

# 思考の図と地：フレーミングによる肯定・否定の非対称性

服部 雅史

## Figure and Ground in Thinking: The Affirmation–negation Asymmetry as a Consequence of Framing

Masaki Hattori

abstract

Logically speaking, affirmation and negation are symmetrical in the sense that they are complements of each other and symbolically exchangeable. Psychologically speaking, however, they are not symmetrical at all. An affirmed event or the occurrence of an incident is generally a target of attention and in the foreground of the mental world. In contrast, a negated event or the nonoccurrence is usually one of others in the background. The contrast between affirmation and negation is not due to the nature of the environment, but is due to the nature of our cognition. The asymmetry is assumed to stem from the way we view the outer world, which is called the figure–ground framing. The effect is similar to some framing effects known in decision-making studies, and the asymmetrical contrast is reminiscent of the notion of figure-and-ground proposed in visual perception. Extending this idea across a wide variety of phenomena in psychology, including reasoning, judgment, decision-making, cognitive development, and social cognition, this study proposes a new theoretical framework.

“Before deciding that question I had grasped the significance of the silence of the dog, for one true inference invariably suggests others. The Simpson incident had shown me that a dog was kept in the stables, and yet, though some one had been in and had fetched out a horse, he had not barked enough to arouse the two lads in the loft. Obviously the midnight visitor was some one whom the dog knew well.”

Sir Arthur Conan Doyle (1893). *Silver Blaze*  
in *The Memoirs of Sherlock Holmes*.

特別な能力を持ったシャーロック・ホームズのような人は別として、一般人にとっては、ものごとの生起に注意を向けるのは容易であるが、通常、ものごとの不生起は見逃されやすい。このこと

は動物にとっても当てはまるようである。Jenkins & Sainsbury (1969, 1970) は、ハトを使った条件づけ学習において、特徴の有無が**非対称的** (asymmetrical) であることを実証した。ハトに何らかの特徴のあるキー (例：ドットが描かれたキー) と特徴のないキー (例：何も描かれていないキー) の二つのキーを提示するとき、特徴のある方をつつくと餌がもらえる事態では学習は容易だが、特徴のない方をつつくと餌がもらえる事態では、同じ弁別学習がずっと困難になる。Newman, Wolff, & Hearst (1980) は、**特徴肯定性効果** (feature-positive effect) と呼ばれるこの現象が大学生にも見られることを示した。6個の記号 (○、□、△、T、☆、X) のうちの4個が描かれた90枚のカードが順次提示されて、各カードが「よい」(good) か「よくない」(not good) かを当てる課題で、△を含むカードが「よい」とされた場合、正解に到達するまでに見た枚数は平均 17.8 枚 ( $n=16$ ) であったが、△を含まないカードが「よい」場合は 62.9 枚 ( $n=16$ ) であった。

特徴の存在と不在、事象の生起と不生起は、論理的には**相補的** (complement) な関係である。一方を命題  $X$  とすれば、もう一方は  $\neg X$  ( $X$  の否定) と表すことができるため、 $Y = \neg X$  と置換すれば形式上は肯定と否定を入れ替えることができる。しかし、特徴肯定性効果は、肯定と否定が心理的には対称ではないことを示している。



図1 ルビンの盃  
(Rubin, 1921 より引用)

肯定と否定の関係は、知覚現象について Rubin (1915/1958, 1921) が指摘した「**図と地**」(figure and ground) の関係に似ている。一般に、肯定 (生起・行為) は注意の対象であり、否定 (不生起・非行為) はそれ以外の背景である。注意の向かない背景事象は曖昧になり、詳細な認知的処理を受けにくい。図と地の概念は、有名な**ルビンの盃**(Rubin's vase; 図1)によって象徴的に表される。Rubin (1915/1958) は、形態の知覚において図と地が異なるはたらきをすることを指摘した。「地は、ある意味で形を持たない」(Rubin, 1915/1958, p. 194) としているが、このことは図2を見るとわかる。放射状の直線でできた十字を図と捉えれば、地の部分の同心円には切れ目がないように見えるが、同心円の部分を図と捉えれば、背景には一様な放射状の直線が広がる (図3)。

本論文では、この「図と地」の概念を知覚的なものから論理的なものへ拡張し、これまで推論や判断の異なる側面とされていた現象が、共通の特性を持つことを指摘する。取り上げる錯誤やバイアスは、心理学の各分野で既成のバイアスとしての地位を確立しているものばかりで、心理学におけるいわば常識であるが、そのような常識的知識を疑い新しい光を当てることが心理学の新展開の契機となることを期待している。

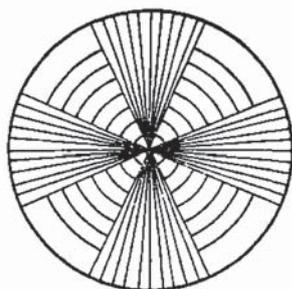


図2 図と地の関係  
(Rubin, 1915/1958 より引用)

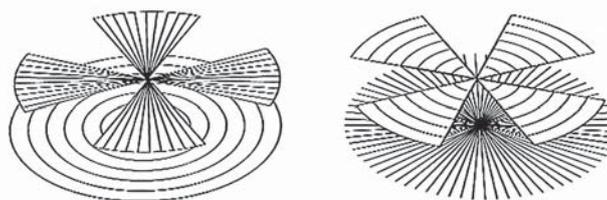


図3 図2の「図と地」の見える図式表現

## I. 図地フレーミングと確率推論

Levin & Gaeth (1988) は、「赤身 75%」と表現された挽き肉は、「脂身 25%」の挽き肉より「高品質」「脂っこくない」と評価されることを示し、**属性フレーミング** (attribution framing) と名付けた。赤身を「図」と見れば、脂身は「地」となり、赤身の肯定的な属性が強調される。逆に脂身を図と見れば、同じ対象でも脂身の否定的な属性が強調されて認識される。同様に、Sanford, Fay, Stewart, & Moxey (2002) は、「25% 乳脂肪を含むヨーグルト」(“containing 25% fat”)の方が「75% 無脂肪のヨーグルト」(“being 75% fat free”)よりも非健康的と判断されることを示した。本研究では、このような肯定・否定の間の非対称性を発生させる心理的枠組みを**図地フレーミング** (figure-ground framing) と呼ぶことにする。

本節では、図地フレーミングが確率推論のエラーやバイアスを説明するのに役立つことを見ていく。ここでは、因果帰納における  $d$  セル無視と基準率錯誤を取り上げる。**因果帰納** (causal induction) とは、たとえば、「イグニッションキーを回す ( $C$ ) とエンジンがスタートする ( $E$ )」といったような二つの事象の間の因果関係を表す規則を、事例の観察からその生起頻度に基づいて形成することをいう。ただし、ここでの因果は厳密なものではなく、認知主体が素朴に方向性を想定していればよいものとする。**原因候補事象** (candidate cause)  $C$  と**結果事象** (effect)  $E$  の生起／不生起の組み合わせ、すなわち**共変** (covariation) は、 $2 \times 2$  分割表 (表1) であらわされるが、共変情報 (セルの事象頻度) は形成される因果規則の強さに影響する。

