

内受容感覚と共感・心の理解の関係に関する理論的検討

— 自己の身体内部の感覚はどのように寄与しているのか？ —

久崎孝浩

Theoretical examination of the relationship between interoception and empathy and understanding minds: How does the sense of self's internal body contribute to empathic response and understanding minds for others?

Takahiro HISAZAKI

[要約] 本論文は、自己の身体内部の感覚である内受容感覚が他者への共感や他者の心の理解といった社会性にどの程度またどのように関わっているのかについて、理論や研究知見に基づいて論考した。まず、内受容感覚が一人称視点といった主観的意識や自己存在感といった自己覚知に寄与し、自他分化に導く可能性を指摘した。しかし一方で、近年、内受容感覚の覚知の精度が高いと模倣・情動伝染・情動的共感の反応も良いことを示唆する研究知見が見受けられ、内受容感覚の覚知の働きを考慮することで、自己覚知を把持しながら他者に適切に共感することを説明できる可能性を検討した。その検討結果に基づいて、内受容感覚の予測や予測誤差を組み込んだミラーニューロンシステムの予測符号化モデルを構築した。また、説得力のある理論や研究知見は乏しいが、他者の心の理解にも内受容感覚の予測が影響する可能性を考慮して、メンタライジングシステムを上位とする、他者の心の理解の予測符号化モデルを構築し、予測と予測誤差に基づいた他者の心の理解のメカニズムの説明を試みた。

キーワード：内受容感覚、主観的意識、自己覚知、予測符号化、共感と心の理解

1. 内受容感覚の基本的な機能は何か？

自己及び他者の心は自身の身体から切り離して理解することができるのであろうか。例えば、ある状況で他者の情動を理解するときに、その他者の情動表出を知覚して何等かのルールや理論に従ってその情動経験を推測しているのであろうか。特にその理解を“共感”(empathy)と呼ぶのであれば、その情動表出に連動して特定の身体内部の感覚が呼び起されなければ、その他者の情動経験の内容を実感するのは難しいのでないだろうか。

近年、この身体内部感覚をめぐる心理的機能が多く議論されている。こうした動向の背後には、私たちが、自身の心の働きや仕組みを意識的に言葉で説明することに限界を感じ、大半の心の活動を占める無意識的なプロセスは身体内部感覚が教えてくれるのではないかという期待があるからなのかもしれない。また、急速に発達する人工知能

(AI) がいずれ人間と同じ感情・情動までももつのではないかと怯えながら、身体による制約のない AI にはそれを実装できないことを期待しているからかもしれない。

内受容感覚には心肺系、消化器系、侵害受容系、内分泌系、免疫系などの多様な生理学的システムが関わっており、各生理学的システムから送られる情報は島皮質 (insula cortex)、体性感覚皮質 (somatosensory cortex)、前帯状皮質 (anterior cingulate cortex) といった脳領域で処理されて、身体の生理学的状態の詳細な表象を形成していると考えられる (Craig, 2002)。その表象のおかげで、私たちに降りかかるストレスに対して素早く変動する“アロスタシス”(allostasis) と、その内部状態はある程度一定に維持されている“ホメオスタシス”(homeostasis) が機能する。アロスタシスとは状況に応じて血圧、体温、血糖値などの最適な目標値が新たに決まる過程であり、ホメオス

タシスとはその目標値に向かって常に血圧, 体温, 血糖値などが調整される過程である。内受容感覚から形成される身体内部状態の表象を目標値に基づいて調整して生体の各システムを安定的に維持する過程がホメオスタシスであり, その表象によって目標値を修正することで身体の各システムを安定的に機能させようとするのがアロスタシスとも言える (Engelen et al., 2023)。このように, 内受容感覚は, 外界の変化に対して私たちの身体の各システムを調整し, 身体内部状態を維持するための役割をもっている。

2. 内受容感覚はどのようにして自己存在感や一人称視点をもたらすか？

2-1. 自己存在感及び一人称視点と内受容感覚とは何か？

ヒトにのみ, 身体各組織の生理学的状態を一括して表象する求心性の伝達経路が確認されている (e.g., Craig et al., 2000)。それ以来, 自分自身を実体として感じる表象がヒトにのみ存在し, それは他の高等類人猿にはないものと考えられている (Craig, 2003)。そうした自己表象やそれを伴った意識的経験に関わるおそらく最も重要な要素が“自分が在る”という感覚, すなわち“自己存在感” (sense of presence) であろう。自分が在る世界が現実あるいは仮想現実 (virtual reality, VR) であれ, イメージされる心的世界であれ, 私たちが自らの意志による何らかの活動で得られた経験を自己のものとして実感する前提として, 自己存在感は必要であろう。VR 技術は近年飛躍的に進歩し, ワイヤレスのヘッドマウントディスプレイを取り付ければ, VR 技術が提示する仮想の世界を認識し, その世界に自分が在ることを実感することができる。自己存在感は定義するならば, ある世界が現実であり, その世界に自分が在るという主観的感覚と言える (Seth et al., 2012)。Seth et al. (2012) は, この自己存在感は内受容感覚状態の予測と, 外受容器官や自己受容器官からの作用を受けた実際の内受容感覚の状態との照合によって生成されるとした。反対に, 内受容感覚状態の予測が実際の状態とずれている場合には自己存在感は薄れるため, 予測のもとになるモデルを更新するか, 実際に内受容感覚状態を変えるために内

受容感覚に関わる組織や臓器に何らかの変化を生じさせるか, その変化を生じさせるための行動が遂行されることになる。

Jicol et al. (2023) は自己存在感が内受容感覚の予測によって成り立つとする仮説に対して1つの具体的な例を挙げている。VR 実験の参加者は, 恐怖の VR を体験する前にそれを知らされると, その VR 体験で生じるであろう恐怖を予測するとともに, その参加者の身体内部では神経生理学的な変化が生じる。その予測される恐怖の強さが実際に VR 環境で生じる恐怖よりも強いまたは弱いならば, その予測される恐怖と実際に生起する恐怖のずれをなくさない限り, VR 環境で自己存在感は生じにくい。そのずれがなく合致すれば, 参加者は VR 環境でも自己存在感を感じる。Jicol et al. (2023) は, この具体的な仮説に基づいて, 内受容感覚を変化させよう喜びや恐れ的情動状態の予測が VR 体験中に自己存在感を喚起させるかを実験的に検討した。その実験で参加者には, VR 体験前にその VR 体験内容から生じるであろう情動の強さが尋ねられ, また VR 体験後に実際に生起した情動の強さや自己存在感の強さが尋ねられた。その結果, 喜びを伴う VR 体験では実際に生起した喜びよりも弱い喜びを予測した (VR 体験での喜びが予測以上だった) 人たちのほうが自己存在感が強く, 反対に, 恐れを伴う VR 体験では実際に生起した恐れよりも強い恐れを予測した (VR 体験で生起した恐れが予測ほどではなかった) 人たちのほうが自己存在感が強いことが明らかになった。これは, 内受容感覚 (とそれに伴う情動) の予測が自己存在感の強さに影響することを示唆するが, 予測される状態と実際に生起した状態がずれていると自己存在感は生じにくいとする本来の仮説を支持しない結果であった。ただし, Seth et al. (2012) や Jicol et al. (2023) の仮説の内容は見直す必要があるかもしれない。予測符号化理論 (Friston, 2010) では, 内的モデルによる知覚の予測はベイズ統計学という事前分布として表現され, 実際に身体器官から来る感覚信号はノイズを伴う確率分布として入力され, その予測と感覚信号の差分として予測誤差が計算される。そして, この予測誤差に基づいて内的モデルの更新のための事後分布が計算されるが, 私たち

が主観的に経験する知覚は、この事後分布（知覚の予測である事前分布が意識されたもの）と考える。しかし、Seth et al. (2012) の提唱する自己存在感の生成モデルでは、内受容感覚の単一的予測ではなく外受容感覚や自己受容感覚の予測も含む複合的な予測が想定されており、その予測と実際に生起する内受容感覚の状態とを照合することで、自己存在感は生成すると仮定している。自己存在感の生成過程を検証するときには内受容感覚の予測だけでなく外受容感覚や自己受容感覚の予測も想定した実験方法を検討するべきであろう。

また、“自分が何かを見ている、感じている、思考している”という意識的経験には“自分が”という“主観的枠組み” (subjective frame) が与えられており、その経験は唯一無二の自分自身のものとして意識される。私たちの脳は無意識的に膨大な情報を処理しているが、その処理の帰結として知覚、認知、情動等の経験は自分自身のものとして単一的に意識されるのである。こうした意識的経験は、物理的環境や精神的環境における独自の身体を中心とした視点で意識されるという意味にとどまるものではなく、経験または脳内情報処理の帰結が自分の単一の意識に帰するという意味をも含む (Zahavi, 2005)。この主観的枠組みをもった意識は“一人称視点” (first-person perspective) と呼ばれてきた (Vogeley & Fink, 2003; Zahavi, 2005)。Azzarini et al. (2019) や Engelen et al. (2023) は、その一人称視点が心臓や胃腸等の内臓からの入力信号 (内受容感覚) の神経的モニタリングによって主観的枠組みが作られることで生じると考えた。

一人称視点が内受容感覚のモニタリングによって生じることを直接示唆するものではないが、知覚・認知の精度や感受性が内受容感覚情報の入力によって向上することを示唆する研究は幾つかある。例えば、Garfinkel et al. (2014) は、心拍情報入力知覚に及ぼす影響を検討している。具体的には、心臓の収縮 (血液が心臓から送られて、動脈の圧受容器が中枢神経系に心拍の強さやタイミングの情報を送信する) 時に中性、恐れ、嫌悪、喜びの4種の表情画像を実験参加者に知覚閾値水準で見せる条件と、心臓の膨張 (圧受容器が働かない拍動と拍動の間) 時に同じ4種の表情画像を

