Japanese Psychological Review
2018, Vol. 61, No. 1, 42–54

心理学におけるモデリングの必要性

竹　澤　正　哲
北海道大学

On the necessity of rigorous modeling in psychology

Masanori TAKEZAWA
Hokkaido University

In the last decade, psychology faced a serious crisis, called the reproducibility problem. To solve the problem, several methodological and institutional changes have been proposed and implemented such as the promotion of replication studies and publication of negative results, the introduction of a preregistration system in academic journals, and the implementation of novel statistical methods for suppressing false-positive results. In this paper, based on a previously proposed formal model of population dynamics of scientific discovery, I first explain that these changes are insufficient for avoiding the reproducibility problem. Based on two examples in social psychology, I also discuss that what is necessary for fundamental reformulation of psychology is rigorous modeling, which has already been widely applied in many scientific fields outside psychology.

Key words: Bayesian modeling, formal model, metascience, reproducibility problem
キーワード：メタ科学，再現可能性問題，フォーマル・モデル，ベイジアン・モデリング

1. はじめに

数十年の後，あるいは数世紀の後，未来の科学史家が21世紀の心理学史について書物を著すならば，その中にはおそらく，次のような一文が記されていることだろう―「心という対象を自然科学の手法で研究する実験心理学は19世紀に誕生し，20世紀に飛躍的な発展を遂げた。だが21世紀に入ると，心理学者は再現可能性と呼ばれる問題に直面する」。

再現可能性問題とは，学術論文や書籍に掲載され，科学者コミュニティや一般社会から「真実である可能性が高い」とみなされている現象が，追試によって再現されないという問題である（詳細は友永・三浦・針生（2016）を始めとする心理学評論第59号1巻の特集号掲載論文を参照されたい）。特に2015年に*Science*誌に掲載された論文は大きな衝撃を与えた（Open Science Collaboration, 2015）。*Psychological Science, Journal of Personality and Social Psychology, Journal of Experimental Psychology: Learning, Memory, and Cognition*といった心理学の主要誌に2008年に掲載された論文の中から選びだされた98の実験について，複数のチームが追試したところ，36%しか実験結果が再現されなかったのである。

原理的には，再現可能性問題は，実証研究を行う全ての領域で発生しうる。だが，2016年に同じく*Science*誌に掲載された論文は，別の意味で衝撃をもたらした。実験経済学における著名な雑誌（*American Economic Review*, *Quarterly Journal of Economics*）に掲載された18の実験について，Open Science Collaboration（2015）と同様の複数チームによる追試が行われた結果，心理学の再現率が36%であったのに対し，実験経済学における再現率は61%だったのである（Camerer et al., 2016）[1)]。

「真実だろう」と受け止められていた現象の多くが，実際には存在しないらしいという事実。それは心理学の評判に関わる由々しき事態である。そのため，この問題が広く認知されるようになる

1)　初稿提出後，さらに衝撃的な論文が公刊された（Cova, Strickland, Abatista, et al., 2018）。実験哲学における40の実験を同様に追試したところ，再現率は70%を超える高さだったのである。

と，問題を解決するための処方について，心理学者の間で活発な議論が巻き起こった。そして実際に心理学における研究慣行に大きな変化が生じている。

本論文の目的は，心理学に対する信頼を取り戻すために必要な最後のピースとして，モデリングという科学的営為の重要性を紹介することにある。まず第2節では，再現可能性問題をメタ科学の観点から分析した研究を紹介し，現在心理学に導入されつつある処方だけでは，心理学を取り巻く危機からの脱却は困難であることを指摘する。続く第3節では，社会心理学における具体例を紹介しながら，実際にモデリングという営みが心理学に大きな変革をもたらしていることを指摘する。そして第4節では，モデリングという営みを導入することの重要性を議論する。

2. メタ科学からみる心理学の危機

2.1 再現可能性問題に対するこれまでの反応

再現できない現象が，確立された知見として公刊され流通してしまう原因は，大きく分けると2つあると言われている。まず第1に出版バイアス（publication bias）である。すでに論文に掲載された知見を追試しても，再現に失敗した場合には結果が公開されないならば，誤った知見が淘汰されずに流通し続ける。第2に心理学において広く見られる，問題のある研究慣行（questionable research practices; QRP）である。たとえば統計的検定の結果が有意になるまでサンプルサイズを増やしたり，仮説を定めずに多くの検定を行って有意になった結果だけを報告することである。Simmons, Nelson, and Simonsohn（2011）は，複数のQRPが組み合わさることで，偽陽性の確率が飛躍的に増大することを示した。

再現可能性問題に対する注目が集まると，ここで挙げた2つの要因を修正しようとする動きが出てきた。多くの学術誌が事前登録型追試論文の投稿を受け付けるようになると同時に，追試研究が積極的に公刊されるようになった。またQRPが広く知られるようになったことで，データセットや研究で使用した全てのマテリアルの公開，メタ分析の奨励など，分析において発生する偽陽性を検出・抑制する試みが広まっている。

QRPを取り除くことは，論文が公刊される前に偽陽性の発生を防ぎ，科学的知見として流通することを水際で防ぐ手段である。追試の奨励や出版バイアスの抑制は，いったん流通してしまった偽陽性である知見を，時間をかけてコミュニティから淘汰する役割を果たす。従って，これら2つの手段が浸透することで，心理学に対する信頼性は大きく回復するはずだと考えられている。

だが，それだけで充分なのだろうか？ この問題を考える上で，McElreath and Smaldino（2015）は深い示唆を与えてくれる。

2.2 科学知識の発見と流通のダイナミクス

一般的に科学の世界では，仮説構築，実験による仮説の検証，論文の出版というプロセスを通して，知識が流通していく。McElreath and Smaldino（2015）は，科学という世界における知識の発見と流通のプロセスを生物学における個体群動態（population dynamics）のモデルによって表現した。これは科学の手法を用いて，科学という世界の仕組みを分析するメタ科学（metascience）と呼ばれる研究の一種である。

仮説構築と実験　まず科学者が実験で検証したい仮説を選ぶと考えよう。仮説は真か偽かいずれかであるが，それは誰の目にも見えないものだと考えよう。実験が行われると「（仮説は）正しい」あるいは「誤り」という結果が得られる。ここでは統計的検定の論理に従い，「本当は真である仮説」を検証した時にネガティブな結果が得られる確率をβ（第二種の過誤あるいは偽陰性），「本当は偽である仮説」を検証した時にポジティブな結果が得られる確率をα（第一種の過誤あるいは偽陽性）とする。

科学者が実験を行う際，すでに公刊された論文に掲載された知見（＝仮説）を追試する場合と，科学者が独自に導出した仮説を検証する場合がある。既に公刊された仮説を実験対象として選ぶ確率をr，新たに仮説を導出して実験する確率を$1-r$としよう。また科学者が独自に導出した仮説が「真である」確率をbとする。QRPが蔓延している場合には第一種の過誤の発生確率αの値が高いことで表現され，また追試が奨励されないという状況はrの値が低いことで表現される。

