

令和 2 年 6 月 2 日現在

機関番号：62615

研究種目：若手研究(B)

研究期間：2017～2019

課題番号：17K12754

研究課題名(和文) ロボット構成論による精神障害における多様な病態の統合的理解

研究課題名(英文) Integrative understanding of various symptoms of psychiatric disorders by constructive approach using robots

研究代表者

村田 真悟 (MURATA, Shingo)

国立情報学研究所・情報学プリンシプル研究系・助教

研究者番号：80778168

交付決定額(研究期間全体)：(直接経費) 3,200,000円

研究成果の概要(和文)：本研究課題は、精神障害における多様な症状を生み出す脳・神経メカニズムをロボット構成論によりシステムレベルで統合的に理解することを目指すものである。そのために、不確実性を考慮した予測誤差最小化が脳の神経回路における基本的な計算原理であるという仮説を提案し、その仮説を階層的な再帰型神経回路モデルで具現化し、ロボットに実装することで検証を行った。ロボットの行動学習実験の結果、不確実性推定の失調が予測誤差信号の減少・増加をもたらし、その結果として知覚や行動における異常が生じることがわかった。さらに、不確実性推定の失調は、階層間の機能的断裂や神経活動の均一化によってもたらされるということが示唆された。

研究成果の学術的意義や社会的意義

本研究は、不確実性を考慮した予測誤差最小化が人間の認知や行動における柔軟性の実現に欠かせない基本的な計算原理であり、その失調によって精神障害における多様な症状が理解できることを示した。これまでの精神障害を対象とした研究において、神経レベル、情報処理レベル、認知・行動レベル等、異なるレベルにおいて様々な知見・理論が提案されてきた。本成果は、神経科学・計算論・身体性認知科学といった観点を統合したロボット構成論的アプローチによって、これらの知見・理論の橋渡しをするものであり、精神障害の統合的理解に貢献することが期待される。

研究成果の概要(英文)：This research project aims at integrative and system-level understanding of the brain and neural mechanisms that produce various symptoms of psychiatric disorders by a constructive approach using robots. We hypothesize that the prediction error minimization considering uncertainty estimation is the fundamental computational principle in the neural networks of the brain. We developed a hierarchical recurrent neural network model based on the principle and evaluated the hypothesis by implementing the model into a robot. Experimental results on robot behavior learning demonstrate that aberrant uncertainty estimation decreases or increases prediction error signals, resulting in abnormalities in perception and action. Furthermore, the results also suggest that the aberrant uncertainty estimation is derived from disconnection between hierarchies and homogenization of neural activities.

研究分野：認知ロボティクス

キーワード：計算論的精神医学 認知ロボティクス 神経回路モデル ニューラルネットワーク 予測符号化

## 様式 C - 19、F - 19 - 1、Z - 19 (共通)

### 1. 研究開始当初の背景

精神障害の脳・神経メカニズムを理解するため、計算論的神経科学の手法を積極的に応用する試みを計算論的精神医学 [Montague et al., 2012] という。2013年には世界初となる計算論的精神医学に関する国際会議 Computational Psychiatry がアメリカで開催され、2017年には同名の論文誌が MIT Press Journals より創刊された。また、2015年にはドイツ Max Planck Institute とイギリス University College London (UCL)による共同研究所 Max Planck UCL Centre for Computational Psychiatry and Ageing Research が設立されるなど、近年その重要性が世界的に注目されている。本分野ではこれまで、強化学習に基づく意思決定過程の理論やベイズ推定に基づく脳の知覚体験様式の理論から正常モデルを構築し、その異常として病態の説明がなされてきた。しかし、認知発達ロボティクス [Asada et al., 2009] にみられるような、個体の身体性や環境との相互作用、発達過程を考慮した研究はほとんど行われていない。

上記背景を踏まえた上で、本研究の課題として次の二点が挙げられる。

#### (1) 正常メカニズムの理解

脳・身体・環境の相互作用を考慮し、正常とされる知覚・行動様式を実現するために必須の神経メカニズムを備えた計算モデルを構築し、実環境においてロボットの行動生成学習実験を行い検証する。