知覚閾値水準で見せる条件を設定し、心臓の収縮・膨張が表情画像の認識に影響を及ぼすかを検証している。その結果、膨張よりも収縮時に見せた恐れ表情画像が認識されやすく、情動的に強く評価されることが分かった。心臓の収縮・膨張のような極短時間周期の内受容感覚変動が知覚・認知の精度を向上させることが示唆された。Park et al. (2014) も、脳波計によって測定される心臓の活動による電位変化 (心拍誘発電位, HEP) が強い人は微かな縞模様を感知しやすいことを明らかにしており、これも内受容感覚入力が知覚の精度を向上させていることを示唆する。また、Marshall et al. (2020) の研究では、表情画像を2回見せた後に微かな縞模様を知覚閾値水準で提示して、縞模様が見えたか否かを判定させる実験を行っている。表情画像の見せ方は“中性→怒り” “中性→中性” “怒り→中性” “怒り→怒り” の4パターンであり、表情画像視認直後に脳波計で心拍誘発電位を測定している。この研究の重要な結果は、“怒り→怒り”の表情画像視認直後で心拍誘発電位が低い人ほど微かな縞模様を感知しやすいということであった。おそらく怒りの表情を2回見た結果として心拍誘発電位が低かった人は迫るであろう脅威に対して心拍変動を予測しており、その予測が脅威をもたらす外的環境への注意機能を亢進させて、実際の縞模様探知の精度をも上げた可能性が考えられる。これは、単なる内受容感覚の入力ではなく、内受容感覚の予測が知覚・認知の精度を向上させた可能性を示唆するものである。

ここまで記した研究は心臓の活動に対する感覚に限定されるが、他の内臓感覚の入力や予測も知覚・認知の精度に影響することが確認されている。例えば、Flexman et al. (1974) は、肺の収縮と膨張である呼吸のリズムが視覚刺激探知の精度に影響することを実証しており、吸気時よりも呼気時に提示された視覚刺激が探知されやすいことを明らかにした。また、参加者の呼吸を記録しながらその参加者が知覚閾値水準で提示される縞模様刺激の位置を感知して回答する実験研究 (Kluger et al., 2021) では、各参加者の正答率と刺激の縞のコントラストの度合いの相互依存度と、正答率と縞のコントラストにさらに呼吸の周期

位相を加えた相互依存度を算出して比較したところ、呼吸の周期位相を加えた相互依存度のほうが有意に大きいことが確認され、知覚・認知の精度が呼吸の周期の影響を受けることが示唆された。

こうした知見を踏まえると、内受容感覚の入力や予測が知覚・認知の精度に影響を及ぼす、あるいは知覚・認知の精度を向上させることは間違いないのであろう。ただし、Azzarini et al. (2019) や Engelen et al. (2023) の仮説である、内受容感覚モニタリングが知覚・認知過程において単一かつ主観的な意識である一人称視点を生成しているということまでは言えない。

2-2. 内受容感覚はどのような神経科学的メカニズムによって自己存在感や一人称視点を生み出すか？

それでは、内受容感覚はどのようなメカニズムで、自己存在感や一人称視点の生成、あるいは知覚・認知の精度に寄与しているのでしょうか。

Seth et al. (2012) は、離人症（自分自身が現実に存在するという実感を失っている状態）の人たちの前部島（Anterior Insula Cortex, AIC）という脳領域の活性が異様に低く（Phillips et al., 2001）、前部島の活性が内受容感覚の意識的知覚に大きく関与しているという知見（Harrison et al., 2017）から、自己存在感は前部島とその関連領域における内受容感覚の予測符号化によって生成されると考えた。前部島で内受容感覚の予測符号化が行われている可能性については、これまでに Paulus & Stein (2006) や Gu et al. (2013) が提唱してきた。Paulus & Stein (2006) は、不安の強い人は身体内部状態の予測と結果の違いを感知して内受容感覚の予測シグナルを更新しやすいため、内受容感覚の予測シグナルを増大させやすく、その分、不安感情や回避行動も強くなると考えた。また、前部島がそうした情報処理において重要な役割を担っているとした。Gu et al. (2013) は、他者の情動状態の理解（共感）において前部島と前帯状皮質が共起するが、特に前部島は前帯状皮質とは機能的に異なり、刺激入力からの情報とより高次レベルの情報を統合する機能を有し、情動状態の知覚において必須であることを研究知見をベース

にして説明している。前部島が、内受容感覚において予測した状態と実際に生じた状態を照合する機能をもっていることは間違いないのであろう。しかし、Seth et al. (2012) の自己存在感の生成モデルに依拠すれば、内受容感覚のみの予測だけでなく外受容感覚・自己受容感覚の予測も関与しており、そうした情報の処理過程にどのような神経メカニズムが存在しているかまではよく分かっていない。

また、一人称視点の生成において内受容感覚の情報処理が寄与する神経メカニズムも全く明らかになっていない。Tallon-Baudy (Azzarini et al., 2019; Engelen et al., 2023; Tallon-Baudy, 2023) は、内臓には一定リズムの運動があり、その内受容感覚シグナルが各脳領域間の時間的相関（本来、時間的相関とは神経細胞間で発火が同期することであり、その同期は神経細胞集団レベルでの情報処理に影響を与えるものであるとされるが、ここでは離れた脳領域間で情報処理のタイミングが同期するという意味で用いている）を引き起こし、それによって離れた脳領域間で符号化された情報が結び付けられて単一の意識的知覚がもたらされるという仮説を打ち立てている。ただし、各内臓の一定リズムの運動周期は異なっている（心臓は1秒以内、肺は5秒以内、胃腸は20秒以内）ため、異なる運動周期の内受容感覚シグナルがどのように各脳領域の情報処理を結びつけているかまでは想定されていない。

3. 内受容感覚は共感性・社会性に関与するのか？

3-1. 内受容感覚は確固とした自己覚知・自己表象を築くのか？

内受容感覚のシグナルやその覚知は自己存在感や一人称視点以外に、“自己覚知” (self-awareness) に影響していることも確認されてきた。自己覚知とは、Zahavi (2005) の言う“前意識的自己覚知” (pre-reflective self-awareness or self-consciousness) と同義である。Zahavi (2005) は、前意識的自己覚知を、本質的に自分自身を対象化していない自己理解であり、その自己理解のあり方に主体-客体という構造はなく、単一の方法で自分自身が意識されるとしている。より具体的に

言えば、それは、“これを動かしているのは自分自身だ”“今これについて考えているのは自分自身だ”という感覚である“主体感”(sense of agency) (Moore & Fletcher, 2012) や“この手は自分自身のものだ”“この感情を経験しているのは自分自身だ”といった感覚である“所有感”(sense of ownership) (Gallagher, 2000) を含む。

その自己覚知を身体基盤とする仮説・理論の多くは、生成される自己覚知の中心に内臓器官やその神経表象を想定してきた。例えば、Damasio (1999) は、意思決定に影響する身体からの信号は自己意識にも影響するとしてソマティックマーカー仮説を提唱した。Craig (2009) は主観的感覚の生成において内受容感覚の神経表象を構成する前部島の神経学的構造や機能に注目している。Seth (Seth et al., 2012; Seth & Tsakiris, 2018) は前述の自己存在感でも記したように、自己覚知は身体と外的世界との相互作用によって成立し、その際に内受容感覚の予測符号化が重要な役割を果たすと考えた。さらに、Owens et al. (2018) は、自律神経系に予測できない変化が生じると知覚的ノイズが大きくなり、そのノイズの大きさが知覚的判断に対する自信を降下させる (Allen et al., 2016) という知見等を根拠として、自己覚知が内受容感覚の予測に基づいて生成されるだけでなく、“内受容感覚による自己の予測”(interoceptive self-inference) によって意識的な“自己内省”(self-reflection) が生じるのではないかと考えた。

自己覚知に内受容感覚のシグナルが影響することを実証している複数の実験研究では、例えば、内受容感覚を精確に覚知しないあるいはできない場合に、“身体所有感”(sense of body ownership) が強く歪められることが明らかとなっており、自己覚知が外受容感覚の影響を受けやすくなることが示唆されている (Ainley & Tsakiris, 2013; Moseley et al., 2008; Tajadura-Jiménez & Tsakiris, 2014; Tsakiris et al., 2011)。また、Babo-Rebelo et al. (2016) の研究では、デフォルトモード・ネットワーク領域の脳部位(腹側被蓋野と腹内側前頭前野)における心臓の拍動に対する神経反応が、自由思考を許されている最中に思考主体として自己を感じている程度(主体性)や思考対象として自己を注視している程度(客

