

推論課題におけるマウストラッキングを用いた方略推定の試み: 計算論モデリングとの比較

Attempts to estimate strategies based on mouse tracking in inference tasks: Comparison with computational modeling

白砂 大[†], 香川 璃奈[‡], 本田 秀仁^{††}

Masaru Shirasuna, Rina Kagawa, Hidehito Honda

[†] 静岡大学, [‡] 産業技術総合研究所, ^{††} 追手門学院大学

Shizuoka University, National Institute of Advanced Industrial Science and Technology, Otemon Gakuin University
m.shirasuna1392@gmail.com

概要

複数の選択肢から 1 つを選ぶ際に人がどのような判断方略(ヒューリスティックまたは知識)を用いたのかは、通常、行動データの計算論モデリングから推定される。本研究では、カーソル軌跡という観測可能な動作から判断方略を推定することを試みた。具体的には、二者択一の推論課題を用いて、個人ごとにモデルに基づいて使用方略を推定し、カーソル軌跡の形状に基づいてクラスタリングした。結果として、使用方略と軌跡との連関は非有意だったが中程度の効果量を示した。マウストラッキングの応用可能性と限界について議論した。
キーワード: ヒューリスティック(heuristic), 計算論モデリング(computational modeling), マウストラッキング(mouse tracking), クラスタリング(clustering)

1. マウストラッキングによる判断方略の推定

複数の選択肢から 1 つを選ぶ際に、人は様々な判断方略を取りうる。例えば、具体的な情報(知識)を持っていれば、それに基づいて判断を行える。一方、知識がなければ、「なじみ深い方を選ぶ」などヒューリスティックな判断を行う。判断時に人がどのような方略を用いていたのかは、通常、行動データの計算論モデリングによって推定される(e.g., Jenny et al., 2014; Pachur et al., 2013)。すなわち、様々な判断方略をそれぞれ記述する認知モデルを構築し、どのモデルが判断者の回答パターンを最もよく説明するかを統計的手法によって分析することで、明らかにしていく。人がどのようなヒューリスティックを用いるかを検証することは、認知科学・意思決定科学における中心トピックの 1 つであり、人が課題に適応する形で判断方略を選び取っているのかという合理性の理解(e.g., Hertwig et al., 2022)や、個人の判断傾向に沿ってパーソナライズされた介入策の検討(e.g., Ingendahl et al., 2021)において重要となる。

モデルベースによる計算論的手法とは別に、より工学的・実験心理学的手法として、判断時の動作に基づく

推定がある。中でも簡易的かつ低コストで実施できるものの 1 つに、マウストラッキングがある(e.g., Freeman, 2018; Freeman & Ambady, 2010)。これは、PC 上での選択課題において、課題提示から選択までの間に判断者が描いたマウスカーソルの軌跡から、当該の試行における認知プロセス(e.g., 迷いの大きさ)を推定するアプローチである。マウストラッキングは、カーソル軌跡という観測可能な指標に基づくため、試行ごとに逐一、判断者の認知傾向を推定できる。同手法はまた、課題遂行中のデータからリアルタイムに推定を行い、迅速な介入を設計できる可能性も秘めている。

しかし計算論モデリングとマウストラッキングは、それぞれ(計算論的)認知科学と工学・実験心理学という別個の領域で主に扱われてきた。そのため、カーソル軌跡に基づく推定が、モデルの当てはめによる推定とどの程度整合的であるかは不明である。カーソル軌跡のプロトタイプとそれに対応する認知プロセスは提案されているものの(e.g., 直線的な軌跡は迷いの少なさを意味する; Wulff et al., 2025)、これらの認知モデルなどによる裏付けは乏しい。マウストラッキングとモデル分析の両結果を比較することは、計算論的および工学的の両アプローチを統合し、判断傾向のリアルタイムな推定の裏付けやそれに基づく介入設計に資するだろう。

