

制止学習アプローチに対する古典的条件づけ研究からの 批判的検討

二瓶正登¹・松本昇²・佐藤友哉³

¹鹿児島純心大学

²信州大学

³新潟大学

Rethinking the inhibitory learning approach: The roles of extinction and associative learning theory

Masato NIHEI¹, Noboru MATSUMOTO² and Tomoya SATO³

¹Kagoshima Immaculate Heart University

²Shinshu University

³Niigata University

The inhibitory learning approach is recognized as the underlying theory of exposure therapy, based on the findings of fear conditioning and extinction. This paper raises questions regarding the validity of two claims in the inhibitory learning approach: (1) exposure therapy is a clinical analog of the extinction procedure in fear conditioning, and (2) the inhibitory learning approach is based on associative learning theories. Little evidence exists to argue that the mechanism of exposure therapy is underpinned by extinction or habituation. Inhibitory learning, as assumed by the inhibitory learning approach, is a unique construct that differs from previous associative models. To bridge the gap between associative theories and the inhibitory learning approach, we introduce a new model and discuss directions for future research.

Key words: exposure therapy, fear conditioning, inhibitory learning approach, associative learning model, extinction

キーワード：エクスポージャー療法、恐怖条件づけ、制止学習アプローチ、連合学習モデル、消去

1. はじめに

不安や恐怖に関連した問題に関する介入として、認知行動療法は非常に有効であることが多くの研究で示されてきた (Norton & Price, 2007)。その中でも特に恐怖や不安を誘発する刺激に対して回避行動や安全確保行動を行わずに曝露を行うことで、その刺激に対する不安や恐怖の減少をもたらすエクスポージャー療法の有効性は広く報告されている。メタ分析において心的外傷後ストレス障害 (Powers et al., 2010)、限局性恐怖症 (Wolitzky-Taylor et al., 2008)、パニック症 (Pompoli et al., 2018)、社交不安症 (Mayo-Wilson et al., 2014) など恐怖や不安を主症状とする精神

障害における症状の改善及び精神的健康の向上に寄与することが多く示されている (Abramowitz, Deacon, & Whiteside, 2019)。これらの知見はエクスポージャー療法が不安や恐怖の減少のみならず、より幅広いアウトカム指標に効果を持つ有用な介入法であることを示している。

エクスポージャー療法の作用機序は多様な観点から説明が行われてきたが、近年では恐怖条件づけにおける消去手続きの効果とみなすことが多い (Craske et al., 2014)。恐怖条件づけとは電気ショックなどの生物学的に重要な害をもたらす、無条件反応 (unconditioned response; UR) を生じさせる嫌悪的な無条件刺激 (unconditioned stimulus; US) と中性的な条件刺激 (conditioned

stimulus; CS) が対提示された時、以前は生じなかった恐怖に関する条件反応 (conditioned response; CR) が CS に対して生じるようになる現象を指し、古典的条件づけ (Pavlov, 1927) の一種である。消去手続きとは条件づけ後に US を伴わず CS のみを反復提示する手続きを指し、一般に CR の減少が確認される。

近年、恐怖条件づけの実験的知見および理論体系に基づいたエクスポージャー療法を理解する枠組みとして、制止学習アプローチ (inhibitory learning approach) が提唱された (Craske, Hermans, & Vervliet, 2018; Craske et al., 2008; Craske et al., 2012; Craske et al., 2014; Craske et al., 2022; 佐藤・田中・前田, 2023)。制止学習アプローチはエクスポージャー療法を消去手続きとみなし諸技法を提案するだけでなく、これらの理論的基盤として学習心理学領域における代表的な説明体系である連合学習諸理論を説明に用いている点が特徴である。

制止学習アプローチの有用性はこれまで多くの知見から支持されてきた。佐藤ら (2023) は制止学習アプローチに基づく介入に関するシステマティックレビューを行い、適格基準に該当した 59 の論文のうち 33 の論文で何らかの介入効果が認められたことを確認している。しかしながら、制止学習アプローチが依拠する理論体系においては基礎心理学的知見や連合学習理論とは異なる説明が導入されている点も散見される。本論文では制止学習アプローチが条件づけに関する基礎的な観点からどのように理解可能であるかについて批判的検討を行い、制止学習アプローチが持つ特徴と課題点を明確化することを目指す。なお、制止学習アプローチは後述の理由で制止検索アプローチ (inhibitory retrieval approach) という名称に変更された (Craske et al., 2022) が、制止学習アプローチという表現は多くの研究でこれまで用いられてきたことも踏まえ、本論文では制止学習アプローチという表現で統一する。

2. 制止学習アプローチ

制止学習アプローチとはエクスポージャー療法の作用機序に対する説明体系であるが、それに際し大きく 2 つの主張を行っている。第一に、エク

スポージャー療法の作用機序を恐怖条件づけにおける消去に求める点である。従来、エクスポージャー療法は拮抗条件づけ (Wolpe, 1958) や馴化・恐怖構造の修正 (Foa & Kozak, 1986) といった多様な観点から説明がなされてきた。制止学習アプローチはこうした説明の妥当性を批判した上で、消去による学習を最適化することがエクスポージャー療法において重要な作業にするべきであると提案している。

