

意思決定におけるバイアス矯正の研究動向

立教大学大学院現代心理学研究科 相馬 正史

立教大学現代心理学部 都築 誉史

Recent trends in research on decision making with respect to debiasing

Masashi Soma (Graduate school of Contemporary Psychology, Rikkyo University), and
Takashi Tsuzuki (College of Contemporary Psychology, Rikkyo University)

Studies on decision making have revealed that we make decisions not only through reason, but also through simple heuristic judgments. Several researchers have pointed out that heuristic judgments tend to lead to errors in decision making. Studies suggest that decision making can be improved by employing various strategies. Debiasing strategies are mainly classified into three categories: motivational, cognitive, and technique strategies. The technique strategies have been studied in the field of economics, while cognitive strategies have been studied in psychology. In this paper, we present an overview of studies on debiasing in psychology research. Finally, we discuss the future of debiasing research.

Key words: heuristic, prescriptive approach, debiasing, consider an opposite

判断と意思決定に関する理論的研究では、最適な解を求め、合理的な決断をとるための技法や理論が提案されている。しかし、実証的な意思決定研究によれば、実際の人々は、非合理で簡便な判断をとる傾向がある。これはヒューリスティクス (heuristics) と呼ばれているが、多くの場合、ヒューリスティクスが判断と意思決定の誤りにつながることが示されてきた。判断の誤りは意思決定バイアス (making decision bias) と呼ばれているが、意思決定バイアスを改善することは重要な研究課題である。本論文では、初めに、古典的な意思決定研究における二つのアプローチ、つまり、(a) 規範的アプローチ (normative approach: 合理的な判断を理論化する研究) と、(b) 記述的アプローチ (descriptive approach: 実際の人々の判断を解明する研究) を紹介する。そして、その古典的研究で論争とされてきた、“本当にヒューリスティクスは誤りなのか?”という論点と、“いかにして判断の誤りをなくせるか?”という論点に

ついてふれる。次に、論争を契機に台頭した、処方的アプローチ (prescriptive approach: 人々の判断を支援する研究) と、バイアス矯正の研究アプローチを概観する。また、意思決定バイアスを総括して説明するモデルや分類について説明し、バイアス矯正方略の研究を紹介する。さらに、バイアス矯正で主に研究されている、考察方略の研究について解説する。最後に、バイアス矯正の今後の展望を述べて結語とする。

規範的アプローチ、記述的アプローチ、 処方的アプローチ

意思決定研究では、主に三つの研究主流がある。一つは、規範的アプローチと呼ばれており、合理的経済人 (homo economicus) を仮定した上で、意思決定の研究が行われてきた。合理的経済人とは、自己利益を最大化しようと行動する人間であり、多くの経済学モデルの基本的な前提となっている。経済学では、この合理的経済人を仮定した

上で、合理的モデルの理論生成、決定分析等による意思決定の改善に、研究の重きをおいてきた。

しかし、数々の実証的な心理学研究によれば、人々は合理的な意思決定を行わないことが確認されている。以下では、心理学研究で見いだされた、非合理的な判断パターンを紹介する (Kahneman, 2011)。

利用可能性ヒューリスティック (availability heuristic) 想起しやすい事柄を優先して評価しやすい意思決定方略のことをいう。あるリスク事例を思い浮かべることが容易であれば、その事例の発生率が高いと判断する。たとえば、航空機の墜落事故の発生率と他の交通事故の発生率とを比べる時、墜落事故はインパクトが強く、鮮明に思い浮かぶため、他の交通事故の生起確率よりも過大に評価されやすいことが知られている。

代表性ヒューリスティック (representative heuristic) 特定のカテゴリーに典型的と思われる事項の確率を過大に評価しやすい意思決定方略をさす。たとえば、“リンダは31歳の独身女性である。非常に知的で、はっきりものを言う。大学時代は哲学を専攻し、学生の頃は社会主義と差別問題に関する活動に深く関わり、核兵器反対のデモにも参加したことがある”という文章を読んだ後では、“彼女は銀行員である”と回答する人よりも、“彼女は銀行員で、女性運動で活動している”と回答する人の方が多い。

係留と調整ヒューリスティック (anchoring and adjustment heuristic) 最初に与えられた情報を基準として、それに調整を加えることで判断し、最初の情報に現れた特定の特徴を極端に重視しやすい意思決定方略をさす。例えば、“国連に属している国のうち、アフリカ大陸にある国家の割合はいくらか”という質問をする前に、“65%よりも大きいと思うか小さいと思うか”と質問された人は、“10%よりも大きいと思うか小さいと思うか”と質問された人よりも、多い割合を回答する。

以上の意思決定ヒューリスティックスの例により、我々は多様な意思決定場面で、非合理的な判断

を下す可能性があることが示唆される。人々が実際にどのような意思決定をしているかについて研究することを記述的アプローチと呼び、非合理的な判断パターンは認知バイアス (cognitive bias) や意思決定バイアス (decision making bias) として扱われている。

規範的アプローチでは、最善で合理的な判断を追求する一方で、記述的アプローチでは、実際の人々の非合理的な判断を研究している。この両者の乖離は、(a) “本当にヒューリスティックスは誤りなのか”という問題と (b) “いかにして判断の誤りをなくすことができるのか”という大きな課題を研究者達に突き付けた。以下では、そうした問題に関する議論を述べる。

本当にヒューリスティックスは誤りなのか

規範的アプローチと記述的アプローチの乖離は、“本当に矛盾が存在するのか”という問題や、“本当にヒューリスティックスは誤りなのか”という問題を研究者たちに突き付けた。