本研究では、マウスカーソルの軌跡という観測可能な動作から、モデルベースで推定された判断方略を予測できる可能性を探索的に検討した。具体的には、二者択一の推論課題を用いて、マウストラッキングと計算論モデリングのそれぞれによる判断方略の推定結果に、どの程度の連関が見られるかを検証した。

2. 方法

2.1. 実験参加者

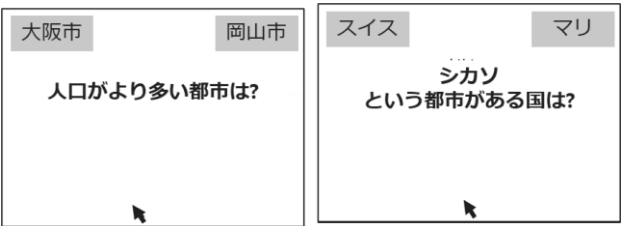
20~59 歳の日本人 104 名が参加した。このうち半数

の 52 名ずつが、それぞれ後述の「人口推定課題」($M_{age} = 37.42, SD_{age} = 13.18, N_{female} = 26$) または「関係比較課題」($M_{age} = 39.15, SD_{age} = 12.75, N_{female} = 26$)に割り当てられた。参加者は、中央サトー(株)に登録しているモニターより募集された。

2.2. 題材・手続き

結果の頑健性を担保するため、いずれもヒューリスティック研究で使われてきた二者択一の推論課題である人口推定課題(Goldstein & Gigerenzer, 2002)と関係比較課題(Shirasuna et al., 2020)の 2 種類を用意した。実験は対面形式で、PC とマウスを用いて行われた。実験課題は PsychoPy 3.0 を用いて作成された。二者択一課題では、1 フレームごとのカーソル位置(xy 座標)が記録された。各問題において、マウスカーソルの初期位置は画面下中央に配置され、問題文が画面中央に、選択肢が画面左上端と右上端にそれぞれ呈示された。

図 1 二者択一課題の例



註: 左は人口推定課題の例、右は関係比較課題の例。

人口推定課題: 2 つの都市が呈示され、人口が多いと思う方を選ばせる課題であった(図 1 左; e.g., 人口の多い都市は? 大阪市 or 岡山市)。この課題は、2 つの選択肢の比較(e.g., 大阪市と岡山市を比べる)が求められる構造であった。本研究では Honda et al. (2017)に使われた 210 問を使用した。二者択一課題の後に、なじみ深さ評定課題と知識課題の 2 つが行われた。なじみ深さ評定課題は、二者択一に使われた地名が 1 つずつ呈示され、それぞれの程度なじみがあるか(0: 全くなじみがない—100: 非常になじみがある)を visual analog scale で評価する課題であった。知識課題は、なじみ深さ評定課題と同じく都市が 1 つずつ呈示され、各都市の人口数を直接回答させる課題であった。

関係比較課題: ある都市が、2 つの国のうちどちらに属すると思うかを選ばせる課題であった(図 1 右; e.g., シカソという都市がある国は? スイス or マリ)。この課題は、選択肢に加えて問題文中の対象物を考慮した比

較(上の例では「シカソとスイス」と「シカソとマリ」を比べる)が求められる課題構造であった。本研究では Shirasuna et al. (2020)に使われた 100 問が用いられた。人口推定課題と同様に、二者択一課題の後にはなじみ深さ評定課題と知識課題が行われた。前者は人口推定課題と同様の手続きであった。後者は、呈示された地名があると思う「国」(都市名が呈示された場合のみ)や「地域」「言語圏」などの属性を回答させる課題であった。

2.3. 分析

人口推定課題、関係比較課題ともに、先行研究で difficult 問題と設定された半数の問題(それぞれ 105 問、50 問)を分析に使用した。これは、easy 問題で大半の参加者が短時間で正答してしまう(非常に似通った回答パターンのみが現れる)ことを避けるためであった。計算論モデリング: 人の回答パターンをよく説明するとされる複数の判断方略のモデルを用いた(表 1; Honda et al., 2017; Shirasuna et al., 2020)。先行研究で提案された、ベイズ情報量規準などに基づく統計的手法(詳細は Jenny et al., 2014 など参照)により、各参加者の回答パターンを最もよく説明するモデルを特定し、そのモデルを「その参加者が使用した方略」と見なした。

