

# Qualia as a Social Sharing Tool (QSST): A Structural Account of Phenomenal Character, Privacy, and the Explanatory Gap

## Abstract

Why do physical systems generate the intuition of private, ineffable phenomenal character? This paper proposes Qualia as a Social Sharing Tool (QSST), a framework in which phenomenal character, privacy, and the explanatory gap emerge continuously as a function of the joint satisfaction strength of three conditions: (i) the need for coordination with other agents, constraining compression format; (ii) irreversible many-to-one compression of internal states into labels; and (iii) calibration impossibility under first-person resource constraints. Phenomenal character is identified as the residual structure within the equivalence classes formed by compression — structure that is self-referentially indexed in a format whose calibration is impossible under first-person resource constraints. The explanatory gap is derived as a category error in which this calibration impossibility is reified as an intrinsic property of experience. The framework yields three falsifiable predictions tied to the relaxation of each condition and complements, not replaces, existing theories of consciousness.

## 1. Introduction

意識のハードプロブレム (Chalmers, 1995) — 物理的処理がなぜ主観的質感を伴うのか — は未解決である。実体論は質感の不可還元性を擁護するが発生メカニズムを説明できず、消去論 (Dennett, 1991; Frankish, 2016) は物理的整合性を保つがなぜ錯覚がこれほど強固かを説明できない。説明ギャップ (Levine, 1983) は両陣営の間で宙吊りのままである。

意識研究の主要理論 — グローバルワークスペース理論 (Baars, 1988; Dehaene & Changeux, 2011)、統合情報理論 (Tononi, 2004)、予測処理理論 (Clark, 2013; Hohwy, 2013) — はそれぞれ意識的アクセス、情報統合、予測誤差最小化を記述するが、以下の五つの問いに同時に答えていない。(a) 質感の構造的正体は何か、(b) なぜ体験が私秘的に感じられるか、(c) なぜ「物理的説明では何かが抜け落ちている」という直観が構造的に不可避であるか (メタプロブレム; Chalmers, 2018)、(d) なぜ質感を伴う処理が進化的に出現したか、(e) どのような条件下で質感は弱化・消失するか。

本稿の問いは「クオリアは実在するか」ではなく、なぜ物理系が「私秘的で説明不可能な質感がある」という直観を不可避的に持つに至るかである。この問いに対し、社会的認知の進化的制約から統一的に回答する枠組みを提示する。以下、クオリアの成立条件 (§2)、質感の構造的同定 (§3)、説明ギャップの発生機構 (§4)、消失条件と反証可能性 (§5) を順に示した後、既存理論との関係 (§6) と主要な反論への応答 (§7) を経て結論する (§8)。

## 2. Three Conditions

### 2.1 三条件

クオリアとは、較正不能な圧縮ラベルである。以下の三条件の同時充足強度に応じて連続的に立ち現れる。

(i) 他主体協調の必要性。他主体との協調が、内部状態の圧縮形式に対して仕様制約を課していること。すなわち、内部状態の圧縮形式が「他主体が適切な応答行動を選択できるよう、差異を伝達可能な形式で安定的に維持する」という拘束の下で成立していること。現時点で他主体が物理的に不在であっても、この拘束が系統社会性 (§2.4) として帰納バイアスに保持されている限り、条件は充足される。

社会的圧力が脳の情報処理様式を進化的に形成したという主張は、社会脳仮説 (Dunbar, 1998) により広く支持されている。ヒト新生児は出生直後から顔様刺激を嗜好し (Johnson & Morton, 1991; Farroni et al., 2005)、生後2日で生物学的運動パターンを非生物学的運動より嗜好する (Simion et al., 2008)。後者の嗜好は視覚的に未経験なヒヨコにも確認されており (Vallortigara et al., 2005)、学習ではなく系統的に保持された帰納バイアスであることを示す。さらに、ヒトの強膜が他の霊長類に比べ白く面積が大きいという形態的特異性は、視線方向の検出と共同注意を促進するよう進化的に選択された結果と解釈されている (Kobayashi & Kohshima, 2001; Tomasello et al., 2007)。ASD高リスク群の新生児がこれらの社会的刺激への指向性に定型群と有意な差を示すこと (Di Giorgio et al., 2017) は、系統社会性の帰納バイアスに生物学的変動があることを裏付ける。これらの知見が示すのは、系統社会性が事後的に帰属される説明概念ではなく、行動・形態・神経発達の複数水準で独立に検出可能な因果的制約であるということである