第二に、エクスポージャー療法の実施による恐怖反応の減少を、連合学習理論に基づき制止学習によるものと説明している点である。連合学習理論とは条件づけや消去手続きの効果を経験に伴って形成された刺激や反応に関する表象間の連合に求める説明体系であり、古典的条件づけに関する諸現象の説明として古くから用いられてきた。連合学習理論における最も著名なモデルは Rescorla-Wagner モデル (RW モデル: Rescorla & Wagner, 1972) であり、制止学習アプローチが中核に据えるモデルの 1 つである。RW モデルとはある条件づけ試行に伴う連合強度の変化量を US の強度と刺激の明瞭度から説明したものであり、以下の式により表現される。

$$\Delta V_A = \alpha_A \beta (\lambda - \Sigma V) \quad (1)$$

この式において V_A は刺激 A が持つ連合強度を指し、 ΔV_A は試行に伴う刺激 A に対する連合強度の変化量である。 ΣV は刺激 A を含むその試行で提示された刺激全ての連合強度の合計を指す。 α_A と β はそれぞれ刺激 A と US の明瞭度であり学習率を規定するパラメータである。連合強度の漸近値を規定する λ は US 強度として解釈される。このモデルは現在の試行で有している連合強度 (V_A) とその試行で生じた結果 (λ) 間の差分が学習量 (ΔV_A) を規定することから、予測誤差モデル、あるいは誤差修正モデルと表現される。RW モデルは古典的条件づけにおける多様な現象の説明を可能にただけでなく、新たな現象の予測も導くことに成功した (Miller, Barnet, & Grahame, 1995)。

制止学習アプローチが中核としているもう 1 つのモデルとして Bouton のモデル (Bouton, 1993) がある。このモデルに基づけば、古典的条件づけにおける獲得と消去では、CS に対して異なる連

合が独立に形成される。獲得時はCRを生じさせる興奮性連合（CS-US連合）が形成される一方で、消去手続き時にはCRを抑制する制止性連合（CS- no US連合）が形成される。すなわち獲得時には興奮性連合が形成されるためCRが出現し、消去時にはそれを抑制する制止性連合が形成されるためCRが減弱する。Boutonのモデルの画期的な点は、この仮定に文脈刺激の役割を考慮した点である。ここでいう文脈刺激とは物理的環境、内的環境、時間間隔も含めたCSとUS以外の環境刺激の総体を指す（Bouton, 2002）。興奮性連合はCSが提示される文脈の影響を受けにくい一方で、制止性連合は強く影響を受ける文脈依存的な連合であるとされる。すなわち消去手続き時の文脈以外でCSが提示された場合、制止性連合は検索されにくくなり、結果としてCRが出現しやすくなる。

制止学習アプローチではこの2つのモデルを活用したエクスポージャー療法の理解と技法の提案を行っている（Craske et al., 2014）。その中核技法となるのが予期妨害（expectancy violation）である。これはクライアントが嫌悪的事象（US）を到来すると強く信じている時に刺激への曝露を行うことで、予期していた結果が実際は到来しないという経験の効果を強めることを目指す手法であり、RWモデルにおける予測誤差のアイデアに基づく。RWモデルでは消去時において生体が持つ連合強度（CSがUSを予測する程度）が強い場合、大きな反応減少が生じることを予測する。制止学習アプローチではこの知見を援用し、可能な限りクライアントが持つ予期を大きくさせた上で刺激への曝露を行うことが制止学習の促進に重要であるとしている。

加えて、制止学習アプローチは、再発の機序および再発防止を目指す方略の説明をBoutonのモデルから援用している。すなわちエクスポージャー療法に伴って形成される制止学習は文脈依存的であるため、介入後に文脈変化が生じるだけで恐怖反応の再発が生じると説明する。基礎研究において消去後に特定の手続きを経験することでCRが再出現する現象は消去からの回復効果（recovery-from-extinction；McConnell & Miller, 2014）として知られており、連合学習の観点からは主にBoutonのモデルから説明されてきた

（Bouton, 2002）。特に消去後の文脈変化（復元効果：Bouton & Bolles, 1979）、USの単独提示（復帰：Rescorla & Heth, 1975）、および時間経過（自発的回復：Pavlov, 1927）に伴うCRの再出現、そして消去後に再度条件づけを行うと消去以前よりも速くCRが獲得される急速な再獲得（Bouton & Swartzentruber, 1989）の4種の現象を消去からの回復効果として指すことが多い（Bouton, 2002）。制止学習アプローチにおいて、この4種はエクスポージャー療法による介入後の恐怖反応の再発の原因であるとみなされる。すなわちエクスポージャー療法を実施した場所からの移動（復元）、日常場面での嫌悪的事象への遭遇（復帰）、介入からの単純な時間経過（自発的回復）、そしてトラウマ事象の再体験（急速な再獲得）によって介入により消失したはずの恐怖反応は容易く回復する。

消去からの回復効果は特定の手続きによって減少することが知られている。例えば複数の物理的文脈で消去手続きを行う（Gunther, Denniston, & Miller, 1998）、あるいは消去手続き時に数回USを提示する（Thomas & Ayres, 2004）といった手続きを導入すると消去からの回復効果は減少し、これらの効果もBoutonのモデルに基づき、主に制止性連合の検索促進の観点から説明がされることが多かった（Nihei et al., 2023）。制止学習アプローチもこの観点を導入し、多くの実験知見に基づく介入再発防止方略を提案しその有効性も検証されている（Craske et al., 2014）。

3. 制止学習アプローチは妥当な説明体系か？

本セクションでは、制止学習アプローチが主張する（a）エクスポージャー療法とは恐怖条件づけの消去手続きである、（b）エクスポージャー療法の作用機序は制止学習によるものである、の2点についてその妥当性を古典的条件づけにおける先行研究および理論との整合性の観点から検証を行う。

