

優秀修士論文概要

Reinforcement processes supporting habitual avoidance behaviour

川原田 弘 人

研究背景

連合学習の二重過程理論 (Dickinson, 1985) は, ヒトを含む動物の自発行動を二種類に分類する。特定の強化子 (結果) を得るための意識的な目的指向行動と, 特定の状況下で半自動的に誘発される習慣行動である。目的指向行動は行動と強化子の関係性の学習 (Response-Outcome: R-O 連合学習) に支えられる。一方, 習慣行動は状況刺激と行動の関係性の学習 (Stimulus-Response: S-R 連合学習) に支えられる。強化子の表象を参照するかどうかは両者の最大の違いである。

この理論的枠組みは, 強化子の価値操作と呼ばれる手続きを用いて検証されてきた。典型的には, 動物に特定の餌を得るためのレバー押し反応を訓練させた後, 餌と内蔵不快感を誘発する薬物を対呈示する。この手続きを通じて餌に対する味覚嫌悪条件づけが学習され, 餌の正の強化子 (報酬) としての価値は大きく低下する。その後テストを行うと, 行動が目的指向, つまり R-O 連合に制御される場合, 餌の価値低下に応じレバー押し反応は減少する (Adams & Dickinson, 1981)。逆に長期間の訓練等により行動が習慣的, つまり S-R 連合に制御されるようになった場合, 餌の価値が低下しても反応は減少しない (Adams, 1982)。強化子 (O) は文字通り S-R 連合には符号化されないためである。

上述の手続きを基に多くの研究が行われてきた (Perez & Dickinson, 2020) が, その対象はあくまで嗜好性行動 (餌等の正の強化子を獲得するための行動) に限られていた。不快・有害な事象 (負の強化子) を避けるための回避行動の習慣化を示した例は存在しない。本研究はラットの電撃回避行動について, 二重過程理論の目的指向—習慣の原理を拡張することを目的とした。

実験 1

本研究の最初の段階として, 負の強化子の価値操作手続きを確立し, 目的指向的な回避行動を示すことを目的とした。恐怖条件づけ研究で確立された方法 (e.g. Rescorla, 1974) を応用し, 電撃回避行動の確立後に異なる強度の電撃に曝露することで負の強化子としての電撃の価値を操作した。

【方法】

被験体・装置. 被験歴のない雄ラット (Long-Evans 系統) 16頭を用いた。標準的なラット用オペラント箱 (MED Associates) に電撃を呈示するためのショックジェネレーター (ENV-414, MED Associates) を接続した。

回避訓練. 各セッション開始と同時に回避期間が始まり, この間に回避反応 (レバー押し) が生じない場合には電撃期間に移行した。電撃期間では0.5mA の電撃が最大3回呈示され, 次の回避期間に移

行した。回避期間または電撃期間中にレバーを押すと安全信号（3 kHzの純音）が呈示され、安全信号の終了と共に次の回避期間に移行した。最大60回の電撃が呈示されるか、1時間が経過した時点でセッションは終了した。

価値操作. 訓練で獲得基準を達成したラットは、価値操作セッションを経験した。レバーと安全信号は呈示されず、ラットは計15回の電撃に強制的に曝露された。電撃の強度は Devalued 群では訓練時より弱い0.25mA、Inflated 群では訓練時より強い1.00mAであった。

テスト. 価値操作の翌日に消去テスト、翌々日に強化テストを行った。消去テストでは電撃は呈示されず、レバー押しに随伴して安全信号のみ呈示された。強化テストは、訓練の最終段階と同一のセッションであった。テストの成績については、相対回避反応率を主たる行動指標とした。これは訓練の最終3日間の回避反応率を基準に、各テストにおける回避反応率について相対値を算出したものである。

【結果・考察】

消去テストでは、Devalued 群の相対回避反応率が Inflated 群より有意に低かった ($t(8.8) = 2.72, p = .02$)。この結果は、回避行動が負の強化子の価値（電撃の強度表象）を参照していることを示唆する。強化テストでは群間の差は見られず ($t(12.3) = 0.41, p = .69$)、強度表象以外の意図しない影響により差が生じた可能性が排除された。

実験2

嗜好性行動の研究では、訓練量の増加に伴い行動が目的指向から習慣に移行することが示されている (e.g., Adams & Dickinson, 1981)。実験2では長期訓練を行い、実験1で確認された目的指向的な回避行動が習慣行動へと移行する、つまり価値操作手続きに対して鈍感になると示すことを目的とした。