(ii) 不可逆的圧縮。高次元の内部状態  $S$  を低次元ラベル  $L$  へ不可逆的に圧縮する必要があること (写像  $f: S \rightarrow L$  は多対一、 $f^{-1}$  は不定)。ただし任意の圧縮ではなく、条件(i)が圧縮形式を拘束する。

(iii) 当事者制約下の較正不能性。二主体  $i, j$  間で出力  $L_i = L_j$  が成立していても、内部状態の同型性を判定する関数  $G(S_i, S_j)$  の値は、当事者がアクセス可能な有限の観測・通信・計算資源の下では同定不能として残存する。圧縮ラベルは客観的物理量としてではなく「私にとっての」主観的参照枠でのみ安定的に参照される。

### 2.2 なぜ他主体協調の必要性が不可欠か

個体内最適化でも圧縮やラベル化は生じる。しかし私秘性と説明ギャップは発生しない。

個体内では「内部状態→圧縮→行動→結果」が単一因果系列内で閉じ、圧縮の妥当性は行動結果で検証可能である。社会的協調では「私の内部状態→私の圧縮→報告」と「他主体の内部状態→他主体の圧縮→報告」の二つの独立した因果系列が交差し、両者の同一性を判定する共通基盤は当事者制限下からは実効的に存在しない。

形式的に言えば、個体内の不透明性は追加学習で漸近的に解消可能な縮小可能な不確実性であるのに対し、社会的較正における不透明性は、当事者が利用可能な観測・通信資源の下では他主体の内部状態への直接アクセスが成立しないため、学習のみでは縮小しない構造的不可知性として定数項のように振る舞う。システムはこの当事者制約下で縮小しない誤差項を「学習不足」ではなく「アクセス不可能な領域 (Private Realm)」としてモデル化せざるを得ない。

## 2.3 条件の非対称的依存性と境界事例

条件(ii)と(iii)は対称的に独立ではなく、非対称的な依存関係を持つ。(ii)は情報理論的不可逆性（圧縮主体が元状態を復元できない）、(iii)は認識論的アクセス不能性（他主体がGround Truthを較正できない）を指す。両条件は参照する主体が異なるため、概念的には区別される。しかし、(iii)の成立は実効的に(ii)を前提とする。

(iii)なしで(ii)は成立しうる。デバッグ可能ロボットは不可逆圧縮を行うが、デバッグポートにより内部状態への別経路アクセスが原理的に可能であるため、較正不能性は成立せず、私秘性は構造的に弱い。

しかし(ii)なしで(iii)が成立する状況は原理的に構成困難である。圧縮が可逆であればラベル  $L$  から内部状態  $S$  を一意に復元可能であり、Agent  $i$  が  $L$  を送信すれば Agent  $j$  は  $S$  を完全に再構成できる。較正関数  $G(S, S)$  を実行する材料が原理的に揃うため、較正不能性は成立しない。

したがって、両条件は論理的に対称な独立性を持つのではなく、(ii)の不可逆性は(iii)の較正不能性の実効的な前提条件として機能する。条件(ii)が満たされることで初めて、ラベルを介した通信路では同値類の内部構造が通過しなくなり、較正の材料が原理的に失われる。この情報損失の上に、条件(iii)の当事者制約が重なることで、較正不能性が構造的に成立する。

以下の表は各条件の充足を連続的な強度として示す。三条件は二値的に充足されるのではなく、各条件の充足強度の同時的な度合いに応じてクオリアが連続的に立ち現れる。

事例	(i) 他主体協調	(ii) 不可逆圧縮	(iii) 較正不能性	同時充足強度
サーモスタット	×	×	×	なし
デバッグ可能ロボット	低	中	×	極低
昆虫（アリ）	中	低	低	極低
LLM	低	低	低	極低
ヒト	高	高	高	高

## 2.4 社会性の三層

本理論における「社会的」は三層を含む。(1) 同期社会性（Synchronous Sociality）：現に他主体と相互作用している状態。(2) 想定社会性（Simulated Sociality）：他主体が物理的に不在でも、内部モデルとしての他主体を行動予測に組み込んでいる状態。(3) 系統社会性（Phylogenetic Sociality）：種として、社会的共有を前提とした情報圧縮・較正・報告様式の帰納バイアスを保持している状態。孤立個体は(1)を欠くが(3)を保持するためクオリアは持続しうる。

