

# 精度最適化は創発的性質なのか

西村正秀

## 1. 序

知覚の認知的侵入可能性とは、知覚が信念や欲求などの高次認知状態によってその形成に直接的な影響を受けうることである。知覚の認知的侵入可能性を否定する論証に「注意シフトに基づく論証」がある<sup>(1)</sup>。この論証によれば、知覚への認知的侵入を示す事例の多くは、認知に駆動された注意による刺激自体の変更に起因するものとして理解できる。その場合、「あるプロセスが感知する刺激の種類が限定されていること」という知覚のモジュール性が維持され、知覚の認知的侵入可能性は否定される<sup>(2)</sup>。近年、注意シフトに基づく論証に対して、注意のバイアス競合モデルの擁護者から反論が提出されている。従来の理論では、注意は認知と知覚の中間プロセスとされてきた。だが、バイアス競合モデルによれば、注意は知覚プロセスの一部であり、認知が注意を駆動する以上、知覚も認知によって直接的に変化することになる。それゆえ、知覚の認知的侵入可能性は肯定される。

現在、知覚の理論として、知覚をベイズ主義的推論と見なす「予測誤差最小化モデル(prediction error minimization model)」が注目を浴びている。このモデルでは、注意は「精度最適化(precision optimization)」として特徴づけられる。何人かの論者によれば、この特徴づけはバイアス競合モデルと同様に、注意を知覚から独立したプロセスではなく、知覚プロセスの一部と見なすものとして解釈可能である。この解釈が正しければ、予測誤差最小化モデルが採用された場合、注意シフトに基づく論証は反駁されると思われる<sup>(3)</sup>。

本稿では、予測誤差最小化モデルに関するこのような見込みを批判的に検討する。最初に、このモデルにおける精度最適化としての注意概念を説明する(第2節)。次に、注意シフトに基づく論証の構造を明らかにした上で、精度最適化としての注意がこの論証をどのように反駁しうるのかを説明する(第3節)。そして、その反駁には十分な説得力がないことを、精度最適化としての注意が持つ存在論的身分の検討を通じて示す(第4～5節)。

## 2. 予測誤差最小化モデルと注意

### 2.1 予測誤差最小化モデル

予測誤差最小化モデルには様々な定式化があるが、その代表の一つはカール・フリスト

ンによる「自由エネルギー原理」に基づいたモデルである(Friston, 2009, 2010)<sup>(4)</sup>。このモデルは現在心理学者のみならず、アンディ・クラークやヤコブ・ホーヴィなどの哲学者によっても支持されている(e.g., Clark, 2013; Hohwy, 2013)。元々、自由エネルギーは19世紀にヘルマン・フォン・ヘルムホルツが唱えた熱力学における状態量を表す概念であるが、予測誤差最小化モデルでは、情報理論における「サプライザル」(予測が外れたときの驚きの程度を表す関数)の上限を構成する状態量を表す。ところで、ヘルムホルツは自由エネルギーの話に加えて、知覚とは感覚入力からその原因である外的対象を無意識に推論するプロセスであるとも考えていた。予測誤差最小化モデルは、この「知覚=推論」というアイデアを継承した、計算主義理論の一種である。感覚入力の原因となる外的対象の候補は複数あり、前者から後者を特定する際には決定不全性の問題(「逆問題」と呼ばれる)が生じる。この問題を解決するために、予測誤差最小化モデルは前者から後者の推論をベイズ主義で説明する。脳は外界に関する様々な仮説を形成し、それらに事前確率を付与する。また、ある仮説の下である感覚入力を与えられる尤度も経験的に学習する。そして、ある感覚入力を与えられた場合の事後確率をベイズの定理にしたがって計算し、事後確率が最も高い仮説を採用する。ただし、現実的には正確な事後確率を求めることは困難なので、脳は事後確率の近似値を計算する。この近似値と正確な事後確率の差は、カルバック=ライブラー・ダイバージェンスで表される。このダイバージェンスにサプライザルを加えたものが自由エネルギーであり、後述する予測誤差に相当する。正しい知覚推論とは予測誤差を最小化するものであり、「自由エネルギー(予測誤差)を最小化せよ」という形で表されるのが自由エネルギー原理である。

予測誤差最小化モデルは、標準的な計算主義理論とは異なる特徴を有している。ここでは、その中から四点を紹介しておこう。第一に、予測誤差最小化モデルは統一的性格を持つ。このモデルでは、知覚だけでなく、信念や記憶や行為など脳の機能すべてが自由エネルギー原理という単一の原理で説明される。この統一的性格が、多くの哲学者をこのモデルに惹きつける主たる要因となっている。

第二に、予測誤差最小化モデルでは、知覚は外界の情報の中から、仮説に合致する情報ではなく、仮説に合致しない誤った情報を伝達する。ベイズ主義的推論では事前確率が付与された仮説から、次に得られる感覚入力の予測が導かれる。だが、その予測が実際の感覚入力と合致することには、情報という観点から見れば大した価値はない。むしろ、価値があるのは予測と実際の感覚入力に誤差がある場合であり、その時には仮説の修正が行われる。この「仮説を修正して予測誤差を最小化する」ことを知覚の本性と見なす点は、「脳は感覚入力から受動的に外界の表象を形成する」と素朴に考える標準的な計算主義理論と

は異なっている。

第三に、予測誤差最小化モデルでは、脳の認知システムは階層化される。外界の仮説は心的表象という形で与えられるが、外界は時間空間の様々なスケールに応じて、多様な仕方では表象される。例えば、庭にあるバラの樹を表象する場合でも、風による揺れなど微細な特徴も含めれば時間的に移ろいやすい表象となり、そのような細部を無視すれば抽象度の高い安定した表象となる。外界の仮説はこのような抽象度の違いに応じて、信念などを含む高次の階層から感覚的な低次の階層に至る様々なレベルで作成される。外界の仮説を事前知識とし、そこからの予測を知覚の基盤とするアイデアは、高次認知状態からのトップダウン・プロセスが知覚の構成要素であることを意味している。このトップダウンによる情報伝達と感覚から上がってくるボトムアップの情報伝達、そして、それらによる仮説のテストは、それぞれの階層間で行われる。

第四に、予測誤差最小化モデルは、知覚と行為を表裏一体と見なす理論である。このモデルによれば、知覚で目指されるのは予測誤差の最小化であった。この目的を達成する方法には大別して二つある。一つは、上述した、与えられた感覚入力を予測と照らし合わせて仮説を修正する方法であり、これは知覚推論である。もう一つは、単に感覚入力を与えられるのを待つのではなく、主体が外界に働きかけて仮説を確証する情報を収集する方法である。知覚表象には、主体がとるべき行為を命令する役割も付与されている。主体は知覚表象の誤り（予測誤差）を最小化するために、その表象の正しさを示す新たな感覚入力を得るように行為するのである。このような行為を命令するプロセスは「行為推論(active inference)」と呼ばれる(Friston et al., 2011)。予測誤差最小化モデルでは、知覚推論と行為推論は、実質的には共通の神経基盤に基づいていると考えられている。

## 2.2 注意＝精度の最適化

前節では、予測誤差最小化モデルの骨格を説明した。本節では、このモデルで注意がどのように特徴づけられるのかを説明する。

最初に、注意について簡単な解説をしておこう。注意とは、心の情報選択メカニズムのことである。20世紀以降、様々な注意のモデルが提案されているが、本稿に最も関係するのは、ロバート・デジモンとジョン・ダンカンによる「バイアス競合モデル」である(Desimone & Duncan, 1995)。脳は複数のタスクを同時遂行してしまうので、特定のタスクを首尾よく達成するために情報を選択しなければならない。バイアス競合モデルでは、この選択は知覚プロセスや行為制御における競合が生み出す創発的性質として特徴づけられる。知覚プロセスは低次階層に行くほど細胞の受容野が小さくなる。プロセスの各階層で、細胞の受