科学的知見の流通　得られた実験結果は論文と

して公刊される。だが出版バイアスとして知られるように，全ての実験結果が公刊される訳ではない。モデルでは「論文として公刊される確率（C）」は，実験によってポジティブな結果が得られた場合（+）と，ネガティブな結果が得られた場合（–）では異なるとみなす。さらに実施された実験が追試である場合（R）と，科学者が独自に導出した新しい仮説を検証した場合（N）とでも，実験結果が公刊される確率が異なるとみなす。そのため「実験結果が公刊される確率」はC_{N+}，C_{N-}，C_{R+}，C_{R-}という，4つの異なるパラメータで表現された。たとえば「仮説と合致しなかった結果は公刊されない」という出版バイアスは，C_{N-}やC_{R-}の値が低いことで表現される。追試結果が公刊されにくい状況はC_{R+}やC_{R-}の値が低いことで表現される。

科学的知見のダイナミクス　科学とは，仮説構築，実験，出版というプロセスが再帰的に繰り返される営みである。従って，このモデルにおいて焦点となるのは，上述のプロセスが再帰的に繰り返された場合，長期的にどのような状態が生じるか，である。

時が経つと，ある知見（仮説）についてポジティブな実験結果とネガティブな実験結果の両方が論文として公刊され，混在する状態が生じるだろう。ここで単純化のために，ポジティブな実験結果の数が，ネガティブな実験結果の数よりも多い仮説を「科学者コミュニティにおいて正しいとみなされている仮説」と考えることにする。

ここで問題となるのは「正しいとみなされている仮説」が，実際に「真である」確率はどのくらいか，である。もしこの確率が低いならば，その科学の領域において正しい知見として流通している仮説は，実は偽である可能性が高いことを意味する。

彼らのモデルには外生的な要因によって値が決定されるパラメータが8つある。この中で，科学の信頼性に大きく影響しているのはどれなのだろうか？

こうしたモデルではパラメータに様々な値を代入して，どのような状態が生じるのか検討できる。彼らは，論文の中で2つのケースを検討している。楽観的なシナリオとして実験の検定力が高く（$1-\beta=0.8$），第一種の過誤の発生率が低く（$\alpha=0.05$），そして「新しく導入される仮説が真である確率（b）」が0.1であるケース。悲観的シナリオとして検定力が低く（$1-\beta=0.6$），第一種の過誤の発生率が高く（$\alpha=0.1$），そして新しい仮説が真である確率が非常に低い（$b=1/1000$）ケースである。なおいずれのシナリオでも出版バイアスは低く設定されていた（$C_{N+}=1$, $C_{N-}=C_{R+}=C_{R-}\fallingdotseq 0.9$）。

本稿において注目したいのは，前者の楽観的なケースである。つまり出版バイアスが抑制されており，ネガティブな実験結果もポジティブな実験結果と同程度に公刊され，またQRPが一掃されることで実験の検定力は高く，第一種の過誤の発生率は低いという理想的な状況である。彼らの分析によれば，この理想的状況においては，ある仮説に関して公刊された実験結果のうち，ポジティブな結果の数がネガティブな数を上回るならば，その仮説が「本当は真」である確率は非常に高く，1に近い値になる。これは，理想的な状況においては，科学という営為によって客観的な真実が見出されることを意味する。

正しい仮説の重要性　この理想的な状況において，もしある1つのパラメータの値が変化しただけで結果が大きく変化するならば，それは科学の信頼性を担保する上で重要な要因であることを意味する。McElreath and Smaldino（2015）の分析結果によれば，興味深いことに，8つのパラメータの中で，科学の信頼性に最も大きく影響するのはb，すなわち「科学者が新しく導入する仮説が真である確率」だったのである。そしてその次に影響を与えていたのは，第一種の過誤の発生確率（α）であった。

bが大きな影響を持つことについては，ベイズ推定における事前確率の影響として理解することができるだろう。ある病気に罹患していることを確かめる検査法が存在しているとしよう。たとえ検査の検定力（$1-\beta$）が高く第一種の過誤の発生確率（α）が低い優秀な検査法であっても，もともと病気に罹患している人の比率が非常に小さいならば，検査結果が偽陽性である確率は飛躍的に高まる。同様に，科学者が独自に導出した仮説の多くが，本当は「誤り」であるならば，実験におけるQRPが改善されても，論文として公刊された「正しい」とみなされる知見の多くが，実は

「偽」である可能性が高くなる。この問題は，再現可能性のコンテクストにおいて以前から指摘されており（Sterne & Smith, 2001; Ioannidis, 2005），心理学評論における再現可能性特集号において池田・平石（2016）が言及している点でもある[2)]。

McElreath and Smaldino（2015）が示したのは，さらにその先に生じる問題である。彼らの楽観的なシナリオは，出版バイアスは抑制されており，また公刊された実験結果が他の科学者によって追試される可能性も充分に高い状況である。だが，そのような状況においてすら，科学者が独自に導出する仮説が偽である確率が高いならば，偽陽性の知見が追試という網の目をくぐり抜けて，その後も流通し続ける可能性が示唆されたのである。

再現可能性問題をきっかけとして，2つの大きな改革が動き出した。第1に追試を奨励し，出版バイアスを抑制する動きである。第2に，実験の実施や分析のプロセスにおけるQRPを取り除くことで，偽陽性の発生率を低減しようとする動きである。いずれも重要な一歩である。だが，McElreath and Smaldino（2015）のモデルによれば，科学者が導出する新しい仮説が真である確率（b）が低いままでは，改革が進んでも，流通する科学的知見の妥当性は大きく毀損されてしまう。「本当は偽である仮説」が最初の実験によって「正しい」と判断され，論文として公刊されてしまったならば，たとえ追試が奨励されたとしても，取り除くことが困難なのである。

2.3 「弱い理論」と心理学

「科学者が新たに導出した仮説が，本当は真である確率」は，理論的には定義できても，実測は困難である。だが，少なくともこの問題について科学者の間には一定の直感的な合意があるだろう。すなわち，物理学のような厳密でフォーマルな理論がある場合には，現象を高い精度で予測できるためbの値は高くなるが，思いつきで当てずっぽうに仮説を導く場合には低い，と。たとえばMcElreath and Smaldino（2015）はbの値が非常に低い領域の一例として，ゲノムワイド関連解析（Genome-wide Associate Study; GWAS）を挙げている。これは，ゲノム全体をカバーする50万から100万近くに及ぶ一塩基多型と表現型（たとえば特定の疾患など）との相関を，総当りで調べる研究手法である。当然のことながら，しかるべき理論に基づいて数個の遺伝子に候補を絞った上で検定をするわけではなく，当てずっぽうに100万回の検定を行うことに等しいため，「本当は真である仮説を導出する確率」という観点からみれば，非常に低い値となるのは当然である。

