

科学研究費助成事業 研究成果報告書

平成 26 年 6 月 18 日現在

機関番号：32689

研究種目：基盤研究(C)

研究期間：2011～2013

課題番号：23530871

研究課題名(和文) 評定判断における意思決定過程の研究

研究課題名(英文) Decision making processes in rating judgments

研究代表者

椎名 乾平 (Shiina, Kenpei)

早稲田大学・教育・総合科学学術院・教授

研究者番号：60187317

交付決定額(研究期間全体)：(直接経費) 1,200,000円、(間接経費) 360,000円

研究成果の概要(和文)：評定尺度法は心理学の基本的ツールであるが、その性質はよくわかっていない。本研究では新しい研究パラダイム(動的評定尺度)を用いて、評定判断における心的過程や意思決定過程を明らかにしようと試みた。結果としてカーソル軌跡の速度や振動が心的状態を反映することが強く示唆され、心理学における有効性が確認できた。

研究成果の概要(英文)：Rating methods are a principal data acquisition procedure for psychology. We do not know, however, the nature of the method very well. Using a new method of Dynamic Rating Scale, this study tried to clarify decision making processes in rating judgment. It was revealed the new method is promising because the velocity and variation of trajectories reflect mental states.

研究分野：社会科学

科研費の分科・細目：心理学 教育心理学

キーワード：評価 評定尺度 意思決定

1. 研究開始当初の背景

人間の判断や意思決定は、それ自身が研究対象となると共に、ブラックボックスとして様々な研究にデータを提供している。例えば、知覚判断における反応時間の分布や平均値はそれ自身研究の対象になると同時に、反応時間生成過程をブラックボックスとして仮定した上で多くの認知研究に基礎データを提供している。

伝統的に心理学で最も使用されているデータ獲得手段はなんといっても「評価尺度」であろうが、評価判断を意思決定過程と捉えた場合、結果として得られる評価値は相当に複雑な意思決定過程の結果と考えるべきである。例えば、紙と鉛筆を用いた評価判断において、決定できずに迷っている受検者はしばしば観察されるし、鉛筆を評価カテゴリーの間で右往左往させるような光景すら散見される。このような overt な行動は、受検者の心的状態を反映していると思われるが、従来はこの直感を実証するのは難しかった。ところで図1に示されるようなコンピュータを用いたインターフェースを使用するならば、カーソルの軌跡を直接記録することができる。申請者はこのような評価尺度を動的評価尺度(DRS, Dynamic Rating Scale)と呼んでおり、その基本的性質について2004年頃から少しずつ調べてきた。DRSによって獲得されるのは、図1上に示されているような軌跡のデータなので、評価判断のリアルタイム処理に対して新しい展望を与えてくれるだろうと期待したからである。

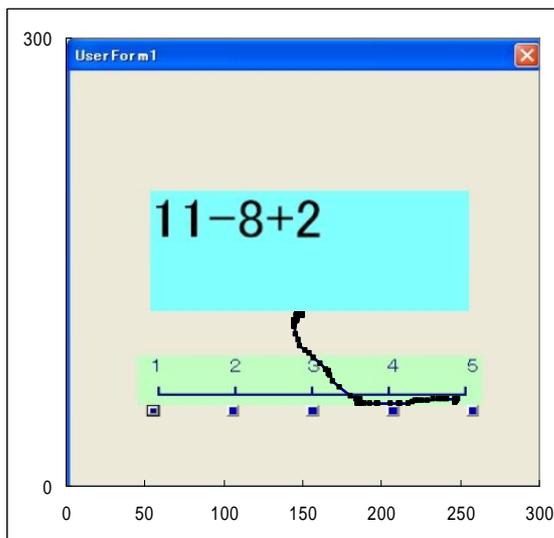


図1 動的評価尺度とカーソル軌跡の例

評価尺度はギリシャ時代から存在し、知覚・生理のような基礎分野でも使われている。しかしながら、評価尺度の基本スペックに対する疑問(例えば、尺度水準は何か? 何を測っているのか? 統計処理上の問題は無いのか? 等々)は繰り返し提起されているものの未解答と言えよう。そこで「測定装置としての評価尺度」について、基本的性質を明らかにし、心理学の足元を見直す契機としたいと考えた。