3.1 エクスポージャー療法は消去手続きか？

制止学習アプローチではエクスポージャー療法が馴化による作用であるという説明を批判し、その代替的な説明として消去手続きによる効果を説

明に用いてきた。馴化とは個体に反復的な刺激を与えることによって生じる反応の減少である (Rankin et al., 2009)。Craske et al. (2012) はエクスポージャー療法の馴化説への批判の1つとして、馴化は一時的な現象であり別の刺激の提示 (脱馴化) によって馴化の効果は容易に取り除かれるため、介入に伴う長期的な効果を説明できないことを挙げている。

しかし刺激を多数回提示する、あるいは馴化と脱馴化、もしくは馴化と自発的回復を繰り返すことで馴化の効果は増強される (Rankin et al., 2009)。この効果は介入時におけるエクスポージャー療法に伴った恐怖刺激への反応減少 (馴化) とセッション間の間隔 (自発的回復) や他の刺激の経験 (脱馴化) に伴う馴化の減弱、および次のセッションでの曝露に伴う反応減少 (馴化) というプロセスと合致している。そのためこの批判は適当でないだろう (エクスポージャー療法における馴化の役割としては Benito & Walther, 2015 を参照)。

馴化説を否定するもう一つの根拠として、馴化を作用機序の中核に据える情動処理理論 (Foa & Kozak, 1986) が重視する恐怖の賦活、セッション内馴化、セッション間馴化という3種の指標と、介入に伴う症状尺度得点や恐怖刺激への回避行動の程度といった治療アウトカムが相関しない点を主張してきた (Craske et al., 2008; Craske et al., 2012)。恐怖の賦活とはあるエクスポージャー療法セッションにおいて賦活された恐怖の最高強度、セッション内馴化とは特定のセッションにおける恐怖反応の減少度合い、セッション間馴化とはあるセッションの最後での恐怖反応と次のセッションの最初での恐怖反応の減少度合いを指す (Craske et al., 2008)。これらの指標は情動処理理論において介入効果をもたらすための中核的な作用機序とみなされてきた (Foa & Kozak, 1987)。

しかしこの批判にも大きな問題がある。前述のとおり、馴化とは刺激の反復提示に伴う刺激への反応減少を指す用語である。エクスポージャー療法の実施に伴う恐怖反応の減少 (セッション内馴化) が多くの研究で一貫して報告されている (Craske et al., 2008) ことを考慮すると、刺激への曝露に伴って馴化が生じているという仮説は支持されていると言える。エクスポージャー療法に伴

う恐怖反応への作用が馴化であるかと、馴化の大きさが治療アウトカムと関連するかはいささか異なる問題である。

この点は制止学習アプローチにも当てはまる。消去手続きも馴化と同様に刺激に対する恐怖の減少をもたらす手続きであり、臨床研究において測定される多様なアウトカム指標全般に影響を及ぼすことは考えにくい。例えば刺激への恐怖反応を消去した場合においても、その刺激に対する回避行動は持続することは広く知られている (Dymond, 2019)。このことは、不安に関連する諸問題は刺激への恐怖反応の減少のみで解決する問題ではないことを示している。すなわち、制止学習アプローチに基づく介入の結果、多様な症状アウトカムが減少したとしても、それが消去手続きの効果のみによると主張することは簡単ではない。この点は介入に伴う恐怖反応の減少が馴化か消去かという議論とは別に行われるべきであろう。

3.2 制止学習アプローチは連合学習理論に基づいているか？

上述の通り、制止学習アプローチでは多様な介入技法が含まれるが、これらの手続きは単一の連合学習理論では説明できない。例えば予期妨害は主に予測誤差モデルの考えに依拠している技法である。連合学習理論における予測誤差モデルとは、主にUSへの予期と実際の結果の差分に基づいて連合強度が変化するモデルの総称を指す。制止学習アプローチでは「クライアントが持つ嫌悪的事象の到来の予期を事前に大きくする」ことを重視するが、曝露に伴う恐怖の表出、あるいは表出された恐怖の減少については重視していない (Craske et al., 2008)。形成された連合強度と表出される恐怖反応が線形に対応しないという仮定は、多くの予測誤差モデルと反する (線形性を仮定しないモデルの例としては Miller & Matzel, 1988)。また、予期妨害の理論的根拠として頻繁に引用される RW モデルは Bouton のモデルと異なり、消去時に制止性の連合が独立して形成されるという仮定は持たない (制止性連合の概念を導入している予測誤差モデルとして Pearce & Hall, 1980)。

一方で文脈に関する諸技法は Bouton のアイデ

アが中心である。しかし、Bouton のモデルは興奮性・制止性連合がどのような機序で形成されるかについて言及しておらず、あくまで形成された連合の検索が文脈により規定されることを述べたモデルである。これは連合の概念では扱いにくい諸手続きに関する文脈の効果も説明するためである (Bouton, 1993)。すなわち Bouton のモデルからは予期妨害の技法は演繹されない。まとめると、制止学習アプローチが仮定する「制止性連合の形成は予測誤差によって生じ、その連合の検索は文脈刺激によって規定される」という連合学習モデルはこれまで存在せず、制止学習アプローチに独自のモデルであることがうかがえる。