Larrick (2004) によると、多くの経済学者や哲学者は、“人々は合理的に行動するため、判断の失敗はランダムまたは不規則なものである。規範的アプローチと記述的アプローチの矛盾は、不適切な実験測定法に基づくものである”と主張してきた。しかし、Stanovich (1999) は、彼らを楽観的であると批判した。初期のバイアス矯正研究は、さまざまな正確な測定法を用い、意思決定バイアスの頑健さを示すことで、古典的な机上の主張を大々的に批判した (Fischhoff, 1982)。多くの研究者は、ヒューリスティックスは重大な判断の欠陥とみなし、意思決定バイアスの改善を試みてきた (Arkes, 1991; Fischhoff, 1982; Kahneman, 2011; Larrick, 2004)。

しかし、人の心理メカニズムの多くは、進化生物学的な意味で、生物学的適応であると考えられる進化心理学者は、ヒューリスティックスを判断の誤りと捉えることに批判的である。彼らは、ヒューリスティックスが、人間が適応的に生活するために進化したものであると述べている (Gigerenzer,

2007)。ドイツの意思決定研究者である Gigerenzer (2007) は、ヒューリスティックスを意思決定バイアスとみなす立場を一貫して批判している。彼は、俊敏で簡潔な判断が、合理的判断と同等、またはそれ以上の威力を発揮することを、数々の研究によって示してきた。例えば、Gigerenzer & Goldstein (1996) は、“最もよい手がかりを選択し、あとは無視せよ”という“最善を尽くす”ヒューリスティックス (take the best heuristic) の存在を指摘している。そして、“二つの町のうち、どちらが大きいのか”というような現実的な状況での推論において、驚くことにこのヒューリスティックスが正確であることが示されている。

“本当にヒューリスティックスは誤りなのか”という問題は、意思決定研究で論争を巻き起こしている。しかし、この論争には、ある研究領域で検討されているヒューリスティックスを、他の研究領域で検討されているものへと、過度に一般化することに由来する誤りがあると考えられる。Myers (2004) は、ヒューリスティック研究を展望し、ヒューリスティックスが有効であるか否かは、環境や研究課題によってそれぞれ異なると主張している。現在では、ヒューリスティックスは意思決定の誤りであることが、多くの研究者に認められている。本論文でも、ヒューリスティックスが意思決定バイアスであり、改善すべきものと想定し、話をすすめる。

バイアス矯正の研究流派

規範アプローチと記述アプローチの乖離は、さらに“いかにして判断の誤りをなくすことができるのか”という問題を想起させた。言い換えると、“意思決定バイアスを矯正できるのか”という問題である。この問題に直面することにより、研究者たちは非合理的判断を改善し、支援する研究に関心を向けた。意思決定バイアスを改善する手続きを、バイアス矯正 (debiasing)、意思決定の支援を行う研究の流れを処方的アプローチ (prescriptive approach) と呼ぶ (Bell, Raiffa, & Tversky, 1988)。この処方的アプローチの提唱以

来、心理学研究では、さまざまなバイアス矯正方略が検討されるようになった。

バイアス矯正の方略が意思決定をどの程度改善するかは、議論の中心である。Stanovich (1999) は、バイアス矯正の研究流派を概観し、(a) 改善主義 (Meliorists)、(b) 弁明主義 (Apologists)、(c) 技法主義 (Technologists) の三つに分類している。バイアス矯正の古典的な研究者である Fischhoff (1982) は改善主義 (Meliorists) の立場をとっており、バイアスを除去するにあたって、“多くの思慮は、合理的な考えに至らないが、経験や訓練を通して改善される”と仮定している (Table 1 を参照)。

改善主義者は、ヒューリスティックスが意思決定の歪みをもたらすため、優れた方略を学習することで、バイアスを改善しようと試みている。それとは反対に、“人工的な課題や抽象的な課題 (例えば、確率推論でのベイズ推定) を用いるために、多くの課題で人々は合理・規範な水準を獲得できない”と主張し、非現実な課題設定を排除しようとする立場を“弁明主義 (Apologists)”と呼ぶ。彼らは Gigerenzer のように、ヒューリスティックスを判断の誤りと見なすことに批判的である。弁明主義者は、ヒューリスティックスが自然な環境・状況において進化論的に適応的であり、自然な状況において最善な結果を得ると主張している。その例として、推論課題の文面を“確率”で表現した場合よりも“頻度”で表現した方が、推論が正確になることを示した研究がある。Sedlmeier (1999) は、推論において、実験参加者に確率ではなく頻度として情報を提示したり、確率から頻度に認識を変える訓練をすることによって、推論判断が改善されることを示した。つまり、環境や課題が抽象的なものではなく、より自然な状況において、ヒューリスティックスは優れたものになること指摘している。また、思考の方略を、抽象的なものではなく、日常生活など具体的なものに置き換えて課題に取り組むことで、パフォーマンスの向上につながることも指摘している。

バイアス矯正の三つ目の立場は、意思決定者が

Table 1
バイアス矯正に対する見方

立場	主張
改善主義	多くの熟慮は合理的な考えに至らないが、経験や教育を通して、バイアスは改善
弁明主義	自然な状況において、ヒューリスティクスは優れるため、自然な状況づくりをすることで、バイアスは改善
技法主義	道具・ツールを使用することで、バイアスは改善

外的技法を使用することによって、バイアスを改善するものである。このアプローチは、“技法主義 (Meliorists)” と呼ばれる。技法主義は、規範的アプローチの理論・技法がベースとなっており、決定ツール（例えば、意思決定支援システム、決定分析、統計モデル）を使用することで、判断が規範基準に近づき、最善な意思決定を得られると仮定している。技法主義の研究は、主に経済学で行われているが、心理学研究においても検討されている（相馬, 2012）。