2. 研究の目的

本研究の目的を簡潔に述べれば、動的評価尺度

を用いて「評価尺度判断過程というブラックボックスを開けてみる」ということに尽きる。カーソル軌跡分析は最近増加しており、意思決定の認知コンポーネントを捕捉可能とされている[2]。しかし、評価尺度の文脈でこの方法論を用いるのは2014年時点では本研究のみである。

また取得されるデータは「カーソルの軌跡」となるが、その解析には既存のデータ解析手法が使用できないため、新たなデータ解析手法の創案も研究の主要部分になる。

3. 研究の方法

本研究ではPCのカーソル軌跡をトレースすることにより(図1), 評価判断過程を解明しようと試みる。このために様々な評価課題が行われた(表1)。総ての課題において、図1のフォームが、カーソル軌跡、最終的な評価判断、およびRTを記録するために使用された。ただし、評価尺度の上部ラベルは、課題の性質により、図1のように1から5までの数値ラベルが用いられる場合と、カテゴリー1の位置に「いいえ」、5の位置に「はい」、3の位置に「わからない」というラベルが置かれる場合があった(後者の場合、カテゴリー2および4の位置にラベルはなかった)。

評価実験は、被験者ペースで行われた。試行の進行は以下の通りである。まず、被験者がフォームの中心に置かれた「スタート」ボタンをクリックすると、ボタンが消去され、フォームの中央の表示ボックスに問題が提示される。被験者は、可能な限り早く「正しい」あるいは「最適な」カテゴリーボタンをクリックするように求められ、最初と最後のクリック間のカーソル軌跡と時間が記録された。すると画面中央に再び「スタート」ボタンが現れ、被験者がこのボタンを押すと次の問題が呼び出された。このようにして各被験者はセッション終了まで課題に解答した。実験プログラムは、Microsoft ExcelのVBAで書かれ、実験は、Excel上で実行された。

使用された課題を表1に示す。これらの課題は、異なる認知領域および自己評価領域から選ばれたものである。被験者にはこれらの課題の内のいくつかを組み合わされて与えられた。ただし課題0(ベンチマーク課題)はすべての被験者に最初に与えられた。被験者一人あたりの試行数は組み合わせに応じて異なっていた。

Table 1: Summary of 16 Tasks

	Items
0) † Benchmark	2
1) Self esteem (Rosenberg)	10
2) Maximization (Schwartz)	8
3) Regret (Schwartz)	8
4) Big 5 Extraversion	5
5) Big 5 Neuroticism	5
6) Big 5 Conscientiousness	5
7) Big 5 Agreeableness	5
8) Big 5 Openness to Experience	5
9) Indeterminacy	15
10) Social desirability (Marlowe-Crowne)	10
11) Life style (e.g., We can be happy without money)	5
12) Opinion (e.g., Japan has rather a bright future)	5
13) † Addition and subtraction (e.g., 3-2+1, 1+3-2.)	30
14) † Division (e.g., 17/13.)	10
15) Geography (France is larger than Japan in land area)	5

Note. † “1-2-3-4-5” Category label was used. Otherwise “No-Don’t know-Yes” label was used.

課題0はベンチマーク課題であり,1から5までのいずれかの番号がランダムに提示され,被験者は,フォーム上の対応するボタンを素早くクリックするように要求された.過去の研究[3]では,この課題での軌跡は典型的な目標指向(単純到達)運動となるため,最小の認知作業負荷を伴うベースライン課題として機能することが示されている.課題1-12は正解がない主観的評価課題で,特に課題1-10は標準化心理尺度である.

課題13 - 15は正解がある正答志向型課題である.課題13では,刺激は3つの数字の加算と減算(e.g., 3-2+1, 1+3-2)からなる簡単な計算式であり,被験者は,迅速に正しい答えの番号をクリックすることを求められた.課題14は,割り算課題であり,被験者は素数を別の数で割り(例えば,13/17)四捨五入して整数化し,対応する数字をクリックするのを要求された.

被験者 すべての早稲田大学の学生だった.

