

[Title]

サリエンスをアフォーダンスとして捉え直す：能動的推論の視点からの、精神病症状の異常サリエンス仮説への示唆

Salience as affordance: implications for aberrant salience hypothesis of psychosis from active inference perspective

[Authors]

吉田 正俊¹、宮園 健吾²、西尾 慶之³、山下 祐一⁴、鈴木 啓介¹

[Affiliations]

¹ 北海道大学人間知・脳・AI 研究教育センター、² 北海道大学文学研究院、³ 東京都立松沢病院 精神科・脳神経内科、⁴ 国立精神・神経医療研究センター 神経研究所

[責任著者]

吉田 正俊
myoshi@chain.hokudai.ac.jp

[Abstract]

サリエンスという言葉には多義性がある。かたや「知覚サリエンス」では知覚的に周りと比べて目立つという特性を表しており、かたや「動機サリエンス」では物体に情動価を付与させる働きとして捉えられている。本論文では「知覚サリエンス」と「動機サリエンス」との関係を明確にすることを目的とする。このことは精神医学において提唱されている ASH を精緻化できるという貢献があることが期待される。精神病症状の異常サリエンス仮説 aberrant salience hypothesis of psychosis (ASH)は精神病症状の発達から回復までを統合的に説明したものとして重要性が高い。しかし重要な問題点としてサリエンスという言葉が曖昧だという点がある。そこでこの論文ではまず現状でのサリエンスの定義を明確にするため、知覚サリエンスと動機サリエンスに分けて考える。そしてこのサリエンスをアフォーダンスとして捉える考え方がふたつの方向性（心理学の哲学と情報理論的な脳のモデル(能動的推論))から支持されることを示す。このような視点からサリエンスを捉え直すことで、サリエンスは epistemic affordance, exploitive affordance, aversive affordance に分類できる。このようにサリエンスを捉え直すことによって、ASH はより精緻なものとなる。つまり ASH においては、epistemic affordance と aversive affordance 両方の亢進を考慮する必要がある。ここで前者は対象への接近を動機づけるが、後者は対象の回避を動機づける。このようなコンフリクトが患者を苦しめ、なんらかの認知的な解決を必要とする。このことは ASH での「サリエンスの誤帰属」という言葉からは見えてこない。最後に以上のことと脳部位との対応について議論し、さらに本論文で提案される考えの応用範囲について議論する。

[Keywords]

Schizophrenia, delusion, visual salience, incentive salience, active inference, free-energy principle

[Introduction]

サリエンスとはなにか

サリエンス salience という言葉は多義的に使われてきた。かたや視覚科学や computer vision の分野では、視覚サリエンス visual salience という言葉で、ある視覚的なアイテムがその周辺と異なっているために突出してみえる、目立つ、という主観的な知覚特性のことを表している (Itti, 2007)¹。このような意味での salience は視覚には限らず、聴覚などの他の感覚モダリティでも想定可能であるので、この論文では、視覚のみに限らない総称として知覚サリエンス perceptual salience と呼ぶことにする。

一方で情動や addiction の研究の分野では、incentive salience という言葉で、ある刺激に魅力的な、欲しいと動機づけるような特性を与える心理的なプロセスのことを表している (Robinson and Berridge, 1993)²。Incentive salience の概念は元々 Kent Berridge によって、欲求の情動的側面 (=want) と主観的な快楽としての側面 (=like) とを区別するために導入された。前者の want (= incentive salience) と後者の like が独立した過程であり、関わる脳部位も異なることが明らかになっている。

Incentive salience は正の情動価をもつものを指すが、負の情動価を持つ fearful / aversive salience (ここでは aversive salience の語を使う) を合せて motivational salience と呼ぶようになっている (Berridge, 2018)。そこでこの論文では、正または負の情動価をもつサリエンスの総称として、動機サリエンス motivational salience の語を用いることにする。以上をまとめて図に示した (Figure 1)。

知覚サリエンスと動機サリエンスの研究はそれぞれがべつのグループによって進められてきたという歴史的経緯があるため、両者がどのくらい関連があるのかという議論はほとんど進められてこなかった。哲学においても、近年になってやっとサリエンスに関する著書 (Salience: A Philosophical Inquiry (Archer, 2022)) が現れるという状況にある。

この論文でわれわれは知覚サリエンスと動機サリエンスに深い関連があることについて議論する。その動機の一つは、知覚サリエンスと動機サリエンスの関連性を明確にすることによって、精神医学において精神病症状 psychosis を説明するために

提唱された、「異常サリエンス仮説 Aberrant Salience Hypothesis」をより精緻なものにできると考えるからだ。

精神病症状の異常サリエンス仮説

ここで Kapur による異常サリエンス仮説 Aberrant Salience Hypothesis (以降 ASH と略する) (Kapur, 2003) について説明しておく。ASH に拠れば、統合失調症その他の精神障害で生じる幻覚や妄想などの精神病症状 psychosis には、ドパミン分泌亢進に伴って生じる異常なサリエンスが主要な役割を果たしている³。Kapur によればサリエンスの帰属とは、出来事や思考が、報酬と罰との結びつきによって、注意をひき、行為を駆り立て、目標指向行動な行動に影響を

¹ “the distinct subjective perceptual quality which makes some items in the world stand out from their neighbors and immediately grab our attention.” (Itti, 2007)

² “Incentive salience is a psychological process that transforms the perception of stimuli, imbuing them with salience, making them attractive, 'wanted', incentive stimuli.” (Robinson and Berridge, 1993)

³ (Kapur, 2003) から引用すると、ドパミン系の亢進が「不適切なサリエンスと動機的重要度を外的及び内的な刺激に付与する (“an exaggerated release of dopamine leads to the assignment of inappropriate salience and motivational significance to external and internal stimuli”)」。そしてそれが「新規で困惑する状態 novel and perplexing state」を引き起こす。そして「気づきと情動性の亢進 heightened awareness and emotionality のステージ」が状況を make sense するための drive となる。

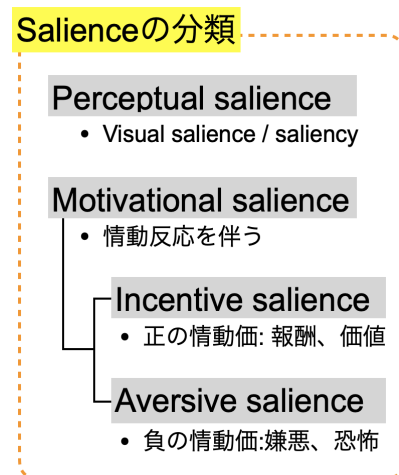


Figure 1

及ぼすプロセスに他ならない。

(Kapur, 2003)の中ではこのサリエンスはドパミン神経によって媒介された動機サリエンスを想定している。統合失調症前駆期において認められる気付き亢進や妄想知覚などの病的知覚体験は、日常的な知覚体験に異常な動機サリエンスが付与されることで生じる現象であると説明できる。統合失調症で高頻度に認められる固定化した内容を持つ幻覚や妄想は、体験者がこの困惑させられる新奇な体験を make sense するために行った認知的説明 cognitive elaboration によって生み出されると考えることができる。このようなケースにおける動機サリエンスは、知覚サリエンスと同様に知覚体験に付与されるものと捉えることができる。では、動機サリエンスと知覚サリエンスの機能は明確に区別可能なのだろうか？それとも両者の間には連続性があるのだろうか？

また、後年の研究からは、このようなサリエンスの変容がかならずしも動機サリエンスに限定されていないことが示唆されている。たとえばサリエンスの変容を定量化する方法として Aberrant Salience Inventory (ASI) (Cicero et al., 2010) という質問リストがある。ASI は 29 個の質問から成り立つが、これらの質問から抽出される 5 つの要素が挙げられている。このうち Factor 1 の Increased Significance と Factor 4 の Heightened Emotionality は motivational salience の亢進として捉えられるが、Factor 2 の Senses Sharpening のように perceptual salience の亢進として捉えられるものもある。

じっさいに統合失調症においては、これまでによく知られている報酬系や知覚(聴覚のみならず視覚でも)での変容があるだけでなく、動機サリエンスの変容(Roiser et al., 2009, 2013)と知覚サリエンスの変容(Yoshida et al., 2015)の両方が報告されている。このことから、ASH における「異常なサリエンスの帰属」には動機サリエンスと知覚サリエンスの両方が関与する可能性がある。

異常サリエンス仮説については、その概念的な曖昧さや、理論的な不明瞭さが指摘されてきた (Ratcliffe and Broome, 2022)⁴。異常サリエンス仮説は、より厳密で正確な説明によって補充されるまでの一時的な placeholder として理解されるべきかもしれない。また Ratcliffe & Broome が指摘するように、単純な異常サリエンス仮説はきめの荒い統合失調症の理解しか与えてくれず、統合失調症患者の経験の微妙な側面や細かな違いを説明できないかもしれない。例えば、統合失調症において、患者はある salient な外的対象に対して (知覚的には) 強く注意がひかれる一方、その対象が (情動的には) 恐ろしく、あるいはよそよそしく感じることもあるかもしれないのだ。

従来の異常サリエンス仮説に対するこのような問題点を克服するためにも、知覚サリエンスと動機サリエンスとの関連性の明確化という本稿のプロジェクトは有意義なものとなるだろう。そこで本論文では以下の構成で議論をすすめる。