体性)と共変することが確認されており、内受容感覚のシグナルが思考の主体性や客体性を強めている可能性が示唆されている。

こうした研究知見や提唱されてきた仮説・理論からすると、内受容感覚シグナルやその覚知の精度の強さは自己覚知を強めることは間違いないであろう。加えて、自己覚知が強まれば、自己存在に対する他者への意識も明確になって自他分化(self-other distinction)も強まることが想像される。もっと言えば、内受容感覚のシグナルは自己覚知を確固としたものにするが、一方で他者との繋がり、他者への共感、他者の心の理解とは無縁であり、それらはむしろ外受容感覚や自己受容感覚をベースにして成り立つものとも考えることができる。例えば、Palmer & Tsakiris (2018) は、予測符号化理論に基づいて、確率論的に生起する自己表象は、内受容感覚の予測と内臓や感覚器からの入力によって、内受容感覚や外受容感覚の予測誤差が統合されることで生じると主張する。そして、安定した自己表象は、外受容感覚の予測誤差の精確さを内受容感覚の予測誤差の精確さが上回ることで生じ、その両者間の精確さのバランスは注意(attention)によって最適化されるという。すなわち、外受容感覚の予測誤差の精度が内受容感覚の予測誤差の精度に比して低いと、内受容感覚の入力が外受容感覚器からの入力(外界知覚)よりも優位になり、比較的安定した内受容感覚をベースとする自己表象が形成されるというのである。

3-2. 内受容感覚は自己と外的環境・他者を隔てるものか？

しかしながら、広く、内受容感覚の研究に目を向けてみると、内受容感覚のシグナルやその覚知が他者との繋がりや他者への共感と緊密な関係にあることが窺え、内受容感覚の機能や内受容感覚と外受容感覚・自己受容感覚の関係について新たに大きな疑問が浮かび上がる。例えば、Fukushima et al. (2011) は、内受容感覚のシグナルを脳内で処理している程度を把握する指標として特定の課題を行っている最中の心拍誘発電位を測定し、それが質問紙で測定される共感性と正の相関関係を示すことを明らかにしている。また、

自身の心拍を精確に捉える程度（内受容感覚の覚知のレベルの高さ）が、他者に対する模倣のしやすさ（Ainley et al., 2014）、他者の喜びや悲しみの表情に対する敏感さ（Terasawa et al., 2014）、他者と目を合わせたときの“表情模倣”（facial mimicry）の生起しやすさ（Imafuku et al., 2020）、物語に登場する人物の情動を理解する能力（Shah et al., 2017）に促進的に関与することが明らかになっている。こうした知見は、他者の動きや表情を模倣したり、他者の情動を理解したりすることに対して、内受容感覚のシグナルやその覚知の精度に関わる何らかのメカニズムが手助けしていることを示唆する。

しかし、Tsakiris (Ainley et al., 2014; Palmer & Tsakiris, 2018) は、内受容感覚の覚知の精度が高い人は確固とした自己表象を形成し、自己と他者を明確に区別することができるため、他者の動きに対する模倣を抑止する傾向も高いと考えている。それは予測符号化理論が論拠になっており、内受容感覚の覚知の精度が高い人は内受容感覚の予測の精度が良いので、実際の身体内部の生理的状态を差し引いた内受容感覚の予測誤差の精度も良く（大平, 2018）、内受容感覚の予測誤差は外受容感覚の予測誤差の精度を上回り、内受容感覚ベースの自己表象は外受容感覚ベースの他者表象と混同されずに他者表象とは区別されるというのである。なお、先でも述べたが、内受容感覚の予測誤差の精度と外受容感覚の予測誤差の精度の両者間バランスは“習慣的な注意”（habits of attention）によって調整されるという（Ainley et al., 2016）。しかしながら、これは、身体に関する複数感覚間の統合に限って成り立つ仮説なのではないだろうか（Ainley et al., 2014）。例えば、Tsakiris らの一連の研究（Ainley & Tsakiris, 2013; Tajadura-Jiménez & Tsakiris, 2014; Tsakiris et al., 2011）や Moseley et al. (2008) の研究結果より、内受容感覚の覚知の精度が低い人は身体所有感が視覚等の外受容感覚の入力によって歪められる可能性を上述したが、そもそも身体所有感は視覚と体性感覚という2種の感覚間の統合によって成立するものである。一方、他者の動きに対する模倣は、他者の動きに関する視覚情報だけに依拠した表象レベルの現象で、自己の動き

に関する感覚情報を必要とせず、複数感覚間の照合によって成り立つものではない。自己表象や自己分化をどのような課題や指標で測定・把握するかが、内受容感覚と自己分化の関係に対する理解のあり方の別れ道になるのではないだろうか。ラバーハンド錯覚のような、自己身体の表象が複数感覚間の統合によって生じる現象の場合、内受容感覚の覚知の精度の高さに基づく内受容感覚入力への優位化によって自己身体の表象が確固としたものになれば、外受容感覚入力の影響を受けにくくなり、明確な自己分化が成り立ったと解釈できる。しかし、他者の動きに対する模倣の場合、ミラーニューロンシステム（Rizzolatti et al., 2001）と呼ばれる脳内の上側頭溝（superior temporal sulcus）、下頭頂葉（inferior parietal lobe）、腹側運動前野（ventral premotor cortex）の活動が関与しており（Iacoboni et al., 1999; Nishitani & Hari, 2002）、他者の表情の模倣においてもミラーニューロンシステム内の腹側運動前野の活動が関与していることが確認されている（Leslie et al., 2004; Likowski et al., 2012; Sato et al., 2015）。他者の動きや表情に対する模倣は、このミラーニューロンシステムを通じて他者身体の表象と自己身体の表象が共有されて自己未分化になることで生じるものと考えられる。さらに詳しく、内受容感覚のシグナルが他者身体の視覚入力・表象及び自己身体の表象の形成にどのように作用し、またミラーニューロンシステム内でどのように機能するのかが明らかになれば、内受容感覚の覚知の精度の高さが自己分化に繋がるのか、それとも自己未分化に寄与するのかを検討することができるであろう。Tsakiris (Ainley et al., 2014; Palmer & Tsakiris, 2018) は、内受容感覚の覚知の精度が高い人は自己と他者を明確に区別ことができ、他者の動きに対する模倣を抑止する傾向が高いという仮説を立てたが、その前に、模倣における内受容感覚のシグナルの詳細な作用や機能を検討しなければならなかったのではないだろうか。

なお、他者の動きに対する模倣を抑止するためには、予測符号化理論で考えれば、自己受容感覚の予測誤差の精度を下げる必要がある（Friston et al., 2011）。実は、Tsakiris も、内受容感覚の覚知の精度が高い人は模倣しやすいという自身の

研究結果 (Ainley et al., 2014) に対して、内受容感覚の覚知の精度が高い人は初めから自己受容感覚の予測誤差の精度も高かったのではないかと考察しており、内受容感覚の予測誤差の精度が自己受容感覚の予測や予測誤差の精度に関与する可能性を暗示する。

3-3. 他者との関わりにおいて内受容感覚は外受容感覚・自己受容感覚やミラーニューロンシステムに関与するのか？

内受容感覚は外受容感覚や自己受容感覚に比べてそのシグナルの発生場所や発生のタイミングを特定しがたく、たいていは意識に到達しない (Ondobake et al., 2017)。それは、予測符号化の考えに基づけば、絶え間なく変化する身体内部のシグナルに対して、その予測を行う内部モデルの予測精度が低く、様々な受容体によってそのシグナルがキャッチされても予測誤差の精度も低いいためだと考えられる。それゆえに、予測や予測誤差の精度の低い内受容感覚は外受容感覚や自己受容感覚と殆ど連携することがないと考えられてきた嫌いがある。また、例えば自分の動きを視認できるような自己受容感覚と外受容感覚の結びつきが内受容感覚にはほぼないため、内受容感覚の精度は自身の内受容感覚の内部モデルによる予測に大きく頼ることになる (Ondobake et al., 2017)。そして、他者の内受容感覚も外受容感覚で確認しにくいいため、他者との関わりでは、他者の内受容感覚に基づいた情動や意図の推測を自己自身の内受容感覚の内部モデルの予測に頼ったとしても、その予測精度は低いため、他者の情動や意図の正確な推測に貢献しないとも考えられ得るのである。

それに対して、例えば、ある対象物を視認しながら操作しようとする際に、その操作の結果として対象物をコントロールできないことが視認されれば、操作に関わる運動の予測が修正されて新たな運動による操作が生み出されるというように、外受容感覚と自己受容感覚は連携して一致する傾向がかなり高い (Ondobake et al., 2017)。他者との関わりにおいては、他者が何らかの対象物をつかもうとする運動とその結果は観察することができて、その視覚情報を自分自身の運動システムにマッピングすれば、他者の運動を実際に模倣した

り他者の運動の背後にある意図を推測したりしやすくなる。

こうした論からすると、内受容感覚は外受容感覚や自己受容感覚と全く連携していないように思われるが、果たしてそうなのであろうか。Stephan et al. (2016) によれば、全ての行動は内受容感覚を安定的に維持するために選択されたり調整されたりするという。具体的には、まず、身体外部環境の刺激に対してその刺激に適応しようとする (アロスタシスの) ために身体内部状態の変化の予測が生じ、その予測によって身体内部状態の目標値が更新される。続いて、その目標値に対して身体内部の生理的状態のシグナルが入力されて、その目標値に対する入力差分 (予測誤差) がゼロになるように (ホメオスタシスのための) 行動が選択・調整され、その差分がゼロになるまで行動選択・調整は継続されるというのである。この Stephan et al. (2016) の考えからすれば、身体外部環境の刺激に対する身体状態変化の予測には外受容感覚の入力や予測が関与し、身体内部状態の目標値と入力差分から生じる行動選択・調整には運動に関わる自己受容感覚の入力や予測が関与していると考えることができる。実際に、内受容感覚と外受容感覚との連携については、上述したように、心臓や肺の動きに関する入力や予測が視覚の精度を向上させることが確認されている (Garfinkel et al., 2014; Kluger et al., 2021; Marshall et al., 2020; Park et al., 2014)。また、内受容感覚と自己受容感覚との連携については、例えば Kleckner et al. (2017) や Megiot et al. (2022) が、アロスタシスにおける内臓 (内受容感覚) と運動 (自己受容感覚) の間の神経シグナル伝達をモデル化して検討している。具体的にそのモデルは、次のとおりである。身体外部環境の刺激に対する適応において、前部・中部帯状皮質、扁桃体背側部、前部島腹側部が内臓活動と身体運動を通じて身体内部状態を変化させるために内臓と筋や腱・関節に予測シグナルを伝達する。一方、これらの脳領域からは身体内部状態 (内受容感覚) の予測シグナルも中部島皮質と後部島皮質の各背側部に送りこまれて、そこでは身体内部から実際に生じる内受容感覚の入力によって内受容感覚の予測誤差が生じる。その予測誤差は、前部・中部