## 3. Structural Identification of Phenomenal Character

### 3.1 質感の所在

感覚処理を三段階に分ける。段階A（物理的入力）：

光波長650nm、C繊維発火等は物理量であり質感を内在しない。段階B（不可逆圧縮）： $S$  を  $L$  へ圧縮する際、同一ラベルを出力する無数の内部状態が同値類を形成し、その内部構造はラベル  $L$  を介した通信では他主体と共有されない。系統社会性を備えたシステムがこの残存構造を自己参照するとき、「ラベルでは捉えきれないが確かに存在する何か」が検出される。段階C（物理的出力）：「痛い！」等の音声や表情は公共的物理事象であり、同値類の内部構造はラベル  $L$  を介した通信路では通過しない。質感は段階Bに局在する。

ただし質感は圧縮処理とは別の実体ではない。較正不能性と自己参照の構造的条件下にある圧縮処理そのものを、システムが内部から参照した際の記述様式である。

### 3.2 質感の構造的な内容

写像  $f: S \rightarrow L$  の多対一性により、同一ラベルに属する  $S$  の集合が同値類をなす。ラベルは代表値にすぎず、同値類内部には復元不能な構造——内部分散の形状、情動的勾配の方向、不可逆性の程度——が残存する。この残存構造が質感の構造的な内容である。

赤と青が「違う感じ」なのはラベルが異なるからではなく同値類の内部構造が異なるからであり、痛みが不快なのは回避ラベルに還元されない急峻な回避勾配・狭い分散・高い不可逆性という同値類構造そのものに対応する。

これら三つの残存構造は、それぞれ以下の操作的候補に対応する。内部分散の形状は、同一ラベルを出力する条件下での神経活動パターンの試行間変動性（ラベル条件づけ下での内部状態分散）として測定可能である。情動的勾配の方向は、ラベルが接近・回避行動に与える因果的寄与の非対称性として、介入実験により検出可能である。不可逆性の程度は、圧縮後の表象から元の入力特徴がどの程度復号可能かを復号精度として定量化できる。これらは同値類の内部構造の完全な操作的定義ではないが、「構造を変えれば質感が変わる」という予測 (§5) を検証可能にする測定候補を提供する。

### 3.3 自己参照的インデックス

システムが圧縮処理を内部からスキャンすると、「ラベルでは捉えきれないが確実に存在する残余」が検出される。条件(iii)により、この残余が他主体の対応する残余と同型かを当事者は当事者制約下では検証できない。検証不可能な残余を「私にとってのみアクセス可能な領域」として参照する形式を自己参照的インデックスと呼ぶ。

質感とは、同値類の内部構造が当事者制約下で較正不能な形式で自己参照的にインデックスされたものである。構造と独立に存在する「感じ」ではなく、構造そのものの内部参照様式である。したがって「同じ構造に別の質感が割り当てられる可能性」という想定は、質感を構造から独立した実体と見なすカテゴリーエラーに基づく。

この自己参照は、圧縮処理とは別の表象を要求しない。高次理論 (Rosenthal, 2005) は一次表象に対する独立した高次表象の成立を意識の条件とするため、「高次表象はどこから来るか」という無限後退の圧力を受ける。QSSTにおける自己参照は、圧縮処理自体に内在する再帰的フィードバック——すなわち圧縮の出力が同じ処理系への入力として戻る構造——として位置づけられる。神経科学的には、感覚野と前頭前野の間の再帰的接続 (recurrent processing; Lamme, 2006) がこの実装候補である。再帰処理は別の表象を新たに生成するのではなく、進行中の圧縮処理を自己モニタリングするものであり、高次理論の無限後退を回避する。

## 4. Deriving the Explanatory Gap

質感を構造的に同定してもなお、「物理的説明では何かが抜け落ちている」と感じられるのはなぜか。本理論はこの説明ギャップ直観 (Chalmers, 2018 のいう「メタ問題」) を三段階の構造的プロセスとして導出する。