#### (2) 病態メカニズムの理解

学習後の正常モデルや学習過程のモデルに損傷を加えることで精神障害モデルを構築し、(1)と同様の実験を行い様々な病態メカニズムを統合的に理解する。

UCLの理論神経科学者 K. Friston による変分推論に基づいた自由エネルギー原理 [Friston, 2010] は現在、脳の計算原理として広く受け入れられており、同原理の計算論的精神医学への応用例も多くある。しかし、抽象的な数理モデルによって記述されているため、簡素な数値実験のみが行われてきた。本研究の特色は、「不確実性を考慮した予測誤差最小化」という計算原理によって自由エネルギー原理を支持し、神経回路モデルによる原理の具現化及び実環境におけるロボットの行動生成学習実験を行う点である。また、学習・発達面を考慮することで、過去の経験に基づくトップダウン予測と現在の感覚に基づくボトムアップ修正の両プロセスに関する異常性の検討が可能であると予想される。このような構成論的アプローチによって広範囲の精神障害における多様な病態メカニズムの統合的理解を試みる点は独創的であり、他に類を見ない。脳(神経回路モデル)・身体(ロボット)・多くの不確実性を有する実環境の相互作用の観点から計算原理を検証することは理論研究の多い本研究分野において大きな意義があると考えられる。

### 2. 研究の目的

本研究課題では、精神障害における多様な症状を生み出す脳・神経メカニズムをシステムレベルで統合的に理解することを目的とし、神経科学・計算論・身体性認知科学の観点を統合したロボット構成論的アプローチにより取り組む。具体的には、不確実性を考慮した予測誤差最小化原理に基づく階層的な再帰型神経回路モデル (Recurrent Neural Network: RNN) を構築し、ヒューマノイドロボットに実装する。まず(1)ロボットの行動生成学習実験を通して、人間の柔軟かつ適応的な認知メカニズムを実現する計算原理の理解を目指す(正常メカニズムの理解)。そして、(2)その正常メカニズムの異常・破綻として精神障害を捉え、多様な症状を生み出すメカニズムの理解を目指す(病態メカニズムの理解)。

### 3. 研究の方法

本研究課題では、まず初めに(1)正常メカニズムの理解を目的として、不確実性を考慮した予測誤差最小化が脳の神経回路における基本的な計算原理であるという仮説を提案する。提案原理を具現化する階層的な RNN を構築し、モデルを搭載したロボットの行動生成学習実験によってこの仮説を検証する。なお、ロボットのタスクとして、実験者とのボールの転がし合いを設定した(図1)。続いて、(2)病態メカニズムの理解を目的として、不確実性に関する推定の破綻が多様な病態メカニズムの基本原理になるという仮説を提案し、精神障害モデルの構築とロボットの行動生成学習実験によってこの仮説を検証する。特に、不確実性に関する推定の破綻がもたらす影響、その破綻をもたらす原因について考察する。

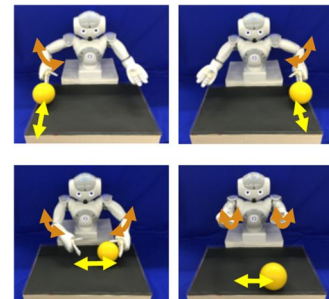


図1. ロボットが学習した行動パターン

#### (1) 正常メカニズムの理解

不確実性を考慮した予測誤差最小化に基づく階層的な RNN の構築とロボット実験による正常モデルの評価を行った。具体的には、これまでに研究代表者が提案してきた Stochastic Continuous-Time RNN (S-CTRNN) に Parametric Bias (PB) と呼ばれる潜在表現を付与することで階層性を持たせた。上位層としての PB は感覚運動情報のパターンの違いを抽象的に表現し、下位層としての S-CTRNN はその詳細を表現することが期待される。本モデルは、ロボットの感覚運動情報が正規分布に従うと仮定し、その平均と分散(不確実性)の予測学習を行う。

学習時は、感覚運動情報の観測及びモデルによって予測された平均と分散を用いて定義される誤差関数（負対数尤度）の最小化を目指して、勾配法によりモデルのパラメータ（結合重み・発火閾値）及びPBの潜在表現の最適化を行う。ここで、誤差関数は不確実性の対数と不確実性を除すことで重み付けられた予測誤差（重み付き予測誤差）の二項の和からなる。