表1 二事象の共変情報

	$E$	$-E$
$C$	$a$	$b$
$-C$	$c$	$d$

注)  $C$ ,  $-C$  はそれぞれ原因候補事象の生起と不生起を、 $E$ ,  $-E$  は結果事象の生起と不生起を表す。

共変に基づく因果帰納の規範的モデルとして、最もよく知られているのが、 $\Delta P$  と呼ばれる次のモデルである。

$$\Delta P = P(E|C) - P(E|\neg C) = \frac{a}{a+b} - \frac{c}{c+d}$$

このモデルは、結果の生起確率を原因が起きた場合と起きなかった場合に分けて、前者の方が後者より高い場合には、その分だけその原因に結果をコントロールする力があるとする考え (Jenkins & Ward, 1965) に基づく。この考え自体は理に合っているように思われるが、 $\Delta P$  は必ずしも人間の判断データによく適合するわけではない (e.g., Mandel & Lehman, 1998; Perales & Shanks, 2007; Ward & Jenkins, 1965)。

Hattori & Oaksford (2007) は、**二要因ヒューリスティック** (dual factor heuristic; DFH) と呼ばれる記述的モデルを新たに提唱し、 $\Delta P$  と DFH の予測が直交する刺激セットを用いた実験によって、 $\Delta P$  にはまったく記述力がないことを明らかにした。DFH モデルは次の通りである。

$$H = \sqrt{P(E|C)P(C|E)} = \frac{a}{\sqrt{(a+b)(a+c)}}$$

このモデルの顕著な特徴は、 $d$  セルを含まないことである。2×2 分割表の四つのセルの中で、 $a$  セル影響力が最も強く、 $d$  セルが最も弱いことは古くからよく知られていた (e.g., Schustack & Sternberg, 1981; Smedslund, 1963)。DFH は、この知見を最も極端な形で取り入れたモデルと見ることができる。

因果帰納において  $d$  セルが無視 (より穏当に言えば軽視) されるのは、 $d$  セルがまさに「地」の事象であるからと考えることができる。因果関係の有無に興味があるとき、通常、原因や結果事象の一方、または両方が起きたときに注意が向く。先の「イグニッションキー」の例では、「キーも回さず、エンジンも動かない」事象が  $d$  セルに相当するが、このように「何も起こらない」状態は、両事象の関係についての推論をトリガーしない。因果の概念には、何かを「引き起こす」ことが含まれるが、それはすなわち「何も起きていない」状態を変化させることである。つまり、因果の概念の背景には、何も起こらない状態を基準とするという暗黙の前提があると考えられる。もちろん、実際には世界のあらゆるところで常に何かが起こっていることから、厳密な意味で何も起こらない状態というのは存在しない。認知的な基準になる状態 (デフォルト状態) を想定するという意味である。この前提こそが、 $d$  セル事象を「地」とみなすことに他ならない。

これと密接に関連する問題として、 $d$  セル頻度のカウント困難性がある。「キーの回転 ( $C$ )」も「エンジンの発動 ( $E$ )」も行為であるが、「キーの非回転 ( $\neg C$ )」や「エンジンの非発動 ( $\neg E$ )」は状態である。行為と行為の組み合わせ事象 ( $a$  セル) や行為と状態の組み合わせ事象 ( $b$  セルと  $c$  セル) の頻度をカウントすることは可能であるが、状態と状態の組み合わせ事象 ( $d$  セル) を 1 回、2 回と数えるのは困難である。せいぜい、ある日のある時刻といったように観測時刻を固定するか、時間間隔を一定の幅に区切った上で状態を記録するくらいしか方法はないだろう。このことは、 $d$  セルだけが情報として特異であることを示しており、 $d$  セルを地とすることの合理的根拠になる。

以上の因果帰納とはまったく異なる文脈で研究されてきた**基準率錯誤** (base rate fallacy) についても、同じように論理的な図と地の観点から解釈することが可能である。Tversky & Kahneman (1980)

は、確率判断課題においてしばしば課題中に示された基準率が無視されることを指摘し、そのためにエラーが起こるとして**基準率無視** (base rate neglect) と呼んだ。しかし、Hattori & Nishida (2009) は、基準率は無視されず、エラーが起こるのはターゲットとなる2事象の生起確率がほぼ等しいと仮定されるためであるとした (**等確率仮説** equiprobability hypothesis)。彼らの実験では、解決者の等確率性の仮定を崩して課題の確率構造が正しく把握されるようにすれば、正答率が劇的に上昇することが示された。しかし、等確率性と矛盾する確率情報が課題の中で与えられても、それがデフォルトの等確率仮定を上書きするために使われない理由については十分に解明されていない。そこで以下では、この点について図地フレーミングの観点から検討する。

基準率錯誤が認められる課題（以下では基準率課題と呼ぶ）は以下のようなものである。これは、後述の実験1, 2の統制条件で使用されたもので、Eddy (1982) に準じている。なお、以下の下線とその番号は、以下の説明のためのもので、実験課題に表示されたわけではない。

(1) 集団検診に参加する40歳の女性が乳がんにかかっている確率は1%です。 (2) 乳がんにかかっているとき、検査で陽性とする確率は80%です。つまり、乳がんにかかっているにもかかわらず20%の確率で陰性と出ます。また、(3) 乳がんでないにもかかわらず、誤って陽性とする確率が9.6%あります。さて、(4) 40歳のある女性が検診で陽性となりました。この女性が実際に乳がんにかかっている確率はどれくらいでしょうか。 (正解：7.6%)

乳がんであることを  $H$ 、検査で陽性と出るとを  $D$  とすると、この課題で与えられる確率情報は、(1)  $P(H) = 0.01$ , (2)  $P(D|H) = 0.8$ , (3)  $P(D|\neg H) = 0.096$  であり、求めるべき確率は (4)  $P(H|D)$  である。この中で、(3)  $P(D|\neg H)$  だけが否定命題（「乳がんでない」）に基づく条件つき確率である。この課題において、「乳がんである」ことは明らかにターゲット事象であるので、その否定は「地」に相当する。よって、これを「地をベースとする確率情報」と呼ぶ。

地をベースとする情報は、否定命題のスコープ（命題が対象とする範囲）に光を当てる必要がある。しかし、Rubin が「視覚的な地に形がない」と表現したように、論理的な地も「形」、すなわち、命題のスコープが曖昧になる。そのため、この情報の扱いには困難が伴うことになると考えられる。その結果、 $P(D)$  の推定が難しくなり、結果として、等確率性が仮定されるのではないだろうか。