【方法】

被験体・装置. 被験歴がない雄 Long-Evans 16頭を用いた。装置は実験1と同一であった。

回避訓練. 獲得基準を達成した後に20セッションの追加訓練を導入した。追加訓練以外は、実験1と同一の手順で行った。

価値操作. 実験1と同一の手順で行った。

テスト. 消去テストは、実験1と同一の手順で行った。強化テストは、訓練ではなく価値操作に用いた強度の電撃（Devalued 群で0.25mA、Inflated 群では1.00mA）を群に応じて呈示した。それ以外の点は、実験1と同一であった。

【結果・考察】

消去テストでは、群間の相対回避反応率に有意差がなかった ($t(10.8) = 0.64, p = .54$)。この結果は、長期訓練により、回避行動が負の強化子の価値を参照しない習慣行動に移行したことを示している。一方強化テストでは、Devalued 群が Inflated 群より有意に低かった ($t(12.0) = 2.63, p = .02$)。この結果は、電撃の強度の差が十分であった、つまりラットが強度の違いを弁別できなかったために消去テストにおける価値操作への鈍感さが生じた可能性を排除するものであった。

実験 3

実験 2 で示された価値操作への鈍感さは、二つの異なる解釈が可能であった。第一に、当初予測した通り、長期間の訓練に伴い電撃回避という負の強化過程自体が習慣化したという解釈である。第二に、強化過程自体が負の強化から正の強化へと移行したという解釈である。安全信号はそれ自体が正の強化子としての性質を獲得することが知られている (Dinsmoor, 2001)。主要な強化過程が電撃回避 (負の強化) から安全信号獲得 (正の強化) へ移行した結果、負の強化子である電撃が参照されなくなった可能性がある。この二つの可能性を検討することを目的に、拮抗条件づけにより安全信号の価値を操作した。この操作は中程度訓練と長期訓練の両条件で実施し、訓練量による強化過程の変化を検討した。

【方法】

被験体・装置. 被験歴がない雄 Long-Evans 24頭を用いた。装置は実験 1 と同一であった。

回避訓練. 中程度訓練 (MT) 条件と長期訓練 (OT) 条件が設定された。MT 条件は実験 1、OT 条件は実験 2 と同一の手順で訓練された。

低価値化. ラットは各訓練条件 (OT・MT) 内で、さらに Devalued 群と Control 群に分けられた。レバーを呈示せず匂い刺激で文脈を変更した点は実験 1 と共通であったが、電撃の強度は訓練時と同じ 0.5mA で統一された。Devalued 群では、安全信号として用いられていた 3 kHz の純音と電撃が対呈示された。Control 群では、明確に非対呈示された。両群とも 15 回の呈示を 1 セッションのみ行った。

テスト. 消去テストでは、電撃だけでなく、安全信号も呈示されなかった。強化テストは、実験 1 と同一の手順で行った。

【結果・考察】

MT 条件の消去テストでは、Devalued 群の相対回避反応率が Control 群より有意に低かった ($t(6.7) = 2.87, p = .03$)。この結果は実験 1 と同様に、回避行動が安全信号の価値も参照していることを示唆している。一方 OT 条件の消去テストでは、群間に有意な差はなかった ($t(8.9) = 1.04, p = .33$)。この結果は実験 2 と同様に、長期訓練によって回避行動が安全信号の低価値化 (価値操作) に鈍感になったことを示唆している。

MT 条件の強化テストでは群間の差は見られなかったが ($t(9.1) = 0.25, p = .81$)、OT 条件の強化テストでは Devalued 群が Control 群より有意に低かった ($t(9.7) = 2.39, p = .04$)。この結果は、訓練が進むにつれて電撃と安全信号の相対的な強化力が逆転したことを反映している可能性がある。

総合考察

本研究は、ラットの電撃回避行動について目的指向—習慣の遷移過程を実験的に検討した。実験 1 では異なる強度の電撃に曝露する価値操作手続きを行い、少なくとも中程度訓練下では、回避行動が負の強化子の価値を参照する目的指向行動であることを明らかにした。また、全般的な活動レベルの低下等が影響を与えた可能性を排除する強化テストの結果と併せて、今回採用した価値操作法の妥当性が示された。実験 2 では追加訓練後に価値操作を行い、回避行動が長期訓練を通じて強化子の価値変化に鈍感な習慣行動になることを示した。また強化テストの結果も、その鈍感さが電撃強度の弁別能力ではなく、