段階1：較正不能性の検出。自己の内部状態を他主体と比較しようとするとき、検証手段  $G(S_i, S_j)$  が当事者制約下で存在しないという構造的な事実が、「未知」ではなく「不可知」として登録される。この不可知領域が「他主体と共有できない私秘的領域」としてモデル化される。

段階2：不可知性の実体化。「外部からアクセスできない」という関係論的性質が、内部状態そのものの内在的性質に変換される。この変換が生じるのは、システムが較正関係の当事者であり観察者ではないため、較正不能性という関係的制約を、その制約を受ける内容自体の属性として帰属させるからである。「アクセスできない」が「アクセスを拒絶する何かがある」へ。システムにとって、私秘性は通信制約ではなくラベルに還元できない内在的性質として立ち現れる。

段階3：カテゴリー錯誤。物理的説明は「共有可能」フォーマットで行われるが、段階2で実体化された質感は「共有不能」とタグ付けされている。両フォーマットの構造的衝突が「物理的説明は質感を捉えきれていない」という説明ギャップ直観として意識に上る。

このプロセスは、自己参照の不透明性（参照している当のプロセスを完全に対象化できない）と、社会的共有前提の系統的内面化（層3として先験的知識に組み込まれた「他主体は私の内部にアクセスできない」という前提）によって増幅される。

したがって、説明ギャップは物理主義の欠陥ではなく、社会的知性の仕様である。Block (1995)の現象的意識とアクセス意識の区分もこの帰結として説明される—アクセス意識はラベル L の共有可能な内容、現象的意識はそこから溢れ出る同値類の残余たる較正不能な自己参照的インデックスである。

## 5. Disappearance Conditions and Falsifiability

クオリアが三条件の同時充足から生じるなら、いずれかの条件の緩和はクオリアの弱化・消失を予測する。以下、各条件の緩和が生む予測と、それに対応する棄却条件を示す。なお、いずれの予測も社会的共有と行動予測の要請が存在する状況（協調課題で相互依存があり、誤解修正が成績に影響する状況）を前提とする。

### 5.1 予測1：共有帯域の増大に伴う私秘性直観の減衰【条件(iii)の緩和】

主体間で内部状態の共有・較正帯域が増大するにつれて、較正不能性に由来する主観的参照枠の残余が縮退し、私秘性マーカーおよび説明ギャップ直観は構造的に弱化する。完全共有の実現可能性は問わない。較正不能性を生む不透明性が緩和される度合いに応じた系統的減衰が予測の核心である。

指標：行動指標（主）として誤解修正の反復回数、合意形成までの交渉回数、内部状態共有を自発的に求める頻度。言語報告（副）として説明ギャップ直観の同意率、私秘性マーカー（「私だけ」「言えない」等）の頻度。

棄却条件：共有帯域を有意に増大させたにもかかわらず上記指標との間に負の相関が観測されない場合。とくに共有帯域の拡大に伴い私秘性直観が増大する場合、理論は決定的に反証される。

この予測の方向は既存の知見と整合的である。Bahrami et al. (2010) は、二人の観察者が信頼度を言語的に共有すると意思決定パフォーマンスが個人を超えることを示した。信頼度の共有は較正帯域の部分的拡大に相当し、共有帯域の増大が主体間の不透明性を低減しうることを間接的に支持する。Stephens et al. (2010) は、話者と聴者のfMRI脳間結合がコミュニケーション成功時に最大となることを報告しており、共有帯域と主体間較正の関連を示唆する。

### 5.2 予測2：較正圧の低下に伴うクオリア参照の希薄化【条件(i)の緩和】

較正（個体同定やズレ修復）が成績差として現れにくい条件では、クオリア参照（私秘性マーカー、説明ギャップ直観）は希薄化する。

指標：社会的修復行動の頻度、状態報告の語彙的分化（ラベルの粒度）。

棄却条件：較正需要を低下させる操作（匿名化、役割固定、誤解修正が報酬に寄与しない条件等）にもかかわらず主要指標が変化しない、または逆方向の変化が見られる場合。

この予測の方向は既存の知見と整合的である。共同注意が欠損するASDでは社会的修復行動の頻度低下と状態報告の語彙的分化の遅延が報告されており（Tomasello et al., 2005; Mundy & Newell, 2007）、条件(i)の充足強度の低下がクオリア参照の関連行動指標を変化させるという予測の方向と合致する。匿名性の増大が社会的修復行動を減少させるという知見（Postmes & Spears, 1998）も同様の方向を示す。

