

信念 (belief) と亜信念 (alief) について

研究動向の概観および Mandelbaum/Danón の論争の検討*

田村 歩

概要

We sometimes take actions that contradict our beliefs or desires—for example, we may hesitate to walk on a glass floor high above the ground even though we are certain that it is completely safe. Or we may hesitate to wear clothes that were previously worn by someone we dislike, even though we are certain that human properties cannot be physically transmitted. Gendler (2008a) proposed the existence of the mental state of ‘alief’ to explain why we can behave in ways that contradict our beliefs and desires. Gendler defines alief as a mental state with associatively-linked content that is representational, affective and behavioral, which is, unlike belief, not a propositional attitude and not reality-sensitive.

There has already been a certain amount of research on this concept, but extremely few Japanese-language publications deal with it as a subject. In light of this situation, I will begin with a detailed introduction to the concept. In addition, I will summarise recent research trends, such as Mandelbaum’s counter-argument that aliefs are indistinguishable from beliefs and other existing mental states in that they all possess propositional content (therefore, there is no reason to assume the existence of aliefs) and Danón’s defence that the propositional content of aliefs is ‘semi-structured’, and in this respect, aliefs can be distinguished from beliefs and other mental states. Lastly, I will present, based on these studies, some issues for future study.

Keywords: 亜信念 信念 想像 命題的内容 半構造化された命題的内容

* CAP Vol. 17 (2026) pp. 1–15. Submitted: 2025.12.16. Accepted: 2026.03.13. Category: 研究ノート. Published: 2026.04.06.

1 導入

人間は、ときに自らの信念[belief]や欲求[desire]に矛盾する行動をとることがある——たとえば、絶対に安全であるとたしかに信じていたとしても、高所にあるガラス張りの床を歩くことを躊躇する。あるいは、安全な食べ物であるとたしかに信じていたとしても、犬の糞にそっくりのお菓子を食べることを躊躇する。あるいは、人種差別や男女差別は避けるべきであるとたしかに信じていたとしても、いざ現実的な場面では差別的な言動をとってしまう(内容的に相違がなくとも、黒人的な名前の書かれた履歴書よりも白人的な名前の書かれた履歴書を重視してしまう)。あるいは、人間の性質が物理的に伝染することはありえないとたしかに信じていたとしても、嫌悪対象者が過去に着用していた完全に洗濯済みの衣服を着用することを躊躇する。

T. Gendler は「亜信念と信念」(2008a)という論文において、このように人間が自らの信念や欲求に矛盾する行動をとる理由を説明するために、「亜信念[alief]」という心的状態の存在を主張した^{*1}。亜信念とは、表象的内容・感情的内容・行動的内容が連想的に結び付いた心的状態であり、特定の刺激によって自動的に活性化され、また信念や想像とは異なり命題的態度ではなく、さらに信念のように現実を敏感に反映するものではないとされる。そしてこの亜信念の概念は、潜在的な偏見における持続性や修正困難性をめぐる議論、また人間の道徳性の進化論的由来をめぐる進化倫理学的議論に大きく影響するものとして注目されている(たとえば、千葉 2021; Fürst 2023; Hung 2023; Kriegel 2015)。

この概念についてはすでに一定程度の研究蓄積がみられるが、これを主題とする邦語文献はきわめて少ない^{*2}という現状に鑑み、本稿ではまずその詳細な紹介を行い、その後の研究動向を概観する。さらに、亜信念はそれが命題的内容を持つという点で信念やその他既存の心的状態とは異ならない[ゆえに亜信念を仮定する理由はない]という E. Mandelbaum の反論、および、亜信念が命題的内容を持つという点では Mandelbaum の主張に同意するものの、命題的内容の概念を強いものと弱いものとに区分し、亜信念は「半構造化された命題的内容」を持っている[ゆえに亜信念とその他の心的状態は区別される]とい

^{*1} Gendler は、当該論文とほぼ同時期に出版された「亜信念の作用(および反作用)」(Gendler 2008b)において、亜信念を導入することの利点として「説明上の有用性」を挙げている。つまり、この事例は本能に、あの事例は習慣に、あるいは忘却や想像によるものである、というように、異なる事例には異なる説明が必要であるという見解——Gendler はこれを「古典的認知主義的な図式[classic cognitivist picture]」(Gendler 2008b: 556)と呼ぶ——は、「信念、想像、習慣、本能の間の特定の重要な差異に対して十分に敏感ではなく、またそれに応じて、それらのカテゴリに分類される諸事例間の特定の重要な類似性に対しても十分に敏感ではない」(Gendler 2008b: 556)のに対して、亜信念に基づく図式は、「諸事例間の特定の重要な類似性に注意を向けさせるとともに、他の認知的状態間の重要な差異にも敏感になることを可能にする」(Gendler 2008b: 556–557)という有用性がある。その一方で、亜信念をめぐる問題は「信念や欲求[...]といった概念を用いない他の枠組みにおいては生じないかもしれない」(Gendler 2008b: 555)と断っている。

^{*2} 亜信念を主題としたほぼ唯一の邦語文献は千葉(2021)である。この論文は、人間における道徳性の進化論的由来を究明する進化倫理学を擁護するために亜信念概念を援用している。また「亜信念」という訳語を提案しているのもこの論文である。千葉によれば、「「亜」とは「なにかに準ずる」ことを表す漢字であるから[...]、「亜信念」と言うことで「信念のようだがそれよりは原始的ななにか」という内容を表すことができ」(千葉 2021: 88)、また「西洋語の“a”に対する由緒正しい漢字での音訳である」(千葉 2021: 88)ことから、「この訳語案は、意味と音声の両面で優れていると考えられる」(千葉 2021: 88)。

う L. Danón の亜信念擁護という論争を検討し、それらを踏まえて今後の課題となりうる点を提示する。

2 亜信念とは何か

2.1 暫定的特徴

亜信念の説明にあたり、Gendler (2008a: 635–636)は、主体が自身の信念に矛盾する言動をとる事例を多く挙げており、ここでは一部を参照したい(記号化は論者による便宜上のものである)。