表 1 本研究で用いた推論モデル

人口推定課題 (Honda et al., 2017 も参照)	
FH	よりなじみ深い都市を選ぶ
KI	人口数が多いと思う都市を選ぶ
関係比較課題 (Shirasuna et al., 2020 も参照)	
FM	「都市のなじみ深さ」に、より近いなじみ深さを持つ国を選ぶ
FH	よりなじみ深い国を選ぶ
KI	「都市の属性」と「国の属性」を比較し、一致する属性がより多い国を選ぶ

註: なじみ深さが等しいなど 2 選択肢の差が見られない場合は、ランダムに選ばれると仮定された。略語は、FM=familiarity matching、FH=familiarity heuristic、KI=knowledge-based inference を示す。なお FM と FH はヒューリスティックを用いた方略、KI は知識を用いた方略である。

マウストラッキング: 各試行で計測されたカーソル軌跡(1 フレームごとの xy 座標)について、終点方向を画面左上部に揃え、軌跡を 100 点の等間隔に分割した(位

置および長さの正規化)。この正規化手法は、マウストラッキングにおける典型的な手続きであった(e.g., Wulff et al., 2025)。そして、軌跡同士の(非)類似度の指標としてユークリッド距離を算出し、その距離行列に基づいて k-means 法による分類を行った。モデルと対応付けるため、クラスタ数は人口推定課題では「2」、関係比較課題では「3」とした。参加者ごとに、各問題(人口推定課題は 105 問、関係比較課題は 50 問)における軌跡がどのクラスタに分類されるかを数え、最も多く分類されたクラスタをその参加者の判断傾向と見なした。例えば、関係比較課題に回答した参加者 A について、50 問中のクラスタ 1、2、3 の観測数がそれぞれ 13、27、10 であった場合、参加者 A の判断傾向はクラスタ 2 に属すると見なした。以上の分析は、R の mousetrap パッケージを用いて行われた。計算論モデリングおよびマウストラッキングの結果に基づき、どの方略の使用者がどのような軌跡を示すかを検証した。

3. 結果と考察

まず、行動データの計算論モデリングにより、各参加者が使用したとされる方略を推定した(図2、棒グラフ)。その結果、知識に基づいて考える人(KI ユーザー)だけでなく、ヒューリスティックを用いる人(FH、FM ユー

ザー)も存在していた。この結果は、実世界の知識を問う推論課題であっても難しい問題ではヒューリスティックが使われやすいという先行研究の知見(e.g., Honda et al., 2017)と整合的であった。

続いて、マウスカーソルの軌跡の形状に基づいてクラスタリングを行った。その結果、直線的で迷いが少ないもの(straight)、直線的だが中程度の迷いが生じたもの(curved)、強い競合があり直前で判断を切り替えたもの(Change of mind; CoM)が確認された(図2、軌跡図)。これらはいずれも、Wulff et al. (2025)で提案された軌跡プロトタイプにも合致するものであった。

以上の結果に基づき、各方略と軌跡との対応関係を検証した(図2、対応表)。その結果、両課題における連関の効果量は 0.25 前後と中程度であり、軌跡と使用方略は強く対応するほどではなかった(人口推定 $\chi^2 = 3.144, p = .109$, Cramer's $V = 0.246$; 関係比較 $\chi^2 = 8.189, p = .080$, Cramer's $V = 0.281$)。すなわち、特定の方略使用者が特定の軌跡を示すわけではないといえる。一見すると、KI ユーザーは慎重な判断を行い、CoM 的な動きを示しやすいように思われる。しかし実際には、CoM な動きもそうでない動きも同程度に示した。一方ヒューリスティック(FM、FH)ユーザーは、常に直線的な動きを示すわけではなく、課題によって異なる傾向を示した(人口推定課題では CoM、関係比較課題では curved