負の強化子の価値の参照の有無から生じていることを支持した。

ただし実験2で観察された電撃の価値操作への鈍感さは、回避行動を支える主たる強化過程が電撃による負の強化から安全信号による正の強化へと移行したことで生じた可能性もあった。この仮説を検討するため、実験3では安全信号の低価値化を行った。中程度訓練下では低価値化で回避行動が減少したが、長期訓練下ではその差が消失した。実験3の結果は、訓練が進むにつれて電撃・安全信号両方の価値が参照されなくなることを示唆している。つまり今回観察された回避行動の習慣化は、負の強化から正の強化への移行ではなく、実験2で事前に予測したようにS-R連合の優位性に特徴づけられる。

一方実験3の強化テストでは、中程度訓練下では群間の差が観察されず、むしろ長期訓練下で低価値化の効果があつた。消去テストと反対の結果が得られた理由としては、例えば訓練の進行に伴う安全信号と電撃の相対的な強化力のバランスの変化が挙げられる。訓練初期は電撃の強化力が相対的に強く、電撃と安全信号の両方が呈示される強化テスト場面において安全信号の低価値化の影響が隠蔽されていたのかもしれない。

本研究は、二重過程理論の枠組みが回避行動にも適用可能であることをはじめて示した。これは回避行動の基盤を明らかにするだけでなく、強迫性障害をはじめとした不適切な回避を症状とする精神疾患への応用の端緒ともなることが期待される。

優秀修士論文概要

カルバック・ライブラー情報量を用いた人の価格判断の
メカニズムに関する研究

高 正 月

目 的

本論文では、価格に対する予想値と提示価格の間のKL情報量 (Kullback-Leibler divergence) を注目度を示す指標とし、竹村 (1998) の心的モノサシ理論から発展した「注目度が高い価格に対して、高いか安いかの判断がより敏感になる傾向がある」という価格判断過程のメカニズムを説明するモデル (SKL モデル, Sum of the Kullback-Leibler Divergence Model) を提案した。また、このモデルを、物価に対する実感を対象とした場合 (研究1) と特定商品の価格判断を対象とした場合 (研究2) の二つの場面で検討することを目的とした。

提案モデル

まず、KL情報量は2つの確率分布間の乖離を示す指標であり、PのQに対するKL情報量は(1)式で求められる。ここで、 $p(\cdot)$ 、 $q(\cdot)$ はそれぞれP、Qの確率密度関数を表す。

$$KL(P, Q) = \int_{-\infty}^{\infty} p(x) \log \frac{p(x)}{q(x)} dx \quad (1)$$

本提案モデルでは、Itti & Baldi (2009) が提案したベイジアン・サプライズを参考に、物価に対する予測値 (商品の価格に対する予想値、以下では参照価格とする) と実際の物価水準 (商品の提示価格) とのギャップが人の注目度に影響を与えたと考えた。なお、このギャップは、当該時点の実際の物価水準 (商品の提示価格) $u_{1,t}$ を平均、その分散 $s_{1,t}^2$ を分散とする正規分布と、当該時点の物価に対する予測値 (商品に対する参照価格) $u_{2,t}$ を平均、 $s_{2,t}^2$ を分散とする正規分布間のKL情報量を用いて表現できると考えた。また、竹村 (1998) の心的モノサシ理論に基づき、注目度が価格に対する感じ方に影響を与えることを仮定し、(2)・(3)式を用いて物価に対する実感の判断値 (商品の価格に対する判断値) $v(u_{1,t})$ を導出できると考えた。このモデルをSKLモデルと称した。

$$v(u_{1,t}) = k \cdot \int KL(N(u_{1,t}, s_{1,t}^2), N(u_{2,t}, s_{2,t}^2)) du_{1,t} \quad (2)$$

$$= k \cdot \left(\frac{u_{1,t}u_{2,t}^2 - u_{1,t}^2u_{2,t} + \frac{u_{1,t}^3}{3} + s_{1,t}^2u_{1,t}}{2s_2^2} + u_{1,t} \ln \frac{s_{2,t}}{s_{1,t}} - \frac{u_{1,t}}{2} \right) \quad (3)$$

研究1

目的 研究1では、人々の物価に対する実感に着目し、SKLモデルを検討することを目的にした。具体的には、物価に対する実感データを分析対象とし、単純差分モデル、対数モデル、そして村上他