4. 研究成果

本研究では「軌跡データ」の解析が必須であるが,既存の標準的な手続きが存在しないため,まず解析法を確立しなければならない.そこで本研究で用いた解析法(多くは新手法)について解説する.
解析法1: 評価反応時間 反応時間は認知心理学等では標準的な測度であるが,評価判断の反応時間を調べた研究は多くはない.後述するように,本研究では逆U字型あるいは逆J字型の評価反応時間パターンが研究対象となった.図2参照.

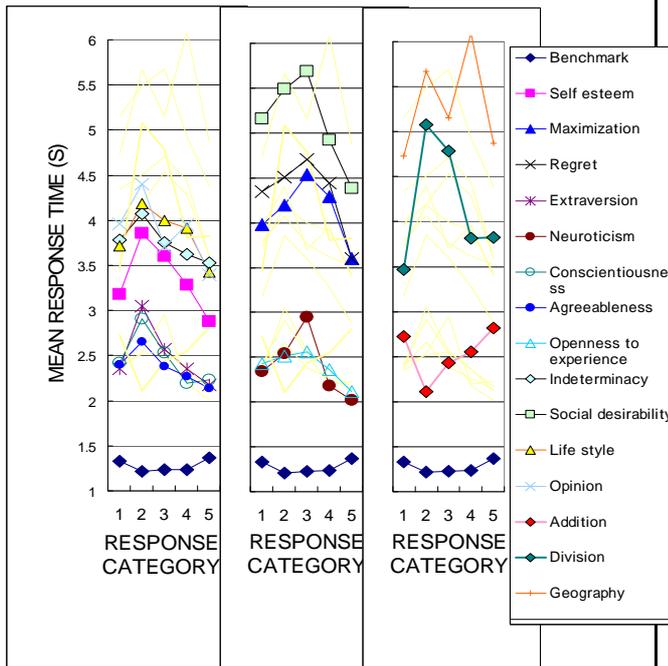


図2 各課題における平均反応時間 横軸は評価カテゴリー 左側のパネルは第2カテゴリーにピークがある逆J字型,中央のパネルは第3カテゴリーにピークがある逆J字型,右のパネルはそれ以外

解析法2: 接線速度(Tangential Velocity)

軌跡を時間インデックス付き2次元ベクトル $(x(t), y(t))$, $0 < t < RT$ と定義する.本研究では原軌跡の

長さはすべて異なるので,まずRTを256で割って時間領域でのステップサイズを定義し,次に軌跡を「標準化」するために,線形補間により位置 (x_i, y_i) , $i = 0, 255$ を推定した.さらに,時刻*i*での軌跡の接線速度(tangential velocity)を:

$$TV_i = \frac{\sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2}}{RT / 256} = \frac{\text{Traveled Distance}}{\text{time step}}$$

とする.接線速度は,軌跡の速度であり,課題特性を反映することになる.

平均接線速度 時間及び,最終的にクリックされた評価カテゴリーの,関数としての平均接線速度曲線(図3)を計算する.例えば,図3aは,ベンチマーク課題における,最終的にクリックされた評価カテゴリーごとの平均接線速度曲線である.括弧内の数字は平均RTである.この課題でのカーソル移動は単純到達運動であり,従って初期の弾道相と終末期の補正制御相が存在することになる.弾道運動は,迅速かつベル形状であり,フィードフォワード制御されるので,初期速度と方向は移動の開始前に決定され,運動コンポーネントと認知コンポーネント間の解離を示唆する[1].

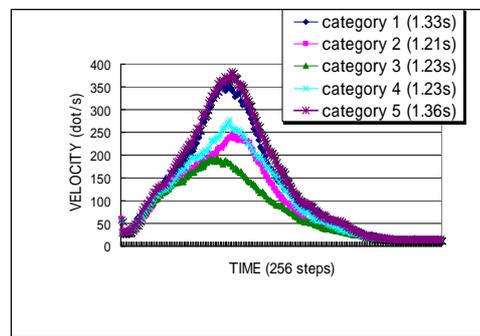


図3a ベンチマーク課題(課題0)