翻って，心理学はどうだろうか？ 池田・平石（2016）は，Eysenck（1985）の「弱い理論」という概念に依拠し，すべからく心理学の理論は「弱い」ため，誤った仮説を導出してしまう危険性が高いと論じている。もっとも，両者の論考を読み進めても，「弱い理論」の定義が曖昧であるため，池田・平石（2016）の指摘が論理的に妥当であるか定かではない。であるにも関わらず，物理学における理論と比較すると，心理学の理論は弱いため，信頼性の低い仮説しか導出できないという指摘には，我々の直感に響くものがある。それでは，この直感の背後にあるものは，一体何なのだろうか？

まず指摘しておきたいのだが，心理学における理論は，物理学と同レベルの精度で量的な予測を可能とする理論ではない（Meehl, 1967）。だがその点に関しては，心理学に限らず社会科学から自然科学に至るまで，科学における多くの分野も同様であり，物理学と同程度に量的予測が可能なレベルに達しているわけではない。実際，Eysenck（1985）の議論においても，その点が問題視されているわけではない。

Eysenck（1985）は，「弱い理論」の特徴としてまず以下の特徴を挙げている。

> “強い理論とは，根拠が確かで，また実験によって立証された数多くの前提群Kにもとづいて構成されている。そのため，新たに得られた実験結果によって，仮説H_1，H_2，…，H_nのどれが支持されるのかを確実に決定することができる。弱い理論には，そのような確かな基盤がない。そのため，新しい実験結果によって，

2) 本稿冒頭で紹介したOpen Science Collaboration（2015）によれば，実は認知心理学の実験よりも社会心理学の実験の方が再現率が低いことが示されている（50% vs. 25%）。Wilson and Wixted（2018）は，この違いが，社会心理学において検証される仮説が，実は偽である確率が高いことに起因していることを，Open Science Collaboration（2015）の再分析を通して指摘している。

仮説Hが誤りであることを証明できるのと同じくらい簡単に，前提群Kが誤りだと証明することができる（p. 27; 筆者訳）”

たとえばここで，再現可能性問題においてしばしば槍玉に挙げられる老人プライミング実験（Bargh, Cheng, & Burrows, 1996）を考えてみよう。老人に関連する単語からなる文章を作ると，歩行速度が減少するという実験結果がある。この現象は，老人という概念が記憶の中で活性化されると，それと強く結びついた行動表象も活性化され，その結果，歩行という実際の行動に影響が及ぶのだと説明されてきた。

ここで，追試によってこの実験結果が再現されないことが確実になったとしよう。この新たな実験結果によって，何が反証されうるのだろうか？老人プライミングに関する説明を吟味すると，大きく分けて2つの要素から構成されている。まず，「ある概念の活性化は，それと強く結合した行動表象にまで波及し，最終的には行動に影響する」という観念運動原理（ideomotor principle）である。だがこの原理のみからは，老人プライミング実験の結果を演繹できない。「老人という概念は，遅い歩みという行動表象と強く結びついている」という前提（あるいは補助命題）が存在して初めて，「老人について考えると歩みが遅くなる」という現象が導かれるのである。したがって，実験結果が再現できないことが明らかになっても，それは観念運動原理という一般的な心理メカニズムが否定されたとも解釈できるし，「老人概念は動きの遅さという行動表象と結合している」という，現象の前提が否定されたと解釈することもできる。

このように，ある実験結果によって，科学者が問題とする仮説（理論）が反証されうるのか，それとも前提が反証されうるのか判然としないという理論構成は，少なくとも社会心理学においては非常にありふれたものである。この意味において，確かに社会心理学は「弱い理論」に満ちあふれている。

さらにEysenck（1985）は，次のように述べている。

“強い理論と対比させた時，弱い理論が持つもう一つの特徴について触れておくべきだろう。強い理論においては，理論の前提が相互に依存しあっている―ある前提を変更しようとすると，他の前提についても変更を加える必要が出てきてしまい，実際には理論全体を放棄しなければならなくなる。弱い理論の場合，そのような相互依存性はあまり見られない。そのため，理論の一部を容易に変更することができる。（p. 28; 筆者訳）”

ここで指摘されている，命題や前提における相互依存性の低さ，いわばフレキシビリティの高い理論構成は，社会心理学だけでなく，心理学全体においても，（領域によって濃淡はあれども）広く見いだすことができるだろう。そして，後述する通り，この問題は信頼性の高い仮説演繹を妨げるだけでなく，様々な問題を確かに引き起こしている。

2.4　正確さと厳密さを備えた理論を目指して

Eysenck（1985）も述べるように，ほぼ全ての科学理論は，最初は正確さや厳密さを犠牲にした，フレキシビリティが高い「弱い理論」から出発する。物理学に代表される「強い理論」へ到達するためには，長い時間をかけて変化し続けていくしかない。故に，現在の心理学が弱い理論しか持たないとしても，そのことによって心理学という領域の存在価値が失われるわけではない。真に評価されるべきは，心理学という科学者コミュニティが，今ある場所から先へ向かって動いているか否か，である。確かに，再現可能性問題をきっかけとして，QRPや出版バイアスを放逐し，追試を推奨する変化が生じた。だが，老人プライミングに見られるようなフレキシビリティの高い理論構成が放置され続ける限り，そして自然科学に見られる正確さと厳密さを備えた理論構築を目指して，心理学が変化しようとしない限り，再現可能性問題という危機から，心理学が本当の意味で脱却することはできないだろう。

それでは，相互依存性が低く，正確さと厳密さに欠いたフレキシブルな理論構成から先へ進むために，今，心理学に必要とされているアクションは何だろうか？心理学における理論には意味がないのだからそれらを捨て去り，物理学や経済学において構築された厳密な理論，たとえば力学や

ゲーム理論を導入して心理学を再構築することなのだろうか？

次節では，特に社会心理学を例として，この問に対する一つの回答を提示したい。それは，モデリングという科学的営為である。

3. 心理学におけるモデリング

3.1 モデルとは何か

科学哲学者ワイスバーグ（Weisberg, 2013）によれば，モデルとは，科学者が関心を持つ「対象の潜在的な表現（p. 270）」であり，モデリングとは，「モデルの構築や分析を通じて，現実世界を間接的に研究する手法（p. 5）」である。自然科学，そして経済学などの分野では，数理モデルを通じて現象を表現することが多い。だが数学的な表現を使うことはモデルの本質ではない。数学を一切使わずに，我々が日常的に用いる自然言語によって現象を表現するモデルもある。心理学における現象の記述は，そのほとんどが言語による現象のモデル化と言って良いだろう。