制止学習アプローチでは消去は制止学習の結果であるとみなされるが、消去がどのようなメカニズムにより達成されるかは用いる説明体系に依存する。この点について、そのメカニズムが制止学習であるとする根拠として上述した消去からの回復効果の存在を挙げている (e.g., Craske et al., 2014)。制止性連合を仮定せずに消去を興奮性連合の減弱により説明するモデル (RW モデルなど) では、これらの現象の説明が困難であるとされてきた (Bouton, 1993)。しかし、制止性連合を仮定しないモデルでもこれらの現象はいくつかの仮定を追加することで説明可能である (McConnell & Miller, 2014)。また、Bouton のモデルもこれらの全てを扱えるわけではなく、その説明に批判も存在する (e.g., Nihei et al., 2023; Riccio, Richardson, & Ebner, 1999)。例えば復元や復帰は文脈条件づけの影響を受ける (e.g., Miller, Laborda, & Polack, 2020) という複数の知見は、文脈刺激を連合の検索手がかりとしてみなす Bouton モデルの仮定と反する。すなわち消去および関連する諸現象は制止性連合の概念を仮定しないと説明できない現象群ではない。

Craske et al. (2022) は制止学習アプローチの名称を検索の概念を重視する目的で制止検索アプローチに変更した。これにより連合の形成過程よりも文脈要因による連合の検索過程を重視する Bouton のモデルの発想に近づいた印象がある。しかし前述の通り Bouton のモデルはあくまで形成された連合が CR として表出される際の表出規則のモデルであり、予期妨害を始めとして恐怖反応の減少を促進する手続きは演繹できない。検索

という表現は連合形成プロセスと表出プロセスが独立していることを示唆するものであるが、例えば RW モデルでは形成された連合強度と表出される反応は線形に対応することが仮定される。このような齟齬を避けるためには、より具体的な手続きに焦点を当てた「消去学習アプローチ」という名称の方が、制止学習アプローチが依拠する現象群を踏まえると適切であるかもしれない。

制止学習アプローチは学習心理学の知見を援用しその作用機序を消去に求める一方で、記憶研究といった他領域の知見の導入も積極的に行っている。例えば制止学習を促進するために刺激の変動性を増加させることを認知心理学における記憶研究から提案している (Craske et al., 2008)。また、情動制御の研究から恐怖減少よりも恐怖耐性を重視するという方針 (Craske et al., 2008) や、ポジティブ情動の喚起・情動ラベリングのような従来の連合学習理論では扱われなかった概念も積極的に取り入れられている (Craske et al., 2014) ことから、エクスポージャー療法の効果を連合学習理論の仮定のみに基づくアプローチとみなすことは適当ではないだろう。

他の問題も山積している。例えば予期妨害では予測誤差を促進するため、事前に認知的介入を行わないといった点は提案当初から現在に至るまで一貫して重視されてきた (Craske et al., 2008; Craske et al., 2022)。これは思考などの認知的内容が予期に影響を与えるという仮定であるが、連合学習理論は主に動物研究を中心に発展してきた経緯もあり、この点を説明できる理論は多くない (これらを扱うモデルとして De Houwer, 2009; Lovibond & Shanks, 2002)。もちろん、Bouton のモデルも RW モデルもこの点は考慮していない。また、予期妨害では驚きをもたらすことが重要であると強調されることが多い (Craske et al., 2012) が、この驚きの解釈には注意が必要である。RW モデルにおける驚きとは「予期していた US の到来と実際の経験の差」を指すものであり、素朴心理学的な「驚き感」のような心的状態ではない。今後、制止学習アプローチにおける予期、あるいは驚きがどのような構成概念であるか、すなわち制止学習の形成と表出される恐怖反応はどのように関係し、認知的な内容や操作が予期にどのような影響を及ぼすかを明確にしていける必要があるだ

ろう。

また、不安症者、あるいは高不安者は分化条件づけにおいて一貫してUSと対提示されない刺激であるCS-へ高い反応を示すことが知られている (Duits et al., 2015) が、これについても制止学習アプローチでは制止学習の障害と説明がなされてきた (Craske et al., 2012)。しかし古くから恐怖条件づけ研究では、CS-への反応は鋭敏化の指標として用いられてきた (澤, 2011)。鋭敏化とは馴化と反対に、強度の高い刺激の反復提示、あるいは生体が興奮状態にある時に刺激を反復提示すると反応強度が増加する現象であり (Domjan, 2015 漆原・坂野訳 2022)、馴化と同様に非連合的現象の1つである。

もし制止学習アプローチが仮定するようにCS-へ制止性の学習が生じるならば、制止性連合の形成において事前の興奮性連合の形成は不要ということになる。これまで制止性連合を仮定してきた諸モデルにおいて、制止性連合は既に形成された興奮性連合を抑制する機能 (Bouton, 1993)、あるいはCSと予期されたUSの非到来に伴う拮抗的な情動反応 (恐怖条件づけであればリラクゼーションなど) を生じさせる内的刺激間の連合 (Pearce & Hall, 1980) とみなされており、この点についても連合学習モデルの仮定とは大きく乖離することとなる。

まとめると、制止学習アプローチで述べられている「制止学習」は条件づけ研究における制止性連合の概念とは乖離しているのみならず、非常に多様な現象をまとめた独自の概念と考える方が適切である。一方で、これまで基礎研究ではあまり取り組まれてこなかった条件づけ研究と記憶研究、感情制御領域の研究知見などの統一を目指す挑戦的な試みである。多くの新たなエクスポージャー療法の技法を生み出しその有効性も報告されていることから、それらの論文が特定の研究グループに偏っているなどの批判はあるものの (佐藤ら, 2023)、臨床的な有用性の観点からは優れた理論といえる。

しかしこの独自の説明体系が理論として科学的に妥当であるかは議論と検証が必要である。特に制止学習アプローチが仮定する「制止学習」がどのような構成概念であるのかを明確にし、関連する諸技法を包括的にまとめるための理論研究は必