帯状皮質、扁桃体背側部、前部島皮質の腹側部にフィードバックされて、内臓活動や身体運動の内部モデルは更新されて、新たな予測シグナルが内臓と筋や腱・関節に送られて、内臓活動と身体運動が変化する。このモデルでは、内受容感覚の予測誤差によって筋や腱・関節等の自己受容感覚の予測が更新されることが想定されており、内受容感覚が自己受容感覚と連動していることの論拠の1つになる。

さて、他者との関わりにおいては、内受容感覚の覚知の精度の高い人は他者の表情に敏感で模倣しやすく、ひいては他者の情動も理解しやすいこと (Ainley et al., 2014; Fukushima et al., 2011; Imafuku et al., 2020; Shah et al., 2017; Terasawa et al., 2014) を先述したが、これは、内受容感覚に関する精度の高い予測や予測誤差のシグナルが、表情模倣、“情動伝染”(emotional contagion)、“情動的共感”(emotional empathy) に大きく関わるミラーニューロンシステムの中で機能しているからではないかと考えられる。

それでは、ミラーニューロンシステムはどのようにして、他者の運動に対する視覚表象と自己の運動表象を直接結びつけているのであろうか。Kilner et al. (2007) は予測符号化の枠組みで、以下のような計算論的モデルを提唱している。他者の運動を見ると、前頭葉のミラーニューロンでは予測的に生成された他者の運動目標から運動指令が予測されて頭頂葉のミラーニューロンに伝達されて、続いてその頭頂葉ミラーニューロンでその運動指令の予測から具体的な運動身体情報が予測されて上側頭溝に送信される。その上側頭溝では、その運動身体情報の予測が実際に目の当たりにしている他者の運動身体情報に対してどれ程ずれているかが計算されて予測誤差が生じる。この予測誤差が今度は逆行して頭頂葉ミラーニューロンに送られて運動指令の表象が更新され、さらにその更新によって運動指令に関する予測誤差が生じて前頭葉ミラーニューロンにフィードバックされて、他者の運動目標の表象も更新される。他者の運動身体情報の視覚入力に予測される自己の運動身体情報とは大抵大きな差異があり、その両者の予測誤差を最小化するために運動指令は大きく調整(運動指令の予測は更新)される (Shimada,

2022) ため、その予測誤差が大きいほど他者の運動に対する模倣が生じやすいと考えられる。

こうしたミラーニューロンシステムに内受容感覚に関わる神経システムが関与することを想定すれば、予測符号化理論という枠組みで、内受容感覚の予測あるいは予測誤差の精度の良さが表情模倣・情動伝染・情動的共感に対して促進的に影響することを説明できるのではないか。試みとして、ミラーニューロンシステムに内受容感覚の予測と予測誤差を組み込んだモデルを Figure 1 に図示した。

(1) 外受容感覚に対する内受容感覚の作用：まず、先述のとおり、心臓や肺の動きに関する入力や予測が視覚覚の精度を向上させることが確認されており (Garfinkel et al., 2014; Kluger et al., 2021; Marshall et al., 2020; Park et al., 2014)、内受容感覚の予測の精度が良いと内受容感覚の予測誤差の精度も良く、その予測誤差は視覚等の外受容感覚の内部モデルが更新される際に用いられると考えられる。その内部モデルの更新によって、身体外部環境にある刺激に対して視覚覚等の外受容感覚は敏感になるものと考えられる。

(2) 内受容感覚に対する外受容感覚の作用：反対に、外受容感覚が内受容感覚に作用する例はないだろうか。Marshall et al. (2017) は、怒りの表情を複数回連続して見せた実験参加者において、脳波計の計測によって心拍誘発電位に反復抑制 (repetition suppression) が生じることを確認している。これは、表情を繰り返し見ることによって視覚的な期待が形成され、それが心臓の動きに関する情報処理に影響したものとする。視覚的な期待は外受容感覚の予測に相当し、心拍誘発電位の大きさが心拍モニタリングの精度を表しているとするなら、外受容感覚の予測が内受容感覚の予測誤差を計算する領域に影響しているということはあるかもしれない。

(3) 自己受容感覚に対する内受容感覚の作用：先述のとおり、Stephan et al. (2016) の言う、内受容感覚の予測誤差を最小化するために全ての行動が選択されたり調整されたりするという見方からすれば、内受容感覚が自己受容感覚に影響を及ぼすことはあると考えてよいだろう。これを実証している研究は確認されていないが、多くの行動

によって身体運動や内臓運動が引き起こされて内受容感覚の入力も変更し、最終的に内受容感覚の予測誤差を最小化するという事は考えられよう。そこで、内受容感覚の予測誤差が自己受容感覚の予測に作用することは想定できる。

(4) 内受容感覚に対する自己受容感覚の作用：一方、内受容感覚に対する自己受容感覚の作用を説明する理論や例証・実証した研究は確認されないが、運動指令に伴って自己受容感覚の予測シグナルが内受容感覚の予測誤差の計算を行う領域に伝達されて、内受容感覚の覚知の精度が引き上げられるということはあるのではないだろうか。例えば、内受容感覚の予測は身体内部状態の情報を提供して“基本的情感” (basic affective feeling) を生み出すと考えられている (Barrett, 2017) が、情動が身体的に表出される直前に、その表出に関わる自己受容感覚の予測が内受容感覚の予測誤差に作用して身体内部感覚を伴う基本的情感を生み出している可能性がある。そこで、自己受容感覚の予測シグナルが内受容感覚の予測誤差に作用するパスも想定した。

(5) 外受容感覚と自己受容感覚の間の相互作用：そもそも外受容感覚と自己受容感覚は緊密で一致する傾向が高いことを先述した (Ondobake et al., 2017)。先で示した Kilner et al. (2007) の、ミラーニューロンシステムを予測符号化で捉える見方では、運動指令の予測から自己の運動身体情報の予測が上側頭溝に送られて、実際に見ている他者の運動身体の見視情報も上側頭溝に送られ、上側頭溝で両者の予測誤差が計算され、その予測誤差が大きければそれは逆行して運動指令の予測が更新されると仮定している。他者の表情・動きに対する反応ではないが、自己生成反応に基づいた運動主体感の研究においても、運動指令から予測される見視入力に対して実際に生じた見視的結果がフィードバックされて、その両者の予測誤差が小さければ運動主体感が生じることが想定されている。その予測誤差が大きければ運動指令の予測が更新されて新たに運動が試みられ、予測誤差が最小化されて運動主体感が感じられたところで運動指令の予測の更新と運動による試行が終了することになる (Wen & Haggard, 2019)。他者の身体運動の見視情報でも、運動主体感の場合と同

じように外受容感覚と自己受容感覚の相互作用があるものと考えてよいのではないだろうか。以上より、おそらくミラーニューロンシステムでは、自己受容感覚の情報処理領域におけるある特定の内部モデルによる予測が、外受容感覚の情報処理領域における予測誤差を計算する領域に送られて、計算されたその予測誤差が自己受容感覚の内部モデルを更新するという緊密な連動が生じているものと想定した。

Figure 1における上記 (1) ~ (5) を想定すれば、内受容感覚の覚知の精度が高い人は内受容感覚の予測誤差の精度が高く、その予測誤差は視覚等の外受容感覚の予測の精度を高めて身体外部環境の刺激に対する外受容器官の精度を高めるとともに、自己受容感覚の予測の精度を高めて筋や関節・腱の反応性も高めることが考えられる。そうであれば、例えば、他者の表情や動きへの視覚入力の精度は高まってその視覚入力に対する予測誤差の精度も高まり、その予測誤差によって自己受容感覚の予測の精度がさらに高まって、他者の表情や動きに対する模倣も生じやすくなるものと思われる。

また、他者の表情や動きの見視入力に対する予測誤差の精度が高まることで視覚の予測の精度も高まり、それが内受容感覚の予測誤差の精度を引き上げて内受容感覚の予測の精度も向上させて、他者の表情や動きに対する身体内部状態を正確に覚知しやすくなることも想定できる。情動は基本的に内受容感覚の入力や予測が基盤にあると考えられる (Barrett, 2017; Craig, 2007)。したがって、上記のような、内受容感覚の影響を想定することによって、ある他者の表情や動きを見た者が、その他者の内受容感覚や自己受容感覚に似た感覚を伴って直接的あるいは代理的にその他者の情動を経験することを説明することができる。ここでいう直接的な経験とは情動伝染のことを、代理的な経験とは情動的共感のことを指す。情動伝染が自己覚知・自己表象が確立しておらず自他の区別なく他者の情動を経験することであるのに対して、情動的共感とは自己覚知・自己表象が確立しているために他者のものとして他者の情動を経験することとして区別される (Asada, 2015)。それゆえ、自己覚知・自己表象と内受容感覚の関係に関する