容野に与えられた諸刺激は神経反応の奪い合いをしている。この競合に勝った刺激が選択され、負けた刺激は無視される。この勝敗を決める要因の一つが、高次認知プロセスで設定された目的によるバイアスである。目的の達成に貢献する刺激を選択するように、高次認知プロセスからトップダウンでバイアスが掛かり、それによって競合の結果が左右される。注意とは、このようなトップダウンのバイアスや下位レベルからのボトムアップのバイアスの下での競合によって刺激が選択されることである。この説明は、注意の存在論的身分を創発的性質として規定する。従来のモデルでは、注意は感覚入力を決定する、知覚から独立のプロセスと見なされてきた。だが、デジモンとダンカンによれば、注意はこのような情報フィルタリングのプロセスではなく、「バイアス競合を通じて生じる刺激の選択」という形で創発する性質であり、知覚プロセスに帰される。注意を創発的性質とする理由は、注意を実現する特定タイプの神経状態が存在しないからである(Mole, 2011a, 2011b, 2012, 2015; Koralus, 2014)。バイアス競合は様々な神経活動によって多重実現される。実は後で論じるように、バイアス競合モデルにおける注意を創発的性質とする解釈には反論がある。しかし、議論の便宜上、とりあえずこの解釈が正しいとしよう。その場合、注意を知覚の一部とする点は、このモデルを他の注意のモデルから区別する大きな特徴となる。

予測誤差最小化モデルが指定する注意は、バイアス競合モデルによるそれと軌を一にしているものとして解釈できる。予測誤差最小化モデルでは、注意は予測誤差の最小化における「精度の最適化」として特徴づけられる(Friston & Stephan, 2007; Feldman & Friston, 2010; Hohwy, 2013)。数学的記述を使わずに言えば、精度とは、ニューロンから送られる信号の信頼性を表す尺度である。事実上、脳内の階層間で情報が伝達される際には、送信された信号に加えて必ず何かしらのノイズが含まれる。「信号の精度が高い」とは、予測誤差がノイズに起因しているのではなく、送信された信号に起因している見込みが高いことを意味する。ところで、精度は文脈相対的である。同じ予測誤差であっても、ノイズが多い環境と少ない環境とでは、精度は前者では低く、後者では高くなる。例えば、明るい照明の下で対象を知覚する場合と暗い照明の下で対象を知覚する場合とでは、得られる感覚入力の信頼性について違いが出る。したがって、脳はどのような環境ならばどのような刺激が得られそうかについて学習し、それを基に予測誤差の精度についての予測も行う。脳はこのような「特定の文脈における予測誤差の精度予測」という二階の知覚推論に基づいて、信号の精度を高めるように働く。この働きが精度の最適化であり、「その文脈で最も登場しやすい刺激に注意を向ける」という形で表現されるわけである。

以上の仕方の特徴づけられる注意概念は、注意に関する様々な現象を統一的に扱うことができる(Clark, 2013; cf. Ransom et al., 2020)<sup>5)</sup>。例えば、注意には外発的注意（特定の刺激

に反応して主体が自動的に向ける注意)と内発的注意(特定のタスクを達成するために主体が意図的に向ける注意)の区別がある。前者は、目立った特徴を持つ刺激は予測誤差をもたらす見込みが高いという形で説明される。後者は、上述したように、特定の文脈における刺激の登場パターンの学習に基づいて、登場が予測される刺激が選択されるという仕方  
方で説明される。ここで、登場しやすい刺激の予測においては、高次認知プロセスからトップダウンで送られる情報がバイアスとして働いていると考えることができる(Marchi, 2020)。その場合、精度の最適化はバイアス競合の勝者を表象する知覚プロセスの性質として理解可能であり、予測誤差最小化モデルにおける注意概念はバイアス競合モデルによるそれと親和的になる(Marchi, 2020; Feldman & Friston, 2010)<sup>(6)</sup>。

### 3. 注意シフトに基づく論証とその反論

本節では、最初に認知的侵入可能性の概念の明確化と注意シフトに基づく論証の再構成を行い、その上で、この論証への反論として「直接性からの論証」を紹介する。そして最後に、予測誤差最小化モデルが直接性からの論証を支持しうることを確認する。

#### 3.1 認知的侵入可能性と注意シフトに基づく論証

近年、知覚の認知的侵入可能性を示唆する経験科学の成果が数多く提出されており、この問題を巡る議論は活性化している。だが、同時に状況は泥沼化もしている。その一因は認知的侵入可能性という概念が論者によって別様に特徴づけられている点にある。本稿では、それらの特徴づけの中から、最も単純な「因果的定義」を使用する<sup>(7)</sup>。

#### 認知的侵入可能性の因果的定義

知覚Pが認知的に侵入可能であるのは、次の二つの条件が満たされる場合であり、かつ、その場合に限る。

- (1) Pは何らかの認知状態Cに因果的に依存している。
- (2) PとCの因果関係は(a)内的かつ(b)心的であり、(c)感覚入力を変化させる中間項を含まない。

この定義に関しては注意点が三つある。第一は、「知覚」の意味についてである。認知的侵入可能性は、知覚内容(知覚経験の表象内容)に関する場合と知覚プロセス(知覚内容の媒体)に関する場合とがある。本稿では特に断りがない限り、「知覚」にこれら両方を含める。第二は、認知状態と知覚の因果プロセスは複数の項で構成されている可能性がある

という点である。特定の感覚モダリティの表象を出力するのが単一のプロセスなのか、複数の下位プロセスが集まって階層的に構成されているのかは、経験的な問いである(cf. Clarke, 2020)。予測誤差最小化モデルを後者の立場と結びつける可能性も、ア・プリオリには否定できない。第三は、条件(2)についてである。(a)と(b)は、認知状態から知覚への因果的影響が主体の脳内だけで起こり、かつ、身体運動などの心的状態以外の作用を介さないという制約である。この制約のおかげで、「友人を探そうとして頭の向きを変更して、その結果、何が見えるかが変化する」といったトリヴィアルな事例を認知的侵入可能性の事例から排除することができる。また、(c)は、内的かつ心的なプロセスの中にも感覚入力の変化を引き起こす項が含まれるかもしれないので、それを排除するために追加された条件である<sup>(8)</sup>。条件(2)は、認知状態が直接的に知覚に因果的影響を与えていることを確保するものだと言える（上述したように、この直接的因果関係は複数の項から構成されうることに留意されたい）。なお、知覚と認知状態が直接的関係に立つことは、前者が後者の表象内容を計算において使用できることを含意している。

以上の定義を前提した上で、次に、注意シフトに基づく論証を再構成しておこう。注意シフトに基づく論証は、認知的侵入可能性の反対者が用いる常套手段である(e.g., Fodor, 1988; Pylyshyn, 1999; Firestone & Scholl, 2016)。ウサギアヒル図のような認知的侵入可能性を示すと見なされている事例の多くは、事前知識などの認知状態が特定の空間領域や対象や性質への注意を引き起こし、その注意によって感覚入力が間接的に決定されることで知覚が生じているものとして解釈できる（ここでは、注意は知覚や認知状態に類した心的状態ではなく、ある種の作用的なプロセスとして理解されていることに留意されたい）。このような「認知状態→注意→知覚」（→は因果的影響）という説明は、「認知状態→知覚」を求める因果的定義の条件(2)に反する。知覚が高次認知状態から何らかの仕方で影響を受けることは正しいとしても（このこと自体は誰も否定しない）、認知状態が決定するのは「どの刺激がどの知覚プロセスに与えられるのか」の決定を担う注意であり、一旦刺激が決定すればその種類の刺激のみを受け入れる知覚プロセスが生じるので、知覚のモジュール性は維持される、というのが注意シフトに基づく論証である。以下、まとめておこう。

#### 注意シフトに基づく論証

- P1: 認知的侵入可能性には、認知状態と知覚の因果関係が直接的であることが必要である。[因果的定義の条件(2)]
- P2: 多くの事例では、認知状態が因果的に直接決定するのは注意であり、知覚の因果的変更は注意を介した間接的な仕方で生じる。

C:           したがって、これらの事例は認知的侵入可能性を示すものではない。

### 3.2 直接性からの論証

注意シフトに基づく論証に対しては、バイアス競合モデルに基づいた反論が何人かの論者によって提出されている(Mole, 2015; Marchi, 2020; Stokes, 2021)。これらの論者による反論は、細部における違いはあるが、基本的には同じ論証である。本節では、その中からダスティン・ストークスによる議論を取り上げて、予測誤差最小化モデルがこの反論と同様の反論を構成できることを確認する。