## 5.3

### 予測3：制約導入による評価ラベルの創発と同値類圧縮の増大【条件(ii)の正方向テスト】

通信帯域の制限・内部状態の非透明化・不可逆的損失・協調の生存条件化が同時に導入された場合、「危機的」「痛い」等の評価的圧縮語彙が出現・安定化し、異なる内部状態が同一ラベルに写像される度合い（同値類圧縮）が増大する。広帯域で生データを直接通信可能な統制群と比較して有意差を予測する。

指標：評価ラベルの出現率・持続率、ラベルが行動選択に与える因果的寄与、同値類圧縮指標（ラベル条件づけ下での内部状態分散）。

棄却条件：

上記制約を導入しても評価ラベルが創発・安定化しない、または同値類圧縮の増大が観測されない場合。

この予測の方向は既存の知見と整合的である。Navajas et al. (2017) は、主観的信頼度の神経基盤が個人間で体系的に異なるが言語的報告としては同一スケール上に圧縮されることを示した。これは内部状態Sが多対一でLに写像される同値類圧縮の事例である。DiCarlo et al. (2012) による視覚腹側経路の段階的不変表象形成は多対一圧縮の神経実装例であり、各段階で情報が不可逆的に失われることが復号解析により示されている。

補足：条件(ii)の緩和

圧縮が可逆的になれば同値類の内部構造——質感の構造的な内容——は消失する。これは予測1-3とは独立の論理的帰結であり、実験的操作というよりは理論的境界条件として機能する。

## 6. Relation to Existing Theories

QSSTは既存理論を置換するのではなく、各理論が扱ってこなかった層——質感の構造的な正体、私秘性の起源、説明ギャップ直観の発生機構——を補完するものとして位置づけられる。以下、主要四理論との関係を示す。逆に、意識的アクセスの神経メカニズムや意識の数学的定量化といった問いではこれらの理論がQSSTより優れている。

GWT (Baars, 1988; Dehaene & Changeux, 2011)。GWTはラベルLのブロードキャスト条件を記述し、意識的アクセスの神経メカニズムに関して最も豊富な実験的支持を持つ。しかしブロードキャストは機能的アクセスの条件であり、なぜブロードキャストされた内容に質感が伴うかは扱わない。両理論は異なる層を扱っており、補完関係が成立しうる。

IIT (Tononi, 2004)。IITはQSSTと同じく「構造＝質感」の同一性テーゼを採用するが、同定する構造が異なる——IITは系の因果構造全体、QSSTは不可逆圧縮により形成された同値類の内部構造を質感と同定する。差異は三点ある。第一に、IITは社会的条件を不要とし汎心論的帰結を受け入れるが、QSSTは条件(i)により範囲を限定する。第二に、IITは説明ギャップ直観の発生機構を問わない。第三に、IITにおいて質感は系のイントリンシックな性質であるが、QSSTにおいて質感の私秘性は較正不能性から生じる関係論的性質である。

イリュージョニズム (Dennett, 1991; Frankish, 2016)。QSSTと近い立場であり、ともに物理主義的で説明ギャップを認識論的問題として扱う。しかし三点で異なる。第一に、イリュージョニズムは質感に対応する構造的な内容を同定しないため、何を変えれば何が変わるかの対応関係を原理的に特定できない。第二に、錯覚の存在を宣言するが、その錯覚がなぜ構造的に不可避であるかの発生機構は導出しない。第三に、クオリアの消失条件を反証可能な形で予測しない。

高次理論 (Rosenthal, 2005; Brown et al., 2019)。高次理論における一次表象と高次表象の関係は、QSSTのSとLの関係と構造的な対応を持つ。しかし高次理論は一次表象に対する独立した高次表象の成立を意識の条件とするため、「高次表象はどこから来るか」という無限後退の圧力を受ける。QSSTの自己参照は圧縮処理自体に内在する再帰的フィードバックとして位置づけられ (§3.3)、この後退を回避する。