事例A: 新品の便器に入ったスープを食べたくない

事例B: 自身が嫌悪する人物が着ていた洗濯済みのシャツを着たくない

事例C: 自身が愛する人物の顔写真にダーツを投げたくない

このとき、主体はそれぞれ、「新品の便器には衛生上の問題はない」や「当該シャツには衛生上の問題はない、かつ、いかなる仕方であれ以前の着用者の性質が当該シャツを介して伝染することはない」や「ダーツによって写真の人物に危害が生じることはありえない」という信念をたしかに有している。にもかかわらず、主体は上記の行動にためらいを覚える(あるいはCの場合、ダーツを投げたとしても、無関係な人物の顔写真の場合に比べ命中率が下がる (Rozin et al. 1986: 705))。このような信念と行動との不一致を説明しうるのが「亜信念」であり、Gendler は以下のように規定している。

典型的な亜信念とは、表象的・感情的・行動的な内容がそれぞれ連想的に結びついた心的状態であり、主体の内的環境または周囲環境の諸特性によって、意識的または無意識的に活性化されるものである。(Gendler 2008a: 642)

つまり亜信念とは、何らかの命題に対する心的な態度(命題的態度)である信念や想像とは異なり——この点については次節で言及する——あくまで心的な状態である。そしてこの状態とは、何らかの対象に関する表象^{*3}と、何らかの感情的ないし情動的な状態の経験と、何らかの運動ルーティンの準備という三つの要素が同時に活性化しそれぞれ結び付いている状態である。たとえば、主体が便器を知覚したとき、便器の表象が得られると同時に、それに対する嫌悪の感情と回避の行動が連想的に活性化されることがありうる(ただしこれはあくまで「典型的な」亜信念の特性であり、Gendler は、表象という概念が適用できないほど活性化が低い場合や、明らかな感情的要素を欠く状態、あるいは運動ルーティンの明確な活性化を含まない状態の亜信念もある可能性を考慮している) (Gendler 2008a: 641; 643–644; 645–646)。加えて、亜信念の内容は、主体の内的環境および主体を取り巻く環境の諸特性によって活性化されるが、それは

^{*3} Gendler (2008a: 643)は、この表象が命題的ないし非命題的、また概念的ないし非概念的でありうることを主張しているが、亜信念が命題的内容を持つか否かは大きな問題である。この問題については第3節で言及する。

意識的な場合もあれば無意識的な場合もあるとされる。つまり、主体がどのような習慣や思考を持ってきたかやどのような環境にあるかによって活性化される亜信念の内容は異なりうるし、活性化された感情や運動準備を主体が意識していない場合もありうる (Gendler 2008a: 644) (たとえば、主体によっては便器に嫌悪の感情を抱かないかもしれないし、写真を傷付けても被写体実物に影響はないと固く信じている主体が自らの愛する者の顔写真にダーツを投げるとき、第三者の顔写真に比べ命中率が落ちるのは、無意識の感情や運動準備によるものであるかもしれない)。

さらに、亜信念は「顕在的なもの、または傾向的なもののいずれかである」(Gendler 2008a: 644)とされる。特定の表象に関する思考を抱き、特定の感情を経験し、特定の運動を準備するという一連の傾向が、主体の内的環境や周囲環境によって実際に活性化されたとき、主体は表象的・感情的・行動的な内容を伴う顕在的な亜信念を持っていることになる。また、主体がそれに遭遇したならば、表象的・感情的・行動的な内容を当該主体に顕在的に亜信させるような〔内的であれ外的であれ潜在的な〕刺激が存在するとき、主体は傾向的な亜信念を持っていることになる (Gendler 2008a: 645)*4。

これらの規定に先の各事例を当てはめると以下ようになる。

亜信念A: 便器だ! (表象的内容) / 汚い! (感情的内容) / 避けよ! (行動的内容)

亜信念B: 不潔なものだ! / 汚染されている! / 触れるな!

亜信念C: 愛する人物に対する有害な行為だ! / 危険で賢明ではない! / 投げるな!

このとき、主体は、「新品の便器には衛生上の問題はない」や「当該シャツには衛生上の問題はない、かつ、いかなる仕方であれ以前の着用者の性質が当該シャツを介して伝染することはない」という信念を持っはいるが、便器に入ったスプーンを目の当たりにすることで、また当該シャツの着用者が自身の敵対者である事実を認知することで、嫌悪の感情および回避行動の活性化が連想的に〔同時に〕引き起こされる。あるいは、「ダーツによって写真の人物に危害が生じることはありえない」という信念を持っはいるが、自分が投げようとしているダーツの的が愛する者の顔写真である事実を認知することで、「かわいそう」や「悲しい」といった感情および中断という行動の活性化が連想的に〔同時に〕引き起こされる。

2.2 亜信念とその他の態度

加えて Gendler (2008a, 647–656) は、亜信念と信念・想像といった他の認知的態度との区別を分析している。決定的な相違は、信念や想像が命題的態度であるのに対して、亜信念はそうではないという点にある。つまり、前者には、命題Pを〔現実上の環境の一部ないし想像上の環境の一部として〕受容するという

*4 最後に暫定的な特徴として、「主体がRを亜信する」という二項関係あるいは「主体がR-A-Bを亜信する」という四項関係の両方があることが挙げられている (Gendler 2008a: 645)。この点については当該論文で詳しく述べられていないが、別の論文では、典型的な亜信念であれば表象的・感情的・行動的内容の三要素が連想的に活性化するが、活性化がそのうちの一つまたは二つに限定される場合にも亜信念を適用できるように、二項関係で表現することが提案されている (Gendler 2008b: 560)。