ストークスの反論は「直接性からの論証」と呼ばれている(Stokes, 2021, chap. 5)<sup>(9)</sup>。この反論は、注意シフトに基づく論証の P2 を否定する論証である。彼によれば、注意シフトに基づく論証は、注意の特徴づけに関して三つの原理を前提している(*Ibid.*, pp. 116-117)。

- (1) スポットライト原理：注意は空間的注意（どの空間領域に焦点を合わせるかという役割を果たす注意）である。
- (2) 行為原理：注意は主体による身体運動や行為を含んでいる。
- (3) 門番原理：注意は認知と知覚の間の中間プロセスである。

ストークスはこれらの原理をすべて批判するが、最も重要なのは原理(3)である<sup>(10)</sup>。ストークスによれば、経験科学の成果は、認知状態に駆動された注意が知覚に変化をもたらす場合、その注意は空間的注意ではなく「非行為者の選択的注意」であることを示唆している。非行為者の選択的注意とは、主体が意図的に制御していない、特定の対象や性質を基盤とした注意のことである。この非行為者の選択的注意は知覚プロセスの一部であるとストークスは主張する。彼はこの主張を支持する理由を二つ挙げている。第一に、選択的注意の神経的相関者は外線条皮質の四次視覚野にも存在しており、また、眼球のサッケードは選択的注意のサインだと考えられているが、これらは通常知覚プロセスの一部と見なされる。第二に、選択的注意を知覚プロセスの一部と見なす理論が実際に存在している。ここで具体的に挙げられているのは、バイアス競合モデルとジェス・プリンツによる作業記憶モデルである<sup>(11)</sup>。もしこれらのモデルが正しければ、認知状態から知覚への因果関係は直接的なものとなる。それゆえ、注意シフトに基づく論証の P2 は否定される。

### 3.3 予測誤差最小化モデルと直接性からの論証

最後に、予測誤差最小化モデルで提案される「精度最適化」としての注意概念が、認知

的侵入可能性についてどのような判定を下しうるかを確認しよう<sup>(12)</sup>。ホーヴィは、精度最適化としての注意は学習率(learning rate)を変化させることによって知覚への認知的侵入を増減させると主張する(Hohwy, 2017, p. 81)。ホーヴィによれば、認知的侵入はベイズ主義的に最適な学習率から予測誤差最小化が逸脱することによって生じる。学習率とは、一回の学習で予測誤差の精度の重みを変更する量を表すパラメータのことである。学習率は主体が経験してきた学習によって変化するので、認知的侵入可能性には個人差がある。ある刺激に関してより大きな注意が払われると、その刺激に対する予測誤差の重みが増すことによって学習率が上がり(つまり、ボトムアップで得られる刺激の情報が重視され)、知覚への認知的侵入は起こりにくくなる。逆に、刺激に対する注意が減じると学習率が下がるので、事前知識が重視されて認知的侵入が起こりやすくなる。また、ホーヴィによれば、精度の予期自体も、感覚入力の精度に関する「高階の信念」である(Hohwy, 2017, p. 81)。もちろん、この信念は偽であることもありうる。その場合、それに基づく推論は錯覚や幻覚をもたらすが、それも認知的侵入可能性を示す事例である。

ここで、精度最適化としての注意が知覚とは独立のプロセスであるならば、予測誤差最小化モデルは注意シフトに基づく論証の餌食となる。しかし、上述したように、精度最適化モデルはバイアス競合モデルの一種として解釈できる。ここから、フランセスコ・マルキは精度最適化モデルを直接性からの論証と結びつけている(Marchi, 2020, chap. 8)。マルキによれば、精度最適化はバイアス競合によって実現される創発的性質として、知覚プロセスに帰される。そうすると、認知状態と知覚に直接的な因果関係が成立するので、「注意シフトに基づく論証」は反駁される。

マルキは、精度の最適化によって認知的侵入可能性が生じている具体例として、最適な刺激条件の下で曖昧な刺激が与えられる「双安定イメージ(bi-stable images)」の知覚を挙げている(Marchi, 2020, chap. 8)。例えば、Liu et al. (2012)の実験では、被験者は中心軸の周りを回転するダンサーのシルエットのアニメーションを、そのイメージの一点に視線を合わせながら見る。このシルエットは、時計回りと反時計回りが途中で入れ替わるように知覚される双安定イメージである。被験者にこのシルエットが双安定イメージであることを告げた上で、今見えている回転向きが逆になるように意図して見ることを要求した場合、イメージが逆回転する割合は、意図的な逆回転の知覚を試みない場合と比べて約2倍となった。これは視線が固定されている「潜在的注意(covert attention)」(身体運動を伴わない注意)の実験であり、同じ対象からの異なる知覚経験が注意のおかげでより早く生じる事例である。この注意は「逆回転を見る」という意図的な認知状態がバイアスとなり、次に登場する性質(逆回転)が予測されることによって実現していると解釈可能であり、認知的侵入



可能性を示す事例として説明できる<sup>(13)</sup>。

#### 4 精度最適化モデルの問題点(1)：遍在性の問題

これまで見てきたように、注意の精度最適化モデルは認知的侵入可能性を支持する理論だと思われる。本節と次節では、このような見立てに対して反論を提示する。

前節で確認したように、予測誤差最小化モデルによる注意シフトに基づく論証の反駁が直接性からの論証に依存しているならば、そもそもの直接性からの論証に説得力があるのを見なければならぬ。この論証は、バイアス競合モデルにおける創発的性質としての注意概念に訴えたものであった。しかし、実はバイアス競合モデルにはいくつかの問題がある。そこで、最初にこれらの問題を確認した後、予測誤差最小化モデルにそれらが継承されるのかを検討しよう。

バイアス競合モデルの問題はいくつか指摘されているが、それらの中で本稿と関係するのは次の二つである。第一の問題は、バイアス競合による注意の説明は広すぎるというものである(Watzl, 2017; Levy, 2019)。モデルの提唱者自身が指摘しているように(Duncan, 2006)、バイアス競合は脳における認知システムの至るところで生じている。しかし、このような脳の中で遍在的に生じているバイアス競合のすべてが注意と関わっているという主張は行き過ぎである。なぜなら、通常は注意が介在していないとされている現象においても、バイアス競合は生じているからである。例えば、網膜においてすら、バイアス競合を実現する神経プロセスは生じている(Watzl, 2017, p. 31)。あるいは、ある人がうわの空で車を運転している場合、それでも車の運転というタスクの達成のために、歩行者の有無を確認する、ハンドルを見たりするなどの知覚は不十分ながらも遂行されており、最低限のバイアス競合は生じているであろう。しかし、一般的にこれは「不注意」に分類される行為だと考えられる<sup>(14)</sup>。

第二の問題は、バイアス競合モデルにおける注意を必ずしも創発的性質と見なさなくてもよいというものである。たしかに、このモデルを提唱したデジモンやダンカン、注意を刺激選択の原因としてのプロセスではなく、バイアス競合の結果として刺激が選択されることと同一視し、それを知覚プロセスの創発的性質と見なした。また、このモデルの擁護者の多くも、この考えを踏襲している。例えば、クリストファー・モールはオーケストラにおいて多様な楽器がユニゾンした結果ある音楽が創発されるように、注意を多様な認知的要素がユニゾンして創発する性質だと主張し、その存在論的身分を何らかの神経プロセスと同定することを拒んでいる(Mole, 2011a)。このような考えの背後にあるのは、本稿第2節で見た「注意は多重実現される性質であり、ある機会に注意を構成したプロセス・