以上の考えをまとめると、以下の三つの仮説になる。第1に、論理的な図地反転は容易には生じにくい。第2に、地は地平（スコープ）が曖昧になるので、「地をベースとする確率情報」、すなわち、「地」とみなされる事象の条件つき確率は使用が困難になる。第3に、確率情報が明示的（効果的）に与えられないとき、ターゲット事象の等確率が仮定される。以上の考えが正しければ、確率情報を地ベースの形態で与えることを避けることにより正答率が上昇するはずである。そこで、この予測を実験によって確認した。

## II. 実験1

従来、基準率無視を起こすと考えられてきた確率判断課題（前述の乳がん問題など）では、否定命題の条件つき確率（地をベースとする確率）が用いられている。もしこの課題の困難さが、人々に基準率を無視する傾向があるからではなく、人々が図地フレーミングによって課題を理解する傾向によるのだとすれば、地ベースの確率情報を別の適切な形態で提示すればエラーが減少するはずであ

る。そこで、この予想を確かめるため実験を行った。

## 方法

**実験参加者および実験計画** 立命館大学の学生 97 人（男性 40 人、女性 57 人、年齢： $M=20.6$ ,  $SD=1.9$ ）が実験に参加した。参加者は、統制条件、実験条件のいずれかに無作為に割り当てられた。

**課題および手続き** 統制条件では、前述の課題を使用した。実験条件では、次の変形課題を使用した。

集団検診に参加する 40 歳の女性が乳がんにかかっている確率は 1% です。<sup>(5)</sup> 検診参加者のうち、乳がんにかかっている、しかも検査で陽性とする人は全体の 0.8% です。つまり、乳がんにかかっているも陰性とする人が全体の 0.2% います。一方、<sup>(6)</sup> 検診参加者全体の 9.5% は、乳がんにかかっているのに陽性とするのがわかっています。

この課題では、地ベース情報を避けるため、(3)  $P(D|\neg H)=0.096$  の代わりに (6)  $P(\neg H, D)=0.095$  という確率情報を与えている。これに伴い、正陽性率 (2)  $P(D|H)=0.8$  を同時確率 (5)  $P(H, D)=0.008$  に置き換え、全体集合を基準とする確率にそろえた。解答に際しては、計算するのではなく直感的に答えるよう求めた。制限時間は設けなかった。

## 結果および考察

回答の中央値は、実験群 (1.5%) の方が統制群 (70.4%) より有意に低かった (Wilcoxon Test)、 $Z=4.30$ ,  $p=1.0 \times 10^{-4}$ ,  $r=.437$ 。よって、仮説が支持された。

しかし、実験条件で正答率が高かったのは、実験条件の課題が単に統制群より計算が簡単だった可能性を排除できない。すなわち、 $P(H, D)$ ,  $P(H, \neg D)$ ,  $P(\neg H, D)$ ,  $P(\neg H, \neg D)$  をそれぞれ  $p_a, p_b, p_c, p_d$  とすると、求めるべき確率は、実験群では  $P(H|D)=p_a/(p_a+p_c)=(5)/[(5)+(6)]$  と容易に計算できるのに対して、統制群では、 $p_a=P(H)P(D|H)=(1) \times (2)$ 、 $p_c=P(D|\neg H)(1-P(H))=(3) \times [1-(1)]$  であるため、 $P(H|D)=p_a/(p_a+p_c)=(1) \times (2)/[(1) \times (2) + (3) \times [1-(1)]]$  と計算する必要がある。この違いが実験群の正答率を高くしたという解釈も可能である。そこで、このことを確認するため、実験 2 を実施した。

## Ⅲ. 実験 2

実験 1 で正答率が上昇したのは、課題の計算が簡単になったためという可能性が排除できない。そこで、オリジナル課題と同程度に計算が複雑な課題を用いて、実験 1 の結果の再現を試みた。

## 方法

**実験参加者および実験計画** 立命館大学の学生 145 人と金沢大学の学生 45 人、合計 190 人の学生（男性 75 人、女性 113 人、不明 2 人、年齢： $M=20.4$ ,  $SD=2.2$ ）が実験に参加した。参加者は、統制条件、実験条件のいずれかに無作為に割り当てられた。

**課題および手続き** 統制条件の課題は、実験 1 と同じものを使用した。実験条件には次のような課題を与えた。

集団検診に参加する40歳の女性が乳がんにかかっている確率は1%です。乳がんにかかっているとき、検査で陽性とする確率は80%です。つまり、乳がんにかかっているにもかかわらず20%の確率で誤って陰性と出ます。また、検査の誤りには偏りがあり、<sup>(7)</sup> 誤って陽性とする人の数は、誤って陰性とする人の47.5倍にも達します。

この課題では、与えられる偽判定オッズを $(7) f = p_c / p_b = 47.5$ とすると、 $p_c = f P(H) (1 - P(D|H))$   
 $= (1) \times (7) \times [1 - (2)]$ ゆえ、 $P(H|D) = p_a / (p_a + p_c) = (1) \times (2) / [(1) \times (2) + (1) \times (7) \times [1 - (2)]]$   
 として計算する必要があり、数式の形と項の数から判断する限り、計算の複雑さは実験1の実験条件と同等以上と考えられる。

### 結果および考察

未回答の1人を分析から除外した。回答の中央値は、実験群(5%)の方が統制群(50%)よりも有意に低かった(Wilcoxon Test)、 $Z = 3.02, p = .0024, r = .220$ 。

以上の結果は、実験1で正答率が上昇したのは、地ベース情報の使用が困難であったためであることを示唆する。ただし、効果量(.220)は実験1の結果(.437)より小さかったことから、課題の計算の複雑さも、正答率の上昇に一定程度は貢献したと考えるのが妥当かもしれない。

## IV. 図地フレーミングと論理

ここまでは、確率的な思考の中で図地フレーミングが重要な役割を果たしていることを示す証拠を確認してきた。以下では、演繹、仮説検証、類似性、選好性(意思決定)、心の理論、社会的推論などの様々な心理学的現象が、図地フレーミングの枠組みで再解釈できることを順に見ていきたい。

肯定と否定の非対称性を最初に実験的に扱ったのは、おそらく Wason (1959) であろう。先駆的な実験によって彼は、文の処理にかかる反応時間と誤答率を指標として、肯定情報の処理が否定情報より容易であることを明らかにした。ある文が真ならば、その文の否定は偽である。逆に、文が偽ならその否定は真である。よって、肯定性と真理値の組み合わせによって、 $2 \times 2 = 4$ 種類の文(真肯定文、偽否定文、偽肯定文、真否定文)を作ることができる。これらの文に情報処理の容易性の違いがあるかどうか調べられた。参加者に図4のような「状況」を見せ、真条件では「文章が状況に一致するように語を選択する」ことを要求した。偽条件では、文章が矛盾するように選択することが求められた。与えられた文は、肯定条件では「4に{黄色/緑色}があり、かつ3に{赤色/黒色}がある」(“There is both Yellow/Green in 4 AND Red/Black in 3”)といったもの、否定条件では「4に{黄色/緑色}があり、かつ3に{赤色/黒色}がある、というわけではない」(“There is NOT both Yellow/Green in 4 AND Red/Black in 3”)といったものであった。24名の参加者には、4種類の文について、それぞれ与えられた二つの選択肢から正しい語を選んで、できるだけ速く正確に○を付けることが要求された(参加者内計画)。実験の結果、平均反応時間が最も短かったのは真肯定文(8.99秒)で、次に偽肯定文(11.09秒)、真否定文(12.58秒)と続き、最も長かったのが偽否定文(15.17秒)であった。つまり、肯定情報の方が否定情報より処理に要する認知的負荷が少ないことが示された。この実験は、肯定と否定の非対称性を明らかにしたものとみることができる。