態度が含まれるのに対して、後者にはPの受容は含まれない。Gendler の挙げる事例を参照しよう。

P. Rozin らの実験において、被験者は、二本の瓶に砂糖が注がれ、そのボトルにそれぞれ「砂糖」と「青酸カリ」と記されたラベルを自分の判断で貼った。このとき、被験者は、自らラベルを貼ったにもかかわらず、「青酸カリ」のラベルが貼られた瓶から砂糖を飲むことを嫌がった(このとき、〈青酸カリだ！危ない！避けよう！〉という亜信念が発生したと考えられる)。本題はここからだが、その四年後の追跡調査において、被験者はまず、二本の茶色の空瓶が目の前に置かれ、それらに砂糖が注がれた。次いで、「青酸カリは含まれておらず、毒ではありません」と記され、「not」という単語と赤い頭蓋骨と骨が印刷されたラベルと、「ショ糖、食卓用砂糖」と記されたラベルを一枚ずつ瓶に貼るよう指示された。そして二つのコップに同じ飲み物が注がれ、加えて先の二本の瓶からそれぞれ粉末を注がれた結果、被験者は、「青酸カリは含まれておらず、毒ではありません」と表示された瓶の砂糖の入ったコップから飲むことを躊躇する傾向が認められた(Rozin, et al. 1986: 705)。

このとき、主体が「瓶に青酸カリが入っている」という信念を持っているとすれば、主体は「瓶に青酸カリが入っている」という命題Pを〔現実において真なるものとして〕受容しているということになる。しかし当該実験において、実験道具はすべて主体に開示されており、この命題を受容する余地はない。また、主体が「瓶に青酸カリが入っている」と想像しているとすれば、やはり主体はPを〔自身の想像において真なるものとして〕受容しているということになる。しかし主体がPを想像していたとすると、当該のコップから液体を飲むことを主体が躊躇した理由の説明が困難になる。この躊躇が想像によるものであるとするならば、1) 主体は以前には瓶に青酸カリが入っていると想像したが、しかし現在はそうしていないからである、つまり主体が過去に想像していた〔ただし現在は想像していない〕ことが何らかの形で現在の行動を説明しているということになる^{*5}か、あるいは、2) 主体が「瓶に青酸カリが入っていない」と想像していることが、何らかの仕方ですその反対の想像——「瓶に青酸カリが入っている」——を含んでおり、そして自分がそれを明らかに信じてはいないことを完全に自覚しているようなこの内容が、主体の運動ルーティンの制御センターとでも言うべきものに忍び込み、主体を躊躇させたということになる。しかし Gendler (2008a: 650)によれば、当該実験における躊躇は、現在には否定されている過去の想像によって生じた動機付けとして説明するよりも、「瓶に青酸カリが入っていない」という信念を持ち、またそのように想像している主体による〈青酸カリだ！危ない！避けよう！〉という亜信念として説明する方がより自然である^{*6}。

上記の実験に基づく事例的説明に加えて、Gendler は、亜信念が受容という態度なしに生ずる理由を示すより一般的な論拠も提示している。Gendler (2008a: 650–652)によれば、その本質として連想の連鎖の活性化を伴う亜信念は、連想を活性化する場合に対する態度に関係なく発生するものであり、非意識的に活性化される場合もあるため、その内容に対してまったく態度を持たないこともありうる。つまり亜信念の

^{*5} もし過去にも現在にもPと想像していれば絶対に飲まないだろうし、過去にも現在にもPではないと想像していれば問題なく飲むだろう(つまり躊躇は生じない)。

^{*6} Gendler はここで「瓶に青酸カリが入っていない」という想像がその反対の想像を含んでいる可能性については詳しく検討していないが、かわりに、亜信念を想像と同一視できないことのより一般的な論拠を以降で提示している(Gendler 2008a: 650–652)。それについては本稿の次段落で紹介する。

文脈では、たとえば「安全に摂取できる」および「毒ではない」という表現のように、互いに同一指示的 [coreferential] であると主体が明示的に認識している表現でさえ完全に置換可能ではありえない (Gendler はこれを「超不透明性 [hyperopacity]」と呼ぶ)。たとえ「安全に摂取できる」および「毒ではない」という表現が、同範囲の物質を指すものだと主体が信じていたとしても、あるいはたとえその信念に焦点を当ててそれを頭の中で鮮明に思い描いていたとしても、これら二つの表現によって活性化される亜信念はまったく異なるものになる可能性がある。反対に、想像ではこのようなことはありえない。主体がPとQとが同義であることを明示的に認識し、PとQの同一指示性に明示的に焦点を当てつつPを想像した場合、Pを想像すると同時にQも想像することになり、この点において亜信念は想像とは異なるのである。

さらに、亜信念が受容を伴わないことは、主体が自由に亜信することができない理由にもなる。信念については、主体がPを信じ、その後Pに反する証拠が提示された場合、主体はその信念を修正することになる (つまり信念は現実を敏感に反映する)。また想像については、主体がPを想像し、その後Pではないと知った場合、主体はその想像を変更することもしないことも可能である (つまり想像は現実を無視しうる)。しかし亜信念はそのどちらでもなく、[反対] 証拠に反応して現実を反映しようとするものでもなければ、現実を完全に無視して自由に生み出されるものでもない。亜信念はあくまで特定の状況下で自動的に活性化^{*7}する心の「状態」であり、命題を証拠に応じて受容する信念や命題を自由に受容する想像といった心の「態度」ではないのである。

2.3 その後の研究動向

この亜信念という概念が提示されて以降、それは進化心理学や認知科学や人類学といった広い分野において用いられるようになった。たとえば U. Kriegel は、認知科学における「二重過程理論」——私たちの思考は二つの異なるシステムによるものであるとする——と親和的であるがゆえに亜信念説は支持されるとする。つまり、二重過程理論では、「高速、効率的、自動的、暗黙的、無意識的、特化的で、行動に直接結び付いているが、柔軟性に欠け、とりわけ監視や制御には適さず (情報が閉ざされているため)、範囲が限定されている」(Kriegel 2015: 476) システム1、および「柔軟で、広範な範囲を持ち、意識的かつ明示的に制御も操作も可能であるが、行動との関連は間接的で遅く非効率的」(Kriegel 2015: 476) なシステム2が想定されているが、この相違は、「典型的な場合、自動化され、比較的制御が難しく、運動出力と構造的に関連し、意識にアクセスできない過程で生成される」(Kriegel 2015: 476) ところの亜信念、および「行動への影響は間接的である一方で、一般的に意識的で制御可能な過程によって生成される」(Kriegel 2015: 476) ところの信念にも当てはまる。これを踏まえて Kriegel は亜信念を進化心理学に応用