学習後のモデルをロボットに搭載し、PBに基づきトップダウンに感覚運動情報及びその不確実性を予測生成し、実際の感覚運動情報を受け取ることで、誤差を算出しPBまでボトムアップに伝達する。正常モデルにおいては、モデルによって推定された不確実性によって適切に重み付けられた予測誤差を最小化するようにPBを修正することで、外部環境の変化に対して柔軟かつ適応的な知覚や行動が実現されると考えられる。

## (2) 病態メカニズムの理解

不確実性に関する推定の破綻がもたらす影響

(1)で構築したモデルによって推定される不確実性を強制的に減少・増加させることで異常モデル（精神障害モデル）を構築した。具体的には、モデルは不確実性として  $v_{t,i} = \exp(u_{t,i})$  ( $t$ は時刻、 $i$ は次元、 $u_{t,i}$ はニューロンの内部ポテンシャル)で表現される分散を推定するが、内部ポテンシャル  $u_{t,i}$ に対してパラメータ  $k = -8, -4, 0, 4, 8$  ( $k = 0$ の場合は正常モデルに対応する)を足し合わせることで不確実性の減少・増加をシミュレートした。(1)で述べたように、誤差関数において、不確実性を除すことで重み付けが行われているため、不確実性の減少は伝達される重み付き予測誤差の増加を意味する。また逆に、不確実性の増加は伝達される重み付き予測誤差の減少を意味する。このように不確実性に関する推定の破綻が生じることで、ボトムアップの重み付き予測誤差の伝達にも異常が生じると考えられ、その際の知覚・行動様式を観察し、正常モデルとの比較を行う。

不確実性に関する推定の破綻をもたらす原因

では、学習後のモデルに対して操作を加えることで、不確実性に関する推定の破綻をシミュレートした。ここでは、学習の過程においてそのような破綻をもたらす原因を検討する。具体的には、以下の二点について検討を行い、発達の精神障害モデルを構築した。

神経回路の階層間の機能的断裂

神経回路の学習時、結合重みの更新をする際に上位層のPBと下位層のS-CTRNNを繋ぐ重みに対して一様乱数を加えることで機能的断裂をシミュレートした。具体的には、学習時に毎回重み  $w_{ij}$ の大きさ  $|w_{ij}|$ に対して  $k$ 倍した値の負値・正値を下限・上限とする一様乱数を加えた。ここで、パラメータ  $k$ は  $k = 0, 0.2, 0.4$ とした ( $k = 0$ の場合は正常モデルに対応する)。

神経活動の均一化

神経回路において各ニューロンが有する発火閾値（バイアス）の影響を検討するため、その分布としてガウス分布を仮定し分散の値を操作することで、発火閾値の分布が異なる複数のモデルを用意した。具体的には、発火閾値を  $N(0, k)$ に従う正規乱数で初期化し、その値を固定した状態で結合重みとPBの学習を行った。ここで、パラメータ  $k$ は  $k = 0.1, 1, 10, 100$ とした。 $k$ が小さい場合は多くの発火閾値がほぼ0となるため、神経活動は入力に依存した過敏なものとなると考えられる。そして  $k$ を大きくしていくことで、発火閾値に多様性が生まれ、神経活動も多様になると考えられる。

## 4. 研究成果

### (1) 正常メカニズムの理解

正常モデルを搭載したロボットは事前にボールのダイナミクスに依存した四つの行動パターン（図1）を学習し、学習後に実験者とボールのやり取りに関するインタラクションを要求された。そのインタラクション中に実験者が行動パターンを切り替えることで外部環境に変化が生じた際も、ロボットは状況に応じた行動を生成することができた。これは、推定された不確実性によって適切に重み付けされた予測誤差が上位層のPBまで伝播し、適切にPBを修正することで実現されたと考えられる。

### (2) 病態メカニズムの理解

不確実性に関する推定の破綻がもたらす影響

(1)で学習済みのモデルを用いて、不確実性の推定値を強制的に減少・増加させた精神障害モデルにおいては、行動切り替えが困難であることや同一行動への固執・硬直といった多様なふるまいが観察された。図2にパラメータ  $k$ と行動の出現頻度の関係を示す。不確実性を減少させたモデルにおいては、正常な場合に比べて過大な重み付き予測誤差が生じることで不必要なPBの修正が行われ、それによって前述の異常行動が生成されたと考えられる。一方、不確実性を増加させたモデルにおいては、正常な場