言語によるモデルの例として，ワイスバーグは，Shepard and Metzler（1971）による図形の心的回転課題の実験論文を取り上げている。ワイスバーグによれば，この有名な論文では，「参加者が頭の中で図形を三次元図形として思い描き，それを実際頭の中で回転させているのだ」と，自然言語によって簡潔に実験結果が説明されている。これも立派な言語によるモデルである。だが，これはモデルとしては不完全であると，ワイスバーグは指摘する。なぜならば，「シェパードとメッツラーのモデルを詳細に記述すれば，「視覚的な入力Vが心的機構Mを作動させ，それがPによって処理されてOを出力する」といったようなメカニズムの詳細が，多数含まれると思われる（p. 25）」からである。言語によるモデルでも現象を詳細かつ厳密に記述していけば，それは数学的なモデルとなんら変わることがないレベルに達する。だが，しばしば言語によって現象をモデル化しようとすると，本来ならば記述されるべき前提やプロセスが省略されてしまう。

細部を省略することの何が悪いのかと思われる読者も多いだろう。だが，言語によるモデルでは，前提やプロセスを省略することによって厳密さが失われ，その結果として誤った結論が導かれてしまうことがある。そして科学者がその誤りに気づかずに時が過ぎ去ることもある。

以下では，社会心理学における2つの事例を紹介したい。いずれも社会心理学における古典的な知見であるが，厳密なモデルという視点を導入することによって，社会心理学という科学者コミュニティが数十年に渡って見落としてきた問題が浮き彫りとなった事例である。

3.2 認知的不協和理論とモデリング

自由選択パラダイムと認知的不協和　自由選択パラダイム（Brehm, 1956）とは，以下のような手続きを用いた実験である。参加者はまず，音楽CDのようなアイテムを複数提示され，それぞれのアイテムに対する好意度をリカート尺度で評価する。次に実験者は，参加者が同じ評価を与えたアイテムを2つ選んで参加者に提示して，そのどちらか好きな方を謝礼として持ち帰ってもらうために自由に選択してもらう。そして最後に，もう一度全てのアイテムに対して参加者は好意度評定を行う。この手続きを用いた実験の多くで，最初は同じ評価をしていたにもかかわらず，自らが選択したアイテムの評価は，選択しなかったアイテムの評価よりも高くなる現象が見られる。

最初は2つのアイテムを同程度に選好していたのに，一方のアイテムを選択するという行動によって，アイテムに対する選好が変化したという現象である。そのため自由選択パラダイムにおけるこうした実験結果は，認知的不協和理論を支持するものだと考えられてきた。

自由選択パラダイムをモデル化する　さてここで，以上の議論をより厳密にモデル化してみよう。まずある個人がアイテム i に対して持つ選好の強さを u_i とする。この選好は個人の心理変数であり，直接観察できないものとしよう。続いて個人が好意度評定に回答する。自らの選好の強さを表現するよう求められた個人は，選好の強さ u_i と完全に一致する数値を質問紙において選択するものとしよう。さてここである参加者は，アイテム1と2の双方に対して，リカート尺度において5という好意度評定をしたとしよう。以上に定義したモデルから，この参加者がアイテム1と2に対して持つ選好の強さは $u_1=u_2=5$ となるはずで

ある。この時，参加者がどちらか一方を必ず選択しなければならず，何らかの理由にもとづいてアイテム 1 を選択したものとする。その後で再び，全てのアイテムに対してリカート尺度で好意度評定を求められた個人は，アイテム 1 に対して 6，アイテム 2 に対して 5 という評定を下したものとしよう。上述の通り，このモデルでは「好意度評価に際し，個人は，自らが持つ選好の強さ u_i と完全に一致する数値を選択する」と定義した。そのため，この実験結果は，個人のアイテム 1 に対する選好 u_1 は 5 から 6 へ増加したことを意味する。

迂遠に思われるかもしれないが，これは自由選択パラダイムにおける現象をワイスバーグが求める厳密さでモデル化したものである。そしてこのように詳細にモデルを構築すると，ここにはいくつかの奇妙な前提があることに気づかされる。最も奇妙なのは，好意度評定を求められた個人は「自らが持つ選好の強さ u_i と完全に一致する数値を選択する」という前提である。仮に，あるアイテムに対する選好の強さは時間を通じて一切変化しないとしても，そもそも選好は観察不可能な心理変数である。好意度評定を求められるたびに，人間が常に自らの心の中にある選好の強さを「5」という整数として知覚し，さらにリカート尺度において常に「5」という反応を選ぶことなどありうるのだろうか？

人間の行動には，必ずエラーが発生しうる。したがって，本当は 2 つのアイテムに対する選好の強さが同一であるとしても，好意度評定を求められるたびに，平均すると 5 を選択するものの，小さな確率で毎回異なる回答を選んでしまうという前提を置くべきではないだろうか？[3] 実は，この前提が妥当であるならば，認知記不協和理論が主張するような行動による選好の変化が起きなくても，上述の実験結果が生じうるのである。もし個人がアイテム 2 ではなくアイテム 1 を選択していたのならば，実際には個人がアイテム 1 に対してアイテム 2 より強い選好を持っていた可能性が高い（$u_1 > u_2$）。一方で，両者に対する選好の強さが近似しているならば，最初の好意度評定においてアイテム 1 とアイテム 2 に対して同じ評価を下す可能性が出てくる。なぜならば好意度評価において小さな確率でエラーによるぶれが生じるからである。だが，実際にはアイテム 2 よりもアイテム 1 を強く選好しているならば，2 回目の好意度評定では，高い確率でアイテム 1 に対してアイテム 2 より高い評価を下すことだろう。つまり，「好意度評価は選好に基づくが，ランダムエラーによって評価する度にぶれが生じる」という前提を置くならば，自由選択パラダイムにおける実験結果は，認知的不協和による選好変化がなくても生じうるのである。

なぜ厳密なモデルが必要なのか？　以上は Chen and Risen（2009, 2010）が数理モデルを用いて表現し，そして Izuma and Murayama（2013）が数値計算シミュレーションを用いて，数理モデルが不得手な科学者でも平易に理解できるように展開したロジックである。もちろん，ここで注目して頂きたいのは，Chen and Risen（2009, 2010）が指摘するまでの 50 年間に渡って，おそらく社会心理学者の誰も，自由選択パラダイムの背後に潜む前提に気付かなかったという点である。「2 回目の評定で変化が生じた」という実験結果によって，認知的不協和が生じたという仮説が支持されるためには，いくつもの暗黙の前提が必要とされる。それは，ここまでの議論を読んできた読者にとっては自明である。だがその自明の理は，厳密さに欠けた言語によるモデルを，数理モデルや数値計算シミュレーションという形で厳密に構築し直さない限り，浮き彫りにされることがなかったのである。そして次に注目して頂きたいのは，こうした厳密なモデル化によって初めて，実験結果が理論を検証し得ない可能性に科学者が気づき，それによって理論と実験の双方向で発展が生じる点である。いずれも，言語によるモデル化で満足し歩みを止めていたのでは生じ得なかった，科学の進化である。

