

# 認知発達ロボティクスから探る「感性」の発達

— 人と感性を共有するロボットを目指して —

Cognitive Developmental Robotics toward Emergence of KANSEI in Robots

長井 志江  
Yukie NAGAI

大阪大学大学院工学研究科  
Graduate School of Engineering, Osaka University

**Keywords:** Cognitive developmental robotics, Emergence of KANSEI, Internal model, Predictive learning

## 1. はじめに

ロボットは「感性」を持つことができるのであろうか？  
そもそも、ロボットにとっての「感性」とはいったい何か？  
本特集号の執筆依頼を受けたとき、最初にこの疑問が頭に浮かんだ。著者はロボットの認知機能やコミュニケーション能力について研究しているが、これまで感性という言葉を用いたことはなく、どちらかといえば、そのような心的機能に関する表現を避けてきたように思う。その理由は、定義の曖昧性にあり、他の心的機能を示す用語にもよく見られることであるが、定式化が非常に困難といった問題が挙げられる。

一般に、感性は外界の刺激に対する心の働きであり、感覚を通して入力される信号に、何らかの意味や価値を見出す能力であると定義される。例えば、感性が豊かな人は、外界で生じている事象の真の意味を理解し、それを様々な視点から評価することができるし、他者の心的状態といった、直接には知覚できない状態にも気づくことができる。また、コミュニケーションにおいて他者と感性を共有することで、そこに新たな共通認識が生まれる。物理的にやりとりされる信号に、内的に共有された価値や意味が付加されることで、より豊かで効果的なコミュニケーションが実現されると期待できる。

本稿では、感性を持ったロボットの実現を目指して、まず、感性を計算論的視点から再考したいと思う。ロボットが、感覚信号に内在する意味や価値をどのようにして発見するのか、そのメカニズムについて考える。そして、認知発達ロボティクス[1-3]の視点から、ロボットが自らの感覚・運動経験を通して、感性を獲得する仕組みについて議論する。ロボットが人と環境との相互作用を通してどのような感性を獲得するのか、また、その獲得された感性が人と何かしら共通性を持ちうるのかについて、自他認知と利他的行動の発達モデル[4-6]を例に述べる。

## 2. 感性の計算論的モデル

ロボットの感性について議論するために、まずは、感性を計算論的視点から捉え直したいと思う。前節で述べたと

おり、感性とは、感覚を通して入力される信号に、何らかの意味や価値を見出す能力である。これを工学的に解釈すれば、単なる信号識別の枠を超えて、信号に内在する仕組みや原理を理解する能力であると捉えることができる。このような能力は、実は、感性に限らず様々な（特に、高次の）認知機能に共通しており、一般には内部モデル[7]と呼ばれている。

### 2.1 社会的認知発達の基盤としての内部モデル

著者は、内部モデルの獲得を通じた社会的認知機能の発達について、計算論的枠組みを提案してきた[10]。図1に、著者が提案する内部モデルの概念図を示す。内部モデル（予測器）は現在（時刻  $t$ ）の感覚信号  $s_i(t)$  と、運動指令  $a_j(t)$  の遠心性コピーをもとに、次時刻  $t+1$  の感覚信号  $\hat{s}_i(t+1)$  と運動指令  $\hat{a}_j(t+1)$  を予測する役割をもつ。実際に  $a_j(t)$  を実行した場合に、身体（感覚運動器）を通して知覚される感覚信号  $s_i(t+1)$  を脳がシミュレートできるように、内部モデルが感覚と運動の対応関係を獲得する。ここで、実際の感覚信号  $s_i(t+1)$  と予測した感覚信号  $\hat{s}_i(t+1)$  の差  $e_i(t+1)$  は予測誤差と呼ばれ、内部モデルの更新や行動生成のトリガとして用いられる（詳細は後述）。従来から、内部モデルは身体の制御や環境の理解に、重要な役割を果たすことが指摘されてきた[7-9]。内部モデルを持つことで、人は目標指向的な運動を生成したり、外界の事象を予測することができるようになる。