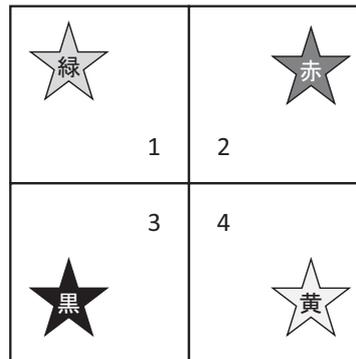


図4 Wason (1959) で使用された「状況」

この実験の翌年に Wason (1960) が考案した 2-4-6 課題では、図地フレーミングの影響がより明確になる。この課題では、しばしば**確認バイアス** (confirmation bias) が観察される。すなわち、参加者は、自分の仮説に合致する事例によって仮説を確認しようとする傾向を持つが、仮説に合致しない事例を使って仮説の真偽を確認しようとはしない。たとえば、三つ組整数を規定する規則について、「上昇する偶数列」という仮説を立てたとき、正事例 (4-6-8 や 10-12-14 など) が仮説に合うことを確かめようとすることは多いが、負事例 (1-8-3 や 51-20-11 など) が仮説に合わないことを確かめようとすることは少ない。われわれが注意を向けるのは、仮説が正しいことや、仮説の正事例、さらには事例に対する肯定的なフィードバックである。仮説が正しくないことや仮説の負事例は背景に退く。すなわち、仮説や事例の肯定的側面が図であり、仮説や事例の否定的側面は地となる。

特に、図地フレーミングとの関係で注意すべき点は、二種類の確認 (表2) の心理的意味が異なるということである。確認バイアスを論じるときに焦点化されるのは、正事例に対して肯定的フィードバックを得る状況であるが、負事例に対して否定的フィードバックを得る場合も確認である。反対に、反証には、正事例に対して否定的フィードバックを得る場合と、負事例に対して肯定的フィードバックを得る場合がある。しかし、同じ確認であっても、正事例検証に対する肯定的フィードバックと負事例検証に対する否定的フィードバックは心理的な意味合いが異なる。「図」としての確認には注意が向いてバイアスの対象になるが、「地」としての確認には注意が向かない。

表2 検証方略と確認・反証の関係

検証方略 (予想結果)	実際の結果	
	支持	不支持
正事例検証 (支持)	確認	反証
負事例検証 (不支持)	反証	確認

なお、Wason (1960) は、このバイアスを確認または**検証** (verification) ということばで紹介したが、確認バイアスは仮説検証「方略」に関するバイアスである。しかし、確認か反証かは検証「方略」ではなく、検証「結果」を指す。よって、混乱を招くため、Klayman & Ha (1987) は、より正確な**正事例検証方略** (positive test strategy) という呼称を提案した。ところが、現在でも確認バイアスという用語は使われている。それは、おそらく直感に訴える響きがあるせいと思われるが、もしそうだとすれば、それも図地フレーミングのせいかもしれない。

図地フレーミングによる確認の非対称性を明確にするのが、ヘンペルのカラス (Hempel, 1945) の

パラドックスである。「カラスならば黒い」という仮説があるとき、「黒いカラス」は仮説を確証する。同様に、「黒くないものはカラスではない」という仮説は、「黒くないカラスでないもの」によって確証される。これらの仮説は論理的に等価な対偶の関係にあるため、「カラスならば黒い」という仮説は、「白い靴」や「赤い傘」によって確証されることになる。この見かけ上のパラドックスを解消するためには、以下に述べるような二つの考え方がある。

第1は、ベイズモデルによる解決で、仮説検証を対立モデルの尤度比較とみなす考え方である (e.g., McKenzie & Mikkelsen, 2000)。いま「黒いカラス」( $D$ )が観察されたとすると、「カラスならば黒い」( $H_1$ )という仮説が本当に正しくて $D$ が観察された場合と、そうではないが(たとえば、本当は「カラスと黒さは独立(無関係)」( $H_0$ )だが)、偶然に $D$ が観察された場合が考えられる。両者を比べると、「カラスの確率」も「黒いものの確率」も小さいときは、前者の方が圧倒的に生起しやすい。ここで、「カラス(あるいは黒いもの)の確率」とは、無作為に何かを抽出したとき、それが偶然にカラス(あるいは黒いもの)である確率を指す。世の中には膨大なものがあるので、これらの確率は非常に小さいこと(稀少性 *rarity*)が期待できる。そのような場合、 $H_1$ という仮説についての期待獲得情報量は、「白い靴」などに比べると $D$ の方が圧倒的に大きくなる。つまり、 $D$ は仮説の確信度を大きく高めるが、「白い靴」は仮説の確信度をほとんど変化させない。そのため、黒いカラスの発見には大きな意味があるが、白い靴の発見にはほとんど意味がない。以上の分析結果を図地フレーミングの考えに則って表現すれば、「稀少性を満たす対象(カラスや黒いもの)は図になり、それ以外の対象が地になる」ということである。

第2の解決は、条件文の確率解釈によるもので、仮説検証を条件文の確率判断とみなす考え方である。近年の条件文推論に関する心理学的研究によれば、条件文「もし $p$ ならば $q$ 」( $p \rightarrow q$ )の確率 $P(p \rightarrow q)$ は条件つき確率 $P(q|p)$ とみなすことができる (e.g., Evans & Over, 2004)。すなわち、「もしカラス( $R$ )ならばそれは黒い( $B$ )」という条件文が正しい確率は、「カラスの集合を全体とみなしたとき、黒いカラスがどれくらいを占めるか $P(B|R)$ 」になる。これを知るためには、白い靴や赤い傘は無関係で、カラス自体を調べる必要がある。全数探索であれサンプリングであれ、カラスを調べない限り、この確率についての情報を得ることはできない。以上の考え方を図地フレーミングの考えに則って表現すれば、「条件文の前件(カラス)と後件(黒いもの)は図になり、それ以外の対象は地になる」ということである。

## V. 図地フレーミングと属性

ここまでは、事象(対象)自体の論理的な肯定と否定を扱ってきたが、対象に付随する特徴(属性)についても類似した議論が成り立つ。ものごとには無数の属性があるが、注意を向けられた属性は重みづけられて処理される。その結果、同じ対象が異なった認知処理を受けることになる。たとえば、リンゴを物体と見れば、色や形などの物理的性質に焦点が当たるが、食べものと見れば、味や栄養などに注意が向く。つまり、焦点化や観点、状況、文脈などの違いによって、類似性や選好性の判断が大きく影響を受けることになる。