^{*7} 亜信念の自動性を説明するにあたり、Gendler は Bargh et al. (1996: 236–237) の実験を援用する。その実験で、あるグループの被験者らは高齢者に関連する用語 (しわ、ビンゴ、退職、等) を含む文章を並べ替える課題を行い、別のグループの被験者らは中立的な用語のみを含む文章を並べ替える課題を行った。実験終了後、被験者らは、エレベータまでの廊下を歩く時間を密かに測定された結果、前者の被験者らは、後者の被験者らよりもかなり長い時間を要した (いわゆるプライミング効果——あらかじめ与えられた刺激ないし情報がその後の行動に影響を及ぼすこと)。このとき、高齢者のステレオタイプを与えられた被験者らは、「高齢である／疲れている／エレベータまで歩くときは注意しよう」と亜信した結果、そのような行動をとったと考えられる (Gendler 2008a: 658–659)。

し、反人種差別的な信念を有する主体がそれに反する行動をとってしまうといういわゆる暗黙の偏見を説明しうる「道徳的亜信念[moral alief]」の存在を主張している(Kriegel 2015: 477)。他にも、K. M. Hodge は、人間が直観的に他者が死後も存在し続けると信じる理由を認知的に説明するために亜信念を援用している。Hodge (2011: 374–376)によれば、死によって精神的・身体的なあらゆる状態が消滅すると考える主体であっても、死者が自身の死を自覚しているというような死後の状態がありうると信じるのは、その主体が「死者は単に不在であるだけであり、いまだ生きている者と同様に活動や状態を維持しうる」と亜信するからである。そして信念とは違って現実を敏感に反映しないという亜信念の特徴を踏まえれば、主体がある者の死の確固たる証拠を得たとしても死後の世界への信仰を持つ——死によってすべてが終わるという信念を持っていたとしても、死者が自分を見守ってくれていると考えてしまう——ことは可能であり、このような亜信念によって、人間は特定の死後の世界観(キリスト教等)を明示的に信じていなくとも何かしらの死後の世界への信仰を持ちうるのである。

しかし、亜信念それ自体を主題とする哲学的研究においては批判が多く向けられている。G. Hubbs は、心を説明する体系には亜信念の概念を追加する必要はない、つまり、〈信念－欲求〉説という二元論を超越すべく亜信念を導入した Gendler の思惑に反して、検討される諸事例はすべて従来の〈信念－欲求〉の枠組みで説明されうると主張する。第一に、一般的に同一の主体に対してそれぞれ矛盾する信念を帰属させることはできないにもかかわらず信念と行動とが一致しない事例については、一時的な忘却および〔意識的〕アクセスの観点から説明されうる。たとえば、ある主体がナイフの収納場所を棚Aから棚Bに自ら移したにもかかわらず、その後しばらくしてナイフを使おうと棚Aを開けてしまうという場合、主体に「ナイフは棚Aにある」という信念と「ナイフは棚Bにある」という信念を同時に帰属させることはできないため、「ナイフ／欲しい！／棚Aから出そう」という亜信念が主体に生じていると Gendler なら主張するだろう。しかし Hubbs (2013: 615–616)によれば、主体がナイフを棚Bに移したことを、つまり「ナイフは棚Bにある」という信念を、一時的に忘却していると考えることでこの事例は説明されうる。つまり、「ナイフは棚Bにある」という信念が一時的に忘却されているならば、主体はその信念にアクセスすることはできず、ゆえに「ナイフは棚Aにある」という矛盾する信念をその主体に帰属させたとしても不合理ではない。このように考えれば、主体は「ナイフは棚Aにある」という信念のもと棚Aに手を伸ばすのであるから、亜信念を用いる必要はなくなる。さらに Hubbs は「である／に見える」の区別のもとに事例を説明する。たとえばある主体が高所の橋を眼前にして、その橋は安全であるとたしかに信じているはずなのに渡ることができない場合、当該主体は、「高所の橋／危ない！／避けよう」という亜信念が主体に生じていると Gendler なら主張するだろう。しかし Hubbs (2013: 616)によれば、当該主体は「この橋は安全である」という信念に加えて「この橋は安全ではないように見える」という信念も有しているのであり、後者の「に見える」が感情反応を引き起こし、それが行動の原因となるのである。であるならばここでもやはり亜信念に頼る必要はない。

また G. Currie および A. Ichino も、Gendler の挙げた諸事例は信念や欲求や想像といった既存のカテゴリで説明可能であり、「Gendler の事例は表象・感情・運動ルーティンを主要な要素とする因果的過程の結果である」(Currie & Ichino 2012: 793)とは言えても、その三要素が「実質的な実体を構成するわけではない」(Currie & Ichino 2012: 790)と批判する。具体的には、Gendler が想像では説明できないとした

Rozin の毒の実験について、「「Pではない」という形式の通達事項を目にすることによってPを想像するよう誘発されるということは十分にありうる」(Currie & Ichino 2012: 792)——たとえば威嚇的なマフィアが私たちの愛する者に「死んでいない」というラベルを貼る場合、反対のことが容易に思い浮かばれる——と主張する。また、人種差別反対の確固とした信念を持つ主体がそれに反する行動をとってしまうという事例についても、「自身の信念の体系全体において一貫性がいまだ達成されていない、あるいは自身の価値観に反する欲求がいまだ存在している、と結論付けるのがより適切」(Currie & Ichino 2012: 793)であり、亜信念に頼らずとも既存の信念や欲求によって説明することができる*8。