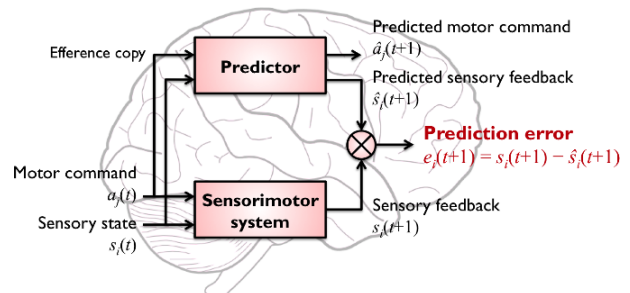


図1：感覚・運動情報の予測器としての内部モデル（[10]より引用）。感性がもつ信号の意味や価値を見出す能力は、内部モデルの予測機能として説明される。

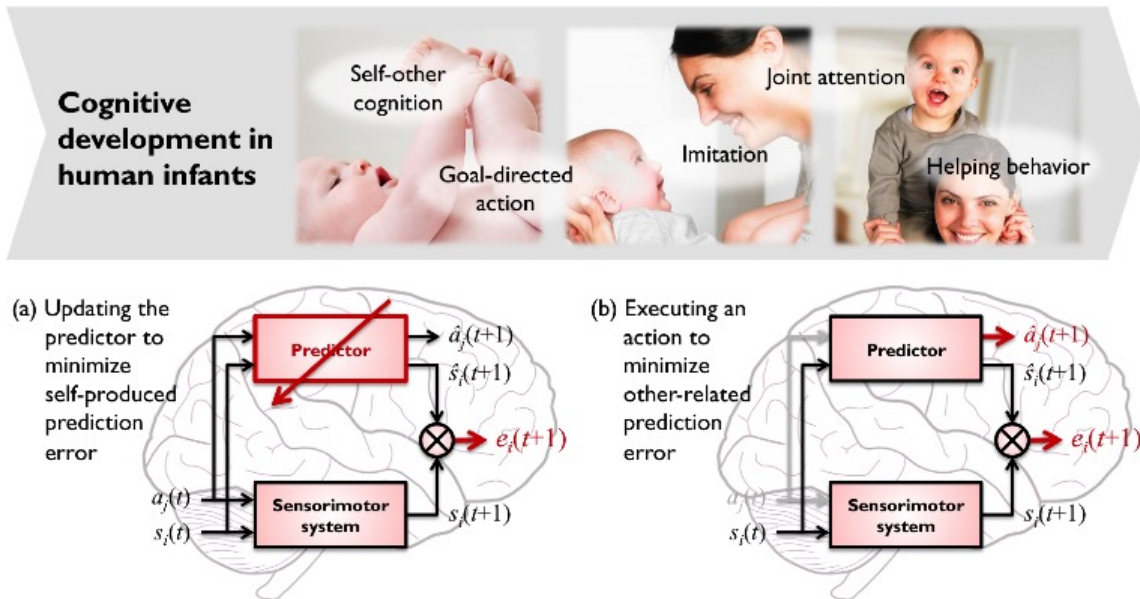


図2：内部モデルの学習とそれを利用した社会的認知機能の発達 ([10]より引用)