トークンが別の機会にも注意を構成するとは限らない」という論点である。モールによれば、注意はプロセスではなく、情報選択の「仕方」として副詞理論的に理解されるべきなのである(Mole, 2011a; Mole & Henry, 2017)。しかし、モール自身も認めているように、バイアス競合モデルには、注意をバイアス競争というプロセスに還元する解釈も存在している(Mole, 2012)。サビーネ・カストナーとレスリー・アンガーライダーによれば、バイアス競合モデルにおける注意は、高次認知状態が生じる前頭部皮質と頭頂皮質から送られる情報を担うプロセスに還元できる(Kastner & Ungerleider, 2001)。また、セバスチャン・ワッツルは、モールによる注意の副詞理論を批判して、注意は、(1)主体が意識レベルで情報選択するプロセスとしてか、あるいは、(2)競合を担う個別の意識下プロセスではなく「バイアス競合」というより抽象的で全体論的な意識下プロセスとして同定できると主張する(Watzl, 2011)。もしこれらのプロセスが知覚から独立したプロセスとして理解できるならば、注意シフトに基づく論証が復活することになる。

議論の便宜上、これらの問題には説得力があるとしよう。では、これらの問題は予測誤差最小化モデルにも適用されるのであろうか。まず、第一の問題から検討しよう。この問題を回避する一つの方法は、バイアス競合の中から注意に関わるものを特定することである。残念ながら、バイアス競合モデルの提唱者はこの点に関して具体的な提案を出していない。だが、予測誤差最小化モデルの提唱者は、この問題に次のように返答するかもしれない。上述したように、ホーヴィは、ノイズの過多や感覚証拠の不確実性を反映して精度の最適化が行われ、その結果、学習率が変動し、知覚への認知的侵入が生じるか否かが決まると主張した。この主張にしたがえば、どのような精度の最適化が生じるかは、部分的にはノイズの過多や感覚証拠の不確実性によって決まるのであり、これはバイアス競合の中から注意に関わるものを特定するための実質的な条件だと言えそうである<sup>(15)</sup>。

しかし、このような反論は二つのリスクをはらんでいる。一つは、注意に関わるバイアス競合プロセスの特定は、本来バイアス競合モデルの強みであった、注意を創発的性質と見なして知覚プロセスの一部と見なすアイデアに反して、注意を特定の独立したプロセスに還元してしまう可能性を持ち込むという点である。これは、上述したバイアス競合モデルの第二の問題が発生しうることを意味する。もう一つは、この反論は精度最適化モデルにおける注意概念が伴う不整合を結局は解決しないという点である。たとえ精度最適化モデルにおいてバイアス競合が生じない事例が認められたとしても、精度最適化が常に認知プロセスの至るところで生じていることは否定されない。実際、ホーヴィは「知覚推論は予期された精度によって常に導かれるので、注意は予期と並んで常に学習率を調節する役割を果たす」と主張している(Hohwy, 2017, p. 81)。そうすると、いくら精度最適化の中で認

知的侵入を引き起こすものと引き起こさないものを区別できたとしても、「精度最適化＝注意」という図式自体は維持されることになる。これは注意概念をトリヴィアルなものとしてしまい、一般的に「不注意」と呼ばれる知覚の事例においても注意が生じていることになるという問題は解決されないままとなる。

以上の議論は、もちろん精度最適化モデルを完全に退けるものではない。第一のリスクに対しても第二のリスクに対しても、「注意はあくまでも程度を許容する概念であり、その中で知覚への認知的侵入を許す外的条件を規定しただけだ」と開き直すことは可能であろう<sup>(16)</sup>。しかし、注意が知覚プロセスの創発的性質であることを否定する第二の問題はより深刻である。そこで次に、この問題が精度最適化モデルに及ぼす影響を考察しよう。

## 5. 精度最適化モデルの問題点(2)：創発的性質の問題

### 5.1 注意はどのようなタイプの創発的性質でありうるのか

精度最適化としての注意を創発的性質ではなくプロセスとして同定できるかという問題を考察するにあたり、最初に「創発的性質」とは何かをもう少し詳しく見ておきたい。創発的性質とは、その基盤となる下位性質に依存しているが、それに還元されるわけではなく、それとは区別された性質として存在するものである。標準的理解では、創発的性質はそれ自身で因果的効力を持つ。しかし、創発的性質のさらなる特徴づけについては、依存関係の在り方などを巡り様々な仕方がある。ここでは、O'Conner (2021)による整理を参考にしながら、本稿に必要な二つの区別を提示しておく。一つは、弱い創発と強い創発の区別である。前者は、創発的性質はそれが依存する基礎的な物理的対象や性質によって完全に構成され、形而上学的に決定されているという立場であり、後者は、創発的性質の中にはそれが依存する物理的対象や性質によっては完全に構成されないものがあることを認める立場である。もう一つは、創発的性質とその下位基盤となる物理的性質の依存関係に関する区別であり、これについては、(1)スーパーヴィニエンス、(2)機能的実現、(3)法則的あるいは因果的依存の三種類がある。

残念ながら、バイアス競合モデルの注意を創発的性質と解釈する論者は、「創発的性質」や「創発」によって自分たちが何を意味しているのかを明示的に説明していない。本稿では、このモデルにおける注意は弱い創発的性質であり、かつ、その依存関係は機能的実現であるという理解を提案する。まず、注意の因果的効力がその神経基盤によって実現されないこと（物理的クロージャーを超え出ること）を明確に主張する論者は見出せない。モールとエアロン・ヘンリーは、「ある人が注意を払うことは、その人の行動的・認知的パフォーマンスの因果的効力や持続や速さや正確性を説明するのに不可欠な役割を果たす」と

主張するが(Mole & Henry, 2022, p. 2)、少なくともトークン・レベルでは、注意の因果的効力がバイアス競合などの神経プロセスや、さらにその物理的基盤の因果的効力を超えるものだと理解する必要はない。それゆえ、注意は弱い創発的性質と見なしうる。次に、依存関係については、(2)の機能的実現と理解するのが安全策である。(3)は強い創発を唱える場合に用いられる関係なので、ここでは無視できる。また、注意を(1)と見なすことはモールが否定している(Mole, 2011a, chap. 3)。スーパーヴィニエンスとは、「性質 A が性質 B にスーパーヴィンするのは、B に変化がなければ A が変化することはできない場合かつその場合に限る」という関係である。モールによれば、例えばトリーズマン的な特徴統合プロセスは、視覚探索の速さを競うタスクでは注意となるのに、他のタスクに没頭している場合は注意とならない。同じ神経プロセスは文脈に応じて注意にも不注意にもなるので、注意は神経プロセスにスーパーヴィンしないというわけである。したがって、本稿では依存関係を機能的実現、すなわち、「創発的性質の機能的役割は下位レベルの物理的性質により実現される」という関係として理解しておく。実際マルキは、注意やバイアス競合は多様な下位基盤によって「実装される(implemented)」(「実現される(realized)」と同義)と表現しており(Marchi, 2020, chap. 5)、これは本稿の理解と整合的である。

## 5.2 精度最適化としての注意はプロセスである

前節で、バイアス競合モデルにおける注意は、もし創発的注意として解釈される場合は機能的実現に基づく弱い創発的性質であることを確認した。これは言わば、注意の機能主義的特徴づけである。バイアス競合モデルにおける注意と類比的に精度最適化モデルにおける注意を創発的性質として理解する場合には、後者にもこの特徴づけが適用されることになる。では、このような適用に説得力はあるのか、それとも、精度最適化としての注意はプロセスとして同定できるのか。

この問いに対して、本稿では、精度最適化としての注意はプロセスとして同定できると答えたい。バイアス競合モデルで注意を創発的性質とする理由は、バイアス競合による刺激の選択が多様な神経プロセスによって多重実現されるからであった。だが、上述のように、バイアス競合は高次の包括的プロセスとして抽出することができる。同じ戦略は精度最適化モデルにも適用できる。この場合、高次プロセスとして同定されるのは「精度最適化」である。実際フリストンは、バイアス競合モデルにおける注意が創発的性質である点を把握しながら、注意を「階層的推論における感覚情報(予測誤差)の精度を表象するシナプス・ゲインを最適化するプロセス」と記述している(Feldman & Friston, 2010, p. 2)。また、クラークも、このフリストンの文言に言及しながら、「注意は...、何らかの現在与え

られているタスクや脅威や機会に対して最も信頼できる感覚情報を提供すると見積もられた予測誤差ユニットへのゲイン...を主体が高める手段あるいはプロセスである」と述べている(Clark, 2016a, pp. 59–60)。