さらにこれらの批判的研究に加えて、近年では亜信念が命題的内容を有するか否かについての論争が生じている(命題的内容の有無の観点から亜信念を否定する研究は 2013 年のものだが、それへの反論は比較的最近提出されたものである)。以下ではこの論争を概観し、残された課題を提示したい。

3 Mandelbaum による亜信念批判: 亜信念は命題的内容を持つか

Gendler の亜信念説に対して、Mandelbaum は命題的内容という観点からその存在を否定する。Mandelbaum (2013: 198)によれば、亜信念には次のジレンマが内在している。

- 1) もし亜信念が命題的内容を持つならば、それは既存の心理学的実体と区別されえず、この場合亜信念という強固な概念は存在しない。
- 2) もし亜信念が命題的内容を持たないならば、それは Gendler が課した説明的役割を果たすことができず、よって亜信念を仮定する理由はない。

3.1 亜信念が命題的内容を持つならば・・・

このジレンマを扱うにあたり、Mandelbaum は、まず亜信念が命題的内容を持つと仮定して議論を進める。Mandelbaum (2013: 199–202)によれば、表象的・感情的・行動的な内容がそれぞれ連想的に結びついているという亜信念の特徴は、他の心的状態にも見出される。たとえば、「私のベッドにトラがいる」という信念は、恐怖反応・逃げるべしという信念・逃避行動の準備を促す運動反応といった行動的および認知的状態を同時に活性化させる傾向にある。さらに Mandelbaum は、既存の心理学的概念である「極小感情価[micro-valence]」を援用することで、信念は亜信念と同様に機能しうると主張する。極小感情価とは概念と関連する微細な感情の断片のことであり、その感情はあまりに微細であるため主体は内省によって

*8 ただし Currie および Ichino は、自身らの研究によって従来の〈信念—欲求〉説の方が正しいと立証できるわけではなくと留保を付している。「心についての説明とそれを支える諸カテゴリは暫定的なものにすぎないとみなされるべきである」(Currie & Ichino 2012: 797)。Currie および Ichino は、当該諸事例を統一的に説明できるような堅固なカテゴリとしての Gendler 的亜信念は否定するが、場合によっては必要とされるかもしれない追加的な表象カテゴリ——「亜信念の親戚」(Currie & Ichino 2012: 797)——の可能性は認めている。

明確に気付くことはない。Lebrecht and Tarr (2010: 966)および Lebrecht et al. (2012: 2–4)の研究によれば、あらゆる概念はこの極小感情価を有しており、より高い認知コストを伴わずに迅速な運動反応を促進する機能を果たしている。たとえば、朝早くにコーヒーを淹れて、たくさんのマグカップの中から一つを選んでコーヒーを入れるとき、マグカップ間の好みの順序を定める意思決定理論的な評価指標[metric]を持つという、認知効率の面でコストの高い手続きが行われるのではなく、各マグカップに極小感情価が存在し、それが運動反応を誘導し、迅速な意思決定プロセスが促進される。この研究に基づき Mandelbaum (2013: 200–201)は、何かしらの概念から構成される信念は、その概念に付随する極小感情価を継承し、さらにその極小感情価に結び付いた運動ルーティンも継承することになる、つまり、主体が持つ信念は価値中立的なものではなく、常に感情と運動ルーティンとに関連付けられており、この点で信念は亜信念と同じように機能しうる、と主張する。

3.2 亜信念が命題的内容を持たないならば・・・

上記の分析が正しいければ、亜信念の存在を認めるためには、つまり亜信念と他の心的状態とを区別するためには、亜信念が命題的内容を持たず単に連想的——〈青酸カリだ！危ない！避けよう！〉というように——であると主張しなければならないことになる。しかし Mandelbaum (2013: 202–204)によれば、亜信念が本質的に連想的であることも、命題的内容を持たないことも不可能である。再度 Rozin の実験を引き合いに出そう。Mandelbaum は、この実験において主体が Gendler の言うように〈青酸カリだ！危ない！避けよう！〉という亜信念を持っていると仮定した場合、この亜信念は一体何を避けるよう主体に伝えているのか、と問う。というのも、もし主体が当該亜信念中の諸概念を単に順にトークン化しているだけであるならば、なぜ主体は他ならぬその瓶[および中身]に対してだけ——窓や床や机にではなく——回避行動を示すのかを説明できないからだ。つまり、ここで活性化されている[回避]行動が当該の瓶に特化している以上、仮定される亜信念は何かしらの仕方でそれらと結合していなければならないが、単に亜信念の内容が当該の瓶に関係しているというだけでは、その亜信念と当該の瓶との結合を説明できない^{*9}。したがって、この結合が成り立つために、亜信念はその内容が構造化されている[命題的内容を持っている]必要があるということになる(「この瓶の中身-は-危険である」のように)。

さらに、Gendler の亜信念は推論を許すように思われるが、であるならばやはり亜信念は命題的内容を持たなければならないことになる。たとえば、Rozin の実験に参加した被験者が、その直後に、他人が「青酸カリ」と表示された瓶から液体を飲むか否かという追跡質問を受けた場合、被験者は他人がその瓶からは飲まないだろうと推論しようと Mandelbaum (2013: 204)は主張する(その瓶には毒が入っている、人は毒

^{*9} この問題を明確にするために Mandelbaum (2013: 203–204)は、連合主義的認知に関する L. Williams および J. Bargh らの実験研究を参照している。この実験は、人が温かさと親しみやすさを関連付ける傾向にある、たとえば、温かいコーヒーカップを持っていると通常よりも親しみやすい行動をとる傾向にあることを示した。そしてこれが意味するのは、被験者は、コーヒーカップそのものに対して親しみやすい行動をとる傾向にあるということではなく、全般的に親しみやすい行動をとる傾向にあるということである(Williams & Bargh 2008)。これに対して Rozin の実験の被験者は、全般的に恐怖を抱いているわけではなく、その恐怖は当該の瓶およびその中身にのみ結合している。