哲学者 Goodman (1972) は、類似性という概念を徹底的に批判した。彼が挙げた八つの観点のうち最も重要と思われるのが、この焦点化の問題である。彼は、空港の搭乗受付における荷物の例を

挙げている (p. 445)。まわりの人は荷物の形、大きさ、色、材質などに注目するが、パイロットは重さを、乗客は行き先と持ち主を気にする。その結果、どの荷物が似ているかは、判断の文脈や判断者の立場によって異なることになる。つまり、状況によって類似性は変化する。結論として、このような捉えどころのない概念は、哲学的分析のツールとして役に立たないとした。

Tversky (1977) も同様の問題を扱ったが、むしろ彼は、最も影響力のある類似性のモデルを構築した。彼のモデルは、類似性判断が特徴に基づいて行われるとしており、そのモデルに基づいて、類似性と相違性が**相補的** (complement) でないことを示した。すなわち、通常、*A* と *B* が似ていればいるほど「類似性」評定値は高く、逆に「相違性」評定値は低くなり、両者を足し合わせると一定の値 (たとえば100) になることが期待されるが、必ずしもそうならないことを実証した。比較する対象をよく知っている場合は、両対象の共通特徴も弁別特徴も多く持つことになるが、対象をあまり知らない場合は共通特徴も弁別特徴も少ないことになる。類似性判断では共通特徴に注意が向けられ、相違性判断では弁別特徴に注意が向く。したがって、よく知っているものがペアになっている場合 (**顕著ペア** prominent pair) は、あまり知らない場合 (**非顕著ペア** nonprominent pair) に比べて、より似ていると判断されると同時に、より異なるとも判断されると予想される。60名の参加者の半分 (類似条件) は、二つずつのペアで構成された四つの国のうち、どちらのペアが「似ているか」を判断した。残りの参加者 (相違条件) は、同じ刺激でどちらが「異なるか」を判断した。いずれの条件でも、約7割の参加者が「セイロン—ネパール」でなく「西ドイツ—東ドイツ」を選んだ。つまり、「西ドイツ—東ドイツ」は、似ていると同時に異なっているとも判断されたことになる。他の顕著ペアについても同様の結果を得た。

ものの属性の処理のされ方が影響を与えるのは、類似性の判断だけではない。Shafir (1993) は、意思決定においても同様の現象を見出した。参加者は、以下の二人の親に対する子どもの単独監護権について判断した。

親 A	親 B
平均的な収入	平均以上の収入
平均的な健康状態	子どもとの非常に親密な関係
平均的な労働時間	極めて活発な社会生活
子どもとの適切な関係	多くの仕事関係の出張
比較的安定した社会生活	軽微な健康上の問題

170名の参加者の半分は、どちらの親に単独監護権を「認めるか (award)」を判断し、残りの半分の参加者は、どちらの親に単独監護権を「与えないか (deny)」を判断した。一方に単独監護権を認めることは、もう一方に単独監護権を与えないことを意味するので、両者は相補的な結果になることが期待される。ところが、「認める」場合は64%が親Bを選び、「与えない」場合は55%が親Bを選んだ。つまり、両者を足し合わせると100を超え ( $64+55=119$ )、相補性が成り立たないことが示された。選択肢の中から対象を「選ぶ」場合には、選ぶための理由になるような肯定的な特徴に注意が向くが、対象を「排除する」場合には、排除するための理由になるような否定的な特徴に注意が向く。その結果、肯定的な特徴と否定的な特徴の両方を合わせ持つ極端な選択肢 (**強化選択肢** enriched option) は、特に目立った特徴のない平均的な選択肢 (**貧弱選択肢** impoverished option) に比

べて、選ぶ場合も排除する場合もターゲットになりやすい。

## VI. 図地フレーミングと知識の呪縛

図地フレーミングは、さらに高階の推論、すなわち信念についての信念についてもはたらくと考えられる。Birch & Bloom (2003) は、心的状態の帰属における非対称性バイアスを明らかにした。彼女らは、スマーティー課題 (smarties task; Perner, Leekam, & Wimmer, 1987) の変形版を使って、就学前の幼児がこのような傾向を持つことを明らかにした。実験では、中に小さなおもちゃが入った容器を用意し、自分と相手がそれぞれ中身を知っている場合と知らない場合の  $2 \times 2 = 4$  状況下の参加者内計画で、3, 4, 5 歳児 (各 16 名) が他者の信念を正しく推論できるかどうかテストした。その結果、自分が中身を知っている場合には、相手が実際に中身を知っているにもかかわらず、「知っている」と答える確率が高かった (それぞれ、69% と 48%) が、自分が中身を知らない場合には、相手が中身を知っていれば正しく「知っている」と答えるが、相手が知らないのに間違えて「知っている」と答える確率は低かった (それぞれ、59% と 26%、以上、すべて 3 歳児の場合だが 4 歳児も同様であり、5 歳児はすべての場合に正答率が高かった)。つまり、幼児は、自分が知っていることは他人も知っていると思ってしまうが、自分が知らないことを他人も知らないと思ってしまうことはないことが明らかになった。

このような結果は、「知識の呪縛 (the curse of knowledge)」という概念によって説明される。知識の呪縛とは、もともとは経済学において提唱されたもので、「十分な情報を与えられた主体が、自分の情報を無視することが自分のためになる場合ですら、それを無視することができない」(Camerer, Loewenstein, & Weber, 1989, p. 1232) ために判断にバイアスが生じることを指す。3 歳以下の幼児は、多くの場合、マクシ課題 (Maxi task; Wimmer & Perner, 1983) やスマーティー課題、サリー・アン課題 (Sally-Ann task; Baron-Cohen, Leslie, & Frith, 1985) などの誤信念課題に正解することができないが、それは、「自分が知っていることは、それを知りえないはずの他人も知っていると思ってしまう」傾向を示している。この傾向は、当初は、Piaget にさかのぼる自己中心性 (egocentrism) として理解されてきたが、次第に、他者のみならず自分の過去の心的状態の推論も同様に子どもは苦手であること (Gopnik & Astington, 1988) や、大人も子どもと同様、他者や過去の自分の心的状態の推論が苦手であること (Birch & Bloom, 2007) が明らかにされてきた。このようなことを踏まえ、Birch & Bloom (2003) は、知識の呪縛による説明を提供した。