3.3　対応バイアスとモデリング

対応バイアス（correspondence bias）とは外的要因に行動の原因を帰属できる状況において，個人の特性という内的要因に原因を帰属するバイアスである。Jones and Harris（1967）の有名な実験では，参加者は，ある人物がディベート・リーダーから指示されてキューバの首相フィデル・カスト

3)　なお，このような前提は，経済学ではランダム効用モデルとしてごく一般的に置かれている。

ロを支持する（あるいは否定する）エッセイを書いたと教示される。そしてエッセイを読んだ後でその人物が持つ態度を推測するよう求められた。実験の結果，作者はエッセイに書かれたのと同じ方向の態度を持つと，参加者は推測していた。カストロ支持（あるいは否定）のエッセイを書くという行動は，リーダーの指示という外的要因に帰属できるはずなのに，作者がエッセイの方向と一致する態度を持つと推測することは論理的には誤りのはずである。対応バイアスは，人間が持つ非合理的な認知バイアスの代表例として，広く一般に知られている。

ここで，なぜこの実験結果から，非合理的な推論が行われたという結論が導かれるのか，社会心理学者が暗黙裡のうちに想定してきたプロセスをモデル化してみよう。まず人間の行動は，外的要因か内的要因のいずれかの原因によって，相互背反的に決定されているとしよう。そして人間は，リーダーからある行動を取るよう命令されると，その行動を好むか否かに関わらず，必ずそれに従うとする。この時，もしある人物がリーダーから指示された行動を取っていたとしても，その事実から，その人物がその行動を好んでいると推論することはできない。

このように明示的にモデル化すると，ここで前提とされているのは決定論的な世界であることがわかるだろう。だが，人間の行動は決定論的な因果によって規定されているという前提は，どれほど妥当なのだろうか？ 同一の状況におかれても，個人によって行動にばらつきが生じるのは当然のことであるだろう。リーダーからの命令があってもサボタージュして従わないこともあるだろう。そもそもリーダーからの命令にも強弱があるはずではないだろうか。こうしたより現実的な状況では，決定論的な世界を前提とした推論ではなく，確率論的な推論を行うことが合理的なはずである。そして，合理的な確率的推論は，ベイズ推定（ベイズ・モデル）によって表現できる。

対応バイアスのベイズ・モデル　Walker, Smith, and Vul（2015）は，こうした考えに基づいて，ここで紹介したような実験状況において，人間がベイズ的に原因帰属を行うモデルを考案した。以下，その要点だけ紹介しよう。まず個人が持つ特性を d というパラメータで表現する。個人が自由に行動を選択できる場合，d が 0 より大きな値を取るほど，その行動を選択する確率が 1 へ近づく。d が 0 の場合，行動を選択する確率は 0.5，そして d が 0 より小さくなるほど行動を選択する確率は 0 へ近づく。同様に，個人に対して行動を取るよう強制する外的要因の強さを s で表すことにする。$s=0$ ならば，外的な圧力が何もないことを意味し，s が 0 より大きくなるほど行動を選択する確率が 1 へと近づく。

このように，人間が行動を選択する確率は，個人の特性（d）と状況からの圧力（s）という 2 つのパラメータの和によって決定されるものとしよう。人間がこうした行動原理に従って振る舞うことを知っている観察者がいたとする。そしてこの観察者は，ある行為者に対して行動を選択するよう強制する外的圧力が加わったことを観察し，さらにその行為者がその行動を選択したことを観察する。ここで観察者が行うのは，この行為者が持っている特性 d の値を推測することである。

Walker et al.（2015）は，合理的なベイジアンである観察者は，次のように推論すると考えた。まず観察者は，事前分布 $P(d)$ を心の中に持っているとする。これは「特性 d は人間という母集団においてどのように分布しているか」について観察者が持つ事前信念である。そして状況の観察を通して，行為者に加わった状況の強さ s の値を査定する。最後に「行為者がその行動を選択した（$a=1$）」という情報を受け取った観察者は，査定した s の値と，自らが持つ事前信念（事前分布）から，事後信念（事後確率）$P(d|s, a=1)$ をベイズの公式に従って推定する。

ここで注意して頂きたいのは，行為者の持つ特性 d をどのように推測するのかは，観察者が持つ事前信念 $P(d)$ と知覚された状況の強さ s という 2 つのパラメータの値によって左右されるが，いずれも Jones and Harris（1967）の実験では測定されていない点である。詳細は論文に譲るが，ウォーカーらは，これら 2 つのパラメータを適当な値に設定すると，ベイジアン合理的な推論者は Jones and Harris（1967）の実験結果と酷似した推論を下すことを数値計算によって示したのである。

これは重大なインプリケーションを持つ。Jones and Harris（1967）の実験結果からは，「人間

は非合理的な推論者である」という結論を一意に導けないことを意味するからである。そうした結論を導くためには，人間は決定論的な推論を行うという前提を置く必要がある。だが，もし人間が確率的な推論を行っているという前提を置くならば，「人間は非合理的な推論者である」という，数十年に渡って社会心理学で受容されてきた仮説が，この実験結果によって反証される可能性が出てくる（付録参照）。

4. 心理学が未来へ進むための道筋

再現可能性問題を巡って生じつつある大きな動き，すなわちQRPの排除や出版バイアスの抑制，追試の奨励は，心理学が信頼に値する科学の一領域となるために，必要不可欠な動きである。本論文の目的は，それだけでは不十分であること，モデリングという多くの領域に浸透している科学の営みを取り込む必要性を指摘することにあった。だが，心理学において厳密なモデルを構築して研究を行うことに対して，懐疑的な態度を取る研究者も多いことだろう。

厳密なモデルの限界　ミクロ経済学者でありゲーム理論家である神取道宏は，厳密な理論モデルの限界を，「街灯の下で鍵を探す男」というメタファーを使って紹介している（神取，1994）。月のない夜，男が街灯の下で探しものをしている。落とした鍵を探しているのだというが，鍵を落としたのは，街灯が当たらない遠く離れた場所だという。なぜこの男は，鍵が落ちているはずのない街灯の下を探し回っているのか？ それは，街灯が当たっている場所以外は何も見えないからだと，男は説明する。

厳密な理論は，明るい街灯のように全てを照らし出す。だが，厳密な理論モデルが表現できる現象には，多くの場合限界がある。もし厳密な理論モデルに固執し続けるならば，本当に重要で科学者が関心を持つべき対象を，研究の俎上に取り上げられないこともあるだろう。おそらく多くの心理学者は，厳密なモデルが持つこうした限界に気づいている。故に，慣れ親しんで来た「弱い理論」に代わって，厳密なモデルを心理学に導入することに抵抗を感じるだろう。