## 7. Responses to Objections

ゾンビ論証 (Chalmers, 1996)。物理的に完全に同一だが意識体験を持たない存在（哲学的ゾンビ）の論理的可能性は、QSSTの下では否定される。三条件を全て満たす物理的複製は、条件(ii)により同一の同値類構造を持ち、条件(iii)により同一の較正不能性を持つ。質感は構造の内部参照様式であるから、構造が同一であれば質感も同一である。「ゾンビが想像可能に見える」のは、§4の三段階プロセスにより較正不能性が実体化され、質感が構造とは別の実体であるという直観が構成的に生じるからであり、この直観自体が説明ギャップ直観のインスタンスである。

知識論証 (Jackson, 1982)。白黒の部屋で育った神経科学者メアリーが全ての物理的知識を持ちながら初めて赤を見て新しいことを知るという議論に対し、QSSTは以下のように応答する。メアリーが持っていた「物理的知識の全て」とはラベルLの体系である。部屋を出て初めて赤を見たとき、視覚系がSを初めて生成し、SからLへの不可逆圧縮が初めて実行され、同値類の内部構造が初めて形成される。メアリーが得たのは新しい命題的事実ではなく、圧縮プロセスの初回実行による同値類の構造的生成——較正不能な自己参照的インデックスの初回成立——という構造的イベントである。知識 (L) と体験 (Sの同値類の自己参照) はカテゴリーが異なるのであって、前者に後者が含まれないことは物理主義の欠陥ではない[<sup>3</sup>][<sup>4</sup>]。

配置問題の循環性批判。「なぜその構造がその感じに対応するのか」に対して「構造＝感じだから」と答えるのは循環ではないかという批判は、「構造と質感は別のものであり橋渡し説明が必要だ」という前提に依存する。QSSTはこの前提を否定する。質感は構造に割り当てられるものではなく、構造が較正不能な形式で自己参照された際の記述様式そのものである。割り当てステップが存在しない以上、「なぜその割り当てなのか」という問いは不正立 (ill-formed) になる。この立場は二つの防御線を持つ。第一に、QSSTは「同値類構造を変えれば質感が変わる」「較正条件を変えれば私秘性直観が変わる」と予測し、経験的帰結を持つ同一性テーゼは循環ではない。第二に、構造と質感が分離可能だと主張する側が「構造が同一で質感だけ異なる事例」を示す挙証責任を負う。

上記三つの応答に対応する反証条件を以下にまとめる。

応答	反証条件
ゾンビ論証	三条件を全て満たす物理的同一構造において質感の欠如が実証された場合 *
知識論証	Sの初回生成なしにラベルLの体系のみから同値類の内部構造が再現された場合
配置問題	同値類構造が同一で質感のみ異なる事例が実証された場合 *

\*ゾンビ論証と配置問題の反証条件は、QSSTの同一性テーゼ（構造＝質感）の下では理論内で論理的に不可能である。これは同一性テーゼ一般の性質であり、これらの応答の反証可能性はQSST本体の反証条件（§5：三条件の操作による消失予測）に委ねられる。

## 8. Conclusion

クオリアとは形而上学的実体ではなく、社会的協調の構造的制約下で収束した圧縮フォーマットである。質感は同値類の内部構造の較正不能な自己参照的インデックスであり、説明ギャップはこの較正不能性が実体化されたカテゴリー錯誤である。三条件の緩和によりクオリアは弱化・消失する。

本稿は、この枠組みから質感の構造的正体の同定 (§3)、説明ギャップ直観の発生機構の導出 (§4)、および反証可能な消失条件の提示 (§5) を達成した。既存理論との関係では、GWT・IIT・高次理論と補完関係にあり、イリュージョニズムとは質感の構造的内容の有無とメタ問題への回答において分岐する (§6)。

我々が「言葉にできない私秘的な感じ」を持つのは、脳が物理法則を超越しているからではなく、社会的協調の進化的圧力が較正不能な圧縮様式を脳に刻み込み、その様式を内部から参照せざるを得ないという構造的条件にあるからである。