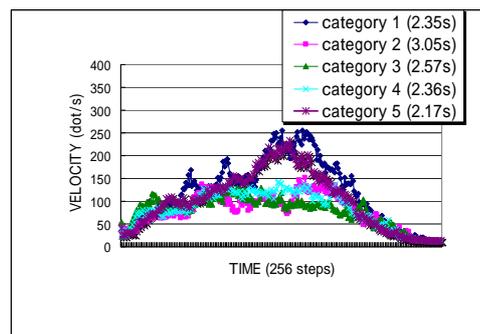


図3b 外向性課題(課題5)

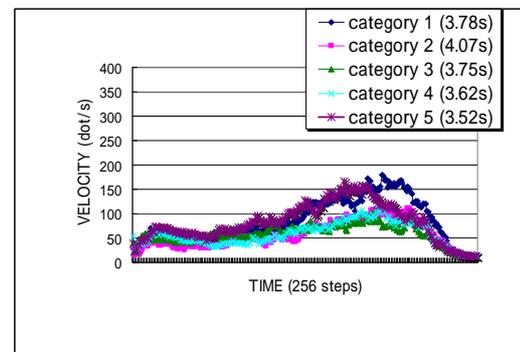


図3c 不決断傾向課題(課題9)

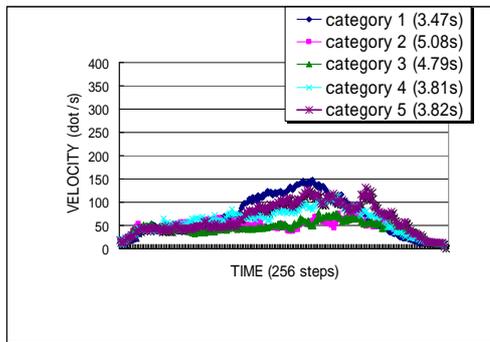


図3d 割り算課題 (課題14)

対照的に、非弾道運動は被験者の意思決定における葛藤と動揺を反映すると思料される。以上より、弾道運動を、「ためらいがない」ことを示すマーカーとして使用することができる。図3a中の曲線は、最初大きなBell型の弾道運動(左のピーク)と、後期の小さな補正運動(右の小さな振動)を見せている。最高速度が200ドット/Sを超え、かつBell型、である曲線は弾道運動を示すと仮に定義する。

この定義に従えば、図3aでカテゴリー1,2,4,5に到着した軌跡は弾道運動であり、カテゴリー3に最終到着した軌跡は非弾道運動だったことになる。また、単純到達運動は割り算課題や不決断傾向課題には存在せず、また「外向性」課題中(図3b)でカテゴリー1および5を目指した軌跡は弾道運動であったことになる。

図3cおよび3dで、すべての曲線が初頭に緩慢に移動しているが、これは問題文の読みや計算に関係していると考えられる。

カーソル移動が速ければ、早く目標に到達するのは当然だから、以上の分析はRTの説明という観点からは、トートロジーに見えるかもしれない。確かに、カーソル運動が常に最適で経済的であればこの批判は有効である。しかし、接線速度には、軌跡が単純に早く動く時も、ある場所で(無駄に)高速で彷徨する時も、共に大きくなるという特性があるので、より詳細な分析が必要となる。

解析法3:接線速度の変動

平均曲線の分析のみでは不十分な場合、速度の変動性に焦点を当てた分析を行う必要が出てくる。そこで以下の量を定義する。

\bar{X}_i 時点*i*での平均接線速度

s_i 時点*i*での接線速度の標準偏差

$CV_i \equiv s_i / \bar{X}_i$ 時点*i*での平均接線速度変動指数(ピアソンの変動指数)。

大きな量は大きく変動する自然な傾向があるので、この指数は、平均の大きさを考慮した、軌跡変動性を示す。さらに平均CV指数は、ある課題であるカテゴリーを最終的に選択した際の、被験者の内的迷いを総合的に示していると考えられる。

解析法4:Stroke Analysis

Shiina[4]は、カーソルの軌跡は、strokeと命名されたパルス状の変動(図4)を含むことが多く、ベンチマーク課題では速度曲線の50%以上がシングルピークであったが、この比率は他の課題でははるかに低いのを報告した(図5)。このstrokeの