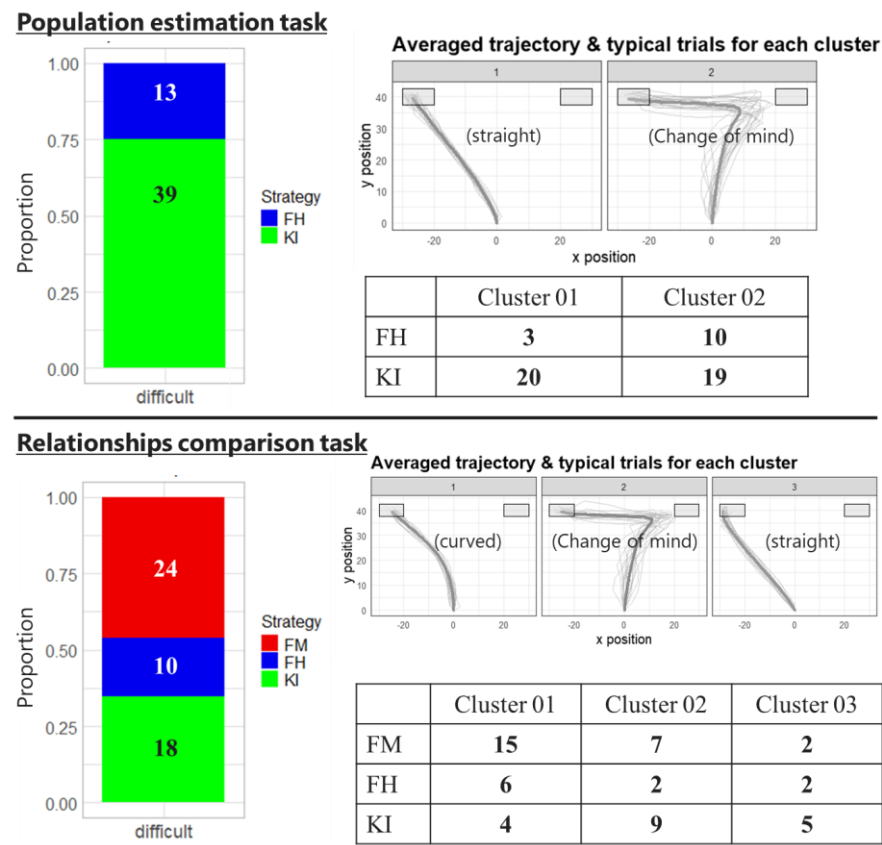


図2 計算論モデリングとマウストラッキングの結果

註: 上半分は人口推定課題、下半分は関係比較課題の結果である。左の棒グラフは、モデル分析から推定された、各方略の使用者数を示す(いずれも $N = 52$; 数値は人数)。右上の軌跡図は、クラスタリングで抽出された軌跡を示す(細線は各試行レベルで観測された軌跡、太線は平均的な軌跡)。横軸と縦軸はそれぞれ画面上の x 座標と y 座標を示す。右下の対応表は、各方略と各クラスタに該当した人数の内訳を示す。

が多数派であった)。これは、ヒューリスティックが単純で直感的な方略なのではなく、課題に適応する形で使われている(Shirasuna et al., 2020)ことを反映しているかもしれない。いずれにせよ本研究の結果は、観測可能なカーソル軌跡のみから個人の判断方略を強く予測できるとは限らないことを示唆している。

4. 総合考察

本研究では、二者択一の推論課題において、マウスカーソルの軌跡から、行動データの計算論モデリングによって推定された判断方略を予測できる可能性を、探索的に検討した。結果として、両者の連関は有意ではなかったが、効果量は中程度であった。

実世界では、デジタル化の進展に伴い、PC デバイスによる課題遂行の場面が増えている。マウストラッキングは、判断者への身体的負荷をかけずに選択に至るまでの認知プロセスや判断傾向をリアルタイムに検知でき、個人に適した形で迅速に介入策を施すといった応用も期待されている(e.g., Stillman et al., 2018)。カーソル軌跡に基づくより精緻な認知プロセスの推定に向けては今後、次の3点の検討が望まれる。

