潮流の中で位置づけられよう。

日本においても、犯罪や殺人の減少は同じ傾向が示されている。また女性の地位の向上について、たとえば女子の四年制大学進学率は、2016年では男子が56%、女子が48%であるが、1970年代初頭では男子が約35%に対して、女子は10%程度(短期大学への進学率を合算しても25%)であった。女子には学問は不要という文化背景があり、また、女性への偏見や蔑視も共有された時代である。たとえば、連合赤軍事件で死刑判決を受けた永田洋子(獄中死)の1982年の判決文では、「女性特有の執拗さ、底意地の悪さ、冷酷な加虐趣味」という文言が登場している。「女性特有の」という点について、もちろん当時も批判されたが、このような公的文書においてこのような文言が用いられるという時代背景は、現代とは異なったものであろう。

4節において、迷信等の恐怖は進化的に新しいシステムによって上書きされにくいと記したが、歴史的にみると迷信は減少している。たとえば、内山(2007)が指摘していることだが、日本において、1965年を境に、キツネに騙されるという話が途絶えるのである。内山は、自然と人間の関係性の変化や宗教観の変化などがあるのではないかと考察しているが、この現象も、進化的に新しいシステムが何らかの方法で、「キツネに騙される」という共同妄想を消し去ってしまったと解釈できる。

## 7. 結論

進化的に古いシステムの出力を新しいシステムは修正できるのか。概して、古いシステムからの出力に、強い感情が伴う場合には修正されにくいという傾向がある。しかし、前節での事例から、進化的に古いシステムからのネガティブな出力も歴史的には修正されつつあるということが伺われる。ただしそれは、修正する側の、同じく進化的に古いシステムである心の理論モジュールの出力による共感など、比較的強い情動を伴った作用によって抑制されやすいと考えられる。つまり、進化的に新しいシステムによる直接的な修正というよりは、拮抗する進化的に古いシステムからの力を借りて、啓蒙活動、高等教育の普及による進化的に新しいシステムの、アルゴリズム的精神や内省的精神によって修正が行われたと推定できる。

それでは、このような進化的に古いシステムからのネガティブな出力は、長期的には抑制されるようにな

るのだろうか。それについてはまだ結論はない。その理由は、攻撃性や差別感情など、行動として抑制はされても、根強く潜伏している可能性があるからである。4節の表現を用いれば、「共存」という状態である。根強く残った進化的に古いシステムからの感情は、抑制されたまま別の形をとる場合もある。それが、たとえばインターネットのような匿名性や隠匿性が高いような状況で出現する暴言であったり、執拗な嫌がらせであったりしている可能性がある。また、社会レベルでは、ヨーロッパにおける極右政党の台頭であったり、米国におけるトランプ現象であったり、また、日本におけるヘイトスピーチであったりするのかもしれない。

まだ明確な方法は確立されていないが、このような歴史記録的な方法から進化的に新しいシステムの働きを推定していく手法は、単に、二重過程理論における合理性の問題ではなく、現代とはどのような時代なのかという現代論に言及し、より快適な社会をどのように作っていくのかという問題にも示唆を与える有望な手法である。

## 参考文献

- [1] Banaji, M. R. & Greenwald, A. G. (2013). *Blindspot*. New York: Delacorte Press. (北村英哉・小林知博 訳 (2015). 『心の中のブラインド・スポット—善良な人々に潜む非意識のバイアス』. 北大路書房)
- [2] Carruthers, P. (2002). The roots of scientific reasoning: Infancy, modularity and the art of tracking. In P. Carruthers, S. Stich, & M. Siegal (eds.), *The cognitive basis of science*. Cambridge, UK: Cambridge University Press. Pp. 73-98.
- [3] Evans, J. St. B. T., & Over, D. E. (1996). *Rationality and reasoning*. Hove, UK: Psychology Press. (山祐嗣 訳)(2000) 『合理性と推理』. ナカニシヤ出版)
- [4] Evans, J. St. B. T., & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Sciences*, 8, 223-241.
- [5] Flynn, J. R. (2012) *Are we getting smarter?: Rising IQ in the twenty-first century*. New York: Cambridge University Press. (水田賢政 訳 (2015). 『なぜ人類のIQは上がり続けているのか?—人種、性別、老化と知能指数』. 大田出版)
- [6] Haidt, J. (2007). The new synthesis of moral psychology. *Science*, 316, 998-1002.
- [7] Haidt, J. (2012). *The righteous mind: Why good people are divided by politics and religion*. New York: Pantheon. (高橋洋 訳(2014). 『社会はなぜ左と右にわかれるのか—対立を超えるための道徳心理学』. 紀伊國屋書店)
- [8] Pinker, S. (2011) *The better angels of our nature: Why violence has declined*. New York: Viking. (幾島幸子・塩原通緒 訳(2015). 『暴力の人類史』. 青土社)
- [9] Stanovich, K. E. (2009) Distinguishing the reflective, algorithmic, and autonomous minds: Is it time for a tri-process theory? In J. St. B. T. Evans & K. Frankish (eds.), *In two minds: Dual processes and beyond*. Oxford: Oxford University Press, Pp. 55-88.
- [10] Stanovich, K. E., & West, R. F. (1998). Individual differences in rational thought. *Journal of Experimental Psychology: General*, 127, 161-188.
- [11] 内山節 (2007). 『日本人はなぜキツネにだまされなくなったのか』. 講談社現代新書