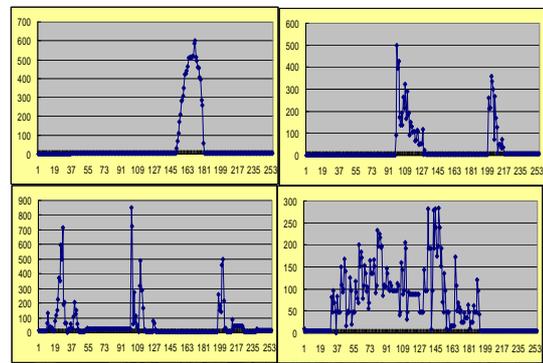


図4 割り算課題で観測された接線速度の例。横軸は時間、縦軸は接線速度。ストローク(パルス)が観測できる

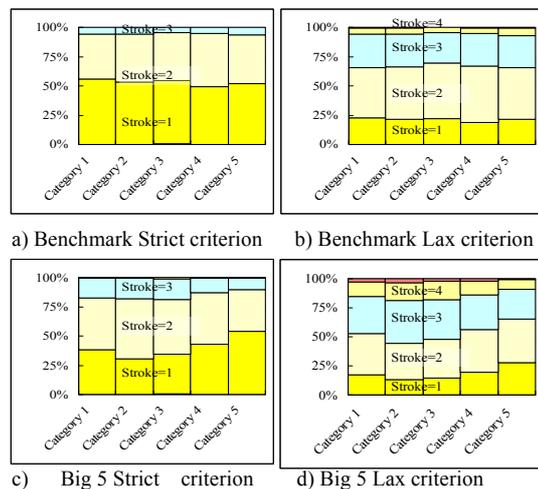


Figure 5: The ratios of stroke numbers for each rating category. Very small ratios (more than 4 or 5 strokes) are invisible in this figure. a) Benchmark Task under strict criterion (12,075 total responses), b) Benchmark Task under lax criterion. In Benchmark Task, Categories 1, 2, 3, 4, and 5 were clicked 2417, 2414, 2415, 2414, and 2415 times, respectively. c) Big 5 Task under strict criterion (14,490 total responses), d) Big 5 Task under lax criterion. Categories 1, 2, 3, 4, and 5 were clicked 2642, 1987, 2289, 3367, and 4205 times, respectively.

分析も有効である。たとえば、図3の平均速度パターンの別解釈は、異なる発生頻度のパルス(ストローク)の融合したものということになるし、CV値の別解釈は、時刻*i*におけるパルスの発生頻度の不安定性の測度であるということになる。弾道的単峰ストローク速度曲線は、内的ためらいの存在を否定するものであり、逆に複数のストロークを持つ軌跡は弾道的ではないという意味で、内的ためらいを示唆することになる。

心理学研究における応用例

以上解析法1-4の解説を行った。動的評定尺度法とこれらの解析法を用いて行った心理学研究の一例を紹介し、動的評定尺度法の有効性を示す。

自己参照型評定判断における逆U字型反応パターンの研究

自己参照型の判断や決定は神経認知科学や社会心理学の重要なトピックであり、代表的な研究パラダイムは、質問紙を用いた性格特性の自己評定である。さて、リッカート型尺度に解答する際、

尺度の両極よりも中ほどのカテゴリに対する RT (反応時間) が長いのが見出されている (図 2 参照)。そこで、「評定判断」でのカーソル軌跡を解析することで、「両極の判断は迷いなしで行われるが、中間部分の判断は conflict resolution を伴うので、結果として逆 U 字効果が生み出される」という説に対して、直接的な証拠が提出できるかどうか試みる。

図 2 は全課題における、最終的にクリックされた評定カテゴリ別の平均 RT を示す。図 2 では 13 の曲線が逆 U 字型であり、3 つの曲線はそうでないのがわかる。逆 U 字パターンの多くは、主観的評定課題 (課題 1 から 12 まで) であり、逆 U 字型でないのは正答志向型であるベンチマーク、加算、地理課題である。