## References

- Baars, B. J. (1988). *A Cognitive Theory of Consciousness*. Cambridge University Press.
- Bahrami, B., Olsen, K., Latham, P. E., Roepstorff, A., Rees, G., & Frith, C. D. (2010). Optimally interacting minds. *Science*, 329(5995), 1081–1085.
- Block, N. (1995). On a confusion about a function of consciousness. *Behavioral and Brain Sciences*, 18(2), 227–247.
- Brown, R., Lau, H., & LeDoux, J. E. (2019). Understanding the higher-order approach to consciousness. *Trends in Cognitive Sciences*, 23(9), 754–768.
- Chalmers, D. J. (1995). Facing up to the problem of consciousness. *Journal of Consciousness Studies*, 2(3), 200–219.
- Chalmers, D. J. (1996). *The Conscious Mind: In Search of a Fundamental Theory*. Oxford University Press.
- Chalmers, D. J. (2018). The meta-problem of consciousness. *Journal of Consciousness Studies*, 25(9-10), 6–61.
- Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36(3), 181–204.
- Dehaene, S., & Changeux, J.-P. (2011). Experimental and theoretical approaches to conscious processing. *Neuron*, 70(2), 200–227.
- Dennett, D. C. (1991). *Consciousness Explained*. Little, Brown and Co.
- Di Giorgio, E., et al. (2017). Difference in visual social predispositions between newborns at low- and high-risk for autism. *Scientific Reports*, 6, 26395.
- DiCarlo, J. J., Zoccolan, D., & Rust, N. C. (2012). How does the brain solve visual object recognition? *Neuron*, 73(3), 415–434.
- Dunbar, R. I. M. (1998). The social brain hypothesis. *Evolutionary Anthropology*, 6(5), 178–190.
- Farroni, T., Johnson, M. H., Menon, E., Zulian, L., Faraguna, D., & Csibra, G. (2005). Newborns' preference for face-relevant stimuli: Effects of contrast polarity. *Proceedings of the National Academy of Sciences*, 102(47), 17245–17250.
- Frankish, K. (2016). Illusionism as a theory of consciousness. *Journal of Consciousness Studies*, 23(11-12), 11–39.
- Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, 11(2), 127–138.
- Goff, P. (2019). *Galileo's Error: Foundations for a New Science of Consciousness*. Pantheon.
- Hohwy, J. (2013). *The Predictive Mind*. Oxford University Press.
- Jackson, F. (1982). Epiphenomenal qualia. *Philosophical Quarterly*, 32(127), 127–136.
- Johnson, M. H., & Morton, J. (1991). *Biology and Cognitive Development: The Case of Face Recognition*. Blackwell.
- Kobayashi, H., & Kohshima, S. (2001). Unique morphology of the human eye and its adaptive meaning: Comparative studies on external morphology of the primate eye. *Journal of Human Evolution*, 40(5), 419–435.
- Lamme, V. A. F. (2006). Towards a true neural stance on consciousness. *Trends in Cognitive Sciences*, 10(11), 494–501.



- Levine, J. (1983). Materialism and qualia: The explanatory gap. *Pacific Philosophical Quarterly*, 64, 354–361.
- Mundy, P., & Newell, L. (2007). Attention, joint attention, and social cognition. *Current Directions in Psychological Science*, 16(5), 269–274.
- Nagel, T. (1974). What is it like to be a bat? *Philosophical Review*, 83(4), 435–450.
- Navajas, J., Hindocha, C., Foda, H., Keramati, M., Latham, P. E., & Bahrami, B. (2017). The idiosyncratic nature of confidence. *Nature Human Behaviour*, 1(11), 810–818.
- Postmes, T., & Spears, R. (1998). Deindividuation and antinormative behavior: A meta-analysis. *Psychological Bulletin*, 123(3), 238–259.
- Rosenthal, D. M. (2005). *Consciousness and Mind*. Oxford University Press.
- Simion, F., Regolin, L., & Bulf, H. (2008). A predisposition for biological motion in the newborn baby. *Proceedings of the National Academy of Sciences*, 105(2), 809–813.
- Stephens, G. J., Silbert, L. J., & Hasson, U. (2010). Speaker–listener neural coupling underlies successful communication. *Proceedings of the National Academy of Sciences*, 107(32), 14425–14430.
- Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, 28(5), 675–691.
- Tomasello, M., Hare, B., Lehmann, H., & Call, J. (2007). Reliance on head versus eyes in the gaze following of great apes and human infants: The cooperative eye hypothesis. *Journal of Human Evolution*, 52(3), 314–320.
- Tononi, G. (2004). An information integration theory of consciousness. *BMC Neuroscience*, 5, 42.
- Vallortigara, G., Regolin, L., & Marconato, F. (2005). Visually inexperienced chicks exhibit spontaneous preference for biological motion patterns. *PLoS Biology*, 3(7), e208.