さらに著者は、この内部モデルの学習と利用の過程で、自他認知や模倣、利他的行動といった多様な社会的認知機能が発達すると提案してきた[10]。図2に乳幼児が生後、数年の間に獲得する社会的認知機能と（上段）、それを実現しているメカニズムを示す（下段）。ここでは便宜上、発達を二段階に分けて考える。最初の段階は、図2(a)に示す、予測誤差の最小化を規範とした内部モデルの学習である。生後まもない乳幼児は感覚・運動経験が乏しく、未熟な内部モデルを持って生まれる。そして、身体バブリングなどの探索行動を通して、自己の身体（内受容）や環境（外受容）からの感覚信号を獲得し、その信号とそれを生成した運動との対応関係を学習していく。この、内部モデルの学習過程で発現するのが、自他認知の能力である。乳幼児は全ての感覚信号に対して予測学習を行うと、予測誤差がゼロに収束する信号群と、そうでない信号群に大別されることに気づく。前者は自己に対応し、後者は他者（非自己）に対応する。自己の身体は、ある運動指令に対して、毎回、同じ感覚信号をフィードバックするため、感覚と運動の間には高い予測性が存在する。これに対して、他者や物体などの非自己は、感覚と運動の対応関係をいくら学習したとしても、不可観測な他者の内的状態（意図や信念など）や文脈の影響を受けるため、多少の予測誤差を残す。つまり、内部モデルを更新することで、乳幼児は自己身体の制御を学習するだけでなく、自己と他者の違いにも気づくことができるようになる。

第二の段階は、図2(b)に示す、他者運動に起因する予測誤差をトリガとした、運動の生成である。これは、模倣や利他的行動といった、社会的機能の創発に関連している。上述したとおり、他者の運動は自己の運動に比べて予測性が低く、内部モデルを学習したとしても、いくらかの予測誤差を残す。特に、他者が動作を失敗した場合などは、大

きな予測誤差が検出される。しかし、自他認知の能力がまだ未熟である場合、乳幼児はこの誤差がどのような原因によって引き起こされているかを知らないため、あくまでも誤差を最小化するように振る舞うと推測される。ここで、著者らが仮説として考えているのは、内部モデルで予測した運動  $\hat{a}_i(t+1)$  を自己の運動として生成することで、予測誤差を最小化するという仕組みである。例えば、他者の運動を観察しているときに、予測誤差をトリガとして  $\hat{a}_i(t+1)$  を出力すると、自己の運動が感覚信号  $\hat{s}_i(t+1)$  に大きく影響するため予測誤差を減少することができる。そして興味深いことに、内部モデルが他者の運動に類似した感覚・運動信号を経験していた場合、 $\hat{a}_i(t+1)$  は他者の運動をなぞるような形、つまり模倣運動として現れる。また、他者が動作の途中でそれを失敗・中断した場合は、 $\hat{a}_i(t+1)$  は他者の次時刻の運動に類似するので、乳幼児がそれを実行することで利他的行動が創発する。この段階で、乳幼児には利他的に振る舞う意図はないが、予測誤差の最小化規範に従うことで、社会的行動の原初が出現しうると考えられる。

## 2.2 内部モデルと感性

ここで、感性の話に戻りたいと思う。先述したとおり、著者は、感性も内部モデルの働きとして説明できると考えている。自他認知や利他的行動が、内部モデルの学習やそれを利用した運動生成の結果として創発するように、感性も内部モデルの獲得を通して発達すると考えている。

例えば、感覚信号の処理において、内部モデルは現在の信号をそのまま識別するのではなく、過去の感覚・運動経験に基づいて、その意味や価値を理解することを可能にする。ある物体の機能を推測しようとした場合に、内部モデルを用いることで、そこに蓄えられた過去の類似物体の知識や、それと関連した運動を参照することができ、物体の

表層的な特徴のみにとらわれない、本質的な機能の推定が実現される。また、他者とのコミュニケーションにおいては、他者が表出する運動や表情を自己の内部モデルに基づいてシミュレートすることができるので、直接は知覚することのできない他者の心的状態に気づいたり、他者の将来の行動を予測して、それに応じた応答をとることが可能になる。つまり、内部モデルがもつ感覚・運動情報の予測能力は、感性の機能として挙げた感覚信号に内在する意味や価値の発見や、他者の心的状態など不可観測な情報への気づきを可能にしていることが理解できる。