先述の論からすれば、内受容感覚の予測誤差の精度が低い場合には自己覚知も弱まって情動伝染が生じやすくなるが、内受容感覚の予測誤差の精度が高い場合には自己覚知は確固としたものになって情動的共感が生じやすくなるということも考えられる。

4. 内受容感覚は他者の心の理解にどのように寄与しているか？

4-1. ミラーニューロンシステムとメンタライジングシステムはどのような関係にあるか？

上述では、内受容感覚の覚知の精度の高さが他者への共感にいかに関与するかについて、理論と知見をもとに予測符号化理論をベースとした説明を試みた。その参照した知見の1つ (Shah et al., 2017) を振り返ると興味深いことに、自身の心拍を知覚する精確さは情動的共感の項目得点との間に有意な正の相関があるのに対し、非情動的共感の項目得点との間には相関関係がなかったのである。この非情動的共感とは“認知的共感” (cognitive empathy) のことであり、“心の理論” (theory of mind) とおよそ同じ定義を共有し、情動的共感とは情動伝染がベースにあるのに対して、認知的共感とは“視点取得” (perspective taking) をベースとしている (Asada, 2015; de Waal, 2008; Shamay-Tsoory et al., 2009)。また、情動的共感では上側頭溝、下頭頂葉、腹側運動前野を含む脳領域で構成されるミラーニューロンシステムが機能し、認知的共感では、側頭葉先端部 (temporal pole)、上側頭溝後部 (posterior superior temporal sulcus)、内側前頭前野 (medial prefrontal cortex) を含む脳領域で構成されるメンタライジングシステムが機能している、それぞれ独立したシステムであることが仮定されてきた (Asada, 2015; Shamay-Tsoory et al., 2009)。もしそうならば、心拍知覚 (内受容感覚) の精度の高さが情動的共感と正の相関関係にあるのに対して、非情動的共感 (認知的共感) とは無相関であることは首肯できる。

しかし、他者の表情や動きからその心を精確に推測することに関して、両システムが独立的に機能すると考えるのは現実的・効率的ではない。現に、様々な研究が、他者の心の推測に関する課題

においてミラーニューロンシステムとメンタライジングシステムの相互作用が重要であることを指摘してきた (Sperduti et al., 2014; Spunt & Lieberman, 2012b; Ciaramidaro et al., 2014; Arioli et al., 2017)。例えば、Sperduti et al. (2014) は、手遊びで他者と相互に模倣し合う時と他者の手遊びを観察する時の脳活動を fMRI (functional magnetic resonance imaging) で観察記録したところ、模倣し合う時にミラーニューロンシステムとメンタライジングシステムの緊密な結びつきが確認された。Sperduti et al. (2014) は、社会的相互作用において2つのシステムが補完的關係にあるとして、ミラーニューロンシステムは自身の行動の準備と他者の行動のシミュレーションを担うが、メンタライジングシステムは他者の意図を推測して相互にやりとりされる行為を共同で調整することに関与すると論じている。また、Arioli et al. (2018) は、相互協力的に振る舞う二人の様子が映った写真と相互に愛情豊かに振る舞う二人の様子が映った写真を実験参加者に見せたときに、実験参加者の (ミラーニューロンシステムとメンタライジングシステムに該当する) 各脳領域間のどこに効果的な結びつきが生じるのかを fMRI で調査した。その結果、相互協力の写真を見るときには上側頭溝後部と上頭頂皮質 (superior parietal cortex)・腹側前運動野との (ミラーニューロンシステム内の) 結びつきが亢進し、愛情豊かな写真を見るときには上側頭溝後部 (ミラーニューロンシステムに含まれる脳領域) と腹内側前頭前野 (ventromedial prefrontal cortex) (メンタライジングシステムに含まれる脳領域) の間の結びつきが亢進することが明らかになった。前者の結果は、写真が相互に協力し合う様子を表しているため、他者の行動目標を共有するのに必要な視覚-運動情報のボトムアップ型処理を担うミラーニューロンシステム内の3つの脳領域が緊密に連携したと解釈できる。後者の結果は、写真が相互に感情や思いを共有している様子を表しており、他者の感情や心を推測するためのトップダウン型処理が誘発されて、ミラーニューロンシステムとメンタライジングシステムが緊密に連携したと考えられる。こうした知見から、メンタライジングシステムはミラーニューロンシステム

の上位にあり、特に他者の心を推測する認知的共感においては、メンタライジングシステムとミラーニューロンシステムの強い連携が必要であることが推察される。

4-2. 他者の心の理解に内受容感覚は必要か？

他者の心の理解にあたる認知的共感、視点取得、心の理論において内受容感覚は必要であろうか。自身の心拍を知覚する精確さは認知的共感とは無関係であるという先に挙げた知見 (Shah et al., 2017) からすれば、必要ではないということになる。しかし実は、内受容感覚と認知的共感の結びつきを示唆する研究知見は幾つかある。その1つ (Grynberg & Pollatos, 2015) は、自身の心拍を知覚する精度が高いほど、痛ましい状況を映した写真を見て経験する同情 (情動的共感) の程度が強く、その状況の痛ましさを評価 (認知的共感) の程度も強いことを明らかにしている。また、内受容感覚の予測が心の理論に貢献することを理論的に説明した Ondobake et al. (2017) は、単に内受容感覚が外受容感覚や自己受容感覚と統合されることで心の理論が働くのではなく、多様相かつ階層型の構造の中で、他者を説明する各様相の感覚情報の予測誤差が求心的にフィードバックされてより高次の概念的レベルの内部モデルが更新されることで、他者の心は推論されるとしている。これは、先述の Ariori et al. (2018) の結果が示唆することと同じで、高次の概念的レベルの内部モデルを含む脳領域がメンタライジングシステムであり、内受容感覚・外受容感覚・自己受容感覚の予測誤差を生み出す脳領域がミラーニューロンシステムに相当すると考えてよいのではないだろうか。それでは、心の理論 (他者の心の理解) のモデルにおいては、内受容感覚はどのように機能・貢献しているのであろうか。1つの例として、暑い夏に母親が見守る中で子どもが水遊びをしていて、その時間が長いので母親が家に戻ろうと誘ったところ、子どもが反発したという場面を想像してみよう。その場面を目撃した者は、子どもの反発する行動からミラーニューロンシステム内で視覚・運動情報が処理されて“家には戻らない”“もっとたくさん水遊びをする”という行動目標を推測することができるだろう。しかし、“なぜ”

“どのような心的表象”がその行動目標を生み出しているかは理解できない。そこに、自己の内受容感覚状態の予測があれば、気温が高く身体内部温度が上昇しやすい状況で水に触れる気持ちよさが予測されて、子どもが水遊びを楽しんでいる“理由”を理解することができる。また、暑い中で水に触れる気持ちよさの予測から、“家に戻りたくない”“もっと水遊びをしたい”という“欲求”を想像することができる。すなわち、内受容感覚の予測は、他者の行動に対して身体感覚的な理由を与えてその行動の目的・目標の理解を推し進めるとともに、具体的な心的表象を与えてその行動から心を的確かつスムーズに理解できるようにするという2つの重要な働きを担っていると考えられるのである。

なお、上記の、子どもが水遊び中止に抵抗する例は、Figure 1のミラーニューロンシステムの予測符号化モデルでは、“以下のことまで”は説明することができる。暑い中で子どもが水遊びをしていて母親の制止に反発する様子を見た観察者においては、最初に、その情報が視覚から入力されて視覚の大きな予測誤差が生み出され、その予測誤差のシグナルによって視覚及び自己受容感覚の内部モデルは大きく更新され、自己受容感覚の予測は子どもの反発する行動や怒り表情の運動表象を形成する。続いて、更新した視覚の内部モデルから生じる予測は、内受容感覚の予測誤差に変化をもたらして内受容感覚の内部モデルを通じて内受容感覚の予測を生んで (内受容感覚の覚知の精度が高い人は特に予測誤差や予測が大きいため)、暑い中で水に触れる身体感覚や気持ちよさが精確に予測される。また内受容感覚の予測誤差は自己受容感覚の予測にも作用して、身体感覚や気持ちよさに関わるシグナルは先の運動表象に入力されて、より具体的になった運動表象によって子どもの反発する行動や怒り表情の意味が理解されやすくなる。しかし、“家に戻りたくない”“もっと水遊びをしたい”といった欲求や思い等の心的表象をミラーニューロンシステム内で生成するには限界がある。他者の行動の目的・目標に合った心的表象の生成には、メンタライジングシステムが不可欠になってくるのである。

