学位請求論文審査の要旨

報告番号　　甲　第　　号

氏名　　山田　航太

論文題目　　学習心理学と行動分析学における数理工学的手法の実践と意義

［審査担当者］

主査　慶應義塾大学文学部教授・大学院社会学研究科委員

博士（農学）　　伊澤　栄一

副査　慶應義塾大学文学部教授・大学院社会学研究科委員

博士（心理学）　　梅田　聡

副査　慶應義塾大学文学部助教

博士（神経科学）　　兎田　幸司

副査　京都大学 大学院医学系研究科准教授

博士（科学）　　濱口　航介

論文要旨

1. 本論文の構成

本博士論文は，動物の報酬を伴う学習によって生じる行動および生理反応の変容と，学習を通して獲得された行動の特性を，数理工学的手法を援用して計測し，理論化することを通して，それらの背後にある心理学的および生物学的なメカニズムを明らかにすることを目的としている。本論文の構成は以下の通りである。

第１章：序論

学習をめぐる3つの学問

学習心理学とパブロフ型条件づけ

行動分析学とオペラント条件づけ

学習心理学と行動分析学の違い

機械学習が学習心理学と行動分析学にもたらすもの

強化学習について

強化学習とパブロフ型条件づけ

強化学習とオペラント条件づけ

計測技術と行動分析学・学習心理学の発展

本研究の位置づけ・目的

第２章：研究１ 瞳孔計測による報酬予測と報酬予測に基づく行動表出の分離

第３章：:研究２ 好奇心駆動型強化学習による消去バーストの制御変数の同定

第４章：研究３ 強化学習によるバウト・休止パターンのシミュレーション

第５章：研究４ 行動ネットワークの構造変化としての習慣形成

第６章：総括

強化学習による分野間の相互乗り入れ

計測技術に期待されるもの

学習心理学と行動分析学における計算論的手法の意義

2. 論文の概要

ヒトを含む生物個体は生存や繁殖のため，自身を取り巻く環境の中で，危険から身を守り，食餌を獲得し，繁殖のために配偶者を獲得する。しかし，環境は決して定常的ではなく，季節のような周期性を有する予測可能な変化や，突如とした天敵の出現，資源の枯渇などの予測不可能な出来事に満ちている。個体は自身の行動を変容させることによって，将来の類似した出来事に直面した時に，適切な行動を選択することができる。こうした経験によって生じる行動の長期的な変容は学習として定義され，経験に伴う時々刻々とした行動の変化や，経験を通して形成された行動を制御する心理学的および生物学的なメカニズムが，様々な動物種を用いた研究によって明らかにされてきた。

動物の学習能力を対象とした研究は古くから心理学が中心となって進められてきたが，現在では神経科学や工学分野へも研究が波及している。こうした学習研究への異分野の参入により，心理学が積み上げてきた知見に後押しされる形で，学習能力の背後にある生物学的なメカニズムや，より高度な理論的な研究が展開されている。一方で，心理学における学習研究は，伝統的な技術に依拠した方法論の上で展開されており，行動や生理指標の計測，そして行動やその背後にあるメカニズムの理論化という側面において，近年の分野横断的研究に対して遅れをとっていると言わざるを得ない状況になっている。

本論文では，こうした学習研究への異分野の参入がもたらした方法論的展開を，心理学において実践することで，心理学における学習研究を異分野へと融合し，発展させる道筋を計測と理論という2つの側面から照らし出す。心理学において，行動や生理指標の計測は実験

の中核をなすものである。動物を対象とした学習研究においては，動物が実験中に従事しうる行動は多岐にわたる。従来の計測技術では，それを生理指標とともに長期に計測することは困難であった。しかし，近年発展した機械学習や新たな実験装置の開発によって，技術的な問題は解決されつつあり，神経科学分野においては，心理学がこれまで扱ってきた行動や生理指標計測の精度や情報量を凌ぐ研究が広がっている。これまでにも心理学分野において学習過程や行動の数理的な記述や理論化がなされてきたが，近年の機械学習モデルや計算論的神経科学によって精緻化された理論的枠組みは，心理学における既存の行動研究に関する知見に新たな説明を与え，異分野との接合を促進することができる手法を数多く有している。本論文では，計測と理論という2つの側面から，従来の心理学が扱ってきた問題に取り組むことを通して，学習過程を経て獲得される行動の神経生物学的メカニズムと，その背後にある生成則を計算論的に明らかにする。