しかし、精度最適化を高次の包括的プロセスとして規定することは、実質的な意味を持つのであろうか。知覚の認知的侵入可能性を問題とする場合に重要なのは、精度最適化を知覚とは独立のプロセスとして、あるいは、「認知状態→注意→知覚」の図式を崩さないように定立できるか否かである。ある意味で、精度最適化を「プロセス」と呼ぶことは容易い(実際、クラークはバイアス競合も「プロセス」と記述している(Clark, 2016a, p. 61))。問題は、そのプロセスがどのような因果的役割を果たしうるのかである。

この問題に対する本稿の答えは、精度最適化は刺激の選択を引き起こすプロセス、あるいは、刺激の選択を準備するプロセスとして理解できるというものである。もし精度最適化が創発的性質ならば、そこには刺激の選択が既に構成的に含まれているので、その機能(因果的効力)に刺激の選択は含まれないはずである<sup>(17)</sup>。それに対して、本稿が提案するのは、精度最適化自体は刺激の選択を含まず、むしろ、刺激の選択を引き起こす、あるいは、その準備をする機能を有しているという解釈である。以下で、この解釈を支持する根拠を挙げる。もしこの議論に説得力があれば、精度最適化は知覚から独立したプロセスになり、注意シフトに基づく論証は復活する。

第一の根拠は、精度の最適化は空間的注意をカバーしているという点である。本稿第2節で見たように、精度最適化モデルは、従来「注意」として扱われてきた多種多様な現象を統一的に扱うことが期待されていた。その現象の中には、高次認知領域からのバイアスを通じて特定の空間領域にスポットライトを当てる「空間的注意」も含まれる。ホーヴィは精度最適化としての注意はスポットライト型ではないと言うが、事実上彼が扱っている現象の中には、ポズナー・パラダイムに代表される空間的注意が含まれており、特定の空間領域の選択が精度最適化と見なされている(Hohwy, 2013, chap. 9)。空間的注意は空間領域の選択までが仕事であり、その後に領域内で刺激が与えられることになるので、「認知状態→注意→知覚」という図式が成立してしまい、認知的侵入可能性の条件は満たされない(Stokes, 2021, chap. 5)。もちろん、精度最適化モデルでは、性質や対象を基盤とした注意も扱われる。しかし、空間的注意が守備範囲に含まれる以上、このモデルでは、注意は刺激の選択を構成的に含んだ創発的性質として一般化できないことになる。

第二の根拠は、高次認知状態からのトップダウン・バイアスに基づく場合は、精度最適化は先行手掛かり(pre-cueing)に潜在的注意が引き起こされるのと同じメカニズムで理解できるという点である。先行手掛かりとは、注意を特定の領域や対象や性質に向けさせるた

めに、主体に手掛かりとなる刺激を事前に与えることである。先行手掛かりは、次にどのような刺激がターゲットとして与えられるかの態勢(baseline)を認知システムに用意させるものであり、ターゲットが与えられるよりも前に注意が向けられるので、注意は刺激の選択を行う役割を果たすことになる。何人かの論者は、先行手掛かりに訴えて認知的侵入可能性を否定している(Raftopoulos, 2009, 2017, 2019; Gatzia & Brogaard, 2017)。また、この解釈はバイアス競合モデルにも適用できる(cf. Zeimbekis & Raftopoulos, 2015, pp. 22–23; 西村, 2017, 63–64 頁)。このモデルでは、競合に負けた刺激は抑制されるが、これはどの刺激が感覚入力として与えられるのかが、受容野を巡るバイアス競合プロセスによって決まることを意味している。それゆえ、このプロセスを注意と同定すれば、注意は先行手掛かりによって刺激選択の準備をする機能を有していることになる。先行手掛かりに訴えて認知的侵入可能性を否定する論者は、精度最適化モデルについては明示的に論じていないが、以下の理由から、この戦略は同モデルにも適用できると考えられる。

第一に、精度最適化モデルの提唱者も、刺激の選択を引き起こすものとして精度最適化を描写している。例えば、フリストンは「この[隠された文脈の]表象を最適化することは刺激間の競合を導き(leads)、推測された文脈を刺激の存在とより整合的にする」(Feldman & Friston, 2010, p. 16 傍点は引用者による)と述べるが、これは精度最適化が刺激の選択を引き起こしていると読むことができる。第二に、予測誤差最小化モデルにおける行為推論は精度最適化が刺激の選択を準備する機能を果たしていることを強く示唆する。本稿第2節で見たように、予測誤差最小化モデルでは、知覚推論と行為推論は表裏一体のものとされる。この行為推論は意志的注意と結びつけられる(Hohwy, 2013, pp. 197–199)。意志的注意とは、ある性質や対象や領域に対して意図的かつ意識的に注意を向けることであり、内発的注意の一種(内発的注意は無意識的なものでありうる)として理解できる。主体は意志的注意による決定に基づいて特定の行為をとる。ホーヴィによれば、意志的注意は外界の仮説に合致する刺激(実際は、逆に予測誤差を生み出す刺激)を見つけるための選択的サンプリングを行うものである。

「ある空間領域への意志的注意...は、外界に関する仮説に適合するように外界が選択的にサンプリングされる状況と見なしうる。その仮説は高い精度で何かがその領域で生じるというものであろう。この仮説はその領域の予測誤差のゲインを上げ、もしそこで実際に何かが生じれば、それがすぐに検出される見込みを高めるであろう。」(Hohwy, 2013, p. 197)

この引用文は、精度最適化としての注意が特定の刺激を検出するように認知システムの態勢を整えさせる機能を果たしていることを示唆している。この態勢の準備が実際の刺激が選択される前のプロセスであることを踏まえると、態勢の準備までが精度最適化プロセスであり、注意は知覚とは独立のプロセスであることになる<sup>(18)</sup>。

### 5.3 反論と再反論

以上のように、精度最適化をプロセスとして同定することは可能であり、また、このプロセスは刺激の選択を引き起こすものとして統合的に理解できる。最後に、この理解に対して予測される反論を二つ取り上げて退けておく<sup>(19)</sup>。

第一の反論は、精度最適化は刺激の選択ではなく表象の選択であり、本稿のように精度最適化を刺激の選択を引き起こすプロセスとして解釈することは的外れではないかというものである。この反論はマルキの議論から導かれる(Marchi, 2020)。予測誤差最小化モデルの提案者であるデジモンとダンカンは、バイアス競合の結果選択されるのは刺激であると主張していた。それに対してマルキは、バイアス競合の結果選択されるのは、厳密に言えば刺激ではなく表象であると主張する(Marchi, 2020, chap. 5)。その理由は、実際に競合して選択されるのは認知システムにおけるプロセスや状態、すなわち、表象媒体であり、表象される刺激自体ではないからである。ここから、バイアス競合モデルにおける注意は、認知状態を対象とする認知状態という意味で「メタ認知(metacognition)」と呼ばれる(Marchi, 2020, chap. 6)。そして、精度最適化モデルはバイアス競合モデルの一形態であるという解釈に基づいて、精度最適化もメタ認知として特徴づけられる。この特徴づけから導き出されるのは、「(1)注意シフトが入力[刺激]選択の一形態であることの否定と、(2)注意が認知的侵入可能性を引き起こす一手段であることを擁護する議論形成の一助となること」である(Marchi, 2020, chap. 6)。なぜなら、精度最適化によって選択されるのは、もはや感覚入力としての刺激ではなくその刺激を表象する状態やプロセスであり、その状態やプロセスの決定は高次認知レベルからのトップダウン情報伝達と感覚レベルからのボトムアップ情報伝達によりなされるので、感覚入力は表象の決定に時間的に先行するものとなるからである。

この反論に対しては、たとえ知覚における選択的注意をメタ認知と見なしても、注意が果たす実質的な機能に変更はないという再反論が可能である。バイアス競合や精度最適化は、認知システムのあらゆる階層で生じている。たしかに、高次の階層では、それらプロセスの対象となるのは認知状態としての表象であり、その選択はトップダウンとボトムアップ両方の情報伝達に基づいているであろう。しかし、最下層の感覚レベルのバイアス競合や精度最適化が常にこれら両方の情報伝達に基づいているとは限らない。ポズナー・パ

