考えるな、（身体を）感じろ
―― 情動の身体性評価説 ――

鵜子 修司・成瀬 翔

はじめに

本稿で検討するプリンツの身体性評価説は、現代の情動（emotion）の理論において、もっとも有望な理論の一つである。結論を先取りすると、プリンツの理論では、情動は「身体性の評価」と「感情価マーカー」という二つの部分で構成された知覚の一種と定義される（Prinz, 2004）。そして、情動において知覚されるのは、有機体の「福利（well-being）」に関連した、環境との「関係的性質（e.g., 喪失、危険）」である。こうした関係的性質はラザルスが「中心的関係主題」と呼ぶものである（Lazarus, 1991）。すなわち、特定の人物や状況がもつ個人的な意味である。従って、プリンツの理論を簡単に要約すると、情動とは、環境への対応として有機体の身体状態に変化が生じたとき、その変化が体性感覚システムの作用により記録（register）される過程である。こうした記録が有機体の生存に有用になる（ことで進化の過程で獲得された）のは、情動は身体変化を記録することにより、実質的には中心的関係主題を表象（represent）しているためである（i.e., 身体性の評価）。この表象に行動の内的強化子と定義される感情価が結びつくことで（i.e., 感情価マーカー）、有機体は様々な中心的関係主題に応じて適切な行動をとるようになる（e.g., 大切なモノの〈喪失〉を避ける、〈危険〉な場所から離れる）。

上記の要約も分かるように、プリンツの理論を構成する説明や概念の多くは、独自のものではなく、むしろ従来の代表的な心理学理論ですでに提唱されてきたものである。特に身体性評価説については、現代でも人気がある情動の「多次元評価説」（このうち、最も洗練された理論の一つがラザルスの提唱する理論である）と、古くはジェイムズ―ランゲ説に由来する情動の「身体説」を調停した理論と言える。この試みがいかにして可能になるのか、およびその試みがもつ価値はプリンツの理論が採用す

る先行理論の仮定だけでなく、プリンツの理論が拒否する仮定を説明することで明らかになるだろう。

そこで、本稿ではプリンツの身体性評価説とその背景となる諸理論との関係を検討する。まず第 1 節においてジェイムズ―ランゲ説に由来する「身体説」とそれに対する批判を概観する。次に、第 2 節において、「多次元評価説」が含む暗黙の理論的仮定と、それらに対するプリンツの立場を簡単に記述する。続く第 3 節では、「ザイアンス対ラザルス論争」を通じて、プリンツが多次元評価説にどのような批判を加えるのかを確認する。第 4 節では、「身体性の評価」という観点からプリンツの議論の構成を確認し、最後に第 5 節では「感情価マーカー」という概念の分析を行う。これらの考察により、プリンツの身体性評価説の構成を明らかにすることを試みたい。

1. ジェイムズ―ランゲ説

ジェイムズ―ランゲ説において、情動は主観的経験、すなわち「感じ（feeling）」と定義される。これは情動に関する日常的な直観に沿う定義(1)だが、この理論の特徴は感じの源泉にある。すなわち、情動における感じとは、身体に生じている変化の知覚であると主張するのである（James, 1884; Lange, 1885）。これは情動の感じと身体変化の順序に関する素朴な理解に反する。つまり、常識的には感じは身体変化に先行するものである（e.g., 悲しいから泣く）。これに対して、この理論では身体変化がむしろ感じに先行している（e.g., 泣くから悲しい）。

ジェイムズ―ランゲ説は、ジェイムズの元学生であった生理学者キャノンの痛烈な批判を受けることになる（Cannon, 1927）。この批判によりジェイムズ―ランゲ説は、標準的な心理学史では（少なくとも一時的には）棄却されることになる（Cornelius, 1996）。キャノンはジェイムズ―ランゲ説への批判として、内臓活動の変化では多様な情動の生起を説明するのに不十分であることを、5 つの根拠を挙げて示している。しかし、これらの根拠がいずれも、情動における内臓活動の必要性への疑いに終始している時点で、キャノンの批判は、かなり単純化されたジェイムズ―ランゲ説への批判となっている。つまり、仮に 5 つの根拠がすべて妥当であったとしても、それは少なくともジェイムズに対する決定的な反証にはなり得ない。なぜなら内臓活動の変化

だけでは説明が困難な事例でも、ジェイムズならば異なる身体変化に訴えることで説明できるからである(2)。

一方で、ジェイムズ―ランゲ説の妥当性を支持する証拠にはどのようなものがあるだろうか。プリンツは、まずジェイムズとランゲ自身が挙げた複数の証拠を検討している。そうした証拠の一つとして、ここでは身体からのフィードバックに関する仮説を取り上げよう。すなわち、身体状態を意図的に変えることが、情動に影響を及ぼす可能性の指摘である(e.g., 意図的に笑顔になることで幸福感が生じる)。ジェイムズとランゲ以後の研究者は、この仮説の妥当性を実験により示している。その一例として、特定の音を含んだ名前をもつ主人公の物語を音読することによって実際に情動が生じうることが示されている(Zajonc, Murphy, & Inglehart, 1989)。また、「顔面フィードバック」として知られている実験結果は、ペンを特定の向きで咥え、実験条件に応じて被験者が異なる表情を自然に作ることによっても情動を生じうることを示している(Strack, Martin, & Stepper, 1988)。さらにプリンツは Damasio et al.,(2000)による神経科学的な証拠を挙げている。それによると、人が情動を経験するときには、身体変化を探知する脳領域が活動しているという。そうした脳領域は情動状態の神経相関項と考えられる。プリンツ(2004)は以上のような諸々の証拠から、情動と身体状態の間には、最低でも規則的な相関があると信じる圧倒的な理由があると結論づける。プリンツ流の表現に倣えば、情動は「身体化(embodiment)されている」のである(3)。

ジェイムズ―ランゲ説を検討する最後に、この理論がプリンツ(2004)の分類では、「身体感じ説」に含まれることを確認しておこう。身体感じ説とは、情動と身体状態の感じを同一視する理論であり、このとき情動は定義からして意識的なものである。これに対して、身体状態の知覚を重視しつつ、情動と感じを同一視までは求めない理論が身体説である。こうした理論を唱える例として、神経学者のダマシオの理論では、脳による情動には身体状態の記録が不可欠であると考えてはいるが、それには非意識的に行われる記録も含まれる。すなわち、非意識的で感じを伴わない情動もありうるのである。すなわち、ダマシオの理論は情動の「身体説」と呼ばれるべきである。プリンツも無意識に情動が生じる可能性を認めているため、身体性評価説の基礎にあるのは、身体感じ説というより、正確には身体説ということになるだろう。身体性評価