(2022) で提案された KL 回帰モデルと比較しながら、SKL モデルのデータへの当てはまりの良さを検討した。

分析対象 物価に対する実感データとして、日本銀行の生活意識アンケート調査における「物価に対する実感」の 2008 年から 2023 年の 16 年間にわたるデータを用いた (日本銀行, 2024)。具体的には、「現在の物価は 1 年前に比べて何 % 程度変化したと思うか?」に対してパーセンテージで回答させた平均値と、「1 年後の物価は現在と比べて何 % 程度変化すると思うか?」に対して同じ要領で回答させた平均値を使用した。実際の物価データとしては、2008 年から 2023 年までの 2020 年基準の日本消費者物価指数 (Consumer Price Index, CPI) の総合指数を用いた (総務省統計局)。

分析方法 上記の「物価に対する実感」データに対して、(3) 式に定数項 l と誤差項 ε を加えた SKL モデルを最小二乗法により当てはめた。ここで、 $v(u_{1,t})$ を t 時点における「現在の物価は 1 年前に比べて何 % 程度変化したと思うか?」の平均回答、 $u_{2,t}$ を $t-1$ 時点の「1 年後の物価は現在と比べて何 % 程度変化すると思うか?」に対する平均回答 (例えば、2009 年の $v(u_{1,t})$ に対して、2008 年の回答を $u_{2,t}$ に用いた)、 $s_{2,t}^2$ を全時点分の「現在の物価は 1 年前に比べて何 % 程度変化したと思うか?」の平均回答の分散とした。そして、 $u_{1,t}$ については、 $u_{2,t}$ の単位と揃えるため、(4) 式のように変換した値を用いた。 $s_{1,t}^2$ にはこの変換後の値の全時点分の分散を代入した。ここで、 CPI_t は t 時点の CPI 総合指数を表す。

$$\frac{CPI_t - CPI_{t-1}}{CPI_{t-1}} \times 100 \quad (4)$$

また、SKL モデルの当てはまりの良さを比較するために、(5) 式の単純差分モデル、(6) 式の対数モデル、および (7) 式の KL 回帰モデルについても当てはめを行った。なお、(7) 式に関しては、AIC に基づいて回帰モデルの変数選択も行った。ここで、(7) 式は村上他 (2022) や高他 (2023) が提案したもので、KL 情報量、KL 情報量の 1 回差分 (速度に相当する $first_KLdiff_t$)、および KL 情報量の 2 回差分 (加速度に相当する $second_KLdiff_t$) を説明変数として取り入れたモデルである。1 回差分 $first_KLdiff_t$ は t 時点の KL 情報量と $t-1$ 時点の KL 情報量の差分として計算され、2 回差分 $second_KLdiff_t$ は t 時点の KL 情報量と $t-1$ 時点の KL 情報量の差分、および $t-1$ 時点の KL 情報量と $t-2$ 時点の KL 情報量の差分との差分として計算される。ここで、 α 、 θ 、 β 、 γ 、 δ はフリーパラメータ、 l は定数項、 ε は誤差項である。

$$y_t = \alpha \cdot (u_{1,t} - u_{2,t}) + l + \varepsilon \quad (5)$$

$$y_t = \theta \cdot \log\left(\frac{u_{1,t}+100}{u_{2,t}+100}\right) + l + \varepsilon \quad (6)$$

$$y_t = y_{t-1} + \beta \cdot KL\left(\left(N(u_{1,t}, s_{1,t}^2), N(u_{2,t}, s_{2,t}^2)\right)\right) + \gamma \cdot first_KLdiff_t + \delta \cdot second_KLdiff_t + l + \varepsilon \quad (7)$$

結果と考察 物価に対する実感データを対象に (3)、(5)-(7) 式の各モデルを当てはめ、モデルの説明率として決定係数 R^2 、予測精度の指標として AIC を求めた。また、実際の物価水準、実感判断値、

各モデルの推定値を比較し、時系列変化を視覚的に検討した。結果から、SKLモデルは単純差分モデルと対数モデルより高い説明力を持つことが示唆された。一方で、KL回帰モデルより一貫して優れているとは言えないものの、データ特性に応じてよりの確に説明できる可能性が示された。

研究2

目的 研究2では、商品の価格判断実験（泉・矢崎，2023）を通じてSKLモデルとそれが示す「注目度が高い価格に対して、高いか安いかの判断がより敏感になる傾向がある」という考えを検討した。