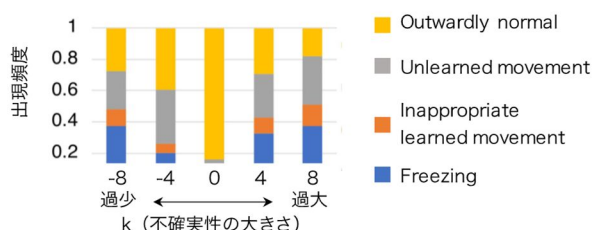


図2. 不確実性の大きさと出現した行動の関係

合に比べて過少な重み付き予測誤差が生じることで本来であれば必要なPBの修正が行われず、それによって前述の異常行動が生成されたと考えられる。これらの結果から、不確実性に関する推定の破綻が、固執的行動生成や対人やりとりにおける欠陥に寄与することが示唆された。

#### 不確実性に関する推定の破綻をもたらす原因

ここでは、ロボットの学習する行動パターンとして図1の上段の二つのみとし、以下の二条件で再度学習し直し検証を行った。

#### 神経回路の階層間の機能的断裂

機能的断裂がある( $k$ が大きい)場合、モデルによって推定される不確実性が増加し、それによって重み付き予測誤差が減少することでロボットは外部環境の変化に適応することが困難であるということがわかった。また、ロボットの行動パターンの違いを符号化する上位層の表現にも影響がみられ、特に機能的断裂が深刻な場合はパターンの違いによらず表現が一樣となり、下位層における反射的行動のみで外部環境の変化に適応することが可能であるということも確認された。図3にパラメータ  $k$  と行動の出現頻度の関係を示す。これらの結果から、機能的断裂によって不確実性に関する推定が正常時よりも増加し、それによって固執的行動生成や対人やりとりにおける欠陥に寄与することが示唆された。

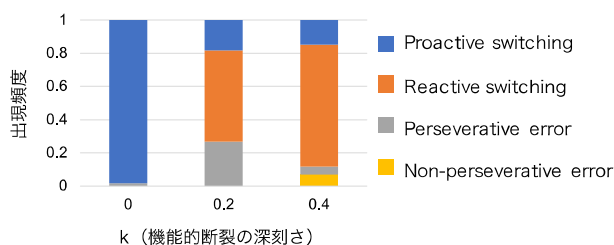


図3. 機能的断裂の深刻さと出現した行動の関係

#### 神経活動の均一化

発火閾値の分布が急峻な( $k$ が小さい)場合、モデルによって推定される不確実性が減少し、ロボットの行動生成は不安定となり、汎化性や柔軟性の低下、同一性保持、不適応な繰り返し行動等をもたらされることがわかった。一方、発火閾値の分布がなだらかな場合( $k$ が大きい場合)、十分な汎化性や行動の柔軟性を確認した。図4にパラメータ  $k$  と行動の汎化性・柔軟性の関係を示す。前述のように、分布が急峻な場合は多くの発火閾値がほぼ0となるため、神経活動は入力に依存した過敏なものとなり、上位層のPBをうまく利用することができず、下位層のみで無理に学習しようとした結果として行動の汎化性や柔軟性が低下したと考えられる。 $k$ が大きくなることで発火閾値の分布がなだらかになり、発火閾値は多様になる。それによって、神経活動には入力に依存しやすいもの、しにくいもの等の多様性が生じ、下位層と上位層のそれぞれにおいて特性にあった情報表現を行うことが可能となり、汎化性や柔軟性が獲得可能になったと考えられる。これらの結果から、ニューロンの発火閾値の分布に依存した神経活動の均一化によって不確実性に関する推定が正常時よりも減少し、それによって固執的行動生成や対人やりとりにおける欠陥に寄与することが示唆された。