説に含まれるジェイムズ―ランゲ説由来の理論的仮定は、情動には身体状態を追跡する過程が含まれることに限られるのである。

2. 多次元評価説

次に、身体性評価説のもう一つの源流である多次元評価説を検討しよう。心理学的な情動理論において、「評価（appraisal）」という用語を流通させたのは、アーノルドの功績とされる（Arnold, 1960）。ここでの「評価」とは、遭遇した環境が自分に与える影響についての評価的判断である。こうした単なる環境の知覚を超えた認知的な評価こそ、情動の真の原因であると主張するのが評価説である。代表的な評価説では、情動の原因となる評価には、いくつかの下位次元があると想定しており、こうした想定をもつ評価説について、プリンツは特に多次元評価説と呼称している。このとき評価的判断は、いくつかの次元に分けられる。アーノルド（1960）は最初に3つの次元を挙げたが、その後の理論家たちによって、より適切な評価次元の内容・数については様々な提案が為されてきた。そうした後続の理論のうちで、プリンツはラザルスが提案する6次元の評価説を「かなり高度に洗練された」理論として評価している（Lazarus, 1991）。

ラザルスによれば、評価は主体と環境の関係についての価値づけである（Table 1）。

Table 1
Lazarus（1991）の提唱する6次元

評価次元		概要
一次評価	(1) 目標との関連	対象との関わりが自己の目標に関連する／関連しない （無関連ならば情動は生じない）
	(2) 目標との適合性	対象との関わりは目下の目標達成を容易にする／阻害する （達成を促進するなら正の情動が、阻害するなら負の情動が生じる)
	(3) 自我関与の類型	対象との関わりにおいて問題になる (at stake) ものは何か （例えば、自我同一性や道徳的価値、人生の目標、他人の幸福など）
二次評価	(4) 責任の所在と功罪	対象との関わりは誰ないし何に責任があるのか 対象との関わりは賞賛と非難のどちらに値するか
	(5) 対処能力	対象との関わりの結果を、自分はどの程度まで処理できるか
	(6) 将来の見込み	今後の見込みが、目標と一致する／一致しない

次元（1）～（3）は環境の情動的な重要性を決めるものであり、また次元（4）～（6）は主体の対処方針に関わるものである（それぞれ一次評価および二次評価と呼ばれる）。6つの次元は、情動が生じるとき我々が実際に行う認知的判断に対応しており、それらの判断結果の組み合わせ次第で、異なる情動が生じる。ラザルス（1991）は6次元から為された評価を「分子評価」と呼び、そうした評価の要約を「モル評価」と呼んでいる。このうちモル評価はラザルス（1991）が「中心的関係主題」と呼ぶ、有機体の福利に関わる、有機体と環境との間にある情動ごとの関係を表している[(4)]（Table 2）。

Table 2
情動と対応する中心的関係主題（Lazarus, 1991）

情動		中心的関係主題
怒り	（Anger）	自分や所有物を貶める攻撃
不安	（Anxiety）	直面している不確かで実存的な脅威
恐怖	（Fright）	直面している直接的・具体的・圧倒的な身の危険
罪悪感	（Guilt）	道徳上の義務に反している
恥	（Shame）	自我理想にかなっていない
悲哀	（Sadness）	取り返しのつかない喪失を経験している
羨望	（Envy）	他人のモノを欲しがる
嫉妬	（Jealousy）	他人からの愛情を失うこと、あるいはそうした脅威をもたらす第三者に向けた憤慨
嫌悪	（Disgust）	気に食わないモノや考えを抱いたり、それに近づきすぎたりする
幸福	（Happiness）	目標の実現へと適度に近づく
誇り	（Pride）	自分自身あるいは自分が連帯意識をもつ他者や集団が、価値あるモノや達成を得ることで自我同一性が高められる
安堵	（Relief）	目標に合致しない悲惨な条件が改善する、または無くなる
希望	（Hope）	最悪を恐れ、ましな状況を切望する
愛情	（Love）	愛を結ぶ関係を望んだり参加したりする。通常は相互的だが、必ずしもそうである必要はない
哀れみ	（Compassion）	他人の苦難に心を動かされ助けたいと望む

多次元評価説において、情動は有機体の適応と関連づけられる。すなわち、情動における身体変化は適応的な行動への準備とみなされる（Arnold, 1960; Lazarus, 1991）。しかし、この見解はダーウィン（1872/2021）以後の情動研究で広く共有されており、実際にジェイムズ（1884）やダマシオ（2000）もダーウィンからの影響を認めている。すなわち、情動と適応の関連について、身体説の支持者らと多次元評価説の支持者らの見解は基本的に一致している。

しかし、多次元評価説が身体説に対して不満を向けるのは、単なる知覚が適応的な身体変化を生じる過程について、明確に述べていない点である（cf. Arnold, 1960）。言い換えると、なぜ我々は状況に応じて適切な身体変化を生じることができるのか、身体説からは明らかでないようにみえるのだ。例えば、山中で野生のクマを知覚すると、大抵の人には逃走や（やり過ごすための）硬直を準備する身体変化が生じるだろ

う。一方で、動物園で檻の中にいるクマを知覚したときには、むしろ接近を準備する身体変化が生じるかもしれない。また野生のクマを知覚する場合も、その人がクマ撃ちの猟師であれば、やはり同様に接近を準備する身体変化が生じるだろう。あるいは、我々は通常の方法では知覚すらできないウイルスに怯え、自宅に籠ることもある。以上で挙げた、情動生起における個人内／間の差や非知覚的な対象の例は、いずれも情動の原因を単なる知覚と想定する身体説からは説明が困難にみえる。しかし、この困難は認知的判断を情動の原因に仮定すれば容易に解消できるだろう（cf. 成瀬・鵜子（近刊））。