実験概要 実験は、価格判断課題と購買意思決定課題の2つで構成された。価格判断課題では、提示された商品の価格を9件法（1：非常に安い～9：非常に高い）で評価させた。一方、購買意思決定課題では、商品を購入するか否かを回答させた。対象はクラッカー、パスタ、ティッシュの3種類の商品で、それぞれ同じ金額の2種類の価格提示順序（cheap系列とexpensive系列）を設定した。各商品について20時点分の価格判断を行わせ、その後、20時点目に購買意思決定課題を実施した。価格提示順序は、新保（2022）の2時点シミュレーション法とSKLモデルに基づき、同じ価格に対して安いと感じやすいと予想される提示順序（cheap系列）と、高いと感じやすいと予想される提示順序（expensive系列）として設定された。さらに、実験中には参加者にアイカメラを装着させ、視線データの測定も行った。

実験対象 実験参加者は大学生54名（男性27名、女性27名）で、年齢は18歳から48歳、平均年齢は21.60歳（標準偏差4.16歳）であった。そのうち、最終的に分析には52名（男性25名、女性27名）のデータを用いた。年齢は18歳から48歳、平均が21.70歳（4.22歳）であった。

分析方法 実験結果を以下の3つの観点から分析した。1つ目は提示価格のexpensive系列とcheap系列が、SKLモデルの予測通りの効果を示すかどうかを確認することであった。2つ目は、SKLモデルが価格判断データにどの程度当てはまるかを評価することであった。3つ目は、各時点の価格が提示されてから判断が決まるまでの反応時間（注視時間）、提示価格に関するKL情報量、および判断値の変化率（評価関数の傾き）の関係性を分析することであった。

結果と考察 まず、19時点目と20時点目の判断値の差、および20時点目の判断値についてt検定を行い、その結果をTable 1に示した。SKLモデルでは、cheap系列において19時点目と20時点目の判断値の差が大きく、20時点目の判断値が小さいことが期待されるが、この効果はパスタの一部においてのみ確認された。また、判断値の平均データに対するSKLモデルの決定係数 R^2 をTable 2に示し、全体的に説明率が高いことが示唆された。最後に、反応時間（注視時間）、KL情報量、判断値の変化率の関係性について無相関検定を実施したところ、SKLモデルが仮定する正の相関関係は支持されなかった。結果から、SKLモデルとその示す「注目度が高い価格に対して、高いか安いかの判断がより敏感になる」という価格判断のメカニズムは、研究2の実験において十分に支持されないことが示唆された。

Table 1 各商品の系列別判断値の平均値と t 検定結果

商品		cheap 系列の 平均値	expensive 系列の 平均値	p 値	効果量
クラッカー	19 時点目と 20 時点目 の判断値の差	1.13	1.12	.85	.03
	20 時点目の判断値	3.67	3.62	.62	.07
パスタ	19 時点目と 20 時点目 の判断値の差	1.10	1.15	.61	.09
	20 時点目の判断値	3.67	3.96	.04	.35
ティッシュ	19 時点目と 20 時点目 の判断値の差	1.29	1.31	.90	.02
	20 時点目の判断値	3.75	3.79	.81	.04

Table 2 商品と系列ごとの SKL モデルの R^2

商品	系列	R^2
クラッカー	cheap	.77
	expensive	.78
パスタ	cheap	.77
	expensive	.76
ティッシュ	cheap	.80
	expensive	.77

総合考察

研究 1 と研究 2 の結果から、SKL モデルは物価に対する実感データや価格判断データを説明するモデルとして一定の可能性を示したものの、その仮定や適用範囲には引き続き課題が残されていることが明らかになった。今後は、より精度の高いデータの収集、実験とモデルの改良を進めることで、SKL モデルの仮定の妥当性や適用範囲をさらに検証する必要がある。これにより、価格判断における注目度と判断の敏感さの関係性がより明確になり、価格判断メカニズムに対する理解が深まることが期待される。

謝辞

本研究に際してご助言およびご協力をいただいた、竹村和久、村上始、川杉桂太、井出野尚の各先生、左浩然、平木琴音、泉麗華、矢崎蒼の各氏に深く感謝申し上げます。

引用文献

- Itti, L., & Baldi, P. (2009). Bayesian surprise attracts human attention. *Vision research (Oxford)*, 49 (10), 1295-1306.
- 竹村 和久 (1998). 状況依存的な意思決定の定性的モデル—心的モノサシ理論による説明— 認知科学, 5(4), 417-434