ラダイムのような先行手掛かりに基づくタスクでは、先にトップダウンによる情報のみに基づいてバイアス競合や精度最適化が行われ、その後で刺激が与えられる。この場合は注意を通じて、その刺激の表象が結果として選択されることになる。たしかに、厳密に言えば、選択される対象は表象であるかもしれないが、実際はその表象を介して何を刺激として選択するかが決定されているので、注意シフトに基づく論証に影響はない。

第二の反論は、神経科学的な反論である。注意シフトに基づく論証を反駁する際に、ストークスが「四次視覚野に存在する選択的注意の神経的相関者や眼球のサッケードは知覚プロセスの一部である」と主張していたことを思い出されたい。もしこの主張が正しければ、注意は知覚プロセスから独立しているとは言えなくなる。そうすると、たとえ精度最適化をプロセスと同定しても、直接性からの論証は維持されて、認知的侵入可能性が成立するのではないか。

この反論については、「知覚から独立している」という表現の意味を検討することで再反論できる。この表現には二通りの解釈がある。一つは、「知覚を担う神経的相関者を含まない」という神経科学的解釈である。もう一つは機能的解釈であり、この場合は、神経的相関者が脳のどの場所に配置されているかは問題ではなく、当該の表現は「知覚の入力と出力に直接的な影響を与えない」、あるいは、「知覚のモジュール性を侵害しない」と換言される。本稿が採用するのは機能的解釈である。かつてストークスは、現状では、認知から知覚へのトップダウンの情報伝達に関する神経基盤が心的状態やその機能と正確に対応している十分な証拠はないと指摘していた(Stokes, 2013, p. 654)。この指摘は現在でも正しいと思う。たとえバイアス競合や精度最適化を担う神経プロセスが四次視覚野などの知覚領域に跨っているとしても、それが先行手掛かりによる刺激選択の準備として機能している限り、知覚のモジュール性は否定されず、認知的侵入可能性を認める必要もない<sup>(20)</sup>。

ここで、「知覚から独立している」という表現を機能的に解釈する本稿の提案をスティーヴン・グロスが提出している「注意命令(attentional command)」説と対比させておくことは有益である(Gross, 2017)。注意命令とは、「そこ(There)」や「今(Now)」のような命令と類比的に理解される、高次認知状態によって引き起こされた心的状態であり、先行手掛かりに基づく内発的注意を捉えようとしたものである。グロスによれば、この注意命令は心的状態(恐らく「表象」だとされる(Gross, 2017, p. 6))としては高次認知状態からの直接的影響を受けるが、その機能はどこに注意を向けるかの命令を与えることに尽きるので、注意の結果として生じる知覚内容には、高次認知状態の表象内容は直接的に反映されない。それゆえ、注意命令は「知覚の中(in perception)」にあるが、知覚状態の一部ではない。このように注意命令とその結果を峻別すれば、知覚への高次認知状態の影響は間接的なものに



留まることになる。このグロスの主張は、本稿が提案する機能的解釈と軌を一としている。

ただし、グロスの主張は、認知的侵入可能性の提唱者に譲歩し過ぎているように思われる。グロスは、注意命令が知覚の中の「表象」である可能性を認めている。注意命令が表象である場合、それが表象する内容は「そこ」などの指標詞的内容となる。だが、この「そこ」は高次認知状態の表象内容を反映したものであり、それによって刺激の存在する場所が選択されるならば、高次認知状態の内容が注意命令を介して知覚内容に直接的な影響を与えることになると思われる<sup>(21)</sup>。これは、「そこ」のような空間的注意ではなく、対象や性質を基盤とした注意の場合に、より明らかである。赤色を探索する先行手掛かりが与えられた場合、注意命令の内容は「赤色」となる（また、この内容は、被験者が持っている赤色に関する高次認知状態の内容によって直接的影響を受けたものである）。この注意命令の内容によって赤色の刺激が感覚入力として与えられるならば、高次認知状態と注意命令と知覚の各内容間に直接的な推論関係が生じてしまう。これはグロスにとっても本稿にとっても問題となる。まず、グロスは、認知的侵入の必要条件に知覚内容と認知内容の推論的關係（あるいは整合性）を含める、認知的侵入可能性の「意味論的定義」を採用している(Gross, 2017)。このような立場にとっては、上述の直接的な推論関係は認知的侵入を成立させることになり都合が悪い。次に、本稿にとっては、もし注意を表象と見なせば、認知状態と注意と知覚の各内容間に直接的な因果関係も生じることになるので、認知的侵入可能性の因果的定義の条件(2)が満たされてしまう。

このような事態を避けるためには、グロスとは異なり、注意を非表象的プロセス（非表象的作用）として知覚と認知の間に措定するべきであろう<sup>(22)</sup>。そもそも選択メカニズム自体を表象と同一視する必然性はない。上述したように、ホーヴィは注意を精度の予期という「高階の信念」を最適化するメカニズムとして捉えていたが、この最適化は表象に対する操作であり、それ自体も表象であると考えer十分な理由はない。同様に、マルキが唱えるメタ認知は「思考の思考」(Marchi, 2020, chap. 6)と表現されており、彼自身は思考を表象的プロセスとして理解しているようだが、ここでの「思考」は、実質的には他の認知状態に対する統制的プロセス（「メタ認知的統制(metacognitive regulation)」(Marchi, 2020, chap. 6)と呼ばれる）という意味で使われているに過ぎず、必ずしもメタ認知自体が表象的である必要はないと考えられる<sup>(23)</sup>。

## 6. 結論

本稿では、予測誤差最小化モデルで措定される精度最適化としての注意が、認知的侵入可能性を否定する「注意シフトに基づく論証」への反論となりうるのかを検討した。その

結果、精度最適化としての注意は注意シフトに基づく論証を反駁しないことが結論付けられた。精度最適化モデルをバイアス競合モデルの一形態と見なしても、これらモデルにおける注意を創発的性質と見なす十分な理由はない。バイアス競合としての注意を創発的性質と見なす理由は、バイアス競合が多様な要素によって多重実現されるからであった。しかし、バイアス競合は高次の包括的プロセスとして同定可能であり、この戦略は精度最適化にも適用可能である。さらに、そのプロセスは刺激の選択という機能を有するものとして整合的に理解可能であり、認知と知覚の中間プロセスとして措定される。