以上のように、これらの種類のアプローチはバイアス矯正に異なる見解を持っているが、バイアス矯正には介入が必要であると見なす点は共通点している。各アプローチから数々のバイアス矯正方略が検討されてきたが、そうした方略については本論文の後半で証明する。

意思決定バイアスの種類・モデル

通常、心理現象のメカニズムや因果関係は、特定の知覚レベルによって説明がなされ、モデルの実証的研究がなされる。しかし、個々の心理現象を総括し、大きくまとまった一つの心理現象として説明がなされることもある。意思決定研究でも、利用可能性ヒューリスティックや代表性ヒューリスティックなどの意思決定バイアスを、一つの大きくまとまった心理プロセスとして把握し、モデ

ル化の試みが行われている。意思決定バイアスのモデル化や分類はいくつか存在するが、ここでは、研究で頻繁に引用されている、Arkes (1991) の意思決定バイアスの分類と、Wilson & Brekke (1994) の意思決定バイアスモデルを紹介する。

Arkes による意思決定バイアスの分類

Arkes (1991) は、意思決定バイアスが、二重過程理論 (dual process theory) で説明されると主張した。二重過程理論とは、人の判断や思考が、二つの思考モードで説明できるという理論をさす (Epstein, 1994; Kahneman, 2003, 2011 : Table 2 を参照)。一つ目は、自動的、無意識的、領域固有的、迅速な処理であり、“システム1” と呼ばれる。二つ目は、論理的、分析的、領域一般的、遅い処理であり、“システム2” と呼ばれる。システム1は無意識的で素早く、自動的な情報処理が行われるとされる。それに対して、システム2では、情報に基づいて意識的で時間をかけた慎重な処理が行われるとされる。システム1は、さまざまな意思決定バイアスと関連があることが、数多くの心理学研究で示されてきた (Arkes, 1991)。これらのヒューリスティクスが生じる要因としては、人が世界を認知する際にはなるべく認知資源を節約しようとする傾向があるためと考えられる (Baumeister, Bratslavsky, Muraven, & Tice, 1998)。

また、Arkes (1991) は、進化論の観点から、

Table 2
二重過程理論のまとめ

システム 1	システム 2
自動的	論理的
無意識的	分析的
領域固有	領域一般
迅速な処理	遅い処理

ヒューリスティックスを含め、成功した適応行動には、生存における利益とコストが伴うと主張している。例として、交感神経系の生理反応をあげる。心拍の増加、発汗、瞳孔の拡大といった生理反応は、危険な状況では、逃げる、攻撃をとるなど素早く対応することができるため、適応性が高い反応であり、その意味では利益を有する。しかし、慢性的な生理反応は、そのコストとして、重大な生理学的なストレス症候群を引き起こすことがわかっている。Arkes (1991) はこの二重過程理論と利益・コストの観点から、意思決定バイアスを大きく三つに分類している (Table 3を参照)。彼の分類の初めの二つは自動的・無意識によるエラー (システム1と関連) である。分類の三つ目は、意識的、熟考によるエラー (システム2と関連) に基づいている。

心理物理学に基づくエラー (psychophysically-based error, system 1) これは、損益と損益に対する評価が非線形な対応関係であるエラーをさす。このエラーの例として、利益・損失の非対称性 (基準値よりプラスの領域ではリスク回避的になり、基準値よりマイナスの領域では、リスク追求的になる) をあげることができる。例えば、5千円と1万円のどちらかを得るかという問題においては敏感であるが、10万円を得るか10万5千円を得るかという問題については、どちらを得ても大して差はないと思う傾向をさす。

連合に基づくエラー (association-based error, system 1) これは、記憶情報へのアクセスに基づき、自動的に引き起こされる判断のエラーをい

う。このエラーの例として、利用可能性ヒューリスティックスをあげることができる。刺激によって引き起こされる心的表象が、知覚・感覚的に、連合する認知を活性化し、連合しない認知を抑制する場合、連合に基づくエラーが発生する。例えば、乗り物事故を想像する場合、飛行機事故は鮮明に連想され、船の事故は連想されにくいことが多い。連合に基づくエラーは、判断において、基本となる情報の処理を偏らせてしまう。

方略に基づくエラー (strategy-based error, system 2) これは劣った決定ルールや方略を使用することによって引き起こされるエラーをさす。方略に基づくエラーは、最善の判断を求める決定方略ではなく、最善に近く、満足のいく判断を求める決定方略を用いる場合に生じる。方略に基づくエラーの例として、辞書編纂型ルール (lexicographic rule) があげられる。辞書編纂型ルールとは、最も重要な属性について、もっとも高い評価を持つ代替案が選択されるが、同順位の代替案があった際は、二番目に重要な属性について最も高い評価を持つ代替案が選択される決定方略をさす。Arkes (1991) は、このエラーは、システム2と関連したバイアスであると主張している。

以上がArkes (1991) による意思決定バイアスの分類である。Arkes (1991) はこの分類から、意思決定バイアスのタイプによって、バイアス矯正方略の効果に影響の違いがみられるため、各バイアスに適したバイアス矯正方略を使用すべきであると主張している。