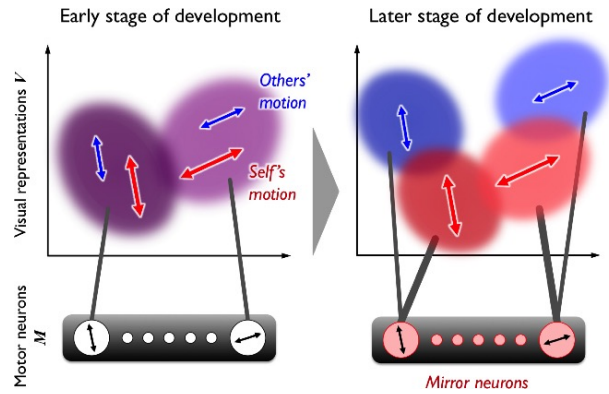
### 3. 内部モデルの学習を通じたロボットの社会的認知発達

感性を内部モデルの働きとして捉えることで、計算論的にモデル化できることが分かってきた。本節では、著者がこれまでに取り組んだ自己認知と利他的行動の発達を例に、内部モデルの学習を通じたロボットの社会的認知発達について述べる。

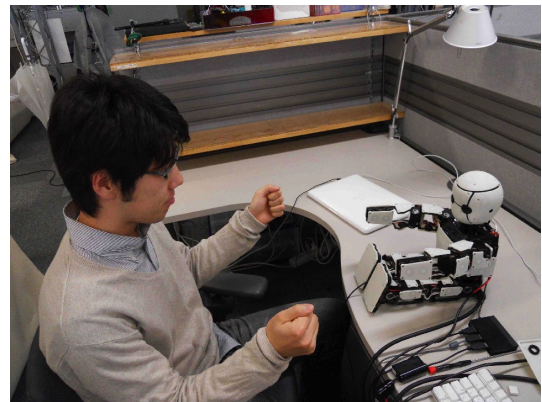
#### 3.1 自己認知の発達とそれを通じたミラーニューロンの創発

一つ目の例は、内部モデルの学習を通じた自己認知の発達である。人間の乳幼児は自己と他者の境界があいまいな状態で生まれ、他者を含む環境との相互作用を通して、徐々に自他を識別するようになると言われている[11-13]。例えば、乳児に自己身体のリアルタイムの映像と、時空間的に操作した（映像を左右／上下反転したり、時間遅れを挿入した）映像を呈示して、どちらをより長く注視するか調べると（選好注視法）、低月齢の乳児では選好性を示さないのに対して、高月齢（条件によって異なるが、4から6ヶ月くらい）の乳児では非自己の映像をより長く見る、つまり、自己と非自己を区別することが知られている。また、乳幼児は随伴性に対して非常に敏感であり[14,15]、自己認知においても随伴性が指標となっていることが示唆されている。

著者らはこれらの知見を参考に、感覚・運動情報の予測学習を通して、ロボットが自己認知の能力を獲得するモデルを提案した[4,5]。ここでの仮説とは、生後まもない乳幼児は感覚・運動能力が未熟なため、自己と他者の予測性の違いに気づくことができない（つまり、自他が未分化である）が、感覚・運動機能の発達とそれと並行した内部モデルの学習により、自他間の予測性の違いに気づくようになり、徐々に自己と他者を識別するようになるというものである。図 3(a) に提案した計算論的発達モデルを、(b) にそれをロボットに実装して実験を行った様子を示す。提案したモデルは、視覚表象（図 3(a)上段）と運動表象（下段）の二つの層からなり、左が発達初期の状態を、右が発達後期の状態を表している。ロボットはこのモデルをもとに、まず運動表象層にある運動ニューロンからランダムに一



(a) 自己認知の計算論的発達モデル。ロボットは視覚表象（上段）と運動表象（下段）の予測性を学習することで、自己と他者を識別し、さらに、それと並行して知覚が徐々に精緻化することで、自他の対応関係をミラーニューロンとして獲得する。