## 謝辞

本稿は JSPS 科研費 15K0199 の助成を受けたものである。

## 註

- (1) この呼び名は Stokes (2012) の「注意シフト解釈」に負う。
- (2) この特徴は「非等方性(anisotropy)」と呼ばれる(Fodor, 1983)。なお、知覚のモジュール性と認知的侵入可能性が両立可能であるか否かは、それぞれの特徴づけによる。Macpherson (2017) や Drayson (2017) は、予測誤差最適化モデルと知覚のモジュール性の両立可能性を主張している。本稿で採用する「知覚からの独立性」の特徴づけ(本稿第 5.3 節参照)では、両者は両立不可能となる。
- (3) 予測誤差最小化モデルは認知的侵入可能性を許容すると論じる哲学者は多い(e.g., Howhy, 2013, 2017; Lupyan, 2015; Marchi, 2020)。注意シフトに基づく論証以外の観点から、予測誤差最小化モデルにおける認知的侵入可能性を批判する議論には、Drayson (2017) や Vance & Stokes (2017) がある。
- (4) 以下の説明は、Howhy (2013), Wiese & Metzinger (2017), Ransom et al. (2020), 乾・阪口 (2020) に負う。
- (5) ただし、この統一的性格には反論もある(e.g., Ransom et al., 2020)。
- (6) ただし、これらのモデル間に論理的関係があるわけではない(Marchi, 2020, chap. 7)。
- (7) 以下の定式化は、Stokes (2015, p. 78) と Marchi (2020, chap. 3) の定義を参照している。なお、認知的侵入可能性の代表的な定義には他に「意味論的定義」と「帰結主義的定義」がある。因果的定義も含めて、どの定義もそれぞれ問題を抱えている。因果的定義と意味論定義の問題点については Stokes (2015) や Raftopoulos & Zeimbekis (2015) を、帰結主義的定義の問題点については西村 (2017) を参照されたい。
- (8) ここで念頭に置いているのは、後で言及される「潜在的注意(covert attention)」の処遇である。註(7)で触れたストークスの定義では、条件(2)の因果関係は(a)と(b)だけを満たすものとされているが、これだと潜在的注意が自動的にその条件を満たすことになりかねない。
- (9) ストークスは「直接性からの反論」に加えて、認知的侵入可能性の帰結主義的定義に基づく反論も提出している(Stokes, 2021, pp. 135–138)。註(7)で述べたように、この定義は問題を含むので本稿では扱わない。
- (10) 実は、ストークスの主張に反して、原理(1)と原理(2)は注意シフトに基づく論証で常に前提されているわけではない。例えば、チャズ・ファイアストーンとブライアン・ショールは、この論証を展開する際に対象を基盤とした注意や性質を基盤とした注意にも言及している(Firestone & Scholl, 2016, p. 12)、また、彼らは、視線の移動などの身体運動を含む「顕在的注意(overt attention)」でも、それを含まない潜在的注意でも、その役割は刺激の決定に収束すると主張している(Firestone & Scholl, 2016, p. 62)。なお、ストークスは潜在的注意を行為的注意、非意志的注意、意識下の注意と同列に扱っているが(Stokes, 2014, p. 25, fn. 40)、本稿第 3.3 節で取り上げるように、潜在的注意には意志的なものもあり、このような扱いに正当性はない。
- (11) 作業記憶モデル(Prinz, 2011)には、その意識に関する含意に批判があるので、本稿では扱わない(cf. Watzl, 2011)。
- (12) 予測誤差最小化モデルの擁護者には、全面的な知覚の認知的侵入可能性を認めて、知覚と認知の区別自体を否定する者もいる(e.g., Clark, 2013)。だが、この主張を否定する経験科学的証拠がある。例えば、ゲ

シユタルト的な知覚の体制化は、トップダウンのバイアスよりも早く検知される(van Leeuwen, 2015)。事実、予測誤差最小化モデルの擁護者の多くは、認知と知覚の区別を認めている。

(13) ホーヴィは注意によって認知的侵入可能性が生じる事例として、Carrasco, Ling & Read (2004)による実験を挙げている(Hohwy, 2013, pp. 135–137)。しかし、ストークスが指摘するように、この事例での注意は知覚レベルで説明できるので、この事例は認知的侵入可能性を示すものとは見なせない(Stokes, 2014)。

(14) この例はジョン・テイラーによる例を改変・援用したものである(Taylor, 2015)。

(15) 厳密には、精度最適化には主体がどれだけその環境に適応しているのかも影響する(Hohwy, 2017, p. 81)。

(16) 不注意や注意散漫も注意の一種であるという議論は、Irving (2016)や Henry (2019)が展開している。

(17) 実はこの点について、バイアス競合モデルの注意を創発的性質と解釈する論者は曖昧である。上で「ある人が注意を払うことは、その人の行動的・認知的パフォーマンスの因果的効力や持続や速さや正確性を説明するのに不可欠な役割を果たす」というモールとヘンリーの文言を引用したが、この文言は、注意を払うことによって刺激が速く正確に選択されることを肯定しているようにも読める。また、デジモンとダンカンも、バイアス競合モデルで空間的注意を扱っているが、後述するように、これは注意が刺激の選択を引き起こす事例として解釈できる(Desimone & Duncan, 1995)。

(18) 実際、ホーヴィは行為推論などで身体的変化が伴われる場合、それと連動して感覚入力が増えるために、認知的侵入可能性はほぼ生じなくなることを認めている(Hohwy, 2013, pp. 119–120)。その上で、彼は感覚入力が増える場合でも認知的侵入可能性を認めるように基準を緩めることを提案している。だが、注意によって感覚入力自体が増えるか否かは、知覚のモジュール性を巡る文脈で認知的侵入可能性の是非が論じられる際の論点そのものであり、そこを緩める提案は受け入れられない。また、ホーヴィは基準を緩めることで得られる長所を強調するが、Vance & Stokes (2017)はその長所を批判している。

(19) 以下で取り上げるもの以外の反論としては、精度最適化は認知システムの多くの階層でシナプス・ゲインを変化させるので、知覚を含めた認知システム全体を変化させるというクラークの議論がある(Clark, 2016b)。だが、この議論に対しては、ファイアストーンとショールが、(1)具体的な経験科学的証拠が不足している、(2)クラークの主張は認知と知覚の区別自体を解消するものだが、そのようなラディカルな主張を受け入れる必要はないという再反論を既に提出している(Firestone & Scholl, 2016, pp. 61–62)。(2)については本稿註(12)も参照されたい。

(20) アタナシオス・ラフトポウロスによる、初期視覚に的を絞った認知的侵入不可能性を擁護する議論(e.g., Raftopoulos, 2019)は、「知覚からの独立性」の神経生理学的解釈と機能的解釈を混同したものと見なすことができよう。なお、彼はバイアス競合モデルにおける注意を創発的性質と見なすアイデアは放棄していないと考えられる。

(21) この点は、指示詞的内容を用いた概念主義擁護の議論と類比的に理解できるであろう。概念主義者によれば、知覚内容に組み込まれる指示詞的内容(例えば、「あれ(that)」)は概念的内容であり、他の表象内容との推論関係を可能にする(McDowell, 1994)。

(22) 実際、モールのようにバイアス競合モデルにおける注意を「情報選択の仕方」として副詞的に解釈する場合は、注意を表象と見なす立場はより苦しくなるであろう。知覚の哲学における標準的な副詞理論では、「感覚の仕方」として副詞的に理解される知覚経験は表象的性質を持たない(Fish, 2010, chap. 3)。

(23) 「直接性からの論証」の擁護者にとっては、メタ認知としての注意が思考(高次認知状態)であることを強調し過ぎれば、注意を知覚プロセスの一部と見なせなくなるという問題もある。

## 文献

- Carrasco, M., Ling, S. & Read, S. (2004). 'Attention Alters Appearance,' *Nature Neuroscience*, 7, 308–313.
- Clark, A. (2013). 'Whatever Next? Predictive Brains, Situated Agents, and the Future of Cognitive Science,' *Behavioural and Brain Sciences*, 36 (3), 181–204.
- (2016a). *Surfing Uncertainty: Prediction, Action, and the Embodied Mind*, Oxford: Oxford University Press.
- (2016b). 'Attention Alters Predictive Processing,' in Firestone & Scholl (2016, pp. 23–24).
- Clarke, S. (2021). 'Cognitive Penetration and Informational Encapsulation: Have We Been Failing the Module?,' *Philosophical Studies*, 178, 2599–2620.
- Desimone, R. & Duncan, J. (1995). 'Neural Mechanisms of Selective Visual Attention,' *Annual Review of Neuroscience*, 18, 193–222.
- Drayson, Z. (2017). 'Modularity and the Predictive Mind,' in Metzinger & Wiese (2017).