Table 3
Arkes によるバイアスの種類とその特徴

	心理物理学に基づくエラー	連合に基づくエラー	方略に基づくエラー
記述・定義	損益に対する評価が非線形な対応関係であるエラー	記憶情報へのアクセスに基づき、自動的に引き起こされる判断のエラー	劣った決定ルールや方略を使用することによって引き起こされるエラー
バイアスの例	現状維持バイアス 結合評価・分離評価での選好逆転 確率荷重関数の曲線	アンカリング 確証バイアス 後知恵バイアス サブリミナルプライミング	肯定的ポジティブテスト方略 辞書編纂型ルール
短所・欠点	線形と想定される尺度に対する非線形変換 刺激に対して一貫しない判断	問題解決における機能的固着 確率判断における結果の過剰 焦点 判断における属性の偏った利用	判断における情報の無視 判断における不適当な情報の統合
システム	システム 1	システム 1	システム 2
サンクコストの例	損失状況の感覚減少効果により、サンクコストの後、加えられた損失はあまり痛みを感じない	初めの投資決定と一致した認知は、記憶の中からアクセスしやすい	過去のコストは、現在の会計においても維持される

バイアスのプロセスモデル

Wilson & Brekke (1994) は、意思決定バイアスやバイアス矯正の研究を概観し、バイアス矯正が一貫して意思決定バイアス（彼らはメンタル汚染 (mental contamination) と呼んでいる）を改善しないこと、バイアス矯正方略によって矯正の効果に違いがみられることに着目し、意思決定バイアスが段階的に発生するプロセスモデルを提唱している (Figure 1 を参照)。具体的には、意思決定バイアスにはいくつかの段階があり、それは (a) 望ましくない判断プロセスへの気づき、(b) バイアスを正す動機の有無、(c) バイアスの方向とバイアスの大きさへの気づき、(d) 反応のコントロール・調整といった段階からなる。各段階は順序立てて進むと仮定されている。さらに、意思決定バ

イアスに陥らないためには、段階ごとにいくつかの目標を達成しなければならない。

始めに、決定者は望ましくない判断プロセスに気づかなければならない。そのプロセスについて内省することにより、バイアスへの気づきが獲得される。そのため、“この問題には、意思決定バイアスが潜んでいる”と人々にほめかすことは、バイアスへの気づきを促し、バイアス矯正の成功につながる。二つ目に、このエラーを正す動機を持つ必要がある。“課題を慎重に行うように”と動機づけることは、決定者が熟考して判断を下すようになるため、バイアス矯正として有効である。三つ目に、バイアスの方向と大きさに気づかなければならない。たとえ、バイアスを改善しようとする動機をもったとしても、バイアスは一貫して

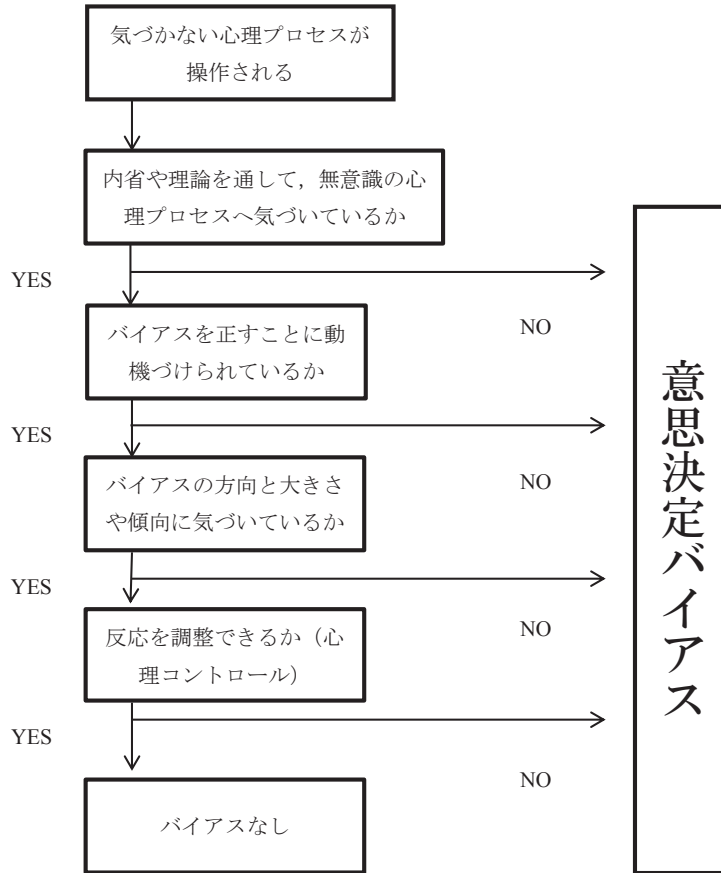


Figure 1. Wilson & Brekke (1994) の意思決定バイアスのプロセスモデル (一部改変)

改善されないことが報告されている (Camerer & Hogarth, 1999)。前述のとおり、係留・調節ヒューリスティックスにおいて、“国連に属している国のうち、アフリカ大陸にある国家の割合はいくらか”という質問をする前に、“65%よりも大きいと思うか小さいと思うか”と質問された人は、“10%よりも大きいと思うか小さいと思うか”と質問された人よりも、多い割合を報告する。このバイアスをなくすためには、特定の方略を用いて自身の反応をコントロールする必要がある。

以上のように、Wilson & Brekke (1994) のモデルは、バイアスに陥る段階的なプロセスを示しているが、それは同時に、バイアスを改善するために必要なプロセスとしてとらえ直すことができ

る。バイアスを改善するためには、(a) バイアスの存在を気づき、(b) バイアスを正す動機を持ち、(c) バイアスの方向や大きさを調整する必要がある。しかし、このモデルは実験的に検証されていないため、実証的な研究が今後の課題である。