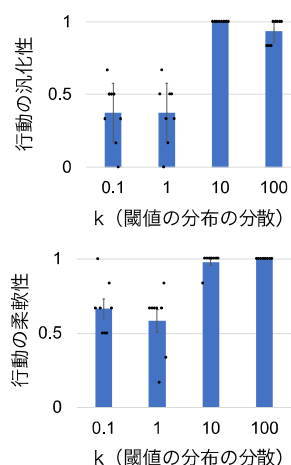


図4. 閾値の分布の分散と行動の汎化性・柔軟性の関係

以上の成果をまとめると、外部環境の変化に対する柔軟かつ適応的な知覚や行動の実現のためには、不確実性を考慮した予測誤差最小化が重要であることが示された。また、不確実性の推定が破綻し、推定される不確実性が増加・減少すると重み付き予測誤差の減少・増加が生じ、それに伴い正常時とは異なる知覚・行動様式が出現する。また、そのような不確実性の増加は神経回路の階層間の機能的断裂によってもたらされ、不確実性の減少は神経活動の均一化によってもたらされる。

これまでの精神障害を対象とした研究において、神経レベル、情報処理レベル、認知・行動レベル等、異なるレベルにおいて様々な知見・理論が提案されてきた。本研究成果は、神経科学・計算論・身体性認知科学といった観点を統合したロボット構成論的アプローチによって、これらの知見・理論の橋渡しをするものであり、精神障害の統合的理解に貢献することが期待される。

#### <引用文献>

- [Montague et al., 2012] R. Montague, R. Dolan, K. Friston, and P. Dayan, "Computational Psychiatry," *Trends in Cognitive Science*, Special Issue on Neuropsychiatric Disorders, Vol. 16, No. 1, pp. 72–80, 2012.
- [Asada et al., 2009] 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.
- [Friston, 2010] K. Friston, "The Free-Energy Principle: A Unified Brain Theory?" *Nature Reviews Neuroscience*, Vol. 11, No. 2, pp. 127–138, 2010.

5. 主な発表論文等

〔雑誌論文〕 計6件（うち査読付論文 6件／うち国際共著 0件／うちオープンアクセス 1件）

1. 著者名 出井 勇人, 村田 真悟, 尾形 哲也, 山下 祐一	4. 巻 62
2. 論文標題 不確実性の推定と自閉スペクトラム症 神経ロボティクス実験による症状シミュレーション	5. 発行年 2020年
3. 雑誌名 精神医学	6. 最初と最後の頁 219-229
掲載論文のDOI (デジタルオブジェクト識別子) 10.11477/mf.1405206009	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Murata Shingo, Yanagida Hikaru, Katahira Kentaro, Suzuki Shinsuke, Ogata Tetsuya, and Yamashita Yuichi	4. 巻 -
2. 論文標題 Large-scale Data Collection for Goal-directed Drawing Task with Self-report Psychiatric Symptom Questionnaires via Crowdsourcing	5. 発行年 2019年
3. 雑誌名 Proceedings of the 2019 IEEE International Conference on Systems, Man, and Cybernetics (SMC 2019)	6. 最初と最後の頁 3839-3845
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/SMC.2019.8914041	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Murata Shingo, Sawa Hiroki, Sugano Shigeki, and Ogata Tetsuya	4. 巻 -
2. 論文標題 Looking Back and Ahead: Adaptation and Planning by Gradient Descent	5. 発行年 2019年
3. 雑誌名 Proceedings of the Ninth Joint IEEE International Conference on Development and Learning and on Epigenetic Robotics (ICDL-EpiRob 2019)	6. 最初と最後の頁 151-156
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/DEVLRN.2019.8850693	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Idei Hayato, Murata Shingo, Chen Yiwen, Yamashita Yuichi, Tani Jun, and Ogata Tetsuya	4. 巻 2
2. 論文標題 A Neurorobotics Simulation of Autistic Behavior Induced by Unusual Sensory Precision	5. 発行年 2018年
3. 雑誌名 Computational Psychiatry	6. 最初と最後の頁 164-182
掲載論文のDOI (デジタルオブジェクト識別子) 10.1162/cpsy_a_00019	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