この説明が自己中心性より優れているのは、第 1 には、上述の実験結果の非対称性を説明できることである。すなわち、「自分より無知の誰かの見え方を認識しようとするときには、知識によってバイアスを受けるが、自分より知識のある誰かの見え方を認識しようとするときには、無知によるバイアスはない」(Birch & Bloom, 2003, p. 285) ことは、少なくとも自己中心性だけでは説明できない。第 2 には、「心の理論 (theory of mind)」(Premack & Woodruff, 1978) の議論に象徴される他者の信念についての推論の問題を、社会的認知を含む他のさまざまな (成人の) 認知的バイアスと同列に、包括的に捉えられることである。知識の呪縛に関連するバイアスとして Birch & Bloom (2004) が挙げているのは、後知恵バイアス (hindsight bias; Fischhoff, 1975)、スポットライト効果 (spotlight effect; Gilovich, Medvec, & Savitsky, 2000)、透明性の錯覚 (illusion of transparency; Gilovich, Savitsky, &

Medvec, 1998)、**ファルス・コンセンサス効果** (false consensus effect; Ross, Greene, & House, 1977) などである。

このような信念に関するさまざまなバイアスは、その少なくとも一部は、同じ認知のしくみ、いわば高階の図地フレーミングによって統合的にとらえられる可能性がある。自分が「知っていること」や「信じていること」には注意を向けることができるが、「知らないこと」や「信じていないこと」に注意を向けるのは難しい。つまり、信念（誤信念も含む）は図になるが、非信念は図になりにくい。信念は、たとえば「『もの  $X$  が場所  $Y$  にある』と信じる」*Believe(Exist( $X, Y$ ))* といった**二階の述語** (second-order predicate) によって表すことができる。さらには、**二次の信念** (second-order belief) と呼ばれる「『太郎はここにあると思っている』と花子は思っている」といった入れ子の信念もある。それに対応するように、フレーミングにも高階のものを想定することが可能で、そこに肯定と否定の非対称性が発生するとみることができる。Birch & Bloom (2004) が挙げている他のバイアス、たとえば後知恵バイアスにも同様のことが起こるかどうかは、実証的に確認する価値があるだろう。後知恵バイアスとは、「事後に振り返るとそれが当然の帰結であったように思う傾向」を指す。何かが起きたとき（たとえば「誰かが亡くなった」とき）には、それが当然起こると思っていた（自分は「亡くなることを予期していた」と感じるが、何も起きなかったときには、何も起きないことを予期していたという思いを抱くことがあまりないとすれば、肯定の信念と否定の信念の間に非対称性が存在することになる。つまり、知識の呪縛は、いわば高階の図地フレーミングによって生じると考えることができるだろう。

## Ⅶ. まとめ：二重フレーム理論の構築に向けて

ここまでで、図地フレーミングが思考のさまざまなバイアスやエラーと関連がある可能性を見つけた。図地フレーミングが広範かつ強力な認知傾向であるとするれば、それには認識論的効用があると考えられるべきであろう。そこで、図地フレーミングの明るい面について考えてみたい。図地フレーミングのメリットとしてもっとも考えやすいのは、認知的処理にかかる負荷の軽減である。因果帰納における図地フレーミングの例として、二要因ヒューリスティックモデルを紹介したが、Hatori & Oaksford (2007) は、 $d$ セルを無視することの適応的側面について、コンピュータ・シミュレーションに基づいて分析している。稀少性（原因も結果も生起確率が小さいこと）を前提とすれば、 $d$ セルを無視しても二事象間の相関関係をある程度の正確さで計算することができる。むしろ、大量の  $d$ セル事例を記憶に保持しておく必要がないため、記憶の負荷が軽減されるというメリットがある。あるいは、記憶に同じ負荷をかけるなら、他のセルのサンプルを稼ぐことができるため、より精度の高い推定が可能になる。つまり、認知的負荷低減と出力の正確さの間のトレードオフの解決として、 $d$ セル無視、つまり、図地フレーミングが存在している可能性があるといえよう。

もう一つ、重要な論点は、図地フレーミングが**フレーム問題** (frame problem; McCarthy & Hayes, 1969; Pylyshyn, 1987) を解決するための、もしかしたら唯一のしくみなのかもしれないということである。フレーム問題とは、かつて人工知能研究が提示した哲学の重要な未解決問題で、有限の処理能力しかもたないエージェントが、時々刻々と変化する環境の中で問題に対処することの本質的困難性を指す。特に、環境の中で、エージェントが何らかの行為をしたり外界で何かが生じたとき、

残りの圧倒的多数のことがらには変化しないが、それらのすべてを逐一記述していたら記述の量が爆発してしまうというのが当初の懸念であった。このような指摘から、動的環境におけるリアルタイムの活動のために必要な記述量や計算量の観点に基づく問題が顕在化した。人工知能に図と地の概念はない。いわば、すべてが図の状態である。この点が、フレーム問題を顕在化させている可能性がある。現実問題として、図地フレーミングを人工知能にどうインプリメントできるかは、現段階では明らかではないが、何らかの形で焦点化（図）と無視（地）のしくみを取り入れない限り、知的な機械の実現は難しいように思われる。

その反面、図地フレーミングには、当然ながら短所もある。図地フレーミングは一種のヒューリスティクスである。ヒューリスティクスの宿命として、状況によってはうまく働かず失敗にいたる。本稿で見てきたさまざまな認知的バイアスは、その結果である。しかし、図地フレーミングがいわゆる「ものの見方」の一つであるならば、見方を変えることによってバイアスを減らすことは可能なのかという疑問が生まれる。その答えのヒントは、「IV. 図地フレーミングと論理」の節で紹介した2-4-6課題を用いた実験結果にある。Tweney et al. (1980) は、この課題の変形版が正答率を劇的に上昇させることを発見した（実験4）。DAX-MED版と呼ばれる変形2-4-6課題では、各事例（三つ組整数）は、DAX（「上昇する数列」）とMED（DAX以外の数列）と呼ばれる二つの相補的なカテゴリーのいずれかに分類された。すなわち、参加者は、各事例が規則に合うか合わないかではなく、DAXであるかMEDであるかのフィードバックを受けた。その結果、最初の報告での正答率が60%（12/20）になった（Tweney et al., 1980の実験3の単一仮説条件では3/20=15%であった）。同様の結果は複数の研究から得られており（e.g., Tukey, 1986; Wetherick, 1962; Wharton, Cheng, & Wickens, 1993）、発見すべき規則が相補的であることがポイントと考えられている（Gale & Ball, 2006）。

このTweney et al. (1980)の実験結果が示唆することは、図地フレーミングの影響力が何らかの要因によって低減され得るということである。対象（たとえば、三つ組整数を分類するカテゴリー）を、ターゲット（たとえば「上昇する数列」）とそれ以外に分ける場合と、二分法的に対等に（DAXとMEDのように）分ける場合では、起動するフレーム（いわば世界の見方）が異なるのではないだろうか。すなわち、肯定・否定表現によって起動されるフレームと、対比的表現によって起動されるフレームは、推論・判断のあり方を変える可能性がある。ここでは、前者をAフレーム（Aは注意性 attentional、連合性 associationalの頭文字）、後者をBフレーム（Bは均衡的・不偏的 balancedの頭文字）と呼ぶことにする。Aフレームは、肯定否定表現と親和性が高く、図地関係を明確化し、図地フレーミング効果を生じさせる。一方、Bフレームは、二項対立的表現と親和性が高く、スコープが明確になるため規範的な判断や推論を引き出しやすい。また、Aフレームは、肯定事象を焦点化して否定事象を背景に追いやるため、肯定事象と否定事象は心理的に可換ではなくなり、非対称性が生まれる。しかし、Bフレームは、両事象を対比的に捉えることを促すため、両者は心理的に可換で対称的である（表3参照）。