バイアス矯正の方略

先に述べたように、バイアス矯正研究は改善主義、弁明主義、技法主義の三つの研究流派があり、さまざまな矯正方略が検討されてきた。バイアス矯正研究の先駆者である Fischhoff (1982) は、意思決定課題の種類により、有効なバイアス矯正を分類している。さらに Fischhoff (1982) は、一般的なバイアス矯正として、(a) バイアスが潜む局

面で警告を出すこと, (b) バイアスがどの方向に作用するかを示すこと, (c) バイアスを決定者に説明すること, (d) 訓練, フィードバック, コーチング, その他の広範囲なトレーニングを提供することといった四つの簡潔な方略を提言している。Larrick (2004) は, バイアス矯正研究を総括し, 研究で検討されてきた矯正方略を (a) 動機づけ方略 (motivational strategies), (b) 認知方略 (cognitive strategies), (c) 技法方略 (technological strategies), の三つに大きく分類している。以下

では, 動機づけ方略, 認知方略, 技法方略について説明する (Table 4を参照)。

“動機づけ方略 (motivational strategies)”とは, 意思決定課題に対して動機づけを高め, 課題のパフォーマンスを上げて, 意思決定バイアスを除去する方略をさす。これは, 課題への取り組む態度を操作する方略である。この方略は, “利得が損失・コストを上回るとき, 合理的な方略が使用される”という仮定が核にある。例えば, 課題を行うことは, 時間の拘束があり, できれば避けたい

Table 4
バイアス矯正の方略

バイアス矯正方略	種類	説明
動機づけ方略	報酬による動機づけ	実験後に高い報酬を与えることを教示することで, 実験の動機を高める方略
	バイアスへの事前警告と熟考の教示	課題遂行について, “注意深く”や“バイアスが潜んでいる”と教示し, 課題遂行の動機づけを高める方略
	説明責任	自分が下した判断や考えを他者に説明し, また他者から評価される手続き
認知方略	考察方略	初めに抱く考えとは反対の考えや, 他の考えを考察する方略
	ルール・表象・バイアスに関する訓練	規範的ルール, 心理表象, バイアスを学習することで, 認知バイアスを改善する手続き
	反実仮想	“～していれば, ～だろうに”という反実仮想を促し, 対立した考察を促す手続き
技法方略	決定分析	KT法, AHP, Raiffaの技法などの決定分析を使用し, バイアスを除去
	回帰モデル	回帰分析など統計モデルを利用することで, ヒューリスティック的な評価を回避
	意思決定支援システム	対話型のコンピュータシステムを使用
	集団意思決定	“個人の誤差は, 集団の成績によって相殺される”という原理を基に, 意思決定課題に集団を利用

ことである。しかし、課題への報酬が高いことを、課題前に教示すれば、課題に対する取り組み方は、合理的に解決されるだろう。動機づけ方略を用いた研究として、(a) 事前警告と熟考の教示 (Epley & Gilovich, 2005; 松崎・古屋, 2010; Schul, 1993), (b) 報酬による動機づけ (Camerer & Hogarth, 1999; Epley & Gilovich, 2005), (c) 説明責任 (Lerner & Tetlock, 1999) がある。しかし、動機づけ方略の研究では、安定したバイアス矯正の効果が得られていないため (Camerer & Hogarth, 1999; Lerner & Tetlock, 1999; Wilson & Brekke, 1994), 動機づけ方略の効果は限定的と考えられる。

“認知方略 (cognitive strategies)”とは、課題への取り組み方を改善する方略をさし、自分自身で行うことが可能な方略である。認知方略 (後述する技法方略も同様に) は、バイアス矯正の前提として、“ヒューリスティックスは不完全な判断であるが、規範的な方略を用いることで改善される”ことを想定している。認知方略として、(a) ルール・表象・バイアスに関する訓練 (Fong, Krantz, & Nisbett, 1986; Nisbett, Krantz, Jepson, & Kunda, 1983; Sedlmeier, 1999), (b) 考察方略 (Hirt, Kardes, & Markman, 2004; Hirt & Markman, 1995), (c) 反実仮想 (Galinsky, 2000; Miller, Markman, Wagner, & Hunt, 2013), をあげることができる。認知方略は、心理学研究で最も頻繁に検討されている方略であり、特に、反対意見の考察 (consider an opposite) や対立した考察 (consider an alternative) が、意思決定バイアスに有効であることが、数多くの研究によって見出されている (Hirt et al., 2004; Hirt & Markman, 1995)。

“技法方略”とは、決定ツールを用いて意思決定バイアスを矯正する方略をいい、前述した技法主義方略がこれに属する。技法方略は一般に、心理学分野よりも、経済学・経営学分野で検討されることが多い。主な技法方略として、(a) 決定分析 (Hammond, Keeney, & Raiffa, 1997), (b) 意思決定支援システム (Edwards & Fasolo, 2001), (c) 集団意思決定 (Soll & Larrick, 2004), (d) 回帰モデル (Dawes, Faust, & Meehl, 1989) の四つをあげ

ることができる。この技法方略は、簡単に利用できるものもあれば、複雑なものもあるため、専門的な知識を有し、厳密な手続きで適用する必要がある。また、技法方略については、決定ツールがバイアス矯正として有効であることを示す研究が少ないため、さらなる実証研究が必要である (相馬, 2012)。

このように、バイアス矯正の方略にはさまざまなものが存在する。一般に、動機づけ方略と認知方略は、主に心理学研究で検討されており、一方、技法方略は経済学研究で検討されている。以下では、バイアス矯正方略で最も効果が確認されている“考察方略 (consider strategy)”について説明する。