研究1では，画像処理と頭部固定装置という新たな解析技術と実験装置を用い，マウスの瞳孔サイズを定量化することを通じて，学習課題中の瞳孔の動態が報酬の予測を反映するかを検証した。ある音刺激の呈示が直後の報酬の到来を予測する条件と，音刺激と報酬が無関係に呈示される条件で，瞳孔サイズの動態を調べた。音刺激によって報酬が予測できる条件では音刺激の提示によって瞳孔サイズが拡大し，予測できない条件では瞳孔サイズは変化を示さなかった。さらに，同課題を遂行中のマウスの行動を強化学習によってモデル化することで，動物内部の報酬予測の動態を推定し，モデルの報酬予測の動態と瞳孔サイズの動態が一致したことから，瞳孔の瞬間的な変化が学習課題中の報酬に対する予測を反映することを明らかにした。

研究2では，「好奇心」を数理的に表現し，モデルの変数として組み込むことで，動物において観察される，報酬が得られなくなることによって生じる一過性の反応性の増加，すなわち消去バーストを説明することを試みた。「好奇心」を餌などの外的報酬とは異なる形で表現し，モデルに実装し，数理シミュレーションを行うことで，消去バーストが生じる環境を同定した。その同定された環境の下で，マウスによる行動実験を行い，行動を解析することで，消去バーストが生じる環境を明らかにした。さらに，同課題中の瞳孔を計測することによって，報酬が急に与えられなくなる，といった環境の変化によって瞳孔の大きさが反映することを明らかにした。この「好奇心」による消去バーストの説明は，自身の行動によって環境の状態に関する情報を得るものであり，環境からの入力に対する動物の受動的な一面だけでなく，環境に対して能動的に働きかける側面を描出した。

研究3と4においては，実験中の動物が従事しうる多様な行動，すなわち他行動に着目して，強化学習モデルを構築することで，バウト－休止パターンや習慣形成といった，学習性の行動の生起則を，行動間の状態遷移とそのマクロな構造によって説明することを試みた。これらのシミュレーションでは，動物の行動を他行動と学習性の行動が織りなすネットワークとして扱うことで，動物の行動のマクロな構造の変化として捉え直し，それによって様々な行動現象が説明されうる可能性を示した。

研究1と2によって，マウスを対象とした行動実験に，深層学習を用いた画像解析による瞳孔計測を取り入れ，強化学習による行動のモデル化によって，瞳孔の大きさが報酬に対する予測や，予測誤差を反映しうることを明らかにした。これは，学習という現象が，単に環境からの入力に応じて生じる受動的なプロセスではなく，報酬の予測性や環境のダイナミクスに応じて，感覚入力や行動の変調を行うことで，能動的に環境へと関わるプロセスである可能性を示唆した。研究3と4では，バウト－休止パターンと習慣形成という現象を題材に，現実の個体が従事しうる実験とは無関係な行動が，学習によって獲得される行動に及ぼす影響を明らかにした。研究3と4は数理シミュレーションに限られるものの，現実の生物個体が従事する多様な行動に着目した学習理論の新たな発展可能性を示すことに成功した。

本博士論文は，数理工学的な手法を取り入れることで，学習に伴う瞬間的な行動・生理的な変化を捉え，個体に学習を可能とする生物学的メカニズムの一端を明らかにし，学習によって獲得された行動が示す様々な特徴が従来とは異なる理論的な枠組みで説明しうることを示した。さらに，機械学習や深層学習をはじめとする数理工学的な手法が，心理学における学習と行動研究において果たす役割を示した。本博士論文は，今後，心理学が神経科学や工学などの異分野と共同して，学習という現象の心理学的・生物学的基盤や，その背後にある計算アルゴリズムを明らかにすることに貢献するものである。

３． 審査要旨

2023年2月10日（金）に三田キャンパスにて公開博士論文審査会が開催された。審査会には，神経行動学，認知神経科学，行動神経科学，システム神経科学を専門とする審査員４名と学生および研究員など計10名が出席し，山田君による博士論文の解説がなされた後，質疑応答が行われた。本博士論文の評価および質疑応答の概要は以下の通りである。

本論文の高く評価できる点は，心理学，生理学，神経科学，機械工学などで用いられてきた様々な技術を援用することで，学習心理学と行動分析学という心理学の分野を，学習科学という学際分野へと発展させる方法論を示した点である。従来の学習心理学および行動分析