また情動の過程に認知的要素が含まれるなら、身体説では情動の多様性を説明できないとするキャノンの批判にも答えられる。こうした批判に対して、身体説を擁護する根拠はあるが （i.e., 自律神経系による特定化）、こうした証拠について懐疑的な研究者も存在する（e.g., Cacioppo, Berntson, Larsen, Poechlmann, & Ito, 2000）。情動の多様性は、それに対応する多様な認知的判断の結果だと考える方が、優れた説明かもしれないのである。ただし、その場合は認知的要素が必ずしも身体変化に先行すると考える必要はない。実際、情動に認知が含まれると主張する理論の中には、身体変化こそが認知的解釈に先行すると主張する理論もあれば（Schachter, 1964）、特定の順序を想定しない理論もある（Damasio, 1994）。いずれにせよ、これらの理論は認知を情動の必要な構成要素とみなすが、情動を単なる認知であるとは考えていない点に注意が必要である。

3. ザイアンス対ラザルス論争

これまで検討してきた情動理論に対して、主として哲学の領域で提唱されてきた情動理論には、情動と認知を同一視する理論もある [5]。プリンツは、前者を「純粋でない認知説」、後者を「純粋な認知説」と呼び分けるが、これらを総称して「認知説」と分類している。以下では多次元評価説を含め、すべての認知説に含まれている暗黙の仮説と、それらに対するプリンツの見解をみていこう。

プリンツによると、哲学的・心理学的な認知説のいずれにおいても、次の 3 つの仮説に対する暗黙のコミットメントが存在する（Prinz, 2004）。

（1）　概念化仮説：情動には概念が必要である。

（2）　非身体性仮説：情動と結びついた認知的要素は身体から独立している。

（3）　評価仮説：情動には有機体の福利に関わる、有機体と環境との関係の表象が含まれる。

これらの仮説の妥当性を検討するためにプリンツ（2004）が参照するのは、1980年代にザイアンスとラザルスの間で生じた、情動における認知の必要性に関する論争である 。プリンツは、ザイアンス（1984）が情動に認知は必要ないと主張する 5 つの根拠と、それぞれに対するラザルス（1984）の応答を検討した上で、結論としてザイアンスに軍配を上げている。

この結論を裏付けるためには、ザイアンスが挙げているうち、最も直接的な根拠を一つ紹介すれば十分だろう。それは、情動が薬物や顔面フィードバックなどの物理的誘因でも引き起こされうるという事実である。これは非身体性仮説への強力な反論になるだろう。つまり、情動の生起に身体から独立した要素は必要ないことを示しているのである。

これに対して、ラザルスを擁護するために、「認知」という用語の定義上の問題を指摘することで反論できるかもしれない。コーネリアス（1996）は、ザイアンス対ラザルスの論争が 2 つの概念の定義に関する見解の相違へ帰結すると結論づけている（Cornelius, 1996）。第一に、両者で「認知的評価」の定義が異なっている。すなわち、ザイアンスが認知的評価を意識的で入念かつ合理的な評価と想定しているのに対し、ラザルスは非意識的でありうるが、迅速で、ときには不合理な評価と想定している。情動生起には刺激を識別する最小限の情報処理が必要であることはザイアンスも認めているため、結局のところ、両者は同じものを認知と呼ぶか否かで争っているのかもしれない。それではどちらの見解が妥当なのだろうか。

プリンツの見解では、ザイアンスとラザルスの両者は共に誤っているという。プリンツにとってはザイアンスとラザルスの両者、あるいは先行研究において提案されてきた他の定義でも、認知の定義としては問題がある（6）。それらに代わり、プリンツ

が提案しているのは、「有機体の制御」に訴えた定義である。有機体の制御とは、特定の心理的システムや神経的構造によるトップダウンの制御を意味する。そして、認知的状態とは有機体の制御下にある心的表象、すなわち概念を含む状態と定義する。一方、知覚状態とは環境の制御下にある心的表象を含む状態である。従って、知覚が記憶に蓄えられた表象は概念であるが、そもそもの知覚状態は概念ではない。

認知的状態には概念が含まれるという点については、認知説の支持者ならば同意しうるだろう。すなわち、情動が行動に複雑化された評価の組み合わせに依存するならば、情動には評価の記述が指示する概念を持つことが必要になる。これは先述した概念化仮説の意味するところである。しかし、プリンツはこの想定が誤っていると指摘する。そのように主張する根拠は、第一に多くの情動は外界により制御されていることが挙げられる。言い換えると、プリンツの定義に従えば、多くの情動は知覚を原因として生じていることである。また第二に、情動に関する神経解剖学の知見も根拠となる。扁桃体に関するルドゥーの研究では、皮質下にある比較的に原始的な構造である扁桃体が、特定の情動反応（特に恐怖）の生起に重要な役割を果たすことが示されている（LeDoux, 1996）。それによれば、ヘビのような対象を知覚するとき、網膜覚像の信号は視神経を通して視床に送られる。次に視床から情報が新皮質の視覚野に送られると、対象の最終的な認識が成立する。しかし、視床は扁桃体へと直接の信号を送る経路も持っている。そこから情報を受け取った扁桃体は、恐怖に関連した身体・行動反応を調整する脳の各部位に出力を送ることで、新皮質を介さず恐怖反応を編成できるのである。現在では認知の座は新皮質にあると考えられており、新皮質を介さず情動を生起しうるなら、情動における認知の必要性は疑わしくなる（cf. 成瀬・鵜子（近刊））。無論、扁桃体のように原始的な皮質下構造でも複雑な概念的評価が可能だと示されたなら、認知説を擁護する強力な根拠となる。しかし、そのような根拠は今のところは無さそうである。以上より、プリンツは概念化仮説が誤っていると結論づける。すなわち、ラザルスはザイアンスからの論難を認知の定義を修正することで回避することはできないのである。

しかし、コーネリアスによれば、ザイアンス対ラザルス論争にみられる定義上の相違はもう一つある（Cornelius, 1996）。それは、そもそもどのような反応を情動として定義するかの違いである。ザイアンスが認知説を批判する根拠として挙げた例には、

嫌悪条件づけや単純接触効果などのように、閾下呈示された刺激に対しても成立しうる情動的な学習がある。すなわち、無意識の学習が情動に影響することをもって、情動が認知なしに生じうることの根拠とみなす。これに対し、ラザルスの反論は例ごとに分かれている。まず嫌悪条件づけについては、先述した認知の定義に訴えた反論を展開している。つまり、無意識に成立する条件づけの過程にも、認知的な評価は介在していると主張する。