1点目は、各試行レベルでの検討である。使用方略は課題内である程度一貫していると想定されるのに対し、カーソル軌跡は試行ごとに様々に観測される。軌跡と判断とをより直接的に対応づけるには、なじみ深い方が選ばれたか否かなど、個人の回答も試行レベルで検討することが望ましい。例えば、「なじみ深い方が選ばれたときに CoM が多く観測される」傾向があれば、人はまずなじみの薄い方に手を伸ばし、選択の直前になじみ深い方に切り替えるというプロセスが示唆される。

2点目は、マウストラッキング手法の適用可能性である。先述のように、使用方略とカーソル軌跡の関係は課題によって異なっていた。この点は、マウストラッキングがどの課題にも普遍的に適用できるとは限らないこと、また結果の解釈の際には課題の性質を考慮すべきであることを示唆していると考えられる。

3点目は、カーソル軌跡の時間的な指標である。今回のクラスタリングは、カーソルの空間的な指標(xy座標)に依存しており、反応時間や速度、加速度などは考慮されていなかった。しかし、同じ直線的動作であっても、どの程度時間をかけたか、各試行のどの時点で速度が最大になったかなどによって、認知プロセスは異なると考えられる(Wulff et al., 2025)。カーソル軌跡の時間

的・空間的の双方の指標を用いることで、人の潜在的な認知をより反映した分析が可能となるだろう。

謝辞

本研究の一部は、JST さきがけ(JPMJPR2313)および科研費(23K25169、23K16249)の支援を受けた。

文献

- Freeman, J. B. (2018). Doing psychological science by hand. *Current Directions in Psych Sci*, 27(5), 315–323.
- Freeman, J. B., & Ambady, N. (2010). MouseTracker: Software for studying real-time mental processing using a computer mouse-tracking method. *Behav Res Meths*, 42, 226–241.
- Goldstein, D., & Gigerenzer, G. (2002). Models of ecological rationality: The recognition heuristic. *Psych Rev*, 109(1), 75–90.
- Hertwig, R., Leuker, C., Pachur, T., Spiliopoulos, L., & Pleskac, T. J. (2022). Studies in Ecological Rationality. *Topics in Cognitive Science*, 14(3), 467–491.
- Honda, H., Matsuka, T., & Ueda, K. (2017). Memory-based simple heuristics as attribute substitution: Competitive tests of binary choice inference models. *Cog Sci*, 41(S5), 1093–1118.
- Ingendahl, M., Hummel, D., Maedche, A., & Vogel, T. (2021). Who can be nudged? Examining nudging effectiveness in the context of need for cognition and need for uniqueness. *J of Cons Behav*, 20(2), 324–336.
- Jenny, M. A., Rieskamp, J., & Nilsson, H. (2014). Inferring conjunctive probabilities from noisy samples: Evidence for the configural weighted average model. *J of Exp Psych: LM & C*, 40(1), 203–217.
- Pachur, T., Hertwig, R., & Rieskamp, J. (2013). Intuitive judgments of social statistics: How exhaustive does sampling need to be? *J of Exp Soc Psych*, 49(6), 1059–1077.
- Shirasuna, M., Honda, H., Matsuka, T., & Ueda, K. (2020). Familiarity-matching: An ecologically rational heuristic for the relationships-comparison task. *Cog Sci*, 44(2), e12806.
- Stillman, P. E., Shen, X., & Ferguson, M. J. (2018). How mouse-tracking can advance social cognitive theory. *Trends in Cognitive Sciences*, 22(6), 531–543.
- Wulff, D. U., Kieslich, P. J., Henninger, F., Haslbeck, J. M. B., & Schulte-Mecklenbeck, M. (2025). Movement tracking of psychological processes: A tutorial using mousetrap. *Behav Res Meths*, 57(11), 307.