「弱い理論」の歴史的な意義　これこそまさにEysenck（1985）が，科学は「弱い理論」から出発するしかないと議論した所以である。Eysenck（1985）によれば，いかなる科学の領域も，最初はフレキシビリティの高い，厳密さに欠ける「弱い理論」から出発しなければならない。不正確で，厳密さに欠ける「弱い理論」は，フレキシブルであるからこそ，多くの現象について考え，議論し，仮説を構築する上で役に立つからである。正確さと厳密さにかける科学には，歴史的な必要性と役割があるのだ。また現代の心理学における理論は，その多くが言語によるモデルである。たとえば，心的回転課題の実験結果を説明するための，Shepard and Metzler（1971）による簡潔な言語によるモデルは，その後の視覚心理学の歴史を見る限り，たとえ不完全であったとしても充分に役割を果たしていた。このように見れば，今なお多くの心理学者が，厳密さに欠けた理論に基づいて仮説を導出し，実験結果を解釈しているとしても，決して一概に非難されるべきではない。

だが「弱い理論」，厳密さに欠けたフレキシビリティの高いモデルは，あくまでも出発点に過ぎない。科学とは，時間をかけて漸進的に進化していく知の体系である。これまでのやり方で生み出された研究にほころびが生じているならば，それを改めて先へ進んでいかなければならない。厳密なモデルが持つ限界を理由にして，いつまでも「弱い理論」の世界に踏みとどまることは，許されないはずである。

厳密なモデルを作る意義　自然科学，認知科学，神経科学，そして経済学といった領域では，モデルとは数理モデルや数値計算モデルを指すことが多い。だが，ワイスバーグが指摘するように，モデルは必ずしも数学を含む必要はない。自然言語だけでメカニズムの詳細を完全に記述できるならば，それは数理モデルとなんら変わることがない厳密さを持つ。だが，自然言語には限界が伴う。3節で見てきたように，言語だけで議論を展開していくと，どこかで必要な前提が見落とされ，長いことそれに科学者が気づかない事態が生じる。また，これはアネクドートに過ぎないのだが，ある数理心理学者が，欧米の著名な社会心理学者の理論を数理モデルで表現しようと試みたところ，頁が進むにつれて概念の定義が次々と変化する上に，論理に破綻が生じるため，厳密なモデ

ル化を断念せざるを得なかったという話を聞いたこともある。

モデリングがしばしば数理的な表現を伴うのは，自然言語だけで思考を展開する時に生じがちな，定義のすり替えや論理の飛躍，必要な前提の見落としが生じないよう，科学者に自覚を迫るからである。数理モデルや数値計算モデルを作るためには，細部まで厳密に設計しなければ先に進めない。これこそ，厳密なモデルを構築することのメリットである。

今心理学に求められているのは，「弱い理論」か，厳密なモデルのいずれかを選ぶことではない。これまでのようにフレキシブルな「弱い理論」による研究を続けることは当然としても，その中に厳密なモデルを漸次取り入れることで議論を精緻化させていくことなのである。

心理学が変化するために必要なこと　繰り返すが，モデリングの本質は，精緻で厳密に論理を展開していくことに他ならない。本論文の目的も，心理学の外部で生み出された特定の理論モデル，たとえば量子力学（Pothos & Busemeyer, 2013）や合理的選択理論，ゲーム理論や進化ゲーム理論を心理学に導入すべきだという主張を行うことではない。心理学における理論構成をより厳密にし，議論の中に隠れた前提をあぶり出すためには，研究がある程度進んだ段階で，厳密なモデルを構築していく必要がある―それが，本論文の主張である。

だが，モデリングを導入する上では，いくつものハードルが存在することだろう。数理的な表現を交えた厳密なモデルを作れるようになるためには，特別な素養がないかぎり，時間をかけてトレーニングを受ける必要がある。いまさらそれだけのコストをかけてモデリングを導入するよりも，フレキシブルな「弱い理論」のままで良いから，ひたすらデータを蓄積した方が個人的には高い成果を挙げられると見積もる研究者も多いだろう。また議論を精緻化するために，どのようなタイプの理論モデルを構築すべきか，研究領域によっても大きく異なる。手本がないところに，新たな道を切り開くのは難しい。こうした問題を考えれば，これまでモデリングに慣れ親しんだことのない研究者が，突然厳密なモデルを導入することは確かに困難であろう。

だが，科学はたったひとりの個人の手によって担われているわけではない。科学とは，世代から世代へと蓄積されながら，時間をかけて累積的に発展していく知の体系である。個人の手にあまるならば，次世代の手に委ねれば良い。そのためには，大学院生やポスドクがモデリングを学べる場を持てるよう配慮すれば良いだけである。「自らに残された時間を考慮すれば，新しい知識を学ぶよりも，今のやり方を続けた方が良い」と判断しても，次世代の人間に同じやり方を続けさせる必要はない。

最後に　モデリングという科学の営みは，心理学の中でも数理心理学などの形を取って，以前から存在してきた。だがこの営みが，さらに心理学の中に広まっていく上で，今は最高のタイミングであるかもしれない。再現可能性問題を端緒として，従来のやり方のまま研究を進めることには限界があるという認識が生まれつつある上に，厳密なモデルに触れる場が心理学の中で広まりつつあるからである。

本特集号において清水（2018）が紹介するように，ベイズ統計モデルは，従来の統計モデルが抱える限界を乗り越え，現象を記述するための自由度の高いデータ解析を可能とする。だがそのためには，Stan のような確率プログラミング言語を用いて，科学者が自らの手で厳密に細部を定義しながら統計モデルを構築していかなければならない。GUI のメニューから選択肢をマウスでクリックするのではなく，研究者がみずからプログラミングを書いていくことは，まさに厳密なモデルを構築することに他ならない。

また本特集号における国里（2018），そして中村（2018）では人間の認知プロセスを厳密に記述するための，ベイズ・モデルが紹介されている。1980 年代に社会心理学において「人間は非合理的な存在である」とする人間観が広まった。だが近年，認知科学や神経科学において，人間の心や脳はベイジアン合理性を達成するシステムであるとする考えが復権し，多くのベイズ・モデルが成功を収めている。その対象は飛躍的に拡大しつつあるため，自らの研究のためにベイズ・モデルに触れる機会も増えていくことだろう（付録参照）。

目的は異なるが，いずれもそれを学ぶことで自然と厳密なモデルやモデリングという考えを身に

つけることができる。科学の様々な領域に浸透した，モデリングという営為に至る道は複数存在する。そしてその中の幾つかの道が，いままさに心理学者の前に広く開かれている。

引用文献

Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype-activation on action. *Journal of Personality and Social Psychology*, *71*, 230–244. http://dx.doi.org/10.1037/0022-3514.71.2.230

Binmore, K. (2008). *Rational Decisions*. Princeton University Press.