しかし、ここで注目したいのは、単純接触効果に向けられた反論である。ラザルスは「選好」が情動ではない場合もあると主張するのである。この反論には確かに説得力がある。例えば、誰かに「好意」を寄せているという場合なら、その選好は一般的に情動とみなされるだろう。一方で、店頭の商品を右端から選ぶことや、腕時計を左腕につけることを「好む」という場合、その選好を情動とみなすべきかどうかは定かでないと思われる(7)。同様に、単純接触効果の実験でみられる文字などへの選好は、情動の含まれない純粋な知的判断か、あるいは単なる感覚なのかもしれない。つまり、ザイアンスの反論の一部は、ラザルスなら情動とはみなさない反応の知見から構成されているのである。

ただし、情動の定義に訴えてもザイアンスの批判すべてに答えられるわけではない。例えば、先述した物理的誘因による情動の生起や、神経解剖学の知見に基づく批判は、ザイアンスやラザルスがともに情動とみなすであろう「楽しさ」や「恐怖」において、非身体性仮説や概念化仮説は支持されないことを示している。また仮にラザルスが、認知を原因としない状態は情動ではないと主張するならば、それは論点先取となる。しかし、情動の定義に下位区分を設けて認知説を修正する場合はどうか。すなわち、原因となる認知が必要な情動と、必要ない情動を区別するのである。実際、こうした下位分類を支持する情動理論は哲学・心理学のどちらにもみられる。例えば、哲学側ではグリフィスによって、嫉妬、罪悪感、誇りなどの「高次認知的情動（higher-cognitive emotions）」と彼が呼ぶ洗練された認知を含むようにみえる情動と、恐怖、怒り、幸福など進化的に獲得された基本的な情動が区別されている（Griffiths, 1997）。前者に含まれる情動は、心理学において「自己意識的感情　（self-conscious affect: Tangney & Fischer, 1995）」と呼ばれる情動と基本的に重なっている。また、後者はエクマンが「感情プログラム（affect program）」と呼ぶリストに合致する（Ekman, 2003）(8)。

こうした情動の下位分類の存在から「情動は一貫したクラスではない」と考えること、言い換えれば「情動は自然種（natural kind）ではない」と考えることを、プリンツは「不統一テーゼ」と呼ぶ。こうした想定とは反対に、プリンツはむしろ情動が自然種であると積極的に主張する立場をとる。その方法として、プリンツは少数の派生的でない情動のリスト、すなわち基本情動の存在を想定し、その他の情動も基本情動からの派生であると主張する。

そうしてプリンツは、そもそも自然種とは何か、基本情動とは何で、どの情動が基本的でありそうか、それらから他の情動への派生はどのように生じるかなどを詳細に検討している。しかし、これらの議論まで詳細に紹介することは、プリンツの情動理論における核心を紹介するという本稿の目的を超える。従って、この問題は別稿にて扱うべきだろう。なお、その際には基本情動という仮説に反対の立場を取り、よりラディカルな社会構築主義を擁護する情動理論（e.g., Barrett & Russell, 2015）を参照することで批判的な検討が可能である。さしあたり本稿で確認しておきたいのは、プリンツの理論がもつ利点は、身体説と認知説の調停だけに留まらず、これらと同じく一見は両立しないようにみえる、自然主義的な理論と構成主義的な理論の調停にまで及ぶという点である。そして、後者の調停が成功しているならば、情動の定義に訴えて認知説を擁護する方針にも見込みがないと言える。高次の認知を必要とするようにみえる情動も、実際には身体性の要素のみから構成されているのである。

4. 身体性の評価

ここまで確認した通り、プリンツは主にザイアンス（1984）の挙げる根拠に従って、認知説が暗黙に含む仮説のうち非身体性仮説と概念化仮説については否定してきた。それでは最後に残った評価仮説についてはどうだろうか。驚くべきことに、プリンツは評価仮説を擁護している。すなわち、概念化仮説と非身体性仮説を否定しても、評価仮説を単独に支持できると言うのである。実際には、多くの評価は非身体的である。しかし、これは評価が原理的に非身体的であることを意味しない。そしてプリンツによれば、情動こそ身体を介した評価、すなわち「身体性の評価 （embodied appraisal）」の例である。この意味で、プリンツの理論は身体説と評価説の複合理論と言える。

プリンツは情動が身体性の評価であると主張するために、まず情動が心的表象であることを確認する(10)。なぜなら、評価とは自分の利害との関わりから対象を表象することであり、評価仮説が正しいならば必然的に情動は心的表象となるためである。ただし情動が心的状態であるのは明白なので、検討されるのは情動が表象であるか否かの問題に限られる。この問題に関してプリンツが採用するのは、ドレツキが提唱する心的表象の理論である（Dretske, 1981, 1986）。ドレツキの理論によれば、心的表象とは以下の 2 つの条件を満たす心的状態である。第一に、情報を担っていること、第二に、誤りうることである。「情報を担う」という表現は、ある状態は、その状態と信頼のおける仕方で共に生じるものについての情報を担うことを意味している。その典型例は因果であり、ある原因が非常に高い確率で特定の結果を生じると言えるなら、その結果は原因に関する情報を担っていると言える。ただし、情報を担うことは、表象することと同義ではない。なぜなら、ある状態が特定の対象を表象していると言えるためには、その対象以外が同じ状態を生じさせた場合、それは誤りだと言える必要があるためである。例えば、ライオンの概念は我々が動物園でライオンを見たときだけでなく、上手く変装した犬を見た場合にも生じうる(9)。この場合、ライオン概念はライオンだけでなく、ライオンによく似た犬の情報をも担っている。しかし、ライオン概念が表象であるならば、この概念がライオン以外に作動させられた場合には「勘違い」であると言えねばならない。

ドレツキは、単なる情報の担い手が表象になるためには、その担い手に特定の情報を担う「機能」が必要になると考えている。すなわち、情報の担い手が表象とみなされるためには、（心的表象の場合なら）学習あるいは進化の過程で、当該の情報を担うために獲得されたと言える必要がある。こう考えることで、ある表象が機能的に不適切な情報を担う場合には、表象の誤りとみなせるのである。以上でドレツキのアイデアを踏まえて心的表象について結論をまとめると、心的表象とは学習や進化により、「特定の何かにより作動させられるために備わった心的状態」ということになる。