考察方略

心理学のバイアス矯正研究では、考察方略が多くの研究で検討されている。これは“初めの判断が間違っている理由は何か”と自身に尋ねる方略である。この考察方略は主に、(a) 反対意見の考察 (consider an opposite), (b) 対立した考察 (consider an alternative) の二つに焦点を当てている。前者は、初めに浮かんだ考えとは反対のものを考察する方略であり、後者は初めに浮かんだ考えとは反対の考え、またはその他の考えを考察する方略である。考察方略は、連合に基づく短絡的な判断を妨害し、考慮してなかった他の証拠に注意を向けさせる、有効な方略である。これに対し、決定者が単純に支持する理由のみ理由をリストアップすることは、意思決定バイアスの改善には役立たないことが見出されている (Hirt & Markman, 1995)。

対立した考察方略は、後知恵バイアス (Sanna, Schwarz, & Stocker, 2002), ハロー効果 (Chandon & Wansink, 2007), 説明バイアス (Hirt et al., 2004; Hirt & Markman, 1995), 社会判断 (Lord, Lepper, & Preston, 1984), 不安と連合した判断 (Bentz, Williamson, & Smith, 1999), 性的偏見 (Miller et al., 2013) など、数ある心理現象で検討されており、バイアス矯正の効果も確認されている。この対立した考察方略がバイアス矯正に最も有効である原因として、Koehler (1991) は、焦点仮説

(focal hypothesis) を提唱している。焦点仮説によれば、決定者が自身の信念に焦点を当てている時、焦点化している事柄と一致した根拠を求めようとする。しかし、対立した考察が、その事柄への焦点標準を壊すため、意思決定バイアスに陥らないと主張している。この仮説は、説明バイアスを、対立した考察方略によって矯正する研究 (Hirt & Markman, 1995) において、支持されている。

考察方略がバイアス矯正に有効であるならば、意見をより多く考察することが、さらに良いバイアス改善を促すのだろうか。Sanna et al. (2002) と Sanna & Schwarz (2003) は、後知恵バイアス課題について、対立した考察が多い (約10個) 人は、少ない (約3個) 人よりも、後知恵バイアスが強く現れることを見出した。また、Hirt et al. (2004) の研究でも、説明バイアス課題において、Sanna et al. (2002) と同様な結果を得ている。Hirt et al. (2004) は、考察を8個行った群は、考察を2個行った群よりも、説明バイアスが大きくなることを見出した。これら二つの結果は、一見すると矛盾を感じるかもしれない。なぜなら、考察する数が多ければ多いほど、情報量が多くなるため、バイアスを改善すると考えられるからである。

上記した実験結果は、“検索容易性 (ease of retrieval)” という概念によって説明がなされている。検索容易性とは、何かを思い出そうとしたときに経験される容易さの感覚である。例えば、r から始まる英単語の実例を思い出すのは簡単だが、3番目の文字がrである英単語の実例を挙げるのは難しいというように、検索容易性は、前述した利用可能性ヒューリスティックスと深くかかわっている。Schwarz, Bless, Strack, Klumpp, Rittenauer, & Simons (1991) は、“自分は積極的である” という見込みを評価させるとき、“この課題の遂行は容易である” という教示を受けた実験参加者群と、“この課題の遂行は困難である” という教示を受けた実験参加者群に分け、検索容易性を操作した。その結果、前者の実験参加者群の方が後者の実験参加者群よりも、“自分は積極

的である” と評価することがわかった。このことは、検索の容易さがその後の判断に影響を及ぼすことを示している。

さらに、Schwarz et al. (1991) は、想起した情報の数が検索容易性を変化させ、見込みの判断に影響を及ぼすことを実験によって示している。この研究では、“自分は積極的である” という見込みを評価する前に、その理由を実験参加者に想起させた。その際に、想起数が過剰に多い実験参加者群と、想起数が少ない実験参加者群とにわけ、見込みの評価を行わせた。その結果、想起数が過剰に多い実験参加者群は、想起数が少ない実験参加者群よりも、自分のことを積極的でないと評価することが見いだされた。この結果は、想起の数が過剰に多い場合、実例の検索に困難を感じ、その主観的経験をもとに、見込みの判断が行われるためであると説明されている。

以上の検索容易性に関する研究知見が示すように、意思決定バイアスの課題で、考察方略を行う場合には、(a) 考察の数と、(b) 考察の内容に注意しなければならない。過剰に多くの理由を考察することは、検索の困難さを感じるため、バイアス矯正に逆効果となることを先行研究は示している。また、“自殺をすることはどうしていけないことなのか” について考えることは、“消防隊が勇敢であると思う理由は何か” について考えることよりも、理由を考察することが困難である。考察すべき課題の内容によっても、方略は影響を受ける (Schwarz et al., 1991)。

また、考察方略には実験手続き上の問題がいくつかある。問題点の一つとして実験者効果をあげることができる。反対意見の考察を用いた研究では、“～することはなぜいけないのか” と考えること自体によって、実験参加者に、実験目的を気づかせてしまい、実験結果に影響を及ぼす可能性がある。つまり、考察方略の実験手続きが実験者効果となり、そのためにバイアスが改善されたのかもしれない。実験者効果を排除するためには、対立した考察方略で検討したり、潜在的なプライミングの操作を行うことで (Hirt et al., 2004)、反

対意見の考察方略を検討する必要がある。二つ目の問題として、リアクタンスの要因があげられる (Brehm, 1966)。意思決定課題において、実験参加者に反対意見の考察を勧めたり、過剰に多くの理由を考察させることは、実験参加者に負担を与え、自由な判断を妨げ、実験に対して反発を感じさせるかもしれない。