須である。このような曖昧な概念を用いた理論化は心理学領域に広く見られるが、これが心理学における再現性問題の要因であるとの指摘もある (竹澤, 2018)。厳密な理論的裏付けが不足している現状では、制止学習アプローチをエクスポージャー療法の効果を説明する理論体系ではなく、基礎心理学における複数の理論やモデルから演繹された技法のパッケージとみなす方が適切である。

4. 解決に向けて

本セクションでは取り上げた2つの問題のうち後者の解決、すなわち制止学習アプローチの仮定に合致する連合理論の構築を目指した筆者らが行った研究を取り上げる。初めに Nihei, Hojo, and Sawa (2021) は定性的モデルである Bouton のモデルにおける消去の仮定に予測誤差モデルを組み込み、定量的なモデルとして表現する試みを行った。このモデルは以下の式で表現される。

$$CR = Ve + Vi \times Relapse \quad (2)$$

$$\Delta Ve = \alpha_{Ve} (\lambda - Ve) \quad (3)$$

$$\Delta Vi = \alpha_{Vi} (0 - Vi + Ve) \quad (4)$$

ここではCS-US対提示が行われた時 (式3) にCRを生成する興奮性の連合 (Ve) が、CSの単独提示が行われた時 (式4) にCRを抑制する制止性連合 (Vi) が、予測誤差モデルに従ってそれぞれ独立に形成される。CRの大きさは両連合の加算によって決定されるが (式2)、その際に制止性連合は文脈の影響を受け検索の度合いが減弱する (*Relapse*)。すなわち興奮性連合はどの文脈でテストされた場合でも検索の減衰は生じないが、制止性連合は *Relapse* (0~1の間を取る) に従って減衰が生じる。これにより制止学習アプローチが仮定する「制止性連合の形成は予測誤差によって生じ、その連合の検索は文脈刺激によって規定される」という連合学習モデルの定式化に成功した。

Nihei et al. (2021) は恐怖条件づけ、消去およびABA復元 (獲得と消去を別の文脈で行った後、獲得文脈でテストをするとCRが再出現する現象) を検証した実験データにこのモデルを適用した (図1)。図1は63名の参加者から獲得、消去、

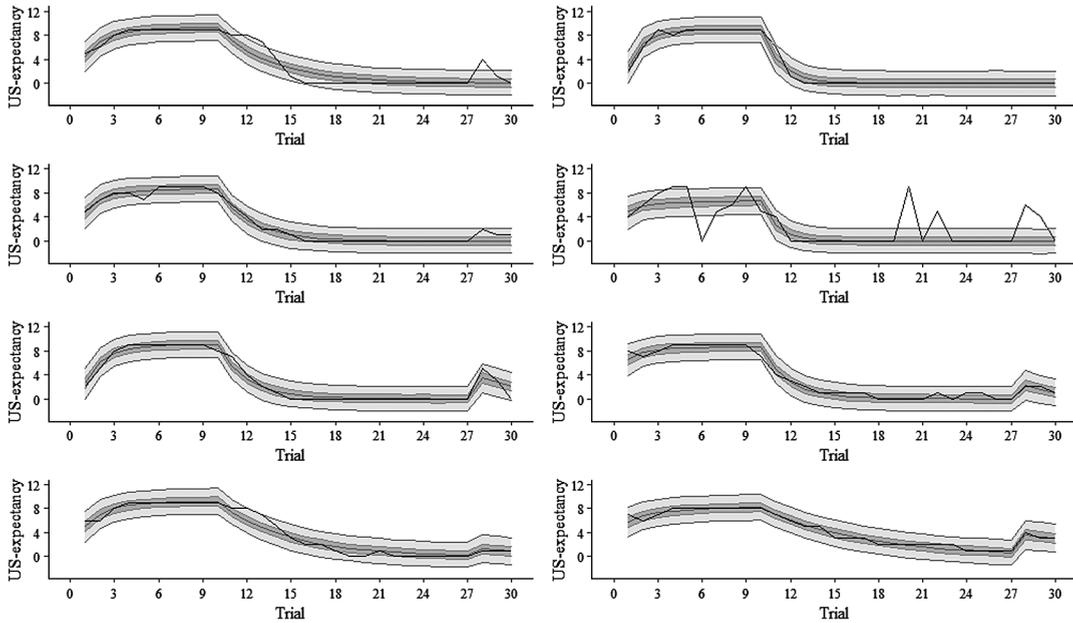


図1 Nihei et al. (2021) における8名の条件反応データとモデルにより推定された事後予測分布
 Note：実線はデータを表し，濃い灰色は50%，薄い灰色は95%信用区間を表す

およびテスト期を全て同じ文脈で行った参加者4名（上段）と消去期のみ異なる文脈で行った参加者4名（下段）のデータ，そしてモデルから推定された事後予測分布を表している。数名の参加者のデータ（右列の上から2番目など）はモデルの区間と乖離が見られたものの，多くの参加者においては概ねモデルの数理的表現と一致するデータが得られたことが確認された。

しかしこのモデルにはいくつかの欠点が存在した。第一に，*Relapse*，すなわち制止性連合の検索を減弱するパラメータの厳密な解釈を行っていない点である。これは消去の文脈依存性がどのような要因によって生じているか，あるいはどのように操作することが出来るかをモデル化出来ていないことを示す。第二に，数理モデル化したことにより Bouton のモデルで扱えていたいくつかの現象が扱えなくなった点である。例えばこのモデルでは急速な再獲得現象を表現することが出来ない。

第三に，Bouton のモデルの欠点を解決できていない点である。例えば Bouton のモデルにおいて復帰は US 単独提示に伴って形成される文脈刺激への連合強度そのものが興奮性連合の検索手が