- Duncan, J. (2006). 'Brain Mechanisms of Attention,' *Quarterly Journal of Experimental Psychology*, 59, 2–27.
- Feldman, H. & Friston, K. J. (2010). 'Attention, Uncertainty, and Free-energy,' *Frontiers in Human Neuroscience*, 4, 1–23.
- Firestone, C. & Scholl, B. J. (2016). 'Cognition Does Not Affect Perception: Evaluating the Evidence for 'Top-down' Effects,' *Behavioural and Brain Sciences*, 39, 1–77.
- Fish, W. (2010). *Philosophy of Perception: A Contemporary Introduction*, New York, Routledge. 邦訳：ウィリアム・フィッシュ(2014, 山田圭一監訳, 『知覚の哲学入門』, 勁草書房.)
- Fodor, J. A. (1983). *The Modularity of Mind*, Cambridge, MA: MIT Press.
- (1988). 'A Reply to Churchland's 'Perceptual Plasticity and Theoretical Neutrality,' in *Philosophy of Science*, 55 (2), 188–198.
- Friston, K. J. (2009). 'The Free-energy Principle: A Rough Guide to the Brain?,' *Trends in Cognitive Sciences*, 13 (7), 293–301.
- (2010) 'The Free-energy Principle: A Unified Brain Theory?,' *Nature Reviews Neuroscience*, 11(2), 127–138.
- Friston, K. J., Mattout, J. & Kilner, J. (2011). 'Action Understanding and Active Inference,' *Biological Cybernetics*, 104, 137–160.
- Friston, K. J. & Stephan, K. E. (2007). 'Free-energy and the Brain,' *Synthese*, 159 (3), 417–458.
- Gatzia, D. E. & Brogaard, B. (2017). 'Pre-cueing, Perceptual Learning and Cognitive Penetration,' *Frontiers in Psychology*, 8, article 739.
- Gross, S. (2017) 'Cognitive Penetration and Attention,' *Frontiers in Psychology*, 8, article 221.
- Henry, A. (2019). *An Agent of Attention: An Inquiry into the Source of Our Control*, Dissertation, University of Toronto.
- Hohwy, J. (2013). *The Predictive Mind*, Oxford: Oxford University Press. 邦訳：ヤコブ・ホーヴィ(2021, 佐藤亮司監訳, 『予測する心』, 勁草書房.)
- (2017). 'Priors in Perception: Top-down Modulation, Bayesian Perceptual Learning Rate, and Prediction Error Minimization,' *Consciousness and Cognition*, 47, 75–85.
- 乾敏郎・阪口豊 (2020). 『脳の大統一理論: 自由エネルギー原理とはなにか』, 岩波書店.
- Irving, Z. C. (2016). 'Mind-wandering Is Unguided Attention: Accounting for the "Purposeful" Wanderer,' *Philosophical Studies*, 173, 547–571.
- Kastner, S., & Ungerleider, L. G. (2000). 'Mechanisms of Visual Attention in the Human Cortex,' *Annual Review of Neuroscience*, 23, 315–341.
- Koralus, P. (2014). 'The Erotetic Theory of Attention: Questions, Focus and Distraction,' *Mind & Language*, 29 (1), 26–50.
- Levy, Y. (2019). 'Is Attending a Mental Process?,' *Mind & Language*, 34, 283–298.
- Liu, C.-H., Tzeng, O. J. L., Hung, D. L., Tseng, P. & Juan, C.-H. (2012). 'Investigation of Bistable Perception with the "Silhouette Spinner": Sit Still, Spin the Dancer with Your Will,' *Vision Research*, 60, 34–39.
- Lupyan, G. (2015). 'Cognitive Penetrability of Perception in the Age of Prediction: Predictive Systems Are Penetrable Systems,' *Review of Philosophy and Psychology*, 6, 547–569.
- Macpherson, F. (2017). 'The Relationship between Cognitive Penetration and Predictive Coding,' *Consciousness and Cognition*, 47, 6–16.
- Marchi, F. (2020). *The Attentional Shaping of Perceptual Experience: An Investigation into Attention and Cognitive Penetrability*, Springer. Kindle ver.
- McDowell, J. (1994). *Mind and World*, Cambridge, MA: Harvard University Press. 邦訳：ジョン・マクダウエル (2012, 神崎繁・河田健太郎・荒畑靖宏・村井忠康訳, 『心と世界』, 勁草書房.)
- Metzinger, T. & Wiese, W. (Eds.) (2017). *Philosophy and Predictive Processing*, Frankfurt am Main: MIND Group, Retrieved, 2023/ 02/ 15, from <https://predictive-mind.net/papers/modularity-and-the-predictive-mind>
- Mole, C. (2011a). *Attention Is Cognitive Unison: An Essay in Philosophical Psychology*, Oxford: Oxford University Press.
- (2011b). 'The Metaphysics of Attention,' in Mole, Smithies & Wu (2011, pp. 60–77).
- (2012). 'Attention,' in E. Margolis, R. Samuels & S. Stich (Eds.), *The Oxford Handbook of Philosophy of Cognitive Science* (2012, pp. 201–221), Oxford: Oxford University Press.
- (2015). 'Attention and Cognitive Penetration,' in Zeimbekis & Raftopoulos (2015, pp. 218–238).
- Mole, C. & Henry, A. (2022). 'What Is Attention? Adverbialist Theories,' *Wiley Interdisciplinary Reviews: Cognitive Science*, e1588, 1–8.
- Mole, C., Smithies, D. & Wu, W. (Eds.) (2011) *Attention: Philosophical and Psychological Essays*, Oxford: Oxford

University Press.

- 西村正秀 (2017). 「運動知覚の認知的侵入可能性」, 『彦根論叢』, 412 号, 52–66.
- O’Conner, T. (2021). ‘Emergent Properties,’ in E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*, Retrieved, 2023/02/15, from <https://plato.stanford.edu/archives/win2021/entries/properties-emergent/>
- Prinz, J. J. (2011) ‘Is Attention Necessary and Sufficient for Consciousness?’, in Mole, Smithies & Wu (2011, pp. 174–204).
- Pylyshyn, Z. W. (1999). ‘Is Vision Continuous with Cognition? The Case for Cognitive Impenetrability of Visual Perception,’ *Behavioural and Brain Sciences*, 22, 341–423.
- Raftopoulos, A. (2009). *Cognition and Perception: How Do Psychology and Neural Sciences Inform Philosophy?*, Cambridge, MA: MIT Press.
- (2017). ‘Pre-cueing, the Epistemic Role of Early Vision, and the Cognitive Impenetrability of Visual Perception,’ *Frontiers in Psychology*, 8, article 1156.
- (2019). *Cognitive Penetrability and the Epistemic Role of Perception*, Palgrave Macmillan.
- Raftopoulos, A. & Zeimbekis, J. (2015). ‘The Cognitive Penetrability of Perception: An Overview,’ in Zeimbekis & Raftopoulos (2015, pp. 1–56).
- Ransom, M., Fazelpour, S., Markovic, J., Kryklywy, J., Thompson, E. T. & Rebecca M. T. (2020). ‘Affect-biased Attention and Predictive Processing,’ *Cognition*, 203, article 104370.
- Stokes, D. (2012) ‘Perceiving and Desiring: A New Look at the Cognitive Penetrability of Experience,’ *Philosophical Studies*, 158, 477–492.
- (2013). ‘Cognitive Penetrability of Perception,’ *Philosophy Compass*, 8/7, 646–663.
- (2014). ‘Cognitive Penetration and the Perception of Art,’ *Dialectica*, 68 (1), 1–34.
- (2015). ‘Toward a Consequentialist Understanding of Cognitive Penetration,’ in Zeimbekis & Raftopoulos (2015, pp. 75–100).
- (2021). *Thinking and Perceiving: On the Malleability of the Mind*, New York: Routledge.
- Taylor, J. (2015). ‘Against Unifying Accounts of Attention,’ *Erkenntnis*, 80 (1), 39–56.
- van Leeuwen, C. (2015). ‘Hierarchical Stages or Emergence in Perceptual Integration?’, in J. Wagemans (Ed.), *The Oxford Handbook of Perceptual Organization* (2015, pp. 969–988), Oxford: Oxford University Press.
- Vance, J. & Stokes, D. (2017). ‘Noise, Uncertainty, and Interest: Predictive Coding and Cognitive Penetration,’ *Consciousness and Cognition*, 47, 86–98.
- Watzl, S. (2011). ‘Review of *Attention Is Cognitive Unison: An Essay in Philosophical Psychology*,’ *Notre Dame Philosophical Reviews*, Retrieved 2023/02/15, from <https://ndpr.nd.edu/>
- (2017). *Structuring Mind*, Oxford: Oxford University Press.
- Wiese, W. & Metzinger, T. (2017). ‘Vanilla PP for Philosophers: A Primer on Predictive Processing,’ in Metzinger & Wiese (2017).
- Zeimbekis, J. & Raftopoulos, A. (Eds.) (2015). *The Cognitive Penetrability of Perception: New Philosophical Perspectives*, Oxford: Oxford University Press.

[滋賀大学教授・哲学]