- 1) まず問題設定 1 として、知覚サリエンスが感覚入力のみで決まる特性であるという考えについて見直しを図る。
- 2) このため、サリエンスをアフォーダンスとして捉え直すことができるという本論の基本テーゼの理論的な概要を解説する。
- 3) 計算論的脳科学の分野で提唱されている「能動的推論」の枠組みからも、サリエンスをアフォーダンスとして捉える考えが支持されることを示す。
- 4) つぎに問題設定 2 として、知覚サリエンスと動機サリエンスの関係を問う。
- 5) 「能動的推論」の枠組みからは、知覚サリエンスと動機サリエンスが「同じ通貨を使う」という意味で統一的に扱えることを示す。
- 6) 問題設定 3 として、知覚サリエンスと動機サリエンスをアフォーダンスとして分類する考えを提案する。そこでは対象に対して接近 approach するか、回避 avoid するかという点からサリエンスが分類しなおされる。
- 7) この視点から、ASH をより精緻にすることを試みる。ASH においては知覚的サリエンスと動機サリエンスの両方が亢進しているように見えるが、アフォーダンスとして捉える立場からは、負の情動価を持ち avoid したい aversive affordance と、でも make sense するために理解するために approach したい epistemic affordance というふたつの矛盾する要素があることを指摘する。
- 8) 最後にこれらの知見と脳部位との対応について議論し、さらに本論文で提案される考え

⁴ “It is not entirely clear from this whether or not an experience of aberrant salience should be regarded as perceptual in nature” (Ratcliffe and Broome, 2022)

の応用範囲について議論する。

[問題設定 1: 知覚サリエンスをアフォーダンスとして捉える]

まず第 1 の問題について議論することにしよう。知覚サリエンスは上記の Itti の定義を含めて一般的には、感覚情報によって規定された心理的特性として捉えられてきた。しかしこれはほんとうに正しいだろうか？ サリエンスは主体の行動によって規定される側面があるのではないだろうか。

ここで手がかりとなるのは、アフォーダンスの概念だ。ギブソンによればアフォーダンスとは環境に実在するものであるとともに、主体にとっての生態学的価値によって規定されるものとされている。知覚サリエンスについても同じように捉えられるのではないだろうか。つまり、生態学的価値という観点からは、知覚サリエンスのことを「注意をアフォードするもの」として捉えることができるのではないだろうか。問題設定 1 では、この問題について議論してゆく。

なお、サリエンスを生態学的価値から捉えるという考え方は、問題設定 2 の知覚サリエンスと動機サリエンスの連続性を議論するところでも、重要な役割を果たすことになる。

知覚的サリエンスと注意の関係

「知覚サリエンスが注意をアフォードする」ということを説明するために、まず注意とは何か、から始める必要がある。ここでは空間のある位置に向けられた「空間的注意」に絞って議論しよう。認知心理学の分野において空間的注意は、その生起過程の違いからボトムアップ性注意とトップダウン性注意に分けられる⁵。ボトムアップ性注意では、視覚刺激そのものの属性によって注意が誘引される。トップダウン性注意は、ある視覚刺激が現在のゴールや報酬と結びついているときに該当する。低レベルの視覚特徴、たとえば輝度や色や動きによって周りとは比べて目立つ、というときの視覚サリエンスは、ボトムアップ性注意を誘引する視覚刺激に該当する⁶。ボトムアップ性の注意は「注意を掴まれる」「attention-grabbing」と表現することができる。トップダウン性の注意は「注意を払う」「paying attention」と表現することができる。視覚サリエンスの理論(Itti and Koch, 2001)によると、視覚情報は脳内で物体認知などの処理と並行して、視覚サリエンスの検出の処理が行われる。視覚サリエンスは、視野のどの位置に「目立つ、目を惹くものがあるか」という情報について 2 次元の空間マップ(サリエンシー・マップ)として脳内に保持される。このような視覚サリエンスの情報をもとに空間的注意が操作される。あるときはサリエンスの高いところに視線を向ける、という行為(overt attention と呼ばれる)が行われ、またあるときは視線を動かさずに、周辺視野にある、サリエンスの高いところに注意を向けるという心的行為(covert attention)が行われる。このようにして知覚的サリエンスが空間的注意を導く guide、という関係にある。

サリエンスの基本テーゼ: サリエンスをアフォーダンスとして捉える

ではつぎに知覚的サリエンスをアフォーダンスとして捉えることができるか議論しよう。そのためには、「知覚サリエンスが空間的注意という行為または心的行為をアフォードする」という言い方が意味を成すかを検討する必要がある。そこでここでは哲学で用いられる議論を援用

⁵ 注意を誘引するものが外的な刺激そのものの特性である exogeneous な注意(ボトムアップ性注意に対応)と、内的なゴールによって誘引される endogeneous な注意(トップダウン性注意に対応)というのが元々の Posner の分類(Posner, 1980)。

⁶ なお、視覚的注意研究のもう一つの流れとして、Ann Triesman の特徴統合理論がある(Treisman and Gelade, 1980)。特徴統合理論では(輝度や傾きなどの)視覚特徴ごとにどこが「ポップアウト」するかを決める前注意過程があって、それが意識的な注意に入ってくるというモデルになっている。特徴統合理論での前注意過程はボトムアップ性注意にだいたい対応し、特徴統合理論での注意はトップダウン性注意にだいたい対応する。そして視覚サリエンスのモデル(サリエンシー・マップ)はこの特徴統合理論から発展して作られている。以上のような歴史的背景があるため、「注意」の語の使用の際には混同しないように気をつける必要がある。

して、「サリエンスを一種のアフォーダンスとして捉える」という基本テーゼを、以下のように定式化する。

基本テーゼ：対象 O（あるいはその性質 P）がエージェント X にとって“salient”あるいは“attention-grabbing”である iff. O（あるいはその性質 P）が X の（視覚的）注意（あるいは注意を伴う諸行動）を afford している

基本テーゼはサリエンスが持ついくつかの基本的な特徴をうまく説明できる。Salient な対象は我々の注意を掴む grab attention が、しかし、これは salient な対象が注意を強制するという意味ではない（すなわち、salient な対象に注意を払わない pay attention ことは可能である）。加えて、salient でない対象に注意を払うことができないという意味でもない（例えば、何かの目的により、salient ではない詳細部分に注意を払うことは可能である）。これらの特徴はアフォーダンスの特徴に類似している。マグカップの取手は掴むという動作を afford するが、しかし、これは取手が掴むという動作を強制するという意味ではない（すなわち、取手を掴まないことは可能である）。加えて、掴みのアフォーダンスを持たない対象を掴むことができないという意味でもない（例えば、何かの目的により、掴みのアフォーダンスを持たない部分、例えばマグカップの底面を強引に掴むことは可能である。）

なお、基本テーゼは、対象がなんらかのアフォーダンスを持つ = salient であるという主張ではない。マグカップの取手は掴むという動作を afford するが、取手は必ずしもとりわけ salient ではないかもしれない。

基本テーゼが意味をなすためには、そもそも「注意についてのアフォーダンス」というアイデアが意味をなすのでなければならない。そのためには注意を一種の行為、すなわち心的行為とみなし、心的行為にまで適用できる一般性のあるアフォーダンスの理論を採用する必要がある。McClelland の Mental Affordance Hypothesis (McClelland, 2020) はまさにその様なアフォーダンスの理論となっている。最も一般的なレベルにおいてアフォーダンスは以下のように定義される。

AFFORDANCE: S が Φ することを x が afford する iff. S が Φ することの機会 opportunity を x が提供する。

- 知覚要件： Φ することを x が afford することを S が知覚する。
- 促進要件： Φ することを x が afford することを S が知覚することが、S が Φ することを促進する（すなわち、自動的に備える）。

McClelland のアフォーダンスは身体的な行為（掴む、投げる、等々）と心的行為（想像する、推理する、等々）の双方に適用される、アフォーダンスの一般理論である。この図式を注意に適用すると以下が得られる。

ATTENTION AFFORDANCE: S が注意することを x が afford する iff. S が (x に) 注意することの機会を x が提供する。

- 知覚要件：(x に) 注意することを x が afford することを S が知覚する。
- 促進要件：(x に) 注意することを x が afford することを S が知覚することが、S が (x に) 注意することを促進する（すなわち、自動的に備える）。

McClelland によるアフォーダンスの定義に従えば、知覚サリエンスについて、その対象へ注意を向けるという心的行為(周辺視野への注意:covert attention)および行為(視線移動による注意:overt attention)を afford するものとして捉えることが可能であることがわかる。

能動的推論とは

サリエンスをアフォーダンスとして捉えるという考え方は、脳を情報理論的に捉える立場である能動的推論 active inference の枠組みからも支持される。ここではまず能動的推論とはなにか説明しよう。そのためには自由エネルギー原理 free energy principle の説明が必要だ。英

UCLのKarl Fristonは、知覚と学習と行動とを統一的に説明する、計算理論的なフレームワークとして「自由エネルギー原理」を提唱している(Friston, 2010)。自由エネルギー原理においては、われわれ人間を含む生物の適応的活動を perception-action cycle として捉える(Figure 2)。

まずわれわれ生物は(たとえば網膜のような)感覚器からの入力を元に信念を更新し、それを(たとえば眼球運動を行う外転筋のような)効果器を通して世界に働きかけ、感覚器の入力を変化させるという、perception-action cycle を想定する。感覚器によって感受した情報と世界についてのモデルからその生物にとっての外界の状態を推定する(知覚)。そしてそれに基づいて世界についてのモデルを更新する(学習)。そのモデルに基づいて外界に介入する(行動)。このとき生物が持っている、外界の状態についての「信念」と世界についての知識である「生成モデル」から、「変分自由エネルギー」が定義できる。適応的な生物ではこの変分自由エネルギーが下がる方向に時間発展する、という想定が「自由エネルギー原理」だ。

自由エネルギー原理においては、行動を変分自由エネルギーを下げるものとして捉える。つまり行動は、周囲の環境をより予測しやすい状態に変化させることで、生物にとってより望ましい感覚入力を得ることが可能になる。このような、世界を理解するための推論として行動を捉える考え方を能動的推論と呼ぶ。