今後の展望

以上がバイアス矯正の研究動向であるが、バイアス矯正の問題として、(a) バイアス矯正の限定的効果の問題と、(b) 使用する側の問題の二つをあげることができる。

バイアス矯正は、バイアスの改善を生み出すが、何人かの研究者はそれが限定的な効果であることを示している。つまり、(a) 決定者のバイアス矯正の使用が特定の文脈のみに限られること、(b) バイアス矯正方略の学習から時間が経つにつれて、バイアスが回復してしまうことが見いだされている (Fishhoff, 1982; Lichtenstein & Fishhoff, 1980; Tversky & Kahneman, 1986)。例えば、自信過剰バイアスの研究では、個人に特化したフィードバックを意思決定者に集中的に与えれば、判断は改善されるが、その効果は短期的なものにとどまることを示している (Lichtenstein & Fishhoff, 1980)。

Tversky & Kahneman (1986) によれば、判断における基本的なバイアスは、時を重ねても自動的に矯正されない。さらに、学習には、正確かつ迅速なフィードバックが必要であるが、現実にそれが利用できる状況はめったにない。Tversky & Kahneman (1986) はその理由として、(a) 意思決定の結果が出るまでには時間がかかり、何らかの結果が出たからといって、その原因をある特定の行動に帰することは容易ではないこと、(b) 意思決定を下す環境には曖昧性があり、フィードバックの信頼性が損なわれること、(c) 通常、もし違う意思決定がなされていたら、どんな結果になっていたかを示す情報はないこと、(d) 最も重要な意思決定はそれぞれがユニークなので、学習

の機会はほとんどないことといった4点をあげている。

さらに Larrick (2004) は、外的な手段を使わず、人々が自力でバイアス矯正できない理由があると指摘している。第一に、決定結果のフィードバックが遅れたり、原因と結果の因果関係が曖昧であり、間違いの存在や原因を特定するのが難しいなど、しばしばバイアスに気づかない。第二に、決定プロセスの評価を行うため、決定結果を用いる場合、不確実な状況下では間違った結論を導くことがある。また Larrick (2004) は、バイアス研究を総括し、最も成功するバイアス矯正方略は、当該バイアスとその文脈に特化している必要があるため、訓練と検証は密接につながっていなければならないと述べている。この指摘は、バイアス矯正方略を自動化して使用できるような訓練が必要であることを示唆している。

また、バイアス矯正の問題として、Larrick (2004) はバイアス矯正の使用の問題点について述べ、決定者がバイアス矯正の説得を嫌がる理由があると主張している。第一に、決定者は決定が間違っているとされたくないし、決定プロセスをコントロールされたくない。第二に、人々は自身の決断に自信過剰をもつ傾向がある。意思決定課題において、実験参加者に外的な方法の使用を勧めることは、自身の意思決定のやり方が無効であることを間接的に説明し、自由な判断を妨げ、実験に対する反発を感じさせるかもしれない (Brehm, 1966)。さらに、バイアス矯正の技法は異質で複雑であるため、実験参加者自身の能力と比べて、バイアス矯正の技法によるメリットはほとんど理解されない。バイアス矯正の使用を促すために、実験参加者に対する説得の技法について検討することも今後の学術的課題である。

バイアス矯正研究の多くは、意思決定課題に対するものであり、感情を伴う意思決定バイアスについてのバイアス矯正や支援も検討すべき事項である。意思決定状況によっては何が最善なのか決め難い状況も存在し (相馬・都築, 2013)、この

ような状況では、満足や不満や後悔など、感情反応が重要となってくる。感情に基づくバイアスの矯正技法について研究するためには、感情が強く働く意思決定にいかにか効果的に対処するかという問題を検討する必要がある。また、本論文では主に、個人の認知レベルでの意思決定バイアスを扱ったが、対人相互作用を伴う意思決定バイアスも、また考察すべき事柄である。藤井（2011）は、社会的ジレンマの解決として、処方箋となる技法を系統的に分類している。今後のバイアス矯正研究として、対人相互作用を伴う意思決定バイアスもまた、バイアス矯正研究の検討すべき課題である。

引用文献

- Arkes, H. R. (1991). Costs and benefits of judgment errors: Implications for debiasing. *Psychological bulletin*, **110**, 486-498.
- Baumeister, R. F., Bratslavsky, E., Muraven, M., & Tice, D. M. (1998). Ego depletion : Is the active self a limited resource? *Journal of personality and social psychology*, **74**, 1252-1265.
- Bell, D. E., Raiffa, H., & Tversky, A. (1988). *Decision making: Descriptive, normative, and prescriptive interactions*. Cambridge University Press.
- Bentz, B. G., Williamson, D. A., & Smith, C. F. (1999). The prediction of negative events associated with anxiety and dietary restraint: A test of the content specificity hypothesis. *Journal of psychopathology and behavioral assessment*, **21**, 97-108.
- Brehm, J. W. (1966). *A theory of psychological reactance*. Academic Press.
- Camerer, C. F., & Hogarth, R. M. (1999). The effects of financial incentives in experiments: A review and capital labor production framework. *Journal of risk and uncertainty*, **19**, 7-42.
- Chandon, P., & Wansink, B. (2007). The biasing health halos of fast – food restaurant health claims: Lower calorie estimates and higher side – dish consumption intentions. *Journal of consumer research*, **34**, 301-314.
- Dawes, R. M., Faust, D., & Meehl, P. E. (1989). Clinical versus actuarial judgment. *Science*, **243**, 1668-1674.
- Edwards, W., & Fasolo, B. (2001). Decision technology. *Annual review of psychology*, **52**, 581-606.
- Epley, N., & Gilovich, T. (2005). When effortful thinking influences judgmental anchoring: differential effects of forewarning and incentives on self-generated and externally provided anchors. *Journal of behavioral decision making*, **18**, 199-212.
- Epstein, S. (1994). Integration of the cognitive and the psychodynamic unconscious. *American psychologist*, **49**, 709-724.
- Fischhoff, B. (1982). Debiasing. In D. Kahneman, P. Slovic, & A. Tversky (eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 422-44). Cambridge: Cambridge University Press.
- Fong, G. T., Krantz, D. H., & Nisbett, R. E. (1986). The effects of statistical training on thinking about everyday problems. *Cognitive psychology*, **18**, 253-292.
- 藤井 聡 (2011). 社会的ジレンマの処方箋—都市・交通・環境問題のための心理— ナカニシヤ出版
(Fujii, S.)
- Galinsky, A. D. (2000). Counterfactuals as behavioral primes: Priming the simulation heuristic and consideration of alternatives. *Journal of experimental social psychology*, **36**, 384-409.
- Gigerenzer, G. (2007). *Gut Feelings : The Intelligence of the unconscious*. New York: Viking Press.
- Gigerenzer, G., & Goldstein, D. G. (1996).