従って、情動が表象であるなら、情動は特定のものによって作動させられており、また情動はその特定のものによって作動させられるために備わったことを示す必要がある。身体説の想定が正しいならば、情動を信頼のおける仕方で引き起こす原因は身体変化である。身体説を検討した段階で確認した通り、情動には身体変化を追跡す

る過程が含まれている。すなわち、情動が生起するとき、我々は自身の身体状態を知覚しているのである。それでは情動は身体変化を表象しているのだろうか。言い換えれば、情動は身体変化を探知するために進化の過程で獲得されたのだろうか。

だが、これは奇妙な想定である。通常の自然主義的立場では、進化の過程で選択される能力は、生存を有利するように働くものと考えられている。しかし、身体変化を知覚することから、どうして生存しやすくなるのかは不明である。もちろん先行する身体変化は、適応的な行動を準備するだろうが、それは情動の利点ではない。またダマシオは、情動に関連した脳部位を損傷した患者から得られる知見から、情動は合理的な意思決定に必要であると主張している（Damasio, 1994）。これは説得力のある根拠に支えられた主張だが、プリンツ曰く、この主張はダマシオの唱える身体説と相容れない。なぜなら、身体変化を知覚する能力が（e.g., 血管収縮を知覚）、良い決定（e.g., 成功する投資）を導くとは考えづらいからである。

この疑問に対して、プリンツは情動が表象するのは身体状態ではなく、身体外部の事物だと考える。プリンツによると、情動は確かに身体変化を「記録」しているが、このとき情動は身体の外にあるもの、すなわち身体変化の原因を「表象」しているのである。それでは、情動に関わる身体変化を信頼のおける仕方で引き起こす原因とは何であろうか。プリンツがその有力な候補とするのは、ラザルス（1991）が提唱する中心的関係主題である。プリンツは中世思想を参照するケニーの指摘に従い、情動の対象（i.e., 情動を誘発する条件）を個別的対象（particular object）と形式的対象（formal object）に区別する（Kenny, 1963）。前者は情動を誘発した出来事そのものであり、後者は出来事に含まれた情動を誘発する性質である。そして、中心的関係主題は形式的対象である。例えば、野生のクマとの遭遇は、一般人には恐怖を、猟師には幸福をもたらすかもしれない。このとき個別的対象にのみ注目すると、情動には信頼のおける原因がないように思われる。しかし、こうした環境と各個人の関係、すなわち中心的関係主題に注目すると、この状況は一般人には身の「危険」をもたらすが、猟師には「目標の実現」をもたらすと考えられる。つまり、個別的対象は同じでも形式的対象は異なりうるのである。

この点を踏まえた上で、ラザルス（1991）が提示する中心的関係主題のリストを確認すると、これは情動が表象する内容を少なくとも近似的にはうまく表している。さ

らに、中心的関係主題を表象することは、有機体の生存を有利にするだろう。それというのも中心的関係主題とは、定義から有機体の福利に関係するため、これを上手く利用することもまた、定義から適応的と言えるのである。従って、情動は中心的関係主題に関する情報を担う機能があると考えても不自然ではない。以上から、プリンツは情動が中心的関係主題を表象していると結論づける。

ここで 2 点の注意が必要である。第一に、ラザルスが提示する中心的関係主題は、元々は複雑な認知的判断の結果を要約したものである。一方で、情動における身体変化は、こうした複雑な構造を持たないように見える。これについてプリンツは、表象されているものが複雑だからといって、表象それ自体が複雑である必要はないとする。ここでプリンツが例に挙げるのは「ファズバスター」なる装置である。これはスピード違反を取り締まる警察のレーダーを探知し、警告音で運転手に知らせるための装置である。つまり、ファズバスターの警告音は「警察のレーダー」を表象しているが、その警告音は「警察」や「レーダー」などに対応する別の音に分解できるわけではない。これと同じく情動もまた、複雑な構造をもつことなく、中心的関係主題を表象しうるのである。

第二に、プリンツは、情動がファズバスターの警告音と同じく構造を持たないとは考えていない。むしろ、情動とは中心的関係主題に対応するための身体変化を記録することで、中心的関係主題を表象する構造になっている。プリンツの表現に従うなら、情動は名目的には身体変化を記録し、実質的には中心的関係主題を表象しているのである。これは心的表象の性質として珍しくはない。例えば、我々が犬を認識する場合を考えると、犬の外見的特徴（e.g., 体毛、吠え声、尻尾）を追跡している。これらの特徴は、どれも犬の本質的な特徴（e.g., 犬ゲノム）ではない。それにも関わらず、これらの特徴を通じて我々は犬を認識している。これが可能になるのは、犬の外見的特徴とは、犬の本質的特徴によって信頼のおける仕方で引き起こされているからである。同様に、まず我々は特定の知覚により特定の身体変化を生じる仕組みを実装している。こうした実装には、進化の過程で獲得した生得的なものも、学習により獲得したものもある。いずれにしても、この身体変化の機能は知覚した事物に対処するための準備である。従って、このときの身体変化は知覚した事物と有機体との関係に関す

る情報を担っている。そして、情動とは身体変化を記録することで、そうした直接に観察できない関係的性質（i.e., 中心的関係主題）を表象しているのである。

以上で、プリンツの情動理論における「身体性の評価」については、この仮説が導かれるまでの議論を簡潔ながら紹介してきた。改めてまとめると、まず我々の体性感覚システムは身体状態の変化を追跡することができる。そして、我々が情動を経験するとき、このシステムに関わる脳領域が活性化している。従って、情動は身体変化の知覚が含まれるだろう。これはジェイムズ―ランゲ説に由来する身体説から得られた知見である。次に、情動を生じる過程に注目すると、そこには必ずしも認知的要素が含まれない。つまり、情動は認知的要素から生じる場合もあるが、単なる知覚から生じる場合も多い（なお知覚と情動の結びつきには、生得的なものもあれば、学習されたものもあるだろう）。また情動の過程に関する神経解剖学的知見も、情動の原因に概念付加的な認知的判断を想定している認知説と相容れない。従って、認知説を構成する主要な仮説は支持されず、認知説は情動の理論としては問題含みである。しかし、認知説に含まれる評価仮説については維持されうる。情動における身体変化は、その身体変化を生じた有機体と環境との関係についての情報を担っている。その内容はラザルス（1991）が示す中心的関係主題の内容と近似するだろう。そして、情動が記録する身体状態の変化は、中心的関係主題の情報を担うためのものであり、有機体は情報を用いることで適応的な立ち回りができる。従って、ドレツキの理論に従えば、情動は身体を介して中心的関係主題を表象（評価）している知覚なのである。