Brehm, J. W. (1956). Post-decision changes in the desirability of choice alternatives. *Journal of Abnormal Social. Psychology*, *52*, 384–389. http://dx.doi.org/10.1037/h0041006

Camerer, C. F., Dreber, A., Forsell, E., Ho, T.-H., Huber, J., Johannesson, M., et al. (2016). Evaluating replicability of laboratory experiments in economics. *Science*, *351*, 1433–1436. http://doi.org/10.1126/science.aaf0918

Chen, M. K., & Risen, J. L. (2009). Is choice a reliable predictor of choice? A comment on Sagarin and Skowronski. *Journal of Experimental Social Psychology*, *45*, 425–427.

Chen, M. K., & Risen, J. L. (2010). How choice affects and reflects preferences: Revisiting the free-choice paradigm. *Journal of Personality and Social Psychology*, *99*, 573–594. http://dx.doi.org/10.1037/a0020217

Cova, F., Strickland, B., Abatista, A. et al. (2018). Estimating the Reproducibility of Experimental Philosophy. http://doi.org/10.17605/OSF.IO/SXDAH

Eysenck, H. J. (1985). The place of theory in a world of facts. In K. B. Madsen & L. Mos (Eds.), *Annals of Theoretical Psychology, Volume 3* (pp. 17–72). New York: Plenum Press. doi: 10.1007/978-1-4613-2487-4_2

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, *11*, 127–138. http://doi.org/10.1038/nrn2787

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, *103*, 650–669. http://dx.doi.org/10.1037/0033-295X.103.4.650

Gilbert, D. T., & Malone, P. S. (1995). The correspondence bias. *Psychological Bulletin*, *117*, 21–38. http://dx.doi.org/10.1037/0033-2909.117.1.21

Griffiths, T. L., & Tenenbaum, J. B. (2006). Optimal predictions in everyday cognition. *Psychological Science*, *17*, 767–773. http://doi.org/10.1111/j.1467-9280.2006.01780.x

Griffiths, T. L., Vul, E., & Sanborn, A. N. (2012). Bridging levels of analysis for probabilistic models of cognition. *Current Directions in Psychological Science*, *21*, 263–268. http://doi.org/10.1177/0963721412447619

池田功毅・平石 界（2016）心理学における再現可能性危機：問題の構造と解決策　心理学評論，*59*, 3–14.

Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine*, *2*, e124. doi: 10.1371/journal.pmed.0020124

Izuma, K., & Murayama, K. (2013). Choice-induced preference change in the free-choice paradigm: A critical methodological review. *Frontiers in Psychology*, *4*. http://doi.org/10.3389/fpsyg.2013.00041

Jones, E. E., & Harris, V. A. (1967). The attribution of attitudes. *Journal of Experimental Social Psychology*, *3*, 1–24. https://doi.org/10.1016/0022-1031(67)90034-0

神取道宏（1994）ゲーム理論による経済学の静かな革命　岩井克人・伊藤元重（編）現代の経済理論（pp. 15–56）東京大学出版会.

Knill, D. C., & Pouget, A. (2004). The Bayesian brain: The role of uncertainty in neural coding and computation. *Trends in Neurosciences*, *27*, 712–719. http://doi.org/10.1016/j.tins.2004.10.007

国里愛彦（2018）臨床心理学と認知モデリング　心理学評論，*61*, 55–66.

McElreath, R., & Smaldino, P. E. (2015). Replication, Communication, and the Population Dynamics of Scientific Discovery. *Plos One*, *10*, e0136088. http://doi.org/10.1371/journal.pone.0136088

Meehl, P. E. (1967). Theory-testing in psychology and physics: A methodological paradox. *Philosophy of Science*, *34*, 103–115. http://doi.org/10.1086/288135

中村國則（2018）高次認知研究におけるベイズ的アプローチ　心理学評論，*61*, 67–85.

Open Science Collaboration (2015). Estimating the reproducibility of psychological science. *Science*, *349*, aac4716–aac4716. http://doi.org/10.1126/science.aac4716

Pothos, E. M., & Busemeyer, J. R. (2013). Can quantum probability provide a new direction for cognitive modeling. *Behavioral and Brain Sciences*, *36*, 255–274. https://doi.org/10.1017/S0140525X12001525

Sanborn, A. N., & Chater, N. (2016). Bayesian Brains without Probabilities. *Trends in Cognitive Sciences*, *20*, 883–893. http://doi.org/10.1016/j.tics.2016.10.003

Savage, L. J. (1954). *The Foundations of Statistics*, Wiley.

Shepard, R. N., & Metzler, J. (1971). Mental Rotation of Three-Dimensional Objects. *Science*, *171*, 701–703. http://doi.org/10.1126/science.171.3972.701

清水裕士（2018）心理学におけるベイズ統計モデリング　心理学評論，*61*, 22–41.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology. *Psychological Science*, *22*, 1359–1366. http://doi.org/10.1177/0956797611417632

Sterne, J. A. C., & Smith, G. D. (2001). Sifting the evidence—What's wrong with significance tests? Another comment on the role of statistical methods. *BMJ (Clinical Research Ed.)*, *322*, 226–231. doi: 10.1136/bmj.322.7280.226

友永雅己・三浦麻子・針生悦子（2016）巻頭言　心理学の再現可能性：我々はどこから来たのか　我々は何者か　我々はどこへ行くのか―特集号の刊行に寄せて―　心理学評論，*59*, 1–2.

Walker, D., Smith, K., & Vul, E. (2015). The "Fundamental Attribution Error" is rational in an uncertain world. *Proceedings of the 37th Annual Meeting of the Cognitive Science Society*. https://mindmodeling.org/cogsci2015/papers/0437/paper0437.pdf

Weisberg, M. (2013). Simulation and Similarity: Using models to understand the world. Oxford University Press. 松王政浩（訳）（2017）科学とモデル：シミュレーションの哲学入門　名古屋大学出版会.