能動的推論から見た視覚サリエンス

では能動的推論において視覚サリエンスはどのように扱われているか。実生活でのわれわれの視覚は、たえず眼や頭を動かしながら視覚シーンを理解するということを実現している。ここからは視線移動を伴う視覚を例に説明してみよう。われわれ人間を含む霊長類では、視野の中心のみ視力が高いため、視線を移動させる眼球運動(サックードと呼ぶ)を繰り返すことによって、視覚シーンの理解が可能となっている。

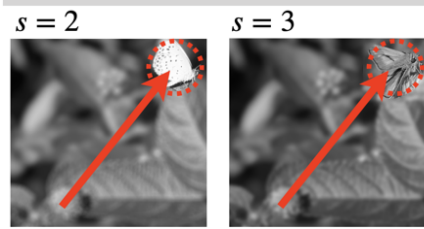
ここで視覚サリエンスの高いところに目を向けるという行動について考えてみよう。Figure 3 左の例では、左下(赤丸)の位置を注視している。このときわれわれの周辺視野はぼやけて見えるので、右上にある物体がなにであるかは未確定だ。

現在: 左下を見ている。
s=1を観測している。



現在のこの部分の
視覚サリエンス

将来: 右上を見たときに
観測されるsの予測



a_1 (右上を見る)という行動をした
ときに期待されるMI(a_1)

$$G(a) = -\mathbb{E}_s [\text{Bayesian surprise}(a)] \\ = -MI(a)$$

Figure 3

的にどれだけ視覚シーンをより正確に推定できるかを、相互情報量 $MI(a_1)$ として評価することができる。他の行動についても同様に相互情報量 MI を評価できる。このとき視覚サリエンスが

⁷ これが「生成モデル」の内実であり、「生成モデル」とは行動が感覚入力をどのように変えるかについての技能知であって、命題的な知識ではない。よって、この「予想」は、意識的な予想である必要はない。

高い場所とは、これらのまだ行っていない行動の中で、情報獲得の大きい場所に向かうものである。そしてこのとき $MI(a)$ が最大化されることで、変分自由エネルギーの期待値 $G(a)$ の最小化が実現される。これが能動的推論の枠組みから捉えた視覚サリエンスだ。

さてこのように視覚サリエンスを捉え直してみるならば、視覚サリエンスとは、従来考えられてきたような、画像のピクセルごとの特性というよりは、いま注視していない場所についての agent が可能な行動についての特性、であるといえる。この捉え方は、[サリエンスの基本テーゼ]で議論した、視覚サリエンスをアフォーダンスの一種として捉える考えと整合的になっている。

より正確に言うなら、これは視覚サリエンスを、空間的なものから時間的驚き(サプライズ)まで含めて拡張した、ベイジアン・サプライズ(Itti and Baldi, 2009)として捉えたものになる。このような意味でのベイジアン・サプライズでは、「空間的に、時間的に目立つ」だけでなく、「興味を引く、知りたいと思う」という意味づけがなされている。

視覚サリエンスが行動と密接に関連していることは、われわれを含む霊長類とは異なって、中心視野と周辺視野に視力の差のない仮想的な生物を考えてみれば、さらに納得がいくはずだ。つまり、そのような動物では、視線を移動させることは視覚シーンのより正確な理解には寄与できない。そして、このような生物では視覚サリエンスという概念自体が無意味になる。つまり、視覚サリエンスという概念は、われわれの網膜センサーが中心視野と周辺視野で感度の違いを持っているという進化によるデザインと不可欠に関わっている。

このように視覚サリエンスを捉えることで、空間的注意との関係も明確になる。空間的注意とはいま利用可能な感覚入力についてのプロセスであるのに対して、視覚サリエンスとはまだ行っていない行動によって将来利用可能となる感覚入力についてのプロセスである、という点で両者は明確に異なっている(Parr and Friston, 2019)⁸。しかし両者は perception-action cycle の一部として、外界の状態についての信念をより正確にするために協調して働く。つまり現在空間的注意を向けていることを元にして、次どこに空間的注意を向けるかを定めるために、知覚サリエンスが評価される。このような関係にあることが、前述の「知覚サリエンスは空間的注意をアフォードする」ことの内実となっている。注意を overt な注意(眼球運動)と covert な注意に分けて扱うならば、知覚サリエンスは overt な注意という行為と covert な注意という心的行為の両方をアフォードしている、と理解することができる。

注意の前運動理論

注意と視覚サリエンスが密接に関連しながら perception-action cycle を回しているという考えは、視覚が受動的に網膜像のコピーを作るものであるというよりはむしろ、行動(この場合サックード)でアクティブにサンプルされるものである、とする active vision の考えとも整合性が高い(Yoshida and Taguchi, 2018)。そしてさらに「注意の前運動理論」 premotor theory of attention とも整合的である。注意の前運動理論とは、注意と行動は同じ制御メカニズムを共有しているというものである(Rizzolatti and Craighero, 2010)。このことは眼球運動の実行(overt attention)と周辺視野へ注意を向ける内的行為(covert attention)それぞれに係る脳部位が大幅にオーバーラップしている(Corbetta et al., 1998)という点からも支持される。また、動物で前頭眼野を電気刺激すると眼球運動が引き起こされるが、刺激を弱くすると眼球運動を起こす代わりに、眼球運動と同じ方向に注意が向く(Moore, 2001)。しかもこのとき注意に関わる視覚連合野の神経応答は増強される。つまり脳内で gain control がじっさいに行われているという証拠がある。

このようにして、注意と行動が同じ制御メカニズムを共有しているのであれば、行動についてのアフォーダンスに対応して(注意のような)心的行動についてのアフォーダンスを想定することは理にかなっているといえる。なお、能動的推論においては、(心的行為を含まない狭義の)アフォーダンスは、感覚運動統合において体性感覚と知覚を同時に予測する前運動野の状態として、知覚的サリエンスとは別ものとして扱われてきた(Friston et al., 2012)。しかしここでのアフォーダンスはドパミン神経による精度の調整によって制御されるものとして捉えられており、注意のメカニズムと多くを共有している。この点でも、視覚刺激が注意という心的行動をア

⁸ "Attention is a process of (precision dependent) gain control." "Salience is associated with planned actions that resolve uncertainty." (Parr and Friston, 2019)

フォードすること(視覚サリエンス)と、視覚刺激が(たとえば)到達行動をアフォードすること(狭義のアフォーダンス)との間に連続性があることは支持される。

問題設定 1 のまとめおよび基本テーゼ再訪

ここまでの議論では、サリエンスを一種のアフォーダンスだとみなす基本テーゼを導入し、その上で、能動的推論の観点から、視覚サリエンスが可能な行動によって規定されるものであること、視覚サリエンスが covert な注意という心的な行動とともに perception-action cycle を構成していること、これらの点が基本テーゼと整合的であることを論じてきた。

これまでの議論は、しかし、もっぱら知覚サリエンスについてのものであった。基本テーゼは、しかし、知覚サリエンス以外にも拡張可能である。(テーゼ右辺における)「注意」または「注意が引き起こす諸行動」の種類の違いに応じて、(テーゼ左辺における)さまざまな種類のサリエンスを規定することができる。すなわち、基本テーゼは、(テーゼ右辺における)異なる詳細に応じて、(テーゼ左辺における)異なる種類のサリエンスの分類が可能となる。以下の[問題設定 3]ではこの考えを用いてサリエンスをアフォードするものによって分類する。この点が我々の提案の大きな強みの一つであり、例えば統合失調症に関する ASH の洗練化に役立つことになる。

次節以降では、サリエンスをアフォーダンスとみなす基本テーゼの枠組みの中で、知覚サリエンスと動機サリエンスを統一的に理解しつつ、しかし、異なる種類のアフォーダンスとして区別することができることを示したい。

[問題設定 2: 知覚サリエンスと動機サリエンスの統一的説明]

ここからは問題設定 2 として、知覚サリエンスと動機サリエンスの連続性、そして統一的説明が可能かどうかについて検討する。この目的のため、能動的推論の枠組みでは動機サリエンスがどのように説明されるかを紹介する。

能動的推論から見た動機サリエンス

ここでは動機サリエンスのうち、正の情動価を持つ incentive salience について考えよう。能動的推論の枠組みにおいては、視覚サリエンスの高いところに目を向けたとき、incentive salience の高いところに目を向けたとき、どちらも変分自由エネルギーを下げることに貢献する。

具体例として、散らかった部屋でコーヒーカップを探す状況を考えてみよう。散らかったさまざまな物体という視覚サリエンスの高い刺激が複数ある中で、コーヒーカップという報酬を探している。現在この人は左下に視線を向けている(Figure 4 上左、赤丸)。ここで 2 つの行動が可能だとする(Figure 4 上右)。(1) 左上の視覚サリエンスの高いものに目を向ける。(2) 右上の動機サリエンスの高いもの(報酬)に目を向ける。

ここで将来の変分自由エネルギーの期待値 $G(a)$ は行動 a の関数であり、Figure 4 下にある 2 つの項の和として計算される⁹。第 1 項の epistemic term は

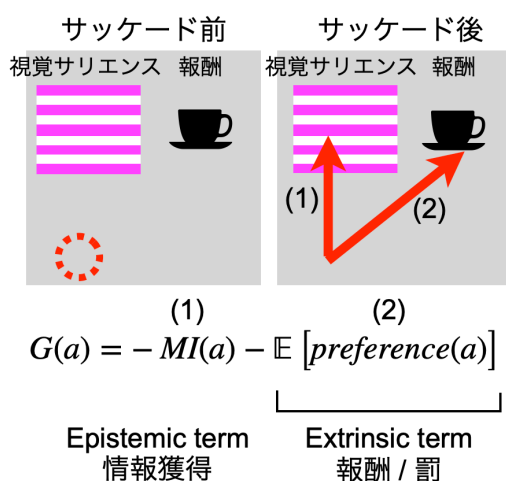


Figure 4

⁹ Friston の論文の初出では、第 1 項は epistemic value と呼び、第 2 項を extrinsic value と呼ぶ(Friston et al., 2015)。これは強化学習においては第 2 項の外的な報酬の最大化のみが目的関数となるのに対して、能動的推論では第 1 項の情報獲得も内的な報酬と捉えることで、強化学習に内発的動機づけを組み込むという意義がある。しかし現在の文脈では、value(価値)という言葉が混乱を生むので、よりニュートラルな「項 term」という語に