(b) 自己認知の発達のためのインタラクション実験。ロボットは人との随伴的なやりとりを通して、自他の対応関係を学習する。

図 3：内部モデルの学習を通じた自己認知の発達（[4, 5]より引用）

つを発火させ、それに対応する運動を生成する。ここでは、腕を上下／左右に動かす運動を用意している。そして運動実行中に自己の腕の動きを視覚上で観察しながら、ロボットは検出したオプティカルフローを視覚表象層でクラスタリングし、そのクラスタと発火させた運動ニューロンとの関係をヘップ則で学習する（視覚表象と運動表象間の結合）。ここで重要なのは、発達初期では自己と他者の動きが同一のクラスタとして認識され、区別されないが、発達とともにクラスタが徐々に分化していくという点である。図 3(b) に示すように、人（他者）はロボット（自己）と対面した状態で、養育者と乳幼児の関係と同様に、ロボットに対してある程度（完全ではない）随伴的に応答すると、発達初期ではロボットの感覚が未熟なため、ロボットは自己の動きと他者の動きの時間的・空間的ずれを検

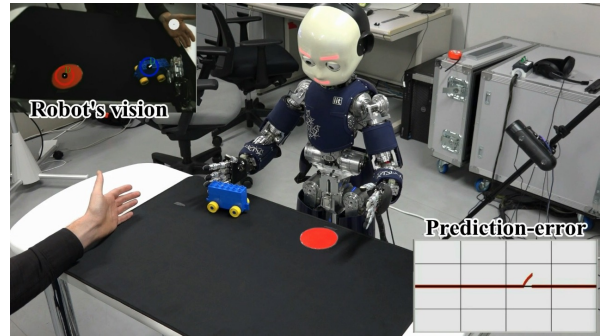
出ることができず、両方を同一視したクラスタを形成してしまう(図3(a)左参照)。そして、このクラスタは同時に発火した運動ニューロンと結合されるので、自己と他者の区別ができない状態を生み出す。一方、発達が進み感覚が精緻化すると、ロボットは他者の動きに内在する時間的・空間的ずれを検出するようになり、自己と他者の動きを別々の視覚表象クラスタとして認識する(図3(a)右参照)。これが、運動ニューロンとの結合強度にも違いを生じ、結合の強いクラスタを予測性の高い自己として、弱いクラスタを予測性の低い他者として認識するようになる。

以上が感覚・運動の予測性の学習に基づく自己認知の発達であるが、本モデルには、もう一つの興味深い特徴が存在する。それは自己認知の発達過程で創発する、ミラーニューロンである。ミラーニューロンは運動野に存在するニューロンで、自己の運動を生成するときだけではなく、それと同じ他者の運動を観察しているときにも発火するという、鏡のような機能をもつ[16,17]。ミラーニューロンを利用することで、人は他者の運動をまるで自己の運動であるかのように認識することができるし、他者の運動の予測や、意図などの心的状態の理解といった、感性が持つとされる能力も実現することができる。著者らが提案したモデルでは、学習の結果、獲得された感覚・運動のマップに、ミラーニューロンと同様の機能が創発した(図3(a)右参照)。視覚・運動表象間の結合を見ると、運動ニューロンが視覚表象層の自己の動きだけではなく、それと等価な他者の動きとも結びついていることが確認できる。これがまさに、ミラーニューロンである。本実験は自己認知能力の発達に焦点を当てているが、本特集号が目にする感性も、ミラーニューロンのような形で内部モデルに獲得されることを示唆している。