逆 J 字パターンの出現 さらに図 2 を良く見ると、YES の回答 (カテゴリ 5 の反応) が NO の回答 (カテゴリ 1 の反応) よりも速く、逆 U というより逆 J の形状をしている曲線が観察できる。13 個の逆 U 字カーブのうち 12 のカーブで逆 J 字パターンを観察することができ、例外は正答志向型である課題 14 (割り算) である。したがって、逆 J パターンは、すべての主観的課題 (正答のない課題) で生起しており、さらに心理尺度 (課題 1-10) はすべて逆 J パターンになっている。平均接線速度を用いた分析の概要は解析法 2 で述べた。本課題の場合、接線速度で検出される弾道運動のみでは、逆 U 字 RT 型効果の粗い説明はできて、逆 J 字型効果の説明は難しい。

平均接線速度変動指数 CV_i を 256 のすべての時点で (時間) 平均すると、図 2 内の各点に対応する軌跡変動性の推定値を総計 80 個 (16 課題 \times 5 カテゴリ) 計算できる。この平均 CV 指数は、ある課題であるカテゴリを選択した際の、被験者の内的迷いを示していると考えられる。この指数及びその他の変数を用いて図 2 の平均 RT を目的変数とする、2 次回帰モデルの当てはめを行い、

$$RT = 0.065 \times (\text{Number of letters of a question}) + 0.285 \times (\text{category number}) - 0.072 \times (\text{category number})^2 + 1.359 \times CV - 0.434$$

を得た。このモデルは主観的評定課題の 12 個の逆 J 字型曲線 (図 2) をうまく説明できることが分かった ($R^2 = .94$, $F(4, 55) = 201.79$, $p < .0001$, 図 6)。

この式では、問題を構成する文字の数が、RT の全体的なレベルを決定し、二次式の部分は、逆 U 字形状を生成、CV の項は、カテゴリ 5 の CV が小さく、中央カテゴリで CV が大きいことにより、逆 U 字を逆 J 字に変換し、全体として逆 J 字パターンを形成する。

さて、弾道の単峰ストローク速度曲線は、内的ためらいの存在を否定するものであり、逆に複数のストロークを持つ軌跡は弾道的ではないという意味で、内的ためらいを示唆するので、単峰ピークの軌跡の割合を、各課題の各カテゴリ (図 7, [4]) について計算した。図はやや雑然としているが、平均化することによって明確な J 字型平均曲線が浮上している (大きな赤四角)。この結果は、内的葛藤が少なければ単峰弾道運動が増加すると