かりとなることを仮定する (Bouton & Nelson, 1998) が，もし文脈刺激が連合強度を持つ場合，その検索を規定する“文脈刺激の文脈刺激”が必要になるはずである。

そこで筆者らはこれらの問題を解決し，かつ消去の促進と消去からの回復効果に関する多くの知見を包括的に説明可能な文脈類似性モデルを提唱した (Nihei et al., 2023)。このモデルは以下の式によって表現される。

$$CR_A = Ve_A * S_A^e + Vi_A * S_A^i \quad (5)$$

$$\Delta Ve_A = \alpha_A^e (\lambda - \Sigma (Ve * S^e) + \Sigma (Vi * S^i)) \quad (6)$$

$$\Delta Vi_A = \alpha_A^i (0 - \Sigma (Ve * S^e) + \Sigma (Vi * S^i)) \quad (7)$$

このモデルの大部分は Nihei et al. (2021) のモデルを引き継いでおり，式6が興奮性連合，式7が制止性連合の変化を表現し，最終的な CR の表出は式5によって規定される。しかし重要な点がいくつか変化している。第一に，検索の減衰を表現するパラメータを，*Relapse*ではなく文脈間の類似性 (*S*) という心理学的概念として定義づけた点である。文脈間の類似性とは現在の文脈と興奮性連合であれば獲得時の文脈，制止性連合であれば消去時の文脈がどの程度類似しているかを表す

構成概念（それぞれ S^e と S^d ）であり，類似性の程度によってその文脈で両連合が検索される度合いが変化する。例えば消去時の文脈とテスト文脈を生体が類似していると認識するほどテスト時の制止性連合の検索は促進され，結果として反応の減弱が維持される。なおこのモデルでは興奮性連合も文脈間の類似性の影響を受けるため，テスト文脈が獲得文脈と類似するほど，反応は増加する。

第二に，式7で表される制止性連合の漸近値が I_e ，すなわち獲得時に形成された興奮性連合の強度ではなく，その文脈において「検索された」興奮性連合となる点である（ $I_e * S^e$ ）。前述の通り興奮性連合は文脈間の類似性の程度により検索強度が減衰されるため，獲得文脈と異なる文脈で消去手続きを行うほど，消去時に検索される興奮性連合の強度は減衰し，結果として形成される制止性連合は弱くなる。この仮定は制止性連合をCSと「予期されたUSの非到来に伴う拮抗的な反応に先行する内的刺激」との連合とみなす Pearce-Hall モデル（Pearce & Hall, 1980）に基づいている。この考えに基づけば，消去手続きによって生じる

拮抗的な反応の強度は，獲得時に経験した興奮性連合そのままではなく，その文脈で興奮性連合が検索された程度に依存するはずである。

文脈類似性モデルでは Bouton のモデルで扱えなかった諸現象も含め，多くの現象を説明することが可能となった。Bouton のモデルでは獲得と消去の文脈が同じ場合（AA）と違う場合（AB），その後に中性文脈へ移動した際に生じる反応の再出現，すなわち復元効果の大きさに差がないことを予測する（ $AAB=ABC$ ）。しかし先行研究において，獲得文脈と消去文脈が同じ場合，違う文脈であるよりも復元効果が小さいことが示されている（ $AAB < ABC$ ；Thomas et al., 2003）。文脈類似性モデルでは上述の通り消去時の文脈が獲得時の文脈と異なるほど形成される制止性連合は弱まることを仮定するため，この現象の発生を予測することが出来る（図2）。ただし文脈類似性モデルを個々のデータに合致させる取り組みは不十分であり，今後その妥当性を確認していく必要がある。

文脈類似性モデルにより予測誤差モデルと Bouton のモデルを矛盾なく融合し，かつエクス

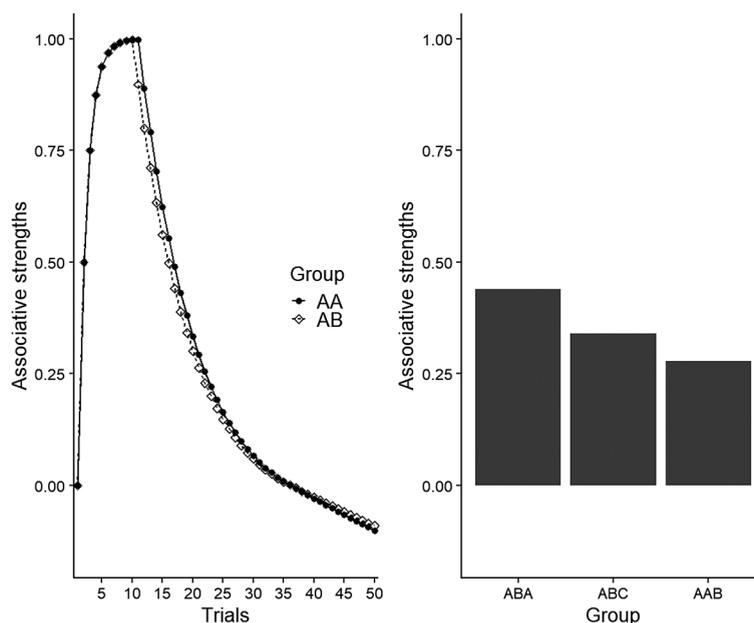


図2 文脈類似性モデルに基づく獲得と消去（左図）および復元効果（右図）のシミュレーション（Nihei et al., 2023の図を一部改変）

ポージャー療法の効果を促進させるための諸技法を包括的に扱えることが可能となった。すなわち制止学習アプローチの基盤理論になり得る可能性を有していると考えられる。ただし本モデルは実験研究の知見に基づいたものであり、モデルの予測とエクスポージャー療法に伴う恐怖反応の減少が一致するかは未知数である。また、従来の条件づけ研究では説明が難しい技法は本モデルでも扱うことは出来ないため、制止学習アプローチの全てを包括的に扱える理論の構築は今後も行われるべきであろう。