前セクションで言及した相互情報量 $MI(a)$ であり、外界の世界についての情報獲得の大きさを表している。よって(1) 視覚サリエンスの高いところに目を向けた場合、視覚シーンについての情報を獲得し、視覚シーンをより正確に推定することが可能となるため、epistemic term は大きくなり、結果として $G(a)$ は小さくなる。

第2項の extrinsic term が前セクションから新しく付加されている(前セクションでは第2項が関わらないので省略していた)。この項は、その生き物にとって望ましい視覚入力 (preference(a)) を得たときに大きくなる。よって(2) 報酬、つまり incentive salience の高いところに目を向けたらいい、extrinsic term は大きくなり、結果として $G(a)$ は小さくなる。ここまでは報酬という正の強化子を考えてきたが、嫌悪刺激についても負の報酬として扱うことで同じ枠組みが使える(Hesp et al., 2021)。たとえば Figure 5 にあるように嫌悪刺激(ヘビ)

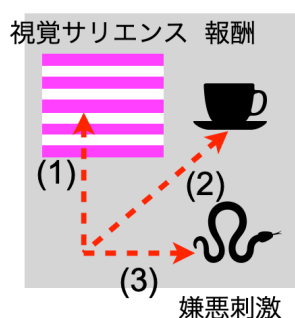


Figure 5

が提示されたとき、aversive salience の高いところから「目を背ける」ことによって extrinsic term は大きくなり、結果として $G(a)$ は小さくなる。

以上のようにして、Figure 5 での(1) 視覚サリエンスの高いところに目を向けること、(2) incentive salience の高いところに目を向けること、(3) aversive salience の高いところから目を背けること、これらすべてがその生物にとってより適応的な知覚状態を得るために行動選択をしているという点で、能動的推論を実現している。じっさいにどちらの行動を選ぶかは、 $G(a)$ を下げるのにどれがより有利かによってきまる。つまり、視覚サリエンスと動機サリエンスは「共通の通貨」を使っている。つまり、視覚サリエンスと動機サリエンスはともに視覚的な対象に付与されるものとして捉えられるが、そのサリエンスの大きさは行動 a が $G(a)$ をどの程度下げるかによって決まる¹⁰。

知覚サリエンスと動機サリエンスが持つ生態学的価値

能動的推論の枠組みによって、知覚サリエンスと動機サリエンスを「注意によって導かれた、適応的な行動を動機づけるもの」として統一的に捉えることができた。これはまさに (McClelland, 2020) が mental affordance hypothesis で提唱した mental affordance の一例となっている。ギブソンの知覚理論では、環境が生物にアフォードするものは、環境にある物理的な情報そのものではなくて、その生物と環境との間で規定される「生態学的な情報」であるとする。能動的推論の立場からは、視覚サリエンスと動機サリエンスがともに生態学的な情報であり、それゆえ連続的なものであると捉えることができる。

能動的推論の立場からは、視覚サリエンスの高いところに目を向けるという行動は、世界の状態についてより正確な推定をすることで変分自由エネルギーを下げていく。よってある生き物にとってある視線移動の視覚サリエンスが高いかどうかは、その生き物の現在の「信念」と世界についてのモデル「生成モデル」の両方に依存する。生成モデルとは、生き物が発達と進化の過程で生存のために獲得したものである。よって視覚サリエンスは外界の視覚入力だけで決まる特性ではなく、その生き物にとっての重要性、生態学的価値が含まれている。

たとえば白背景に黒い丸があると目立つ。視覚サリエンスが高い。でもそれはその生き物が視

置き換えている。

¹⁰ Extrinsic term 中の preference(a) は望ましい感覚入力の確率分布であるため、preference 自体は価値の大きさを表す。つまり、報酬では高く、嫌悪刺激では低い。しかしそのとき選ばれる行動 a を考慮するならば、報酬のときに収穫行動を行う、嫌悪刺激のときに回避行動を行う、どちらも extrinsic term はも正の大きい値を取る。このため extrinsic term は動機サリエンスを表すと解釈することが妥当となる。

覚によって行動する生き物であるからだ。このようにして視覚サリエンスにおいて扱われる低レベルの視覚特徴(明るさ、色、傾き、動き)ですら、生物が進化と発達で構築した生成モデルに依存する、というよりは生成モデルの一部となっている。

われわれ人間や他の哺乳類にとっては「顔」という視覚刺激は「目立つ」。しかしじつは顔刺激自体は低レベルの視覚特徴の「視覚サリエンス」は高くない。つまり顔の例ではわれわれ生物が顔を見つけることが発達の段階で重要であるために、顔が目立つように学習してきたのだ。

また連合学習理論においては、ある物体が価値と連合することで、注意を向ける対象となっていく過程をモデル化している。たとえば新規性が高いことによって獲得されたサリエンスは“acquired salience”としてモデル化されている(Esber and Haselgrove, 2011)。

このように「低レベルの視覚特徴による視覚サリエンス」、「顔のような発達、進化レベルで獲得されたサリエンス」、「そのつど報酬との連合で獲得されたサリエンス」、これらはどれも生物にとっての生態学的な意味を持ったものである。そしてこれらは明確に区別可能なものというよりは、進化的に獲得されたものか、それとも個体特有の経験に依存したものか、というスペクトラムによって連続的につながっている。

問題設定 2 まとめ

以上の議論では、知覚サリエンスと動機サリエンスは能動的推論の枠組みからは適応的な行動選択を動機づけるものとして統一的に説明した。また、生態学的価値からも、その価値の獲得過程や正負を考慮することで連続したものとして捉えられることが明らかになった。

[問題設定 3: 知覚サリエンスと動機サリエンスをアフォーダンスとして捉える]

問題設定 1 と問題設定 2 から得られたことを併せて考えると、知覚サリエンスだけでなく動機サリエンスもある種のアフォーダンスとして考えられることができる。ではこれらのサリエンスはどのような行動を動機づける(アフォードする)か。本セクションではこの問題にとりくむ。

サリエンスをアフォーダンスによって分類し直す

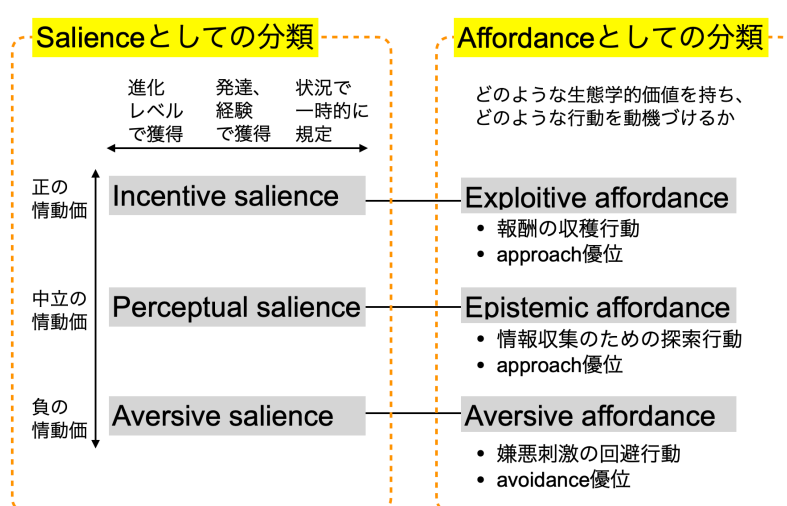


Figure 6

これらのサリエンスを「どのような行動を動機づけているか」という点から分類してみよう。どのサリエンスも知覚的経験に付与されるものであり、一義的にはサリエンスは注意をアフォードしている。そしてこの注意がどのような行動を促進するかによってサリエンスを分類することが可能となる。まず perceptual salience の場合はどうか。能動的推論によれば perceptual salience とは、外界についての信念をより正確にできるような行動(たとえばサッカー)によって規定されるものだった。

これは眼球運動だけに限らない。そのほかにも「聞き耳を立てる」や「対象に接近する」これらの行動によって選択される感覚入力には perceptual salience が高いと言える。このように考えると、perceptual salience が促進するのは「情報探索のための行動」であるといえるだろう。

う。Perceptual salience の高いところに向かう行動は期待自由エネルギーG(a)の第1項 epistemic term を最大化する。そこで本論文では情報探索を促進するものを epistemic affordance と呼ぶことにする¹¹。

それでは incentive salience の場合はどうか。能動的推論によれば報酬のある場所の感覚入力をサンプルする行動を選ぶことで、extrinsic term を最大化して、変分自由エネルギーの期待値の最小化を実現するのだった。このようにして考えると、incentive salience が促進するのは「報酬の収穫のための行動」であるといえるだろう。これが報酬の収穫を促進し、Epistemic affordance が情報収集を促進するということは、強化学習でいう explore or exploit (探索と活用)をこの2つで実現していることになる。そこで本論文では報酬の収穫を促進するものを exploitive affordance と呼ぶことにする。