5. 感情価マーカー

前述のように、プリンツは情動が自然種であると考えている。すなわち、情動には必ず身体性の評価が含まれるという一貫性がある。しかし、プリンツによると、情動を構成する一貫した要素には、「感情価（valence）」なるものも含まれる。通常、感情価とは情動を正（positive）と負（negative）に分ける概念である。前者の典型例には「喜び」や「安らぎ」があり、後者の典型例には「恐怖」や「悲しみ」が挙げられるだろう。神経科学的な研究からは、情動の正負それぞれに共通した基盤が存在する

ことが示されており（e.g., Elliott, Friston, & Dolan, 2000; Lane, Chua, & Dolan, 1999）、これは感情価が単なる研究者による恣意的な分類ではなく、実在する区分であることを示唆する（ただし、感情価の正確な解剖学的基礎は未解明である）。しかし問題は、何が情動を正負に分けるのかである。つまり、情動の正負をより説明的な用語に置き換えることはできるのか。このとき最も直観的な特徴づけは、快／不快の感じに訴えるものだろう（e.g., Frijda, 1993）。だが、この方針に従うと無意識の情動はありえないことになる。先述したように、プリンツは無意識の情動はありうると考えているので、この方針は支持できない。またプリンツの内観に従えば、実際のところ、感情価ごとに独特の感じが存在するという想定は疑わしい。部分的に共通の身体変化がある情動には、その感じにも共通する部分がありうるが（プリンツは快の感じには共通する核として喜びの情動があると考えている）、独自の身体変化による情動には共通の感じはないように思われるのだ（例えば、典型的に負の情動と考えられる恐怖や悲しみには、それぞれ全く異なる感じがある）。このように内観に訴えた議論は決定的なものではないが、プリンツは独自に情動的意識に関する理論を提示することで、感情価の感じを否定する明確な根拠を論じてもいる。しかし、これを紹介することも、本稿の範囲を超えるだろう。従って、さしあたり本稿では内観に訴えた反論に満足するものとしよう。

感情価を特徴づける理論は他にもあるが、これらのうちでプリンツが有力とみなすのは、学習理論の道具立てを用いたグレイの理論である（Gray, 1981, 1982, 1987）。これは情動と強化子を結びける理論である。学習理論において「強化子」は、有機体が示す反応の確率を上げたり（i.e., 正の強化子）、下げたり（i.e., 負の強化子）する刺激のことを指している。グレイによると各強化子に対する行動反応は、脳内のBAS（行動接近系: behavioral approach system）とBIS（行動抑制系: behavioral inhibition system）により編成されている（11）。そしてグレイは、このBAS／BISの活動水準と感情価を同一視する。すなわち、感情価は強化子に対する反応調整だと主張するのである。例えば、正の情動は接近行動を促すもので、負の情動は抑制行動を促すものとなる。

しかし、プリンツは、このグレイの主張自体に対して同意していないという点に注意が必要である。その理由の一つは、この主張は情動の感情価が固定的であるという

結論を導くためである。しかし実際には、ある情動が異なる感情価を持つことは十分にありうる。例えば、驚きには宝くじの当選のような正の驚きもあれば、有名人の訃報のような負の驚きもあるだろう。また典型的な負の情動である恐怖や怒りが、デモ行進のような積極的な行動を促すこともある。その反対に、喜びに浸って足を止めることもあるかもしれない。

プリンツは感情価を強化子への反応と同一視する代わりに、感情価を強化子そのものと同一視することを提案している。一般的に、学習理論における強化子には報酬や罰として働く外的刺激が想定されている。しかし、何が報酬となり何が罰となるかは、生得的に決定されるだけでなく、古典的条件づけを通して学習することができる。その結果、給餌に先行して鳴らされるベルの音のように、ほとんど何でも強化子になりうる。このように強化子が多様な形態をとりうることから、脳には強化子を追跡し続ける方法が必要となる。すなわち、どの外的刺激が報酬／罰であるかをラベリングするような仕組みが必要となる。そしてプリンツは、こうしたラベルとなる内的強化子こそが、感情価の正体であると考える。つまり情動の正負は、正の内的強化子（IPR: inner positive reinforcer）と負の内的強化子（INR: inner negative reinforcer）にそれぞれ対応するのである。IPR と INR はそれぞれが内的命令として機能する。すなわち、IPR は「これを増やせ！」、INR は「これを減らせ！」という指令となり、この指令は将来の反応確率に影響するマーカーとして機能する。例えば、過去に投資で手痛い失敗をした記憶に INR が付けば、その人は将来において無茶な投資を控えるようになるだろう。無論、投資の成功体験に IPR が付けば、将来において積極的な投資をするようになる。つまり、感情価とは身体性の評価を将来に役立てる仕組みに他ならないのである。

しかし、情動には身体性の評価が含まれるとしても、そこには対処が含まれていないことに注意しなくてはならない。例えば、恐怖が危険を表象していても、依然として危険に対して無頓着でありうるのである。しかし、恐怖と負の強化子が一緒になると、危険を避けるという適応的な行動が導かれるようになる。このとき内的強化子説ならば、恐怖から接近／抑制のいずれも生じうることも問題なく説明できる。なぜなら、危険を避けるために適切な行動は状況により様々でありうる。山中でクマのような大型動物の足跡を発見したときは、いち早くその場から逃走すべきだろう。だがす

でに遭遇している場合は、じっと動かずやり過ごすことが有効になるかもしれない。いずれにせよ、その状況を満たす情動は恐怖であり（12）、それらの行動は危険の回避を目的にしているのである。

結びに

本稿ではプリンツの身体評価仮説を検討してきた。ここまでの議論のプリンツの理論の要点を再確認すると、情動を構成するのは「身体性の評価」と「感情価マーカー」に他ならない。プリンツの理論では、情動は前者のみで個別化され、また情動の感じも前者のみに依存している。しかし、前者だけでは情動は生存において有益にならない。情動が身体を介して中心的関係主題を表象するのは、その表象を感情価マーカーという内的強化子と組み合わせて、適応的な行動を導くことに利用できるからである。