5. 結 論

本論文では制止学習アプローチの妥当性について古典的条件づけおよび連合学習諸理論の観点から検討する試みを行った。この結果、制止学習アプローチは臨床的に有用な枠組みである一方で、中核である制止学習の概念は基礎心理学的には根拠が不十分、あるいは曖昧な点が散見された。このことから、制止学習アプローチが仮定する説明体系について改善の必要があるといえる。

はじめにエクスポージャー療法の作用機序が消去か馴化かという点について議論を行った。制止学習アプローチが提唱されて以降、エクスポージャー療法の作用機序は消去手続きであることが強調されるようになった。しかし、消去とする証拠および馴化説を批判する根拠には不十分な点も存在するため、今後どちらがより妥当な説明であるかの議論と検証は行われるべきであろう。ただしエクスポージャー療法が消去手続きであると仮定し、そこから新しい技法や介入方法を提案することは科学的な取り組みとして重要である。制止学習アプローチが提唱されて以降、多くの消去研究との橋渡し研究が行われ、実際に介入効果をもたらしてきた。すなわち消去であると断言はできないものの、そのような仮定の下で理論化を行い、その理論から予測される新たな技法の効果を検証することは非常に意義があると思われる。

次に、制止学習アプローチと連合学習理論の乖離について議論を行い、制止学習アプローチにおける「制止学習」は連合学習理論における概念とは異なる独自の構成概念であることを主張した。この点については文脈類似性モデルを構築するこ

とで問題の一部は解決されたが、制止学習と関連づけられてきた認知心理学や感情研究など条件づけ研究以外の領域をどのように扱うかは議論を必要とする。すなわち制止学習アプローチが抱える科学的な説明体系としての問題は依然として山積している。

従来、エクスポージャー療法に限らず、新しく提唱された治療に関する説明体系や治療的アプローチの臨床的有用性を効果研究などから評価する研究が重点的に行われる一方で、説明体系そのものの妥当性の検証は不十分であったように思われる (Hardy & Llewelyn, 2015)。科学的な臨床心理学およびその実践を目指すにあたっては、効果研究のみならず臨床心理学における理論研究の発展が必要である。

文 献

- Abramowitz, J. S., Deacon, B. J., & Whiteside, S. P. (2019). *Exposure therapy for anxiety: principles and practice* (2nd ed.). The Guilford Press.
- Benito, K. G., & Walther, M. (2015). Therapeutic process during exposure: Habituation model. *Journal of Obsessive-Compulsive and Related Disorders*, 6, 147–157.
- Bouton, M. E. (1993). Context, time, and memory retrieval in the interference paradigms of Pavlovian learning. *Psychological Bulletin*, 114, 80–99.
- Bouton, M. E. (2002). Context, ambiguity, and unlearning: sources of relapse after behavioral extinction. *Biological Psychiatry*, 52, 976–986.
- Bouton, M. E., & Bolles, R. C. (1979). Role of conditioned contextual stimuli in reinstatement of extinguished fear. *Journal of Experimental Psychology: Animal Behavior Processes*, 5, 368–378.
- Bouton, M. E., & Nelson, J. B. (1998). The role of context in classical conditioning: Some implications for cognitive behavior therapy. In O'Donohue W, (eds) *Learning and Behavior Therapy*. Needham Heights, MA: Allyn and Bacon, pp. 59–84.
- Bouton, M. E., & Swartztruber, D. (1989). Slow reacquisition following extinction: Context, encoding, and retrieval mechanisms. *Journal of Experimental Psychology: Animal Behavior Processes*, 15, 43–53.
- Craske, M. G., Hermans, D., & Vervliet, B. (2018). State-of-the-art and future directions for extinction as a translational model for fear and anxiety. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 373, 20170025.
- Craske, M. G., Kircanski, K., Zelikowsky, M., Mystkowski,