Aversive salience の場合はどうか。能動的推論によれば aversive salience は負の報酬と捉えることで incentive salience と同様の扱いをすることができる。つまり、嫌悪刺激という負の報酬をサンプルしないような行動を取ることによって extrinsic term を最大化することができる。よって aversive salience が促進するのは「嫌悪刺激からの回避行動」¹²であるといえるだろう。そこで本論文では回避行動を促進するものを aversive affordance と呼ぶことにする。なお、アフォーダンスは従来、道具のようなポジティブな価値を持つものに対する行動を動機づけるものとして捉えられてきたが、これをネガティブな価値を持つものに対する行動へと拡張することが近年提唱されている(Proust, 2015)¹³。

以上をまとめると、それぞれのサリエンスとそれが注意を介して促進する行動を Figure 6 のように分類することができる。このような形で知覚サリエンスも動機サリエンスも、一義的には注意をアフォードすることで、さらなる行動が促進される、と理解することができる¹⁴。

3つのアフォーダンスの関係

ところでこれらのサリエンス/アフォーダンスは必ずしも独立して働くわけではない。3つのアフォーダンスがどのように協調して働くか明らかにするために、報酬刺激および嫌悪刺激に対する応答がどのように行われるか考えてみよう。まず報酬刺激を提示した場合を考えよう (Figure 7 上)。この図で縦軸は対象の情動価についての認識を表す。横軸は右向きに時間経過を表す。

主体はまず対象を知るために、対象に目を向ける、聞き耳を立てる、接近するなどの行動をするだろう。これは epistemic affordance に動機づけられた行動と捉えることができる。もしこの対象が正の情動価を持つ報酬であることがわかれば、その対象に接近して収穫する行動を取るだろう。これは exploitive affordance に動機づけられた行動と捉えることができる。正の情動価を持つ対象については、epistemic affordance と exploitive affordance の両方があることがわかる。

それでは負の価値がある対象(嫌悪刺激)の場合はどうなるか。主体はまず対象を知るために

¹¹ "Epistemic affordance" という語は本論文とは別の文脈で、心の哲学の研究者である Mohan Matthen が用いている。"If we extend J. J. Gibson's notion of an affordance—an object or situation apt for certain actions—to include epistemic actions of co-classification and the like, then the thesis is that perceptual states offer information about affordances, including—crucially—epistemic affordances" p.233 in (Matthen, 2005)

¹² たとえばヘビという嫌悪刺激に出会ったら、ヘビから逃げるだけでなく、ヘビを追い払うのも適応的な行動だ。よって aversive salience は「fight-or-flight を動機づける」という表現も可能だ。しかしここでは「嫌悪刺激という負の報酬をサンプルしないような行動」の総称としてここでは「回避行動」と表現している。

¹³ "an affordance is a perceived utility, which can be positive (something to approach and grasp) or negative (something to avoid and from which to flee)" (p.7) (Proust, 2015)

¹⁴ ここで「動機サリエンスが直接的に収穫行動や回避行動をアフォードする」と考えない理由は、サリエンスが知覚経験に付与されるものという定義からの帰結である。嫌悪刺激を例に取ってみよう。回避行動をアフォードするのはあくまでも「嫌悪刺激の知覚」であって、「嫌悪刺激のサリエンス」ではない。「嫌悪刺激のサリエンス」が「嫌悪刺激の知覚」に付与するものとは、注意を誘導して回避行動を促進することである、このように考えるからだ。一方で、動機サリエンスが脳内でどのように行動に影響を与えるのかは、実験的検証が必要な問題である。この点については[脳の解剖学的な対応による妥当性]において議論した。

なお、まったく逆の方向性として、「嫌悪刺激(の知覚)」自体は回避行動をアフォードしない、「嫌悪刺激のサリエンス」が回避行動をアフォードする、という立場もありうる。こちらは Berridge の incentive salience の定義(「ある刺激に欲しいと動機づけるような特性」とは整合性がある。しかしこのとき、「嫌悪刺激(の知覚)」自体がアフォードするものは、その刺激を物体としてどのように掴むことが可能価というものだけになる。こちらは生態学的価値としてのアフォーダンスの概念と整合的でない。

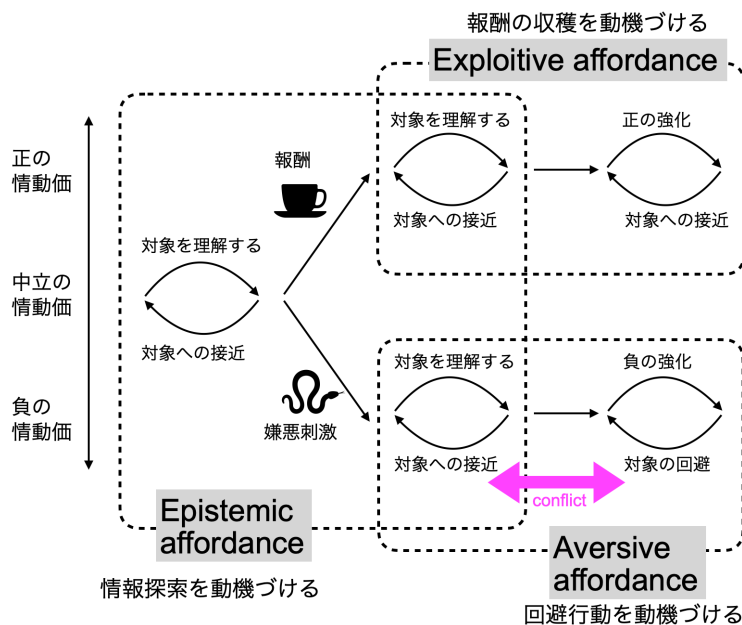


Figure 7

になった。

epistemic affordance に動機づけられた行動を起こす(Figure 7 下)。対象が嫌悪刺激であることが判明したならば、つぎはただちにこの刺激を回避する行動、つまり aversive affordance に動機づけられた行動を起こす。このようにして、負の情動価を持つ対象については、epistemic affordance と aversive affordance の両方があることがわかる。このようにして、報酬の場合も嫌悪刺激の場合も、まず epistemic affordance によって対象への情報収集を動機づけたうえで、ひきつづき報酬の場合は収穫行動、嫌悪刺激の場合は回避行動が動機づけられる。このような関係にあることが明確

Exploitive affordance と aversive affordance のコンフリクト

ここで正の価値と負の価値とで非対称性があることが明らかになる。正の価値の場合では、epistemic affordance と exploitive affordance の間にはコンフリクトがない。どちらも対象への接近を動機づけるという意味では類似性がある。しかし負の価値の場合では、epistemic affordance と aversive affordance の間にはコンフリクトがある(Figure 7 右下のマゼンタ矢印)。つまり、epistemic affordance は対象への接近を動機づけるのに対して、aversive affordance は対象の回避を動機づける。このため、負の価値を持つ対象について動機づけられる行動はより巧みな調節が必要となる。つまり、未知の対象について epistemic affordance が動機づけられて情報収集をしたとして、対象が負の価値の持つことが判明したらただちに aversive affordance に動機づけられた回避行動をすることが適応的であることになる。本論文ではサリエンスを生態学的な価値から捉え直すことで、その対象が動機づける行動の違いから分類することを試みた。この試みをもたらす重要な知見として「exploitive affordance と aversive affordance の非対称性」を得ることができた。この知見はこれまでにサリエンスが扱われてきたさまざまな分野で役に立つ可能性がある。ここでは精神医学への貢献の例として、本論文の introduction で採りあげた、異常サリエンス仮説 ASH に応用してみることにしよう。

[精神医学に対する示唆]

異常サリエンス仮説におけるサリエンスとアフォーダンス

[Aberrant Salience Hypothesis of Schizophrenia]のサブセクションで議論したように、ASH における「不適切なサリエンスの付与」のサリエンスの実体は明確でない。ここに動機サリエンスと知覚サリエンスの両方が関与する可能性がある。さらに、もしこのサリエンスの亢進の経験が、負の情動を引き起こす、忌避すべき刺激であるならば、それを回避すればよいだけで

あるが、その「新規で困惑する状態」perplexing state を認知的に説明して make sense しようとする点には、サリエンスの亢進の経験が「興味を惹く」という要素もあるように見える、ということを指摘した。

このような一見パラドキシカルに見える状況も、ここまで議論してきた incentive salience と aversive salience の非対称性を考慮すれば、納得のいく説明を与えることができる。つまり、ASH におけるサリエンスの変容とは、epistemic affordance と aversive affordance の両方の亢進であると考えれば、このパラドキシカルに見える状況も納得がいく。つまり、統合失調症においてドパミン神経の活動亢進が epistemic affordance と aversive affordance の両方を亢進させる。これによって epistemic affordance (対象への接近を動機づける) と aversive affordance (対象の回避を動機づける) との間のコンフリクトはさらに強まる。これが ASH で想定されている、ネガティブな情動の増強経験(「ある対象が不吉な意味を持つように思える」「自分の恐怖や痛みや不安を説明できない」)と同時に興味の増強経験(「普段なら気にもとめないような人や事象に興味を抱く」「偶然にみえる状況や事象になんらかの意味を見出したくなる」)の両方が起こることに対応している(Cicero et al., 2010)。このような「新規で困惑する状態」novel and perplexing state が、妄想のような認知的解決を必要とするほどまでに危機的状況であることは、epistemic affordance と aversive affordance のコンフリクトが強まったと考えれば、納得がいくのではないだろうか。

この新しい考えは、ASH をどのように実証するべきかについても手がかりを与えてくれる。これまでの ASH の実証を目指した研究で使われる行動課題として Salience Attribution Task (SAT)(Roiser et al., 2009)がある。この課題は視覚刺激と報酬の連合を被検者が学習して報酬を最大化する課題だが、視覚刺激セットは2つの次元(たとえば色と形状)を持っている。片方の次元(たとえば形状)が報酬と関連していて、もう片方の次元(たとえば色)は報酬と関連がない。しかし統合失調症患者では無関係の次元に意味づけをしてこれに固執するということが知られている。この課題で検証されているのは incentive salience に基づいた意思決定の過程である。しかし、もしこの論文での議論が正しいのであれば、ASH の直接的実証のためには、epistemic affordance と aversive affordance それぞれの変容について直接検証できる課題の開発が必要になると予測できる。