特に、因果帰納においては、両フレームは推論モードの違いに直結する。Aフレームによって因果関係を捉えるときは、原因候補事象Cと結果事象Eの間の関連性を検出するモードになると考えられる。この**関連性モード**（relevance mode）においては、関連性がない状態、すなわち両事象が独立の状態がデフォルトであり、その基準と比べてどれくらい共変関係がみられるかが問題となる。関連性の指標としては四分点相関係数 $\phi$ を使うことが可能であろう。そのとき、CとEの稀少性を仮定すれば、 $\lim_{d \rightarrow \infty} \phi$ （dセル値を無限大に発散させたときの $\phi$ の極限）としてのDFHモデル（Hattori &

表3 検証方略と確証・反証の関係

	A フレーム	B フレーム
	肯定事象の焦点化	比較の観点
	非対称（心理的に非可換）	対称（心理的に可換）
	関連性モード	差異化モード
	<i>d</i> セル無視	<i>d</i> セル考慮
	単極的因果性（有／無）	双極的因果性（抑制／無関連／促進）
認知システム	ヒューリスティック的	分析的
外界への侵襲性	観察	介入
否定の方法	明示的否定	潜在的否定
記述特性	属性	行為
動態性	静的	動的
コミットメント	コミット無し	コミット有り
能動性	受動的	能動的
ターゲット事象の基準率	稀少	非稀少

Oaksford, 2007) が記述力を持つことが予想できる。その場合、出力は単極性（ゼロ～正）となる。つまり、関連性があるかないか、あるとすればどれくらい強いかを問題とすることになる。一方、B フレームは、*C* が存在する場合としない場合を対比的に見ることを促すため、*C* のコントロールによって *E* の生起にどう変化が出るかを検出するモードになると予想される。「I. 図地フレーミングと確率推論」の節で紹介した  $\Delta P$  は、この差異化モード (differentiation mode) に適合するため、 $\Delta P$  がモデルとして記述力を増すことが予想される。また、*C* が存在する場合にむしろ *E* の発生が抑制される効果の検出も比較的容易になると予想されるため、出力は双極性（負～ゼロ～正）になる傾向があると考えられる。さらに言えば、関連性モードは観察 (observation) だけでも十分に機能するが、差異化モードは *C* の生起と不生起を自らコントロールする介入 (intervention) 事態の方が馴染みやすい。

以上、図地フレーミングの枠組みを提案し、このフレーミングによる効果を引き起こすフレームについて考察してきた。この理論的枠組みは、視覚認知という比較的低次の認知機能の概念から出発し、演繹、帰納、判断、意思決定、問題解決などの高次の認知機能のみならず、社会的認知や認知発達の研究で扱われてきた問題も包括する。これらの広範な現象が、本当に共通の認知のしくみに基づくものなのか、それとも単に似ているように見えるだけで実際にはまったく異なるしくみに基づくものなのかは、今後の研究成果を待たねばならない。このような枠組みを提示することの意義は、そのことによって推進される研究の広がり、説明がもたらす影響力の大きさによって判断されるべきであろう。そのためには、今後、さまざまな現象の共通性の裏づけを検討していく必要がある。ある現象について明らかになったことが、別の現象についての未発見の側面について予測を可能にするかどうかという検討である。たとえば、Birch & Bloom (2004) は、知識の呪縛による非対称性の神経基盤として、認知的抑制による仮説を提案しているが、誤信念課題はもとより、視覚的課題から基準率錯誤や透明性の錯誤まで、この同じ神経基盤に基づくものなのかどうかを吟味

していくといったことも有効かもしれない。さらには、今後、A/B 二つのフレームの心理的実在性も実証的に確認していく必要があるだろう。

## 謝辞

本研究は、日本学術振興会二国間交流事業共同研究 CHORUS プログラム J121000148（平成 23–25 年）および科学研究費補助金 22500247（平成 22–25 年）の助成を受けて実施された。本研究のアイデアは、同プログラムに参加する日仏メンバー、特に、服部郁子氏、高橋達二氏、David Over 氏との議論によって深めることができた。ここに記して謝意を表する。

## 文献

- Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a “theory of mind”? *Cognition*, *21*, 37–46. doi: 10.1016/0010-0277(85)90022-8
- Birch, S. A. J., & Bloom, P. (2003). Children are cursed: An asymmetric bias in mental-state attribution. *Psychological Science*, *14*, 283–286. doi: 10.1111/1467-9280.03436
- Birch, S. A. J., & Bloom, P. (2004). Understanding children’s and adults’ limitations in mental state reasoning. *Trends in Cognitive Sciences*, *8*, 255–260. doi: 10.1016/j.tics.2004.04.011
- Birch, S. A. J., & Bloom, P. (2007). The curse of knowledge in reasoning about false beliefs. *Psychological Science*, *18*, 382–386. doi: 10.1111/j.1467-9280.2007.01909.x
- Camerer, C. F., Loewenstein, G., & Weber, M. (1989). The curse of knowledge in economic settings: An experimental analysis. *Journal of Political Economy*, *97*, 1232–1254.
- Eddy, D. M. (1982). Probabilistic reasoning in clinical medicine: Problems and opportunities. In D. Kahneman, P. Slovic & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 249–267). Cambridge, UK: Cambridge University Press.
- Evans, J. St. B. T., & Over, D. E. (2004). *If*. New York, NY: Oxford University Press.
- Fischhoff, B. (1975). Hindsight is not equal to foresight: The effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance*, *1*, 288–299. doi: 10.1037/0096-1523.1.3.288
- Gale, M., & Ball, L. J. (2006). Dual-goal facilitation in Wason’s 2-4-6 task: What mediates successful rule discovery? *The Quarterly Journal of Experimental Psychology*, *59*, 873–885. doi: 10.1080/02724980543000051
- Gilovich, T., Medvec, V. H., & Savitsky, K. (2000). The spotlight effect in social judgment: An egocentric bias in estimates of the salience of one’s own actions and appearance. *Journal of Personality and Social Psychology*, *78*, 211–222. doi: 10.1037/0022-3514.78.2.211
- Gilovich, T., Savitsky, K., & Medvec, V. H. (1998). The illusion of transparency: Biased assessments of other’s ability to read one’s emotional states. *Journal of Personality and Social Psychology*, *75*, 332–346. doi: 10.1037/0022-3514.75.2.332
- Goodman, N. (1972). *Problems and projects*. Indianapolis, IN: Bobbs-Merrill.
- Gopnik, A., & Astington, J. W. (1988). Children’s understanding of representational change and its relation to the understanding of false belief and the appearance–reality distinction. *Child Development*, *59*, 26–37. doi: 10.2307/1130386