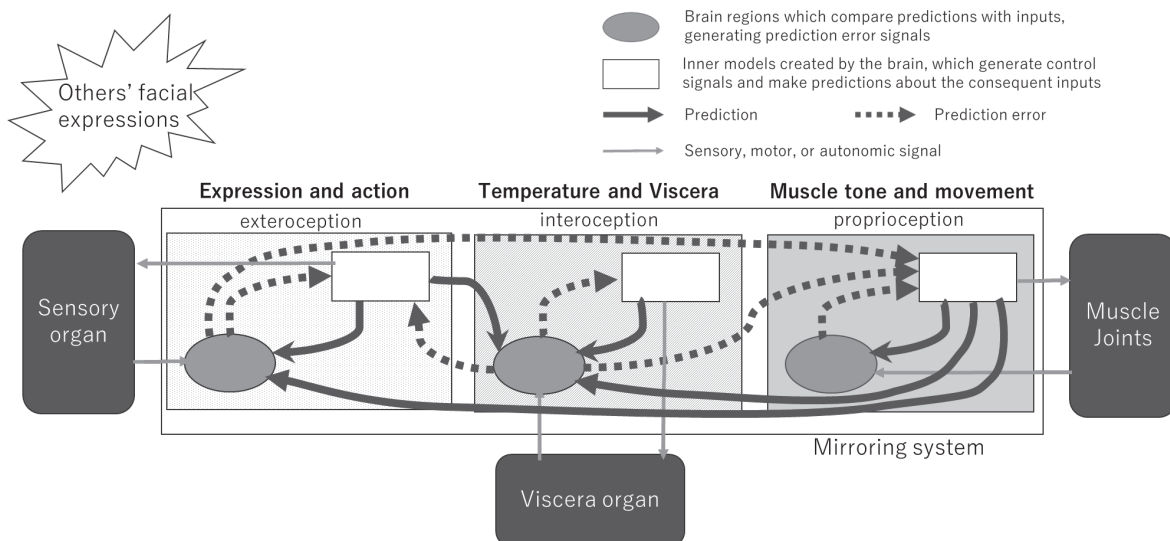


Figure 1 The predictive coding model of mirror neuron system

4-3. 内受容感覚の予測を伴った他者の心の理解をモデル化できるか？

他者の心の理解を予測符号化理論の枠組みで説明するモデルは、Ariori et al. (2018) の結果の示唆する内容や Ondobake et al. (2017) の考えに依拠すれば、ミラーニューロンシステムの上位にメンタライジングシステムを配した構造がおそらく適する。では、ミラーニューロンシステムとメンタライジングシステムはどのように連携しているのでしょうか。これについて具体的な理論や証左が現在無く、Figure 2 のようなモデルを試行的に構成した。以下に、その具体的な仮説を述べる。メンタライジングシステム内に1対の心的表象の予測誤差を計算する領域と予測を生み出す内部モデルをシンプルに想定するが、心的表象を生成するためにはまずその予測誤差計算領域に、ミラーニューロンシステム内の外受容感覚・内受容感覚・自己受容感覚の予測（表象）がシグナルとして入力される必要がある。それによってメンタライジングシステム内に予測誤差が生じて、心的表象を予測する内部モデルが更新されて外受容感覚・内受容感覚・自己受容感覚の各表象に合った心的表象が予測される。その心的表象の予測が実際に外受容感覚・内受容感覚・自己受容感覚の表象と整合していて腑に落ちるものであれば、その心的表象の予測はメンタライジングシステム内の予測誤

差を最小化するだけでなく、外受容感覚・内受容感覚・自己受容感覚の各予測誤差をも最小化して、心的表象も外受容感覚・内受容感覚・自己受容感覚の表象も更新されなくなり、他者の心の推測は終息する。

他者の行動が自分も経験したことがある馴染み深いものでかつ自分の経験と区別され、また他者や他者が置かれている状況に関する情報・知識が豊かであれば、その他者の心は的確に推測されて直ぐに終息する。他者の心の推測はその意味で文脈依存적であり（磯村, 2018）、その推測が難しい場合には、自分も持っている経験と知識を動員して心的表象を予測しても関係する各予測誤差を最小化できずに、幾度も、心的表象の内部モデルの更新によって新たな心的表象が生み出されることになる。例えば、寒い雨の中で人が少し険しい表情で走っている様子を見た者であれば、ミラーニューロンシステム内では“不快感を避ける目的で急いでいる”ことが理解され、さらにメンタライジングシステム内では経験や知識が手伝って心的表象の内部モデルは1回の更新で済んで、“雨に濡れたくない”“体を冷やしたくない”という心的表象の予測によって心の推測はすぐに終わる。しかし、寒い中で人が険しい表情をして何度も水を浴びて震えながら何かを唱えている様子を見た場合、その行動の文化的意味を知らない者で

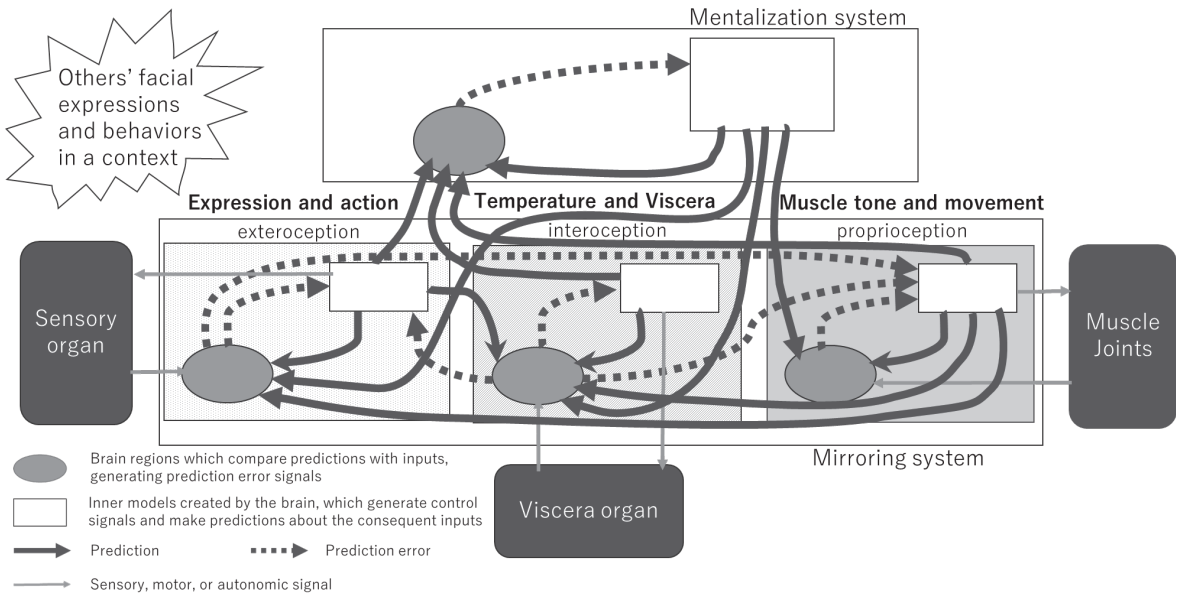


Figure 2 The predictive coding model of understanding others' minds

あれば、まずミラーニューロンシステム内で、視覚における予測誤差から自己受容感覚において水を浴びて唱える運動や苦しい表情や体の震えの予測が生じるが、水を浴び続けることによる身体感覚や痛みなどの予測誤差のシグナルは先の自己受容感覚の予測には部分的にしか貢献せず、苦しい表情や体の震えが身体感覚や痛みによるものと理解できても、水を浴びて唱える行動の理由は理解できない。さらに、メンタライジングシステム内では、経験や知識がない限り、水を浴びて唱え続ける行動の視覚表象、寒い中で水を浴びることによる身体感覚や痛みの表象、苦しい表情や体の震えや水を浴びて唱える行動に関する運動表象それぞれに整合した1つの心的表象を予測するのが難しいため、視覚・内受容感覚・自己受容感覚の各予測誤差は最小化されずに心的表象は繰り返し予測されるか、予測誤差を最小化できない違和感を残しながら心的表象の予測を諦めるということになる。もし水行・水垢離に関する経験や知識があれば、“大切な願いを叶えたい”“苦痛に耐えれば大切な願いが叶うと信じている”という心的表象が予測されて、各予測誤差は最小化されるであろう。

以上のように、一定の経験や知識がなければ、メンタライジングシステムのレベルで他者の行動からの的確に心は推測されず、心的表象は繰り返し

更新されて予測は終息しないことになる。そして、Figure 2に示すように、内受容感覚の予測誤差の精度の高さがミラーニューロンシステムの機能に一定の貢献をし、またその予測誤差に基づく予測の精度の高さがメンタライジングシステムの機能を一定レベルで充進させることも分かる。こうして考えると、他者への的確な共感と心の理解には、身体内部変化を豊富にもたらし内受容感覚の予測の精度を上げるような発達の環境とともに、身体内部変化への気づきを引き出して内受容感覚の予測誤差の精度を上げるような経験や他者との関わりが必要なのであろう。

5. 結び

近年、心理学や心理臨床学の領域では内受容感覚が刮目され、内受容感覚の覚知の精度を上げることの重要性や意味が示唆されるようになってきた。特に、内受容感覚の覚知の精度が良いと自己覚知が明確になり、自他分化をもたらす可能性や共感性を高める可能性が強調されるようになってきた。まず、本論では、関係する理論と研究成果を参照して論考を行ったが、その2つの可能性はともに示唆された。

また、内受容感覚の覚知の指標の多くが心臓の動きに基づくものであり、肺や胃腸等の他の内臓

の動きに基づいた方法や指標によって調査・実験を進めた研究は殆どなかったため、他の内臓の動きに基づいた指標やそれらと心臓の動きを合わせた統合的な指標によって内受容感覚の影響を検討する必要があるだろう。ただし、各種内臓の動きのリズムは異なるため、これらがバインドされて1つの内受容感覚としてモニタリングされて機能しているのか、それとも個別にモニタリングされて個々に機能しているのかをまずは検討する必要があると思われる。