このような情動の有用性は認知能力の観点からも裏打ちされる。すなわち、ヒトは発達した認知能力を持つが、多くの動物はこれを持たない。しかし、そのような動物も（あるいはヒトも）、情動があれば複雑な環境において適応的な行動を迅速に選択できるだろう。すなわち、情動は高度な認知の前に動物に備わった環境の評価システムなのである。

註

(1) パンクセップは、英語圏で行われた調査において、多くの人は情動に関わる様々な要素のうち感じを最も重視しているという点を指摘する（Panksepp, 2000）。管見によれば、こうした事情は本邦でも同様と思われる。

(2) 実際、神経科学者のダマシオは、身体変化にホルモンの分泌まで含めることでジェイムズを援護している（Damasio, 1994）。さらに、ダマシオは身体変化を表象する脳中枢が活性化さえすれば、実際に身体変化が生じていない場合でも、情動が生じうるのだと主張する。ダマシオによれば、人間には身体変化に関わる脳領

域へ内因性の刺激を送る経路があり（ダマシオはこれを「あたかもループ（as if loop)」と呼ぶ)、この経路により実際の身体変化がなくとも、情動的な状態をシミュレートできるのである。以上より、身体変化の範囲を広げれば、ジェイムズ—ランゲ説はキャノン流の批判からも十分に擁護されうると考えられる。なおプリンツは目下の理論形成において、幅広い範囲で身体変化（e.g., 表情、心拍数、ホルモンの分泌）を捉えている（Prinz, 2004)。また体性感覚　(somatosensory）も、これに準じた幅広いの感覚を指して用いられる。すなわち、皮膚感覚や固有感覚だけではなく、内受容感覚まで含めた用語である。

(3) プリンツ（1991）が挙げる以外にも、ジェイムズ—ランゲ説を支持する根拠はある。例えば、キャノン（1927）の批判に反して、近年の研究では少なくとも一部の情動について、自律神経系の活動パターンから特定できる可能性が示唆されている（Levenson, Ekman, & Friesen, 1990; Stephens, Christie, & Friedman, 2010)。あるいは内受容感覚への気づき（e.g., 心拍の感覚）と情動との相関を示した研究も存在する（Barrett, Quigley, Bliss-Moreau, & Aronson, 2004; Pollatos, Gramann, & Schandry, 2007; Wines, Mezzacappa, & Katkin, 2000)。ただし、「自律神経系による特定化」に関しては、社会構築主義的理論によるメタ分析の結果が好ましくないという批判もなされており、論争が続いている。この問題は身体的評価仮説と社会構築主義が対立する重要な論点になりうるが、これについては紙幅の関係上、別稿にて検討したい。

(4) 後述のように、プリンツは情動に認知が必要とする理論を厳しく批判しており、従って分子評価について詳しく説明する必要はないだろう。一方で、プリンツは「情動は中心的関係主題を表象する」という形で、ラザルス（1991）が提唱する中心的関係主題という概念を理論に組み込んでいる。ここで「表象する」が何を意味しているかについても後述に譲るとして、ここではプリンツの理論にとって、中心的関係主題が重要な役割をもつことを確認しておきたい。

(5) プリンツによれば、そうした理論の最初期のものはストア派においてみられる。

(6) このような評価する根拠として、プリンツは先行研究で提唱された定義を検討しているが、本稿では議論の詳細については割愛する。

(7) 無論、プーチン大統領のように時計を右腕に好む人物も存在するが、この場合に

おいても事情は同様である。

(8) 情動の具体例に下位分類を設ける方針の検討については、成瀬・鵜子（近刊）を参照していただきたい。

(9) 2020 年のスペインでは、実際に通りを歩く犬をライオンだと見間違えた人々からの通報騒ぎがあった。

(10) しかし近年になり、プリンツは情動を表象とみなさない立場に転向している (Shargel & Prinz, 2018)。すなわち、情動は有機体から独立して世界に存在する性質の中心的関係主題を表象する知覚ではなく、むしろ身体状態に依存した命令的アフォーダンスを産出（enact）する行為だと主張している。しかし、こうした転向は必ずしも身体性評価説の（ひいては本稿の）価値を失わせるものではないだろう。というのも、近年のプリンツがもつ意見とは関係なく、身体性評価説は未だ有望な情動理論であり、また近年の現象学的な立場にも、身体化という要素は引き継がれているからである。

(11) 紙幅の関係上、システムの神経相関項に関する提案については割愛する。

(12) 身体状態が異なるのに同じ情動なのかという疑問は当然のものだろう。しかし、プリンツの身体評価仮説では、情動の特定は、表象の実質的内容と名目的内容の両面から可能と考えているので、実質的内容（e.g., 危険）が同じなら、名目的内容（e.g., 逃走／硬直）が異なっても同じ情動（少なくとも近縁種の情動）だと主張できる。なお社会構築主義理論は、この点をもって身体説に対する反論を展開するだろう。

参考文献

Arnold, M. B. (1960). *Emotion and Personality*. New York: Columbia University Press.

Barrett, L. F., & Russell, J. A. (Eds.). (2015). *The Psychological Construction of Emotion*. New York: Guilford.

Barrett, L. F., Quigley, K. S., Bliss-Moreau, E., & Aronson, K. R. (2004). Interoceptive Sensitivity and Self-Reports of Emotional Experience. *Journal of Personality and Social Psychology,* 87(5), 684–697.

Cacioppo, J., BernstoD, G., Larson, J., Roehlmann, K., & Ito, T. (2000). The Psychophysiology of Emotion. In M. Lewis & J. Haviland-Jones (Eds.), *Handbook of Emotions* (2nd ed., pp. 173-191). New York: Guilford Press.

Cannon, W. B. (1927). The James-Lange Theory of Emotions: A Critical Examination and an Alternative Theory. *American Journal of Psychology*, *39*(1), 106-124.

Cornelius, R. R. (1996). *The Science of Emotion: Research and Tradition in the Psychology of Emotions.* New Jersey: Prentice Hall.

Damasio, A. R. (1994). *Descartes' Error: Emotion, Reason, and the Human Brain*. New York: Putnam.