1. 著者名 Murata Shingo, Masuda Wataru, Tomioka Saki, Ogata Tetsuya, and Sugano Shigeki	4. 巻 10613
2. 論文標題 Mixing Actual and Predicted Sensory States Based on Uncertainty Estimation for Flexible and Robust Robot Behavior	5. 発行年 2017年
3. 雑誌名 Lecture Notes in Computer Science (LNCS)	6. 最初と最後の頁 11-18
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/978-3-319-68600-4_2	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Idei Hayato, Murata Shingo, Chen Yiwen, Yamashita Yuichi, Tani Jun, and Ogata Tetsuya	4. 巻 -
2. 論文標題 Reduced behavioral flexibility by aberrant sensory precision in autism spectrum disorder: A neurorobotics experiment	5. 発行年 2017年
3. 雑誌名 Proceedings of the Seventh Joint IEEE International Conference on Development and Learning and on Epigenetic Robotics (ICDL-EpiRob 2017)	6. 最初と最後の頁 271-276
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/DEVLRN.2017.8329817	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計9件（うち招待講演 0件 / うち国際学会 3件）

1. 発表者名 出井勇人, 村田真悟, 山下祐一, 尾形哲也
2. 発表標題 神経興奮性の変調による過学習と行動切り替えの困難さ RNNとロボットを用いた発達障害の知見の橋渡し
3. 学会等名 発達神経科学学会 第8回学術集会
4. 発表年 2019年

1. 発表者名 Murata Shingo, Hirano Kai, Higashi Naoto, Kumagaya Shin-ichiro, Yamashita Yuichi, and Ogata Tetsuya
2. 発表標題 Analysis of Imitative Interactions between Typically Developed or Autistic Participants and a Robot with a Recurrent Neural Network
3. 学会等名 The Ninth Joint IEEE International Conference on Development and Learning and on Epigenetic Robotics (ICDL-EpiRob 2019) (国際学会)
4. 発表年 2019年

1. 発表者名 盧承彩, 村田真悟, 澤弘樹, 尾形哲也, 菅野重樹
2. 発表標題 インタラクション生成に必要な情報の選択モデル 情報の不確実性に着目したモデルの検証
3. 学会等名 日本機械学会ロボティクス・メカトロニクス講演会2019
4. 発表年 2019年

1. 発表者名 村田真悟, 平野加依, 東直人, 熊谷晋一郎, 山下祐一, 尾形哲也
2. 発表標題 RNNを備えたロボットと定型発達者/自閉スペクトラム症者の模倣インタラクションの解析
3. 学会等名 発達神経科学学会 第7回大会
4. 発表年 2018年

1. 発表者名 宗近亮弥, 出井勇人, 村田真悟, 山下祐一, 尾形哲也
2. 発表標題 神経階層間の機能的断裂がもたらす感覚不確実性の推定異常 再帰型神経回路モデルを用いた精神疾患モデリング
3. 学会等名 発達神経科学学会 第7回大会
4. 発表年 2018年

1. 発表者名 Idei Hayato, Murata Shingo, Yamashita Yuichi, and Ogata Tetsuya
2. 発表標題 Altered Sense of Self Induced by Functional Disconnection in a Hierarchical Neural Network: A Neuro-Robotics Study
3. 学会等名 International Consortium on Hallucination Research and Related Symptoms Kyoto Satellite Meeting (ICHR 2018 KYOTO) (国際学会)
4. 発表年 2018年

1. 発表者名 村田真悟, 東直人, 平野加依, 山下祐一, 菅野重樹, 尾形哲也
2. 発表標題 RNNを備えたロボットと人間の模倣インタラクションの解析
3. 学会等名 発達神経科学学会 第6回大会
4. 発表年 2017年

1. 発表者名 出井勇人, 村田真悟, 山下祐一, 尾形哲也
2. 発表標題 ニューロンの興奮性の均一化による汎化と適応性の低下 人型ロボットを用いた精神障害モデリング
3. 学会等名 発達神経科学学会 第6回大会
4. 発表年 2017年

1. 発表者名 Idei Hayato, Murata Shingo, Yamashita Yuichi, and Ogata Tetsuya
2. 発表標題 Altered Behavioral Flexibility and Generalization Induced by Reduced Heterogeneity of Intrinsic Neuronal Excitability: A Neurorobotics Study
3. 学会等名 WPA XVII World Congress of Psychiatry (国際学会)
4. 発表年 2017年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
---------	---------------------------	-----------------------	----