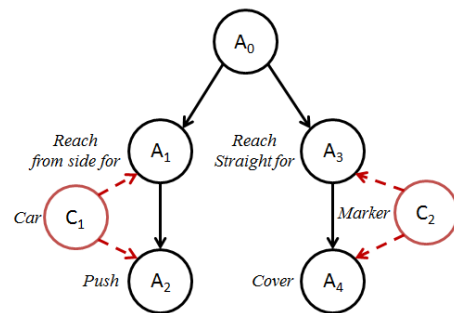
### 3.2 利他的行動の創発

二つ目の例は、内部モデルを用いた運動の予測と生成による、利他的行動の創発である。発達心理学研究によると、乳幼児は14ヶ月頃からすでに、自発的な利他的行動を示すことが知られている[18,19]。例えば、実験者が物を床に落とすと、それを拾って実験者に渡したり、実験者が両手に本を抱え、本棚の扉を開けられず困っていると、実験者の代わりに扉を開けるといった行動が観察されている。ここで注目すべき点は、乳幼児が実験者から手助けを求めるような信号(視線や発話)や、利他的行動に対しての報酬を受けているわけではないのに、自発的に実験者を助けようとする点である。これは、別の言い方をすると、感性の豊かさから生じる行動とも捉えられ、利他的行動がどのようなメカニズムに基づいて創発しているのかを議論することで、感性の発達メカニズムにも迫れると期待できる。

著者らは、内部モデルを利用した運動の予測と、他者起因の予測誤差をトリガとした運動の生成が、利他的行動の創発につながると考えた[6]。図4(a)にロボットを使った実



(a) 利他的行動の創発のためのインタラクション実験。ロボットは他者運動に起因する予測誤差を最小化するために、内部モデルで予測した運動を生成する。



(b) 確率遷移モデルで表現されたロボットの内部モデル。C<sub>i</sub>が感覚信号を、A<sub>j</sub>が運動指令を表している。

図4：内部モデルを用いた運動の予測と生成による利他的行動の創発 ([6]より引用)

験の様子を、(b)に確率遷移モデルで表現したロボットの内部モデルを示す。ロボットはあらかじめ、環境に置かれた二つの物体(青い車、赤いマーカ)に対するアフォーダンスを、感覚・運動情報の予測学習を通して内部モデルに獲得する。例えば、青い車に対しては、それを押すことで車を前進させ、赤いマーカに関しては、手をマーカの上に置くことで、マーカを隠すことができる。本実験で用いた内部モデルでは、感覚信号(図4(b)のC<sub>i</sub>)と運動指令(A<sub>j</sub>)の確率遷移として、このアフォーダンスを表現している。そして、この内部モデルをもとに、他者とのインタラクションの中で感覚と運動の予測を行う。図4(a)では、実験者が手を伸ばして青い車を押そうとしているが、車が遠いためなかなか動作目標を達成することができない。するとロボットは、動作観察中の予測誤差をリアルタイムで計算し(図4(a)右下参照)、それが一定の閾値を越えたところで、内部モデルが予測した運動を生成して、予測誤差を最小化しようとする。先述したとおり、ロボットには予測誤差を最小化する動機しか存在しないが、それによって生成され

た運動が、結果的に他者の動作目標を達成することになり、利他的行動が創発するという仕組みである。

#### 4. おわりに

本稿では、感性を持ったロボットの実現を目指して、認知発達ロボティクスの視点から、感性の計算論的モデルを提案した。感性が持つ、感覚信号に内在する意味や価値の発見や、他者の心的状態などの不可観測な情報への気づきといった能力は、感覚・運動情報の予測性に基づく内部モデルの働きとして説明することができる。自他認知と利他的行動の創発を例にしたロボットの実験で、これらの能力が、環境や他者との相互作用を通して獲得されることが示された。

しかし、ここで獲得された内部モデルは感覚・運動情報の一次の予測性のみを評価しており、より高次の予測性を獲得するようなモデルの提案が望まれる。予測性がより高次の抽象化した表象で評価されることで、感性がより豊かなものになると期待される。本稿で提案したモデルが、人と共通した感性を持つロボットの設計の一助になれば幸いである。