サリエンスの獲得、強化過程の寄与

異常サリエンス仮説では、ドパミン神経の活動亢進から不適切なサリエンスの付与が起こる、という説明以上のことは与えられていない。われわれの図式は異常サリエンスの獲得過程を明らかにすることに寄与できるかもしれない。上記の図式(Figure 7)では生得的な刺激に対する反応を取り上げていた。しかし異常サリエンス仮説における「不適切なサリエンスの付与」を捉えるために、サリエンスを獲得、強化する過程を含めてみよう。

新しい図式(Figure 8)では、見ただけでは中身がわからないものが対象となった例を表している。たとえばピンクの包み紙にはかならず甘いお菓子が入っているとするとする(Figure 8 上)。このとき初回は中立の情動価をもつ対象に(epistemic affordance によって)接近して、甘いお菓子を食べることで正の強化が行われる。これを繰り返すことで、ピンクの包み紙を見ただけでそれに接近して収穫したいという exploitive affordance を獲得するだろう。

次の例では、オレンジの包み紙にかならずワサビの塊が入っているとするとする(Figure 8 下)。このとき初回は中立の情動価をもつ対象に(epistemic affordance によって)接近して、ワサビの塊を食べることで正の罰が与えられる。これを繰り返すことで、オレンジの包み紙を見ただけでその対象を回避したいという aversive affordance を獲得するだろう(これは負の強化と呼ばれる)。異常サリエンス仮説における「不適切なサリエンスの付与」とは、このようなサリエンスの獲得、強化過程が適切に働いていないことと解釈することができる。そこでは epistemic affordance の亢進が不可欠な役割を果たすであろうことが、本論文で提示した図式からは予測できる。

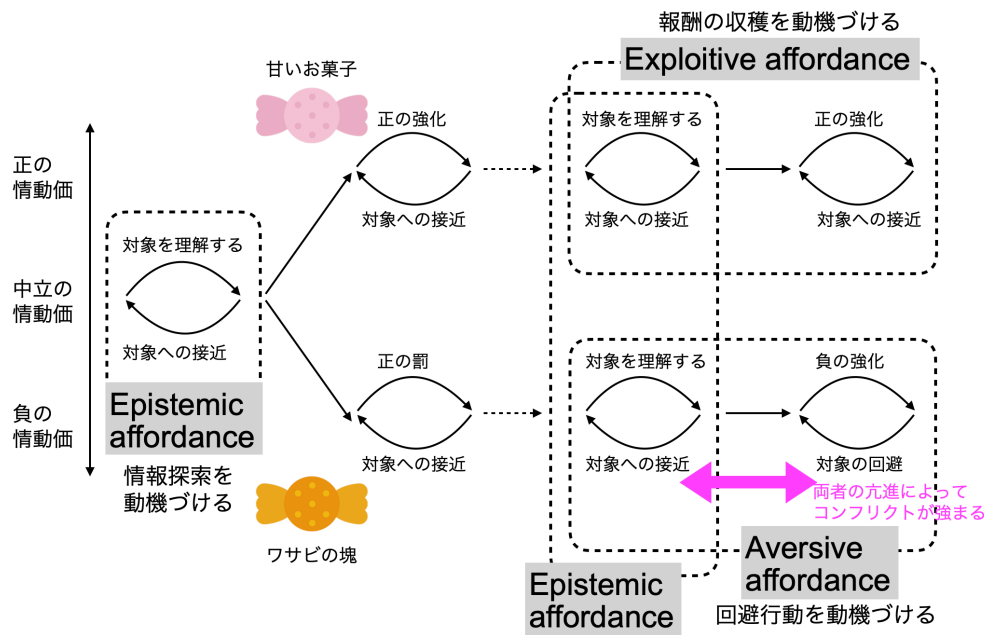


Figure 8

精神病症状の予測誤差精度モデルとの関連

能動的推論の立場からは精神病症状に関する説明モデルが提案されている(Adams et al., 2013; Sterzer et al., 2018)。これらは予測誤差の精度に変容があることが精神病症状を引き起こすことを提唱している。統合失調症患者でのイメージング研究から、幻聴の有無と聴覚情報処理での予測の精度とが関連していることが明らかになっている(Powers et al., 2017)。予測誤差の精度モデルはASHと重なり合う部分も大きい。予測誤差の精度の変容は、前述の通りベイジアンサプライズを過大評価することでサリエンスの変容を起こす。

しかし、既存のASHも予測誤差の精度モデルも、現状では精神病症状が起きるときのダイナミクスの片面しか説明していないように見える。つまり、ASHにおいてはサリエンスの亢進経験を説明するために妄想という認知的説明に飛びつく jumping-to-conclusion (JTC) bias とともに、妄想を否定する証拠が提示されたときにそれを受け入れることができない bias against disconfirmatory evidence (BADE)がある。このような2つの(一見)相容れない認知バイアスがなぜ起きるのかを、既存のASHも予測誤差の精度モデルも説明できていない。

本論文が提唱した epistemic affordance と aversive affordance とのコンフリクトは、このような2つの認知バイアスについて説明できるかもしれない。つまり、epistemic affordance の亢進があるため、世界についての新しいモデルとしての妄想を採用するという信念の更新が起こる。これが JTC bias を説明する。一方で、aversive affordance の亢進があるため、新しく採用した世界についてのモデルを廃棄することに情動的な忌避感を持つ。これが BADE を説明する。このようにしてASHにおける妄想の獲得と維持という一見不合理であるような現象について、epistemic affordance と aversive affordance とのコンフリクトによる説明が可能となる。

本論文で提唱した epistemic affordance と aversive affordance のコンフリクトは、予測誤差最小化の理論を用いてさらに解像度を上げることができるかもしれない。知覚や行動は五感を用いた体外のセンサーによる外受容感覚 exteroceptive sense での予測誤差を最小化するものであり、情動とそれにもとづいた行動は体内のセンサーによる内受容感覚 interoceptive sense での予測誤差を最小化するものと捉えられる(Seth and Friston, 2016)。このように考えるなら、epistemic affordance に動機づけられた探索行動とは、exteroceptive な予測誤差の最小化のこととなる。同様に、exploitive/aversive affordance に動機づけられた収穫/回避行動とは、interoceptive な予測誤差の最小化のこととなる。よって epistemic affordance と aversive affordance のコンフリクトとは、exteroceptive な予測誤差の最小化と

interoceptive な予測誤差の最小化の重みづけを適切に行うことができていない状態と捉えることができる。この予測は神経科学的研究によって検証可能だろう。

ASH と予測誤差精度モデルの疾患特異性

既存の ASH および予測誤差精度モデルに関する未解決問題の一つとして、サリエンスの亢進や予測誤差の過大評価が統合失調症などで見られる精神病症状に特異的といえるかという問題がある。統合失調症では聴覚情報処理だけでなく視覚情報処理についても様々な変容があることが知られている(Butler et al., 2008; Silverstein and Keane, 2011; Silverstein, 2016)。一方で感覚過敏は自閉症スペクトラム障害 ASD の主要な特徴である。また、低レベル視覚特徴のサリエンスに対する反応の亢進は ASD でも見られる(Wang et al., 2015)。また感覚情報処理における予測誤差精度の変容は ASD でも見られる(Lawson et al., 2017)。このような意味での epistemic affordance の亢進は必ずしも精神病症状のみで見られるものではないかもしれない。いっぽうで motivational salience の一般的な亢進は精神病症状よりは addiction やその他の精神疾患でこそより重要になるものと考えられる。以上のことから考えても、精神病症状に特異的な要素を探すならば、epistemic affordance と aversive affordance とのコンフリクトは非常に強力な候補となるのではないだろうか。

離人症と現実感喪失症

本論文での論点は、精神病症状と ASH 以外での精神疾患にも役立つかもしれない。たとえば離人症 depersonalization、現実感喪失症 derealization を考えてみよう。現実感喪失症は日常の体験から現実らしさが失われる精神状態とされる。「世界と自分のあいだに膜があるような感じ」「テレビを見ているような感じ」という主観報告から、個々の対象の知覚というより、外界の知覚経験を修飾する「いま・ここ性」のようなものが喪失している状態と考えられる。離人症・現実感喪失症では、環境や自分自身を対象とする主観経験から情動的な修飾がされていない状態とされることもある。このようなことから、離人症・現実感喪失症の患者の知覚には、知覚に情動的な価値を付加している motivational salience が失われている状態と説明できるかもしれない。実際に、離人症・現実感喪失症に情動を喚起するような刺激を提示されたときに皮膚電位が健常者より低いことが示されている(Horn et al., 2020)。Sierra と Berrios の二重過程説(Sierra and Berrios, 1998)では、離人症では情動の低下(hypo-emotionality)が外界への警戒状態の亢進(hyper-vigilance)の両方が起きていると提案されている。このような hyper-vigilance と epistemic affordance には関係があるかもしれない。これは今後の課題としたい。

脳の解剖学的な対応による妥当性

この論文で議論された事項を脳の解剖学的位置にマップするには膨大な数の論文への言及が必要となり、本論文が目的とすることではない。ここでは本論文が提示した、知覚サリエンスと動機サリエンスの連続性に関する部分だけ手短かにまとめておきたい。