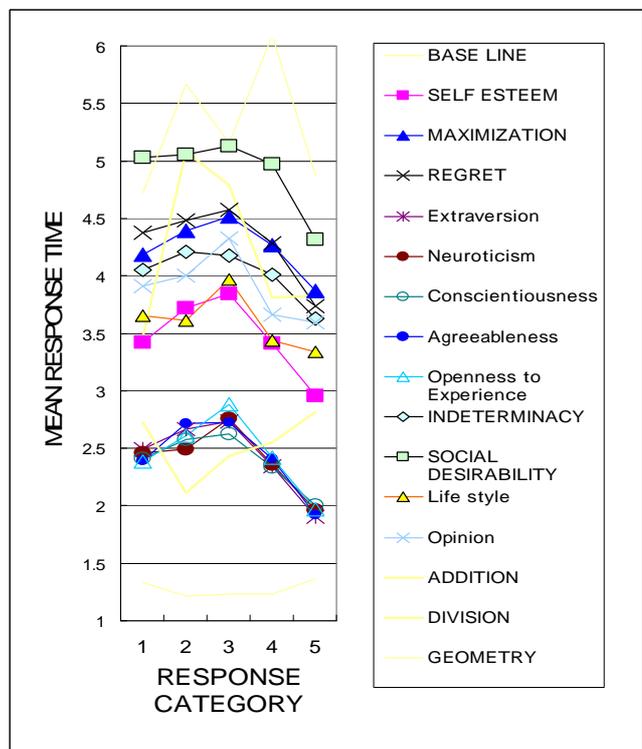


図 6 2 次式あてはめで予想された図 2 のデータ

という見方を強く支持しており、弾道運動が短い RT を誘発し、逆 J 字曲線を出現させたと考えられる。

まとめ 1) 逆 J 字型 RT 効果は、主観的課題で生起する。2) 接線速度解析は、弾道運動が逆 J 字効果の原因であろうことを示唆した。3) 軌跡変動指数 (CV) から、逆 J 字型 RT パターンを説明する二次回帰モデル (図 6) を生成できた。4) ストロークの分析は、非躊躇マーカーとして単峰性弾道軌跡が、短い RT および逆 J 字効果の原因であること示唆した。すなわち、被験者の内的葛藤が、複数のストロークと軌跡振動を引き起こし、逆 J 字効果をもたらす。5) この効果の究極の理由は不明であるが、次のセクションで可能な説明を示す。

一貫性仮説による逆 J 字型 RT 効果の説明

尺度の両極の評定は記憶に依存するのに対して、中間の評定は処理プロセスに依存するため、逆 J(U) 字効果が起こると想定できる。すなわち「はい」の回答は意味記憶からの直接検索により最も高速になる一方、中間のカテゴリは、エピソード記憶を使用して自己と特性との一致度を計算し、高一致度も、低一致度も得られなかった場合に使用されるので、最も遅くなる。「いいえ」の反応は、意味記憶の直接検索と、一貫性計算の両方が働いているので、中間の速さになる。

この仮説によれば、評定判断から生じる数値は評定カテゴリによって質的に異なることになる。逆 J(U) 字効果の出現は、記憶研究者にとっては驚くべきことではないが、精神測定理論にとっては説明困難である。

ただし、一貫性仮説による説明には問題がある。なぜなら刺激文が「あなたは勇敢である」の時、意味記憶上に「私はまあまあ勇敢です」という命題が直接貯蔵されているなら、素早く中程度の評定カテゴリを選択できるはずだからである。したがって、一貫性に基づく解釈を検証するには、記憶に貯蔵されている中間的命題が少数なのを示すか、または中程度の評定は、決定時の葛藤によって阻害される傾向があることを、直接的に証明する必要がある。動的評定尺度の貢献は、中央のカテゴリに向かう軌跡の変動性が高いのに

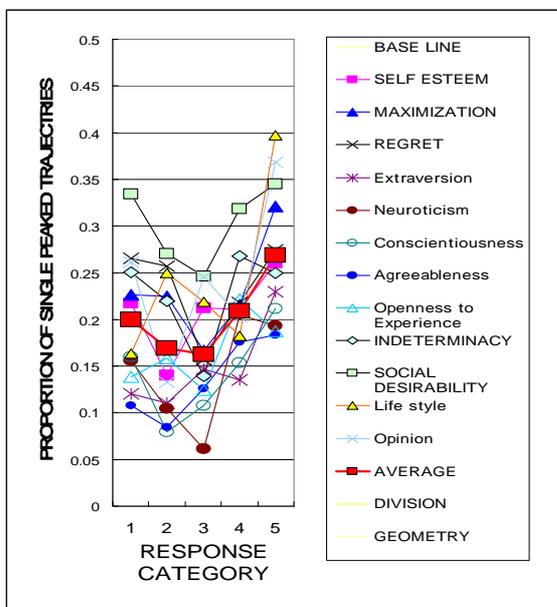


図7 課題と評定カテゴリごとの単峰軌跡割合対して、両極のカテゴリに向かう軌跡は安定しているため、逆U字効果を生み出すのだという予想に対する、行動的証拠を提示していることである。以上のように、動的評定尺度は心理学に有用な知見を与えることが明らかになった。

心理測定学に対する提言

本研究の結果は、評定尺度法で生成される数値は、異なる内的過程を経て得られる場合があると考えられるので、直接比較可能でない可能性があるのを示唆する。

最後に、本研究を行う際、評定尺度法データから計算するピアソンの相関係数に特異な性質があるのを発見したことを付記しておく。

主要文献 [1]Elliott, D., Helsen, W. F., & Chua, R. (2001). A century later: Woodworth's (1899) two-component model of goal-directed aiming. *Psychological Bulletin*, 127, 342-357. [2]Freeman, J. B., Dale, R., & Farmer, T. A. (2011). Hand in motion reveals mind in motion. *Frontiers in Psychology*, 2, 59. [3] 椎名乾平 (2008). カーソルの軌跡を用いて評定判断過程を分析する *学術研究 - 教育心理学編 -*, 56, 1-9. [4] Shiina, K. (2011). Tracing the Process of Rating Decisions through Cursor Movements. *Proceedings of the 33rd Annual Conference of the Cognitive Science Society*