そして、本論では、内受容感覚の覚知によって模倣・情動伝染・情動的共感が生じやすいことを示唆する研究知見をまとめ、内受容感覚がミラーニューロンシステムにどのように関わっているかを予測符号化の枠組みでモデル化した。また、認知的共感や心の理解にも内受容感覚が関与している可能性やミラーニューロンシステムの上位でメンタライジングシステムがトップダウン的に機能している可能性を想定し、予測符号化理論に基づいて内受容感覚を含めた心の理解のモデル化を試みた。これらのモデルについては、各感覚の予測や予測誤差のシグナルが実際に感覚間でまたはシステム間でどのように伝達されているかを示唆する証拠が不十分な中、明確な根拠のないままその構築を試行した。何よりも、感覚間でまたシステム間での予測や予測誤差のシグナルの伝達パターンの存在を示唆する研究の進展が重要になってくる。特に、身体内部環境をモニタリングする内受容感覚において、その予測や予測誤差がどのように生じて機能するのが明らかになり、それをAIに実装すれば、親身になってより具体的に人に共感したり心を理解したりするロボットの実現も夢ではないだろう。しかし、その実現には、複雑な機構をもつ身体内部環境やその環境内での多様なシグナルに対するモニタリングをどのようにモデル化するか、また、それに関わるどのようなデータをどのような方法でサンプリングするかという大きな壁が立ちただかっていることは間違いない。

文献

Ainley, V., Apps, M. A., Fotopoulou, A., & Tsakiris, M. (2016). 'Bodily precision': A

predictive coding account of individual differences in interoceptive accuracy. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences*, 371 (1708), 20160003. <https://doi.org/10.1098/rstb.2016.0003>

Ainley, V., Brass, M., & Tsakiris, M. (2014). Heartfelt imitation: High interoceptive awareness is linked to greater automatic imitation. *Neuropsychologia*, 60, 21-28. <https://doi.org/10.1016/j.neuropsychologia.2014.05.010>

Ainley, V., & Tsakiris, M. (2013). Body conscious? Interoceptive awareness, measured by heartbeat perception, is negatively correlated with self-objectification. *PLoS ONE* 8 (2), e55568. <https://doi.org/10.1371/journal.pone.0055568>

Allen, M., Frank, D., Schwarzkopf, D. S., Fardo, F., Winston, J. S., Hauser, T. U., & Rees, G. (2016). Unexpected arousal modulates the influence of sensory noise on confidence. *eLife*, 5, e18103. <https://doi.org/10.7554/eLife.18103>

Arioli, M., Perani, D., Cappa, S., Proverbio, A. M., Zani, A., Falini, A., & Canessa, N. (2018). Affective and cooperative social interactions modulate effective connectivity within and between the mirror and mentalizing systems. *Human Brain Mapping*, 39 (3), 1412-1427. <https://doi.org/10.1002/hbm.23930>

Asada, M. (2015). Towards artificial empathy: How can artificial empathy follow the developmental pathway of natural empathy?. *International Journal of Social Robotics*, 7, 19-33. <https://doi.org/10.1007/s12369-014-0253-z>

Azzalini, D., Rebollo, I., & Tallon-Baudry, C. (2019). Visceral signals shape brain dynamics and cognition. *Trends in Cognitive Sciences*, 23 (6), 488-509. <https://doi.org/10.1016/j.tics.2019.03.007>

Babo-Rebelo, M., Richter, C. G., & Tallon-Baudry, C. (2016). Neural responses to heartbeats in the default network encode the self in spontaneous thoughts. *The Journal of Neuroscience*, 36 (30), 7829-7840. <https://doi.org/10.1523/JNEUROSCI.0262-16.2016>

Barrett, L. F. (2017). *How emotions are made: The secret life of the brain*. Houghton Mifflin Harcourt.

- Ciaramidaro, A., Becchio, C., Colle, L., Bara, B. G., & Walter, H. (2014). Do you mean me? Communicative intentions recruit the mirror and the mentalizing system. *Social Cognitive and Affective Neuroscience*, 9 (7), 909-916. <https://doi.org/10.1093/scan/nst062>
- Craig, A. D. (2002). How do you feel? Interoception: the sense of the physiological condition of the body. *Nature Reviews Neuroscience*, 3 (8), 655-666. <https://doi.org/10.1038/nrn894>.
- Craig, A. D. (2003). Interoception: the sense of the physiological condition of the body. *Current Opinion in Neurobiology*, 13 (4), 500-505. [https://doi.org/10.1016/s0959-4388\(03\)00090-4](https://doi.org/10.1016/s0959-4388(03)00090-4).
- Craig, A. D. (2009). How do you feel—now? The anterior insula and human awareness. *Nature Reviews Neuroscience*, 10 (1), 59-70. <https://doi.org/10.1038/nrn2555>
- Craig, A. D., Chen, K., Bandy, D., & Reiman, E. M. (2000). Thermosensory activation of insular cortex. *Nature Neuroscience*, 3 (2), 184-190. <https://doi.org/10.1038/72131>
- de Waal, F. B. (2008). Putting the altruism back into altruism: the evolution of empathy. *Annual review of psychology*, 59, 279-300.
- Engelen, T., Solcà, M., & Tallon-Baudry, C. (2023). Interoceptive rhythms in the brain. *Nature Neuroscience*, 26 (10), 1670-1684. <https://doi.org/10.1038/s41593-023-01425-1>.
- Flexman, J., Demaree, R., & Simpson, D. (1974). Respiratory phase and visual signal detection. *Perception & Psychophysics*, 16, 337-339. <https://link.springer.com/article/10.3758/BF03203952>
- Friston, K. (2010). The free-energy principle: a unified brain theory?. *Nature Review Neuroscience*, 11, 127-138 (2010). <https://doi.org/10.1038/nrn2787>
- Friston, K., Mattout, J., & Kilner, J. (2011). Action understanding and active inference. *Biological Cybernetics*, 104, 137-160. <https://doi.org/10.1007/s00422-011-0424-z>
- Fukushima, H., Terasawa, Y., & Umeda, S. (2011). Association between interoception and empathy: Evidence from heartbeat-evoked brain potential. *International Journal of Psychophysiology*, 79 (2), 259-265. <https://doi.org/10.1016/j.ijpsycho.2010.10.015>
- Garfinkel, S. N., Minati, L., Gray, M. A., Seth, A. K., Dolan, R. J., & Critchley, H. D. Fear from the heart: sensitivity to fear stimuli depends on individual heartbeats. (2014). *The Journal of Neuroscience*, 34 (19), 6573-6582. <https://doi.org/10.1523/jneurosci.3507-13.2014>
- Gallagher, S. (2000). Philosophical conceptions of the self: Implications for cognitive science. *Trends in Cognitive Sciences*, 4 (1), 14-21. [https://doi.org/10.1016/S1364-6613\(99\)01417-5](https://doi.org/10.1016/S1364-6613(99)01417-5)
- Grynberg, D., & Pollatos, O. (2015). Perceiving one's body shapes empathy. *Physiology & Behavior*, 140, 54-60. <https://doi.org/10.1016/j.physbeh.2014.12.026>
- Gu, X., Hof, P. R., Friston, K. J., & Fan, J. (2013). Anterior insular cortex and emotional awareness. *The Journal of Comparative Neurology*, 521 (15), 3371-3388. <https://doi.org/10.1002/cne.23368>
- Harrison, N. A., Gray, M. A., Gianaros, P. J., & Critchley, H. D. (2010). The embodiment of emotional feelings in the brain. *The Journal of Neuroscience*, 30 (38), 12878-84. <https://doi.org/10.1523/JNEUROSCI.1725-10.2010>
- Iacoboni, M., Woods, R.P., Brass, M., Bekkering, H., Mazziotta, J.C., & Rizzolatti, G. (1999). Cortical mechanisms of human imitation. *Science*, 286 (5449), 2526-2528. <https://doi.org/10.1126/science.286.5449.2526>
- Imafuku, M., Fukushima, H., Nakamura, Y., Myowa, M., & Koike, S. (2020). Interoception is associated with the impact of eye contact on spontaneous facial mimicry. *Scientific Report*, 10 (1), 19866. <https://doi.org/10.1038/s41598-020-76393-8>
- 磯村拓哉 (2018). 自由エネルギー原理の解説：知覚・行動・他者の思考の推論 日本神経回路学会誌, 25 (3), 71-85. <https://doi.org/10.3902/jnns.25.71>
- Isomura, T., & Watanabe, K. (2020). Direct gaze enhances interoceptive accuracy. *Cognition*, 195, 104113. <https://doi.org/10.1016/j.cognition.2019.104113>
- Jicol, C., Cheng, H. Y., Petrini, K., & O'Neill, E. (2023). A predictive model for understanding the role of emotion for the formation of presence in virtual reality. *PLoS ONE*, 18 (3), e0280390. <https://doi.org/10.1371/journal.pone.0280390>