#### 参 考 文 献

- 1) M. Asada, K. Hosoda, Y. Kuniyoshi, H. Ishiguro, T. Inui, Y. Yoshikawa, M. Ogino, and C. Yoshida, "Cognitive Developmental Robotics: A Survey," *IEEE Transactions on Autonomous Mental Development*, vol. 1, no. 1, pp. 12–34, 2009.
- 2) M. Lungarella, G. Metta, R. Pfeifer, and G. Sandini, "Developmental robotics: a survey," *Connection Science*, vol. 15, no. 4, pp. 151–190, 2003.
- 3) A. Cangelosi and M. Schlesinger, *Developmental Robotics*. The MIT Press, 2015.
- 4) Y. Nagai, Y. Kawai, and M. Asada, "Emergence of Mirror Neuron System: Immature vision leads to self-other correspondence," in *Proc. of the 1st Joint IEEE International Conference on Development and Learning and on Epigenetic Robotics*, 2011.
- 5) Y. Kawai, Y. Nagai, and M. Asada, "Perceptual Development Triggered by its Self-Organization in Cognitive Learning," in *Proc. of the 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2012, pp. 5159–5164.
- 6) J. Baraglia, Y. Nagai, and M. Asada, "Prediction Error Minimization for Emergence of Altruistic Behavior," in *Proc. of the 4th IEEE International Conference on Development and Learning and on Epigenetic Robotics*, 2014, pp. 273–278.
- 7) D. M. Wolpert, Z. Ghahramani, and M. I. Jordan, "An internal model for sensorimotor integration," *Science*, vol. 269, no. 5232, pp. 1880–1882, 1995.
- 8) K. Friston, J. Kilner, and L. Harrison, "A free energy principle for the brain," *Journal of Physiology Paris*, vol. 100, no. 1-3, pp. 70–87, 2006.
- 9) R. P. N. Rao and D. H. Ballard, "Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive field effects," *Nature Neuroscience*, vol. 2, no. 1, pp. 79–87, 1999.
- 10) Y. Nagai and M. Asada, "Predictive Learning of Sensorimotor Information as a Key for Cognitive Development," in *Proc. of the IROS 2015 WS on Sensorimotor Contingencies for Robotics*, 2015.
- 11) L. E. Bahrack and J. S. Watson, "Detection of Intermodal Proprioceptive-Visual Contingency as a Potential Basis of Self-Perception in Infancy," *Developmental Psychology*, vol. 21, no. 6, pp. 963–973, 1985.
- 12) P. Rochat and R. Morgan, "Spatial Determinants in the Perception of Self-Produced Leg Movements by 3- to 5-Month-Old Infants," *Developmental Psychology*, vol. 31, no. 4, pp. 626–636, 1995.
- 13) P. Rochat and T. Striano, "Who's in the mirror? Self-other discrimination in specular images by four- and nine-month-old infants," *Child Development*, vol. 73, no. 1, pp. 35–46, 2002.
- 14) P. Rochat, "Social contingency detection and infant development," *Bulletin of the Menninger Clinic*, vol. 65, no. 3, pp. 347–360, 2001.
- 15) J. Colombo, "Infants' detection of contingency: A cognitive neuroscience perspective," *Bulletin of the Menninger Clinic*, vol. 65, no. 3, pp. 321–334, 2001.
- 16) G. Rizzolatti, L. Fadiga, V. Gallese, and L. Fogassi, "Premotor cortex and the recognition of motor actions," *Cognitive Brain Research*, vol. 3, no. 2, pp. 131–141, 1996.
- 17) G. Rizzolatti, L. Fogassi, and V. Gallese, "Neurophysiological mechanisms underlying the understanding and imitation of action," *Nature Reviews Neuroscience*, vol. 2, pp. 661–670, 2001.
- 18) F. Warneken, F. Chen, and M. Tomasello, "Cooperative Activities in Young Children and Chimpanzees," *Child development*, vol. 77, no. 3, pp. 640–663, 2006.
- 19) F. Warneken and M. Tomasello, "The roots of human altruism," *British Journal of Psychology*, vol. 100, pp. 455–471, 2009.