まず ASH と関連して統合失調症患者およびハイリスク群でのドパミン神経の変容についての最新の知見については(Howes et al., 2020)から辿ることができる。近年の重要な仕事としては、薬物治療を行っていない初発エピソード精神病患者を対象にした機能的 MRI において、知覚サリエンスと動機サリエンスの両方で黒質および腹側被蓋野の応答が低下するというものがある(Knolle et al., 2018)。この研究ではミスマッチ応答の実験パラダイムを用いて、恐怖の表情などの動機サリエンスを持つ刺激だけでなく、情動的にはニュートラルな、新規性のみを持つ刺激(能動的推論においてはベイジアンサプライズを持っている)でも応答が低下することを報告している。この研究は知覚サリエンスと動機サリエンスの連続性が脳の機能レベルでも見られることを示唆している。

同様な知見は動物を対象とした研究からも明らかになっている。マカクザルの前頭基底部の神

経活動は、ミスマッチ応答の実験パラダイムにおいて、刺激が報酬と結びついた incentive salience だけでなく、刺激の新規性や知覚的なサプライズに基づいたサリエンスによっても応答することを明らかにしている(Zhang et al., 2019)。また、マカクザルのドーパミン神経はかつては報酬予測誤差を反映するものと考えられていたが、それとはべつのニューロン集団として、刺激のサプライズを反映するものがあることを明らかにしている(Matsumoto and Hikosaka, 2009)。これはつまり、報酬だけでなく、ネガティブな刺激(顔に向けての空気噴射刺激)においても一過性に活動するドーパミン神経があるということだ。また、マウスの側坐核でのドーパミン放出は報酬予測誤差を反映しているというよりは、知覚されたサリエンスを反映しているという報告がある(Kutlu et al., 2021)。

顔刺激に関する扁桃体の応答についても、動機サリエンス(Cunningham and Brosch, 2012)と知覚サリエンス(Santos et al., 2011; Gonzalez Andino and Grave de Peralta Menendez, 2012)の両方に関わっているという証拠がある。

視覚サリエンスに関連する脳部位についての総説も数多くあるが、最新の知見については(Veale et al., 2017)から辿ることができる。またこの総説は大脳皮質だけでなく、皮質下、とくに上丘を介するサリエンス計算回路に注目している点に特色がある。ラットやマカクザルの研究から、視覚サリエンスは大脳皮質を介さずに上丘から直接黒質ドーパミン神経へと伝えられることが明らかになっている(Comoli et al., 2003; Dommett et al., 2005; Takakuwa et al., 2017)。このような皮質下の回路が知覚サリエンスと動機サリエンスとをつなぐミッシングリンクとなっているかもしれない。

また、本論文と同様に知覚サリエンスと動機サリエンスを統合してASHを説明する試みとして(Miyata, 2019)がある。ここでは salience network に関する議論がなされている。知覚サリエンスと動機サリエンスが注意を誘導することで、強化や意思決定などのさまざまな行動を促進するという図式はたとえば(Schultz, 2015)などで提唱されている。じっさいに、知覚サリエンスだけでなく、動機サリエンス(incentive salience と aversive salience の両方)が注意を促進することについては、近年の心理学とイメージングを組み合わせた研究から明らかになっている(Bachman et al., 2020; Kim et al., 2021)¹⁵。

[Discussion]

この論文における議論の妥当性に関連して2点記しておく。

本議論における能動的推論の位置づけ

この論文では、能動的推論の考えを紹介しながら、知覚サリエンスが可能な行動の評価と関連していること、それゆえに知覚サリエンスを心的行動のアフォーダンスとして捉える考えと整合性があることを示した。また、知覚サリエンスと動機サリエンスとが変分自由エネルギーの期待値最小化からは「共通の通貨」を使っているという意味で連続性があることについて示してきた。

能動的推論はこれまでの脳科学のさまざまな知見と整合性があることから有望な考えであるとは言えるが、それでも能動的推論じたいは変分自由エネルギーの最小化という仮説から出てきたフレームワークであって、能動的推論と整合的であることだけではその考えの正しさを保証しない。むしろここで重要なのは、能動的推論の枠組みに特異的な計算論モデルそのものではなく、能動的推論の結果として達成される、適応的な生物の知覚と行動への視点こそが意味を持つ。よって本論文で提唱された、知覚サリエンスと動機サリエンスをアフォーダンスから捉える視点、epistemic affordance と aversive affordance のコンフリクトという考え方によって、能動的推論はそのアイデアの元とはなっているが、能動的推論の実証自体には依存しな

¹⁵ なお、これらの研究では知覚サリエンスについて physical salience と呼ぶものが多い。Physical salience と呼ぶ場合には、サリエンスの大きさを感覚刺激の intensity と同一視する傾向がある。同様な傾向は連合学習一般のサリエンスの扱いでも見られる。本論文では知覚サリエンスはあくまでも時間的、空間的に周りから突出しているものであり、intensity そのものではないことを協調しておく。この違いは、たとえば白地に黒い丸があったとき、黒い丸は輝度は低い、視覚サリエンスは高い、という例を考えれば明確になる。

い。それとはべつに本論文で提出された考えは将来の実験的な検証を必要とする。たとえば[異常サリエンス仮説 ASH におけるサリエンスとアフォーダンス]で提案したように、SAT に代わるサリエンス変容の行動テストを確立できたならば、本論文の議論の正しさは増強されるだろう。

情動と感情の問題

本論文では、能動的推論の枠組みに依拠しながら、ネガティブな情動について負の報酬としてとりあつた。しかし ASH が想定しているような、統合失調症の前駆期にサリエンスが亢進することで起こるネガティブな情動経験をこのような(意識下であり得る)情動と同一視してよいかは自明ではない。しかしこの問題は(ダマジオの語法に準拠するならば)無意識であり得る情動 emotion と意識経験である感情 feeling の関係という未解決問題に関わるため、本論文では踏み込まなかった。サリエンスの亢進が妄想という認知的な解決策を必要とするほどに苦しいものであることを考えるならば、ネガティブな情動経験をどう扱うかは ASH を精緻化するためには避けて通れない問題であり、これについては将来の課題としたい。

[Conclusion]

本論文では、知覚サリエンスが感覚入力のみで決まる特性であるという考えについて見直しを図るため、哲学での議論を援用して、サリエンスをアフォーダンスとして捉え直すことができるという本論の基本テーゼを提案した。「能動的推論」の枠組みからも、知覚サリエンスをアフォーダンスとして捉える考えが支持されることを示した。また、「能動的推論」の枠組みからは、知覚サリエンスと動機サリエンスが「同じ通貨を使う」という意味で統一的に扱えることを示した。これらをもとに、知覚サリエンスと動機サリエンスをアフォーダンスとして分類する考えを提案した。この視点から、ASH をより精緻にすることを試みた。ASH においては知覚的サリエンスと動機サリエンスの両方が亢進しているように見えるが、アフォーダンスとして捉える立場からは、負の情動価を持ち avoid したい aversive affordance と、でも make sense するために理解するために approach したい epistemic affordance というふたつの矛盾する要素があることを指摘した。

本論文は、心理学の哲学と神経計算論と精神医学を交錯させた結果として生まれた。今後さらにサリエンスの正体を明らかにするためには現象学的なアプローチが有効になるだろう。たとえばフッサール現象学には「受動的綜合」という概念があるが、これはサリエンスの無意識な(先反省的な)形成過程の分析となっているだけでなく、知覚サリエンスと動機サリエンスのつながりを示唆したものとなっている¹⁶。統合失調症患者のサリエンス変容の経験に直面するためにも、前駆期の自我の変容についての現象学的なアプローチ(Sass and Parnas, 2003)が有効であると考えられる。

もう一点、われわれが今後探求する必要があると考えているのは、サリエンスをフィーリング feeling と捉える方向性だ。最後にこの点について簡潔に述べよう。本論文では、サリエンスが注意をアフォードして各種行動を促進する、という図式でサリエンスを行動に影響を与えるものとして扱った。しかしサリエンスを対象に帰属し、経験する過程については、さらなる研究が必要となるだろう。ASH でのサリエンスの変容では、サリエンスが特定の対象に帰属されることを想定しているが、統合失調症の前駆期では、妄想気分 delusional mood と表現されるような、より対象が明確でない変容が起きていることが知られている。サリエンスの変容には「普段はどうでもよいものが気になる」というような誤帰属だけでなく、対象に帰属させることを失敗した結果としての、漠然とした違和感のような形での妄想気分も含まれるのかもしれない。このようなサリエンスの変容は、フィーリングの概念を導入することでより明確にできる

¹⁶ たとえば「赤い円筒と青い円筒が現れたら、それらは形の点では互いに結びつくが、色の点ではコントラストを成す。これは「受動的綜合」とよばれる原初的なはたらきに属する「連合」の現象」p.112 (田口, 2014)、それから「感覚的なもの(…)は同質的な物同士がまとまり、異質的なものとのコントラストを生み出して(…)「際立って」くる。この際立ちが「私」を「触発」する(…)その触発が十分強ければ、「私」がいわば目覚めて、「感覚的なもの」を「意味」へと仕上げる」p.131 (谷, 2009)

だろう。とくに対象が不明確なサリエンスの変容については、実存的フィーリング existential feeling (Ratcliffe, 2009)の概念が有用かもしれない。これらについては今後の課題としたい。

[利益相反の開示]

すべての著者で、本論文の内容に関して開示すべき利益相反はない。

[著者の貢献]

研究の発案: 吉田、宮園。証拠の収集と解釈: 吉田、宮園、西尾、山下、鈴木。主要な執筆者: 吉田。すべての著者が論文の執筆と改定に加わり、最終原稿の投稿を承認した。

[Acknowledgment]

本研究は JSPS 科研費 20H00001, 22H02936、AMED 課題番号 JP19dm0207069 の支援のもとで行われた。

[References]