- Kilner, J. M., Friston, K. J., & Frith, C. D. (2007). Predictive coding: An account of the mirror neuron system. *Cognitive Processing*, 8 (3), 159-166. <https://doi.org/10.1007/s10339-007-0170-2>
- Kleckner, I. R., Zhang, J., Touroutoglou, A., Chanes, L., Xia, C., Simmons, W. K., Quigley, K. S., Dickerson, B. C., & Barrett, L. F. (2017). Evidence for a large-scale brain system supporting allostasis and interoception in humans. *Nature human behaviour*. <https://doi.org/10.1038/s41562-017-0069>
- Kluger, D. S., Balestrieri, E., Busch, N. A., & Gross, J. (2021). Respiration aligns perception with neural excitability. *eLife*, 10, e70907. <https://doi.org/10.7554/eLife.70907>
- Leslie, K.R., Johnson-Frey, S.H., & Grafton, S.T. (2004). Functional imaging of face and hand imitation: towards a motor theory of empathy. *NeuroImage*, 21 (2), 601-607. <https://doi.org/10.1016/j.neuroimage.2003.09.038>
- Likowski, K. U., Mühlberger, A., Gerdes, A. B., Wieser, M. J., Pauli, P., & Weyers, P. (2012). Facial mimicry and the mirror neuron system: Simultaneous acquisition of facial electromyography and functional magnetic resonance imaging. *Frontiers in Human Neuroscience*, 6, 214. <https://doi.org/10.3389/fnhum.2012.00214>
- Marshall, A. C., Gentsch, A., Jelinčić, V., & Schütz-Bosbach, S. (2017). Exteroceptive expectations modulate interoceptive processing: repetition-suppression effects for visual and heartbeat evoked potentials. *Scientific Reports*, 7 (1), 16525. <https://doi.org/10.1038/s41598-017-16595-9>
- Marshall, A. C., Gentsch, A., & Schütz-Bosbach, S. (2020). Interoceptive cardiac expectations to emotional stimuli predict visual perception. *Emotion*, 20 (7), 1113-1126. <https://doi.org/10.1037/emo0000631>
- Migeot J. A., Duran-Aniotz C. A., Signorelli, C. M., Piguet, O., & Ibáñez, A. (2022). A predictive coding framework of allostatic-interoceptive overload in frontotemporal dementia. *Trends in Neuroscience*, 45 (11), 838-853. <https://doi.org/10.1016/j.tins.2022.08.005>
- Moseley, G. L., Olthof, N., Venema, A., Don, S., Wijers, M., Gallace, A., & Spence, C. (2008). Psychologically induced cooling of a specific body part caused by the illusory ownership of an artificial counterpart. *PNAS Proceedings of the National Academy of Sciences of the United States of America*, 105 (35), 13169-13173. <https://doi.org/10.1073/pnas.0803768105>
- Nishitani, N., & Hari, R. (2002). Viewing lip forms: cortical dynamics. *Neuron*, 36 (6), 1211-1220. [https://doi.org/10.1016/S0896-6273\(02\)01089-9](https://doi.org/10.1016/S0896-6273(02)01089-9)
- 大平英樹 (2018). 内受容感覚の予測的符号化—福島論文へのコメント— 心理学評論, 61 (3), 322-329. https://doi.org/10.24602/sjpr.61.3_322
- Ondobaka, S., Kilner, J., & Friston, K. (2017). The role of interoceptive inference in theory of mind. *Brain and Cognition*, 112, 64-68. <https://doi.org/10.1016/j.bandc.2015.08.002>
- Owens, A. P., Allen, M., Ondobaka, S., & Friston, K. J. (2018). Interoceptive inference: From computational neuroscience to clinic. *Neuroscience and Biobehavioral Reviews*, 90, 174-183. <https://doi.org/10.1016/j.neubiorev.2018.04.017>
- Palmer, C. E., & Tsakiris, M. (2018). Going at the heart of social cognition: Is there a role for interoception in self-other distinction? *Current Opinion in Psychology*, 24, 21-26. <https://doi.org/10.1016/j.copsyc.2018.04.008>
- Park, H. D., Correia, S., Ducorps, A., & Tallon-Baudry, C. (2014). Spontaneous fluctuations in neural responses to heartbeats predict visual detection. *Nature Neuroscience*, 17, 612-618. <https://doi.org/10.1038/nn.3671>
- Paulus, M. P., & Stein, M. B. (2006). An Insular View of Anxiety. *Biological Psychiatry*, 60 (4), 383-387. <https://doi.org/10.1016/j.biopsych.2006.03.042>
- Phillips, M. L., Medford, N., Senior, C., Bullmore, E. T., Suckling, J., Brammer, M. J., Andrew, C., Sierra, M., Williams, S. C., & David, A. S. (2001). Depersonalization disorder: thinking without feeling. *Psychiatry Research: Neuroimaging*, 108 (3), 145-160. [https://doi.org/10.1016/S0925-4927\(01\)00119-6](https://doi.org/10.1016/S0925-4927(01)00119-6)
- Rizzolatti, G., Fogassi, L., & Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature reviews. Neuroscience*, 2 (9), 661-70. <https://doi.org/10.1038/35090060>
- Sato, W., Kochiyama, T., & Uono, S. (2015).

- Spatiotemporal neural network dynamics for the processing of dynamic facial expressions. *Scientific Reports*, 5, 12432. <https://doi.org/10.1038/srep12432>
- Seth, A. K., Suzuki, K., & Critchley, H. D. (2012). An interoceptive predictive coding model of conscious presence. *Frontier in Psychology*, 2, 395. <https://doi.org/10.3389/fpsyg.2011.00395>.
- Seth, A. K., & Tsakiris, M. (2018). Being a beast machine: The somatic basis of selfhood. *Trends in Cognitive Sciences*, 22 (11), 969-981. <https://doi.org/10.1016/j.tics.2018.08.008>
- Shah, P., Catmur, C., & Bird, G. (2017). From heart to mind: Linking interoception, emotion, and theory of mind. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior*, 93, 220-223. <https://doi.org/10.1016/j.cortex.2017.02.010>
- Shamay-Tsoory, S. G., Aharon-Peretz, J., & Perry, D. (2009). Two systems for empathy: a double dissociation between emotional and cognitive empathy in inferior frontal gyrus versus ventromedial prefrontal lesions. *Brain*, 132, 617-627.
- Shimada S. (2022). Multisensory and sensorimotor integration in the embodied self: Relationship between self-body recognition and the mirror neuron system. *Sensors*, 22 (13), 5059. <https://doi.org/10.3390/s22135059>
- Sperduti, M., Guionnet, S., Fossati, P., & Nadel, J. (2014). Mirror neuron system and mentalizing system connect during online social interaction. *Cognitive Processing*, 15 (3), 307-316. <https://doi.org/10.1007/s10339-014-0600-x>
- Spunt, R. P., & Lieberman, M. D. (2013). The busy social brain: Evidence for automaticity and control in the neural systems supporting social cognition and action understanding. *Psychological Science*, 24 (1), 80-86. <https://doi.org/10.1177/0956797612450884>
- Stephan, K. E., Manjaly, Z.-M., Mathys, C. D., Weber, L. A., Paliwal, S., Gard, T., Stephan, K. E., Manjaly, Z. M., Mathys, C. D., Weber, L. A. E., Paliwal, S., Gard, T., Tittgemeyer, M., Fleming, S. M., Haker, H., Seth, A. K., & Petzschner, F. H. (2016). Allostatic self-efficacy: A metacognitive theory of dyshomeostasis-induced fatigue and depression. *Frontiers in Human Neuroscience*, 10, 550. <https://doi.org/10.3389/fnhum.2016.00550>
- Tajadura-Jiménez, A., & Tsakiris, M. (2014). Balancing the “inner” and the “outer” self: Interoceptive sensitivity modulates self-other boundaries. *Journal of Experimental Psychology: General*, 143 (2), 736-744. <https://doi.org/10.1037/a0033171>
- Tallon-Baudry C. (2023). Interoception: Probing internal state is inherent to perception and cognition. *Neuron*, 111 (12), 1854-1857. <https://doi.org/10.1016/j.neuron.2023.04.019>
- Terasawa, Y., Moriguchi, Y., Tochizawa, S., & Umeda, S. (2014). Interoceptive sensitivity predicts sensitivity to the emotions of others. *Cognition and Emotion*, 28 (8), 1435-1448. <https://doi.org/10.1080/02699931.2014.888988>
- Tsakiris, M., Tajadura-Jiménez, A., & Costantini, M. (2011). Just a heartbeat away from one's body: interoceptive sensitivity predicts malleability of body-representations. *Proceedings of the Royal Society B*, 278, 2470-2476. <https://doi.org/10.1098/rspb.2010.2547>
- Vogeley, K., & Fink, G. R. (2003). Neural correlates of the first-person-perspective. *Trends in Cognitive Sciences*, 7 (1), 38-42. [https://doi.org/10.1016/S1364-6613\(02\)00003-7](https://doi.org/10.1016/S1364-6613(02)00003-7).
- Wen, W., & Haggard, P. (2019). Prediction error and regularity detection underlie two dissociable mechanisms for computing the sense of agency. *Cognition*, 195. <https://doi.org/10.1016/j.cognition.2019.104074>
- Zahavi, D. (2005). *Subjectivity and Selfhood: Investigating the first-person perspective*. Cambridge, MA: The MIT Press.

(2024.1.26受稿 2024.2.22受理)

Theoretical examination of the relationship between interoception and empathy and understanding minds: How does the sense of self's internal body contribute to empathic response and understanding minds for others?

Takahiro HISAZAKI

This paper examined how and to what extent interoception is related to sociality, such as empathic response and understanding the mind for others. First, several theories and research findings suggested that interoception contributes to subjective frame of conscious experience and self-awareness, which can lead to the discrimination between self and others. Recent research findings, however, have suggested that interoceptive awareness related to responsivity of imitation, emotional contagion, and emotional empathy. Therefore, the mechanism of empathizing with others while discriminating between self and others was discussed by reviewing and examining the influence of interoception on exteroception and proprioception and the mirror neuron system. These discussions resulted in constructing a predictive coding model of the mirror neuron system that incorporates the prediction and prediction error of interoception. Moreover, the possibility that the prediction of interoception influences the understanding of others' minds was considered challengingly, leading to construction of a predictive coding model of the understanding of others' minds including the mentalizing system as the higher level of system than the mirror neuron system.

Key words: interoception, subjective frame of conscious experience, self-awareness, predictive coding, empathy and understanding minds