Damasio, A. R., Grabowski, T. J., Bechara, A., Damasio, H., Ponto, L. L. B., Parvizi, J., & Hichwa, R. D. (2000). Subcortical and Cortical Brain Activity during the Feeling of Self-Generated Emotions. *Nature Neuroscience,* 3, 1049-1056.

Darwin, C. (n.d.). *The Expression of the Emotions in Man and Animals.* Chapter, III: Project Gutenberg. Retrieved from eBook Collection (EBSCOhost), EBSCOhost (accessed February 24, 2021) (Original work published 1872)

Dretske, F. (1981). *Knowledge and the Flow of Information.* Cambridge: MIT Press.

———. (1986). Misrepresentation. In R. Bogdan (Ed.), *Belief: Form, Content and Function* (pp. 17-36). Oxford: Oxford University Press.

Ekman, P. (2003). *Emotions Revealed: Recognizing Faces and Feelings to Improve Communication and Emotional Life.* New York: Times Books.

Elliott, R., Friston, K. J., & Dolan, R. J. (2000). Dissociable Neural Responses in Human Reward Systems. *Journal of Neuroscience,* 20, 6159-6165.

Frijda, N. H. (1993). Moods, Emotion Episodes, and Emotions. In M. Lewis & J. M. Haviland (Eds.), *Handbook of Emotions* (pp. 381-404). New York: Guilford Press.

Gray, J. A. (1982). *The Neuropsychology of Anxiety*. New York: Oxford University Press.

———. (1987). *The Psychology of Fear and Stress* (2nd ed.). Cambridge, UK: Cambridge University Press.

———. (1991). The Neuropsychology of Temperament. In J. Strelau & A. Angleitner (Eds.), *Explorations in Temperament* (pp. 105-128). New York: Plenum Press.

Griffiths, P. E. (1997). *What Emotions Really Are*. Chicago: University of Chicago Press.

James, W. (1884). What is an Emotion? *Mind*, *19*, 188-205.

Kenny, A. (1963). *Action, Emotion and Will*. London: Routledge & Kegan Paul.

Lane, R. D., Fink, G. R., Chau, P. M. L., & Dolan, R. J. (1997). Neural Activation during Selective Attention to Subjective Emotional Responses. *NeuroReport*, 8, 3969-3972.

Lange, C. G. (1885/1922). The Emotions: A Psychophysiological Study. In C. G. Lange & W. James (Eds.), *The Emotions* (pp. 33-90). Baltimore: Williams and Wilkins.

Lazarus, R. S. (1984). *On the Primacy of Cognition. American Psychologist, 39*(2), 124–129.

———. (1991). *Emotion and Adaptation*. New York: Oxford University Press.

Lazarus, R. S., & Alfert, E. (1964). Short-circuiting of threat by experimentally altering cognitive appraisal. *The Journal of Abnormal and Social Psychology*, *69*(2), 195.

LeDoux, J. (1996). *The emotional brain*. New York: Simon & Schuster.

Levenson, R. W. Ekman, P., & Friesen, W. V. (1990). Voluntary facial action generates emotion specific autonomic nervous system activity. *Psychophysiology*, 27, 363-384.

Panksepp, J. (2000). Emotions as Natural Kinds within the Mammalian Brain. In M. Lewis & J. Haviland-Jones (Eds). *Handbook of Emotions* (2nd. Ed., pp-137-156). New York: Guilford Press.

Pollatos, O., Gramann, K., & Schandry, R. (2007). Neural Systems Connecting Interoceptive Awareness and Feelings. *Human Brain Mapping,* 28(1):9-18.

Prinz, J. J. (2004). *Gut Reactions: A Perceptual Theory of Emotion.* Oxford, UK: Oxford University Press.

Roseman, I. J., Spindel, M. S., & Jose, P. E. (1990). Appraisals of Emotion-Eliciting Events: Testing a Theory of Discrete Emotions. *Journal of Personality and Social Psychology*, *59*(5), 899.

Schachter, S. (1964). The Interaction of Cognitive and Physiological Determinants of Emotional State. In L. Berkowitz. (Ed.), *Advances in Experimental Social Psychology* (pp. 49-80). New York: Academic Press.

Scherer, K. R. (1993). Studying the Emotion-Antecedent Appraisal Process: An Expert System Approach. *Cognition & Emotion*, *7*(3-4), 325-355.

Shargel, D., & Prinz, J. J. (2018). An Enactivist Theory of Emotional Content. In the Ontology of Emotions, H. Naar & F. Teroni (Eds.), *The Ontology of Emotions* (pp. 110-129), Cambridge: Cambridge University Press.

Smith, C. A., & Lazarus, R. S. (1993). Appraisal Components, Core Relational Themes, and the Emotions. *Cognition & Emotion*, *7*(3-4), 233-269.

Speisman, J. C., Lazarus, R. S., Mordkoff, A., & Davison, L. (1964). Experimental Reduction of Stress Based on Ego-Defense Theory. *The Journal of Abnormal and Social Psychology*, *68*(4), 367.

Stephens, C. L., Christie, I. C., & Friedman, B. H. (2010). Autonomic Specificity of Basic Emotions: Evidence from Pattern Classification and Cluster Analysis. *Biol Psychol,* 84(3):463-73.

Strack, F., Martin, L. L., & Stepper, S. (1988). Inhibiting and Facilitating Conditions of the Human Smile: a Nonobtrusive Test of the Facial Feedback Hypothesis. *Journal of Personality and Social Psychology,* 54(5), 768.

Tangney, J. P., & Fischer, K. W. (Eds.). (1995). *Self-Conscious Emotions: The Psychology of Shame, Guilt, Embarrassment, and Pride*. New York: Guilford Press.

Wines, S., Mezzacappa, E. S., & Katkin, E. S. (2000). Heartbeat Detection and the Experience of Emotions. *Cognition and Emotion*, 14(3), 417-427.

Zajonc, R. B. (1984). On the Primacy of Affect. *American Psychologist, 39*(2), 117–123.

Zajonc, R. B., Murphy, S. T., & Inglehart, M. (1989). Feeling and Facial Efference: Implications of the Vascular Theory of Emotion. *Psychological Review, 96*(3), 395.

成瀬 翔・鵜子 修司（近刊）「情動に認知はいらない？――ザイアンス vs. ラザルス論争の再検討」名古屋哲学研究会編『哲学と現代』第 36 号、頁数未定