Wilson, B. M., & Wixted, J. T. (2018). The Prior Odds of Testing a True Effect in Cognitive and Social Psychology. Advances in Methods and Practices in Psychological Science, 251524591876712–12. http://doi.org/10.1177/2515245918767122

付　録

本特集号の中村（2018）で紹介されているように，人間の意思決定や推論がベイズの定理に従うことを示した研究は認知科学において増えつつあり（Griffiths & Tenenbaum, 2006），3.3節で紹介したWalker et al.（2017）のモデルもその流れの中で登場したものである。また近年は神経科学においても，視知覚や感覚運動制御の振る舞いがベイズの定理に従うことが明らかとなり（Knill & Pouget, 2004），ベイズとは脳の機能を表現する統一原理だという考えも登場してきた（Friston, 2010）。この付録では，ベイズ・モデルが持つ意義について紹介をしたい。なぜならば，モデリングという科学の営みを心理学に導入するにあたって，多くの心理学者が直面するであろう問題を考える上で，ベイズ・モデルは格好の教材となるからである。

人間の心はベイジアンか？

3.3節で問われていたのは「人間の原因帰属は合理的か」という問いであった。これは意思決定研究において典型的に探求されるタイプの問いである。中村（2018）も指摘する通り，ベイズ・モデルを用いた研究では，「人間の意思決定はベイジアン合理性という基準から照らして適切であるのか（規範的アプローチ）」，「人間の認知はベイジアン合理的なモデルによって表現できるのか（記述的アプローチ）」が問われる。

そのような問いに関心を持つならば，ベイズ・モデルを用いた研究には意義がある。だが，多くの心理学者は，人間が推論や決定を行う際の心理プロセスやメカニズムに関心を持つだろう。その立場から見れば，ベイズ・モデルは非現実的で無意味なものに思える。実際，ベイズ・モデルは2つの理由において，非現実的である。

第1にベイズ推定は，非常に複雑な計算処理を必要とする。特に実際の研究にベイズ・モデルを適用する場合，モデルが複雑になるため多くの情報と計算が必要とされるのだが，モデルが想定する計算量は現実の人間には処理不可能なレベルに達する（Gigerenzer & Goldstein, 1996; Gigerenzer, Todd & the ABC research group, 1999）。第2に，ベイジアン合理性の研究者自身が明言しているように（Binmore, 2008; Savage, 1954），ベイズ推定は現実の複雑な世界には適用できない。なぜならばベイジアン合理的な行為者は，「将来起きる出来事」を全て認識していることが想定されているからである。実験室では，参加者に全ての情報を与えることで，そのような状況を設定できるかもしれない。だが，我々が暮らす現実の世界では，誰にも予期し得ない出来事が常に生じうる。前者は小さい世界（small worlds），後者は大きな世界（large worlds）と呼ばれるが，原理的には大きな世界にベイズを適用することはできない（Binmore, 2004）。

ベイズ・モデルの複雑な数式を見た後で，こうした議論を聞けば「厳密なモデルを使った研究とは，非現実的に複雑なモデルを作って実験データに当てはめているだけの数学遊びではないか。そのような研究には意味などない」と感じる人も出てくるだろう。

だがそうした感想は，記述モデル（descriptive models）とプロセスモデル（process models）の混同から生じる誤解である。

記述モデルとプロセスモデル

そもそも記述モデルとは，その名が指し示す通

り，現象を記述することを目的としており，その内部で想定された処理が実際に行われている必要はない。たとえ現実の人間には不可能な計算処理を仮定するベイズ・モデルであっても，それによって現象が記述できるならば，そこに意味を見出す人々には価値ある研究対象となる。

一方，人間の意思決定をベイズによって記述できたならば，次に以下のような問いが立てられるべきだろう―「ではなぜ，現実の人間には計算不可能なはずのベイズ・モデルが，人間の意思決定をうまく記述できるのか？ どのような情報処理が心や脳の中で行われるならば，あたかもベイジアン合理的な計算が行われたかのような結果が生み出されるのか？」(Griffiths, Vul, & Sanborn, 2012)。

それを担うのがプロセスモデルである(Gigerenzer et al., 1999)。たとえば3節で紹介した対応バイアスの場合にも，「人間の原因帰属には非合理的なバイアスがある」ことが指摘された後に，「ではどのような心理プロセスによって，バイアスの生じた判断が生み出されるのか」が探求されている（Gilbert & Malone, 1995)。

Sanborn and Chater（2016）は，ベイジアン合理性を生み出す心理プロセスについて，興味深い可能性を指摘している。一般的にベイズ推定では，事前分布と尤度から事後分布を解析的に求める。だが，すでに指摘した通り，現実の複雑な問題にベイズ推定を適用する場合，モデルが複雑になってしまうため，たとえコンピュータを使っても事後分布を解析的に求めることが不可能となる。統計学においてベイズ統計が利用される際，この問題を解決するために用いられるのがマルコフ連鎖モンテカルロ（MCMC）と呼ばれる手法である。一言で表すならば，数式を計算して解を求めるのではなく，コンピュータにサイコロを振らせて特定の条件を満たすデータのサンプリングを繰り返し，事後分布を近似的に求める手法である。

Sanborn and Chater（2016）は，人間の脳や心が，MCMCのようなサンプリングのプロセスを通して，近似的にベイズ推定を行っているのではないかと指摘している。人間の心は，確率分布の積分計算は苦手であっても，具体的な事例を無数に想像したり，記憶の中から多くの事例を想起することが得意である。これはコンピュータが，特定の条件に合致するデータを繰り返しサンプリングするプロセスとよく似ている。このように考えれば，人間の心が日常的に行っている自然な情報処理のプロセスによって，あたかもベイジアン合理的な推論が生み出され得る。

厳密なモデルが持つ様々な役割

「厳密なモデル」の例としてベイズ・モデルを取り上げるならば，モデリングは自分の研究には関係ない，訳に立たないと感じる人もでてくることだろう。だが，そもそもベイズのような記述モデルは，意思決定から推論，視知覚から運動といったレベルの異なる対象を，ベイズという統一原理によって接合する役割を持っている。ベイズという枠組みを通すことによって，一見すると関係を持たないように見える現象を統合的に理解することに貢献する。だが記述モデルには限界がある。ある現象がベイズ・モデルで記述できないことが明らかとなった場合，なぜそうなるのかを説明できないからである。実際，古典的な意思決定や推論の研究を見渡せば，人間の推論がベイジアン合理性を満たさない事例を数多く見て取ることができる。人間は，いついかなる場面でも常にベイジアン合理的な推論を下すわけではない。

一方，プロセスモデルは，詳細で具体的であるために適用可能な対象は狭まる。その代わりに，「いつどのような条件の下で，人間の心がベイズ的な振る舞いを示すのか」を説明し，予測できる。実際，Sanborn and Chater（2016）は，人間の心がデータのサンプリングというプロセスによって近似的に事後分布を推定しているとみなすことで，人間の推論において古典的な非合理的バイアスが生じる条件を特定できると論じている。

本論文では，自然言語のみに頼った議論の曖昧さを排除するために厳密なモデルが有効であることを中心的に論じてきた（特に3.2節)。だがひとくちに「厳密なモデル」といっても，様々な種類がある上に，その役割は様々である。この付録では特に，心理学とその関連分野でモデルが導入される際，特に誤解されやすい問題―記述モデルとプロセスモデルの違いと，その相補的関係―について，ベイズ・モデルを例として紹介した。

― 2018. 2. 16　受稿，2018. 3. 7　受理 ―