学における実験および理論の発展の背景には，実験制御・計測機器の画一化とその普及があった。実験装置の普及により知見の蓄積が進んだ反面，画一化された実験装置が計測対象としていない種類の行動，装置が開発された当時では制御および計測が不可能であった時間および空間スケールの行動は，分析の対象から漏れ落ちてしまっていた。その結果，学習心理学および行動分析学において，新たな現象の発見，異分野の技術の導入，それに伴う理論の発展が停滞していたことは否めない。山田君は，そのような学習心理学および行動分析学の歴史的経緯と，それらの学問を取り巻く現状を鑑みた上で，これまで計測対象外とされてきた種類の行動および高精度の生理的反応の計測技術を導入し，ソフトウェアとハードウェアの両側面において独自の実験系を構築することで，本博士論文の研究に取り組んできた。山田君は，本論文の研究1，2において，自らが開発した独自の行動実験系に薬理学的操作を併用することで，学習課題における行動とそれを担う生物学的な構造の一端を明らかにした。研究3，4では，従来の研究では計測対象外であった様々な行動を変数として新たに仮定することにより，機械学習および強化学習を用いたモデルによるシミュレーションを行うことを通して，従来の学習心理学および行動分析学で扱われてきた行動の構造を，複数の行動の状態遷移やネットワーク構造の結合が変容する過程として説明することに成功している。

研究1の実験では，深層学習を用いた解析手法を導入することにより，頭部固定マウスにおけるパヴロフ型条件づけ課題遂行中のマウスの瞳孔サイズの定量化を行っている。瞳孔サイズの測定は，自律神経系活動の変化を非侵襲的な計測によって推定しうる技術であり，行動と生理学的反応の関係を調べる上で有用な手法である。これまでの学習心理学あるいは行動分析学における実験では，動物が自由に行動する場面で実施される研究が多く，そうした実験事態においては，瞳孔サイズを正確に計測することは困難であった。山田君は，神経科学の研究分野で用いられ始めていた頭部固定による行動実験系に着目し，ハードウェアとソフトウェアの両面から独自の実験系を構築することを通して，行動課題を遂行中のマウスの瞳孔サイズの計測に成功した。同君が確立したこの実験系は，開眼度合いや眼球運動、表情の検出などの他の生理指標の計測にも応用することが可能である。その成果は，学習心理学や行動分析学といった心理学分野だけでなく，精神疾患を予測するバイオマーカーの探索などの基礎医学分野における動物モデル研究への応用も容易であり，今後の発展が期待される点でも評価できる。

学習心理学に先端的な計測・制御技術を導入することで，学習理論の実験検証を試みる山田君の取り組みは，独自に確立した瞳孔計測技術を薬理学的操作と融合させた研究1の実験として結実している。「学習は予測と実際の結果の誤差によって生じる」という学習の基本

原理に関する理論は，心理学におけるレスコルラ=ワグナー・モデル，あるいは機械工学における強化学習の根幹をなす考え方であり，現代における学習の基本理論である。この理論を実験によって検証するためには，報酬やその価値の予測を含むような実験課題を動物に行い，予測と結果の差に基づいて動物の行動変化が生じるかを確認する必要がある。従来の行動神経科学における行動実験では，報酬やその価値の予測と相関する生理反応が報告されていたものの，その生理反応が報酬やその価値の情報を符号化しているのか，あるいは報酬情報に基づく運動表出なのかを分離することが困難であった。山田君はドーパミンD2受容体の拮抗薬を用いた薬理操作実験によってマウスの運動を抑制し，報酬の予測についての情報とそれに伴う運動表出に関する情報を分離することに成功した。一方で，審査員からは，実験に使用した薬理学的な影響が運動抑制だけに限らない可能性が指摘された。山田君からは，そうした薬物の制約も理解した上で，運動を確実に抑制する目的で同実験を行ったという趣旨の解説がなされた。さらに同君からは，実験結果の信頼性と頑健性を確保する上では，より選択性の高い阻害薬を用いた検証が必要であるが，それによって本論文の結果が覆る可能性は低いことが丁寧に説明された。他の質疑では，今回の結果を踏まえ「予測」の定義を再検討することで，学習に関する実験と理論を発展させる機会なのではないかという助言的指摘もなされ，同君も同意していた。こうした議論は，同君の研究によって炙り出された新たな問題であり，今後の発展を期待させる点として評価できる。