5. 主な発表論文等

〔雑誌論文〕(計7件)

Kenpei SHIINA 2011 Trajectory-Variability and Inverted-U Effects in Rating Decisions. **European Perspectives on Cognitive Science : Proceedings of the European Conference on Cognitive Science** 査読有
<http://nbu.bg/cogs/eurocogsci2011/proceedings/pdfs/EuroCogSci-paper187.pdf>

Kenpei SHIINA 2011 Tracing the Process of Rating Decisions through Cursor Movements. **Proceedings of the 33rd Annual Conference of the Cognitive Science Society.** 査読有
<http://csjarchive.cogsci.rpi.edu/Proceedings/2011/papers/0154/paper0154.pdf>

大内善広・上田卓司・椎名乾平・岡田いずみ 2012

カテゴリ数の異なる順序カテゴリ尺度同士の相関係数の性質 (*学術研究 - 人文科学・社会科学編 -*, 60, 93-103).

椎名乾平 2012 評定尺度法とカーソル運動 *知能と情報*, 24, 858-870. 査読有

https://www.jstage.jst.go.jp/article/jsoft/24/4/24_858/pdf

椎名乾平 2013 七つの正規分布 *心理学評論* 56, 7-34. 査読有

椎名乾平 2014.2 PC カーソルの軌跡で心理状態を推測する ~ 自己参照型評定の場合 ~ *信学技報*, vol. 113, no. 426, HCS2013-76, 37-42.

Kenpei SHIINA 2014 Inverted-U and Inverted-J Effects in Self-Referenced Decisions. **Proceedings of the 36th Annual Conference of the Cognitive Science Society.** 査読有

〔学会発表〕(計8件)

椎名乾平 2011.9 カーソルは心の窓になるか? 日本認知科学会第28回大会 東京大学 (査読有)

Kenpei Shiina, Saori Kubo, Yoshihiro Ouchi, Takashi Ueda 2012.7 An Aberrant Behavior of Pearson's Correlation Coefficient when Two Variables have Unequal Number of Categories 77th International Meeting of the Psychometric Society, Lincoln:Nebraska (査読有)

Kenpei Shiina 2012.11 A Cursor-tracing Paradigm For Rating Decisions: Another window on mind? Society for Judgment and Decision Making 33rd Annual Conference. Minneapolis : Minnesota. (査読有)

椎名乾平 2012.11 性格評定における逆U字型反応時間について(3)-実は逆J字型反応時間なのかもしれない- 日本教育心理学会第54回大会 琉球大学

Kenpei Shiina, Saori Kubo, Yoshihiro Ouchi, Takashi Ueda 2013.7 A paradoxical fall of r and reliability by increasing the number of rating categories . 78th International Meeting of the Psychometric Society, Arnhem : The Netherland (査読有)

椎名乾平, 久保沙織, 大内善広, 上田卓司 2013.9 信頼性係数のパラドックス-尺度のカテゴリ数を増やすと信頼性係数が減ることがある 日本心理学会第77回大会 北海道医療大学

日本心理学会シンポジウム 2013.9 評定尺度はどのような測定装置なのか? - 評定尺度法を問いなおす - 企画代表者 椎名乾平 話題提供者 久保沙織(早稲田大学)妻藤 真彦(美作大学)村上隆(中京大学)日本心理学会第77回大会 北海道医療大学

Kenpei Shiina, Saori Kubo, Takashi Ueda 2014 A Queer Behavior of Pearson's Correlation Coefficient . ICAP2014 Paris (査読有)

〔図書〕(計 件)〔産業財産権〕なし

〔その他〕なし

6. 研究組織

(1)研究代表者 椎名乾平 (SHIINA Kenpei)
 早稲田大学 教育・総合科学学術院 教授
 研究者番号: 60187317

(2)研究分担者 なし (3)連携研究者 なし