- J., Chowdhury, N., & Baker, A. (2008). Optimizing inhibitory learning during exposure therapy. *Behaviour Research and Therapy*, 46, 5–27.
- Craske, M. G., Liao, B., Brown, L., & Vervliet, B. (2012). Role of Inhibition in Exposure Therapy. *Journal of Experimental Psychopathology*, 3, 322–345.
- Craske, M. G., Treanor, M., Conway, C. C., Zbozinek, T., & Vervliet, B. (2014). Maximizing exposure therapy: An inhibitory learning approach. *Behaviour Research and Therapy*, 58, 10–23.
- Craske, M. G., Treanor, M., Zbozinek, T. D., & Vervliet, B. (2022). Optimizing exposure therapy with an inhibitory retrieval approach and the OptEx Nexus. *Behaviour Research and Therapy*, 152, 104069.
- De Houwer, J. (2009). The propositional approach to associative learning as an alternative for association formation models. *Learning & Behavior*, 37, 1–20.
- Domjan, M. (2015) The principles of learning and behavior, 7th edition. Cengage Learning. 漆原宏次・坂野雄二 (監訳) (2022) ドムヤンの学習と行動の原理 原著第7版 北大路書房.
- Duits, P., Cath, D. C., Lissek, S., Hox, J. J., Hamm, A. O., Engelhard, I. M., ... Baas, J. M. P. (2015). Updated meta-analysis of classical fear conditioning in the anxiety disorders. *Depression and Anxiety*, 32, 239–253.
- Dymond, S. (2019). Overcoming avoidance in anxiety disorders: The contributions of Pavlovian and operant avoidance extinction methods. *Neuroscience and Biobehavioral Reviews*, 98, 61–70.
- Foa, E. B., & Kozak, M. J. (1986). Emotional processing of fear: exposure to corrective information. *Psychological Bulletin*, 99, 20–35.
- Gunther, L. M., Denniston, J. C., & Miller, R. R. (1998). Conducting exposure treatment in multiple contexts can prevent relapse. *Behaviour Research and Therapy*, 36, 75–91.
- Hardy, G. E., & Llewelyn, S. (2015). Introduction to Psychotherapy Process Research. In O. C. G. Gelo, A. Pritz, & B. Rieken (Eds.), *Psychotherapy Research*. Springer Vienna, pp. 183–194.
- Lovibond, P. F., & Shanks, D. R. (2002). The role of awareness in Pavlovian conditioning: Empirical evidence and theoretical implications. *Journal of Experimental Psychology: Animal Behavior Processes*, 28, 3–26.
- Mayo-Wilson, E., Dias, S., Mavranzouli, I., Kew, K., Clark, D. M., Ades, A. E., & Pilling, S. (2014). Psychological and pharmacological interventions for social anxiety disorder in adults: a systematic review and network meta-analysis. *The Lancet Psychiatry*, 1, 368–376.
- McConnell, B. L., & Miller, R. R. (2014). Associative accounts of recovery-from-extinction effects. *Learning and Motivation*, 46, 1–15.
- Miller, R. R., Barnet, R. C., & Grahame, N. J. (1995). Assessment of the Rescorla-Wagner model. *Psychological Bulletin*, 117, 363–386.
- Miller, R. R., Laborda, M. A., & Polack, C. W. (2020). Extinction training can make the extinction context a stimulus-specific inhibitor: A potential mechanism of experimental renewal. *Learning and Motivation*, 70, 101623.
- Miller, R. R., & Matzel, L. D. (1988). The Comparator Hypothesis: A Response Rule for The Expression of Associations. In *Psychology of Learning and Motivation* (pp. 51–92).
- Nihei, M., Hojo, D., & Sawa, K. (2021). The renewal effect in fear conditioning with aversive facial expression and negative sentences as unconditioned stimuli. *Learning and Motivation*, 74, 101725.
- Nihei, M., Hojo, D., Tanaka, T., & Sawa, K. (2023). A model for recovery-from-extinction effects in Pavlovian conditioning and exposure therapy. *Learning & Behavior*.
- Norton, P. J., & Price, E. C. (2007). A meta-analytic review of adult cognitive-behavioral treatment outcome across the anxiety disorders. *Journal of Nervous and Mental Disease*, 195, 521–531.
- Pavlov, I. P. (1927). *Conditioned reflexes*. (G.W. Anrep, Trans.). London: Oxford University Press. (Original work published 1926)
- Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review*, 87, 532–552.
- Pompili, A., Furukawa, T. A., Efthimiou, O., Imai, H., Tajika, A., & Salanti, G. (2018). Dismantling cognitive-behaviour therapy for panic disorder: A systematic review and component network meta-analysis. *Psychological Medicine*, 48, 1945–1953.
- Powers, M. B., Halpern, J. M., Ferenschak, M. P., Gillihan, S. J., & Foa, E. B. (2010). A meta-analytic review of prolonged exposure for posttraumatic stress disorder. *Clinical Psychology Review*, 30, 635–641.
- Rankin, C. H., Abrams, T., Barry, R. J., Bhatnagar, S., Clayton, D. F., Colombo, J., ... Thompson, R. F. (2009). Habituation revisited: An updated and revised description of the behavioral characteristics of habituation. *Neurobiology of Learning and Memory*, 92, 135–138.
- Rescorla, R. A., & Heth, C. D. (1975). Reinstatement of fear to an extinguished conditioned stimulus. *Journal of Experimental Psychology: Animal Behavior Processes*, 1, 88–96.
- Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W.F. Prokasy (eds.), *Classical conditioning II: current*

- research and theory* (pp. 64–99). New York: Appleton-Century-Crofts.
- Riccio, D. C., Richardson, R., & Ebner, D. L. (1999). The contextual change paradox is still unresolved: Comment on Bouton, Nelson, and Rosas (1999). *Psychological Bulletin*, *125*, 187–189.
- 佐藤友哉・田中恒彦・前田駿太（2023）制止学習アプローチに基づいたエクスポージャーの研究動向心理学研究, *94*, 83–97.
- 澤 幸祐（2011）古典的条件づけ 廣中直行（編）心理学研究法3 学習・動機・情動（pp. 15–38）誠信書房.
- 竹澤正哲（2018）心理学におけるモデリングの必要性心理学評論, *61*, 42–54.
- Thomas, B. L., & Ayres, J. J. B. (2004). Use of the ABA fear renewal paradigm to assess the effects of extinction with co-present fear inhibitors or excitors: Implications for theories of extinction and for treating human fears and phobias. *Learning and Motivation*, *35*, 22–52.
- Wolitzky-Taylor, K. B., Horowitz, J. D., Powers, M. B., & Telch, M. J. (2008). Psychological approaches in the treatment of specific phobias: A meta-analysis. *Clinical Psychology Review*, *28*, 1021–1037.
- Wolpe, J. (1958). *Psychotherapy by reciprocal inhibition*. Stanford University Press; Stanford, CA

— 2023. 2. 20 受稿, 2023. 9. 27 受理 —