- Reasoning the fast and frugal way: Models of bounded rationality. *Psychological review*, **103**, 650-669.
- Hammond, J. S., Keeney, R. L., & Raiffa, H. (1997). *Smart choices: A practical guide to making better life decisions*. New York: Broadway Books.
- Hirt, E. R., Kardes, F. R., & Markman, K. D. (2004). Activating a mental simulation mind-set through generation of alternatives: Implications for debiasing in related and unrelated domains. *Journal of experimental social psychology*, **40**, 374-383.
- Hirt, E. R., & Markman, K. D. (1995). Multiple explanation: A consider an alternative strategy for debiasing judgments. *Journal of personality and social psychology*, **69**, 1069-1086.
- Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American psychologist*, **58**, 697-720.
- Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.
- Koehler, D. J. (1991). Explanation, imagination, and confidence in judgment. *Psychological Bulletin*, **110**, 499-519.
- Larrick, R. P. (2004). Debiasing, in D. J. Koehler and N. Harvey (eds.) *Blackwell handbook of judgment and decision making*, Oxford, England, Blackwell Publishers.
- Lerner, J. S., & Tetlock, P. E. (1999). Accounting for the effects of accountability. *Psychological Bulletin*, **125**, 225-275.
- Lichtenstein, S., & Fischhoff, B. (1980). Training for calibration. *Organizational behavior and human performance*, **26**, 149-171.
- Lord, C. G., Lepper, M. R., & Preston, E. (1984). Considering the opposite: A corrective strategy for social judgment. *Journal of personality and social psychology*, **47**, 1231-1243.
- 松崎圭佑・古屋 健 (2010). 推論・判断課題における二重過程説の検討 群馬大学教育学部紀要人文・社会科学編, **59**, 181-194. (Matsuzaki, K., & Furuya, T.)
- Miller, A. K., Markman, K. D., Wagner, M. M., & Hunt, A. N. (2013). Mental simulation and sexual prejudice reduction: The debiasing role of counterfactual thinking. *Journal of applied social psychology*, **43**, 190-194.
- Myers, D. G. (2004). *Intuition: Its powers and perils*. Yale University Press.
- Nisbett, R. E., Krantz, D. H., Jepson, D., & Kunda, Z. (1983). The use of statistical heuristics in everyday reasoning. *Psychological review*, **90**, 339-363.
- Sanna, L. J., & Schwarz, N. (2003). Debiasing the hindsight bias: The role of accessibility experiences and (mis)attributions. *Journal of experimental social psychology*, **39**, 287-295.
- Sanna, L. J., Schwarz, N., & Stocker, S. L. (2002). When debiasing backfires: Accessible content and accessibility experiences in debiasing hindsight. *Journal of experimental psychology*, **28**, 497-502.
- Schul, Y. (1993). When warning succeeds: The effect of warning on success in ignoring invalid information. *Journal of experimental social psychology*, **29**, 42-62.
- Schwarz, N., Bless, H., Strack, F., Klumpp, G., Rittenauer, S. H., & Simons, A. (1991). Ease of retrieval as information: Another look at the availability heuristic. *Journal of personality and social psychology*, **61**, 195-202.
- Sedlmeier, P. (1999). *Improving statistical reasoning: Theoretical models and practical implications*. Mahwah, NJ: Erlbaum.
- Soll, J. B., & Larrick, R. P. (2004). *Strategies for revising judgment: How, and how well, do people use others' opinions?* Unpublished Manuscript.
- 相馬正史 (2012). 処方的アプローチにおいて利用できる意思決定ツールとその展望 立教大学心理学研究, **54**, 91-100. (Soma, M. (2012). In prescriptive approach

available decision tool and that future. *Rikkyo Psychological Research*, **54**, 91-100.)

相馬正史・都築誉史 (2013). 道徳ジレンマ状況における意思決定研究の動向 立教大学心理学研究, **55**, 67-78.

(Soma, M., & Tsuzuki, T. (2013). Recent trends in research on decision making with respect to the moral dilemma problem. *Rikkyo Psychological Research*, **55**, 67-78.)

Stanovich, K. E. (1999). *Who is rational? Studies of*

individual differences in reasoning. Mahwah, NJ: Erlbaum.

Tversky, A., & Kahneman, D. (1986). Rational choice and the framing of decisions. *The journal of business*, **59**, S251-S278.

Wilson, T. D., & Brekke, N. (1994). Mental contamination and mental correction: Unwanted influences on judgments and evaluations. *Psychological bulletin*, **116**, 117-142.

——2013.10.3受稿, 2013.12.16受理——