を飲むことを好まない、ゆえに人はその瓶から飲まないだろう、という無意識的な推論の連鎖を通じて)。すなわち、「それは危険な青酸カリである、だから回避せよ」から、たとえば「他人は「青酸カリ」と表示された瓶を避けたがるだろう」や「その瓶は持ち上げられても中身は変わらないだろう」といった意味論的に関連する他の思考が推論されるのであり、そしてこれらの推論は命題状態からしか導出されえないのだから、したがって信念に似た命題状態が存在しなければならない。すなわち、亜信念の存在を仮定する必要はない。

4 Danón による反論：亜信念における半構造化された命題的内容

4.1 亜信念と信念との相違点

Danón (2021: 8509)は、Mandelbaum が見落としている信念と亜信念の相違として、亜信念は主体の行動に直接かつ自動的に影響を与えると考えられるが、一般的に信念はそれ自体では行動を引き起こすことはできないという点を挙げている。具体的には、信念が行動を引き起こすためにはある種の推論プロセスにおいて欲求等の他の命題的態度と組み合わせられなければならない、そして同一の信念であってもそれが推論によって他のどのような心的状態と結び付くかで無数の異なる行動を引き起こす可能性があるという一般的な見解 (Millikan 2004; Fodor 1987; Davidson 1963) に基づき、たとえば主体が「捕食動物が近くにいる」という信念を持っている場合、逃げ切れる自信があれば走るだろうし、あるいは近辺に複数の強力な味方がいると認識すれば助けを求めて叫ぶだろうし、あるいは自分がそれに勝てるだけ十分に強いと判断すればその場に留まって戦うだろうが、それに対して亜信念は、主体が他のいかなる心的状態にあるかにかかわらず、一つの特定の運動ルーティン群と堅固に関連付けられていると主張する。加えて Danón (2021: 8509)は、亜信念は現在の知覚刺激によって常に堅固で体系的な仕方で生起するが、信念はそれ以外にも推論や証言や過去の経験の記憶等によって獲得される場合もあるという相違も指摘している。

4.2 半構造化された命題的内容

さらに Danón は、Mandelbaum が呈した亜信念における結合の問題と推論の問題とを解決するために、亜信念が持ちうる命題的内容の在り方を分析する。その際に Danón (2021: 8510–8511)は、「文のような構造を持つ」——つまり、概念で構成され、それら概念は自由かつ自発的に再結合して新たな命題的内容を形成でき、特定の状況への適用とは独立にその意味が把握可能であり、そして普遍量子子および存在量子子を許容する——という厳しい要件にではなく、「あるものが特定の仕方で存在することを表象し、その結果、真ないし偽である、あるいはより穏便に言えば正しいないし間違いであると評価されうる」という緩い要件に依拠し、命題的内容の亜種として「半構造化された命題的内容」の概念を提唱する。そしてこれが亜信念の持ちうる命題的内容である。

信念の場合、その内容が概念から構成されているだけでなく、構成要素たるそれら概念を分離し、異なる方法で再結合することが可能である。つまり、主体が「ガラス製の床は危険である」という信念を持っている場合、主体は「ガラス」「床」「危険」という概念をこの特定の内容から分離し、他の概念と再結合できる。これに対して亜信念の場合、その内容の構成要素である概念の一部を他の部分から分離することが不可能であると考えられる。たとえば、主体が特定のaについて、それがFであるとみなさずには考えることができないという場合、主体は概念aを用いるたびに、特定のaをFという性質を持つものとして表象することを避けられない。換言すれば、主体が概念aを用いるたびに、不可避免的にそれを「aはFである」という内容に埋め込むことになる。この点で亜信念の内容は、完全に構造化された内容——自由に分離・再結合可能な概念単位のみで構成されるもの——ではないということになる (Danón 2021: 8512–8513)。

ただし、主体が概念aを概念Fから分離できない場合でも、「aはFである」と考える際、主体は完全に非構造的な内容——より小さな表象単位に分解不可能な内容——を持っているわけではない。というのも、この内容全体から分離可能で、他の内容に再配置されても常に同一の安定した意味的貢献をないうる少なくとも一つの構成要素的概念単位 (属性概念 F) を特定できるからである。たとえばグランドキャニオン・スカイウォークに足を踏み入れるとき、主体は「グランドキャニオン・スカイウォークは危険である」と亜信しうが、この場合、「グランドキャニオン・スカイウォーク」という概念は、「危険」という性質概念から分離不可能である——当該状況下ではこの亜信念は自動的に形成され、主体は自発的にこれを抑制することはできない——ものの、「危険」という属性概念それ自体は、「青酸カリ」と表示された瓶といった他の対象に帰属させられて新たな亜信念を生起させることも可能である。要するに、亜信念の内容は、全体的な構造を持たない内容と完全に構造化された内容の間の領域を占めており、自由に分離・再配置可能な小さな部分と、より大きな表象的構造に不可逆的に埋め込まれた部分から構成されているのである (Danón 2021: 8513–8514) *10。

そして Danón によれば、上記の仮説には Mandelbaum による結合の問題と推論の問題とを免れうるという利点がある。結合の問題については、Rozin の実験の場合、「この瓶の中身は危険である」という半構造化された命題的内容を持つ亜信念を想定すれば良い。この内容は、「この瓶の中身」という概念が「危険」という属性概念と不可分であると同時に、「危険」という属性概念それ自体はこの内容から分離可能であり意味を保持したまま他の内容に再配置されうるという点で、半構造化されている。このとき主体は、たしかに瓶の中身に「危険」という性質を帰属させており、他の物体には帰属させてはいない (Danón 2021:

*10 ではなぜ亜信念においては概念同士が堅固に結合しているのか。その理由として Danón (2021: 8514–8515) は二つの仮説を挙げている。第一に、人間はそうように自動的に表象する生得的な傾向を有するというものである (トラと対峙したとき、その危険性について証拠を収集したり思案したりする時間を浪費せずに、瞬時にそれを危険であると表象する傾向は有利であったと考えられる)。第二に、主体がaを知覚するたびに、aはFであると信じるという自動的で硬直的な傾向を経験によって獲得するというものである (ある生物が、aはFであることを繰り返し経験し、aに概念Fを何度も帰属させるようになると、その生物はそのような帰属を行うことに慣れ、その結果、その生物は、その事実を示す証拠がない場合でも、aを知覚するたびに、aがFであるという表象を維持する堅固で自動的な傾向を獲得しうると考えられる)。ただし後者の仮説については、習慣によって半構造化された命題的内容が生成されない事例が多くあることを断っている (毎日同じマグカップでコーヒーを飲んでいるからといって、そのマグカップがコーヒーで満たされていると自動的に表象することはない)。

8516)。

また推論の問題についても、亜信念における半構造化された命題的内容は、原理的に、様々な種類の論理的推論に使用することができると考えられる。なぜなら、半構造化された内容も、信念における構造化された内容と同じく、事物が特定の仕方で存在することを表象するという点で真理適合性を有しているし、論理接続詞や量子子がこれらの内容の構成要素として機能しない理由はないからだ。たとえば Rozin の実験の場合、他者に危険な食べ物を与えるべきではないという信念を持ち、また「青酸カリ」と表示された瓶の中身は危険であるという亜信念を持っている主体は、その瓶を客に提供しない方が良いと結論付けることがありうるだろう (Danón 2021: 8516)。

さらに Danón は、半構造化された内容を亜信念に付与することで、なぜ亜信念は信念とは違って反対証拠に対する高い無反応性を有する^{*11}のかを説明できる利点があると主張する。Rozin の実験の被験者が「青酸カリ」と表示された瓶の中身は危険だと亜信しており、そしてその後、当該亜信念と矛盾する証拠を獲得した——たとえば、誰かがその瓶から中身を摂取しても悪影響を受けないのを見たことによって——と仮定しよう。このとき、主体は当該亜信念と矛盾する信念を形成する可能性がある。しかし亜信念が半構造化された内容を持つ場合、そのような証拠は亜信念を改変する力を持たない。この亜信念における「「青酸カリ」と表示された瓶の中身」という概念は、「危険」という属性概念と不可分であるため、主体は、「青酸カリ」と表示された瓶を認識し表象している限り、他の情報を得てもその瓶を危険なものとして表象し続けることになる (Danón 2021: 8517–8518)。

5 問題の所在

最後に、上記の論争を踏まえ今後の課題となりうる観点を提示したい。

第一に、信念が行動を引き起こすには他の心的態度が必要であり、状況によって様々な行動がありうるのに対して、亜信念は一つの特定の運動ルーティン群と堅固に関連付けられているという Danón の主張は、信念は極小感情価を持つ諸概念で構成されるため行動を引き起こしうるという Mandelbaum の主張の決定的な反論になりうるか否かという点である。たしかに Danón の言うとおり、主体が「捕食動物が近くにいる」という信念を持つ場合、状況によって「逃げる」や「助けを求めて叫ぶ」や「戦う」といった複数の行動が可能である。しかし、もしあらゆる概念が〔極小〕感情価を持つとするならば、「捕食動物が近くにいる」という信念は捕食動物 (たとえばトラ) という概念における感情価——「対峙するのは望ましくない」等——を引き継いでいるのだから、主体がその信念を有した時点で回避行動が促進されることになる。このとき、実際に「逃げる」か「戦う」か等の複数の行動がありうるとしても、それは亜信念も同じである。主体が「〔青酸カリの表示のある〕この瓶の中身は危険である」という亜信念を抱くときには同時に回避行動が促進されることになるが、しかし Gendler の説明に従えばそれはあくまで促進であって実行ではないのだから、主体が実際にその中身を摂取しないのか、あるいは躊躇しつつ摂取するのかまでは決定されていないと考え

^{*11} 一般的な理解では、信念は、それに反する証拠が提示された場合、変更されうるが、亜信念はそうではない。この特徴は Gendler (2008b: 565–566; 570–571) 自身も指摘しているが、明確な理由は提示していない。

られる。であるならば、Mandelbaum の反論を崩すには、すべての概念が極小感情価を有しており、それらによって構成される信念もその極小感情価を継承しているという前提を否定しうるかが争点となるだろう。

第二に、信念が理由や証拠に対する反応性を有するのに対して、亜信念はそうではない、という見解——既述のとおり Danón は自身の「半構造化された命題的内容」説の長所として、この見解の説明が容易になるという点を挙げている——は妥当であるか否かという点である。J. C. Jenson は、既存の心理学的研究の成果に基づき、「[反対証拠に対する反応性の有無という]この区別は過大評価されている」(Jenson 2015: 15)と主張する。L. Ross らの研究によれば、本物の自殺メモと偽の自殺メモを複数提示され、本物を識別するよう求められた被験者のうち、ランダムに選択された一部の被験者は 25 件中 24 件を正しく識別していたと伝えられ、他の被験者は 25 件中 10 件のみを正しく識別していたと伝えられたが、実験の最後に、伝えられた結果は完全に偽りであったと伝えられたところ、成功のフィードバックを受けた被験者は、そのテストにおいて自身は他者よりも正確に識別することができ、また今後のテストでもそうであると考えた(Ross et al. 1975)。つまり、前者の被験者はテストの結果を通知されたことで「私は本物の自殺メモと偽の自殺メモとをかなり正確に識別することができる」という信念を抱いたが、その後に反対証拠を提示されたにもかかわらず、その信念を維持しているのであって、これは信念が必ずしも証拠に反応するわけではないことを示唆している。加えて Jenson は、反対に、亜信念が証拠に反応しうることを示唆するものとして B. Stewart らの実験を挙げている。その実験で被験者は、まず「白人」か「アフリカ系アメリカ人」どちらかの顔写真をミリ秒単位で見せられ、次に「銃」か「工具」どちらかの写真をミリ秒単位で見せられるという工程を無作為の順序で繰り返す(黒人→工具→白人→銃→白人→工具……のように)。このとき、被験者は「アフリカ系アメリカ人」の顔写真の後に見せられた「工具」の写真を「銃」と誤答することが多く、被験者には自動的ステレオタイプ化が生じている。しかし、「黒人の顔に対して〈安全〉と考えて反応することを固く決意するため、「私は絶対に黒人の顔に〈安全〉と考えて反応したい」と心の中で繰り返してください」と指示された被験者には、この自動的ステレオタイプ化は生じなかった(Stewart and Payne 2008)。これは、特定の思考習慣を植え付けることで暗黙の偏見や自動的なステレオタイプ化の影響を軽減できることを証明しており、一部の亜信念も他の要因に反応しうることを示唆している(Jenson 2015: 16)。この研究が正しいければ、理由や証拠に対する反応性の有無によって亜信念と信念という二つの実体を措定することの妥当性が棄損される可能性がある。むしろ、単一の統一された「信念」という実体を措定し、それで説明されえないものを説明するために別の実体(亜信念)を仮定するのではなく、そもそも「信念」という素朴概念を完全に放棄し、複数の異なるタイプの心的状態の存在を容認する(Jenson 2015: 20)という方針も選択肢となるだろう。