- Hattori, M., & Nishida, Y. (2009). Why does the base rate appear to be ignored? The equiprobability hypothesis. *Psychonomic Bulletin & Review*, *16*, 1065–1070. doi: 10.3758/PBR.16.6.1065
- Hattori, M., & Oaksford, M. (2007). Adaptive non-interventional heuristics for covariation detection in causal induction: Model comparison and rational analysis. *Cognitive Science*, *31*, 765–814. doi: 10.1080/03640210701530755
- Hempel, C. G. (1945). Studies in the Logic of Confirmation (I.). *Mind*, *54*, 1–26. doi: 10.2307/2250886
- Jenkins, H. M., & Sainsbury, R., S. (1969). The development of stimulus control through differential reinforcement. In N. J. Mackintosh & W. K. Honig (Eds.), *Fundamental issues in associative learning*. Halifax, Nova Scotia, Canada: Dalhousie University Press.
- Jenkins, H. M., & Sainsbury, R., S. (1970). Discrimination learning with the distinctive feature on positive or negative trials. In D. I. Mostofsky (Ed.), *Attention: Contemporary theory and analysis*. New York: Appleton-Century-Crofts.
- Jenkins, H. M., & Ward, W. C. (1965). Judgment of contingency between responses and outcomes. *Psychological Monographs: General and Applied*, *79*, 1–17. (Whole No. 594)
- Klayman, J., & Ha, Y.-W. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review*, *94*, 211–228. doi: 10.1037/0033-295X.94.2.211
- Levin, I. P., & Gaeth, G. J. (1988). How consumers are affected by the framing of attribute information before and after consuming the product. *Journal of Consumer Research*, *15*, 374–378.
- Mandel, D. R., & Lehman, D. R. (1998). Integration of contingency information in judgments of cause, covariation, and probability. *Journal of Experimental Psychology: General*, *127*, 269–285. doi: 10.1037/0096-3445.127.3.269
- McCarthy, J., & Hayes, P. (1969). Some philosophical problems from the standpoint of the artificial intelligence. In D. Michie (Ed.), *Machine intelligence* (Vol. 4, pp. 463–502). Edinburgh, UK: Edinburgh University Press.
- McKenzie, C. R. M., & Mikkelsen, L. A. (2000). The psychological side of Hempel's paradox of confirmation. *Psychonomic Bulletin & Review*, *7*, 360–366.
- Newman, J. P., Wolff, W. T., & Hearst, E. (1980). The feature-positive effect in adult human subjects. *Journal of Experimental Psychology: Human Learning and Memory*, *6*, 630–650. doi: 10.1037/0278-7393.6.5.630
- Perales, J. C., & Shanks, D. R. (2007). Models of covariation-based causal judgment: A review and synthesis. *Psychonomic Bulletin & Review*, *14*, 577–596. doi: 10.3758/BF03196807
- Perner, J., Leekam, S. R., & Wimmer, H. (1987). Three-year-olds' difficulty with false belief: The case for a conceptual deficit. *British Journal of Developmental Psychology*, *5*, 125–137. doi: 10.1111/j.2044-835X.1987.tb01048.x
- Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, *1*, 515–526. doi: 10.1017/S0140525X00076512
- Pylshyn, Z. W. (1987). *Robot's dilemma: The frame problem in artificial intelligence*. Norwood, NJ: Ablex.
- Ross, L., Greene, D., & House, P. (1977). The false consensus effect: An egocentric bias in social perception and attribution processes. *Journal of Experimental Social Psychology*, *13*, 279–301. doi: 10.1016/0022-1031(77)90049-X
- Rubin, E. (1915/1958). Figure and ground. In D. C. Beardslee & M. Wertheimer (Eds.), *Readings in perception* (pp. 194–203). Princeton, NJ: D. Van Nostrand. English translation of key sections from Rubin's dissertation.
- Rubin, E. (1921). *Visuell wahrgenommene Figuren: Studien in psychologischer Analyse [Visually perceived figures: Studies in psychological analysis]*. København: Gyldendalske Boghandel.
- Sanford, A. J., Fay, N., Stewart, A., & Moxey, L. (2002). Perspective in statements of quantity, with

- implications for consumer psychology. *Psychological Science*, *13*, 130–134.
- Schustack, M. W., & Sternberg, R. J. (1981). Evaluation of evidence in causal inference. *Journal of Experimental Psychology: General*, *110*, 101–120. doi: 10.1037/0096-3445.110.1.101
- Shafir, E. B. (1993). Choosing versus rejecting: Why some options are both better and worse than others. *Memory & Cognition*, *21*, 546–556.
- Smedslund, J. (1963). The concept of correlation in adults. *Scandinavian Journal of Psychology*, *4*, 165–173. doi: 10.1111/j.1467-9450.1963.tb01324.x
- Tukey, D. D. (1986). A philosophical and empirical analysis of subjects' modes of inquiry in Wason's 2-4-6 task. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, *38*, 5–33. doi: 10.1080/14640748608401583
- Tversky, A. (1977). Features of similarity. *Psychological Review*, *84*, 327–352. doi: 10.1037/0033-295X.84.4.327
- Tversky, A., & Kahneman, D. (1980). Causal schemas in judgments under uncertainty. In M. Fishbein (Ed.), *Progress in social psychology* (Vol. 1). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Tweney, R. D., Doherty, M. E., Worner, W. J., Pliske, D. B., Mynatt, C. R., Gross, K. A., et al. (1980). Strategies of rule discovery in an inference task. *The Quarterly Journal of Experimental Psychology*, *32*, 109–123. doi: 10.1080/00335558008248237
- Ward, W. C., & Jenkins, H. M. (1965). The display of information and the judgment of contingency. *Canadian Journal of Psychology*, *19*, 231–241.
- Wason, P. C. (1959). The processing of positive and negative information. *The Quarterly Journal of Experimental Psychology*, *11*, 92–107. doi: 10.1080/17470215908416296
- Wason, P. C. (1960). On the failure to eliminate hypotheses in a conceptual task. *The Quarterly Journal of Experimental Psychology*, *12*, 129–140.
- Wetherick, N. E. (1962). Eliminative and enumerative behaviour in a conceptual task. *Quarterly Journal of Experimental Psychology*, *14*, 246–249. doi: 10.1080/17470216208416542
- Wharton, C. M., Cheng, P. W., & Wickens, T. D. (1993). Hypothesis-testing strategies: Why two goals are better than one. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, *46A*, 743–758. doi: 10.1080/14640749308401037
- Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, *13*, 103–128. doi: 10.1016/0010-0277(83)90004-5

(本学文学部教授)