研究2では，頭部固定マウスにおけるオペラント条件づけの実験系によって，行動分析学の分野で消去バーストと呼ばれて研究されてきた現象の再現に取り組んだ。消去バーストとは，反応によって予測された結果がもたらされなかったときに，その直後の反応が一時的に増加する現象である。反応が消去されているにもかかわらず，反応数が増加するこの現象は，通常の学習理論では説明が難しい。学習心理学および行動分析学の教科書においても紹介される知見でありながら，実験室場面での再現例は少ない。山田君は，独自に確立した頭部固定マウスのオペラント条件づけの実験系を用いて，予測される報酬が与えられなくなった直後にマウスのレバー押し反応が増加することを実験的に検証し，実験下における消去バーストの再現に成功した。その上で，数理モデルを用いて消去バーストを生成する要因を検討することを通して，好奇心駆動型の強化学習を実装したモデルによって行動の予測が可能であることを立証した。さらに，研究1で確立した計測系を用い，消去バーストが生じるタイミングで瞳孔の拡大が生じていることも明らかにしている。審査員からは，研究1と同様に，瞳孔の拡大は運動の影響もあるため，結果の解釈には慎重さが必要であるとの指摘もあった。また，研究1と研究2で見られた瞳孔拡大のメカニズムは同じなのかという問いかけがなさ

れ，山田君からは，今後さらなる考察を続けることで，異なる行動課題で動物が表出する行動の背後にある一般法則を見出したいとの回答がなされた。研究２の全体を通し，数理モデルによるシミュレーション，動物実験による検証，生理指標の計測を融合した研究は，新しい学習心理学と行動分析学の方法論を提示している点で評価できる。

研究3の数理モデルによるシミュレーション実験は，行動分析学的な研究において，実験中の動物が従事している様々な行動，特に「反応」として定義され定量化されるレバー押しやキーつつきといった行動だけでなく，それら以外の「他行動」に着目し，強化学習による数理モデルを構築することで，バウト－休止パターンといった学習性の行動の持つ特徴を，他行動と学習性の行動間の状態遷移やネットワーク結合の変化によって説明しようとする試みである。これらのシミュレーションでは，動物の行動を他行動と学習性の行動が織りなすネットワークとして捉えることで，動物の行動のマクロな側面を捉えて，それによって様々な行動現象が説明されうる可能性を示した。山田君が研究3の遂行において確立したバウト－休止パターンの解析を含めた方法論は，事象が生じた後の行動を平均化することに依拠した解析を行なってきた近年の神経科学における単一ニューロン活動記録の解析などにおいても非常に注目されており，実験心理学において広く用いられている反応数や反応潜時の解析では見出すことができない，行動間あるいは行動と生理反応の間にある隠れた関係を調べる上で優れた解析方法を例示する研究として，高く評価できる。

研究4の数理モデルによるシミュレーション実験においては，ネットワーク解析を学習心理学の理論検証に導入することで，行動の習慣化に関する新しい理論を提唱している。習慣化は，結果の予測に基づいた目標志向的行動が，同じ場面で繰り返されることで，結果を予測する手がかり刺激に対して自動的に表出されるようになる現象とされてきた。習慣化は，行動分析学で主に研究されてきたオペラント条件づけにおいて観察される現象でありながら，その背後には，学習心理学で主に研究されてきたパヴロフ型条件づけの機構が存在している点が興味深い。行動の説明原理に関して哲学を異にする両分野は融合が進まず，それゆえ習慣化は，画一的な説明が繰り返されてきた。こうした背景に対し，研究４は，強化学習やネットワークモデルなどの工学分野における学習理論を援用することによって，習慣化を説明しようとする大胆な試みである。研究4では，強化学習を用いた数理シミュレーションにより，エージェントの行動をネットワークとして捉え，習慣形成をネットワークの構造の変化として説明し，実験場面において観察されてきた習慣形成の数々の結果の再現に見事に成功している。既存の理論では，「目標志向的行動」が「習慣」に推移していく連続的過程をどのように扱うのかが明確ではなく，本研究4はそれを数理モデル的に明確したものとして，

新しい学習理論の構築に資するものであり高く評価できる。一方で，研究4における数理シミュレーションでは，エージェントが最短経路の探索を行うという前提が存在する。この前提の妥当性については審査員から指摘があったが，山田君自身もこの問題点は認識していること，ただし数理シミュレーション研究において条件設定は不可避であること，論文には掲載していない複数条件でのシミュレーションを試みた上で，同前提を採用したことが回答としてなされた。この点は，論文本体において十分な考察がなされており，かつ，この問題自体が研究2における好奇心の研究展開にも繋がっており，研究の性質を鑑みると恣意的というものではなく，むしろ適切な姿勢として評価できる。

審査委員および博士論文審査会の参加者から指摘された点は，いずれも今後の研究において議論を深め，発展させることが可能な有益なものであった。本論文を構成する個々の研究の水準は非常に高く，その成果は既に3本の筆頭および責任著者として受理されている国際学術論文に結実している。以上の点を鑑み，審査員一同は本論文が博士（心理学）の学位を授与するに値するものと判断する。