6 結語

以上の概観から分かるように、亜信念は、それが実在するのであれば人種差別や偏見等を論じる進化心理学・倫理学の発展に大きく貢献しうることはたしかである一方で、いまだ堅固な心理学的カテゴリとし

ての地位を認められているとは言い難い。しかし、新たなカテゴリとして「亜信念」が提示されたことで、〈信念－欲求〉説という従来の説明的枠組みや、信念等の既存のカテゴリそのものの再考を促す契機が生じたことは明らかである。Currie および Ichino (2012: 797)の言うように、「脳活動の画像を見て「これは信念の発火であって結局のところ亜信念ではない」と断言できるわけではない」一方で、Kriegel (2012: 476)の言うように、「今後数年間で哲学者や認知科学者たちによってさらに多くのことが述べられるのは間違いない」だろう。亜信念に関する研究は発展途上であり、本稿が本邦における亜信念研究の進展の一助になれば幸いである。

参考文献

- [1] Bargh, J., Chen, M. & Burrows, L. (1996). “The Automaticity of Social Behavior”. *Journal of Personality and Social Psychology* 71(2): 230–44.
- [2] Currie, G. & Ichino, A. (2012). “Aliefs Don’t Exist, Though Some of their Relatives Do”. *Analysis* 72(4): 788–798.
- [3] 千葉将希 (2021) 「進化倫理学の新たな擁護——道徳的亜信念からの論証」『社会と倫理』 36 号、87–105 頁.
- [4] Danón, L. (2021). “The Content of Aliefs”. *Synthese* 198: 8503–8520.
- [5] Davidson, D. (1963). “Actions, Reasons and Causes”. *The Journal of Philosophy* 60(23): 685–700.
- [6] Fodor, J. (1987). *Psychosemantics*. Cambridge, MA: MIT Press.
- [7] Fürst, M. (2023). “Implicit bias and qualiefs”. *Inquiry* 68(10): 4027–4060.
- [8] Gendler, T. (2008a). “Alief and Belief”. *Journal of Philosophy* 105(10): 634–663.
- [9] ———. (2008b). “Alief in Action (and Reaction)”. *Mind & Language* 23(5): 552–585.
- [10] Hodge, K. M. (2011). “On Imagining the Afterlife”. *Journal of Cognition and Culture* 11(3): 367–389.
- [11] Hubbs, G. (2013). “Alief and Explanation”. *Metaphilosophy* 44(5): 604–620.
- [12] Hung, T. W. (2023). “Why Human Prejudice is so Persistent: A Predictive Coding Analysis”. *Social Epistemology* 37(6): 779–797.
- [13] Jenson, J. C. (2015). “The Belief Illusion”. *The British Journal for the Philosophy of Science* 67(4): 965–995.
- [14] Kriegel, U. (2012). “Moral Motivation, Moral Phenomenology and the Alief/Belief Distinction”. *Australasian Journal of Philosophy* 90(3): 469–486.
- [15] Kriegel, U. (2015). *The Varieties of Consciousness*. New York: Oxford University Press.
- [16] Lebrecht, S., Bar, M., Feldman Barrett, M., & Tarr, M. J. (2012). “Micro-valences: Perceiving Affective Valence in Everyday Objects”. *Frontiers in Psychology* 3: Article 107, 1–5.
- [17] Lebrecht, S., & Tarr, M. (2010). “Defining an Object’s Microvalence Through Implicit Measures”.

- Journal of Vision* 10(7): 966.
- [18] Mandelbaum, E. (2013). “Against Alief”. *Philosophical Studies* 165(1): 197–211.
- [19] Millikan, R. (2004). *Varieties of Meaning*. Cambridge, MA: MIT Press.
- [20] Ross, L., Lepper, M. R. & Hubbard, M. (1975). “Perseverance in Self-perception and Social Perception: Biased Attributional Processes in the Debriefing Paradigm”. *Journal of Personality and Social Psychology* 32: 880–892.
- [21] Rozin, P., Millman, L. & Nemeroff, C. (1986). “Operation of the Laws of Sympathetic Magic in Disgust and Other Domains”. *Journal of Personality and Social Psychology* 50(4): 703–771.
- [22] Stewart, B. & Payne, K. (2008). “Bringing Automatic Stereotyping under Control: Implementation Intentions as Efficient Means of Thought Control”. *Personality and Social Psychology Bulletin* 34: 1332–1345.
- [23] Williams, L. & Bargh, J. (2008). “Experiencing Physical Warmth Promotes Interpersonal Warmth”. *Science* 322: 606–607.

著者情報

田村 歩（茨城工業高等専門学校）