- Adams RA, Stephan KE, Brown HR, Frith CD, Friston KJ (2013) The computational anatomy of psychosis. *Front Psychiatry* 4:47.
- Archer S ed. (2022) *Saliency: A Philosophical Inquiry*. Routledge.
- Bachman MD, Wang L, Gamble ML, Woldorff MG (2020) Physical Saliency and Value-Driven Saliency Operate through Different Neural Mechanisms to Enhance Attentional Selection. *J Neurosci* 40:5455–5464.
- Berridge KC (2018) Evolving Concepts of Emotion and Motivation. *Front Psychol* 9:1647.
- Butler PD, Silverstein SM, Dakin SC (2008) Visual Perception and Its Impairment in Schizophrenia. *Biol Psychiatry* 64:40–47.
- Cicero DC, Kerns JG, McCarthy DM (2010) The Aberrant Saliency Inventory: A New Measure of Psychosis Proneness. *Psychol Assess* 22:688–701.
- Comoli E, Coizet V, Boyes J, Bolam JP, Canteras NS, Quirk RH, Overton PG, Redgrave P (2003) A direct projection from superior colliculus to substantia nigra for detecting salient visual events. *Nat Neurosci* 6:974–980.
- Corbetta M, Akbudak E, Conturo TE, Snyder AZ, Ollinger JM, Drury HA, Linenweber MR, Petersen SE, Raichle ME, Van Essen DC, Shulman GL (1998) A common network of functional areas for attention and eye movements. *Neuron* 21:761–773.
- Cunningham WA, Brosch T (2012) Motivational Saliency: Amygdala Tuning From Traits, Needs, Values, and Goals. *Curr Dir Psychol Sci* 21:54–59.
- Dommett E, Coizet V, Blaha CD, Martindale J, Lefebvre V, Walton N, Mayhew JEW, Overton PG, Redgrave P (2005) How visual stimuli activate dopaminergic neurons at short latency. *Science* (80-) 307:1476–1479.
- Esber GR, Haselgrove M (2011) Reconciling the influence of predictiveness and uncertainty on stimulus saliency: a model of attention in associative learning. *Proceedings Biol Sci* 278:2553–2561.
- Friston K (2010) The free-energy principle: a unified brain theory? *Nat Rev Neurosci* 11:127–138.
- Friston K, Rigoli F, Ognibene D, Mathys C, FitzGerald T, Pezzulo G (2015) Active inference and epistemic value. *Cogn Neurosci* 6:187–214.
- Friston KJ, Shiner T, FitzGerald T, Galea JM, Adams R, Brown H, Dolan RJ, Moran R, Stephan KE, Bestmann S (2012) Dopamine, affordance and active inference. *PLoS Comput Biol* 8:e1002327.
- Gonzalez Andino SL, Grave de Peralta Menendez R (2012) Coding of saliency by ensemble bursting in the amygdala of primates. *Front Behav Neurosci* 6:38.

- Hesp C, Smith R, Parr T, Allen M, Friston KJ, Ramstead MJD (2021) Deeply Felt Affect: The Emergence of Valence in Deep Active Inference. *Neural Comput* 33:398–446.
- Horn M, Fovet T, Vaiva G, Thomas P, Amad A, D'Hondt F (2020) Emotional response in depersonalization: A systematic review of electrodermal activity studies. *J Affect Disord* 276:877–882.
- Howes OD, Hird EJ, Adams RA, Corlett PR, McGuire P (2020) Aberrant Salience, Information Processing, and Dopaminergic Signaling in People at Clinical High Risk for Psychosis. *Biol Psychiatry* 88:304–314.
- Itti L (2007) Visual salience. *Scholarpedia* 2:3327.
- Itti L, Baldi P (2009) Bayesian surprise attracts human attention. *Vision Res* 49:1295–1306.
- Itti L, Koch C (2001) Computational modelling of visual attention. *Nat Rev Neurosci* 2:194–203.
- Kapur S (2003) Psychosis as a state of aberrant salience: A framework linking biology, phenomenology, and pharmacology in schizophrenia. *Am J Psychiatry* 160:13–23.
- Kim H, Nanavaty N, Ahmed H, Mathur VA, Anderson BA (2021) Motivational Salience Guides Attention to Valuable and Threatening Stimuli: Evidence from Behavior and Functional Magnetic Resonance Imaging. *J Cogn Neurosci* 33:2440–2460.
- Knolle F, Ermakova AO, Justicia A, Fletcher PC, Bunzeck N, Düzel E, Murray GK (2018) Brain responses to different types of salience in antipsychotic naïve first episode psychosis: An fMRI study. *Transl Psychiatry* 8:1–13.
- Kutlu MG, Zachry JE, Melugin PR, Cajigas SA, Chevee MF, Kelly SJ, Kutlu B, Tian L, Siciliano CA, Calipari ES (2021) Dopamine release in the nucleus accumbens core signals perceived saliency. *Curr Biol* 31:4748-4761.e8.
- Lawson RP, Mathys C, Rees G (2017) Adults with autism overestimate the volatility of the sensory environment. *Nat Neurosci* 20:1293–1299.
- Matsumoto M, Hikosaka O (2009) Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature* 459:837–841.
- Matthen M (2005) Seeing, Doing, and Knowing: A Philosophical Theory of Sense Perception. Oxford: Oxford University Press.
- McClelland T (2020) The Mental Affordance Hypothesis. *Mind* 129:401–427.
- Miyata J (2019) Toward integrated understanding of salience in psychosis. *Neurobiol Dis* 131:104414.
- Moore T (2001) Control of eye movements and spatial attention. *Proc Natl Acad Sci* 98:1273–1276.
- Parr T, Friston KJ (2019) Attention or salience? *Curr Opin Psychol* 29:1–5.
- Posner MI (1980) Orienting of attention. *Q J Exp Psychol* 32:3–25.
- Powers AR, Mathys C, Corlett PR (2017) Pavlovian conditioning-induced hallucinations result from overweighting of perceptual priors. *Science* 357:596–600.
- Proust J (2015) The Representational Structure of Feelings. *Open MIND* 31:1–26.
- Ratcliffe M (2009) Existential Feeling and Psychopathology. *Access* 16:179–194.
- Ratcliffe M, Broome MR (2022) Beyond 'salience' and 'affordance': Understanding anomalous experiences of significant possibilities. In: *Salience*, pp 50–69. London: Routledge.
- Rizzolatti G, Craighero L (2010) Premotor theory of attention. *Scholarpedia* 5:6311 Available at: http://www.scholarpedia.org/article/Premotor_theory_of_attention.
- Robinson TE, Berridge KC (1993) The neural basis of drug craving: an incentive-sensitization theory of addiction. *Brain Res Brain Res Rev* 18:247–291.
- Roiser JP, Howes OD, Chaddock CA, Joyce EM, McGuire P (2013) Neural and behavioral correlates of aberrant salience in individuals at risk for psychosis. *Schizophr Bull* 39:1328–1336.
- Roiser JP, Stephan KE, den Ouden HEM, Barnes TRE, Friston KJ, Joyce EM (2009) Do patients with schizophrenia exhibit aberrant salience? *Psychol Med* 39:199–209.
- Santos A, Mier D, Kirsch P, Meyer-Lindenberg A (2011) Evidence for a general face salience signal in human amygdala. *Neuroimage* 54:3111–3116.
- Sass LA, Parnas J (2003) Schizophrenia, Consciousness and the Self. *Schizophr Bull* 29:427–444.
- Schultz W (2015) Neuronal Reward and Decision Signals: From Theories to Data. *Physiol Rev* 95:853–951.
- Seth AK, Friston KJ (2016) Active interoceptive inference and the emotional brain. *Philos Trans R Soc Lond B Biol Sci* 371.
- Sierra M, Berrios GE (1998) Depersonalization: neurobiological perspectives. *Biol Psychiatry* 44:898–908.
- Silverstein SM (2016) Visual Perception Disturbances in Schizophrenia: A Unified Model. In: *Nebraska Symposium on Motivation*, pp 77–132.

- Silverstein SM, Keane BP (2011) Perceptual organization impairment in schizophrenia and associated brain mechanisms: Review of research from 2005 to 2010. *Schizophr Bull* 37:690–699.
- Sterzer P, Adams RA, Fletcher P, Frith C, Lawrie SM, Muckli L, Petrovic P, Uhlhaas P, Voss M, Corlett PR (2018) The Predictive Coding Account of Psychosis. *Biol Psychiatry* 84:634–643.
- Takakuwa N, Kato R, Redgrave P, Isa T (2017) Emergence of visually-evoked reward expectation signals in dopamine neurons via the superior colliculus in V1 lesioned monkeys. *Elife* 6:1–17.
- Treisman AM, Gelade G (1980) A feature-integration theory of attention. *Cogn Psychol* 12:97–136.
- Veale R, Hafed ZM, Yoshida M (2017) How is visual salience computed in the brain? Insights from behaviour, neurobiology and modelling. *Philos Trans R Soc B Biol Sci* 372:20160113.
- Wang S, Jiang M, Duchesne XM, Laugeson EA, Kennedy DP, ADOLPHS R, Zhao Q (2015) Atypical visual saliency in autism spectrum disorder quantified through model-based eye tracking. *Neuron* 88:1–13.
- Yoshida M, Miura K, Hashimoto R, Fujimoto M, Yamamori H, Yasuda Y, Ohi K, Fukunaga M, Takeda M, Isa T (2015) Saliency-guided eye movement during free-viewing in schizophrenic patients. *J Vis* 15:61.
- Yoshida M, Taguchi S (2018) Free Energy Principle and Visual Consciousness. *The Brain & Neural Networks* 25:53–70.
- Zhang K, Chen CD, Monosov IE (2019) Novelty, Saliency, and Surprise Timing Are Signaled by Neurons in the Basal Forebrain. *Curr Biol* 29:134–142.e3.
- 田口茂 (2014) 現象学という思考. 筑摩書房.
- 谷徹 (2009) 現象学と間文化性. *現代思想* 2009年12月臨時増刊 37:126–141.